State-Sum Models of Piecewise Linear Quantum Gravity

This book gives a description of state-sum quantum gravity models which are based on triangulations of a smooth spacetime manifold. It contains detailed descriptions of Regge quantum gravity, spin-foam models and spin-cube models. Some other similar models, like the dynamical triangulations models, are only briefly described, since the sum over the spacetime triangulations is outside the scope of this book.

The book also contains a detailed description of the approach where the piecewise linear (PL) manifold corresponding to a smooth manifold triangulation is considered as the basic structure of the spacetime. Hence the PL structure is not an auxiliary tool used to define the gravitational path integral for a smooth spacetime, but it is taken as a physical property of the spacetime. Consequently, it is straightforward to construct a finite gravitational path integral. Another consequence is that the problems of determination of the classical limit and how to calculate the quantum corrections can be solved by using the effective action method. The smooth manifold limit problem is then replaced by the problem of a smooth manifold approximation for the effective action, which can be obtained by using the standard quantum field theory with a physical cutoff.

Some physical effects of a PL spacetime quantum gravity theory are also described, one of which is that the cosmological constant spectrum contains the observed value.

A short exposition of higher gauge theory is also given, which is a promising way to generalize a gauge symmetry by using the concept of a 2-group. A 2-group is a categorical generalization of a group, and by using this approach one can construct the spin-cube models of quantum gravity.

World Scientific
www.worldscientific.com
13233 hc

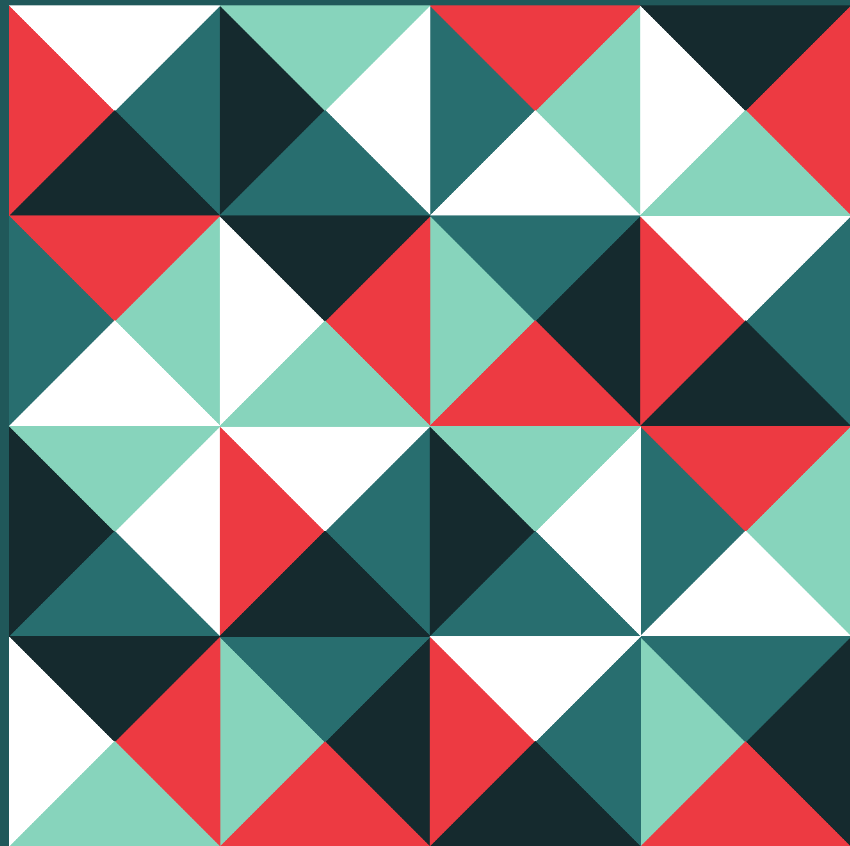
ISBN 978-981-126-931-8

State-Sum Models of Piecewise Linear Quantum Gravity

Miković
Vojinović

State-Sum Models of Piecewise Linear Quantum Gravity

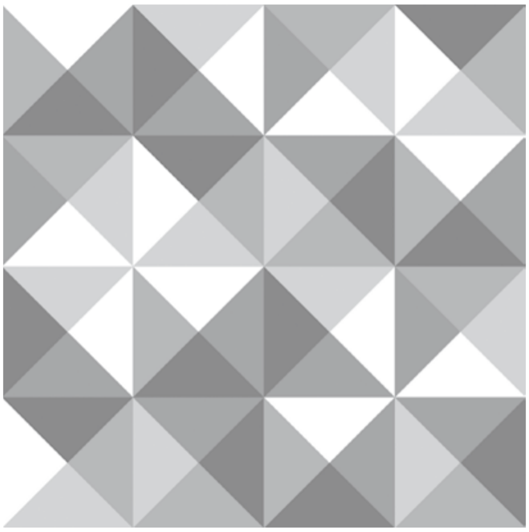
Aleksandar Miković
Marko Vojinović



World Scientific

State-Sum Models of Piecewise Linear Quantum Gravity

State-Sum Models of Piecewise Linear Quantum Gravity



Aleksandar Miković
Lusófona University, Portugal

Marko Vojinović
University of Belgrade, Serbia

World Scientific
NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI • TOKYO

Published by

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Control Number: 2023012301

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

STATE-SUM MODELS OF PIECEWISE LINEAR QUANTUM GRAVITY

ISBN 978-981-126-931-8 (hardcover)
ISBN 978-981-126-932-5 (ebook for institutions)
ISBN 978-981-126-933-2 (ebook for individuals)

For any available supplementary material, please visit
https://www.worldscientific.com/worldscibooks/10.1142/13233#t=suppl

Printed in Singapore

Preface

This book describes a novel approach in the study of quantum gravity (QG) state-sum models, which is based on the application of the effective action method from quantum field theory. Related to that is a study of the effect of a non-trivial path-integral (PI) measure on the PI finiteness, as well as a study on the dependence of the semi-classical expansion of the effective action on the PI measure.

Another novelty is a detailed study of the idea that the spacetime at small distances is not a smooth manifold but a piecewise linear (PL) manifold corresponding to a triangulation of a smooth manifold. This is a radical departure from the standard approach in PLQG, where the PL structure, i.e. the triangulation, is assumed to be non-physical and an auxiliary tool serving to define a QG theory on a smooth manifold. The main advantage of this paradigm shift is that finite QG path integrals can be constructed, while the semi-classical limit can be explored by using the effective action formalism. A smooth spacetime is then interpreted as an approximation to a PL manifold when the maximal edge length is small and the number of spacetime cells is large. The corresponding effective action can be then approximated by the usual QFT effective action with a cutoff, where the cutoff is determined by the average edge length in the spacetime triangulation. A further consequences of the idea that the spacetime is a PL manifold is that the cosmological constant has a continuous spectrum, and that the spectrum contains the observed value of the cosmological constant. We also describe some implications for quantum cosmology.

A description of higher gauge theory formulation of general relativity is also given, since the corresponding state-sum models do not suffer from the problems found in the spin-foam models of QG. These new state-sum models are called spin-cube models, and they are categorical generalizations of the spin-foam models, since one labels the edges, the triangles and the tetrahedra in a triangulation with representations of a 2-group, which is a categorical generalization of a group.

A major part of the book is devoted to the results obtained by the authors in the period from 2009 to 2016, and some more recent results have been also included. The book contains descriptions of the main PLQG approaches, but the emphasis is on a more detailed description of the Regge PLQG and the corresponding effective action. Our book can serve as an introductory text for a further research, so that it can be useful for young researchers, as well as for other researchers who are interested in this area.

We would like to thank John Barrett, Louis Crane, Laurent Freidel, Renate Loll, Steven Carlip, Ignatios Antoniadis and Hermann Nicolai for conversations over the years, who helped us to clarify our ideas.

Lisbon, March 2023

Aleksandar Miković and Marko Vojinović

Contents

universe

MDPI

Article

Henneaux–Teitelboim Gauge Symmetry and Its Applications to Higher Gauge Theories

Mihailo Đorđević †, Tijana Radenković †, Pavle Stipsić † and Marko Vojinović *,†

Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia; mdjordjevic@ipb.ac.rs (M.Đ.); rtijana@ipb.ac.rs (T.R.); pstipsic@ipb.ac.rs (P.S.)
* Correspondence: vmarko@ipb.ac.rs
† These authors contributed equally to this work.

Abstract: When discussing the gauge symmetries of any theory, the Henneaux–Teitelboim transformations are often underappreciated or even completely ignored, due to their on-shell triviality. Nevertheless, these gauge transformations play an important role in understanding the structure of the full gauge symmetry group of any theory, especially regarding the subgroup of diffeomorphisms. We give a review of the Henneaux–Teitelboim transformations and the resulting gauge group in the general case and then discuss its role in the applications to the class of topological theories called *nBF* models, relevant for the constructions of higher gauge theories and quantum gravity.

Keywords: gauge symmetry; trivial gauge transformations; *nBF* theory; Chern–Simons theory; diffeomorphism symmetry

check for updates

Citation: Đorđević, M.; Radenković, T.; Stipsić, P.; Vojinović, M. Henneaux–Teitelboim Gauge Symmetry and Its Applications to Higher Gauge Theories. *Universe* **2023**, *9*, 281. https://doi.org/10.3390/universe9060281

Received: 28 April 2023
Revised: 6 June 2023
Accepted: 7 June 2023
Published: 9 June 2023

1. Introduction

In modern theoretical physics, gauge symmetries play a very prominent role. The two most-fundamental theories we have, which describe almost all observed phenomena in nature—namely Einstein's theory of general relativity and the Standard Model of elementary particle physics—are gauge theories. From Maxwell's electrodynamics to various approaches to quantum gravity, gauge theories play a central role, and gauge symmetry represents one of their most-important aspects. In light of this, there is one class of gauge transformations that is often slightly neglected in the literature, due to their specific nature and properties.

In order to introduce this particular gauge symmetry in the most-elementary way possible, let us look at the following simple example. Every action $S[\phi_1, \phi_2]$, which depends on the fields $\phi_1(x)$ and $\phi_2(x)$, is invariant under the following gauge transformation:

$$\delta_0\phi_1(x) = \epsilon(x)\frac{\delta S}{\delta\phi_2(x)}\,, \qquad \delta_0\phi_2(x) = -\epsilon(x)\frac{\delta S}{\delta\phi_1(x)}\,, \tag{1}$$

as one can see by calculating the variation of the action:

$$\delta S[\phi_1, \phi_2] = \frac{\delta S}{\delta\phi_1}\delta_0\phi_1 + \frac{\delta S}{\delta\phi_2}\delta_0\phi_2 = 0\,. \tag{2}$$

This gauge symmetry exists for every action that is a functional of at least two fields, irrespective of any other gauge symmetry that the action may or may not have. In the literature, this symmetry is often called *trivial* gauge symmetry, since the form variations of the fields are identically zero on-shell. This is in contrast to all other gauge symmetries, which perform some nontrivial change of the fields on-shell.

It should be noted that, being trivial on-shell, the above transformations cannot play a role in obtaining any predictions about observables in a given theory, due to the intrinsic on-shell nature of the physical observables. For example, in practical situations of scattering experiments and measurements of cross-sections, this trivial symmetry is irrelevant. Nevertheless, when constructing a new theory, in general, the off-shell properties of the theory are important. As a typical example, path integral quantization prescription depends not only on the classical equations of motion, but on the whole action of the theory. In this sense, while these trivial transformations are not relevant for making predictions, they do have methodological relevance and value in theory construction, despite their on-shell triviality.

For example, these transformations in fact represent a very important part of the gauge symmetry for any theory and play a crucial role in various contexts, such as in the Batalin–Vilkovisky formalism (see [1] for a review and also the original papers [2–6]), or when discussing the diffeomorphism symmetry of the BF-like class of theories [7–11]. Furthermore, in general, a commutator of two ordinary gauge transformations will remain an ordinary gauge transformation only up to the above trivial transformations, meaning that the latter are important for the algebraic closure of all gauge transformations into a group.

To the best of our knowledge, the most-complete treatment and discussion of the above gauge transformations can be found in the book [12] by Marc Henneaux and Claudio Teitelboim. Therefore, in this paper, we opted to call them Henneaux–Teitelboim (HT) transformations. This naming can also be justified with the paper [7] by Gary Horowitz (published two years before the book [12]), where the author attributes these transformations to Henneaux and Teitelboim in a footnote and thanks them "for explaining this to me".

Regarding terminology, we should also note that we use the terms "gauge symmetry" and "gauge transformations" with a certain level of charity. Namely, one could argue that there are two distinct types of local symmetries—those that are obtained by a localization procedure from a corresponding global symmetry group (the procedure of "gauging" a global symmetry) and those that are intrinsically local, not obtained by any such localization procedure. It is not known whether HT symmetry belongs to the former or the latter class, since a global symmetry whose localization would give rise to HT transformations has not yet been shown to exist. Either way, in the literature, there is no established terminology that distinguishes the two classes of symmetries, and most often, both are called "gauge symmetries". Therefore, in what follows, for a lack of better terminology, we will adhere to this practice and describe HT transformations as a gauge symmetry.

In some of the modern approaches to the problem of quantum gravity based on the spinfoam formalism of loop quantum gravity [13,14], as well as in other applications of the so-called higher gauge theory (see [15] for a review and [16] for an application to quantum gravity), the description of gauge symmetry is being extended from the notion of a Lie group to different algebraic structures, called 2-groups, 3-groups, and in general, n-groups [17–27]. In this context, it is important to revisit and study the specific class of HT gauge symmetries, since they provide a nontrivial insight into the properties of these more general algebraic structures, as well as the physics behind the symmetries they describe.

The purpose of this paper is to provide a review of HT transformations in general and then discuss their properties and applications in two concrete models—the Chern–Simons theory and the $3BF$ theory. The Chern–Simons case is simple enough to serve as an illustrative toy example, while the $3BF$ theory represents a basis for the construction of a realistic theory of quantum gravity with matter within the context of the spinfoam formalism (see also [16,28–32]), discussing that its HT symmetry represents an important stepping stone towards the goal of a more realistic theory. The main result of this work represents a clarification of the structure of the gauge symmetry of a pure topological $3BF$ action, as well as the corresponding symmetry for the constrained $2BF$ action, which is classically equivalent to Einstein's general relativity. We also discuss in detail the relationship between diffeomorphism symmetry and the HT symmetry for the Chern–Simons and $3BF$ theories and offer some conceptual suggestions regarding the notion of gauge symmetry as it is being used in the literature.

The layout of the paper is as follows. In Section 2, we give a review of the general theory of HT transformations and their main properties. Section 3 is devoted to the example of HT symmetry in Chern–Simons theory, which is convenient due to its simplicity. In Section 4, we discuss the main case of HT symmetry in the $3BF$ and $2BF$ theories, which are important for applications in quantum gravity models. Finally, Section 5 contains an overview of the results, future research directions, and some concluding remarks.

The notation and conventions in the paper are as follows. When important, we assume the $(-,+,+,+)$ signature of the spacetime metric. The Greek indices from the middle of the alphabet, $\lambda,\mu,\nu,\dots$, represent spacetime indices and take values $0,1,\dots,D-1$, where D is the dimension of the spacetime manifold $\mathcal{M}_D$ under consideration. The Greek indices from the beginning of the alphabet, $\alpha,\beta,\gamma,\dots$, represent group indices, as well as Latin indices $a,b,c,\dots$ and uppercase Latin indices $A,B,C,\dots$ and $I,J,K,\dots$. All these indices will be assigned to various Lie groups under consideration. Lowercase Latin indices from the middle of the alphabet, $i,j,k,\dots$, are generic and will be used to count all fields in a given theory or for some other purpose depending on the context. Throughout the paper, we denote the space of algebra-valued differential p-forms as

$$\mathcal{A}^p(\mathcal{M},\mathfrak{a}) \equiv \Lambda^p(\mathcal{M})\otimes\mathfrak{a}\,,$$

where $\Lambda^p(\mathcal{M})$ is the ordinary space of differential p-forms over the manifold $\mathcal{M}$, while $\mathfrak{a}$ is some Lie algebra.

2. Review of HT Symmetry

We begin by studying some basic general properties of HT transformations. After the definition, we demonstrate that the group of HT transformations represents a normal subgroup of the *total* gauge group of a given theory, and we discuss the triviality of HT transformations and that they exhaust all possible trivial transformations. Finally, before moving on to concrete theories, we study the subtleties of the dependence of HT symmetry on the choice of the action.

2.1. Definition of HT Transformations

Given an action $S[\phi^i]$ as a functional of fields $\phi^i(x)$ ($i\in\{1,\dots,N\}$ where we assume $N\geqslant 2$), the infinitesimal HT transformation is defined as

$$\phi^i(x)\to\phi'^i(x)=\phi^i(x)+\delta_0\phi^i(x)\,, \tag{3}$$

where the form variations of the fields are defined as

$$\delta_0\phi^i(x)=\epsilon^{ij}(x)\frac{\delta S}{\delta\phi^j(x)}\,. \tag{4}$$

The variation of the action under HT transformations then gives

$$\delta S=\frac{\delta S}{\delta\phi^i}\delta_0\phi^i=\frac{\delta S}{\delta\phi^i}\frac{\delta S}{\delta\phi^j}\epsilon^{ij}\,. \tag{5}$$

If the HT parameters are chosen to be antisymmetric,

$$\epsilon^{ij}(x)=-\epsilon^{ji}(x)\,, \tag{6}$$

the variation of the action (5) is identically zero, and HT transformations (4) represent a gauge symmetry of the theory.

The most-striking thing in the above definition is the fact that we did not specify the action in any way. Aside from the assumption $N\geqslant 2$, which excludes only actions describing a single real scalar field, every action is invariant with respect to the HT transformations. In other words, *HT transformations are a gauge symmetry of essentially every theory*.

The second striking property of the definition is that the form variations of fields become zero on-shell, according to (4). In this sense, the HT symmetry is sometimes called *trivial symmetry*, in contrast to ordinary gauge symmetries that a theory may have, which transform the fields in a nontrivial way on-shell. Triviality is also the reason why HT gauge symmetry does not feature in any way in the Hamiltonian analysis of a theory, so only the presence of ordinary gauge symmetries can be deduced from the Hamiltonian formalism.

2.2. HT Symmetry Group and Its Properties

There are two general properties that can be formulated for HT transformations. The first is that HT transformations form a normal subgroup within the full group of gauge symmetries, while the second is that HT transformations exhaust the set of all possible trivial transformations. The consequence of these properties is that one can always write the total symmetry group of any theory as

$$\mathcal{G}_{\text{total}} = \mathcal{G}_{\text{nontrivial}} \ltimes \mathcal{G}_{HT} \,, \tag{7}$$

where $\mathcal{G}_{\text{nontrivial}}$ is the symmetry group of ordinary gauge transformations (if there are any), $\mathcal{G}_{HT}$ is the HT symmetry group, and the symbol $\ltimes$ stands for a semidirect product. One can also reformulate (7) as

$$\mathcal{G}_{\text{nontrivial}} = \mathcal{G}_{\text{total}} / \mathcal{G}_{HT} \,, \tag{8}$$

so that the group of ordinary gauge symmetries is represented as a quotient group.

The easiest way to demonstrate (7) is to prove that the Lie algebra corresponding to $\mathcal{G}_{HT}$ represents an ideal within the Lie algebra corresponding to $\mathcal{G}_{\text{total}}$. To that end, pick an arbitrary form variation of fields that represents a symmetry of the action and write it in the form

$$\hat{\delta}_0 \phi^i(x) = F^i(x) \,, \qquad \text{such that} \qquad \hat{\delta} S = \frac{\delta S}{\delta \phi^i} F^i \equiv 0 \,. \tag{9}$$

Then, using (4), we can take concatenated variations of this form variation and the HT form variation as

$$\delta_0 \hat{\delta}_0 \phi^i = \frac{\delta F^i}{\delta \phi^j} \frac{\delta S}{\delta \phi^k} \epsilon^{jk} \,,$$

and

$$\hat{\delta}_0 \delta_o \phi^i = \frac{\delta}{\delta \phi^k} \left(\epsilon^{ij} \frac{\delta S}{\delta \phi^j} \right) F^k = \frac{\delta \epsilon^{ij}}{\delta \phi^k} \frac{\delta S}{\delta \phi^j} F^k + \epsilon^{ij} \frac{\delta}{\delta \phi^j} \left(\frac{\delta S}{\delta \phi^k} F^k \right) - \epsilon^{ij} \frac{\delta S}{\delta \phi^k} \frac{\delta F^k}{\delta \phi^j} \,.$$

The term in the second parentheses is zero by (9), so the commutator of two-form variations becomes

$$[\delta_0 \,, \hat{\delta}_0] \phi^i = \left(\epsilon^{jk} \frac{\delta F^i}{\delta \phi^j} - \epsilon^{ji} \frac{\delta F^k}{\delta \phi^j} - \frac{\delta \epsilon^{ik}}{\delta \phi^j} F^j \right) \frac{\delta S}{\delta \phi^k} \,, \tag{10}$$

which is again an HT transformation, since the expression in the parentheses is antisymmetric with respect to indices i, k. Therefore, the commutator is always an element of HT algebra, which means that HT algebra itself is an ideal of the total symmetry algebra. At the Lie group level, this translates into (7).

The second general property is the statement that there are no other trivial transformations beside the HT transformations. Assuming that some transformation described by the form variation $\bar{\delta}_0 \phi^i$ is a gauge symmetry of the action that vanishes on-shell, i.e., that it satisfies

$$\frac{\delta S}{\delta \phi^i} \bar{\delta}_0 \phi^i = 0 \,, \qquad \text{and} \qquad \bar{\delta}_0 \phi^i \approx 0 \,,$$

then one can prove that this transformation is an HT transformation, i.e., there exists a choice of antisymmetric HT parameters ϵ^{ij} such that the form variation $\bar{\delta}_0 \phi^i$ is of type (4):

$$\bar{\delta}_0 \phi^i = \epsilon^{ij} \frac{\delta S}{\delta \phi^j} \,. \tag{11}$$

Provided certain suitable regularity conditions for the action S, this statement can be rigorously formulated as a theorem. However, we omitted the proof since it is technical and off topic for the purposes of this paper. The interested reader can find the details of both the theorem and the proof in [12], Appendix 10.A.2.

To sum up, the first property (10) tells us that one can always factorize the total gauge symmetry group into the form (7), while the second property (11) guarantees that the quotient group (8) contains only nontrivial gauge transformations. This factorization of the total symmetry group is a key result that lays the groundwork for any subsequent analysis of HT transformations in particular and gauge symmetry in general.

2.3. Dependence of HT Symmetry on the Action

The final property of HT transformations that needs to be discussed is their dependence on the choice of the action. Suppose we are given some action $S_{\text{old}}[\phi^i]$, where $i \in \{1, \dots, N\}$, which has the corresponding HT transformation described as in (4):

$$\delta_0^{\text{old}} \phi^i = \epsilon^{ij} \frac{\delta S_{\text{old}}}{\delta \phi^j} \,. \tag{12}$$

Now, suppose that we modify that action into another one, $S_{\text{new}}[\phi^i, \chi^k]$, where $k \in \{N + 1, \dots, N + M\}$, by adding an extra term to the old action:

$$S_{\text{new}}[\phi^i, \chi^k] = S_{\text{old}}[\phi^i] + S_{\text{extra}}[\phi^i, \chi^k] \,. \tag{13}$$

Here, χ^j are additional fields that may be introduced into the new action. The HT transformation corresponding to the new action can be written in the block-matrix form, made of blocks of sizes N and M, as follows:

$$\begin{pmatrix} \delta_0^{\text{new}} \phi^i \\ \delta_0^{\text{new}} \chi^k \end{pmatrix} = \left(\begin{array}{c|c} \epsilon^{ij} & \zeta^{il} \\ \hline \theta^{kj} & \psi^{kl} \end{array}\right) \begin{pmatrix} \dfrac{\delta S_{\text{new}}}{\delta \phi^j} \\ \dfrac{\delta S_{\text{new}}}{\delta \chi^l} \end{pmatrix} , \qquad \begin{array}{l} i, j \in \{1, \dots, N\} \,, \\ k, l \in \{N+1, \dots, N+M\} \,. \end{array} \tag{14}$$

Here, $\epsilon = -\epsilon^T$ is an antisymmetric $N \times N$ block of parameters ϵ^{ij}, ζ is a rectangular $N \times M$ block of parameters ζ^{il}, θ is a rectangular $M \times N$ block such that $\theta = -\zeta^T$, and finally, $\psi = -\psi^T$ is an antisymmetric $M \times M$ block of parameters ψ^{kl}. Overall, the total parameter matrix is antisymmetric, as required by (6).

The question one can now study is what is the relation between the two HT gauge symmetry groups $\mathcal{G}_{HT}^{\text{old}}$ and $\mathcal{G}_{HT}^{\text{new}}$ that correspond to the two actions. In practice, this question is most often relevant in cases when one introduces the piece S_{extra} as a gauge-fixing term, whose purpose is to break the ordinary gauge symmetry down to its subgroup:

$$G_{\text{nontrivial}}^{\text{new}} \subset G_{\text{nontrivial}}^{\text{old}} \,.$$

Naively, one might expect a similar relationship between the HT symmetry groups, $\mathcal{G}_{HT}^{\text{new}} \subset \mathcal{G}_{HT}^{\text{old}}$. However, looking at (12) and (14), this is obviously wrong. Namely, if $M \geqslant 1$, the HT symmetry of the new action is *larger* than the HT symmetry of the old action. Counting the number of independent parameters of both, one easily sees that

$$\dim(\mathcal{G}_{HT}^{\text{old}}) = \frac{N(N-1)}{2} \,, \qquad \dim(\mathcal{G}_{HT}^{\text{new}}) = \frac{(N+M)(N+M-1)}{2} \,,$$

so that the only possible relationship would be the opposite, $\mathcal{G}_{HT}^{\text{old}} \subset \mathcal{G}_{HT}^{\text{new}}$. However, in fact, this can also be shown to be wrong. Namely, one can choose the extra parameters ζ, θ and ψ to be zero in (14), reducing it to the form that is formally similar to (12):

$$\delta_0^{\text{new}} \phi^i = \epsilon^{ij} \frac{\delta S_{\text{new}}}{\delta \phi^j} \,.$$

However, taking into account the relationship (13) between the two actions, the HT transformation takes the form

$$\delta_0^{\text{new}} \phi^i = \epsilon^{ij} \frac{\delta S_{\text{old}}}{\delta \phi^j} + \epsilon^{ij} \frac{\delta S_{\text{extra}}}{\delta \phi^j} \,,$$

which is explicitly different from (12), due to the presence of the term S_{extra} in the action. Therefore, the gauge group $\mathcal{G}_{HT}^{\text{old}}$ is not a subgroup of $\mathcal{G}_{HT}^{\text{new}}$ either.

The overall conclusion is that introducing additional terms to the action changes the total gauge symmetry in a nontrivial way. On the one hand, the ordinary gauge symmetry group typically becomes *smaller* due to explicit symmetry breaking by the extra term. On the other hand, the HT gauge symmetry group may become *larger* if the extra term contains additional fields, but either way becomes *different*, as a consequence of the very presence of the extra term. Given this, one can conclude that the *total* symmetry groups for the two actions will always be mutually different:

$$\mathcal{G}_{\text{total}}^{\text{new}} = \mathcal{G}_{\text{nontrivial}}^{\text{new}} \ltimes \mathcal{G}_{HT}^{\text{new}} \qquad \neq \qquad \mathcal{G}_{\text{total}}^{\text{old}} = \mathcal{G}_{\text{nontrivial}}^{\text{old}} \ltimes \mathcal{G}_{HT}^{\text{old}} \,.$$

Specifically, one cannot claim that the group $\mathcal{G}_{\text{total}}^{\text{old}}$ is being broken down into $\mathcal{G}_{\text{total}}^{\text{new}}$ as its subgroup; such a relationship may hold exclusively for the quotient groups of ordinary gauge transformations.

In the next two sections, we will turn to explicit examples of all general properties and features of the HT symmetry that have been discussed above. Moreover, we will also discuss some additional particular properties, such as the fact that some nontrivial gauge subgroups of $\mathcal{G}_{\text{total}}$ are not simultaneously subgroups of $\mathcal{G}_{\text{nontrivial}}$, which is a consequence of the semidirect product in (7). One such example will be the diffeomorphism symmetry in the Chern–Simons and $3BF$ actions.

Let us conclude this section with one conceptual comment. Throughout the literature, the typical practice is to always take the quotient between the total and HT symmetry groups as in (8), in order to isolate the nontrivial gauge transformations, and call the latter simply as the "gauge symmetry" of a theory. This approach is in fact advocated for in [12]. However, we believe that this practice can be misleading and that one should instead describe the group $\mathcal{G}_{\text{total}}$ as "the gauge symmetry" of a theory, explicitly including the HT subgroup as a legitimate gauge symmetry group. Namely, despite the fact that it is often called "trivial", the consequences of its presence in $\mathcal{G}_{\text{total}}$ are far from trivial. Granted, it may often be enough to discuss the gauge symmetry on-shell, and then, one can indeed calculate all symmetry transformations only "up to equations of motion", with no mention of the HT subgroup. However, whenever one needs to discuss the gauge transformations off-shell, the HT subgroup simply cannot be ignored anymore. Typical situations include the Batalin–Vilkovisky formalism [1], various generalizations of gauge symmetry in the context of higher gauge theories and quantum gravity [33], and even the traditional contexts such as the Coleman–Mandula theorem [34]. The situations in which HT transformations play a significant role may be rare, but nevertheless, they tend to be important. Thus, in our opinion, it would be prudent to always be aware that, for any given theory, its total gauge symmetry group is in fact bigger, and more feature-rich, than just the group of ordinary gauge transformations that are typically discussed in the literature.

3. HT Symmetry in Chern–Simons Theory

As an illustrative example of the general properties of HT symmetry from the previous section, let us discuss the HT transformations for the simple case of the Chern–Simons theory. The Chern–Simons theory represents an excellent toy example since it is well known in the literature and most readers should be familiar with it.

Given any Lie group G, its corresponding Lie algebra $\mathfrak{g}$, and a three-dimensional manifold $\mathcal{M}_3$, the Chern–Simons theory can be defined as a topological field theory over a trivial principal bundle $G \to \mathcal{M}_3$, given by the action:

$$S_{CS} = \int_{\mathcal{M}_3} \langle A \wedge \mathrm{d}A \rangle_{\mathfrak{g}} + \frac{1}{3} \langle A \wedge [A \wedge A] \rangle_{\mathfrak{g}} \,. \tag{15}$$

Here, $A \in \mathcal{A}^1(\mathcal{M}_3, \mathfrak{g})$ is a $\mathfrak{g}$-valued connection one-form over a manifold $\mathcal{M}_3$, and $\langle \_ , \_ \rangle_{\mathfrak{g}}$ is a G-invariant symmetric nondegenerate bilinear form on $\mathfrak{g}$. One often rewrites the Chern–Simons action within the framework of the enveloping algebra of $\mathfrak{g}$, introducing the notion of a *trace* as

$$\mathrm{Tr}(XY) \equiv \langle X, Y \rangle_{\mathfrak{g}} \,,$$

for every $X, Y \in \mathfrak{g}$. Then, the Chern–Simons action can be rewritten as

$$S_{CS} = \int_{\mathcal{M}_3} \mathrm{Tr}\left(A \wedge \mathrm{d}A + \frac{2}{3} A \wedge A \wedge A \right) , \tag{16}$$

where, for the second term, one employs the identity $\mathrm{Tr}(X[Y, Z]) = \mathrm{Tr}(XYZ) - \mathrm{Tr}(XZY)$ for every $X, Y, Z \in \mathfrak{g}$.

The gauge symmetry of the Chern–Simons action consists of G-gauge transformations, determined with the parameters $\epsilon_{\mathfrak{g}}{}^I(x)$. Using the basis of generators T_I to expand the connection A into components as

$$A = A^I{}_\mu(x) \, \mathrm{d}x^\mu \otimes T_I \,,$$

the form variation of the connection components $A^I{}_\mu$ corresponding to gauge transformations can then be written as

$$\delta_0 A^I{}_\mu = \partial_\mu \epsilon_{\mathfrak{g}}{}^I - f_{JK}{}^I \epsilon_{\mathfrak{g}}{}^J A^K{}_\mu \,, \tag{17}$$

where $f_{JK}{}^I$ are the structure constants corresponding to the generators T_I. Therefore, the gauge symmetry of the Chern–Simons theory is usually quoted as the initially chosen Lie group G:

$$\mathcal{G}_{CS} = G \,. \tag{18}$$

However, as we have seen in the previous section, this is not the complete set of gauge transformations, and the *total* gauge group should in fact be

$$\mathcal{G}_{\text{total}} = \mathcal{G}_{CS} \ltimes \mathcal{G}_{HT} \,. \tag{19}$$

Let us define the HT transformations for the Chern–Simons action (15). If we denote the dimension of the Lie algebra $\mathfrak{g}$ as $\dim(\mathfrak{g}) = p$, the number of independent field components $A^I{}_\mu$ is $N = 3p$. The HT transformation is then defined with the HT parameters $\epsilon^{IJ}{}_{\mu\nu}(x)$ as

$$\delta_0 A^I{}_\mu = \epsilon^{IJ}{}_{\mu\nu} \frac{\delta S}{\delta A^J{}_\nu} \,. \tag{20}$$

The requirement that the variation of the action vanishes:

$$\delta S = \frac{\delta S}{\delta A^I{}_\mu} \frac{\delta S}{\delta A^J{}_\nu} \epsilon^{IJ}{}_{\mu\nu} = 0 \,,$$

enforces the antisymmetry restriction on the HT parameters:

$$\epsilon^{IJ}{}_{\mu\nu} = -\epsilon^{JI}{}_{\nu\mu} \,.$$

Note that this equation can be satisfied in two different ways—the parameters can be either antisymmetric with respect to group indices IJ and symmetric with respect to spacetime indices $\mu\nu$, or vice versa. We, therefore, have two possible choices for their symmetry properties. The first possibility is defined as

$$\epsilon^{IJ}{}_{\mu\nu} = \epsilon^{IJ}{}_{\nu\mu} = -\epsilon^{JI}{}_{\mu\nu} = -\epsilon^{JI}{}_{\nu\mu}\,, \tag{21}$$

while the second possibility is defined as

$$\epsilon^{IJ}{}_{\mu\nu} = \epsilon^{JI}{}_{\mu\nu} = -\epsilon^{IJ}{}_{\nu\mu} = -\epsilon^{JI}{}_{\nu\mu}\,. \tag{22}$$

Varying the action, one obtains an explicit form of the HT transformation:

$$\delta_0 A^I{}_\mu = \epsilon^{IJ}{}_{\mu\nu}\varepsilon^{\nu\rho\sigma}\left(\partial_\rho A_{J\sigma} - \partial_\sigma A_{J\rho} + f_{KLJ}\, A^K{}_\rho A^L{}_\sigma\right). \tag{23}$$

In order to demonstrate that HT transformations have highly nontrivial implications, despite being trivial on-shell, it is instructive to discuss diffeomorphisms. Namely, looking at the action (15), one expects that the theory has diffeomorphism symmetry, since it is formulated in a manifestly covariant way using differential forms. However, one can check that diffeomorphisms are not a subgroup of the ordinary gauge symmetry group $\mathcal{G}_{CS}$ given by (18), but nevertheless can be obtained as a subgroup of the total gauge group (19). In other words, one can demonstrate that

$$Diff(\mathcal{M}_3) \not\subset \mathcal{G}_{CS}\,, \qquad \text{but} \qquad Diff(\mathcal{M}_3) \subset \mathcal{G}_{\text{total}} = \mathcal{G}_{CS} \ltimes \mathcal{G}_{HT}\,.$$

Let us examine this in detail. The diffeomorphism transformation

$$x^\mu \to x'^\mu = x^\mu + \xi^\mu(x)\,, \tag{24}$$

determined by the parameter $\xi^\mu(x)$ represents a subgroup $Diff(\mathcal{M})$ of the full gauge symmetry of some given action, if for every field $\phi(x)$ in the theory and every choice of diffeomorphism parameters $\xi^\mu(x)$, there exists a choice of the gauge parameters $\epsilon^{\text{gauge}}(x)$ and the HT parameters $\epsilon^{\text{HT}}(x)$, such that:

$$\delta_0{}^{\text{diff}}\phi = \delta_0{}^{\text{gauge}}\phi + \delta_0{}^{\text{HT}}\phi\,. \tag{25}$$

In other words, if a theory has diffeomorphism symmetry, the diffeomorphism form variations of all the fields in the theory should be expressible in terms of their ordinary gauge and HT form variations.

In the case of Chern–Simons theory, this can be demonstrated explicitly. If one chooses the gauge parameters $\epsilon_{\mathfrak{g}}{}^I$ and the HT parameters $\epsilon^{IJ}{}_{\mu\nu}$ as

$$\epsilon_{\mathfrak{g}}{}^I = -\xi^\lambda A^I{}_\lambda\,, \qquad \epsilon^{IJ}{}_{\mu\nu} = -\frac{1}{2}\xi^\lambda \varepsilon_{\lambda\mu\nu} g^{IJ}\,, \tag{26}$$

where g^{IJ} is the inverse of $g_{IJ} \equiv \langle T_I\,, T_J\rangle_{\mathfrak{g}}$, one can apply Equations (25) using (17) and (23) to reproduce precisely the well-known diffeomorphism form variation of the connection $A^I{}_\mu$:

$$\delta_0{}^{\text{diff}} A^I{}_\mu = -A^I{}_\lambda \partial_\mu \xi^\lambda - \xi^\lambda \partial_\lambda A^I{}_\mu\,. \tag{27}$$

Therefore, as expected, despite the fact that $Diff(\mathcal{M}_3) \not\subset \mathcal{G}_{CS}$, one obtains that $Diff(\mathcal{M}_3) \subset \mathcal{G}_{\text{total}} = \mathcal{G}_{CS} \ltimes G_{HT}$. Note that the choice of HT parameters in (26) is nontrivial, which emphasizes the role of HT transformations and the fact that the full group of gauge symmetries is $\mathcal{G}_{\text{total}}$ rather than $\mathcal{G}_{CS}$. As we shall see in the next section, this property is not specific only to the Chern–Simons theory.

4. HT Symmetry in 3*BF* Theory

After discussing the Chern–Simons theory as a toy example, we move to the more important case of the 3*BF* theory. This theory is relevant for building models of quantum gravity; see [8,20,21,33,35]. Therefore, it is important to study its gauge symmetry and, in particular, the role of HT transformations.

4.1. Review of the 3BF Theory

Analogous to the fact that Chern–Simons theory is a topological theory based on a Lie group and a 3-dimensional manifold, the $3BF$ theory is also a topological theory based on a notion of a three-group and a 4-dimensional manifold. The notion of a three-group represents a categorical generalization of the notion of a group, in the context of higher gauge theory (HGT); see [15] for a review and motivation. For the purpose of defining the $3BF$ theory, we are interested in particular in a strict Lie three-group, which is known to be isomorphic to a so-called Lie two-crossed module; see [17–19] for details.

A Lie two-crossed module, denoted as $(L \xrightarrow{\delta} H \xrightarrow{\partial} G\,,\triangleright\,,\{\_\,,\_\}_{\rm pf})$, is an algebraic structure specified by three Lie groups G, H, and L, together with the homomorphisms $\delta : L \to H$ and $\partial : H \to G$, an action $\triangleright$ of the group G on all three groups, and a G-equivariant map, called the Peiffer lifting:

$$\{\_\,,\_\}_{\rm pf} : H \times H \to L\,.$$

In order for this structure to form a two-crossed module, the structure constants of algebras $\mathfrak{g}$, $\mathfrak{h}$, and $\mathfrak{l}$ (the Lie algebras corresponding to the Lie groups G, H, and L, respectively), as well as the maps ∂ and δ, the action $\triangleright$, and the Peiffer lifting, must satisfy certain axioms; see [20] for details.

Given a two-crossed module and a four-dimensional compact and orientable spacetime manifold $\mathcal{M}_4$, one can introduce the notion of a trivial principal three-bundle, in analogy with the notion of a trivial principal bundle constructed from an ordinary Lie group and a manifold; see [15]. Then, one can introduce the notion of a three-connection, an ordered triple (α, β, γ), where α, β, and γ are algebra-valued differential forms, $\alpha \in \mathcal{A}^1(\mathcal{M}_4, \mathfrak{g})$, $\beta \in \mathcal{A}^2(\mathcal{M}_4, \mathfrak{h})$, and $\gamma \in \mathcal{A}^3(\mathcal{M}_4, \mathfrak{l})$; see [17–19]. The corresponding fake hree-curvature $(\mathcal{F}, \mathcal{G}, \mathcal{H})$ is defined as:

$$\begin{gathered} \mathcal{F} = {\rm d}\alpha + \alpha \wedge \alpha - \partial\beta\,, \qquad \mathcal{G} = {\rm d}\beta + \alpha \wedge^{\triangleright} \beta - \delta\gamma\,, \\ \mathcal{H} = {\rm d}\gamma + \alpha \wedge^{\triangleright} \gamma + \{\beta \wedge \beta\}_{\rm pf}\,. \end{gathered} \tag{28}$$

Then, for a four-dimensional manifold $\mathcal{M}_4$, one can define the gauge-invariant topological $3BF$ action, based on the structure of a two-crossed module $(L \xrightarrow{\delta} H \xrightarrow{\partial} G\,,\triangleright\,,\{\_\,,\_\}_{\rm pf})$, by the action

$$S_{3BF} = \int_{\mathcal{M}_4} \langle B \wedge \mathcal{F} \rangle_{\mathfrak{g}} + \langle C \wedge \mathcal{G} \rangle_{\mathfrak{h}} + \langle D \wedge \mathcal{H} \rangle_{\mathfrak{l}}\,, \tag{29}$$

where $B \in \mathcal{A}^2(\mathcal{M}_4, \mathfrak{g})$, $C \in \mathcal{A}^1(\mathcal{M}_4, \mathfrak{h})$, and $D \in \mathcal{A}^0(\mathcal{M}_4, \mathfrak{l})$ are Lagrange multipliers and $\mathcal{F} \in \mathcal{A}^2(\mathcal{M}_4, \mathfrak{g})$, $\mathcal{G} \in \mathcal{A}^3(\mathcal{M}_4, \mathfrak{h})$, and $\mathcal{H} \in \mathcal{A}^4(\mathcal{M}_4, \mathfrak{l})$ represent the fake three-curvature given by Equation (28). The forms $\langle \_\,,\_ \rangle_{\mathfrak{g}}$, $\langle \_\,,\_ \rangle_{\mathfrak{h}}$, and $\langle \_\,,\_ \rangle_{\mathfrak{l}}$ are G-invariant symmetric nondegenerate bilinear forms on $\mathfrak{g}$, $\mathfrak{h}$, and $\mathfrak{l}$, respectively. The action (29) is an example of the so-called higher gauge theory.

By choosing the three bases of generators $\tau_\alpha \in \mathfrak{g}$, $t_a \in \mathfrak{h}$, and $T_A \in \mathfrak{l}$ of the three respective Lie algebras, one can expand all fields in the theory into components as

$$\begin{aligned} B &= \frac{1}{2} B^{\alpha}{}_{\mu\nu}(x)\, {\rm d}x^\mu \wedge {\rm d}x^\nu \otimes \tau_\alpha\,, & \alpha &= \alpha^{\alpha}{}_{\mu}(x)\, {\rm d}x^\mu \otimes \tau_\alpha\,, \\ C &= C^{a}{}_{\mu}(x)\, {\rm d}x^\mu \otimes t_a\,, & \beta &= \frac{1}{2} \beta^{a}{}_{\mu\nu}(x)\, {\rm d}x^\mu \wedge {\rm d}x^\nu \otimes t_a\,, \\ D &= D^A(x) T_A\,, & \gamma &= \frac{1}{3!} \gamma^{A}{}_{\mu\nu\rho}(x)\, {\rm d}x^\mu \wedge {\rm d}x^\nu \wedge {\rm d}x^\rho \otimes T_A\,. \end{aligned}$$

One can also make use of the following notation for the components of all maps present in the theory, in the same three bases:

$$
\begin{array}{llll}
[\tau_\alpha\,,\tau_\beta] = f_{\alpha\beta}{}^{\gamma}\tau_\gamma\,, & g_{\alpha\beta} = \langle\tau_\alpha\,,\tau_\beta\rangle_{\mathfrak{g}}\,, & \tau_\alpha \rhd \tau_\beta = \rhd_{\alpha\beta}{}^{\gamma}\tau_\gamma\,, & \delta T_A = \delta_A{}^a t_a\,,\\
[t_a\,,t_b] = f_{ab}{}^{c}t_c\,, & g_{ab} = \langle t_a\,,t_b\rangle_{\mathfrak{h}}\,, & \tau_\alpha \rhd t_a = \rhd_{\alpha a}{}^{b}t_b\,, & \partial t_a = \partial_a{}^{\alpha}\tau_\alpha\,,\\
[T_A\,,T_B] = f_{AB}{}^{C}T_C\,, & g_{AB} = \langle T_A\,,T_B\rangle_{\mathfrak{l}}\,, & \tau_\alpha \rhd T_A = \rhd_{\alpha A}{}^{B}T_B\,, & \{t_a\,,t_b\}_{\rm pf} = X_{ab}{}^{A}T_A\,.
\end{array}
$$

The complete gauge symmetry of the 3BF action was studied in [8] using the techniques of Hamiltonian analysis. It consists of five types of gauge transformations, G-, H-, L-, M-, and N-gauge transformations, determined with the independent parameters $\epsilon_{\mathfrak{g}}{}^{\alpha}(x)$, $\epsilon_{\mathfrak{h}}{}^{a}{}_{\mu}(x)$, $\epsilon_{\mathfrak{l}}{}^{A}{}_{\mu\nu}(x)$, $\epsilon_{\mathfrak{m}}{}^{\alpha}{}_{\mu}(x)$, and $\epsilon_{\mathfrak{n}}{}^{a}(x)$, respectively. The form variations of the fields B, C, D, α, β, and γ, obtained in [8] are given as follows:

$$
\begin{array}{rcl}
\delta_0 B^{\alpha}{}_{\mu\nu} &=& f_{\beta\gamma}{}^{\alpha}\epsilon_{\mathfrak{g}}{}^{\beta}B^{\gamma}{}_{\mu\nu} + 2C_{a[\mu|}\epsilon_{\mathfrak{h}}{}^{b}{}_{|\nu]}\rhd_{\beta b}{}^{a}g^{\alpha\beta} - D_A\rhd_{\beta B}{}^{A}\epsilon_{\mathfrak{l}}{}^{B}{}_{\mu\nu}g^{\alpha\beta} - 2\nabla_{[\mu|}\epsilon_{\mathfrak{m}}{}^{\alpha}{}_{|\nu]}\\
&& +\beta_{b\mu\nu}\rhd_{\beta a}{}^{b}\epsilon_{\mathfrak{n}}{}^{a}g^{\alpha\beta}\,,\\
\delta_0 C^{a}{}_{\mu} &=& \rhd_{\alpha b}{}^{a}\epsilon_{\mathfrak{g}}{}^{\alpha}C^{b}{}_{\mu} + 2D_A X_{(ab)}{}^{A}\epsilon_{\mathfrak{h}}{}^{b}{}_{\mu} - \partial^{a}{}_{\alpha}\epsilon_{\mathfrak{m}}{}^{\alpha}{}_{\mu} - \nabla_{\mu}\epsilon_{\mathfrak{n}}{}^{a}\,,\\
\delta_0 D^{A} &=& \rhd_{\alpha B}{}^{A}\epsilon_{\mathfrak{g}}{}^{\alpha}D^{B} + \delta^{A}{}_{a}\epsilon_{\mathfrak{n}}{}^{a}\,,\\
\delta_0 \alpha^{\alpha}{}_{\mu} &=& -\partial_{\mu}\epsilon_{\mathfrak{g}}{}^{\alpha} - f_{\beta\gamma}{}^{\alpha}\alpha^{\beta}{}_{\mu}\epsilon_{\mathfrak{g}}{}^{\gamma} - \partial_a{}^{\alpha}\epsilon_{\mathfrak{h}}{}^{a}{}_{\mu}\,,\\
\delta_0 \beta^{a}{}_{\mu\nu} &=& \rhd_{\alpha b}{}^{a}\epsilon_{\mathfrak{g}}{}^{\alpha}\beta^{b}{}_{\mu\nu} - 2\nabla_{[\mu|}\epsilon_{\mathfrak{h}}{}^{a}{}_{|\nu]} + \delta_A{}^{a}\epsilon_{\mathfrak{l}}{}^{A}{}_{\mu\nu}\,,\\
\delta_0 \gamma^{A}{}_{\mu\nu\rho} &=& \rhd_{\alpha B}{}^{A}\epsilon_{\mathfrak{g}}{}^{\alpha}\gamma^{B}{}_{\mu\nu\rho} + 3!\beta^{a}{}_{[\mu\nu}\epsilon_{\mathfrak{h}}{}^{b}{}_{\rho]}X_{(ab)}{}^{A} + \nabla_{\mu}\epsilon_{\mathfrak{l}}{}^{A}{}_{\nu\rho} - \nabla_{\nu}\epsilon_{\mathfrak{l}}{}^{A}{}_{\mu\rho} + \nabla_{\rho}\epsilon_{\mathfrak{l}}{}^{A}{}_{\mu\nu}\,.
\end{array}
\qquad (30)
$$

The gauge transformations (30) form a group $\mathcal{G}_{3BF}$:

$$
\mathcal{G}_{3BF} = \tilde{G} \ltimes (\tilde{H}_L \ltimes (\tilde{N} \times \tilde{M}))\,, \qquad (31)
$$

where $\tilde{G}$ denotes the group of G-gauge transformations, the H-gauge transformations together with the L-gauge transformations form the group $\tilde{H}_L$, while $\tilde{M}$ and $\tilde{N}$ are the groups of M- and N-gauge transformations, respectively. All these groups are determined from the structure of the initial chosen two-crossed module that defines the theory; see [8] for details.

However, as we have seen in the general theory in Section 2 and in the example of the Chern–Simons theory in Section 3, the symmetry group $\mathcal{G}_{3BF}$ determined by the Hamiltonian analysis does not include HT transformations, and therefore, the *total* gauge group should in fact be

$$
\mathcal{G}_{\rm total} = \mathcal{G}_{3BF} \ltimes \mathcal{G}_{HT}\,. \qquad (32)
$$

4.2. Explicit HT Transformations

Let us explicitly define the HT transformations for the 3BF action (29). If we denote the dimensions of the Lie algebras $\mathfrak{g}, \mathfrak{h}, \mathfrak{l}$ as

$$
\dim(\mathfrak{g}) = p\,, \qquad \dim(\mathfrak{h}) = q\,, \qquad \dim(\mathfrak{l}) = r\,,
$$

the number of independent field components in the theory can be counted according to the following table:

$B^{\alpha}{}_{\mu\nu}$	$C^{a}{}_{\mu}$	D^{A}	$\alpha^{\alpha}{}_{\mu}$	$\beta^{a}{}_{\mu\nu}$	$\gamma^{A}{}_{\mu\nu\rho}$
$6p$	$4q$	r	$4p$	$6q$	$4r$

The total number of independent field components is, therefore,

$$
N = 6p + 4q + r + 4p + 6q + 4r = 10p + 10q + 5r\,.
$$

Let ϕ^i denote all field components, where $i = 1, 2, \ldots, N$. We can write the fields schematically as a column-matrix with six blocks:

$$\phi^i = \begin{pmatrix} B^{\alpha}{}_{\mu\nu} \\ C^{a}{}_{\mu} \\ D^{A} \\ \alpha^{\alpha}{}_{\mu} \\ \beta^{a}{}_{\mu\nu} \\ \gamma^{A}{}_{\mu\nu\rho} \end{pmatrix}.$$

The HT transformation is then defined via the parameters $\epsilon^{ij}(x)$ as

$$\delta_0 \phi^i = \epsilon^{ij} \frac{\delta S}{\delta \phi^j}.$$

The requirement that the variation of the action vanishes enforces the antisymmetry restriction on the parameters, $\epsilon^{ij} = -\epsilon^{ji}$, for all $i, j \in \{1, \ldots, N\}$. These transformations can be represented more explicitly as a tensorial 6×6 block-matrix equation, in the following form:

$$\begin{pmatrix} \delta_0 B^{\alpha}{}_{\mu\nu} \\ \delta_0 C^{a}{}_{\mu} \\ \delta_0 D^{A} \\ \delta_0 \alpha^{\alpha}{}_{\mu} \\ \delta_0 \beta^{a}{}_{\mu\nu} \\ \delta_0 \gamma^{A}{}_{\mu\nu\rho} \end{pmatrix} = \left(\begin{array}{c|c|c|c|c|c} \epsilon^{\alpha\beta}{}_{\mu\nu\sigma\lambda} & \epsilon^{\alpha b}{}_{\mu\nu\sigma} & \epsilon^{\alpha B}{}_{\mu\nu} & \epsilon^{\alpha\beta}{}_{\mu\nu\sigma} & \epsilon^{\alpha b}{}_{\mu\nu\sigma\lambda} & \epsilon^{\alpha B}{}_{\mu\nu\sigma\lambda\xi} \\ \hline \mu^{a\beta}{}_{\mu\sigma\lambda} & \epsilon^{ab}{}_{\mu\sigma} & \epsilon^{aB}{}_{\mu} & \epsilon^{a\beta}{}_{\mu\sigma} & \epsilon^{ab}{}_{\mu\sigma\lambda} & \epsilon^{aB}{}_{\mu\sigma\lambda\xi} \\ \hline \mu^{A\beta}{}_{\sigma\lambda} & \mu^{Ab}{}_{\sigma} & \epsilon^{AB} & \epsilon^{A\beta}{}_{\sigma} & \epsilon^{Ab}{}_{\sigma\lambda} & \epsilon^{AB}{}_{\sigma\lambda\xi} \\ \hline \mu^{\alpha\beta}{}_{\mu\sigma\lambda} & \mu^{\alpha b}{}_{\mu\sigma} & \mu^{\alpha B}{}_{\mu} & \epsilon^{\alpha\beta}{}_{\mu\sigma} & \epsilon^{\alpha b}{}_{\mu\sigma\lambda} & \epsilon^{\alpha B}{}_{\mu\sigma\lambda\xi} \\ \hline \mu^{a\beta}{}_{\mu\nu\sigma\lambda} & \mu^{ab}{}_{\mu\nu\sigma} & \mu^{aB}{}_{\mu\nu} & \mu^{a\beta}{}_{\mu\nu\sigma} & \epsilon^{ab}{}_{\mu\nu\sigma\lambda} & \epsilon^{aB}{}_{\mu\nu\sigma\lambda\xi} \\ \hline \mu^{A\beta}{}_{\mu\nu\rho\sigma\lambda} & \mu^{Ab}{}_{\mu\nu\rho\sigma} & \mu^{AB}{}_{\mu\nu\rho} & \mu^{A\beta}{}_{\mu\nu\rho\sigma} & \mu^{Ab}{}_{\mu\nu\rho\sigma\lambda} & \epsilon^{AB}{}_{\mu\nu\rho\sigma\lambda\xi} \end{array}\right) \begin{pmatrix} \frac{1}{2}\frac{\delta S}{\delta B^{\beta}{}_{\sigma\lambda}} \\ \frac{\delta S}{\delta C^{b}{}_{\sigma}} \\ \frac{\delta S}{\delta D^{B}} \\ \frac{\delta S}{\delta \alpha^{\beta}{}_{\sigma}} \\ \frac{1}{2}\frac{\delta S}{\delta \beta^{b}{}_{\sigma\lambda}} \\ \frac{1}{3!}\frac{\delta S}{\delta \gamma^{B}{}_{\sigma\lambda\xi}} \end{pmatrix}. \tag{33}$$

The coefficients multiplying the variations of the action in the column on the right-hand side are there to compensate the overcounting of the independent field components. Due to the antisymmetry of HT parameters, all μ blocks (below the diagonal) are determined in terms of the ϵ blocks (above the diagonal), as follows. For the first column of the parameter matrix in (33), we have:

$$\begin{gathered} \mu^{b\alpha}{}_{\sigma\mu\nu} = -\epsilon^{\alpha b}{}_{\mu\nu\sigma}, \qquad \mu^{B\alpha}{}_{\mu\nu} = -\epsilon^{\alpha B}{}_{\mu\nu}, \qquad \mu^{\beta\alpha}{}_{\sigma\mu\nu} = -\epsilon^{\alpha\beta}{}_{\mu\nu\sigma}, \\ \mu^{b\alpha}{}_{\sigma\lambda\mu\nu} = -\epsilon^{\alpha b}{}_{\mu\nu\sigma\lambda}, \qquad \mu^{B\alpha}{}_{\sigma\lambda\xi\mu\nu} = -\epsilon^{\alpha B}{}_{\mu\nu\sigma\lambda\xi}. \end{gathered} \tag{34}$$

For the second column, we have:

$$\begin{gathered} \mu^{Ba}{}_{\mu} = -\epsilon^{aB}{}_{\mu}, \qquad \mu^{\beta a}{}_{\sigma\mu} = -\epsilon^{a\beta}{}_{\mu\sigma}, \\ \mu^{ba}{}_{\sigma\lambda\mu} = -\epsilon^{ab}{}_{\mu\sigma\lambda}, \qquad \mu^{Ba}{}_{\sigma\lambda\xi\mu} = -\epsilon^{aB}{}_{\mu\sigma\lambda\xi}. \end{gathered} \tag{35}$$

The μ parameters in the third column are determined via:

$$\mu^{\beta A}{}_{\sigma} = -\epsilon^{A\beta}{}_{\sigma}, \qquad \mu^{bA}{}_{\sigma\lambda} = -\epsilon^{Ab}{}_{\sigma\lambda}, \qquad \mu^{BA}{}_{\sigma\lambda\xi} = -\epsilon^{AB}{}_{\sigma\lambda\xi}, \tag{36}$$

while the remaining μ parameters in the fourth and fifth columns are determined as:

$$\mu^{b\alpha}{}_{\sigma\lambda\mu} = -\epsilon^{\alpha b}{}_{\mu\sigma\lambda}, \qquad \mu^{B\alpha}{}_{\sigma\lambda\xi\mu} = -\epsilon^{\alpha B}{}_{\mu\sigma\lambda\xi}, \qquad \mu^{Ba}{}_{\sigma\lambda\xi\mu\nu} = -\epsilon^{aB}{}_{\mu\nu\sigma\lambda\xi}. \tag{37}$$

Finally, in addition to all these, the parameters in the blocks on the diagonal also have to satisfy certain antisymmetry relations, specifically:

$$\begin{aligned} \epsilon^{\alpha\beta}{}_{\mu\nu\sigma\lambda} = -\epsilon^{\beta\alpha}{}_{\sigma\lambda\mu\nu}\,, \qquad \epsilon^{ab}{}_{\mu\sigma} = -\epsilon^{ba}{}_{\sigma\mu}\,, \qquad \epsilon^{AB} = -\epsilon^{BA}\,, \\ \epsilon^{\alpha\beta}{}_{\mu\sigma} = -\epsilon^{\beta\alpha}{}_{\sigma\mu}\,, \qquad \epsilon^{ab}{}_{\mu\nu\sigma\lambda} = -\epsilon^{ba}{}_{\sigma\lambda\mu\nu}\,, \qquad \epsilon^{AB}{}_{\mu\nu\rho\sigma\lambda\xi} = -\epsilon^{BA}{}_{\sigma\lambda\xi\mu\nu\rho}\,. \end{aligned} \tag{38}$$

Like in the example of the Chern–Simons theory from the previous section, these antisymmetry relations can be satisfied in various multiple ways. All those possibilities are allowed, as long as the identities (38) are satisfied. The final ingredient in (33) is the expressions for the variation of the action with respect to the fields, and these are given as follows:

$$\begin{aligned} \frac{\delta S}{\delta B^{\beta}{}_{\nu\rho}} &= \frac{1}{2}\varepsilon^{\nu\rho\sigma\tau}\mathcal{F}_{\beta\sigma\tau}\,, \\ \frac{\delta S}{\delta C^{b}{}_{\rho}} &= \frac{1}{3!}\varepsilon^{\rho\sigma\tau\lambda}\mathcal{G}_{b\sigma\tau\lambda}\,, \\ \frac{\delta S}{\delta D^{B}} &= \frac{1}{4!}\varepsilon^{\sigma\tau\lambda\xi}\mathcal{H}_{B\sigma\tau\lambda\xi}\,, \\ \frac{\delta S}{\delta \alpha^{\beta}{}_{\rho}} &= \frac{1}{2}\varepsilon^{\rho\tau\lambda\xi}\left(\nabla_{\tau}B_{\beta\lambda\xi} - \rhd_{\beta a}{}^{b}C_{b\tau}\beta^{a}{}_{\lambda\xi} + \frac{1}{3}\rhd_{\beta B}{}^{A}D_{A}\gamma^{B}{}_{\tau\lambda\xi}\right), \\ \frac{\delta S}{\delta \beta^{b}{}_{\nu\rho}} &= \varepsilon^{\nu\rho\sigma\tau}\left(\nabla_{\sigma}C_{b\tau} - \frac{1}{2}\partial_{b}{}^{\alpha}B_{\alpha\sigma\tau} + X_{(ab)}{}^{A}D_{A}\beta^{b}{}_{\sigma\tau}\right), \\ \frac{\delta S}{\delta \gamma^{B}{}_{\mu\nu\rho}} &= \varepsilon^{\mu\nu\rho\sigma}\left(\nabla_{\sigma}D_{B} + \delta_{B}{}^{a}C_{a\sigma}\right). \end{aligned} \tag{39}$$

4.3. Diffeomorphisms

As in the case of the Chern–Simons theory, it is instructive to discuss diffeomorphism symmetry. The $3BF$ action (29) obviously is diffeomorphism invariant, since it is formulated in a manifestly covariant way, using differential forms. However, one can check that the diffeomorphisms are not a subgroup of the gauge symmetry group $\mathcal{G}_{3BF}$ given by Equation (31), but nevertheless can be obtained as a subgroup of the total gauge group (32):

$$Diff(\mathcal{M}_4) \not\subset \mathcal{G}_{3BF}\,, \qquad \text{but} \qquad Diff(\mathcal{M}_4) \subset \mathcal{G}_{\text{total}} = \mathcal{G}_{3BF} \ltimes \mathcal{G}_{HT}\,. \tag{40}$$

Let us demonstrate this. Like in the Chern–Simons case, we want to demonstrate that the form variation of all fields corresponding to diffeomorphisms can be obtained as a suitable combination of the form variations for the ordinary gauge transformations (30) and the HT transformations (33). In other words, for an arbitrary choice of the diffeomorphism parameters $\xi^{\mu}(x)$ from (24), Equation (25) should hold in the case of the $3BF$ theory as well:

$$\delta_0{}^{\text{diff}}\,\phi = \delta_0{}^{\text{gauge}}\phi + \delta_0{}^{\text{HT}}\phi\,. \tag{41}$$

Indeed, this can be shown by a suitable choice of parameters. Regarding the parameters of the gauge transformations (30), the appropriate choice is given as:

$$\begin{aligned} \epsilon_{\mathfrak{g}}{}^{\alpha} = \xi^{\lambda}\alpha^{\alpha}{}_{\lambda}\,, \qquad \epsilon_{\mathfrak{h}}{}^{a}{}_{\mu} = -\xi^{\lambda}\beta^{a}{}_{\mu\lambda}\,, \qquad \epsilon_{\mathfrak{l}}{}^{A}{}_{\mu\nu} = -\xi^{\lambda}\gamma^{A}{}_{\mu\nu\lambda}\,, \\ \epsilon_{\mathfrak{m}}{}^{\alpha}{}_{\mu} = -\xi^{\lambda}B^{\alpha}{}_{\mu\lambda}\,, \qquad \epsilon_{\mathfrak{n}}{}^{a} = \xi^{\lambda}C^{a}{}_{\lambda}\,. \end{aligned} \tag{42}$$

Regarding the parameters of the HT transformations (33), we chose the following special case, with the majority of the parameters equated to zero:

$$\begin{pmatrix} \delta_0 B^{\alpha}{}_{\mu\nu} \\ \delta_0 C^{a}{}_{\mu} \\ \delta_0 D^{A} \\ \delta_0 \alpha^{\alpha}{}_{\mu} \\ \delta_0 \beta^{a}{}_{\mu\nu} \\ \delta_0 \gamma^{A}{}_{\mu\nu\rho} \end{pmatrix} = \left(\begin{array}{c|c|c|c|c|c} 0 & 0 & 0 & \epsilon^{\alpha\beta}{}_{\mu\nu\sigma} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \epsilon^{ab}{}_{\mu\sigma\lambda} & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & \epsilon^{AB}{}_{\sigma\lambda\xi} \\ \hline \mu^{\alpha\beta}{}_{\mu\sigma\lambda} & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & \mu^{ab}{}_{\mu\nu\sigma} & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & \mu^{AB}{}_{\mu\nu\rho} & 0 & 0 & 0 \end{array}\right) \begin{pmatrix} \frac{1}{2}\frac{\delta S}{\delta B^{\beta}{}_{\sigma\lambda}} \\ \frac{\delta S}{\delta C^{b}{}_{\sigma}} \\ \frac{\delta S}{\delta D^{B}} \\ \frac{\delta S}{\delta \alpha^{\beta}{}_{\sigma}} \\ \frac{1}{2}\frac{\delta S}{\delta \beta^{b}{}_{\sigma\lambda}} \\ \frac{1}{3!}\frac{\delta S}{\delta \gamma^{B}{}_{\sigma\lambda\xi}} \end{pmatrix} . \tag{43}$$

Of course, due to antisymmetry, the nonzero μ blocks take negative values of the corresponding ϵ blocks, in accordance with (34), (35), and (36). The three independent nonzero ϵ blocks are chosen as

$$\epsilon^{\alpha\beta}{}_{\mu\nu\sigma} = \xi^{\rho} g^{\alpha\beta} \varepsilon_{\mu\nu\sigma\rho} \,, \qquad \epsilon^{ab}{}_{\mu\sigma\lambda} = \xi^{\rho} g^{ab} \varepsilon_{\rho\mu\sigma\lambda} \,, \qquad \epsilon^{AB}{}_{\sigma\lambda\xi} = \xi^{\rho} g^{AB} \varepsilon_{\sigma\lambda\xi\rho} \,. \tag{44}$$

Finally, substituting (42) and (44) into (30) and (43), respectively, and then substituting all those results into (41), after a certain amount of work, one obtains precisely the standard form variations corresponding to diffeomorphisms:

$$\begin{aligned} \delta_0^{\text{diff}} B^{\alpha}{}_{\mu\nu} &= -B^{\alpha}{}_{\lambda\nu}\partial_{\mu}\xi^{\lambda} - B^{\alpha}{}_{\mu\lambda}\partial_{\nu}\xi^{\lambda} - \xi^{\lambda}\partial_{\lambda}B^{\alpha}{}_{\mu\nu} \,, \\ \delta_0^{\text{diff}} C^{a}{}_{\mu} &= -C^{a}{}_{\lambda}\partial_{\mu}\xi^{\lambda} - \xi^{\lambda}\partial_{\lambda}C^{a}{}_{\mu} \,, \\ \delta_0^{\text{diff}} D^{A} &= -\xi^{\lambda}\partial_{\lambda}D^{A} \,, \\ \delta_0^{\text{diff}} \alpha^{\alpha}{}_{\mu} &= -\alpha^{\alpha}{}_{\lambda}\partial_{\mu}\xi^{\lambda} - \xi^{\lambda}\partial_{\lambda}\alpha^{\alpha}{}_{\mu} \,, \\ \delta_0^{\text{diff}} \beta^{a}{}_{\mu\nu} &= -\beta^{a}{}_{\lambda\nu}\partial_{\mu}\xi^{\lambda} - \beta^{a}{}_{\mu\lambda}\partial_{\nu}\xi^{\lambda} - \xi^{\lambda}\partial_{\lambda}\beta^{a}{}_{\mu\nu} \,, \\ \delta_0^{\text{diff}} \gamma^{A}{}_{\mu\nu\rho} &= -\gamma^{A}{}_{\lambda\nu\rho}\partial_{\mu}\xi^{\lambda} - \gamma^{A}{}_{\mu\lambda\rho}\partial_{\nu}\xi^{\lambda} - \gamma^{A}{}_{\mu\nu\lambda}\partial_{\rho}\xi^{\lambda} - \xi^{\lambda}\partial_{\lambda}\gamma^{A}{}_{\mu\nu\rho} \,. \end{aligned} \tag{45}$$

This establishes both relations (40), as we set out to demonstrate. We note again that the HT transformations play a crucial role in obtaining the result, since we had to choose the parameters (44) in a nontrivial manner.

4.4. Symmetry Breaking in 2BF Theory

Let us now turn to the topic of symmetry breaking and the way it influences HT transformations. To that end, we studied the topological $2BF$ action, which is a special case of the $3BF$ action (29) without the last term:

$$S_{2BF} = \int_{\mathcal{M}_4} \langle B \wedge \mathcal{F} \rangle_{\mathfrak{g}} + \langle C \wedge \mathcal{G} \rangle_{\mathfrak{h}} \,. \tag{46}$$

In order to be even more concrete, let us fix a two-crossed module structure with the following choice of groups:

$$G = SO(3,1) \,, \qquad H = \mathbb{R}^4 \,, \qquad L = \{e\} \,.$$

In other words, we interpret group G as the Lorentz group, group H as the spacetime translations group, while group L is trivial, for simplicity. This choice corresponds to the so-called Poincaré two-group; see [16] for details. Since the generators of the Lorentz group can be conveniently counted using the antisymmetric combinations of indices from the group of translations, instead of the G-group indices α, we shall systematically write $[ab] \in \{01, 02, 03, 12, 13, 23\}$, where $a, b \in \{0, 1, 2, 3\}$ are H-group indices, and the brackets denote antisymmetrization. With a further change in notation from the connection 1-form α to the spin-connection 1-form ω, the curvature 2-form $\mathcal{F}(\alpha)$ to $R(\omega)$, and interpreting the Lagrange multiplier 1-form C as the tetrad 1-form e, the $2BF$ action can be rewritten in new notation as

$$S_{2BF} = \int_{\mathcal{M}_4} B^{[ab]} \wedge R_{[ab]} + e^a \wedge \mathcal{G}_a \,. \tag{47}$$

The ordinary gauge symmetry group for this action has a form similar to (31):

$$\mathcal{G}_{2BF} = \tilde{G} \ltimes (\tilde{H} \ltimes (\tilde{N} \times \tilde{M}))\,, \tag{48}$$

while the total group of gauge symmetries is extended by the HT transformations, so that

$$\mathcal{G}_{\text{total}} = \mathcal{G}_{2BF} \ltimes \mathcal{G}_{HT}\,. \tag{49}$$

The explicit HT transformations are written as a tensorial 4×4 block-matrix equation, in the form

$$\begin{pmatrix} \delta_0 B^{[ab]}{}_{\mu\nu} \\ \delta_0 e^a{}_\mu \\ \delta_0 \omega^{[ab]}{}_\mu \\ \delta_0 \beta^a{}_{\mu\nu} \end{pmatrix} = \left(\begin{array}{c|c|c|c} \epsilon^{[ab][cd]}{}_{\mu\nu\sigma\lambda} & \epsilon^{[ab]c}{}_{\mu\nu\sigma} & \epsilon^{[ab][cd]}{}_{\mu\nu\sigma} & \epsilon^{[ab]c}{}_{\mu\nu\sigma\lambda} \\ \hline \mu^{a[cd]}{}_{\mu\sigma\lambda} & \epsilon^{ac}{}_{\mu\sigma} & \epsilon^{a[cd]}{}_{\mu\sigma} & \epsilon^{ac}{}_{\mu\sigma\lambda} \\ \hline \mu^{[ab][cd]}{}_{\mu\sigma\lambda} & \mu^{[ab]c}{}_{\mu\sigma} & \epsilon^{[ab][cd]}{}_{\mu\sigma} & \epsilon^{[ab]c}{}_{\mu\sigma\lambda} \\ \hline \mu^{a[cd]}{}_{\mu\nu\sigma\lambda} & \mu^{ac}{}_{\mu\nu\sigma} & \mu^{a[cd]}{}_{\mu\nu\sigma} & \epsilon^{ac}{}_{\mu\nu\sigma\lambda} \end{array}\right) \begin{pmatrix} \frac{1}{4}\frac{\delta S}{\delta B^{[cd]}{}_{\sigma\lambda}} \\ \frac{\delta S}{\delta e^c{}_\sigma} \\ \frac{1}{2}\frac{\delta S}{\delta \omega^{[cd]}{}_\sigma} \\ \frac{1}{2}\frac{\delta S}{\delta \beta^c{}_{\sigma\lambda}} \end{pmatrix}, \tag{50}$$

where the usual antisymmetry rules apply. Here, we have

$$\begin{aligned} \frac{\delta S}{\delta B^{[cd]}{}_{\sigma\lambda}} &= \varepsilon^{\mu\nu\sigma\lambda} R_{[cd]\mu\nu}\,, \\ \frac{\delta S}{\delta \omega^{[cd]}{}_\sigma} &= \varepsilon^{\sigma\mu\nu\rho}\left(\nabla_\mu B_{[cd]\nu\rho} - e_{[c|\mu}\beta_{|d]\nu\rho}\right), \\ \frac{\delta S}{\delta e^c{}_\sigma} &= \frac{1}{2}\varepsilon^{\sigma\mu\nu\rho}\nabla_\mu \beta_{c\nu\rho}\,, \\ \frac{\delta S}{\delta \beta^c{}_{\sigma\lambda}} &= \varepsilon^{\mu\nu\sigma\lambda}\nabla_\mu e_{c\nu}\,. \end{aligned} \tag{51}$$

The $2BF$ action (46) is topological, in the sense that it has no local propagating degrees of freedom. In this sense, it does not represent a theory of any realistic physics. In order to construct a more realistic theory, one proceeds by introducing the so-called *simplicity constraint* term into the action, which changes the equations of motion of the theory so that it does have nontrivial degrees of freedom. An example is the action

$$S_{GR} = \int_{\mathcal{M}_4} B^{[ab]} \wedge R_{[ab]} + e^a \wedge \nabla\beta_a - \lambda_{[ab]} \wedge \left(B^{[ab]} - \frac{1}{16\pi l_p^2}\varepsilon^{abcd} e_c \wedge e_d\right), \tag{52}$$

where the new constraint term features another Lagrange multiplier two-form $\lambda_{[ab]}$. By virtue of the simplicity constraint, the theory becomes equivalent to general relativity, in the sense that the corresponding equations of motion reduce to vacuum Einstein field equations (see [16] for the analysis and proof). In this sense, constraint terms of various types are important when building more realistic theories; see [20] for more examples.

However, adding the simplicity constraint term also changes the gauge symmetry of the theory. In particular, it breaks the gauge group $\mathcal{G}_{2BF}$ from (48) down to one of its subgroups, so that the symmetry group of the action S_{GR} is

$$\mathcal{G}_{GR} \subset \mathcal{G}_{2BF}\,. \tag{53}$$

This is expected and unsurprising. What is less obvious, however, is that the group of HT transformations $\tilde{\mathcal{G}}_{HT}$ of the action S_{GR} *is not a subgroup* of the HT group $\mathcal{G}_{HT}$ of the original action S_{2BF}:

$$\tilde{\mathcal{G}}_{HT} \not\subset \mathcal{G}_{HT}\,, \tag{54}$$

which implies that

$$\mathcal{G}^{GR}_{\text{total}} \not\subset \mathcal{G}^{2BF}_{\text{total}}\,, \tag{55}$$

despite (53).

Let us demonstrate this. Since the action (52) features an additional field $\lambda^{[ab]}{}_{\mu\nu}(x)$, the HT transformations (50) have to be modified to take this into account and obtain the following 5×5 block-matrix form:

$$\begin{pmatrix} \delta_0 B^{[ab]}{}_{\mu\nu} \\ \delta_0 e^a{}_\mu \\ \delta_0 \omega^{[ab]}{}_\mu \\ \delta_0 \beta^a{}_{\mu\nu} \\ \delta_0 \lambda^{[ab]}{}_{\mu\nu} \end{pmatrix} = \left(\begin{array}{c|c|c|c|c} \epsilon^{[ab][cd]}{}_{\mu\nu\sigma\lambda} & \epsilon^{[ab]c}{}_{\mu\nu\sigma} & \epsilon^{[ab][cd]}{}_{\mu\nu\sigma} & \epsilon^{[ab]c}{}_{\mu\nu\sigma\lambda} & \zeta^{[ab][cd]}{}_{\mu\nu\sigma\xi} \\ \hline \mu^{a[cd]}{}_{\mu\sigma\lambda} & \epsilon^{ac}{}_{\mu\sigma} & \epsilon^{a[cd]}{}_{\mu\sigma} & \epsilon^{ac}{}_{\mu\sigma\lambda} & \zeta^{a[cd]}{}_{\mu\sigma\xi} \\ \hline \mu^{[ab][cd]}{}_{\mu\sigma\lambda} & \mu^{[ab]c}{}_{\mu\sigma} & \epsilon^{[ab][cd]}{}_{\mu\sigma} & \epsilon^{[ab]c}{}_{\mu\sigma\lambda} & \zeta^{[ab][cd]}{}_{\mu\sigma\xi} \\ \hline \mu^{a[cd]}{}_{\mu\nu\sigma\lambda} & \mu^{ac}{}_{\mu\nu\sigma} & \mu^{a[cd]}{}_{\mu\nu\sigma} & \epsilon^{ac}{}_{\mu\nu\sigma\lambda} & \zeta^{a[cd]}{}_{\mu\nu\sigma\xi} \\ \hline \theta^{[ab][cd]}{}_{\mu\nu\sigma\lambda} & \theta^{[ab]c}{}_{\mu\nu\sigma} & \theta^{[ab][cd]}{}_{\mu\nu\sigma} & \theta^{[ab]c}{}_{\mu\nu\sigma\lambda} & \psi^{[ab][cd]}{}_{\mu\nu\sigma\xi} \end{array}\right) \begin{pmatrix} \frac{1}{4}\frac{\delta S_{GR}}{\delta B^{[cd]}{}_{\sigma\lambda}} \\ \frac{\delta S_{GR}}{\delta e^c{}_\sigma} \\ \frac{1}{2}\frac{\delta S_{GR}}{\delta \omega^{[cd]}{}_\sigma} \\ \frac{1}{2}\frac{\delta S_{GR}}{\delta \beta^c{}_{\sigma\lambda}} \\ \frac{1}{4}\frac{\delta S_{GR}}{\delta \lambda^{[cd]}{}_{\sigma\xi}} \end{pmatrix}, \tag{56}$$

where

$$\begin{aligned} \frac{\delta S_{GR}}{\delta B^{[cd]}{}_{\sigma\lambda}} &= \varepsilon^{\mu\nu\sigma\lambda}\Big(R_{[cd]\mu\nu} - \lambda_{[cd]\mu\nu}\Big)\,, \\ \frac{\delta S_{GR}}{\delta \omega^{[cd]}{}_{\sigma}} &= \varepsilon^{\sigma\mu\nu\rho}\Big(\nabla_\mu B_{[cd]\nu\rho} - e_{[c|\mu}\beta_{|d]\nu\rho}\Big)\,, \\ \frac{\delta S_{GR}}{\delta e^c{}_\sigma} &= \frac{1}{2}\varepsilon^{\sigma\mu\nu\rho}\Big(\nabla_\mu \beta_{c\nu\rho} + \frac{1}{8\pi l_p^2}\varepsilon_{abcd}\lambda^{[ab]}{}_{\mu\nu}e^d{}_\rho\Big)\,, \\ \frac{\delta S_{GR}}{\delta \beta^c{}_{\sigma\lambda}} &= \varepsilon^{\mu\nu\sigma\lambda}\nabla_\mu e_{c\nu}\,, \\ \frac{\delta S_{GR}}{\delta \lambda^{[cd]}{}_{\sigma\xi}} &= -\varepsilon^{\sigma\xi\mu\nu}\Big(B_{[cd]\mu\nu} - \frac{1}{8\pi l_p^2}\varepsilon_{abcd}e^a{}_\mu e^b{}_\nu\Big)\,. \end{aligned} \tag{57}$$

We can now investigate the differences in the form of HT transformations for the topological and constrained theory. First, comparing (56) to (50), we see that the HT transformations in the constrained theory feature *more gauge parameters* than are present in the topological theory. Namely, compared to S_{2BF}, the action S_{GR} features an extra Lagrange multiplier two-form $\lambda^{[ab]}$, which extends the matrix of HT parameters from 4×4 blocks to 5×5 blocks, and, therefore, introduces the new parameters ζ and ψ (and θ, which are the negative of ζ due to antisymmetry). This means that the group $\tilde{\mathcal{G}}_{HT}$ for the constrained theory is *larger* than the group $\mathcal{G}_{HT}$ for the topological theory. On the one hand, this immediately proves (54) and, consequently, (55). On the other hand, one can ask the opposite question—given that $\tilde{\mathcal{G}}_{HT}$ is larger than $\mathcal{G}_{HT}$, is the latter maybe a subgroup of the former?

The answer to this question is negative:

$$\mathcal{G}_{HT} \not\subset \tilde{\mathcal{G}}_{HT}\,, \tag{58}$$

which together with (54) implies our final conclusion:

$$\mathcal{G}_{HT} \neq \tilde{\mathcal{G}}_{HT}\,. \tag{59}$$

In order to demonstrate (58), we can try to set all extra parameters ζ, ψ, and θ to zero in (56), reducing it to the same form as (50). This would naively suggest that $\mathcal{G}_{HT}$ indeed is a subgroup of $\tilde{\mathcal{G}}_{HT}$. However, upon closer inspection, we can observe that this is not true, since the functional derivatives (57) are different from (51). Namely, even taking into account that the choice $\zeta = \psi = \theta = 0$ eliminates the fifth equation from (57), the first four equations are still different from their counterparts (51) because of the presence of the Lagrange multiplier $\lambda^{[ab]}$ in the action. The Lagrange multiplier is a field in the theory, and generically, it is not zero, since it is determined by the equation of motion:

$$\lambda^{[ab]}{}_{\mu\nu} = R^{[ab]}{}_{\mu\nu}\,.$$

Therefore, the HT transformations (56) in fact cannot be reduced to the HT transformations (50) by setting the extra parameters equal to zero, which proves (58) and (59).

The overall consequences from the above analysis are as follows. The topological action S_{2BF} has a large ordinary gauge group $\mathcal{G}_{2BF}$ and a small HT symmetry group $\mathcal{G}_{HT}$. When one changes the action to S_{GR} by adding a simplicity constraint term, two things happen—the ordinary gauge group breaks down to its subgroup $\mathcal{G}_{GR}$, so that it becomes smaller, while the HT symmetry group *grows larger* to a completely different group $\tilde{\mathcal{G}}_{HT}$. In effect, the *total* gauge groups for the two actions are intrinsically different:

$$\mathcal{G}^{2BF}_{\text{total}} = \mathcal{G}_{2BF} \ltimes \mathcal{G}_{HT} \qquad \neq \qquad \mathcal{G}^{GR}_{\text{total}} = \mathcal{G}_{GR} \ltimes \tilde{\mathcal{G}}_{HT} \,,$$

in the sense that neither is a subgroup of the other. This conclusion is often overlooked in the literature, which mostly puts emphasis on the symmetry breaking of the ordinary gauge group down to its subgroup.

Let us state here, without proof, that the action (52) represents an example of a non-topological action, for which one can also demonstrate a property analogous to (40), that diffeomorphisms are not a subgroup of its ordinary gauge group, but are a subgroup of the total gauge group. Simply put, given that the simplicity constraint term in (52) breaks the ordinary gauge symmetry group $\mathcal{G}_{2BF}$ into its subgroup $\mathcal{G}_{GR}$ (see (53)), one can expect that diffeomorphisms are not a subgroup of $\mathcal{G}_{GR}$, since they are not a subgroup of the larger group $\mathcal{G}_{2BF}$ of the topological action (46). Nevertheless, since the action (52) is written in a manifestly covariant form, diffeomorphisms are certainly a symmetry of the action and, thus, must be a subgroup of the total gauge group $\mathcal{G}^{GR}_{\text{total}} = \mathcal{G}_{GR} \ltimes \tilde{\mathcal{G}}_{HT}$, in line with the statement analogous to (40). We leave the details of the proof as an exercise for the reader. The point of this analysis was to demonstrate that the interplay (40) between diffeomorphisms and the HT symmetry is a generic property of a large class of actions, including the physically relevant ones, and not limited to examples of topological theories such as the Chern–Simons or nBF models.

As the last comment, let us remark that, in fact, almost all conclusions discussed for the cases of the Chern–Simons, $3BF$, and $2BF$ theories are not really specific to these concrete cases. One can easily generalize our analysis to any other theory, and the conclusions should remain unchanged, except maybe in some corner cases.

5. Conclusions

Let us review the results. In Section 2, we gave a short overview of HT gauge symmetry and discussed its most-important general properties. First, the HT group is a normal subgroup of the total group of gauge symmetries of any given action. Second, HT transformations exhaust all "trivial" (i.e., vanishing on-shell) symmetries, in the sense that there are no trivial symmetries that are not of the HT type. Finally, adding additional terms into the action substantially changes the HT group, often enlarging it. This may be considered a counterintuitive result, since usually adding additional terms in the action serves the purpose of fixing the gauge and, thus, is meant to reduce the gauge symmetry, rather than to enlarge it.

After these general results, in Section 3, we discussed the HT symmetry of the Chern–Simons action, which is a convenient toy example that neatly displays the general features from Section 2. Special attention was given to the issue of diffeomorphisms, and it was shown that, while they are not a subgroup of the ordinary gauge group of the Chern–Simons action, they nevertheless do represent a proper subgroup of the total gauge symmetry, and the HT subgroup plays a nontrivial role in demonstrating this.

Section 4 was devoted to the study of HT symmetry in the $2BF$ and $3BF$ theories, which are relevant for the constructions of realistic quantum gravity models within the generalized spinfoam approach and higher gauge theory. After a brief review and introduction to the notion of three-groups and the $3BF$ theory, appropriate HT transformations were explicitly constructed, complementing the ordinary group of gauge symmetries of the $3BF$ action based on a given three-group. This gave us the total gauge symmetry group for this class of theories. We again discussed the issue of diffeomorphisms and demonstrated again that they are a subgroup of the total gauge group, without being a subgroup of the ordinary gauge group, just like in the case of the Chern–Simons theory. Finally, we introduced a completely concrete example of the $2BF$ theory based on the Poincaré two-group, which becomes classically equivalent to Einstein's general relativity when one introduces the additional term into the action, called the simplicity constraint. As argued in general in Section 2, the presence of this constraint breaks the ordinary gauge group down into its subgroup, while simultaneously enlarging the HT group, since it introduces an additional Lagrange multiplier field into the action. This represents an explicit example of the general statement from Section 2 that the total gauge symmetry group changes nontrivially, as opposed to simply breaking down to its subgroup.

It should be noted that the analysis and results discussed here do not cover everything that can be said about HT symmetry. Among the topics not covered, one can mention the question of an explicit form of finite HT transformations, as opposed to infinitesimal ones. Can one write down finite HT transformations in closed form, either for some conveniently chosen action or maybe even in general? A related topic is the explicit evaluation of the commutator of two HT transformations, or equivalently, the structure constants of the HT Lie algebra, or in yet other words, the multiplication rule in the group $\mathcal{G}_{HT}$. Is the group Abelian or not and for which choices of the action? Finally, one would also like to know the topological properties of the group $\mathcal{G}_{HT}$, i.e., its global structure. All these are potentially interesting topics for future research.

As a particularly interesting topic for future research, we should mention the nontrivial change of the HT symmetry group when additional terms are being added to the action. In Section 4.4, we briefly demonstrated that HT symmetry does change in a nontrivial way, on the example action (52). Nevertheless, the precise properties and the physical interpretation of this change are yet to be studied in full and for a general choice of the action. This topic is the subject of ongoing research.

Finally, we would like to reiterate the differences in two possible approaches to the notion of "the gauge symmetry" of a theory. The overwhelmingly common approach throughout the literature is to factor out the HT group and work only with the ordinary, nontrivial gauge group as the relevant symmetry. Admittedly, this approach does feature a certain level of appeal due to its simplicity and economy, since it does not have to deal with HT symmetry at all. Nevertheless, there are important situations where this is not enough, and one really needs to take into account the *total* gauge symmetry group, which includes HT transformations. As a rule, these situations always involve the gauge symmetry off-shell, either for the purpose of quantization or otherwise. A typical example is the Batalin–Vilkovisky formalism, where one needs to explicitly keep track of HT transformations throughout the whole analysis. Another situation, which was discussed here in more detail, is the question of diffeomorphism symmetry, where HT transformations are required in order to prove that diffeomorphisms are a symmetry of the theory even off-shell. This is especially relevant for building quantum gravity models. Finally, the third scenario would be the discussion of the Coleman–Mandula theorem. One of the main assumptions of the theorem is that the Poincaré group is a subgroup of the full symmetry group of the theory. Given this assumption, and a number of other assumptions, the theorem implies that the full symmetry group must be a direct product of the Poincaré subgroup and the internal symmetry subgroup. In certain cases of theories (such as the $3BF$ action), the full symmetry group is not explicitly expressed as such a direct product, and moreover, it is not obvious that the Poincaré group is a subgroup of the full symmetry group to begin with. Therefore, in order to verify whether the above assumption of the theorem is satisfied, one needs to inspect if the Poincaré group is or is not a subgroup of the full symmetry group. At this point, one may run into a scenario similar to diffeomorphisms: the Poincaré group may fail to be a subgroup of the ordinary gauge group, but still be a subgroup of the total gauge group, once the HT symmetry is taken into account. In this sense, HT symmetry may become relevant for the proper analysis and application of the Coleman–Mandula theorem in certain contexts. This topic is the subject of ongoing research [34].

All of the above arguments suggest that it may be prudent to abandon the common approach of factoring out the HT group and instead adopt the description of the symmetry with the total gauge group, which includes HT transformations on equal footing as the ordinary gauge transformations. In the long run, this may be a conceptually cleaner approach. However, either way, we believe that HT symmetry is relevant for the overall symmetry structure of a theory and that better understanding of its properties can add value to and benefit research.

Author Contributions: Conceptualization, M.V.; investigation, M.Đ., T.R., P.S. and M.V.; writing—original draft preparation, M.Đ., T.R., P.S. and M.V.; writing—review and editing, M.Đ., T.R., P.S. and M.V. All authors have read and agreed to the published version of the manuscript.

Funding: All authors were supported by the Ministry of Science, Technological development and Innovations of the Republic of Serbia. In addition, T.R. and M.V. were supported by the Science Fund of the Republic of Serbia, Grant 7745968, "Quantum Gravity from Higher Gauge Theory 2021"—QGHG-2021. The contents of this publication are the sole responsibility of the authors and can in no way be taken to reflect the views of the Science Fund of the Republic of Serbia.

Data Availability Statement: No new data were created nor analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: The authors would like to thank Igor Prlina for discussions and suggestions when writing this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gomis, J.; Paris, J.; Samuel, S. Antibracket, antifields and gauge-theory quantization. *Phys. Rep.* **1995**, *259*, 1. [CrossRef]
2. Batalin, I.A.; Vilkovisky, G.A. Gauge Algebra and Quantization. *Phys. Lett. B* **1981**, *102*, 27. [CrossRef]
3. Batalin, I.A.; Vilkovisky, G.A. Feynman Rules for Reducible Gauge Theories. *Phys. Lett. B* **1983**, *120*, 166. [CrossRef]
4. Batalin, I.A.; Vilkovisky, G.A. Quantization of Gauge Theories with Linearly Dependent Generators. *Phys. Rev. D* **1983**, *28*, 2567; Erratum in *Phys. Rev. D* **1984**, *30*, 508. [CrossRef]
5. Batalin, I.A.; Vilkovisky, G.A. Closure of the Gauge Algebra, Generalized Lie Algebra Equations and Feynman Rules. *Nucl. Phys.* **1984**, *B234*, 106. [CrossRef]
6. Batalin, I.A.; Vilkovisky, G.A. Existence Theorem for Gauge Algebra. *J. Math. Phys.* **1985**, *26*, 172. [CrossRef]
7. Horowitz, G.T. Exactly Soluble Diffeomorphism Invariant Theories. *Commun. Math. Phys.* **1989**, *125*, 417. [CrossRef]
8. Radenković, T.; Vojinović, M. Gauge symmetry of the 3*BF* theory for a generic semistrict Lie 3-group. *Class. Quant. Grav.* **2022**, *39*, 135009. [CrossRef]
9. Celada, M.; González, D.; Montesinos, M. BF gravity. *Class. Quant. Grav.* **2016**, *33*, 213001. [CrossRef]
10. Girelli, F.; Pfeiffer, H.; Popescu, E.M. Topological Higher Gauge Theory—From BF to BFCG theory. *J. Math. Phys.* **2008**, *49*, 032503. [CrossRef]
11. Martins, J.F.; Miković, A. Lie crossed modules and gauge-invariant actions for 2-BF theories. *Adv. Theor. Math. Phys.* **2011**, *15*, 1059. [CrossRef]
12. Henneaux, M.; Teitelboim, C. *Quantization of Gauge Systems*; Princeton University Press: Princeton, NJ, USA, 1991.
13. Rovelli, C. *Quantum Gravity*; Cambridge University Press: Cambridge, UK, 2004.
14. Rovelli, C.; Vidotto, F. *Covariant Loop Quantum Gravity*; Cambridge University Press: Cambridge, UK, 2014.
15. Baez, J.C.; Huerta, J. An Invitation to Higher Gauge Theory. *Gen. Relativ. Gravit.* **2011**, *43*, 2335. [CrossRef]
16. Miković, A.; Vojinović, M. Poincaré 2-group and quantum gravity. *Class. Quant. Grav.* **2012**, *29*, 165003. [CrossRef]
17. Martins, J.F.; Picken, R. The fundamental Gray 3-groupoid of a smooth manifold and local 3-dimensional holonomy based on a 2-crossed module. *Differ. Geom. Appl. J.* **2011**, *29*, 179. [CrossRef]
18. Wang, W. On 3-gauge transformations, 3-curvatures and Gray-categories. *J. Math. Phys.* **2014**, *55*, 043506. [CrossRef]
19. Saemann, C.; Wolf, M. Six-Dimensional Superconformal Field Theories from Principal 3-Bundles over Twistor Space. *Lett. Math. Phys.* **2014**, *104*, 1147. [CrossRef]
20. Radenković, T.; Vojinović, M. Higher Gauge Theories Based on 3-groups. *J. High Energy Phys.* **2019**, *10*, 222. [CrossRef]
21. Miković, A.; Vojinović, M. Standard Model and 4-groups. *Europhys. Lett.* **2021**, *133*, 61001. [CrossRef]
22. Hidaka, Y.; Nitta, M.; Yokokura, R. Higher-form symmetries and 3-group in axion electrodynamics. *Phys. Lett. B* **2020**, *808*, 135672. [CrossRef]

23. Hidaka, Y.; Nitta, M.; Yokokura, R. Global 3-group symmetry and 't Hooft anomalies in axion electrodynamics. *J. High Energy Phys.* **2021**, *1*, 173. [CrossRef]
24. Song, D.; Wu, K.; Yang, J. 3-form Yang-Mills based on 2-crossed modules. *J. Geom. Phys.* **2022**, *178*, 104537. [CrossRef]
25. Song, D.; Wu, K.; Yang, J. Higher form Yang-Mills as higher BFYM theories. *Eur. Phys. J. C* **2022**, *82*, 1034. [CrossRef]
26. Hidaka, Y.; Nitta, M.; Yokokura, R. Topological axion electrodynamics and 4-group symmetry. *Phys. Lett. B* **2021**, *823*, 136762. [CrossRef]
27. Hidaka, Y.; Nitta, M.; Yokokura, R. Global 4-group symmetry and 't Hooft anomalies in topological axion electrodynamics. *Prog. Theor. Exp. Phys.* **2022**, *2022*, 04A109. [CrossRef]
28. Miković, A.; Vojinović, M. A finiteness bound for the EPRL/FK spin foam model. *Class. Quant. Grav.* **2013**, *30*, 035001. [CrossRef]
29. Baez, J.C. An Introduction to Spin Foam Models of Quantum Gravity and BF Theory. *Lect. Notes Phys.* **2000**, *543*, 25.
30. Baratin, A.; Freidel, L. A 2-categorical state sum model. *J. Math. Phys.* **2015**, *56*, 011705. [CrossRef]
31. Miković, A.; Vojinović, M. Solution to the cosmological constant problem in a Regge quantum gravity model. *Europhys. Lett.* **2015**, *110*, 40008. [CrossRef]
32. Asante, S.K.; Dittrich, B.; Girelli, F.; Riello, A.; Tsimiklis, P. Quantum geometry from higher gauge theory. *Class. Quant. Grav.* **2020**, *37*, 205001. [CrossRef]
33. Radenković, T.; Vojinović, M. Topological invariant of 4-manifolds based on a 3-group. *J. High Energy Phys.* **2022**, *07*, 105. [CrossRef]
34. Đorđević, M.; Vojinović, M. Higher gauge theory and Coleman-Mandula theorem for 3-groups. 2023, *in preparation*.
35. Radenković, T.; Vojinović, M. Hamiltonian Analysis for the Scalar Electrodynamics as 3BF Theory. *Symmetry* **2020**, *12*, 620. [CrossRef]

Operational interpretation of the vacuum and process matrices for identical particles

Ricardo Faleiro[1], Nikola Paunković[1,2], and Marko Vojinović[3]

[1]Instituto de Telecomunicações, Avenida Rovisco Pais 1, 1049-001, Lisboa, Portugal
[2]Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001, Lisboa, Portugal
[3]Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia

This work overviews the single-particle two-way communication protocol recently introduced by del Santo and Dakić (dSD), and analyses it using the process matrix formalism. We give a detailed account of the importance and the operational meaning of the interaction of an agent with the vacuum — in particular its role in the process matrix description. Our analysis shows that the interaction with the vacuum should be treated as an operation, on equal footing with all other interactions. This raises the issue of counting such operations in an operational manner. Motivated by this analysis, we apply the process matrix formalism to capped Fock spaces using the framework of second quantisation, in order to characterise protocols with an indefinite number of identical particles.

1 Introduction

In recent years there have been advances in quantum information theory related to new techniques for discussing quantum circuits and quantum computation. One of those techniques is the recently developed process matrix formalism [1]. This formalism is general enough to describe all known quantum processes, in particular the superposed orders of operations in a quantum circuit. Moreover, its most prominent feature is that it allows for a description of more general situations of indefinite causal orders of spacetime points. A formal example of such a process has been introduced and discussed in [1], leading to the violation of the so-called *causal inequalities*. The latter represent device-independent conditions that need to be satisfied in order for a given process to have a well-defined causal order. It is still an open question whether such a process is physical and can be realised in nature. Also, a lot of attention in the literature has been devoted to the *quantum switch* operation, which has been discussed through both theoretical descriptions [2, 3, 4, 5] and experimental implementations [6, 7, 8].

One of the interesting aspects of the quantum switch is that it gives rise to a superposition of orders of quantum operations. In a recent work [9], the difference between the superposition of orders of quantum operations and the superposition of causal orders in spacetime was discussed in detail, and it was demonstrated that the latter can in principle be realised only in the context of quantum gravity (see also [10, 11, 12]). The detailed analysis of the causal structure of the quantum switch has revealed one important qualitative aspect of the process matrix description — in order to properly account for the causal structure of an arbitrary process, it is *necessary* to introduce the notion of the quantum vacuum as a possible physical state. Otherwise, the naive application of the process matrix formalism may suggest a misleading conclusion that quantum switch implementations in flat spacetime feature genuine superpositions of spacetime causal orders. This demonstrates the importance of the concept of vacuum in quantum information processing. Regarding the general role of the vacuum in quantum circuits and optical experiments, see [13] and [14, 15], respectively, and the references therein.

Ricardo Faleiro: ricardofaleiro@tecnico.ulisboa.pt
Nikola Paunković: npaunkov@math.tecnico.ulisboa.pt
Marko Vojinović: vmarko@ipb.ac.rs

Simultaneously with these developments, another interesting quantum process has been recently proposed [16] by del Santo and Dakić — dSD protocol (see also subsequent theoretical [17, 18] and experimental work [19]). As it turns out, while this process enables Alice and Bob to guess each other's input bits with certainty by exchanging a single particle only once, it cannot be correctly described within the process matrix formalism without the introduction of the interaction between the vacuum and the apparatus as an *operation*. Thus, it represents an additional motivation to introduce the vacuum state into the process matrix formalism, independent of the reasons related to the description of the quantum switch process.

Moreover, while the dSD protocol employs only one photon, it is also relevant for multiphoton processes, which opens the question of the treatment of identical particles within the process matrix formalism. Also, taking into account the presence of the vacuum state, one is steered towards the application of the abstract process matrix formalism to systems with variable number of identical particles, to the second quantisation and ultimately generalisation to quantum field theory (QFT). See also a related work on the causal boxes framework [20].

In this work we give a detailed description and treatment of dSD protocol within the process matrix formalism. We analyse in detail the role of the vacuum in the protocol and the formalism, and its operational interpretation. Specifically, our aim is to discuss the following question:

Is the interaction with the vacuum an operation, or not?

Our analysis of dSD protocol leads to a conclusion that the interaction with the vacuum should be considered an operation. The alternative would mean that one could extract information from the system at the final moment of the protocol without performing an operation at all. Since the same physical situation should always be described in the same way, we conclude that the interaction with the vacuum should be treated as an operation, and thus as a resource, in all quantum information protocols. This includes the optical implementation of the quantum switch protocol, leading one to infer that it features *four, rather than just two* operations, as was claimed in a number of papers [3, 4, 5, 6, 7, 8]. In addition, we make use of the dSD protocol as an illuminative example to apply the process matrix formalism to multipartite systems of identical particles.

The paper is organised as follows. In Section 2 we give a short overview of the process matrix formalism and the dSD protocol. Section 3 is devoted to the process matrix formalism description of dSD protocol, and to the discussion of the *operational* role and importance of the vacuum state for its description. In particular, in Subsection 3.4 we present the argument for our main conclusion, namely that the interaction with the vacuum should be considered an operation. In Section 4 we provide the basic rules for the application of the process matrix formalism to identical multiparticle systems. Section 5 is devoted to the summary of our results, discussion and prospects for future research. The Appendix contains various technical details of the calculations.

2 State of the art

In this section, we present an overview of the relevant background results. First, we give a short introduction to the process matrix formalism, and then we present the dSD protocol. This overview is not intended to be complete or self-contained, but merely of informative type. The reader should consult the literature for more details.

2.1 The process matrix formalism

The process matrix formalism is based on an idea of a set of laboratories, interacting with the outside world by exchanging quantum systems. Each laboratory is assumed to be spatially local in the sense that one can consider its size negligible for the problem under discussion. Inside the laboratory, it is assumed that the ordinary laws of quantum theory hold. The laboratory interacts with the outside world by receiving an *input* quantum system and by sending an *output* quantum system. Inside the laboratory, the input and output quantum systems are being manipulated using the notion of an *instrument*, denoting the most general operation one can perform over quantum systems. Each interaction is also assumed to be localised in time, such that each operation of a given laboratory has a separate spacetime point assigned to it (see Subsection 3.4 for a discussion of time delocalised laboratories and operations [21]). Thus, we introduce the notion of a *gate*, which represents the action of an instrument at a given spacetime point (see Section 2 of [9]); for simplicity, both the gate and its corresponding spacetime point will be denoted by the same symbol, G. By G_I and G_O we denote the Hilbert spaces of the input and the output quantum systems, respectively. These Hilbert spaces are assumed to be finite-dimensional or trivial. The action of the instrument is described by an operator, $\mathcal{M}^G_{x,a} : G_I \otimes G_I^* \to G_O \otimes G_O^*$, which may depend on some classical input information a available to the gate G, and some readouts x of eventual measurement results that may take place in G. Thus, the instrument maps a generic mixed input state ρ_I into the output state $\rho_O = \mathcal{M}^G_{x,a}(\rho_I)$.

Given such a setup, one defines a *process*, denoted $\mathcal{W}$, as a functional over the instruments of all gates, as

$$p(x, y, \dots | a, b, \dots) = \mathcal{W}(\mathcal{M}^{G^{(1)}}_{x,a} \otimes \mathcal{M}^{G^{(2)}}_{y,b} \otimes \dots),$$

where $p(x, y, \dots | a, b, \dots)$ represents the probability of obtaining measurement results $x, y, \dots$, given the inputs $a, b, \dots$. In order for the right-hand side to be interpreted as a probability distribution, the process $\mathcal{W}$ must satisfy three basic axioms,

$$\begin{aligned} &\mathcal{W} \geq 0\,, \\ \operatorname{Tr} &\mathcal{W} = \textstyle\prod_i \dim G_O^{(i)}\,, \\ &\mathcal{W} = \mathcal{P}_G(\mathcal{W})\,, \end{aligned} \tag{1}$$

where $\mathcal{P}_G$ is a certain projector onto a subspace of $\bigotimes_i \left(G_I^{(i)} \otimes G_O^{(i)} \right)$ which, together with the second requirement, ensures the normalisation of the probability distribution (see [3] for details).

In order to have a computationally manageable formalism, one often employs the Choi-Jamiołkowski (CJ) map over the instrument operations, such that a given operation $\mathcal{M}^G_{x,a}$ is being represented by a matrix,

$$\begin{aligned} M^G_{x,a} = &\left[\left(\mathcal{I} \otimes \mathcal{M}^G_{x,a} \right) (|\mathbb{1}\rangle\rangle\langle\langle\mathbb{1}|) \right]^T \\ &\in (G_I \otimes G_O) \otimes (G_I \otimes G_O)^*, \end{aligned} \tag{2}$$

where

$$|\mathbb{1}\rangle\rangle \equiv \sum_i |i\rangle \otimes |i\rangle \in G_I \otimes G_I \tag{3}$$

is the so-called *transport vector*, representing the non-normalised maximally entangled state, and $\mathcal{I}$ is the identity operator. Then, one can describe the process $\mathcal{W}$ using the *process matrix* W to write

$$\begin{aligned} &p(x, y, \dots | a, b, \dots) = \\ &\qquad \operatorname{Tr} \left[(M^{G_1}_{x,a} \otimes M^{G_2}_{y,b} \otimes \dots) W \right]. \end{aligned} \tag{4}$$

Finally, if an instrument $\mathcal{M}^G_{x,a}$ is linear, one can also use a corresponding "vector" notation (see Appendix A.1 in [3]),

$$|(\mathcal{M}^G_{x,a})^*\rangle\rangle \equiv \left[\mathcal{I} \otimes (\mathcal{M}^G_{x,a})^* \right] |\mathbb{1}\rangle\rangle \in G_I \otimes G_O\,, \tag{5}$$

so that

$$M^G_{x,a} = |(\mathcal{M}^G_{x,a})^*\rangle\rangle\langle\langle(\mathcal{M}^G_{x,a})^*|\,. \tag{6}$$

In cases where all instruments are linear, and in addition the process matrix W is a one-dimensional projector, one can introduce the the corresponding *process vector* $|W\rangle\rangle$, such that $W = |W\rangle\rangle\langle\langle W|$, and rewrite (4) in the form:

$$\begin{aligned} &p(x, y, \dots | a, b, \dots) = \\ &\left\| \left(\langle\langle \mathcal{M}^{G^{(1)}*}_{x,a} | \otimes \langle\langle \mathcal{M}^{G^{(2)}*}_{y,b} | \otimes \dots \right) |W\rangle\rangle \right\|^2. \end{aligned} \tag{7}$$

2.2 The dSD protocol

In a recent paper [16], del Santo and Dakić have introduced a protocol which allows two agents to guess each other's input bits with certainty by exchanging a single particle only once. The protocol goes as follows. Initially, a single particle is prepared in a superposition state of being sent to Alice and being sent to Bob. Upon receiving the particle, both Alice and Bob perform unitary operations on it, encoding their bits of information, a and b, respectively, about the outcomes of their coin tosses. They do this by changing the local phase of the particle by $(-1)^a$ and $(-1)^b$. The particle is subsequently forwarded to a beam splitter, and after that again to Alice and Bob, who now measure the presence or absence of the particle.

This way, the state of the particle stays in *coherent* superposition of different paths in a Mach-Zehnder interferometer. The interference of its paths gives rise to *deterministic* outcome that depends on the relative phase $e^{i\phi} = (-1)^{a \oplus b}$ between the two branches: in case $\phi = 0$, the particle will end up in Alice's laboratory, while otherwise it will end in Bob's. Thus, knowing their own inputs and the outcomes of their local measurements, both agents can determine each other's inputs, allowing for two-way communication using only one particle. This is clearly impossible in classical physics, demonstrating yet another example of the advantage of quantum over classical strategies.

The crucial aspect of the protocol lies in the fact that the *absence* of the particle represents a useful piece of information for an agent. This gives rise to the notion of the *vacuum state* as a carrier of information, playing the central role in the protocol. Thus, in order to describe the protocol using the process matrix formalism, one has to incorporate the notion of the vacuum in the formalism itself. We show this in detail in the next section.

It is interesting to note that the crucial role of the vacuum plays an important part not only in the dSD protocol, but also in a completely different setup that has been discussed a lot in recent literature, namely the *quantum switch* [2]. As analysed in detail in [9], if one takes care to distinguish the two temporal positions of a given laboratory and introduces the notion of a vacuum explicitly, one can demonstrate that the optical implementations of the quantum switch in flat spacetime do not feature any superposition of causal orders induced by the spacetime metric. Instead, it was argued that superpositions of spacetime causal orders can be present only within the context of a theory of quantum gravity. As we shall see below, the notion of the interaction with the vacuum will prove essential to the process matrix description of the dSD protocol as well.

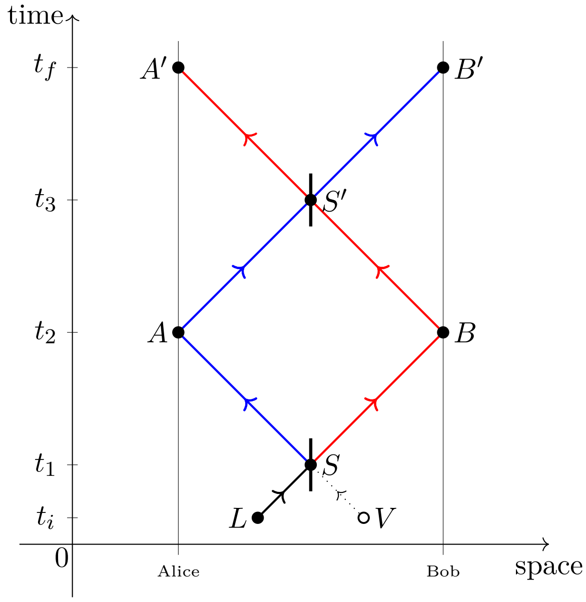


Figure 1: The complete spacetime diagram of the process corresponding to the dSD protocol.

3 Process matrix description of the dSD protocol

3.1 The spacetime diagram

We begin by drawing the spacetime diagram of the process corresponding to the dSD protocol (see Figure 1).

At the initial time t_i the laser L creates a photon and shoots it towards the beam splitter S, which at time t_1 performs the Hadamard operation and entangles it with the incoming vacuum state (described by the dotted arrow from the grey gate V). The entangled state of the photon and the vacuum continues on towards Alice's and Bob's gates A and B, respectively. At time t_2, Alice and Bob generate their random bits a and b, and encode them into the phase of the incoming photon-vacuum system. The system then proceeds to the beam splitter S' which again performs the Hadamard operation at time t_3. The photon-vacuum system then proceeds to the gates A' and B', where it is measured at time t_f by Alice and Bob, respectively. Note that the spatial distance Δl between Alice and Bob is precisely equal to the time distance between the generation of the random bits and the final measurements,

$$\Delta l = c(t_f - t_2)\,,$$

so that a single photon has time to traverse the space between Alice and Bob *only once*. Also, note that the gate V, which generates the vacuum state, corresponds to a "trivial instrument", since the vacuum does not require any physical device to be generated. Nevertheless, the vacuum is still a legitimate physical state of the EM field, so the appropriate gate V has to be formally introduced and accounted for in the process matrix formalism calculations.

3.2 Formulation of the process matrix

Based on the spacetime diagram, we formulate the process matrix description as follows. All spacetime points, where interaction between the EM field and some apparatus may happen, are assigned a gate and an operation which describes the interaction. Each gate has an input and output Hilbert space, as follows:

$$
\begin{array}{llll}
L: & \mathbb{C} \to L_O\,, & A: & A_I \to A_O\,,\\
V: & \mathbb{C} \to V_O\,, & B: & B_I \to B_O\,,\\[1em]
S: & S_I \to S_O\,, & A': & A'_I \to \mathbb{C}\,,\\
S': & S'_I \to S'_O\,, & B': & B'_I \to \mathbb{C}\,.
\end{array}
$$

The initial gates, L and V, have trivial input spaces and nontrivial output spaces. The final gates, A' and B', have trivial output spaces and nontrivial input spaces. The gates A and B have nontrivial input and output spaces. Each nontrivial space is isomorphic to $\mathcal{H}_1 \oplus \mathcal{H}_0 \subset \mathcal{F}$, where $\mathcal{H}_0$ and $\mathcal{H}_1$ are the vacuum and single-excitation subspaces of the Fock space $\mathcal{F}$ in perturbative QED. Namely, by design of the dSD protocol, Alice and Bob may exchange at most one photon, which means that multiparticle Hilbert subspaces of the Fock space can be omitted. Moreover, the resulting probability distribution of the experiment outcomes does not in principle depend on the frequency or the polarisation of the photon in use, so we can approximate the single-excitation space as a one-dimensional Hilbert space. Given that the vacuum Hilbert space $\mathcal{H}_0$ is by definition one-dimensional, we can write

$$
\mathcal{H}_0 = \text{span}\{|0\rangle\} \equiv \mathbb{C}\,, \qquad \mathcal{H}_1 \approx \text{span}\{|1\rangle\} \equiv \mathbb{C}\,,
$$

so that $\mathcal{H}_0 \oplus \mathcal{H}_1 \equiv \mathbb{C} \oplus \mathbb{C}$. Here, $|0\rangle$ and $|1\rangle$ denote the states of the vacuum and the photon in the occupation number basis of the Fock space. Therefore, we have

$$
L_O \cong V_O \cong A_I \cong A_O \cong B_I \cong B_O \cong A'_I \cong B'_I \cong \mathbb{C} \oplus \mathbb{C}\,.
$$

Finally, the input and output spaces of beam splitters S and S' are "doubled", since a beam splitter operates over two inputs to produce two outputs. In particular,

$$
\begin{array}{lll}
S_I = S_I^L \otimes S_I^V\,, & & S'_I = S_I'^A \otimes S_I'^B\,,\\
S_O = S_O^A \otimes S_O^B\,, & & S'_O = S_O'^A \otimes S_O'^B\,,
\end{array}
$$

where again

$$
S_I^L \cong S_I^V \cong S_O^A \cong S_O^B \cong S_I'^A \cong S_I'^B \cong S_O'^A \cong S_O'^B \cong \mathbb{C} \oplus \mathbb{C}\,.
$$

With all relevant Hilbert spaces defined, we formulate the action of each gate, using the CJ map in the form (5). The gates L and V simply generate the photon and the vacuum,

$$
|L^*\rangle\!\rangle^{L_O} = |1\rangle^{L_O}\,, \qquad |V^*\rangle\!\rangle^{V_O} = |0\rangle^{V_O}\,, \tag{8}
$$

where $*$ is the complex conjugation. The action of the beam splitters is

$$
|S^*\rangle\!\rangle^{S_I S_O} = \left[I^{S_I S_I} \otimes (H^*)^{S_O S_I} \right] |\mathbb{1}\rangle\!\rangle^{S_I S_I}\,, \tag{9}
$$

and

$$
|S'^*\rangle\!\rangle^{S'_I S'_O} = \left[I^{S'_I S'_I} \otimes (H^*)^{S'_O S'_I} \right] |\mathbb{1}\rangle\!\rangle^{S'_I S'_I}\,, \tag{10}
$$

where the Hadamard operator for S is defined as

$$
\begin{aligned}
H^{S_O S_I} = {} & \frac{1}{\sqrt{2}} \left(|1\rangle^{S_O^A} |0\rangle^{S_O^B} + |0\rangle^{S_O^A} |1\rangle^{S_O^B} \right) \langle 1|^{S_I^L} \langle 0|^{S_I^V} \\
& + \frac{1}{\sqrt{2}} \left(|1\rangle^{S_O^A} |0\rangle^{S_O^B} - |0\rangle^{S_O^A} |1\rangle^{S_O^B} \right) \langle 0|^{S_I^L} \langle 1|^{S_I^V}\,,
\end{aligned}
$$

and analogously for $H^{S'_O S'_I}$. The unit operator is denoted as I. Next, in the gates A and B, Alice and Bob generate their random bits a and b, and encode them into the phase of the photon. The corresponding actions are defined as

$$
|A^*\rangle\!\rangle^{A_I A_O} = \left[I^{A_I A_I} \otimes (A^*)^{A_O A_I} \right] |\mathbb{1}\rangle\!\rangle^{A_I A_I}\,, \tag{11}
$$

and

$$
|B^*\rangle\!\rangle^{B_I B_O} = \left[I^{B_I B_I} \otimes (B^*)^{B_O B_I} \right] |\mathbb{1}\rangle\!\rangle^{B_I B_I}, \tag{12}
$$

where

$$
A^{A_O A_I} = (-1)^a |1\rangle^{A_O} \langle 1|^{A_I} \oplus |0\rangle^{A_O} \langle 0|^{A_I}\,,
$$

and

$$
B^{B_O B_I} = (-1)^b |1\rangle^{B_O} \langle 1|^{B_I} \oplus |0\rangle^{B_O} \langle 0|^{B_I}\,.
$$

Finally, the gates A' and B' describe Alice's and Bob's measurement of the incoming state in the occupation number basis,

$$
|A'^*\rangle\!\rangle^{A'_I} = |a'\rangle\,, \qquad |B'^*\rangle\!\rangle^{B'_I} = |b'\rangle\,, \tag{13}
$$

where their respective measurement outcomes a' and b' take values from the set $\{0, 1\}$, depending on whether the vacuum or the photon has been measured, respectively.

After specifying the actions of the gates, the last step is the construction of the process vector

$|W_{dSD}\rangle\rangle$ itself. The dSD protocol assumes that the state of the photon remains unchanged during its travel between the gates. Therefore, the process vector will be a tensor product of transport vectors (3), one for each line connecting two gates in the spacetime diagram. The input and output spaces of the gates connected by the line determine the spaces of the corresponding transport vector. Thus, the total process vector is:

$$\begin{aligned}&|W_{dSD}\rangle\rangle = \\ &\quad |\mathbb{1}\rangle\rangle^{L_O S_I^L} |\mathbb{1}\rangle\rangle^{V_O S_I^V} |\mathbb{1}\rangle\rangle^{S_O^A A_I} |\mathbb{1}\rangle\rangle^{S_O^B B_I} \\ &\quad |\mathbb{1}\rangle\rangle^{A_O S_I'^A} |\mathbb{1}\rangle\rangle^{B_O S_I'^B} |\mathbb{1}\rangle\rangle^{S_O'^A A_I'} |\mathbb{1}\rangle\rangle^{S_O'^B B_I'} .\end{aligned} \tag{14}$$

3.3 Evaluation of the probability distribution

Now that the process vector and the operations of all gates have been specified in detail, we can evaluate the probability distribution

$$p(a', b'|a, b) = \|\mathcal{M}\|^2 , \tag{15}$$

where the probability amplitude $\mathcal{M}$ is obtained by taking the scalar product of $|W_{dSD}\rangle\rangle$ with the tensor product of all gates, see (7). It is most instructive to perform the computation iteratively, taking the partial scalar product of $|W_{dSD}\rangle\rangle$ with each gate, one by one. The explicit calculation of each step is based on two lemmas from Appendix A.

We begin by taking the partial scalar product of (14) and the preparation gates (8). Using Lemma 1 from Appendix A, we obtain:

$$\begin{aligned}&\left(\langle\langle L^*|^{L_O} \otimes \langle\langle V^*|^{V_O}\right) |W_{dSD}\rangle\rangle = \\ &\qquad |1\rangle^{S_I^L} |0\rangle^{S_I^V} |\mathbb{1}\rangle\rangle^{S_O^A A_I} |\mathbb{1}\rangle\rangle^{S_O^B B_I} \\ &\qquad |\mathbb{1}\rangle\rangle^{A_O S_I'^A} |\mathbb{1}\rangle\rangle^{B_O S_I'^B} |\mathbb{1}\rangle\rangle^{S_O'^A A_I'} |\mathbb{1}\rangle\rangle^{S_O'^B B_I'} .\end{aligned}$$

Next we take the partial scalar product with the beam splitter S gate operation (9). Using Lemma 2 from Appendix A, we obtain:

$$\begin{aligned}&\left(\langle\langle S^*|^{S_I S_O} \otimes \langle\langle L^*|^{L_O} \otimes \langle\langle V^*|^{V_O}\right) |W_{dSD}\rangle\rangle = \\ &\qquad \frac{1}{\sqrt{2}} \left(|1\rangle^{A_I} |0\rangle^{B_I} + |0\rangle^{A_I} |1\rangle^{B_I}\right) \\ &\qquad |\mathbb{1}\rangle\rangle^{A_O S_I'^A} |\mathbb{1}\rangle\rangle^{B_O S_I'^B} |\mathbb{1}\rangle\rangle^{S_O'^A A_I'} |\mathbb{1}\rangle\rangle^{S_O'^B B_I'} .\end{aligned}$$

Now we apply the Alice's gate operation (11) to obtain:

$$\begin{aligned}&\Big(\langle\langle A^*|^{A_I A_O} \otimes \langle\langle S^*|^{S_I S_O} \otimes \\ &\quad \langle\langle L^*|^{L_O} \otimes \langle\langle V^*|^{V_O}\Big) |W_{dSD}\rangle\rangle = \\ &\qquad \frac{1}{\sqrt{2}} \left((-1)^a |1\rangle^{S_I'^A} |0\rangle^{B_I} + |0\rangle^{S_I'^A} |1\rangle^{B_I}\right) \\ &\qquad |\mathbb{1}\rangle\rangle^{B_O S_I'^B} |\mathbb{1}\rangle\rangle^{S_O'^A A_I'} |\mathbb{1}\rangle\rangle^{S_O'^B B_I'} .\end{aligned}$$

Similarly, applying Bob's gate (12) we get:

$$\begin{aligned}&\Big(\langle\langle B^*|^{B_I B_O} \otimes \langle\langle A^*|^{A_I A_O} \otimes \langle\langle S^*|^{S_I S_O} \otimes \\ &\quad \langle\langle L^*|^{L_O} \otimes \langle\langle V^*|^{V_O}\Big) |W_{dSD}\rangle\rangle = \\ &\quad \frac{1}{\sqrt{2}} \left((-1)^a |1\rangle^{S_I'^A} |0\rangle^{S_I'^B} + (-1)^b |0\rangle^{S_I'^A} |1\rangle^{S_I'^B}\right) \\ &\quad |\mathbb{1}\rangle\rangle^{S_O'^A A_I'} |\mathbb{1}\rangle\rangle^{S_O'^B B_I'} .\end{aligned}$$

The next step is the application of the second beam splitter gate (10). After a little bit of algebra, the result is:

$$\begin{aligned}&\left(\langle\langle S'^*|^{S_I' S_O'} \otimes \langle\langle B^*|^{B_I B_O} \otimes \langle\langle A^*|^{A_I A_O} \otimes \langle\langle S^*|^{S_I S_O} \otimes \langle\langle L^*|^{L_O} \otimes \langle\langle V^*|^{V_O}\right) |W_{dSD}\rangle\rangle = \\ &\qquad \frac{(-1)^a + (-1)^b}{2} |1\rangle^{A_I'} |0\rangle^{B_I'} + \frac{(-1)^a - (-1)^b}{2} |0\rangle^{A_I'} |1\rangle^{B_I'} .\end{aligned}$$

Finally, applying the measurement gates (13), we obtain the complete probability amplitude,

$$\begin{aligned}\mathcal{M} \;=\; & \frac{(-1)^a + (-1)^b}{2} \delta_{a'1} \delta_{b'0} \\ & + \frac{(-1)^a - (-1)^b}{2} \delta_{a'0} \delta_{b'1} ,\end{aligned}$$

and substituting this into (15), we obtain the desired probability distribution of the dSD process:

$$\begin{aligned}p(a', b'|a, b) \;=\; & \frac{1 + (-1)^{a+b}}{2} \delta_{a'1} \delta_{b'0} \\ & + \frac{1 - (-1)^{a+b}}{2} \delta_{a'0} \delta_{b'1} .\end{aligned}$$

From the probability distribution we can now conclude that there are two distinct possibilities:

either Alice detects the photon and Bob does not, $a' = 1, b' = 0$, or vice versa, $a' = 0, b' = 1$. In the first case, because total probability must be equal to one, we have

$$\frac{1+(-1)^{a+b}}{2} = 1\,, \qquad \frac{1-(-1)^{a+b}}{2} = 0\,.$$

The only solution to these equations is $a = b$, which means that Alice and Bob have initially generated equal bits. Since both know the probability distribution and their own bit, they both know each other's bit as well, with certainty. In the second case, when Bob detects the photon, we instead have

$$\frac{1+(-1)^{a+b}}{2} = 0\,, \qquad \frac{1-(-1)^{a+b}}{2} = 1\,,$$

and the only solution is $a \neq b$, meaning that Alice and Bob have initially generated opposite bits. Again, both parties know the probability distribution and their own bit, and therefore each other's bit as well, with certainty.

In order to formalise this result, one can also introduce the parity $\pi \equiv a \oplus b$ and rewrite the probability distribution in the form

$$\begin{aligned} p(a', b'|\pi) \;\; = \;\; & \frac{1+(-1)^{\pi}}{2}\delta_{a'1}\delta_{b'0} \\ & + \frac{1-(-1)^{\pi}}{2}\delta_{a'0}\delta_{b'1}\,. \end{aligned} \tag{16}$$

Thus, if Alice detects the photon, then π is even, while if Bob detects the photon, π must be odd. In both cases, they can "guess" each other's bits with certainty by calculating

$$x = \pi \oplus a\,, \qquad y = \pi \oplus b\,,$$

where x is Alice's prediction of the value of Bob's bit, and y is Bob's prediction of the value of Alice's bit. Therefore, the probability of guessing each other's input bit is

$$p_{\text{success}} \equiv p(x = b \wedge y = a) = 1\,. \tag{17}$$

3.4 Analysis of the process matrix description — operational interpretation of the vacuum

After we have given the detailed process matrix description of the dSD protocol and derived the result (17), we analyse in more detail the role of the vacuum in the formalism, giving its operational interpretation.

In order to clarify the exposition, let us give an overview of the argument, as follows:

- In the next paragraph below, we analyse the role of the vacuum in the dSD protocol, and conclude that the interaction with the vacuum should be regarded as an operation, on the same footing with all other interactions.
- In the following four paragraphs, we discuss the optical implementation of the quantum switch protocol, which also features interactions between the agents and the vacuum. Since the same physical situation should always be described in the same way, we conclude that the interaction with the vacuum should be treated as an operation in this protocol as well. Thus, the protocol features a total of four, rather than two, operations.
- Finally, in the remaining three paragraphs, we discuss the alternative point of view, namely that the interaction with the vacuum is not regarded as an operation. This is the case in the method for counting operations proposed in [22]. We conclude that it would then mean that in the dDS protocol an agent could extract information from the system at t_f without performing an operation at all.

In the dSD protocol four operations (gates), A, A', B and B' are performed (see Figure 1). Note that for each choice of input bits a and b one of the two operations performed, A' and B', is of a special form: it represents the absence of the particle. This gives rise to an *operational interpretation* of the vacuum state as a carrier of information, playing the central role in the protocol — the very interaction between the apparatus and the vacuum (the absence of a particle) plays *exactly* the same role in this protocol as any other operation, i.e., not detecting a particle ("seeing the vacuum") is an operation on its own. From the mathematical point of view, supported by the structure of the process vector (14) that explicitly features the vacuum state, it is perfectly natural to consider the interaction between the apparatus and the vacuum state on equal footing with the interaction between the apparatus and the field excitation (i.e., the particle). Both interactions equally represent operations. Therefore, one should regard the interaction with the vacuum as a resource, in the same way as the interaction with the particle.

Let us now consider the optical quantum switch, a similar protocol in which the notion of

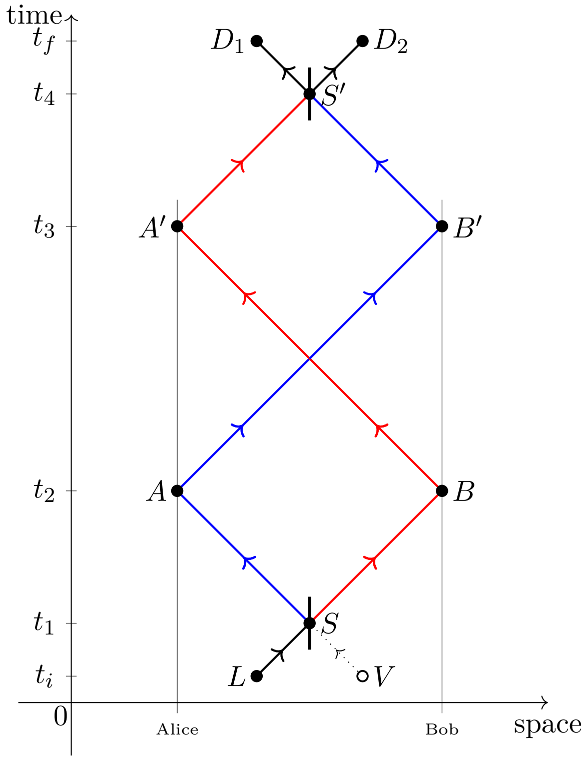


Figure 2: The complete spacetime diagram of the process corresponding to the optical quantum switch. Upon receiving the photon, Alice rotates its polarisation by the unitary U. Analogously, Bob performs rotation V on the photon entering his lab.

the vacuum also plays a role. Current optical implementations of the quantum switch feature four spacetime points, the same as the dSD protocol [9, 10, 11, 12], thus having the similar type of the spacetime schematic description, see Figure 2. However, by introducing the notion of time delocalised operations it was argued that the optical switch implements only two operations, U in spacetime points A or A', and V in spacetime points B or B' [21]. Nevertheless, the optical switch features the same apparatus-vacuum interaction as the one from the dSD protocol: whenever the particle is in, say, the blue branch, and the operations U and V are applied at spacetime points A and B', respectively, Alice's and Bob's labs experience the interaction with the vacuum at spacetime points B and A' (and analogously for the red branch). Therefore, the treatment of the vacuum in the optical quantum switch is mutually incoherent with the treatment of the vacuum in the dSD protocol.

Our analysis can thus serve as motivation for a search towards a more coherent treatment of the vacuum within the operational approach, since the same physical situation — interaction between the apparatus and the vacuum — is currently treated differently in the descriptions of the two protocols.

One might consider the following possible chain of inference. From the examples of both the quantum switch and the dSD protocol, we have that unitary operations (be it "genuine rotations" U and V, as well as phase flips $\pm I$) are considered to be operations. From the example of the dSD protocol, we see that the interaction with the vacuum is an operation as well. Further, in reference [21] it was argued that the optical switch features two "time-delocalised operations", U and V. Thus, by the same token, it follows that within this operational approach the optical switch should feature two additional "time-delocalised operations": interactions with the vacuum, one performed by Alice, and the other by Bob (see Appendix B). Therefore, the protocol features a total of four, rather than two, operations. Note that this is a possible treatment of the vacuum, which still features superposition of orders of operations U and V in the optical switch.

It is obvious that the interaction with the vacuum plays a prominent role in achieving the goal of the dSD protocol — communication between Alice and Bob. But interactions with the vacuum are also crucial in the optical switch. Indeed, without those operations, it would be *impossible* to achieve superposition of orders of operations U and V in flat spacetime with fixed causal order of spacetime points [9].

In [22] the so-called "flag" systems were introduced to count the number of operations performed in a lab without destroying the superposition, which count only one operation per each lab of the optical switch. Note though that using this method, which effectively counts the number of times a particle enters the lab, one would count three rather than four operations in the dSD protocol. This means that either the method is not appropriate, or in fact the dSD protocol features three, instead of four operations. In the case of the former, it would be useful to introduce a formal operational definition of a general method of counting operations, given that the above "flag" method cannot count interactions with the vacuum. In the case of the latter, it would mean that one could extract the information from the system at t_f without performing an operation at all.

Indeed, if the interaction between the vacuum and the apparatus would not be considered an operation, an issue with formulating the process vector for the dSD protocol would arise. The one we formulated in (14) contains input and output Hilbert spaces associated with the interaction between the vacuum and the detectors. It is not possible to formulate a process matrix for the dSD protocol that would feature three operations, without the mentioned interaction with the vacuum. Namely, depending on the choice of input bits a and b, the photon will end up either in Alice's or Bob's lab, rendering it impossible to know in advance which of the two agents is supposed to perform the final operation. Thus, it is not possible to formulate a process matrix which features only one operation at the final moment t_f. Note that the process matrices themselves were introduced as the main tool for describing quantum processes in the operational approach. In other words, the impossibility of formulating the main operational tool for the dSD protocol without introducing the interaction with the vacuum as an operation, suggests that the latter should be considered as an operation in that protocol.

Note that, if the dSD and the optical switch protocols featured incoherent mixtures of the two possible paths instead of coherent superpositions, then one could formulate the corresponding process matrices without treating the interaction with the vacuum as an operation, indeed without even mentioning the vacuum at all. These would be purely classical processes, which would not feature any interference effects. In general, ommiting the vacuum is a natural point of view in classical physics. However, if one wants to describe quantum physics, the notions of the vacuum and its interaction with the apparatus are unavoidable.

4 Identical particles

The above analysis shows that the vacuum state plays a physically relevant role in transmitting information, and cannot be ignored. From the point of view of QFT this is a perfectly natural state of affairs, but from the point of view of quantum mechanics (QM) it is not, since the notion of vacuum as a physical state does not exist in QM a priori, and needs to be explicitly introduced by hand. Moreover, in QFT one can naturally study systems of indefinite number of identical particles. Therefore, as a first step towards the generalization of the process matrix formalism to QFT, we apply the existing abstract process matrix formalism to the representation of the second quantization.

In this section, we give basic elements of the process matrix formalism, when applied to systems of identical particles. In order to avoid working with (anti-)symmetrised vectors of multi-particle states that contain non-physical entanglement whenever two or more identical particles are fully distinguishable (say, one photon is in Alice's, and another in Bob's lab), we will use the representation of the second quantisation in which the effects of particle statistics are governed by the creation and annihilation (anti-)commutation rules. First, we need to move from the single-particle Hilbert spaces associated to the gates and the process matrix to the corresponding capped Fock spaces.

To each gate G, we assign the input/output Fock spaces, $\mathcal{G}_{I/O}$, given in terms of the vacuum state $|0\rangle$ and the single-particle Hilbert spaces $G_{I/O}$. The single-particle input Hilbert space is given as

$$G_I = \text{span}\{|i\rangle = a_i^\dagger|0\rangle \mid i = 1, 2, \dots d_I\},$$

such that its creation and annihilation operators satisfy the standard (anti-)commutation relations,

$$[a_i^\dagger, a_j^\dagger]_\pm = [a_i, a_j]_\pm = 0\,, \quad [a_i, a_j^\dagger]_\pm = \delta_{ij}\,, \tag{18}$$

where $[\_\,,\_]_+$ stands for anti-commutator, and $[\_\,,\_]_-$ for commutator. The overall bosonic input Fock space is then

$$\mathcal{G}_I = \bigoplus_{\ell=0}^{\infty} G_I(\ell), \tag{19}$$

where $G_I(0) = \text{span}\{|0\rangle\}$ is the zero-particle, $G_I(1) = G_I$ the single-particle, and

$$G_I(\ell) = \{[(a_1^\dagger)^{s_1}...(a_{d_I}^\dagger)^{s_{d_I}}]|0\rangle \mid s_1 + ... + s_{d_I} = \ell\}$$

are the ℓ-particle orthogonal subspaces of the input Fock space. For fermions, each $s_i \in \{0, 1\}$, and the orthogonal sum in Equation (19) goes until d_I, instead of ∞. For a given gate, the output Fock space $\mathcal{G}_O$ is defined analogously, and we denote its creation and annihilation operators as $\tilde{a}_i^\dagger$

and $\tilde{a}_i$, respectively, in order to distinguish them from the corresponding operators in $\mathcal{G}_I$.

Our formalism is constructed for quantum circuits which consist of finite number of gates. This means that we work in the approximation of a finite number of spacetime points, as opposed to the standard QFT where one works with an uncountably infinitely many spacetime points. Thus, given the algebra (18) for the creation and annihilation operators at a single gate, the full algebra across all gates is normalised to a Kronecker delta, instead of the standard Dirac delta function. Moreover, the operators in (18) are operators in coordinate space, as opposed to the momentum space operators which are standard in QFT, since they create and annihilate modes at a given gate (i.e., a given spacetime point), instead of modes with a given momentum. Taking into account our assumption of finite number of gates, the single-particle Hilbert spaces $G_{I/O}$ are finite-dimensional, i.e., $d_{I/O} \in \mathbb{N}$. Since the gates are distinguishable, the modes assigned to different gates *always* (anti-)commute.

We restrict ourselves to the Minkowski spacetime, so that the global Poincaré symmetry implies that the vacuum state $|0\rangle$ is identical across different gates, as well as between input and output Fock spaces for a given gate. In this sense, each gate is assumed to be stationary in some inertial reference frame, since the Fock spaces of non-inertial gates would be subject to the Unruh effect. We leave the discussion of non-inertial gates and spacetimes with more general geometries for future work.

Once the Fock spaces have been defined, we pass on to the process matrix description of gate operations. Since a process matrix has to satisfy the normalisation rule (1), the corresponding input and output spaces have to be finite-dimensional. To that end, we restrict ourselves to capped Fock spaces, which contain only a finite number of elements in the sum (19), denoted $N \in \mathbb{N}$. Together with the fact that $d_{I/O}$ is finite, it follows that the capped Fock spaces are finite-dimensional. A gate operation is represented via a CJ isomorphism of the corresponding operator between the input and the output capped Fock spaces, defined in equation (2),

$$M = \Big[(\mathcal{I} \otimes \mathcal{M}) \, (|\mathbb{1}\rangle\!\rangle \langle\!\langle \mathbb{1}|) \Big]^T , \qquad (20)$$

where the transport vector

$$|\mathbb{1}\rangle\!\rangle = \sum_{k=0}^{N} |\mathbb{1}_k\rangle\!\rangle \,, \qquad (21)$$

is given in terms of k-transport vectors defined as

$$|\mathbb{1}_k\rangle\!\rangle = \sum \Big[\prod_{i=1}^{d} \frac{(a_i^\dagger)^{s_i}}{\sqrt{s_i!}} \Big] \otimes \Big[\prod_{i=1}^{d} \frac{(a_i^\dagger)^{s_i}}{\sqrt{s_i!}} \Big] |0\rangle \,, \quad (22)$$

where the sum is taken over all s_i satisfying the constraint $s_1 + ... + s_d = k$.

One special case of the general formula (20) is the case where gates destroy all coherence between k-particle sectors, for example by measuring the number of particles,

$$M = \sum_{k=0}^{N} \Big[(\mathcal{I} \otimes \mathcal{M}_k) \, (|\mathbb{1}_k\rangle\!\rangle \langle\!\langle \mathbb{1}_k|) \Big]^T , \qquad (23)$$

where $\mathcal{M}_k$ represents the k-particle operator for the gate. The above gate represents a classical mixture of operations on each k-particle sector, as opposed to coherent superpositions of them.

Another special case of (20), which does preserve the coherence between k-particle sectors, is represented by linear operations. For a linear gate operation, one can analogously use the "vector" formalism, and the generalisation of the CJ vector (5). With a slight abuse of notation, using $\mathcal{M}$ to denote the operator instead of its superoperator, we can now write

$$\begin{aligned} |\mathcal{M}^*\rangle\!\rangle &= \Big[\mathcal{I} \otimes \mathcal{M}^*\Big] |\mathbb{1}\rangle\!\rangle \\ &= \sum_{k,k',k''=0}^{N} \Big[\mathcal{I}_k \otimes \mathcal{M}^*_{k'}\Big] |\mathbb{1}_{k''}\rangle\!\rangle \\ &= \sum_{k=0}^{N} \Big[\mathcal{I}_k \otimes \mathcal{M}^*_k\Big] |\mathbb{1}_k\rangle\!\rangle \,, \end{aligned}$$

since it is assumed that by definition

$$\Big[\mathcal{I}_k \otimes \mathcal{M}^*_{k'}\Big] |\mathbb{1}_{k''}\rangle\!\rangle \equiv 0 \,, \qquad k'' \notin \{k, k'\} \,.$$

Now, using (6) one can rewrite (20) into the form

$$\begin{aligned} M &= |\mathcal{M}^*\rangle\!\rangle \langle\!\langle \mathcal{M}^*| \\ &= \sum_{k,k'=0}^{N} \Big[\mathcal{I}_k \otimes \mathcal{M}^*_k\Big] |\mathbb{1}_k\rangle\!\rangle \langle\!\langle \mathbb{1}_{k'}| \Big[\mathcal{I}_{k'} \otimes \mathcal{M}^*_{k'}\Big]^\dagger , \end{aligned}$$

which is clearly different from the case (23), since it contains off-diagonal elements which preserve coherence between k-particle sectors. One concrete example of this special case is the dSD protocol, discussed in the previous Section. Another example is a single-particle unitary operator

$$U = \sum_{i,j} u_{ij} \tilde{a}_i^\dagger a_j \,.$$

Then, its capped Fock-space generalisation is given as

$$\mathcal{M} = \sum_{k=0}^{N} \mathcal{M}_k = |0\rangle\langle 0| + \sum_{k=1}^{N} \frac{1}{k!} : U^{\otimes k} : ,$$

where $: U^{\otimes k} :$ is the normal ordering of $U^{\otimes k}$.

Given the capped Fock spaces and actions of instruments in all gates, a process matrix is defined in the same way as in Section 2, according to Eq. (4). A process matrix maps the tensor product of output spaces for all gates into the tensor product of input spaces for all gates. For example, if the process under consideration is a quantum circuit (see Section 2 of [9]), the corresponding process matrix can be represented as a tensor product of transport vectors, each corresponding to a wire connecting two gates. Transport vectors are defined in the same way as (21), where in (22) the first set of creation operators corresponds to the input space of the wire, while the second set corresponds to its output space. Given that a wire is connecting two gates, its input and output spaces correspond to the output and input subspaces of the two gates, respectively. A gate can in general have multiple incoming or outgoing wires attached to it. Therefore, its input (output) space is a tensor product of all output (input) spaces of the corresponding wires.

5 Conclusions

5.1 Summary of the results

In this work we have presented a detailed account of the dSD protocol, formulating it within the process matrix formalism. Analysing the role of the vacuum state in the dSD protocol and its process matrix description, we gave the operational interpretation of the vacuum. Our analysis shows that the interaction with the vacuum should be treated as an operation, on equal footing with all other interactions, thus representing a resource in quantum information protocols (including, for example, [23, 24]). As a consequence, the optical implementation of the quantum switch protocol features four rather than just two operations, in contrast to what was claimed in the literature [3, 4, 5, 6, 7, 8]. Furthermore, we have applied the process matrix formalism to the second quantisation framework restricted to capped Fock spaces, providing the description of systems of identical particles.

5.2 Discussion

The first important point of this work is the necessity of explicitly introducing the interaction with the vacuum as a legitimate operation in the dSD protocol, on equal footing with any other operation. Indeed, the very lack of detection of the particle in the protocol provides an equal amount of information as its detection (*explicit* interaction). As a consequence, instead of interpreting the absence of particle as noninteraction, one should interpret it as the interaction between the vacuum and the apparatus, and thus as an operation. Including the interaction with the vacuum as an operation poses a question of the method of counting operations in a given protocol, since the operations corresponding to the interaction with the vacuum cannot be counted.

The introduction of the vacuum into the process matrix formalism gives a natural motivation to extend the latter to the case of identical particles, both bosons and fermions, which is the second important point of this work. However, note that while employing the formalism of second quantisation, our construction still features only a discrete number of gates. This discreteness means that we still work in *particle ontology* (i.e., mechanics). Nevertheless, our construction is an important first step towards defining the process matrix formalism in *field ontology*, i.e., fully fledged QFT.

5.3 Future lines of investigation

As mentioned in the discussion, a natural next line of investigation would be a generalisation of the process matrix formalism to full, or at least perturbative, QFT. This would include an analysis of non-inertial gates and the corresponding Unruh effect. In addition, a mathematically rigorous formulation of the axioms for the process matrix description in Fock spaces is also an important topic to be addressed. While the primary interest in process matrices lies in their application to higher order processes [25, 26], their generalisation to QFT would also be of great interest. Finally, addressing in more detail the interaction between the agent and the vacuum within the operational approach is an interesting topic of future research.

Acknowledgments

The authors wish to thank Borivoje Dakić and Flavio del Santo for useful discussions.

RF acknowledges support from DP-PMI and Fundação para a Ciência e Tecnologia (FCT) through Grant PD/BD/128636/2017, from FCT/MCTES through national funds and when applicable EU funds under the project UIDB/50008/2020, and the QuantaGENOMICS project, through the EU H2020 QuantERA II Programme, Grant Agreement No 101017733.

NP's work was partially supported by SQIG – Security and Quantum Information Group of Instituto de Telecomunicações, by Programme (COMPETE 2020) of the Portugal 2020 framework [Project Q.DOT with Nr. 039728 (POCI-01-0247-FEDER-039728)] and the Fundação para a Ciência e a Tecnologia (FCT) through national funds, by FEDER, COMPETE 2020, and by Regional Operational Program of Lisbon, under UIDB/50008/2020 (actions QuRUNNER, QUESTS), Projects QuantumMining POCI-01-0145-FEDER-031826, PREDICT PTDC/CCI-CIF/29877/2017, CERN/FIS-PAR/0023/2019, QuantumPrime PTDC/EEI-TEL/8017/2020, as well as the FCT Estímulo ao Emprego Científico grant no. CEECIND/04594/2017/CP1393/CT000.

MV was supported by the Ministry of Science, Technological Development and Innovations of the Republic of Serbia, by the bilateral scientific cooperation between Portugal and Serbia through the project "Symmetries and Quantization - 2020-2022", no. 337-00-00227/2019-09/57 supported by the Portuguese Foundation for Science and Technology (FCT), Portugal, and the Ministry of Education, Science and Technological Development of the Republic of Serbia, and by the Science Fund of the Republic of Serbia, grant 7745968, "Quantum Gravity from Higher Gauge Theory 2021" — QGHG-2021. The contents of this publication are the sole responsibility of the authors and can in no way be taken to reflect the views of the Science Fund of the Republic of Serbia.

References

[1] O. Oreshkov, F. Costa, and Č. Brukner. "Quantum correlations with no causal order". Nature Communications **3**, 1092 (2012).

[2] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron. "Quantum computations without definite causal structure". Physical Review A **88**, 022318 (2013).

[3] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner. "Witnessing causal nonseparability". New Journal of Physics **17**, 102001 (2015).

[4] J. Bavaresco, M. Araújo, Č. Brukner, and M. T. Quintino. "Semi-device-independent certification of indefinite causal order". Quantum **3**, 176 (2019).

[5] P. Allard Guérin, G. Rubino, and Č. Brukner. "Communication through quantum-controlled noise". Physical Review A **99**, 062317 (2019).

[6] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. A. Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther. "Experimental superposition of orders of quantum gates". Nature Communications **6**, 7913 (2015).

[7] G. Rubino, L. A Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther. "Experimental verification of an indefinite causal order". Science Advances **3**, e1602589 (2017).

[8] G. Rubino, L. A. Rozema, F. Massa, M. Araújo, M.a Zych, Č. Brukner, and P. Walther. "Experimental entanglement of temporal order". Quantum **6**, 621 (2022).

[9] N. Paunković and M. Vojinović. "Causal orders, quantum circuits and spacetime: distinguishing between definite and superposed causal orders". Quantum **4**, 275 (2020).

[10] V. Vilasini and R. Renner. "Embedding cyclic causal structures in acyclic spacetimes: no-go results for process matrices" (2022). arXiv:2203.11245.

[11] N. Ormrod, A. Vanrietvelde, and J. Barrett. "Causal structure in the presence of sectorial

constraints, with application to the quantum switch" (2022). arXiv:2204.10273.

[12] V. Vilasini. "An introduction to causality in quantum theory (and beyond)". Master's thesis. ETH. Zürich, Switzerland (2017).

[13] J. Åberg. "Operations and single-particle interferometry". Physical Review A **70**, 012103 (2004).

[14] G. Chiribella and H. Kristjánsson. "Quantum Shannon theory with superpositions of trajectories". Proceedings of the Royal Society A **475**, 20180903 (2019).

[15] H. Kristjánsson, W. Mao, and G. Chiribella. "Witnessing latent time correlations with a single quantum particle". Physical Review Research **3**, 043147 (2021).

[16] F. Del Santo and B. Dakić. "Two-way communication with a single quantum particle". Physical Review Letters **120**, 060503 (2018).

[17] F. Del Santo and B. Dakić. "Coherence equality and communication in a quantum superposition". Physical Review Letters **124**, 190501 (2020).

[18] B. Ghosh, A. Mukherjee, and S. Aravinda. "Advantage in two-way communication using non-classical states of light" (2019). arXiv:1904.04596.

[19] F. Massa, A. Moqanaki, Ä. Baumeler, F. Del Santo, J. A. Kettlewell, B. Dakić, and P. Walther. "Experimental two-way communication with one photon". Advanced Quantum Technologies **2**, 1900050 (2019).

[20] C. Portmann, C. Matt, U. Maurer, R. Renner, and B. Tackmann. "Causal boxes: Quantum information-processing systems closed under composition". IEEE Transactions on Information Theory **63**, 3277–3305 (2017).

[21] O. Oreshkov. "Time-delocalized quantum subsystems and operations: on the existence of processes with indefinite causal structure in quantum mechanics". Quantum **3**, 206 (2019).

[22] M. Araújo, F. Costa, and Č. Brukner. "Computational advantage from quantum-controlled ordering of gates". Physical Review Letters **113**, 250402 (2014).

[23] L.-Y. Hsu, C.-Y. Lai, Y.-C. Chang, C.-M. Wu, and R.-K. Lee. "Carrying an arbitrarily large amount of information using a single quantum particle". Physical Review A **102**, 022620 (2020).

[24] Y. Zhang, X. Chen, and E. Chitambar. "Building multiple access channels with a single particle". Quantum **6**, 653 (2022).

[25] P. Perinotti. "Causal structures and the classification of higher order quantum computations". Chapter in "Time in Physics", pages 103–127. Springer International Publishing. Cham (2017).

[26] A. Bisio and P. Perinotti. "Theoretical framework for higher-order quantum theory". Proceedings of the Royal Society A **475**, 20180706 (2019).

A Two lemmas for the process matrix evaluation

Lemma 1. Let $|\Psi^*\rangle\!\rangle^{X_O} = |\Psi^*\rangle^{X_O}$ represent a gate which has no input, while it prepares the state $|\Psi\rangle \in X_O$ as its output. Then, the scalar product of that vector and the transport vector $|\mathbb{1}\rangle\!\rangle^{X_O Y_I}$ is given as:

$$^{X_O}\langle\!\langle \Psi^* | \mathbb{1} \rangle\!\rangle^{X_O Y_I} = |\Psi\rangle^{Y_I} .$$

Proof. Using the fact that the transport vector is an unnormalized maximally entangled state, the explicit calculation goes as follows:

$$\begin{aligned} ^{X_O}\langle\!\langle \Psi^* | \mathbb{1} \rangle\!\rangle^{X_O Y_I} &= \langle \Psi^* |^{X_O} \sum_k |k\rangle^{X_O} |k\rangle^{Y_I} \\ &= \sum_k \langle \Psi^* | k \rangle \, |k\rangle^{Y_I} \\ &= \sum_k \langle k | \Psi \rangle \, |k\rangle^{Y_I} \\ &= |\Psi\rangle^{Y_I} , \end{aligned}$$

where we have used the unit decomposition $I = \sum_k |k\rangle\langle k|$ and the fact that $\langle \Psi^* | k \rangle = \langle \Psi | k \rangle^* = \langle k | \Psi \rangle$.

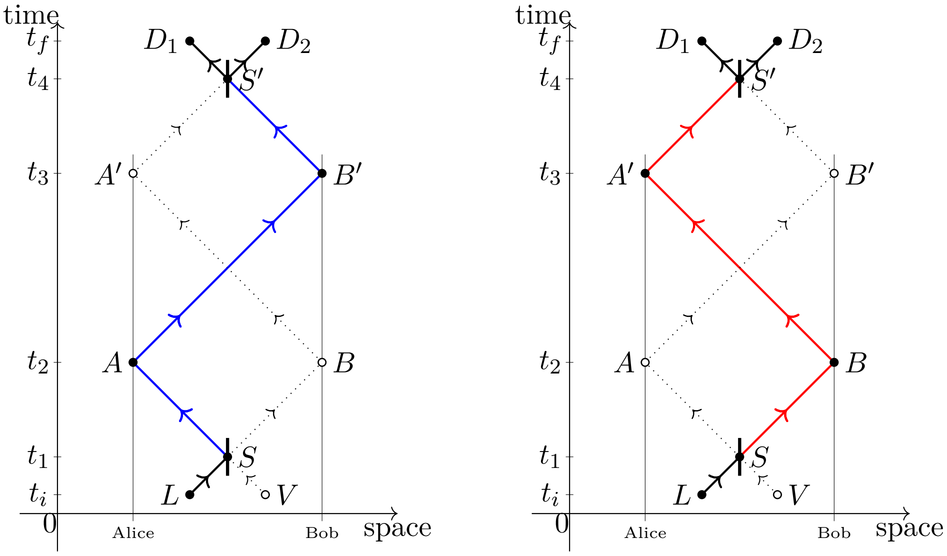


Figure 3: Two branches of the coherent superposition of the optical switch protocol.

Lemma 2. Let

$$|U^*\rangle\!\rangle^{X_I X_O} = \left[I^{X_I X_I} \otimes (U^*)^{X_O X_I} \right] |\mathbb{1}\rangle\!\rangle^{X_I X_I}$$

represent a gate which performs the operation $U : X_I \to X_O$, and let $|W\rangle\!\rangle = |\Psi\rangle^{X_I} |\mathbb{1}\rangle\!\rangle^{X_O Y_I}$. Then the scalar product of the two is

$$_{X_I X_O}\langle\!\langle U^* | W \rangle\!\rangle = \Big(U|\Psi\rangle \Big)^{Y_I}.$$

Proof. Again using the expansion of the transport vectors as unnormalized maximally entangled states, the explicit calculation goes as follows:

$$\begin{aligned}
_{X_I X_O}\langle\!\langle U^* | W \rangle\!\rangle &= \langle\!\langle \mathbb{1} |^{X_I X_I} \left[I^{X_I X_I} \otimes (U^T)^{X_I X_O} \right] |\Psi\rangle^{X_I} |\mathbb{1}\rangle\!\rangle^{X_O Y_I} \\
&= \sum_k \langle k|^{X_I} \langle k|^{X_I} \left[I^{X_I X_I} \otimes (U^T)^{X_I X_O} \right] |\Psi\rangle^{X_I} \sum_m |m\rangle^{X_O} |m\rangle^{Y_I} \\
&= \sum_{k,m} \Big(\langle k|^{X_I} I^{X_I X_I} |\Psi\rangle^{X_I} \Big) \Big(\langle k|^{X_I} (U^T)^{X_I X_O} |m\rangle^{X_O} \Big) |m\rangle^{Y_I} \\
&= \sum_{k,m} \Big(\langle k|\Psi\rangle \Big) \Big(\langle m|U|k\rangle \Big) |m\rangle^{Y_I} \\
&= \sum_m \langle m|U \left(\sum_k |k\rangle\langle k| \right) |\Psi\rangle \; |m\rangle^{Y_I} \\
&= \sum_m \langle m|U|\Psi\rangle \; |m\rangle^{Y_I} \\
&= \Big(U|\Psi\rangle \Big)^{Y_I},
\end{aligned}$$

where we have again used the unit decomposition and the fact that $\langle k|U^T|m\rangle = \langle m|U|k\rangle$.

B Time-delocalised operations in the optical switch

Figure 3 depicts two branches coherently superposed in the optical switch. The left diagram represents the branch in which the photon first enters Alice's lab, and then Bob's. On the right, the photon first visits Bob's lab, and then Alice's. Whenever the photon enters Alice's lab, she applies unitary U (in A, left diagram, or A', right diagram), while Bob interacts with the vacuum (in B, left diagram, or B', right diagram). Analogously, whenever the photon enters Bob's lab, he applies unitary V (in B, right diagram, or B', left diagram), while Alice interacts with the vacuum (in A, right diagram, or A', left diagram).

Since applying the unitaries in a quantum protocol are operations, and since in the optical switch they are applied by Alice (U) and Bob (V) at two different times, we say that the optical switch features two time-delocalised operations U (at A and A') and V (at B and B').

Since the interaction with the vacuum in the dSD protocol is an operation, and since in the optical switch it is applied by Alice and Bob at two different times, one can say that the optical switch features two time-delocalised operations of the interaction with the vacuum (at A and A', as well as at B and B').

universe

MDPI

Article

Equivalence Principle in Classical and Quantum Gravity

Nikola Paunković [1,†] and Marko Vojinović [2,*,†]

[1] Instituto de Telecomunicações and Departamento de Matemática, Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais 1, 1049-001 Lisboa, Portugal
[2] Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia
* Correspondence: vmarko@ipb.ac.rs
† These authors contributed equally to this work.

Abstract: We give a general overview of various flavours of the equivalence principle in classical and quantum physics, with special emphasis on the so-called weak equivalence principle, and contrast its validity in mechanics versus field theory. We also discuss its generalisation to a theory of quantum gravity. Our analysis suggests that only the strong equivalence principle can be considered fundamental enough to be generalised to a quantum gravity context since all other flavours of equivalence principle hold only approximately already at the classical level.

Keywords: equivalence principle; general relativity; quantum gravity

1. Introduction

Quantum mechanics (QM) and general relativity (GR) are the two cornerstones of modern physics. Yet, merging them together in a quantum theory of gravity (QG) is still elusive despite the nearly century-long efforts of vast numbers of physicists and mathematicians. While the majority of the attempts were focused on trying to formulate the full theory of quantised gravity, such as string theory, loop quantum gravity, non-commutative geometry, and causal set theory, to name a few, a number of recent studies embraced a rather more modest approach by exploring possible consequences of basic features and principles of QM and GR, and their status, in a tentative theory of QG. Acknowledging that the superposition principle, as a defining characteristic of any quantum theory, must be featured in QG as well, led to a number of papers studying gravity-matter entanglement [1–7], genuine indefinite causal orders [8–15], quantum reference frames [16–20] and deformations of Lorentz symmetry [21–25], to name a few major research directions. Exploring spatial superpositions of masses, and in general gravitational fields, led to the analysis of the status of various versions of the equivalence principle, and their exact formulations in the context of QG. In particular, in [26], it was shown that the weak equivalence principle (WEP) should generically be violated in the presence of a specific type of superpositions of gravitational fields, describing small quantum fluctuations around a dominant classical geometry. On the other hand, a number of recent studies propose generalisations of WEP to QG framework (see for example [16,20,27–31]), arguing that it remains satisfied in such scenarios, a result *seemingly* at odds with [26] (for details, see the discussion from Section 5).

The modern formulation of WEP is given in terms of a *test particle* and it's *trajectory*: it is a *theorem* within the mathematical formulation of GR stating that the trajectory of a test particle satisfies the so-called geodesic equation [32–46], while it is *violated* within the context of QG, as shown in [26]. In this paper, we present a brief overview of WEP in GR and a critical analysis of the notions of particle and trajectory in both classical and quantum mechanics, as well as in the corresponding field theories.Our analysis demonstrates that WEP, as well as all other flavours of the equivalence principle (EP) aside from the strong one (SEP), hold only approximately. From this we conclude that neither WEP nor any other flavour of EP (aside from SEP) can be considered a viable candidate for generalisation to the quantum gravity framework.

Citation: Paunković, N.; Vojinović, M. Equivalence Principle in Classical and Quantum Gravity. *Universe* **2022**, *8*, 598. https://doi.org/10.3390/universe8110598

Academic Editors: Arundhati Dasgupta and Alfredo Iorio

Received: 1 October 2022
Accepted: 10 November 2022
Published: 12 November 2022

The paper is organised as follows. In Section 2, we give a brief historical overview of various flavours of the equivalence principle, with a focus on WEP. In Section 3, we analyse the notion of a trajectory in classical and quantum mechanics, while in Section 4 we discuss the notion of a particle in field theory and QG. Finally, in the Conclusion, we briefly review and discuss our results, and present possible future lines of research.

2. Equivalence Principle in General Relativity

The equivalence principle is one of the most fundamental principles in modern physics. It is one of the two cornerstone building blocks for GR, the other being the principle of general relativity. While its importance is well understood in the context of gravity, it is often underappreciated in the context of other fundamental interactions. In addition, there have been numerous studies and everlasting debates about whether EP holds also in quantum physics, if it should be generalised to include quantum phenomena or not, etc. Finally, EP has been historically formulated in a vast number of different ways, which are often not mutually equivalent, leading to a lot of confusion about the actual statement of the principle and its physical content [47–53]. Given the importance of EP, and the amount of confusion around it, it is important to try and help clarify these issues.

The equivalence principle is best introduced by stating its purpose—in its traditional sense, the purpose of EP is to *prescribe the interaction between gravity and all other fields in nature, collectively called matter* (by "matter" we assume not just fermionic and scalar fields, but also gauge vector bosons, i.e., nongravitational interaction fields). This is important to state explicitly since EP is often mistakenly portrayed as a property of gravity alone, without any reference to matter. In a more general, less traditional, and often not appreciated sense, the purpose of EP is to prescribe the interaction between *any gauge field* and all other fields in nature (namely fermionic and scalar matter, as well as other gauge fields, including gravity), which we will reflect on briefly in the case of electrodynamics below.

Given such a purpose, let us for the moment concentrate on the gravitational version of EP, and provide its modern formulation, as it is known and understood today. The statement of the equivalence principle is the following:

> *The equations of motion for matter coupled to gravity remain locally identical to the equations of motion for matter in the absence of gravity.*

This kind of statement requires some unpacking and comments.

- When comparing the equations of motion in the presence and in the absence of gravity, the claim that they remain identical may naively suggest that gravity does not influence the motion of matter in any way whatsoever. However, on closer inspection, the statement is that the two sets of equations remain *locally* identical, emphasising that the notion of locality is a crucial feature of the EP. While equations of motion are already local in nature (since they are usually expressed as partial differential equations of finite order), the actual interaction between matter and gravity enters only when *integrating* those equations to find a solution (see Appendix A for a detailed example).
- In order to compare the equations of motion for matter in the presence of gravity to those in its absence, the equations themselves need to be put in a suitable form (typically expressed in general curvilinear coordinates, as tensor equations). The statement of EP relies on a theorem that this can always be achieved, first noted by Erich Kretschmann [54].
- Despite being dominantly a statement about the interaction between matter and gravity, EP also implicitly suggests that the best way to describe the gravitational field is as a property of the geometry of spacetime, such as its metric [55]. In that setup, EP can be reformulated as a statement of *minimal coupling* between gravity and matter, stating that equations of motion for matter may depend on the spacetime metric and its first derivatives, but not on its (antisymmetrised) second derivatives, i.e., the *spacetime curvature does not explicitly appear in the equations of motion for matter*.

- The generalisation of EP to other gauge fields is completely straightforward, by replacing the role of gravity with some other gauge field, and suitably redefining what matter is. For example, in electrodynamics, the EP can be formulated as follows:

  > *The equations of motion for matter coupled to the electromagnetic field remain locally identical to the equations of motion for matter in the absence of the electromagnetic field.*

  This statement can also be suitably reformulated as the minimal coupling between the electromagnetic (EM) field and matter, i.e., coupling matter to the electromagnetic potential A_μ but not to the corresponding field strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. This is in fact the standard way the EM field is coupled to matter (see Appendix A for an illustrative example). Even more generally, the gauge field sector of the whole Standard Model of elementary particles (SM) is built using the minimal coupling prescription, meaning that the suitably generalised version of the EP actually prescribes the interaction between matter and all fundamental interactions in nature, namely strong, weak, electromagnetic and gravitational. In this sense, EP is a cornerstone principle for the whole fundamental physics, as we understand it today.

Of course, much more can be said about the statement of EP, its consequences, and various other details. However, in this work, our attention will focus on the so-called *weak equivalence principle* (WEP), which is a reformulation of EP applied to matter which consists of mechanical particles. To that end, it is important to understand various flavours and reformulations of EP that have appeared through history.

As with any deep concept in physics, EP has been expressed historically through a painstaking cycle of formulating it in a precise way, understanding the formulation, understanding the drawbacks of that formulation, coming up with a better formulation, and repeating. In this sense, EP, as quoted above, is a modern product of long and meticulous refinement over several generations of scientists. Needless to say, each step in that process made its way into contemporary physics textbooks, leading to a plethora of different formulations of EP that have accumulated in the literature over the years. This can bring about a lot of confusion about what EP actually states [47–50] since various formulations from old and new literature may often be not merely phrased differently, but in fact substantively inequivalent. To that end, let us comment on several of the most common historical statements of EP (for a more detailed historical overview and classification, see [56,57]), and their relationship with the modern version:

- *Equality of gravitational and inertial mass.* This is one of the oldest variants of EP, going back to Newton's law of universal gravitation. The statement claims that the "gravitational charge" of a body is the same as the body's resistance to acceleration, in the sense that the mass appearing on the left-hand side of Newton's second law of motion exactly cancels the mass appearing in Newton's gravitational force law on the right-hand side. This allows one to relate it to the modern version of EP, in the sense that a suitably accelerated observer could rewrite Newton's law of motion as the equation for a free particle, exploiting the cancellation of the "intertial force" and the gravitational force on the right-hand side of the equation. Unfortunately, this version of EP is intrinsically nonrelativistic, and applicable only in the context of Newtonian gravity since already in GR the source of gravity becomes the full stress-energy tensor of matter fields, rather than just the total mass. Finally, this principle obviously fails when applied to photons, as demonstrated by the gravitational bending of light.
- *Universality of free fall.* Going back all the way to Galileo, this statement claims that the interaction between matter and gravity does not depend on any intrinsic property of matter itself, such as its mass, angular momentum, chemical composition, temperature, or any other property, leading to the idea that gravity couples universally (i.e., in the same way) to all matter. Formulated from experimental observations by Galileo, its validity is related to the quality of experiments used to verify it. As we shall see below,

in a precise enough setting, one can experimentally observe direct coupling between the angular momentum of a body and spacetime curvature [32–46], invalidating the statement.

- *Local equality between gravity and inertia.* Often called Einstein's equivalence principle, the statement claims that a local and suitably isolated observer cannot distinguish between accelerating and being at rest in a uniform gravitational field. While this statement is much closer in spirit to the modern formulation of EP, it obscures the crucial aspect of the principle — coupling of matter to gravity. Namely, in this formulation, it is merely implicit that the only way an observer can *attempt to distinguish* gravity from inertia is by making local experiments using some form of *matter*, i.e., studying the equations of motion of matter in the two situations and trying to distinguish them by observing whether or not matter behaves differently. Moreover, the statement is often discussed in the context of mechanics, arguing that any given particle does not distinguish between gravity and inertia. This has two main pitfalls—first, the reliance on particles is very misleading (as we will discuss below in much more detail), and second, it implicitly suggests that gravity and inertia are the same phenomenon, which is completely false. Namely, inertia can be understood as a specific form of gravity, but a general gravitational field cannot be simulated by inertia, since inertia cannot account for tidal effects of inhomogeneous configurations of gravity.
- *Weak equivalence principle.* Stating that the equations of motion of particles do not depend on spacetime curvature, or equivalently, that the motion of a free particle is always a geodesic trajectory in spacetime, WEP is in fact an application of modern EP to mechanical point-like particles (i.e., test particles). One can argue that, as far as the notion of a point-like particle is a well-defined concept in physics, WEP is a good principle. Nevertheless, as we will discuss below in detail, the notion of a point-like particle is an idealisation that does not actually have any counterpart in reality, in either classical or quantum physics. Regarding a realistic particle (with nonzero size), WEP *never holds*, due to the explicit effect of gravitational tidal forces across the particle's size. In this sense, WEP can be considered at best an *approximate* principle, which can be assumed to hold only in situations where particle size can be approximated to zero.
- *Strong equivalence principle.* This version of the principle states that the equations of motion of all fundamental fields in nature do not depend on spacetime curvature (see [55], Section 16.2, page 387). To the best of our knowledge so far, fields are indeed the most fundamental building blocks in modern physics (such as SM), while the strength of the gravitational field is indeed described by spacetime curvature (as in GR). In this sense, the statement of SEP is actually an instance of EP applied to field theory, and as such equivalent to the modern statement of EP. So far, all our knowledge of natural phenomena is consistent with the validity of SEP.

As can be seen from the above review, various formulations of EP are both mutually inequivalent and have different domains of applicability. Specifically, only SEP holds universally, while all other flavours of EP hold only approximately. In the remainder of the paper, we focus on the study of WEP since recently it gained a lot of attention in the literature [20,27–29,31], primarily in the context of its generalisation to a "quantum WEP", and in the context of a related question of particle motion in a quantum superposition of different gravitational configurations, the latter being a scenario that naturally arises in QG. Since WEP is stated in terms of a test particle and its trajectory, in order to try and generalise it to the scope of QG one should first analyse these two notions in classical and quantum mechanics and field theory in more detail.

3. The Notion of Trajectory in Classical and Quantum Mechanics

A trajectory of a physical system in three-dimensional space is a set of points that form a line, usually parameterised by time. More formally, a trajectory is a set $\{(x(t), y(t), z(t)) \in$

$\mathbb{R}^3 | t \in [t_i, t_f] \subset \mathbb{R} \wedge t_i < t_f\}$, given by three smooth functions $x, y, z : \mathbb{R} \mapsto \mathbb{R}$. Depending on the nature of the system, the choice of points that form its trajectory may vary.

In classical mechanics, one often considers an ideal "point-like particle" localised in one spatial point $(x(t), y(t), z(t))$ at each moment of time t, in which case the choice of the points forming the system's trajectory is obvious. In the case of systems continuously spread over certain volumes ("rigid bodies", or "objects") or composite systems consisting of several point-like particles or bodies, it is natural to consider their centres of mass as points that form the trajectory. While this definition is natural, widely accepted, and formally applicable to any classical mechanical system, there are cases in which the very notion of a trajectory loses its intuitive, as well as useful, meaning.

Consider for example an electrical dipole, i.e., a system of two point-like particles with equal masses and opposite electrical charges, separated by the distance $\ell(t)$. As long as this distance stays "small" and does not vary significantly with time, the notion of a trajectory of a dipole, defined as the set of centres of mass of the two particles, does meet our intuition, and can be useful. Informally, if the trajectories of each of the two particles are "close" to each other, they can be approximated, and consequently represented, by the trajectory of the system's centre of mass. However, if the separate trajectories of the two particles diverge, one going to the "left", and another to the "right", one could hardy talk of a trajectory of such a composite system, although the set of locations of its centres of mass is still well defined. In fact, the dipole itself ceases to make physical sense when the distance between its constituents is large.

Moving to the realm of quantum mechanics, due to the superposition principle, even the ideal point-like particles do not have a well-defined position, which is further quantified by the famous Heisenberg uncertainty relations. Thus, the trajectory of point-like particles (and any system that in a given regime can be approximated to be point-like) is defined as a set of expectation values of the position operator. Like in the case of composite classical systems, here as well the definition of a trajectory of a point-like particle is mathematically always well defined, yet for a very similar reason is applicable only to certain cases. Namely, in order to give a useful meaning to the above definition of trajectory, the system considered must be *well localised*. Consider for example the double-slit experiment, in which the point-like particle is highly delocalised so that we say that *its trajectory is not well defined*, even though the set of the expectation values of the position operator is.

We see that, while in mechanics both the notions of a particle and its trajectory are rather straightforward and always well defined, the latter make sense only if our system is well localised in space (see for example [58], where the authors analyse the effects of wave-packet spreading to the notion of a trajectory).

4. The Notion of a Particle in Field Theory

While in classical mechanics a point-like particle is always well localised, we have seen that in the quantum case one must introduce an additional constraint in order for it to be considered localised—the particle should be represented by a wave-packet. The source for this requirement lies in the fact that quantum particles, although mechanical, are represented by a *wavefunction*. Thus, it is only to be expected that when moving to the realm of the field ontology, the notion of a particle becomes even more involved and technical.

In field theory, the fundamental concept is the *field*, rather than a particle. The notion of a particle is considered a derived concept, and in fact in QFT one can distinguish two vastly different phenomena that are called "particles".

The first notion of a particle is an elementary excitation of a free field. For example, the state

$$|\Psi\rangle = \hat{a}^\dagger(\vec{k})|0\rangle \,,$$

is called a *single particle state* of the field, or a *plane-wave-particle*. It has the following properties:

- It is an eigenstate of the *particle number operator* for the eigenvalue 1.

- It has a sharp value of the momentum $\vec{k}$, and corresponds to a completely delocalised plane wave configuration of the field.
- It has no centre of mass, and no concept of "position" in space since the "position operator" is not a well-defined concept for the field.
- States of this kind are said to describe *elementary particles*, understood as asymptotic free states of past and future infinity, in the context of the *S*-matrix for scattering processes. An example of a real scalar particle of this type would be the *Higgs particle*. For fields of other types (Dirac fields, vector fields, etc.) examples would be an *electron*, a *photon*, a *neutrino*, an asymptotically free *quark*, and so on. Essentially, all particles tabulated in the Standard Model of elementary particles are of this type.

Note that all the above notions are defined within the scope of free field theory, and do not carry over to interacting field theory. In other words, free field theory is a convenient idealisation, which does not really reflect realistic physics. One should therefore understand the concept of a plane-wave-particle in this sense, merely as a convenient mathematical approximation. Moreover, the particle number operator is not an invariant quantity, as demonstrated by the Unruh effect. We should also emphasise that in an interacting QFT, the proper way to understand the notion of a particle is as a localised wave-packet, interacting with its virtual particle cloud, which does have a position in space and whose momentum is defined through its group velocity. In this sense, the particle as a wave-packet could be better interpreted as a kink, discussed below.

The second notion of the particle in field theory is a bound state of fields, also called a *kink solution*. This requires an interacting theory since interactions are necessary to form bound states. This kind of configuration of fields has the following properties:

- It is not an eigenstate of the particle number operator, and the expectation value of this operator is typically different from 1.
- It is usually well localised in space, and does not have a sharp value of momentum.
- As long as the kink maintains a stable configuration (i.e., as long as it does not decay), one can in principle assign to it the concept of *size*, and as a consequence also the concepts of *centre of mass*, *position in space*, and *trajectory*. In this sense, a kink can play the role of a test particle.
- States of this kind are said to describe *composite particles*. Given an interacting theory such as the Standard Model, under certain circumstances quarks and gluons form bound states called a *proton* and a *neutron*. Moreover, protons and neutrons further form bound states called *atomic nuclei*, which together with electrons and photons form *atoms*, *molecules*, and so on.

For a kink, the notions of centre of mass, position in space and size are described only as classical concepts, i.e., as expectation values of certain field operators, such as the stress-energy tensor. Moreover, given the nonzero size of the kink, its centre of mass and position are not uniquely defined, even classically, since in relativity different observers would assign different points as the centre of mass.

Given the two notions of particles in QFT, one can describe two different corresponding notions of WEP. In principle, one first needs to apply SEP in order to couple the matter fields to gravity, at the fundamental level. Assuming this is completed, the motions of both the plane-wave-particles and kink particles can be derived from the combined set of Einstein's equations and matter field equations, without any appeal to any notion of WEP. In this sense, once the trajectory of the particle in the background gravitational field has been determined from the field equations, one can verify *as a theorem* whether the particle satisfies WEP or not.

Specifically, in the case of a matter field coupled to general relativity such that it locally resembles a plane wave, one can apply the WKB approximation to demonstrate that the wave 4-vector $k^{\mu}(x)$, orthogonal to the wavefront at its every point $x \in \mathbb{R}^4$, will satisfy a geodesic equation,

$$k^{\mu}(x)\nabla_{\mu}k^{\lambda}(x) = 0\,. \tag{1}$$

However, given that the plane-wave-particle is completely delocalised in space, the fact that the wave 4-vector satisfies the geodesic equation could hardly be interpreted as "the particle following a geodesic trajectory", and thus obeying WEP. Indeed, identifying the vector field orthogonal to the wavefront to the notion of "particle's trajectory" is at best an abuse of terminology.

Next, in the case of the kink particle coupled to general relativity, one assumes the configuration of the background gravitational field is such that the particle maintains its structure and that its size can be completely neglected. One can then apply the procedure given in [26,32–46] to demonstrate that the 4-vector $u^\mu(\tau)$, tangent to the kink's world line (i.e., its trajectory), will satisfy a geodesic equation ($\tau \in \mathbb{R}$ represents kink's proper time),

$$u^\mu(\tau)\nabla_\mu u^\lambda(\tau) = 0\,. \tag{2}$$

Thus, one concludes that the kink obeys WEP as a *theorem* in field theory, without the necessity to actually postulate it.

Note the crucial difference between Equations (1) and (2)—while the former features 4-dimensional variable x, the latter is given in terms of only 1-dimensional proper time τ. This reflects the fact that the plane-wave-particle is a highly delocalised object, with no well-defined position and trajectory, while the kink is a highly localised object, with a well-defined position and trajectory. As a consequence, WEP can be formulated only for the kink, and not for the plane-wave particle.

In the case of the kink, it is also important to emphasise that the zero-size approximation of the kink is crucial. Namely, without this assumption, the particle will feel the tidal forces of gravity across its size, effectively coupling its angular momentum $J^{\mu\nu}(\tau)$ to the curvature of the background gravitational field [32–46] (see also [59] for a more refined analysis of tidal effects). This will give rise to an equation of motion for the kink of the form

$$u^\mu(\tau)\nabla_\mu u^\lambda(\tau) = R^\lambda{}_{\mu\rho\sigma} u^\mu(\tau) J^{\rho\sigma}(\tau)\,, \tag{3}$$

which features explicit coupling to curvature (absent from (2)) and thus fails to obey WEP. In this sense, for realistic kink solutions WEP is *always violated*, and can be considered to hold only as an approximation when the size of the particle can be completely neglected compared to the radius of curvature of the background gravitational field. If in addition the kink has negligible total energy, it can be used as a point-like test particle.

In the above discussion, while matter fields are described as quantum, using QFT, the background gravitational field is considered to be completely classical. It should therefore not be surprising that WEP may fail to hold if one allows the gravitational field to be quantum, such as matter fields, and one needs to revisit all steps of the above analysis from the perspective of QG. In fact, the case of the kink particle has been studied in precisely this scenario [26], and it has been shown that if the background gravitational field is in a specific type of quantum superposition of different configurations, the kink will fail to obey WEP even in the zero size approximation. Simply put, the equation of motion for the kink will contain extra terms due to the interference effects between superposed configurations of gravity, giving rise to an effective force that pushes the kink off the geodesic trajectory. Moreover, of course, similar to the case of classical gravity, the resulting conclusion is a *theorem*, which follows from the fundamental field equations of the theory. One of the assumptions of that theorem is that the field equations allow for kink solutions in the first place. Namely, it is entirely possible that in quantum gravity particles cannot be localised at all, as opposed to the classical case where such an approximation can be feasible. If that is the case, then one cannot even formulate (i.e., generalise from classical theory) the notion of WEP in a quantum gravity setup. However, one can instead assume that kink solutions do exist, as was performed in [26], where a particular superposition of gravitational fields was considered, describing small quantum fluctuations around a dominant classical geometry. It was then argued that such superpositions are compatible with the approximation of a well-defined localised particle (see the discussion around Equations (2.2) and (3.15), as well as Section 3.4 of that paper). As it turns out, even in such cases the trajectory of the kink fails to obey WEP. Therefore, the generalisations of WEP and other approximate versions of EP are not the best candidates for analysing the properties of quantum gravity.

Moreover, the assumption of a well-defined notion of a particle in the QG framework can also be supported from the point of view of nonrelativistic limit. Namely, in [4,5] an experiment was proposed in which the effects of QG fluctuations are expected to be observable, by measuring the motion of nonrelativistic particles. Furthermore, an extension of this experiment was also suggested [60], which aims to determine the potential difference between gravitational and inertial masses of those particles in such a setup. In fact, the relation between the two types of masses in the nonrelativistic limit has also been previously analysed in [26], predicting their difference due to quantum fluctuations of geometry. In this sense, the notion of a kink should make sense even in the QG setup, at least in the nonrelativistic limit.

For the case of the plane-wave-particle travelling through the superposed background of two gravitational field configurations, the analysis of the equation of motion for the wave-vector field $k^{\mu}(x)$, in the sense of [26,32–46], has not been performed so far (to the best of our knowledge). However, in principle, one can expect a similar interference term to appear in the WKB analysis, and give rise to a non-geodesic equation for the wave 4-vector as well. In this sense, it is to be expected that generically even the wavefronts of such plane-wave-particles would fail to obey WEP.

5. Conclusions and Discussion

In this paper, we give an overview of the equivalence principle and its various flavours formulated historically, with a special emphasis on the weak equivalence principle. We performed a critical analysis of the notions of particle and trajectory in various frameworks of physics, showing that the notion of a point-like particle and its trajectory are not always well defined. This in turn suggests that WEP might not be the best starting point for generalisation to QG, as we argue in more detail below.

As discussed in Section 4, in [26] it was shown that if superpositions of states of gravity and matter are allowed, WEP can be violated. It is important to note that the cases considered in [26] feature a specific type of superposition of three groups of states: the first consists of a single so-called dominant state—a classical state whose expectation values of the metric and the stress-energy tensors satisfy Einstein field equations; the second consists of states similar to the dominant one, with arbitrary coefficients; and the third consists of states quasi-orthogonal to the dominant one, but with negligible coefficients. Only then one may talk of a (well-localised) trajectory of the test particle in the overall superposed state and consequently about the straightforward generalisation of the classical WEP to the realm of QG. Considering that for the dominant state, being classical, the trajectory of the test particle follows the corresponding geodesic, we see that in the superposed state its trajectory would *deviate from the geodesic of the dominant state*, thus violating WEP. Note that, as discussed in Section 4, this deviation, in addition to classically weighted trajectories of the individual branches, also features purely quantum (i.e., off-diagonal) interference terms.

On the other hand, a number of recent studies propose generalisations of WEP to QG framework, arguing that it remains satisfied in such scenarios, a result *seemingly* at odds with [26]. For example, in [29–31], the authors consider superpositions of an arbitrary number of classical quasi-orthogonal states with arbitrary coefficients, arguing that since WEP is valid in each classical branch, it is valid in its superposition as well. If taken as a *definition* of what it means that a certain principle is satisfied in a superposition of different quantum states, then the above statement is manifestly true. As such, being a definition, it tells little about physics—it merely rephrases one statement ("principle A is separately satisfied in all branches of a superposition") with another, simpler ("principle A is satisfied in a superposition"). Namely, note that in [29,30], such a generalised version of EP plays no functional role in the analyses conducted in those papers. What does play a functional role is the statement of one version of classical EP (specifically, local equality between gravity and inertia) applied to each particular state in a superposition. All physically relevant (and otherwise interesting) conclusions of the two papers could be equally obtained without ever talking about the generalised EP. In addition, in [31] EP itself is not even the main focus of the paper, and its generalisation is just introduced in analogy to the analysis of the conservation laws, which is itself an interesting topic. On the other hand, in the case of weakly superposed gravitational fields, such as in proposed experiments [4,5], the violation of the equality of inertial and gravitational masses is to be expected [26,60]. Moreover, following the spirit of the above definition, one could be misled to conclude that the notions of the particle's position and trajectory are always well-defined, as long as they are defined in each (quasi-classical) branch of the superposition.

An alternative approach to the generalisation of EP to the quantum domain was proposed in [16,20,27,28]. In those works, the authors discuss the coupling of a spatially delocalised wave-particle to gravity, with the aim of generalising such a scenario to QG. To that end, they prove a theorem which essentially states that for such a delocalised wave-particle, even when it is entangled with the gravitational field, one can always find a quantum reference frame transformation, such that in the vicinity of a given spacetime point one has a locally inertial coordinate system. The theorem employs the novel techniques of quantum reference frames (QRF) to generalise to the quantum domain the well-known result from differential geometry, that in the infinitesimal neighbourhood of any spacetime point one can always choose a locally inertial coordinate system.

The authors then employ the theorem to generalise one flavour of EP to the quantum domain. Specifically, even if the wave-particle is entangled with the gravitational field, one can use the appropriate QRF transformation to switch to a locally inertial coordinate system, and then in that system "all the (nongravitational) laws of physics must take on their familiar non-relativistic form". Here, to the best of our understanding, the phrase "non-gravitational laws of physics" refers to the equations of motion for a quantum-mechanical wave-particle, while "non-relativistic form" means that these equations of motion take the same form as in special-relativistic context.

Our understanding is that the above wave-particle generalisation of EP lies somewhere "in between" mechanics and field theory, i.e., it is in a sense stronger than WEP, which discusses particles, but weaker than SEP, which discusses full-blown matter fields. Since it refers to wave-particles rather than kinks, our analysis of WEP and its reliance on the particle trajectory does not apply to this version of EP.

The methodology in [16,20,27,28] is that one should try to generalise even approximate flavours of EP, as a stopgap result in a bigger research programme, in the hope that they may still shed some light on QG. This is of course a legitimate methodology, and from that point of view these kinds of generalisations of EP to the quantum domain represent interesting results. Nevertheless, we also believe it would be preferable to formulate a generalisation of SEP, and in a way which does not appeal to reference frames at all, since that would be closer to the essence of the statement of EP, as discussed in Section 2.

To conclude, our analysis suggests that, instead of trying to generalise various approximate formulations of EP, one should rather talk of operationally verifiable statements regarding the (in)equality of gravitational and inertial masses, possible deviation from the geodesic motion of test particles, the universality of free fall, etc., and study other principles and their possible generalisations to QG, such as SEP (see Section 4.2 in [26]), background independence, quantum nonlocality, and so on.

Author Contributions: Investigation, N.P. and M.V.; writing—original draft preparation, N.P. and M.V.; writing—review and editing, N.P. and M.V. All authors have read and agreed to the published version of the manuscript.

Funding: NP's work was partially supported by SQIG—Security and Quantum Information Group of Instituto de Telecomunicações, by Programme (COMPETE 2020) of the Portugal 2020 framework [Project Q.DOT with Nr. 039728 (POCI-01-0247-FEDER-039728)] and the Fundação para a Ciência e a Tecnologia (FCT) through national funds, by FEDER, COMPETE 2020, and by Regional Operational Program of Lisbon, under UIDB/50008/2020 (actions QuRUNNER, QUESTS), Projects QuantumMining POCI-01-0145-FEDER-031826, PREDICT PTDC/CCI-CIF/29877/2017, CERN/FIS-PAR/0023/2019, QuantumPrime PTDC/EEI-TEL/8017/2020, as well as the FCT Estímulo ao Emprego Científico grant no. CEECIND/04594/2017/CP1393/CT000. MV was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, and by the Science Fund of the Republic of Serbia, grant 7745968, "Quantum Gravity from Higher Gauge Theory 2021"—QGHG-2021. The contents of this publication are the sole responsibility of the authors and can in no way be taken to reflect the views of the Science Fund of the Republic of Serbia.

Institutional Review Board Statement: Not applicable.

Acknowledgments: The authors wish to thank Časlav Brukner, Flaminia Giacomini, Chiara Marletto and Vlatko Vedral for useful discussions. MV is also indebted to Milovan Vasilić and Igor Salom for clarifications regarding the notion of symmetry localisation.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Here, we provide a detailed example of the two applications of the EP. First, we discuss the gravitational EP and apply it to a real scalar field, giving all mathematical details and discussing various related aspects such as locality, symmetry localisation, and so on. Then, we turn to the application of the gauge field generalisation of EP, using electrodynamics as an example. We describe how one can couple matter to an EM field, mimicking the previous gravitational example, and emphasize the analogy between the gravitational and EM case at each step. Note also that the non-Abelian gauge fields can be studied in exactly the same way. Finally, we discuss the case of test particles, and the violation of the WEP in both gravitational and electromagnetic cases.

Throughout this section, we assume that the Minkowski metric $\eta_{\mu\nu}$ has signature $(-,+,+,+)$.

Appendix A.1. The Gravitational Case

Let us begin with an example of a real scalar field in Minkowski spacetime, and apply the equivalence principle by coupling it to gravity. The equation of motion in this case is the ordinary Klein–Gordon equation,

$$\left(\eta^{\mu\nu}\partial_\mu\partial_\nu - m^2\right)\phi(x) = 0\,. \tag{A1}$$

As it stands, it describes the free scalar field in Minkowski spacetime, in an inertial coordinate system. In order to couple it to gravity (in the framework of GR), we first rewrite this equation into an arbitrary curvilinear coordinate system, as

$$\left(\tilde{g}^{\mu\nu}\tilde{\nabla}_\mu\tilde{\nabla}_\nu - m^2\right)\phi(\tilde{x}) = 0\,. \tag{A2}$$

Here the covariant derivative $\tilde{\nabla}_\mu$ is defined in terms of the Levi-Civita connection,

$$\tilde{\Gamma}^\lambda{}_{\mu\nu} = \frac{1}{2}\tilde{g}^{\lambda\sigma}\left(\partial_\mu\tilde{g}_{\nu\sigma} + \partial_\nu\tilde{g}_{\mu\sigma} - \partial_\sigma\tilde{g}_{\mu\nu}\right), \tag{A3}$$

which is in turn defined in terms of the curvilinear Minkowski metric $\tilde{g}_{\mu\nu}$. Note that the tilde symbol denotes the fact that this metric has been obtained by a coordinate transformation $\tilde{x}^\mu = \tilde{x}^\mu(x)$ from the Minkowski metric in an inertial coordinate system, $\eta_{\mu\nu}$,

$$\tilde{g}_{\mu\nu} = \frac{\partial x^\rho}{\partial\tilde{x}^\mu}\frac{\partial x^\sigma}{\partial\tilde{x}^\nu}\eta_{\rho\sigma}\,, \tag{A4}$$

and, therefore, if one were to evaluate the Riemann curvature tensor using $\tilde{g}_{\mu\nu}$ and $\tilde{\Gamma}^{\lambda}{}_{\mu\nu}$, according to the equation

$$R^{\lambda}{}_{\rho\mu\nu} = \partial_{\mu}\tilde{\Gamma}^{\lambda}{}_{\rho\nu} - \partial_{\nu}\tilde{\Gamma}^{\lambda}{}_{\rho\mu} + \tilde{\Gamma}^{\lambda}{}_{\sigma\mu}\tilde{\Gamma}^{\sigma}{}_{\rho\nu} - \tilde{\Gamma}^{\lambda}{}_{\sigma\nu}\tilde{\Gamma}^{\sigma}{}_{\rho\mu}\,, \tag{A5}$$

one would obtain that $R^{\lambda}{}_{\mu\nu\rho} = 0$ at every point in spacetime since transforming into a different coordinate system cannot induce the curvature of spacetime.

Now one can apply EP (in this example specifically SEP) in order to couple the scalar field to gravity. The statement of SEP is that, in the presence of a gravitational field (i.e., in curved spacetime), the equation of motion for the scalar field should locally retain the same form as in the absence of the gravitational field (i.e., in flat spacetime). Since Equation (A2) depends only on the field at a given spacetime point and its first and second derivatives at the same point, the equation is in fact local—it is defined within an infinitesimal neighbourhood of a single point. Given this, EP states that the corresponding equation of motion in the presence of gravity should have precisely the same form:

$$\left(g^{\mu\nu}\nabla_{\mu}\nabla_{\nu} - m^2\right)\phi(x) = 0\,. \tag{A6}$$

The absence of the tilde now denotes the fact that the covariant derivative ∇_{μ} is defined in terms of a generic Levi-Civita connection $\Gamma^{\lambda}{}_{\mu\nu}$ which is in turn defined in terms of a generic metric $g_{\mu\nu}$, which does not necessarily satisfy (A4). In other words, EP postulates that the Equation (A6) now holds even in curved spacetime since for a generic metric and connection, the Riemann curvature tensor need not be equal to zero everywhere. The interaction between the scalar field and gravity, as postulated by EP and implemented in Equation (A6), is also known in the literature as the *minimal coupling* prescription [61].

In order to convince oneself that the preparation step of transforming (A1) to (A2) is trivial in the sense that it does not introduce any substantial modification of (A1), one can additionally demonstrate that (A6) is in fact locally equivalent to (A1) as well. Namely, according to a theorem in differential geometry (see for example the end of Chapter 85 in [62]), at any specific spacetime point x_0 one can choose the locally inertial coordinate system, in which the generic metric $g_{\mu\nu}$, the corresponding connection $\Gamma^{\lambda}{}_{\mu\nu}$ and consequently also the covariant derivative ∇_{μ} take their usual Minkowski values,

$$g_{\mu\nu}(x_0) = \eta_{\mu\nu}\,, \qquad \Gamma^{\lambda}{}_{\mu\nu}(x_0) = 0\,, \qquad \nabla_{\mu}\Big|_{x=x_0} = \partial_{\mu}\,, \tag{A7}$$

so that in the infinitesimal neighbourhood of the point x_0 Equation (A6) obtains the form precisely equal to (A1).

However, note that when *integrating* (A6), one must take into account that spacetime is curved since integration is a nonlocal operation, and the locally inertial coordinate system cannot eliminate spacetime curvature. Therefore, the *solutions* of (A6) will in general be *different* from solutions of (A1), indicating the physical interaction of the scalar field with gravity, despite the fact that the form of the equation of motion is identical in both cases.

Another thing that should be emphasised is that EP itself is not a mathematical theorem, but rather a principle with physical content, since it can be either satisfied or violated. Specifically, we could have prescribed a different coupling of the scalar field to gravity, such that in curved spacetime its equation of motion takes for example the form

$$\left(g^{\mu\nu}\nabla_{\mu}\nabla_{\nu} - m^2 + R^2 + K^2\right)\phi(x) = 0\,, \tag{A8}$$

where $R \equiv R^{\mu\nu}{}_{\mu\nu}$ and $K \equiv R^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma}$ are the curvature scalar and Kretschmann invariant, respectively. This equation is not equivalent to (A2) and there is no coordinate system in which it can take the form (A1) since R and K are invariants. In this sense, (A8) is an example of a scalar field coupled to gravity such that EP is violated. This type of interaction between matter and gravity is also known in the literature as *non-minimal coupling* [61].

Finally, we should note that the transformation from (A1) to (A2) amounts to what is also known in the literature as *symmetry localisation* [61]. In particular, one can verify that (A1) remains invariant with respect to the group $\mathbb{R}^4$ of global translations,

$$x^\mu \to \tilde{x}^\mu = x^\mu + \zeta^\mu \,, \qquad (\zeta \in \mathbb{R}^4)\,, \tag{A9}$$

while (A2) remains invariant with respect to the group $Diff(\mathbb{R}^4)$ of spacetime diffeomorphisms, obtained by localisation of the translational symmetry group,

$$x^\mu \to \tilde{x}^\mu = x^\mu + \zeta^\mu(x) \equiv \tilde{x}^\mu(x)\,, \tag{A10}$$

which represent general curvilinear coordinate transformations, used in (A4). One can explicitly verify that all three Equations (A2), (A6) and (A8) remain invariant with respect to local translations (A10) while describing no coupling to gravity, coupling to gravity that satisfies EP, and coupling to gravity that violates EP, respectively. In this sense, contrary to a common misconception (often stated in the literature) that symmetry localisation gives rise to interactions, one can say that the process of symmetry localisation *does not* introduce nor prescribe interactions in any way whatsoever. In particular, a direct counterexample is the Equation (A4), which manifestly *does* obey local translational symmetry, while it *does not* give rise to any interaction whatsoever (see below for the analogous counterexample in electrodynamics).

Appendix A.2. The Electromagnetic Case

Let us begin with an example of a Dirac field in Minkowski spacetime, and apply the generalised equivalence principle by coupling it to the EM field. The equation of motion in this case is the ordinary Dirac equation,

$$(i\gamma^\mu \partial_\mu - m)\psi(x) = 0\,, \tag{A11}$$

where γ^μ are standard Dirac gamma matrices, satisfying the anticommutator identity of the Clifford algebra $\{\gamma^\mu, \gamma^\nu\} = -2\eta^{\mu\nu}$. As it stands, Equation (A11) describes the free Dirac field, not coupled to an EM field in any way. Note that it is invariant with respect to global $U(1)$ transformations, defined as

$$\psi \to \psi' = e^{-i\lambda}\psi\,, \qquad e^{-i\lambda} \in U(1)\,, \qquad \lambda \in \mathbb{R}\,. \tag{A12}$$

In order to couple it to standard Maxwell electrodynamics, we first rewrite this equation into a form which is invariant with respect to local $U(1)$ transformations,

$$\psi \to \psi' = e^{-i\lambda(x)}\psi\,, \qquad \partial_\mu \to \tilde{\mathcal{D}}_\mu = \partial_\mu + i\partial_\mu\lambda(x)\,, \tag{A13}$$

so that the equation takes the form

$$(i\gamma^\mu \tilde{\mathcal{D}}_\mu - m)\psi(x) = 0\,, \tag{A14}$$

Note that here, $\tilde{\mathcal{D}}$ denotes the covariant derivative with respect to the "pure gauge" connection

$$\tilde{A}_\mu \equiv \partial_\mu\lambda(x)\,, \tag{A15}$$

where $\lambda(x)$ denotes the arbitrary gauge function. Moreover, note that (A11) is analogous to (A1), (A14) is analogous to (A2), while the global and local $U(1)$ gauge transformations (A12) and (A13) are EM analogues of the global and local spacetime translations (A9) and (A10) from the gravitational case. Finally, note that if one were to evaluate the electromagnetic Faraday field strength tensor using $\tilde{A}_\mu$ from (A15), according to the equation

$$F_{\mu\nu} = \partial_\mu \tilde{A}_\nu - \partial_\nu \tilde{A}_\mu\,, \tag{A16}$$

one would obtain that $F_{\mu\nu} = 0$ at every point in spacetime since the potential that is a pure gauge cannot induce an EM field. Here (A16) is analogous to (A5).

Once the Dirac equation in the form (A14) is in hand, one can apply the electromagnetic generalisation of EP in order to couple the Dirac field to an EM field. The statement of EP, in this case, is that in the presence of an EM field, the equation of motion for the Dirac field should locally retain the same form as in the absence of the EM field. Since Equation (A14) depends only on the field at a given spacetime point and its first derivatives at the same point, it is therefore defined within an infinitesimal neighbourhood of a single point—in other words, it is local. Given this, electromagnetic EP states that the corresponding equation of motion in the presence of EM field should have precisely the same form (the analogue of (A6)):

$$\big(i\gamma^\mu \mathcal{D}_\mu - m\big)\psi(x) = 0\,. \tag{A17}$$

The absence of the tilde now denotes the fact that the covariant derivative $\mathcal{D}_\mu \equiv \partial_\mu + iA_\mu$ is defined in terms of a generic $U(1)$ connection A_μ which does not necessarily satisfy (A15), but does obey the usual gauge transformation rule,

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu \lambda(x)\,. \tag{A18}$$

In other words, electromagnetic EP postulates that the Equation (A17) holds even in the presence of an EM field since for a generic connection A_μ the Faraday tensor may not be equal to zero everywhere. The interaction between the Dirac field and the EM field as postulated by the electromagnetic EP and implemented in Equation (A17) is again known in the literature as the *minimal coupling* prescription [61,63].

If one wishes to convince oneself that the preparation step of transforming (A11) to (A14) is trivial in the sense that it does not introduce any substantial modification of (A11), one can additionally demonstrate that (A17) is in fact locally equivalent to (A11). To do this, at any specific spacetime point x_0 one can choose the following $U(1)$ gauge,

$$\lambda(x) = -A_\mu(x_0)x^\mu\,, \tag{A19}$$

so that, according to (A18)

$$A'_\mu(x) = A_\mu(x) - \partial_\mu\big(A_\nu(x_0)x^\nu\big) \qquad \Rightarrow \qquad A'_\mu(x_0) = 0\,, \qquad \mathcal{D}_\mu\Big|_{x=x_0} = \partial_\mu\,. \tag{A20}$$

This choice of gauge is the EM analogue of the choice of a locally inertial coordinate system (A7). Substituting this into the primed version of (A17) and evaluating the whole equation at $x = x_0$, it reduces precisely to the form (A11) in the infinitesimal neighbourhood at that point, despite the presence of nonzero EM field.

Again note that when *integrating* (A17), one must take into account that the EM field is nonzero since integration is a nonlocal operation, and the choice of gauge (A19) eliminates the EM potential from (A17) only at the point x_0, while the Faraday tensor is gauge invariant. Therefore, the *solutions* of (A17) will in general be *different* from solutions of (A11), indicating the physical interaction of the Dirac field with EM field, despite the fact that the form of the equation of motion for the Dirac field is identical in both cases.

As in the case of gravity, we should emphasise that the electromagnetic EP is not a mathematical theorem, but rather a principle with physical content, since it can be either satisfied of violated. Specifically, we could have prescribed a different coupling of the Dirac field to electrodynamics, such that in the presence of an EM field its equation of motion takes for example the form (analogue of (A8))

$$\big(i\gamma^\mu \mathcal{D}_\mu - m + I_1 + I_2\big)\psi(x) = 0\,, \tag{A21}$$

where $I_1 \equiv F^{\mu\nu}F_{\mu\nu}$ and $I_2 \equiv \varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}$ are the two fundamental invariants of the Faraday tensor. This equation is not equivalent to (A14), and there exists no local $U(1)$ gauge in which it could take the form (A11), since I_1 and I_2 are invariants. In this sense, (A21) is an example of a Dirac field coupled to the EM field such that the electromagnetic EP is violated. This is also known in the literature as *non-minimal coupling* [61,63].

Finally, we should also note that the transformation from (A11) to (A14) amounts to what is also known in the literature as *symmetry localisation* [61,63]. Specifically, one can explicitly verify that all three Equations (A14), (A17) and (A21) remain invariant with respect to local $U(1)$ gauge transformations, while describing no coupling to an EM field, coupling to an EM field that satisfies the electromagnetic EP, and coupling to an EM field that violates electromagnetic EP, respectively. In this sense, one can again say that the process of symmetry localisation *does not* introduce nor prescribe interactions in any way whatsoever. In the case of electrodynamics and other gauge theories, this is quite often misrepresented in literature—the step of symmetry localisation is silently joined together with the step of applying the electromagnetic version of EP; thus, in the end, giving rise to an interacting theory, and the resulting presence of the interaction is then mistakenly attributed to the localisation of symmetry, rather than to the application of EP. Similar to the gravitational case above, the equation of motion (A14) is an explicit counterexample to such an attribution, since it *does* have local $U(1)$ symmetry, but *does not* have any interaction with an EM field.

Appendix A.3. The Test Particle Case

The last topic we should address is the context in which the statement of electromagnetic EP is compatible with the existence of the Lorentz force law, acting on charged test particles. Namely, one often distinguishes the motion of a test particle in a gravitational field from a motion of a test particle in an EM field, by comparing the geodesic Equation (2)

$$u^{\mu}(\tau)\nabla_{\mu}u^{\lambda}(\tau) = 0\,, \tag{A22}$$

where u^{μ} is the 4-velocity of the test particle, with the Lorentz force equation

$$u^{\mu}(\tau)\nabla_{\mu}u^{\lambda}(\tau) = \frac{q}{m}F^{\lambda\rho}\,u_{\rho}(\tau)\,, \tag{A23}$$

where q/m is the charge-to-mass ratio of a test particle moving in an external EM field, described by the Faraday tensor $F_{\mu\nu}$. A typical conclusion one draws from this comparison is that the interaction with the EM field gives rise to a "real force", while the interaction with the gravitational field does not.

However, it is highly misleading to compare (A22) to (A23) in the first place. Namely, as we have discussed in detail in Section 4, in field theory the notion of a particle can be defined only approximately, and this applies equally for electrodynamics as well as for gravity. Specifically, given the example discussed above, of a Dirac field coupled to an EM field via Equation (A17), we have seen that in the infinitesimal neighbourhood of a given point x_0 one can completely gauge away any presence of the coupling to EM field from (A17). In this sense, the notion of a test particle that satisfies (A23) cannot be identified with an idealised point-particle, that has exactly zero size. Instead, the realistic test particle is a wave-packet configuration of a Dirac field (a kink), and as such has a small but nonzero size. As it evolves, the different parts of the wave-packet are subject to interaction with the EM potential A_{μ} at *different* points of spacetime, giving rise to an effective non-minimal coupling with the Faraday tensor $F_{\mu\nu}$. This is completely analogous to the case of a test particle with small but nonzero size interacting with spacetime curvature due to tidal forces—both effects are equally nonlocal since both kinks have nonzero size. On the other hand, a test particle that satisfies (A22) represents an idealised point-particle (a leading order approximation in the multipole expansion of the matter field), i.e., a kink which thus has precisely zero size.

In this sense, the Lorentz force Equation (A23) rather ought to be compared with the Papapetrou Equation (3),

$$u^{\mu}(\tau)\nabla_{\mu}u^{\lambda}(\tau) = R^{\lambda}{}_{\mu\rho\sigma}\,u^{\mu}(\tau)J^{\rho\sigma}(\tau)\,. \tag{A24}$$

Indeed, one can see quite a reasonable analogy between (A23) and (A24). There are of course small technical differences due to the precise nature of the coupling to various moments of the kink, but nevertheless, the two equations are strikingly similar. Given this, while one can still draw the conclusion that the interaction of a kink with the EM field gives rise to a "real force", one can draw precisely the same conclusion for the interaction of a kink with the gravitational field. There is no distinction between gravity and the other gauge interactions at this level—all four interactions in nature (strong, weak, electromagnetic and gravitational) are equally "real". In addition, all four interactions satisfy EP at the fundamental field theory level (i.e., in the sense of strong generalised EP), while at the level of mechanics, a corresponding weak generalised EP is manifestly violated in all four cases.

References

1. Kay, B.S. Decoherence of macroscopic closed systems within Newtonian quantum gravity. *Class. Quantum Gravity* **1998**, *15*, L89. [CrossRef]
2. Oniga, T.; Wang, C.H.T. Quantum gravitational decoherence of light and matter. *Phys. Rev. D* **2016**, *93*, 044027. [CrossRef]
3. Bruschi, D.E. On the weight of entanglement. *Phys. Lett. B* **2016**, *754*, 182. [CrossRef]
4. Bose, S.; Mazumdar, A.; Morley, G.W.; Ulbricht, H.; Toroš, M.; Paternostro, M.; Geraci, A.A.; Barker, P.F.; Kim, M.S.; Milburn, G. Spin entanglement witness for quantum gravity. *Phys. Rev. Lett.* **2017**, *119*, 240401. [CrossRef]
5. Marletto, C.; Vedral, V. Gravitationally Induced Entanglement between Two Massive Particles is Sufficient Evidence of Quantum Effects in Gravity. *Phys. Rev. Lett.* **2017**, *119*, 240402.
6. Marletto, C.; Vedral, V. When can gravity path-entangle two spatially superposed masses? *Phys. Rev. D* **2018**, *98*, 046001. [CrossRef]
7. Paunković, N.; Vojinović, M. Gauge protected entanglement between gravity and matter. *Class. Quantum Gravity* **2018**, *35*, 185015. [CrossRef]
8. Oreshkov, O.; Costa, F.; Brukner, Č. Quantum correlations with no causal order. *Nat. Commun.* **2012**, *3*, 1092. [CrossRef]
9. Araújo, M.; Branciard, C.; Costa, F.; Feix, A.; Giarmatzi, C.; Brukner, Č. Witnessing causal nonseparability. *New J. Phys.* **2015**, *17*, 102001. [CrossRef]
10. Vilasini, V. An Introduction to Causality in Quantum Theory (and Beyond). Master's Thesis, ETH, Zürich, Switzerland, 2017.
11. Oreshkov, O. Time-delocalized quantum subsystems and operations: On the existence of processes with indefinite causal structure in quantum mechanics. *Quantum* **2019**, *3*, 206. [CrossRef]
12. Paunković, N.; Vojinović, M. Causal orders, quantum circuits and spacetime: Distinguishing between definite and superposed causal orders. *Quantum* **2020**, *4*, 275. [CrossRef]
13. Vilasini, V.; Colbeck, R. General framework for cyclic and fine-tuned causal models and their compatibility with space-time. *Phys. Rev. A* **2022**, *106*, 032204. [CrossRef]
14. Vilasini, V.; Renner, R. Embedding cyclic causal structures in acyclic spacetimes: No-go results for process matrices. *arXiv* **2022**, arXiv:2203.11245.
15. Ormrod, N.; Vanrietvelde, A.; Barrett, J. Causal structure in the presence of sectorial constraints, with application to the quantum switch. *arXiv* **2022**, arXiv:2204.10273.
16. Giacomini, F.; Castro-Ruiz, E.; Brukner, Č. Quantum mechanics and the covariance of physical laws in quantum reference frames. *Nat. Commun.* **2019**, *10*, 494. [CrossRef] [PubMed]
17. Vanrietvelde, A.; Höhn, P.A.; Giacomini, F.; Castro-Ruiz, E. A change of perspective: Switching quantum reference frames via a perspective-neutral framework. *Quantum* **2020**, *4*, 225. [CrossRef]
18. Krumm, M.; Höhn, P.A.; Müller, M.P. Quantum reference frame transformations as symmetries and the paradox of the third particle. *Quantum* **2021**, *5*, 530. [CrossRef]
19. Ahmad, S.A.; Galley, T.D.; Höhn, P.A.; Lock, M.P.E.; Smith, A.R.H. Quantum Relativity of Subsystems. *Phys. Rev. Lett.* **2022**, *128*, 170401. [CrossRef]
20. de la Hamette, A.C.; Kabel, V.; Castro-Ruiz, E.; Brukner, Č. Falling through masses in superposition: Quantum reference frames for indefinite metrics. *arXiv* **2021**, arXiv:2112.11473.
21. Colladay, D.; Kostelecký, V.A. Lorentz-violating extension of the standard model. *Phys. Rev. D* **1998**, *58*, 116002. [CrossRef]
22. Kostelecký, V.A.; Russell, N. Data tables for Lorentz and *CPT* violation. *Rev. Mod. Phys.* **2011**, *83*, 11–31.
23. Amelino-Camelia, G. Particle-dependent deformations of Lorentz symmetry. *Symmetry* **2012**, *4*, 344–378. [CrossRef]
24. Amelino-Camelia, G.; Palmisano, M.; Ronco, M.; D'Amico, G. Mixing coproducts for theories with particle-dependent relativistic properties. *Int. J. Mod. Phys. D* **2020**, *29*, 2050017. [CrossRef]
25. Torri, M.D.C.; Antonelli, V.; Miramonti, L. Homogeneously Modified Special relativity (HMSR). *Eur. Phys. J. C* **2019**, *79*, 808.
26. Pipa, F.; Paunković, N.; Vojinović, M. Entanglement-induced deviation from the geodesic motion in quantum gravity. *J. Cosmol. Astropart. Phys.* **2019**, *2019*, 57. [CrossRef]
27. Giacomini, F.; Brukner, Č. Einstein's Equivalence principle for superpositions of gravitational fields. *arXiv* **2020**, arXiv:2012.13754.

28. Giacomini, F.; Brukner, Č. Quantum superposition of spacetimes obeys Einstein's equivalence principle. *AVS Quantum Sci.* **2022**, *4*, 015601.
29. Marletto, C.; Vedral, V. On the testability of the equivalence principle as a gauge principle detecting the gravitational t^3 phase. *Front. Phys.* **2020**, *8*, 176.
30. Marletto, C.; Vedral, V. Sagnac interferometer and the quantum nature of gravity. *J. Phys. Commun.* **2021**, *5*, 051001.
31. Marletto, C.; Vedral, V. The quantum totalitarian property and exact symmetries. *AVS Quantum Sci.* **2022**, *4*, 015603. [CrossRef]
32. Einstein, A.; Infeld, L.; Hoffmann, B. The Gravitational Equations and the Problem of Motion. *Ann. Math.* **1938**, *39*, 65. [CrossRef]
33. Mathisson, M. Neue mechanik materieller systemes. *Acta Phys. Pol.* **1937**, *6*, 163.
34. Papapetrou, A. Spinning test-particles in general relativity, I. *Proc. R. Soc. A* **1951**, *209*, 248.
35. Tulczyjev, W. Equations of motion of rotating bodies in general relativity theory. *Acta Phys. Pol.* **1959**, *18*, 393.
36. Taub, A.H. Motion of Test Bodies in General Relativity. *J. Math. Phys.* **1964**, *5*, 112. [CrossRef]
37. Dixon, G. A covariant multipole formalism for extended test bodies in general relativity. *Nuovo Cim.* **1964**, *34*, 317. [CrossRef]
38. Dixon, G. Classical theory of charged particles with spin and the classical limit of the Dirac equation. *Nuovo Cim.* **1965**, *38*, 1616. [CrossRef]
39. Dixon, G. Dynamics of extended bodies in general relativity. I. Momentum and angular momentum. *Proc. R. Soc. A* **1970**, *314*, 499.
40. Dixon, G. Dynamics of extended bodies in general relativity - II. Moments of the charge-current vector. *Proc. R. Soc. A* **1970**, *319*, 509.
41. Dixon, G. The definition of multipole moments for extended bodies. *Gen. Relativ. Gravit.* **1973**, *4*, 199. [CrossRef]
42. Yasskin, P.B.; Stoeger, W.R. Propagation equations for test bodies with spin and rotation in theories of gravity with torsion. *Phys. Rev. D* **1980**, *21*, 2081. [CrossRef]
43. Nomura, K.; Shirafuji, T.; Hayashi, K. Spinning Test Particles in Spacetime with Torsion. *Prog. Theor. Phys.* **1991**, *86*, 1239. [CrossRef]
44. Nomura, K.; Shirafuji, T.; Hayashi, K. Semiclassical particles with arbitrary spin in the Riemann-Cartan space-time. *Prog. Theor. Phys.* **1992**, *87*, 1275. [CrossRef]
45. Vasilić, M.; Vojinović, M. Classical spinning branes in curved backgrounds. *JHEP* **2007**, *7*, 28. [CrossRef]
46. Vasilić, M.; Vojinović, M. Spinning branes in Riemann-Cartan spacetime. *Phys. Rev. D* **2008**, *78*, 104002. [CrossRef]
47. Accioly, A.; Paszko, R. Conflict between the Classical Equivalence Principle and Quantum Mechanics. *Adv. Stud. Theor. Phys.* **2009**, *3*, 65.
48. Longhi, S. Equivalence principle and quantum mechanics: Quantum simulation with entangled photons. *Opt. Lett.* **2018**, *43*, 226. [CrossRef]
49. Chowdhury, P.; Home, D.; Majumdar, A.S.; Mousavi, S.V.; Mozaffari, M.R.; Sinha, S. Strong quantum violation of the gravitational weak equivalence principle by a non-Gaussian wave packet. *Class. Quantum Gravity* **2012**, *29*, 025010. [CrossRef]
50. Rosi, G.; D'Amico, G.; Cacciapuoti, L.; Sorrentino, F.; Prevedelli, M.; Zych, M.; Brukner, Č.; Tino, G.M. Quantum test of the equivalence principle for atoms in coherent superposition of internal energy states. *Nat. Commun.* **2017**, *8*, 15529. [CrossRef]
51. Zych, M.; Brukner, Č. Quantum formulation of the Einstein Equivalence Principle. *Nat. Phys.* **2018**, *14*, 1027. [CrossRef]
52. Anastopoulos, C.; Hu, B.L. Equivalence principle for quantum systems: Dephasing and phase shift of free-falling particles. *Class. Quantum Gravity* **2018**, *35*, 035011. [CrossRef]
53. Hardy, L. Implementation of the Quantum Equivalence Principle. In *Progress and Visions in Quantum Theory in View of Gravity*; Finster, F., Giulini, D., Kleiner, J., Tolksdorf, J., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 189–220.
54. Kretschmann, E. Über den physikalischen Sinn der Relativitätspostulate, A. Einsteins neue und seine ursprüngliche Relativitätstheorie. *Ann. Phys.* **1918**, *358*, 575–614.
55. Misner, C.W.; Thorne, K.S.; Wheeler, J.A. *Gravitation*; W. H. Freeman and Co.: San Francisco, CA, USA, 1973.
56. Okon, E.; Callender, C. Does Quantum Mechanics Clash with the Equivalence Principle—And Does it Matter? *Eur. J. Phil. Sci.* **2011**, *1*, 133. [CrossRef]
57. Casola, E.D.; Liberati, S.; Sonego, S. Nonequivalence of equivalence principles. *Am. J. Phys.* **2015**, *83*, 39. [CrossRef]
58. Viola, L.; Onofrio, R. Testing the equivalence principle through freely falling quantum objects. *Phys. Rev. D* **1997**, *55*, 455–462. [CrossRef]
59. Plyatsko, R. Gravitational ultrarelativistic spin-orbit interaction and the weak equivalence principle. *Phys. Rev. D* **1998**, *58*, 084031.
60. Bose, S.; Mazumdar, A.; Schut, M.; Toroš, M. Entanglement witness for the weak equivalence principle. *arXiv* **2022**, arXiv:2203.11628.
61. Blagojević, M. *Gravitation and Gauge Symmetries*; Institute of Physics Publishing: Bristol, UK, 2002.
62. Landau, L.D.; Lifshitz, E.M. *The Classical Theory of Fields*, 4th ed.; Butterworth-Heinemann: Oxford, UK, 1980.
63. Peskin, M.E.; Schroeder, D.V. *An Introduction to Quantum Field Theory*; Addison-Wesley Publishing Co.: Boston, MA, USA, 1995.

Published for SISSA by Springer

Received: *January 17, 2022*
Accepted: *June 9, 2022*
Published: *July 18, 2022*

Topological invariant of 4-manifolds based on a 3-group

T. Radenković[1] and M. Vojinović

Institute of Physics, University of Belgrade,
Pregrevica 118, 11080 Belgrade, Serbia

E-mail: rtijana@ipb.ac.rs, vmarko@ipb.ac.rs

Abstract: We study a generalization of 4-dimensional BF-theory in the context of higher gauge theory. We construct a triangulation independent topological state sum Z, based on the classical $3BF$ action for a general 3-group and a 4-dimensional spacetime manifold $\mathcal{M}_4$. This state sum coincides with Porter's TQFT for $d = 4$ and $n = 3$. In order to verify that the constructed state sum is a topological invariant of the underlying 4-dimensional manifold, its behavior under Pachner moves is analyzed, and it is obtained that the state sum Z remains the same. This paper is a generalization of the work done by Girelli, Pfeiffer, and Popescu for the case of state sum based on the classical $2BF$ action with the underlying 2-group structure.

Keywords: Differential and Algebraic Geometry, Models of Quantum Gravity, Topological Field Theories, Gauge Symmetry

ArXiv ePrint: 2201.02572

[1]Corresponding author.

https://doi.org/10.1007/JHEP07(2022)105

Contents

JHEP07(2022)105

1 Introduction

Within the Loop Quantum Gravity framework, one studies the nonperturbative quantization of gravity, both canonically and covariantly, see [1–4] for an overview and a comprehensive introduction. The covariant approach focuses on defining the path integral for the gravitational field by considering a triangulation of a spacetime manifold and specifying the path integral as a discrete state sum of the gravitational field configurations living on the simplices in the triangulation. This quantization technique is usually referred to as the *spinfoam quantization method*, and it can be divided into three major steps:

1. first, one writes the classical action $S[g]$ as a topological BF-like action plus simplicity constraints,

2. then one uses the algebraic structure underlying the topological sector of the action to define a topological state sum Z,

3. and finally, one deforms the topological state sum by imposing simplicity constraints, thus promoting it into a path integral for a physical theory.

Spinfoam models for gravity are usually constructed by constraining the topological gauge theory known as BF theory, obtaining the Plebanski formulation of general relativity [5]. For example, in 3 dimensions, the prototype spinfoam model is known as the Ponzano-Regge model [6]. In 4 dimensions there are multiple models, such as the Barrett-Crane model [7, 8], the Ooguri model [9], and the most sophisticated EPRL/FK model [10, 11] (see also [12–14]). All these models aim to define a viable theory of a quantum gravitational field alone, without matter fields. The attempts to include matter fields have had limited success [15], mainly because the mass terms cannot be expressed in the theory due to the absence of the tetrad fields from the topological BF sector of the theory.

In order to overcome this problem, a new approach has been developed within the framework of *higher gauge theory* (for a review of higher gauge theory, see [16, 17], and for its applications in physics see [18–29]). Within higher gauge theory formalism, one generalizes the BF action, based on some Lie group, to an $2BF$ action based on the 2-group structure. Within this approach [30], one rewrites the action for general relativity as a constrained $2BF$ action, such that the tetrad fields are present in the topological sector. This result opened up the possibility to couple all matter fields to gravity in a straightforward way. Nevertheless, the matter fields could not be naturally expressed using the underlying algebraic structure of a 2-group, rendering the spinfoam quantization method only half-implementable, since the matter sector of the classical action could not be expressed as a topological term plus a simplicity constraint, which means that the steps 2 and 3 above could not be performed for the matter sector of the action.

This final issue has recently been resolved in [31], where one more step in the categorical ladder is performed in order to generalize the underlying algebraic structure from a 2-group to a 3-group (see also [32] for the 4-group formulation). This generalization then naturally gives rise to the so-called $3BF$ action, which proves to be suitable for a unified description of both gravity and matter fields. The first step of the spinfoam quantization program is carried out in [31] where the suitable gauge 3-groups have been specified, and the corresponding constrained $3BF$ actions constructed so that the desired classical dynamics of the gravitational and matter fields are obtained. A reader interested in the construction of the constrained $2BF$ actions describing the Yang-Mills field and Einstein-Cartan gravity, and $3BF$ actions describing the Klein-Gordon, Dirac, Weyl, and Majorana fields, each coupled to gravity in the standard way, is referred to [30, 31].

In this paper, we focus our attention on the second step of the spinfoam quantization program: we will construct a triangulation independent topological state sum Z, based on the classical $3BF$ action for a general 3-group and a 4-dimensional spacetime manifold $\mathcal{M}_4$. This state sum coincides with Porter's TQFT [33, 34] for $d = 4$ and $n = 3$. In order to verify that the constructed state sum is topological, we analyze its behavior under Pachner moves [35]. Pachner moves are local changes of a triangulation that preserve topology, such that any two triangulations of the same manifold are connected by a finite number of Pachner moves. In 4 dimensions, there are five different Pachner moves: the $3-3$ move,

$4-2$ move, and $5-1$ move, and their inverses. After defining the state sum, we calculate its behavior under these Pachner moves. We obtain that the state sum Z remains the same, proving that it is a topological invariant of the underlying 4-dimensional manifold. This construction thus completes the second step of the quantization procedure. Our result paves the way for the third step of the covariant quantization procedure and a formulation of a quantum theory of gravity and matter by imposing the simplicity constraints on the state sum. We leave the third step for future work.

The layout of the paper is as follows. In section 2 we review the pure and the constrained nBF theories describing some of the physically relevant models — the constrained $2BF$ actions describing the Yang-Mills field and Einstein-Cartan gravity, and constrained $3BF$ actions describing the Klein-Gordon and Dirac fields coupled to Yang-Mills fields and gravity in the standard way. In section 3, we review the relevant algebraic tools involved in the description of higher gauge theory, 2-crossed modules, and 3-gauge theory. Starting from the notion of Lie 3-groups, we generalize the integral picture of gauge theory to a 3-gauge theory that involves curves, surfaces, and volumes labeled with elements of non-Abelian groups. In section 4, we define the discrete state sum model of topological higher gauge theory in dimension $d=4$. The model is defined for any closed and oriented combinatorial 4-dimensional manifold $\mathcal{M}_4$. The proof that the state sum is invariant under the Pachner moves and thus independent of the chosen triangulation is presented in appendix B.

Notations and conventions throughout the paper are as follows. The local Lorentz indices are denoted by the Latin letters $a, b, c, \ldots$, that take values $0,1,2,3$, and are raised and lowered using the Minkowski metric η_{ab} with signature $(-,+,+,+)$. The spacetime indices are denoted by the Greek letters $\mu, \nu, \ldots$, and are raised and lowered by the spacetime metric $g_{\mu\nu} = \eta_{ab} e^a{}_\mu e^b{}_\nu$, where $e^a{}_\mu$ denotes the tetrad fields. If G is a finite group, $\int_G dg = 1/|G| \sum_{g\in G}$ denotes the normalized sum over all group elements, while δ_G denotes the corresponding δ-distribution on G. The δ-distribution is defined for every element $g \in G$ such that $\delta_G(g) = |G|$ if g is the unit element of the group, i.e., $g = e$, and $\delta_G(g) = 0$ if it is not, i.e., $g \neq e$. If G is a Lie group, $\int_G dg$ and δ_G denote the Haar measure and the δ-distribution on G, respectively. The set of all k-simplices, $0 \leq k \leq d$, is denoted by Λ_k. The set of vertices Λ_0 is finite and ordered, and every k-simplex is labeled by $(k+1)$-tuples of vertices $(i_0 \ldots i_k)$, where $i_0, \ldots, i_k \in \Lambda_0$ such that $i_0 < \cdots < i_k$.

2 Review of the classical theory

2.1 Topological nBF theories

For a given Lie group G whose Lie algebra $\mathfrak{g}$ is equipped with the G-invariant symmetric nondegenerate bilinear form $\langle \_\,, \_\rangle_{\mathfrak{g}}$, and for a given 4-dimensional spacetime manifold $\mathcal{M}_4$, one can introduce the BF action as

$$S_{BF} = \int_{\mathcal{M}_4} \langle B \wedge F \rangle_{\mathfrak{g}} \,, \tag{2.1}$$

where 2-form $F \equiv \mathrm{d}\alpha + \alpha \wedge \alpha$ is the curvature for the $\mathfrak{g}$-valued connection 1-form $\alpha \in \mathcal{A}^1(\mathcal{M}_4\,, \mathfrak{g})$ and 2-form $B \in \mathcal{A}^2(\mathcal{M}_4\,, \mathfrak{g})$ is an $\mathfrak{g}$-valued Lagrange multiplier. Varying the

JHEP07(2022)105

action (2.1) with respect to the Lagrange multiplier B and the connection α, one obtains the equations of motion of the theory,

$$F = 0\,, \qquad \nabla B \equiv \mathrm{d}B + \alpha \wedge B = 0\,. \tag{2.2}$$

From the first equation of motion, one sees that α is a flat connection, which then, together with the second equation of motion, implies that B is constant. Therefore, the theory given by the BF action has no local propagating degrees of freedom, i.e., the theory is topological. For more details about the BF theory see [5, 36, 37].

Within the framework of Higher Gauge Theory, by passing from the notion of a gauge group to the notion of a gauge 2-group, one defines the categorical generalization of the BF action, called the $2BF$ action. A 2-group has a naturally associated notion of a 2-connection $(\alpha\,,\beta)$, described by the usual $\mathfrak{g}$-valued 1-form $\alpha \in \mathcal{A}^1(\mathcal{M}_4\,,\mathfrak{g})$ and an $\mathfrak{h}$-valued 2-form $\beta \in \mathcal{A}^2(\mathcal{M}_4\,,\mathfrak{h})$, where $\mathfrak{h}$ is a Lie algebra of the Lie group H. The 2-connection gives rise to the so-called fake 2-curvature $(\mathcal{F},\mathcal{G})$, where $\mathcal{F}$ is a $\mathfrak{g}$-valued fake curvature 2-form $\mathcal{F} \in \mathcal{A}^2(\mathcal{M}_4\,,\mathfrak{g})$ and $\mathcal{G}$ is an $\mathfrak{h}$-valued curvature 3-form $\mathcal{G} \in \mathcal{A}^3(\mathcal{M}_4\,,\mathfrak{h})$, defined as

$$\mathcal{F} = \mathrm{d}\alpha + \alpha \wedge \alpha - \partial\beta\,, \qquad \mathcal{G} = \mathrm{d}\beta + \alpha \wedge^{\triangleright} \beta\,. \tag{2.3}$$

Representing the 2-group as a crossed-module $(H \stackrel{\partial}{\to} G\,, \triangleright)$, and seeing the next section for the definition and notation, one introduces a $2BF$ action using the fake 2-curvature (2.3) as

$$S_{2BF} = \int_{\mathcal{M}_4} \langle B \wedge \mathcal{F} \rangle_{\mathfrak{g}} + \langle C \wedge \mathcal{G} \rangle_{\mathfrak{h}}\,, \tag{2.4}$$

where the 2-form $B \in \mathcal{A}^2(\mathcal{M}_4\,,\mathfrak{g})$ and the 1-form $C \in \mathcal{A}^1(\mathcal{M}_4\,,\mathfrak{h})$ are Lagrange multipliers, and $\langle \_\,,\_ \rangle_{\mathfrak{g}}$ and $\langle \_\,,\_ \rangle_{\mathfrak{h}}$ denote the G-invariant symmetric nondegenerate bilinear forms for the algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively. Similarly as in the case of the BF theory, varying the $2BF$ action (2.4) with respect to the Lagrange multipliers B and C one obtains the equations of motion,

$$\mathcal{F} = 0\,, \qquad \mathcal{G} = 0\,, \tag{2.5}$$

i.e., the conditions that the curvature 2-form $\mathcal{F}$ and the curvature 3-form $\mathcal{G}$ vanish, while varying with respect to the connections α and β one obtains

$$\nabla B + C \wedge^{\mathcal{T}} \beta = 0\,, \qquad \nabla C - \partial(B) = 0\,. \tag{2.6}$$

Similar to the case of the BF action, the $2BF$ action defines a topological theory, i.e., a theory with no propagating degrees of freedom, see [38–41] for review and references.

Continuing the categorical ladder one step further, one can generalize the $2BF$ action to the $3BF$ action, by passing from the notion of a 2-group to the notion of a 3-group. Representing the 3-group with a 2-crossed module $(L \stackrel{\delta}{\to} H \stackrel{\partial}{\to} G\,, \triangleright\,, \{\_\,,\_\}_{\mathrm{p}})$, and seeing next section for definition and notation, one can define a 3-connection as an ordered triple (α, β, γ), where α, β, and γ are appropriate algebra-valued differential forms, $\alpha \in \mathcal{A}^1(\mathcal{M}_4, \mathfrak{g})$, $\beta \in \mathcal{A}^2(\mathcal{M}_4, \mathfrak{h})$, and $\gamma \in \mathcal{A}^3(\mathcal{M}_4, \mathfrak{l})$. The corresponding fake 3-curvature

$(\mathcal{F},\mathcal{G},\mathcal{H})$ is defined as:

$$\begin{aligned}\mathcal{F} &= \mathrm{d}\alpha + \alpha\wedge\alpha - \partial\beta\,, \qquad\qquad \mathcal{G} = \mathrm{d}\beta + \alpha\wedge^{\rhd}\beta - \delta\gamma\,,\\ \mathcal{H} &= \mathrm{d}\gamma + \alpha\wedge^{\rhd}\gamma + \{\beta\wedge\beta\}_{\mathrm{p}}\,.\end{aligned} \tag{2.7}$$

Then, similar to the construction of BF and $2BF$ actions, one defines the $3BF$ action as

$$S_{3BF} = \int_{\mathcal{M}_4} \langle B\wedge\mathcal{F}\rangle_{\mathfrak{g}} + \langle C\wedge\mathcal{G}\rangle_{\mathfrak{h}} + \langle D\wedge\mathcal{H}\rangle_{\mathfrak{l}}\,, \tag{2.8}$$

where $\mathfrak{g}$, $\mathfrak{h}$, and $\mathfrak{l}$ denote the Lie algebras corresponding to the Lie groups G, H, and L and the forms $\langle\_,\_\rangle_{\mathfrak{g}}$, $\langle\_,\_\rangle_{\mathfrak{h}}$, and $\langle\_,\_\rangle_{\mathfrak{l}}$ are G-invariant symmetric nondegenerate bilinear forms on $\mathfrak{g}$, $\mathfrak{h}$, and $\mathfrak{l}$, respectively. The variables $B\in\mathcal{A}^2(\mathcal{M}_4,\mathfrak{g})$, $C\in\mathcal{A}^1(\mathcal{M}_4,\mathfrak{h})$, and $D\in\mathcal{A}^0(\mathcal{M}_4,\mathfrak{l})$ are Lagrange multipliers, and their associated equations of motion are the conditions that the 3-curvature $(\mathcal{F},\mathcal{G},\mathcal{H})$ vanishes,

$$\mathcal{F}=0\,,\qquad \mathcal{G}=0\,,\qquad \mathcal{H}=0\,. \tag{2.9}$$

Additionally, varying with respect to the 3-connection variables α, β, and γ one gets:

$$\nabla B + C\wedge^{\mathcal{T}}\beta - D\wedge^{\mathcal{S}}\gamma = 0\,, \tag{2.10}$$

$$\nabla C - \partial(B) - D\wedge^{(\chi_1+\chi_2)}\beta = 0\,, \tag{2.11}$$

$$\nabla D + \delta(C) = 0\,. \tag{2.12}$$

For further details see [22, 42, 43] for the definition of the 3-group, and [31] for the definition of the pure $3BF$ action.

All the above actions are topological, in the sense that they do not contain any local propagating degrees of freedom [44, 45]. In this sense, they are not very interesting for the description of realistic physics, which should feature nontrivial dynamics. Nevertheless, by choosing the convenient underlying 2-crossed module structure and imposing the appropriate simplicity constraints onto the degrees of freedom present in the $3BF$ action, one can obtain the nontrivial classical dynamics of the gravitational and matter fields, as we will see in the following subsection.

2.2 Models with relevant dynamics

Let us review how one can employ the n-group structure to introduce the topological nBF actions corresponding to gravity and matter fields, as well as the form of the appropriate simplicity constraints to be imposed on these fields to obtain the classical dynamics.

First we review the most important constrained $2BF$ actions. We begin by rewriting general relativity as a constrained $2BF$ action based on the underlying Poincaré 2-group. The Poincaré 2-group is equivalent to a crossed module $(H\xrightarrow{\partial}G,\rhd)$, where the groups are choosen as $G=\mathrm{SO}(3,1)$ and $H=\mathbb{R}^4$, and the map ∂ is trivial. The action $\rhd$ is a natural action of $\mathrm{SO}(3,1)$ on $\mathbb{R}^4$, defined as

$$M_{ab}\rhd P_c = \eta_{[bc}P_{a]}\,, \tag{2.13}$$

where M_{ab} and P_a are the generators of groups $\mathrm{SO}(3,1)$ and $\mathbb{R}^4$, respectively. The action $\triangleright$ of $\mathrm{SO}(3,1)$ on itself is given via conjugation, by definition of a crossed module. Then, Poincaré 2-group gives rise to the 2-connection (α, β), given by the algebra-valued differential forms

$$\alpha = \omega^{ab} M_{ab}\,, \qquad \beta = \beta^a P_a\,, \tag{2.14}$$

where we have interpreted the connection 1-form α^{ab} as the ordinary spin connection ω^{ab}. Also, the corresponding 2-curvature $(\mathcal{F}, \mathcal{G})$ is given as

$$\begin{aligned} \mathcal{F} &= (\mathrm{d}\omega^{ab} + \omega^a{}_c \wedge \omega^{cb}) M_{ab} \equiv R^{ab} M_{ab}\,, \\ \mathcal{G} &= (\mathrm{d}\beta^a + \omega^a{}_b \wedge \beta^b) P_a \qquad \equiv \nabla\beta^a P_a \equiv G^a P_a\,, \end{aligned} \tag{2.15}$$

where we can recognize the standard Riemann curvature 2-form R^{ab} in $\mathcal{F}$. Having these variables in hand, one defines $2BF$ action (2.4) for the Poincaré 2-group as

$$S_{2BF} = \int_{\mathcal{M}_4} B^{ab} \wedge R_{ab} + e_a \wedge \nabla\beta^a\,. \tag{2.16}$$

Here, the crucial insight is that the Lagrange multiplier fields C^a can be identified with the tetrads [30], since one can show that 1-forms C^a transform in the same way as the tetrad 1-forms e^a under the Lorentz transformations and diffeomorphisms. One can now construct the action for general relativity by simply adding the additional simplicity constraint term to the action (2.16):

$$S = \int_{\mathcal{M}_4} B^{ab} \wedge R_{ab} + e_a \wedge \nabla\beta^a - \lambda_{ab} \wedge \left(B^{ab} - \frac{1}{16\pi l_p^2} \varepsilon^{abcd} e_c \wedge e_d \right). \tag{2.17}$$

Here λ_{ab} is a Lagrange multiplier 2-form associated to the simplicity constraint term, and l_p is the Planck length. It is straightforward to show that the corresponding equations of motion reduce to vacuum Einstein field equations. Thus the action (2.17) is classically equivalent to general relativity. The construction of the action (2.17) is analogous to the Plebanski model, where general relativity is constructed by adding a simplicity constraint to the BF theory based on the Lorentz group. However, one clear advantage of this model over the Plebanski model is that the tetrads are explicitly present in the topological sector of the action. Upon the covariant quantization, tetrads are therefore fundamental, off-shell quantities, in contrast to the Plebanski model where they appear only on-shell, as solutions of the classical equations of motion. The off-shell presence of the tetrads facilitates the straightforward coupling of the matter fields to gravity, and thus overcomes the problems present in the spinfoam models [15].

The Poincaré 2-group can be easily extended to include the coupling of the $\mathrm{SU}(N)$ Yang-Mills fields to gravity [31]. To achieve this, one constructs the crossed module $(H \xrightarrow{\partial} G, \triangleright)$, where the groups are chosen as $G = \mathrm{SO}(3,1) \times \mathrm{SU}(N)$ and $H = \mathbb{R}^4$, while the map ∂ remains trivial, as before. The action $\triangleright$ of the group G on H is such that the $\mathrm{SO}(3,1)$ subgroup acts on $\mathbb{R}^4$ via the vector representation (2.13), while the action of the $\mathrm{SU}(N)$ subgroup is trivial,

$$\tau_I \triangleright P_a = 0\,, \tag{2.18}$$

where τ_I are the SU(N) generators. This crossed module yields the 2-connection (α, β), where algebra-valued 1-form α and algebra valued 2-form β are defined as follows,

$$\alpha = \omega^{ab} M_{ab} + A^I \tau_I \,, \qquad \beta = \beta^a P_a \,, \tag{2.19}$$

where we can identify the gauge connection 1-form A^I. This connection gives rise to the 2-curvature $(\mathcal{F}, \mathcal{G})$, where $\mathcal{F}$ as defined as

$$\mathcal{F} = R^{ab} M_{ab} + F^I \tau_I \,, \qquad F^I \equiv \mathrm{d}A^I + f_{JK}{}^I A^J \wedge A^K \,, \tag{2.20}$$

while the curvature $\mathcal{G}$ for β remains the same as before. Given these variables, the Lagrange multiplier B in the first term of the topological action (2.4) also splits into two pieces corresponding to the direct product of the group G, giving

$$S_{2BF} = \int_{\mathcal{M}_4} B^{ab} \wedge R_{ab} + B^I \wedge F_I + e_a \wedge \nabla \beta^a \,, \tag{2.21}$$

where 2-form $B^I \in \mathcal{A}^2(\mathcal{M}_4 \,, \mathfrak{su}(N))$ is the second piece of the Lagrange multiplier. To obtain the non-trivial classical dynamics for gravity and the Yang-Mills field, we add the appropriate simplicity constraint terms to the action (2.21), and construct the constrained $2BF$ action:

$$\begin{aligned} S = \int_{\mathcal{M}_4} & B^{ab} \wedge R_{ab} + B^I \wedge F_I + e_a \wedge \nabla \beta^a - \lambda_{ab} \wedge \left(B^{ab} - \frac{1}{16\pi l_p^2} \varepsilon^{abcd} e_c \wedge e_d \right) \\ & + \lambda^I \wedge \left(B_I - \frac{12}{g} M_{abI} e^a \wedge e^b \right) + \zeta^{abI} \left(M_{abI} \varepsilon_{cdef} e^c \wedge e^d \wedge e^e \wedge e^f - g_{IJ} F^J \wedge e_a \wedge e_b \right) . \end{aligned} \tag{2.22}$$

Here, the first row is the topological sector and the familiar simplicity constraint for gravity from (2.17), while the second row contains the appropriate simplicity constraints for Yang Mills field, featuring the Lagrange multipliers λ^I and ζ^{abI}. The action (2.22) provides two dynamical equations — the equation for A^I,

$$\nabla_\rho F^{I\rho\mu} \equiv \partial_\rho F^{I\rho\mu} + \Gamma^\rho{}_{\lambda\rho} F^{I\lambda\mu} + f_{JK}{}^I A^J{}_\rho F^{K\rho\mu} = 0 \,, \tag{2.23}$$

where $\Gamma^\lambda{}_{\mu\nu}$ is the standard Levi-Civita connection, and an equation for e^a which is the Einstein field equation with the SU(N) gauge field source term,

$$R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R = 8\pi l_p^2 \, T^{\mu\nu} \,, \qquad T^{\mu\nu} \equiv -\frac{1}{4g} \left(F_{\rho\sigma}{}^I F^{\rho\sigma}{}_I g^{\mu\nu} + 4 F^{\mu\rho}{}_I F_\rho{}^{\nu I} \right) . \tag{2.24}$$

In this way, we see that both gravity and gauge fields can be successfully represented within a unified framework of higher gauge theory, based on a 2-group structure. A generalization from SU(N) Yang-Mills case to the more complicated cases, such as SU(3) × SU(2) × U(1), is straightforward.

Let us now review how one can use the 3-group structure and the corresponding constrained $3BF$ theory to describe general relativity coupled to Klein-Gordon and Dirac fields. To describe a single real Klein-Gordon field coupled to gravity, one begins by specifying a 2-crossed module $(L \xrightarrow{\delta} H \xrightarrow{\partial} G \,, \triangleright \,, \{\_\,,\_\}_\mathrm{p})$, as follows. The groups are given as

$G = \mathrm{SO}(3,1)$, $H = \mathbb{R}^4$, and $L = \mathbb{R}$. The group G acts on H via the vector representation, and on L via the trivial representation. The maps ∂ and δ are chosen to be trivial, as well as the Peiffer lifting. Given this choice of a 2-crossed module, the 3-connection $(\alpha\,,\beta\,,\gamma)$ takes the form

$$\alpha = \omega^{ab} M_{ab}\,, \qquad \beta = \beta^a P_a\,, \qquad \gamma = \gamma \mathbb{I}\,, \tag{2.25}$$

where $\mathbb{I}$ is the sole generator of the Lie group L. This 3-connection gives rise to the fake 3-curvature $(\mathcal{F}, \mathcal{G}, \mathcal{H})$,

$$\mathcal{F} = R^{ab} M_{ab}\,, \qquad \mathcal{G} = \nabla \beta^a P_a\,, \qquad \mathcal{H} = \mathrm{d}\gamma\,. \tag{2.26}$$

The importance of the $3BF$ theory for this choice of the 2-crossed module lies in the fact that the Lagrange multiplier D can transform as a scalar with respect to Lorentz symmetry, $M_{ab} \triangleright \mathbb{I} = 0$, and it transforms as a scalar with respect to diffeomorphisms since D is also a 0-form. In other words, one can interpret the Lagrange multiplier D to be a real scalar field, $D \equiv \phi$, and write the topological $3BF$ action (2.8) as:

$$S_{3BF} = \int_{\mathcal{M}_4} B^{ab} \wedge R_{ab} + e_a \wedge \nabla \beta^a + \phi\, \mathrm{d}\gamma\,. \tag{2.27}$$

In order to obtain the Klein-Gordon field ϕ of mass m coupled to gravity in the standard way, the appropriate simplicity constraints are imposed, and the constrained $3BF$ action takes the form:

$$\begin{aligned} S = \int_{\mathcal{M}_4} & B^{ab} \wedge R_{ab} + e_a \wedge \nabla \beta^a + \phi\, \mathrm{d}\gamma - \lambda_{ab} \wedge \left(B^{ab} - \frac{1}{16\pi l_p^2} \varepsilon^{abcd} e_c \wedge e_d \right) \\ & + \lambda \wedge \left(\gamma - \frac{1}{2} H_{abc} e^a \wedge e^b \wedge e^c \right) + \Lambda^{ab} \wedge \left(H_{abc} \varepsilon^{cdef} e_d \wedge e_e \wedge e_f - \mathrm{d}\phi \wedge e_a \wedge e_b \right) \\ & - \frac{1}{2 \cdot 4!} m^2 \phi^2 \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d\,. \end{aligned} \tag{2.28}$$

The first row is the topological sector (2.27) and the simplicity constraint for gravity from the action (2.17), the second row contains two new simplicity constraints featuring the Lagrange multiplier 1-forms λ and Λ^{ab} and the 0-form H_{abc}, and the third row features the mass term for the scalar field. The action (2.28) has two dynamical equations of motion — the equation for the scalar field ϕ is the covariant Klein-Gordon equation,

$$\left(\nabla_\mu \nabla^\mu - m^2 \right) \phi = 0\,, \tag{2.29}$$

while the equation for the tetrads e^a is the Einstein field equation with the scalar field source term,

$$R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R = 8\pi l_p^2\, T^{\mu\nu}\,, \quad T^{\mu\nu} \equiv \partial^\mu \phi \partial^\nu \phi - \frac{1}{2} g^{\mu\nu} \left(\partial_\rho \phi \partial^\rho \phi + m^2 \phi^2 \right)\,. \tag{2.30}$$

We see that the obtained theory is classically equivalent to general relativity coupled to a scalar field. Most importantly, one sees that the choice of the group L dictates the matter content of the theory, while the action $\triangleright$ of G on L specifies the transformation properties of the matter fields.

Finally, in order to describe the Dirac field coupled to Einstein-Cartan gravity, the 2-crossed module $(L \xrightarrow{\delta} H \xrightarrow{\partial} G\,, \triangleright\,, \{\_\,,\_\}_{\mathrm{p}})$ has to be chosen as follows. The groups are $G = \mathrm{SO}(3,1)$. $H = \mathbb{R}^4$, and $L = \mathbb{R}^8(\mathbb{G})$, where $\mathbb{G}$ is the algebra of complex Grassmann numbers. The maps ∂, δ, and the Peiffer lifting are trivial, as before. The action of the group G on H is via vector representation, and on L via spinor representation, in the following way. Denoting the eight generators of the Lie group $\mathbb{R}^8(\mathbb{G})$ as P_α and P^α, where the bispinor index α takes the values $1, \dots, 4$, the action $\triangleright$ of G on L is given explicitly as

$$M_{ab} \triangleright P_\alpha = \frac{1}{2}(\sigma_{ab})^\beta{}_\alpha P_\beta\,, \qquad M_{ab} \triangleright P^\alpha = -\frac{1}{2}(\sigma_{ab})^\alpha{}_\beta P^\beta\,, \tag{2.31}$$

where $\sigma_{ab} = \frac{1}{4}[\gamma_a, \gamma_b]$, and γ_a are the usual Dirac matrices. This choice of the 2-crossed module gives rise to the 3-connection $(\alpha\,, \beta\,, \gamma)$, defined as

$$\alpha = \omega^{ab} M_{ab}\,, \qquad \beta = \beta^a P_a\,, \qquad \gamma = \gamma^\alpha P_\alpha + \bar{\gamma}_\alpha P^\alpha\,, \tag{2.32}$$

where the 3-connection 3-forms γ^α and $\bar{\gamma}_\alpha$ should not be confused with the Dirac matrices γ_a due to different types of indices. The 3-curvature $(\mathcal{F}\,, \mathcal{G}\,, \mathcal{H})$ is given as:

$$\begin{aligned} &\mathcal{F} = R^{ab} M_{ab}\,, \qquad\qquad \mathcal{G} = \nabla \beta^a P_a\,, \\ &\mathcal{H} = \Big(\mathrm{d}\gamma^\alpha + \frac{1}{2}\omega^{ab}(\sigma_{ab})^\alpha{}_\beta \gamma^\beta\Big) P_\alpha + \Big(\mathrm{d}\bar{\gamma}_\alpha - \frac{1}{2}\omega^{ab}\bar{\gamma}_\beta(\sigma_{ab})^\beta{}_\alpha\Big) P^\alpha \equiv (\overset{\rightarrow}{\nabla}\gamma)^\alpha P_\alpha + (\bar{\gamma}\overset{\leftarrow}{\nabla})_\alpha P^\alpha\,. \end{aligned} \tag{2.33}$$

As in the case of the scalar field, the choice of the group L and action $\triangleright$ of G on L dictates the matter content of the theory and its transformation properties. The group L prescribes that D contains eight independent real anticommuting matter fields as its components. Then, since D is a 0-form and it transforms according to the spinorial representation of $\mathrm{SO}(3,1)$, these eight real Grassmann-valued fields can be identified with the four complex Dirac bispinor fields, and one can write the corresponding topological $3BF$ action as:

$$S_{3BF} = \int_{\mathcal{M}_4} B^{ab} \wedge R_{ab} + e_a \wedge \nabla \beta^a + (\bar{\gamma}\overset{\leftarrow}{\nabla})_\alpha \psi^\alpha + \bar{\psi}_\alpha (\overset{\rightarrow}{\nabla}\gamma)^\alpha\,. \tag{2.34}$$

In order to obtain the action that gives us the dynamics of Einstein-Cartan theory of gravity coupled to a Dirac field, we add the following simplicity constraints:

$$\begin{aligned} S = \int_{\mathcal{M}_4} & B^{ab} \wedge R_{ab} + e_a \wedge \nabla \beta^a + (\bar{\gamma}\overset{\leftarrow}{\nabla})_\alpha \psi^\alpha + \bar{\psi}_\alpha (\overset{\rightarrow}{\nabla}\gamma)^\alpha - \lambda_{ab} \wedge \Big(B^{ab} - \frac{1}{16\pi l_p^2}\varepsilon^{abcd} e_c \wedge e_d\Big) \\ & - \lambda^\alpha \wedge \Big(\bar{\gamma}_\alpha - \frac{i}{6}\varepsilon_{abcd} e^a \wedge e^b \wedge e^c (\bar{\psi}\gamma^d)_\alpha\Big) + \bar{\lambda}_\alpha \wedge \Big(\gamma^\alpha + \frac{i}{6}\varepsilon_{abcd} e^a \wedge e^b \wedge e^c (\gamma^d \psi)^\alpha\Big) \\ & - \frac{1}{12} m\, \bar{\psi}\psi\, \varepsilon_{abcd} e^a \wedge e^b \wedge e^c \wedge e^d + 2\pi i l_p^2\, \bar{\psi}\gamma_5 \gamma^a \psi\, \varepsilon_{abcd} e^b \wedge e^c \wedge \beta^d. \end{aligned} \tag{2.35}$$

The topological sector is in the first row, as well as the gravitational simplicity constraint, the second row contains the new simplicity constraints for the Dirac field, while the third row contains the mass term for the Dirac field and a term that ensures the correct coupling between the torsion and the spin of the Dirac field. Varying the action (2.35), one obtains the following dynamical equations of motion — the equations for ψ and $\bar{\psi}$ which are the standard covariant Dirac equation and its conjugate,

$$(i\gamma^a e^\mu{}_a \overrightarrow{\nabla}_\mu - m)\psi = 0\,, \quad \bar{\psi}(i\overleftarrow{\nabla}_\mu e^\mu{}_a \gamma^a + m) = 0\,, \tag{2.36}$$

and the differential equation of motion for e^a which is the Einstein field equation with a Dirac field source term,

$$R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R = 8\pi l_p^2\, T^{\mu\nu}\,, \quad T^{\mu\nu} \equiv \frac{i}{2}\bar{\psi}\gamma^a \overleftrightarrow{\nabla}^\nu e^\mu{}_a \psi - \frac{1}{2}g^{\mu\nu}\bar{\psi}\Big(i\gamma^a\overleftrightarrow{\nabla}_\rho e^\rho{}_a - 2m\Big)\psi\,, \tag{2.37}$$

where $\overleftrightarrow{\nabla} = \overrightarrow{\nabla} - \overleftarrow{\nabla}$. Moreover, one obtains the desired equation of motion for the torsion,

$$T_a \equiv \nabla e_a = 2\pi l_p^2 s_a\,, \quad s_a = i\varepsilon_{abcd}e^b \wedge e^c \bar{\psi}\gamma_5\gamma^d\psi\,, \tag{2.38}$$

where s_a is the Dirac spin 2-form. The equations of motion (2.36), (2.37), and (2.38) are precisely the equations of motion of the Einstein-Cartan-Dirac theory.

The natural presence of a scalar and Dirac field in the $3BF$ action is an essential property of the specific choices of the 3-group structures in a 4-dimensional spacetime, just like the existence of the tetrad field e^a in the topological $2BF$ action is an essential property of the $2BF$ action and the Poincaré 2-group. In this way, both the scalar field and the Dirac field appear in the topological sector of the action, making the quantization procedure feasible. Similarly, one can introduce Weyl and Majorana fields as well, see [31].

3 A review of 2-groups and 3-groups

As we have seen in the previous section, the gauge symmetry of 3-gauge theory is described by an algebraic structure known as a 3-*group*. In this section, we present the relevant definition of the 3-group, and we briefly explain how this structure is used to equip curves, surfaces, and volumes with holonomies. The results obtained in this section are necessary for the construction of the topological invariant, which will be studied in section IV.

3.1 3-Groups

In the category theory, a 2-group is defined as a 2-category consisting of only one object, where all the morphisms and 2-morphisms are invertible. It has been shown that every strict 2-group is equivalent to a crossed module $(H \overset{\partial}{\to} G\,, \rhd)$.

A *pre-crossed module* $(H \overset{\partial}{\to} G\,, \rhd)$ of groups G and H, is given by a group map $\partial : H \to G$, together with a left action $\rhd$ of G on both groups, by automorphisms, such that the group G acts on itself via conjugation, i.e. , for each $g_1, g_2 \in G$,

$$g_1 \rhd g_2 = g_1 g_2 g_1^{-1}\,,$$

and for each $h_1\,, h_2 \in H$ and $g \in G$ the following identity holds:

$$g\partial h g^{-1} = \partial(g \rhd h)\,.$$

JHEP07(2022)105

In a pre-crossed module the *Peiffer commutator* is defined as:

$$\langle h_1\,, h_2\rangle_{\rm p} = h_1 h_2 h_1^{-1} \partial(h_1) \triangleright h_2^{-1}\,. \tag{3.1}$$

A pre-crossed module is said to be a *crossed module* if all of its Peiffer commutators are trivial, which is to say that the *Peiffer identity* is satisfied:

$$(\partial h_1) \triangleright h_2 = h_1 h_2 h_1^{-1}\,. \tag{3.2}$$

Continuing the categorical generalization one step further, one can generalize the notion of a 2-group to the notion of a 3-group. Similar to the definition of a group and a 2-group within the category theory formalism, a 3-group is defined as a 3-category with only one object, where all morphisms, 2-morphisms, and 3-morphisms are invertible. Moreover, in analogy with how a crossed module encodes a strict 2-group, it has been proved that a semistrict 3-group — Gray group is equivalent to a 2-crossed module [42, 46].

A 2-*crossed module* $(L \overset{\delta}{\to} H \overset{\partial}{\to} G,\, \triangleright,\, \{\_\,,\,\_\}_{\rm p})$ is a chain complex of groups, given by three groups G, H, and L, together with maps ∂ and δ,

$$L \overset{\delta}{\to} H \overset{\partial}{\to} G\,,$$

such that $\partial\delta = 1_G$, an action $\triangleright$ of the group G on all three groups, and a map $\{\_\,,\_\}_{\rm p}$ called the *Peiffer lifting*:

$$\{\_\,,\_\}_{\rm p} : H \times H \to L\,.$$

The maps ∂ and δ, and the Peiffer lifting are G-equivariant, i.e., for each $g \in G$ and $h \in H$

$$g \triangleright \partial(h) = \partial(g \triangleright h)\,, \qquad g \triangleright \delta(l) = \delta(g \triangleright l)\,,$$

and for each h_1, $h_2 \in H$ and $g \in G$:

$$g \triangleright \{h_1\,, h_2\}_{\rm p} = \{g \triangleright h_1,\, g \triangleright h_2\}_{\rm p}\,.$$

The action of the group G on the groups H and L is a smooth left action by automorphisms, i.e., for each $g, g_1, g_2 \in G,\;\; h_1, h_2 \in H,\;\; l_1, l_2 \in L$ and $k \in H, L$,

$$g_1 \triangleright (g_2 \triangleright k) = (g_1 g_2) \triangleright k\,, \qquad g \triangleright (h_1 h_2) = (g \triangleright h_1)(g \triangleright h_2)\,, \qquad g \triangleright (l_1 l_2) = (g \triangleright l_1)(g \triangleright l_2)\,.$$

The action of the group G on itself is again via conjugation. Further, the following identities are satisfied:

$$\delta(\{h_1, h_2\}_{\rm p}) = \langle h_1\,, h_2\rangle_{\rm p}\,, \qquad \forall h_1, h_2 \in H\,; \tag{3.3a}$$

$$[l_1, l_2] = \{\delta(l_1)\,, \delta(l_2)\}_{\rm p}\,, \qquad \forall l_1\,, l_2 \in L\,, \quad \text{where } [l,k] = lkl^{-1}k^{-1}; \tag{3.3b}$$

$$\{h_1 h_2, h_3\}_{\rm p} = \{h_1, h_2 h_3 h_2^{-1}\}_{\rm p} \partial(h_1) \triangleright \{h_2, h_3\}_{\rm p}\,, \qquad \forall h_1, h_2, h_3 \in H\,; \tag{3.3c}$$

$$\{h_1, h_2 h_3\}_{\rm p} = \{h_1, h_2\}_{\rm p}\{h_1, h_3\}_{\rm p}\{\langle h_1, h_3\rangle_{\rm p}^{-1}\,, \partial(h_1) \triangleright h_2\}_{\rm p}\,, \qquad \forall h_1, h_2, h_3 \in H\,; \tag{3.3d}$$

$$\{\delta(l), h\}_{\rm p}\{h, \delta(l)\}_{\rm p} = l(\partial(h) \triangleright l^{-1})\,, \qquad \forall h \in H\,, \qquad \forall l \in L\,. \tag{3.3e}$$

In a 2-crossed module the structure $(L \xrightarrow{\delta} H, \rhd')$ is a crossed module, with action of the group H on the group L defined for each $h \in H$ and $l \in L$ as:

$$h \rhd' l = l \, \{\delta(l)^{-1}, \, h\}_{\mathrm{p}} \,,$$

and it follows that the Peiffer identity is satisfied for each $l_1, l_2 \in L$:

$$\delta(l_1) \rhd' \, l_2 = l_1 \, l_2 \, l_1^{-1} \,.$$

However, the structure $(H \xrightarrow{\partial} G \,, \rhd)$ in the general case does not form a crossed module, but a pre-crossed module, and for each $h, h' \in H$ the Peiffer commutator does not necessarily vanish.

The following identities hold, for each $h_1, h_2, h_3 \in H$ [42]:

$$\{h_1 h_2, h_3\}_{\mathrm{p}} = (h_1 \rhd' \{h_2, h_3\}_{\mathrm{p}})\{h_1, \partial(h_2) \rhd h_3\}_{\mathrm{p}} \,, \tag{3.4}$$

$$\{h_1, h_2 h_3\}_{\mathrm{p}} = \{h_1, h_2\}_{\mathrm{p}}(\partial(h_1) \rhd h_2) \rhd' \{h_1, h_3\}_{\mathrm{p}} \,, \tag{3.5}$$

and are of prime importance for the proof of the Pachner moves invariance. By using the condition (3.3e) of the definition of a 2-crossed module, it follows that for each $h \in H$ and $l \in L$ the following identity holds:

$$\{h, \delta(l)^{-1}\}_{\mathrm{p}} = (h \rhd' l^{-1})(\partial(h) \rhd l) \,. \tag{3.6}$$

Moreover, for each $h_1, h_2 \in H$,

$$\{h_1, h_2\}_{\mathrm{p}}^{-1} = h_1 \rhd' \{h_1^{-1}, \partial(h_1) \rhd h_2\}_{\mathrm{p}} \,, \tag{3.7}$$

$$\{h_1, h_2\}_{\mathrm{p}}^{-1} = \partial(h_1) \rhd \{h_1^{-1}, h_1 h_2 h_1^{-1}\}_{\mathrm{p}} \,, \tag{3.8}$$

$$\{h_1, h_2\}_{\mathrm{p}}^{-1} = (h_1 h_2 h_1^{-1}) \rhd' \{h_1, h_2^{-1}\}_{\mathrm{p}} \,, \tag{3.9}$$

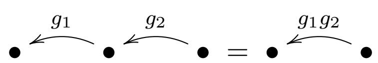
$$\{h_1, h_2\}_{\mathrm{p}}^{-1} = (\partial(h_1) \rhd h_2) \rhd' \{h_1, h_2^{-1}\}_{\mathrm{p}} \,. \tag{3.10}$$

A reader interested in more details about 3-groups is referred to [43].

3.2 3-gauge theory

In this subsection, we will describe how the language of 3-gauge theory can be used in order to define compositions of labeled paths, surfaces, and volumes. In a 3-gauge theory, one labels geometric objects at three levels. Curves are labeled by elements of G. Their composition and orientation reversal is defined as in conventional gauge theory. In addition, surfaces are labeled with elements of H, and volumes are labeled with the elements of L. The reader interested in the formulation of a 2-gauge theory is referred to [47].

Curves are labeled with the elements of G, and the elements are composed as in the ordinary gauge theory, i.e. , for each $g_1, g_2 \in G$,

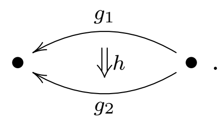


,

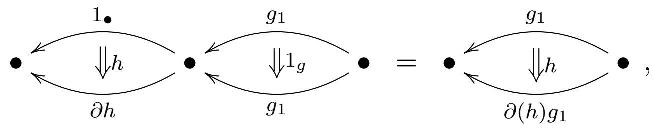
the composition of the elements results in the element $g_1 g_2 \in G$. The orientation of a curve can be reversed if it is labeled by the inverse element g^{-1} instead.

JHEP07(2022)105

Surfaces are labeled with the elements $h \in H$. For each surface, we choose two reference points on the boundary, and split the boundary into two curves, the source curve labeled with $g_1 \in G$, and the target curve labeled with $g_2 \in G$, as demonstrated in the diagram



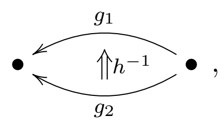
The 2-arrow $h \in H$ maps the curve $g_1 \in G$ to the curve $\partial(h) g_1 \in G$,



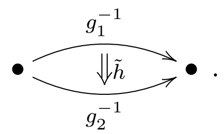
so that the label $h \in H$ of the surface is required to satisfy the following condition:

$$\partial(h) = g_2 g_1^{-1} \,. \tag{3.11}$$

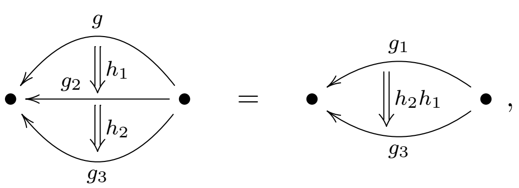
The orientation of the surface can be reversed and labeled with the inverse element instead,



while the orientation reversal of the curves leads to the surface element labeled with $\tilde{h} = g_1^{-1} \triangleright h^{-1}$:



One can now compose 2-morphisms vertically. Let us denote the source and the target of the k-arrow ($k = 1, 2$) of the 2-morphism h as $\partial_k^-(h)$ and $\partial_k^+(h)$, respectively. Then, the vertical composition of 2-morphisms (g_1, h_1) and (g_2, h_2), when they are compatible, i.e., when $\partial_2^+(h_1) = \partial_2^-(h_2)$,



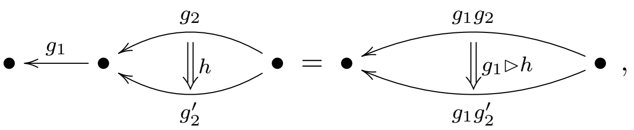
results in a 2-morphism $(g_1, h_2 h_1)$,

$$(g_2, h_2) \#_2 (g_1, h_1) = (g_1, h_2 h_1) \,. \tag{3.12}$$

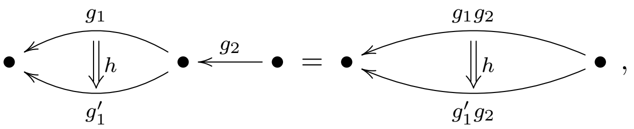
An important operation is known as whiskering. One can whisker a 2-morphism h with a morphism g_1 by attaching the whisker g_1 to the surface h from the left, i.e., such

JHEP07(2022)105

that $\partial_1^-(g_1) = \partial_1^+(h)$,

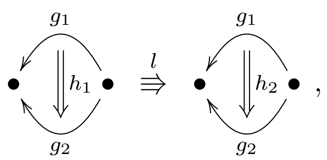


which results in the 2-morphism with the source curve g_1g_2 and target curve g_1g_2', carrying the label $g_1 \triangleright h$. Similarly, by attaching whisker g_2 to a surface h from the right, i.e. , such that $\partial_1^-(h) = \partial_1^+(g_2)$,



one obtains the 2-morphism with the source curve g_1g_2 and target curve $g_1'g_2$, carrying the label h.

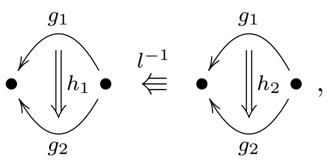
The volumes are labeled with the elements $l \in L$. Let us denote the source and the target of the k-arrow ($k = 1, 2, 3$) of the 3-morphism l as $\partial_k^-(l)$ and $\partial_k^+(l)$, respectively. For each volume, we split the boundary into two surfaces, the source surface labeled with $\partial_3^-(l) = h_1$ and the target surface labeled with $\partial_3^+(l) = h_2$. On the common boundary of the source and target surface, we choose two reference points, and split the boundary into two curves, the source curve labeled with $\partial_2^-(l) = g_1$ and the target curve labeled with $\partial_2^+(l) = g_2$, as demonstrated in the diagram below



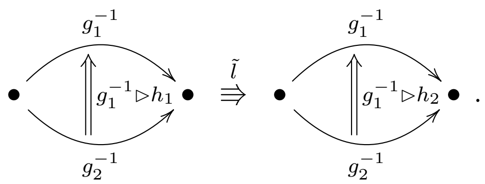
so that the volume label $l \in L$ is required to satisfy the following condition:

$$\delta(l) = h_2 h_1^{-1} \,. \tag{3.13}$$

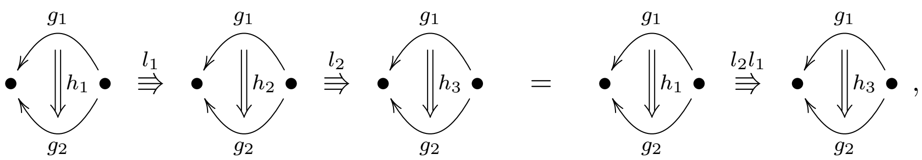
The orientation of the volume can be reversed if one labels it with the inverse element l^{-1}:



while the orientation reversal of the curves and surfaces leads to the surface element labeled with $\tilde{l} = g_1^{-1} \triangleright l$,



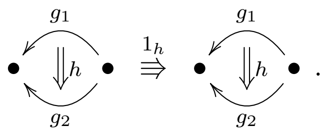
One can compose two 3-morphisms via the *upward composition* (visualizing a third axis, orthogonal to the plane of the paper, as the direction up). The upward composition of 3-morphisms (g_1, h_1, l_1) and (g_1, h_2, l_2), when they are compatible, i.e., when $\partial_3^+(l_1) = \partial_3^-(l_2)$,



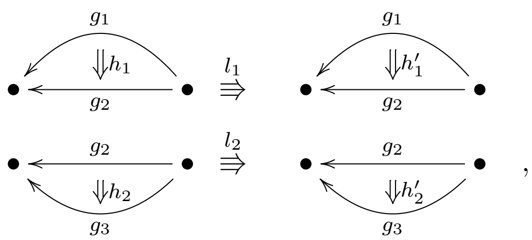
results in a 3-morphism $(g_1, h_1, l_2 l_1)$:

$$(g_1, h_2, l_2) \#_3 (g_1, h_1, l_1) = (g_1, h_1, l_2 l_1) \,. \tag{3.14}$$

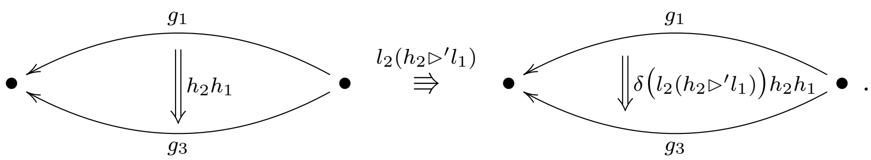
The upward composition of 3-morphisms is associative, and for every $h \in H$ there is a 3-morphism that is an identity for the upward composition of 3-morphisms



The vertical composition of two 3-morphisms (g_1, h_1, l_1) and (g_2, h_2, l_2), when they are compatible, i.e., when $\partial_2^+(l_1) = \partial_2^-(l_2)$,



results in a 3-morphism $(g_1, h_2 h_1, l_2(h_2 \rhd' l_1))$,



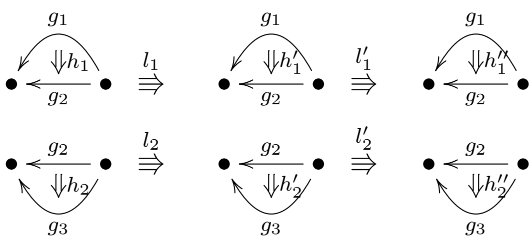
One can write, for (g_1, h_1, l_1) and (g_2, h_2, l_2),

$$(g_2, h_2, l_2) \#_2 (g_1, h_1, l_1) = (g_1, h_2 h_1, l_2(h_2 \rhd' l_1)) \,. \tag{3.15}$$

The vertical composition of 3-morphisms is an associative operation. Composition of 3-morphisms is invariant under the change of order of upward composition and vertical composition of 3-morphisms, i.e.,

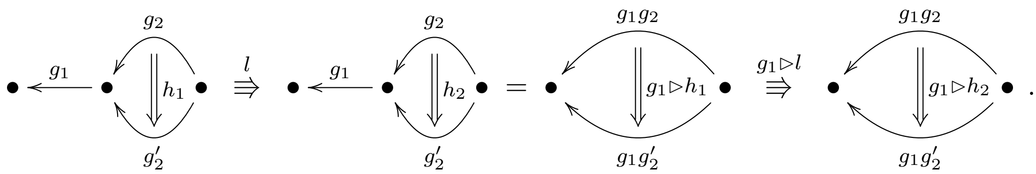
$$\begin{aligned} &((g_2, h_2', l_2') \#_3 (g_2, h_2, l_2)) \#_2 ((g_1, h_1', l_1') \#_3 (g_1, h_1, l_1)) \\ =&((g_2, h_2', l_2') \#_2 (g_1, h_1', l_1')) \#_3 ((g_2, h_2, l_2) \#_2 (g_1, h_1, l_1)) \,, \end{aligned} \tag{3.16}$$

which is demonstrated in the diagram notation, where the diagram



uniquely determines the 3-morphism. The proof of the equation (3.16) is given in the appendix A.

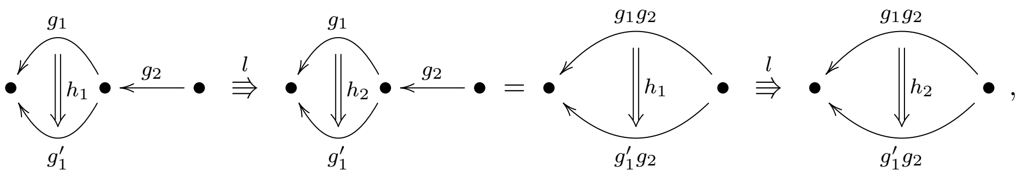
One can whisker the 3-morphisms with morphisms and 2-morphisms. Whiskering of a 3-morphism by a morphism from the left is the composition of a volume $l \in L$ and curve $g_1 \in G$ from the left, when they are compatible, i.e., when $\partial_1^+(l) = \partial_1^-(g_1)$,



The composition results in a 3-morphism:

$$g_1 \#_1 (g_2, h_1, l) = (g_1 g_2, g_1 \triangleright h_1, g_1 \triangleright l)\,. \tag{3.17}$$

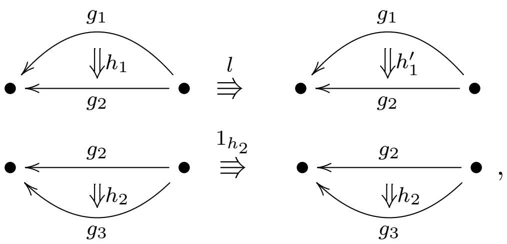
Similarly, one can whisker a 3-morphism by a morphism from the right, when they are compatible, i.e., $\partial_1^-(l) = \partial_1^+(g_2)$,



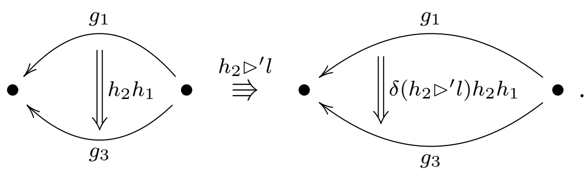
which results in the 3-morphism:

$$(g_1, h_1, l) \#_1 g_2 = (g_1 g_2, h_1, l)\,. \tag{3.18}$$

Whiskering of a 3-morphism with a 2-morphisms from below, when they are compatible, i.e., $\partial_2^+(l) = \partial_2^-(h_2)$, is formed as a vertical composition of 3-morphisms (g_1, h_1, l) and $(g_2, h_2, 1_{h_2})$,



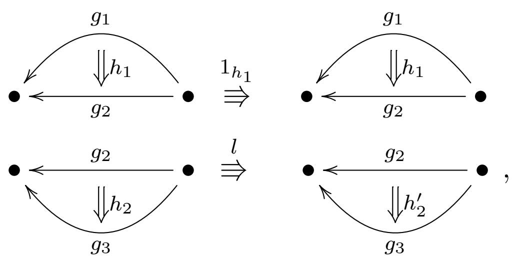
which results in a 3-morphism



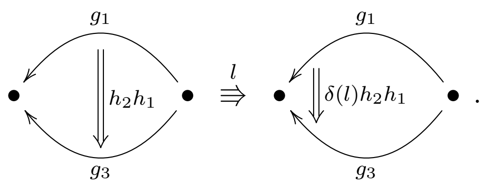
One writes,

$$(g_2, h_2)\#_2(g_1, h_1, l) = (g_1, h_2h_1, h_2 \triangleright' l)\,. \tag{3.19}$$

Whiskering a 3-morphism by 2-morphism from above, when they are compatible, i.e., when $\partial_2^-(l) = \partial_2^+(h_1)$, is formed as a vertical composition of 3-morphisms $(g_1, h_1, 1_{h_1})$ and (g_2, h_2, l),



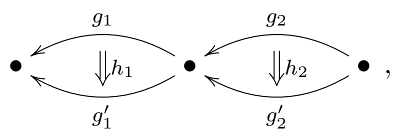
which results in a 3-morphism,



One obtains

$$(g_2, h_2, l)\#_2(g_1, h_1) = (g_1, h_2h_1, l)\,. \tag{3.20}$$

The interchanging 3-arrow is the horizontal composition of two 2-morphisms h_1 and h_2, when they are compatible, i.e., when $\partial_1^-(h_1) = \partial_1^+(h_2)$,

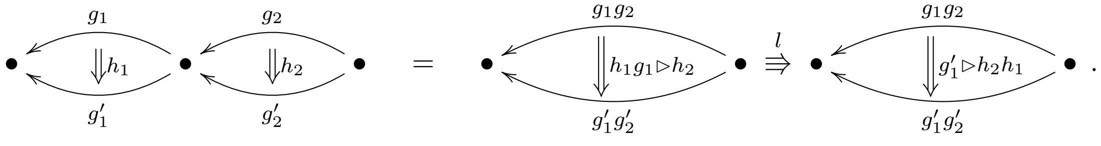


that results in a 3-morphism l, with source surface

$$\partial_3^-(l) = ((g_1, h_1)\#_1 g_2')\#_2(g_1\#_1(g_2, h_2))\,,$$

and target surface

$$\partial_3^+(l) = (g_1'\#_1(g_2, h_2))\#_2((g_1, h_1)\#_1 g_2)\,,$$



One obtains,

$$(g_1,h_1)\#_1(g_2,h_2)=(g_1g_2,h_1g_1\rhd h_2,l)\,, \tag{3.21}$$

where the 3-morphism l is Peiffer lifting $\{h_1,g_1\rhd h_2\}_{\mathrm{p}}^{-1}$. Using the condition (3.13), one obtains

$$((\partial(h_1)g_1)\rhd h_2)h_1=\delta(l)h_1(g_1\rhd h_2)\,, \tag{3.22}$$

and from the definition of the Peiffer commutator, the identity (3.1), and the property (3.3a) of the 2-crossed module, i.e. , $\delta(\{h_1,h_2\}_{\mathrm{p}})=\langle h_1\,,h_2\rangle_{\mathrm{p}}$, one obtains

$$\delta(l)^{-1}=h_1g_1\rhd h_2h_1^{-1}(\partial(h_1)g_1)\rhd h_2{}^{-1}=\langle h_1,g_1\rhd h_2\rangle_{\mathrm{p}}=\delta(\{h_1,g_1\rhd h_2\}_{\mathrm{p}})\,. \tag{3.23}$$

Given any collection of curves, surfaces, and volumes, a configuration of 3-gauge theory is an assignment of elements of G to the curves, elements of H to the surfaces, and elements of L to volumes so that the following conditions hold:

1. For each surface labeled by $h\in H$, one has that $\partial(h)=g_2g_1^{-1}$ where g_1 and g_2 are the source and target curve, respectively;

2. For each volume labeled by $l\in L$, one has that $\delta(l)=h_2h_1^{-1}$, where h_1 and h_2 are the source and target surface, respectively;

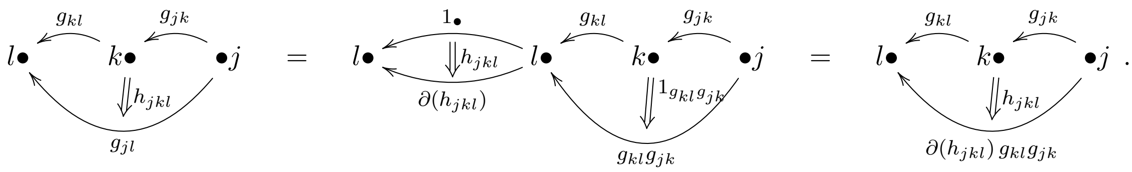
3. For each 4-simplex labeled by $(jk\ell mn)\in\Lambda_4$, the volume holonomy around it is trivial.

The defined configurations can be viewed as the classical configurations of 3-gauge theory or, in a path integral quantum theory, these are the configurations over which one integrates in the path integral.

3.3 Gauge invariant quantities

In subsection 3.2, we have introduced a number of operations by which we can combine labeled paths, surfaces, and volumes, in order to calculate the composition of elementary paths, surfaces, and volumes, to arbitrarily large ones. In this subsection, we will make use of these compositions in order to construct gauge invariant quantities that are associated with closed paths, surfaces, and volumes. In Lemmas 3.1, 3.2, and 3.3, this procedure is used for the boundary path of a triangle, the boundary surface of a tetrahedron, and the boundary volume of the 4-simplex. The result of Lemma 3.1 is already derived for the case of 2-groups and remains unchanged in the 3-gauge theory, see [38]. The higher flatness condition for the boundary surface of a tetrahedron derived in [38], is generalized for the case of 3-groups is Lemma 3.2. One of the main results of the paper is Lemma 3.3 where we derived the higher flatness condition for the boundary volume of the 4-simplex.

Lemma 3.1. Let us consider a triangle, $(jk\ell)$. The edges $(jk)\,,j<k$, are labeled by group elements $g_{jk}\in G$ and the triangle $(jk\ell)\,,j<k<\ell$, by element $h_{jk\ell}\in H$. Consider the diagram (3.24).



(3.24)

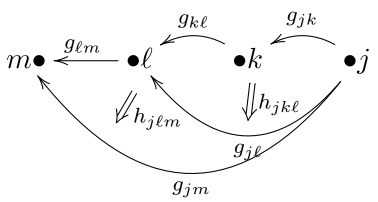
The curve $\gamma_1 = g_{k\ell}g_{jk}$ is the source and the curve $\gamma_2 = g_{j\ell}$ is the target of the surface morphism $\Sigma : \gamma_1 \to \gamma_2$, labeled by the group element $h_{jk\ell}$, i.e.,

$$g_{j\ell} = \partial(h_{jk\ell})g_{k\ell}g_{jk} \,. \tag{3.25}$$

Lemma 3.2. Let us consider a tetrahedron, $(jk\ell m)$. The edges $(jk)\,, j<k$, are labeled by group elements $g_{jk} \in G$ and the triangles $(jk\ell)\,, j<k<\ell$, by elements $h_{jk\ell} \in H$, and the tetrahedron $(jk\ell m)\,, j<k<\ell<m$ by the group element $l_{jk\ell m} \in L$. We have oriented the triangles $(jk\ell)$ so that they have the source is $g_{k\ell}g_{jk}$ and the target is $g_{j\ell}$, i.e. $g_{j\ell} = \partial(h_{jk\ell})g_{k\ell}g_{jk}$.

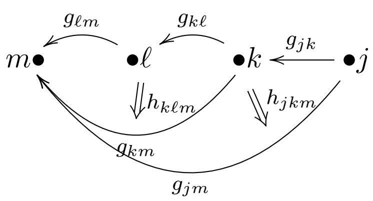
Let us first cut the tetrahedron surface along the edge (jm). This determines the ordering of the vertical composition of the constituent surfaces. One just has to make sure that all surfaces are composable, i.e., they have the suitable reference points and the correct orientation in order to compose them vertically.

Consider the diagram (3.26). We first move the curve from $g_{k\ell}g_{jk}$ to the curve $g_{j\ell}$. At this stage, one cannot compose the result with the triangle $(j\ell m)$, and one first has to whisker it from the left by $g_{\ell m}$. Now the two morphisms are vertically composable, and this moves the curve to g_{jm}. The following 2-morphism is obtained



$$\begin{aligned} &= (g_{\ell m}g_{j\ell}, h_{j\ell m})\#_2(g_{\ell m}\#_1(g_{k\ell}g_{jk}, h_{jk\ell})) \\ &= (g_{\ell m}g_{k\ell}g_{jk}, h_{j\ell m}(g_{\ell m} \triangleright h_{jk\ell}))\,. \end{aligned} \tag{3.26}$$

Let us then consider the diagram (3.27). We first move the curve from $g_{\ell m}g_{k\ell}$ to the curve g_{km}. At this stage, one cannot compose the result with the triangle (jkm), and one first has to whisker it from the right by g_{jk}. Now the two morphisms are vertically composable, and this moves the curve to g_{jm}. The following 2-morphism is obtained



$$\begin{aligned} &= (g_{km}g_{jk}, h_{jkm})\#_2((g_{\ell m}g_{k\ell}, h_{k\ell m})\#_1 g_{jk}) \\ &= (g_{\ell m}g_{k\ell}g_{jk}, h_{jkm}h_{k\ell m})\,. \end{aligned} \tag{3.27}$$

The two surfaces have the same source and target, $\Sigma_1 : g_{\ell m}g_{k\ell}g_{jk} \to g_{jm}$ and $\Sigma_2 : g_{\ell m}g_{k\ell}g_{jk} \to g_{jm}$. Now, transition from the surface shown on the diagram (3.26) to the surface shown on the diagram (3.27) is given by the volume morphism $\mathcal{V} : \Sigma_1 \to \Sigma_2$

JHEP07(2022)105

determined by the group element $l_{jk\ell m}$, i.e. ,

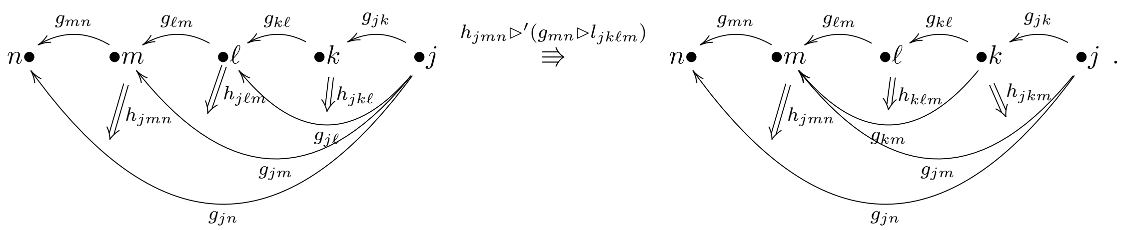
$$(g_{\ell m}g_{k\ell}g_{jk}, h_{jkm}h_{k\ell m}) = (g_{\ell m}g_{k\ell}g_{jk}, \delta(l_{jk\ell m})h_{j\ell m}(g_{\ell m} \rhd h_{jk\ell}))\,, \tag{3.28}$$

that gives the relation,

$$h_{jkm}h_{k\ell m} = \delta(l_{jk\ell m})h_{j\ell m}(g_{\ell m} \rhd h_{jk\ell})\,. \tag{3.29}$$

Lemma 3.3. Let us consider a 4-simplex, $(jk\ell mn)$. The edges $(jk)\,, j < k$, are labeled by group elements $g_{jk} \in G$, the triangles $(jk\ell)\,, j < k < \ell$, by elements $h_{jk\ell} \in H$, and the tetrahedrons $(jk\ell m)\,, j < k < \ell < m$, by the group element $l_{jk\ell m} \in L$. We have oriented the triangles $(jk\ell)$ so that the source curve is $g_{k\ell}g_{jk}$ and the target curve is $g_{j\ell}$, i.e. , $g_{j\ell} = \partial(h_{jk\ell})g_{k\ell}g_{jk}$, and the tetrahedrons $(jk\ell m)$ so that the source surface is $h_{j\ell m}(g_{\ell m} \rhd h_{jk\ell})$ and the target surface is $h_{jkm}h_{k\ell m}$, i.e. , $h_{jkm}h_{k\ell m} = \delta(l_{jk\ell m})h_{j\ell m}(g_{\ell m} \rhd h_{jk\ell})$.

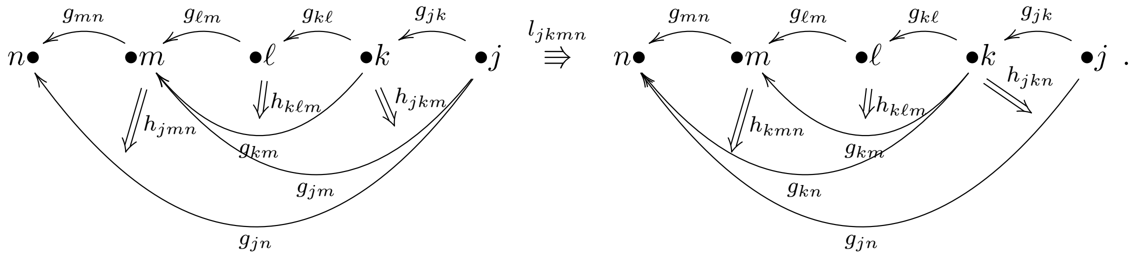
Let us first cut the 4-simplex volume along the surface $h_{jmn}g_{mn} \rhd (h_{j\ell m}g_{\ell m} \rhd h_{jk\ell})$. This surface determines the ordering of the vertical composition of the constituent volumes. We have to make sure that all volumes are composable, i.e. , they have the suitable reference points and the correct orientation in order to compose them vertically. First, let us consider the diagram (3.30). We first move the surface from $h_{j\ell m}g_{\ell m} \rhd h_{jk\ell}$ to surface $h_{jkm}h_{k\ell m}$ with the 3-arrow $l_{jk\ell m}$. To compose the resulting 3-morphism with the surface h_{jmn} one must first whisker it from the left with g_{mn}. The obtained 3-morphism $(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, g_{mn} \rhd (h_{j\ell m}g_{\ell m} \rhd h_{jk\ell}), g_{mn} \rhd l_{jk\ell m})$ can be whiskered from below with the 2-morphism $(g_{mn}g_{jm}, h_{jmn})$, and the resulting 3-morphism is $(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, h_{jmn}g_{mn} \rhd (h_{j\ell m}g_{\ell m} \rhd h_{jk\ell}), h_{jmn} \rhd' (g_{mn} \rhd l_{jk\ell m}))$, with the source surface $h_{jmn}g_{mn} \rhd (h_{j\ell m}g_{\ell m} \rhd h_{jk\ell})$ and the target surface $h_{jmn}g_{mn} \rhd (h_{jkm}h_{k\ell m})$,



(3.30)

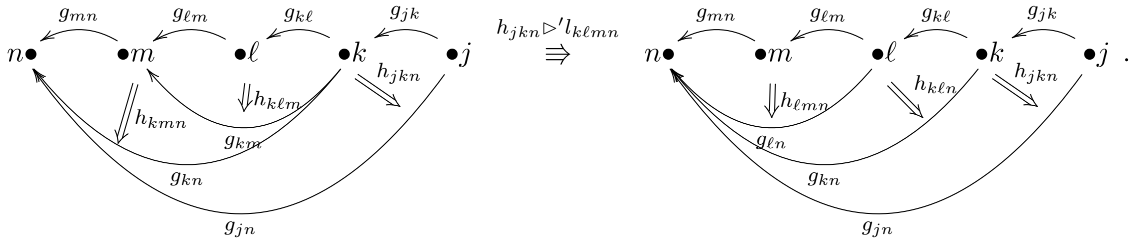
Let us move the surface to $h_{jkn}h_{kmn}g_{m\ell} \rhd h_{k\ell m}$, see diagram (3.31). To do that, we consider the 3-morphism $(g_{mn}g_{km}g_{jk}, h_{jmn}g_{mn} \rhd h_{jkm}, l_{jkmn})$ with the source surface $h_{jmn}g_{mn} \rhd h_{jkm}$ and target surface $h_{jkn}h_{kmn}$. This 3-morphism can be whiskered from above with the 2-morphism $(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, g_{mn} \rhd h_{k\ell m})$, and the obtained 3-morphism is $(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, h_{jmn}\, g_{mn} \rhd (h_{jkm}h_{k\ell m}), l_{jkmn})$, with the source surface $h_{jmn}g_{mn} \rhd (h_{jkm}h_{k\ell m})$ and target surface

$h_{jkn}h_{kmn}g_{mn} \rhd h_{k\ell m}$,



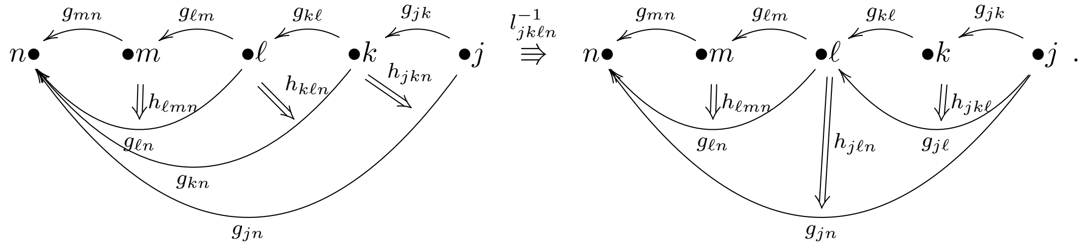
(3.31)

Next, we want to move the surface $h_{jkn}h_{kmn}g_{mn} \rhd h_{k\ell m}$ to surface $h_{jkn}h_{k\ell n}h_{\ell mn}$, as shown on the diagram (3.32). We whisker the 3-morphism $(g_{mn}g_{\ell m}g_{k\ell}, h_{kmn}g_{mn} \rhd h_{k\ell m}, l_{k\ell mn})$, with the source surface $h_{kmn}g_{mn} \rhd h_{k\ell m}$ and target surface $h_{k\ell n}h_{\ell mn}$, with the morphism g_{jk} from the right, obtaining the 3-morphism $(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, h_{kmn}g_{mn} \rhd h_{k\ell m}, l_{k\ell mn})$. Now, we whisker this 3-morphism with the 2-morphism $(g_{kn}g_{jk}, h_{jkn})$ from below, and we obtain the 3-morphism $(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, h_{jkn}h_{kmn}g_{mn} \rhd h_{k\ell m}, h_{jkn} \rhd' l_{k\ell mn})$,



(3.32)

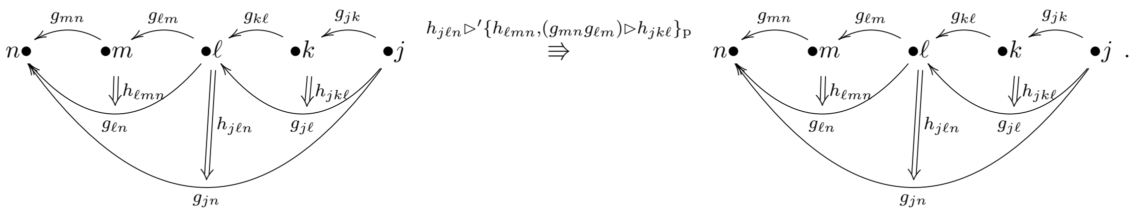
The mapping of the surface $h_{jkn}h_{k\ell n}h_{\ell mn}$ to the surface $h_{j\ell n}g_{\ell n} \rhd h_{jk\ell}h_{\ell mn}$ in shown on the diagram (3.33). The 3-morphism with the appropriate source and target is constructed by whiskering the 3-morphism $(g_{\ell n}g_{k\ell}g_{jk}, h_{jkn}h_{k\ell n}, l^{-1}_{jk\ell n})$ with 2-morphism $(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, h_{\ell mn})$ from above. The obtained 3-morphism is $(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, h_{jkn}h_{k\ell n}h_{\ell mn}, l^{-1}_{jk\ell n})$,



(3.33)

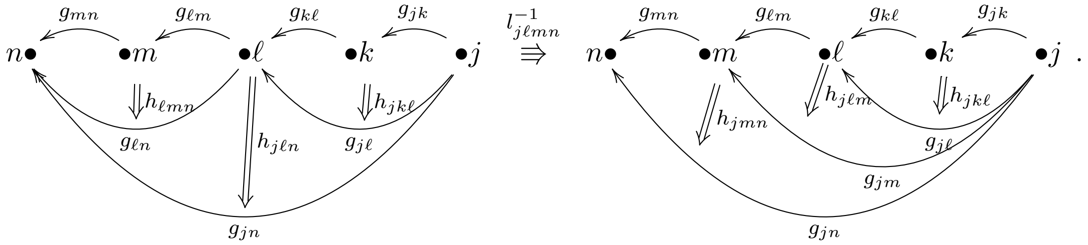
Next we map the surface $h_{j\ell n}g_{\ell n} \rhd h_{jk\ell}h_{\ell mn}$ to the surface $h_{j\ell n}h_{\ell mn}(g_{mn}g_{\ell m}) \rhd h_{jk\ell}$, see the diagram (3.34). We use the inverse interchanging 2-arrow composition to map the surface $g_{\ell n} \rhd h_{jk\ell}h_{\ell mn}$ to the surface $h_{\ell mn}(g_{mn}g_{\ell m}) \rhd h_{jk\ell}$, resulting in the 3-morphism $(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, g_{\ell n} \rhd h_{jk\ell}h_{\ell mn}, \{h_{\ell mn}, (g_{mn}g_{\ell m}) \rhd h_{jk\ell}\}_{\mathrm{p}})$. Next, we whisker the obtained 3-morphism with the 2-morphism $(g_{\ell n}g_{j\ell}, h_{j\ell n})$ from below. The obtained 3-morphism with the appropriate source and target surfaces is $(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, h_{j\ell n}g_{\ell n} \rhd h_{jk\ell}h_{\ell mn}, h_{j\ell n} \rhd'$

$\{h_{\ell mn}, (g_{mn}g_{\ell m}) \triangleright h_{jk\ell}\}_{\mathrm{p}})$,



(3.34)

Finally, we construct the 3-morphism that maps the surface $h_{j\ell n}h_{\ell mn}(g_{mn}g_{\ell m}) \triangleright h_{jk\ell}$ to the starting surface $h_{jmn}g_{mn} \triangleright (h_{j\ell m}g_{\ell m} \triangleright h_{jk\ell})$. To obtain the 3-morphism with the appropriate source and target surfaces we first move the surface $h_{j\ell n}h_{\ell mn}$ to the surface $h_{jmn}g_{mn} \triangleright h_{j\ell m}$ with the 3-arrow $(g_{mn}g_{\ell m}g_{j\ell}, h_{j\ell n}h_{\ell mn}, l^{-1}_{j\ell mn})$. Next, we whisker the 3-morphism $(g_{mn}g_{\ell m}g_{j\ell}, h_{j\ell n}h_{\ell mn}, l^{-1}_{j\ell mn})$ with the 2-morphism $(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, (g_{mn}g_{\ell m}) \triangleright h_{jk\ell})$ from above. The obtained 3-morphism $(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, h_{j\ell n}h_{\ell mn}(g_{mn}g_{\ell m}) \triangleright h_{jk\ell}, l^{-1}_{j\ell mn})$ moves the surface to the starting surface, as shown on the diagram (3.35),



(3.35)

After the upward composition of the 3-morphisms given by the diagrams (3.30)–(3.35), the obtained 3-morphism is:

$$
\begin{aligned}
&(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, h_{j\ell n}h_{\ell mn}(g_{mn}g_{\ell m}) \triangleright h_{jk\ell}, l^{-1}_{j\ell mn})\#_3\\
&(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, g_{\ell n} \triangleright h_{jk\ell}h_{\ell mn}, h_{j\ell n} \triangleright' \{h_{\ell mn}, (g_{mn}g_{\ell m}) \triangleright h_{jk\ell}\}_{\mathrm{p}})\#_3\\
&(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, h_{jkn}h_{k\ell n}h_{\ell mn}, l^{-1}_{jk\ell n})\#_3\\
&(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, h_{jkn}h_{kmn}g_{m\ell} \triangleright h_{k\ell m}, h_{jkn} \triangleright' l_{jkmn})\#_3\\
&(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, h_{jmn}g_{mn} \triangleright (h_{jkm}h_{k\ell m}), l_{jkmn})\#_3\\
&(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, h_{jmn}g_{mn} \triangleright (h_{j\ell m}g_{\ell m} \triangleright h_{jk\ell}), h_{jmn} \triangleright' (g_{mn} \triangleright l_{jk\ell m}))\\
={}&(g_{mn}g_{\ell m}g_{k\ell}g_{jk}, h_{jmn}g_{mn} \triangleright (h_{j\ell m}g_{\ell m} \triangleright h_{jk\ell}), l^{-1}_{j\ell mn}\, h_{j\ell n} \triangleright' \{h_{\ell mn}, (g_{mn}g_{\ell m}) \triangleright h_{jk\ell}\}_{\mathrm{p}}\\
&l^{-1}_{jk\ell n}(h_{jkn} \triangleright' l_{k\ell mn})l_{jkmn}h_{jmn} \triangleright' (g_{mn} \triangleright l_{jk\ell m}))\,.
\end{aligned}
\tag{3.36}
$$

The obtained 3-morphism is the identity morphism with source and target surface $\mathcal{V}_1 = \mathcal{V}_2 = h_{jmn}g_{mn} \triangleright (h_{j\ell m}g_{\ell m} \triangleright h_{jk\ell})$, i.e. ,

$$
l^{-1}_{j\ell mn}\, h_{j\ell n} \triangleright' \{h_{\ell mn}, (g_{mn}g_{\ell m}) \triangleright h_{jk\ell}\}_{\mathrm{p}}\, l^{-1}_{jk\ell n}(h_{jkn} \triangleright' l_{k\ell mn})l_{jkmn}h_{jmn} \triangleright' (g_{mn} \triangleright l_{jk\ell m}) = e\,. \tag{3.37}
$$

JHEP07(2022)105

4 Quantization of the topological $3BF$ theory

In conventional BF theory, one chooses the action in such a way that the theory does not depend on any background field, but only the spacetime manifold. The classical field equations of the theory require the gauge connection to be flat, i.e., in terms of the holonomy variables, that any null-homotopic closed curve corresponds to the identity of the gauge group. In the framework of higher gauge theory, specifically 2-gauge theory, one generalizes this idea by imposing the *higher flatness condition* requiring that the surface holonomy around the boundary 2-sphere of any 3-ball be trivial instead. One can continue further categorical generalization by choosing a 3-group structure to describe the gauge symmetry of the theory, and formulate a $3BF$ theory whose equations of motion impose a higher flatness condition for a 3-curvature $(\mathcal{F},\mathcal{G},\mathcal{H})$. In this section, a combinatorial description of such model for any triangulation of any smooth manifold of dimension $d=4$ is presented. This model coincides with Porter's abstract definition of a TQFT [33] for $d=4$ and $n=3$, which is itself a generalization of Yetter's work [48, 49].

Let us show how to construct a state sum model from the classical action (2.8) by the usual heuristic spinfoam quantization procedure. We consider the path integral for the action S_{3BF},

$$Z=\int \mathcal{D}\alpha\,\mathcal{D}\beta\,\mathcal{D}\gamma\,\mathcal{D}B\,\mathcal{D}C\,\mathcal{D}D\,\exp\left(i\int_{M_4}\langle B\wedge\mathcal{F}\rangle_{\mathfrak{g}}+\langle C\wedge\mathcal{G}\rangle_{\mathfrak{h}}+\langle D\wedge\mathcal{H}\rangle_{\mathfrak{l}}\right). \tag{4.1}$$

The formal integration over the Lagrange multipliers B, C, and D leads to:

$$Z=\mathcal{N}\int\mathcal{D}\alpha\,\mathcal{D}\beta\,\mathcal{D}\gamma\;\delta(\mathcal{F})\delta(\mathcal{G})\delta(\mathcal{H})\,. \tag{4.2}$$

Similarly to conventional gauge theory, the connection 1-form $\alpha\in\mathcal{A}^1(\mathcal{M}_4,\mathfrak{g})$ is discretized by colouring the edges $\epsilon=(jk)\in\Lambda_1$ of the triangulation with group elements $g_\epsilon\in G$. The connection 2-form $\beta\in\mathcal{A}^2(\mathcal{M}_4\,,\mathfrak{h})$ is represented by group elements $h_\Delta\in H$ coloring the triangles $\Delta=(jk\ell)\in\Lambda_2$. The connection 3-form $\gamma\in\mathcal{A}^3(\mathcal{M}_4\,,\mathfrak{l})$ is represented by group elements $l_\tau\in L$ coloring the tetrahedrons $\tau=(jk\ell m)\in\Lambda_3$.

The path integral measures of (4.1) are discretized by replacing

$$\int\mathcal{D}\alpha\quad\mapsto\quad\prod_{(jk)\in\Lambda_1}\int_G dg_{jk}\,, \tag{4.3}$$

$$\int\mathcal{D}\beta\quad\mapsto\quad\prod_{(jk\ell)\in\Lambda_2}\int_H dh_{jk\ell}, \tag{4.4}$$

$$\int\mathcal{D}\gamma\quad\mapsto\quad\prod_{(jk\ell m)\in\Lambda_3}\int_L dl_{jk\ell m}\,, \tag{4.5}$$

where dg_{jk}, $dh_{jk\ell}$, and $dl_{jk\ell m}$ denote integration with respect to the Haar measures of G, H, and L, respectively. The vanishing fake curvature condition is discretized on each triangle $(jkl)\in\Lambda_2$ by discretizing $\delta(\mathcal{F})$. When passing from a smooth manifold to its triangulation, the δ distribution is defined over the appropriate set of simplices as follows,

$$\delta(\mathcal{F})=\prod_{(jk\ell)\in\Lambda_2}\delta_G(g_{jk\ell})\,, \tag{4.6}$$

where for each $(jkl) \in \Lambda_2$ the δ-function $\delta_G(g_{jkl})$ is given by:

$$\delta_G(g_{jk\ell}) = \delta_G\big(\partial(h_{jk\ell})\, g_{k\ell}\, g_{jk}\, g_{j\ell}^{-1}\big)\,. \tag{4.7}$$

Similarly, on the triangulated manifold the condition $\delta(\mathcal{G})$ on the fake curvature 3-form reads

$$\delta(\mathcal{G}) = \prod_{(jk\ell m)\in\Lambda_3} \delta_H(h_{jk\ell m})\,, \tag{4.8}$$

where for every tetrahedron $(jk\ell m) \in \Lambda_3$ one has:

$$\delta_H(h_{jk\ell m}) = \delta_H\big(\delta(l_{jk\ell m}) h_{j\ell m}\,(g_{\ell m} \triangleright h_{jk\ell})\, h_{k\ell m}^{-1}\, h_{jkm}^{-1}\big)\,. \tag{4.9}$$

Finally, the condition $\delta(\mathcal{H})$ is discretized as

$$\delta(\mathcal{H}) = \prod_{(jk\ell mn)\in\Lambda_4} \delta_L(l_{jk\ell mn})\,, \tag{4.10}$$

where for each 4-simplex $(jk\ell mn) \in \Lambda_4$ one has:

$$\delta_L(l_{jk\ell mn}) = \delta_L\big(l_{j\ell mn}^{-1}\, h_{j\ell n} \triangleright' \{h_{\ell mn}, (g_{mn} g_{\ell m}) \triangleright h_{jk\ell}\}_{\mathrm{p}}\, l_{jk\ell n}^{-1} (h_{jkn} \triangleright' l_{k\ell mn}) l_{jkmn} h_{jmn} \triangleright' (g_{mn} \triangleright l_{jk\ell m})\big)\,. \tag{4.11}$$

The identities (4.7), (4.9), and (4.11) are the results of Lemmas 3.1, 3.2, and 3.3, respectively.

After substituting the expressions for discretized measures (4.3)–(4.5) and δ-functions (4.6), (4.8), and (4.10) into the equation (4.2) one obtains:

$$Z = \mathcal{N} \prod_{(jk)\in\Lambda_1} \int_G dg_{jk} \prod_{(jk\ell)\in\Lambda_2} \int_H dh_{jk\ell} \prod_{(jk\ell m)\in\Lambda_3} \int_L dl_{jk\ell m} \Big(\prod_{(jk\ell)\in\Lambda_2} \delta_G(g_{jk\ell})\Big)\Big(\prod_{(jk\ell m)\in\Lambda_3} \delta_H(h_{jk\ell m})\Big)\Big(\prod_{(jk\ell mn)\in\Lambda_4} \delta_L(l_{jk\ell mn})\Big). \tag{4.12}$$

By inserting (4.7), (4.9), and (4.11) into (4.12), we obtain an explicit expression for the state sum over a given triangulation of the manifold $\mathcal{M}_4$. This expression can be made independent of the triangulation if one appropriately chooses the constant factor $\mathcal{N}$, obtained after the integration over the Lagrange multipliers B, C, and D. This is done by requiring that the state sum is invariant under the Pachner moves, which leads us to the appropriate form of the constant factor $\mathcal{N}$, as given by the definition 4.1.

Definition 4.1. Let $\mathcal{M}_4$ be a compact and oriented combinatorial d-manifold, $d = 4$, and $(L \xrightarrow{\delta} H \xrightarrow{\partial} G\,, \triangleright\,, \{\_\,,\_\}_{\mathrm{pf}})$ be a 2-crossed module. The state sum of *topological higher gauge theory* is defined by

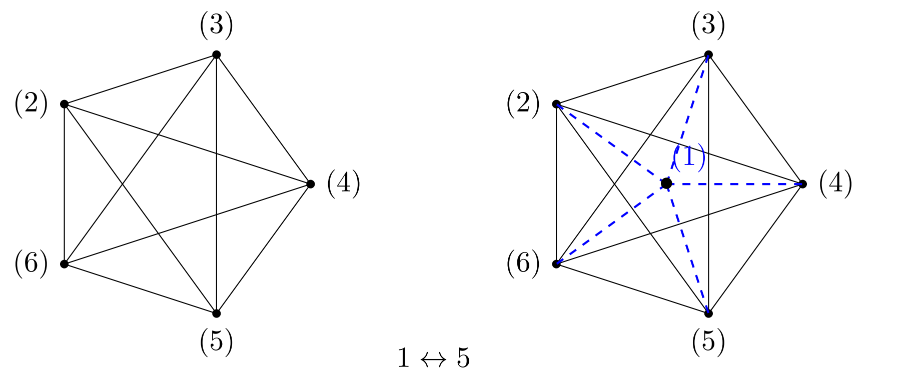
$$\begin{aligned} Z = {} & |G|^{-|\Lambda_0|+|\Lambda_1|-|\Lambda_2|} |H|^{|\Lambda_0|-|\Lambda_1|+|\Lambda_2|-|\Lambda_3|}\, |L|^{-|\Lambda_0|+|\Lambda_1|-|\Lambda_2|+|\Lambda_3|-|\Lambda_4|} \\ & \times\Big(\textstyle\prod_{(jk)\in\Lambda_1} \int_G dg_{jk}\Big)\Big(\prod_{(jk\ell)\in\Lambda_2} \int_H dh_{jk\ell}\Big)\Big(\prod_{(jk\ell m)\in\Lambda_3} \int_L dl_{jk\ell m}\Big) \\ & \times\Big(\textstyle\prod_{(jk\ell)\in\Lambda_2} \delta_G\big(\partial(h_{jk\ell})\, g_{k\ell}\, g_{jk}\, g_{j\ell}^{-1}\big)\Big)\Big(\prod_{(jk\ell m)\in\Lambda_3} \delta_H\big(\delta(l_{jk\ell m}) h_{j\ell m}\,(g_{\ell m} \triangleright h_{jk\ell})\, h_{k\ell m}^{-1}\, h_{jkm}^{-1}\big)\Big) \\ & \times\Big(\textstyle\prod_{(jk\ell mn)\in\Lambda_4} \delta_L\big(l_{j\ell mn}^{-1}\, h_{j\ell n} \triangleright' \{h_{\ell mn}, (g_{mn} g_{\ell m}) \triangleright h_{jk\ell}\}_{\mathrm{p}}\, l_{jk\ell n}^{-1} (h_{jkn} \triangleright' l_{k\ell mn}) l_{jkmn} h_{jmn} \triangleright' (g_{mn} \triangleright l_{jk\ell m})\big)\Big)\,. \end{aligned} \tag{4.13}$$

Here we integrate over $g_{jk} \in G$ for every edge $(jk) \in \Lambda_1$, over $h_{jk\ell} \in H$ for every triangle $(jk\ell) \in \Lambda_2$ and over l_{jklm} for every tetrahedron $(jk\ell m) \in \Lambda_3$. The δ-distributions under the integral impose the following conditions. First, the condition that $\partial(h_{jk\ell})\, g_{k\ell}\, g_{jk} = g_{j\ell}$ for each triangle $(jk\ell) \in \Lambda_2$, i.e., that each surface label $h_{jk\ell}$ has got the appropriate source and target, see Lemma 3.1. Second, the condition that $h_{jkm}\, h_{k\ell m} = \delta(l_{jk\ell m}) h_{j\ell m}\, (g_{\ell m} \rhd h_{jk\ell})$ for each tetrahedron $(jk\ell m) \in \Lambda_3$, i.e., that each volume label $l_{jk\ell m}$ has got the appropriate source and target, see Lemma 3.2. Finally, the condition that the volume holonomy around every 4-simplex $(jk\ell mn) \in \Lambda_4$ is trivial, i.e., that $l^{-1}_{j\ell mn}\, h_{j\ell n} \rhd' \{h_{\ell mn}, (g_{mn} g_{\ell m}) \rhd h_{jk\ell}\}_{\rm p}\, l^{-1}_{jk\ell n} (h_{jkn} \rhd' l_{k\ell mn}) l_{jkmn} h_{jmn} \rhd' (g_{mn} \rhd l_{jk\ell m})$ is equal to the neutral element of the group L for each 4-simplex $(jk\ell mn) \in \Lambda_4$, see Lemma 3.3.

Theorem 4.2. Let $\mathcal{M}_4$ be a closed and oriented combinatorial 4-manifold and $(L \xrightarrow{\delta} H \xrightarrow{\partial} G\,, \rhd\,, \{\_\,,\_\}_{\rm pf})$ be a 2-crossed module. The state sum (4.13) is invariant under Pachner moves.

The statements of Pachner move invariance are formulated in the following subsections, while corresponding proofs are given in the appendix B.

4.1 Pachner move $1 \leftrightarrow 5$



$1 \leftrightarrow 5$

Let us verify that the state sum (4.13) is invariant under $1-5$ Pachner move. Since the partition function is independent of the total order of vertices, let us fix the ordering and verify the move in only one case. Let us denote the vertices of the 4-simplex on the left hand side of the $1-5$ Pachner move as (23456). Then, adding a vertex 1 on the right hand side of the Pachner move one obtains five 4-simplices $M_4 = \{(13456), (12456), (12356), (12346), (12345)\}$. On the r.h.s. there are tetrahedrons $M_3 = \{(1234), (1235), (1236), (1245), (1246), (1256), (1345), (1346), (1356), (1456)\}$, triangles $(jk\ell) \in M_2 = \{(123), (124), (125), (126), (134), (135), (136), (145), (146), (156)\}$, edges $(jk) \in M_1 = \{(12), (13), (14), (15), (16)\}$ and vertices $(j) \in M_0 = \{(1)\}$. All other simplices are present on both sides of the move.

	$\|\Lambda_0\|$	$\|\Lambda_1\|$	$\|\Lambda_2\|$	$\|\Lambda_3\|$	$\|\Lambda_4\|$
l.h.s.	5	10	10	5	1
r.h.s.	6	15	20	15	5

Table 1. Number of vertices $|\Lambda_0|$, edges $|\Lambda_1|$, triangles $|\Lambda_2|$, tetrahedrons $|\Lambda_3|$, and 4-simplices $|\Lambda_4|$ on both sides of the $1 \leftrightarrow 5$ move.

If the $1-5$ Pachner move does not change the state sum (4.13), then the state sum of the right hand side,

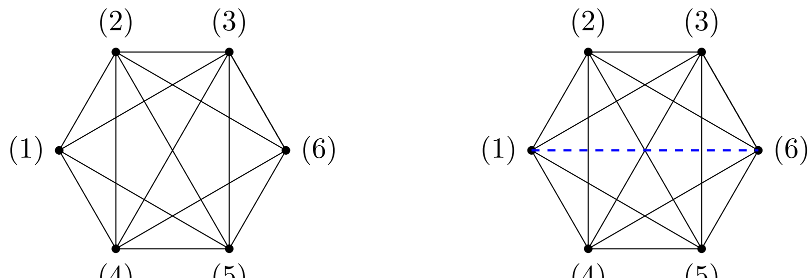
$$Z^{1\leftrightarrow 5}_{\text{right}} = |G|^{-11}|H|^{-4}|L|^{-1} \int_{G^5} \prod_{(jk)\in M_1} dg_{jk} \int_{H^{10}} \prod_{(jk\ell)\in M_2} dh_{jk\ell} \int_{L^{10}} \prod_{(jklm)\in M_3} dl_{jklm}$$
$$\cdot \left(\prod_{(jk\ell)\in M_2} \delta_G(g_{jk\ell}) \right) \left(\prod_{(jk\ell m)\in M_3} \delta_H(h_{jk\ell m}) \right) \left(\prod_{(jk\ell mn)\in M_4} \delta_L(l_{jk\ell mn}) \right) Z_{\text{remainder}}\,, \tag{4.14}$$

should be equal to the state sum of the left hand side,

$$Z^{1\leftrightarrow 5}_{\text{left}} = |G|^{-5}|H|^{0}|L|^{-1}\delta_L(l_{23456}) Z_{\text{remainder}}\,. \tag{4.15}$$

Here, the prefactors $|G|^{-|\Lambda_0|+|\Lambda_1|-|\Lambda_2|}$, $|H|^{|\Lambda_0|-|\Lambda_1|+|\Lambda_2|-|\Lambda_3|}$, and $|L|^{-|\Lambda_0|+|\Lambda_1|-|\Lambda_2|+|\Lambda_3|-|\Lambda_4|}$ are $|G|^{-11}|H|^{-4}|L|^{-1}$ on the r.h.s. and $|G|^{-5}|H|^{0}|L|^{-1}$ on the l.h.s., as obtained by counting the numbers of the k-simplices on both sides of the $1-5$ move, shown in the table 1. The $Z_{\text{remainder}}$ denotes the part of the state sum that is the same on both sides of the move, and thus irrelevant for the proof of invariance. The proof that $Z_{\text{left}} = Z_{\text{right}}$ is given in the appendix B.

4.2 Pachner move $2 \leftrightarrow 4$



$2 \leftrightarrow 4$

In order to verify the state sum (4.13) invariance under $2-4$ Pachner move, we order the vertices in such a way that on the l.h.s. of the move we have two 4-simplices $M_4^{\text{left}} = \{(23456),(12345)\}$, while on the r.h.s. we have four 4-simplices $M_4^{\text{right}} = \{(12346),(12356),(12456),(13456)\}$. On the l.h.s. we have one tetrahedron $M_3^{\text{left}} = \{(2345)\}$, whereas on the r.h.s. there are six tetrahedrons $M_3^{\text{right}} = \{(1236),(1246),(1256),(1346),(1356),(1456)\}$. All other tetrahedrons appear on both sides of the move. On the r.h.s. there are triangles $M_2^{\text{right}} = \{(126),(136),(146),(156)\}$, and one edge $M_1^{\text{right}} = \{(16)\}$, while the rest of the triangles and edges appear on both sides of the move.

JHEP07(2022)105

	$\vert\Lambda_0\vert$	$\vert\Lambda_1\vert$	$\vert\Lambda_2\vert$	$\vert\Lambda_3\vert$	$\vert\Lambda_4\vert$
l.h.s.	6	14	16	9	2
r.h.s.	6	15	20	14	4

Table 2. Number of vertices $|\Lambda_0|$, edges $|\Lambda_1|$, triangles $|\Lambda_2|$, tetrahedrons $|\Lambda_3|$, and 4-simplices $|\Lambda_4|$ on both sides of the $2 \leftrightarrow 4$ move.

On the l.h.s. there is the state sum,

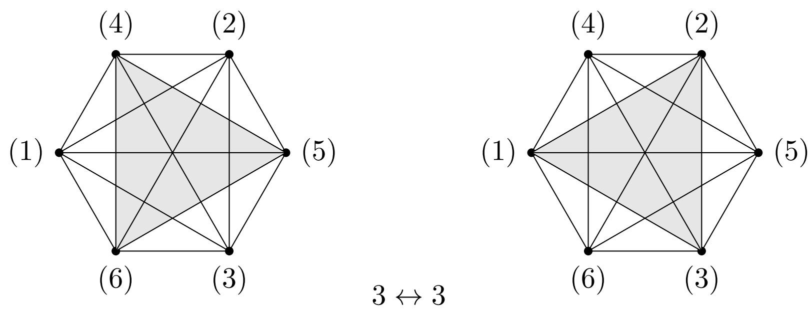
$$Z^{2\leftrightarrow 4}_{\text{left}} = |G|^{-8}|H|^{-1}|L|^{-1} \int_L dl_{2345}\delta_H(h_{2345}) \Bigg(\prod_{(jk\ell mn)\in M_4^{\text{left}}} \delta_L(l_{jk\ell mn}) \Bigg) Z_{\text{remainder}} \,, \tag{4.16}$$

whereas on the r.h.s. the state sum reads:

$$Z^{2\leftrightarrow 4}_{\text{right}} = |G|^{-11}|H|^{-3}|L|^{-1} \int_G dg_{16} \int_{H^4} dh_{126}dh_{136}dh_{146}dh_{156} \int_L dl_{1236}dl_{1246}dl_{1256}dl_{1346}dl_{1356}dl_{1456}$$
$$\Bigg(\prod_{(jk\ell)\in M_2^{\text{right}}} \delta_G(g_{jk\ell}) \Bigg) \Bigg(\prod_{(jk\ell m)\in M_3^{\text{right}}} \delta_H(h_{jk\ell m}) \Bigg) \Bigg(\prod_{(jk\ell mn)\in M_4^{\text{right}}} \delta_L(l_{jk\ell mn}) \Bigg) Z_{\text{remainder}} . \tag{4.17}$$

Here the prefactors $|G|^{-8}|H|^{-1}|L|^{-1}$ on the l.h.s. and $|G|^{-11}|H|^{-3}|L|^{-1}$ on the r.h.s. are obtained by counting the numbers of k-simplices on both sides of the $2-4$ move, as shown in the table 2. The term $Z_{\text{remainder}}$ denotes the part of the state sum that is identical on both sides of the move, as before. The proof that $Z_{\text{left}} = Z_{\text{right}}$ is given in the appendix B.

4.3 Pachner move $3 \leftrightarrow 3$



$3 \leftrightarrow 3$

In order to verify the state sum invariance under $3-3$ Pachner move, we order the vertices in such a way that on the l.h.s. of the $3-3$ move, we have three 4-simplices $M_4^{\text{left}} = \{(23456), (13456), (12456)\}$, whereas on the r.h.s. we have the 4-simplices $M_4^{\text{right}} = \{(12356), (12346), (12345)\}$. On the l.h.s. there are tetrahedrons $M_3^{\text{left}} = \{(1456), (2456), (3456)\}$, and on the r.h.s. $M_3^{\text{right}} = \{(1234), (1235), (1236)\}$. One notices that the six tetrahedrons form the common boundary of both sides of the move, whereas on each side there are three tetrahedrons shared by two 4-simplices. On the l.h.s. one has the triangle $M_2^{\text{left}} = \{(456)\}$ and on the r.h.s. the triangle $M_3^{\text{right}} = \{(123)\}$. All other triangles appear on both sides of the move.

Therefore on the l.h.s. there is the state sum,

$$Z^{3\leftrightarrow 3}_{\text{left}} = \int_H dh_{456} \int_{L^3} dl_{1456} dl_{2456} dl_{3456} \delta_G(g_{456})\, \delta_H(h_{3456}) \delta_H(h_{2456}) \delta_H(h_{1456}) \\ \delta_L(l_{23456}) \delta_L(l_{13456}) \delta_L(l_{12456}) Z_{\text{remainder}}\,, \tag{4.18}$$

whereas on the r.h.s. the state sum reads

$$Z^{3\leftrightarrow 3}_{\text{right}} = \int_H dh_{123} \int_{L^3} dl_{1234} dl_{1235} dl_{1236} \delta_G(g_{123})\, \delta_H(h_{1234}) \delta_H(h_{1235}) \delta_H(h_{1236}) \\ \delta_L(l_{12356}) \delta_L(l_{12346}) \delta_L(l_{12345}) Z_{\text{remainder}}\,. \tag{4.19}$$

The numbers of k-simplices agree on both sides of the $3-3$ move for all k, and the prefactors play no role in this case, therefore they are part of the $Z_{\text{remainder}}$. The proof that $Z_{\text{left}} = Z_{\text{right}}$ is given in the appendix B.

We obtain that the state sum given by the definition 4.1 is invariant under all three Pachner moves, and thus independent of triangulation of the underlying 4-dimensional manifold (see appendix B for the proof).

5 Conclusions

Let us summarize the results of the paper. In section 2 we reviewed the pure the constrained $2BF$ actions describing the Yang-Mills field and Einstein-Cartan gravity, and constrained $3BF$ actions describing the Klein-Gordon and Dirac fields coupled to Yang-Mills fields and gravity in the standard way. In section 3, we reviewed the relevant algebraic tools involved in the description of higher gauge theory, 2-crossed modules, and 3-gauge theory and generalized the integral picture of an ordinary gauge theory to a 3-gauge theory that involves curves, surfaces, and volumes labeled with elements of non-Abelian groups. We have also proved three key results, stated in Lemmas 3.1, 3.2, and 3.3, which are crucial for the construction of the invariant state sum. In section 4, we have presented the two main results of the paper. First, we constructed a triangulation independent state sum Z of a topological higher gauge theory for a general 3-group and a 4-dimensional spacetime manifold $\mathcal{M}_4$. Second, we proved the theorem that the constructed state sum is indeed independent of the choice of triangulation, i.e., that it is a genuine topological invariant.

The constructed state sum coincides with Porter's TQFT [33, 34] for $d = 4$ and $n = 3$. The proof that the state sum is invariant under the local changes of triangulation called the Pachner moves and thus independent of the chosen triangulation is presented in appendix B. It is obtained that the state sum is invariant under all five different Pachner moves: the $3-3$ move, $4-2$ move, and $5-1$ move, and their inverses. The state sum constructed this way can be thought of as a combinatorial construction of a topological quantum field theory (TQFT) in the sense of Atiyah's axioms, a topic that is beyond the scope of this paper and will be studied in a future work.

In order to finish the second step of the spinfoam quantization procedure, however, the generalizations of the Peter-Weyl and Plancharel theorems to 2-groups and 3-groups are required, which so far represent open problems. Namely, these theorems should provide a decomposition of a function on a 3-group into a sum over the corresponding irreducible representations of a 3-group. In this way, the spectrum of labels for the simplices, i.e. , the domain of values of the fields living on the simplices of the triangulation, would be specified. Nonetheless, one can still try to guess the irreducible representations of 3-groups, as was done for example in the case of 2-groups in the spincube model of quantum gravity [30], or obtain the state sum using other techniques, see for example [50–52]).

However, if one wants to describe a real physical theory, i.e. , the theory which contains local propagating degrees of freedom, one needs to construct the nontopological state sum, with the non-trivial dynamics. To do so, once the topological state sum is constructed, the final third step of the spinfoam quantization procedure is to impose the constraints that deform the topological theory into a realistic theory of gravity coupled to matter fields (as defined in [31]) at the quantum level. We leave the construction of the constrained state sum model for future work.

In addition to the above topics, there are also many other possible applications of the invariant state sum, both in physics and mathematics.

Acknowledgments

This research was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia (MPNTR), and by the Science Fund of the Republic of Serbia, Program DIASPORA, No. 6427195, SQ2020. The contents of this publication are the sole responsibility of the authors and can in no way be taken to reflect the views of the Science Fund of the Republic of Serbia.

A Proof of the invariance identity

Let us prove the identity (3.16). Using the definitions of the upward composition (3.14) and the vertical composition (3.15) of the 3-morphisms, one obtains that the left-hand side of the equation (3.16) is equal to:

$$\begin{aligned}((g_2,h'_2,l'_2)\#_3(g_2,h_2,l_2))\#_2((g_1,h'_1,l'_1)\#_3(g_1,h_1,l_1)) &= (g_2,h_2,l'_2l_2)\#_2(g_1,h_1,l'_1l_1)\\ &= (g_1,h_2h_1,l'_2l_2\,h_2\rhd'(l'_1l_1))\,.\end{aligned} \tag{A.1}$$

The right-hand side of the equation (3.16) is equal to:

$$\begin{aligned}((g_2,h'_2,l'_2)\#_2(g_1,h'_1,l'_1))\#_3((g_2,h_2,l_2)\#_2(g_1,h_1,l_1)) &= (g_1,h'_2h'_1,l'_2h'_2\rhd' l'_1)\#_3(g_1,h_2h_1,l_2h_2\rhd' l_1)\\ &= (g_1,h_2h_1,l'_2\,h'_2\rhd' l'_1\,l_2\,h_2\rhd' l_1)\quad (h'_2=\delta(l_2)h_2)\\ &= (g_1,h_2h_1,l'_2\,(\delta(l_2)h_2)\rhd' l'_1\,l_2\,h_2\rhd' l_1)\quad \text{eq. (A.3)}\\ &= (g_1,h_2h_1,l'_2\,\delta(l_2)\rhd'(h_2\rhd' l'_1)\,l_2\,h_2\rhd' l_1)\quad \text{(Peiffer identity)}\\ &= (g_1,h_2h_1,l'_2\,l_2(h_2\rhd' l'_1)l_2^{-1}\,l_2\,h_2\rhd' l_1)\quad (l_2^{-1}l_2=e)\\ &= (g_1,h_2h_1,l'_2\,l_2h_2\rhd' l'_1\,h_2\rhd' l_1)\quad \text{eq. (A.4)}\\ &= (g_1,h_2h_1,l'_2\,l_2h_2\rhd'(l'_1l_1))\,,\end{aligned} \tag{A.2}$$

where in the third and sixth line we use the identities

$$(h_1 h_2) \triangleright' l = h_1 \triangleright' (h_2 \triangleright' l), \qquad \forall h_1, h_2 \in H, \qquad \forall l \in L\,, \tag{A.3}$$
$$h \triangleright' (l_1 l_2) = h \triangleright' l_1 \; h \triangleright' l_2, \qquad \forall h \in H, \qquad \forall l_1, l_2 \in L\,. \tag{A.4}$$

This proves the equation (3.16).

B Proof of Pachner move invariance

In this section, a self contained proof in terms of Pachner moves that the partition function (4.13) is independent of the chosen triangulation is presented.

B.1 Pachner move $1 \leftrightarrow 5$

On the *left hand side of the move* there is the integrand $\delta_L(l_{23456})$:

$$\delta_L(l_{23456}) = \delta_L(l_{2346}{}^{-1}(h_{236} \triangleright' l_{3456}) l_{2356} h_{256} \triangleright' (g_{56} \triangleright l_{2345}) l_{2456}{}^{-1} h_{246} \triangleright' \{h_{456}, (g_{56} g_{45}) \triangleright h_{234}\}_{\mathrm{p}}). \tag{B.1}$$

Let us examine the *right hand side of the move*, given by the equation (4.14). First, one integrates out g_{12} using $\delta_G(g_{123})$, g_{13} using $\delta_G(g_{134})$, g_{14} using $\delta_G(g_{145})$, and g_{15} using $\delta_G(g_{156})$, and obtains:

$$\begin{aligned} g_{12} &= g_{23}^{-1}\, \partial(h_{123})^{-1}\, g_{13}\,, \\ g_{13} &= g_{34}^{-1}\, \partial(h_{134})^{-1}\, g_{14}\,, \\ g_{14} &= g_{45}^{-1}\, \partial(h_{145})^{-1}\, g_{15}\,, \\ g_{15} &= g_{56}^{-1}\, \partial(h_{156})^{-1}\, g_{16}\,. \end{aligned} \tag{B.2}$$

One integrates out h_{123} using $\delta_H(h_{1234})$, h_{124} using $\delta_H(h_{1245})$, h_{125} using $\delta_H(h_{1256})$, h_{134} using $\delta_H(h_{1345})$, h_{135} using $\delta_H(h_{1356})$, and h_{145} using $\delta_H(h_{1456})$, and obtains:

$$\begin{aligned} h_{123} &= g_{34}^{-1} \triangleright h_{134}^{-1}\, g_{34}^{-1} \triangleright \delta(l_{1234})^{-1}\, g_{34}^{-1} \triangleright h_{124}\, g_{34}^{-1} \triangleright h_{234}\,, \\ h_{124} &= g_{45}^{-1} \triangleright h_{145}^{-1}\, g_{45}^{-1} \triangleright \delta(l_{1245})^{-1}\, g_{45}^{-1} \triangleright h_{125}\, g_{45}^{-1} \triangleright h_{245}\,, \\ h_{125} &= g_{56}^{-1} \triangleright h_{156}^{-1}\, g_{56}^{-1} \triangleright \delta(l_{1256})^{-1}\, g_{56}^{-1} \triangleright h_{126}\, g_{56}^{-1} \triangleright h_{256}\,, \\ h_{134} &= g_{45}^{-1} \triangleright h_{145}^{-1}\, g_{45}^{-1} \triangleright \delta(l_{1345})^{-1}\, g_{45}^{-1} \triangleright h_{135}\, g_{45}^{-1} \triangleright h_{345}\,, \\ h_{135} &= g_{56}^{-1} \triangleright h_{156}^{-1}\, g_{56}^{-1} \triangleright \delta(l_{1356})^{-1}\, g_{56}^{-1} \triangleright h_{136}\, g_{56}^{-1} \triangleright h_{356}\,, \\ h_{145} &= g_{56}^{-1} \triangleright h_{156}^{-1}\, g_{56}^{-1} \triangleright \delta(l_{1456})^{-1}\, g_{56}^{-1} \triangleright h_{146}\, g_{56}^{-1} \triangleright h_{456}\,. \end{aligned} \tag{B.3}$$

The δ-functions on the group G now read $\delta_G(e)^6$. Let us show this. First, for $\delta_G(g_{124})$ one obtains

$$\begin{aligned} \delta_G(g_{124}) &= \delta_G\Big(\partial(h_{124})\, g_{24}\, g_{12}\, g_{14}^{-1}\Big) \\ &= \delta_G\Big(\partial(h_{124})\, g_{24}\, g_{23}^{-1}\, \partial(h_{123})^{-1}\, g_{13}\, g_{14}^{-1}\Big) \\ &= \delta_G\Big(\partial(h_{124})\, g_{24}\, g_{23}^{-1}\, g_{34}^{-1}\, \partial(h_{234})^{-1} \partial(h_{124})^{-1} \partial(h_{134})\, g_{34}\, g_{13}\, g_{14}^{-1}\Big) \\ &= \delta_G\Big(\partial(h_{124})\, g_{24}\, g_{23}^{-1}\, g_{34}^{-1}\, (g_{34}\, g_{23}^{-1}\, g_{24}^{-1})\, \partial(h_{124})^{-1}\, e\Big) \\ &= \delta_G(e)\,, \end{aligned} \tag{B.4}$$

Next, for δ-function $\delta_G(g_{125})$ one obtains,

$$\begin{aligned}
\delta_G(g_{125}) &= \delta_G\Big(\partial(h_{125})\, g_{25}\, g_{12}\, g_{15}^{-1}\Big)\,,\\
&= \delta_G\Big(\partial(h_{125})\, g_{25}\, g_{23}^{-1}\, \partial(h_{123})^{-1}\, g_{13}\, g_{15}^{-1}\Big)\\
&= \delta_G\Big(\partial(h_{125})\, g_{25}\, g_{23}^{-1}\, g_{34}^{-1}\, \partial(h_{234})^{-1}\partial(h_{124})^{-1}\partial(h_{134})\, g_{34}\, g_{13}\, g_{15}^{-1}\Big)\\
&= \delta_G\Big(\partial(h_{125})\, g_{25}\, g_{23}^{-1}\, g_{34}^{-1}\, \partial(h_{234})^{-1} g_{45}^{-1}(\partial(h_{245})^{-1}\partial(h_{125})^{-1}\partial(h_{145}))\, g_{45} g_{14}\, g_{15}^{-1}\Big)\\
&= \delta_G\Big(\partial(h_{125})\, g_{25}\, g_{23}^{-1}\, g_{34}^{-1}\, (g_{34}\, g_{23}^{-1}\, g_{24}^{-1}) g_{45}^{-1} (g_{45}\, g_{24}^{-1}\, g_{25}^{-1})\partial(h_{125})^{-1} e\Big)\\
&= \delta_G(e)\,.
\end{aligned} \tag{B.5}$$

Similarly, $\delta_G(g_{126})$ becomes

$$\begin{aligned}
\delta_G(g_{126}) &= \delta_G(\partial(h_{126}) g_{26} g_{12} g_{16}^{-1})\\
&= \delta_G(\partial(h_{126}) g_{26} g_{23}^{-1} \partial(h_{123})^{-1} g_{13} g_{16}^{-1})\\
&= \delta_G(\partial(h_{126}) g_{26} g_{23}^{-1} g_{34}^{-1} \partial(h_{234})^{-1} \partial(h_{124})^{-1} \partial(h_{134}) g_{34} g_{13} g_{16}^{-1})\\
&= \delta_G(\partial(h_{126}) g_{26} g_{23}^{-1} g_{34}^{-1} \partial(h_{234})^{-1} g_{45}^{-1} (\partial(h_{245})^{-1} \partial(h_{125})^{-1} \partial(h_{145})) g_{45} \partial(h_{134}) g_{34} g_{13} g_{16}^{-1})\\
&= \delta_G(\partial(h_{126}) g_{26} g_{23}^{-1} g_{34}^{-1} \partial(h_{234})^{-1} g_{45}^{-1} (\partial(h_{245})^{-1} g_{56}^{-1} \partial(h_{256})^{-1} \partial(h_{126})^{-1} \partial(h_{156}) g_{56}\\
&\quad \partial(h_{145})) g_{45} g_{14} g_{16}^{-1})\\
&= \delta_G(\partial(h_{126}) g_{26} g_{23}^{-1} g_{34}^{-1} (g_{34} g_{23}^{-1} g_{24}^{-1}) g_{45}^{-1} (g_{45} g_{24}^{-1} g_{25}^{-1}) g_{56}^{-1} (g_{56} g_{25}^{-1} g_{26}^{-1}) \partial(h_{126})^{-1}\\
&\quad (g_{16} g_{15}^{-1} g_{56}^{-1}) g_{56} g_{15} g_{16}^{-1})\\
&= \delta_G(e),
\end{aligned} \tag{B.6}$$

and $\delta_G(g_{135})$ now reads,

$$\begin{aligned}
\delta_G(g_{135}) &= \delta_G\Big(\partial(h_{135})\, g_{35}\, g_{13}\, g_{15}^{-1}\Big)\,,\\
&= \delta_G\Big(\partial(h_{135})\, g_{35}\, g_{34}^{-1}\, \partial(h_{134})^{-1}\, g_{14}\, g_{15}^{-1}\Big)\\
&= \delta_G\Big(\partial(h_{135})\, g_{35}\, g_{34}^{-1}\, g_{45}^{-1}\, \partial(h_{345})^{-1}\partial(h_{135})^{-1}\partial(h_{145})\, g_{45}\, g_{14}\, g_{15}^{-1}\Big)\\
&= \delta_G\Big(\partial(h_{135})\, g_{35}\, g_{34}^{-1}\, g_{45}^{-1}\, \partial(h_{345})^{-1}\partial(h_{135})^{-1}\, \partial(h_{145})\, g_{45}\, g_{45}^{-1}\, \partial(h_{145})^{-1}\, g_{15}\, g_{15}^{-1}\Big)\\
&= \delta_G\Big(\partial(h_{135})\, g_{35}\, g_{34}^{-1}\, g_{45}^{-1}\, (g_{45}\, g_{34}^{-1}\, g_{35}^{-1})\partial(h_{135})^{-1}\Big)\\
&= \delta_G(e)\,,
\end{aligned} \tag{B.7}$$

while $\delta_G(g_{136})$ reads:

$$\begin{aligned}\delta_G(g_{136}) &= \delta_G(\partial(h_{136})\, g_{36}\, g_{13}\, g_{16}^{-1}) \\ &= \delta_G(\partial(h_{136})\, g_{36}\, g_{34}^{-1}\, \partial(h_{134})^{-1}\, g_{14}\, g_{16}^{-1}) \\ &= \delta_G(\partial(h_{136})\, g_{36}\, g_{34}^{-1}\, g_{45}^{-1}\, \partial(h_{345})^{-1}\partial(h_{135})^{-1}\partial(h_{145})\, g_{45}\, g_{14}\, g_{16}^{-1}) \\ &= \delta_G(\partial(h_{136})\, g_{36}\, g_{34}^{-1}\, g_{45}^{-1}\, \partial(h_{345})^{-1} g_{56}^{-1}(\partial(h_{356})^{-1}\partial(h_{136})^{-1}\partial(h_{156}))\, g_{56}\partial(h_{145})\, g_{45}\, g_{14}\, g_{16}^{-1}) \\ &= \delta_G(\partial(h_{136})\, g_{36}\, g_{34}^{-1}\, g_{45}^{-1}\, (g_{45}\, g_{34}^{-1}\, g_{35}^{-1}) g_{56}^{-1} (g_{56}\, g_{35}^{-1}\, g_{36}^{-1})\partial(h_{136})^{-1} e) \\ &= \delta_G(e)\,. \end{aligned} \tag{B.8}$$

Finally, the δ-function $\delta_G(g_{146})$ reads:

$$\begin{aligned}\delta_G(g_{146}) &= \delta_G\Big(\partial(h_{146})\, g_{46}\, g_{14}\, g_{16}^{-1}\Big) \\ &= \delta_G\Big(\partial(h_{146})\, g_{46}\, (g_{45}^{-1}\, \partial(h_{145})^{-1}\, g_{15})\, g_{16}^{-1}\Big) \\ &= \delta_G\Big(\partial(h_{146})\, g_{46}\, g_{45}^{-1}\, \partial(h_{145})^{-1}\, (g_{56}^{-1}\, \partial(h_{156})^{-1}\, g_{16})\, g_{16}^{-1}\Big) \\ &= \delta_G\Big(\partial(h_{146})\, g_{46}\, g_{45}^{-1}\, g_{56}^{-1}\partial(h_{456})^{-1}\partial(h_{146})^{-1}\partial(h_{156}) g_{56}\, (g_{56}^{-1}\, \partial(h_{156})^{-1}\, g_{16})\, g_{16}^{-1}\Big) \\ &= \delta_G(e)\,. \end{aligned} \tag{B.9}$$

Next, one integrates out l_{1235} using $\delta_L(l_{12345})$, l_{1236} using $\delta_L(l_{12346})$, l_{1246} using $\delta_L(l_{12456})$, and l_{1346} using $\delta_L(l_{13456})$, and obtains

$$l_{1235} = (h_{125} \rhd' l_{2345}) l_{1245} h_{145} \rhd' (g_{45} \rhd l_{1234}) l_{1345}^{-1}\, h_{135} \rhd' \{h_{345}, (g_{45} g_{34}) \rhd h_{123}\}_{\mathrm{p}}\,, \tag{B.10}$$

$$l_{1236} = (h_{126} \rhd' l_{2346}) l_{1246} h_{146} \rhd' (g_{46} \rhd l_{1234}) l_{1346}^{-1}\, h_{136} \rhd' \{h_{346}, (g_{46} g_{34}) \rhd h_{123}\}_{\mathrm{p}}\,, \tag{B.11}$$

$$l_{1246} = (h_{126} \rhd' l_{2456}) l_{1256} h_{156} \rhd' (g_{56} \rhd l_{1245}) {l_{1456}}^{-1}\, h_{146} \rhd' \{h_{456}, (g_{56} g_{45}) \rhd h_{124}\}_{\mathrm{p}}\,, \tag{B.12}$$

$$l_{1346} = (h_{136} \rhd' l_{3456}) l_{1356} h_{156} \rhd' (g_{56} \rhd l_{1345}) {l_{1456}}^{-1}\, h_{146} \rhd' \{h_{456}, (g_{56} g_{45}) \rhd h_{134}\}_{\mathrm{p}}\,. \tag{B.13}$$

Let us now show that the remaining δ-functions on the group H equal $\delta_H(e)^4$. First, $\delta_H(h_{1235})$ becomes:

$$\begin{aligned}\delta_H(h_{1235}) =& \delta_H(\delta(l_{1235}) h_{135} (g_{35} \rhd h_{123}) h_{235}^{-1} h_{125}^{-1}) \\ =& \delta_H\Big(\delta((h_{125} \rhd' l_{2345}) l_{1245} h_{145} \rhd' (g_{45} \rhd l_{1234}) l_{1345}^{-1} h_{135} \rhd' \{h_{345}, (g_{45} g_{34}) \rhd h_{123}\}_{\mathrm{p}}) h_{135} (g_{35} \rhd h_{123}) h_{235}^{-1} h_{125}^{-1}\Big) \\ =& \delta_H\Big((h_{125} \delta(l_{2345}) h_{125}^{-1} \delta(l_{1245}) h_{145} (g_{45} \rhd \delta(l_{1234})) h_{145}^{-1} \delta(l_{1345})^{-1} h_{135} \delta(\{h_{345}, (g_{45} g_{34}) \rhd h_{123}\}_{\mathrm{p}}) h_{135}^{-1}) \\ & h_{135} (g_{35} \rhd h_{123}) h_{235}^{-1} h_{125}^{-1}\Big) \\ =& \delta_H\Big(h_{235} h_{345} (g_{45} \rhd h_{234}^{-1}) h_{245}^{-1} h_{125}^{-1} h_{125} h_{245} (g_{45} \rhd h_{124}^{-1}) h_{145}^{-1} h_{145} (g_{45} \rhd (h_{124} h_{234} (g_{34} \rhd h_{123}^{-1}) h_{134}^{-1})) \\ & h_{145}^{-1} (h_{145} (g_{45} \rhd h_{134}) h_{345}^{-1} h_{135}^{-1}) h_{135} \delta(\{h_{345}, (g_{45} g_{34}) \rhd h_{123}\}_{\mathrm{p}}) h_{135}^{-1} h_{135} (g_{35} \rhd h_{123}) h_{235}^{-1}\Big) \\ =& \delta_H(h_{345} ((g_{45} g_{34}) \rhd h_{123}^{-1}) h_{345}^{-1} \delta(\{h_{345}, (g_{45} g_{34}) \rhd h_{123}\}_{\mathrm{p}}) (g_{35} \rhd h_{123}). \end{aligned} \tag{B.14}$$

Here, one uses the following identity

$$\delta\{h_1\,, h_2\}_{\mathrm{p}} (\partial(h_1) \rhd h_2) h_1 h_2^{-1} h_1^{-1} = e\,. \tag{B.15}$$

Substituting $g_{35} = \partial(h_{345})g_{45}g_{34}$, and applying the (B.15) identity for $h_1 = h_{345}$ and $h_2 = (g_{45}g_{34}) \triangleright h_{123}$, one obtains

$$\delta_H(h_{1235}) = \delta_H(e). \tag{B.16}$$

Similarly, one obtains for $\delta_H(h_{1236})$:

$$\begin{aligned}
\delta_H(h_{1236}) =& \delta_H(\delta(l_{1236})h_{136}(g_{36}\triangleright h_{123})h_{236}^{-1}h_{126}^{-1}) \\
=& \delta_H\Big(\delta((h_{126}\triangleright' l_{2346})l_{1246}h_{146}\triangleright'(g_{46}\triangleright l_{1236}l_{1346}^{-1}h_{136}\triangleright'\{h_{346},(g_{46}g_{34})\triangleright h_{123}\}_{\rm p})h_{136}(g_{36}\triangleright h_{123})h_{236}^{-1}h_{126}^{-1}\Big) \\
=& \delta_H\Big((h_{126}\delta(l_{2346})h_{126}^{-1}\delta(l_{1246})h_{146}(g_{46}\triangleright\delta(l_{1234}))h_{146}^{-1}\delta(l_{1346})^{-1}h_{136}\delta(\{h_{346},(g_{46}g_{34})\triangleright h_{123}\}_{\rm p})h_{136}^{-1}) \\
& h_{136}(g_{36}\triangleright h_{123})h_{236}^{-1}h_{126}^{-1}\Big) \\
=& \delta_H\Big(h_{236}h_{346}(g_{46}\triangleright h_{234}^{-1})h_{246}^{-1}h_{126}^{-1}h_{126}h_{246}(g_{46}\triangleright h_{124}^{-1})h_{146}^{-1}h_{146}(g_{46}\triangleright(h_{124}h_{234}(g_{34}\triangleright h_{123}^{-1})h_{134}^{-1})) \\
& h_{146}^{-1}(h_{146}(g_{46}\triangleright h_{134})h_{346}^{-1}h_{136}^{-1})h_{136}\delta(\{h_{346},(g_{46}g_{34})\triangleright h_{123}\}_{\rm p})h_{136}^{-1}h_{136}(g_{36}\triangleright h_{123})h_{236}^{-1}\Big) \\
=& \delta_H(h_{346}((g_{46}g_{34})\triangleright h_{123}^{-1})h_{346}^{-1}\delta(\{h_{346},(g_{46}g_{34})\triangleright h_{123}\}_{\rm p})(g_{36}\triangleright h_{123}).
\end{aligned} \tag{B.17}$$

Substituting $g_{36} = \partial(h_{346})g_{46}g_{34}$, and applying the (B.15) identity for $h_1 = h_{346}$ and $h_2 = (g_{46}g_{34}) \triangleright h_{123}$, one obtains

$$\delta_H(h_{1236}) = \delta_H(e)\,. \tag{B.18}$$

Similarly, one obtains that $\delta_H(h_{1246}) = \delta_H(h_{1346}) = \delta_H(e)$. The remaining δ-function on the group L $\delta_L(l_{12356})$ reads:

$$\delta_L(l_{12356}) = \delta_L\big({l_{1236}}^{-1}(h_{126}\triangleright' l_{2356})l_{1256}h_{156}\triangleright'(g_{56}\triangleright l_{1235}){l_{1356}}^{-1}h_{136}\triangleright'\{h_{356},(g_{56}g_{35})\triangleright h_{123}\}_{\rm p}\big)\,. \tag{B.19}$$

After substituting the equations (B.10), (B.11), (B.12), and (B.13), one obtains:

$$\begin{aligned}
\delta_L(l_{12356}) =& \delta_L\Big(h_{136}\triangleright'\{h_{346},(g_{46}g_{34})\triangleright h_{123}\}_{\rm p}^{-1}(h_{136}\triangleright' l_{3456})l_{1356}h_{156}\triangleright'(g_{56}\triangleright l_{1345}){l_{1456}}^{-1} \\
& h_{146}\triangleright'\{h_{456},(g_{56}g_{45})\triangleright h_{134}\}_{\rm p}h_{146}\triangleright'(g_{46}\triangleright l_{1234})^{-1}h_{146}\triangleright'\{h_{456},(g_{56}g_{45})\triangleright h_{124}\}_{\rm p}^{-1}l_{1456} \\
& h_{156}\triangleright'(g_{56}\triangleright l_{1245})^{-1}l_{1256}^{-1}(h_{126}\triangleright' l_{2456})^{-1}(h_{126}\triangleright' {l_{2346}}^{-1})(h_{126}\triangleright' l_{2356})l_{1256} \\
& h_{156}\triangleright'(g_{56}\triangleright((h_{125}\triangleright' l_{2345})l_{1245}h_{145}\triangleright'(g_{45}\triangleright l_{1234})l_{1345}^{-1}h_{135}\triangleright'\{h_{345},(g_{45}g_{34})\triangleright h_{123}\}_{\rm p})) \\
& {l_{1356}}^{-1}h_{136}\triangleright'\{h_{356},(g_{56}g_{35})\triangleright h_{123}\}_{\rm p}\Big)\,.
\end{aligned} \tag{B.20}$$

Using the identity (3.4) the delta function $\delta_L(l_{12356})$ becomes:

$$\begin{aligned}
\delta_L(l_{12356}) =& \delta_L\Big((h_{136}\triangleright' l_{3456})l_{1356}h_{156}\triangleright'(g_{56}\triangleright l_{1345}){l_{1456}}^{-1} \\
& h_{146}\triangleright'\{h_{456},(g_{56}g_{45})\triangleright h_{134}\}_{\rm p}h_{146}\triangleright'(g_{46}\triangleright l_{1234})^{-1}h_{146}\triangleright'\{h_{456},(g_{56}g_{45})\triangleright h_{124}\}_{\rm p}^{-1}l_{1456} \\
& \delta(h_{156}\triangleright'(g_{56}\triangleright l_{1245})^{-1})\triangleright'\Big((\delta(l_{1256})^{-1}h_{126})\triangleright'(l_{2456}^{-1}l_{2346}^{-1}l_{2356})h_{156}\triangleright'(g_{56}\triangleright(h_{125}\triangleright' l_{2345}))\Big) \\
& h_{156}\triangleright'(g_{56}\triangleright(h_{145}\triangleright'(g_{45}\triangleright l_{1234})l_{1345}^{-1})){l_{1356}}^{-1}(h_{136}h_{346})\triangleright'\{h_{346}^{-1}h_{356}g_{56}\triangleright h_{345}, \\
& (g_{56}g_{45}g_{34})\triangleright h_{123}\}_{\rm p}\Big)\,.
\end{aligned} \tag{B.21}$$

Commuting the elements, one obtains

$$
\begin{aligned}
&\delta_L(l_{12356})=\delta_L\Big((h_{156}\rhd'(g_{56}\rhd\delta(l_{1245})^{-1})\delta(l_{1256})^{-1}h_{126})\rhd'(l_{2456}^{-1}l_{2346}^{-1}l_{2356}h_{256}\rhd'(g_{56}\rhd l_{2345}))\\
&h_{156}\rhd'(g_{56}\rhd(h_{145}\rhd'(g_{45}\rhd l_{1234})l_{1345}^{-1}))l_{1356}{}^{-1}(h_{136}h_{346})\rhd'\{h_{346}^{-1}h_{356}g_{56}\rhd h_{345},(g_{56}g_{45}g_{34})\rhd h_{123}\}_{\mathrm{p}}\\
&h_{136}\rhd' l_{3456}l_{1356}h_{156}\rhd'(g_{56}\rhd l_{1345})(\delta(l_{1456})^{-1}h_{146})\rhd'(\{h_{456},(g_{56}g_{45})\rhd h_{134}\}_{\mathrm{p}})\\
&(\delta(l_{1456})^{-1}h_{146})\rhd'((g_{46}\rhd l_{1234})^{-1})(\delta(l_{1456})^{-1}h_{146})\rhd'\{h_{456},(g_{56}g_{45})\rhd h_{124}\}_{\mathrm{p}}^{-1}\Big).
\end{aligned}
\tag{B.22}
$$

The tetrahedron (3456) is part of the integrand on both sides of the move, so using the condition (4.9) for $\delta_H(h_{3456})$ one can write $h_{346}^{-1}h_{356}g_{56}\rhd h_{345} = h_{346}^{-1}\rhd'\delta(l_{3456})^{-1}h_{456}$. Then, using the identity (3.4) one obtains that

$$
\begin{aligned}
\{h_{346}^{-1}h_{356}g_{56}\rhd h_{345},(g_{56}g_{45}g_{34})\rhd h_{123}\}_{\mathrm{p}} &= \{h_{346}^{-1}\rhd'\delta(l_{3456})^{-1}h_{456},(g_{56}g_{45}g_{34})\rhd h_{123}\}_{\mathrm{p}}\\
&=\big(h_{346}^{-1}\rhd'\delta(l_{3456})^{-1}\big)\rhd'\{h_{456},(g_{56}g_{45}g_{34})\rhd h_{123}\}_{\mathrm{p}}\\
&\quad\{h_{346}^{-1}\rhd'\delta(l_{3456})^{-1},(g_{46}g_{34})\rhd h_{123}\}_{\mathrm{p}}\\
&=h_{346}^{-1}\rhd' l_{3456}^{-1}\{h_{456},(g_{56}g_{45}g_{34})\rhd h_{123}\}_{\mathrm{p}}\\
&\quad\big((g_{46}g_{34})\rhd h_{123}h_{346}^{-1}\big)\rhd' l_{3456}\,,
\end{aligned}
\tag{B.23}
$$

where in the last row the definition of the action $\rhd'$ is used. Substituting the equation (B.23) in the equation (B.22) one obtains

$$
\begin{aligned}
\delta_L(l_{12356})=\delta_L\Big(&(h_{156}\rhd'(g_{56}\rhd\delta(l_{1245})^{-1})\delta(l_{1256})^{-1}h_{126}\delta(l_{2456})^{-1})\rhd'\Big(l_{2346}^{-1}l_{2356}h_{256}\rhd'(g_{56}\rhd l_{2345})l_{2456}^{-1}\Big)\\
&h_{156}\rhd'(g_{56}\rhd(h_{145}\rhd'(g_{45}\rhd l_{1234})))(h_{156}\rhd'(g_{56}\rhd\delta(l_{1345})^{-1})\delta(l_{1356})^{-1}h_{136}\delta(l_{3456})^{-1}h_{346})\rhd'\\
&(\{h_{456},(g_{56}g_{45}g_{34})\rhd h_{123}\}_{\mathrm{p}}((g_{46}g_{34})\rhd h_{123})\rhd' l_{3456})(\delta(l_{1456})^{-1}h_{146})\rhd'(\{h_{456},(g_{56}g_{45})\rhd h_{134}\}_{\mathrm{p}})\\
&(\delta(l_{1456})^{-1}h_{146})\rhd'((g_{46}\rhd l_{1234})^{-1})(\delta(l_{1456})^{-1}h_{146})\rhd'\{h_{456},(g_{56}g_{45})\rhd h_{124}\}_{\mathrm{p}}^{-1}\Big).
\end{aligned}
\tag{B.24}
$$

Commuting the element l_{3456} to the end of the expression, one obtains

$$
\begin{aligned}
\delta_L(l_{12356})=\delta_L(&(h_{156}\rhd'(g_{56}\rhd\delta(l_{1245})^{-1})\delta(l_{1256})^{-1}h_{126}\delta(l_{2456})^{-1})\rhd'(l_{2346}^{-1}l_{2356}h_{256}\rhd'(g_{56}\rhd l_{2345})l_{2456}^{-1})\\
&h_{156}\rhd'(g_{56}\rhd(h_{145}\rhd'(g_{45}\rhd l_{1234})))(h_{156}\rhd'(g_{56}\rhd\delta(l_{1345})^{-1})\delta(l_{1356})^{-1}h_{136}\delta(l_{3456})^{-1}h_{346})\rhd'\\
&(\{h_{456},(g_{56}g_{45}g_{34})\rhd h_{123}\}_{\mathrm{p}})(\delta(l_{1456})^{-1}h_{146})\rhd'(\{h_{456},(g_{56}g_{45})\rhd h_{134}\}_{\mathrm{p}})\\
&(\delta(l_{1456})^{-1}h_{146})\rhd'((g_{46}\rhd l_{1234})^{-1})(\delta(l_{1456})^{-1}h_{146})\rhd'\{h_{456},(g_{56}g_{45})\rhd h_{124}\}_{\mathrm{p}}^{-1}\\
&(h_{156}g_{56}\rhd h_{145}h_{246}g_{46}\rhd h_{234}h_{346}^{-1})\rhd' l_{3456})).
\end{aligned}
\tag{B.25}
$$

Acting to the whole expression with $(h_{156}\rhd'(g_{56}\rhd\delta(l_{1245})^{-1})\delta(l_{1256})^{-1}h_{126}\delta(l_{2456})^{-1})^{-1}\rhd'$, one obtains,

$$
\begin{aligned}
\delta_L(l_{12356})=\delta_L(&l_{2346}^{-1}l_{2356}h_{256}\rhd'(g_{56}\rhd l_{2345})l_{2456}^{-1}(h_{246}h_{456}(g_{56}g_{45})\rhd h_{124}^{-1})\rhd'\\
&\Big((g_{56}g_{45})\rhd l_{1234}((g_{56}g_{45})\rhd h_{134}h_{456}^{-1})\rhd'\{h_{456},(g_{56}g_{45}g_{34})\rhd h_{123}\}_{\mathrm{p}}\\
&h_{456}^{-1}\rhd'\{h_{456},(g_{56}g_{45})\rhd h_{134}\}_{\mathrm{p}}h_{456}^{-1}\rhd g_{46}\rhd l_{1234}^{-1}(h_{456}^{-1}g_{46}\rhd h_{124})\rhd'\{h_{456},(g_{56}g_{45})\rhd h_{124}^{-1}\}_{\mathrm{p}}\Big)\\
&(h_{246}g_{46}\rhd h_{234}h_{346}^{-1})\rhd' l_{3456}.
\end{aligned}
\tag{B.26}
$$

Using the identity (3.5) for $\{h_{456},(g_{56}g_{45})\rhd(h_{134}g_{34}\rhd h_{123})\}_{\mathrm{p}}$,

$$
\{h_{456},(g_{56}g_{45})\rhd(h_{134}g_{34}\rhd h_{123})\}_{\mathrm{p}} = \{h_{456},(g_{56}g_{45})\rhd h_{134}\}_{\mathrm{p}}(g_{46}\rhd h_{134})\rhd'\{h_{456},(g_{56}g_{45}g_{34})\rhd h_{123}\}_{\mathrm{p}}\,,
\tag{B.27}
$$

one obtains:

$$\delta_L(l_{12356}) = \delta_L\big(l_{2346}^{-1} l_{2356} h_{256} \rhd' (g_{56} \rhd l_{2345}) l_{2456}^{-1} h_{246} \rhd' \Big((h_{456}(g_{56}g_{45}) \rhd h_{124}^{-1}) \rhd' \Big((g_{56}g_{45}) \rhd l_{1234} h_{456}^{-1} \rhd' \{h_{456}, (g_{56}g_{45}) \rhd (h_{134} g_{34} \rhd h_{123})\}_{\mathrm{p}} h_{456}^{-1} \rhd g_{46} \rhd l_{1234}^{-1} \Big) \{h_{456}, (g_{56}g_{45}) \rhd h_{124}^{-1}\}_{\mathrm{p}} \Big) (h_{246} g_{46} \rhd h_{234} h_{346}^{-1}) \rhd' l_{3456}\big) . \tag{B.28}$$

Using the identity (3.5) for $\{h_{456}, (g_{56}g_{45}) \rhd (h_{124}^{-1} \delta(l_{1234}) h_{134} g_{34} \rhd h_{123})\}_{\mathrm{p}}$ one obtains the terms featuring l_{1234} cancel, i.e. ,

$$\begin{aligned}\delta_L(l_{12356}) =& \delta_L\big(l_{2346}^{-1} l_{2356} h_{256} \rhd' (g_{56} \rhd l_{2345}) l_{2456}^{-1} \\ & h_{246} \rhd' \{h_{456}, (g_{56}g_{45}) \rhd (h_{124}^{-1} \delta(l_{1234}) h_{134} g_{34} \rhd h_{123})\}_{\mathrm{p}} (h_{246} g_{46} \rhd h_{234} h_{346}^{-1}) \rhd' l_{3456} \\ =& \delta_L\big({l_{2346}}^{-1} l_{2356} h_{256} \rhd' (g_{56} \rhd l_{2345}) {l_{2456}}^{-1} h_{246} \rhd' \{h_{456}, (g_{56}g_{45}) \rhd h_{234}\}_{\mathrm{p}} (\delta(l_{2346})^{-1} h_{236}) \rhd' l_{3456})\big) \\ =& \delta_L(l_{23456}),\end{aligned} \tag{B.29}$$

the delta function $\delta_L(l_{12356})$ on the r.h.s. reduces to the delta function $\delta_L(l_{23456})$ of the l.h.s. The integrations over l_{1234}, l_{1245}, l_{1256}, l_{1345}, l_{1356}, and l_{1456} are trivial, and finally one obtains,

$$r.h.s. = \delta_G(e)^6 \delta_H(e)^4 \delta_L(l_{23456}) = |G|^6 |H|^4 \delta_L(l_{23456}) . \tag{B.30}$$

The prefactors $|G|^{-11}|H|^{-4}|L|^{-1}$ on the r.h.s. and $|G|^{-5}|H|^{0}|L|^{-1}$ on the l.h.s., compensate for left-over factors.

B.2 Pachner move $2 \leftrightarrow 4$

On the left hand side of the move one has the following integrals and the integrand,

$$\int_L dl_{2345} \delta_H(h_{2345}) \delta_L(l_{23456}) \delta_L(l_{12345}). \tag{B.31}$$

Integrating out l_{2345} using $\delta_L(l_{12345})$, one obtains

$$l_{2345} = {h_{125}}^{-1} \rhd' (l_{1235} h_{135} \rhd' \{h_{345}, (g_{45}g_{34}) \rhd h_{123}\}_{\mathrm{p}}^{-1} l_{1345} h_{145} \rhd' (g_{45} \rhd l_{1234})^{-1} l_{1245}^{-1}) . \tag{B.32}$$

The δ-function $\delta_H(h_{2345})$ now reads,

$$\begin{aligned}\delta_H(h_{2345}) &= \delta_H\big(\delta(l_{2345}) h_{245} \, (g_{45} \rhd h_{234}) \, h_{345}^{-1} \, h_{235}^{-1}\big) \\ &= \delta_H\Big({h_{125}}^{-1} \delta(l_{1235}) h_{135} \delta(\{h_{345}, (g_{45}g_{34}) \rhd h_{123}\}_{\mathrm{p}}^{-1}) h_{135}^{-1} \delta(l_{1345}) h_{145} (g_{45} \rhd \delta(l_{1234}))^{-1} h_{145}^{-1} \\ &\quad \delta(l_{1245})^{-1} h_{125} h_{245} \, (g_{45} \rhd h_{234}) \, h_{345}^{-1} \, h_{235}^{-1}\big) .\end{aligned} \tag{B.33}$$

Using the identity (4.9) for the tetrahedrons (1235), (1345), (1234), and (1245), the equation (B.33) reduces to:

$$\begin{aligned}\delta_H(h_{2345}) =& \delta_H\Big({h_{125}}^{-1} h_{125} \, h_{235} \, (g_{35} \rhd h_{123}^{-1}) \, h_{135}^{-1} h_{135} \delta(\{h_{345}, (g_{45}g_{34}) \rhd h_{123}\}_{\mathrm{p}}^{-1}) h_{135}^{-1} h_{135} \, h_{345} \, (g_{45} \rhd h_{134}^{-1}) \\ & h_{145}^{-1} h_{145} g_{45} \rhd (h_{134}(g_{34} \rhd h_{123}) h_{234}^{-1} h_{124}^{-1}) h_{145}^{-1} h_{145} (g_{45} \rhd h_{124}) h_{245}^{-1} h_{125}^{-1} h_{125} h_{245} \, (g_{45} \rhd h_{234}) \, h_{345}^{-1} \, h_{235}^{-1}\big) \\ =& \delta_H\Big((g_{35} \rhd h_{123}^{-1}) \, \delta(\{h_{345}, (g_{45}g_{34}) \rhd h_{123}\}_{\mathrm{p}}^{-1}) \, h_{345} \, (g_{45}g_{34}) \rhd h_{123}) \, h_{345}^{-1}) .\end{aligned} \tag{B.34}$$

JHEP07(2022)105

Here, one uses the following identity

$$\delta\{h_1\,,h_2\}_{\rm p}(\partial(h_1)\rhd h_2)h_1h_2^{-1}h_1^{-1}=e\,, \tag{B.35}$$

for $h_1 = h_{345}$ and $h_2 = (g_{45}g_{34}) \rhd h_{123}$, and the identity $g_{35} = \partial(h_{345})g_{45}g_{34}$, and obtains

$$\delta_H(h_{2345})=\delta_H(e)\,. \tag{B.36}$$

The remaining δ-function $\delta_L(l_{23456})$, reads

$$\delta_L(l_{23456})=\delta_L(l_{2346}{}^{-1}(h_{236}\rhd' l_{3456})l_{2356}h_{256}\rhd'(g_{56}\rhd l_{2345})l_{2456}{}^{-1}h_{246}\rhd'\{h_{456},(g_{56}g_{45})\rhd h_{234}\}_{\rm p})\,. \tag{B.37}$$

Substituting the equation (B.33), one obtains

$$\begin{aligned}\delta_L(l_{23456})=\delta_L\Big(&l_{2346}{}^{-1}(h_{236}\rhd' l_{3456})l_{2356}h_{256}\rhd'\Big(g_{56}\rhd(h_{125}{}^{-1}\rhd'(l_{1235}h_{135}\rhd'\{h_{345},(g_{45}g_{34})\rhd h_{123}\}_{\rm p}^{-1}\\&l_{1345}h_{145}\rhd'(g_{45}\rhd l_{1234})^{-1}l_{1245}^{-1}))\Big)l_{2456}{}^{-1}h_{246}\rhd'\{h_{456},(g_{56}g_{45})\rhd h_{234}\}_{\rm p}\Big)\,.\end{aligned} \tag{B.38}$$

Commuting the elements one obtains

$$\begin{aligned}\delta_L(l_{23456})=\delta_L\Big(&l_{2456}{}^{-1}l_{2346}{}^{-1}l_{2356}(h_{256}g_{56}\rhd h_{125}{}^{-1})\rhd' g_{56}\rhd l_{1235}(h_{256}g_{56}\rhd h_{125}{}^{-1}g_{56}\rhd h_{135})\rhd'\\&\Big((g_{35}\rhd h_{123}h_{356}^{-1})\rhd' l_{3456})\{g_{56}\rhd h_{345},(g_{56}g_{45}g_{34})\rhd h_{123}\}_{\rm p}^{-1}\\&(g_{56}\rhd h_{345}(g_{56}g_{45})\rhd(h_{123}h_{234}^{-1})h_{456}^{-1})\rhd'\{h_{456},(g_{56}g_{45})\rhd h_{234}\}_{\rm p}\Big)\\&(h_{256}g_{56}\rhd h_{125}{}^{-1})\rhd' g_{56}\rhd l_{1345}(h_{256}g_{56}\rhd h_{125}{}^{-1}g_{56}\rhd h_{145})\rhd'((g_{56}g_{45})\rhd l_{1234})^{-1}\\&(h_{256}g_{56}\rhd h_{125}{}^{-1})\rhd' g_{56}\rhd l_{1245}^{-1}\Big)\,.\end{aligned} \tag{B.39}$$

Finally, the l.h.s. reads:

$$\text{l.h.s.}=\delta_H(e)\delta_L(l_{23456})=|H|\delta_L(l_{23456})\,. \tag{B.40}$$

Let us now examine the right hand side of the move, i.e. , the integral (4.17). First, one integrates out g_{16} using $\delta_G(g_{126})$, and obtains

$$g_{16}=\partial(h_{126})\,g_{26}\,g_{12}\,. \tag{B.41}$$

Next, one integrates out h_{126} using $\delta_H(h_{1236})$, h_{136} using $\delta_H(h_{1346})$, and h_{146} using $\delta_H(h_{1456})$, and obtains

$$\begin{aligned}h_{126}&=\delta(l_{1236})h_{136}\,(g_{36}\rhd h_{123})\,h_{236}^{-1}\,,\\h_{136}&=\delta(l_{1346})h_{146}\,(g_{46}\rhd h_{134})\,h_{346}^{-1}\,,\\h_{146}&=\delta(l_{1456})h_{156}\,(g_{56}\rhd h_{145})\,h_{456}^{-1}\,.\end{aligned} \tag{B.42}$$

The remaining δ-functions on the group G reduces to $\delta_G(e)^3$. The δ-function $\delta_G(g_{136})$

$$\delta_G(g_{136})=\delta_G\big(\partial(h_{136})\,g_{36}\,g_{13}\,g_{16}^{-1}\big)\,, \tag{B.43}$$

after substituting the equation (B.41) reads:

$$\delta_G(g_{136}) = \delta_G\big(\partial(h_{136})\, g_{36}\, g_{13}\, g_{12}^{-1} g_{26}^{-1} \partial(h_{126})^{-1}\big)\,. \tag{B.44}$$

Using the equations (B.42) for h_{126}, and h_{136}, and h_{146}, and the identity $\partial(\delta l) = 0$ for every element $l \in L$, the δ-function $\delta_G(g_{136})$ reduces to $\delta_G(e)$ after implementing the identity (4.7) for the triangles (156), (145), (456) (134), (346), (236), and (123). Similarly, one obtains $\delta_G(g_{146}) = \delta_G(g_{156}) = \delta_G(e)$.

One integrates out l_{1236} using $\delta_L(l_{12346})$ and obtains

$$l_{1236} = (h_{126} \rhd' l_{2346}) l_{1246} h_{146} \rhd' (g_{46} \rhd l_{1234}) l_{1346}^{-1}\, h_{136} \rhd' \{h_{346}, (g_{46}g_{34}) \rhd h_{123}\}_{\mathrm{p}}, \tag{B.45}$$

l_{1246} using $\delta_L(l_{12456})$ and obtains

$$l_{1246} = (h_{126} \rhd' l_{2456}) l_{1256} h_{156} \rhd' (g_{56} \rhd l_{1245}) {l_{1456}}^{-1}\, h_{146} \rhd' \{h_{456}, (g_{56}g_{45}) \rhd h_{124}\}_{\mathrm{p}}, \tag{B.46}$$

and l_{1346} using $\delta_L(l_{13456})$ and obtains

$$l_{1346} = (h_{136} \rhd' l_{3456}) l_{1356} h_{156} \rhd' (g_{56} \rhd l_{1345}) {l_{1456}}^{-1}\, h_{146} \rhd' \{h_{456}, (g_{56}g_{45}) \rhd h_{134}\}_{\mathrm{p}}. \tag{B.47}$$

The remaining δ-functions on H reduce on $\delta_H(e)^3$, similarly as in the case of $1-5$ Pachner move, i.e. , one obtains $\delta_H(h_{1256}) = \delta_H(h_{1356}) = \delta_H(h_{1456}) = \delta_H(e)$. For the remaining δ-function $\delta_L(l_{12356})$,

$$\delta_L(l_{12356}) = \delta_L\Big({l_{1236}}^{-1}(h_{126} \rhd' l_{2356}) l_{1256} h_{156} \rhd' (g_{56} \rhd l_{1235}) {l_{1356}}^{-1} h_{136} \rhd' \{h_{356}, (g_{56}g_{35}) \rhd h_{123}\}_{\mathrm{p}}\Big)\,, \tag{B.48}$$

one obtains, after substituting the equations (B.45), (B.46), and (B.47), the following

$$\begin{aligned}\delta_L(l_{12356}) = \delta_L\Big(& h_{136} \rhd' \{h_{346}, (g_{46}g_{34}) \rhd h_{123}\}_{\mathrm{p}}{}^{-1} l_{1346} h_{146} \rhd' (g_{46} \rhd l_{1234})^{-1} l_{1246}^{-1} (h_{126} \rhd' l_{2346})^{-1} \\ & (h_{126} \rhd' l_{2356}) l_{1256} h_{156} \rhd' (g_{56} \rhd l_{1235}) {l_{1356}}^{-1} h_{136} \rhd' \{h_{356}, (g_{56}g_{35}) \rhd h_{123}\}_{\mathrm{p}}\Big) \\ = \delta_L\Big(& (h_{126} \rhd' l_{2456})^{-1} (h_{126} \rhd' l_{2346})^{-1} (h_{126} \rhd' l_{2356}) (h_{256} g_{56} \rhd {h_{125}}^{-1}) \rhd' l_{1235} \\ & \delta(l_{1256}) \rhd' \Big(\delta(l_{1356})^{-1} \rhd' (h_{136} \rhd' \{h_{356}, (g_{56}g_{35}) \rhd h_{123}\}_{\mathrm{p}} (h_{136}h_{346}) \rhd' \{h_{346}^{-1}, g_{36} \rhd h_{123}\}_{\mathrm{p}} \\ & (h_{136} \rhd' l_{3456})) h_{156} \rhd' (g_{56} \rhd l_{1345}) {l_{1456}}^{-1}\, h_{146} \rhd' \{h_{456}, (g_{56}g_{45}) \rhd h_{134}\}_{\mathrm{p}} h_{146} \rhd' (g_{46} \rhd l_{1234})^{-1} \\ & h_{146} \rhd' \{h_{456}, (g_{56}g_{45}) \rhd h_{124}\}_{\mathrm{p}}^{-1} l_{1456} h_{156} \rhd' (g_{56} \rhd l_{1245})^{-1}\Big)\Big)\,.\end{aligned} \tag{B.49}$$

Commuting the elements in order to match the l.h.s. of the move, i.e. , the δ-function given by the equation (B.39), and using the identity (3.4), i.e. ,

$$\{h_{346}^{-1} h_{356}, (g_{56}g_{35}) \rhd h_{123}\}_{\mathrm{p}} = h_{346}^{-1} \rhd' \{h_{356}, (g_{56}g_{35}) \rhd h_{123}\}_{\mathrm{p}} \{h_{346}^{-1}, g_{36} \rhd h_{123}\}_{\mathrm{p}}\,, \tag{B.50}$$

one obtains

$$\begin{aligned}\delta_L(l_{12356}) = \delta_L\Big(& (h_{126} \rhd' l_{2456})^{-1} (h_{126} \rhd' l_{2346})^{-1} (h_{126} \rhd' l_{2356}) (h_{126} h_{256} g_{56} \rhd {h_{125}}^{-1}) \rhd' l_{1235} \\ & \delta(l_{1256}) \rhd' \Big(\delta(l_{1356})^{-1} \rhd' ((h_{136}h_{346}) \rhd' \{h_{346}^{-1} h_{356}, (g_{56}g_{35}) \rhd h_{123}\}_{\mathrm{p}} (h_{136} \rhd' l_{3456})) \\ & h_{156} \rhd' (g_{56} \rhd l_{1345}) (\delta(l_{1456})^{-1}\, h_{146}) \rhd' (\{h_{456}, (g_{56}g_{45}) \rhd h_{134}\}_{\mathrm{p}} (g_{46} \rhd l_{1234})^{-1} \\ & \{h_{456}, (g_{56}g_{45}) \rhd h_{124}\}_{\mathrm{p}}^{-1}) h_{156} \rhd' (g_{56} \rhd l_{1245})^{-1}\Big)\Big)\,.\end{aligned} \tag{B.51}$$

Using the identity (3.4) again one rewrites the following term as

$$
\begin{aligned}
&(h_{136}h_{346}) \rhd' \{h_{346}^{-1}h_{356}, (g_{56}g_{35}) \rhd h_{123}\}_{\mathrm{p}}(h_{136} \rhd' l_{3456}) = \\
&(h_{136}h_{346}) \rhd' \{h_{346}^{-1} \rhd' \delta(l_{3456})^{-1}h_{456}g_{56} \rhd h_{345}^{-1}, (g_{56}g_{35}) \rhd h_{123}\}_{\mathrm{p}}(h_{136} \rhd' l_{3456}) = \\
&(h_{136} \rhd' \delta(l_{3456})^{-1}h_{136}h_{346}) \rhd' (\{h_{456}g_{56} \rhd h_{345}^{-1}, (g_{56}g_{35}) \rhd h_{123}\}_{\mathrm{p}}((g_{46}g_{34}) \rhd h_{123}h_{346}^{-1}) \rhd' l_{3456}^{-1}) \,,
\end{aligned}
\tag{B.52}
$$

and substituting it in the equation (B.51) the δ-function becomes:

$$
\begin{aligned}
\delta_L(l_{12356}) = \delta_L\Big(&(h_{126} \rhd' l_{2456})^{-1}(h_{126} \rhd' l_{2346})^{-1}(h_{126} \rhd' l_{2356})(h_{126}h_{256}g_{56} \rhd h_{125}{}^{-1}) \rhd' l_{1235} \\
&\delta(l_{1256}) \rhd' \Big((\delta(l_{1356})^{-1}h_{136} \rhd' \delta(l_{3456})^{-1}h_{136}h_{346}) \rhd' \\
&(\{h_{456}g_{56} \rhd h_{345}^{-1}, (g_{56}g_{35}) \rhd h_{123}\}_{\mathrm{p}}((g_{46}g_{34}) \rhd h_{123}h_{346}^{-1}) \rhd' l_{3456})) \\
&(h_{156}g_{56} \rhd h_{135}g_{56} \rhd (h_{345}g_{45} \rhd h_{134}^{-1})h_{456}^{-1}) \rhd' (\{h_{456}, (g_{56}g_{45}) \rhd h_{134}\}_{\mathrm{p}}(g_{46} \rhd l_{1234})^{-1} \\
&\{h_{456}, (g_{56}g_{45}) \rhd h_{124}\}_{\mathrm{p}}^{-1})\Big)(h_{126}h_{256}g_{56} \rhd h_{125}{}^{-1}) \rhd' (h_{156} \rhd' (g_{56} \rhd l_{1345})(g_{56} \rhd l_{1245})^{-1})\Big) \,.
\end{aligned}
\tag{B.53}
$$

Commuting the elements l_{3456} and $\{h_{456}g_{56} \rhd h_{345}, (g_{56}g_{35}) \rhd h_{123}\}_{\mathrm{p}}$, and using the identity (3.4) to rewrite this Peiffer lifting, one obtains

$$
\begin{aligned}
\delta_L(l_{12356}) = \delta_L\Big(&(h_{126} \rhd' l_{2456})^{-1}(h_{126} \rhd' l_{2346})^{-1}(h_{126} \rhd' l_{2356})(h_{126}h_{256}g_{56} \rhd h_{125}{}^{-1}) \rhd' l_{1235} \\
&(h_{126}h_{256}g_{56} \rhd h_{125}{}^{-1}h_{135}(g_{56}g_{35}) \rhd h_{123}g_{56} \rhd h_{356}^{-1}) \rhd' g_{56} \rhd l_{3456} \\
&(h_{126}h_{256}g_{56} \rhd h_{125}^{-1}g_{56} \rhd h_{135}g_{56} \rhd h_{345}) \rhd' \Big(\{g_{56} \rhd h_{345}^{-1}, (g_{56}g_{35}) \rhd h_{123}\}_{\mathrm{p}} \\
&h_{456}^{-1} \rhd' \{h_{456}, (g_{56}g_{45}g_{34}) \rhd h_{123}\}_{\mathrm{p}}((g_{56}g_{45}) \rhd h_{134}^{-1}h_{456}^{-1}) \rhd' (\{h_{456}, (g_{56}g_{45}) \rhd h_{134}\}_{\mathrm{p}}(g_{46} \rhd l_{1234})^{-1} \\
&\{h_{456}, (g_{56}g_{45}) \rhd h_{124}\}_{\mathrm{p}}^{-1})\Big)(h_{126}h_{256}g_{56} \rhd h_{125}{}^{-1}) \rhd' (h_{156} \rhd' (g_{56} \rhd l_{1345})(g_{56} \rhd l_{1245})^{-1})\Big) \,.
\end{aligned}
\tag{B.54}
$$

After the similar transformations as in the case of $1-5$ move, commuting the element l_{1234} so that the order of the elements matches the order in the expression (B.39), and acting to the whole expression with h_{126}^{-1} one obtains

$$
\begin{aligned}
\delta_L(l_{12356}) = \delta_L\Big(&l_{2456}{}^{-1}l_{2346}{}^{-1}l_{2356}(h_{256}g_{56} \rhd h_{125}{}^{-1}) \rhd' g_{56} \rhd l_{1235}(h_{256}g_{56} \rhd h_{125}{}^{-1}g_{56} \rhd h_{135}) \rhd' \\
&\Big((g_{35} \rhd h_{123}h_{356}^{-1}) \rhd' l_{3456})\{g_{56} \rhd h_{345}, (g_{56}g_{45}g_{34}) \rhd h_{123}\}_{\mathrm{p}}^{-1}(g_{56} \rhd h_{345}(g_{56}g_{45}) \rhd (h_{123}h_{234}^{-1})h_{456}^{-1}) \rhd' \\
&\{h_{456}, (g_{56}g_{45}) \rhd h_{234}\}_{\mathrm{p}}\Big)(h_{256}g_{56} \rhd h_{125}{}^{-1}) \rhd' g_{56} \rhd l_{1345} \\
&(h_{256}g_{56} \rhd h_{125}{}^{-1}g_{56} \rhd h_{145}) \rhd' ((g_{56}g_{45}) \rhd l_{1234})^{-1}(h_{256}g_{56} \rhd h_{125}{}^{-1}) \rhd' g_{56} \rhd l_{1245}^{-1}\Big) \,.
\end{aligned}
\tag{B.55}
$$

which is precisely the equation (B.39). The remaining integration over the element h_{156} of the group H and remaining integration over the three elements of the group L, l_{1246}, l_{1256}, and l_{1356}, are trivial, yielding the result on the r.h.s. to:

$$
\text{r.h.s.} = \delta_G(e)^3 \, \delta_H(e)^3 \, \delta_L(l_{12356}) = |G|^3 \, |H|^3 \, \delta_L(l_{12356}) \,. \tag{B.56}
$$

The prefactors are $|G|^{-8}|H|^{-1}|L|^{-1}$ on the l.h.s., and $|G|^{-11}|H|^{-3}|L|^{-1}$ on the r.h.s. compensate for the left-over factors.

JHEP07(2022)105

B.3 Pachner move $3 \leftrightarrow 3$

Let us first investigate the r.h.s. of the move. First, one integrates out the l_{1235}, exploiting $\delta_L(l_{12345})$ and obtains

$$l_{1235} = (h_{125} \triangleright' l_{2345}) l_{1245} h_{145} \triangleright' (g_{45} \triangleright l_{1234}) {l_{1345}}^{-1} \, h_{135} \triangleright' \{h_{345}, (g_{45} g_{34}) \triangleright h_{123}\}_{\mathrm{p}}, \tag{B.57}$$

and one integrates out l_{1236}, exploiting $\delta_L(l_{12356})$ and obtains

$$l_{1236} = (h_{126} \triangleright' l_{2356}) l_{1256} h_{156} \triangleright' (g_{56} \triangleright l_{1235}) {l_{1356}}^{-1} h'_{136} \triangleright \{h_{356}, (g_{56} g_{35}) \triangleright h_{123}\}_{\mathrm{p}}. \tag{B.58}$$

Next, one integrates out h_{123}, exploiting $\delta_H(l_{1234})$ and obtains:

$$h_{123} = g_{34}^{-1} \triangleright h_{134}^{-1} \, g_{34}^{-1} \triangleright \delta(l_{1234})^{-1} \, g_{34}^{-1} \triangleright h_{124} \, g_{34}^{-1} \triangleright h_{234}. \tag{B.59}$$

The δ-function $\delta_G(g_{123})$, when using the equation (B.59) reads

$$\delta_G(g_{123}) = \delta_G\big(g_{34}^{-1} \triangleright \partial(h_{134})^{-1} \, g_{34}^{-1} \triangleright \partial(\delta(l_{1234}))^{-1} \, g_{34}^{-1} \triangleright \partial(h_{124}) \, g_{34}^{-1} \triangleright \partial(h_{234}) \, g_{23} \, g_{12} \, g_{13}^{-1}\big) , \tag{B.60}$$

which then using the condition $\partial\delta = 0$, reduces to

$$\delta_G(g_{123}) = \delta_G\big(\partial(h_{134})^{-1} \, \partial(h_{124}) \, \partial(h_{234}) \, g_{34}^{-1} g_{23} \, g_{12} \, g_{13}^{-1} g_{34}\big) . \tag{B.61}$$

Using the condition (4.7) for the triangles (134), (124), and (234), it finally reduces to

$$\delta_G(g_{123}) = \delta_G(e) . \tag{B.62}$$

For the δ-function $\delta_H(h_{1235})$, one obtains, after using the equation (B.57):

$$\begin{aligned} \delta_H(h_{1235}) = \delta_H\Big(&(h_{125}\delta(l_{2345})h_{125}^{-1})\delta(l_{1245})(h_{145}(g_{45} \triangleright \delta(l_{1234}))h_{145}^{-1})\delta(l_{1345})^{-1} \\ &h_{135} \triangleright' \{h_{345}, g_{35} \triangleright h_{123}\}_{\mathrm{p}} \, h_{135}((g_{35}{g_{34}}^{-1}) \triangleright (h_{134}^{-1} \, \delta(l_{1234})^{-1} \, h_{124} \, h_{234})) \, h_{235}^{-1} \, h_{125}^{-1}\Big) . \end{aligned} \tag{B.63}$$

Using the δ-functions $\delta_L(h_{2345})$, $\delta_L(h_{1245})$, and $\delta_L(h_{1345})$, that appear on both sides of the move, and are thus part of the integrand,

$$\begin{aligned} \delta(l_{2345}) &= h_{235} \, h_{345} \, (g_{45} \triangleright h_{234}^{-1}) \, h_{245}^{-1} \, , \\ \delta(l_{1245}) &= h_{125} \, h_{245} \, (g_{45} \triangleright h_{124}^{-1}) \, h_{145}^{-1} \, , \\ \delta(l_{1345})^{-1} &= h_{145} \, (g_{45} \triangleright h_{134}) \, h_{345}^{-1} \, h_{135}^{-1} \, , \end{aligned} \tag{B.64}$$

one obtains:

$$\begin{aligned} \delta_H(h_{1235}) = \delta_H\Big(&h_{125} h_{235} \, h_{345} \, (g_{45} \triangleright h_{234}^{-1}) \, h_{245}^{-1} h_{125}^{-1} h_{125} \, h_{245} \, (g_{45} \triangleright h_{124}^{-1}) \, h_{145}^{-1} h_{145} (g_{45} \triangleright \delta(l_{1234})) h_{145}^{-1} \\ &h_{145} \, (g_{45} \triangleright h_{134}) \, h_{345}^{-1} \, h_{135}^{-1} h_{135} \triangleright \delta(\{h_{345}, (g_{45} g_{34}) \triangleright h_{123}\}_{\mathrm{p}}) \\ &h_{135} \, ((g_{35}{g_{34}}^{-1}) \triangleright (h_{134}^{-1} \delta(l_{1234})^{-1} \, h_{124} \, h_{234})) \, h_{235}^{-1} \, h_{125}^{-1}\Big) \\ = \delta_H\Big(&h_{345} (g_{45} g_{34}) \triangleright h_{123}^{-1} \, h_{345}^{-1} \delta(\{h_{345}, (g_{45} g_{34}) \triangleright h_{123}\}_{\mathrm{p}}) \, (g_{35} \triangleright h_{123})\Big) . \end{aligned} \tag{B.65}$$

JHEP07(2022)105

Substituting $g_{35} = \partial(h_{345})g_{45}g_{34}$, and applying the identity

$$\delta\{h_1\,, h_2\}_{\rm p}(\partial(h_1) \rhd h_2)h_1 h_2^{-1} h_1^{-1} = e\,, \tag{B.66}$$

for $h_1 = h_{345}$ and $h_2 = (g_{45}g_{34}) \rhd h_{123}$, one obtains

$$\delta_H(h_{1235}) = \delta_H(e). \tag{B.67}$$

Similarly, one obtains that $\delta_H(h_{1236}) = \delta_H(e)$. The remaining δ-function $\delta_H(l_{12346})$ reads

$$\delta_L(l_{12346}) = \delta_L(l_{1236}{}^{-1}(h_{126} \rhd' l_{2346})l_{1246}h_{146} \rhd' (g_{46} \rhd l_{1234})l_{1346}{}^{-1}\, h_{136} \rhd' \{h_{346}, (g_{46}g_{34}) \rhd h_{123}\}_{\rm p})\,. \tag{B.68}$$

After substituting the equation (B.58), and then the equation (B.57), one obtains:

$$\begin{aligned}\delta_L(l_{12346}) = {} & \delta_L(h_{136} \rhd' \{h_{356}, (g_{56}g_{35}) \rhd h_{123}\}_{\rm p}^{-1} l_{1356}h_{156} \rhd' (g_{56} \rhd l_{1235})^{-1} l_{1256}^{-1} h_{126} \rhd' l_{2356}^{-1} \\ & (h_{126} \rhd' l_{2346})l_{1246}h_{146} \rhd' (g_{46} \rhd l_{1234})l_{1346}{}^{-1} h_{136} \rhd' \{h_{346}, (g_{46}g_{34}) \rhd h_{123}\}_{\rm p}) \\ = {} & \delta_L(h_{136} \rhd' \{h_{356}, (g_{56}g_{35}) \rhd h_{123}\}_{\rm p}^{-1} l_{1356} \\ & h_{156} \rhd' (g_{56} \rhd ((h_{125} \rhd' l_{2345})l_{1245}h_{145} \rhd' (g_{45} \rhd l_{1234})l_{1345}{}^{-1}\, h_{135} \rhd' \{h_{345}, (g_{45}g_{34}) \rhd h_{123}\}_{\rm p}))^{-1} \\ & l_{1256}^{-1} h_{126} \rhd' l_{2356}^{-1}(h_{126} \rhd' l_{2346})l_{1246}h_{146} \rhd' (g_{46} \rhd l_{1234})l_{1346}{}^{-1} h_{136} \rhd' \{h_{346}, (g_{46}g_{34}) \rhd h_{123}\}_{\rm p})\,.\end{aligned} \tag{B.69}$$

After commuting the elements, i.e. , using the Peiffer identity for the crossed module $(L \stackrel{\delta}{\rightarrow} H, \rhd')$, one obtains

$$\begin{aligned}\delta_L(l_{12346}) = {} & \delta_L(h_{136} \rhd' \{h_{356}, (g_{56}g_{35}) \rhd h_{123}\}_{\rm p}^{-1} \\ & (\delta(l_{1356})h_{156}g_{56} \rhd h_{135}) \rhd' g_{56} \rhd \{h_{345}, (g_{45}g_{34}) \rhd h_{123}\}_{\rm p}^{-1} l_{1356}h_{156} \rhd' (g_{56} \rhd l_{1345}) \\ & (h_{156}g_{56} \rhd h_{145}) \rhd' ((g_{56}g_{45}) \rhd l_{1234})^{-1} h_{156} \rhd' (g_{56} \rhd l_{1245})^{-1}(h_{156}g_{56} \rhd h_{125}) \rhd' (g_{56} \rhd l_{2345}^{-1}) l_{1256}^{-1} \\ & h_{126} \rhd' l_{2356}^{-1}(h_{126} \rhd' l_{2346})l_{1246}h_{146} \rhd' (g_{46} \rhd l_{1234})l_{1346}{}^{-1} h_{136} \rhd' \{h_{346}, (g_{46}g_{34}) \rhd h_{123}\}_{\rm p}) \\ = {} & \delta_L((\delta(l_{1346})^{-1}h_{136}) \rhd' \{h_{346}, (g_{46}g_{34}) \rhd h_{123}\}_{\rm p}(\delta(l_{1346})^{-1}h_{136}) \rhd' \{h_{356}, (g_{56}g_{35}) \rhd h_{123}\}_{\rm p}^{-1} \\ & ((\delta(l_{1346})^{-1}\delta(l_{1356})h_{156}g_{56} \rhd h_{135}) \rhd' g_{56} \rhd \{h_{345}, (g_{45}g_{34}) \rhd h_{123}\}_{\rm p}^{-1} \\ & (\delta(l_{1346})^{-1}\delta(l_{1356})h_{156} \rhd' (g_{56} \rhd \delta(l_{1345}))h_{156}g_{56} \rhd h_{145}) \rhd' ((g_{56}g_{45}) \rhd l_{1234})^{-1} l_{1346}^{-1} \\ & l_{1356}h_{156} \rhd' (g_{56} \rhd l_{1345})h_{156} \rhd' (g_{56} \rhd l_{1245})^{-1}(h_{156}g_{56} \rhd h_{125}) \rhd' (g_{56} \rhd l_{2345}^{-1}) \\ & l_{1256}^{-1} h_{126} \rhd' l_{2356}^{-1}(h_{126} \rhd' l_{2346})l_{1246}h_{146} \rhd' (g_{46} \rhd l_{1234}))\,.\end{aligned} \tag{B.70}$$

Using the identity (3.7) one obtains that

$$\{h_{346}, (g_{46}g_{34}) \rhd h_{123}\}_{\rm p} = h_{346} \rhd' \{h_{346}^{-1}, g_{36} \rhd h_{123}\}_{\rm p}^{-1}\,. \tag{B.71}$$

Using a variant of the identity (3.4), i.e. , that

$$\{h_1 h_2 h_3, h_4\}_{\rm p}^{-1} = \{h_1, \partial(h_2 h_3) \rhd h_4\}_{\rm p}^{-1} h_1 \rhd' \{h_2, \partial(h_2) \rhd h_4\}_{\rm p}^{-1} (h_1 h_2) \rhd' \{h_3, h_4\}_{\rm p}^{-1}\,, \tag{B.72}$$

one obtains that

$$\begin{aligned}\{h_{346}^{-1}\, h_{356}\,(g_{56} \rhd h_{345}), (g_{56}g_{45}g_{34}) \rhd h_{123}\}_{\rm p}^{-1} = {} & \{h_{346}^{-1}, (g_{46}g_{34}) \rhd h_{123}\}_{\rm p}^{-1} h_{346}^{-1} \rhd' \{h_{356}, (g_{56}g_{35}) \rhd h_{123}\}_{\rm p}^{-1} \\ & (h_{346}^{-1} h_{356}) \rhd' \{g_{56} \rhd h_{345}, (g_{56}g_{45}g_{34}) \rhd h_{123}\}_{\rm p}^{-1}\,,\end{aligned} \tag{B.73}$$

rendering the expression (B.70) to

$$\delta_L(l_{12346}) = \delta_L((h_{146}g_{46} \rhd h_{134}) \rhd' \{h_{346}^{-1} h_{356} (g_{56} \rhd h_{345}), (g_{56}g_{45}g_{34}) \rhd h_{123}\}_{\mathrm{p}}^{-1} (\delta(l_{1346})^{-1}\delta(l_{1356})h_{156} \rhd' (g_{56} \rhd \delta(l_{1345}))h_{156}g_{56} \rhd h_{145}) \rhd' ((g_{56}g_{45}) \rhd l_{1234})^{-1} l_{1346}^{-1}l_{1356}h_{156} \rhd' (g_{56} \rhd l_{1345})h_{156} \rhd' (g_{56} \rhd l_{1245})^{-1}(h_{156}g_{56} \rhd h_{125}) \rhd' (g_{56} \rhd l_{2345}^{-1})l_{1256}^{-1} h_{126} \rhd' l_{2356}^{-1}(h_{126} \rhd' l_{2346})l_{1246}h_{146} \rhd' (g_{46} \rhd l_{1234}))\,. \tag{B.74}$$

Substituting the equation (B.59), and using the identity (3.5), one obtains that the expression,

$$\begin{aligned}\{h_{346}^{-1} h_{356} (g_{56} \rhd h_{345}), (g_{56}g_{45}g_{34}) \rhd h_{123}\}_{\mathrm{p}}^{-1} &= \{h_{346}^{-1} h_{356} (g_{56} \rhd h_{345}), (g_{56}g_{45}) \rhd ((h_{134}^{-1} \rhd' \delta(l_{1234})^{-1})h_{134}^{-1}h_{124}h_{234}\}_{\mathrm{p}}^{-1} \\ &= (g_{46} \rhd (h_{134}^{-1} \rhd' \delta(l_{1234})^{-1})) \rhd' \{h_{346}^{-1} h_{356} (g_{56} \rhd h_{345}), (g_{56}g_{45}) \rhd \\ &\quad (h_{134}^{-1}h_{124}h_{234})\}_{\mathrm{p}}^{-1}\{h_{346}^{-1} h_{356} (g_{56} \rhd h_{345}), (g_{56}g_{45}) \rhd (h_{134}^{-1} \rhd' \delta(l_{1234})^{-1})\}_{\mathrm{p}}^{-1}\,,\end{aligned} \tag{B.75}$$

using the identity (3.9), i.e. , that

$$\{h_{346}^{-1} h_{356} (g_{56} \rhd h_{345}), (g_{56}g_{45}) \rhd (h_{134}^{-1} \rhd' \delta(l_{1234})^{-1})\}_{\mathrm{p}}^{-1} = g_{46} \rhd (h_{134}^{-1} \rhd' l_{1234}^{-1})(h_{346}^{-1} h_{356} (g_{56} \rhd h_{345})) \rhd' ((g_{56}g_{45}) \rhd (h_{134}^{-1} \rhd' l_{1234}))\,, \tag{B.76}$$

reduces to

$$\begin{aligned}\{h_{346}^{-1} h_{356} (g_{56} \rhd h_{345}), (g_{56}g_{45}g_{34}) \rhd h_{123}\}_{\mathrm{p}}^{-1} = g_{46} \rhd (h_{134}^{-1} \rhd' \delta(l_{1234})^{-1}) \\ \{h_{346}^{-1} h_{356} (g_{56} \rhd h_{345}), (g_{56}g_{45}) \rhd (h_{134}^{-1}h_{124}h_{234})\}_{\mathrm{p}}^{-1} \\ (h_{346}^{-1} h_{356} (g_{56} \rhd h_{345})) \rhd' ((g_{56}g_{45}) \rhd (h_{134}^{-1} \rhd' l_{1234}))\,.\end{aligned} \tag{B.77}$$

Substituting this result in the expression (B.74) the terms featuring l_{1234} cancel, and finally the delta function $\delta_L(l_{12346})$ reads:

$$\delta_L(l_{12346}) = \delta_L((h_{146}g_{46} \rhd h_{134}) \rhd' \{h_{346}^{-1} h_{356} (g_{56} \rhd h_{345}), (g_{56}g_{45}) \rhd (h_{134}^{-1}h_{124}h_{234})\}_{\mathrm{p}}^{-1}l_{1346}^{-1}l_{1356} h_{156} \rhd' (g_{56} \rhd l_{1345})h_{156} \rhd' (g_{56} \rhd l_{1245})^{-1}(h_{156}g_{56} \rhd h_{125}) \rhd' (g_{56} \rhd l_{2345}^{-1}) l_{1256}^{-1}h_{126} \rhd' l_{2356}^{-1}(h_{126} \rhd' l_{2346})l_{1246})\,. \tag{B.78}$$

One obtains that the integration over l_{1234} is trivial, and the r.h.s. of the move finally reads

$$\text{r.h.s.} = \delta_G(e)\delta_H(e)^2\delta_L(h_{156} \rhd' (g_{56} \rhd l_{1245})^{-1}\, h_{156} \rhd' (g_{56} \rhd (h_{125} \rhd' l_{2345}))^{-1}\, l_{1256}^{-1} h_{126} \rhd' l_{2356}^{-1}(h_{126} \rhd' l_{2346})l_{1246}(h_{146}g_{46} \rhd h_{134})\rhd' \{h_{346}^{-1} h_{356} (g_{56} \rhd h_{345}), (g_{56}g_{45}) \rhd (h_{134}^{-1}h_{124}h_{234})\}_{\mathrm{p}}^{-1}l_{1346}^{-1}l_{1356}h_{156} \rhd' (g_{56} \rhd l_{1345})\,. \tag{B.79}$$

The integral of the l.h.s. reads

$$\int_H dh_{456} \int_{L^3} dl_{1456}dl_{2456}dl_{3456}\delta_G(g_{456})\, \delta_H(h_{3456})\delta_H(h_{2456})\delta_H(h_{1456})\delta_L(l_{23456})\delta_L(l_{13456})\delta_L(l_{12456})\,. \tag{B.80}$$

First, one integrates out the l_{1456}, exploiting $\delta_L(l_{13456})$ and obtains

$$l_{1456} = h_{146} \rhd \{h_{456}, (g_{56}g_{45}) \rhd h_{134}\}{l_{1346}}^{-1}(h_{136} \rhd' l_{3456})l_{1356}h_{156} \rhd' (g_{56} \rhd l_{1345}). \tag{B.81}$$

Next, one integrates out the l_{2456}, exploiting $\delta_L(l_{23456})$ and obtains

$$l_{2456} = h_{246} \rhd \{h_{456}, (g_{56}g_{45}) \rhd h_{234}\} {l_{2346}}^{-1}(h_{236} \rhd' l_{3456}) l_{2356} h_{256} \rhd' (g_{56} \rhd l_{2345}) . \quad \text{(B.82)}$$

Next, one integrates out h_{456}, exploiting $\delta_H(h_{3456})$ and obtains

$$h_{456} = h_{346}^{-1}\, \delta(l_{3456}) h_{356}\, (g_{56} \rhd h_{345}) . \quad \text{(B.83)}$$

Using the equation (B.83), one obtains that

$$\delta_G(g_{456}) = \delta_G\big(\partial(h_{346})^{-1}\, \partial(h_{356})\, g_{56} \rhd \partial(h_{345})\, g_{56}\, g_{45}\, g_{46}^{-1}\big) , \quad \text{(B.84)}$$

which, using the identity (4.7) for triangles (346), (356), and (345), reduces to:

$$\delta_G(g_{456}) = \delta_G(e) . \quad \text{(B.85)}$$

Similarly as done for the right-hand side of the move, one shows that $\delta_H(h_{1456})$, when using the equation (B.81), and $\delta_H(h_{2456})$, when using the equation (B.82), reduce to $\delta_H(e)^2$. The remaining $\delta_L(l_{12456})$ now reads

$$\delta_L(l_{12456}) = \delta_L({l_{1246}}^{-1}(h_{126} \rhd' l_{2456}) l_{1256} h_{156} \rhd' (g_{56} \rhd l_{1245}) {l_{1456}}^{-1}\, h_{146} \rhd \{h_{456}, (g_{56}g_{45}) \rhd h_{124}\}_{\mathrm{p}}) . \quad \text{(B.86)}$$

Substituting the equations (B.81) and (B.82), one obtains

$$\begin{aligned}\delta_L(l_{12456}) = \delta_L({l_{1246}}^{-1}(h_{126} \rhd' (h_{246} \rhd \{h_{456},(g_{56}g_{45}) \rhd h_{234}\}_{\mathrm{p}} {l_{2346}}^{-1}(h_{236} \rhd' l_{3456}) l_{2356}\\ h_{256} \rhd' (g_{56} \rhd l_{2345}))) l_{1256} h_{156} \rhd' (g_{56} \rhd l_{1245}) h_{156} \rhd' (g_{56} \rhd l_{1345})^{-1} l_{1356}^{-1}(h_{136} \rhd' l_{3456})^{-1}\\ l_{1346} h_{146} \rhd \{h_{456},(g_{56}g_{45}) \rhd h_{134}\}_{\mathrm{p}}^{-1} h_{146} \rhd \{h_{456},(g_{56}g_{45}) \rhd h_{124}\}_{\mathrm{p}}).\end{aligned} \quad \text{(B.87)}$$

After commuting the elements, i.e. , using the Peiffer identity for the crossed module $(L \xrightarrow{\delta} H, \rhd')$, one obtains

$$\begin{aligned}\delta_L(l_{12456}) = \delta_L((\delta(l_{1246})^{-1} h_{126} h_{246}) \rhd \{h_{456}, (g_{56}g_{45}) \rhd h_{234}\}_{\mathrm{p}} (\delta(l_{1246})^{-1} h_{126} \rhd \delta(l_{2346})^{-1} h_{126} h_{236}) \rhd' l_{3456}\\ l_{1246}^{-1} h_{126} \rhd' {l_{2346}}^{-1} h_{126} \rhd' l_{2356} (h_{126} h_{256}) \rhd' (g_{56} \rhd l_{2345}))\\ l_{1256} h_{156} \rhd' (g_{56} \rhd l_{1245}) h_{156} \rhd' (g_{56} \rhd l_{1345})^{-1} l_{1356}^{-1} l_{1346} (\delta(l_{1346})^{-1} h_{136}) \rhd' {l_{3456}}^{-1}\\ h_{146} \rhd \{h_{456}, (g_{56}g_{45}) \rhd h_{134}\}_{\mathrm{p}}^{-1}\, h_{146} \rhd \{h_{456}, (g_{56}g_{45}) \rhd h_{124}\}_{\mathrm{p}}) .\end{aligned} \quad \text{(B.88)}$$

Using the identity (3.10) for the inverse of the element $\{h_{456}, (g_{56}g_{45}) \rhd h_{134}\}_{\mathrm{p}}^{-1}$, and then the variant of the identity (3.5), i.e. , that is,

$$\{h_1, h_2h_3h_4\}_{\mathrm{p}} = \{h_1, h_2\}_{\mathrm{p}}(\partial(h_1) \rhd h_2) \rhd' \{h_1, h_3\}_{\mathrm{p}}(\partial(h_1) \rhd (h_2h_3)) \rhd' \{h_1, h_4\}_{\mathrm{p}} , \quad \text{(B.89)}$$

one obtains

$$\begin{aligned}\{h_{456}, (g_{56}g_{45}) \rhd (h_{134}^{-1} h_{124} h_{234})\}_{\mathrm{p}} = \{h_{456}, (g_{56}g_{45}) \rhd h_{134}^{-1}\}_{\mathrm{p}}(g_{46} \rhd h_{134}^{-1}) \rhd' \{h_{456}, (g_{56}g_{45}) \rhd h_{124}\}_{\mathrm{p}}\\ (g_{46} \rhd (h_{134}^{-1} h_{124})) \rhd' \{h_{456}, (g_{56}g_{45}) \rhd h_{124}\}_{\mathrm{p}} ,\end{aligned} \quad \text{(B.90)}$$

rendering the equation (B.88) to

$$\begin{aligned}\delta_L(l_{12456}) = \delta_L\big(&(\delta(l_{1246})^{-1}h_{126} \triangleright \delta(l_{2346})^{-1}h_{126}h_{236}) \triangleright' l_{3456} \\ &l_{1246}^{-1}h_{126} \triangleright' l_{2346}{}^{-1}h_{126} \triangleright' l_{2356}(h_{126}h_{256}) \triangleright' (g_{56} \triangleright l_{2345})\big) \\ &l_{1256}h_{156} \triangleright' (g_{56} \triangleright l_{1245})h_{156} \triangleright' (g_{56} \triangleright l_{1345})^{-1}l_{1356}^{-1}l_{1346}(\delta(l_{1346})^{-1}h_{136}) \triangleright' l_{3456}{}^{-1} \\ &(h_{146}g_{46} \triangleright h_{134}) \triangleright' \{h_{456}, (g_{56}g_{45}) \triangleright (h_{134}^{-1}h_{124}h_{234})\}_{\mathrm{p}}) \,.\end{aligned} \tag{B.91}$$

Using the equation (B.83), and the identities (3.4) and (3.6), similarly as for the r.h.s. of the move, one obtains that the terms featuring l_{3456} cancel, i.e. , the delta function $\delta_L(l_{12456})$ reads

$$\begin{aligned}\delta_L(l_{12456}) = \delta_L(&l_{1246}^{-1}h_{126} \triangleright' l_{2346}{}^{-1}h_{126} \triangleright' l_{2356}(h_{126}h_{256}) \triangleright' (g_{56} \triangleright l_{2345}))l_{1256}h_{156} \triangleright' (g_{56} \triangleright l_{1245}) \\ &h_{156} \triangleright' (g_{56} \triangleright l_{1345})^{-1}l_{1356}^{-1}l_{1346}(h_{146}g_{46} \triangleright h_{134}) \triangleright' \{h_{456}, (g_{56}g_{45}) \triangleright (h_{134}^{-1}h_{124}h_{234})\}_{\mathrm{p}}) \,.\end{aligned} \tag{B.92}$$

It follows that the integral over l_{3456} is now trivial and l.h.s. of the move finally reduces to:

$$\begin{aligned}\text{l.h.s.} = \delta_G(e)\delta_H(e)^2\delta_L\big(&h_{126} \triangleright' l_{2346}l_{1246}(h_{146}g_{46} \triangleright h_{134}) \triangleright' \{h_{456}, (g_{56}g_{45}) \triangleright (h_{134}^{-1}h_{124}h_{234})\}_{\mathrm{p}}^{-1} \\ &l_{1346}^{-1}\, l_{1356}\, h_{156} \triangleright' (g_{56} \triangleright l_{1345})\, h_{156} \triangleright' (g_{56} \triangleright l_{1245})^{-1}(h_{156}g_{56} \triangleright h_{125}) \triangleright' (g_{56} \triangleright l_{2345})^{-1} \\ &l_{1256}^{-1}\, h_{126} \triangleright' l_{2356}^{-1}\big) \,.\end{aligned} \tag{B.93}$$

The expressions (B.79) and (B.86) are the same, which proves the invariance of the state sum (4.1) under the Pachner move $3-3$. The numbers of k-simplices agree on both sides of the $3-3$ move for all k, and the prefactors play no role in this case.

References

[1] C. Rovelli, *Quantum Gravity*, Cambridge University Press, Cambridge, U.K. (2004).

[2] C. Rovelli, *Zakopane lectures on loop gravity*, *PoS* **QGQGS2011** (2011) 003 [arXiv:1102.3660] [INSPIRE].

[3] T. Thiemann, *Modern Canonical Quantum General Relativity*, Cambridge University Press, Cambridge, U.K. (2007).

[4] C. Rovelli and F. Vidotto, *Covariant Loop Quantum Gravity*, Cambridge University Press, Cambridge, U.K. (2014).

[5] J.F. Plebanski, *On the separation of Einsteinian substructures*, *J. Math. Phys.* **18** (1977) 2511 [INSPIRE].

[6] G. Ponzano and T. Regge, *Spectroscopic and Group Theoretical Methods in Physics. Racah Memorial Volume*, North-Holland, Amsterdam, Netherlands (1968) pp. 75–103.

[7] J.W. Barrett and L. Crane, *Relativistic spin networks and quantum gravity*, *J. Math. Phys.* **39** (1998) 3296 [gr-qc/9709028] [INSPIRE].

JHEP07(2022)105

[8] J.W. Barrett and L. Crane, *A Lorentzian signature model for quantum general relativity*, *Class. Quant. Grav.* **17** (2000) 3101 [gr-qc/9904025] [INSPIRE].

[9] H. Ooguri, *Topological lattice models in four-dimensions*, *Mod. Phys. Lett. A* **7** (1992) 2799 [hep-th/9205090] [INSPIRE].

[10] J. Engle, E. Livine, R. Pereira and C. Rovelli, *LQG vertex with finite Immirzi parameter*, *Nucl. Phys. B* **799** (2008) 136 [arXiv:0711.0146] [INSPIRE].

[11] L. Freidel and K. Krasnov, *A New Spin Foam Model for 4d Gravity*, *Class. Quant. Grav.* **25** (2008) 125018 [arXiv:0708.1595] [INSPIRE].

[12] A. Miković and M. Vojinović, *Effective action for EPRL/FK spin foam models*, *J. Phys. Conf. Ser.* **360** (2012) 012049 [arXiv:1110.6114] [INSPIRE].

[13] A. Miković and M. Vojinović, *A finiteness bound for the EPRL/FK spin foam model*, *Class. Quant. Grav.* **30** (2013) 035001 [arXiv:1101.3294] [INSPIRE].

[14] A. Miković and M. Vojinović, *Solution to the Cosmological Constant Problem in a Regge Quantum Gravity Model*, *EPL* **110** (2015) 40008 [arXiv:1407.1394] [INSPIRE].

[15] E. Bianchi, M. Han, C. Rovelli, W. Wieland, E. Magliaro and C. Perini, *Spinfoam fermions*, *Class. Quant. Grav.* **30** (2013) 235023 [arXiv:1012.4719] [INSPIRE].

[16] J.C. Baez and J. Huerta, *An Invitation to Higher Gauge Theory*, *Gen. Rel. Grav.* **43** (2011) 2335 [arXiv:1003.4485] [INSPIRE].

[17] Z. Li, *A Global Geometric Approach to Parallel Transport of Strings in Gauge Theory*, arXiv:1910.14230 [INSPIRE].

[18] C. Sämann and M. Wolf, *Non-Abelian Tensor Multiplet Equations from Twistor Space*, *Commun. Math. Phys.* **328** (2014) 527 [arXiv:1205.3108] [INSPIRE].

[19] B. Jurčo, C. Sämann and M. Wolf, *Semistrict Higher Gauge Theory*, *JHEP* **04** (2015) 087 [arXiv:1403.7185] [INSPIRE].

[20] Y. Hidaka, M. Nitta and R. Yokokura, *Higher-form symmetries and 3-group in axion electrodynamics*, *Phys. Lett. B* **808** (2020) 135672 [arXiv:2006.12532] [INSPIRE].

[21] Y. Hidaka, M. Nitta and R. Yokokura, *Global 3-group symmetry and 't Hooft anomalies in axion electrodynamics*, *JHEP* **01** (2021) 173 [arXiv:2009.14368] [INSPIRE].

[22] C. Säemann and M. Wolf, *Six-Dimensional Superconformal Field Theories from Principal 3-Bundles over Twistor Space*, *Lett. Math. Phys.* **104** (2014) 1147 [arXiv:1305.4870] [INSPIRE].

[23] D. Song, K. Lou, K. Wu, J. Yang and F. Zhang, *3-form Yang-Mills based on 2-crossed modules*, *J. Geom. Phys.* **178** (2022) 104537 [arXiv:2108.12852] [INSPIRE].

[24] D. Song, K. Lou, K. Wu and J. Yang, *Higher form Yang-Mills as higher BFYM theories*, arXiv:2109.13443 [INSPIRE].

[25] Y. Hidaka, M. Nitta and R. Yokokura, *Topological axion electrodynamics and 4-group symmetry*, *Phys. Lett. B* **823** (2021) 136762 [arXiv:2107.08753] [INSPIRE].

[26] Y. Hidaka, M. Nitta and R. Yokokura, *Global 4-group symmetry and 't Hooft anomalies in topological axion electrodynamics*, *PTEP* **2022** (2022) 04A109 [arXiv:2108.12564] [INSPIRE].

JHEP07(2022)105

[27] B. Jurčo, C. Sämann and M. Wolf, *Higher Groupoid Bundles, Higher Spaces, and Self-Dual Tensor Field Equations*, Fortsch. Phys. **64** (2016) 674 [arXiv:1604.01639] [INSPIRE].

[28] C. Sämann and M. Wolf, *Supersymmetric Yang-Mills Theory as Higher Chern-Simons Theory*, JHEP **07** (2017) 111 [arXiv:1702.04160] [INSPIRE].

[29] B. Jurčo, T. Macrelli, L. Raspollini, C. Sämann and M. Wolf, *L_∞-Algebras, the BV Formalism, and Classical Fields*, Fortsch. Phys. **67** (2019) 1910025 [arXiv:1903.02887] [INSPIRE].

[30] A. Miković and M. Vojinović, *Poincaré 2-group and quantum gravity*, Class. Quant. Grav. **29** (2012) 165003 [arXiv:1110.4694] [INSPIRE].

[31] T. Radenković and M. Vojinović, *Higher Gauge Theories Based on 3-groups*, JHEP **10** (2019) 222 [arXiv:1904.07566] [INSPIRE].

[32] A. Miković and M. Vojinović, *Standard Model and 4-groups*, EPL **133** (2021) 61001 [arXiv:2008.06354] [INSPIRE].

[33] T. Porter, *Topological quantum field theories from homotopy n-types*, J. London Math. Soc. **58** (1998) 723.

[34] T. Porter, *Interpretations of Yetter's notion of G-coloring: simplicial fibre bundles and non-abelian cohomology*, J. Knot Th. Ramif. **5** (1996) 687.

[35] U. Pachner, *PL homeomorphic manifolds are equivalent by elementary shellings*, Europ. J. Combinat. **12** (1991) 129 [arXiv:1095161].

[36] M. Celada, D. González and M. Montesinos, *BF gravity*, Class. Quant. Grav. **33** (2016) 213001 [arXiv:1610.02020] [INSPIRE].

[37] J.C. Baez, *An Introduction to Spin Foam Models of BF Theory and Quantum Gravity*, Lect. Notes Phys. **543** (2000) 25 [gr-qc/9905087] [INSPIRE].

[38] F. Girelli, H. Pfeiffer and E.M. Popescu, *Topological Higher Gauge Theory - from BF to BFCG theory*, J. Math. Phys. **49** (2008) 032503 [arXiv:0708.3051] [INSPIRE].

[39] J.F. Martins and A. Miković, *Lie crossed modules and gauge-invariant actions for 2-BF theories*, Adv. Theor. Math. Phys. **15** (2011) 1059 [arXiv:1006.0903] [INSPIRE].

[40] A. Miković and M.A. Oliveira, *Canonical formulation of Poincaré BFCG theory and its quantization*, Gen. Rel. Grav. **47** (2015) 58 [arXiv:1409.3751] [INSPIRE].

[41] A. Miković, M.A. Oliveira and M. Vojinović, *Hamiltonian analysis of the BFCG theory for a strict Lie 2-group*, arXiv:1610.09621 [INSPIRE].

[42] J.F. Martins and R. Picken, *The fundamental Gray 3-groupoid of a smooth manifold and local 3-dimensional holonomy based on a 2-crossed module*, arXiv:0907.2566 [INSPIRE].

[43] W. Wang, *On 3-gauge transformations, 3-curvatures, and Gray-categories*, J. Math. Phys. **55** (2014) 043506 [arXiv:1311.3796] [INSPIRE].

[44] T. Radenković and M. Vojinović, *Hamiltonian Analysis for the Scalar Electrodynamics as 3BF Theory*, Symmetry **12** (2020) 620 [arXiv:2004.06901] [INSPIRE].

[45] T. Radenković and M. Vojinović, *Gauge symmetry of the 3BF theory for a generic semistrict Lie three-group*, Class. Quant. Grav. **39** (2022) 135009 [arXiv:2101.04049] [INSPIRE].

JHEP07(2022)105

[46] D. Conduché, *Modules croisés généralisés de longueur 2*, in proceedings of the *Luminy conference on algebraic K-theory (Luminy, 1983)*, *J. Pure Appl. Algebra* **34**, (1984) 155.

[47] F. Girelli and H. Pfeiffer, *Higher gauge theory: Differential versus integral formulation*, *J. Math. Phys.* **45** (2004) 3949 [hep-th/0309173] [INSPIRE].

[48] D.N. Yetter, *Topological quantum field theories associated to finite groups and crossed G-sets*, *J. Knot Theory Ramifications* **1** (1992) 1.

[49] D.N. Yetter, *TQFT's from homotopy 2-types*, *J. Knot Theory Ramifications* **2** (1993) 113.

[50] A. Baratin and L. Freidel, *A 2-categorical state sum model*, *J. Math. Phys.* **56** (2015) 011705 [arXiv:1409.3526] [INSPIRE].

[51] S.K. Asante, B. Dittrich, F. Girelli, A. Riello and P. Tsimiklis, *Quantum geometry from higher gauge theory*, *Class. Quant. Grav.* **37** (2020) 205001 [arXiv:1908.05970] [INSPIRE].

[52] F. Girelli, M. Laudonio and P. Tsimiklis, *Polyhedron phase space using 2-groups: κ-Poincaré as a Poisson 2-group*, arXiv:2105.10616 [INSPIRE].

JHEP07(2022)105

IOP Publishing Classical and Quantum Gravity

Class. Quantum Grav. **39** (2022) 135009 (51pp) https://doi.org/10.1088/1361-6382/ac6b78

Gauge symmetry of the 3*BF* theory for a generic semistrict Lie three-group

Tijana Radenković* **and Marko Vojinović**

Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia

E-mail: rtijana@ipb.ac.rs

Received 17 January 2022, revised 22 April 2022
Accepted for publication 28 April 2022
Published 8 June 2022

CrossMark

Abstract

The higher category theory can be employed to generalize the *BF* action to the so-called 3*BF* action, by passing from the notion of a gauge group to the notion of a gauge three-group. In this work we determine the full gauge symmetry of the 3*BF* action. To that end, the complete Hamiltonian analysis of the 3*BF* action for an arbitrary semistrict Lie three-group is performed, by using the Dirac procedure. The Hamiltonian analysis is the first step towards a canonical quantization of a 3*BF* theory. This is an important stepping-stone for the quantization of the complete standard model of elementary particles coupled to Einstein–Cartan gravity, formulated as a 3*BF* action with suitable simplicity constraints. We show that the resulting gauge symmetry group consists of the familiar *G*-, *H*-, and *L*-gauge transformations, as well as additional *M*- and *N*-gauge transformations, which have not been discussed in the existing literature.

Keywords: quantum gravity, higher gauge theory, higher category theory, three-group, *BF* action, 3*BF* action, gauge symmetry

Contents

*Author to whom any correspondence should be addressed.

1. Introduction

Among the most important open problems in contemporary theoretical physics is the problem of quantization of the gravitational field. Within the framework of loop quantum gravity (LQG), one of the most prominent candidates for the quantum theory of gravity, the study of nonperturbative quantization has evolved in two directions: the canonical and the covariant approach. See [1–4] for an overview and a comprehensive introduction to the theory.

The *covariant quantization* approach focuses on defining the gravitational path integral of the theory:

$$Z_{\text{gr}} = \int \mathcal{D}g\, \mathrm{e}^{\mathrm{i}S_{\text{gr}}[g]}. \tag{1}$$

In order to give the rigorous definition of the path integral, the classical action of the theory S_{gr} is written as a sum of the topological BF action, i.e. the action with no propagating degrees of freedom, and the part featuring the simplicity constraints, i.e. sum of products of Langrange multipliers and the corresponding simplicity constraints imposed on the variables of the topological part of the action. Next, one defines the path-integral of the topological theory given by the BF action, using the topological quantum field theory (TQFT) formalism. Once a path-integral is defined for the topological sector, it is deformed into a non-topological theory, by imposing the simplicity constraints. This quantization technique is known as the *spinfoam quantization* method.

The spinfoam quantization procedure has been successfully employed in various theories, including the three-dimensional topological Ponzano–Regge model of quantum gravity [5], the four-dimensional topological Ooguri model [6], the Barrett–Crane model of gravity in four dimensions [7–9], and others. The most successful among these is the renowned EPRL/FK model [10, 11], which had been specifically formulated to correspond to the quantum theory of gravity obtained by the *canonical loop quantization*, where a state of the gravitational field is described by the so-called *spin network*.

However, note that all mentioned models, formulated as constrained *BF* actions, are theories of pure gravity, without matter fields. Recently, as an endeavor to formulate a theory that unifies all the known interactions, one interesting new avenue of research has been opened, based on a categorical generalization of the *BF* action in the context of higher gauge theory (HGT) formalism [12]. One novel candidate discussed in the literature [13], uses the three-group structure to formulate the 3*BF* action as a categorical generalization of the *BF* theory. Then, modifying the pure 3*BF* action by adding the appropriate simplicity constraints, one obtains the *constrained* 3*BF* action, describing the theory of all the fields present in the standard model coupled in a standard way to Einstein–Cartan gravity.

Once the appropriate classical theory has been constructed, one needs to quantize it by constructing a topological state sum Z using the algebraic structure underlying the topological sector of the constrained 3*BF* action, i.e. the underlying two-crossed module. This construction has been recently carried out in [14], where a triangulation independent state sum Z of a topological HGT for an arbitrary two-crossed module and a four-dimensional closed and orientable spacetime manifold $\mathcal{M}_4$ is defined. Once the topological state sum is formulated, one could proceed to modify the amplitudes of the state sum in order to impose the simplicity constraints and obtain the state sum describing the full theory. In this way one would finally arrive at the rigorous definition of a path integral given by the equation (1).

In addition to the covariant approach, one can also study the constrained 3*BF* action, using the *canonical quantization*. This approach focuses on defining the quantum theory via a triple $(\mathcal{H}, \mathcal{A}, W)$, i.e. the Hilbert space of states $\mathcal{H}$, the algebra of observables $\mathcal{A}$, and the dynamics W given by the transition amplitudes. Specifically, in canonical LQG, the algebra of fields that are promoted to the quantum operators is chosen to be the algebra based on the holonomies of the gravitational connection. However, in the case of the 3*BF* theory, the notion of connection is generalized to the notion of three-connection, which makes its canonical quantization approach an interesting avenue of research. The first step toward the canonical quantization of the theory is the Hamiltonian analysis, resulting in the algebra of first-class and second-class constraints. The first-class constraints become conditions on the physical states determining the Hilbert space, while the Hamiltonian constraint determines the dynamics.

The results presented in this paper are the natural continuation of the results presented in [13]. The main result is the calculation of the full symmetry group of the pure 3*BF* action. To that end, the complete Hamiltonian analysis of the 3*BF* action for a semistrict Lie three-group is performed by using the Dirac procedure (see [15] for an overview and a comprehensive introduction to the Hamiltonian analysis). It is a generalization of the Hamiltonian analysis of a 2*BF* action performed in [16–19], and of the Hamiltonian analysis for the special case of a two-crossed module corresponding to the theory of scalar electrodynamics, carried out in [20]. The analysis of the Hamiltonian structure of the theory gives us the algebra of first-class and second-class constraints present in the theory. As usual, the first-class constraints generate gauge transformations, which do not change the physical state of the system. Using the Castellani's procedure, one can find the generator of the gauge transformations in the theory on a spatial hypersurface. Then, the results obtained by this method are generalized to the whole spacetime. The complete gauge symmetry, consisting of five types of finite gauge transformations, along with the proofs that they are indeed the gauge symmetries of 3*BF* action, is presented. With these results in hand, the structure of the full gauge symmetry group is analyzed, and its corresponding Lie algebra is determined.

The obtained results give rise to a connection between the gauge symmetry group of the 3*BF* action, and its underlining three-group structure, establishing a *duality* between the two. This analysis is an important step towards the study of the gauge symmetry group of the theory of gravity with matter, formulated as the constrained 3*BF* action [13], as well as its canonical quantization. Furthermore, it is important for the overall understanding of the physical meaning of the three-group structure and its interpretation as the underlining symmetry of the pure 3*BF* action, which represents a basis for the constrained 3*BF* action describing the physical theory.

The layout of the paper is as follows. In section 2, we give a brief overview of *BF* and 2*BF* theories, and introduce the 3*BF* action. Section 3 contains the Hamiltonian analysis for the 3*BF* theory. In subsection 3.1, the canonical structure of the theory is obtained, while in subsection 3.2 the resulting first-class and second-class constraints present in the theory, as well as the algebra of constraints, are presented. In the subsection 3.3 we analyze the Bianchi identities (BI) that the first-class constraints satisfy, which enforce restrictions in the sense of Hamiltonian analysis, and reduce the number of independent first-class constraints present in the theory. We then proceed with the counting of the physical degrees of freedom. Finally, this section concludes with the subsection 3.4 where we construct the generator of the gauge symmetries for the topological theory, based on the calculations done in section 3.2.

Section 4 contains the main results of our paper and is devoted to the analysis of the symmetries of the 3*BF* action. Having results of the subsection 3.4 in hand, we find the form variations of all variables and their canonical momenta, and use that result to determine all gauge transformations of the theory. This is done in four steps. The subsection 4.1 deals with the gauge group G, and the corresponding G-gauge transformations. In subsection 4.2 we discuss the gauge group $\tilde{H}_L$ which consists of the H-gauge and L-gauge transformations (familiar from [21]), while the subsection 4.3 examines the novel M-gauge and N-gauge transformations which also arise in the theory. The results of the subsections 4.1–4.3 are summarized in subsection 4.4, where the complete structure of the symmetry group is presented, including its Lie algebra. Our concluding remarks are given in section 5, containing a summary and a discussion of the obtained results, as well as possible future lines of investigation. The appendices contain various technical details concerning three-groups, additional relations of the constraint algebra, the computation of the generator of gauge symmetries, form-variations of all fields and momenta, and some other technical details.

Our notation and conventions are as follows. Spacetime indices, denoted by the mid-alphabet Greek letters $\mu, \nu, \dots$, are raised and lowered by the spacetime metric $g_{\mu\nu}$. The spatial part of these is denoted with lowercase mid-alphabet Latin indices $i, j, \dots$, and the time component is denoted with 0. The indices that are counting the generators of groups G, H, and L are denoted with initial Greek letters $\alpha, \beta, \dots$, lowercase initial Latin letters $a, b, c, \dots$, and uppercase Latin indices $A, B, C, \dots$, respectively. The antisymmetrization over two indices is denoted as $A_{[a_1|a_2\dots a_{n-1}|a_n]} = \frac{1}{2}\left(A_{a_1a_2\dots a_{n-1}a_n} - A_{a_na_2\dots a_{n-1}a_1}\right)$, while the total antisymmetrization is denoted as $A_{[a_1\dots a_n]} = \frac{1}{n!}\sum_{\sigma\in S_n}(-1)^{\text{sgn}(\sigma)}A_{a_{\sigma(1)}\dots a_{\sigma(n)}}$. Likewise, the symmetrization over two indices is denoted as $A_{(a_1|a_2\dots a_{n-1}|a_n)} = \frac{1}{2}\left(A_{a_1a_2\dots a_{n-1}a_n} + A_{a_na_2\dots a_{n-1}a_1}\right)$, while the total symmetrization is denoted as $A_{(a_1\dots a_n)} = \frac{1}{n!}\sum_{\sigma\in S_n}A_{a_{\sigma(1)}\dots a_{\sigma(n)}}$. We work in the natural system of units, defined by $c = \hbar = 1$ and $G = l_\text{p}^2$, where l_p is the Planck length. All additional notation and conventions used throughout the paper are explicitly defined in the text where they appear.

2. The 3*BF* theory

Given a Lie group G and its corresponding Lie algebra $\mathfrak{g}$, one can introduce the so-called *BF* action as

$$S_{BF} = \int_{\mathcal{M}_4} \langle B \wedge F \rangle_{\mathfrak{g}}, \tag{2}$$

where $F \equiv \mathrm{d}\alpha + \alpha \wedge \alpha$ is the curvature two-form for the algebra-valued connection one-form $\alpha \in \mathcal{A}^1(\mathcal{M}_4, \mathfrak{g})$ on a trivial principal G-bundle over a four-dimensional compact and orientable spacetime manifold $\mathcal{M}_4$, and $B \in \mathcal{A}^2(\mathcal{M}_4, \mathfrak{g})$ is a Lagrange multiplier two-form. The $\langle \_, \_ \rangle_{\mathfrak{g}}$ denotes the G-invariant bilinear symmetric nondegenerate form on $\mathfrak{g}$. For more details see [22–24].

Varying the action (2) with respect to the Lagrange multiplier B and the connection α, one obtains the equations of motion,

$$F = 0, \qquad \nabla B \equiv \mathrm{d}B + \alpha \wedge B = 0. \tag{3}$$

These equations of motion imply that α is a flat connection, while the Lagrange multiplier B is a constant field. Therefore, the theory given by the *BF* action has no local propagating degrees of freedom, i.e. the theory is topological.

Within the framework of HGT, one can define the categorical generalization of the *BF* action to the so-called 2*BF* action, by passing from the notion of a gauge group to the notion of a gauge two-group, see [25–27]. In the category theory, a two-group is defined as a two-category consisting of only one object, where all the morphisms and two-morphisms are invertible. It has been shown that every strict two-group is equivalent to a crossed module $(H \xrightarrow{\partial} G, \triangleright)$, where G and H are groups, δ is a homomorphism from H to G, while $\triangleright : G \times H \to H$ is an action of G on H. Given a crossed-module $(H \xrightarrow{\partial} G, \triangleright)$, one can introduce a generalization of the *BF* action, the so-called 2*BF* action [25, 26]:

$$S_{2BF} = \int_{\mathcal{M}_4} \langle B \wedge \mathcal{F} \rangle_{\mathfrak{g}} + \langle C \wedge \mathcal{G} \rangle_{\mathfrak{h}}, \tag{4}$$

where the two-form $B \in \mathcal{A}^2(\mathcal{M}_4, \mathfrak{g})$ and the one-form $C \in \mathcal{A}^1(\mathcal{M}_4, \mathfrak{h})$ are Lagrange multipliers, and $\mathfrak{h}$ is a Lie algebra of the Lie group H. The variables $\mathcal{F} \in \mathcal{A}^2(\mathcal{M}_4, \mathfrak{g})$ and $\mathcal{G} \in \mathcal{A}^3(\mathcal{M}_4, \mathfrak{h})$ define the *fake* two-*curvature* $(\mathcal{F}, \mathcal{G})$ for the two-connection (α, β) on a trivial principal two-bundle over a four-dimensional compact and oriented spacetime manifold $\mathcal{M}_4$. See [28] for a rigorous definition. Here the two-connection (α, β) is given by $\mathfrak{g}$-valued one-form $\alpha \in \mathcal{A}^1(\mathcal{M}_4, \mathfrak{g})$ and an $\mathfrak{h}$-valued two-form $\beta \in \mathcal{A}^2(\mathcal{M}_4, \mathfrak{h})$:

$$\mathcal{F} = \mathrm{d}\alpha + \alpha \wedge \alpha - \partial\beta, \qquad \mathcal{G} = \mathrm{d}\beta + \alpha \wedge^{\triangleright} \beta. \tag{5}$$

The two-curvature $(\mathcal{F}, \mathcal{G})$ is called *fake*, because of the additional term $\partial\beta$, see [12]. Also, $\langle \_, \_ \rangle_{\mathfrak{g}}$ and $\langle \_, \_ \rangle_{\mathfrak{h}}$ denote the G-invariant bilinear symmetric nondegenerate forms for the algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively. See [25, 26] for review and references. Varying the 2*BF* action (4) with respect to variables B and C one obtains the equations of motion

$$\mathcal{F} = 0, \qquad \mathcal{G} = 0, \tag{6}$$

while varying with respect to α and β one obtains

$$\mathrm{d}B_\alpha - f_{\alpha\beta}{}^{\gamma} B_\gamma \wedge \alpha^\beta - \triangleright_{\alpha a}{}^{b} C_b \wedge \beta^a = 0, \tag{7}$$

$$dC_a - \partial_a{}^\alpha B_\alpha + \triangleright_{\alpha a}{}^b C_b \wedge \alpha^\alpha = 0. \tag{8}$$

Here, the coefficients $f_{\alpha\beta}{}^\gamma$ are the structure constants of the algebra $\mathfrak{g}$, $\triangleright_{\alpha a}{}^b$ are the coefficients of the action $\triangleright$ of the algebra $\mathfrak{g}$ on $\mathfrak{h}$, while $\partial_a{}^\alpha$ are the coefficients of the map ∂, given in the bases of algebras $\mathfrak{g}$ and $\mathfrak{h}$ (see the equations (10)–(12) below). Similarly to the case of the *BF* action, the 2*BF* action defines a topological theory, i.e. a theory with no propagating degrees of freedom, see [16, 19].

Continuing the categorical generalization one step further, one can generalize the notion of a two-group to the notion of a three-group. Similarly to the definition of a group and a two-group within the category theory formalism, a three-group is defined as a three-category with only one object, where all morphisms, two-morphisms, and three-morphisms are invertible. Moreover, analogously as a strict two-group is equivalent to a crossed-module, it has been proved that a semistrict three-group is equivalent to a two-crossed module [29].

A Lie two-crossed module, denoted as $(L \xrightarrow{\delta} H \xrightarrow{\partial} G, \triangleright, \{\_,\_\}_{\mathrm{pf}})$ (see appendix A for the precise definition), is an algebraic structure specified by three Lie groups G, H, and L, together with the homomorphisms $\delta : L \to H$ and $\partial : H \to G$, an action $\triangleright$ of the group G on all three groups, and a G-equivariant map, called the Peiffer lifting:

$$\{\_,\_\}_{\mathrm{pf}} : H \times H \to L.$$

In order for this structure to be a three-group, the structure constants of algebras $\mathfrak{g}$, $\mathfrak{h}$, and $\mathfrak{l}$, together with the maps ∂ and δ, the action $\triangleright$, and the Peiffer lifting, must satisfy certain axioms, see [13]. Here $\mathfrak{g}$, $\mathfrak{h}$, and $\mathfrak{l}$ denote the Lie algebras corresponding to the Lie groups G, H, and L.

Analogously to the definition of a two-connection given in [28], one can define a three-connection as follows. Given a two-crossed module and a four-dimensional compact and orientable spacetime manifold $\mathcal{M}_4$, one can introduce a trivial principal three-bundle using the two-crossed module as a fiber over the base manifold $\mathcal{M}_4$. See [21, 29] for the precise definition of a corresponding three-holonomy. This gives rise to a three-connection, which can be represented as an ordered triple (α, β, γ), where α, β, and γ are algebra-valued differential forms, $\alpha \in \mathcal{A}^1(\mathcal{M}_4, \mathfrak{g})$, $\beta \in \mathcal{A}^2(\mathcal{M}_4, \mathfrak{h})$, and $\gamma \in \mathcal{A}^3(\mathcal{M}_4, \mathfrak{l})$. The corresponding fake three-curvature $(\mathcal{F}, \mathcal{G}, \mathcal{H})$ is defined as:

$$\mathcal{F} = d\alpha + \alpha \wedge \alpha - \partial\beta, \quad \mathcal{G} = d\beta + \alpha \wedge^\triangleright \beta - \delta\gamma, \quad \mathcal{H} = d\gamma + \alpha \wedge^\triangleright \gamma + \{\beta \wedge \beta\}_{\mathrm{pf}}. \tag{9}$$

Similarly as in the case of the 2*BF* theory, the three-curvature $(\mathcal{F}, \mathcal{G}, \mathcal{H})$ is called *fake*, because of the additional terms $\partial\beta$, $\delta\gamma$, and $\{\beta \wedge \beta\}_{\mathrm{pf}}$. Fixing the bases in algebras $\mathfrak{g}$, $\mathfrak{h}$, and $\mathfrak{l}$ as $\tau_\alpha \in \mathfrak{g}$, $t_a \in \mathfrak{h}$, and $T_A \in \mathfrak{l}$, one defines the structure constants

$$[\tau_\alpha, \tau_\beta] = f_{\alpha\beta}{}^\gamma\, \tau_\gamma, \qquad [t_a, t_b] = f_{ab}{}^c\, t_c, \qquad [T_A, T_B] = f_{AB}{}^C\, T_C, \tag{10}$$

maps $\partial : H \to G$ and $\delta : L \to H$ as

$$\partial(t_a) = \partial_a{}^\alpha\, \tau_\alpha, \qquad \delta(T_A) = \delta_A{}^a\, t_a, \tag{11}$$

and an action of $\mathfrak{g}$ on the generators of $\mathfrak{g}$, $\mathfrak{h}$, and $\mathfrak{l}$ as

$$\tau_\alpha \triangleright \tau_\beta = f_{\alpha\beta}{}^\gamma\, \tau_\gamma, \qquad \tau_\alpha \triangleright t_a = \triangleright_{\alpha a}{}^b\, t_b, \qquad \tau_\alpha \triangleright T_A = \triangleright_{\alpha A}{}^B\, T_B, \tag{12}$$

respectively. To define the Peiffer lifting in a basis, one specifies the coefficients $X_{ab}{}^A$:

$$\{t_a,\, t_b\}_{\mathrm{pf}} = X_{ab}{}^A T_A. \tag{13}$$

Writing the curvature in the bases of the corresponding algebras and differential forms

$$\mathcal{F} = \frac{1}{2}\mathcal{F}^{\alpha}{}_{\mu\nu}\tau_{\alpha}\mathrm{d}x^{\mu}\wedge\mathrm{d}x^{\nu}, \quad \mathcal{G} = \frac{1}{3!}\mathcal{G}^{a}{}_{\mu\nu\rho}t_{a}\mathrm{d}x^{\mu}\wedge\mathrm{d}x^{\nu}\wedge\mathrm{d}x^{\rho},$$
$$\mathcal{H} = \frac{1}{4!}\mathcal{H}^{A}{}_{\mu\nu\rho\sigma}T_{A}\mathrm{d}x^{\mu}\wedge\mathrm{d}x^{\nu}\wedge\mathrm{d}x^{\rho}\wedge\mathrm{d}x^{\sigma},$$

one obtains the corresponding components:

$$\begin{aligned}
\mathcal{F}^{\alpha}{}_{\mu\nu} &= \partial_{\mu}\alpha^{\alpha}{}_{\nu} - \partial_{\nu}\alpha^{\alpha}{}_{\mu} + f_{\beta\gamma}{}^{\alpha}\alpha^{\beta}{}_{\mu}\alpha^{\gamma}{}_{\nu} - \beta^{a}{}_{\mu\nu}\partial_{a}{}^{\alpha},\\
\mathcal{G}^{a}{}_{\mu\nu\rho} &= \partial_{\mu}\beta^{a}{}_{\nu\rho} + \partial_{\nu}\beta^{a}{}_{\rho\mu} + \partial_{\rho}\beta^{a}{}_{\mu\nu}\\
&\quad + \alpha^{\alpha}{}_{\mu}\beta^{b}{}_{\nu\rho}\rhd_{\alpha b}{}^{a} + \alpha^{\alpha}{}_{\nu}\beta^{b}{}_{\rho\mu}\rhd_{\alpha b}{}^{a} + \alpha^{\alpha}{}_{\rho}\beta^{b}{}_{\mu\nu}\rhd_{\alpha b}{}^{a} - \gamma^{A}{}_{\mu\nu\rho}\delta_{A}{}^{a},\\
\mathcal{H}^{A}{}_{\mu\nu\rho\sigma} &= \partial_{\mu}\gamma^{A}{}_{\nu\rho\sigma} - \partial_{\nu}\gamma^{A}{}_{\rho\sigma\mu} + \partial_{\rho}\gamma^{A}{}_{\sigma\mu\nu} - \partial_{\sigma}\gamma^{A}{}_{\mu\nu\rho}\\
&\quad + 2\beta^{a}{}_{\mu\nu}\beta^{b}{}_{\rho\sigma}X_{(ab)}{}^{A} - 2\beta^{a}{}_{\mu\rho}\beta^{b}{}_{\nu\sigma}X_{(ab)}{}^{A} + 2\beta^{a}{}_{\mu\sigma}\beta^{b}{}_{\nu\rho}X_{(ab)}{}^{A}\\
&\quad + \alpha^{\alpha}{}_{\mu}\gamma^{B}{}_{\nu\rho\sigma}\rhd_{\alpha B}{}^{A} - \alpha^{\alpha}{}_{\nu}\gamma^{B}{}_{\rho\sigma\mu}\rhd_{\alpha B}{}^{A} + \alpha^{\alpha}{}_{\rho}\gamma^{B}{}_{\sigma\mu\nu}\rhd_{\alpha B}{}^{A}\\
&\quad - \alpha^{\alpha}{}_{\sigma}\gamma^{B}{}_{\mu\nu\rho}\rhd_{\alpha B}{}^{A}.
\end{aligned} \tag{14}$$

Then, similarly to the construction of BF and $2BF$ actions, one can define the gauge invariant topological $3BF$ action, with the underlying structure of a three-group. For the four-dimensional compact and orientable manifold $\mathcal{M}_4$ and the two-crossed module $(L \xrightarrow{\delta} H \xrightarrow{\partial} G, \rhd, \{\_,\_\}_{\mathrm{pf}})$, that gives rise to three-curvature (9), one defines the $3BF$ action as

$$S_{3BF} = \int_{\mathcal{M}_4} \langle B\wedge\mathcal{F}\rangle_{\mathfrak{g}} + \langle C\wedge\mathcal{G}\rangle_{\mathfrak{h}} + \langle D\wedge\mathcal{H}\rangle_{\mathfrak{l}}, \tag{15}$$

where $B\in\mathcal{A}^2(\mathcal{M}_4,\mathfrak{g})$, $C\in\mathcal{A}^1(\mathcal{M}_4,\mathfrak{h})$, and $D\in\mathcal{A}^0(\mathcal{M}_4,\mathfrak{l})$ are Lagrange multipliers. The forms $\langle\_,\_\rangle_{\mathfrak{g}}$, $\langle\_,\_\rangle_{\mathfrak{h}}$, and $\langle\_,\_\rangle_{\mathfrak{l}}$ are G-invariant bilinear symmetric nondegenerate forms on $\mathfrak{g}$, $\mathfrak{h}$, and $\mathfrak{l}$, respectively. Note that in the case of a semisimple Lie algebra, a natural choice for this bilinear form is the Killing form. However, one can also choose it differently, and moreover for a solvable Lie algebra one can introduce a non-trivial bilinear form, despite the fact that the Killing form is degenerate in this case. Fixing the basis in algebras $\mathfrak{g}$, $\mathfrak{h}$, and $\mathfrak{l}$, as defined in (10), the forms $\langle\_,\_\rangle_{\mathfrak{g}}$, $\langle\_,\_\rangle_{\mathfrak{h}}$, and $\langle\_,\_\rangle_{\mathfrak{l}}$ map pairs of basis vectors of algebras $\mathfrak{g}$, $\mathfrak{h}$, and $\mathfrak{l}$, to the metrics on their vector spaces, $g_{\alpha\beta}$, g_{ab}, and g_{AB}:

$$\langle\tau_{\alpha},\tau_{\beta}\rangle_{\mathfrak{g}} = g_{\alpha\beta}, \qquad \langle t_a, t_b\rangle_{\mathfrak{h}} = g_{ab}, \qquad \langle T_A, T_B\rangle_{\mathfrak{l}} = g_{AB}. \tag{16}$$

As the symmetric maps are nondegenerate, the inverse metrics $g^{\alpha\beta}$, g^{ab}, and g^{AB} are well defined, and are used to raise and lower indices of the corresponding algebras.

Varying the action (15) with respect to Lagrange multipliers B^{α}, C^{a}, and D^{A} one obtains the equations of motion

$$\mathcal{F}^{\alpha} = 0, \qquad \mathcal{G}^{a} = 0, \qquad \mathcal{H}^{A} = 0, \tag{17}$$

while varying with respect to the three-connection variables α^α, β^a, and γ^A one gets:

$$\mathrm{d}B_\alpha - f_{\alpha\beta}{}^\gamma B_\gamma \wedge \alpha^\beta - \rhd_{\alpha a}{}^b C_b \wedge \beta^a + \rhd_{\alpha B}{}^A D_A \wedge \gamma^B = 0, \tag{18}$$

$$\mathrm{d}C_a - \partial_a{}^\alpha B_\alpha + \rhd_{\alpha a}{}^b C_b \wedge \alpha^\alpha + 2X_{(ab)}{}^A D_A \wedge \beta^b = 0, \tag{19}$$

$$\mathrm{d}D_A - \rhd_{\alpha A}{}^B D_B \wedge \alpha^\alpha + \delta_A{}^a C_a = 0. \tag{20}$$

For further details see [21, 29, 30] for the definition of the three-group, and [13] for the definition of the pure 3BF action.

Choosing the convenient underlying two-crossed module structure and imposing the appropriate simplicity constraints onto the degrees of freedom present in the 3BF action, one can obtain the non-trivial classical dynamics of the gravitational and matter fields. A reader interested in the construction of the constrained 2BF actions describing the Yang–Mills field and Einstein–Cartan gravity, and 3BF actions describing the Klein–Gordon, Dirac, Weyl and Majorana fields coupled to gravity in the standard way, is referred to [13, 27]. One can also introduce higher dimensional, nBF actions, see for example [31]. Various properties of these models have been studied in [32–34]. Naturally, if one is interested in theories defined on a four-dimensional spacetime manifold, there is an upper limit on the order of the differential forms one can use to construct a n-connection, and in four dimensions that is $n = 3$.

3. Hamiltonian analysis of the 3BF theory

In this section, the canonical structure of the theory is presented, with the resulting first-class and second-class constraints present in the theory. The algebra of Poisson brackets between all, the first-class and the second-class constraints, is obtained. We will use this result to calculate the total number of degrees of freedom in the theory, and in order to do that, we will have to analyse the BI that the first-class constraints satisfy, which enforce restrictions in the sense of Hamiltonian analysis. They reduce the number of independent first-class constraints present in the theory, thus increasing the number of degrees of freedom. We will obtain that the pure 3BF theory is topological, i.e. there are no local propagating degrees of freedom. Finally, we will finish this section with the construction of the generator of gauge symmetries of the 3BF action, which is used to calculate the form-variations of all the variables and their canonical momenta. This result will be crucial for finding the gauge symmetries of 3BF action, which will be a topic of section 4.

3.1. Canonical structure and Hamiltonian

Assuming that the spacetime manifold $\mathcal{M}_4$ is globally hyperbolic, the Lagrangian on a spatial foliation Σ_3 of spacetime $\mathcal{M}_4$ corresponding to the 3BF action (15) is given as:

$$L_{3BF} = \int_{\Sigma_3} \mathrm{d}^3\vec{x}\, \epsilon^{\mu\nu\rho\sigma} \left(\frac{1}{4} B^\alpha{}_{\mu\nu}\, \mathcal{F}^\beta{}_{\rho\sigma}\, g_{\alpha\beta} + \frac{1}{3!} C^a{}_\mu\, \mathcal{G}^b{}_{\nu\rho\sigma}\, g_{ab} + \frac{1}{4!} D^A \mathcal{H}^B{}_{\mu\nu\rho\sigma} g_{AB} \right). \tag{21}$$

For the Lagrangian (21), the canonical momenta corresponding to all variables $B^{\alpha}{}_{\mu\nu}$, $\alpha^{\alpha}{}_{\mu}$, $C^{a}{}_{\mu}$, $\beta^{a}{}_{\mu\nu}$, D^{A}, and $\gamma^{A}{}_{\mu\nu\rho}$ are:

$$
\begin{aligned}
\pi(B)_{\alpha}{}^{\mu\nu} &= \frac{\delta L}{\delta \partial_0 B^{\alpha}{}_{\mu\nu}} = 0, \\
\pi(\alpha)_{\alpha}{}^{\mu} &= \frac{\delta L}{\delta \partial_0 \alpha^{\alpha}{}_{\mu}} = \frac{1}{2}\epsilon^{0\mu\nu\rho} B_{\alpha\nu\rho}, \\
\pi(C)_{a}{}^{\mu} &= \frac{\delta L}{\delta \partial_0 C^{a}{}_{\mu}} = 0, \\
\pi(\beta)_{a}{}^{\mu\nu} &= \frac{\delta L}{\delta \partial_0 \beta^{a}{}_{\mu\nu}} = -\epsilon^{0\mu\nu\rho} C_{a\rho}, \\
\pi(D)_{A} &= \frac{\delta L}{\delta \partial_0 D^{A}} = 0, \\
\pi(\gamma)_{A}{}^{\mu\nu\rho} &= \frac{\delta L}{\delta \partial_0 \gamma^{A}{}_{\mu\nu\rho}} = \epsilon^{0\mu\nu\rho} D_{A}.
\end{aligned}
\tag{22}
$$

These momenta give rise to the six primary constraints of the theory, since none of them can be inverted for the time derivatives of the variables,

$$
\begin{aligned}
P(B)_{\alpha}{}^{\mu\nu} &\equiv \pi(B)_{\alpha}{}^{\mu\nu} \approx 0, \\
P(\alpha)_{\alpha}{}^{\mu} &\equiv \pi(\alpha)_{\alpha}{}^{\mu} - \frac{1}{2}\epsilon^{0\mu\nu\rho} B_{\alpha\nu\rho} \approx 0, \\
P(C)_{a}{}^{\mu} &\equiv \pi(C)_{a}{}^{\mu} \approx 0, \\
P(\beta)_{a}{}^{\mu\nu} &\equiv \pi(\beta)_{a}{}^{\mu\nu} + \epsilon^{0\mu\nu\rho} C_{a\rho} \approx 0, \\
P(D)_{A} &\equiv \pi(D)_{A} \approx 0, \\
P(\gamma)_{A}{}^{\mu\nu\rho} &\equiv \pi(\gamma)_{A}{}^{\mu\nu\rho} - \epsilon^{0\mu\nu\rho} D_{A} \approx 0.
\end{aligned}
\tag{23}
$$

Employing the following fundamental Poisson brackets,

$$
\begin{aligned}
\{\, B^{\alpha}{}_{\mu\nu}(\vec{x})\,,\, \pi(B)_{\beta}{}^{\rho\sigma}(\vec{y})\,\} &= 2\delta^{\alpha}_{\beta}\delta^{\rho}_{[\mu|}\delta^{\sigma}_{|\nu]}\,\delta^{(3)}(\vec{x}-\vec{y}), \\
\{\, \alpha^{\alpha}{}_{\mu}(\vec{x})\,,\, \pi(\alpha)_{\beta}{}^{\nu}(\vec{y})\,\} &= \delta^{\alpha}_{\beta}\delta^{\nu}_{\mu}\,\delta^{(3)}(\vec{x}-\vec{y}), \\
\{\, C^{a}{}_{\mu}(\vec{x})\,,\, \pi(C)_{b}{}^{\nu}(\vec{y})\,\} &= \delta^{a}_{b}\delta^{\nu}_{\mu}\,\delta^{(3)}(\vec{x}-\vec{y}), \\
\{\, \beta^{a}{}_{\mu\nu}(\vec{x})\,,\, \pi(\beta)_{b}{}^{\rho\sigma}(\vec{y})\,\} &= 2\delta^{a}_{b}\,\delta^{\rho}_{[\mu|}\delta^{\sigma}_{|\nu]}\,\delta^{(3)}(\vec{x}-\vec{y}), \\
\{\, D^{A}(\vec{x})\,,\, \pi(D)_{B}(\vec{y})\,\} &= \delta^{A}_{B}\,\delta^{(3)}(\vec{x}-\vec{y}), \\
\{\, \gamma^{A}{}_{\mu\nu\rho}(\vec{x})\,,\, \pi(\gamma)_{B}{}^{\sigma\tau\xi}(\vec{y})\,\} &= 3!\delta^{A}_{B}\,\delta^{\sigma}_{[\mu}\delta^{\tau}_{\nu}\delta^{\xi}_{\rho]}\,\delta^{(3)}(\vec{x}-\vec{y}),
\end{aligned}
\tag{24}
$$

one obtains the *algebra of primary constraints*:

$$
\begin{aligned}
\{ P(B)_\alpha{}^{jk}(\vec{x}) \,,\, P(\alpha)_\beta{}^{i}(\vec{y}) \} &= \epsilon^{0ijk}\, g_{\alpha\beta}(\vec{x})\, \delta^{(3)}(\vec{x}-\vec{y}), \\
\{ P(C)_a{}^{k}(\vec{x}) \,,\, P(\beta)_b{}^{ij}(\vec{y}) \} &= -\epsilon^{0ijk}\, g_{ab}(\vec{x})\, \delta^{(3)}(\vec{x}-\vec{y}), \\
\{ P(D)_A(\vec{x}) \,,\, P(\gamma)_B{}^{ijk}(\vec{y}) \} &= \epsilon^{0ijk}\, g_{AB}(\vec{x})\, \delta^{(3)}(\vec{x}-\vec{y}).
\end{aligned}
\tag{25}
$$

Note that all other Poisson brackets vanish. The *canonical, on-shell Hamiltonian* is given by the following expression:

$$
\begin{aligned}
H_{\mathrm{c}} = \int_{\Sigma_3} \mathrm{d}^3\vec{x} \left[\frac{1}{2}\pi(B)_\alpha{}^{\mu\nu}\, \partial_0 B^\alpha{}_{\mu\nu} + \pi(\alpha)_\alpha{}^{\mu}\, \partial_0 \alpha^\alpha{}_\mu + \pi(C)_a{}^{\mu}\, \partial_0 C^a{}_\mu \right. \\
\left. + \frac{1}{2}\pi(\beta)_a{}^{\mu\nu}\, \partial_0 \beta^a{}_{\mu\nu} + \pi(D)_A\, \partial_0 D^A + \frac{1}{3!}\pi(\gamma)_A{}^{\mu\nu\rho}\, \partial_0 \gamma^A{}_{\mu\nu\rho} \right] - L.
\end{aligned}
\tag{26}
$$

Employing the definition of the curvature components (14), the Hamiltonian (26) can be written as the sum of terms that are equal to the product of the primary constraints and time derivatives of the variables, and the remainder. As the primary constraints are zero on-shell, the terms multiplying the time derivatives vanish, and the canonical Hamiltonian becomes:

$$
\begin{aligned}
H_{\mathrm{c}} = -\int_{\Sigma_3} \mathrm{d}^3\vec{x}\, \epsilon^{0ijk} \left[\frac{1}{2} B_{\alpha 0i}\, \mathcal{F}^\alpha{}_{jk} + \frac{1}{6} C_{a0}\, \mathcal{G}^a{}_{ijk} + \beta^a{}_{0i} \left(\nabla_j C_{ak} - \frac{1}{2}\partial_a{}^\alpha B_{\alpha\, jk} + \beta^b{}_{jk}\, D_A\, X_{(ab)}{}^A \right) \right. \\
\left. + \frac{1}{2}\alpha^\alpha{}_0 \left(\nabla_i B_{\alpha jk} - C_{ai} \rhd_{\alpha b}{}^a\, \beta^b{}_{jk} + \frac{1}{3} D_A \rhd_{\alpha B}{}^A\, \gamma^B{}_{ijk} \right) + \frac{1}{2}\gamma^A{}_{0ij} \left(\nabla_k D_A + C_{ak}\, \delta_A{}^a \right) \right].
\end{aligned}
\tag{27}
$$

Adding to the canonical Hamiltonian the product of the Lagrange multipliers λ and the primary constraints, for every primary constraint, one gets the *total, off-shell Hamiltonian*:

$$
\begin{aligned}
H_{\mathrm{T}} = H_{\mathrm{c}} + \int_{\Sigma_3} \mathrm{d}^3\vec{x} \left[\frac{1}{2}\lambda(B)^\alpha{}_{\mu\nu} P(B)_\alpha{}^{\mu\nu} + \lambda(\alpha)^\alpha{}_\mu P(\alpha)_\alpha{}^\mu + \lambda(C)^a{}_\mu P(C)_a{}^\mu + \frac{1}{2}\lambda(\beta)^a{}_{\mu\nu} P(\beta)_a{}^{\mu\nu} \right. \\
\left. + \lambda(D)^A P(D)_A + \frac{1}{3!}\lambda(\gamma)^A{}_{\mu\nu\rho} P(\gamma)_A{}^{\mu\nu\rho} \right].
\end{aligned}
\tag{28}
$$

3.2. Consistency conditions and algebra of constraints

In order for primary constraints to be preserved during the evolution of the system, they must satisfy the consistency conditions,

$$
\dot{P} \equiv \{ P \,,\, H_{\mathrm{T}} \} \approx 0,
\tag{29}
$$

for every primary constraint P. Imposing this condition on primary constraints $P(B)_{\alpha}{}^{0i}$, $P(\alpha)_{\alpha}{}^{0}$, $P(C)_{a}{}^{0}$, $P(\beta)_{a}{}^{0i}$, and $P(\gamma)_{A}{}^{0ij}$, one obtains the secondary constraints $\mathcal{S}$,

$$
\begin{aligned}
\mathcal{S}(\mathcal{F})_{\alpha}{}^{i} &\equiv \frac{1}{2}\epsilon^{0ijk}\mathcal{F}_{\alpha jk} \approx 0,\\
\mathcal{S}(\nabla B)_{\alpha} &\equiv \frac{1}{2}\epsilon^{0ijk}\left(\nabla_{[i}B_{\alpha\, jk]} - C_{a[i}\,\triangleright_{\alpha b}{}^{a}\,\beta^{b}{}_{jk]} + \frac{1}{3}D_{A}\,\triangleright_{\alpha B}{}^{A}\,\gamma^{B}{}_{ijk}\right) \approx 0,\\
\mathcal{S}(\mathcal{G})_{a} &\equiv \frac{1}{6}\epsilon^{0ijk}\mathcal{G}_{aijk} \approx 0,\\
\mathcal{S}(\nabla C)_{a}{}^{i} &\equiv \epsilon^{0ijk}\left(\nabla_{[j|}C_{a|k]} - \frac{1}{2}\partial_{a}{}^{\alpha}B_{\alpha\, jk} + \beta^{b}{}_{jk}\,D_{A}\,X_{(ab)}{}^{A}\right) \approx 0,\\
\mathcal{S}(\nabla D)_{A}{}^{ij} &\equiv \epsilon^{0ijk}\left(\nabla_{k}D_{A} + C_{ak}\,\delta_{A}{}^{a}\right) \approx 0,
\end{aligned}
\tag{30}
$$

while in the case of the constraints $P(\alpha)_{\alpha}{}^{k}$, $P(B)_{\alpha}{}^{jk}$, $P(\beta)_{a}{}^{jk}$, $P(C)_{a}{}^{k}$, $P(\gamma)_{A}{}^{ijk}$, and $P(D)_{A}$ the corresponding consistency conditions determine the following Lagrange multipliers:

$$
\begin{aligned}
\lambda(B)_{\alpha ij} &\approx \nabla_{i}B_{\alpha 0j} - \nabla_{j}B_{\alpha 0i} + C_{a0}\beta^{b}{}_{ij}\triangleright_{\alpha b}{}^{a} + C_{bi}\triangleright_{\alpha}{}^{b}{}_{a}\beta^{a}{}_{0j}\\
&\quad - C_{bj}\triangleright_{\alpha}{}^{b}{}_{a}\beta^{a}{}_{0i} + g_{\beta\gamma}{}^{\alpha}\alpha^{\beta}{}_{0}B^{\gamma}{}_{ij} + D_{B}\gamma^{A}{}_{0ij}\triangleright_{\alpha}{}^{B}{}_{A},\\
\lambda(\alpha)^{\alpha}{}_{i} &\approx \nabla_{i}\alpha^{\alpha}{}_{0} + \partial_{a}{}^{\alpha}\beta^{a}{}_{0i},\\
\lambda(C)^{a}{}_{i} &\approx \nabla_{i}C^{a}{}_{0} + C^{b}{}_{i}\triangleright_{\alpha}{}^{a}{}_{b}\alpha^{\alpha}{}_{0} - 2\beta_{b0i}D_{A}X^{(ba)A} + B_{\alpha 0i}\partial^{a\alpha},\\
\lambda(\beta)^{a}{}_{ij} &\approx \nabla_{i}\beta^{a}{}_{0j} - \nabla_{j}\beta^{a}{}_{0i} - \beta^{b}{}_{ij}\triangleright_{\alpha b}{}^{a}\alpha^{\alpha}{}_{0} + \gamma^{A}{}_{0ij}\delta_{A}{}^{a},\\
\lambda(D)_{A} &\approx \alpha^{\alpha}{}_{0}D_{B}\triangleright_{\alpha A}{}^{B} - C_{a0}\delta_{A}{}^{a},\\
\lambda(\gamma)^{A}{}_{ijk} &\approx -2\beta^{a}{}_{0i}\beta^{b}{}_{jk}X_{(ab)}{}^{A} + 2\beta^{a}{}_{0j}\beta^{b}{}_{ik}X_{(ab)}{}^{A} - 2\beta^{a}{}_{0k}\beta^{b}{}_{ij}X_{(ab)}{}^{A}\\
&\quad - \alpha^{\alpha}{}_{0}\,\triangleright_{\alpha B}{}^{A}\gamma^{B}{}_{ijk} + \nabla_{i}\gamma^{A}{}_{0jk} - \nabla_{j}\gamma^{A}{}_{0ik} + \nabla_{k}\gamma^{A}{}_{0ij}.
\end{aligned}
\tag{31}
$$

Note that the rest of the Lagrange multipliers

$$
\lambda(B)^{\alpha}{}_{0i}, \qquad \lambda(\alpha)^{\alpha}{}_{0}, \qquad \lambda(C)^{a}{}_{0}, \qquad \lambda(\beta)^{a}{}_{0i}, \qquad \lambda(\gamma)^{A}{}_{0ij}, \tag{32}
$$

remain undetermined.

Further, as the secondary constraints must also be preserved during the evolution of the system, the consistency conditions of secondary constraints must be enforced. However, no tertiary constraints arise from these conditions (see equation (B.1) in appendix B), leading the iterative procedure to an end. Finally, the total Hamiltonian can be written in the following form:

$$
\begin{aligned}
H_{\mathrm{T}} = \int_{\Sigma_3} \mathrm{d}^3\vec{x} \Bigg[& \lambda(B)^{\alpha}{}_{0i}\,\Phi(B)_{\alpha}{}^{i} + \lambda(\alpha)^{\alpha}\,\Phi(\alpha)_{\alpha} + \lambda(C)^{a}{}_{0}\,\Phi(C)_{a} + \lambda(\beta)^{a}{}_{0i}\,\Phi(\beta)_{a}{}^{i} \\
& + \frac{1}{2}\lambda(\gamma)^{A}{}_{0ij}\Phi(\gamma)_{A}{}^{ij} - B_{\alpha 0i}\,\Phi(\mathcal{F})^{ai} - \alpha_{\alpha 0}\,\Phi(\nabla B)^{\alpha} - C_{a0}\,\Phi(\mathcal{G})^{a} \\
& - \beta_{a0i}\,\Phi(\nabla C)^{ai} - \frac{1}{2}\gamma_{A0ij}\,\Phi(\nabla D)^{Aij} \Bigg],
\end{aligned}
\tag{33}
$$

where

$$
\begin{aligned}
\Phi(B)_{\alpha}{}^{i} &= P(B)_{\alpha}{}^{0i}, \\
\Phi(\alpha)_{\alpha} &= P(\alpha)_{\alpha}{}^{0}, \\
\Phi(C)_{a} &= P(C)_{a}{}^{0}, \\
\Phi(\beta)_{a}{}^{i} &= P(\beta)_{a}{}^{0i}, \\
\Phi(\gamma)_{A}{}^{ij} &= P(\gamma)_{A}{}^{0ij}, \\
\Phi(\mathcal{F})^{\alpha i} &= \mathcal{S}(\mathcal{F})^{\alpha i} - \nabla_{j}P(B)^{\alpha ij} - P(C)_{a}{}^{i}\partial^{a\alpha}, \\
\Phi(\mathcal{G})_{a} &= \mathcal{S}(\mathcal{G})_{a} + \nabla_{i}P(C)_{a}{}^{i} - \frac{1}{2}\beta_{bij}\rhd_{\alpha}{}^{b}{}_{a}P(B)^{\alpha ij} + P(D)^{A}\delta_{Aa}, \\
\Phi(\nabla C)_{a}{}^{i} &= \mathcal{S}(\nabla C)_{a}{}^{i} - \nabla_{j}P(\beta)_{a}{}^{ij} + C_{bj}\rhd_{\alpha}{}^{b}{}_{a}P(B)^{\alpha ij} \\
&\quad - \partial_{a}{}^{\alpha}P(\alpha)_{\alpha}{}^{i} + 2D_{A}X_{(ab)}{}^{A}P(C)^{bi} + \beta^{b}{}_{jk}X_{(ab)}{}^{A}P(\gamma)_{A}{}^{ijk}, \\
\Phi(\nabla B)_{\alpha} &= \mathcal{S}(\nabla B)_{\alpha} + \nabla_{i}P(\alpha)_{\alpha}{}^{i} - \frac{1}{2}f_{\alpha\gamma}{}^{\beta}B_{\beta ij}P(B)^{\gamma ij} - C_{bi}\rhd_{\alpha a}{}^{b}P(C)^{ai} \\
&\quad - \frac{1}{2}\beta_{bij}\rhd_{\alpha a}{}^{b}P(\beta)^{aij} - P(D)^{A}D_{B}\rhd_{\alpha A}{}^{B} + \frac{1}{3!}P(\gamma)_{A}{}^{ijk}\gamma^{B}{}_{ijk}\rhd_{\alpha B}{}^{A}, \\
\Phi(\nabla D)_{A}{}^{ij} &= \mathcal{S}(\nabla D)_{A}{}^{ij} + \nabla_{k}P(\gamma)_{A}{}^{ijk} - P(\beta)_{a}{}^{ij}\delta_{A}{}^{a} - P(B)^{\alpha ij}\rhd_{\alpha}{}^{B}{}_{A}D_{B},
\end{aligned}
\tag{34}
$$

are the first-class constraints. The second-class constraints in the theory are:

$$
\begin{aligned}
&\chi(B)_{\alpha}{}^{jk} = P(B)_{\alpha}{}^{jk}, \quad \chi(C)_{a}{}^{i} = P(C)_{a}{}^{i}, \quad \chi(D)_{A} = P(D)_{A}, \\
&\chi(\alpha)_{\alpha}{}^{i} = P(\alpha)_{\alpha}{}^{i}, \quad \chi(\beta)_{a}{}^{ij} = P(\beta)_{a}{}^{ij}, \quad \chi(\gamma)_{A}{}^{ijk} = P(\gamma)_{A}{}^{ijk}.
\end{aligned}
\tag{35}
$$

The PB algebra of the first-class constraints is given by

$$
\begin{aligned}
&\{\, \Phi(\mathcal{F})^{\alpha i}(\vec{x})\,,\, \Phi(\nabla B)_{\beta}(\vec{y})\,\} = f_{\beta\gamma}{}^{\alpha}\, \Phi(\mathcal{F})^{\gamma i}(\vec{x})\, \delta^{(3)}(\vec{x}-\vec{y}),\\
&\{\, \Phi(\nabla B)_{\alpha}(\vec{x})\,,\, \Phi(\nabla B)_{\beta}(\vec{y})\,\} = f_{\alpha\beta}{}^{\gamma}\, \Phi(\nabla B)_{\gamma}(\vec{x})\, \delta^{(3)}(\vec{x}-\vec{y}),\\
&\{\, \Phi(\mathcal{G})^{a}(\vec{x})\,,\, \Phi(\nabla C)_{b}{}^{i}(\vec{y})\,\} = -\rhd_{\alpha b}{}^{a}\, \Phi(\mathcal{F})^{\alpha i}(\vec{x})\, \delta^{(3)}(\vec{x}-\vec{y}),\\
&\{\, \Phi(\nabla C)_{a}{}^{i}(\vec{x})\,,\, \Phi(\nabla C)_{b}{}^{j}(\vec{y})\,\} = -2X_{(ab)}{}^{A}\, \Phi(\nabla D)_{A}{}^{ij}(\vec{x})\, \delta^{(3)}(\vec{x}-\vec{y}),\\
&\{\, \Phi(\mathcal{G})^{a}(\vec{x})\,,\, \Phi(\nabla B)_{\alpha}(\vec{y})\,\} = \rhd_{\alpha b}{}^{a}\, \Phi(\mathcal{G})^{b}(\vec{x})\, \delta^{(3)}(\vec{x}-\vec{y}),\\
&\{\, \Phi(\nabla C)^{ai}(\vec{x})\,,\, \Phi(\nabla B)_{\alpha}(\vec{y})\,\} = \rhd_{\alpha b}{}^{a}\, \Phi(\nabla C)^{bi}(\vec{x})\, \delta^{(3)}(\vec{x}-\vec{y}),\\
&\{\, \Phi(\nabla B)_{\alpha}(\vec{x})\,,\, \Phi(\nabla D)_{A}{}^{ij}(\vec{y})\,\} = \rhd_{\alpha A}{}^{B} \Phi(\nabla D)_{B}{}^{ij}(\vec{x}) \delta^{(3)}(\vec{x}-\vec{y}).
\end{aligned}
\tag{36}
$$

The algebra between the first and the second class constraints is given in the appendix B, equation (B.2).

With the algebra of the constraints in hand, one can proceed to calculate the generator of gauge symmetries of the action. The generator will be used to calculate the form-variations of all the variables and their canonical momenta, which will help us find the finite gauge symmetries of the action. Additionally, we can determine the number of independent parameters of gauge transformations, since usually all the first class constraints generate unphysical transformations of dynamical variables, i.e. that to each parameter of the gauge symmetry there corresponds one first-class constraint. However, before we embark on the construction of the symmetry generator, we will devote some attention to the number of local propagating degrees of freedom in the theory, in order to determine if the $3BF$ action is topological or not.

3.3. Number of degrees of freedom

In this subsection, we will show that the structure of the constraints implies that there are no local degrees of freedom in a $3BF$ theory. To that end, let us first specify all the BI present in the theory.

The two-form curvatures corresponding to one-forms α and C, given by

$$
F^{\alpha} = \mathrm{d}\alpha^{\alpha} + f_{\beta\gamma}{}^{\alpha}\, \alpha^{\beta} \wedge \alpha^{\gamma}, \qquad T^{a} = \mathrm{d}C^{a} + \rhd_{\alpha b}{}^{a}\, \alpha^{\alpha} \wedge C^{b}, \tag{37}
$$

satisfy the BI:

$$
\epsilon^{\lambda\mu\nu\rho}\, \nabla_{\mu} F^{\alpha}{}_{\nu\rho} = 0, \tag{38}
$$

$$
\epsilon^{\lambda\mu\nu\rho} \left(\nabla_{\mu} T^{a}{}_{\nu\rho} - \rhd_{\alpha b}{}^{a} F^{\alpha}{}_{\mu\nu} C^{b}{}_{\rho} \right) = 0. \tag{39}
$$

Similarly, the three-form curvatures corresponding to two-forms B and β, given by

$$
S^{\alpha} = \mathrm{d}B^{\alpha} + f_{\beta\gamma}{}^{\alpha}\, \alpha^{\beta} \wedge B^{\gamma}, \qquad G^{a} = \mathrm{d}\beta^{a} + \rhd_{\alpha b}{}^{a}\, \alpha^{\alpha} \wedge \beta^{\beta}, \tag{40}
$$

Table 1. The fields present in the $3BF$ theory.

$\alpha^{\alpha}{}_{\mu}$	$\beta^{a}{}_{\mu\nu}$	$\gamma^{A}{}_{\mu\nu\rho}$	$B^{\alpha}{}_{\mu\nu}$	$C^{a}{}_{\mu}$	D^{A}
$4p$	$6q$	$4r$	$6p$	$4q$	r

Table 2. Second-class constraints in the $3BF$ theory.

$\chi(B)_{\alpha}{}^{jk}$	$\chi(C)_{a}{}^{i}$	$\chi(D)_{A}$	$\chi(\alpha)_{\alpha}{}^{i}$	$\chi(\beta)_{a}{}^{ij}$	$\chi(\gamma)_{A}{}^{ijk}$
$3p$	$3q$	r	$3p$	$3q$	r

satisfy the BI:

$$\epsilon^{\lambda\mu\nu\rho}\left(\frac{2}{3}\nabla_{\lambda}S^{\alpha}{}_{\mu\nu\rho}-f_{\beta\gamma}{}^{\alpha}F^{\beta}{}_{\lambda\mu}B^{\gamma}{}_{\nu\rho}\right)=0, \tag{41}$$

$$\epsilon^{\lambda\mu\nu\rho}\left(\frac{2}{3}\nabla_{\lambda}G^{a}{}_{\mu\nu\rho}-\triangleright_{\alpha b}{}^{a}F^{\alpha}{}_{\lambda\mu}\beta^{b}{}_{\nu\rho}\right)=0. \tag{42}$$

Finally, defining the one-form curvature for D,

$$Q^{A}=\mathrm{d}D^{A}+\triangleright_{\alpha B}{}^{A}\alpha^{\alpha}\wedge D^{B}, \tag{43}$$

one can write the corresponding BI for Q^{A}:

$$\epsilon^{\lambda\mu\nu\rho}\left(\nabla_{\nu}Q^{A}{}_{\rho}-\frac{1}{2}\triangleright_{\alpha B}{}^{A}F^{\alpha}{}_{\nu\rho}D^{B}\right)=0. \tag{44}$$

These BI play an important role in determining the number of degrees of freedom present in the theory.

As the general theory states, if there are N fields in the theory, F independent first-class constraints per space point, and S independent second-class constraints per space point, the number of independent field components, i.e. the number of the physical degrees of freedom present in the theory, is given by:

$$n=N-F-\frac{S}{2}. \tag{45}$$

Let p denote the dimensionality of the group G, q the dimensionality of the group H, and r the dimensionality of the group L. Determining the number of fields present in the $3BF$ theory, by counting the field components listed in table 1, one obtains $N=10(p+q)+5r$. Similarly, one determines the number of independent components of the second-class constraints by counting the components listed in table 2 and obtains $S=6(p+q)+2r$. However, when counting the number of the first-class constraints F one notes they are not all mutually independent. Namely, one can prove the following identities, as a consequence of the BI.

Taking the derivative of $\Phi(\mathcal{F})_{\alpha}{}^{i}$ one obtains

$$\nabla_{i}\Phi(\mathcal{F})^{\alpha i}+\partial_{a}{}^{\alpha}\Phi(\mathcal{G})^{a}=\frac{1}{2}\epsilon^{0ijk}\nabla_{i}F^{\alpha}{}_{jk}-\frac{1}{2}f_{\beta\gamma}{}^{\alpha}\mathcal{F}^{\beta}{}_{ij}P(B)^{ij}. \tag{46}$$

This relation gives

$$\nabla_{i}\Phi(\mathcal{F})^{\alpha i}+\partial_{a}{}^{\alpha}\Phi(\mathcal{G})^{a}=0, \tag{47}$$

since the first term on the right-hand side of (46) is zero off-shell because $\epsilon^{ijk}\nabla_i F^a{}_{jk} = 0$ are the $\lambda = 0$ components of BI (38), and the second term on the right-hand side is also zero off-shell, since it is a product of two constraints:

$$\frac{1}{2} f_{\beta\gamma}{}^{\alpha}\mathcal{F}^{\beta}{}_{ij}P(B)^{ij} = \frac{1}{2} f_{\beta\gamma}{}^{\alpha}\epsilon_{0ijk}\mathcal{S}(\mathcal{F})^{\beta k}P(B)^{ij} = 0. \tag{48}$$

The relation (47) means that p components of the first-class constraints $\Phi(\mathcal{F})^{\alpha i}$ and $\Phi(\mathcal{G})^a$ are not independent of the others. Furthermore, taking the derivative of $\Phi(\nabla C)_a{}^i$ one obtains

$$\begin{aligned}
&\nabla_i\Phi(\nabla C)_a{}^i + C_{bi}\rhd_{\alpha a}{}^b\Phi(\mathcal{F})^{\alpha i} + \partial_a{}^{\alpha}\Phi(\nabla B)_{\alpha} - \beta^b{}_{ij}X_{(ab)}{}^A\Phi(\nabla D)_A{}^{ij} - 2D_A X_{(ab)}{}^A\,\Phi(\mathcal{G})^b \\
&\quad = \frac{1}{2}\epsilon^{0ijk}\left(\nabla_i T_{ajk} - \rhd_{\alpha b}{}^a F^{\alpha}{}_{jk}C^b{}_i\right) - \frac{1}{2}\epsilon^{0ijk}\rhd_{\alpha a}{}^b\,P(B)^{\alpha}{}_{ij}\,S(\nabla C)_{bk} \\
&\qquad + \epsilon^{0ijk}X_{(ab)}{}^A\,P(C)^b{}_i\,S(\nabla D)_{A\,jk} + \frac{1}{3}\epsilon^{0ijk}X_{(ab)}{}^A\,P(\gamma)^A{}_{ijk}\,S(\mathcal{G})^b + \frac{1}{2}\epsilon^{0ijk}\rhd_{\alpha a}{}^b\,P(\beta)^b{}_{ij}\,S(\mathcal{F})^{\alpha}{}_k.
\end{aligned} \tag{49}$$

Noting that the right-hand side of (49) is zero off-shell as the $\lambda = 0$ components of the BI (39), and the remaining terms on the right-hand side are zero off-shell as products of two constraints, one obtains the following relation:

$$\nabla_i\Phi(\nabla C)_a{}^i + C_{bi}\rhd_{\alpha a}{}^b\Phi(\mathcal{F})^{\alpha i} + \partial_a{}^{\alpha}\Phi(\nabla B)_{\alpha} - \beta^b{}_{ij}X_{(ab)}{}^A\Phi(\nabla D)_A{}^{ij} - 2D_A X_{(ab)}{}^A\Phi(\beta)^b = 0. \tag{50}$$

This relation means that q components of the constraints $\Phi(\nabla C)_a{}^i$, $\Phi(\mathcal{F})^{\alpha i}$, $\Phi(\nabla B)_{\alpha}$, $\Phi(\nabla D)_A{}^{ij}$, and $\Phi(\beta)^b$, are not independent of the others, further lowering the number of the independent first-class constraints. Finally, the following relation is satisfied

$$\begin{aligned}
&\nabla^j\Phi(\nabla D)^A{}_{ij} - \rhd_{\alpha B}{}^A D^B\Phi(\mathcal{F})_{\alpha i} - \delta^A{}_a\Phi(\nabla C)^a{}_i \\
&\quad = \epsilon_{0ijk}\left(\nabla^j Q_A{}^k + \frac{1}{2}\rhd_{\alpha A}{}^B F^{\alpha jk}D_B\right) + \frac{1}{2}\epsilon^{0jkl}\rhd_{\alpha B}{}^A\,P(\gamma)^B{}_{ijk}\,S(\mathcal{F})_{\alpha l} \\
&\qquad - \frac{1}{2}\epsilon^{0jkl}\rhd_{\alpha B}{}^A\,P(B)^{\alpha}{}_{ij}\,S(\nabla D)^B{}_{kl}.
\end{aligned} \tag{51}$$

Since the first term on the right-hand side is precisely the $\lambda = 0$ component of the BI (44), while the second and third terms are equal to zero as products of two constraints, this gives:

$$\nabla^j\Phi(\nabla D)^A{}_{ij} - \rhd_{\alpha B}{}^A D^B\Phi(\mathcal{F})_{\alpha i} - \delta^A{}_a\Phi(\nabla C)^a{}_i = 0. \tag{52}$$

This relation suggests that $3r$ components of the primary constraints $\Phi(\nabla D)^A{}_{ij}$, $\Phi(\mathcal{F})_{\alpha i}$, and $\Phi(C)^a{}_i$ are not independent of the others. However, this is slightly misleading, since the covariant derivative of the BI (44) is automatically satisfied as a consequence of the BI (38),

$$\epsilon^{\lambda\mu\nu\rho}D^B\rhd_{\alpha B}{}^A\nabla_{\mu}F^{\alpha}{}_{\nu\rho} = 0, \tag{53}$$

Table 3. First-class constraints in the 3*BF* theory.

$\Phi(B)_\alpha{}^i$	$\Phi(C)_a$	$\Phi(\alpha)_\alpha$	$\Phi(\beta)_a{}^i$	$\Phi(\gamma)_A{}^{ij}$	$\Phi(\mathcal{F})^{\alpha i}$	$\Phi(\mathcal{G})^a$	$\Phi(\nabla C)^{ai}$	$\Phi(\nabla B)^\alpha$	$\Phi(\nabla D)_A{}^{ij}$
$3p$	q	p	$3q$	$3r$	$3p-p$	q	$3q-q$	p	$3r-2r$

which means that there are in fact only $2r$ components of the constraint (52). A formal proof of this statement would involve evaluating the Wronskian of all first-class constraints, and is out of the scope of this paper.

The number of independent components of first-class constraints is determined by counting the components listed in table 3, and then subtracting the number of independent relations (47), (50) and (52).

Bearing the previous analysis in mind, one obtains the number of independent first-class constraints:

$$F = 8(p+q) + 6r - p - q - 2r = 7(p+q) + 4r.$$

Finally, using the definition (45), the number of degrees of freedom in the 3*BF* theory is:

$$n = 10(p+q) + 5r - 7(p+q) - 4r - \frac{6(p+q)+2r}{2} = 0. \tag{54}$$

Therefore, there are no local propagating degrees of freedom in a 3*BF* theory.

3.4. Symmetry generator

The unphysical transformations of dynamical variables are often referred to as gauge transformations. The gauge transformations are *local*, meaning that the parameters of the transformations are arbitrary functions of space and time. We shall now construct the generator of all gauge symmetries of the theory governed by the total Hamiltonian (33), using the Castellani's algorithm (see chapter 5 in [15] for a comprehensive overview of the procedure). The details of the construction are given in appendix C, and the following result is obtained

$$\begin{aligned} G = \int_{\Sigma_3} \mathrm{d}^3\vec{x} \Big(& (\nabla_0 \epsilon_{\mathfrak{g}}{}^\alpha)(\tilde{G}_1)_\alpha + \epsilon_{\mathfrak{g}}{}^\alpha(\tilde{G}_0)_\alpha + (\nabla_0 \epsilon_{\mathfrak{h}}{}^a{}_i)(\tilde{H}_1)_a{}^i + \epsilon_{\mathfrak{h}}{}^a{}_i(\tilde{H}_0)_a{}^i + \frac{1}{2}(\nabla_0 \epsilon_{\mathfrak{l}}{}^A{}_{ij})(\tilde{L}_1)_A{}^{ij} \\ & + \frac{1}{2}\epsilon_{\mathfrak{l}}{}^A{}_{ij}(\tilde{L}_0)_A{}^{ij} + (\nabla_0 \epsilon_{\mathfrak{m}}{}^\alpha{}_i)(\tilde{M}_1)_\alpha{}^i + \epsilon_{\mathfrak{m}}{}^\alpha{}_i(\tilde{M}_0)_\alpha{}^i + (\nabla_0 \epsilon_{\mathfrak{n}}{}^a)(\tilde{N}_1)_a + \epsilon_{\mathfrak{n}}{}^a(\tilde{N}_0)_a \Big), \end{aligned} \tag{55}$$

where

$$
\begin{aligned}
(\tilde{G}_1)_\alpha &= -\Phi(\alpha)_\alpha, \\
(\tilde{G}_0)_\alpha &= -\left(f_{\alpha\gamma}{}^{\beta} B_{\beta 0i}\Phi(B)^{\gamma i} + C_{a0} \rhd_{\alpha b}{}^{a}\Phi(C)^{b0} + \beta_{a0i} \rhd_{\alpha b}{}^{a}\Phi(\beta)^{b0i} \right. \\
&\qquad \left. - \frac{1}{2}\gamma^{A}{}_{0ij} \rhd_{\alpha A}{}^{B}\Phi(\gamma)_B{}^{ij} - \Phi(\nabla B)_\alpha \right), \\
(\tilde{H}_1)_a{}^{i} &= -\Phi(\beta)_a{}^{i}, \\
(\tilde{H}_0)_a{}^{i} &= C_{b0} \rhd_{\alpha a}{}^{b}\Phi(B)^{\alpha i} - 2\beta^{b}{}_{0j} X_{(ab)}{}^{A}\Phi(\gamma)_A{}^{ij} + \Phi(\nabla C)_a{}^{i}, \\
(\tilde{L}_1)_a{}^{ij} &= \Phi(\gamma)_A{}^{ij}, \\
(\tilde{L}_0)_a{}^{ij} &= -\Phi(\nabla D)_A{}^{ij}, \\
(\tilde{M}_1)_\alpha{}^{i} &= -\Phi(B)_\alpha{}^{i}, \\
(\tilde{M}_0)_\alpha{}^{i} &= \Phi(\mathcal{F})_\alpha{}^{i}, \\
(\tilde{N}_1)_a &= -\Phi(C)_a, \\
(\tilde{N}_0)_a &= \beta_{b0i} \rhd_{\alpha a}{}^{b}\Phi(B)^{\alpha i} + \Phi(\mathcal{G})_a,
\end{aligned}
\tag{56}
$$

and $\epsilon_{\mathfrak{g}}{}^{\alpha}$, $\epsilon_{\mathfrak{h}}{}^{a}{}_{i}$, $\epsilon_{\mathfrak{l}}^{A}{}_{ij}$, $\epsilon_{\mathfrak{m}}{}^{\alpha}{}_{i}$, and $\epsilon_{\mathfrak{n}}{}^{a}$ are the independent parameters of the gauge transformations.

The obtained gauge generator (55) is then employed to calculate the form variations of variables and their corresponding canonical momenta, denoted as $A(t, \vec{x})$, using the following equation,

$$
\delta_0 A(t, \vec{x}) = \{A(t, \vec{x}), G\}. \tag{57}
$$

The form variations of all fields and canonical momenta are given in appendix E, equation (E.2), while the algebra of the generators (56) is obtained in the appendix B, equations (B.4)–(B.10). However, one must bear in mind that the gauge generator (55) is the generator of the symmetry transformations on a slice of spacetime, i.e. on a hypersurface Σ_3. Having in hand all these results, specifically the form variations of all variables and their canonical momenta (E.2), we can determine the full gauge symmetry of the theory, which will be done in the next section.

4. Symmetries of the 3*BF* action

In order to systematically describe all gauge transformations of the 3*BF* action, we will discuss in turn each set of gauge parameters $\epsilon_{\mathfrak{g}}{}^{\alpha}$, $\epsilon_{\mathfrak{h}}{}^{a}{}_{i}$, $\epsilon_{\mathfrak{l}}^{A}{}_{ij}$, $\epsilon_{\mathfrak{m}}{}^{\alpha}{}_{i}$, and $\epsilon_{\mathfrak{n}}{}^{a}$, appearing in (55). The subsection 4.1 deals with the gauge group G, and the G-gauge transformations, which are already familiar from the ordinary *BF* theory. In subsection 4.2 we discuss the gauge group $\tilde{H}_L$ which consists of the *H*-gauge and *L*-gauge transformations, familiar from the previous literature [21], while the subsection 4.3 examines the *M*-gauge and *N*-gauge transformations which are also present in the theory. Finally, the results of the subsections 4.1–4.3 will be summarized in the subsection 4.4, where we will present the complete structure of the gauge symmetry group.

4.1. Gauge group G

First, consider the infinitesimal transformation with the parameter $\epsilon_{\mathfrak{g}}{}^{\alpha}$, given by the form variations

$$
\begin{aligned}
&\delta_0\alpha^{\alpha}{}_{\mu} = -\partial_\mu\epsilon_{\mathfrak{g}}{}^{\alpha} - f_{\beta\gamma}{}^{\alpha}\alpha^{\beta}{}_{\mu}\epsilon_{\mathfrak{g}}{}^{\gamma}, \quad \delta_0 B^{\alpha}{}_{\mu\nu} = f_{\beta\gamma}{}^{\alpha}\epsilon_{\mathfrak{g}}{}^{\beta}B^{\gamma}{}_{\mu\nu},\\
&\delta_0\beta^{a}{}_{\mu\nu} = \rhd_{\alpha b}{}^{a}\epsilon_{\mathfrak{g}}{}^{\alpha}\beta^{b}{}_{\mu\nu}, \quad \delta_0 C^{a}{}_{\mu} = \rhd_{\alpha b}{}^{a}\epsilon_{\mathfrak{g}}{}^{\alpha}C^{b}{}_{\mu},\\
&\delta_0\gamma^{A}{}_{\mu\nu\rho} = \rhd_{\alpha B}{}^{A}\epsilon_{\mathfrak{g}}{}^{\alpha}\gamma^{B}{}_{\mu\nu\rho}, \quad \delta_0 D^{A} = \rhd_{\alpha B}{}^{A}\epsilon_{\mathfrak{g}}{}^{\alpha}D^{B},
\end{aligned}
\tag{58}
$$

which is analogous to writing the transformation as:

$$
\begin{aligned}
&\alpha \to \alpha' = \alpha - \nabla\epsilon_{\mathfrak{g}}, \quad B \to B' = B - [B, \epsilon_{\mathfrak{g}}],\\
&\beta \to \beta' = \beta + \epsilon_{\mathfrak{g}} \rhd \beta, \quad C \to C' = C + \epsilon_{\mathfrak{g}} \rhd C,\\
&\gamma \to \gamma' = \gamma + \epsilon_{\mathfrak{g}} \rhd \gamma, \quad D \to D' = D + \epsilon_{\mathfrak{g}} \rhd D.
\end{aligned}
\tag{59}
$$

Based on these infinitesimal transformations, one can extrapolate the finite symmetry transformations, defined in the theorem 1.

Theorem 1 (G-gauge transformations). *In the* 3*BF theory for the two-crossed module* $(L \xrightarrow{\delta} H \xrightarrow{\partial} G, \rhd, \{\_,\_\}_{\mathrm{pf}})$*, the following transformation is a gauge symmetry,*

$$
\begin{aligned}
&\alpha \to \alpha' = \mathrm{Ad}_g\alpha + g\mathrm{d}g^{-1}, \quad B \to B' = gBg^{-1},\\
&\beta \to \beta' = g \rhd \beta, \quad C \to C' = g \rhd C,\\
&\gamma \to \gamma' = g \rhd \gamma, \quad D \to D' = g \rhd D,
\end{aligned}
\tag{60}
$$

where $g = \exp(\epsilon_{\mathfrak{g}} \cdot \hat{G}) = \exp(\epsilon_{\mathfrak{g}\alpha}\hat{G}^{\alpha}) \in G$*, and* $\epsilon_{\mathfrak{g}} : \mathcal{M}_4 \to \mathfrak{g}$ *is the parameter of the transformation.*

Proof. Note that if one considers an element of the group, $g \in G$, the transformations of the theorem 1 give rise to the following three-curvature transformation

$$
\mathcal{F} \to \mathcal{F}' = g\mathcal{F}g^{-1}, \qquad \mathcal{G} \to \mathcal{G}' = g \rhd \mathcal{G}, \qquad \mathcal{H} \to \mathcal{H}' = g \rhd \mathcal{H},
\tag{61}
$$

and the invariance of the 3*BF* action under this transformation follows from the *G*-invariance of the symmetric bilinear forms on $\mathfrak{g}$, $\mathfrak{h}$, and $\mathfrak{l}$. _

Let us consider two subsequent infinitesimal *G*-gauge transformations, determined by the small parameters $\epsilon_{\mathfrak{g}\,1}^{\alpha}$ and $\epsilon_{\mathfrak{g}\,2}^{\beta}$. To calculate the commutator between the generators of the *G*-gauge transformations, we will make use of the Baker–Campbell–Hausdorff (BCH) formula in the case when the parameters of the transformations are small

$$\mathrm{e}^{\epsilon_{\mathfrak{g}}{}^{\alpha}{}_{1}\hat{G}_{\alpha}}\mathrm{e}^{\epsilon_{\mathfrak{g}}{}^{\beta}{}_{2}\hat{G}_{\beta}} = \mathrm{e}^{\epsilon_{\mathfrak{g}}{}^{\alpha}{}_{1}\hat{G}_{\alpha}+\epsilon_{\mathfrak{g}}{}^{\beta}{}_{2}\hat{G}_{\beta}+\frac{1}{2}\epsilon_{\mathfrak{g}}{}^{\alpha}{}_{1}\,\epsilon_{\mathfrak{g}}{}^{\beta}{}_{2}\,[\hat{G}_{\alpha},\hat{G}_{\beta}]+O(\epsilon_{\mathfrak{g}}{}^{3})}, \tag{62}$$

from which it follows:

$$\mathrm{e}^{\epsilon_{\mathfrak{g}}{}^{\alpha}{}_{1}\hat{G}_{\alpha}}\mathrm{e}^{\epsilon_{\mathfrak{g}}{}^{\beta}{}_{2}\hat{G}_{\beta}} - \mathrm{e}^{\epsilon_{\mathfrak{g}}{}^{\beta}{}_{2}\hat{G}_{\beta}}\mathrm{e}^{\epsilon_{\mathfrak{g}}{}^{\alpha}{}_{1}\hat{G}_{\alpha}} = \epsilon_{\mathfrak{g}}{}^{\alpha}{}_{1}\,\epsilon_{\mathfrak{g}}{}^{\beta}{}_{2}\,[\hat{G}_{\alpha},\hat{G}_{\beta}] + O(\epsilon_{\mathfrak{g}}{}^{3}). \tag{63}$$

Using the equation (63), we obtain that the generators of the *G*-gauge transformations defined in the theorem 1 satisfy the following commutation relations:

$$[\hat{G}_{\alpha},\hat{G}_{\beta}] = f_{\alpha\beta}{}^{\gamma}\hat{G}_{\gamma}, \tag{64}$$

where $f_{\alpha\beta}{}^{\gamma}$ are the structure constants of the algebra $\mathfrak{g}$. By noting that there exists an isomorphism between generators $\hat{G}_{\alpha} \cong \tau_{\alpha}$, one establishes that the group of the *G*-gauge transformations from the theorem 1 is the same as the group *G* of the two-crossed module $(L \xrightarrow{\delta} H \xrightarrow{\partial} G, \triangleright, \{\_,\_\}_{\mathrm{pf}})$. This is an important result, which will not be true for the remaining symmetry transformations, as we shall see below.

4.2. The gauge group $\tilde{H}_L$

Let us now consider the form variations of the variables corresponding to the parameter $\epsilon_{\mathfrak{h}}{}^{a}{}_{i}$. For example, one can see from the equation (E.2) that the form-variation of the variables $\alpha^{\alpha}{}_{0}$ and $\alpha^{\alpha}{}_{i}$ are:

$$\delta_0\alpha^{\alpha}{}_{o} = 0, \qquad \delta_0\alpha^{\alpha}{}_{i} = -\partial_{a}{}^{\alpha}\epsilon_{\mathfrak{h}}{}^{a}{}_{i}. \tag{65}$$

Taking into account that the action of the generator (55) gives the symmetry transformations on one hypersurface Σ_3 with the time component of the parameter equal to zero, $\epsilon_{\mathfrak{h}}{}^{a}{}_{0} = 0$, one can extrapolate that for parameter of the spacetime gauge transformations $\epsilon_{\mathfrak{h}}{}^{a}{}_{\mu}$, the form-variation of the variable $\alpha^{\alpha}{}_{\mu}$ is given as:

$$\delta_0\alpha^{\alpha}{}_{\mu} = -\partial_{a}{}^{\alpha}\epsilon_{\mathfrak{h}}{}^{a}{}_{\mu}, \tag{66}$$

and similarly for the rest of the variables. Thus, the infinitesimal symmetry transformations in the whole spacetime corresponding to the parameter $\epsilon_{\mathfrak{h}}{}^{a}{}_{\mu}$ are given by the form variations:

$$\begin{aligned}
&\delta_0\alpha^{\alpha}{}_{\mu} = -\partial_{a}{}^{\alpha}\epsilon_{\mathfrak{h}}{}^{a}{}_{\mu}, \quad \delta_0 B^{\alpha}{}_{\mu\nu} = 2C_{a[\mu|}\epsilon_{\mathfrak{h}}{}^{b}{}_{|\nu]}\triangleright_{\beta b}{}^{a}g^{\alpha\beta},\\
&\delta_0\beta^{a}{}_{\mu\nu} = -2\nabla_{[\mu|}\epsilon_{\mathfrak{h}}{}^{a}{}_{|\nu]}, \quad \delta_0 C^{a}{}_{\mu} = 2D_A X_{(ab)}{}^{A}\epsilon_{\mathfrak{h}}{}^{b}{}_{\mu},\\
&\delta_0\gamma^{A}{}_{\mu\nu\rho} = 3!\beta^{a}{}_{[\mu\nu}\epsilon_{\mathfrak{h}}{}^{b}{}_{\rho]}X_{(ab)}{}^{A}, \quad \delta_0 D = 0.
\end{aligned} \tag{67}$$

For these infinitesimal transformations one obtains the finite symmetry transformations given in theorem 2.

Theorem 2 (H-gauge transformations). *In the* 3*BF theory for the two-crossed module* $(L \xrightarrow{\delta} H \xrightarrow{\partial} G, \triangleright, \{\_,\_\}_{\mathrm{pf}})$*, the following transformation is a symmetry:*

$$\alpha \to \alpha' = \alpha - \partial\epsilon_{\mathfrak{h}}, \quad \beta \to \beta' = \beta - \nabla'\epsilon_{\mathfrak{h}} - \epsilon_{\mathfrak{h}} \wedge \epsilon_{\mathfrak{h}},$$
$$\gamma \to \gamma' = \gamma + \{\beta', \epsilon_{\mathfrak{h}}\}_{\mathrm{pf}} + \{\epsilon_{\mathfrak{h}}, \beta\}_{\mathrm{pf}}, \; B \to B' = B - C' \wedge^{\mathcal{T}} \epsilon_{\mathfrak{h}} - \epsilon_{\mathfrak{h}} \wedge^{\mathcal{D}} \epsilon_{\mathfrak{h}} \wedge^{\mathcal{D}} D,$$
$$C \to C' = C - D \wedge^{\mathcal{X}_1} \epsilon_{\mathfrak{h}} - D \wedge^{\mathcal{X}_2} \epsilon_{\mathfrak{h}}, \quad D \to D' = D. \tag{68}$$

where $\epsilon_{\mathfrak{h}} \in \mathcal{A}^1(\mathcal{M}_4, \mathfrak{h})$ *is an arbitrary* $\mathfrak{h}$*-valued one-form, and* ∇' *denotes the covariant derivative with respect to the connection* α'*. The maps* $\mathcal{T}$*,* $\mathcal{D}$*,* $\mathcal{X}_1$*, and* $\mathcal{X}_2$ *are defined in appendix* D.

Proof. Note that the three-curvature transforms as

$$\begin{aligned} &\mathcal{F} \to \mathcal{F}' = \mathcal{F}, \\ &\mathcal{G} \to \mathcal{G}' = \mathcal{G} - \mathcal{F} \wedge^{\triangleright} \epsilon_{\mathfrak{h}}, \\ &\mathcal{H} \to \mathcal{H}' = \mathcal{H} + \{\mathcal{G}', \epsilon_{\mathfrak{h}}\}_{\mathrm{pf}} - \{\epsilon_{\mathfrak{h}}, \mathcal{G}\}_{\mathrm{pf}}. \end{aligned} \tag{69}$$

Taking into account the transformations of the three-curvature (69) and the transformations of the Lagrange multipliers, the action S_{3BF} transforms as:

$$\begin{aligned} S'_{3BF} = S_{3BF} + \int_{\mathcal{M}_4} \big(&-\langle C' \wedge^{\mathcal{T}} \epsilon_{\mathfrak{h}}, \mathcal{F}\rangle_{\mathfrak{g}} - \langle \epsilon_{\mathfrak{h}} \wedge^{\mathcal{D}} \epsilon_{\mathfrak{h}} \wedge^{\mathcal{D}} D, \mathcal{F}\rangle_{\mathfrak{g}} \\ &- \langle C', \mathcal{F} \wedge^{\triangleright} \epsilon_{\mathfrak{h}}\rangle_{\mathfrak{h}} - \langle D \wedge^{\mathcal{X}_1} \epsilon_{\mathfrak{h}}, \mathcal{G}\rangle_{\mathfrak{h}} - \langle D \wedge^{\mathcal{X}_2} \epsilon_{\mathfrak{h}}, \mathcal{G}\rangle_{\mathfrak{h}} \\ &+ \langle D, \{\mathcal{G}, \epsilon_{\mathfrak{h}}\}_{\mathrm{pf}}\rangle_{\mathfrak{l}} - \langle D, \{\mathcal{F} \wedge^{\triangleright} \epsilon_{\mathfrak{h}}, \epsilon_{\mathfrak{h}}\}_{\mathrm{pf}}\rangle_{\mathfrak{l}} - \langle D, \{\epsilon_{\mathfrak{h}}, \mathcal{G}\}_{\mathrm{pf}}\rangle_{\mathfrak{l}} \big) . \end{aligned} \tag{70}$$

Using the definitions of the maps $\mathcal{T}, \mathcal{D}$, $\mathcal{X}_1$, and $\mathcal{X}_2$, given in appendix D, one sees that the terms in the parentheses cancel, specifically the first term with the third, second with seventh, fourth with eighth, and fifth with the sixth term. ■

The H-gauge transformations do not form a group. Namely, one can check that the two consecutive H-gauge transformations do not give a transformation of the same kind, i.e. the closure axiom of the group is not satisfied. This is analogous to the well-known structure of Lorentz group, where boost transformations are not closed, and thus do not form a group. Indeed, one must consider both rotations and boosts to obtain the set of transformations that forms the Lorentz group. In the case of the H-gauge transformations, we will show that together with the H-gauge transformations one needs to consider the transformations corresponding to the parameter $\epsilon_{\mathfrak{l}}{}^{A}{}_{ij}$. From the equation (E.2) one reads the form-variations on a space hypersurface Σ_3 corresponding to this parameter. Similarly as it is done in the case of the H-gauge transformations, one extrapolates that the form-variations for all the variables corresponding to the parameter $\epsilon_{\mathfrak{l}}{}^{A}{}_{\mu\nu}$ are given as:

$$
\begin{aligned}
&\delta_0 \alpha^{\alpha}{}_{\mu} = 0, \\
&\delta_0 B^{\alpha}{}_{\mu\nu} = -D_A \triangleright_{\beta B}{}^{A} \epsilon_{\mathfrak{l}}{}^{B}{}_{\mu\nu} g^{\alpha\beta}, \\
&\delta_0 \beta^{a}{}_{\mu\nu} = \delta_A{}^{a} \epsilon_{\mathfrak{l}}^{A}{}_{\mu\nu}, \\
&\delta_0 C^{a}{}_{\mu} = 0, \quad \delta_0 \gamma^{A}{}_{\mu\nu\rho} = \nabla_\mu \epsilon_{\mathfrak{l}}^{A}{}_{\nu\rho} - \nabla_\nu \epsilon_{\mathfrak{l}}^{A}{}_{\mu\rho} + \nabla_\rho \epsilon_{\mathfrak{l}}^{A}{}_{\mu\nu}, \\
&\delta_0 D^{A} = 0.
\end{aligned}
\tag{71}
$$

These infinitesimal transformations correspond to the finite symmetry transformations defined in theorem 3.

Theorem 3 (*L*-gauge transformations). *In the* 3*BF theory for the two-crossed module* $(L \xrightarrow{\delta} H \xrightarrow{\partial} G, \triangleright, \{\_,\_\}_{\mathrm{pf}})$, *the following transformation is a symmetry*

$$
\begin{aligned}
&\alpha \to \alpha' = \alpha, &&B \to B' = B + D\wedge^{\mathcal{S}} \epsilon_{\mathfrak{l}}, \\
&\beta \to \beta' = \beta + \delta\epsilon_{\mathfrak{l}}, &&C \to C' = C, \\
&\gamma \to \gamma' = \gamma + \nabla\epsilon_{\mathfrak{l}}, &&D \to D' = D,
\end{aligned}
\tag{72}
$$

where $\epsilon_{\mathfrak{l}} \in \mathcal{A}^2(\mathcal{M}_4, \mathfrak{l})$ *is an arbitrary* $\mathfrak{l}$*-valued two-form, and the map* $\mathcal{S}$ *is defined in appendix* D.

Proof. Note that the three-curvature transforms as

$$
\begin{aligned}
&\mathcal{F} \to \mathcal{F}' = \mathcal{F}, \\
&\mathcal{G} \to \mathcal{G}' = \mathcal{G}, \\
&\mathcal{H} \to \mathcal{H}' = \mathcal{H} + \mathcal{F}\wedge^{\triangleright} \epsilon_{\mathfrak{l}}.
\end{aligned}
\tag{73}
$$

Taking into account the transformations (73) and the transformations of the Lagrange multipliers, the action transforms as:

$$
S'_{3BF} = S_{3BF} + \int_{\mathcal{M}_4} \left(\langle D\wedge^{\mathcal{S}} \epsilon_{\mathfrak{l}}, \mathcal{F} \rangle_{\mathfrak{g}} + \langle D, \mathcal{F}\wedge^{\triangleright} \epsilon_{\mathfrak{l}} \rangle_{\mathfrak{l}} \right). \tag{74}
$$

According to the definition of the map $\mathcal{S}$, the terms in the parentheses cancel. _

Let us denote the generators of the *H*-gauge transformations given by the theorem 2 and the *L*-gauge transformations given by the theorem 3 as $\hat{H}_a{}^{\mu}$ and $\hat{L}_A{}^{\mu\nu}$, respectively. As we have commented above, one can now check that the transformations defined in the theorem 2, i.e. the *H*-gauge transformations, do not form a group. If one performs two consecutive *H*-gauge transformations, defined with parameters $\epsilon_{\mathfrak{h}1}$ and $\epsilon_{\mathfrak{h}2}$, one obtains

$$
\mathrm{e}^{\epsilon_{\mathfrak{h}1}\cdot\hat{H}} \mathrm{e}^{\epsilon_{\mathfrak{h}2}\cdot\hat{H}} - \mathrm{e}^{\epsilon_{\mathfrak{h}2}\cdot\hat{H}} \mathrm{e}^{\epsilon_{\mathfrak{h}1}\cdot\hat{H}} = 2\left(\{\epsilon_{\mathfrak{h}1} \wedge \epsilon_{\mathfrak{h}2}\}_{\mathrm{pf}} - \{\epsilon_{\mathfrak{h}2} \wedge \epsilon_{\mathfrak{h}1}\}_{\mathrm{pf}}\right) \cdot \hat{L}, \tag{75}
$$

where $\epsilon_{\mathfrak{h}} \cdot \hat{H} = \epsilon_{\mathfrak{h}}{}^{a}{}_{\mu} \hat{H}_a{}^{\mu}$ and $\epsilon_{\mathfrak{l}} \cdot \hat{L} = \frac{1}{2} \epsilon_{\mathfrak{l}}{}^{A}{}_{\mu\nu} \hat{L}_A{}^{\mu\nu}$. Using the equation analogous to BCH formula (63), one obtains that the commutator of the generators of two *H*-gauge transformations is the generator of an L-gauge transformation (see appendix F for the details of the calculation):

$$\left[\hat{H}_a{}^{\mu}, \hat{H}_b{}^{\nu}\right] = 2X_{(ab)}{}^{A}\hat{L}_A{}^{\mu\nu}. \tag{76}$$

Next, note that the transformations defined in theorem 3 are the linear transformations, and the two subsequent L-gauge transformations give one L-gauge transformation with the parameter $\epsilon_{\mathfrak{l}1} + \epsilon_{\mathfrak{l}2}$. Formally, one can write the previous statement as

$$\mathrm{e}^{\epsilon_{\mathfrak{l}1}\cdot\hat{L}}\mathrm{e}^{\epsilon_{\mathfrak{l}2}\cdot\hat{L}} = \mathrm{e}^{(\epsilon_{\mathfrak{l}1}+\epsilon_{\mathfrak{l}2})\cdot\hat{L}}, \tag{77}$$

which leads to the conclusion that the generators of the L-gauge transformations are mutually commuting:

$$\left[\hat{L}_A{}^{\mu\nu}, \hat{L}_B{}^{\rho\sigma}\right] = 0. \tag{78}$$

Thus, the L-gauge transformations form an abelian group, which will be denoted as $\tilde{L}$. According to the index structure of the parameters and generators, we can conclude that the group $\tilde{L}$ is isomorphic to $\mathbb{R}^{6r}$, where r is the dimension of the group L:

$$\tilde{L} \cong \mathbb{R}^{6r}. \tag{79}$$

Our analogy with the case of the Lorentz group can once again prove useful, since the closure of the L-gauge transformations resembles the fact that the composition of two rotations is a rotation. The abelian group $\tilde{L}$ should not be confused with the non-abelian group L of the two-crossed module $(L \xrightarrow{\delta} H \xrightarrow{\partial} G, \triangleright, \{\_,\_\}_{\mathrm{pf}})$.

Let us now examine the relationship between H-gauge transformations and L-gauge transformations. The following result,

$$\mathrm{e}^{\epsilon_{\mathfrak{h}}\cdot\hat{H}}\mathrm{e}^{\epsilon_{\mathfrak{l}}\cdot\hat{L}} = \mathrm{e}^{\epsilon_{\mathfrak{l}}\cdot\hat{L}}\mathrm{e}^{\epsilon_{\mathfrak{h}}\cdot\hat{H}}, \tag{80}$$

leads to the conclusion that the commutator of generators of the H-gauge transformations and generators of the L-gauge transformations vanishes:

$$\left[\hat{H}_a{}^{\mu}, \hat{L}_A{}^{\nu\rho}\right] = 0. \tag{81}$$

From the closure of the algebra (76), (78) and (81), one can conclude that the H-gauge transformations together with the L-gauge transformations form a group, which will be denoted as $\tilde{H}_L$. Lastly, the action of the group G on the H-gauge and L-gauge transformations is examined by calculating the expressions:

$$[\epsilon_{\mathfrak{g}}\cdot\hat{G}, \epsilon_{\mathfrak{h}}\cdot\hat{H}] = (\epsilon_{\mathfrak{g}} \triangleright \epsilon_{\mathfrak{h}})\cdot\hat{H}, \qquad [\epsilon_{\mathfrak{g}}\cdot\hat{G}, \epsilon_{\mathfrak{l}}\cdot\hat{L}] = (\epsilon_{\mathfrak{g}} \triangleright \epsilon_{\mathfrak{l}})\cdot\hat{L}, \tag{82}$$

which lead to the following commutators:

$$\begin{aligned}\left[\hat{G}_\alpha, \hat{H}_a{}^{\mu}\right] &= \triangleright_{\alpha a}{}^{b}\,\hat{H}_b{}^{\mu}, \\ \left[\hat{G}_\alpha, \hat{L}_A{}^{\mu\nu}\right] &= \triangleright_{\alpha A}{}^{B}\,\hat{L}_B{}^{\mu\nu}.\end{aligned} \tag{83}$$

Theorems 1–3 represent the G-, H-, and L-gauge transformations, which are already familiar from the previous literature (see for example [21, 30]).

4.3. The gauge groups M and N

Next, consider the infinitesimal transformation with the parameter ${\epsilon_{\mathfrak{m}}}^{\alpha}{}_{i}$, given by the form variations in appendix E. In a similar manner as done in the previous subsection, one establishes that the form variations obtained as a result of the Hamiltonian analysis are transformations on one hypersurface Σ_3, from which one can guess the symmetry in the whole spacetime. Keeping in mind that the variations on the hypersurface have the time component of the parameter set to ${\epsilon_{\mathfrak{m}}}^{\alpha}{}_{0} = 0$, one extrapolates the form-variations of the whole spacetime for the parameter ${\epsilon_{\mathfrak{m}}}^{\alpha}{}_{\mu}$ to be:

$$
\begin{aligned}
&\delta_0 \alpha^{\alpha}{}_{\mu} = 0,\\
&\delta_0 B^{\alpha}{}_{\mu\nu} = -2\nabla_{[\mu|}{\epsilon_{\mathfrak{m}}}^{\alpha}{}_{|\nu]},\\
&\delta_0 \beta^{a}{}_{\mu\nu} = 0,\\
&\delta_0 C^{a}{}_{\mu} = -\partial^{a}{}_{\alpha}{\epsilon_{\mathfrak{m}}}^{\alpha}{}_{\mu},\\
&\delta_0 \gamma^{A}{}_{\mu\nu\rho} = 0,\\
&\delta_0 D^{A} = 0.
\end{aligned}
\tag{84}
$$

Based on this result, one obtains the finite symmetry transformations in the whole spacetime, as defined in theorem 4, which we will refer to as the *M*-gauge transformations.

Theorem 4 (M-gauge transformations). *In the 3BF theory for the two-crossed module* $(L \xrightarrow{\delta} H \xrightarrow{\partial} G, \triangleright, \{\_,\_\}_{\mathrm{pf}})$*, the following transformation is a symmetry*

$$
\begin{aligned}
\alpha &\to \alpha' = \alpha,\\
B &\to B' = B - \nabla\epsilon_{\mathfrak{m}},\\
\beta &\to \beta' = \beta,\\
C^{a} &\to C'^{a} = C^{a} - \partial^{a}{}_{\alpha}{\epsilon_{\mathfrak{m}}}^{\alpha},\\
\gamma &\to \gamma' = \gamma,\\
D &\to D' = D,
\end{aligned}
\tag{85}
$$

where $\epsilon_{\mathfrak{m}} \in \mathcal{A}^1(\mathcal{M}_4, \mathfrak{g})$ *is an arbitrary* $\mathfrak{g}$*-valued one-form.*

Proof. Consider the transformation of the $3BF$ action under the transformations of the variables defined in the theorem 4. One obtains:

$$
S'_{3BF} = S_{3BF} + \int_{\mathcal{M}_4} \mathrm{d}^4x\, \epsilon^{\mu\nu\rho\sigma} \left(-\frac{1}{2}(\nabla_{\mu}{\epsilon_{\mathfrak{m}}}^{\alpha}{}_{\nu})\mathcal{F}_{\alpha\rho\sigma} - \frac{1}{3!}\partial^{a}{}_{\alpha}{\epsilon_{\mathfrak{m}}}^{\alpha}{}_{\mu}\mathcal{G}_{a\nu\rho\sigma} \right). \tag{86}
$$

Using the definition of three-curvature, given by the expressions (14), one obtains:

$$
\begin{aligned}
S'_{3BF} = S_{3BF} + \int_{\mathcal{M}_4} \mathrm{d}^4x\, \epsilon^{\mu\nu\rho\sigma} \bigg(&-\frac{1}{2}(\nabla_{\mu}{\epsilon_{\mathfrak{m}}}^{\alpha}{}_{\nu}) \left(F_{\alpha\rho\sigma} - \partial^{a}{}_{\alpha}\beta_{a\rho\sigma}\right)\\
&- \frac{1}{3!}\partial^{a}{}_{\alpha}{\epsilon_{\mathfrak{m}}}^{\alpha}{}_{\mu} \left(3\nabla_{\nu}\beta_{a\rho\sigma} - \delta^{A}{}_{a}\gamma_{A\nu\rho\sigma}\right) \bigg).
\end{aligned}
\tag{87}
$$

Taking into account that the second and the third term cancel, while the last term is zero because of the identity (A.1), the expression reduces to:

$$S'_{3BF} = S_{3BF} - \frac{1}{2}\int_{\mathcal{M}_4} \mathrm{d}^4x\, \epsilon^{\mu\nu\rho\sigma} \epsilon_{\mathfrak{m}}{}^{\alpha}{}_{\mu} \nabla_{\nu} F_{\alpha\rho\sigma}. \tag{88}$$

Finally, the term $\epsilon^{\mu\nu\rho\sigma}\nabla_{\nu}F_{\alpha\rho\sigma} = 0$ is the BI (38). One concludes that the action S_{3BF} is invariant under the transformation defined in theorem 4. □

Note that the transformations defined in theorem 4 are linear transformations, and the two subsequent M-gauge transformations give one M-gauge transformation with the parameter $\epsilon_{\mathfrak{m}1} + \epsilon_{\mathfrak{m}2}$. Denoting the generators of the M-gauge transformations as $\hat{M}_{\alpha}{}^{\mu}$, one can now write the previous statement formally as:

$$\mathrm{e}^{\epsilon_{\mathfrak{m}1}\cdot\hat{M}}\mathrm{e}^{\epsilon_{\mathfrak{m}2}\cdot\hat{M}} = \mathrm{e}^{(\epsilon_{\mathfrak{m}1}+\epsilon_{\mathfrak{m}2})\cdot\hat{M}}, \tag{89}$$

where $\epsilon_{\mathfrak{m}} \cdot \hat{M} = \epsilon_{\mathfrak{m}}{}^{\alpha}{}_{\mu}\hat{M}_{\alpha}{}^{\mu}$, leading to the conclusion that:

$$\left[\hat{M}_{\alpha}{}^{\mu}, \hat{M}_{\beta}{}^{\nu}\right] = 0. \tag{90}$$

Thus, the M-gauge transformations form an abelian group, which will be denoted as $\tilde{M}$. According to the index structure of its parameters and generators, we see that this group is isomorphic to $\mathbb{R}^{4p}$, where p is the dimension of the group G:

$$\tilde{M} \cong \mathbb{R}^{4p}. \tag{91}$$

Next, one can examine the relationship of M-gauge transformations with the G, H, and L-gauge transformations defined in the previous subsections. Specifically, considering the G-gauge symmetry generators, one finds

$$[\epsilon_{\mathfrak{g}} \cdot \hat{G}, \epsilon_{\mathfrak{m}} \cdot \hat{M}] = (\epsilon_{\mathfrak{g}} \rhd \epsilon_{\mathfrak{m}}) \cdot \hat{M}, \tag{92}$$

obtaining the result:

$$\left[\hat{G}_{\alpha}, \hat{M}_{\beta}{}^{\mu}\right] = f_{\alpha\beta}{}^{\gamma}\hat{M}_{\gamma}{}^{\mu}. \tag{93}$$

Considering the H- and L-gauge transformations, one obtains

$$\mathrm{e}^{\epsilon_{\mathfrak{h}}\cdot\hat{H}}\mathrm{e}^{\epsilon_{\mathfrak{m}}\cdot\hat{M}} = \mathrm{e}^{\epsilon_{\mathfrak{m}}\cdot\hat{M}}\mathrm{e}^{\epsilon_{\mathfrak{h}}\cdot\hat{H}}, \tag{94}$$

$$\mathrm{e}^{\epsilon_{\mathfrak{l}}\cdot\hat{L}}\mathrm{e}^{\epsilon_{\mathfrak{m}}\cdot\hat{M}} = \mathrm{e}^{\epsilon_{\mathfrak{m}}\cdot\hat{M}}\mathrm{e}^{\epsilon_{\mathfrak{l}}\cdot\hat{L}}, \tag{95}$$

leading to the conclusion that the generators of the M-gauge transformations commute with both the generators of H-gauge transformations and the generators of the L-gauge transformations:

$$\left[\hat{H}_{a}, \hat{M}_{\alpha}{}^{\mu}\right] = 0, \qquad \left[\hat{L}_{A}{}^{\mu\nu}, \hat{M}_{\alpha}{}^{\rho}\right] = 0. \tag{96}$$

Finally, examining the infinitesimal transformation corresponding to the parameter $\epsilon_{\mathfrak{n}}{}^{a}$, given by the form-variations as calculated in (E.2),

$$
\begin{aligned}
&\delta_0\alpha^{\alpha}{}_{\mu} = 0,\\
&\delta_0 B^{\alpha}{}_{\mu\nu} = \beta_{b\mu\nu} \triangleright_{\alpha' a}{}^{b}\epsilon_{\mathfrak{n}}{}^{a} g^{\alpha\alpha'},\\
&\delta_0\beta^{a}{}_{\mu\nu} = 0,\\
&\delta_0 C^{a}{}_{\mu} = -\nabla_{\mu}\epsilon_{\mathfrak{n}}{}^{a},\\
&\delta_0\gamma^{A}{}_{\mu\nu\rho} = 0,\\
&\delta_0 D^{A} = \delta^{A}{}_{a}\epsilon_{\mathfrak{n}}{}^{a}.
\end{aligned} \tag{97}
$$

one obtains the theorem 5, the symmetry transformations which will be referred to as *N*-gauge transformations. Note that the *N*-gauge transformations are simultaneously the transformations in the whole spacetime, since the parameter does not carry spacetime indices.

Theorem 5 (N-gauge transformations). *In the* 3*BF theory for the two-crossed module* $(L \xrightarrow{\delta} H \xrightarrow{\partial} G, \triangleright, \{\_,\_\}_{\mathrm{pf}})$*, the following transformation is a symmetry*

$$
\begin{aligned}
\alpha &\to \alpha' = \alpha,\\
B &\to B' = B - \beta \wedge^{\mathcal{T}} \epsilon_{\mathfrak{n}},\\
\beta &\to \beta' = \beta,\\
C &\to C' = C - \nabla\epsilon_{\mathfrak{n}},\\
\gamma &\to \gamma' = \gamma,\\
D^{A} &\to D'^{A} = D^{A} + \delta^{A}{}_{a}\epsilon_{\mathfrak{n}}{}^{a},
\end{aligned} \tag{98}
$$

where $\epsilon_{\mathfrak{n}} : \mathcal{M}_4 \to \mathfrak{h}$ *is an arbitrary* $\mathfrak{h}$*-valued zero-form.*

Proof. Under the transformations defined in theorem 5, the action is transformed as follows:

$$
\begin{aligned}
S'_{3BF} = S_{3BF} + \int_{\mathcal{M}_4} \mathrm{d}x^4 \epsilon^{\mu\nu\rho\sigma} \bigg(&\frac{1}{4}\beta_{b\mu\nu} \triangleright_{\alpha a}{}^{b}\epsilon_{\mathfrak{n}}{}^{a}\mathcal{F}^{\alpha}{}_{\rho\sigma} - \frac{1}{3!}(\nabla_{\mu}\epsilon_{\mathfrak{n}}{}^{a})\mathcal{G}_{a\nu\rho\sigma}\\
&+ \frac{1}{4!}\delta^{A}{}_{a}\epsilon_{\mathfrak{n}}{}^{a}\mathcal{H}_{A\,\mu\nu\rho\sigma}\bigg).
\end{aligned} \tag{99}
$$

Using the expressions for the three-curvature defined in (9), one obtains

$$
\begin{aligned}
S'_{3BF} = S_{3BF} + \int_{\mathcal{M}_4} \mathrm{d}x^4 \epsilon^{\mu\nu\rho\sigma} \bigg(&\frac{1}{4}\beta_{b\mu\nu} \triangleright_{\alpha a}{}^{b}\epsilon_{\mathfrak{n}}{}^{a}\left(F^{\alpha}{}_{\rho\sigma} - \partial_{\mathrm{c}}{}^{\alpha}\beta^{c}{}_{\rho\sigma}\right)\\
&- \frac{1}{3!}(\nabla_{\mu}\epsilon_{\mathfrak{n}}{}^{a})\left(3\nabla_{\nu}\beta_{a\rho\sigma} - \delta^{A}{}_{a}\gamma_{A\,\nu\rho\sigma}\right)\\
&+ \frac{1}{4!}\delta^{A}{}_{a}\epsilon^{a}\left(4\nabla_{\mu}\gamma_{A\,\nu\rho\sigma} + 6X_{(bc)A}\beta^{b}{}_{\mu\nu}\beta^{c}{}_{\rho\sigma}\right)\bigg).
\end{aligned} \tag{100}
$$

Here, after one partial integration the last term in the first row of the equation (100) cancels with the first term in the second row, while taking into account the identity

$$\frac{1}{2}\epsilon^{\mu\nu\rho\sigma}(\nabla_\nu\nabla_\mu\epsilon_{\mathfrak{n}}{}^a)\beta_{a\rho\sigma} = -\frac{1}{4}\epsilon^{\mu\nu\rho\sigma}\beta_{b\rho\sigma}\rhd_{\alpha a}{}^b\epsilon_{\mathfrak{n}}{}^a F^\alpha{}_{\mu\nu}, \tag{101}$$

the first term and the third term also cancel, leading to the following expression:

$$S'_{3BF} = S_{3BF} + \int_{\mathcal{M}_4} \mathrm{d}x^4\epsilon^{\mu\nu\rho\sigma}\left(\frac{1}{4}\epsilon_{\mathfrak{n}a}\rhd_{\alpha(b|}{}^a\partial_{|c)}{}^\alpha\beta^b{}_{\mu\nu}\beta^c{}_{\rho\sigma}\right.$$
$$\left.+\frac{1}{4}\epsilon_{\mathfrak{n}a}\delta_A{}^a X_{(bc)}{}^A\beta^b{}_{\mu\nu}\beta^c{}_{\rho\sigma}\right). \tag{102}$$

Here, the remaining two terms vanish because of the symmetrized form of the identity (A.6):

$$\rhd_{\alpha(b|}{}^a\partial_{|c)}{}^\alpha + \delta_A{}^a X_{(bc)}{}^A = f_{(bc)}{}^a = 0,$$

as a consequence of the antisymmetry of the structure constants. One concludes that the S_{3BF} action is invariant under the transformations defined in theorem 5. □

The N-gauge transformations defined in theorem 5 define the group which will be denoted as $\tilde{N}$. Note that these transformations are also linear, and the composition of two N-gauge transformations gives one N-gauge transformation with the parameter $\epsilon_{\mathfrak{n}1}+\epsilon_{\mathfrak{n}2}$. The generators of the group $\tilde{N}$ will be denoted with $\hat{N}_a$, and one can write these results as:

$$\mathrm{e}^{\epsilon_{\mathfrak{n}1}\cdot\hat{N}}\mathrm{e}^{\epsilon_{\mathfrak{n}2}\cdot\hat{N}} = \mathrm{e}^{(\epsilon_{\mathfrak{n}1}+\epsilon_{\mathfrak{n}2})\cdot\hat{N}}, \tag{103}$$

where $\epsilon_{\mathfrak{n}}\cdot\hat{N} = \epsilon_{\mathfrak{n}}^a\hat{N}_a$, leading to the conclusion that:

$$[\hat{N}_a, \hat{N}_b] = 0. \tag{104}$$

It follows that the group $\tilde{N}$ is abelian, and the index structure of parameters and generators indicates that it is isomorphic to $\mathbb{R}^q$, where q is the dimension of the group H. Therefore,

$$\tilde{N} \cong \mathbb{R}^q. \tag{105}$$

Next, one can examine the relationship of the N-gauge transformations with the G, H, L, and M-gauge transformations. First, considering the G-gauge transformations one obtains:

$$[\epsilon_{\mathfrak{g}}\cdot\hat{G}, \epsilon_{\mathfrak{n}}\cdot\hat{N}] = (\epsilon_{\mathfrak{g}}\rhd\epsilon_{\mathfrak{n}})\cdot\hat{N}, \tag{106}$$

from which it follows:

$$[\hat{G}_\alpha, \hat{N}_a] = \rhd_{\alpha a}{}^b\hat{N}_b. \tag{107}$$

Let us now examine the relationship between N-gauge transformations and H-gauge transformations, calculating the following expression:

$$\mathrm{e}^{\epsilon_{\mathfrak{h}}\cdot\hat{H}}\mathrm{e}^{\epsilon_{\mathfrak{n}}\cdot\hat{N}} - \mathrm{e}^{\epsilon_{\mathfrak{n}}\cdot\hat{N}}\mathrm{e}^{\epsilon_{\mathfrak{h}}\cdot\hat{H}} = -(\epsilon_{\mathfrak{n}}\wedge^{\mathcal{T}}\epsilon_{\mathfrak{h}})\cdot\hat{M}, \tag{108}$$

where the proof is given in appendix F. One obtains that the commutator between the generators of H-gauge transformation and N-gauge transformation is the generator of M-gauge transformation:

$$\left[\hat{H}_a{}^\mu, \hat{N}^b\right] = \rhd_{\alpha a}{}^b\hat{M}^{\alpha\mu}. \tag{109}$$

Analogously, one can check that the following is satisfied

$$e^{\epsilon_{\mathfrak{l}}\cdot\hat{L}}e^{\epsilon_{\mathfrak{n}}\cdot\hat{N}} = e^{\epsilon_{\mathfrak{n}}\cdot\hat{N}}e^{\epsilon_{\mathfrak{l}}\cdot\hat{L}}, \qquad e^{\epsilon_{\mathfrak{m}}\cdot\hat{M}}e^{\epsilon_{\mathfrak{n}}\cdot\hat{N}} = e^{\epsilon_{\mathfrak{n}}\cdot\hat{N}}e^{\epsilon_{\mathfrak{m}}\cdot\hat{M}}, \tag{110}$$

leading to the conclusion that the generators of L-gauge, M-gauge, and N-gauge transformations mutually commute, i.e.

$$\left[\hat{M}_{\alpha}{}^{\mu}, \hat{N}_a\right] = 0, \qquad \left[\hat{L}_A{}^{\mu\nu}, \hat{N}_a\right] = 0. \tag{111}$$

This concludes the calculation of the algebra of generators.

4.4. Structure of the symmetry group

Summarizing the results of the previous subsections, one can write the algebra of the generators of the full gauge symmetry group as follows.

- The algebra $\mathfrak{g}$ of the group G of the two-crossed module $(L \xrightarrow{\delta} H \xrightarrow{\partial} G, \triangleright, \{\_,\_\}_{\mathrm{pf}})$:

$$[\hat{G}_\alpha, \hat{G}_\beta] = f_{\alpha\beta}{}^{\gamma}\hat{G}_\gamma. \tag{112}$$

- The algebra of the group $\tilde{H}_L$ consisting of the generators of H- and L-gauge transformations:

$$\begin{aligned} &[\hat{H}_a{}^{\mu}, \hat{H}_b{}^{\nu}] = 2X_{(ab)}{}^{A}\hat{L}_A{}^{\mu\nu}, \\ &[\hat{L}_A{}^{\mu\nu}, \hat{L}_B{}^{\rho\sigma}] = 0, \\ &[\hat{H}_a{}^{\mu}, \hat{L}_A{}^{\nu\rho}] = 0. \end{aligned} \tag{113}$$

- The algebra of the generators of M-gauge transformations:

$$[\hat{M}_{\alpha}{}^{\mu}, \hat{M}_{\beta}{}^{\nu}] = 0. \tag{114}$$

- The algebra of the generators of N-gauge transformations:

$$[\hat{N}_a, \hat{N}_b] = 0. \tag{115}$$

- The commutators between the generators of the groups $\tilde{M}$ and $\tilde{N}$:

$$[\hat{M}_{\alpha}{}^{\mu}, \hat{N}_a] = 0. \tag{116}$$

- The action of the generators of the group $\tilde{H}_L$ on the generators of M- and N-gauge transformations:

$$\begin{aligned} &[\hat{H}_a{}^{\mu}, \hat{N}^b] = \triangleright_{\alpha a}{}^{b}\hat{M}^{\alpha\mu}, \\ &[\hat{H}_a{}^{\mu}, \hat{M}_{\alpha}{}^{\nu}] = 0, \\ &[\hat{L}_A{}^{\nu\rho}, \hat{M}_{\alpha}{}^{\mu}] = 0, \\ &[\hat{L}_A{}^{\mu\nu}, \hat{N}_a] = 0. \end{aligned} \tag{117}$$

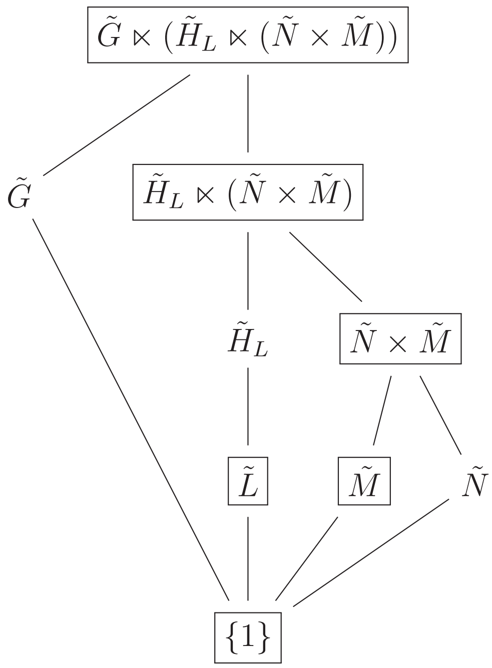


Figure 1. Relevant subgroups of the symmetry group $\mathcal{G}_{3BF}$. The invariant subgroups are boxed.

- The action of the generators of the group G on the generators of H-, L-, M-, and N-gauge transformations:

$$
\begin{aligned}
\left[\hat{G}_\alpha, \hat{H}_a{}^{\mu}\right] &= \triangleright_{\alpha a}{}^{b}\, \hat{H}_b{}^{\mu}, \\
\left[\hat{G}_\alpha, \hat{L}_A{}^{\mu\nu}\right] &= \triangleright_{\alpha A}{}^{B} \hat{L}_B{}^{\mu\nu}, \\
\left[\hat{G}_\alpha, \hat{M}_\beta{}^{\mu}\right] &= f_{\alpha\beta}{}^{\gamma} \hat{M}_\gamma{}^{\mu}, \\
\left[\hat{G}_\alpha, \hat{N}_a\right] &= \triangleright_{\alpha a}{}^{b}\, \hat{N}_b.
\end{aligned}
\tag{118}
$$

Based on the equations (112)–(118), one can investigate the symmetry group structure. On the Hesse-like diagram shown in figure 1, we have included only the relevant subgroups of the whole symmetry group $\mathcal{G}_{3BF}$, where the invariant subgroups are boxed.

Let us remember that the subgroup is an *invariant subgroup*, or equivalently a *normal subgroup*, if it is invariant under conjugation by members of the group of which it is a subgroup. Formally, one says the group H is an invariant subgroup of the group G if H is a subgroup of G, i.e. $H \leqslant G$, and for all $h \in H$ and $g \in G$, the conjugation of the element of H with the element of G is an element of H, i.e. $\exists h' \in H$ such that $ghg^{-1} = h'$. On the level of algebra, the corresponding object is an *ideal*. Formally written, an algebra A is a subalgebra of an algebra L with respect to the multiplication in L, i.e. $[A, A] \subset A$. Then, a subalgebra A of L is an *ideal* in L if its elements, multiplied with any element of the algebra, give again an element of the subalgebra, i.e. $[A, L] \subset A$.

With the above definitions in mind, note first that the groups $\tilde{L}$, $\tilde{M}$, and $\tilde{N}$, are subgroups of the full symmetry group $\mathcal{G}_{3BF}$. The groups $\tilde{L}$ and $\tilde{M}$ are invariant subgroups, since the only nontrivial commutators between the generators $\hat{L}_A{}^{\mu\nu}$, and $\hat{M}_\alpha{}^\mu$, are with the generators of the group $\tilde{G}$, and are equal to some linear combinations of the generators of $\tilde{L}$, and $\tilde{M}$, respectively. The group $\tilde{N}$ is not an invariant subgroup, since the commutator between the generators $\hat{N}_a$ and $\hat{H}_a{}^\mu$ are linear combinations of the generators $\hat{M}_\alpha{}^\mu$. However, the generators of the groups $\tilde{N}$ and $\tilde{M}$ are mutually commuting, and the group $\tilde{N}$ is an invariant subgroup of the product of the groups $\tilde{M}$ and $\tilde{N}$, which makes this product a direct product. The obtained group $\tilde{N} \times \tilde{M}$ is an invariant subgroup of the whole symmetry group.

On the other hand, we saw that the H-gauge transformations together with the L-gauge transformations form the group $\tilde{H}_L$. This group is not an invariant subgroup of the whole symmetry group $\mathcal{G}_{3BF}$, because of the commutator of the generators $\hat{H}_a{}^\mu$ and $\hat{N}_b$. Similarly as before, one can join these two subgroups, of which one is invariant and one is not, using a semidirect product, to obtain a subgroup $\tilde{H}_L \ltimes (\tilde{N} \times \tilde{M})$, that will as a result be an invariant subgroup of the complete symmetry group $\mathcal{G}_{3BF}$. Here, the product is semidirect because the group $\tilde{H}_L$ is not an invariant subgroup of the group $\tilde{H}_L \ltimes (\tilde{N} \times \tilde{M})$, due to the commutator between the generators $\hat{H}_a{}^\mu$ and $\hat{N}_b$.

Finally, following the same line of reasoning, one adds the G-gauge transformations and obtains the complete gauge symmetry group $\mathcal{G}_{3BF}$ as:

$$\mathcal{G}_{3BF} = \tilde{G} \ltimes (\tilde{H}_L \ltimes (\tilde{N} \times \tilde{M})). \tag{119}$$

This concludes the analysis of the group of gauge symmetries for the 3BF action.

5. Conclusions

5.1. Summary of the results

Let us summarize the results of the paper. In section 2, we have introduced a generalization of the BF theory in the framework of higher category theory, the 3BF theory. Section 3 contains the Hamiltonian analysis for the 3BF theory. In subsection 3.1, the basic canonical structure and the total Hamiltonian are obtained, while in subsection 3.2 the complete Hamiltonian analysis of the 3BF theory is performed, resulting in the first-class and second-class constraints of the theory, as well as their Poisson brackets. In the subsection 3.3 we have discussed the BI and also the generalized BI, since they enforce restrictions and reduce the number of independent first-class constraints present in the theory, and having those identities in mind, the counting of the dynamical degrees of freedom has been performed. As expected, it was established that the considered 3BF action is a topological theory. Finally, this section concludes with the subsection 3.4 where we have constructed the generator of the gauge symmetries for the topological theory, based on the calculations done in section 3.2, and we have found the form-variations for all the variables and their canonical momenta, listed in the appendix E, equation (E.2).

In section 4, the main results of our paper are presented. With the material of the subsection 3.2 in hand, after obtaining the form variations of all variables and their canonical momenta, we proceeded to find all the gauge symmetries of the theory. The subsection 4.1 examined the gauge group G, and the G-gauge transformations. In subsection 4.2 we discussed the gauge group $\tilde{H}_L$ which gives the H-gauge and L-gauge transformations, while in the subsection 4.3 we analyzed the M-gauge and N-gauge transformations which represent a novel result. The results of the subsections 4.1–4.3 are summarized in subsection 4.4, where the complete structure of the symmetry group had been presented. The known G-, H-, and L-gauge transformations have been rigorously defined in theorems 1–3, while the two novel M- and N-gauge transformations, have been defined in theorems 4 and 5. The Lie algebra of the full gauge symmetry group $\mathcal{G}_{3BF}$ has also been obtained.

5.2. Discussion

One of the most important consequences of our results is the relationship between a two-crossed module and a symmetry group of the corresponding $3BF$ action, which we denoted as a *duality*. In particular, from the Lie algebra of the symmetry group $\mathcal{G}_{3BF}$ one sees that the structure constants depend on the choices of groups G, H, and L of the two-crossed module, on the action $\triangleright$, and on the symmetric part of the Peiffer lifting. However, $\mathcal{G}_{3BF}$ does not depend on the antisymmetric part of the Peiffer lifting, nor on the homomorphisms ∂ and δ. This means that in principle one can have several different two-crossed modules dual to the same symmetry group. Therefore, the term 'duality' is used in a loose sense, since there is no one-to-one correspondence between a two-crossed module and a symmetry group of the corresponding $3BF$ action. In addition, this result allows one to implement a strategy for the construction of a two-crossed module, by first specifying the choice of the group $\mathcal{G}_{3BF}$, and then supplying the additional information about the homomorphisms and the antisymmetric part of the Peiffer lifting, in a way that satisfies all axioms in the definition of a two-crossed module.

Another important topic for discussion is the following. From the fact that the $3BF$ action is formulated in a manifestly covariant way, using differential forms, it should be obvious that the diffeomorphisms are a symmetry of the theory. However, by looking at the structure of the gauge group $\mathcal{G}_{3BF}$, one does not immediately see whether $\mathrm{Diff}(\mathcal{M}_4, \mathbb{R})$ is its subgroup. In fact, this issue is subtle, and it deserves some discussion.

It is easy to see that every action, which depends on at least two fields $\phi_1(x)$ and $\phi_2(x)$, is invariant under the following transformation, determined by the Henneaux–Teitelboim (HT) parameter ϵ^{HT} (see [35] for details and naming),

$$\delta_0{}^{\mathrm{HT}}\phi_1 = \epsilon^{\mathrm{HT}}(x)\frac{\delta S}{\delta\phi_2}, \qquad \delta_0{}^{\mathrm{HT}}\phi_2 = -\epsilon^{\mathrm{HT}}(x)\frac{\delta S}{\delta\phi_1}, \tag{120}$$

which can be easily verified by calculating the variation of the action:

$$\delta^{\mathrm{HT}}S[\phi_1,\phi_2] = \frac{\delta S}{\delta\phi_1}\delta_0{}^{\mathrm{HT}}\phi_1 + \frac{\delta S}{\delta\phi_2}\delta_0{}^{\mathrm{HT}}\phi_2 = 0. \tag{121}$$

Since this invariance is present even in theories with no gauge symmetry, it is not associated with constraints, and thus not present in the generator of gauge symmetries (55), see [35] for details.

Now, let us consider the diffeomorphism transformation

$$x^\mu \to x'^\mu = x^\mu + \xi^\mu(x), \tag{122}$$

where the parameter $\xi^\mu(x)$ is an arbitrary function, which we will consider to be infinitesimal. Also, let us denote all parameters of the gauge group collectively as $\epsilon_i(x)$. If diffeomorphisms are a symmetry of the action, then for every field $\phi(x)$ in the theory, and every parameter of the diffeomorphisms $\xi^\mu(x)$, there should exist a choice of the parameters $\epsilon_i(x)$ and $\epsilon^{\rm HT}(x)$, such that:

$$(\delta_0{}^{\rm gauge} + \delta_0{}^{\rm HT} + \delta_0{}^{\rm diff})\,\phi = 0. \tag{123}$$

In other words, if the diffeomorphisms are a symmetry of the theory, their form variations should be expressible as gauge form variations combined with HT form variations:

$$\delta_0{}^{\rm diff}\,\phi = -\delta_0{}^{\rm gauge}\phi - \delta_0{}^{\rm HT}\phi. \tag{124}$$

In our case, the $3BF$ action depends on the fields $\alpha^\alpha{}_\mu$, $\beta^a{}_{\mu\nu}$, $\gamma^A{}_{\mu\nu\rho}$, $B^\alpha{}_{\mu\nu}$, $C^a{}_\mu$, and D^A. The HT parameters $\epsilon^{{\rm HT}\alpha\beta}{}_{\mu\nu\rho}$, $\epsilon^{{\rm HT}ab}{}_{\mu\nu\rho}$, and $\epsilon^{{\rm HT}AB}{}_{\mu\nu\rho}$ are defined via the following form variations, analogous to (120):

$$\begin{aligned}
\delta_0{}^{\rm HT}\alpha^\alpha{}_\mu &= \frac{1}{2}\epsilon^{{\rm HT}\alpha\beta}{}_{\mu\nu\rho}\frac{\delta S}{\delta B^\beta{}_{\nu\rho}},\\
\delta_0{}^{\rm HT}B^\alpha{}_{\mu\nu} &= -\epsilon^{{\rm HT}\alpha\beta}{}_{\rho\mu\nu}\frac{\delta S}{\delta \alpha^\beta{}_\rho},\\
\delta_0{}^{\rm HT}\beta^a{}_{\mu\nu} &= \epsilon^{{\rm HT}ab}{}_{\mu\nu\rho}\frac{\delta S}{\delta C^b{}_\rho},\\
\delta_0{}^{\rm HT}C^a{}_\mu &= -\frac{1}{2}\epsilon^{{\rm HT}ab}{}_{\nu\rho\mu}\frac{\delta S}{\delta \beta^b{}_{\nu\rho}},\\
\delta_0{}^{\rm HT}\gamma^A{}_{\mu\nu\rho} &= \epsilon^{{\rm HT}AB}{}_{\mu\nu\rho}\frac{\delta S}{\delta D^B},\\
\delta_0{}^{\rm HT}D^A &= -\frac{1}{3!}\epsilon^{\rm HTAB}{}_{\mu\nu\rho}\frac{\delta S}{\delta \gamma^B{}_{\mu\nu\rho}},
\end{aligned} \tag{125}$$

while the gauge parameters $\epsilon_{\mathfrak{g}}{}^\alpha$, $\epsilon_{\mathfrak{h}}{}^a{}_\mu$, $\epsilon_{\mathfrak{l}}{}^A{}_{\mu\nu}$, $\epsilon_{\mathfrak{m}}{}^\alpha{}_\mu$, and $\epsilon_{\mathfrak{n}}^a$ are defined in theorems 1–5. Given these, there indeed exists a choice of these parameters, such that (123) is satisfied for all fields. Specifically, if one chooses the gauge parameters as

$$\begin{aligned}
\epsilon_{\mathfrak{g}}{}^\alpha &= -\xi^\lambda\alpha^\alpha{}_\lambda,\\
\epsilon_{\mathfrak{h}}{}^a{}_\mu &= \xi^\lambda\beta^a{}_{\mu\lambda},\\
\epsilon_{\mathfrak{l}}^A{}_{\mu\nu} &= \xi^\lambda\gamma^A{}_{\mu\nu\lambda},\\
\epsilon_{\mathfrak{m}}{}^\alpha{}_\mu &= \xi^\lambda B^\alpha{}_{\mu\lambda},\\
\epsilon_{\mathfrak{n}}{}^a &= -\xi^\lambda C^a{}_\lambda,
\end{aligned} \tag{126}$$

and the HT parameters as

$$\begin{aligned}
\epsilon^{{\rm HT}\alpha\beta}{}_{\mu\nu\rho} &= \xi^\lambda g^{\alpha\beta}\epsilon_{\mu\nu\rho\lambda},\\
\epsilon^{{\rm HT}ab}{}_{\mu\nu\rho} &= \xi^\lambda g^{ab}\epsilon_{\lambda\mu\nu\rho},\\
\epsilon^{{\rm HT}AB}{}_{\mu\nu\rho} &= \xi^\lambda g^{AB}\epsilon_{\mu\nu\rho\lambda},
\end{aligned} \tag{127}$$

one can obtain, using (124), precisely the standard form variations corresponding to diffeomorphisms:

$$\begin{aligned}
\delta_0{}^{\text{diff}}\alpha^{\alpha}{}_{\mu} &= -\partial_\mu\xi^\lambda\alpha^\alpha{}_\lambda - \xi^\lambda\partial_\lambda\alpha^\alpha{}_\mu, \\
\delta_0{}^{\text{diff}}\beta^{a}{}_{\mu\nu} &= -\partial_\mu\xi^\lambda\beta^a{}_{\lambda\nu} - \partial_\nu\xi^\lambda\beta^a{}_{\mu\lambda} - \xi^\lambda\partial_\lambda\beta^a{}_{\mu\nu}, \\
\delta_0{}^{\text{diff}}\gamma^{A}{}_{\mu\nu\rho} &= -\partial_\mu\xi^\lambda\gamma^A{}_{\lambda\nu\rho} - \partial_\nu\xi^\lambda\gamma^A{}_{\mu\lambda\rho} - \partial_\rho\xi^\lambda\gamma^A{}_{\mu\nu\lambda} - \xi^\lambda\partial_\lambda\gamma^A{}_{\mu\nu\rho}, \\
\delta_0{}^{\text{diff}}B^{\alpha}{}_{\mu\nu} &= -\partial_\mu\xi^\lambda B^\alpha{}_{\lambda\nu} - \partial_\nu\xi^\lambda B^\alpha{}_{\mu\lambda} - \xi^\lambda\partial_\lambda B^\alpha{}_{\mu\nu}, \\
\delta_0{}^{\text{diff}}C^{a}{}_{\mu} &= -\partial_\mu\xi^\lambda C^a{}_\lambda - \xi^\lambda\partial_\lambda C^a{}_\mu, \\
\delta_0{}^{\text{diff}}D^{A} &= -\xi^\lambda\partial_\lambda D^A.
\end{aligned} \tag{128}$$

This establishes that diffeomorphisms are indeed contained in the full gauge symmetry group $\mathcal{G}_{3BF}$, up to the HT transformations, which are always a symmetry of the theory.

5.3. Future lines of investigation

Based on the results obtained in this work, one can imagine various additional topics for further research.

First, since we have obtained that the pure 3*BF* theory is a topological theory, it does not describe a realistic physical theory which ought to contain local propagating degrees of freedom. To build a realistic physical theory, one introduces the degrees of freedom by imposing the simplicity constraints on the topological action. In our previous work [13], we have formulated the classical actions that manifestly distinguish the topological sector from the simplicity constraints, for all the fields present in the standard model coupled to Einstein–Cartan gravity. Specifically, we have defined the constrained 2*BF* actions describing the Yang–Mills field and Einstein–Cartan gravity, and also the constrained 3*BF* actions describing the Klein–Gordon, Dirac, Weyl and Majorana fields coupled to gravity in the standard way. The natural continuation of this line of research would be the Hamiltonian analysis of all such constrained 3*BF* models of gravity coupled to various matter fields, and the study of their canonical quantization.

On the other hand, as an alternative to the canonical quantization, one may choose the spinfoam quantization approach, and define the path integral of the theory as the state sum for the Regge-discretized 3*BF* action. The topological nature of the 3*BF* action, together with the structure of the gauge three-group, should ensure that such a sum is a topological invariant, i.e. that it is triangulation independent. This construction was recently carried out in [14], where the 3*BF* state sum for a general two-crossed module and a closed and orientable four-dimensional manifold $\mathcal{M}_4$ is defined. Unfortunately, in order to rigorously define this state sum, one needs the higher category generalizations of the Peter–Weyl and Plancherel theorems, from ordinary groups to the cases of two-groups and three-groups. These theorems ought to determine the domains of various labels living on simplices of the triangulation, as a consequence of the representation theory of three-groups. Until these mathematical results are obtained, one can try to guess the appropriate structure of the irreducible representations of a three-group and construct the topological invariant Z for the 3*BF* topological action, in analogy with what was done in the case of 2*BF* theory, see [25, 27]. Once the topological state sum is obtained, one can proceed to impose the simplicity constraints, and thus construct the state sum corresponding to the tentative quantum theory of gravity with matter. The classical action for gravity and matter is formulated in [13] in a way that explicitly distinguishes between the topological sector and the simplicity constraints sector of the action, making the procedure of imposing the constraints straightforward.

Next, it would be useful to investigate in more depth the mathematical structure and properties of the simplicity constraints, in particular their role as the gauge fixing conditions for the symmetry group $\mathcal{G}_{3BF}$. The simplicity constraints should explicitly break the symmetry group $\mathcal{G}_{3BF}$ to the subgroup corresponding to the constrained $3BF$ theory, which may then be further spontaneously broken by the Higgs mechanism.

One of the results obtained in this work is a duality between the gauge symmetry group of the $3BF$ action, $\mathcal{G}_{3BF}$, and the underlining three-group, i.e. the two-crossed module $(L \xrightarrow{\delta} H \xrightarrow{\partial} G, \triangleright, \{\_,\_\}_{\mathrm{pf}})$. This duality should be better understood. On one hand, the group $\mathcal{G}_{3BF}$ can provide further insight into the construction of the TQFT state sum, i.e. a topological invariant corresponding to the underlining three-group structure. On the other hand, this duality is interesting from the perspective of pure mathematics, since it can provide deeper insight in the structure of three-groups. In addition, one could expect that the $3BF$ theory would have a three-group of higher gauge symmetries, but it is not obvious if the five types of gauge transformations can form a three-group structure or not. This is an important topic for future research.

Finally, in [31] it was pointed out that it may be useful to make one more step in the categorical generalization, and consider a $4BF$ theory as a description of a quantum gravity model with matter fields. One could then calculate the gauge group of the $4BF$ action, and compare the results with the results obtained for the $3BF$ theory.

The list is not conclusive, and there may be many other interesting topics to study.

Acknowledgments

This research was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia (MPNTR), and by the Science Fund of the Republic of Serbia, Program DIASPORA, Nos. 6427195, SQ2020. The contents of this publication are the sole responsibility of the authors and can in no way be taken to reflect the views of the Science Fund of the Republic of Serbia.

Data availability statement

No new data were created or analysed in this study.

Appendix A. Two-crossed module

Definition (Differential two-crossed module). A differential two-crossed module is given by an exact sequence of Lie algebras:

$$\mathfrak{l} \xrightarrow{\delta} \mathfrak{h} \xrightarrow{\partial} \mathfrak{g},$$

together with left action $\triangleright$ of $\mathfrak{g}$ on $\mathfrak{g}$, $\mathfrak{h}$, and $\mathfrak{l}$, by derivations, and on itself via adjoint representation, and a $\mathfrak{g}$-equivariant bilinear map called the **Peiffer lifting**:

$$\{\_,\_\}_{\mathrm{pf}} : \mathfrak{h} \times \mathfrak{h} \to \mathfrak{l}.$$

Fixing the basis in the algebras as $T_A \in \mathfrak{l}$, $t_a \in \mathfrak{h}$ and $\tau_\alpha \in \mathfrak{g}$:

$$[T_A, T_B] = f_{AB}{}^C\, T_C, \qquad [t_a, t_b] = f_{ab}{}^c\, t_c, \qquad [\tau_\alpha, \tau_\beta] = f_{\alpha\beta}{}^\gamma\, \tau_\gamma,$$

one defines the maps ∂ and δ as:

$$\partial(t_a) = \partial_a{}^\alpha\, \tau_\alpha, \qquad \delta(T_A) = \delta_A{}^a\, t_a,$$

and the action of $\mathfrak{g}$ on the generators of $\mathfrak{l}$, $\mathfrak{h}$, and $\mathfrak{g}$ is, respectively:

$$\tau_\alpha \rhd T_A = \rhd_{\alpha A}{}^B\, T_B, \qquad \tau_\alpha \rhd t_a = \rhd_{\alpha a}{}^b\, t_b, \qquad \tau_\alpha \rhd \tau_\beta = \rhd_{\alpha\beta}{}^\gamma\, \tau_\gamma.$$

The coefficients $X_{ab}{}^A$ are introduced as:

$$\{t_a,\, t_b\}_{\mathrm{pf}} = X_{ab}{}^A T_A.$$

The maps ∂ and δ satisfy the following identity:

$$\partial_a{}^\alpha\, \delta_A{}^a = 0. \tag{A.1}$$

Note that when η is a $\mathfrak{g}$-valued differential form and ω is $\mathfrak{l}$-, $\mathfrak{h}$-, or $\mathfrak{g}$-valued differential form, the previous action is defined as:

$$\eta \wedge^\rhd \omega = \eta^\alpha \wedge \omega^A \rhd_{\alpha A}{}^B\, T_B,$$
$$\eta \wedge^\rhd \omega = \eta^\alpha \wedge \omega^a \rhd_{\alpha a}{}^b\, t_b,$$
$$\eta \wedge^\rhd \omega = \eta^\alpha \wedge \omega^\beta f_{\alpha\beta}{}^\gamma\, \tau_\gamma,$$

where the forms are multiplied via the wedge product $\wedge$, while the generators of G act on the generators of the three groups via the action $\rhd$.

The following identities are satisfied:

(i) In the differential crossed module $(L \xrightarrow{\delta} H, \rhd')$ the action $\rhd'$ of $\mathfrak{h}$ on $\mathfrak{l}$ is defined for each $\underline{h} \in \mathfrak{h}$ and $\underline{l} \in \mathfrak{l}$ as:

$$\underline{h} \rhd' \underline{l} = -\{\delta(\underline{l}),\, \underline{h}\}_{\mathrm{pf}},$$

or written in the basis where $t_a \rhd' T_A = \rhd'{}_{aA}{}^B T_B$ the previous identity becomes:

$$\rhd'{}_{aA}{}^B = -\delta_A{}^b X_{ba}{}^B; \tag{A.2}$$

(ii) The action of $\mathfrak{g}$ on itself is via adjoint representation:

$$\rhd_{\alpha\beta}{}^\gamma = f_{\alpha\beta}{}^\gamma; \tag{A.3}$$

(iii) The action of $\mathfrak{g}$ on $\mathfrak{h}$ and $\mathfrak{l}$ is equivariant, i.e. the following identities are satisfied:

$$\partial_a{}^\beta f_{\alpha\beta}{}^\gamma = \rhd_{\alpha a}{}^b \partial_b{}^\gamma, \qquad \delta_A{}^a \rhd_{\alpha a}{}^b = \rhd_{\alpha A}{}^B \delta_B{}^b; \tag{A.4}$$

(iv) The Peiffer lifting is $\mathfrak{g}$-equivariant, i.e. for each $\underline{h}_1, \underline{h}_2 \in \mathfrak{h}$ and $\underline{g} \in \mathfrak{g}$:

$$\underline{g} \rhd \{\underline{h}_1, \underline{h}_2\}_{\mathrm{pf}} = \{\underline{g} \rhd \underline{h}_1, \underline{h}_2\}_{\mathrm{pf}} + \{\underline{h}_1,\, \underline{g} \rhd \underline{h}_2\}_{\mathrm{pf}},$$

or written in the basis:

$$X_{ab}{}^{B} \triangleright_{\alpha B}{}^{A} = \triangleright_{\alpha a}{}^{c} X_{cb}{}^{A} + \triangleright_{\alpha b}{}^{c} X_{ac}{}^{A}; \tag{A.5}$$

(v) $\delta(\{\underline{h}_1, \underline{h}_2\}_{\mathrm{pf}}) = \langle \underline{h}_1, \underline{h}_2 \rangle_{\mathrm{p}}, \ \forall \ \underline{h}_1, \underline{h}_2 \in \mathfrak{h}$.

The map $(\underline{h}_1, \underline{h}_2) \in \mathfrak{h} \times \mathfrak{h} \to \langle \underline{h}_1, \underline{h}_2 \rangle_{\mathrm{p}} \in \mathfrak{h}$ is bilinear $\mathfrak{g}$-equivariant map called the **Peiffer paring**, i.e. all $\underline{h}_1, \underline{h}_2 \in \mathfrak{h}$ and $\underline{g} \in \mathfrak{g}$ satisfy the following identity:

$$\underline{g} \triangleright \langle \underline{h}_1, \underline{h}_2 \rangle_{\mathrm{p}} = \langle \underline{g} \triangleright \underline{h}_1, \underline{h}_2 \rangle_{\mathrm{p}} + \langle \underline{h}_1, \underline{g} \triangleright \underline{h}_2 \rangle_{\mathrm{p}}.$$

Fixing the basis the identity becomes:

$$X_{ab}{}^{A} \delta_{A}{}^{c} = f_{ab}{}^{c} - \partial_{a}{}^{\alpha} \triangleright_{\alpha b}{}^{c}; \tag{A.6}$$

(vi) $[\underline{l}_1, \underline{l}_2] = \{\delta(\underline{l}_1), \delta(\underline{l}_2)\}_{\mathrm{pf}}, \ \forall \ \underline{l}_1, \underline{l}_2 \in \mathfrak{l}$, i.e.

$$f_{AB}{}^{C} = \delta_{A}{}^{a} \delta_{B}{}^{b} X_{ab}{}^{C}; \tag{A.7}$$

(vii) $\{[\underline{h}_1, \underline{h}_2], \underline{h}_3\}_{\mathrm{pf}} = \partial(\underline{h}_1) \triangleright \{\underline{h}_2, \underline{h}_3\}_{\mathrm{pf}} + \{\underline{h}_1, [\underline{h}_2, \underline{h}_3]\}_{\mathrm{pf}} - \partial(\underline{h}_2) \triangleright \{\underline{h}_1, \underline{h}_3\}_{\mathrm{pf}} - \{\underline{h}_2, [\underline{h}_1, \underline{h}_3]\}_{\mathrm{pf}}, \ \forall \ \underline{h}_1, \underline{h}_2, \underline{h}_3 \in \mathfrak{h}$, i.e.

$$\begin{aligned} \{[\underline{h}_1, \underline{h}_2], \underline{h}_3\}_{\mathrm{pf}} &= \{\partial(\underline{h}_1) \triangleright \underline{h}_2, \underline{h}_3\}_{\mathrm{pf}} - \{\partial(\underline{h}_2) \triangleright \underline{h}_1, \underline{h}_3\}_{\mathrm{pf}} \\ &\quad - \{\underline{h}_1, \delta\{\underline{h}_2, \underline{h}_3\}_{\mathrm{pf}}\}_{\mathrm{pf}} + \{\underline{h}_2, \delta\{\underline{h}_1, \underline{h}_3\}_{\mathrm{pf}}\}_{\mathrm{pf}}, \end{aligned} \tag{A.8}$$

$$f_{ab}{}^{d} X_{dc}{}^{B} = \partial_{a}{}^{\alpha} X_{bc}{}^{A} \triangleright_{\alpha A}{}^{B} + X_{ad}{}^{B} f_{bc}{}^{d} - \partial_{b}{}^{\alpha} \triangleright_{\alpha A}{}^{B} X_{ac}{}^{A} - X_{bd}{}^{B} f_{ac}{}^{d}; \tag{A.9}$$

(viii) $\{\underline{h}_1, [\underline{h}_2, \underline{h}_3]\}_{\mathrm{pf}} = \left\{\delta\{\underline{h}_1, \underline{h}_2\}_{\mathrm{pf}}, \underline{h}_3\right\}_{\mathrm{pf}} - \left\{\delta\{\underline{h}_1, \underline{h}_3\}_{\mathrm{pf}}, \underline{h}_2\right\}_{\mathrm{pf}}, \ \forall \ \underline{h}_1, \underline{h}_2, \underline{h}_3 \in \mathfrak{h}$, i.e.

$$X_{ad}{}^{A} f_{bc}{}^{d} = X_{ab}{}^{B} \delta_{B}{}^{d} X_{dc}{}^{A} - X_{ac}{}^{B} \delta_{B}{}^{d} X_{db}{}^{A}; \tag{A.10}$$

(ix) $\{\delta(\underline{l}), \underline{h}\}_{\mathrm{pf}} + \{\underline{h}, \delta(\underline{l})\}_{\mathrm{pf}} = -\partial(\underline{h}) \triangleright \underline{l}, \ \forall \ \underline{l} \in \mathfrak{l}, \ \forall \ \underline{h} \in \mathfrak{h}$, i.e.

$$\delta_{A}{}^{a} X_{ab}{}^{B} + \delta_{A}{}^{a} X_{ba}{}^{B} = -\partial_{b}{}^{\alpha} \triangleright_{\alpha A}{}^{B}. \tag{A.11}$$

A reader interested in more details about three-groups is referred to [21, 30].

The structure constants satisfy the Jacobi identities

$$\begin{aligned} f_{\alpha\gamma}{}^{\delta} f_{\beta\epsilon}{}^{\gamma} &= 2 f_{\alpha[\beta|}{}^{\gamma} f_{\gamma|\epsilon]}{}^{\delta}, \\ f_{ad}{}^{c} f_{be}{}^{d} &= 2 f_{a[b|}{}^{d} f_{d|e]}{}^{c}, \\ f_{AD}{}^{C} f_{BE}{}^{D} &= 2 f_{A[B|}{}^{D} f_{D|E]}{}^{C}. \end{aligned} \tag{A.12}$$

Also, the following relations are useful:

$$f_{\beta\gamma}{}^{\alpha} \triangleright_{\alpha b}{}^{a} = 2 \triangleright_{[\beta|c}{}^{a} \triangleright_{|\gamma]b}{}^{c}, \qquad f_{\beta\gamma}{}^{\alpha} \triangleright_{\alpha B}{}^{A} = 2 \triangleright_{[\beta|C}{}^{A} \triangleright_{|\gamma]B}{}^{C}. \tag{A.13}$$

Appendix B. Additional relations of the constraint algebra

In this appendix the useful technical results used in the subsection 3.2 are given. First, since the secondary constraints, given by the equation (30), must be preserved during the evolution of the system, the consistency conditions of secondary constraints must be enforced. However, no tertiary constraints arise from these conditions, since one obtains the following PB:

$$
\begin{aligned}
\{\mathcal{S}(\mathcal{F})^{\alpha i}, H_{\mathrm{T}}\} &= f_{\beta\gamma}{}^{\alpha}\mathcal{S}(\mathcal{F})^{\beta i}\alpha^{\gamma}{}_{0},\\
\{\mathcal{S}(\nabla B)_{\alpha}, H_{\mathrm{T}}\} &= f_{\beta\gamma\alpha}B^{\gamma}{}_{0k}\mathcal{S}(\mathcal{F})^{\beta k} + f_{\beta\alpha}{}^{\gamma}\alpha^{\beta}{}_{0}\mathcal{S}(\nabla B)_{\gamma} + C_{a0}\rhd_{\alpha b}{}^{a}\mathcal{S}(\mathcal{G})^{b}\\
&\quad - \rhd_{\alpha a}{}^{b}\beta^{a}{}_{0k}\mathcal{S}(\nabla C)_{b}{}^{k} + \frac{1}{2}\rhd_{\alpha}{}^{B}{}_{A}\gamma^{A}{}_{0jk}\mathcal{S}(\nabla D)_{B}{}^{jk},\\
\{\mathcal{S}(\mathcal{G})^{a}, H_{\mathrm{T}}\} &= \rhd_{\alpha b}{}^{a}\beta^{b}{}_{0k}\mathcal{S}(\mathcal{F})^{\alpha k} - \alpha^{\alpha}{}_{0}\rhd_{\alpha b}{}^{a}\mathcal{S}(\mathcal{G})^{b},\\
\{\mathcal{S}(\nabla C)_{a}{}^{i}, H_{\mathrm{T}}\} &= C_{b0}\rhd_{\alpha}{}^{b}{}_{a}\mathcal{S}(\mathcal{F})^{\alpha i} + \rhd_{\alpha a}{}^{b}\alpha^{\alpha}{}_{0}\mathcal{S}(\nabla C)_{b}{}^{i} + 2X_{(ab)}{}^{A}\beta^{b}{}_{0j}\mathcal{S}(\nabla D)_{A}{}^{ij},\\
\{\mathcal{S}(\nabla D)_{A}{}^{ij}, H_{\mathrm{T}}\} &= \alpha^{\alpha}{}_{0}\rhd_{\alpha A}{}^{B}\mathcal{S}(\nabla D)_{B}{}^{ij}.
\end{aligned}
\tag{B.1}
$$

The PB between the first-class constraints, given by the equation (34), and the second-class constraints, given by the equation (35), are given by:

$$
\begin{aligned}
\{\, \Phi(\mathcal{F})^{\alpha i}(\vec{x})\,,\, \chi(\alpha)_{\beta j}(\vec{y})\,\} &= -f_{\beta\gamma}{}^{\alpha}\, \chi(B)^{\gamma ij}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}),\\
\{\, \Phi(\mathcal{G})^{a}(\vec{x})\,,\, \chi(\alpha)_{\alpha}{}^{i}(\vec{y})\,\} &= \rhd_{\alpha b}{}^{a}\, \chi(C)^{bi}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}),\\
\{\, \Phi(\mathcal{G})^{a}(\vec{x})\,,\, \chi(\beta)_{b}{}^{ij}(\vec{y})\,\} &= -\rhd_{\alpha b}{}^{a}\, \chi(B)^{\alpha ij}(\vec{x})\,\delta^{(3)}(x-y),\\
\{\, \Phi(\nabla C)^{ai}(\vec{x})\,,\, \chi(\alpha)_{\alpha}{}^{j}(\vec{y})\,\} &= -\rhd_{\alpha b}{}^{a}\, \chi(\beta)^{bij}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}),\\
\{\, \Phi(\nabla C)^{ai}(\vec{x})\,,\, \chi(\beta)_{b}{}^{jk}(\vec{y})\,\} &= 2X^{(ac)A}g_{bc}\, \chi(\gamma)_{A}{}^{ijk}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}),\\
\{\, \Phi(\nabla C)^{ai}(\vec{x})\,,\, \chi(C)_{b}{}^{j}(\vec{y})\,\} &= \rhd_{\alpha b}{}^{a}\, \chi(B)^{\alpha ij}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}),\\
\{\, \Phi(\nabla C)^{ai}(\vec{x})\,,\, \chi(D)_{A}(\vec{y})\,\} &= 2X^{(ab)}{}_{A}\, \chi(C)_{b}{}^{i}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}),\\
\{\, \Phi(\nabla B)^{\alpha}(\vec{x})\,,\, \chi(\alpha)_{\beta}{}^{i}(\vec{y})\,\} &= f_{\beta\gamma}{}^{\alpha}\, \chi(\alpha)^{\gamma i}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}),\\
\{\, \Phi(\nabla B)^{\alpha}(\vec{x})\,,\, \chi(\beta)_{a}{}^{ij}(\vec{y})\,\} &= g^{\alpha\beta}\rhd_{\beta a}{}^{b}\, \chi(\beta)_{b}{}^{ij}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}),\\
\{\, \Phi(\nabla B)^{\alpha}(\vec{x})\,,\, \chi(\gamma)_{A}{}^{ijk}(\vec{y})\,\} &= g^{\alpha\beta}\rhd_{\beta A}{}^{B}\, \chi(\gamma)_{B}{}^{ijk}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}),\\
\{\, \Phi(\nabla B)^{\alpha}(\vec{x})\,,\, \chi(B)_{\beta}{}^{ij}(\vec{y})\,\} &= f\beta\gamma^{\alpha}\, \chi(B)^{\gamma ij}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}).\\
\{\, \Phi(\nabla B)^{\alpha}(\vec{x})\,,\, \chi(C)_{a}{}^{i}(\vec{y})\,\} &= -\rhd^{\alpha b}{}_{a}\, \chi(C)_{b}{}^{i}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}).\\
\{\, \Phi(\nabla B)^{\alpha}(\vec{x})\,,\, \chi(D)_{A}(\vec{y})\,\} &= g^{\alpha\beta}\rhd_{\beta A}{}^{B}\, \chi(D)_{B}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}),\\
\{\, \Phi(\nabla D)^{Aij}(\vec{x})\,,\, \chi(\alpha)_{\alpha}{}^{k}(\vec{y})\,\} &= \rhd_{\alpha B}{}^{A}\chi(\gamma)^{Bijk}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}),\\
\{\, \Phi(\nabla D)^{Aij}(\vec{x})\,,\, \chi(D)_{B}(\vec{y})\,\} &= -\rhd_{\alpha B}{}^{A}\chi(B)^{\alpha ij}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}).
\end{aligned}
\tag{B.2}
$$

Finally, it is useful to calculate PB between the first-class constraints, given by the equation (34), and the total Hamiltonian, given by the equation (33):

$$
\begin{aligned}
\{\Phi(\mathcal{F})^{\alpha i},\, H_{\mathrm{T}}\} &= f_{\beta\gamma}{}^{\alpha}\Phi(\mathcal{F})^{\beta i}\alpha^{\gamma}{}_{0},\\
\{\Phi(\nabla B)_{\alpha},\, H_{\mathrm{T}}\} &= f_{\beta\gamma\alpha}B^{\gamma}{}_{0k}\Phi(\mathcal{F})^{\beta k} + f_{\beta\alpha}{}^{\gamma}\alpha^{\beta}{}_{0}\Phi(\nabla B)_{\gamma} + C_{a0}\triangleright_{\alpha b}{}^{a}\Phi(\mathcal{G})^{b}\\
&\quad - \triangleright_{\alpha a}{}^{b}\beta^{a}{}_{0k}\Phi(\nabla C)_{b}{}^{k} + \frac{1}{2}\triangleright_{\alpha}{}^{B}{}_{A}\gamma^{A}{}_{0jk}\Phi(\nabla D)_{B}{}^{jk},\\
\{\Phi(\mathcal{G})^{a},\, H_{\mathrm{T}}\} &= \triangleright_{\alpha b}{}^{a}\beta^{b}{}_{0k}\Phi(\mathcal{F})^{\alpha k} - \alpha^{\alpha}{}_{0}\triangleright_{\alpha b}{}^{a}\Phi(\mathcal{G})^{b},\\
\{\Phi(\nabla C)_{a}{}^{i},\, H_{\mathrm{T}}\} &= C_{b0}\triangleright_{\alpha}{}^{b}{}_{a}\Phi(\mathcal{F})^{\alpha i} + \triangleright_{\alpha a}{}^{b}\alpha^{\alpha}{}_{0}\Phi(\nabla C)_{b}{}^{i} + 2X_{(ab)}{}^{A}\beta^{b}{}_{0j}\Phi(\nabla D)_{A}{}^{ij},\\
\{\Phi(\nabla D)_{A}{}^{ij},\, H_{\mathrm{T}}\} &= \alpha^{\alpha}{}_{0}\triangleright_{\alpha A}{}^{B}\Phi(\nabla D)_{B}{}^{ij}.
\end{aligned}
\tag{B.3}
$$

The calculated PB brackets given by the equation (B.3) will be useful for calculation of the generator of gauge symmetries (55). With these results one can proceed to the construction of the gauge symmetry generator on one hypersurface Σ_3 given in the equation (55), and ultimately obtain the finite gauge symmetry of the whole spacetime.

The PB algebra of gauge symmetry generators $(\tilde{M}_0)_{\alpha}{}^{i}$, $(\tilde{M}_1)_{\alpha}{}^{i}$, $(\tilde{G}_0)_{\alpha}$, $(\tilde{G}_1)_{\alpha}$, $(\tilde{H}_0)_{a}{}^{i}$, $(\tilde{H}_1)_{a}{}^{i}$, $(\tilde{N}_0)_{a}$, $(\tilde{N}_1)_{a}$, $(\tilde{L}_0)_{A}{}^{ij}$, and $(\tilde{L}_1)_{A}{}^{ij}$, as defined in (56), is:

$$\{(\tilde{G}_0)_{\alpha}(\vec{x}), (\tilde{G}_0)_{\beta}(\vec{y})\} = f_{\alpha\beta}{}^{\gamma}(\tilde{G}_0)_{\gamma}\,\delta^{(3)}(\vec{x}-\vec{y}), \tag{B.4}$$

$$\{(\tilde{H}_0)_{a}{}^{i}(\vec{x}), (\tilde{H}_0)_{b}{}^{j}(\vec{y})\} = 2X_{(ab)}{}^{A}(\tilde{L}_0)_{A}{}^{ij}\,\delta^{(3)}(\vec{x}-\vec{y}), \tag{B.5}$$

$$\{(\tilde{H}_0)_{a}{}^{i}(\vec{x}), (\tilde{H}_1)_{b}{}^{j}(\vec{y})\} = 2X_{(ab)}{}^{A}(\tilde{L}_1)_{A}{}^{ij}\,\delta^{(3)}(\vec{x}-\vec{y}), \tag{B.6}$$

$$\{(\tilde{H}_0)_{a}{}^{i}(\vec{x}), (\tilde{N}_0)^{b}(\vec{y})\} = \triangleright_{\alpha a}{}^{b}(\tilde{M}_0)^{\alpha i}\,\delta^{(3)}(\vec{x}-\vec{y}), \tag{B.7}$$

$$\{(\tilde{H}_1)_{a}{}^{i}(\vec{x}), (\tilde{N}_0)^{b}(\vec{y})\} = \triangleright_{\alpha a}{}^{b}(\tilde{M}_1)^{\alpha i}\,\delta^{(3)}(\vec{x}-\vec{y}), \tag{B.8}$$

$$\{(\tilde{H}_0)_{a}(\vec{x}), (\tilde{N}_1)^{bi}(\vec{y})\} = \triangleright_{\alpha a}{}^{b}(\tilde{M}_1)^{\alpha i}\,\delta^{(3)}(\vec{x}-\vec{y}), \tag{B.9}$$

$$\{(\tilde{G}_0)_{\alpha}(\vec{x}), (\tilde{M}_0)_{\beta}{}^{i}(\vec{y})\} = f_{\alpha\beta}{}^{\gamma}(\tilde{M}_0)_{\gamma}{}^{i}\,\delta^{(3)}(\vec{x}-\vec{y}), \tag{B.10}$$

$$\{(\tilde{G}_0)_{\alpha}(\vec{x}), (\tilde{M}_1)_{\beta}{}^{i}(\vec{y})\} = f_{\alpha\beta}{}^{\gamma}(\tilde{M}_1)_{\gamma}{}^{i}\,\delta^{(3)}(\vec{x}-\vec{y}), \tag{B.11}$$

$$\{(\tilde{G}_0)_{\alpha}(\vec{x}),\, (\tilde{H}_1)_{a}{}^{i}(\vec{y})\} = \triangleright_{\alpha a}{}^{b}\,(\tilde{H}_1)_{b}{}^{i}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}), \tag{B.12}$$

$$\{(\tilde{G}_0)_{\alpha}(\vec{x}),\, (\tilde{H}_0)_{a}{}^{i}(\vec{y})\} = \triangleright_{\alpha a}{}^{b}\,(\tilde{H}_0)_{b}{}^{i}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}), \tag{B.13}$$

$$\{(\tilde{G}_0)_{\alpha}(\vec{x}),\, (\tilde{N}_1)_{a}(\vec{y})\} = \triangleright_{\alpha a}{}^{b}\,(\tilde{N}_1)_{b}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}), \tag{B.14}$$

$$\{(\tilde{G}_0)_{\alpha}(\vec{x}),\, (\tilde{N}_0)_{a}(\vec{y})\} = \triangleright_{\alpha a}{}^{b}\,(\tilde{N}_0)_{b}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}), \tag{B.15}$$

$$\{(\tilde{G}_0)_{\alpha}(\vec{x}),\, (\tilde{L}_0)_{A}{}^{ij}(\vec{y})\} = \triangleright_{\alpha A}{}^{B}(\tilde{L}_0)_{B}{}^{ij}(\vec{x})\,\delta^{(3)}(\vec{x}-\vec{y}). \tag{B.16}$$

The gauge symmetry group has the following structure. First, the groups $\tilde{M}_1 \times \tilde{M}_0$, $\tilde{N}_1 \times \tilde{N}_0$ and $\tilde{L}_1 \times \tilde{L}_0$ with the corresponding algebras $\mathfrak{a}_1$, $\mathfrak{a}_2$ and $\mathfrak{a}_3$, respectively, where:

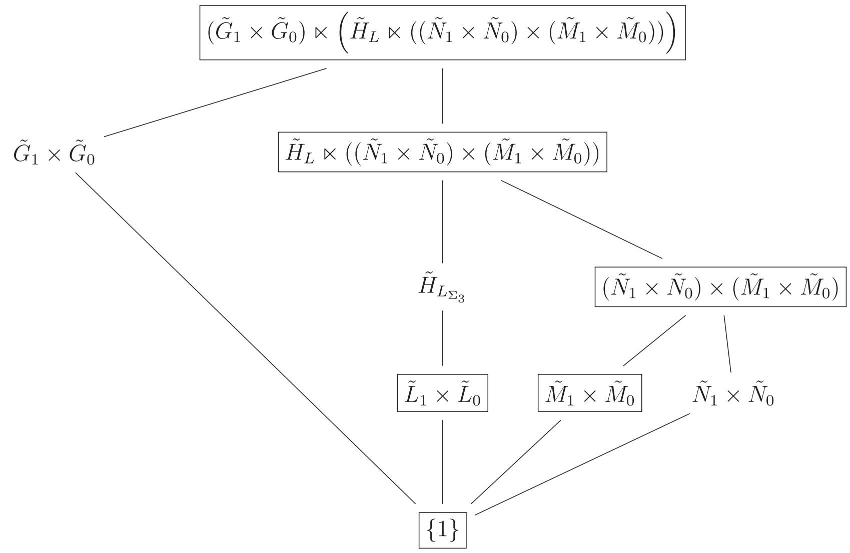


Figure B1. The symmetry group $\mathcal{G}_{\Sigma_3}$ of the Poisson bracket algebra in the phase space. The invariant subgroups are boxed.

$$
\begin{aligned}
\mathfrak{a}_1 &= \text{span}\{(\tilde{M}_1)_\alpha{}^i\} \oplus \text{span}\{(\tilde{M}_0)_\alpha{}^i\}, \\
\mathfrak{a}_2 &= \text{span}\{(\tilde{N}_1)_a\} \oplus \text{span}\{(\tilde{N}_0)_a\}, \\
\mathfrak{a}_3 &= \text{span}\{(\tilde{L}_1)_A{}^{ij}\} \oplus \text{span}\{(\tilde{L}_0)_A{}^{ij}\},
\end{aligned}
\tag{B.17}
$$

are the subgroups of the full symmetry group $\tilde{\mathcal{G}}_{\Sigma_3}$. Besides, the subgroups $\tilde{L}_1 \times \tilde{L}_0$ and $\tilde{M}_1 \times \tilde{M}_0$ are the invariant subgroups. The group $\tilde{N}_1 \times \tilde{N}_0$ is not an invariant subgroup of the whole symmetry group, since the Poisson brackets $\{(\tilde{H}_0)_a{}^i(\vec{x}), (\tilde{N}_0)^b(\vec{y})\}$ and $\{(\tilde{H}_1)_a{}^i(\vec{x}), (\tilde{N}_0)^b(\vec{y})\}$ are equal to some linear combinations of the generators of $\tilde{M}_1 \times \tilde{M}_0$. Nevertheless, one can form a direct product $(\tilde{N}_1 \times \tilde{N}_0) \times (\tilde{M}_1 \times \tilde{M}_0)$, since the generators of these groups are mutually commuting, giving a group which is an invariant subgroup of the complete symmetry group.

Next, consider a subgroup $\tilde{H}_{L_{\Sigma_3}}$ determined by the algebra spanned by the generators $(\tilde{L}_1)_A{}^{ij}$, $(\tilde{L}_0)_A{}^{ij}$, $(\tilde{H}_1)_a{}^i$, and $(\tilde{H}_0)_a{}^i$. This group is not invariant subgroup of the whole symmetry group, because of the PB $\{(\tilde{H}_0)_a{}^i(\vec{x}), (\tilde{N}_0)^b(\vec{y})\}$ and $\{(\tilde{H}_1)_a{}^i(\vec{x}), (\tilde{N}_0)^b(\vec{y})\}$, due to the same argument as before. Now, one can join these two subgroups, of which one is invariant and one is not, using a semidirect product into an invariant subgroup $H_L \ltimes ((N_1 \times N_0) \times (M_1 \times M_0))$, determined by the algebra $\mathfrak{a}_4$:

$$
\mathfrak{a}_4 = \text{span}\{(\tilde{M}_0)_\alpha{}^i, (\tilde{M}_1)_\alpha{}^i, (\tilde{H}_0)_a{}^i, (\tilde{H}_1)_a{}^i, (\tilde{N}_0)_a, (\tilde{N}_1)_a, (\tilde{L}_0)_A{}^{ij}, (\tilde{L}_1)_A{}^{ij}\}.
$$

Finally, following the same line of reasoning, one adds the group $\tilde{G}_1 \times \tilde{G}_0$ and obtains the full gauge symmetry group $\tilde{\mathcal{G}}_{\Sigma_3}$ to be equal to:

$$\tilde{\mathcal{G}}_{\Sigma_3} = (\tilde{G}_1 \times \tilde{G}_0) \ltimes \left(\tilde{H}_L \ltimes ((\tilde{N}_1 \times \tilde{N}_0) \times (\tilde{M}_1 \times \tilde{M}_0))\right).$$

The complete symmetry group structure is shown in the figure B1 appendix B. Here, the invariant subgroups of the whole symmetry group are boxed.

Appendix C. Construction of the symmetry generator

When one substitutes the generators (56) into the equation (55), one obtains the gauge generator of the 3*BF* theory in the following form

$$\begin{aligned} G = -\int_{\Sigma_3} \mathrm{d}^3\vec{x} \Bigg(& (\nabla_0 \epsilon_{\mathfrak{m}}{}^{\alpha}{}_i)\Phi(B)_\alpha{}^i - \epsilon_{\mathfrak{m}}{}^{\alpha}{}_i \Phi(\mathcal{F})_\alpha{}^i + (\nabla_0 \epsilon_{\mathfrak{g}}{}^{\alpha})\Phi(\alpha)_\alpha \\ & + \epsilon_{\mathfrak{g}}{}^{\alpha} \Big(f_{\alpha\gamma}{}^{\beta} B_{\beta 0 i}\Phi(B)^{\gamma i} + C_{a0}\rhd_{\alpha b}{}^a \Phi(C)^{b0} + \beta_{a0i}\rhd_{\alpha b}{}^a \Phi(\beta)^{b0i} \\ & - \frac{1}{2}\gamma^A{}_{0ij}\rhd_{\alpha A}{}^B \Phi(\gamma)_B{}^{ij} - \Phi(\nabla B)_\alpha \Big) + (\nabla_0 \epsilon_{\mathfrak{n}}{}^a)\Phi(C)_a \\ & - \epsilon_{\mathfrak{n}}{}^a \left(\beta_{b0i}\rhd_{\alpha a}{}^b \Phi(B)^{\alpha i} + \Phi(\mathcal{G})_a + (\nabla_0 \epsilon_{\mathfrak{h}}{}^a{}_i)\Phi(\beta)_a{}^i \right) \\ & - \epsilon_{\mathfrak{h}}{}^a{}_i \Big(C_{b0}\rhd_{\alpha a}{}^b \Phi(B)^{\alpha i} - 2\beta^b{}_{0j} X_{(ab)}{}^A \Phi(\gamma)_A{}^{ij} + \Phi(\nabla C)_a{}^i \Big) \\ & - \frac{1}{2}(\nabla_0 \epsilon_{\mathfrak{l}}{}^A{}_{ij})\Phi(\gamma)_A{}^{ij} + \frac{1}{2}\epsilon_{\mathfrak{l}}{}^A{}_{ij}\Phi(\nabla D)_A{}^{ij} \Bigg), \end{aligned} \tag{C.1}$$

where $\epsilon_{\mathfrak{g}}{}^{\alpha}$, $\epsilon_{\mathfrak{h}i}{}^{a}$, $\epsilon^A_{\mathfrak{l}ij}$, $\epsilon_{\mathfrak{m}i}{}^{\alpha}$, and $\epsilon_{\mathfrak{n}}{}^a$ are the independent parameters of the gauge transformations.

The generator of gauge transformations (C.1) in 3*BF* theory given by the action (15), is obtained by the Castellani's procedure, requiring the following requirements to be met

$$G_1 = C_{\mathrm{PFC}}, \tag{C.2}$$

$$G_0 + \{G_1, H_{\mathrm{T}}\} = C_{\mathrm{PFC}}, \tag{C.3}$$

$$\{G_0, H_{\mathrm{T}}\} = C_{\mathrm{PFC}}, \tag{C.4}$$

where C_{PFC} denotes some first-class constraints, and assuming that the generator has the following structure:

$$\begin{aligned} G = \int_{\Sigma_3} \mathrm{d}^3\vec{x} \Bigg(& \dot{\epsilon}_{\mathfrak{m}}{}^{\alpha}{}_i (G_1)_{\mathfrak{m}\alpha}{}^i + \epsilon_{\mathfrak{m}}{}^{\alpha}{}_i (G_0)_{\mathfrak{m}\alpha}{}^i + \dot{\epsilon}_{\mathfrak{g}}{}^{\alpha}(G_1)_{\mathfrak{g}\alpha} + \epsilon_{\mathfrak{g}}{}^{\alpha}(G_0)_{\mathfrak{g}\alpha} \\ & + \dot{\epsilon}_{\mathfrak{h}}{}^a{}_i (G_1)_{\mathfrak{h}a}{}^i + \epsilon_{\mathfrak{h}}{}^a{}_i (G_0)_{\mathfrak{h}a}{}^i + \dot{\epsilon}_{\mathfrak{n}}{}^a (G_1)_{\mathfrak{n}a} + \epsilon_{\mathfrak{n}}{}^a (G_0)_{\mathfrak{n}a} \\ & + \frac{1}{2}\dot{\epsilon}_{\mathfrak{l}}{}^A{}_{ij}(G_1)_{\mathfrak{l}A}{}^{ij} + \frac{1}{2}\epsilon_{\mathfrak{l}}{}^A{}_{ij}(G_0)_{\mathfrak{l}A}{}^{ij} \Bigg). \end{aligned} \tag{C.5}$$

The first step of Castellani's procedure, imposing the set of conditions

$$\begin{aligned}(G_1)_{\mathfrak{m}\alpha}{}^{i} &= C_{\text{PFC}},\\ (G_1)_{\mathfrak{g}\alpha} &= C_{\text{PFC}},\\ (G_1)_{\mathfrak{h}a}{}^{i} &= C_{\text{PFC}},\\ (G_1)_{\mathfrak{n}a} &= C_{\text{PFC}},\\ (G_1)_{\mathfrak{l}A}{}^{ij} &= C_{\text{PFC}},\end{aligned} \tag{C.6}$$

is satisfied with a natural choice:

$$\begin{aligned}(G_1)_{\mathfrak{m}\alpha}{}^{i} &= -\Phi(B)_{\alpha}{}^{i},\\ (G_1)_{\mathfrak{g}\alpha} &= -\Phi(\alpha)_{\alpha},\\ (G_1)_{\mathfrak{h}a}{}^{i} &= -\Phi(C)_{\alpha}{}^{i},\\ (G_1)_{\mathfrak{n}a} &= -\Phi(\beta)_{a},\\ (G_1)_{\mathfrak{l}A}{}^{ij} &= \Phi(\gamma)_{A}{}^{ij}.\end{aligned} \tag{C.7}$$

It remains to determine the five generators G_0.

The Castellani's second condition for the generator $(G_0)_{\mathfrak{m}\alpha}{}^{i}$ gives:

$$\begin{aligned}(G_0)_{\mathfrak{m}\alpha}{}^{i} - \{\Phi(B)_{\alpha}{}^{i}, H_{\text{T}}\} &= (C_{\text{PFC}})_{\alpha}{}^{i},\\ (G_0)_{\mathfrak{m}\alpha}{}^{i} - \Phi(\mathcal{F})_{\alpha}{}^{i} &= (C_{\text{PFC}})_{\alpha}{}^{i},\end{aligned} \tag{C.8}$$

that is $(G_0)_{\mathfrak{m}\alpha}{}^{i} = (C_{\text{PFC}})_{\alpha}{}^{i} + \Phi(\mathcal{F})_{\alpha}{}^{i}$. Subsequently, from the Castellani's third condition it follows

$$\begin{aligned}&\{(G_0)_{\mathfrak{m}\alpha}{}^{i}, H_{\text{T}}\} = (C_{\text{PFC1}})_{\alpha}{}^{i},\\ &\{(C_{\text{PFC}})_{\alpha}{}^{i} + \Phi(\mathcal{F})_{\alpha}{}^{i}, H_{\text{T}}\} = (C_{\text{PFC1}})_{\alpha}{}^{i},\\ &\{(C_{\text{PFC}})_{\alpha}{}^{i}, H_{\text{T}}\} - f_{\beta\gamma\alpha}\alpha^{\beta}{}_{0}\Phi(\mathcal{F})^{\gamma i} = (C_{\text{PFC1}})_{\alpha}{}^{i},\end{aligned} \tag{C.9}$$

which gives

$$(C_{\text{PFC}})_{\alpha}{}^{i} = f_{\beta\gamma\alpha}\alpha^{\beta}{}_{0}\Phi(B)^{\gamma i}.$$

It follows that the generator is:

$$(G_0)_{\mathfrak{m}\alpha}{}^{i} = f_{\beta\gamma\alpha}\alpha^{\beta}{}_{0}\Phi(B)^{\gamma i} + \Phi(\mathcal{F})_{\alpha}{}^{i}. \tag{C.10}$$

The Castellani's second condition for the generator $(G_0)_{\mathfrak{g}\alpha}$ gives:

$$\begin{aligned}(G_0)_{\mathfrak{g}\alpha} - \{\Phi(\alpha)_{\alpha}, H_{\text{T}}\} &= (C_{\text{PFC}})_{\alpha},\\ (G_0)_{\mathfrak{g}\alpha} - \Phi(\nabla B)_{\alpha} &= (C_{\text{PFC}})_{\alpha},\end{aligned} \tag{C.11}$$

that is $(G_0)_{\mathfrak{g}\alpha} = (C_{\text{PFC}})_\alpha + \Phi(\nabla B)_\alpha$. Subsequently, from the Castellani's third condition it follows

$$\begin{aligned}
&\{(G_0)_{\mathfrak{g}\alpha}, H_{\text{T}}\} = (C_{\text{PFC1}})_\alpha, \\
&\{(C_{\text{PFC}})_\alpha + \Phi(\nabla B)_\alpha, H_{\text{T}}\} = (C_{\text{PFC1}})_\alpha, \\
&\{(C_{\text{PFC}})_\alpha, H_{\text{T}}\} + B_{\beta 0i} f_{\alpha\gamma}{}^\beta \Phi(\mathcal{F})^{\gamma i} - \alpha^\beta{}_0 f_{\alpha\beta}{}^\gamma \Phi(\nabla B)_\gamma + C_{a0} \triangleright_{\alpha b}{}^a \Phi(\mathcal{G})^b \\
&\qquad + \beta_{a0i} \triangleright_{\alpha b}{}^a \Phi(\nabla C)^{bi} - \frac{1}{2}\gamma^A{}_{0ij} \triangleright_{\alpha A}{}^B \Phi(\nabla D)_B{}^{ij} = (C_{\text{PFC1}})_\alpha,
\end{aligned} \tag{C.12}$$

which gives

$$\begin{aligned}
(C_{\text{PFC}})_\alpha = &-B_{\beta 0i} f_{\alpha\gamma}{}^\beta \Phi(B)^{\gamma i} + \alpha^\beta{}_0 f_{\alpha\beta}{}^\gamma \Phi(\alpha)_\gamma - C_{a0} \triangleright_{\alpha b}{}^a \Phi(C)^b \\
&- \beta_{a0i} \triangleright_{\alpha b}{}^a \Phi(\beta)^{bi} + \frac{1}{2}\gamma^A{}_{0ij} \triangleright_{\alpha A}{}^B \Phi(\gamma)_B{}^{ij}.
\end{aligned}$$

It follows that the generator is:

$$\begin{aligned}
(G_0)_{\mathfrak{g}\alpha} = &-B_{\beta 0i} f_{\alpha\gamma}{}^\beta \Phi(B)^{\gamma i} + \alpha^\beta{}_0 f_{\alpha\beta}{}^\gamma \Phi(\alpha)_\gamma - C_{a0} \triangleright_{\alpha b}{}^a \Phi(C)^b \\
&- \beta_{a0i} \triangleright_{\alpha b}{}^a \Phi(\beta)^{bi} + \frac{1}{2}\gamma^A{}_{0ij} \triangleright_{\alpha A}{}^B \Phi(\gamma)_B{}^{ij} + \Phi(\nabla B)_\alpha.
\end{aligned} \tag{C.13}$$

The Castellani's second condition for the generator $(G_0)_{\mathfrak{n}a}$ gives

$$\begin{aligned}
(G_0)_{\mathfrak{n}a} - \{\Phi(C)_a, H_{\text{T}}\} &= (C_{\text{PFC}})_a, \\
(G_0)_{\mathfrak{n}a} - \Phi(\mathcal{G})_a &= (C_{\text{PFC}})_a,
\end{aligned} \tag{C.14}$$

that is $(G_0)_{\mathfrak{n}a} = (C_{\text{PFC}})_a + \Phi(\mathcal{G})_a$. Subsequently, from the Castellani's third condition it follows

$$\begin{aligned}
&\{(G_0)_{\mathfrak{n}a}, H_{\text{T}}\} = (C_{\text{PFC1}})_a, \\
&\{(C_{\text{PFC}})_a + \Phi(\mathcal{G})_a, H_{\text{T}}\} = (C_{\text{PFC1}})_a, \\
&\{(C_{\text{PFC}})_a, H_{\text{T}}\} + \alpha^\alpha{}_0 \triangleright_{\alpha a}{}^b \Phi(\mathcal{G})_b - \beta_{b0i} \triangleright_{\alpha a}{}^b \Phi(\mathcal{F})^{\alpha i} = (C_{\text{PFC1}})_a,
\end{aligned} \tag{C.15}$$

which gives

$$(C_{\text{PFC}})_a = -\alpha^\alpha{}_0 \triangleright_{\alpha a}{}^b \Phi(C)_b + \beta_{b0i} \triangleright_{\alpha a}{}^b \Phi(B)^{\alpha i}.$$

It follows that the generator is:

$$(G_0)_{\mathfrak{n}a} = -\alpha^\alpha{}_0 \triangleright_{\alpha a}{}^b \Phi(C)_b + \beta_{b0i} \triangleright_{\alpha a}{}^b \Phi(B)^{\alpha i} + \Phi(\mathcal{G})_a.$$

The Castellani's second condition for the generator $(G_0)_{\mathfrak{h}a}{}^i$ gives:

$$\begin{aligned}
&(G_0)_{\mathfrak{h}a}{}^i - \{\Phi(\beta)_a{}^i, H_{\text{T}}\} = (C_{\text{PFC}})_a{}^i, \\
&(G_0)_{\mathfrak{h}a}{}^i - \Phi(\nabla C)_a{}^i = (C_{\text{PFC}})_a{}^i,
\end{aligned} \tag{C.16}$$

that is $(G_0)_{\mathfrak{h}a}{}^i = (C_{\rm PFC})_a{}^i + \Phi(\nabla C)_a{}^i$. Subsequently, from the Castellani's third condition it follows

$$
\begin{aligned}
&\{(G_0)_{\mathfrak{h}a}{}^i, H_{\rm T}\} = (C_{\rm PFC1})_a{}^i,\\
&\{(C_{\rm PFC})_a{}^i + \Phi(\nabla C)_a{}^i, H_{\rm T}\} = (C_{\rm PFC1})_a{}^i,\\
&\{(C_{\rm PFC})_a{}^i, H_{\rm T}\} + \alpha^\alpha{}_0 \triangleright_{\alpha a}{}^b \Phi(\nabla C)_b{}^i - C_{b0} \triangleright_{\alpha a}{}^b \Phi(\mathcal{F})^{\alpha i} + 2\beta^b{}_{0j} X_{(ab)}{}^A \Phi(\nabla D)_A{}^{ij} = (C_{\rm PFC1})_a{}^i,
\end{aligned}
$$

which gives

$$
(C_{\rm PFC})_a{}^i = -\alpha^\alpha{}_0 \triangleright_{\alpha a}{}^b \Phi(\beta)_b{}^i + C_{b0} \triangleright_{\alpha a}{}^b \Phi(B)^{\alpha i} - 2\beta^b{}_{0j} X_{(ab)}{}^A \Phi(\gamma)_A{}^{ij}.
$$

It follows that the generator is:

$$
(G_0)_{\mathfrak{h}a}{}^i = -\alpha^\alpha{}_0 \triangleright_{\alpha a}{}^b \Phi(\beta)_b{}^i + C_{b0} \triangleright_{\alpha a}{}^b \Phi(B)^{\alpha i} - 2\beta^b{}_{0j} X_{(ab)}{}^A \Phi(\gamma)_A{}^{ij} + \Phi(\nabla C)_a{}^i.
$$

The Castellani's second condition for the generator $(G_0)_{\mathfrak{l}A}{}^{ij}$ gives:

$$
\begin{aligned}
&(G_0)_{\mathfrak{l}A}{}^{ij} + \{\Phi(\gamma)_A{}^{ij}, H_{\rm T}\} = (C_{\rm PFC})_A{}^{ij},\\
&(G_0)_{\mathfrak{l}A}{}^{ij} + \Phi(\nabla D)_A{}^{ij} = (C_{\rm PFC})_A{}^{ij},
\end{aligned}
\tag{C.17}
$$

that is $(G_0)_{\mathfrak{l}A}{}^{ij} = (C_{\rm PFC})_A{}^{ij} - \Phi(\nabla D)_A{}^{ij}$. Subsequently, from the Castellani's third condition it follows:

$$
\begin{aligned}
&\{(G_0)_{\mathfrak{l}A}{}^{ij}, H_{\rm T}\} = (C_{\rm PFC1})_A{}^{ij},\\
&\{(C_{\rm PFC})_A{}^{ij} - \Phi(\nabla D)_A{}^{ij}, H_{\rm T}\} = (C_{\rm PFC1})_A{}^{ij},\\
&\{(C_{\rm PFC})_A{}^{ij}, H_{\rm T}\} - \alpha^\alpha{}_0 \triangleright_{\alpha A}{}^B \Phi(\nabla D)_B{}^{ij} = (C_{\rm PFC1})_A{}^{ij},
\end{aligned}
\tag{C.18}
$$

which gives

$$
(C_{\rm PFC})_A{}^{ij} = \alpha^\alpha{}_0 \triangleright_{\alpha A}{}^B \Phi(\gamma)_B{}^{ij}.
$$

It follows that the generator is:

$$
(G_0)_{\mathfrak{l}A}{}^{ij} = \alpha^\alpha{}_0 \triangleright_{\alpha A}{}^B \Phi(\gamma)_B{}^{ij} - \Phi(\nabla D)_A{}^{ij}.
\tag{C.19}
$$

At this point, it is useful to summarize the results, and introduce the new notation:

$$
\begin{aligned}
\dot{\epsilon}_{\mathfrak{m}}{}^\alpha{}_i (G_1)_{\mathfrak{m}\alpha}{}^i + \epsilon_{\mathfrak{m}}{}^\alpha{}_i (G_0)_{\mathfrak{m}\alpha}{}^i &= -\nabla_0 \epsilon_{\mathfrak{m}}{}^\alpha{}_i \Phi(B)_\alpha{}^i + \epsilon_{\mathfrak{m}}{}^\alpha{}_i \Phi(\mathcal{F})_\alpha{}^i\\
&= \nabla_0 \epsilon_{\mathfrak{m}}{}^\alpha{}_i (\tilde{M}_1)_\alpha{}^i + \epsilon_{\mathfrak{m}}{}^\alpha{}_i (\tilde{M}_0)_\alpha{}^i.
\end{aligned}
\tag{C.20}
$$

Note that the time derivative of the parameter combines with some of the other terms into a covariant derivative in the time directions.

For the second part of the total generator one obtains:

$$
\begin{aligned}
&{}^{\alpha}\dot{\epsilon}_{\mathfrak{g}}(G_1)_{\mathfrak{g}\alpha} + \epsilon_{\mathfrak{g}}{}^{\alpha}(G_0)_{\mathfrak{g}\alpha} \\
&\quad = -{}^{\alpha}\dot{\epsilon}_{\mathfrak{g}}\Phi(\alpha)_{\alpha} - \epsilon_{\mathfrak{g}}{}^{\alpha}\Bigg(B_{\beta 0 i} f_{\alpha\gamma}{}^{\beta}\Phi(B)^{\gamma i} - \alpha^{\beta}{}_{0} f_{\alpha\beta}{}^{\gamma}\Phi(\alpha)_{\gamma} + C_{a0}\rhd_{\alpha b}{}^{a}\Phi(C)^{b} \\
&\quad + \beta_{a0i}\rhd_{\alpha b}{}^{a}\Phi(\beta)_{b}{}^{i} - \frac{1}{2}\gamma^{A}{}_{0ij}\rhd_{\alpha A}{}^{B}\Phi(\gamma)_{B}{}^{ij} - \Phi(\nabla B)_{\alpha}\Bigg) \\
&\quad = -\nabla_0\epsilon_{\mathfrak{g}}{}^{\alpha}\Phi(\alpha)_{\alpha} - \epsilon_{\mathfrak{g}}{}^{\alpha}\Bigg(B_{\beta 0 i} f_{\alpha\gamma}{}^{\beta}\Phi(B)^{\gamma i} + C_{a0}\rhd_{\alpha b}{}^{a}\Phi(C)^{b} \\
&\quad + \beta_{a0i}\rhd_{\alpha b}{}^{a}\Phi(\beta)_{b}{}^{i} - \frac{1}{2}\gamma^{A}{}_{0ij}\rhd_{\alpha A}{}^{B}\Phi(\gamma)_{B}{}^{ij} - \Phi(\nabla B)_{\alpha}\Bigg) \\
&\quad = \nabla_0\epsilon_{\mathfrak{g}}{}^{\alpha}(\tilde{G}_1)_{\alpha} + \epsilon_{\mathfrak{g}}{}^{\alpha}(\tilde{G}_0)_{\alpha}.
\end{aligned}
\tag{C.21}
$$

Furthermore, it follows:

$$
\begin{aligned}
\dot{\epsilon}_{\mathfrak{h}i}^{a}(G_1)_{\mathfrak{h}a}{}^{i} + \epsilon_{\mathfrak{h}}{}^{a}{}_{i}(G_0)_{\mathfrak{h}a}{}^{i} &= -\nabla_0\epsilon_{\mathfrak{h}}{}^{a}{}_{i}\Phi(\beta)_{\alpha}{}^{i} + \epsilon_{\mathfrak{h}}{}^{a}{}_{i}\big(C_{b0}\rhd_{\alpha a}{}^{b}\Phi(B)^{\alpha i} \\
&\quad - 2\beta^{b}{}_{0j}X_{(ab)}{}^{A}\Phi(\gamma)_{A}{}^{ij} + \Phi(\nabla C)_{a}{}^{i}\big) \\
&= \nabla_0\epsilon_{\mathfrak{h}}{}^{a}{}_{i}(\tilde{H}_1)_{a}{}^{i} + \epsilon_{\mathfrak{h}}{}^{a}{}_{i}(\tilde{H}_0)_{a}{}^{i},
\end{aligned}
\tag{C.22}
$$

$$
\begin{aligned}
\dot{\epsilon}_{\mathfrak{n}}{}^{a}(G_1)_{\mathfrak{n}a} + \epsilon_{\mathfrak{n}}{}^{a}(G_0)_{\mathfrak{n}a} &= -\nabla_0\epsilon_{\mathfrak{n}}{}^{a}\Phi(C)_{a} + \epsilon_{\mathfrak{n}}{}^{a}(\beta_{b0i}\rhd_{\alpha a}{}^{b}\Phi(B)^{\alpha i} + \Phi(\mathcal{G})_{a}) \\
&= \nabla_0\epsilon_{\mathfrak{n}}{}^{a}(\tilde{N}_1)_{a} + \epsilon_{\mathfrak{n}}{}^{a}(\tilde{N}_0)_{a}.
\end{aligned}
\tag{C.23}
$$

Finally, one gets:

$$
\begin{aligned}
\frac{1}{2}\dot{\epsilon}_{\mathfrak{l}ij}^{A}(G_1)_{\mathfrak{l}A}{}^{ij} + \frac{1}{2}\epsilon_{\mathfrak{l}}^{A}{}_{ij}(G_0)_{\mathfrak{l}A}{}^{ij} &= \frac{1}{2}\dot{\epsilon}_{\mathfrak{l}ij}^{A}\Phi(\gamma)_{A}{}^{ij} + \frac{1}{2}\epsilon_{\mathfrak{l}}^{A}{}_{ij}\alpha^{\alpha}{}_{0}\rhd_{\alpha A}{}^{B}\Phi(\gamma)_{B}{}^{ij} \\
&\quad - \frac{1}{2}\epsilon_{\mathfrak{l}}^{A}{}_{ij}\Phi(\nabla D)_{A}{}^{ij} \\
&= \frac{1}{2}\nabla_0\epsilon_{\mathfrak{l}}^{A}{}_{ij}\Phi(\gamma)_{A}{}^{ij} - \frac{1}{2}\epsilon_{\mathfrak{l}}^{A}{}_{ij}\Phi(\nabla D)_{A}{}^{ij} \\
&= \frac{1}{2}\nabla_0\epsilon_{\mathfrak{l}}^{A}{}_{ij}(\tilde{L}_1)_{A}{}^{ij} + \frac{1}{2}\epsilon_{\mathfrak{l}}^{A}{}_{ij}(\tilde{L}_0)_{A}{}^{ij}.
\end{aligned}
\tag{C.24}
$$

Appendix D. Definitions of maps $\mathcal{T}$, $\mathcal{S}$, $\mathcal{D}$, $\mathcal{X}_1$, and $\mathcal{X}_2$

Given G-invariant symmetric non-degenerate bilinear forms in $\mathfrak{g}$ and $\mathfrak{h}$, one can define a bilinear antisymmetric map $\mathcal{T} : \mathfrak{h}\times\mathfrak{h}\to\mathfrak{g}$ by the rule:

$$
\langle \mathcal{T}(\underline{h}_1, \underline{h}_2), \underline{g}\rangle_{\mathfrak{g}} = -\langle \underline{h}_1, \underline{g}\rhd \underline{h}_2\rangle_{\mathfrak{h}}, \quad \forall\ \underline{h}_1, \underline{h}_2 \in \mathfrak{h},\ \forall\ \underline{g}\in\mathfrak{g}.
$$

Written in basis:

$$\mathcal{T}(t_a, t_b) = \mathcal{T}_{ab}{}^{\alpha} \tau_{\alpha},$$

where the components of the map $\mathcal{T}$ are:

$$\mathcal{T}_{ab}{}^{\alpha} = -g_{ac} \rhd_{\beta b}{}^{c} g^{\alpha\beta}.$$

See [26] for more properties and the construction of $2BF$ invariant topological action using this map.

The transformations of the Lagrange multipliers and the $3BF$ invariant topological action is defined via maps

$$\mathcal{S} : \mathfrak{l} \times \mathfrak{l} \to \mathfrak{g}, \qquad \mathcal{X}_1 : \mathfrak{l} \times \mathfrak{h} \to \mathfrak{h}, \qquad \mathcal{X}_2 : \mathfrak{l} \times \mathfrak{h} \to \mathfrak{h}, \quad \mathcal{D} : \mathfrak{h} \times \mathfrak{h} \times \mathfrak{l} \to \mathfrak{g},$$

as it is defined in [13]. The map $\mathcal{S} : \mathfrak{l} \times \mathfrak{l} \to \mathfrak{g}$ is defined by the rule:

$$\langle \mathcal{S}(\underline{l}_1, \underline{l}_2), \underline{g} \rangle_{\mathfrak{g}} = -\langle \underline{l}_1, \underline{g} \rhd \underline{l}_2 \rangle_{\mathfrak{l}}, \quad \forall\ \underline{l}_1,\ \forall\ \underline{l}_2 \in \mathfrak{l},\ \forall\ \underline{g} \in \mathfrak{g}.$$

Written in the basis:

$$\mathcal{S}(T_A, T_B) = \mathcal{S}_{AB}{}^{\alpha} \tau_{\alpha},$$

the defining relation for $\mathcal{S}$ becomes:

$$\mathcal{S}_{AB}{}^{\alpha} = -\rhd_{\beta[B}{}^{C}\ g_{A]C} g^{\alpha\beta}.$$

Given two $\mathfrak{l}$-valued forms η and ω, one can define a $\mathfrak{g}$-valued form:

$$\omega \wedge^{\mathcal{S}} \eta = \omega^A \wedge \eta^B \mathcal{S}_{AB}{}^{\alpha} \tau_{\alpha}.$$

Using this map, the transformations of the Lagrange multipliers under L-gauge are defined in [13].

Further, to define the transformations of the Lagrange multipliers under H-gauge transformations the bilinear map $\mathcal{X}_1 : \mathfrak{l} \times \mathfrak{h} \to \mathfrak{h}$ is defined:

$$\langle \mathcal{X}_1(\underline{l}, \underline{h}_1), \underline{h}_2 \rangle_{\mathfrak{h}} = -\langle \underline{l}, \{\underline{h}_1, \underline{h}_2\} \rangle_{\mathfrak{l}}, \quad \forall\ \underline{h}_1, \underline{h}_2 \in \mathfrak{h},\ \forall\ \underline{l} \in \mathfrak{l},$$

and bilinear map $\mathcal{X}_2 : \mathfrak{l} \times \mathfrak{h} \to \mathfrak{h}$ by the rule:

$$\langle \mathcal{X}_2(\underline{l}, \underline{h}_2), \underline{h}_1 \rangle_{\mathfrak{h}} = -\langle \underline{l}, \{\underline{h}_1, \underline{h}_2\} \rangle_{\mathfrak{l}}, \quad \forall\ \underline{h}_1, \underline{h}_2 \in \mathfrak{h},\ \forall\ \underline{l} \in \mathfrak{l}.$$

As far as the bilinear maps $\mathcal{X}_1$ and $\mathcal{X}_2$ one can define the coefficients in the basis as:

$$\mathcal{X}_1(T_A, t_a) = \mathcal{X}_{1\ Aa}{}^b\, t_b, \qquad \mathcal{X}_2(T_A, t_a) = \mathcal{X}_{2\ Aa}{}^b\, t_b.$$

When written in the basis the defining relations for the maps $\mathcal{X}_1$ and $\mathcal{X}_2$ become:

$$\mathcal{X}_{1Ab}{}^c = -X_{ba}{}^B g_{AB} g^{ac}, \qquad \mathcal{X}_{2Ab}{}^c = -X_{ab}{}^B g_{AB} g^{ac}.$$

Given $\mathfrak{l}$-valued differential form ω and $\mathfrak{h}$-valued differential form η, one defines a $\mathfrak{h}$-valued form as:

$$\omega \wedge^{\mathcal{X}_1} \eta = \omega^A \wedge \eta^a \mathcal{X}_{1Aa}{}^b t_b, \qquad \omega \wedge^{\mathcal{X}_2} \eta = \omega^A \wedge \eta^a \mathcal{X}_{2Aa}{}^b t_b.$$

Finally, a trilinear map $\mathcal{D} : \mathfrak{h} \times \mathfrak{h} \times \mathfrak{l} \to \mathfrak{g}$ is needed:

$$\langle \mathcal{D}(\underline{h}_1, \underline{h}_2, \underline{l}), \underline{g} \rangle_{\mathfrak{g}} = -\langle \underline{l}, \{\underline{g} \triangleright \underline{h}_1, \underline{h}_2\} \rangle_{\mathfrak{l}}, \quad \forall\ \underline{h}_1, \underline{h}_2 \in \mathfrak{h},\ \forall\ \underline{l} \in \mathfrak{l},\ \forall\ \underline{g} \in \mathfrak{g},$$

One can define the coefficients of the trilinear map as:

$$\mathcal{D}(t_a, t_b, T_A) = \mathcal{D}_{abA}{}^{\alpha} \tau_{\alpha},$$

and the defining relation for the map $\mathcal{D}$ expressed in terms of coefficients becomes:

$$\mathcal{D}_{abA}{}^{\beta} = -\triangleright_{\alpha a}{}^c X_{cb}{}^B g_{AB} g^{\alpha\beta}.$$

Given two $\mathfrak{h}$-valued forms ω and η, and $\mathfrak{l}$-valued form ξ, the g-valued form is given by the formula:

$$\omega \wedge^{\mathcal{D}} \eta \wedge^{\mathcal{D}} \xi = \omega^a \wedge \eta^b \wedge \xi^A \mathcal{D}_{abA}{}^{\beta} \tau_{\beta}.$$

With these maps in hand, the transformations of the Lagrange multipliers under H-gauge transformations are defined, see [13].

Appendix E. Form-variations of all fields and momenta

The obtained gauge generator (55) is employed to calculate the form variations of variables and their corresponding canonical momenta, denoted as $A(t, \vec{x})$, using the following equation,

$$\delta_0 A(t, \vec{x}) = \{A(t, \vec{x}), G\}. \tag{E.1}$$

The computed form variations are given as follows:

$$
\begin{aligned}
\delta_0 B^{\alpha}{}_{0i} &= -\nabla_0 \epsilon_{\mathfrak{m}i}{}^{\alpha} + f_{\beta\gamma}{}^{\alpha} \epsilon_{\mathfrak{g}}{}^{\beta} B^{\gamma}{}_{0i} \\
&\quad + \epsilon_{\mathfrak{n}}{}^{a} \rhd_{\alpha a}{}^{b} \beta_{b0i} + \epsilon_{\mathfrak{h}i}{}^{a} \rhd_{\alpha a}{}^{b} C_{b0}, \\
\delta_0 B^{\alpha}{}_{ij} &= -2\nabla_{[i|} \epsilon_{\mathfrak{m}|j]}{}^{\alpha} + f_{\beta\gamma}{}^{\alpha} \epsilon_{\mathfrak{g}}{}^{\beta} B^{\gamma}{}_{ij} - \epsilon_{\mathfrak{l}ij}{}^{A} \rhd_{\alpha A}{}^{B} D_B \\
&\quad + \epsilon_{\mathfrak{n}}{}^{a} \rhd_{\alpha a}{}^{b} \beta_{bij} + 2\epsilon_{\mathfrak{h}}{}^{a}{}_{[j|} \rhd_{\alpha a}{}^{b} C_{b|i]}, \\
\delta_0 \alpha^{\alpha}{}_{0} &= -\nabla_0 \epsilon_{\mathfrak{g}}{}^{\alpha}, \\
\delta_0 \alpha^{\alpha}{}_{i} &= -\nabla_i \epsilon_{\mathfrak{g}}{}^{\alpha} - \partial_a{}^{\alpha} \epsilon_{\mathfrak{h}i}{}^{a}, \\
\delta_0 C^{a}{}_{0} &= -\nabla_0 \epsilon_{\mathfrak{n}}{}^{a} + \epsilon_{\mathfrak{g}}{}^{\alpha} \rhd_{\alpha b}{}^{a} C^{b}{}_{0}, \\
\delta_0 C^{a}{}_{i} &= -\nabla_i \epsilon_{\mathfrak{n}}{}^{a} + \epsilon_{\mathfrak{g}}{}^{\alpha} \rhd_{\alpha b}{}^{a} C^{b}{}_{i} \\
&\quad - \epsilon_{\mathfrak{m}i}{}^{\alpha} \partial^{a}{}_{\alpha} + 2\epsilon_{\mathfrak{h}}{}^{b}{}_{i} D_A X_{(bc)}{}^{A} g^{ac}, \\
\delta_0 \beta^{a}{}_{0i} &= -\nabla_0 \epsilon_{\mathfrak{h}i}{}^{a} + \epsilon_{\mathfrak{g}}{}^{\alpha} \rhd_{\alpha b}{}^{a} \beta_{b0i}, \\
\delta_0 \beta^{a}{}_{ij} &= -2\nabla_{[i|} \epsilon_{\mathfrak{h}|j]}{}^{a} + \epsilon_{\mathfrak{g}}{}^{\alpha} \rhd_{\alpha b}{}^{a} \beta^{b}{}_{ij} + \epsilon_{\mathfrak{l}ij}{}^{A} \delta_A{}^{a}, \\
\delta_0 \gamma^{A}{}_{0ij} &= \epsilon_{\mathfrak{g}}{}^{\alpha} \gamma^{B}{}_{0ij} \rhd_{\alpha B}{}^{A} + \nabla_0 \epsilon_{\mathfrak{l}ij}{}^{A} \\
&\quad - 4\epsilon_{\mathfrak{h}}{}^{a}{}_{[i|} \beta^{b}{}_{0|j]} X_{(ab)}{}^{A}, \\
\delta_0 \gamma^{A}{}_{ijk} &= \epsilon_{\mathfrak{g}}{}^{\alpha} \gamma^{B}{}_{ijk} \rhd_{\alpha B}{}^{A} + \nabla_i \epsilon_{\mathfrak{l}jk}{}^{A} \\
&\quad - \nabla_j \epsilon_{\mathfrak{l}ik}{}^{A} + \nabla_k \epsilon_{\mathfrak{l}ij}{}^{A} + 3!\epsilon_{\mathfrak{h}[i}{}^{a} \beta^{b}{}_{jk]} X_{(ab)}{}^{A}, \\
\delta_0 D^{A} &= \epsilon_{\mathfrak{n}}{}^{a} \delta_a{}^{A} + \epsilon_{\mathfrak{g}}{}^{\alpha} D^{B} \rhd_{\alpha B}{}^{A},
\end{aligned}
$$

$$
\begin{aligned}
\delta_0 \pi(B)_{\alpha}^{0i} &= f_{\alpha\beta}{}^{\gamma} \epsilon_{\mathfrak{g}}{}^{\beta} \pi(B)_{\gamma}{}^{0i}, \\
\delta_0 \pi(B)_{\alpha}{}^{ij} &= f_{\alpha\beta}{}^{\gamma} \epsilon_{\mathfrak{g}}{}^{\beta} \pi(B)_{\gamma}{}^{ij}, \\
\delta_0 \pi(\alpha)_{\alpha}^{0} &= f_{\alpha\beta}{}^{\gamma} \epsilon_{\mathfrak{m}}{}^{\beta}{}_{i} \pi(B)_{\gamma}{}^{0i} + f_{\alpha\beta}{}^{\gamma} \epsilon_{\mathfrak{g}}{}^{\beta} \pi(\alpha)_{\gamma}{}^{0} \\
&\quad + \rhd_{\alpha b}{}^{a} \epsilon_{\mathfrak{n}}{}^{b} \pi(C)_a^{0} + \rhd_{\alpha b}{}^{a} \epsilon_{\mathfrak{h}}{}^{b}{}_{i} \pi(\beta)_a^{i} \\
&\quad - \frac{1}{2} \rhd_{\alpha B}{}^{A} \epsilon_{\mathfrak{l}}{}^{B}{}_{ij} \pi(\gamma)_{A}{}^{0ij}, \\
\delta_0 \pi(\alpha)_{\alpha}^{i} &= f_{\alpha\beta}{}^{\gamma} \epsilon_{\mathfrak{m}}{}^{\beta}{}_{j} \pi(B)_{\gamma}{}^{ij} + f_{\alpha\beta}{}^{\gamma} \epsilon_{\mathfrak{g}}{}^{\beta} \pi(\alpha)_{\gamma}{}^{i} \\
&\quad + \rhd_{\alpha b}{}^{\alpha} \epsilon_{\mathfrak{n}}{}^{b} \pi(C)_a^{i} + \rhd_{\alpha b}{}^{\alpha} \epsilon_{\mathfrak{h}j}{}^{b} \pi(\beta)_a^{ij} \\
&\quad - \frac{1}{2} \rhd_{\alpha B}{}^{A} \epsilon_{\mathfrak{l}jk}{}^{B} \pi(\gamma)_A^{ijk} - \epsilon^{0ijk} \nabla_j \epsilon_{\mathfrak{m}\alpha k}, \\
&\quad - \frac{1}{2} \epsilon^{0ijk} \epsilon_{\mathfrak{n}}{}^{a} \rhd_{\alpha b}{}^{a} \beta^{b}{}_{jk}, \\
\delta_0 \pi(C)_a^{0} &= -\epsilon_{\mathfrak{g}}{}^{\alpha} \rhd_{\alpha a}{}^{b} \pi(C)_b^{0} + \epsilon_{\mathfrak{h}bi} \rhd_{\alpha a}{}^{b} \pi(B)^{\alpha 0i}, \\
\delta_0 \pi(C)_a^{i} &= -\epsilon_{\mathfrak{g}}{}^{\alpha} \rhd_{\alpha a}{}^{b} \pi(C)_b^{i} + \epsilon_{\mathfrak{h}bj} \rhd_{\alpha a}{}^{b} \pi(B)^{\alpha ij}, \\
\delta_0 \pi(\beta)_a{}^{0i} &= -\epsilon_{\mathfrak{g}}{}^{\alpha} \rhd_{\alpha a}{}^{b} \pi(\beta)_b{}^{0i} + \epsilon_{\mathfrak{n}b} \rhd_{\alpha a}{}^{b} \pi(B)^{\alpha 0i} \\
&\quad - 2\epsilon_{\mathfrak{h}j}{}^{b} X_{(ab)}{}^{A} \pi(\gamma)_A^{0ij}, \\
\delta_0 \pi(\beta)_a{}^{ij} &= -\epsilon_{\mathfrak{g}}{}^{\alpha} \rhd_{\alpha a}{}^{b} \pi(\beta)_b{}^{ij} + \epsilon_{\mathfrak{n}b} \rhd_{\alpha a}{}^{b} \pi(B)^{\alpha ij} \\
&\quad - 2\epsilon_{\mathfrak{h}k}{}^{b} X_{(ab)}{}^{A} \pi(\gamma)_A{}^{ijk} \\
&\quad + \epsilon^{0ijk} \nabla_k \epsilon_{\mathfrak{n}a} + \epsilon^{0ijk} \epsilon_{\mathfrak{h}k}{}^{a} \partial_{a\alpha}, \\
\delta_0 \pi(\gamma)_A{}^{0ij} &= -\epsilon_{\mathfrak{g}}{}^{\alpha} \rhd_{\alpha A}{}^{B} \pi(\gamma)_B{}^{0ij}, \\
\delta_0 \pi(\gamma)_A{}^{ijk} &= -\epsilon_{\mathfrak{g}}{}^{\alpha} \rhd_{\alpha A}{}^{B} \pi(\gamma)_B{}^{ijk} + \epsilon^{oijk} \delta_{aA} \epsilon_{\mathfrak{n}}{}^{a}, \\
\delta_0 \pi(D)_A &= -2\epsilon_{\mathfrak{h}i}{}^{a} X_{(ab)A} \pi(C)^{bi} \\
&\quad - \frac{1}{2} \epsilon_{\mathfrak{l}B}{}^{ij} \rhd_{\alpha A}{}^{B} \pi(B)_{0ij}^{\alpha} \\
&\quad - \epsilon_{\mathfrak{g}}{}^{\alpha} \rhd_{\alpha A}{}^{B} \pi(D)_B
\end{aligned}
\tag{E.2}
$$

Appendix F. Symmetry algebra calculations

To obtain the structure of the symmetry group of the 3*BF* action, as presented in the subsection 4.4, one has to calculate the commutators between the generators of all the symmetries, i.e. the *G*-, *H*-, *L*-, *M*-, and *N*-gauge symmetries. This process is described in the subsections 4.1–4.3, while details of the calculation which are not straightforward will be given in the following.

F.1. Commutator [H, H]

Let us derive the commutator of the generators of the *H*-gauge transformations, i.e. the equation (76). After transforming the variables under *H*-gauge transformations for the parameter $\epsilon_{\mathfrak{h}1}$ one obtains the following

$$\alpha' = \alpha - \partial\epsilon_{\mathfrak{h}1}, \tag{F.1}$$

$$\beta' = \beta - \overset{\alpha-\partial\epsilon_{\mathfrak{h}1}}{\nabla}\epsilon_{\mathfrak{h}1} - \epsilon_{\mathfrak{h}1}\wedge\epsilon_{\mathfrak{h}1}, \tag{F.2}$$

$$\gamma' = \gamma + \{\beta - \overset{\alpha-\partial\epsilon_{\mathfrak{h}1}}{\nabla}\epsilon_{\mathfrak{h}1} - \epsilon_{\mathfrak{h}1}\wedge\epsilon_{\mathfrak{h}1}, \epsilon_{\mathfrak{h}1}\}_{\mathrm{pf}} + \{\epsilon_{\mathfrak{h}1}, \beta\}_{\mathrm{pf}}, \tag{F.3}$$

$$B' = B - (C - D\wedge^{\mathcal{X}_1}\epsilon_{\mathfrak{h}1} - D\wedge^{\mathcal{X}_2}\epsilon_{\mathfrak{h}1})\wedge^{\mathcal{T}}\epsilon_{\mathfrak{h}1} - \epsilon_{\mathfrak{h}1}\wedge^{\mathcal{D}}\epsilon_{\mathfrak{h}1}\wedge^{\mathcal{D}}D, \tag{F.4}$$

$$C' = C - D\wedge^{\mathcal{X}_1}\epsilon_{\mathfrak{h}1} - D\wedge^{\mathcal{X}_2}\epsilon_{\mathfrak{h}1}, \tag{F.5}$$

$$D' = D, \tag{F.6}$$

and transforming the variables once more for the parameter $\epsilon_{\mathfrak{h}2}$ one obtains:

$$\begin{aligned}
\alpha'' &= \alpha - \partial\epsilon_{\mathfrak{h}1} - \partial\epsilon_{\mathfrak{h}2},\\
\beta'' &= \beta - \overset{\alpha-\partial\epsilon_{\mathfrak{h}1}}{\nabla}\epsilon_{\mathfrak{h}1} - \epsilon_{\mathfrak{h}1}\wedge\epsilon_{\mathfrak{h}1} - \overset{\alpha-\partial\epsilon_{\mathfrak{h}1}-\partial\epsilon_{\mathfrak{h}2}}{\nabla}\epsilon_{\mathfrak{h}2} - \epsilon_{\mathfrak{h}2}\wedge\epsilon_{\mathfrak{h}2},\\
\gamma'' &= \gamma + \{\beta - \overset{\alpha-\partial\epsilon_{\mathfrak{h}1}}{\nabla}\epsilon_{\mathfrak{h}1} - \epsilon_{\mathfrak{h}1}\wedge\epsilon_{\mathfrak{h}1}, \epsilon_{\mathfrak{h}1}\}_{\mathrm{pf}} + \{\epsilon_{\mathfrak{h}1}, \beta\}_{\mathrm{pf}}\\
&\quad + \{\beta - \overset{\alpha-\partial\epsilon_{\mathfrak{h}1}}{\nabla}\epsilon_{\mathfrak{h}1} - \epsilon_{\mathfrak{h}1}\wedge\epsilon_{\mathfrak{h}1} - \overset{\alpha-\partial\epsilon_{\mathfrak{h}1}-\partial\epsilon_{\mathfrak{h}2}}{\nabla}\epsilon_{\mathfrak{h}2} - \epsilon_{\mathfrak{h}2}\wedge\epsilon_{\mathfrak{h}2}, \epsilon_{\mathfrak{h}2}\}_{\mathrm{pf}}\\
&\quad + \{\epsilon_{\mathfrak{h}2}, \beta - \overset{\alpha-\partial\epsilon_{\mathfrak{h}1}}{\nabla}\epsilon_{\mathfrak{h}1} - \epsilon_{\mathfrak{h}1}\wedge\epsilon_{\mathfrak{h}1}\}_{\mathrm{pf}},\\
B'' &= B - (C - D\wedge^{\mathcal{X}_1}\epsilon_{\mathfrak{h}1} - D\wedge^{\mathcal{X}_2}\epsilon_{\mathfrak{h}1})\wedge^{\mathcal{T}}\epsilon_{\mathfrak{h}1} - \epsilon_{\mathfrak{h}1}\wedge^{\mathcal{D}}\epsilon_{\mathfrak{h}1}\wedge^{\mathcal{D}}D\\
&\quad - (C - D\wedge^{\mathcal{X}_1}\epsilon_{\mathfrak{h}1} - D\wedge^{\mathcal{X}_2}\epsilon_{\mathfrak{h}1} - D\wedge^{\mathcal{X}_1}\epsilon_{\mathfrak{h}2} - D\wedge^{\mathcal{X}_2}\epsilon_{\mathfrak{h}2})\wedge^{\mathcal{T}}\epsilon_{\mathfrak{h}2}\\
&\quad - \epsilon_{\mathfrak{h}2}\wedge^{\mathcal{D}}\epsilon_{\mathfrak{h}2}\wedge^{\mathcal{D}}D,\\
C'' &= C - D\wedge^{\mathcal{X}_1}\epsilon_{\mathfrak{h}1} - D\wedge^{\mathcal{X}_2}\epsilon_{\mathfrak{h}1} - D\wedge^{\mathcal{X}_1}\epsilon_{\mathfrak{h}2} - D\wedge^{\mathcal{X}_2}\epsilon_{\mathfrak{h}2},\\
D'' &= D.
\end{aligned} \tag{F.7}$$

It is easy to see that for variables $\alpha^{\alpha}{}_{\mu}$, $C^{a}{}_{\mu}$ and D^{A} the following is obtained:

$$\begin{aligned} e^{\epsilon_{\mathfrak{h}1}\cdot H}e^{\epsilon_{\mathfrak{h}2}\cdot H}\alpha^{\alpha}{}_{\mu} &= e^{\epsilon_{\mathfrak{h}2}\cdot H}e^{\epsilon_{\mathfrak{h}1}\cdot H}\alpha^{\alpha}{}_{\mu},\\ e^{\epsilon_{\mathfrak{h}1}\cdot H}e^{\epsilon_{\mathfrak{h}2}\cdot H}C^{a}{}_{\mu} &= e^{\epsilon_{\mathfrak{h}2}\cdot H}e^{\epsilon_{\mathfrak{h}1}\cdot H}C^{a}{}_{\mu},\\ e^{\epsilon_{\mathfrak{h}1}\cdot H}e^{\epsilon_{\mathfrak{h}2}\cdot H}D^{A} &= e^{\epsilon_{\mathfrak{h}2}\cdot H}e^{\epsilon_{\mathfrak{h}1}\cdot H}D^{A}. \end{aligned} \tag{F.8}$$

For the remaining variables, $\beta^{a}{}_{\mu\nu}$, $\gamma^{A}{}_{\mu\nu\rho}$ and $B^{\alpha}{}_{\mu\nu}$, after subtracting (appendix F.1) and the corresponding equation where $\epsilon_{\mathfrak{h}1} \leftrightarrow \epsilon_{\mathfrak{h}2}$, one obtains:

$$\begin{aligned} \left(e^{\epsilon_{\mathfrak{h}1}\cdot H}e^{\epsilon_{\mathfrak{h}2}\cdot H} - e^{\epsilon_{\mathfrak{h}2}\cdot H}e^{\epsilon_{\mathfrak{h}1}\cdot H}\right)\frac{1}{2}\beta^{a}{}_{\mu\nu} &= \partial_{b}{}^{\alpha}\epsilon_{\mathfrak{h}2}{}^{b}{}_{[\mu|}\epsilon_{\mathfrak{h}1}{}^{c}{}_{|\nu]}\rhd_{\alpha c}{}^{a} - \partial_{b}{}^{\alpha}\epsilon_{\mathfrak{h}1}{}^{b}{}_{[\mu|}\epsilon_{\mathfrak{h}2}{}^{c}{}_{|\nu]}\rhd_{\alpha c}{}^{a}\\ &= 2\delta_{A}{}^{a}X_{(bc)}{}^{A}\epsilon_{\mathfrak{h}1}{}^{b}{}_{[\mu|}\epsilon_{\mathfrak{h}2}{}^{c}{}_{|\nu]}\\ &= \delta_{A}{}^{a}(\{\epsilon_{\mathfrak{h}1}\wedge\epsilon_{\mathfrak{h}2}\}_{\mathrm{pf}} - \{\epsilon_{\mathfrak{h}2}\wedge\epsilon_{\mathfrak{h}1}\}_{\mathrm{pf}})^{A}_{\mu\nu},\\ \left(e^{\epsilon_{\mathfrak{h}1}\cdot H}e^{\epsilon_{\mathfrak{h}2}\cdot H} - e^{\epsilon_{\mathfrak{h}2}\cdot H}e^{\epsilon_{\mathfrak{h}1}\cdot H}\right)\frac{1}{3!}\gamma^{A}{}_{\mu\nu\rho} &= 2(\partial_{[\mu}\epsilon_{\mathfrak{h}1}{}^{a}_{\nu})\epsilon_{\mathfrak{h}2}{}^{b}_{\rho]}X_{(ab)}{}^{A} + 2\epsilon_{\mathfrak{h}1}{}^{a}_{[\nu}(\partial_{\mu}\epsilon_{\mathfrak{h}2}{}^{b}_{\rho]})X_{(ab)}{}^{A}\\ &\quad + 2\alpha^{\alpha}{}_{[\mu}\epsilon_{\mathfrak{h}1}{}^{a}_{\nu}\epsilon_{\mathfrak{h}2}{}^{b}_{\rho]}X_{(ab)}{}^{B}\rhd_{\alpha B}{}^{A}\\ &= \nabla_{[\mu}(\{\epsilon_{\mathfrak{h}1}\wedge\epsilon_{\mathfrak{h}2}\}_{\mathrm{pf}} - \{\epsilon_{\mathfrak{h}2}\wedge\epsilon_{\mathfrak{h}1}\}_{\mathrm{pf}})^{A}{}_{\nu\rho]},\\ \left(e^{\epsilon_{\mathfrak{h}1}\cdot H}e^{\epsilon_{\mathfrak{h}2}\cdot H} - e^{\epsilon_{\mathfrak{h}2}\cdot H}e^{\epsilon_{\mathfrak{h}1}\cdot H}\right)\frac{1}{2}B^{\alpha}{}_{\mu\nu} &= D^{A}\epsilon_{\mathfrak{h}2}{}^{a}_{[\mu|}\epsilon_{\mathfrak{h}1}{}^{b}_{|\nu]}(X_{1Aa}{}^{c} + X_{2Aa}{}^{c})\mathcal{T}_{cb}{}^{\alpha}\\ &\quad - D^{A}\epsilon_{\mathfrak{h}1}{}^{b}_{[\mu|}\epsilon_{\mathfrak{h}2}{}^{a}_{|\nu]}(X_{1Ab}{}^{c} + X_{2Ab}{}^{c})\mathcal{T}_{ca}{}^{\alpha}\\ &= -2D_{A}\epsilon_{\mathfrak{h}1}{}^{a}_{[\mu|}\epsilon_{\mathfrak{h}2}{}^{b}_{|\nu]}\left(X_{(ac)}{}^{A}\rhd_{\alpha b}{}^{c} + X_{(bc)}{}^{A}\rhd_{\alpha a}{}^{c}\right)\\ &= -2D_{A}\epsilon_{\mathfrak{h}1}{}^{a}_{[\mu|}\epsilon_{\mathfrak{h}2}{}^{b}_{|\nu]}X_{(ab)}{}^{B}\rhd_{\alpha B}{}^{A}\\ &= \left(D\wedge^{\mathcal{S}}(\{\epsilon_{\mathfrak{h}1}\wedge\epsilon_{\mathfrak{h}2}\}_{\mathrm{pf}} - \{\epsilon_{\mathfrak{h}2}\wedge\epsilon_{\mathfrak{h}1}\}_{\mathrm{pf}}\right)^{\alpha}_{\mu\nu}. \end{aligned} \tag{F.9}$$

Comparing (F.8) and (F.9) with (72), one concludes that the commutator of two H-gauge transformations is the L-gauge transformation with the parameter $\epsilon_{\mathfrak{l}}^{A}{}_{\mu\nu} = 4\epsilon_{\mathfrak{h}1}{}^{a}{}_{[\mu|}\epsilon_{\mathfrak{h}2}{}^{b}{}_{|\nu]}X_{(ac)}{}^{A}$:

$$e^{\epsilon_{\mathfrak{h}1}\cdot H}e^{\epsilon_{\mathfrak{h}2}\cdot H} - e^{\epsilon_{\mathfrak{h}2}\cdot H}e^{\epsilon_{\mathfrak{h}1}\cdot H} = 2\left(\{\epsilon_{\mathfrak{h}1}\wedge\epsilon_{\mathfrak{h}2}\}_{\mathrm{pf}} - \{\epsilon_{\mathfrak{h}2}\wedge\epsilon_{\mathfrak{h}1}\}_{\mathrm{pf}}\right)\cdot\hat{L}. \tag{F.10}$$

F.2. Commutator [H, N]

Let us calculate the commutator between the generators of H-gauge transformation and N-gauge transformation, i.e. derive the equation (109). This is done by calculating the expressions

$$\left(e^{\epsilon_{\mathfrak{h}}\cdot H}e^{\epsilon_{\mathfrak{n}}\cdot N} - e^{\epsilon_{\mathfrak{n}}\cdot N}e^{\epsilon_{\mathfrak{h}}\cdot H}\right)A, \tag{F.11}$$

for all variables A present in the theory. It is easy to see that for variables $\alpha^{\alpha}{}_{\mu}$, $\beta^{a}{}_{\mu\nu}$, $\gamma^{A}{}_{\mu\nu\rho}$, and D^{A} the following is obtained:

$$
\begin{aligned}
\mathrm{e}^{\epsilon_{\mathfrak{h}}\cdot H}\mathrm{e}^{\epsilon_{\mathfrak{n}}\cdot N}\alpha^{\alpha}{}_{\mu} &= \mathrm{e}^{\epsilon_{\mathfrak{n}}\cdot N}\mathrm{e}^{\epsilon_{\mathfrak{h}}\cdot H}\alpha^{\alpha}{}_{\mu},\\
\mathrm{e}^{\epsilon_{\mathfrak{h}}\cdot H}\mathrm{e}^{\epsilon_{\mathfrak{n}}\cdot N}\beta^{a}{}_{\mu\nu} &= \mathrm{e}^{\epsilon_{\mathfrak{n}}\cdot N}\mathrm{e}^{\epsilon_{\mathfrak{h}}\cdot H}\beta^{a}{}_{\mu\nu},\\
\mathrm{e}^{\epsilon_{\mathfrak{h}}\cdot H}\mathrm{e}^{\epsilon_{\mathfrak{n}}\cdot N}\gamma^{A}{}_{\mu\nu\rho} &= \mathrm{e}^{\epsilon_{\mathfrak{n}}\cdot N}\mathrm{e}^{\epsilon_{\mathfrak{h}}\cdot H}\gamma^{A}{}_{\mu\nu\rho},\\
\mathrm{e}^{\epsilon_{\mathfrak{h}}\cdot H}\mathrm{e}^{\epsilon_{\mathfrak{n}}\cdot N}D^{A} &= \mathrm{e}^{\epsilon_{\mathfrak{n}}\cdot N}\mathrm{e}^{\epsilon_{\mathfrak{h}}\cdot H}D^{A}.
\end{aligned}
\tag{F.12}
$$

For the remaining variables, $B^{\alpha}{}_{\mu\nu}$ and $C^{a}{}_{\mu}$, after the H-gauge transformation one obtains the following:

$$
B' = B - (C - D\wedge^{\chi_1}\epsilon_{\mathfrak{h}} - D\wedge^{\chi_2}\epsilon_{\mathfrak{h}})\wedge^{\tau}\epsilon_{\mathfrak{h}} - \epsilon_{\mathfrak{h}}\wedge^{\mathcal{D}}\epsilon_{\mathfrak{h}}\wedge^{\mathcal{D}}D, \tag{F.13}
$$

$$
C' = C - D\wedge^{\chi_1}\epsilon_{\mathfrak{h}} - D\wedge^{\chi_2}\epsilon_{\mathfrak{h}}. \tag{F.14}
$$

Next, transforming those variables with N-gauge transformation one obtains:

$$
\begin{aligned}
B'' &= B' - \beta'\wedge^{\mathcal{T}}\epsilon_{\mathfrak{n}}\\
&= B - (C - D\wedge^{\chi_1}\epsilon_{\mathfrak{h}} - D\wedge^{\chi_2}\epsilon_{\mathfrak{h}})\wedge^{\tau}\epsilon_{\mathfrak{h}} - \epsilon_{\mathfrak{h}}\wedge^{\mathcal{D}}\epsilon_{\mathfrak{h}}\wedge^{\mathcal{D}}D\\
&\quad - (\beta - \overset{\{\alpha^{\alpha} - \partial_a{}^{\alpha}\epsilon_{\mathfrak{h}}^{a}\}}{\nabla}\epsilon_{\mathfrak{h}} - \epsilon_{\mathfrak{h}}\wedge\epsilon_{\mathfrak{h}})\wedge^{\mathcal{T}}\epsilon_{\mathfrak{n}},\\
C'' &= C' - \overset{\{\alpha^{\alpha} - \partial_a{}^{\alpha}\epsilon_{\mathfrak{h}}^{a}\}}{\nabla}\epsilon_{\mathfrak{n}}\\
&= C - D\wedge^{\chi_1}\epsilon_{\mathfrak{h}} - D\wedge^{\chi_2}\epsilon_{\mathfrak{h}} - \overset{\{\alpha^{\alpha} - \partial_a{}^{\alpha}\epsilon_{\mathfrak{h}}^{a}\}}{\nabla}\epsilon_{\mathfrak{n}}.
\end{aligned}
\tag{F.15}
$$

Let us now exchange the order of transformations, and first transform the variables with N-gauge transformation,

$$
B^{\cdot} = B - \beta\wedge^{\mathcal{T}}\epsilon_{\mathfrak{n}}, \tag{F.16}
$$

$$
C^{\cdot} = C - \nabla\epsilon_{\mathfrak{n}}, \tag{F.17}
$$

and then with H-gauge transformation:

$$
\begin{aligned}
B^{\cdot\cdot} &= B^{\cdot} - (C^{\cdot} - D^{\cdot}\wedge^{\chi_1}\epsilon_{\mathfrak{h}} - D^{\cdot}\wedge^{\chi_2}\epsilon_{\mathfrak{h}})\wedge^{\tau}\epsilon_{\mathfrak{h}} - \epsilon_{\mathfrak{h}}\wedge^{\mathcal{D}}\epsilon_{\mathfrak{h}}\wedge^{\mathcal{D}}D^{\cdot}\\
&= B - \beta\wedge^{\mathcal{T}}\epsilon_{\mathfrak{n}} - \big(C - \nabla\epsilon_{\mathfrak{n}} - (D + \delta\epsilon_{\mathfrak{n}})\wedge^{\chi_1}\epsilon_{\mathfrak{h}}\\
&\quad - (D + \delta\epsilon_{\mathfrak{n}})\wedge^{\chi_2}\epsilon_{\mathfrak{h}}\big)\wedge^{\tau}\epsilon_{\mathfrak{h}} - \epsilon_{\mathfrak{h}}\wedge^{\mathcal{D}}\epsilon_{\mathfrak{h}}\wedge^{\mathcal{D}}(D + \delta\epsilon_{\mathfrak{n}}),\\
C^{\cdot\cdot} &= C^{\cdot} - D^{\cdot}\wedge^{\chi_1}\epsilon_{\mathfrak{h}} - D^{\cdot}\wedge^{\chi_2}\epsilon_{\mathfrak{h}}\\
&= C - \nabla\epsilon_{\mathfrak{n}} - (D + \delta\epsilon_{\mathfrak{n}})\wedge^{\chi_1}\epsilon_{\mathfrak{h}} - (D + \delta\epsilon_{\mathfrak{n}})\wedge^{\chi_2}\epsilon_{\mathfrak{h}}.
\end{aligned}
\tag{F.18}
$$

After subtracting (F.15) and (F.18) one obtains:

$$\begin{aligned}
\left(e^{\epsilon_{\mathfrak{h}}\cdot H}e^{\epsilon_{\mathfrak{n}}\cdot N} - e^{\epsilon_{\mathfrak{n}}\cdot N}e^{\epsilon_{\mathfrak{h}}\cdot H}\right) B^{\alpha} &= \nabla\epsilon_{\mathfrak{n}}^{a} \wedge {\epsilon_{\mathfrak{h}}}^{b}{\mathcal{T}_{ab}}^{\alpha} + {\delta^{A}}_{a}\epsilon_{\mathfrak{n}}^{a}{\epsilon_{\mathfrak{h}}}^{b} \wedge {\epsilon_{\mathfrak{h}}}^{d}{X_{1Ab}}^{c}{\mathcal{T}_{cd}}^{\alpha} \\
&\quad + {\delta^{A}}_{a}\epsilon_{\mathfrak{n}}^{a}{\epsilon_{\mathfrak{h}}}^{b} \wedge {\epsilon_{\mathfrak{h}}}^{d}{X_{2Ab}}^{c}{\mathcal{T}_{cd}}^{\alpha} - {\epsilon_{\mathfrak{h}}}^{a} \wedge {\epsilon_{\mathfrak{h}}}^{b}{\delta_{A}}^{c}{\epsilon_{\mathfrak{n}}}^{c}{D_{Aab}}^{\alpha}, \\
&\quad - \nabla\epsilon_{\mathfrak{h}}^{a} \wedge \epsilon_{\mathfrak{n}}^{b}{\mathcal{T}_{ab}}^{\alpha} + {\partial_{a}}^{\beta}\epsilon_{\mathfrak{h}}^{a} \triangleright_{\beta c}{}^{b}\epsilon_{\mathfrak{h}}^{c}\epsilon_{\mathfrak{n}}^{d}{\mathcal{T}_{bd}}^{\alpha} - \epsilon_{\mathfrak{h}}^{a} \wedge \epsilon_{\mathfrak{h}}^{b}{f_{ab}}^{c}\epsilon_{\mathfrak{n}}^{d}{\mathcal{T}_{cd}}^{\alpha}, \\
\left(e^{\epsilon_{\mathfrak{h}}\cdot H}e^{\epsilon_{\mathfrak{n}}\cdot N} - e^{\epsilon_{\mathfrak{n}}\cdot N}e^{\epsilon_{\mathfrak{h}}\cdot H}\right) C^{c} &= -({\delta^{A}}_{a}\epsilon_{\mathfrak{n}}^{a}) \wedge \epsilon_{\mathfrak{h}}^{b}{X_{1Ab}}^{c} - ({\delta^{A}}_{a}\epsilon_{\mathfrak{n}}^{a}) \wedge \epsilon_{\mathfrak{h}}^{b}{X_{2Ab}}^{c} - {\partial_{a}}^{\beta}\epsilon_{\mathfrak{h}}^{a} \triangleright_{\beta b}{}^{c}\epsilon_{\mathfrak{n}}^{b},
\end{aligned} \tag{F.19}$$

where after using the definitions of the maps $\mathcal{T}$, $\mathcal{D}$, χ_1, and χ_2 one obtains the result

$$\begin{aligned}
\left(e^{\epsilon_{\mathfrak{h}}\cdot H}e^{\epsilon_{\mathfrak{n}}\cdot N} - e^{\epsilon_{\mathfrak{n}}\cdot N}e^{\epsilon_{\mathfrak{h}}\cdot H}\right) B^{\alpha} &= \nabla\epsilon_{\mathfrak{n}}^{a} \wedge {\epsilon_{\mathfrak{h}}}^{b}{\mathcal{T}_{ab}}^{\alpha} - \nabla\epsilon_{\mathfrak{h}}^{a} \wedge \epsilon_{\mathfrak{n}}^{b}{\mathcal{T}_{ab}}^{\alpha} \\
&= \nabla(\epsilon_{\mathfrak{n}} \wedge^{\mathcal{T}} \epsilon_{\mathfrak{h}})^{\alpha}, \\
\left(e^{\epsilon_{\mathfrak{h}}\cdot H}e^{\epsilon_{\mathfrak{n}}\cdot N} - e^{\epsilon_{\mathfrak{n}}\cdot N}e^{\epsilon_{\mathfrak{h}}\cdot H}\right) C^{c} &= {\partial^{c}}_{\alpha}(\epsilon_{\mathfrak{n}} \wedge^{\mathcal{T}} \epsilon_{\mathfrak{h}})^{\alpha},
\end{aligned} \tag{F.20}$$

Comparing (F.12) and (F.20) with (85), one obtains that:

$$\left(e^{\epsilon_{\mathfrak{h}}\cdot H}e^{\epsilon_{\mathfrak{n}}\cdot N} - e^{\epsilon_{\mathfrak{n}}\cdot N}e^{\epsilon_{\mathfrak{h}}\cdot H}\right) = -(\epsilon_{\mathfrak{n}} \wedge^{\mathcal{T}} \epsilon_{\mathfrak{h}}) \cdot M. \tag{F.21}$$

ORCID iDs

Tijana Radenković https://orcid.org/0000-0002-2310-5281
Marko Vojinović https://orcid.org/0000-0001-6977-4870

References

[1] Rovelli C 2011 Zakopane lectures on loop gravity (arXiv:1102.3660)
[2] Rovelli C 2004 *Quantum Gravity* (Cambridge: Cambridge University Press)
[3] Thiemann T 2007 *Modern Canonical Quantum General Relativity* (Cambridge: Cambridge University Press)
[4] Rovelli C and Vidotto F 2014 *Covariant Loop Quantum Gravity* (Cambridge: Cambridge University Press)
[5] Ponzano G and Regge T 1968 Semiclassical limit of Racah coefficients *Spectroscopic and Group Theoretical Methods in Physics: Racah Memorial* (Amsterdam: North-Holland Publ. Co.) pp 1–58
[6] Ooguri H 1992 Topological lattice models in four dimensions *Mod. Phys. Lett.* A **7** 279
[7] Barrett J W and Crane L 1998 Relativistic spin networks and quantum gravity *J. Math. Phys.* **39** 3296
[8] Barrett J W and Crane L 2000 A Lorentzian signature model for quantum general relativity *Class. Quantum Grav.* **17** 3101
[9] Crane L and Sheppeard M D 2003 Two-categorical Poincaré representations and state sum applications (arXiv:math/0306440)
[10] Engle J, Livine E, Pereira R and Rovelli C 2008 LQG vertex with finite Immirzi parameter *Nucl. Phys.* B **799** 136
[11] Freidel L and Krasnov K 2008 A new spin foam model for 4D gravity *Class. Quantum Grav.* **25** 125018
[12] Baez J C and Huerta J 2011 An invitation to higher gauge theory *Gen. Relativ. Gravit.* **43** 2335
[13] Radenković T and Vojinović M 2019 Higher gauge theories based on three-groups *J. High Energy Phys.* JHEP10(2019)222
[14] Radenković T and Vojinović M 2022 Topological invariant of four-manifolds based on a three-group (arXiv:2201.02572)

[15] Blagojević M 2002 *Gravitation and Gauge Symmetries* (Bristol: Institute of Physics Publishing)
[16] Miković A, Oliveira M A and Vojinović M 2016 Hamiltonian analysis of the BFCG theory for a generic Lie two-group (arXiv:1610.09621)
[17] Miković A, Oliveira M A and Vojinović M 2019 Hamiltonian analysis of the BFCG formulation of general relativity *Class. Quantum Grav.* **36** 015005
[18] Miković A, Oliveira M A and Vojinović M 2016 Hamiltonian analysis of the BFCG theory for the Poincaré two-group *Class. Quantum Grav.* **33** 065007
[19] Miković A and Oliveira M A 2015 Canonical formulation of Poincare BFCG theory and its quantization *Gen. Relativ. Gravit.* **47** 58
[20] Radenković T and Vojinović M 2020 Hamiltonian analysis for the scalar electrodynamics as 3*BF* theory *Symmetry* **12** 620
[21] Wang W 2014 On three-gauge transformations, three-curvatures, and Gray-categories *J. Math. Phys.* **55** 043506
[22] Plebanski J F 1977 On the separation of Einsteinian substructures *J. Math. Phys.* **12** 2511
[23] Baez J C 2000 An introduction to spin foam models of *BF* theory and quantum gravity *Lect. Notes Phys.* **543** 25
[24] Celada M, González D and Montesinos M 2016 *BF* gravity *Class. Quantum Grav.* **33** 213001
[25] Girelli F, Pfeiffer H and Popescu E M 2008 Topological higher gauge theory: from *BF* to BFCG theory *J. Math. Phys.* **49** 032503
[26] Martins J and Miković A 2011 Lie crossed modules and gauge-invariant actions for two-*BF* theories *Adv. Theor. Math. Phys.* **15** 1059
[27] Miković A and Vojinović M 2012 Poincaré two-group and quantum gravity *Class. Quantum Grav.* **29** 165003
[28] Baez J and Schreiber U 2004 Higher gauge theory: two-connections on two-bundles (arXiv: hep-th/0412325)
[29] Faria Martins J and Picken R 2011 The fundamental Gray three-groupoid of a smooth manifold and local three-dimensional holonomy based on a two-crossed module *Differ. Geom. Appl.* **29** 179–206
[30] Saemann C and Wolf M 2014 Six-dimensional superconformal field theories from principal three-bundles over twistor space *Lett. Math. Phys.* **104** 1147
[31] Miković A and Vojinović M 2021 Standard model and four-groups *Europhys. Lett.* **133** 61001
[32] Miković A and Vojinović M 2012 Effective action for EPRL/FK spin foam models *J. Phys.: Conf. Ser.* **360** 012049
[33] Miković A and Vojinović M 2013 A finiteness bound for the EPRL/FK spin foam model *Class. Quantum Grav.* **30** 035001
[34] Miković A and Vojinović M 2015 Solution to the cosmological constant problem in a Regge quantum gravity model *Europhys. Lett.* **110** 40008
[35] Horowitz G T 1989 Exactly soluble diffeomorphism invariant theories *Commun. Math. Phys.* **125** 417–37

Journal of Physics: Conference Series

PAPER • **OPEN ACCESS**

Higher category theory and n-groups as gauge symmetries for quantum gravity

To cite this article: Bojan Nikolić *et al* 2023 *J. Phys.: Conf. Ser.* **2667** 012019

Journal of Physics: Conference Series **2667** (2023) 012019 doi:10.1088/1742-6596/2667/1/012019

Higher category theory and n-groups as gauge symmetries for quantum gravity

Bojan Nikolić, Danijel Obrić, Tijana Radenković, Igor Salom and Marko Vojinović

Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia

E-mail: vmarko@ipb.ac.rs

Abstract. Higher category theory can be employed to generalize the notion of a gauge group to the notion of a gauge n-group. This novel algebraic structure is designed to generalize notions of connection, parallel transport and holonomy from curves to manifolds of dimension higher than one. Thus it generalizes the concept of gauge symmetry, giving rise to a topological action called nBF action, living on a corresponding n-principal bundle over a spacetime manifold. Similarly as for the Plebanski action, one can deform the topological nBF action by adding appropriate simplicity constraints, in order to describe the correct dynamics of both gravity and matter fields. Specifically, one can describe the whole Standard Model coupled to gravity as a constrained $3BF$ or $4BF$ action. The split of the full action into a topological sector and simplicity constraints sector is adapted to the spinfoam quantization technique, with the aim to construct a full model of quantum gravity with matter. In addition, the properties of the gauge n-group structure open up a possibility of a nontrivial unification of all fields. An n-group naturally contains additional novel gauge groups which specify the spectrum of matter fields present in the theory, in a similar way to the ordinary gauge group that prescribes the spectrum of gauge vector bosons in the Yang-Mills theory. The presence and the properties of these new gauge groups has the potential to explain fermion families, and other structure in the matter spectrum of the theory.

1. Introduction

The formulation of a quantum theory of gravity represents one of the fundamental open problems in modern theoretical physics. Among the many approaches to this problem, some have developed into vast research frameworks, such as Loop Quantum Gravity, which aims to formulate a model of quantum gravity (QG) in a nonperturbative fashion, both canonically and covariantly [1, 2, 3]. The covariant approach aims to give a tentative rigorous definition of the path integral for the gravitational field,

$$Z = \int \mathcal{D}g \; e^{iS[g]} \,. \tag{1}$$

One of the essential assumptions is a triangulation of a spacetime manifold, and the path integral is introduced as a discrete state sum of the gravitational field configurations, living on the simplicial complex structure. This approach to quantization of gravity is usually called the *spinfoam* quantization method. It is performed via the following three steps:

(1) one reformulates the classical action $S[g]$ as a constrained BF action, separating the topological BF part and the constraint part of the action;
(2) one employs the underlying Lie group structure of the BF sector of the action, in order to define a triangulation-independent state sum Z;
(3) finally, one deforms the topological state sum by applying the simplicity constraints, and therefore redefining it into a triangulation-dependent state sum, which plays the role of a definition for the path integral (1).

This type of quantization prescription has been implemented in a number of cases, for various choices of the gravitational action, of the Lie group, and of the spacetime dimension. Historically the first spinfoam model was the Ponzano-Regge model [4], defined in 3 spacetime dimensions. In 4 dimensions multiple models have been formulated, differing in the choice of the Lie group and the way one imposes the simplicity constraints [5, 6, 7, 8, 9]. While all these models do represent definitions of the gravitational path integral, none of them are able to include matter fields in a seamless way. Introducing the latter into a spinfoam QG model has so far had only limited success [10], predominantly due to the lack of the tetrad fields in the topological part of the model.

Recently, a new approach has been developed to address the issue of matter fields, which employs the framework of *higher gauge theory* (see [11] for a review). Specifically, one uses the notion of a *categorical ladder* to generalize the BF action (based on a Lie group) to a $2BF$ action (based on the so-called 2-group structure), and further to a $3BF$ action (based on a 3-group structure). A convenient choice of the *Poincaré 2-group* gives rise to the needed tetrad fields in the topological sector of the action [12], while an additional extension to the 3-group naturally introduces the matter fields (fermions and scalars) into the model [13]. The steps of the categorical ladder and their corresponding structures are summarized as follows:

categorical structure	algebraic structure	linear structure	topological action	degrees of freedom
Lie group	Lie group	Lie algebra	BF theory	gauge fields
Lie 2-group	Lie crossed module	differential Lie crossed module	$2BF$ theory	tetrad fields
Lie 3-group	Lie 2-crossed module	differential Lie 2-crossed module	$3BF$ theory	scalar and fermion fields

The main aim of this work is to provide a short review of the classical pure BF, $2BF$ and $3BF$ actions, in order to demonstrate the categorical ladder procedure and the construction of higher gauge theories. In other words, we mainly focus on the step 1 of the spinfoam quantization programme, with a very short review of step 2 of the programme.

The layout of the paper is as follows. Section 2 deals with first three examples of nBF theories, namely BF, $2BF$ and $3BF$ actions, and their construction using the categorical ladder. After this, in Section 3 we briefly present an application of a $3BF$ theory to the Standard Model of elementary particles coupled to Einstein-Cartan gravity. As it turns out, the scalar and fermion fields are *naturally associated to a new gauge group*, generalizing the role of an ordinary gauge group in the Yang-Mills theory. This opens up a possibility of an algebraic classification of matter fields, and (more speculatively) a possibility of the explanation of the three fermion families. Finally, Section 4 contains some discussion and our conclusions.

The notation and conventions are as follows. Spacetime indices are denoted by the Greek letters $\mu, \nu, \dots$, and are raised and lowered by the spacetime metric $g_{\mu\nu} = \eta_{ab} e^a{}_\mu e^b{}_\nu$, where $e^a{}_\mu$ are the tetrad fields. The inverse tetrad is denoted as $e^\mu{}_a$. The local Lorentz indices are denoted by the Latin letters $a, b, c, \dots$, take values $0, 1, 2, 3$, and are raised and lowered using the Minkowski metric η_{ab} with signature $(-, +, +, +)$. All other indices that appear in the paper depend on the context, and their use is explicitly defined in the text where they appear. We work in the natural system of units where $c = \hbar = 1$, and $G = l_p^2$, where l_p is the Planck length. The exterior product in the space of differential forms is denoted with the standard "wedge" symbol, $\wedge$.

2. nBF theories

We begin by giving a short review of nBF theories, for $n = 1, 2, 3$, which represent the most interesting cases for physics.

2.1. BF theory

A BF theory and its various applications in physics are already well known in the literature, see for example [14, 15, 16], so here we merely give a brief definition. Given a Lie group G, and its corresponding Lie algebra as $\mathfrak{g}$, one defines the BF action in the form (we discuss only the 4-dimensional spacetime manifolds $\mathcal{M}_4$):

$$S_{BF} = \int_{\mathcal{M}_4} \langle B \wedge \mathcal{F} \rangle_{\mathfrak{g}} \,. \tag{2}$$

Here, $\mathcal{F} \equiv \mathrm{d}\alpha + \alpha \wedge \alpha$ is the curvature 2-form for the $\mathfrak{g}$-valued connection 1-form $\alpha \in \Lambda^1(\mathcal{M}_4) \otimes \mathfrak{g}$, while $B \in \Lambda^2(\mathcal{M}_4) \otimes \mathfrak{g}$ is a $\mathfrak{g}$-valued Lagrange multiplier 2-form. Also, $\langle \_ \,, \_ \rangle_{\mathfrak{g}}$ denotes a G-invariant nondegenerate symmetric bilinear form over $\mathfrak{g}$.

Varying the action (2) with respect to B and α, one obtains the equations of motion:

$$\mathcal{F} = 0 \,, \qquad \nabla B \equiv \mathrm{d}B + \alpha \wedge B = 0 \,. \tag{3}$$

The first equation implies that α is a flat connection, in the sense that $\alpha = 0$ up to gauge transformations. The second equation then implies that B is covariantly constant. From these one can deduce that there are no local propagating degrees of freedom, and therefore the theory is said to be *topological*.

2.2. 2BF theory

Once we have introduced the BF model, we procceed to first step of the *categorical ladder*, generalizing the algebraic notion of a group to the notion of a 2-group. This leads to the generalization of the BF theory to the $2BF$ theory, also sometimes called $BFCG$ theory [11, 17, 18, 19].

The categorical ladder is a procedure of generalizing various notions in mathematics, using the framework of category theory, and works as follows. One starts from the notion of a group as an algebraic structure, and notes that it can be understood as a category with only one object and invertible morphisms [11]. Then, one employs the fundamental idea that a category can be generalized to the so-called *higher categories*, which have not only objects and morphisms, but also 2-morphisms (maps between morphisms), 3-morphisms (maps between 2-morphisms), and so on. This tower of n-categories is known as the *categorical ladder*. Applying the construction to groups, it is straightforward to introduce the notion of a *2-group* as a 2-category consisting of only one object, where all the morphisms and all 2-morphisms are invertible. It was demonstrated that every strict 2-group is equivalent to a *crossed module* $(H \stackrel{\partial}{\to} G \,, \triangleright)$, see [13] for detailed definitions. Here G and H are groups, ∂ is a homomorphism from H to G, while $\triangleright : G \times H \to H$ is an action of G on H.

Just like an ordinary Lie group G has a naturally associated connection α and gives rise to a BF theory, a Lie 2-group has a naturally associated 2-connection $(\alpha \,, \beta)$, described by the usual $\mathfrak{g}$-valued 1-form $\alpha \in \Lambda^1(\mathcal{M}_4) \otimes \mathfrak{g}$ and an $\mathfrak{h}$-valued 2-form $\beta \in \Lambda^2(\mathcal{M}_4) \otimes \mathfrak{h}$, where $\mathfrak{h}$ is a

Lie algebra of the Lie group H. This 2-connection gives rise to the so-called *fake 2-curvature* $(\mathcal{F},\mathcal{G})$, defined as

$$\mathcal{F} = \mathrm{d}\alpha + \alpha\wedge\alpha - \partial\beta\,, \qquad \mathcal{G} = d\beta + \alpha\wedge^{\triangleright}\beta\,. \tag{4}$$

Here $\alpha\wedge^{\triangleright}\beta$ means that α and β are multiplied as forms using $\wedge$, and simultaneously multiplied as algebra elements using $\triangleright$, see [13]. The curvature pair $(\mathcal{F},\mathcal{G})$ is called "fake" due of the presence of the additional term $\partial\beta$ in the definition of $\mathcal{F}$ [11].

Using the structure of a 2-group, or equivalently the crossed module, one can introduce the so-called $2BF$ action, as a generalization of the BF action, as follows [17, 18]:

$$S_{2BF} = \int_{\mathcal{M}_4} \langle B\wedge\mathcal{F}\rangle_{\mathfrak{g}} + \langle C\wedge\mathcal{G}\rangle_{\mathfrak{h}}\,. \tag{5}$$

Here the 2-form $B\in\Lambda^2(\mathcal{M}_4)\otimes\mathfrak{g}$ and the 1-form $C\in\Lambda^1(\mathcal{M}_4)\otimes\mathfrak{h}$ are Lagrange multipliers. Also, $\langle\_\,,\_\rangle_{\mathfrak{g}}$ and $\langle\_\,,\_\rangle_{\mathfrak{h}}$ denote the G-invariant nondegenerate symmetric bilinear forms over the Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively. As a consequence of the axiomatic structure of a crossed module (see [13]), the bilinear form $\langle\_\,,\_\rangle_{\mathfrak{h}}$ is H-invariant as well. See [17, 18] for review and references.

The equations of motion for a $2BF$ theory are an extension of the equations of motion of a BF theory. Varying with respect to B and C one obtains

$$\mathcal{F}^{\alpha} = 0\,, \qquad \mathcal{G}^{a} = 0\,, \tag{6}$$

while varying with respect to α and β one obtains the equations for the multipliers,

$$\nabla B + C\wedge^{\mathcal{T}}\beta = 0\,, \qquad \nabla C - \partial B = 0\,. \tag{7}$$

Here the map $\mathcal{T}$ is defined in [13]. A rigorous Hamiltonian analysis of the model demonstrates that in this case as well there are no local propagating degrees of freedom [20, 21] (see also [22]). Therefore the $2BF$ theory is also topological.

2.3. 3BF theory

When constructing more realistic (nontopological) models by adding constraints to BF and $2BF$ models, it becomes apparent that the group G with a constrained BF action can successfully describe ordinary gauge vector bosons, while the so-called Poincaré 2-group with a constrained $2BF$ action can successfully describe general relativity. However, neither of these can suitably accomodate matter fields, such as fermions or scalars. Nevertheless, it turns out that this can be remedied if we make one further step in the categorical ladder, passing from the notion of a 2-group to the notion of a 3-group. As we shall see in the next Section, the notion of a 3-group will prove to be an excellent structure for the description of all fields that are present in the Standard Model, coupled to Einstein-Cartan gravity. Moreover, a 3-group contains one more gauge group, which is novel and specifies the spectrum of scalar and fermion fields present in the theory. This is an unexpected and beautiful result, absent from ordinary gauge theory.

Applying the categorical ladder once more, one can introduce the notion of a 3-group in the framework of higher category theory, as a 3-category with only one object where all the morphisms, 2-morphisms and 3-morphisms are invertible. Also, the equivalence between a 2-group and a crossed module has been generalized to the equivalence between a strict 3-group and a 2-*crossed module* [23]. A Lie 2-crossed module, denoted as $(L\stackrel{\delta}{\to}H\stackrel{\partial}{\to}G\,,\triangleright\,,\{\_\,,\_\})$, is an algebraic structure specified by three Lie groups G, H and L, together with the homomorphisms δ and ∂, an action $\triangleright$ of the group G on all three groups, and a G-equivariant map

$$\{\_\,,\_\}: H\times H\to L\,.$$

called the Peiffer lifting. The maps ∂, δ, $\triangleright$ and the Peiffer lifting satisfy certain axioms, so that the resulting structure is equivalent to a 3-group [13].

Based on a given 2-crossed module $(L \overset{\delta}{\to} H \overset{\partial}{\to} G\,, \triangleright\,, \{\_\,,\_\})$, one can introduce a gauge invariant topological $3BF$ action over the manifold $\mathcal{M}_4$ as follows. Denoting $\mathfrak{g}$, $\mathfrak{h}$ and $\mathfrak{l}$ as Lie algebras corresponding to the groups G, H and L, respectively, the Lie 3-group structure allows one to introduce a 3-connection (α, β, γ) given by the algebra-valued differential forms $\alpha \in \Lambda^1(\mathcal{M}_4) \otimes \mathfrak{g}$, $\beta \in \Lambda^2(\mathcal{M}_4) \otimes \mathfrak{h}$ and $\gamma \in \Lambda^3(\mathcal{M}_4) \otimes \mathfrak{l}$. The corresponding fake 3-curvature $(\mathcal{F}\,, \mathcal{G}\,, \mathcal{H})$ is then defined as

$$\begin{gathered} \mathcal{F} = \mathrm{d}\alpha + \alpha \wedge \alpha - \partial\beta\,, \qquad \mathcal{G} = \mathrm{d}\beta + \alpha \wedge^{\triangleright} \beta - \delta\gamma\,, \\ \mathcal{H} = \mathrm{d}\gamma + \alpha \wedge^{\triangleright} \gamma + \{\beta \wedge \beta\}\,, \end{gathered} \tag{8}$$

see [23, 24] for details. Note that γ is a 3-form, while its corresponding field strength $\mathcal{H}$ is a 4-form, requiring that the spacetime manifold be at least 4-dimensional. Also, for this reason, going beyond 3-groups and 4-groups in the categorical ladder does not have many applications in realistic 4-dimensional physics. A $3BF$ action is defined as

$$S_{3BF} = \int_{\mathcal{M}_4} \langle B \wedge \mathcal{F} \rangle_{\mathfrak{g}} + \langle C \wedge \mathcal{G} \rangle_{\mathfrak{h}} + \langle D \wedge \mathcal{H} \rangle_{\mathfrak{l}}\,, \tag{9}$$

where $B \in \Lambda^2(\mathcal{M}_4) \otimes \mathfrak{g}$, $C \in \Lambda^1(\mathcal{M}_4) \otimes \mathfrak{h}$ and $D \in \Lambda^0(\mathcal{M}_4) \otimes \mathfrak{l}$ are Lagrange multipliers valued in the respective algebras. Note that exclusively in 4 spacetime dimensions the Lagrange multiplier D corresponding to $\mathcal{H}$ is a 0-form, i.e. a scalar function. As before, the bilinear forms $\langle\_\,,\_\rangle_{\mathfrak{g}}$, $\langle\_\,,\_\rangle_{\mathfrak{h}}$ and $\langle\_\,,\_\rangle_{\mathfrak{l}}$ are G-invariant, nondegenerate and symmetric, over the algebras $\mathfrak{g}$, $\mathfrak{h}$ and $\mathfrak{l}$, respectively.

The equations of motion can be obtained by varying the action with respect to the multipliers B, C and D,

$$\mathcal{F} = 0\,, \qquad \mathcal{G} = 0\,, \qquad \mathcal{H} = 0\,, \tag{10}$$

and by varying with respect to the connections α, β and γ,

$$\nabla B + C \wedge^{\mathcal{T}} \beta - D \wedge^{\mathcal{S}} \gamma = 0\,, \qquad \nabla C - \partial B - D \wedge^{(\mathcal{X}_1 + \mathcal{X}_2)} \beta = 0\,, \qquad \nabla D + \delta C = 0\,. \tag{11}$$

See [13] for the detailed definitions of the maps $\mathcal{T}$, $\mathcal{S}$, $\mathcal{X}_1$ and $\mathcal{X}_2$.

3. The Standard Model 3-group

At this point we are finally ready to construct a realistic classical action, featuring the full Standard Model of elementary particles coupled to Einstein-Cartan gravity. The action is based on a so-called Standard Model 3-group, which is a 2-crossed module $(\, L \overset{\delta}{\to} H \overset{\partial}{\to} G\,,\, \triangleright\,,\, \{\_\,,\_\}\,)$ with a following choices for the Lie groups:

$$G = SO(3,1) \times SU(3) \times SU(2) \times U(1)\,, \qquad H = \mathbb{R}^4\,,$$

$$L = \mathbb{C}^4 \times \mathbb{G}^{64} \times \mathbb{G}^{64} \times \mathbb{G}^{64}\,.$$

We choose the group G as a product of the Lorentz group and the usual internal gauge symmetry group of the Standard Model. The group H is chosen to be the group of spacetime translations, motivated by the Poincaré 2-group construction [12]. Finally, we choose the group L as a product of $\mathbb{C}^4$ accouning for the doublet of complex scalar fields, and three copies of the 64-dimensional Grassmann algebra $\mathbb{G}^{64}$, representing three families of fermions. The maps δ, ∂ and $\{\_\,,\_\}$ are trivial, while the map $\triangleright$ is chosen in a natural way, in accord with the usual action of the gauge group G onto translations and various components of matter fields. It is defined in detail in [13].

Once the 3-group has been completely specified, the corresponding action can be written as a $3BF$ action with suitable constraint terms, as follows:

$$
\begin{array}{rcll}
S & = & \displaystyle\int \overbrace{B_\alpha \wedge F^\alpha + B^{[ab]} \wedge R_{[ab]}}^{\langle B\wedge\mathcal{F}\rangle} + \overbrace{e_a \wedge \nabla\beta^a}^{\langle C\wedge\mathcal{G}\rangle} + \overbrace{\phi^A (\nabla\gamma)_A + \bar\psi_A (\overset{\rightarrow}{\nabla}\gamma)^A - (\bar\gamma \overset{\leftarrow}{\nabla})_A \psi^A}^{\langle D\wedge\mathcal{H}\rangle} & 3BF \\
& - & \displaystyle\int \lambda_{[ab]} \wedge \left(B^{[ab]} - \frac{1}{16\pi l_p^2}\, \varepsilon^{[ab]cd}\, e_c \wedge e_d \right) + \frac{1}{96\pi l_p^2}\, \Lambda\, \varepsilon_{abcd}\, e^a \wedge e^b \wedge e^c \wedge e^d & \text{GR and CC} \\
& + & \displaystyle\int \lambda^\alpha \wedge \left(B_\alpha - 12\, C_\alpha{}^\beta M_{\beta ab}\, e^a \wedge e^b \right) + \zeta^{\alpha ab} \left(M_{\alpha ab}\, \varepsilon_{cdef}\, e^c \wedge e^d \wedge e^e \wedge e^f - F_\alpha \wedge e_a \wedge e_b \right) & \text{YM} \\
& + & \displaystyle\int \lambda^A \wedge \left(\gamma_A - H_{abcA}\, e^a \wedge e^b \wedge e^c \right) + \Lambda^{abA} \wedge \left(H_{abcA}\, \varepsilon^{cdef}\, e_d \wedge e_e \wedge e_f - (\nabla\phi)_A \wedge e_a \wedge e_b \right) & \text{Higgs} \\
& - & \displaystyle\int \frac{1}{12}\, \chi \left(\phi^A \phi_A - v^2 \right)^2 \varepsilon_{abcd}\, e^a \wedge e^b \wedge e^c \wedge e^d & \text{Higgs potential} \\
& + & \displaystyle\int \bar\lambda_A \wedge \left(\gamma^A + \frac{i}{6}\, \varepsilon_{abcd}\, e^a \wedge e^b \wedge e^c \left(\gamma^d \psi \right)^A \right) - \lambda^A \wedge \left(\bar\gamma_A - \frac{i}{6}\, \varepsilon_{abcd}\, e^a \wedge e^b \wedge e^c \left(\bar\psi \gamma^d \right)_A \right) & \text{Dirac} \\
& - & \displaystyle\int \frac{1}{12}\, Y_{ABC}\, \bar\psi^A \psi^B \phi^C\, \varepsilon_{abcd}\, e^a \wedge e^b \wedge e^c \wedge e^d & \text{Yukawa} \\
& + & \displaystyle\int 2\pi i\, l_p^2\, \bar\psi_A \gamma_5 \gamma^a \psi^A\, \varepsilon_{abcd}\, e^b \wedge e^c \wedge \beta^d\,. & \text{spin-torsion}
\end{array}
$$

Here the first row represents the topological $3BF$ part, while the remaining rows represent various constraint terms, each corresponding to one sector of the theory. Taking all together, the equations of motion obtained from the action S are equivalent to the full set of equations of motion for all Standard Model fields, coupled to the Einstein-Cartan theory of gravity.

The key novelty of the above structure is the role of the group L, which prescribes the spectrum of scalar and fermion fields present in the theory, via the $\langle D\wedge\mathcal{H}\rangle$ term in the topological sector of the action.

4. Conclusions

Let us summarize the results of the paper. In Section 2 we have introduced the nBF theories for $n = 1, 2, 3$, and explanied in brief terms how the categorical ladder procedure can be applied to generalize the notion of a group to the notions of a 2-group and a 3-group, which represent more powerful ways to describe the gauge symmetry of a physical theory. These structures were employed in Section 3 to construct the constrained $3BF$ action for the Standard Model of elementary particles coupled to the Einstein-Cartan gravity in the usual way. Within that framework, the spectrum of scalar and fermion fields happens to be determined by a *new gauge group*, in a way similar to that of the ordinary gauge group determining the spectrum of gauge vector bosons in Yang-Mills theory. This opens up a very interesting possibility of applying the structure of a 3-group to classify matter fields, and possibly gain some insight into why there are three families of fermions.

These results complete the first step of the spinfoam quantization programme, as outlined in the Introduction. The second step has also been performed in [25], for a general case of a Lie 2-crossed module $(L \overset{\delta}{\to} H \overset{\partial}{\to} G\,, \triangleright\,, \{\_\,,\_\})$. The resulting state sum is a novel topological invariant of a 4-dimensional manifold, and has the following form:

$$
\begin{aligned}
Z = {} & |G|^{-|\Lambda_0|+|\Lambda_1|-|\Lambda_2|}|H|^{|\Lambda_0|-|\Lambda_1|+|\Lambda_2|-|\Lambda_3|}\,|L|^{-|\Lambda_0|+|\Lambda_1|-|\Lambda_2|+|\Lambda_3|-|\Lambda_4|} \\
& \times \prod_{(jk)\in\Lambda_1} \int_G dg_{jk} \prod_{(jk\ell)\in\Lambda_2} \int_H dh_{jk\ell} \prod_{(jk\ell m)\in\Lambda_3} \int_L dl_{jk\ell m} \\
& \times \prod_{(jk\ell)\in\Lambda_2} \delta_G\Big(\partial(h_{jk\ell})\, g_{k\ell}\, g_{jk}\, g_{j\ell}^{-1}\Big) \prod_{(jk\ell m)\in\Lambda_3} \delta_H\Big(\delta(l_{jk\ell m})h_{j\ell m}\,(g_{\ell m}\rhd h_{jk\ell})\, h_{k\ell m}^{-1}\, h_{jkm}^{-1}\Big) \\
& \times \prod_{(jk\ell mn)\in\Lambda_4} \delta_L\Big(l_{j\ell mn}^{-1}\, h_{j\ell n} \rhd' \{h_{\ell mn}, (g_{mn}g_{\ell m)} \rhd h_{jk\ell}\}_{\mathrm{p}}\, l_{jk\ell n}^{-1}(h_{jkn}\rhd' l_{k\ell mn})l_{jkmn}h_{jmn}\rhd' (g_{mn}\rhd l_{jk\ell m})\Big).
\end{aligned}
\tag{12}
$$

Here $g_{ij}, h_{ijk}, l_{ijkl}$ are elements from groups G, H, L, respectively, which are assigned to simplices of the triangulation whose vertices are numerated by indices $i, j, \ldots$ In other words, g_{ij} are assigned to edges, h_{ijk} are assigned to triangles, and l_{ijkl} are assigned to tetrahedra of the simplicial complex representing a compact 4-manifold, which has a total number of Λ_0 vertices, Λ_1 edges, Λ_2 triangles, Λ_3 tetrahedra, and Λ_4 4-simplices.

Of course, when building a realistic theory, we are in fact not interested in a topological theory, but instead in a theory which contains local propagating degrees of freedom. Thus the state sum Z should be appropriately deformed. This is the task of step 3 of the spinfoam quantization programme, by imposing the simplicity constraints on Z. The classical action from Section 3 manifestly distinguishes the topological sector from the simplicity constraints. Imposing those constraints should thus complete the spinfoam quantization programme, and would ultimately lead us to a tentative model of quantum gravity with matter, by providing a rigorous definition for the path integral

$$Z = \int \mathcal{D}g \int \mathcal{D}\phi \; e^{iS[g,\phi]} \,, \tag{13}$$

which is a generalization of (1) in the sense that it contains matter fields as well as gravity, at the quantum level.

In addition to the construction of a full quantum theory of gravity, there are also many additional possible studies of the classical constrained $3BF$ action. For example, a full Hamiltonian analysis of the $3BF$ action has been done for the example of scalar electrodynamics [26], and then also for a general choice of a Lie 3-group [27], and the complete gauge symmetry group has been discussed in detail [27, 28]. Also, it is worth looking into the idea of imposing the simplicity constraints using a spontaneous symmetry breaking mechanism, and some work has already begun in this area. Finally, one can also study in more depth the mathematical structure and properties of the simplicity constraints. The list is not conclusive, and there may be many other interesting topics to study.

Acknowledgments
This research was supported by the Ministry of Science, Technological Development and Innovations of the Republic of Serbia, and by the Science Fund of the Republic of Serbia, grant 7745968, "Quantum Gravity from Higher Gauge Theory 2021" — QGHG-2021. The contents of this publication are the sole responsibility of the authors and can in no way be taken to reflect the views of the Science Fund of the Republic of Serbia.

References

[1] Rovelli C 2004 *Quantum Gravity* (Cambridge University Press: Cambridge)
[2] Rovelli C and Vidotto F 2014 *Covariant Loop Quantum Gravity* (Cambridge University Press: Cambridge)
[3] Thiemann T 2007 *Modern Canonical Quantum General Relativity* (Cambridge University Press: Cambridge)
[4] Ponzano G and Regge T 1968 *Semiclassical limit of Racah coefficients* (Spectroscopic and Group Theoretical Methods in Physics) ed F Bloch (Amsterdam: North Holland)

[5] Barrett J W and Crane L 1998 *J. Math. Phys.* **39** 3296
[6] Barrett J W and Crane L 2000 *Class. Quant. Grav.* **17** 3101
[7] Ooguri H 1992 *Mod. Phys. Lett. A* **7** 279
[8] Engle J, Livine E R, Pereira R and Rovelli C 2008 *Nucl. Phys.* **B799** 136
[9] Freidel L and Krasnov K 2008 *Class. Quant. Grav.* **25** 125018
[10] Bianchi E, Han M, Magliaro E, Perini C, Rovelli C and Wieland W 2013 *Class. Quantum Grav.* **30** 235023
[11] Baez J C and Huerta J 2011 *Gen. Relativ. Gravit.* **43** 2335
[12] Miković A and Vojinović M 2012 *Class. Quant. Grav.* **29** 165003
[13] Radenković T and Vojinović M 2019 *JHEP* **10** 222
[14] Celada M, González D and Montesinos M 2016 *Class. Quant. Grav.* **33** 213001
[15] Rovelli C 2011 Zakopane lectures on loop gravity *Preprint* arXiv:1102.3660
[16] Plebanski J F 1977 *J. Math. Phys.* **18** 2511
[17] Girelli F, Pfeiffer H and Popescu E M 2008 *Jour. Math. Phys.* **49** 032503
[18] Martins J F and Miković A 2011 *Adv. Theor. Math. Phys.* **15** 1059
[19] Crane L and Sheppeard M D 2003 2-categorical Poincare Representations and State Sum Applications *Preprint* math/0306440
[20] Miković A and Oliveira M A 2015 *Gen. Relativ. Gravit.* **47** 58
[21] Miković A, Oliveira M A and Vojinović M 2016 Hamiltonian analysis of the BFCG theory for a generic Lie 2-group *Preprint* arXiv:1610.09621
[22] Miković A, Oliveira M A and Vojinović M 2019 *Class. Quant. Grav.* **36** 015005
[23] Martins J F and Picken R 2011 *Differ. Geom. Appl. Journal* **29** 179
[24] Wang W 2014 *Jour. Math. Phys.* **55** 043506
[25] Radenković T and Vojinović M 2022 *JHEP* **7** 105
[26] Radenković T and Vojinović M 2020 *Symmetry* **12** 620
[27] Radenković T and Vojinović M 2022 *Class. Quant. Grav.* **39** 135009
[28] Djordjević M, Radenković T, Stipsić P and Vojinović M 2023 *Universe* **9** 281

Infrastructure for Simulating n-Dimensional Simplicial Complexes

Dušan Cvijetić, Nenad Korolija, and Marko Vojinović

***Abstract*—We present an infrastructure for simulating simplicial comple es. The classes for storing the structure of simplicial comple es and simplices are e plained in detail, for arbitrary dimension.**

The implementation is tested using functions for seeding simplicial comple es and for printing them on the screen. eside these functions, the supporting classes and the function for assigning unique identifiers and screen coordinates is also e plained.

esults of simulation show that there are potentials for the simulator to be used for big data problems, although appropriate e perimental results are still being collected. uture wor includes parallelizing the e ecution of the simulator using supercomputing architectures.

***Index Terms*— implicial comple ; triangulation; manifold; algebraic topology.**

I. Introduction

A manifold is one of the fundamental concepts in mathematics [1], and its importance in applications in physics, technology and engineering cannot be overstated. Virtually all modern physics describes the world using *field theory* [2], in which all physical quantities (fields) are represented as functions over some manifold (for example, spacetime). In technology, manifolds appear in all forms and guises, whenever one needs to deal with curved surfaces --- from civil engineering to graphics in video games.

While most of the interest in science and engineering revolves around *smooth* manifolds, for the purpose of studying manifolds using numerical techniques, the attention focuses on the so called *piecewise-linear* manifolds [3], which can intuitively be imagined as a structure made out of small flat cells called *simplices,* arranged like bricks into a structure which models a manifold. The procedure of approximating a smooth manifold with a piecewise-linear one is commonly called *triangulation,* see Fig. 1.

Within the framework of algebraic topology, the formal mathematical structure which describes piecewise-linear manifolds is called a *simplicial comple* . For the purpose of this article, we provide an informal descriptive definition of a simplicial complex, without mathematical rigour. A simplicial complex is a combinatorial structure, containing the information about *simplices* of various dimensions that make up a complex, and the information about how simplices are connected to each other. A *k-simple* is an elementary building block of a simplicial complex. It is an elementary geometrical "cell" of dimension k, which is being used to build simplices of higher dimension, and the entire simplicial complex. For $k = 0$, the simplex is called a *erte* , it is represented geometrically as a single point, and has no internal structure. The $k = 1$ simplex is called an *edge,* geometrically represented as a single straight line, having two vertices at its boundary. For $k = 2$, the simplex is a *triangle,* having three boundary edges and three vertices. The case $k = 3$ describes a *tetrahedron*, having four boundary triangles, six edges and four vertices. One can go further into higher dimensions: $k = 4$ represents a simplex called *pentachoron* – it is a 4-dimensional figure, having five boundary tetrahedra, 10 triangles, 10 edges and five vertices. In general, one can introduce a k-simplex for arbitrary dimension k, also called *le el.*

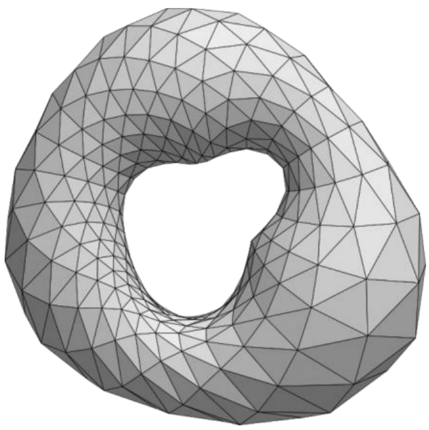


Fig. 1. Simplicial complex of a torus (source: Wikipedia).

Given a set of simplices, one can "glue them up" into a bigger geometrical structure, called simplicial complex. In order to describe a manifold of dimension D, a simplicial complex is constructed by gluing a set of D-simplices by identifying their common boundary (D-1)-simplices. Naturally, this implies the identification of all corresponding sub-simplices of level $k < D-1$ as well. The resulting simplicial complex is homeomorphic to a piecewise-linear manifold of dimension D.

The most important information about the simplicial complex, aside from its dimension D, is the data that tells one

Dušan Cvijetić is a student of the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: dusancvijetic2000 @ gmail.com).

Nenad Korolija is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: nenadko @ etf.bg.ac.rs).

Marko Vojinović is with the Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Pregrevica, Serbia (e-mail: vmarko @ ipb.ac.rs).

which simplices are glued to which. This gives rise to a notion of a *neighborhood* of a k-simplex, which is a set of all simplices which contain a given simplex as its sub-simplex (called super-neighbors) and simplices which are contained in a given simplex (called sub-neighbors). Each k-simplex (for $0 \leq k \leq D$) in the complex has its set of neighbors, where by definition a simplex is not a neighbor of itself (this is convenient to avoid infinite loops when traversing a complex). The neighborhood structure of the entire complex determines the *topology* of the corresponding manifold.

While manifolds of various topologies are important in their own right in mathematics, the applications in physics and engineering typically introduce functions over manifolds, such as distances, areas and volumes, temperature, electric and magnetic fields, etc. In the language of simplicial complexes, these functions are commonly called *colors*, and are assigned to simplices of various level k within the complex. Given a k-simplex, one can assign to it multiple colors, representing the value of a given function when evaluated on the k-simplex. A prototype example of colors is the geometry of a simplicial complex: each k-simplex is assigned its "size" according to its geometry --- each 1-simplex (an edge) is assigned a real number representing its length, each 2-simplex (a triangle) is assigned a real number representing its area, tetrahedra are assigned volumes, and so on. Other examples are abound --- vertices can be assigned a temperature, edges can be assigned vectors of electric field, and so on. Depending on the problem at hand, one may or may not impose relationships between various colors, such as that the area of a triangle is consistent with the length of its edges, or similar. These relationships are collectively called *constraints*.

In most everyday applications, one is interested in manifolds of dimension 1 and 2 (curves and surfaces). However, within the context of theoretical physics, one often needs to deal with manifolds of higher dimension – most commonly 3, 4, 5, 10, 11 and 26, while more sporadically anything in between and above. One of the typical scenarios is *uantum gra ity* [4,5], a vast research area of fundamental theoretical physics, where the notion of spacetime is described as a piecewise-linear manifold of dimension D=4 or higher [6,7]. In order to apply numerical techniques to study the manifolds in such research disciplines, it is necessary to formulate and implement structures and algorithms which describe colored simplicial complexes of arbitrarily large dimension, in a uniform and optimal way. In what follows, we describe one such implementation, which is purposefully designed to mimic the mathematical structure of a simplicial complex as close as possible, while simultaneously providing efficient numerical techniques for the manipulation and study of such structures.

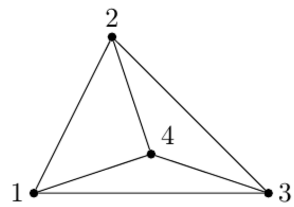
I. N-Dimensional Simplicial Complexes

This section describes the structure of simplicial complexes, and explains an example C++ implementation of classes for storing simplicial complexes.

Simplicial complexes consist of k-simplices at different levels. Given a simplicial complexes of dimension D, these elements include k-simplices for each level from zero to D. Elements at level zero are vertices, elements at level one are edges, elements on level two are triangles, etc. Finally, there are elements of highest level D. The representative source code of class for simplicial complexes is given in Algorithm 3 from the Appendix. The source code is pruned from comments and unnecessary functionalities for the presentation of the simulator.

K-simplex stores the level it has, the dimension of the simplicial complex it belongs to, neighboring elements and colors assigned to it.

Neighboring elements of a k-simplex are defined as k-simplices that this k-simplex is touching. Since these can be on various levels, the structure of neighbors is the same as for the simplicial complex. Therefore, the two main classes are mutually connected.




```
Printing SimpComp tetrahedron, D = 3
Simplices k = 0:
1, 2, 3, 4
Simplices k = 1:
(1-2), (1-3), (1-4), (2-4), (2-3), (3-4)
Simplices k = 2:
(1-2-3), (1-3-4), (1-2-4), (2-3-4)
Simplices k = 3:
(1-2-3-4)
```

Fig. 2. Tetrahedron and a corresponding output of the simplicial complexes simulator.

One possible implementation of the neighboring elements is to store only neighbors from one level above, and one level beneath (first sub-neighbors and first super-neighbors). The lower- and higher-level neighbors can be deduced following the structure of the first neighbors. However, we have opted for storing neighbors from all levels, giving us the opportunity to divide the structure onto multiple computing nodes and run the code in parallel. At current state, the simulator is running on a single CPU.

The instructions a CPU is executing are repeated over and over again, which makes this simulator suitable for acceleration using the dataflow paradigm [8,9]. The effort required for programming such architectures is higher than for conventional von Neumann architectures [10], but the simulator is suitable for transforming the C++ source code automatically [11]. Executing multiple simplicial complex operations in parallel requires appropriate scheduling

 ISBN 978-86-7466-930-3

techniques [12].

Each k-simplex (including all vertices, edges, triangles, etc.) can be colored with different types of color. Example colors include:

- k-simplex name,
- unique identifier of k-simplex,
- boundary color of k-simplex,
- screen coordinates.

These colors are included in our simplicial complex simulator, but the structure of the simulator allows adding additional user defined colors.

The representative source code of the class for k-simplices is given in Algorithm 4 from the Appendix. Just like it is the case with simplicial complexes, this source code is pruned for better clarity.

For simulation purposes, we have developed functions for seeding simplicial complexes at various levels, as it will be explained in the following section. In addition, coloring and printing simplicial complexes is also implemented. Pretty printing (or compact printing) prints k-simplices at all levels, where k-simplices of level higher than zero are printed as tuples consisting of unique identifiers (IDs) of their vertices. Fig. 2 shows an example tetrahedron (i.e. simplicial complex of dimension D = 3 consisting of a single 3-simplex and its sub-simplices) whose vertices are colored with unique identifiers that auto-increment after each assignment of the unique color to a vertex. Details of the implementation of compact printing is also explained in this manuscript.

Screen coordinates can be attached to vertices of the tetrahedron. Therefore, it can be drawn on the screen. However, there is no need to assign coordinates. They are just a convenient way to show an object on a screen. Similarly, there is no need to assign unique ID to any vertex. In the previous example, if a vertex with unique ID four would not have a unique ID assigned to it, the tetrahedron could still be printed out, but with word "Simplex" being printed out in place of number four.

II. Seeding Simplicial Complexes

This section describes seeding simplicial complexes using C++ implementation of function *seed single edge* . The example source code for seeding a single edge is used for demonstrating purposes.

The process of seeding simplicial complexes will be explained using the source code shown in Algorithm 1. The source code is pruned from comments and unnecessary statements. Seeding a simplicial complex consists of the following steps, and statements in Algorithm 1 follow the same principle in the same order:

- creating an empty simplicial complex of given dimension,
- creating k-simplices for storing vertices and simplices of higher levels,
- connecting vertices at each level with vertices on higher and lower levels.

Adding a neighbor to a k-simplex is a symmetric operation. This means that both k-simplices (the calling one and the one given as an argument) are neighbors to each other. All functions of the simulator are written in a robust manner, checking the validity of input parameters.

Note that multiple colors can be assigned to each k-simplex, which is left out of consideration in this algorithm for better clarity.

III. Coloring and Pretty Printing K-simplices

This section describes coloring and pretty printing simplicial complexes. These functions might work in pair, but are not necessarily connected.

A. Coloring -simplices

Coloring k-simplices will be explained using Algorithm 2 by coloring vertices of an edge with boundary colors. First, vertices have to be created as k-simplices of level zero. Then, colors have to be created for all vertices. Finally, colors need to be pushed back to the vector of colors that each k-simplex has.

Algorithm 1: Seeding a single edge.

```
SimpComp* seed_single_edge(string name){
    SimpComp *edge = new SimpComp(
            name, 1);
    KSimplex *v1 =
            edge->create_ksimplex(0);
    KSimplex *v2 =
            edge->create_ksimplex(0);
    KSimplex *e1 =
            edge->create_ksimplex(1);
    v1->add_neighbor(e1);
    v2->add_neighbor(e1);
    return edge;
}
```

Algorithm 2: Coloring vertices with boundary color.

```
KSimplex *v1 =
        edge->create_ksimplex(0);
KSimplex *v2 =
        edge->create_ksimplex(0);
Color *c1 = new BoundaryColor(true);
Color *c2 = new BoundaryColor(true);
v1->colors.push_back(c1);
v2->colors.push_back(c2);
```

Following colors are currently available:

- unique ID colors
- boundary colors
- screen coordinate colors.

Additionally, user is allowed to construct a custom color and use it within the simulator. The source code of the simulator is organized as a library, and user is allowed to extend it by using the library.

Unique ID colors are predominantly used for pretty printing simplicial complexes. They are implemented by a class inherited from the basic color class. Two main fields include

 ISBN 978-86-7466-930-3

static integer number, and an integer number. The first represents the current maximum of a unique color ID that is in use, and the second one is the color of a given k-simplex.

Unlike unique ID colors, boundary colors have special meaning. Each k-simplex may contain boundary color, but it does not have to. A simplicial complex can have boundaries on k-simplices of one level lower than the dimension of the simplicial complex. For example, a triangle can have edges as boundaries.

Screen coordinate colors are used for drawing simplicial complexes on a screen. The basic graphical user interface is under development.

B. Pretty Printing -simplices

Printing k-simplices includes printing of all of the fields that *Simple* class contains. This includes printing all of the neighborhood elements the k-simplex has. This is usually overwhelming for a user. Therefore, pretty printing is designed to print unique ID colors of each k-simplex in most readable way authors could think of.

Function *Simple print compact* is responsible for pretty printing. It assigns to the pointer to the unique ID a value returned by a function *get uni ueI* that returns either nullptr if a k-simplex doesn't have a unique ID, or a pointer to the color.

If there is no unique ID color assigned to a k-simplex, the output consists solely of word "Simplex". Otherwise, *print compact* function is called for a color that the pointer points to. Further, the following procedure is repeated, if level k is greater than zero and there are neighboring elements for all neighbors. A set of integer values is constructed, and then function *print ertices in parentheses s* is called for neighbors, adding unique IDs to the set. This way, printing sorted values is achieved, along with avoiding duplicate values. Sample output of a simplicial complex pretty printing is shown in Fig. 1.

IV. CONCLUSION

We have demonstrated how one can implement in code the structure of a simplicial complex of arbitrary dimension, in a way that is faithful to its combinatorial definition, and perform the most basic operations on it, like instantiating, coloring and printing.

The implementation of the basic classes of the code described in this work represents a fundamental basic building block for a more versatile software collection that aims to construct, manipulate and study the properties of simplicial complexes of arbitrary dimension. Future extensions of the software library will include the functions which implement attaching additional simplices to a boundary of a complex, performing Pachner moves [13] which transform a given complex into a different one without changing its topology, and functions for manipulating the colors and evaluating various mathematical constructions that include them. Note that the experimental data regarding the parallelization is yet to be collected (see the accompanying paper [14]).

The resulting software collection will feature the generality and versatility that aim for applications both in pure mathematics (algebraic topology research) and theoretical physics (quantum gravity, field theory), but also with potential applications in other disciplines of engineering and industry, wherever the analysis and the study of geometry of manifolds and curved surfaces may be relevant.

APPENDIX

Algorithm 3: Declaration of SimpComp class.

```
class SimpComp{
public:
    SimpComp(int dim);
    SimpComp(string s, int dim);
    ~SimpComp();
    int count_number_of_simplexes(
            int level);
    void print(string space = "");
    bool all_uniqueID(int level);
    void collect_vertices(set<int> &s);
    void print_set(set<int> &s);
    void print_vertices_in_parentheses(
            set<int> &s);
    void print_compact();
    // Creating new KSimplex at level k:
    KSimplex* create_ksimplex(int k);
    void print_sizes();

    string name;
    int D;
    // An element at each level
    // is a list or vector
    // of KSimplex pointers
    // to KSimplex on that level:
    vector< vector<KSimplex *> >
            elements;
};
```

Algorithm 4: Declaration of KSimplex class.

```
class KSimplex{
public:
    KSimplex();
    KSimplex(int k, int D);
    ~KSimplex();
    bool find_neighbor(KSimplex *k1);
    void add_neighbor(KSimplex *k1);
    void print(string space = "");
    UniqueIDColor* get_uniqueID();
    void print_compact();

    int k; // level
    int D; // dimension
    SimpComp *neighbors;
    vector<Color *> colors;
};
```

ACKNOWLEDGMENT

DC and NK were partially supported by the School of

 ISBN 978-86-7466-930-3

Electrical Engineering, University of Belgrade, Serbia. NK was partially supported by the Institute of Physics Belgrade, contract no. 0801-1264/1. MV was supported by the Science Fund of the Republic of Serbia, grant no. 7745968, “Quantum gravity from higher gauge theory” – QGHG-2021. All authors were partially supported by the Ministry of Education, Science, and Technological Development of the Republic of Serbia.

REFERENCES

[1] M. W. Hirsch, *ifferential Topology*, New York, USA: Springer Verlag, 1976.

[2] A. Hobson, “There are no particles, there are only fields”, *Amer. our. Phys.* **1**, 211-223 (2013).

[3] E. H. Spanier, *Algebraic Topology*, New York, USA: Springer Verlag, 1966.

[4] C. Rovelli, *uantum Gra ity*, Cambridge, UK: Cambridge University Press, 2004.

[5] C. Rovelli and F. Vidotto, *Co ariant oop uantum Gra ity*, Cambridge, UK: Cambridge University Press, 2014.

[6] T. Radenković and M. Vojinović, “Higher Gauge Theories Based on 3-Groups”, *EP* **1** , 222 (2019).

[7] A. Miković and M. Vojinović, “Standard Model and 4-Groups”, *Europhys. ett.* **1** , 61001 (2021).

[8] B. Lee and A. R. Hurson, “Issues in dataflow computing,” *Ad ances in computers*, Elsevier, , 285-333 (1993).

[9] V. Milutinovic, J. Salom, D. Veljovic, N. Korolija, D. Markovic, and L. Petrovic, “Transforming applications from the control flow to the dataflow paradigm,” *ataflow supercomputing essentials*, Springer, Cham, 107-129 (2017).

[10] J. Popovic, D. Bojic, and N. Korolija, “Analysis of task effort estimation accuracy based on use case point size,” *IET Software*, **( )**, 166-173 (2015).

[11] N. Korolija, J. Popović, M. Cvetanović, and M. Bojović, “Dataflow-based parallelization of control-flow algorithms,“ *Ad ances in computers*, Elsevier, **1 4**, 73-124 (2017).

[12] N. Korolija, D. Bojić, A. R. Hurson, and V. Milutinovic, “A runtime job scheduling algorithm for cluster architectures with dataflow accelerators,” *Ad ances in computers*, Elsevier, **12** (2022).

[13] U. Pachner, “PL homeomorphic manifolds are equivalent by elementary shellings”, *Eur. our. Combinat.* **12**, 129-145 (1991).

[14] D. Cvijetić, N. Korolija and M. Vojinović, “Possibilities for Parallelizing Simplicial Complexes Simulation”, IcETRAN 2022, Novi Pazar, Republic of Serbia, June 6-9, 2022, Belgrade: Društvo za ETRAN, Beograd: Akademska misao (2022).

 ISBN 978-86-7466-930-3

Possibilities for Parallelizing Simplicial Complexes Simulation

Dušan Cvijetić, Nenad Korolija, and Marko Vojinović

***Abstract*—This manuscript presents potentials for parallelizing simulation of simplicial comple es. The implementation of most important fields and methods of classes for storing simplicial comple es and -simplices is followed by wrapper classes for simplicial comple es and -simplices respectively. Infrastructure for communication between essage Passing Interface (PI) processes along with helper functions is e plained further in the manuscript. nce multiple data are prepared to be sent from each PI process to other PI processes, sending and receiving is performed in the bac ground. ecause of the stall introduced by using PI directives, the amount of data to be transmitted is ma imized by processing multiple operations over simplicial comple es in parallel. This requires the method for loc ing simplicial comple es and -simplices by the owner PI process until all the requests are processed. oc ing mechanism and supporting simplicial comple class actions regarding loc ing is not in the scope of this manuscript.**

***Index Terms*— implicial comple ; -simple ; triangulation; manifold; PI; parallelization.**

I. Introduction

In modern theoretical physics, a lot of problems are too complicated for study using analytical methods, and one needs to resort to numerical techniques. Among those problems, an especially important class deals with evaluation of functions over simplicial complexes. A simplicial complex [1] is a piecewise-linear approximation of a smooth spacetime manifold [2] and is typically 4-dimensional or higher. Functions over a simplicial complex represent physical fields on spacetime, and one commonly employs path integral evaluations of such structures to extract expectation values of observables. For example, in Lattice Quantum Chromo-dynamics, one employs such numerical techniques to predict the theoretical values for the masses of elementary particles called hadrons [3]. Also, in Causal Dynamical Triangulations approach to quantum gravity [4,5], one uses these techniques to evaluate spectral dimension of spacetime, and study various properties of phase space of triangulated manifolds. Finally, in the Regge Quantum Gravity approach [6,7,8] one can study the entanglement properties of matter fields and gravity described by the Hartle-Hawking wavefunction [9,10], again using the techniques of numerical evaluation of path integrals over simplicial complexes.

It goes without saying that all such calculations are exceptionally expensive in computation time. Typically, one develops custom-made code, heavily optimized to solve precisely one specific problem, and executes it over months-long periods on hardware dedicated for high performance computing (HPC), usually clusters with thousands of work nodes. Such enormous calculational efforts are usually unavoidable due to the nature of the problems that need to be solved.

Nevertheless, at least for one class of such problems, it may be possible to construct a more general algorithm and structures which would provide a common basis for solving an all-encompassing class of problems using the same underlying software, while intrinsically exploiting the parallelization possibilities of the code itself and the distributed nature of the underlying hardware. Our aim is to develop such a generic software library, which could be used to solve a whole host of physics problems in the same way and optimize it for parallelized HPC environments. In this work we present the first steps towards the construction of such a library. This approach of developing common code for a whole class of problems has not been attempted so far because research teams are usually concentrated on solving only one specific problem and opt to construct custom code for that problem. However, in our opinion, a generic software library, which would provide support for a whole class of problems simultaneously, would open new avenues for numerical research, since one could use the same code to study new, yet unexplored problems as well as old well-known ones.

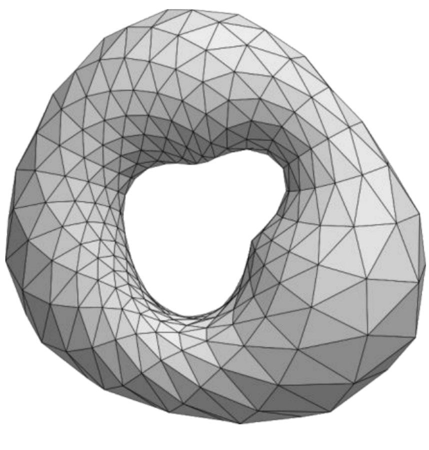


Fig. 1. Simplicial complex of a torus (source: Wikipedia).

The fundamental structure which lies at the core of the whole numerical method is the notion of a *simplicial comple* . A simplicial complex is a combinatorial structure which is easiest to understand as a generic lattice-like mesh,

Dušan Cvijetić is a student of the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: dusancvijetic2000 @ gmail.com).

Nenad Korolija is with the School of Electrical Engineering, University of Belgrade, 73 Bulevar kralja Aleksandra, 11020 Belgrade, Serbia (e-mail: nenadko @ etf.bg.ac.rs).

Marko Vojinović is with the Institute of Physics, University of Belgrade, Pregrevica 118, 11080 Pregrevica, Serbia (e-mail: vmarko @ ipb.ac.rs).

 ISBN 978-86-7466-930-3

whose cells are called *simplices*, and are connected to each other along their boundaries to form the simplicial complex of a given dimension. The purpose of the whole structure is to approximate the smooth spacetime manifold with a discrete structure which is more convenient for numerical methods.

The most elementary simplex is a simplex of level zero, often called *0-simple* or *erte* – it is just a dimensionless point with no structure. Next is the *1-simple* , also called an *edge* – it is a one-dimensional line with two vertices at its boundaries. At level two we have the *2-simple* or *triangle*, whose boundary are three edges and their vertices. The *3-simple* , also known as the *tetrahedron*, has the boundary made of four triangles and their edges and vertices. The procedure of constructing simplices can be done for arbitrary dimension, giving rise to the notion of a *k-simple* , whose level (i.e. natural dimension of space in which it is defined) is equal to any positive integer k. The most commonly used example is the *4-simple* , also called *pentachoron* – a 4-dimensional figure whose boundary consists of 5 tetrahedra, 10 triangles, 10 edges and 5 vertices. In most applications in physics, the spacetime manifold is considered to be 4-dimensional, and it is cut into a lattice-like structure made of 4-simplices, which are glued together along their boundary tetrahedra. The resulting structure is a simplicial complex of dimension 4. Fig. 1 depicts an intuitive example of a 2-dimensional simplicial complex of a torus.

Given a simplicial complex, one typically wants to introduce functions that are evaluated on it. These are commonly called *colors* and are assigned via their values to each k-simplex within in the complex. In other words, some colors live on vertices, some on edges, some on triangles, and so on. The colors are a natural discretization of the notion of a *field* over a manifold. For example, just like electric and magnetic fields have a value at each point of a smooth spacetime, analogously the colors have values at each k-simplex in the simplicial complex.

Depending on the type of the problem at hand, algorithms that are used to evaluate required quantities on a simplicial complex can vary in complexity, from conceptually simple Monte Carlo integration techniques, to vastly complicated traversal and ray-tracing algorithms, to various methods for solving functional partial differential equations. Due to the variability of the complexity of all these algorithms, dictated by the nature of the problem at hand, it is helpful to develop the underlying software simulator to exploit the parallelization avenues that are intrinsic to the simplicial complexes and k-simplices themselves, so that the simulator can exploit parallel hardware environments even for algorithms that are themselves hard to parallelize. This helps the code developer with overall optimization and application to HPC hardware architectures. In what follows, we shall demonstrate a set of possible approaches to these intrinsic parallelization techniques.

II. N-DIMENSIONAL SIMPLICIAL COMPLEXES

This section describes data structures used in the simulator of simplicial complexes from the point of view of their suitability for parallelizing the simulator execution. Data demanding structures are of main interest for optimizing the communication between processing units. Along with those, data that describes the structure and needs to be updated on multiple processing units will be described in detail. Further, the amount of data that needs to be exchanged and the frequency of expected changes will be compared to the pyramid, where top elements demand less memory, but require more often communication.

The parallelization is simulated using the MPI framework. The simulator is implemented in C++, and, as a result, the parallelization framework is built on top of the simulator. As improving the simulator of simplicial complexes is an ongoing process, the possibility for accelerating the computation is simulated based on the requirements.

Simplicial complexes are formed out of k-simplices at various levels. Simplicial complexes at level zero represent vertices. The structure of each vertex is stored in *Simple* class. Simplicial complexes at level one represent edges. Each edge consists of two vertices. As it is the case with vertices, information about edges are also kept in a *Simple* class. However, while vertices can be independent of other vertices, representing separate simplicial complexes, each edge must have at least two vertices defined as neighbors. Neighbor of an k-simplex is defined also as a k-simplex that the first k-simplex relies on. Neighboring relation is symmetrical. Therefore, if two vertices are neighbors of an edge, edge is also the neighbor of both vertices. Further, edges can form a triangle. By analogy, neighbors of triangle are three edges, but also the triangle is neighbor of these edges. The neighboring relation spans more than one level up or down. The triangle has also three vertices as neighbors and the opposite.

Simplicial complex representing a triangle consists of a k-simplex representing a triangle along with all neighbors of the triangle. Simplicial complex class is used for storing information about simplicial complexes. As it has elements field that is a pointer to pointer of k-simplices, it is also used for keeping neighbors of each k-simplex.

III. PARALLELIZING SIMPLICIAL COMPLEXES SIMULATION

Parallelizing operations over simplicial complexes is implemented by splitting the structure over multiple MPI processes. First, we can consider a single simplicial complex system, as the most general approach. If no screen coordinates for k-simplices are assigned, we can artificially assign this type of color, so that we can present k-simplices in 2D space. Further, we can imagine multiple planes, where each plane is responsible for keeping k-simplices of one dimension. This way, we can consider n-dimensional simplicial complex as a pyramid that we observe from the bird's eye view. Now we could have a bottom-up approach, where k-simplices of dimension zero are divided onto MPI processes based on their screen coordinates. Going up, each MPI process would store higher dimensional k-simplices that have those that are one level below as their neighbors. When a k-simplex has neighbors on one level below that belong to multiple MPI processes, this k-simplex gets copied to all MPI processes involved. Finally, all MPI processes would keep the highest-level k-simplex. In the case of multiple simplicial complexes, they could be split over MPI processes based on the same bottom-up approach.

 ISBN 978-86-7466-930-3

The notion of determining the MPI process where a k-simplex is located is hidden by using wrapper functions, so that the calculation operations are performed as if all k-simplices would have been on the same MPI process, i.e. as if the simulation was executed serially. Each wrapper function can keep either a pointer to the structure, if it exists on the same MPI process, and the ID used for finding the structure on the owner MPI process.

Algorithm 1 describes the most important aspects of simplicial complex classes. First, a basic *SimpComp* class is given, followed by the wrapper class *irtualSimpComp* used for parallelization.

Algorithm 1: Declaration of simplicial complex classes.

```
class SimpComp{
public:
    SimpComp(int dim);
    SimpComp(string s, int dim);
    ~SimpComp();
    // Creating new KSimplex
    // at level k:
    VirtualKSimplex* create_ksimplex(
        int k);
    void update_owner(int owner);

    string name;
    int D;
    vector< vector<
        VirtualKSimplex *> > elements;
};
class VirtualSimpComp{
public:
    SimpComp *find_simpcomp;

    int id;
    int ownerRank;
    SimpComp *simpComp;
};
```

Algorithm 2 describes the most important aspects of k-simplices classes. A basic *Simple* class is followed by the wrapper class *irtual Simple* used for parallelization.

Algorithm 2: Declaration of k-simplex classes.

```
class KSimplex{
public:
    KSimplex();
    KSimplex(int k, int D);
    ~KSimplex();
    bool find_neighbor(
            VirtualKSimplex *k1);
    void add_neighbor(
            VirtualKSimplex *k1);

    int k; // level
    int D; // dimension
    VirtualSimpComp *neighbors;
    vector<Color *> colors;
};
class VirtualKSimplex{
public:
    KSimplex *find_ksimplex();
    int id;
    int ownerRank;
    KSimplex *ksimplex;
};
```

In both algorithms, wrapper functions store a pointer to the base class object, if such exists on a local MPI process. Otherwise, the value is *nullptr*, and the data is searched for on the so called *owner ank* based on unique identifier called *id*. Owner of this k-simplex can issue multiple requests while it holds a lock.

IV. INFRASTRUCTURE FOR COMMUNICATION BETWEEN MPI PROCESSES

The communication between MPI processes is organized as follows. Each MPI process is preparing the data to be sent to other MPI processes. Order of operations prepared for other MPI processes is not important. All requests to other MPI processes for processing are packed in to_rank vector of vectors of unsigned char.

Each type of primitive data is serialized into the array of unsigned characters as it will be explained in the following section. Each prepared byte is pushed to the back of the vector of unsigned characters. Once all the data is prepared, the data is sent to other MPI processes in the background using *MPI Isend* directive. If a reference to the vector of array of unsigned characters is called *ec*, the pointer to the array is obtained by calling member function *data* of vector class from standard template library. After issuing all *MPI Isend* directives, waiting for each of sending to finish is achieved using *MPI Wait*.

Similarly receiving the data from other MPI processes is implemented in the background using *MPI Irec* , followed by *MPI Wait*, once the data is needed for the processing. The data is received into array of unsigned characters, that is further packed into vector of vectors of unsigned characters called from_rank for simple processing.

V. MPI SUPPORTING FUNCTIONS

As already mentioned, variables are serialized into the array of unsigned characters using the following syntax:

```
*( (__typeof__ (variable) *) (array + nArray) ) = variable;
nArray += sizeof(variable);
```

Here, *array* is array of unsinged characters where the data stored in the variable is serialized, and *nArray* is the number serialized bytes in the array.

Similarly, a variable is read and prepared into the to_rank using the following syntax:

```
__typeof__ (variable) temp_var = variable; \
int nBytes = sizeof(temp_var); \
for(int iByte = 0; iByte < nBytes; iByte++) \
   to_rank[rankNumber].push_back(
         ((unsigned char *) &temp_var) [iByte] );
```

This can be further optimized, but the optimization is out of the scope of this research.

 ISBN 978-86-7466-930-3

The communication between MPI processes is continued for as long as any MPI process requires further communication with other MPI processes. This is achieved using the following source code, where the MPI process that requires further communication sets variable *to send* to one:

```
int to_receive = 0; // A rank required communication
MPI_Allreduce(&to_send, &to_receive, 1, MPI_INT,
      MPI_SUM, MPI_COMM_WORLD);
```

After *MPI Allreduce* is executed, all MPI processes will have the information whether they have to communicate further in *to recei e* variable.

VI. Parallelization Possibilities Using Dataflow Paradigm

This simulator issues the same set of computer architecture instructions repeatedly. As in majority simulator of physical phenomena, the number of instructions is dependent on the precision of the model and is limited by the computing resources and the total simulation time requirement. These conditions are exactly what is required for a program to be suitable for acceleration using the dataflow paradigm [11]. Programming dataflow architectures requires programming skills that are higher than those needed for programming conventional von Neumann architectures. One of the possibilities is to write a program in a VHDL. More suitable solution to most of the programmers would be to exploit the framework that enables writing source code in a Java-like language, which gets automatically translated into the FPGA image [12,13]. Even in this case, the effort needed for programming such architectures is higher [14]. Besides programming dataflow architecture for the simplicial complex simulator, appropriate scheduling scheme is also needed for efficient running of multiple jobs simultaneously [15].

As the number of operations that can be applied to simplicial complexes can lead to several days' simulation time or even more, having in mind the aging and the probability of failure of supercomputing nodes [16], we have decided to write restarts after given number of simulations defined by the user, so that the calculation can continue from the last stored state.

VII. Conclusion

In this work we have presented the basics of the paralellization techniques that can be applied to the structure of a simplicial complex, which underlies a host of research problems in theoretical physics (see also our accompanying paper [17]). These problems tend to be computationally extremely expensive, and the common underlying software that enables parallelization at the level of the basic data structure can possibly go a long way towards optimization of code for numerical study using heavily parallel hardware platforms such as HPC clusters. In particular, the simplicial complex naturally allows for various aspects of parallelization, and we have described the basic classes, corresponding MPI communication infrastructure, supporting functions and the dataflow paradigm employed for the construction.

One should note that our work represents just a first step towards a full working software implementation, and much more effort is needed to properly implement, optimize and test the resulting code in real world environments. All that is the topic for future work. In particular, the data regarding the experimental evaluation, which would compare the proposed parallelization method to ordinary sequential methods still needs to be gathered and analyzed. Nevertheless, this first step is fundamental, and it is conceptually important since it represents a paradigm in which parallelization is implemented dominantly at the level of the simplicial complex as the underlying data structure, rather than at the level of the particular algorithm that aims to solve some particular problem using these data structures.

Finally, we note that our code, once properly developed, may possibly find applications not just in theoretical physics, but also in other disciplines of science, technology and engineering.

Acknowledgment

DC and NK were partially supported by the School of Electrical Engineering, University of Belgrade, Serbia. NK was partially supported by the Institute of Physics Belgrade, contract no. 0801-1264/1. MV was supported by the Science Fund of the Republic of Serbia, grant no. 7745968, "Quantum gravity from higher gauge theory" – QGHG-2021. All authors were partially supported by the Ministry of Education, Science, and Technological Development of the Republic of Serbia.

References

[1] E. H. Spanier, *Algebraic Topology*, New York, USA: Springer Verlag, 1966.

[2] M. W. Hirsch, *ifferential Topology*, New York, USA: Springer Verlag, 1976.

[3] S. Durr, Z. Fodor, J. Frison, C. Hoelbling, R. Hoffmann, S. D. Katz, S. Krieg, T. Kurth, L. Lellouch, T. Lippert, K. K. Szabo and G. Vulvert, "Ab-initio Determination of Light Hardon Masses", *Science* **22**, 1224-1227 (2008).

[4] J. Ambjorn, A. Goerlich, J. Jurkiewicz and R. Loll, "Nonperturbative Quantum Gravity", *Phys. ep.* **51** , 127 (2012).

[5] M. Vojinović, "Causal dynamical triangulations in the spincube model of quantum gravity", *Phys. e .* **4**, 024058 (2016).

[6] T. Radenković and M. Vojinović, "Higher Gauge Theories Based on 3-Groups", *EP* **1** , 222 (2019).

[7] A. Miković and M. Vojinović, "Standard Model and 4-Groups", *Europhys. ett.* **1** , 61001 (2021).

[8] A. Miković and M. Vojinović. "Quantum gravity for piecewise flat spacetimes", *SFIN* **I**, 267 (2018).

[9] N. Paunković and M. Vojinović, "Gauge protected entanglement between gravity and matter", *Class. uant. Gra .* **5**, 185015 (2018).

[10] J. B. Hartle and S. W. Hawking, "Wave function of the Universe", *Phys. e .* **2** , 2960 (1983).

[11] B. Lee and A. R. Hurson, "Issues in dataflow computing," *Ad ances in computers*, Elsevier, , 285-333 (1993).

[12] V. Milutinovic, J. Salom, D. Veljovic, N. Korolija, D. Markovic, and L. Petrovic, "Transforming applications from the control flow to the dataflow paradigm," *ataflow supercomputing essentials*, Springer, Cham, 107-129 (2017).

[13] N. Korolija, J. Popović, M. Cvetanović, and M. Bojović, "Dataflow-based parallelization of control-flow algorithms," *Ad ances in computers*, Elsevier, 1 4, 73-124 (2017).

[14] J. Popovic, D. Bojic, and N. Korolija, "Analysis of task effort estimation accuracy based on use case point size," *IET Software*, **( )**, 166-173 (2015).

[15] N. Korolija, D. Bojić, A. R. Hurson, and V. Milutinovic, "A runtime job scheduling algorithm for cluster architectures with dataflow accelerators," *Ad ances in computers*, Elsevier, **12** (2022).

[16] K. Huang, Y. Liu, N. Korolija, J. M. Carulli, and Y. Makris, "Recycled IC detection based on statistical methods," *IEEE transactions on computer-aided design of integrated circuits and systems*, **4( )**, 947-960 (2015).

[17] D. Cvijetić, N. Korolija and M. Vojinović, "Infrastructure for Simulating n-Dimensional Simplicial Complexes," IcETRAN 2022, Novi Pazar, Republic of Serbia, June 6-9, 2022, Belgrade: Društvo za ETRAN, Beograd: Akademska misao (2022).

ISBN 978-86-7466-930-3