Научном већу

Института за физику у Београду

Прегревица 118, 11 080 Земун

Београд, Србија

Предмет Реизбор у звање истраживач сарадник

Молим Научно веће Института за физику у Београду да покрене поступак за мој реизбор у звање истраживач сарадник.

Прилажем:

1. Молбу за реизбор у звање
2. Мишљење руководиоца пројекта
3. Биографију
4. Списак и фотокопије научних радова
5. Преглед научне активности
6. Потврду о упису на докторске академске студије
7. Фотокопију дипломе
8. Потврду о прихваћеној теми докторске дисертације

С поштовањем, Бојана Благојевић

Београд, 8.10.2018.

МИШЉЕЊЕ РУКОВОДИОЦА ПРОЈЕКТА

Бојана Благојевић, дипломирани физичар, је ангажована на пројекту основних истраживања ОИ171004 (ATLAS експеримент и физика честица на LHC енергијама) Министарства просвете, науке и технолошког развоја Републике Србије, а од 1. септембра 2017. и на ERC-2016-CoG:725741, где се бави теоријском физиком кварк-глуонске плазме (QGP), под менторством др Магдалене Ђорђевић, научног саветника Института за физику. Од јуна 2014. до новембра 2017. године, била је ангажована и на пројекту SNSF SCOPES IZ73Z0-152297, под менторством др Марка Ђорђевића (ванредног професора на Биолошком факултету Универзитета у Београду) и др Магдалене Ђорђевић, где се бави теоријским проучавањем имуног система бактерија и регулације експресије гена. У оквиру досадашњих истраживања објавила је четири научна рада **M21** категорије и један **M21a** категорије, а још један **M21** рад у припреми тренутно је под (позитивном) рецензијом у Physical Review C часопису. Поред тога, има и два саопштења са међународних скупова штампана у целини **M33** и предавање по позиву штампано у целини на домаћој конференцији са међународним учешћем **M61.** Такође је учествовала у изради поглавља у Истакнутој монографији међународног значаја **M11**. На основу изложеног, предлажем Научном већу Института за физику у Београду да Бојану Благојевић поново изабере у звање истраживач сарадник.

Предлажем комисију за реизбор у звање у следећем саставу:

др Магдалена Ђорђевић, научни саветник Института за физику

др Лидија Живковић, научни саветник Института за физику

проф. др Маја Бурић, редовни професор, Физички факултет

др Игор Салом, научни сарадник Института за физику

др Лидија Живковић, научни саветник

Руководилац пројекта

Биографија кандидата

Бојана Благојевић је рођена у Приједору, Босна и Херцеговина, 24. 8. 1984. године. У Добоју је завршила основну школу и гимназију Јован Дучић, као ђак генерације. Јуна 2003. године била је победник Физичке олимпијаде Босне и Херцеговине, а августа 2003. учествовала је на 34. Међународној физичкој олимпијади у Тајпеху, Тајван.

Основне академске студије на Физичком факултету Универзитета у Београду, смер Теоријска и експериментална физика, започиње 2003. године и завршава их 2013. године са просечном оценом 10,00, одбранивши дипломски рад на тему *Продукција, масе и распади суперсиметричних честица у оквиру cMSSM модела на LHC-у*. Током основних академских студија Бојана Благојевић је награђена *Eurobank EFG школарином*, која се додељује најбољим студентима завршне године државних факултета за остварене изванредне резултате током студија. Током академске 2006/2007. године била је стипендиста фонда *Проф. др Ђорђе Живановић* као један од најбољих студената III године физике на Физичком факултету. Дипломски рад је израђен у Лабораторији за физику високих енергија Института за физику у Београду, а израдом рада руководила је др Марија Врањеш Милосављевић, виши научни сарадник са Института за физику.

Академске 2013/2014 године уписује докторске академске студије на Физичком факултету Универзитета у Београду, ужа научна област Физика честица и поља. Ментор докторских студија је др Магдалена Ђорђевић, научни саветник Института за физику у Београду. На Институту за физику је запослена од 30. 06. 2013. године на пројекту „ATLAS експеримент и физика честица на LHC енергијама" Министарства просвете, науке и технолошког развоја Републике Србије.

До сада је публиковала научни један **M21a** категорије, четири рада **M21** категорије, као и још један **M21** рад који је тренутно под (позитивном) рецензијом у Phys. Rev. C часопису; два саопштења са међународних скупова штампана у целини **M33** и предавање по позиву штампано у целини на домаћој конференцији са међународним учешћем **M61**. Такође је учествовала у изради поглавља у Истакнутој монографији међународног значаја **M11**.

Списак радова:

M21a:

1. Magdalena Djordjevic, Marko Djordjevic and Bojana Blagojevic, *RHIC and LHC jet suppression in non-central collisions*, Phys. Lett. B **737**, 298-302 (2014).

M21:

1. Bojana Blagojevic and Magdalena Djordjevic, *Importance of different energy loss effects in jet suppression at RHIC and LHC*, J. Phys. G **42**, 075105 (2015).; такође истакнут и у LabTalk-у.
2. Magdalena Djordjevic, Bojana Blagojevic and Lidija Zivkovic, *Mass tomography at different momentum ranges in quark-gluon plasma*, Phys. Rev. C **94**, 044908 (2016).

3. Andjela Rodic, Bojana Blagojevic, Evgeny Zdobnov, Magdalena Djordjevic and Marko Djordjevic, *Understanding key features of bacterial restriction-modification systems through quantitative modeling*, BMC Systems Biology, **11**:377, (2017).
4. Andjela Rodic, Bojana Blagojevic, Magdalena Djordjevic, Konstantin Severinov and Marko Djordjevic, *Features of CRISPR-Cas Regulation Key to Highly Efficient and Temporally-Specific crRNA Production*, Front. Microbiol., **8**, 2139 (2017).
5. Bojana Blagojevic, Magdalena Djordjevic and Marko Djordjevic, *Calculating hard probe radiative energy loss beyond soft-gluon approximation: how valid is the approximation?*, arXiv:nucl-th/1804.07593 (under review in Phys. Rev. C) (2018).

M11:

1. Andjela Rodic, Bojana Blagojevic, Marko Djordjevic (2018) *Systems Biology of Bacterial Immune Systems: Regulation of Restriction-Modification and CRISPR-Cas Systems.* In: Rajewsky N., Jurga S., Barciszewski J. (eds) Systems Biology. RNA Technologies. Springer, Cham.

M61:

1. Jelena Guzina, Andjela Rodic, Bojana Blagojevic and Marko Djordjevic, *Modeling and bioinformatics of bacterial immune systems: understanding regulation of CRISPR/Cas and restriction-modification systems*, Biologia Serbica: **39** pp. 112-122 (2017).

M33:

1. Bojana Blagojevic and Magdalena Djordjevic, *Energy loss in jet suppression - what effects matter?*, J. Phys. Conf. Ser. **612**, 012006 (2015).
2. Bojana Blagojevic and Magdalena Djordjevic, *Modeling jet-medium interactions at RHIC and LHC - which energy loss effect is crucial*, J. Phys. Conf. Ser. **668**, 012044 (2016).

Преглед научне активности

Бојана Благојевић је започела свој истраживачки рад у Лабораторији за физику високих енергија Института за физику у Београду, у новембру 2012. године, а запослена је од 30. 06. 2013. године. Ангажована на пројекту основних истраживања ОН171004 (ATLAS експеримент и физика честица на LHC енергијама) Министарства просвете, науке и технолошког развоја Републике Србије, а од 1. 09. 2017. и на ERC-2016-CoG:725741, где се бави теоријском физиком кварк-глуонске плазме (QGP) под менторством др Магдалене Ђорђевић, научног саветника Института за физику у Београду. QGP је ново стање материје у којем су кваркови и глуони асимптотски слободни, и које је креирано у ултра-релативистичким сударима тешких јона у RHIC (Relativistic Heavy Ion Collider, Brookhaven National Laboratory) и LHC (Large Hadron Collider, CERN) експериментима. Конкретно, Бојана Благојевић ради на теоријским предвиђањима пригушења, тј. супресије (промене дистрибуција енергије) високо енергијских честица, на поређењу добијених предвиђања са експерименталним резултатима, као и аналитичком унапређењу модела енергијских губитака. Од јуна 2014. године ангажована је и на пројекту SNSF SCOPES IZ73Z0-152297, под менторством др Марка Ђорђевића, ванредног професора на Биолошком факултету Универзитета у Београду, и др Магдалене Ђорђевић, где се бави теоријским проучавањем

имуног система бактерија и регулације експресије гена код рестрикционо-модификационих система, као и CRISPR/Cas-а.

До сада објављени радови су наведени у одељку Списак радова, док је истраживања презентовала на доленаведеним конференцијама.

Током септембра 2014. године Бојана Благојевић је била учесник конференције *Hot Quarks 2014, Workshop for young scientists on the physics of ultrarelativistic nucleus-nucleus collisions* у Лас Неграсу, Андалузија, Шпанија, где је одржала предавање под називом *Energy loss in jet suppression - what effects matter?*. Током јуна и јула 2015. године била је полазник летње школе *Helmholtz International Summer School, Dubna International Advanced School of Theoretical Physics, Dense Matter 2015*, која је одржана у Дубни, Русија. Такође у Дубни је, током јула 2015. учествовала на конференцији *Strangeness in Quark Matter*, где је одржала предавање под називом *Importance of different energy loss effects in jet suppression at RHIC and LHC*.

Током јуна 2016. године била је учесник конференције *Belgrade BioInformatics Conference BelBi 2016*, у Београду, Србија, где је одржала предавање под називом *Explaining regulatory features of bacterial R-M systems through theoretical modeling*. Током августа 2016. године кандидаткиња је била полазник школе *Young Scientists School, Systems Biology and Bioinformatics SBB 2016* у Новосибирску, Русија, где је одржала предавање под називом *The role of different regulatory features in achieving safe and efficient R-M system establishment*. Такође у Новосибирску је, током августа и септембра 2016. године, учествовала на конференцији *The 10th International Multiconference Bioinformatics of Genome Regulation and Structure/Systems Biology BGRS/SB 2016*, где је презентовала постер под називом *Design of bacterial restriction-modification systems: relating the system architecture with its dynamical response*. Крајем децембра 2016. године, заједно са Анђелом Родић, одржала је предавање на Биоинформатичком семинару Математичкг факултета Универзитета у Београду под називом: *CRISPR/Cas and restriction-modification systems: modeling dynamics of bacterial immune system expression*.

У мају 2018. године учествовала је на конференцији *The 27th International Conference on Ultrarelativistic Nucleus-Nucleus Collisions, Quark Matter 2018*, у Венецији, Италија, где је презентовала постер под називом *Soft-gluon approximation in calculating radiative energy loss of high $p_\perp$ particles - is it well-founded?*. Током јуна исте године била је учесник конференције *Belgrade BioInformatics Conference BelBi 2018*, у Београду, Србија, где је презентовала е-постер под називом *Defining dynamical property observables which ensure efficient restriction-modification systems establishment in bacterial host*. Септембра 2018. године учествовала је на конференцији *Hot Quarks 2018, Workshop for young scientists on the physics of ultrarelativistic nucleus-nucleus collisions* на острву Тексел у Холандији, где је одржала предавање под називом *Testing reliability of the soft-gluon approximation in calculating radiative energy loss of high $p_\perp$ particles*, док је почетком октобра исте године била учесник конференције *Hard Probes 2018: International Conference on Hard and Electromagnetic Probes of High-Energy Nuclear Collisions*, где је одржала предавање под насловом *Hard probe radiative energy loss beyond soft-gluon approximation*.
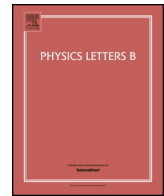
RHIC and LHC jet suppression in non-central collisions

Magdalena Djordjevic [a],[*], Marko Djordjevic [b], Bojana Blagojevic [a]

[a] *Institute of Physics Belgrade, University of Belgrade, Serbia*
[b] *Faculty of Biology, University of Belgrade, Serbia*

ARTICLE INFO

ABSTRACT

Understanding properties of QCD matter created in ultra-relativistic heavy-ion collisions is a major goal of RHIC and LHC experiments. An excellent tool to study these properties is high-momentum hadron suppression of light and heavy flavor observables. Utilizing this tool requires accurate suppression predictions for different experiments, probes and experimental conditions, and their unbiased comparison with experimental data. With this goal, we here extend our dynamical energy loss formalism towards generating predictions for non-central collisions; the formalism takes into account both radiative and collisional energy loss, dynamical (as opposed to static) scattering centers, finite magnetic mass, running coupling and uses no free parameters in comparison with experimental data. Specifically, we here generate predictions for all available centrality ranges, for both LHC and RHIC experiments, and for four different probes (charged hadrons, neutral pions, D mesons and non-prompt J/ψ). We obtain good agreement with all available non-central data, and also generate predictions for suppression measurements that will soon become available. Finally, we discuss implications of the obtained good agreement with experimental data with different medium models that are currently considered.

1. Introduction

High-momentum hadron suppression [1] of light and heavy observables provides an excellent tool [2–4] for studying properties of QCD matter created in ultra-relativistic heavy ion collisions. Mapping these properties is also a major goal of RHIC and LHC experiments, which requires comparison of hadron suppression measurements with corresponding theoretical predictions. To ensure the unbiased comparison with experimental data, it is necessary to generate predictions for different experiments, experimental probes and experimental conditions, within the same theoretical model. With a major goal of generating these predictions, we developed dynamical energy loss formalism, that *i)* allows treating, at the same time, both light and heavy partons, *ii)* is computed in dynamical QCD medium (i.e. takes into account recoil of the medium constituents), *iii)* includes both collisional [5] and radiative [6,7] energy losses, computed within the *same* theoretical framework, *iv)* includes realistic finite size effects, i.e. the fact that experimentally created QCD medium has finite size, and that the jets are produced inside the medium, *v)* includes finite magnetic mass effects [8] and running coupling [9]. We further integrated this formalism into numerical procedure which also includes multi-gluon fluctuations [10], path length fluctuations [11]

and most up-to-date jet production [12,13] and fragmentation functions [14]; the procedure allows generating predictions with no free parameters used in comparison with experimental data.

We previously applied the computational procedure outlined above for generating predictions in *most central* collisions for a number of different probes at LHC [9]. These predictions showed a very good agreement with experimental data; however, a comprehensive comparison also requires generating predictions for non-central collisions at RHIC and LHC. With this goal, we here extend the formalism towards generating predictions for different centrality ranges. We consequently generate the suppression predictions for all available centrality ranges, for both RHIC and LHC experiments and for four different probes – specifically for charged hadrons, D mesons and non-prompt J/ψ at LHC and neutral pions at RHIC. Such comprehensive comparison allows testing some of important assumptions behind our current understanding of the created QCD matter, such as ranges of validity for different medium models.

2. Theoretical framework

The numerical procedure for calculating high-momentum hadron suppression for central collisions is outlined in detail in [9]. We below first briefly list the main steps in this procedure and then describe the extension of the procedure that is necessary for generating the predictions for non-central collisions:

i) Energy loss calculations: Our model takes into account both radiative and collisional contributions to jet energy loss. Specifically, the radiative energy loss calculations present a state-of-the-art extension of a well-known DGLV model [15,16] towards a finite size dynamical medium [6,7], finite magnetic mass [8] and running coupling [9]. These extensions are further discussed below.

ii) Dynamical scattering centers: To calculate the radiative energy loss, we use finite size dynamical energy loss formalism. This formalism removes a ubiquitous assumption of static scattering centers [17] and takes into account that the medium constituents are in reality dynamical, i.e. moving particles; similarly, the unrealistic assumption of infinite medium is also removed. Calculations of the jet energy loss in dynamical medium are done by using two-hard-thermal-loop approach. In contrast to the static energy loss, where only the electric contribution appears in the final result, both electric and magnetic contributions appear in the dynamical case. This then directly leads to the question of finite magnetic mass, which we further discuss below.

iii) Magnetic mass: In pQCD energy loss calculations – including our (initial) dynamical energy loss formalism [6,7] – magnetic mass is taken to be zero. However, different non-perturbative approaches suggest a non-zero magnetic mass at RHIC and LHC (see e.g. [18–21]). To address this issue, we generalized the dynamical energy loss calculations to the case of finite magnetic mass. Introducing the finite magnetic mass is described in detail in [8], where the finite magnetic mass is introduced through generalized sum-rules.

iv) Running coupling: Introducing the running coupling is described in detail in [9]. One should note that the obtained result is infrared safe and moreover of a moderate value. There is consequently no need to introduce an artificial cutoff as is commonly done elsewhere with the running coupling.

v) Suppression procedure: We further integrated the energy loss formalism outlined above into a numerical procedure that includes: light and heavy flavor production [12,13], path-length [11] and multigluon [10] fluctuations, up-to-date fragmentation functions [14] for light and heavy flavor and the decay of heavy mesons to single electrons and J/ψ. In the calculations, as a start point we use an effective temperature of 304 MeV for 0–40% centrality Pb + Pb collisions at LHC (as extracted by ALICE [22]) and effective temperature of 221 MeV for 0–20% centrality Au + Au collisions at RHIC (as extracted by PHENIX [23]). The other parameter values are specified in the next section, while the details of the procedure are provided in [9]. Note that we use no free parameters in comparison with the data, i.e. all the parameters that we use correspond to standard literature values.

To extend the computational procedure outlined above to non-central collisions, we start by obtaining the path-length distributions for different centrality ranges from [24]. Furthermore, we determine the temperature for each centrality region according to [15] $T^3 \sim \frac{\frac{dN_g}{dy}}{V} \rightarrow T = c\left(\frac{\frac{dN_g}{dy}}{N_{part}}\right)^{1/3}$ (more details will be provided in [25]), where $\frac{dN_g}{dy}$ is gluon rapidity density, V is the volume of created medium, and we take that $V \sim N_{part}$ (number of participants for a given collision). Furthermore, c is a constant for a specific system/collider energy, and $\frac{\frac{dN_g}{dy}}{N_{part}}$ is directly proportional to experimentally measured charged particle multiplicity per participant pair $\left(\frac{\frac{dN_{ch}}{dy}}{N_{part}/2}\right)$, which is measured for both RHIC [26] and LHC [27] and across different centralities. The constants c can be fixed through ALICE measurement of effective temperature for 0–40% centrality at 2.76 TeV Pb + Pb collisions at LHC, and through PHENIX measurement of effective temperature for 0–20% centrality at 200 GeV Au + Au collisions at RHIC (see above).

3. Numerical results

In this section, we concentrate on 200 GeV Au + Au collisions at RHIC and 2.76 TeV Pb + Pb collisions at LHC, and present our suppression predictions for light and heavy flavor observables. We proceed by considering a QGP with $n_f = 2.5$ effective light quark flavors for RHIC and $n_f = 3$ for LHC. Perturbative QCD scale is taken to be $\Lambda_{QCD} = 0.2$ GeV. For the light quarks we assume that their mass is dominated by the thermal mass $M = \mu_E/\sqrt{6}$, where the temperature dependent Debye mass $\mu_E(T)$ is obtained from [28]. Magnetic mass μ_M is taken as $0.4\mu_E < \mu_M < 0.6\mu_E$ [18–21], and the gluon mass is $m_g = \mu_E/\sqrt{2}$ [29]. For the charm (bottom) mass we use $M = 1.2$ GeV ($M = 4.75$ GeV). Path-length distribution and temperatures for different centralities are computed according to the procedure outlined in the previous section. Parton production, fragmentation functions and decays, which are used in the numerical calculations, are specified in [9]. Note that, on each panel of every figure, the gray region corresponds to the range of $0.4 < \mu_M/\mu_E < 0.6$, where the upper (lower) boundary of each band corresponds to $\mu_M/\mu_E = 0.6$ ($\mu_M/\mu_E = 0.4$).

We start by generating predictions for momentum dependence of hadron suppression at LHC experiments, for different centrality regions, which are shown in Figs. 1 and 2. Each panel in these figures shows a fixed centrality region (0–5%, 5–10%, 10–20%, etc.) and for each of these centrality regions, momentum dependence of R_{AA} is shown. Fig. 1 shows predictions for charged hadron R_{AA} and their comparison with relevant ALICE and CMS experimental data at 2.76 TeV Pb + Pb collisions at LHC. In Fig. 2 predictions for D meson R_{AA} are shown; predictions for 0–10% and 30–50% are compared with the available ALICE data, where a very good agreement can also be seen. Note that predictions for 30–50% region were generated *before* the experimental data – that are now shown in the figure – became available [34]. The experimental data for the rest of the predictions (the other two panels in Fig. 2) are expected to become available soon.

In Fig. 3, we show equivalent predictions as in Figs. 1 and 2, but for RHIC measurements of neutral pions at 200 GeV Au + Au collisions. Each panel shows predictions for different (fixed) centrality bin, which are compared with experimental data. Similarly as for LHC measurements, we see a very good agreement between the theoretical predictions and RHIC data.

In Fig. 4, instead of fixing the centrality ranges (as in Figs. 1–3), we fix the momentum regions and explore how R_{AA} changes for different centrality values (i.e. number of participants). The predictions are generated for both RHIC and LHC experiments, and for various probes. Specifically, we compare our predictions with experimental data for neutral pions at RHIC and charged hadrons, D mesons and non-prompt J/ψ at LHC. One can see that we here also obtain a robust agreement with the experimental data.

4. Conclusions

We here generated suppression predictions for all available centrality ranges, for both RHIC and LHC, and for diverse experimental probes. These predictions were generated by the same theoretical formalism and within the same numerical procedure. Furthermore, all the predictions within the same experiment (i.e. within RHIC and within LHC) were generated with the same parameter set, which corresponds to standard literature values, and with no free parameters used in comparison with experimental data. We obtained good agreement of the theoretical predictions with the diverse experimental measurements, for all momentum ranges larger than 10 GeV.
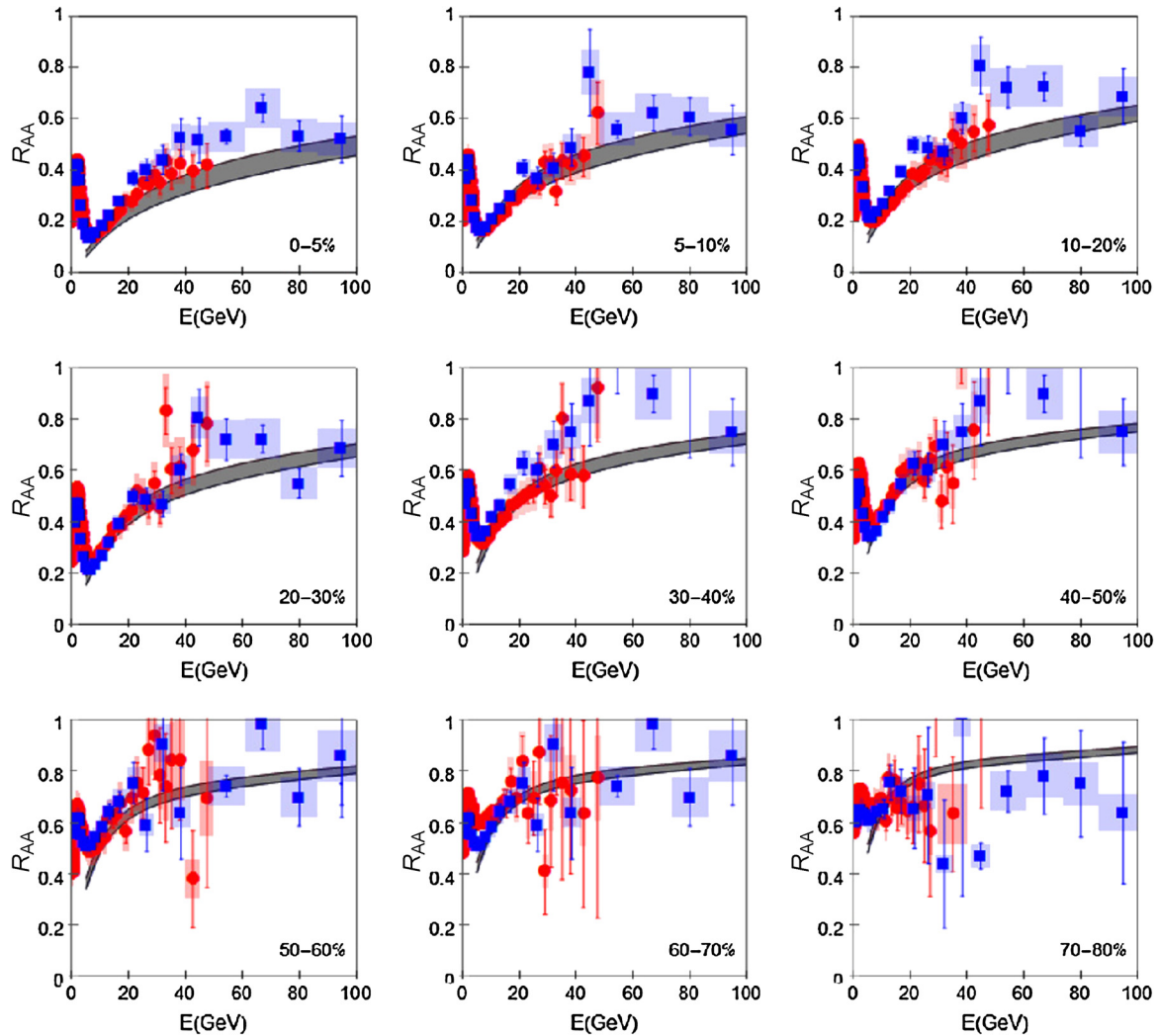
Fig. 1. Theory *vs.* experimental data for momentum dependence of charged hadron R_{AA} for different centrality bins at LHC. The panels show the comparison of charged hadron suppression predictions with experimentally measured R_{AA} for charged particles at 2.76 Pb + Pb collisions at LHC, for different (fixed) centrality ranges. Red circles and blue squares correspond to ALICE [30] and CMS [31] experimental data, respectively. In the lower right corner of each panel we denote the centrality for which the data and the predictions are presented. Note that, on the third and the fourth panels, CMS data for centrality bin 10–30% are shown. Similarly, on the fifth and the sixth panels, CMS data for centrality bin 30–50% are shown, on the seventh and the eight panels, CMS data for centrality bin 50–70% are shown, and on the ninth panel CMS data for 70–90% in centrality are shown.
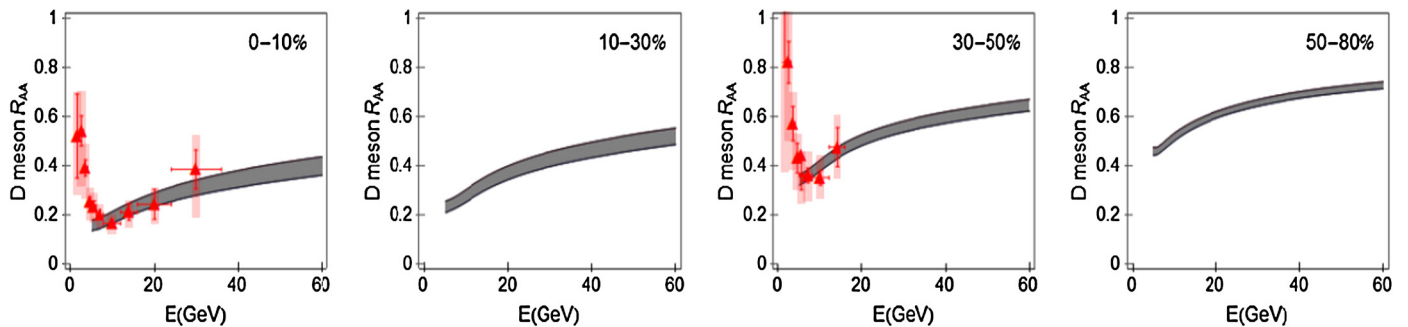

Fig. 2. Theory *vs.* experimental data for momentum dependence of D meson R_{AA} for different centrality bins at LHC. The left panel shows the comparison of D meson suppression predictions with D meson R_{AA} at 0–7.5% central 2.76 Pb + Pb collisions at LHC [32] (the red triangles). The other three panels show the theoretical predictions for D meson R_{AA} for centrality bins 10–30%, 30–50% and 50–80%, respectively. In the third panel (30–50% centrality region), the predictions are compared with ALICE preliminary data [33] that recently became available.

The robust agreement discussed above has interesting implications for ranges of validity of different medium models, which are incorporated in different approaches to hadron suppression predictions. As discussed in the Introduction, our calculations employ state-of-the-art method for energy loss calculations and numerical

procedure for suppression calculations, but do not explicitly take into account the medium evolution (i.e. the evolution is taken into account through effective/average medium parameters). This is in contrast to a number of other approaches (see e.g. [39–42]), which simplify the energy loss to a various degree, in order to more
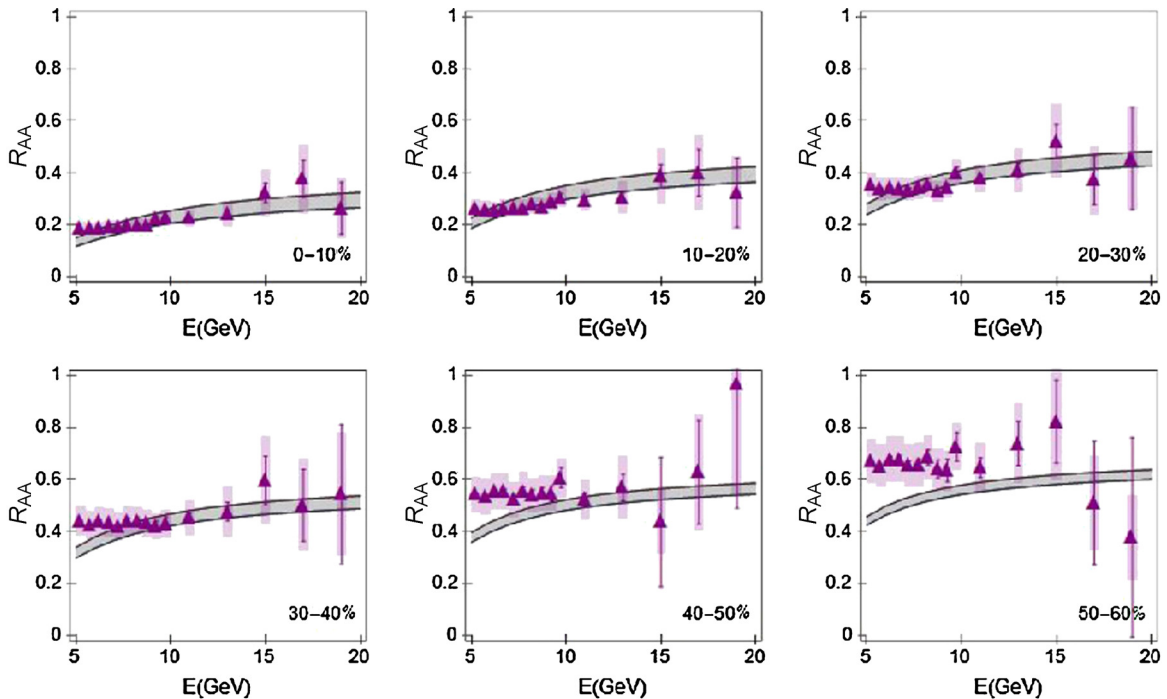
Fig. 3. Theory *vs.* experimental data for momentum dependence of neutral pion R_{AA} for different centrality bins at RHIC. The panels show the comparison of neutral pion suppression predictions with R_{AA} for neutral pions at 200 GeV Au + Au collisions at RHIC, for different (fixed) centrality regions. Purple triangles correspond to PHENIX [35] data. In the lower right corner of each panel we denote the centrality for which the data and the predictions are presented.
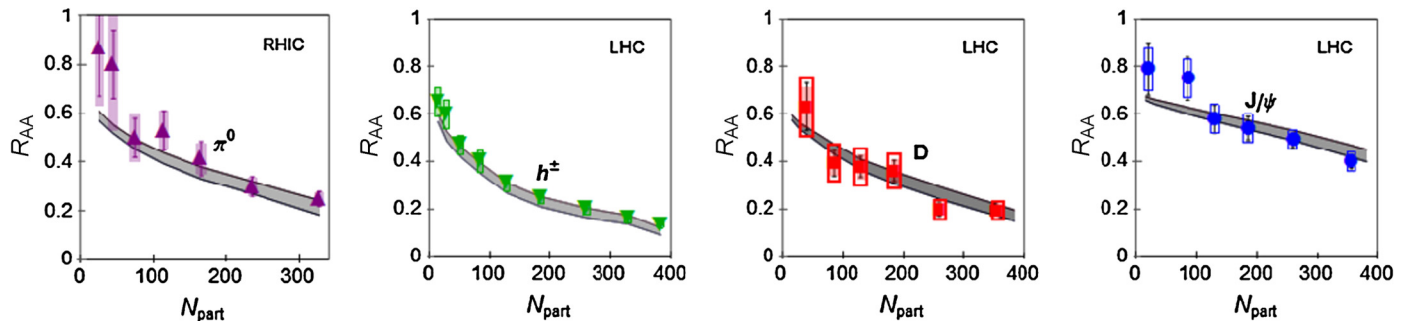

Fig. 4. Theory *vs.* experimental data for participant dependence of light and heavy flavor R_{AA} at RHIC and LHC. The first panel compares theoretical predictions with experimental data for participant dependence of π^0 R_{AA} [35] at 200 GeV Au + Au collisions at RHIC, where π^0 momentum is larger than 7 GeV. The second, third and fourth panels compare theoretical predictions with experimental data for participant dependence of, respectively, $h^{\pm}$ [36], D meson [37] and non-prompt J/ψ [38] R_{AA} at 2.76 TeV Pb + Pb collisions at LHC. The jet momentum ranges for the second, the third and the forth panels are, respectively, 6–12 GeV, 8–16 GeV and 6.5–30 GeV.

explicitly incorporate the evolving medium. Consequently, the obtained robust agreement with the experimental data above 10 GeV, strongly suggests that expansion of the medium does not play a major role in explaining *angular averaged* high momentum hadron suppression data. We hypothesize that the reason behind this result is that hard probes have a sufficiently large amount of energy, while the created medium is sufficiently short, so that the angular averaged suppression of the outcoming hadrons is only sensitive to the average properties of the created medium. This hypothesis is actually in line with previous work [43–47], which used different formalism – and explored lower centrality regions – but also obtained that angular averaged high momentum hadron suppression results are not sensitive to explicitly including the medium expansion. What however remains to be tested is validity of this simplification for jet energy ranges outside of those tested here, and whether the framework employed here can also reasonably explain angular differential suppression observables (such as high-momentum v_2); note that these observables are expected to be more sensitive to the medium evolution than the angular averaged

suppression studied here [46]. Such analysis could simplify theoretical predictions and facilitate intuitive understanding of complex experimental data.

Acknowledgements

References

[1] J.D. Bjorken, FERMILAB-PUB-82-059-THY, 1982, pp. 287, 292.
[2] N. Brambilla, et al., preprint, arXiv:hep-ph/0412158, 2004.
[3] M. Gyulassy, Lect. Notes Phys. 583 (2002) 37.
[4] D. d'Enterria, B. Betz, Lect. Notes Phys. 785 (2010) 285.
[5] M. Djordjevic, Phys. Rev. C 74 (2006) 064907.

[6] M. Djordjevic, Phys. Rev. C 80 (2009) 064909.
[7] M. Djordjevic, U. Heinz, Phys. Rev. Lett. 101 (2008) 022302.
[8] M. Djordjevic, M. Djordjevic, Phys. Lett. B 709 (2012) 229.
[9] M. Djordjevic, M. Djordjevic, Phys. Lett. B 734 (2014) 286.
[10] M. Gyulassy, P. Levai, I. Vitev, Phys. Lett. B 538 (2002) 282.
[11] S. Wicks, W. Horowitz, M. Djordjevic, M. Gyulassy, Nucl. Phys. A 784 (2007) 426.
[12] M. Cacciari, S. Frixione, N. Houdeau, M.L. Mangano, P. Nason, G. Ridolfi, J. High Energy Phys. 1210 (2012) 137.
[13] Z.B. Kang, I. Vitev, H. Xing, Phys. Lett. B 718 (2012) 482;
R. Sharma, I. Vitev, B.W. Zhang, Phys. Rev. C 80 (2009) 054902.
[14] D. de Florian, R. Sassot, M. Stratmann, Phys. Rev. D 75 (2007) 114010.
[15] M. Gyulassy, P. Levai, I. Vitev, Nucl. Phys. B 594 (2001) 371.
[16] M. Djordjevic, M. Gyulassy, Nucl. Phys. A 733 (2004) 265.
[17] M. Gyulassy, X.N. Wang, Nucl. Phys. B 420 (1994) 583;
X.N. Wang, M. Gyulassy, M. Plumer, Phys. Rev. D 51 (1995) 3436.
[18] Yu. Maezawa, et al., WHOT-QCD Collaboration, Phys. Rev. D 81 (2010) 091501;
Yu. Maezawa, et al., WHOT-QCD Collaboration, PoS Lattice 194 (2008).
[19] A. Nakamura, T. Saito, S. Sakai, Phys. Rev. D 69 (2004) 014506.
[20] A. Hart, M. Laine, O. Philipsen, Nucl. Phys. B 586 (2000) 443.
[21] D. Bak, A. Karch, L.G. Yaffe, J. High Energy Phys. 0708 (2007) 049.
[22] M. Wilde, for theALICE Collaboration, Nucl. Phys. A 904–905 (2013) 573c.
[23] M. Csanad, arXiv:1101.1282.
[24] A. Dainese, Eur. Phys. J. C 33 (2004) 495.
[25] M. Djordjevic, M. Djordjevic, in preparation.
[26] S.S. Adler, et al., PHENIX Collaboration, Phys. Rev. C 71 (2005) 034908.
[27] K. Aamodt, et al., ALICE Collaboration, Phys. Rev. Lett. 106 (2011) 032301.
[28] A. Peshier, arXiv:hep-ph/0601119, 2006.
[29] M. Djordjevic, M. Gyulassy, Phys. Rev. C 68 (2003) 034914.
[30] B. Abelev, et al., ALICE Collaboration, arXiv:1208.2711.
[31] S. Chatrchyan, et al., CMS Collaboration, Eur. Phys. J. C 72 (2012) 1945.
[32] J. Otwinowski, for theALICE Collaboration, PoS ConfinementX 170 (2012).
[33] A. Festanti, for the ALICE Collaboration, Heavy-flavour production and nuclear modification factor in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV with ALICE, oral presentation at Quark Matter 2014, Darmstadt, Germany.
[34] M. Djordjevic, Theoretical predictions of jet suppression: a systematic comparison, oral presentation at Quark Matter 2014, Darmstadt, Germany.
[35] A. Adare, et al., PHENIX Collaboration, Phys. Rev. C 87 (2013) 034911, arXiv:1204.1526.
[36] K. Aamodt, et al., ALICE Collaboration, Phys. Lett. B 720 (2013) 52.
[37] B. Abelev, et al., ALICE Collaboration, J. High Energy Phys. 9 (2012) 112.
[38] S. Chatrchyan, et al., CMS Collaboration, J. High Energy Phys. 5 (2012) 063.
[39] S.A. Bass, C. Gale, A. Majumder, C. Nonaka, G.-Y. Qin, T. Renk, J. Ruppert, Phys. Rev. C 79 (2009) 024901.
[40] B. Betz, M. Gyulassy, arXiv:1402.3419.
[41] Z. Xu, C. Greiner, Phys. Rev. C 71 (2005);
Z. Xu, C. Greiner, Phys. Rev. C 76 (2007).
[42] P.B. Gossiaux, M. Nahrgang, M. Bluhm, T. Gousset, J. Aichelin, Nucl. Phys. A 904–905 (2013) 992c.
[43] T. Renk, J. Ruppert, C. Nonaka, S.A. Bass, Phys. Rev. C 75 (2007) 031902.
[44] D. Molnar, D. Sun, Nucl. Phys. A 910–911 (2013) 486–489.
[45] D. Molnar, D. Sun, arXiv:1405.4848.
[46] T. Renk, Phys. Rev. C 85 (2012) 044903.
[47] T. Renk, H. Holopainen, U. Heinz, C. Shen, Phys. Rev. C 83 (2011) 014910.

PAPER

Importance of different energy loss effects in jet suppression at the RHIC and the LHC

To cite this article: Bojana Blagojevic and Magdalena Djordjevic 2015 *J. Phys. G: Nucl. Part. Phys.* **42** 075105

View the article online for updates and enhancements.

Related content

- Explaining the fine hierarchy in pion and kaon suppression at LHC: importance of fragmentation functions
  Magdalena Djordjevic and Marko Djordjevic

- Magnetic and electric contributions to the energy loss in a dynamical QCD medium
  M Djordjevic

- Nuclear suppression of light hadrons and single electrons at the RHIC and LHC
  B G Zakharov

Recent citations

- Radiative heavy quark energy loss in an expanding viscous QCD plasma
  Sreemoyee Sarkar *et al*

- Dynamical energy loss as a novel Quark–Gluon Plasma tomographic tool
  Magdalena Djordjevic

- Modeling jet-medium interactions at RHIC and LHC - which energy loss effect is crucial?
  B Blagojevic and M Djordjevic

Importance of different energy loss effects in jet suppression at the RHIC and the LHC

Bojana Blagojevic and Magdalena Djordjevic

Institute of Physics Belgrade, University of Belgrade, Serbia

E-mail: magda@ipb.ac.rs

CrossMark

Abstract

Jet suppression is considered to be an excellent probe of quantum chromodynamic (QCD) matter created in ultra-relativistic heavy ion collisions. Our theoretical predictions of jet suppression, which are based on our recently developed dynamical energy loss formalism, show a robust agreement with various experimental data, which spans across different probes, experiments (Relativistic Heavy Ion Collider (RHIC) and Large Hadron Collider (LHC)) and experimental conditions (i.e. all available centrality regions). This formalism includes several key ingredients, such as the inclusion of dynamical scattering centers, a finite size QCD medium, collisional energy loss, finite magnetic mass and running coupling. While these effects have to be included based on theoretical grounds, it is currently unclear what their individual importance is in accurately interpreting the experimental data; in particular because other approaches to suppression predictions commonly neglect some—or all—of these effects. To address this question, we study the relative importance of these effects in obtaining accurate suppression predictions for D mesons (clear energy loss probe) at top RHIC and LHC energies. We obtain that several different ingredients are responsible for accurate predictions, i.e. robust agreement with the data is a cumulative effect of all the ingredients, though inclusion of the dynamical scattering centers has the largest relative importance.

Keywords: energy loss, jet suppression, quark-gluon plasma, charm quarks

(Some figures may appear in colour only in the online journal)

1. Introduction

The suppression of high transverse momentum light and heavy flavor observables [1] is considered to be an excellent probe of QCD matter created in ultra-relativistic heavy ion collisions at the RHIC and the LHC. One of the major goals of these experiments is mapping

the quark–gluon plasma (QGP) properties, which requires a comparison of the available suppression data with theoretical predictions [2–4]. Such a comparison tests different theoretical models and provides an insight into the underlying QGP physics. It is generally considered that the crucial ingredient for reliable suppression predictions is an accurate energy loss calculation.

Therefore, we previously developed the dynamical energy loss formalism, which includes the following effects: (i) dynamical scattering centers; (ii) a QCD medium of a finite size [5, 6]; (iii) both radiative [5, 6] and collisional [7] energy losses; (iv) finite magnetic mass [8] and (v) running coupling [9]. This energy loss formalism is based on the pQCD calculations in a finite size and optically thin dynamical QCD medium, and has been incorporated into a numerical procedure [9] that allows generating state-of-the art suppression predictions.

These predictions are able to explain the heavy flavor puzzle (the fact that, contrary to pQCD expectations, both light and heavy flavor probes have very similar experimentally measured R_{AA}) at both the RHIC [10] and the LHC [11] and, in general, show a very good agreement with the available suppression data at these experiments, for a diverse set of probes [9, 10] and centrality regions [12].

However, such good agreement between the predictions and the experimental data raises the question of which energy loss effects are responsible for the accurate predictions. In other words, is there a single dominant energy loss effect responsible for the good agreement, or is this agreement the result of a superposition of several smaller improvements? This issue is moreover important, given the fact that various pQCD approaches [13–22] to the energy loss calculations neglect some (or many) of these effects.

Consequently, here we address the importance of different energy loss ingredients in the suppression calculations. For this purpose, it would be optimal to have a probe that is sensitive only to the energy loss, i.e. for which fragmentation and decay functions do not play a role. The D meson suppression is such a probe, since the fragmentation functions do not modify bare charm quark suppression, as previously shown in [10, 11]. To explore different energy loss approximations, which have been used in suppression predictions, we concentrate on the D meson suppression in central 200 GeV Au+Au collisions at the RHIC and 2.76 TeV Pb+Pb collisions at the LHC. While high momentum D meson suppression data are not available at the RHIC—the RHIC measurements extend up to 6 GeV—such data are available at the LHC, which is useful as a baseline for assessing the importance of different effects.

Our approach is to systematically include different energy loss effects. In particular, we first compare the relative importance of radiative and collisional contribution to the D meson suppression predictions to assess the adequacy of the historically widely used static approximation. We then investigate the importance of including the dynamical scattering centers, followed by the collisional energy loss and the finite size (LPM) effect. Finally, we also address the importance of including the finite magnetic mass and the running coupling.

2. Theoretical and computational frameworks

In this section we first provide a brief overview of the computational framework and our dynamical energy loss formalism. As mentioned above, this formalism leads to a very good agreement with the suppression experimental data across different probes, collision energies and centrality regions [9, 10, 12]. We also introduce how the energy loss expression is modified, as different ingredients are excluded from this formalism. Note that, in section 3, we will for clarity address different energy loss effects in a reverse order, i.e. we will start

from the static approximation and systematically include all the effects, as such a historically-driven approach is more comprehensible and easier to follow.

For studying the importance of different energy loss effects, we will use angular averaged nuclear modification factor R_{AA}, which is well established as a sensitive observable for the interaction of high-momentum particles with the QCD medium. The nuclear modification factor R_{AA} is defined as the ratio of the quenched $A + A$ spectrum to the $p + p$ spectrum, scaled by the number of binary collisions N_{bin}:

$$R_{AA}(p_T) = \frac{\mathrm{d}N_{AA}/\mathrm{d}p_T}{N_{\text{bin}}\, \mathrm{d}N_{pp}/\mathrm{d}p_T}. \tag{1}$$

Furthermore, since angular averaged R_{AA} was previously shown to be sensitive almost entirely to the average properties (temperature) of the evolving medium (in distinction to elliptic flow, v_2, which is considered highly sensitive to the details of the medium evolution) [23, 24], angular averaged R_{AA} can be taken as a 'nearly pure' test of the jet–medium interactions. Due to this, we do not consider the effects of the medium evolution in this study, but provide a detailed study of the importance of different jet–medium effects. For this purpose, we model the medium by assuming an effective temperature of 304 MeV at the LHC (as extracted by ALICE [25]) and effective temperature of 221 MeV at the RHIC (as extracted by PHENIX [26]).

In order to obtain the quenched spectra, we use generic pQCD convolution [9, 27]:

$$\frac{E_f \mathrm{d}^3\sigma}{\mathrm{d}p_f^3} = \frac{E_i \mathrm{d}^3\sigma(Q)}{\mathrm{d}p_i^3} \otimes P\left(E_i \to E_f\right). \tag{2}$$

In equation (2) Q stands for charm quarks and $\frac{E_i \mathrm{d}^3\sigma(Q)}{\mathrm{d}p_i^3}$ denotes the initial charm quark spectrum computed at next-to-leading order [28]. $P(E_i \to E_f)$ is the energy loss probability, which includes both radiative and collisional energy losses in a finite size dynamical QCD medium, multi-gluon [29] and path length [27, 30] fluctuations. The path length distributions are extracted from [30]. Distinction from equation (1) from [9], in our calculations we do not use the fragmentation function $D(Q \to H_Q)$ of the charm quark into the D meson (H_Q), because fragmentation does not alter bare charm quark suppression [10, 11], nor do we use decay function $f(H_Q \to e)$, because D mesons are directly measured in the experiments.

The expression for the radiative energy loss in a finite size dynamical QCD medium [5, 6], obtained from hard-thermal-loop (HTL) approximation, at first order in opacity is given by:

$$\frac{\Delta E_{\text{rad}}}{E} = \frac{C_R \alpha_S}{\pi} \frac{L}{\lambda} \int \mathrm{d}x \frac{\mathrm{d}^2 k}{\pi} \frac{\mathrm{d}^2 q}{\pi}\, v(\mathbf{q}) \left(1 - \frac{\sin \frac{(\mathbf{k} + \mathbf{q})^2 + \chi}{xE^+}L}{\frac{(\mathbf{k} + \mathbf{q})^2 + \chi}{xE^+}L} \right)$$
$$\times \frac{2(\mathbf{k} + \mathbf{q})}{(\mathbf{k} + \mathbf{q})^2 + \chi} \left(\frac{(\mathbf{k} + \mathbf{q})}{(\mathbf{k} + \mathbf{q})^2 + \chi} - \frac{\mathbf{k}}{\mathbf{k}^2 + \chi} \right). \tag{3}$$

In equation (3), $v(\mathbf{q})$ denotes the effective cross section defined below, L is the length of the finite size QCD medium, E is the jet energy, $\mathbf{k}$ is the transverse momentum of the radiated gluon, while $\mathbf{q}$ is the transverse momentum of the exchanged (virtual) gluon and x represents the longitudinal momentum fraction of the jet carried away by the emitted gluon. The color factor is $C_R = \frac{4}{3}$. $\chi = M_c^2 x^2 + m_g^2$, where $m_g = \mu_E/\sqrt{2}$ is the effective (asymptotic) mass for

gluon with the hard momenta $k \gtrsim T$, while μ_E is Debye (electric) screening mass and $M_c = 1.2$ GeV is the charm quark mass. λ is the mean free path in the QCD medium and in the dynamical case is given by $\dfrac{1}{\lambda_{\mathrm{dyn}}} = 3\alpha_S T$. In the incoherent limit [5],

$$\frac{\sin \dfrac{(\mathbf{k} + \mathbf{q})^2 + \chi}{xE^+} L}{\dfrac{(\mathbf{k} + \mathbf{q})^2 + \chi}{xE^+} L} \to 0.$$

The effective cross section, with the included finite magnetic mass [8], is given by the equation below, where μ_M is the magnetic screening mass:

$$v(\mathbf{q}) = \frac{\mu_E^2 - \mu_M^2}{\left(\mathbf{q}^2 + \mu_E^2\right)\left(\mathbf{q}^2 + \mu_M^2\right)}. \tag{4}$$

Note that, in the case when magnetic mass is equal to zero, the above expression reduces to a well-known HTL effective cross section [5, 18]:

$$v(\mathbf{q}) = \frac{\mu_E^2}{\mathbf{q}^2\left(\mathbf{q}^2 + \mu_E^2\right)}. \tag{5}$$

Non-perturbative approaches [31–35] suggest that at the RHIC and the LHC the range of magnetic to electric mass ratio is $0.4 < \mu_M/\mu_E < 0.6$. We therefore use these values in equation (4) when generating suppression predictions in the case of the finite magnetic mass. In the case of zero magnetic mass, we use equation (5) above.

The collisional energy loss is calculated in accordance with [7], i.e. we use equation (14) from that reference for the finite size QCD medium and equation (16) for the incoherent limit.

The running coupling is introduced according to [9] and is defined as in [36]:

$$\alpha_S\left(Q^2\right) = \frac{4\pi}{\left(11 - 2/3n_f\right)\ln\left(Q^2/\Lambda_{\mathrm{QCD}}^2\right)}, \tag{6}$$

where Λ_{QCD} is the perturbative QCD scale ($\Lambda_{\mathrm{QCD}} = 0.2$ GeV) and $n_f = 2.5$ ($n_f = 3$) for the RHIC (LHC) is the number of the effective light quark flavors. In the case of the running coupling, Debye mass μ_E [37] is obtained by self-consistently solving the equation:

$$\frac{\mu_E^2}{\Lambda_{\mathrm{QCD}}^2}\ln\left(\frac{\mu_E^2}{\Lambda_{\mathrm{QCD}}^2}\right) = \frac{1 + n_f/6}{11 - 2/3n_f}\left(\frac{4\pi T}{\Lambda_{\mathrm{QCD}}}\right)^2. \tag{7}$$

Otherwise, when the running coupling is not included, fixed values of the strong coupling constant $\alpha_S = \frac{g^2}{4\pi} = 0.3$ for the RHIC ($\alpha_S = 0.25$ for the LHC) [38] and Debye mass $\mu_E = gT$ are used.

Transition from the dynamical to the static [20] approximation in the case of the radiative energy loss is determined through the following two changes and according to [6]. The mean free path is altered as:

$$\frac{1}{\lambda_{\mathrm{stat}}} = \frac{1}{\lambda_g} + \frac{1}{\lambda_q} = 6\frac{1.202}{\pi^2}\frac{1 + n_f/4}{1 + n_f/6}3\alpha_S T = c(n_f)\frac{1}{\lambda_{\mathrm{dyn}}}, \tag{8}$$

where $c(n_f = 2.5) \approx 0.84$ is a slowly increasing function of n_f that varies between $c(0) \approx 0.73$ and $c(\infty) \approx 1.09$ and the effective cross section changes to:

$$v(\mathbf{q})_{\text{stat}} = \frac{\mu_E^2}{\left(\mathbf{q}^2 + \mu_E^2\right)^2}. \tag{9}$$

3. Results and discussion

In this section, we concentrate on central 200 GeV Au+Au collisions at the RHIC and 2.76 TeV Pb+Pb collisions at the LHC, and investigate how different energy loss ingredients affect the D meson suppression predictions. Regarding the LHC, for which the high momentum D meson R_{AA} data are available [39], we compare our calculations with experimental data in order to visually investigate, both qualitatively and quantitatively, the importance of individual effects in explaining the data.

We will start the analysis from the static approximation, which has been historically the first approach to the energy loss calculations. After investigating the adequacy of the static approximation, we will address the importance of including the dynamical scattering centers, the collisional energy loss and the finite size effect. Finally, we will also investigate the importance of finite magnetic mass and the running coupling.

We therefore start from the static approximation, where we use a fixed value of the strong coupling constant $\alpha_S = \frac{g^2}{4\pi} = 0.3$ at the RHIC ($\alpha_S = 0.25$ at the LHC) and Debye screening mass $\mu_E \approx gT$. Note that these values are used in figures 1–4 and figure 6. Also, note that magnetic mass effect is not included ($\mu_M = 0$) in figures 1–5, while the finite magnetic mass is considered in figures 6 and 7. The running coupling is considered in figures 5 and 7. The finite size QCD medium is considered in each figure, whereas figure 4 investigates the significance of the finite size effect.

To test the adequacy of the widely used static approximation (modeled by the Yukawa potential) [40], we compare the relative importance of radiative and collisional energy loss contributions to the suppression predictions. Namely, in the static approximation, collisional energy loss has to be equal to zero, i.e. the static approximation implies that collisional energy loss can be neglected compared to radiative energy loss. However, in figure 1, we see that the suppression due to collisional energy loss is comparable—or even larger—compared to the radiative energy loss suppression.

This, then, clearly shows that the static approximation is not an adequate one for the D meson suppression calculations, and that the collisional energy loss has to be taken into account in the suppression predictions. Therefore, a number of the approaches which take only radiative energy loss (for an overview see [41])–and some that take only collisional energy loss (e.g. [42, 43]) are clearly not adequate. This can also be directly observed in the right panel of figure 1, where we see that the static approximation leads to a strong disagreement with the data, i.e. to two to three times smaller suppression than the one observed experimentally. Consequently, we will below first test the importance of including the dynamical effects in radiative energy loss (figure 2) and then also test the importance of collisional energy loss within such a dynamical medium (figure 3).

Therefore, in figure 2, we compare the D meson suppression obtained from radiative energy loss only in the static framework, with the one in the dynamical framework. We observe a large difference in the two suppressions, with a significant suppression increase in the dynamical case. Consequently, the dynamical energy loss effect has to be taken into
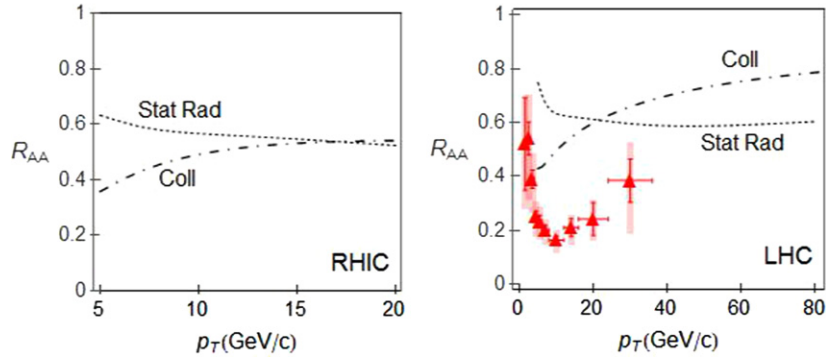
Figure 1. Static radiative versus collisional energy loss suppression. D meson suppression predictions, as a function of transverse momentum, are shown for radiative energy loss only in a static QCD medium (dotted curve), and for collisional energy loss only in a dynamical QCD medium (dot-dashed curve). Left (right) panel corresponds to the RHIC (the LHC) case. Right panel also shows the D meson R_{AA} data in 0–7.5% central 2.76 TeV Pb+Pb collisions at the LHC [39] (red triangles). Debye mass is $\mu_E = gT$, coupling constant is $\alpha_S = 0.3$ ($\alpha_S = 0.25$) for the RHIC (LHC) and finite magnetic mass effect is not included (i.e. $\mu_M = 0$).
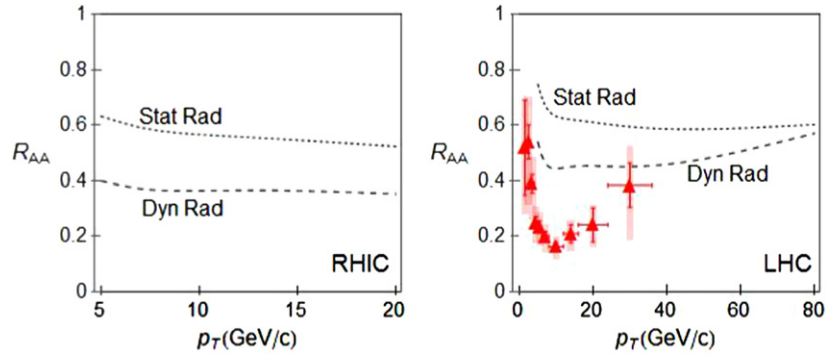

Figure 2. Radiative energy loss suppressions in a static *vs* dynamical QCD medium. D meson suppression predictions are shown, as a function of transverse momentum, assuming only radiative energy loss in static (dotted curve) and in dynamical (dashed curve) QCD medium. Left (right) panel corresponds to the RHIC (the LHC) case. Right panel also shows the D meson R_{AA} data in 0–7.5% central 2.76 TeV Pb+Pb collisions at LHC [39] (red triangles). Debye mass is $\mu_E = gT$, coupling constant is $\alpha_S = 0.3$ ($\alpha_S = 0.25$) for the RHIC (the LHC) and no finite magnetic mass effect is included (i.e. $\mu_M = 0$).

account at the RHIC, as there are no momenta within the RHIC jet momentum range where static approximation becomes adequate. At the LHC, the results indicate that, for jet momentum ranges larger than 100 GeV/c, the static approximation to *radiative* energy loss may become valid, in general agreement with [5, 6, 13, 14]; note, however, that the dynamical effect has to be taken into account even for these momenta, as the collisional energy loss, which is zero in the static approximation, gives a significant contribution to the jet suppression (see the right panel in figure 1). However, despite the fact that inclusion of the
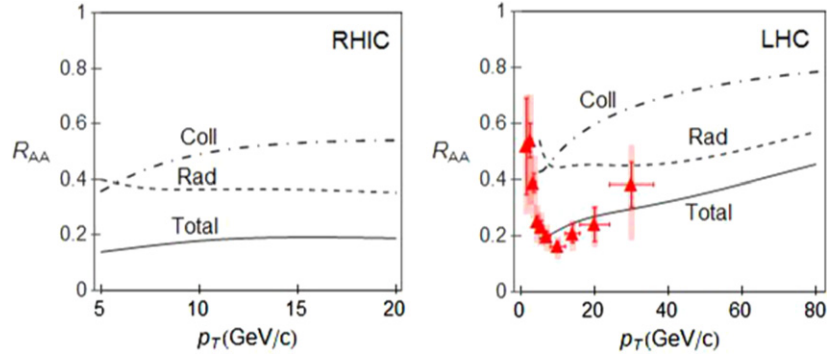
Figure 3. Radiative versus collisional energy loss suppressions in a dynamical QCD medium. D meson suppression predictions are shown, as a function of transverse momentum, for radiative (dashed curve), collisional (dot-dashed curve) and radiative + collisional (solid curve) energy loss. Left (right) panel corresponds to the RHIC (the LHC) case. Right panel also shows the D meson R_{AA} data in 0–7.5% central 2.76 TeV Pb+Pb collisions at the LHC [39] (red triangles). Debye mass is $\mu_E = gT$, coupling constant is $\alpha_S = 0.3$ ($\alpha_S = 0.25$) for the RHIC (the LHC) and no finite magnetic mass effect is included (i.e. $\mu_M = 0$).

Figure 4. Finite size effect on R_{AA}. D meson suppression predictions are shown, as a function of transverse momentum, with (solid curve) and without (dashed curve) finite size effect. Upper (lower) panels correspond to the RHIC (the LHC) case. Left, central and right panel show, respectively, the finite size effect on radiative, collisional and total (radiative + collisional) energy loss in a dynamical QCD medium. Debye mass is $\mu_E = gT$, coupling constant is $\alpha_S = 0.3$ ($\alpha_S = 0.25$) for the RHIC (the LHC) and no finite magnetic mass effect is included (i.e. $\mu_M = 0$).
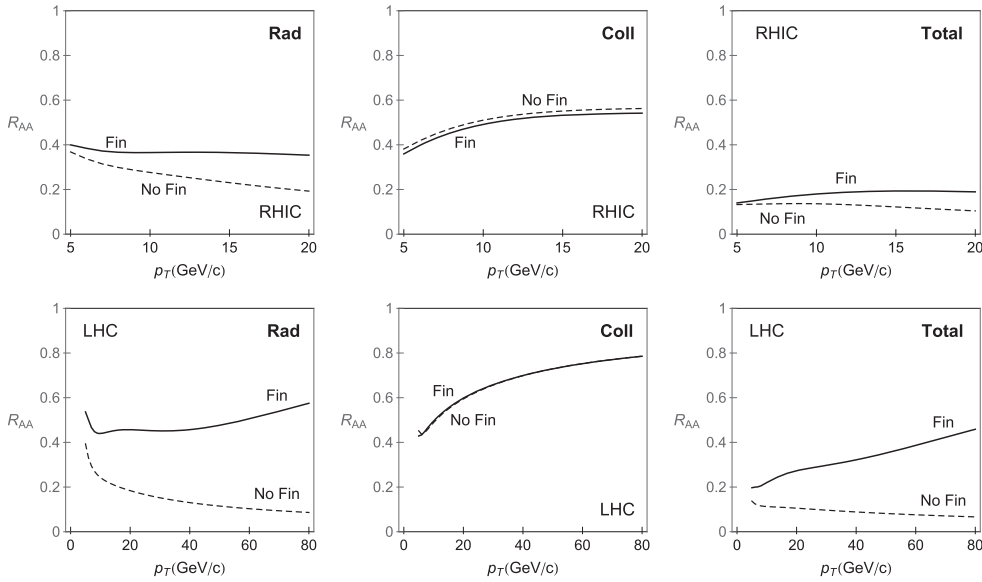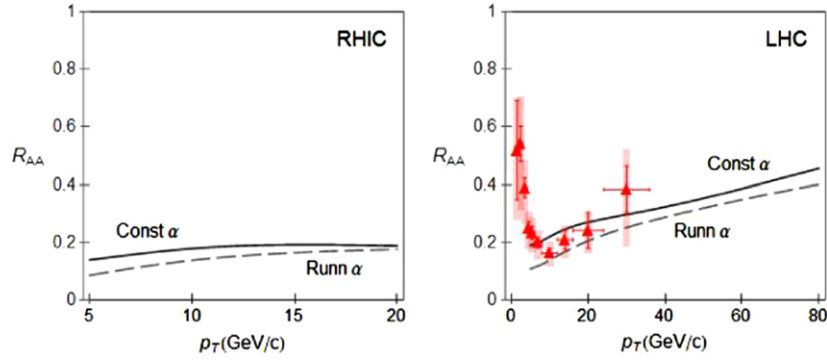
Figure 5. Running coupling effect on R_{AA}. D meson suppression predictions are shown, as a function of transverse momentum, with constant coupling $\alpha_S = 0.3$ ($\alpha_S = 0.25$) for the RHIC (the LHC) (solid curve) and with running coupling (dashed curve). No finite magnetic mass effect is included (i.e. $\mu_M = 0$). In both cases radiative + collisional contributions in dynamical QCD medium are included. Left (right) panel corresponds to the RHIC (the LHC) case. Right panel also shows the D meson R_{AA} data in 0-7.5% central 2.76 TeV Pb+Pb collisions at the LHC [39] (red triangles).

Figure 6. Magnetic mass effect on R_{AA}. D meson suppression predictions are shown, as a function of transverse momentum, for radiative + collisional energy loss in dynamical QCD medium, with (gray band) and without (solid curve) magnetic mass. Left (right) panel corresponds to the RHIC (the LHC) case. Right panel also shows the D meson R_{AA} data in 0–7.5% central 2.76 TeV Pb+Pb colli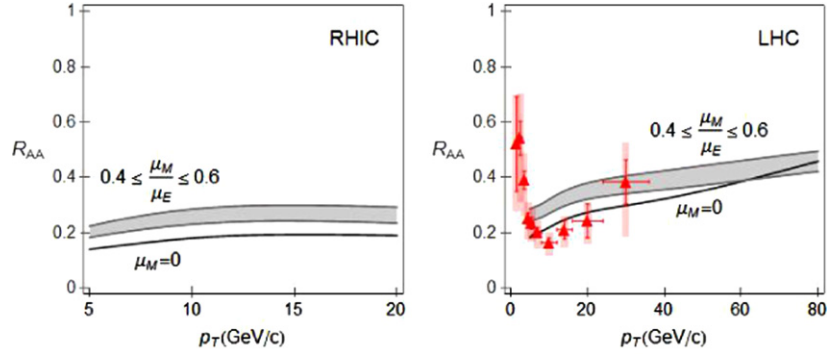sions at the LHC [39] (red triangles). Debye mass is $\mu_E = gT$ and coupling constant is $\alpha_S = 0.3$ ($\alpha_S = 0.25$) for the RHIC (the LHC). Upper (lower) boundary of each band corresponds to $\mu_M/\mu_E = 0.6$ ($\mu_M/\mu_E = 0.4$).

dynamical effect significantly increases the suppression compared to the static approximation, from the right panel in figure 2 we see that, at least below 50 GeV/c, radiative energy loss alone is not able to neither quantitatively nor qualitatively (see the shape of the curve) explain the experimental data, which leads to the conclusion that including only radiative energy loss to model the jet–medium interaction is clearly not adequate.

Furthermore, the results shown in figure 2 imply the question of whether a collisional energy loss is still relevant in the dynamical QCD medium, as suppression due to radiative energy loss significantly increases in the dynamical QCD medium. To address this question,
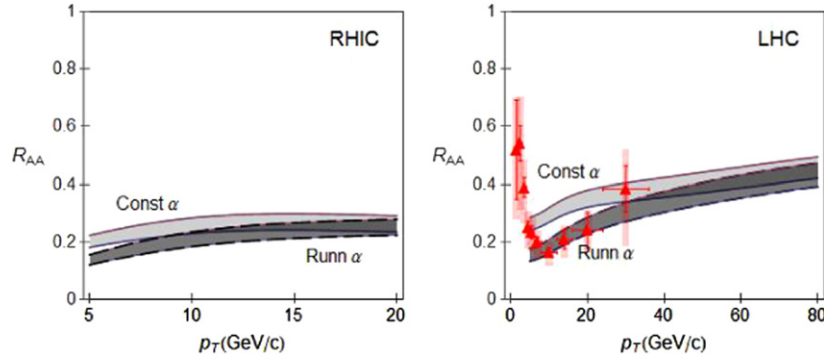
Figure 7. Running coupling and magnetic mass effect on R_{AA}. D meson suppression predictions are shown, as a function of transverse momentum, with the constant coupling $\alpha_S = 0.3$ ($\alpha_S = 0.25$) for the RHIC (the LHC) (light gray band) and with the running coupling (dark gray band). In both cases radiative + collisional contributions in dynamical QCD medium are included. Upper (lower) boundary of each band corresponds to $\mu_M/\mu_E = 0.6$ ($\mu_M/\mu_E = 0.4$). Left (right) panel corresponds to the RHIC (the LHC) case. Right panel also shows the D meson R_{AA} data in 0–7.5% central 2.76 TeV Pb+Pb collisions at the LHC [39] (red triangles).

in figure 3 we compare the D meson suppressions resulting from collisional and radiative energy loss, both calculated in the dynamical QCD medium. We observe that, even when the dynamical effect is accounted for, suppressions from both radiative and collisional contributions are important (consistent with the claims in [7, 44, 45]). This further underscores that collisional energy loss has to be included in the D meson suppression predictions at both the RHIC and the LHC. Moreover, we see that including the collisional contribution increases D meson suppression by up to factor of two compared to the suppression resulting only from dynamical radiative energy loss. Consistent with this observation, we see that the total suppression is significantly larger than either of the two contributions—radiative alone or collisional alone—so that they have to be taken into account jointly for accurate predictions. Furthermore, our main observation from figure 3 is that inclusion of the dynamical effect results in a (rough) agreement with the experimental data, which leads to the conclusion that the dynamical effect is the main/necessary ingredient for accurate description of the jet–medium interactions.

Since we showed that collisional and radiative energy losses are important, we will further investigate how they are affected by the finite size (LPM) effect, as it is commonly considered that this effect is not important for heavy flavor at the RHIC. In figure 4, we separately investigate the finite size effect for radiative (the left panels), collisional (the central panels) and radiative plus collisional (the right panels) energy loss; the top and the bottom panels correspond to the RHIC and the LHC cases, respectively.

We see that for D mesons at both the RHIC and the LHC, the finite size effect is indeed negligible for collisional energy loss, but that they are significant for both radiative and total energy loss suppressions. That is, we see that neglecting LPM effect can lead to as much as two times larger suppression at the RHIC and several times larger suppression at the LHC. In figure 4 we also observe that, LPM effect leads to qualitatively different suppression dependence on momenta, as this effect can lead to a decrease—rather than increase—of suppression with jet momentum. Consequently, the LPM effect has to be taken into account in heavy flavor suppression predictions at both the RHIC and the LHC.

We next consider how the running coupling [9] affects the R_{AA}. Therefore, in figure 5 we compare the D meson suppression predictions obtained by using the fixed value of strong coupling constant, with the predictions when the running coupling is accounted, as a function of the transverse momentum. From figure 5 we observe that the running coupling leads to an increase in the suppression by almost a factor of two at lower jet momenta, while it makes no significant difference at higher jet momenta. Note that such an unequal contribution notably changes the shape of the suppression pattern, so that accounting for the running coupling for D mesons at both the RHIC and the LHC is also important. Furthermore, when comparing the predictions with available (LHC) experimental data (see the right panel of figure 5), we see that inclusion of running coupling leads to a somewhat worse agreement with experimental data, compared to the predictions with constant coupling; we will however see below that inclusion of *both* the running coupling and the finite magnetic mass improves the predictions.

We next investigate the significance of taking into account the finite magnetic mass in the suppression calculations. Namely, all previous energy loss calculations assumed zero magnetic mass, in accordance with the perturbative QCD. However, different non-perturbative approaches [31–35] reported a non-zero magnetic mass at the RHIC and the LHC, which indicates that the finite magnetic mass has to be included in the radiative energy loss calculations [8].

Hence in figure 6 we compare the D meson suppression predictions with and without the finite magnetic mass included in the suppression calculations. To investigate the importance of magnetic mass only, we do not include running coupling in this figure, i.e. we assume the constant coupling. Figure 6 shows that the inclusion of the finite magnetic mass effect leads to a notable $\sim 30\%$ decrease in the suppression. Consequently, the finite magnetic mass effect is also important. Furthermore, when comparing the predictions with available (LHC) experimental data (see the right panel of figure 6), we see that the effect of the inclusion of magnetic mass runs in the opposite direction from the inclusion of running coupling, and also in itself leads to a worse agreement with experimental data (compared to predictions with zero magnetic mass). From this and the previous figure (i.e. Figures 5 and 6), one can conclude that inclusion of the individual improvements in the energy loss calculations—in particular the running coupling alone, or the magnetic mass alone—does not necessarily lead to the improvement in the agreement between the predictions and the data.

Consequently, we finally consider how the inclusion of both the running coupling [9] and the magnetic mass affects R_{AA}. Therefore, in figure 7 we use the finite value of magnetic mass, and compare the D meson suppression predictions with fixed value of strong coupling constant, with those when the running coupling is used, as a function of transverse momentum. We see that these two effects, taken together, lead to a very good agreement with the experimental data, i.e. to both quantitative and qualitative improvement compared to the case in figure 3. This illustrates possible synergy in including different energy loss effects: taken individually, the running coupling and the finite magnetic mass lead to worse agreement with the experimental data, but taken together they notably improve the agreement. Therefore, detailed study of parton energy loss, as well as inclusion of all important medium effects, may be necessary to correctly model the interactions of high-momentum particles with the QCD medium.

4. Conclusion

Since our dynamical energy loss formalism led to a robust agreement with the experimentally measured nuclear modification factor for different experiments, probes and experimental

conditions (i.e. centrality ranges) [9, 10, 12], we investigated how different energy loss effects contribute to such a good agreement. In particular, we aimed at determining whether such a good agreement is a consequence of a single dominant effect or of several smaller improvements. We investigated this issue for the case of D mesons, whose suppression patterns are not modified by the fragmentation functions, so that they present a clear energy loss probe. We used an approach where we started from the simplest reasonable (and historically justified) energy loss model—which includes only radiative energy loss—and then sequentially added different model improvements. This approach both allows investigating the importance of different energy loss ingredients and obtaining the historical perspective on how the energy loss model has been improved. In particular, we studied the importance of the transition from the static to the dynamical framework and of including collisional energy loss, the finite size effect, the finite magnetic mass and the running coupling. As an overall conclusion, we found that the most important effect in modeling jet–medium interactions is the inclusion of the dynamical effect, i.e. modeling the medium constituents as dynamical (moving) particles, instead of the commonly used static scattering centers. However, for a fine agreement with the data, we find that each energy loss effect is important, and that the robust agreement between the theoretical predictions and the experimental data is a cumulative effect of all these improvements. As an outlook, the presented results suggest that further improvements in the energy loss model may be significant for accurately explaining the data even outside of the energy ranges and observables that we have tested so far. Therefore, we expect that data from the upcoming RHIC and LHC runs will help testing—or even further constraining—model calculations at higher transverse momentum.

Acknowledgments

References

[1] Bjorken J D 1982 *FERMILAB-PUB-82-059-THY* 287–92
[2] Brambilla N *et al* 2014 *Eur. Phys. J.* C **74** 2981
[3] Gyulassy M 2002 *Lect. Notes Phys.* **583** 37–9
[4] d'Enterria D and Betz B 2010 *Lect. Notes Phys.* **785** 285
[5] Djordjevic M 2009 *Phys. Rev.* C **80** 064909
[6] Djordjevic M and Heinz U 2008 *Phys. Rev. Lett.* **101** 022302
[7] Djordjevic M 2006 *Phys. Rev.* C **74** 064907
[8] Djordjevic M and Djordjevic M 2012 *Phys. Lett.* B **709** 229
[9] Djordjevic M and Djordjevic M 2014 *Phys. Lett.* B **734** 286
[10] Djordjevic M and Djordjevic M 2014 *Phys. Rev.* C **90** 034910
[11] Djordjevic M 2014 *Phys. Rev. Lett.* **112** 042302
[12] Djordjevic M, Djordjevic M and Blagojevic B 2014 *Phys. Lett.* B **737** 298
[13] Baier R, Dokshitzer Y L, Mueller A H, Peigne S and Schiff D 1997 *Nucl. Phys.* B **483** 291
   Baier R, Dokshitzer Y L, Mueller A H, Peigne S and Schiff D 1997 *Nucl. Phys.* B **484** 265
[14] Baier R, Dokshitzer Y L, Mueller A H and Schiff D 1998 *Phys. Rev.* C **58** 1706
[15] Zakharov B G 1996 *JETP Lett.* **63** 952
   Zakharov B G 1997 *JETP Lett.* **65** 615
[16] Wiedemann U A 2000 *Nucl. Phys.* B **588** 303
[17] Armesto N, Salgado C A and Wiedemann U A 2004 *Phys. Rev.* D **69** 114003

[18] Arnold P B, Moore G D and Yaffe L G 2001 *J. High Energy Phys.* JHEP11(2001)057
     Arnold P B, Moore G D and Yaffe L G 2002 *J. High Energy Phys.* JHEP06(2002)030
     Arnold P B, Moore G D and Yaffe L G 2003 *J. High Energy Phys.* JHEP01(2003)030
[19] Gyulassy M, Levai P and Vitev I 2000 *Nucl. Phys.* B **571** 197
     Gyulassy M, Levai P and Vitev I 2001 *Nucl. Phys.* B **594** 371
[20] Djordjevic M and Gyulassy M 2004 *Nucl. Phys.* A **733** 265–98
[21] Wang X N and Guo X F 2001 *Nucl. Phys.* A **696** 788
[22] Majumder A and van Leeuwen M 2011 *Prog. Part. Nucl. Phys.* A **66** 41
[23] Molnar D and Sun D 2014 *Nucl. Phys.* A **932** 140
     Molnar D and Sun D 2013 *Nucl. Phys.* A **910-911** 486
[24] Renk T 2012 *Phys. Rev.* C **85** 044903
[25] Wilde M (for the ALICE Collaboration) 2013 *Nucl. Phys.* A **904–5** 573c
[26] Adare A *et al* 2010 *Phys. Rev. Lett.* **104** 132301
[27] Wicks S, Horowitz W, Djordjevic M and Gyulassy M 2007 *Nucl. Phys.* A **784** 426
[28] Kang Z B, Vitev I and Xing H 2012 *Phys. Lett.* B **718** 482
[29] Gyulassy M, Levai P and Vitev I 2002 *Phys. Lett.* B **538** 282
[30] Dainese A 2004 *Eur. Phys. J.* C **33** 495
[31] Maezawa Y *et al* WHOT-QCD Collaboration 2010 *Phys. Rev.* D **81** 091501
[32] Maezawa Y *et al* (WHOT-QCD Collaboration) 2008 *Proc. of Sci. Lattice* **2008** 194
[33] Nakamura A, Saito T and Sakai S 2004 *Phys. Rev.* D **69** 014506
[34] Hart A, Laine M and Philipsen O 2000 *Nucl. Phys.* B **586** 443
[35] Bak D, Karch A and Yaffe L G 2007 *J. High Energy Phys.* JHEP08(2007)049
[36] Field R 1995 *Applications of Perturbative QCD* (Cambridge, Massachussetts: Perseus Books)
[37] Peshier A 2006 arXiv:hep-ph/0601119
[38] Betz B and Gyulassy M 2012 *Phys. Rev.* C **86** 024903
[39] Grelli A (ALICE Collaboration) 2013 *Nucl. Phys.* A **904–5** 635c
     Abelev B *et al* 2012 *J. High Energy Phys.* JHEP09(2012)112
[40] Gyulassy M and Wang X N 1994 *Nucl. Phys.* B **420** 583
     Wang X N, Gyulassy M and Plumer M 1995 *Phys. Rev.* D **51** 3436
[41] Armesto N *et al* 2012 *Phys. Rev.* C **86** 064904
[42] Thoma M H 1991 *Phys. Lett.* B **273** 128
[43] Mustafa M G and Thoma M H 2005 *Acta Phys. Hung.* A **22** 93
[44] Mustafa M G 2005 *Phys. Rev.* C **72** 014905
[45] Dutt-Mazumder A K, Alam J, Roy P and Sinha B 2005 *Phys. Rev.* D **71** 094016

Mass tomography at different momentum ranges in quark-gluon plasma

Magdalena Djordjevic,[*] Bojana Blagojevic, and Lidija Zivkovic

Institute of Physics Belgrade (University of Belgrade), Pregrevica 118, 11080 Zemun, Belgrade, Serbia
(Received 9 June 2016; revised manuscript received 29 July 2016; published 17 October 2016)

We here show that at lower momentum (i.e., $p_\perp \sim 10\,\mathrm{GeV}$) single particle suppression for different types of probes exhibit a clear mass hierarchy, which is a direct consequence of the differences in the energy loss, rather than the differences in the initial distributions. On the other hand, we predict that the mass hierarchy is not expected at high momentum (i.e., $p_\perp \sim 100\,\mathrm{GeV}$); i.e., while we surprisingly predict that suppression for charged hadrons will be somewhat *smaller* than the suppression for heavy mesons, we find that this difference will be a consequence of fragmentation functions, not the finite mass effects. That is, apart from the fragmentation functions, the probes of different masses exhibit nearly the same suppression in the high momentum region. We also argue that the same insensitivity on the probe types also appears for jets. In particular, the experimental data in the momentum regions where they exist for both types of probes, show similar suppressions of charged hadrons and inclusive jet data. Interestingly, we also find that our state-of-the-art suppression predictions for high momentum single particles are also in agreement with the jet suppression data, where the reasons behind this agreement yet remain to be understood. Finally, the available jet data also show (though with large error bars) an overlap between b jets (heavy) and inclusive jets (light) probes. Consequently, our results suggest that single particles in the momentum region below 50 GeV present an excellent tool for mass tomography, while high momentum single particles and (possibly) jets are somewhat insensitive to the details of the interaction with quark-gluon plasma.

I. INTRODUCTION

Quark-gluon plasma (QGP) is a new state of matter [1,2] consisting of interacting quarks, antiquarks, and gluons. Such new state of matter is created in ultrarelativistic heavy ion collisions at Relativistic Heavy Ion Collider (RHIC) and Large Hadron Collider (LHC). Rare high momentum probes, which are created in such collisions and which transverse QGP, are excellent probes of this extreme form of matter [3–5]. As these probes have different masses and consequently interact with the medium in a different manner, such mass tomography allows investigating properties of the interactions with the medium [6–8]. Furthermore, as higher momentum ranges become increasingly available at the LHC experiments, there are both different probes and a wide range of their momentum, which become available for such mass tomography. However, there is now a question which exactly probes, and momentum ranges, are optimal for such tomography, i.e., will lead to different behavior that can provide new information about interactions with the medium. To address this question, we will in this paper concentrate on the nuclear modification factor (R_{AA}), as suppression is traditionally considered to be an excellent observable for mass tomography.

As an example, it was previously widely expected that such clear distinction between the suppression patterns will be provided by the measurements of charged hadron (light) and D-meson (heavy) probes (see, e.g., [7,8]). However, as shown by both the experimental data [9,10] and theoretical predictions [11], these two probes have the same suppression at least in the momentum region between 10 and 50 GeV, which is a consequence of a serendipitous interplay between

energy loss and fragmentation functions. Below 10 GeV, there exists a small difference in the R_{AA}s between D mesons and charged hadrons; however, this difference in the suppressions is both small and somewhat influenced by the fragmentation functions [11], so it is, unfortunately, not suitable for extracting any reliable conclusions. Furthermore, at high momentum, recent jet measurements indicate (though with large error bars) the same suppression for b jets [12], and inclusive (light) jets [13,14]. Consequently, there is a nontrivial question of what exactly probes and momentum ranges can be used for obtaining new information on probe-medium interactions. Addressing this will, in turn, allow optimally exploiting experimental efforts and provide further tests of our understanding of QCD matter. Systematically testing the mass tomography effects, for different probes, and at wide momentum ranges, will be the main goal of this paper.

To achieve this goal, we will here use our state-of-the-art dynamical energy loss formalism [15,16], which removes a widely used static approximation and takes into account interactions of the probe with the moving (dynamical) medium constituents. Through this, it consistently treats both radiative [15,16] and collisional [17] energy loss, which was shown to be crucial for quantitatively and qualitatively explaining the experimental data [18]. Additionally, the formalism also takes into account finite magnetic mass [19] and running coupling [20], and is integrated in an up-to-date numerical procedure, which includes path-length [21] and multigluon [22] fluctuations. The formalism was previously shown to be consistent with the wide range of suppression data corresponding to different probes and experimental conditions [11,20,23]. Importantly, *no free parameters* are used in comparing predictions with the experimental data. The same parameter set, corresponding to the standard literature values, will be used in this paper, so that the generated

_____

[*]Corresponding author: magda@ipb.ac.rs

predictions will be also constrained by an agreement with a wealth of previous data.

We will here generate single particle R_{AA} predictions at both lower momentum (i.e., $p_\perp \sim 10$ GeV) and high momentum (i.e., $p_\perp \sim 100$ GeV) regions. Our predictions are applicable for both 2.76 TeV and 5.02 TeV collision energies, because we here predict the same suppression at these two collision energies for light flavor, while we previously [24] predicted the same suppression at these energies for heavy flavor. Comparing these predictions with single particle R_{AA} measurements will allow investigating how suppression depends on the mass hierarchy in different momentum regions, particularly because high precision single particle R_{AA} measurements are (or will soon become) available at both lower and high momentum ranges. In the high momentum range, we will generate predictions for 5.02-TeV collision energy, where preliminary experimental data are currently becoming available. The high momentum single particle predictions are not available for 2.76 TeV, so in this range, we will compare our single particle predictions for 5.02 TeV (which are also applicable to 2.76 TeV; see above) with the available jet measurements. The comparison of the single particle predictions with the available jet data is motivated by the fact that, in the momentum region where both (limited) single particle and jet R_{AA} data exist, these two observables show the same suppression within the error bars, as we present below. This observation leads to a question of how the leading particle R_{AA} predictions, done with state-of-the-art dynamical energy loss model, compare with the whole jet R_{AA}, which we will here address. Consequently, we will here provide a systematic investigation of how the predicted suppression depends on the probe type, the momentum and collision energy range, and how these predictions compare with various available data.

II. METHODS

The numerical framework for generating suppression predictions is presented in detail in [20]. We below briefly outline the main steps in this procedure.

We study the angular averaged nuclear modification factor R_{AA}, which is established as an excellent probe for studying the interaction of high-momentum particles with QGP. R_{AA} is the ratio of the quenched spectrum in $A + A$ collisions to the spectrum in $p + p$ collisions, scaled by the number of binary collisions N_{bin}:

$$R_{AA}(p_\perp) = \frac{dN_{AA}/dp_\perp}{N_{\text{bin}} dN_{pp}/dp_\perp}. \tag{1}$$

To calculate the quenched spectra of light and heavy probes, we use the generic pQCD convolution, which consists of the following steps:

$$\frac{E_f d^3\sigma}{dp_{\perp,f}^3} = \frac{E_i d^3\sigma(Q)}{dp_{\perp,i}^3} \otimes P(E_i \to E_f) \otimes \tag{2}$$

$$\otimes D(Q \to H_Q) \otimes f(H_Q \to J/\psi). \tag{3}$$

Here "i" and "f" subscripts correspond, respectively, to "initial" and "final," E is energy, $p_\perp$ is transversal momentum,

Q denotes partons (quarks and gluons), and the terms in the equation correspond to the following:

(i) $E_i d^3\sigma(Q)/dp_{\perp,i}^3$ denotes the initial parton spectrum. For light quarks and gluons, the spectrum is extracted from [25], while for charm and bottom quarks, the spectrum is extracted from [26].

(ii) $P(E_i \to E_f)$ is the energy loss probability. The probability is generalized to include both collisional [17] and radiative [15,16] energy loss in the same framework (i.e., realistic finite size dynamical QCD medium), which abolishes the widely used approximation of static scattering centers. It is also recently improved to include path-length [21] and multigluon [22] fluctuations, as well as running coupling [20] and finite magnetic mass [19].

(iii) $D(Q \to H_Q)$ is the parton to hadron H_Q fragmentation function. For light hadrons, and D and B mesons we use DSS [27], BCFY [28], and KLP [29] fragmentation functions, respectively. Note, however, that for heavy flavor, fragmentation functions do not influence the suppression of heavy mesons [11]. That is, heavy meson R_{AA} is a true probe of heavy quark R_{AA}.

(iv) For nonprompt J/Ψ, we also have to include the decay of B meson to J/ψ, which is represented by the function $f(H_Q \to J/\psi)$ and obtained according to [26].

To generate the suppression predictions for light and heavy flavor observables in Pb+Pb collisions at the LHC experiments, we used the following parameter set: QGP with perturbative QCD scale of $\Lambda_{\text{QCD}} = 0.2$ GeV and effective light quark flavors $n_f = 3$. In the calculations, as a starting point we use an effective temperature of 304 MeV for 0%–40% centrality Pb+Pb collisions at the LHC [30] experiments (as extracted by ALICE); the average temperature for every centrality region is then determined according to the procedure given in [23]. Also, for every centrality region, we use different path-length distributions, which are provided to us by [31]. The light quark mass is assumed to be dominated by the thermal mass $M = \mu_E/\sqrt{6}$, where temperature-dependent Debye mass μ_E is obtained from [32]. The gluon mass is $m_g = \mu_E/\sqrt{2}$ [33], while the charm and the bottom masses are $M = 1.2$ GeV and $M = 4.75$ GeV, respectively. Magnetic to electric mass ratio is $0.4 < \mu_M/\mu_E < 0.6$, as extracted from several independent nonperturbative QCD calculations [34–37], so the uncertainty in the predictions, presented in the next section, will come from this range of screening masses ratio. Note that our model uses no free parameters in comparison with the experimental data, that is, all the parameters correspond to standard literature values.

III. NUMERICAL RESULTS

In this section, we will generate predictions which correspond to the probe suppression at both lower (~ 10 GeV) and high (~ 100 GeV) momentum ranges. At high-momentum ranges, we will also compare the single particle and jet measurements with each other, and with the generated theoretical predictions. The predictions will be generated both as
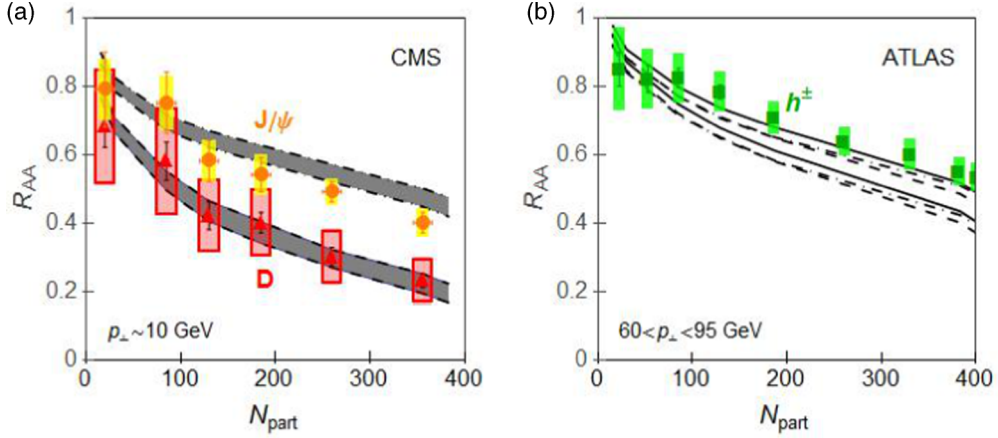
FIG. 1. R_{AA} vs N_{part} for single particles at the 2.76 TeV Pb+Pb collisions at the LHC experiments. (a) Theoretical predictions for R_{AA} vs N_{part} are compared with CMS experimental data for D mesons [38] (red triangles) and nonprompt J/ψ [39] (orange circles) in, respectively, $8 < p_\perp < 16$ GeV and $6.5 < p_\perp < 30$ GeV momentum regions. Gray bands with dashed, and dot-dashed boundaries, respectively, correspond to predictions for D mesons and nonprompt J/Ψ in corresponding momentum regions. (b) Theoretical predictions for $h^\pm$, and D and B mesons R_{AA} vs N_{part} in $60 < p_\perp < 95$ GeV momentum region are, respectively, provided as white bands with full, dashed, and dot-dashed boundaries. The $h^\pm$ predictions are compared with ATLAS (green squares) [40] $h^\pm$ experimental data in the same momentum region. On each panel, the upper (lower) boundary of each gray (or white) band corresponds to $\mu_M/\mu_E = 0.6$ ($\mu_M/\mu_E = 0.4$).

a function of probe momentum and the number of participants and for both light and heavy flavor observables.

We first show predictions for the suppression dependence for single particles on the number of participants at 2.76 TeV Pb+Pb collision energy. In Fig. 1(a) we compare predictions with the data in the lower momentum range ($p_\perp \sim 10$ GeV), while in Fig. 1(b) we compare predictions with the data in the high momentum range ($p_\perp \sim 100$ GeV). Note that the formalism is developed under the assumption that $M^2/E^2 \ll 1$, so, for all types of quarks (both light and heavy), our predictions are valid for $p_\perp \gtrsim 10$ GeV. The predictions in Fig. 1(a) are generated for J/Ψ and D mesons, and, compared with the corresponding CMS experimental data [38], D-meson data from ALICE [41], not shown for figure representation, are consistent with CMS D-meson data. Also, the charged hadrons (light probes) are not shown in Fig. 1(a) for clarity, as it was previously shown that both experimental data [9,10] and theoretical predictions [11] largely overlap with those for D mesons. Because charged hadrons are indirect and complex probes, composed of both light quarks and gluons with the nontrivial effect of fragmentation functions [11], for mass tomography it is clearly beneficial to, whenever possible, concentrate on D mesons (clear charm quark probes [11]) instead of charged hadrons. In Fig. 1(b), the theoretical predictions for charged hadrons, and D and B mesons are generated and shown together with the ATLAS charged hadron experimental data.

A clear distinction in predictions between lower and high $p_\perp$ ranges are observed. In addition, for lower $p_\perp$ [Fig. 1(a)], it is obvious that the light and heavy flavor suppressions are significantly different. On the other hand, in the high $p_\perp \sim 100$ GeV range [Fig. 1(b)], the predictions for all the probes (both light and heavy) almost overlap with each other. From the pQCD perspective, a reason for similar suppressions at high momentum is that the mass of the probe becomes small compared to its momentum, so the relevance of mass

effects should also become small in this region. However, while plausible (expected) from the pQCD perspective, this prediction can be quite distinct in other approaches, as, e.g., AdS-CFT predicts a clear suppression mass hierarchy, even for high momentum regions [42,43].

The experimental data shown in Fig. 1 are in good agreement with the generated theoretical predictions. Moreover, these data also confirm the predicted qualitative distinction between the light or charm and bottom suppressions at the lower momentum region. At the higher momentum range, such comparison between the light and heavy flavor experimental data cannot be made, as the corresponding single particle data are currently available only for charged hadrons. Therefore, the overlap of the light and heavy flavor suppressions at high momentum ranges, provides a clear prediction to be tested by the upcoming experiments.

For understanding the difference between D-meson and nonprompt J/Ψ suppressions, we studied the importance of different contributions to this suppression difference. Regarding this, note that it was considered that this difference may largely originate from the differences in the initial distributions between charm and bottom quarks, rather than the difference in their energy losses [44]. We show in Fig. 2(a) that this is not the case, i.e., the contribution to the suppression difference from the initial distributions is small, while the contribution from the different energy loss is substantially larger. Note, however, that the contributions shown in Fig. 2(a) are not pure effects of initial distribution and energy loss. This is because J/Ψ suppression is not a direct probe of b quarks, i.e., it includes a decay from B mesons to J/Ψ, i.e., $f(B \rightarrow J/\Psi)$. Consequently, to exclude the decay contribution from these two effects, in Fig. 2(b) we show the same contributions but with the B mesons (clear b quark probe) instead of J/Ψ. We see that, in the case of B mesons, the energy loss contribution becomes even larger, while the initial distribution
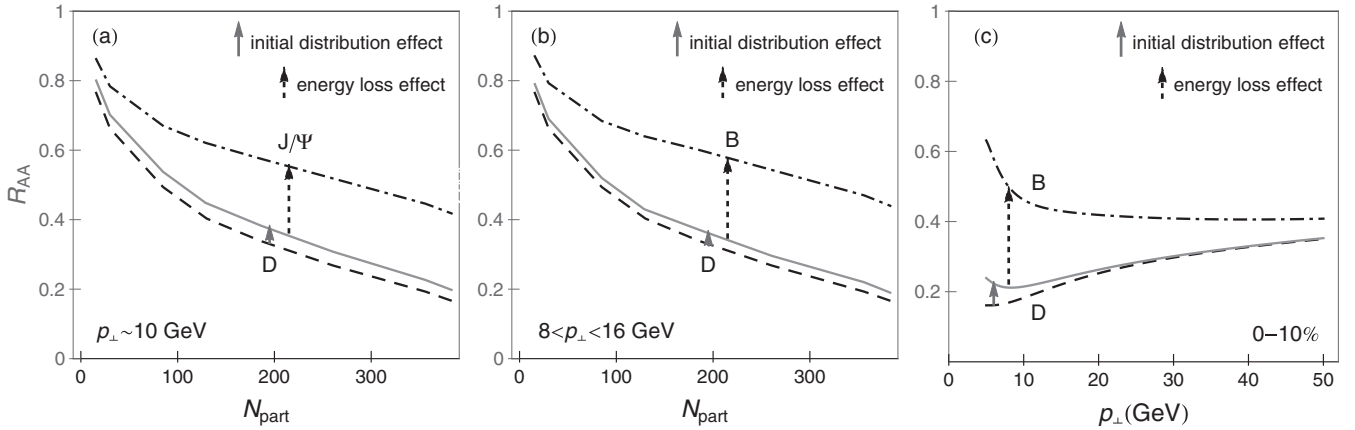
FIG. 2. Suppression contributions at lower momentum. (a) Theoretical predictions for R_{AA} vs N_{part} are compared for D mesons (dashed curve, $8 < p_\perp < 16\,\text{GeV}$ momentum region) and nonprompt J/Ψ (dot-dashed curve, $6.5 < p_\perp < 30\,\text{GeV}$ momentum region). Gray curve shows the analogous nonprompt J/Ψ predictions, if the originating bottom quark would have the same energy loss as charm quark in QGP. (b) Theoretical predictions for R_{AA} vs N_{part} are compared for D and B mesons in $8 < p_\perp < 16\,\text{GeV}$ momentum region. Comparing (a) and (b) shows the effect of decay functions to the contributions analyzed in (a). (c) Theoretical predictions for R_{AA} vs $p_\perp$ are compared for D and B mesons. In (b) and (c), the curve legend is the same as in (a) with distinction that now B replaces J/Ψ. In each panel, the full arrow points to the contribution of the different initial distributions to the difference in the suppression between the D meson and nonprompt J/Ψ (or B meson), while the dashed arrow points to the contribution of the different energy losses to the difference between D-meson and the nonprompt J/Ψ (or B-meson) suppression. Magnetic-to-electric mass ratio is set to $\mu_M/\mu_E = 0.4$.

becomes even smaller. Therefore, the strong mass dependence of the suppression, which is observed and predicted at lower momentum ranges, is clearly a consequence of the differences in the energy loss, rather than the consequence of the initial distributions or decay. Furthermore, we show in Fig. 2(c) that there is no momentum region in which initial distribution makes a significant effect on the suppression difference between different types of probes. Therefore, studying the difference between D- and B-meson suppression patterns in the lower momentum region is not influenced by the production, fragmentation, and the decay, and therefore allows directly assessing how different probes interact with QGP.

While there are currently only limited data for single particles at high momentum, these types of measurements are expected to become increasingly available at 5.02-TeV Pb+Pb collisions at the LHC. At 5.02-TeV collision energy, the R_{AA} measurements for charged hadrons are expected to become available up to $p_\perp \sim 400\,\text{GeV}$, for D mesons the R_{AA} measurements might be available up to $p_\perp \sim 200\,\text{GeV}$, while for bottom mesons the measurements are expected up to $p_\perp \sim 100\,\text{GeV}$ [45] and possibly even higher. It is, therefore, useful providing single particle R_{AA} predictions in the high-momentum region, and studying the effects of high $p_\perp$ mass tomography.

With this goal, in Fig. 3, we provide predictions for charged hadrons, and D and B mesons R_{AA} at 5.02-TeV 0%–10% central Pb+Pb collisions at the LHC. From Fig. 3(a), we see that at $p_\perp \sim 100\,\text{GeV}$, all types of probes show similar suppressions, as supported by Fig. 1(c). However, for $p_\perp > 100\,\text{GeV}$, we also observe that, while D and B (i.e., heavy) meson R_{AA}s become almost identical, the $h^\pm$ (i.e., light hadron) R_{AA} shows a surprising tendency for a *lower* suppression compared to heavy mesons. Moreover, we see that the difference between light and heavy meson suppression

increase with increasing $p_\perp$, leading to more than 10% higher R_{AA} for $h^\pm$ compared to heavy mesons at $p_\perp > 150\,\text{GeV}$. That is, our observation is that, contrary to the 2.76-TeV collision energy where overlap between $h^\pm$ and the D meson R_{AA} was observed for the entire momentum region where both data are available ($p_\perp < 50\,\text{GeV}$), we here predict that increasing momentum (above $p_\perp$ of 100 GeV) will lead to the separation in the R_{AA} of these two observables, but in a different direction than intuitively expected.

A naive conclusion from this prediction would be that, for highly energetic partons, the light partons start to lose notably less energy compared to heavy partons, which is not in accordance with pQCD, as discussed just after Fig. 1. To further investigate this issue, in Fig. 3(b), we compare R_{AA} predictions for bare quarks, i.e., for up, charm, and bottom quarks. We here observe that for $p_\perp > 100\,\text{GeV}$, and in accordance with pQCD, finite mass effects for all types of quarks become negligible, leading to the same suppressions for both light and heavy flavor partons. However, from Fig. 3(c), we see that the nonintuitive result observed in Fig. 3(a), is a consequence of the fragmentation function effect on the light partons that compose the charged hadrons. That is, the effect of fragmentation functions on the light quarks is to decrease their suppression [noted by the vertical arrow in Fig. 3(c)]; the gluon contribution (partially) compensate this effect (as discussed in [11]), but for $p_\perp > 100\,\text{GeV}$, the gluon contribution, and therefore the gluon compensation effect, is small. Because of this, we conclude that, if our predicted larger R_{AA} for $h^\pm$ compared to heavy flavor (D and B) in the high $p_\perp$ region is indeed experimentally observed, this increase will be a pure consequence of the fragmentation function effect, and therefore not related to the mass tomography in the QGP.

Moreover, the predictions presented in Fig. 3 show that the mass tomography effects can be clearly observed below
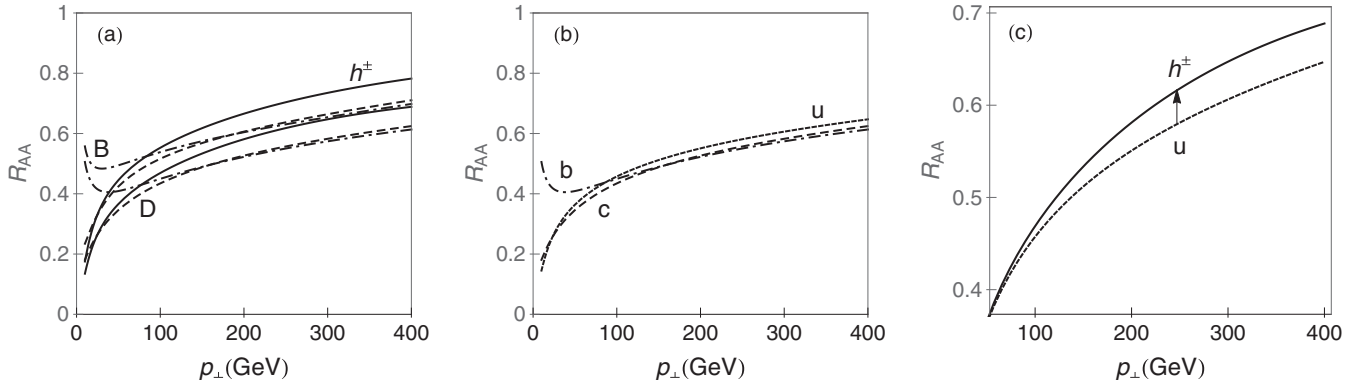
FIG. 3. R_{AA} vs $p_\perp$ for single particles at the 5.02-TeV Pb+Pb 0%–10% central collisions at the LHC. (a) Theoretical predictions for $h^\pm$, and D and B mesons R_{AA} vs $p_\perp$ are, respectively, given as white bands with full, dashed, and dot-dashed boundaries. The upper (lower) boundary of each band corresponds to $\mu_M/\mu_E = 0.6$ ($\mu_M/\mu_E = 0.4$). (b) Theoretical predictions for bare quark R_{AA} vs $p_\perp$ are shown for u (dotted curve), c (dashed curve), and b (dot-dashed curve). μ_M/μ_E ratio is set to 0.4. (c) Theoretical predictions for R_{AA} vs $p_\perp$ are compared for u (dotted curve) with $h^\pm$ (full curve). μ_M/μ_E ratio is set to 0.4.

50 GeV. In particular, we see that below 50 GeV, bottom suppression significantly differs compared to charm and light probes. On the other hand, such a distinction does not appear for high probe momentum (above 50 GeV) where all the suppression predictions nearly overlap (apart from the fragmentation functions effect discussed above). Furthermore, we also showed that indirect bottom probes (i.e., nonprompt J/Ψ) lower the dead-cone [8] effect compared to the clear B-meson probes. Consequently, we propose that one should concentrate on the lower momentum range and on the difference between the B-meson suppression on one side, and D-meson and charged hadron suppression on the other side, for observing significant mass tomography effects.

Furthermore, in [24], we have shown that, for heavy flavor, the R_{AA} predictions for 5.02 TeV and 2.76 TeV overlap with each other, because of interplay between energy loss and initial distributions. In Fig. 4, we show that the same conclusion is valid for charged hadrons as well. We therefore conclude that all the predictions and observations presented in this paper are valid for both 2.76-TeV and 5.02-TeV collision energies.

With regard to this, we note that, while high $p_\perp$ data are not available for single particles at the currently available 2.76-TeV collisions, high $p_\perp$ data are abundant for jets. As our theoretical predictions for single particle R_{AA} data at 5.02 TeV are also applicable for 2.76-TeV collision energy, it is tempting to compare these predictions with the available jet data at 2.76-TeV collision energy. Before comparing single particle predictions with the jet data, we address the same comparison with the experimental data, i.e., we start by asking how the single particle data and the jet measurements correspond to each other, in the momentum range where both are available. Consequently, in Fig. 5, we compare the available experimental data for charged hadrons (the green squares and circles) and inclusive jets (the blue squares and circles). In Fig. 5(a), we show the comparison of the measured suppression dependencies on the probe momentum (for the similar, fixed centrality region), while in Fig. 5(b), we show the comparison of the measured suppression dependencies on the number of participants (for the similar, fixed momentum

region). Therefore, one can see that similar suppressions are observed for single particles and jets, i.e., while the inclusive jets show a somewhat higher suppression compared to charged hadrons, they are the same within the error bars.

The results presented above then motivate us to investigate how our bare quark (i.e., leading particle) suppression predictions, done with the dynamical energy loss, agree with the jet suppression measurements. To that end, in Fig. 6, we show our predictions of R_{AA} vs $p_\perp$ for the light (full curve), charm (dashed curve), and bottom (dot-dashed curve) probes. These leading particle predictions are shown together with inclusive jets from the ATLAS experiments [14] [Fig. 6(a)]) and with both inclusive jets [13] and b jets [12] from CMS

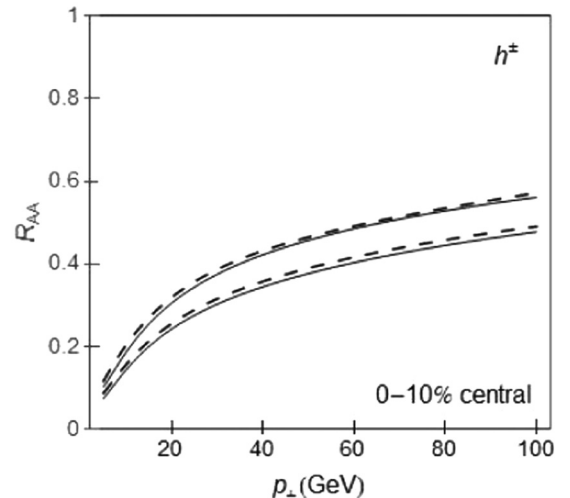FIG. 4. Comparison of R_{AA} predictions for charge hadrons at 2.76 and 5.02 TeV. Charged hadron suppression predictions, as a function of transverse momentum, are shown. R_{AA} predictions at 5.02 TeV (2.76 TeV) 0%–10% central Pb+Pb collisions at the LHC are presented as white bands with full (dashed) boundaries. The upper (lower) boundary of each band corresponds to $\mu_M/\mu_E = 0.6$ ($\mu_M/\mu_E = 0.4$).
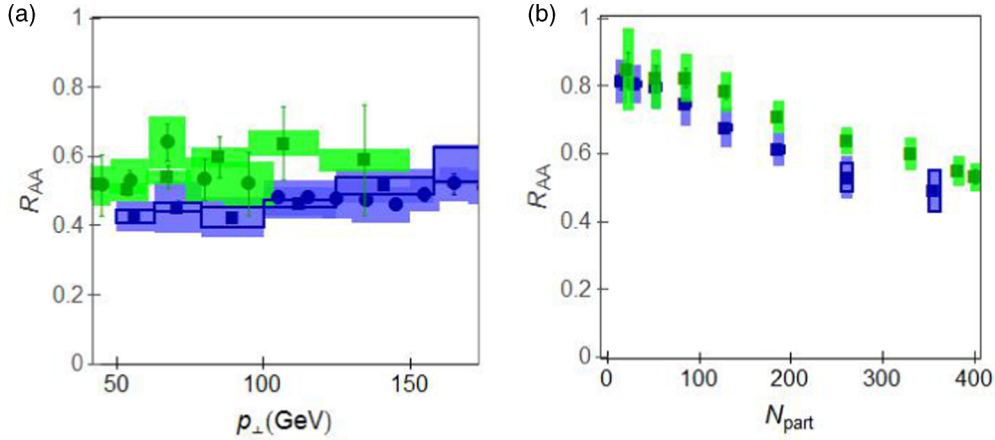
FIG. 5. Comparison of single particle and jet suppression data at the LHC experiments. (a) R_{AA} vs $p_\perp$ experimental data are compared for inclusive jets from ATLAS [14] (blue squares) and CMS [13] (blue circles) and charged hadrons [40] from ATLAS (green squares) and CMS [46] (green circles). ATLAS jet data correspond to 0%–10% centrality, while the other data correspond to 0%–5% centrality. (b) R_{AA} vs N_{part} ATLAS experimental data are compared for inclusive jets [14] (blue squares, $63 < p_\perp < 80 \,\text{GeV}$) and charged hadrons [40] (green squares, $65 < p_\perp < 90 \,\text{GeV}$).

[Fig. 6(b)]. The predictions for both light and heavy probes are in a good agreement with the available jet measurements. This, together with the near overlap of the single particle and the jet suppression data shown in Fig. 5, therefore suggests that the leading particle predictions agree well with the jet R_{AA} measurements.

There are few other important conclusions: (i) Above 50 GeV, we predict almost the same suppressions for the light, charm, and bottom quarks [see also Fig. 3(b)]; (ii) this prediction, extrapolated from the single particle predictions to the light and b jets, is in agreement with the measured experimental data. Because charm jet suppression is not yet measured, our result that the charm suppression overlaps with the light and bottom suppressions, likely suggests that c-jet R_{AA} will overlap with both unidentified and b-jet R_{AA}s.

Finally, the similar conclusion is also obtained if the suppression is analyzed as a function of the number of participants (Fig. 7). In particular, we also see that R_{AA} vs N_{part} single particle predictions for all three types of probes nearly overlap with each other and explain well the inclusive and b-jets data, which are also shown in the figure. Finally, the overlap of the suppression predictions is also consistent with the overlap in the data—similarly, as shown in Fig. 6, the case of the charm jets is a new prediction to be tested by the future measurements.

IV. CONCLUSION

In this paper, we systematically explored the mass tomography effects, which can be observed for different probes and in the wide momentum range corresponding to the span of the available experimental data. The predictions of the suppression dependence from both the momentum and the number of participants were generated and compared with the
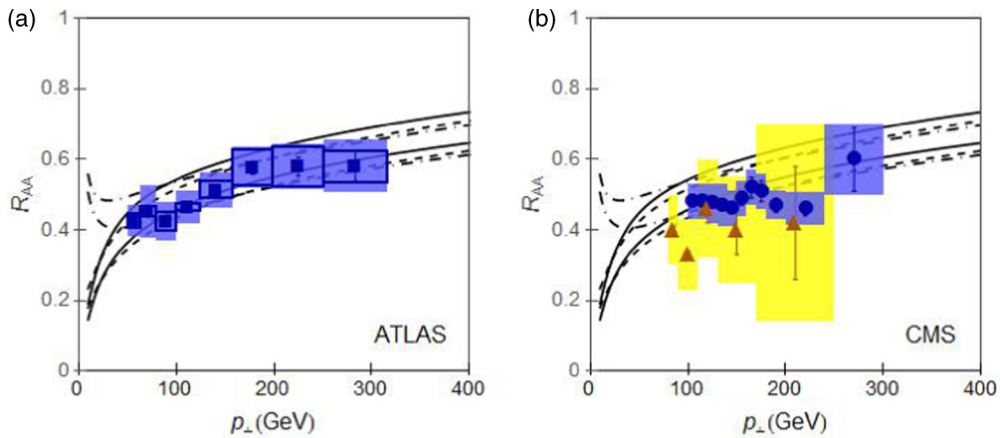
FIG. 6. Single particle suppression predictions vs jet data. (a) Theoretical R_{AA} vs $p_\perp$ predictions for single particles are compared with 0%–10% centrality ATLAS experimental data for inclusive jets [14] (blue squares). (b) R_{AA} vs $p_\perp$ single particle predictions are compared with CMS experimental data for inclusive jets [13] (blue circles, 0%–5% centrality) and b jets [12] (orange triangles, 0%–10% centrality). On each panel, white bands with dashed, dot-dashed, and full boundaries, respectively, correspond to charm, bottom, and light flavor predictions, and the upper (lower) boundary of each band corresponds to $\mu_M/\mu_E = 0.6$ ($\mu_M/\mu_E = 0.4$).
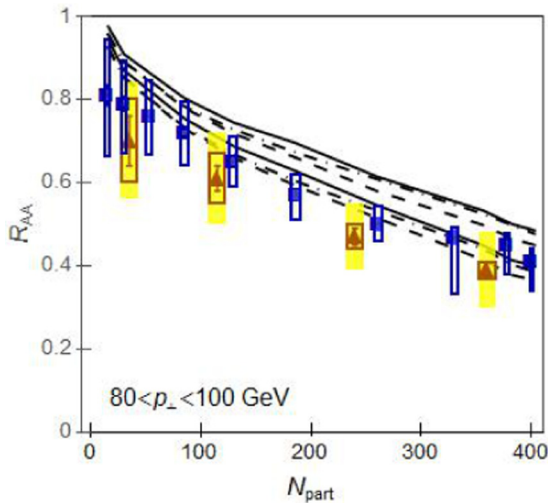
FIG. 7. Single particle suppression predictions vs jet data. Single particle predictions for R_{AA} vs N_{part} are compared with ATLAS data for inclusive jets [14] (blue squares, $80 < p_\perp < 100$ GeV momentum region) and CMS data for b jets [12] (orange triangles, $80 < p_\perp < 90$ GeV momentum region). White bands with dashed, dot-dashed, and full boundaries, respectively, correspond to charm, bottom, and light flavor predictions for $80 < p_\perp < 100$ GeV. The upper (lower) boundary of each band corresponds to $\mu_M/\mu_E = 0.6$ ($\mu_M/\mu_E = 0.4$).

available single particle and jet measurements. As a result, we obtained both the agreement of the theoretical results with the available data, and generated new predictions to be tested in the upcoming experiments, as we briefly summarized below.

For the single particle predictions, we obtained that significant mass tomography effects can be noticed below 50 GeV, related to the difference between the bottom and the charm and light suppressions. While this difference is sometimes attributed to different initial distributions for the charm and bottom quarks, we here showed that this effect is almost entirely a consequence of the differences in the respective energy losses (i.e., the dead-cone effect), while the initial distribution contribution to the difference is almost negligible.

Furthermore, at the existing 2.76-TeV collision energy, we showed that the leading particle predictions agree well with the jet measurements. Moreover, the experimental results show that there is a reasonable overlap between the single

particle and jet suppression experimental data. These findings are interesting, particularly because our suppression approach does not include the features such as jet reconstruction [47] [which are considered crucial for accurate description of (di)jet suppression [48–55]], but includes an advanced dynamical energy loss description for the leading parton. Therefore, the agreement between the single particle and jet R_{AA} measurements, both with respect to the experimental data and the theoretical predictions is currently unclear, and even if this agreement turns out to be accidental, understanding it may provide an important outlook for the future research.

Finally, we here provide clear predictions for the upcoming experimental data at 5.02-TeV collision energy: For the single particle data, we predict that, at the high momentum range $p_\perp > 100$ GeV, B- and D-meson (and likely c- and b-jet) R_{AA} data will nearly overlap with each other. On the other hand, our predictions for $h^\pm$ R_{AA} unintuitively suggest a tendency for lower suppression compared to heavy mesons. We, however, show that this lower suppression is a pure consequence of the fragmentation function effect on $h^\pm$, while the finite mass effect is negligible in this region. Finally, we predicted significant mass tomography effects related to B-meson suppression below 50 GeV. As discussed above, these predictions also provide specific guidelines on where future experimental efforts related to this goal should be concentrated. For example, given these results, we think that it is clearly beneficial to concentrate further efforts on improving b probe data in the relevant momentum region; this can include both directly measuring B mesons instead of nonprompt J/Ψ, reducing the uncertainties, as well as increasing the number of available measurements for this important probe. With regards to this, note that the CMS experiment already published its measurement of the nuclear modification factor for fully reconstructed B mesons in p+Pb collisions [56], while such measurements in Pb+Pb collisions are expected to become available soon from ALICE.

ACKNOWLEDGMENTS

[1] J. C. Collins and M. J. Perry, Phys. Rev. Lett. **34**, 1353 (1975).

[2] G. Baym and S. A. Chin, Phys. Lett. B **62**, 241 (1976).

[3] M. Gyulassy and L. McLerran, Nucl. Phys. A **750**, 30 (2005).

[4] E. V. Shuryak, Nucl. Phys. A **750**, 64 (2005).

[5] B. Jacak and P. Steinberg, Phys. Today **63**(5), 39 (2010).

[6] J. D. Bjorken, FERMILAB-PUB-82-059-THY, 287 (1982), 292

[7] M. Djordjevic, M. Gyulassy, and S. Wicks, Phys. Rev. Lett. **94**, 112301 (2005).

[8] Yu. L. Dokshitzer and D. Kharzeev, Phys. Lett. B **519**, 199 (2001).

[9] B. Abelev *et al.* (ALICE collaboration), Phys. Lett. B **720**, 52 (2013).

[10] J. Otwinowski (ALICE Collaboration), PoS ConfinementX 170 (2012).

[11] M. Djordjevic, Phys. Rev. Lett. **112**, 042302 (2014).

[12] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Rev. Lett. **113**, 132301 (2014).

[13] G. I. Veres, (CMS Collaboration), Nucl. Phys. A **904-905**, 146c (2013).

[14] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. Lett. **114**, 072302 (2015).

[15] M. Djordjevic, Phys. Rev. C **80**, 064909 (2009).

[16] M. Djordjevic and U. Heinz, Phys. Rev. Lett. **101**, 022302 (2008).

[17] M. Djordjevic, Phys. Rev. C **74**, 064907 (2006).

[18] B. Blagojevic and M. Djordjevic, J. Phys. G **42**, 075105 (2015).

[19] M. Djordjevic, Phys. Lett. B **709**, 229 (2012).

[20] M. Djordjevic and M. Djordjevic, Phys. Lett. B **734**, 286 (2014).

[21] S. Wicks, W. Horowitz, M. Djordjevic, and M. Gyulassy, Nucl. Phys. A **784**, 426 (2007).

[22] M. Gyulassy, P. Levai, and I. Vitev, Phys. Lett. B **538**, 282 (2002).

[23] M. Djordjevic, M. Djordjevic, and B. Blagojevic, Phys. Lett. B **737**, 298 (2014).

[24] M. Djordjevic and M. Djordjevic, Phys. Rev. C **92**, 024918 (2015).

[25] Z. B. Kang, I. Vitev, and H. Xing, Phys. Lett. B **718**, 482 (2012); R. Sharma, I. Vitev, and B. W. Zhang, Phys. Rev. C **80**, 054902 (2009).

[26] M. Cacciari, S. Frixione, N. Houdeau, M. L. Mangano, P. Nason, and G. Ridolfi, J. High Energy Phys. 10 (2012) 137.

[27] D. de Florian, R. Sassot, and M. Stratmann, Phys. Rev. D **75**, 114010 (2007).

[28] M. Cacciari and P. Nason, J. High Energy Phys. 09 (2003) 006; E. Braaten, K.-M. Cheung, S. Fleming, and T. C. Yuan, Phys. Rev. D **51**, 4819 (1995).

[29] V. G. Kartvelishvili, A. K. Likhoded, and V. A. Petrov, Phys. Lett. B **78**, 615 (1978).

[30] M. Wilde (ALICE Collaboration), Nucl. Phys. A **904-905**, 573c (2013).

[31] A. Dainese, Eur. Phys. J. C **33**, 495 (2004).

[32] A. Peshier, arXiv:hep-ph/0601119 (2006).

[33] M. Djordjevic and M. Gyulassy, Phys. Rev. C **68**, 034914 (2003).

[34] Y. Maezawa, S. Aoki, S. Ejiri, T. Hatsuda, N. Ishii, K. Kanaya, N. Ukita, and T. Umeda (WHOT-QCD Collaboration), Phys. Rev. D **81**, 091501(R) (2010).

[35] A. Nakamura, T. Saito, and S. Sakai, Phys. Rev. D **69**, 014506 (2004).

[36] A. Hart, M. Laine, and O. Philipsen, Nucl. Phys. B **586**, 443 (2000).

[37] D. Bak, A. Karch, and L. G. Yaffe, J. High Energy Phys. 08 (2007) 049.

[38] CMS Collaboration (B. Hong at Quark Matter 2015)

[39] S. Chatrchyan *et al.* (CMS Collaboration), J. High Energy Phys. 05 (2012) 063.

[40] G. Aad *et al.* (ATLAS Collaboration), J. High Energy Phys. 09 (2015) 050.

[41] B. Abelev *et al.* (ALICE Collaboration), J. High Energy Phys. 09 (2012) 112.

[42] W. A. Horowitz, Phys. Rev. D **91**, 085019 (2015).

[43] W. A. Horowitz, Nucl. Phys. A **904-905**, 186c (2013).

[44] J. Adam *et al.* (ALICE Collaboration), J. High Energy Phys. 11 (2015) 205.

[45] Yen-Jie Lee (private communication).

[46] S. Chatrchyan *et al.* (CMS Collaboration), Eur. Phys. J. C **72**, 1945 (2012).

[47] M. Cacciari, G. P. Salam, and G. Soyez, Eur. Phys. J. C **72**, 1896 (2012).

[48] J. Casalderrey-Solana, J. G. Milhano, and U. A. Wiedemann, J. Phys. G **38**, 035006 (2011).

[49] C. Young, B. Schenke, S. Jeon, and C. Gale, Phys. Rev. C **84**, 024907 (2011).

[50] Y. He, I. Vitev, and B. W. Zhang, Phys. Lett. B **713**, 224 (2012).

[51] F. Senzel, O. Fochler, J. Uphoff, Z. Xu, and C. Greiner, J. Phys. G **42**, 115104 (2015).

[52] X. N. Wang and Y. Zhu, Phys. Rev. Lett. **111**, 062301 (2013).

[53] G. Y. Qin and B. Muller, Phys. Rev. Lett. **106**, 162302 (2011); **108**, 189904 (2012).

[54] G. L. Ma, Phys. Rev. C **87**, 064901 (2013).

[55] K. C. Zapp, F. Krauss, and U. A. Wiedemann, J. High Energy Phys. 03 (2013) 080.

[56] V. Khachatryan *et al.* (CMS Collaboration), Phys. Rev. Lett. **116**, 032301 (2016).

BMC Systems Biology

CrossMark

Understanding key features of bacterial restriction-modification systems through quantitative modeling

Andjela Rodic[1,2†], Bojana Blagojevic[3†], Evgeny Zdobnov[4], Magdalena Djordjevic[3] and Marko Djordjevic[1*]

Abstract

Background: Restriction-modification (R-M) systems are rudimentary bacterial immune systems. The main components include restriction enzyme (R), which cuts specific unmethylated DNA sequences, and the methyltransferase (M), which protects the same DNA sequences. The expression of R-M system components is considered to be tightly regulated, to ensure successful establishment in a naïve bacterial host. R-M systems are organized in different architectures (convergent or divergent) and are characterized by different features, i.e. binding cooperativities, dissociation constants of dimerization, translation rates, which ensure this tight regulation. It has been proposed that R-M systems should exhibit certain dynamical properties during the system establishment, such as: *i*) a delayed expression of R with respect to M, *ii*) fast transition of R from "OFF" to "ON" state, *iii*) increased stability of the toxic molecule (R) steady-state levels. It is however unclear how different R-M system features and architectures ensure these dynamical properties, particularly since it is hard to address this question experimentally.

Results: To understand design of different R-M systems, we computationally analyze two R-M systems, representative of the subset controlled by small regulators called 'C proteins', and differing in having convergent or divergent promoter architecture. We show that, in the convergent system, abolishing any of the characteristic system features adversely affects the dynamical properties outlined above. Moreover, an extreme binding cooperativity, accompanied by a very high dissociation constant of dimerization, observed in the convergent system, but absent from other R-M systems, can be explained in terms of the same properties. Furthermore, we develop the first theoretical model for dynamics of a divergent R-M system, which does not share any of the convergent system features, but has overlapping promoters. We show that *i*) the system dynamics exhibits the same three dynamical properties, *ii*) introducing any of the convergent system features to the divergent system actually diminishes these properties.

Conclusions: Our results suggest that different R-M architectures and features may be understood in terms of constraints imposed by few simple dynamical properties of the system, providing a unifying framework for understanding these seemingly diverse systems. We also provided predictions for the perturbed R-M systems dynamics, which may in future be tested through increasingly available experimental techniques, such as re-engineering R-M systems and single-cell experiments.

Keywords: Restriction-modification, Transcription regulation, Bacterial immune systems, Biophysical modeling, Gene expression dynamics

* Correspondence: dmarko@bio.bg.ac.rs
†Equal contributors
[1]Institute of Physiology and Biochemistry, Faculty of Biology, University of Belgrade, Studentski trg 16, 11000 Belgrade, Serbia
Full list of author information is available at the end of the article

Background

Restriction-modification systems are rudimentary bacterial immune systems, whose main components are the restriction enzyme (R), and the methyltransferase (M). We here consider Type II restriction-modification (R-M) systems [1], where R cuts the same DNA sequences that are protected by M. Consequently, R and M act, respectively, as a toxic molecule and its antidote, and analogies of R-M and toxin-antitoxin systems are often made [2]. R-M present rudimentary "bacterial immune systems", as they protect the host bacterial cell against infection by foreign DNA, such as viruses (bacteriophages) [3–6]. The protection mechanism is straightforward, as the foreign DNA entering bacterial cell is unmethylated, and is consequently cut (destroyed) by R. On the other hand, the host DNA is methylated due to presence of M, and is therefore not cut by R, which prevents autoimmunity. In fact, many bacteriophages are under pressure from R-M systems with whom they have common hosts [7, 8], and have developed different mechanisms to avoid restriction [9–11]. Consequently, expression of the toxic molecule and its antidote provides an effective protection of the bacterial cell against foreign DNA infection [12].

R-M systems are often mobile [2, 12, 13], spreading from one bacterial host to the other, so that a bacterial host, which initially did not contain the R-M system (a naïve host), can acquire it through horizontal transfer. Expression of R and M was directly observed in single cells only very recently, for the Esp1396I system [14], and it is still unclear how different R-M system features affect this expression. It is however assumed that R-M expression has to be tightly regulated during its establishment in a naïve host [15]. For example, as the naïve host genome is initially unmethylated, R must be, and where tested actually is, expressed after a delay with respect to M, so that the host's genomic DNA can be protected before the appearance of R [14, 16, 17]. To ensure such tight regulation, a significant subset of R-M systems contains a third gene, which expresses the control protein (C) [5, 6, 18–23]. C is a transcription factor, which regulates expression of genes in R-M system, including its own expression. In fact, C is typically co-transcribed with R from a common promoter (CR promoter), while M is transcribed from a separate promoter (M promoter) [5, 6, 24].

With respect to the organization of the transcription units, two different architectures are exhibited, which correspond to the convergent (Fig. 1a), and the divergent (Fig. 1b) orientation of CR and M promoters [5, 6, 14, 20, 21, 23, 25, 26]. Despite R-M systems being known for few decades now, with numerous biotechnological uses of restriction enzymes, control of expression of these systems has been insufficiently studied. Two relatively well studied examples are AhdI (a representative of the convergent architecture) [6], and EcoRV (a divergent architecture

representative) [5]. For both systems, the core promoters (binding sites of RNA polymerase), and the binding sites of C protein, are experimentally mapped. In addition, for AhdI system, the transcription activity of CR promoter was measured as a function of C protein amount. We previously showed that a thermodynamic model of CR promoter regulation provides a good agreement with this measurement [6]. We also recently showed [14] that a similar thermodynamic model, coupled with a dynamical model of transcript and protein synthesis, can reasonably explain the dynamics of the enzyme synthesis measured by single-cell experiments in another convergent R-M system (Esp1396I). This strongly suggests that quantitative modeling presented here can realistically explain R-M system transcription control. Additionally, thermodynamical modeling of transcription regulation was successfully applied to a number of different biological problems [27–30], while dynamical modeling was applied to explain both more and less complex gene circuits including control of other convergent R-M systems [31–33].

As we detail below on the example of AhdI (convergent system), and EcoRV (divergent system), it is experimentally firmly established that R-M systems exhibit both different architectures, and different features that characterize their gene expression regulation [1, 15]. On the other hand, the regulation should yield the same three dynamical properties, so that the host genome is protected, while the system is efficiently established. In particular, as discussed above, there would have to be a significant expression of M before R is expressed, to ensure that the host genome is protected. Furthermore, once the host genome is protected, the system should likely turn to "ON" state as rapidly as possible, so that the host genome becomes "immune" to the virus infections – this would then require that after an initial delay, R is rapidly generated. Finally, we also previously proposed that, once the toxic molecule (R) reaches a steady-state, its fluctuations should be low – otherwise a high fluctuation in the toxic molecule (R) may not be matched by the antidote (M), which could destroy the host genome [34].

It is however unclear how the diverse system features and architectures, relate with the constraints on the dynamical response of the system stated above. Experimentally, one could, in principle, address this issue by mutating the relevant features (or introducing them in the system where they do not exist), and then measuring how the resulting system dynamics is perturbed. This would however be very hard, as the system would have to be extensively experimentally mutated and/or redesigned, and the resulting protein dynamics measured in-vivo during the system establishment. In that respect, note that the in-vivo dynamics of R and M expression were directly observed for only two Type II systems – in PvuII via nearly simultaneous introduction into a culture using bacteriophage M13 [17], and in Esp1396I, via
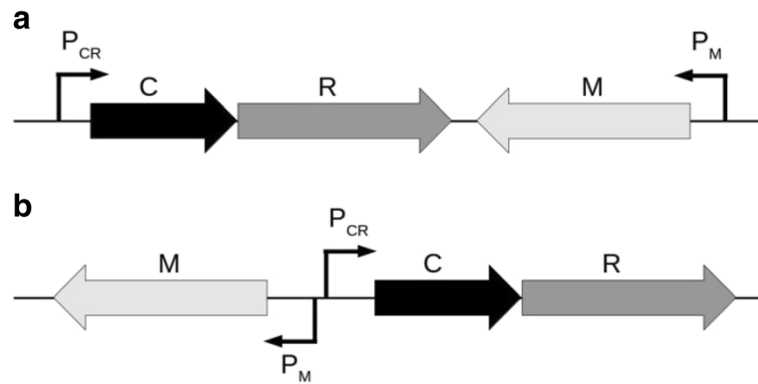
Fig. 1 Typical gene arrangement and promoter orientation in convergent and divergent R-M systems. **a** Convergent systems, a representative of which is AhdI, where other studied systems encoding C protein include Esp1396I, Kpn2I, Csp231I, PvuII [14, 23, 47–49]. Note that C and R genes are transcribed together from P_{CR} promoter. Transcription of M is exhibited from the separate P_M promoter. **b** Divergent systems, a representative of which is EcoRV, where BamHI is another studied divergent system that encods C protein [20]. C and R genes are also co-transcribed, but now share a common promoter region with M gene. In EcoRV the two divergent promoters (P_{CR} and P_M) have overlapping RNA polymerase binding sites

transformation followed by single cell analysis [14]. Even in these cases, the measurements are done only on the wild-type (wt) system, i.e. perturbations were not introduced in the system.

Therefore, the main purpose of this paper is to investigate the relationship between different system features/architectures, and the dynamical properties which the system is expected to exhibit during its establishment. In particular, it is our hypothesis that the diverse features exhibited in R-M systems may largely be explained in terms of the three dynamical properties discussed above. To start testing this hypothesis, we will here biophysically model the control of AhdI and EcoRV, and assess the resulting dynamics when the characteristic system features are either perturbed (in AhdI case) or (artificially) introduced (in EcoRV case) in the system. This is analogous to a classical approach in molecular biology, where the system is analyzed by mutating its main features, or introducing new features in the system where they do not exist, and consequently observing what effect these perturbations have on the presumed system function. The difference is that we here analyze the system *computationally* instead of *experimentally*, where we build on the fact that we previously showed that the modeling approach that we employ here can reasonably explain the available equilibrium measurements [6], and the available single cell experiments [14]. Therefore, the ability of the modeling to explain the measured wild-type data in R-M systems provides a reasonable confidence that our predictions for the perturbed system will also be realistic. Moreover, with the advancement of sophisticated experimental approaches, such as single cell experiments, or possibility to reengineer the system, there comes a prospect of directly experimentally testing these predictions in the future.

Specifically, we will here start by reviewing the relevant experimental information for AhdI and EcoRV

systems (the structure of their promoter regions and their regulatory features), which will provide a bases for our theoretical modeling. We will then quantify the general principles discussed above, i.e. introduce what we here call the dynamical property observables, which will allow us quantifying the delay between R and M, how fast the system makes the transition from OFF to ON state, and the stability of R steady-state levels. We will then investigate if abolishing the characteristic features of AhdI also diminishes these observables, i.e. negatively affects the dynamical properties discussed above. Furthermore, we will also study if these dynamical properties also apply to the system (EcoRV) where AhdI features are absent, but a new feature is present (the overlapping promoters). We will then ask what happens if the AhdI features are (computationally) introduced in wild-type EcoRV system, where they originally do not exist. That is, we will investigate if introducing these features leads to (at least) some of the three dynamic property observables being diminished – therefore explaining why they are absent from EcoRV. Overall, we will here systematically investigate how perturbing (or introducing new) features in two characteristic R-M systems affects the resulting system dynamics.

Methods
In the first subsection, we provide in detail the experimentally available information on AhdI (the convergent system) and EcoRV (the divergent system), on which we base our quantitative modeling. The main properties of the model, including the observables through which we assess the system dynamical properties, are provided in the second subsection. We note that the model itself is provided in details in Additional files 1 and 2, where all

the parameters (including their experimental/theoretical support) are listed.

Experimentally determined configurations of AhdI and EcoRV

For AhdI, the positions of different promoter elements (C protein and RNAP binding sites) were experimentally mapped for both CR and M promoters [6] (see Fig. 1a). In addition, the binding affinities and the transcription activities for both the wild type and mutant systems (where C protein binding sites were mutated) were measured [6]. These measured values, together with the standard literature values for the kinetic parameters (the translation and the degradation rates), were used to parameterize the model, as provided in detail in Additional file 1.

As indicated in Fig. 2a, C binds to CR promoter, regulating both its own transcription and the transcription of R [6, 19]. C binds to promoter DNA as a dimer, where binding to the distal binding site (configuration K_3), when C is present at relatively low concentration, leads to transcription activation, as C dimer bound to this position recruits RNAP binding to the promoter (configuration K_5). On the other hand, when C is present at high concentration, C dimer bound to the distal binding site recruits another C dimer to the proximal binding site (the tetramer configuration, K_4), thus repressing the transcription, as RNAP cannot bind to the promoter. Note that the configuration in which C dimer binds only to the proximal binding site (equivalent to K_3) is not shown, as the binding affinity to the proximal binding

site is much lower compared to the distal binding site, making this configuration much less probable. As for M gene, its transcription is controlled by a negative feedback loop, i.e. M methylates specific sites in its own core promoter thereby repressing the transcription (Fig. 2b).

There are three features which characterize control of AhdI expression [6]. First, there is a very high cooperativity in binding of the C protein dimers to the distal and the proximal positions in CR promoter, so that C dimer bound only to the distal site (K_3 configuration) exists only very transiently in the wild-type (wt) AhdI system. That is, in the absence of RNAP, a C dimer bound to the distal position immediately recruits another C dimer to the proximal binding site. Second, the C dissociation constant of dimerization for AhdI is very high, so that almost all C protein in the solution is in the form of monomers. Finally, C protein is translated from a leaderless transcript (i.e. a transcript which does not contain a ribosome binding site), which was in *E. coli* shown to be associated with lower translation initiation rate [35, 36].

For EcoRV, CR and M promoters are divergently oriented, as schematically shown in Fig. 1b. Consequently, the promoter elements are located in the intergenic region that separates CR and M genes, and these elements are also experimentally mapped [5]. Some of the binding affinities were also measured [5], while the others were eliminated by rescaling the equations (see Additional file 2) – note that we can rescale the equations, as we are interested only in the relative protein amounts. The kinetic parameters (the
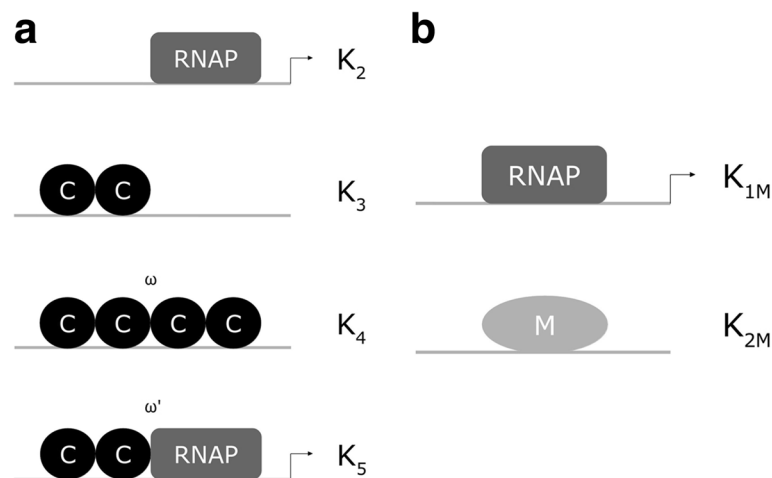
Fig. 2 AhdI R-M system promoter regions. The arrangement of the promoter elements for AhdI CR and M promoters is based on the experimental information provided in [6]. The regions which are schematically shown correspond to (**a**) P_{CR} promoter. Circles indicate C monomers, the rectangles indicates RNAP, while the arrows indicate transcriptionally active configurations. $K_2 - K_5$ denote the dissociation constants (see Additional file 1) corresponding to different promoter configurations (K_1 denotes the dissociation constant of dimerization), where ω and ω' denote, respectively, the binding cooperativity between the two C dimers bound to DNA, and between C dimer bound to the distal binding site and RNAP. C binds to the promoter as either dimer (K_3) or tetramer (K_4). The bound dimer recruits RNAP to the promoter (K_5). On the other hand, the tetramer configuration corresponds to the repression, as it prevents RNAP binding to the promoter. **b** Transcription is repressed by DNA methylation due to M binding [6], i.e. M methylates specific sites in M promoter that overlap RNAP binding site – for simplicity this is in the figure represented as M being bound to the promoter DNA

translation and the degradation rates), correspond to the standard literature values, and are taken to be the same as for AhdI (with the exception of C translation rate, see below).

In contrast to AhdI, the main feature of EcoRV is the partially overlapping CR and M core promoters, as schematically shown in Fig. 3. Consequently, RNAP cannot simultaneously bind to and initiate transcription from both P_M and P_{CR}. Moreover, the characteristic features of AhdI are not found in EcoRV [5]. In particular, while the transcription control of the CR promoter by C protein is similar as in AhdI, the main difference is that the large cooperativity between the C dimers at the distal and the proximal binding site is now absent, in fact it was found in EcoRV that the two dimers bind to DNA with no cooperativity [5]. Furthermore, the transcription from P_M is not directly influenced by C protein binding, i.e. C binding does not directly affect RNAP binding to P_M. However, the influence of C on P_M transcription is indirect, as the regulation by C of RNAP binding to P_{CR}, also affects when RNAP can bind to P_M. Consequently, while in AhdI transcription of CR and M was independent from each other, in EcoRV we have a more complex system where their transcription is strongly coupled. Similar regulation through overlapping CR and M core promoters is also found in CfrBI R-M system [26, 37]. Finally, C transcript is not leaderless in EcoRV, so the feature which was associated with lower translation initiation rate in *E. coli*, and which is present in AhdI, is now absent from EcoRV.

Modeling AhdI and EcoRV dynamics

We model R and M synthesis upon introducing AhdI and EcoRV in naïve bacterial hosts. The models are based on the experimental knowledge of AhdI and EcoRV transcription regulation, which is summarized in Figs. 2 and 3, respectively. The models are provided in detail in Additional files 1 and 2, and are briefly based on:

(i) A thermodynamic model, which takes into account the activation and the repression of CR promoter by C, and the repression of M gene by its own product (which was experimentally shown in [6]). The model assumes that the promoter transcription activity is proportional to the equilibrium binding probability of RNAP to promoter, which is a general assumption initially proposed by the classical Shea-Ackers approach [38].

(ii) Equations that predict how the transcription activity of CR and M promoters depends on C-protein concentration, which further allows modeling the dynamics of transcript and protein expression. That is, the modeled transcription activities provide the main
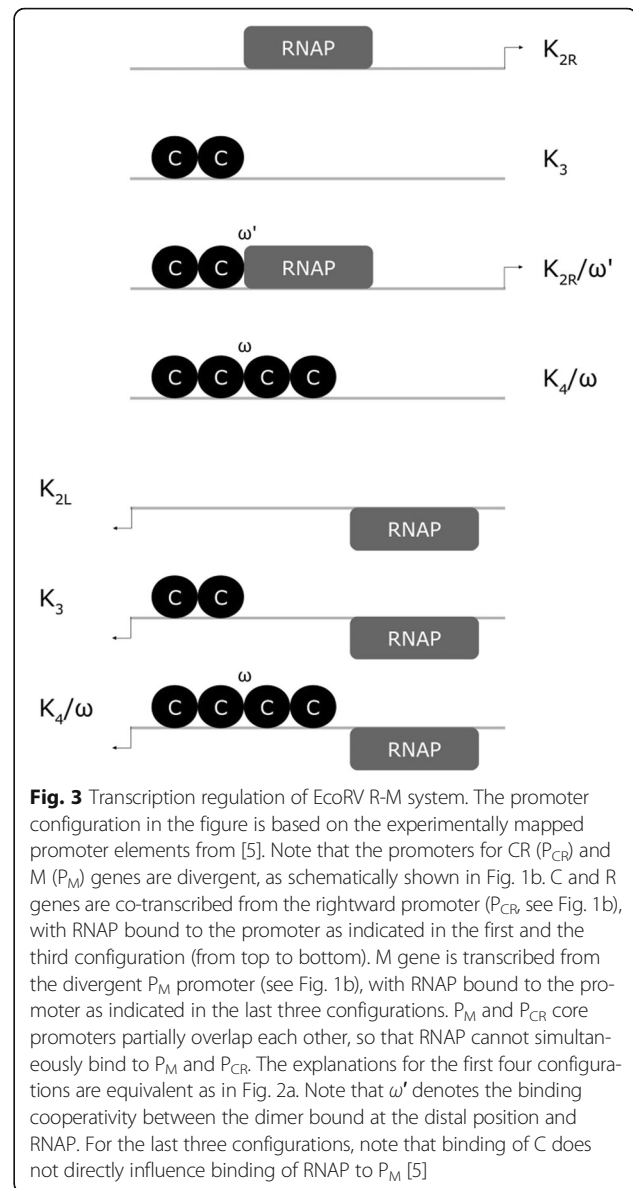

Fig. 3 Transcription regulation of EcoRV R-M system. The promoter configuration in the figure is based on the experimentally mapped promoter elements from [5]. Note that the promoters for CR (P_{CR}) and M (P_M) genes are divergent, as schematically shown in Fig. 1b. C and R genes are co-transcribed from the rightward promoter (P_{CR}, see Fig. 1b), with RNAP bound to the promoter as indicated in the first and the third configuration (from top to bottom). M gene is transcribed from the divergent P_M promoter (see Fig. 1b), with RNAP bound to the promoter as indicated in the last three configurations. P_M and P_{CR} core promoters partially overlap each other, so that RNAP cannot simultaneously bind to P_M and P_{CR}. The explanations for the first four configurations are equivalent as in Fig. 2a. Note that ω' denotes the binding cooperativity between the dimer bound at the distal position and RNAP. For the last three configurations, note that binding of C does not directly influence binding of RNAP to P_M [5]

input for a kinetic model, which calculates R, C and M transcript and protein synthesis. Also, note that R-M systems are characterized by very high expression of R and M proteins [14] so that on the order of thousands of molecules are present in the cell. Consequently, the system is expected to be well in the limit where deterministic modeling can be used to realistically describe the system.

We previously showed that such modeling can well explain the wild-type measurements for AhdI [6] - in particular the measured dependence of the transcription activity on C protein concentration – as well as the most recent measurements in single-cell experiments allowing directly observing the dynamics of R and M synthesis

[14]. Our aim here is to computationally analyze how systematically abolishing individual system features affects the system's dynamics, focusing on the following properties:

  i. the time delay between R and M accumulation,
  ii. the transition speed of the system from "OFF" to "ON",
  iii. the stability of R steady-state levels.

For this, we will introduce observables (which we call the dynamical property observables) that can quantify these properties. To reasonably define them, it is useful to visualize the predicted system dynamics, and the stability of R steady-state levels in wild-type AhdI system, which is shown in Fig. 4 and calculated from Eqs. (1.12), (1.22) and (1.24)–(1.27) (see Additional file 1).

The first dynamical property observable (delay)

From Fig. 4a, we see that the system features lead to a significant delay in the expression of R compared to M, in accordance with the first dynamical property. To quantify how the delay changes upon perturbing these features, we introduce the first dynamical property observable, which corresponds to the ratios of the shaded areas in the perturbed system and in wt AhdI, at an initial interval post-system entry.
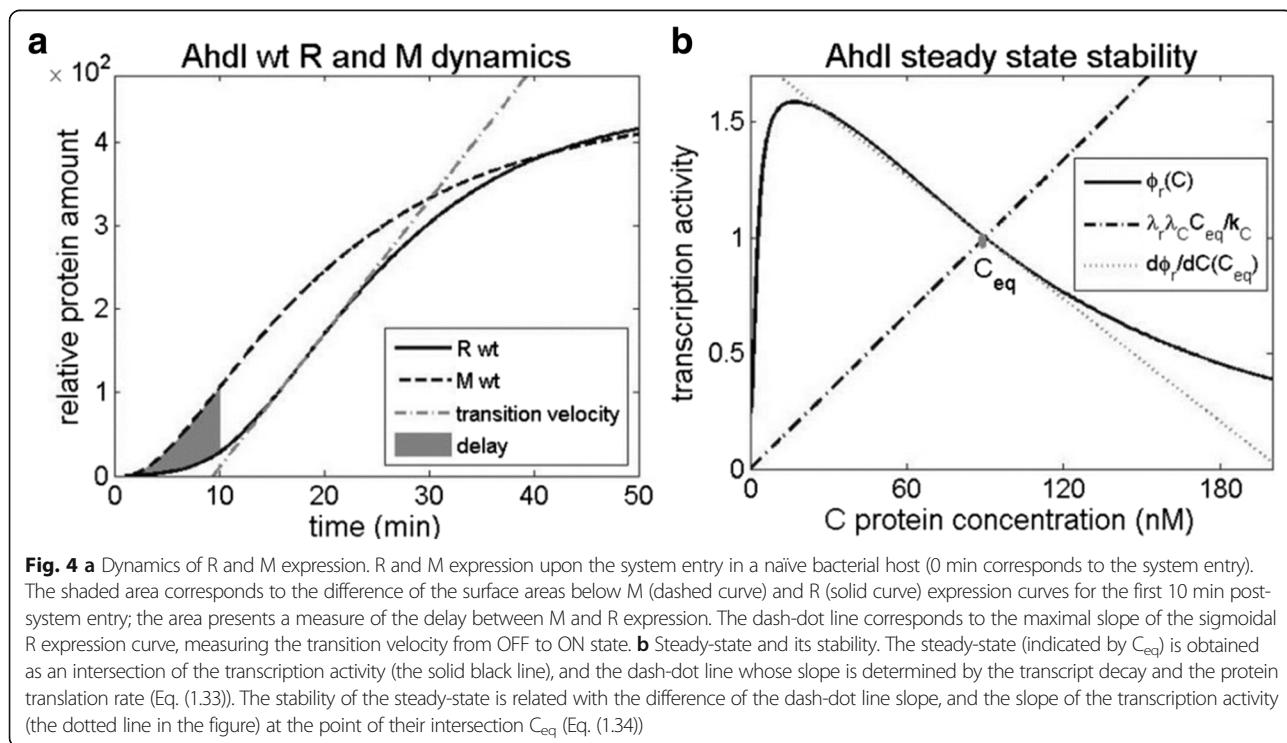
The second dynamical property observable (OFF to ON transition speed)

Furthermore, in Fig. 4a, we see that R expression curve has a sigmoidal shape. Consequently, the maximal slope of this curve (indicated in the figure) provides a reasonable measure of transition velocity from "OFF" (low R value) to "ON" (high R value) state. Therefore, as the second dynamical property observable, we introduce the maximal slope of this curve. The changes of this slope will allow assessing how the transition velocity – which determines the time window between the host genome being methylated, and the cell being protected against viruses – will be affected when the system features are perturbed.

The third dynamical property observable (R steady-state level stability)

Finally, the third dynamical property relates with fluctuations of the toxic R molecule, which we propose should be small in the steady-state [34]. The fluctuations are directly related with the stability of the steady-state, so that smaller fluctuations imply larger steady-state stability, which we introduce as the third dynamical property observable.

Different (in-silico) perturbations of the wild-type system – i.e. gradually abolishing the existing or introducing new features – will be introduced in either the thermodynamic model, or in the kinetic equations (see Additional files 1 and 2).

Fig. 4 a Dynamics of R and M expression. R and M expression upon the system entry in a naïve bacterial host (0 min corresponds to the system entry). The shaded area corresponds to the difference of the surface areas below M (dashed curve) and R (solid curve) expression curves for the first 10 min post-system entry; the area presents a measure of the delay between M and R expression. The dash-dot line corresponds to the maximal slope of the sigmoidal R expression curve, measuring the transition velocity from OFF to ON state. **b** Steady-state and its stability. The steady-state (indicated by C_{eq}) is obtained as an intersection of the transcription activity (the solid black line), and the dash-dot line whose slope is determined by the transcript decay and the protein translation rate (Eq. (1.33)). The stability of the steady-state is related with the difference of the dash-dot line slope, and the slope of the transcription activity (the dotted line in the figure) at the point of their intersection C_{eq} (Eq. (1.34))

Results and discussion

We will start by gradually abolishing the three characteristic AhdI features introduced above, and assess how this will affect the dynamical property observables. We will next model the dynamics of EcoRV establishment in a naïve bacterial host, to see if the proposed dynamical properties also apply to a system with different architecture and transcription regulation features. This will provide, to our knowledge, the first quantitative model of a divergent R-M system control, and an opportunity to assess dynamics of R and M expression, which was up to now not experimentally observed for the divergent systems. Finally, we will in-silico introduce to EcoRV the regulation features that exist in AhdI, but are not found in EcoRV, to investigate how this effects the dynamical property observables, and why these features are not present in EcoRV.

Perturbing AhdI system features

The three characteristic AhdI features are the high C subunit dissociation constant of dimerization, the large cooperativity between C dimers bound at the distal and the proximal position, and the low C transcript translation initiation rate. It was previously discussed that these features serve to limit the amount of the synthesized toxic molecule (R) [6]. However, it is not clear that this amount per-se should be limited, as a too small steady-state amount of R may compromise the immune response – i.e. it can lead to the virus genome being protected by M before it can be destroyed by R [39]. As we discussed above, it would be very hard to experimentally investigate the effect of these AhdI features on the system dynamics, this can be readily predicted from the model that we formulated above.

Decreasing the dissociation constant of dimerization

The dissociation constant of dimerization K_1 is very high for AhdI, leading to almost all C subunits being present as monomers in solution [6, 40] – e.g. for another convergent R-M system (Esp1396I), the measured dissociation constant of dimerization was found to be significantly (four times) lower [41]. We start by gradually decreasing this high dissociation constant of dimerization, in the range that corresponds to the wild-type (all monomers in the solution) to the opposite limit of lower K_1, in which only dimers are present in the solution. In Fig. 5a, we see that this perturbation has a significant effect on R synthesis dynamics – note that the M dynamics curve, which is also indicated in the figure for reference, is not affected by perturbing the three characteristic AhdI features. One can observe the three main effects from Fig. 5a: The decrease of the delay between R and M expression, the slower transition from OFF to ON state, and the decrease in the steady-state

level of R. The first two effects are further quantified in Fig. 5b and c, as discussed below.

In Fig. 5b, we see that decreasing K_1 leads to a significant, more than twofold, decrease in the relative delay between R and M expression. This perturbation can then significantly impact the ability of the system to protect the host genome from being cut during R-M establishment, with the necessary lag also depending on the specific activity of the M protein and the propensity for R to nick hemimethylated sites. Furthermore, in Fig. 5c we see that decreasing K_1 also leads to a significantly slower transition from OFF to ON state, so that the maximal slope is decreased for almost two-fold. Therefore, decreasing the wt dissociation constant of dimerization also significantly impacts the time window in which the host will be protected from foreign DNA infection. However, perturbing K1 has no significant effect on the steady-state stability of R levels (Fig. 5d). Overall, decreasing the high dissociation constant of dimerization characteristic for wt AhdI, has a significant adverse effect on two of the three proposed design principles.

Increasing C protein translation rate

In AhdI C transcript is leaderless [6], which was in *E. coli* [35, 36] shown to be associated with a significantly smaller translation initiation rate – consequently in [6] a five times smaller C transcript translation rate k_C, compared to R and M was assumed. We now test the effect of perturbing this system feature, i.e. increasing k_C towards those of R and M transcripts, which is shown in Fig. 6. We see that the main effect of this perturbation is on decreasing the steady-state level of R and the delay between R and M expression (for ~ 40%), as shown in Fig. 6a-b. Intuitively, this can be understood that by a more efficient C transcript translation, C accumulates faster, facilitating the formation of the activating and the repressing complexes on the CR promoter, so that R is expressed with a smaller delay, and reaches the lower steady-state level. On the other hand, the effect on the other two design-observables, i.e. on the transition velocity and the stability of R steady-state levels, is rather small (Fig. 6c-d). Consequently, increasing the low C transcript translation rate adversely affects one of the dynamical property observables, i.e. the delayed expression of R with respect to M, which is considered crucial for the protection of the host genome.

Decreasing cooperativity in the dimer binding

A rather drastic feature of AhdI is a very large cooperativity ω in binding of the two dimers to the distal and the proximal position in the promoter [6], which is either not present (EcoRV) [5], or significantly smaller (Esp1396I) [41], in other R-M systems. We therefore investigate how gradually abolishing this high cooperativity affects the
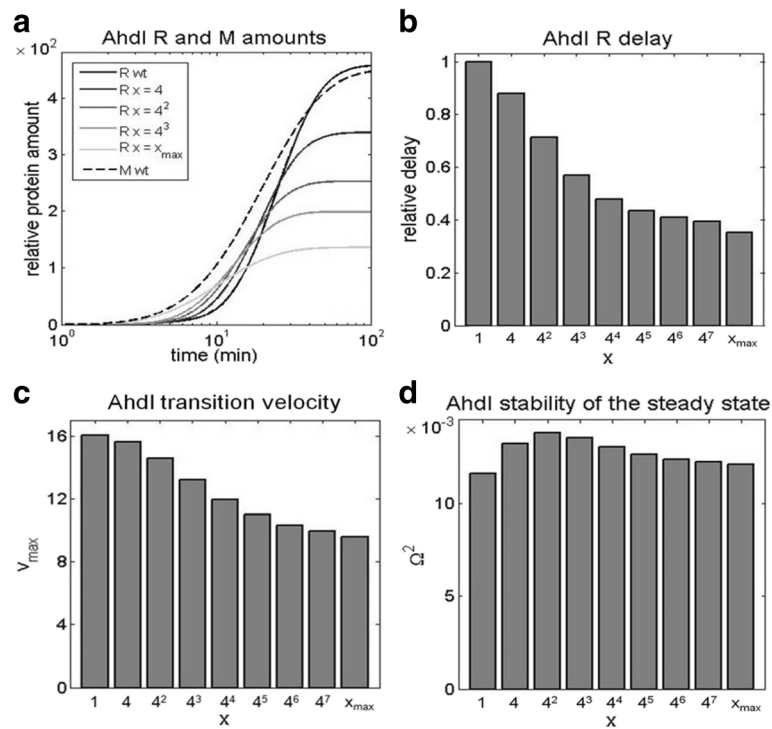
Fig. 5 Decreasing AhdI dissociation constant of dimerization. K_1 is decreased from the high value corresponding to mostly monomers in the solution, to the low value corresponding to mostly dimers in the solution, and the effect is assessed on **a** The dynamics of the protein synthesis. The black line corresponds to all monomers in the solution (wt), while the light gray line corresponds to all dimers in the solution. The curves in-between (in different shades of gray) correspond to the gradually decreasing values of K_1. The relative protein amounts for a wt system (on the vertical axis) are derived from in-vitro transcription activity measurements in [6]. x indicated in the legend corresponds to the relative decrease of K_1 (e.g. x = 4 is a four-fold decrease). **b** The first dynamical property observable, corresponding to the relative delay of R with respect to M expression. The delay is normalized with respect to the wild type (corresponding to one). **c** The second dynamical property observable, corresponding to the transition velocity from "OFF" to "ON" state, represented by the maximal slope of the R expression curve. **d** The third dynamical property observable, corresponding to the stability of R steady-state levels (see Methods)

system dynamics and the design observables. In Fig. 7a, we see that abolishing ω affects only the late dynamics of R, so that the first two dynamical properties are not affected (and not shown in Fig. 7). On the other hand, we see that the steady-state amount of R significantly increases as the cooperativity ω decreases. This can be intuitively understood by the fact that perturbing the cooperativity affects only the efficiency of forming the repressor tetramer complex. As the probability of forming this complex is proportional to C^4 (see Additional file 1), it becomes significant only in the later period, when a large enough amount of C is synthesized. Furthermore, in accordance with the perturbation affecting the late dynamics, from Fig. 7b, we see that decreasing the cooperativity significantly impacts the stability of R steady-state levels, leading to its 50% decrease.

Importantly, the first two AhdI features (the large dissociation constant of dimerization, and the small C translation initiation rate) have an opposite effect on the steady-state amount of R, as compared to the large

cooperativity in C dimer binding. That is, while we showed that the first two features significantly increase the steady-state R amount, the third feature (the large cooperativity) significantly decreases it. On the other hand, all three features generally have the same effect on the three dynamical properties that we consider, i.e. abolishing these features either decreases the values of the dynamical property observables (making the corresponding dynamical property less optimal), or do not significantly affect them. This can then explain the extremely large binding cooperativity that was experimentally observed, as on the one side it allows controlling the steady-state amount of the toxic protein due to the opposite effect from the other two features, while at the same time working together with the first two features to ensure more optimal dynamical properties. In particular, note that both the large dissociation constant of dimerization and the large binding cooperativity significantly increase the stability of R steady-state levels, while having a significant - but opposite – effects on the steady-state R amounts.
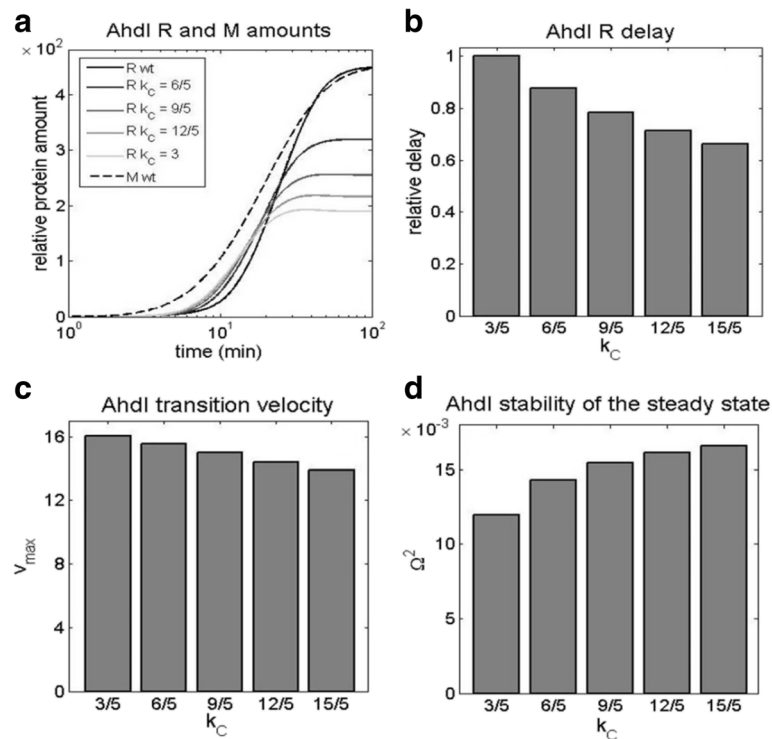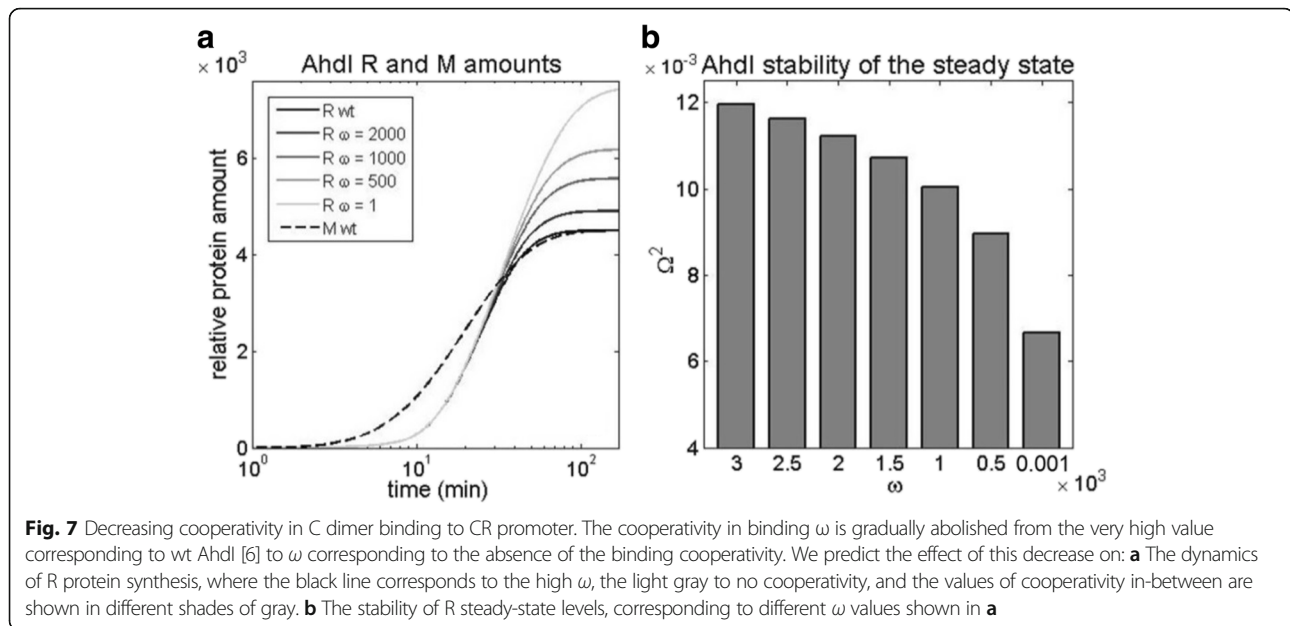
Fig. 6 Increasing C transcript translation rate: k_C is increased from the lower value (3/5 1/min) as taken in [6] to the value which equals those for R and M transcripts (3 1/min). The effect of this decrease is assessed for: **a** The dynamics of the protein synthesis, with the black curve corresponding to the lowest (wt) k_C, and the light gray curve corresponding to the highest k_C (which equals those of R and M transcripts). The curves in different shades of gray correspond to the gradually increasing k_C values. **b** The relative delay (normalized with respect to wt) of R with respect to M expression. **c** The maximal slope of the R expression curve, reflecting the transition velocity from "OFF" to "ON" state. **d** The stability of R steady-state levels, is shown on the vertical axis

EcoRV wild-type dynamics

EcoRV is an example of R-M system with a divergent organization of CR and M transcription units. Overlapping CR and M promoters is the most distinctive feature of this system (presenting its main difference with respect to AhdI), which is, together with C protein binding, responsible for control of EcoRV transcription. That is, high occupancy of M promoter by RNAP, prevents RNAP binding to CR promoter, leading to lower CR transcription activity, and vice versa. In modeling the gene expression regulation, we consider that CR promoter transcription is controlled by C, while C binding has little to none direct effect on M promoter transcription activity, as shown in [5]. In distinction to AhdI [6], which shows an extremely high cooperativity in C dimer binding, no coperativity was found in EcoRV [5]. We also assume that C dissociation constant of dimerization is significantly lower than the relevant range of C concentration, so that the majority of C molecules in solution exist as dimers. Note that in another R-M system (Esp1396I), which has a much lower cooperativity in C dimer binding compared to AhdI, a significantly lower dissociation constant of dimerization is also observed

[41]. Finally, in distinction to AhdI, C transcript in EcoRV is not leaderless, so for EcoRV we assume that C has the same translation initiation rate as R and M.

Consequently, EcoRV does not have the three features that control transcription in AhdI, but has instead another characteristic feature, i.e. the overlapping CR and M promoters. We therefore ask if EcoRV, with different architecture and the regulation features, can also meet the three dynamical properties that we consider. To that end, we modeled the synthesis of R and M during the system establishment in wild-type EcoRV, under the assumptions stated above, and following the scheme of the transcription configurations shown in Fig. 3. The model is provided in detail in Additional file 2, and is based on the same thermodynamics assumptions as the one for AhdI dynamics. To our knowledge, this presents the first model of expression dynamics for a divergent R-M system, which has a more complex regulation due to overlapping nature of their promoters. This model moreover presents the first opportunity to assess the dynamics of R and M synthesis for a divergent R-M system, as, to our knowledge, either their regulation or their expression dynamics was not previously measured.

Fig. 7 Decreasing cooperativity in C dimer binding to CR promoter. The cooperativity in binding ω is gradually abolished from the very high value corresponding to wt AhdI [6] to ω corresponding to the absence of the binding cooperativity. We predict the effect of this decrease on: **a** The dynamics of R protein synthesis, where the black line corresponds to the high ω, the light gray to no cooperativity, and the values of cooperativity in-between are shown in different shades of gray. **b** The stability of R steady-state levels, corresponding to different ω values shown in **a**

The predictions for R and M accumulation in wild-type EcoRV are shown by the full black curve (for R) and by the black dashed curve (for M), in Fig. 8 below. From the figure we see that, regardless of lacking the characteristic AhdI regulatory features, the synthesis of R and M is well in accordance with the three dynamical properties. Namely, by comparing Fig. 4 (the dynamics of AhdI) with the EcoRV dynamics, we see that: *i*) the time delay for EcoRV is even larger compared to AhdI, *ii*) there is a clear switch-like behavior of R expression in EcoRV, i.e. the speed of transition from "OFF" to "ON" state is comparable to the one in AhdI, iii) the system reaches the steady-state level ($\Omega^2 > 0$), where the reached stabilities of R steady-state levels are comparable (compare Fig. 5d with Fig. 8c). Therefore, we see that the design principles which we showed are inherent to AhdI R-M system, are retained in EcoRV R-M system, despite the apparent distinction in gene expression regulation.

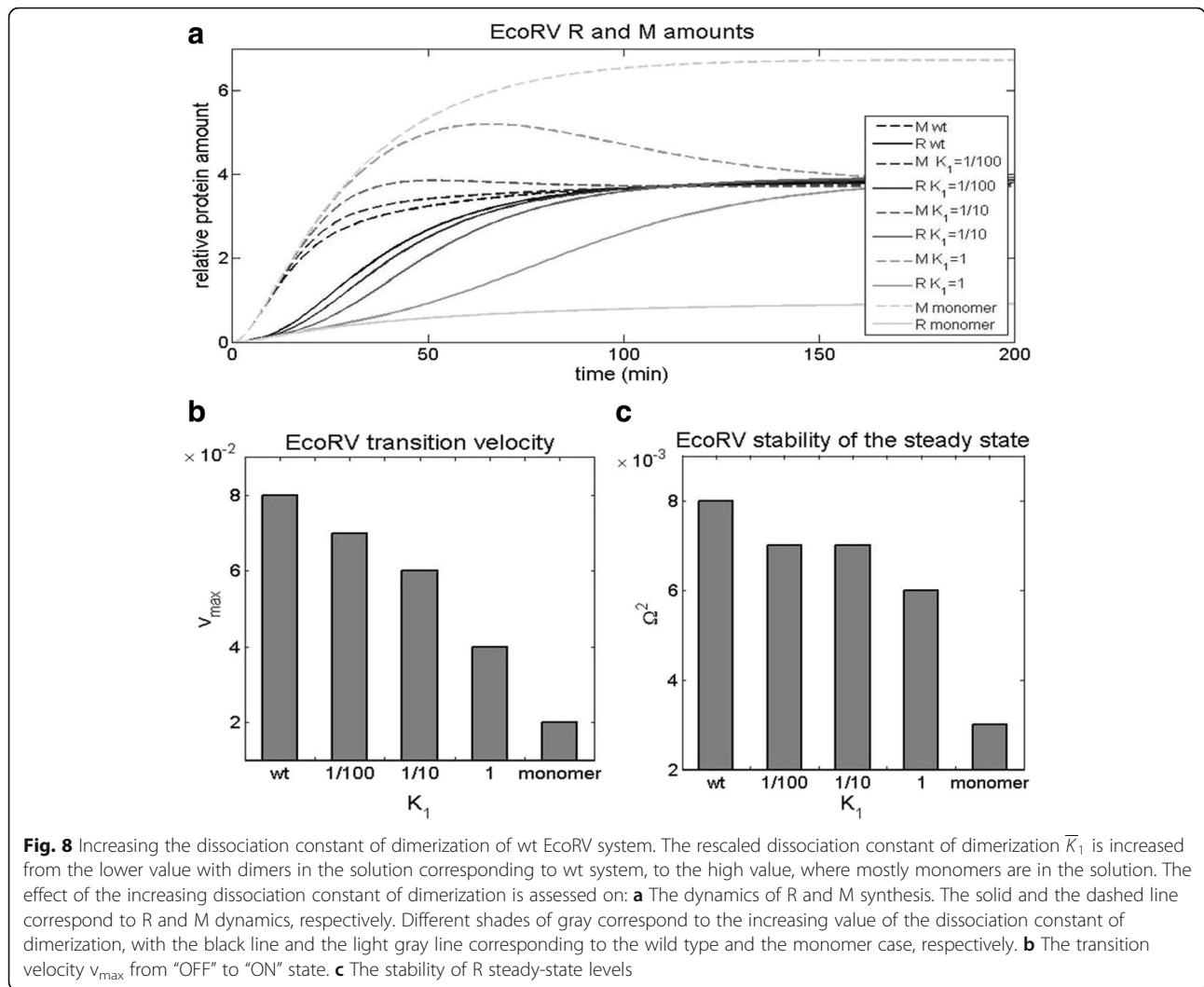**Introducing AhdI control features to EcoRV**
Next, there is a question of why the characteristic AhdI features are absent from EcoRV. That is, could we get even more optimal design-observables if AhdI control features are introduced in wild-type EcoRV? Therefore, we next use our model, to individually introduce each of the three control features of AhdI, on the top of the existing wt EcoRV regulation (i.e. the overlapping promoters). Specifically, in the wild-type EcoRV, we will perturb: *i*) the dissociation constant of dimerization towards the high values characteristic for AhdI, *ii*) cooperativity in C dimer binding to the promoter, also

towards the high values observed in AhdI, *iii*) C protein translation rate k_C, towards the low values characteristic for leaderless AhdI C transcripts.

Introducing the high dissociation constant of dimerization to EcoRV
We first perturb the wt EcoRV system by increasing the rescaled equilibrium dissociation constant of dimerization $\overline{K}_1$ (see Fig.8 and Additional file 2), which corresponds to a gradual transition from the solution containing mostly C dimers to the solution containing mostly C monomers. Note that the dynamics of both R and M expression is now affected by the perturbation, in distinction to AhdI where only R expression is changed. This is because CR and M promoters overlap in EcoRV, so that changing transcription from one promoter, necessarily impacts transcription from the other.

We observe that this perturbation does not significantly affect the early accumulation of R and M (during the first ~10 min), but that the dynamics at later times is significantly affected (see Fig. 8a). In particular, we see that increasing the dissociation constant of dimerization leads to a significantly slower switch from "OFF" to "ON" state, so that the transition velocity decreases as much as four times (Fig. 8b). Furthermore, in Fig. 8c, we see that increasing $\overline{K}_1$ also significantly decreases the stability of R steady-state levels Ω^2, which drops almost three times. Consequently, introducing the high dissociation constant of dimerization to EcoRV, which is characteristic for AhdI, has a significant adverse effect on two of the three dynamical properties.
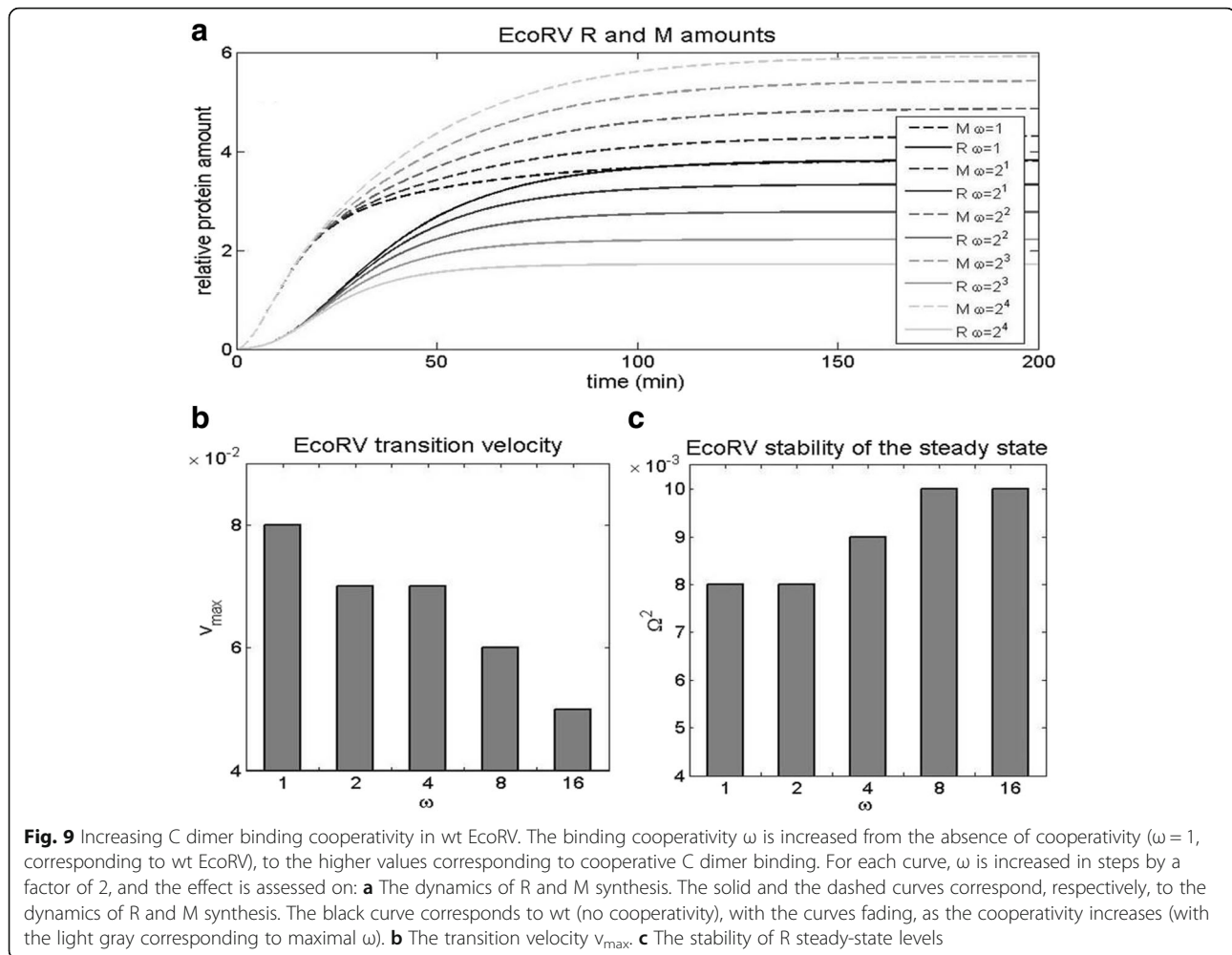
Fig. 8 Increasing the dissociation constant of dimerization of wt EcoRV system. The rescaled dissociation constant of dimerization $\overline{K}_1$ is increased from the lower value with dimers in the solution corresponding to wt system, to the high value, where mostly monomers are in the solution. The effect of the increasing dissociation constant of dimerization is assessed on: **a** The dynamics of R and M synthesis. The solid and the dashed line correspond to R and M dynamics, respectively. Different shades of gray correspond to the increasing value of the dissociation constant of dimerization, with the black line and the light gray line corresponding to the wild type and the monomer case, respectively. **b** The transition velocity v_{max} from "OFF" to "ON" state. **c** The stability of R steady-state levels

Introducing the high C dimer binding cooperativity

We next modify wt EcoRV by increasing the cooperativity ω of C dimer binding to the proximal and the distal binding site, while keeping the other wt EcoRV features unchanged. Note that the experimental measurements in wt EcoRV show an absence of C dimer binding cooperativity ($\omega = 1$) [5], as opposed to the extremely large binding cooperativity that is observed in AhdI [6]. In Fig. 9, we see that increasing ω has the following effects: *i*) the time delay remains nearly the same (Fig. 9a), *ii*) the transition velocity decreases (Fig. 9b), where we see that increasing ω for a relatively moderate factor (2^4), leads to a significant (somewhat less than twofold) decrease of v_{max}, *iii*) stability of R steady-state levels slightly increases. Consequently, we see that perturbing wt EcoRV cooperativity towards the higher values characteristic for AhdI, has a significant adverse effect on one of the dynamical properties (the transition velocity), while not significantly affecting the other two.

Decreasing C translation rate in EcoRV

Finally, we perturb wt EcoRV by decreasing C transcript translation rate k_C, towards the value characteristic for AhdI. Note that C transcript is leaderless in AhdI [6], which is not the case for EcoRV [5], so that we assume the same translation rate for all three transcripts (C, R and M) in EcoRV, while k_C is taken as five times lower in AhdI according to [6]. In Fig. 10a we observe that decreasing k_C does not impact the initial R and M accumulation (during the first ~10 min). On the other hand, at later times the perturbation significantly decreases both the transition velocity that decreases two times (see Fig. 10b), and the stability of R steady-state levels that decreases somewhat less than twofold (see Fig. 10c). Consequently, we see that again two of the three dynamical properties are significantly adversely affected by introducing a control feature from AhdI.

Overall, introducing AhdI characteristic features to EcoRV has a significant adverse effect on at least one of
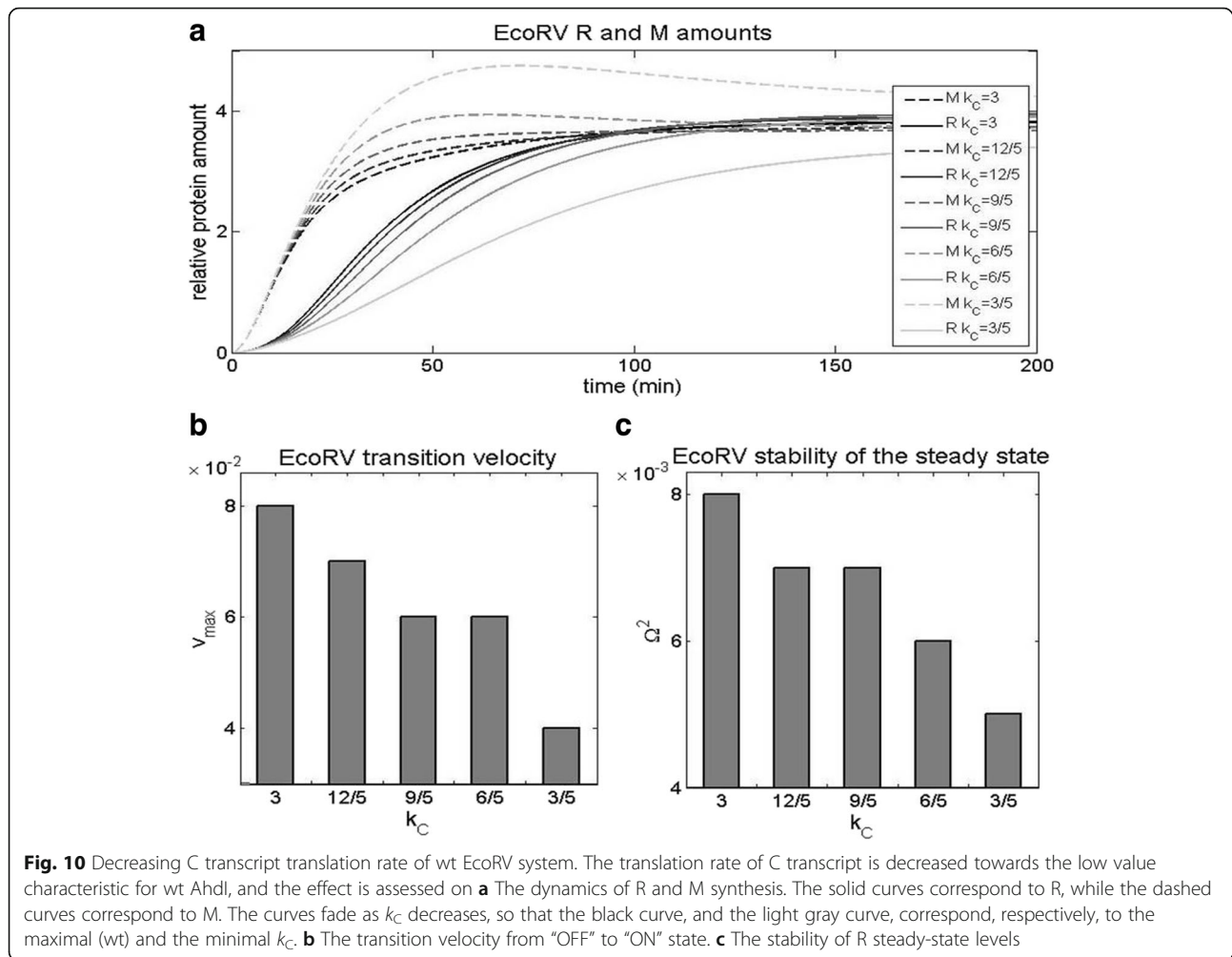
Fig. 9 Increasing C dimer binding cooperativity in wt EcoRV. The binding cooperativity ω is increased from the absence of cooperativity (ω = 1, corresponding to wt EcoRV), to the higher values corresponding to cooperative C dimer binding. For each curve, ω is increased in steps by a factor of 2, and the effect is assessed on: **a** The dynamics of R and M synthesis. The solid and the dashed curves correspond, respectively, to the dynamics of R and M synthesis. The black curve corresponds to wt (no cooperativity), with the curves fading, as the cooperativity increases (with the light gray corresponding to maximal ω). **b** The transition velocity v_{max}. **c** The stability of R steady-state levels

the dynamical properties, which may explain why those features are not found in EcoRV. Additionally, perturbing EcoRV wt parameters towards the AhdI values (Figs. 8a, 9a and 10a) changes M to R ratio in the same direction for each introduced feature (consistently increasing the ratio). This is in distinction to AhdI, where the high cooperativity of C dimer binding has an opposite effect on this ratio, compared to the other two features. Consequently, we argue that another reason for why the characteristic AhdI features are not observed in EcoRV, is because they do not allow balancing the amounts of R and M in the host cell.

Conclusion

R-M systems are characterized by different architectures and control features. We here test a hypothesis that these diverse features can be explained by constraints imposed by few dynamical properties. We started from a relatively well studied AhdI system, and computationally abolished three of its characteristic control features, showing that this has a clear adverse effect on the three

dynamical properties. We then modeled a system with different architecture (EcoRV), and showed that its expression dynamics also satisfies the same properties. The EcoRV model has significance in its own right, as the expression dynamics of the divergent R-M systems was, to our knowledge, not studied before, either theoretically or experimentally. Finally, we computationally introduced to EcoRV the control features that exist in AhdI, and showed that this diminishes at least some of the proposed dynamical properties, consistent with the fact that these features do not appear in wt EcoRV. Moreover, increasing the binding cooperativity has the same effect on M to R ratio in EcoRV as increasing the dissociation constant of dimerization, or lowering the translation rate, which prevents balancing M to R ratio upon introducing these perturbations – this then provides another argument for why AhdI control features are absent from wt EcoRV.

Furthermore, dynamical properties proposed here can provide an explanation for a surprisingly large value of the cooperativity in C protein binding, accompanied by

Fig. 10 Decreasing C transcript translation rate of wt EcoRV system. The translation rate of C transcript is decreased towards the low value characteristic for wt AhdI, and the effect is assessed on **a** The dynamics of R and M synthesis. The solid curves correspond to R, while the dashed curves correspond to M. The curves fade as k_C decreases, so that the black curve, and the light gray curve, correspond, respectively, to the maximal (wt) and the minimal k_C. **b** The transition velocity from "OFF" to "ON" state. **c** The stability of R steady-state levels

the large dissociation constant of dimerization that was observed in wt AhdI. We here showed that these two features have an opposite effect on the steady-state levels of the toxic molecule (R), allowing balancing the steady-state R amount, while at the same time leading to more optimal dynamical properties. In support of this proposal, a similar convergent system with lower binding cooperativity (Esp1396I) was also found to have a lower value of the dissociation constant of dimerization. As a prediction, it will be interesting to test if, in other R-M systems, the value of the dissociation constant of dimerization and the binding cooperativity are also related in this way.

Overall, this work provides an example that the system properties that may appear "random" or even surprising (such as the extremely large binding cooperativity) may be explained by constraints imposed by few general principles (in this case the system dynamical properties). Additionally, some of these system properties may serve other functions, e.g. the leaderless C transcripts might be related with a need for preferential translation under specific

physiological conditions [42]. Analyzing other R-M systems can further test relation of the system features with the simple dynamical properties, where the main obstacle is that their transcription regulation is generally not well studied. In particular, investigating up to now poorly understood linear R-M systems, which have different architecture compared to the convergent and the divergent systems studied here, and which do not encode C proteins – but may exhibit control by antisense RNAs or at the level of translation initiation efficiency - may be particularly useful [43, 44]. As a further outlook, it will be interesting investigating if properties of other bacterial immune systems, such as recently discovered CRISPR/Cas systems [45], can also be explained by similar dynamical properties [34]. With that respect note that CRISPR/Cas is more advanced, i.e. adaptive bacterial immune system, which retains a memory of the past infections incorporated as spacers in the CRISPR array [46].

Also, in this work we follow a standard approach in molecular biology, where features of the system are perturbed/mutated (which is here done in-silico), and the

effect of these perturbations on the presumed system function is assessed. In addition to such "single mutations", a computational equivalent of "double" or "triple" mutations can be exhibited, where more than one system feature would be simultaneously perturbed. This would address the question if perturbations in one feature, can be rescued by also perturbing the other feature(s), which is related to the system robustness. While this question is out of the scope of this work, it also provides an interesting outlook for future research.

Finally, the recent advancement of experimental techniques, such as single-cell experiments, allows directly observing the protein dynamics during the system establishment. While in principle arduous, it would be interesting to experimentally observe how the relevant dynamics is perturbed when some of the key system features are abolished. This would then directly put to test some of the prediction from the computational modelling, which we provided here.

Additional files

Additional file 1: Model of AhdI regulation and dynamics. (PDF 415 kb)

Additional file 2: Model of EcoRV regulation and dynamics. (DOCX 198 kb)

Abbreviations
C: Control protein; M: Methyltransferase; R: Restriction enzyme; R-M: Restriction-modification; RNAP: RNA polymerase; wt: Wild-type

Availability of data and materials
The datasets used and/or analyzed during the current study available from the corresponding author on reasonable request.

Authors' contributions
MJD, MRD and EZ conceived the work. AR and BB performed the analysis. All the authors interpreted the results. MJD, AR and BB wrote the paper, with the help of MRD and EZ. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Consent for publication
Not applicable

Ethics approval and consent to participate
Not applicable

Author details
[1]Institute of Physiology and Biochemistry, Faculty of Biology, University of Belgrade, Studentski trg 16, 11000 Belgrade, Serbia. [2]Multidisciplinary PhD program in Biophysics, University of Belgrade, Belgrade, Serbia. [3]Institute of Physics Belgrade, University of Belgrade, Belgrade, Serbia. [4]Department of Genetic Medicine and Development, University of Geneva and Swiss Institute of Bioinformatics, Geneva, Switzerland.

Published: 24 February 2017

References
1. Pingoud A, Wilson GG, Wende W. Type II restriction endonucleases—a historical perspective and more. Nucleic Acids Res. 2014;42(12):7489–527.
2. Mruk I, Kobayashi I. To be or not to be: regulation of restriction–modification systems and other toxin–antitoxin systems. Nucleic Acids Res. 2014;42(1):70-86.
3. Gingeras TR, Brooks JE. Cloned restriction/modification system from Pseudomonas aeruginosa. Proc Natl Acad Sci. 1983;80(2):402–6.
4. Kiss A, Posfai G, Keller CC, Venetianer P, Roberts RJ. Nudeotide sequence of the BsuRI restriction-modification system. Nucleic Acids Res. 1985;13(18):6403–21.
5. Semenova E, Minakhin L, Bogdanova E, Nagornykh M, Vasilov A, Heyduk T, Solonin A, Zakharova M, Severinov K. Transcription regulation of the EcoRV restriction-modification system. Nucleic Acids Res. 2005;33(21):6942–51.
6. Bogdanova E, Djordjevic M, Papapanagiotou I, Heyduk T, Kneale G, Severinov K. Transcription regulation of the type II restriction-modification system AhdI. Nucleic Acids Res. 2008;36(5):1429–42.
7. Krüger D, Bickle TA. Bacteriophage survival: multiple mechanisms for avoiding the deoxyribonucleic acid restriction systems of their hosts. Microbiol Rev. 1983;47(3):345.
8. Tock MR, Dryden DT. The biology of restriction and anti-restriction. Curr Opin Microbiol. 2005;8(4):466–72.
9. Korona R, Korona B, Levin BR. Sensitivity of naturally occurring coliphages to type I and type II restriction and modification. Microbiology. 1993;139(6):1283–90.
10. Makino O, Saito H, Ando T. Bacillus subtilis-phage φ1 overcomes host-controlled restriction by producing BamNx inhibitor protein. Mol Gen Genet MGG. 1980;179(3):463–8.
11. Takahashi I, Marmur J. Replacement of thymidylic acid by deoxyuridylic acid in the deoxyribonucleic acid of a transducing phage for Bacillus subtilis. 1963.
12. Kobayashi I. Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. Nucleic Acids Res. 2001;29(18):3742–56.
13. Jeltsch A, Pingoud A. Horizontal gene transfer contributes to the wide distribution and evolution of type II restriction-modification systems. J Mol Evol. 1996;42(2):91–6.
14. Morozova N, Sabantsev A, Bogdanova E, Fedorova Y, Maikova A, Vedyaykin A, Rodic A, Djordjevic M, Khodorkovskii M, Severinov K. Temporal dynamics of methyltransferase and restriction endonuclease accumulation in individual cells after introducing a restriction-modification system. Nucleic Acids Res. 2016;44(2):790–800.
15. Nagornykh M, Bogdanova E, Protsenko A, Solonin A, Zakharova M, Severinov K. Regulation of gene expression in a type II restriction-modification system. Russ J Genet. 2008;44(5):523–32.
16. Ichige A, Kobayashi I. Stability of EcoRI restriction-modification enzymes in vivo differentiates the EcoRI restriction-modification system from other postsegregational cell killing systems. J Bacteriol. 2005;187(19):6612–21.
17. Mruk I, Blumenthal RM. Real-time kinetics of restriction–modification gene expression after entry into a new host cell. Nucleic Acids Res. 2008;36(8):2581–93.
18. Ball NJ, McGeehan J, Streeter S, Thresh S-J, Kneale G. The structural basis of differential DNA sequence recognition by restriction–modification controller proteins. Nucleic acids research 2012;40(20):10532–42.
19. McGeehan JE, Papapanagiotou I, Streeter SD, Kneale GG. Cooperative binding of the C.AhdI controller protein to the C/R promoter and its role in endonuclease gene expression. J Mol Biol. 2006;358(2):523–31.
20. Sohail A, Ives CL, Brooks JE. Purification and characterization of C·BamHI, a regulator of the BamHI restriction-modification system. Gene. 1995;157(1):227–8.
21. Sorokin V, Severinov K, Gelfand MS. Systematic prediction of control proteins and their DNA binding sites. Nucleic Acids Res. 2009;37(2):441–51.

22. Ives CL, Nathan PD, Brooks JE. Regulation of the BamHI restriction-modification system by a small intergenic open reading frame, bamHIC, in both Escherichia coli and Bacillus subtilis. J Bacteriol. 1992;174(22): 7194–201.

23. Tao T, Bourne J, Blumenthal R. A family of regulatory genes associated with type II restriction-modification systems. J Bacteriol. 1991;173(4):1367–75.

24. Česnavičienė E, Mitkaitė G, Stankevičius K, Janulaitis A, Lubys A. Esp1396I restriction–modification system: structural organization and mode of regulation. Nucleic Acids Res. 2003;31(2):743–9.

25. Karyagina A, Shilov I, Tashlitskii V, Khodoun M, Vasil'ev S, Lau PC, Nikolskaya I. Specific binding of SsoII DNA methyltransferase to its promoter region provides the regulation of SsoII restriction-modification gene expression. Nucleic Acids Res. 1997;25(11):2114–20.

26. Zakharova M, Minakhin L, Solonin A, Severinov K. Regulation of RNA polymerase promoter selectivity by covalent modification of DNA. J Mol Biol. 2004;335(1):103–11.

27. Vilar JM, Saiz L. Systems biophysics of gene expression. Biophys J. 2013; 104(12):2574–85.

28. Tolkunov D, Morozov AV. Genomic studies and computational predictions of nucleosome positions and formation energies. Adv Protein Chem Struct Biol. 2010;79:1–57.

29. Gertz J, Siggia ED, Cohen BA. Analysis of combinatorial cis-regulation in synthetic and genomic promoters. Nature. 2009;457(7226):215–8.

30. Seo Y-J, Chen S, Nilsen-Hamilton M, Levine HA. A mathematical analysis of multiple-target SELEX. Bull Math Biol. 2010;72(7):1623–65.

31. Aguda B, Friedman A. Models of cellular regulation. New York: Oxford University Press; 2008.

32. Alon U. An introduction to systems biology: design principles of biological circuits. London: CRC press; 2006.

33. Williams K, Savageau MA, Blumenthal RM. A bistable hysteretic switch in an activator–repressor regulated restriction–modification system. Nucleic Acids Res. 2013;41(12);6045–57.

34. Djordjevic M. Modeling bacterial immune systems: strategies for expression of toxic - but useful - molecules. Biosystems. 2013;112(2):139–44.

35. O'Donnell SM, Janssen GR. The initiation codon affects ribosome binding and translational efficiency in Escherichia coli of cI mRNA with or without the 5′ untranslated leader. J Bacteriol. 2001;183(4):1277–83.

36. Shell SS, Wang J, Lapierre P, Mir M, Chase MR, Pyle MM, Gawande R, Ahmad R, Sarracino DA, Ioerger TR. Leaderless transcripts and small proteins are common features of the mycobacterial translational landscape. PLoS Genet. 2015;11(11):e1005641.

37. Beletskaya IV, Zakharova MV, Shlyapnikov MG, Semenova LM, Solonin AS. DNA methylation at the CfrBI site is involved in expression control in the CfrBI restriction–modification system. Nucleic Acids Res. 2000;28(19):3817–22.

38. Shea MA, Ackers GK. The OR control system of bacteriophage lambda. A physical-chemical model for gene regulation. J Mol Biol. 1985;181(2): 211–30.

39. Enikeeva FN, Severinov KV, Gelfand MS. Restriction–modification systems and bacteriophage invasion: Who wins? J Theor Biol. 2010;266(4):550–9.

40. Streeter SD, Papapanagiotou I, McGeehan JE, Kneale GG. DNA footprinting and biophysical characterization of the controller protein C.AhdI suggests the basis of a genetic switch. Nucleic Acids Res. 2004;32(21):6445–53.

41. Bogdanova E, Zakharova M, Streeter S, Taylor J, Heyduk T, Kneale G, Severinov K. Transcription regulation of restriction-modification system Esp1396I. Nucleic Acids Res. 2009;37(10):3354–66.

42. Vesper O, Amitai S, Belitsky M, Byrgazov K, Kaberdina AC, Engelberg-Kulka H, Moll I. Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in Escherichia coli. Cell. 2011;147(1):147–57.

43. Mruk I, Liu Y, Ge L, Kobayashi I. Antisense RNA associated with biological regulation of a restriction–modification system. Nucleic Acids Res. 2011;39(13): 5622–32.

44. Nagornykh M, Zakharova M, Protsenko A, Bogdanova E, Solonin AS, Severinov K. Regulation of gene expression in restriction-modification system Eco29kI. Nucleic Acids Res. 2011;39(11):4653–63.

45. Barrangou R, Marraffini LA. CRISPR-Cas systems: prokaryotes upgrade to adaptive immunity. Mol Cell. 2014;54(2):234–44.

46. Al-Attar S, Westra ER, van der Oost J, Brouns SJ. Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes. Biol Chem. 2011; 392(4):277–89.

47. Rezulak M, Borsuk I, Mruk I: Natural C-independent expression of restriction endonuclease in a C protein-associated restriction-modification system. Nucleic Acids Res. 2016;44(6):2646–60.

48. Lubys A, Jurenaite S, Janulaitis A. Structural organization and regulation of the plasmid-borne type II restriction-modification system Kpn2I from Klebsiella pneumoniae RFL2. Nucleic Acids Res. 1999;27(21):4228–34.

49. Knowle D, Lintner RE, Touma YM, Blumenthal RM. Nature of the promoter activated by C. PvuII, an unusual regulatory protein conserved among restriction-modification systems. J Bacteriol. 2005;187(2):488–97.

Features of CRISPR-Cas Regulation Key to Highly Efficient and Temporally-Specific crRNA Production

Andjela Rodic [1,2], Bojana Blagojevic [3], Magdalena Djordjevic [3], Konstantin Severinov [4,5] and Marko Djordjevic [1]*

[1] Faculty of Biology, Institute of Physiology and Biochemistry, University of Belgrade, Belgrade, Serbia, [2] Multidisciplinary PhD Program in Biophysics, University of Belgrade, Belgrade, Serbia, [3] Institute of Physics Belgrade, University of Belgrade, Belgrade, Serbia, [4] Waksman Institute of Microbiology, Rutgers University, Piscataway, NJ, United States, [5] Skolkovo Institute of Science and Technology, Skolkovo, Russia

Bacterial immune systems, such as CRISPR-Cas or restriction-modification (R-M) systems, affect bacterial pathogenicity and antibiotic resistance by modulating horizontal gene flow. A model system for CRISPR-Cas regulation, the Type I-E system from *Escherichia coli*, is silent under standard laboratory conditions and experimentally observing the dynamics of CRISPR-Cas activation is challenging. Two characteristic features of CRISPR-Cas regulation in *E. coli* are cooperative transcription repression of *cas* gene and CRISPR array promoters, and fast non-specific degradation of full length CRISPR transcripts (pre-crRNA). In this work, we use computational modeling to understand how these features affect the system expression dynamics. Signaling which leads to CRISPR-Cas activation is currently unknown, so to bypass this step, we here propose a conceptual setup for *cas* expression activation, where *cas* genes are put under transcription control typical for a restriction-modification (R-M) system and then introduced into a cell. Known transcription regulation of an R-M system is used as a proxy for currently unknown CRISPR-Cas transcription control, as both systems are characterized by high cooperativity, which is likely related to similar dynamical constraints of their function. We find that the two characteristic CRISPR-Cas control features are responsible for its temporally-specific dynamical response, so that the system makes a steep (switch-like) transition from OFF to ON state with a time-delay controlled by pre-crRNA degradation rate. We furthermore find that cooperative transcription regulation qualitatively leads to a cross-over to a regime where, at higher pre-crRNA processing rates, crRNA generation approaches the limit of an infinitely abrupt system induction. We propose that these dynamical properties are associated with rapid expression of CRISPR-Cas components and efficient protection of bacterial cells against foreign DNA. In terms of synthetic applications, the setup proposed here should allow highly efficient expression of small RNAs in a narrow time interval, with a specified time-delay with respect to the signal onset.

Keywords: CRISPR-Cas activation, pre-crRNA processing, CRISPR regulation, crRNA generation, biophysical modeling

INTRODUCTION

CRISPR-Cas are adaptive immune systems, which defend prokaryotic cells against foreign DNA, including viruses and plasmids. A CRISPR-Cas system consists of a CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) array and associated *cas* genes (Makarova et al., 2006; Barrangou et al., 2007; Brouns et al., 2008; Hille and Charpentier, 2016). CRISPR arrays consist of identical direct repeats (R) of about 30 bp in length, interspaced with spacers (S) of similar length and variable sequence. Spacer sequences are often complementary to fragments of viral or plasmid DNA. A match between a CRISPR spacer and invading phage (bacterial virus) sequence provides immunity to infection (Barrangou et al., 2007; Hille and Charpentier, 2016). The entire CRISPR locus is initially transcribed as a long transcript (called pre-crRNA) (Pougach et al., 2010; Pul et al., 2010), which is further processed by Cas proteins to small protective CRISPR RNAs (called crRNAs) (Brouns et al., 2008; Pougach et al., 2010; Djordjevic et al., 2012). crRNAs are responsible for recognition and, together with Cas proteins, inactivation of invading foreign genetic elements (Brouns et al., 2008; Al-Attar et al., 2011). Cas proteins also take part in CRISPR adaptation, which is a process in which new spacers from viral genomes are inserted in CRISPR array. **Figure 1** shows a schematic gene diagram for Type I-E CRISPR-Cas from *E. coli*, (Mojica and Diez-Villasenor, 2010; Patterson et al., 2017), which we consider in this paper. The *cas* genes and the CRISPR array are transcribed from separate promoters, which are located inside of the intergenic regions here denoted by IGLB and L (the leader sequence), respectively (see **Figure 1**; Pougach et al., 2010; Pul et al., 2010).

Promoters for *cas* operon and the CRISPR array are repressed in Type I-E CRISPR-Cas in *E. coli* (Pougach et al., 2010; Pul et al., 2010; Westra et al., 2010), which makes this system silent under standard conditions. Consequently, to generate crRNAs that can protect the bacterial cell, CRISPR-Cas has to be activated. Thus, to understand the system function it is crucial to understand the main features that control dynamics of CRISPR-Cas activation (Mojica and Diez-Villasenor, 2010; Richter et al., 2012; Patterson et al., 2017). However, approaching this problem experimentally is complicated due to the following:

i. It requires direct experimental observation of *in vivo* dynamics of molecular species (proteins or RNA) in a cell (see e.g., Morozova et al., 2015).
ii. The signaling which leads to system induction is currently unclear (Ratner et al., 2015; Patterson et al., 2017), e.g., even a viral infection, an obvious trigger, is not sufficient to activate the system.
iii. To understand the roles of the key system features in its response/dynamics these features would have to be perturbed, which may require extensive reengineering of the system.

A complementary approach is to use mathematical/biophysical modeling to assess how different features of CRISPR-Cas expression affect system dynamics. Moreover, *in silico* analysis allows one to study alternative system architectures, and/or to perturb the natural system (see e.g., Rodic et al., 2017), which in turn allows understanding the role of its key regulatory features.

Experimental research has led to a consistent picture of the main CRISPR-Cas regulatory features in closely related *E. coli* and *Salmonella enterica* (Pul et al., 2010; Westra et al., 2010; Medina-Aparicio et al., 2011). Under standard conditions, promoters for both CRISPR array and *cas* genes are repressed by global regulators (H-NS and LRP). Repression by these regulators is highly cooperative, as their binding is nucleated at certain position, and then extends along the DNA through cooperative interactions between repressor molecules (Bouffartigues et al., 2007). Additional regulators, such as CRP, may also be involved in the repression of *cas* operon (Yang et al., 2014). While the exact signaling mechanism remains unclear, this repression must be relieved upon appropriate external signal (e.g., envelope stress that may signal bacteriophage invasion), through the action of transcription activators (LexA, LeuO, and BaeR-S are likely involved) (Richter et al., 2012; Patterson et al., 2017). In particular, for Type I-E CRISPR-Cas in *E. coli*, it was shown that cooperative repression by H-NS can be relieved by elevated amount of LeuO (Pul et al., 2010; Westra et al., 2010). Thus, highly cooperative repression, which is abolished by transcription activators, emerges as a major feature of CRISPR-Cas transcription control in *E. coli* and its relatives.

Another crucial mechanism in CRISPR-Cas expression is pre-crRNA transcript processing (Brouns et al., 2008; Pougach et al., 2010). Experiments in *E. coli*, reported that overexpression of
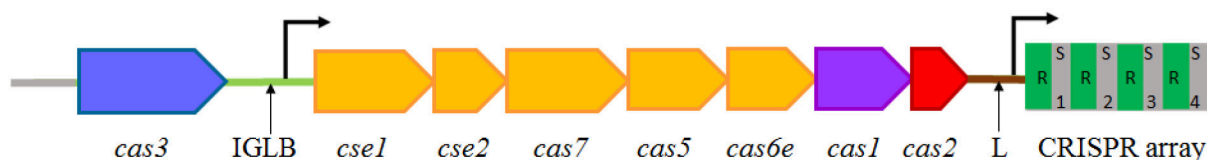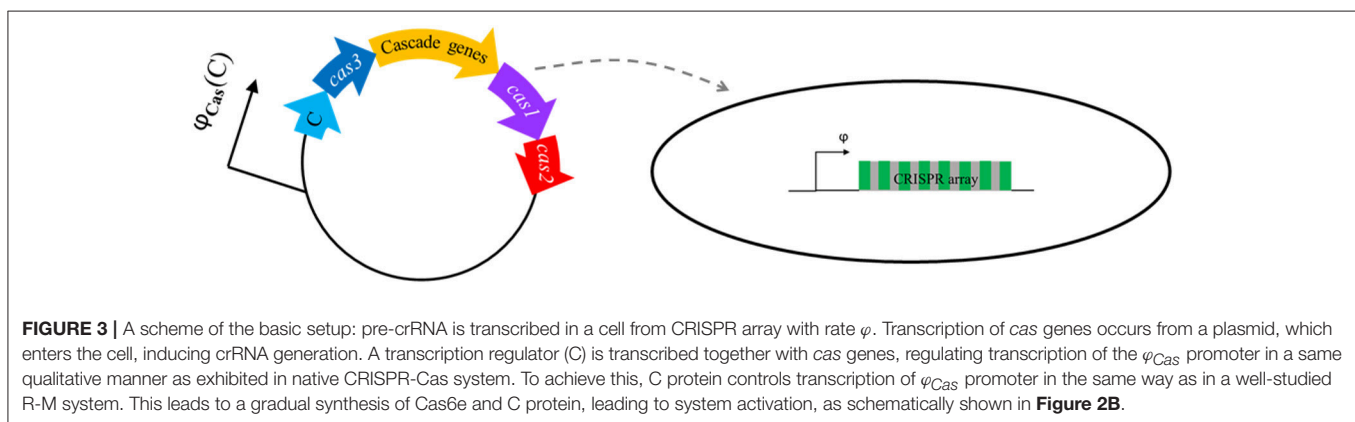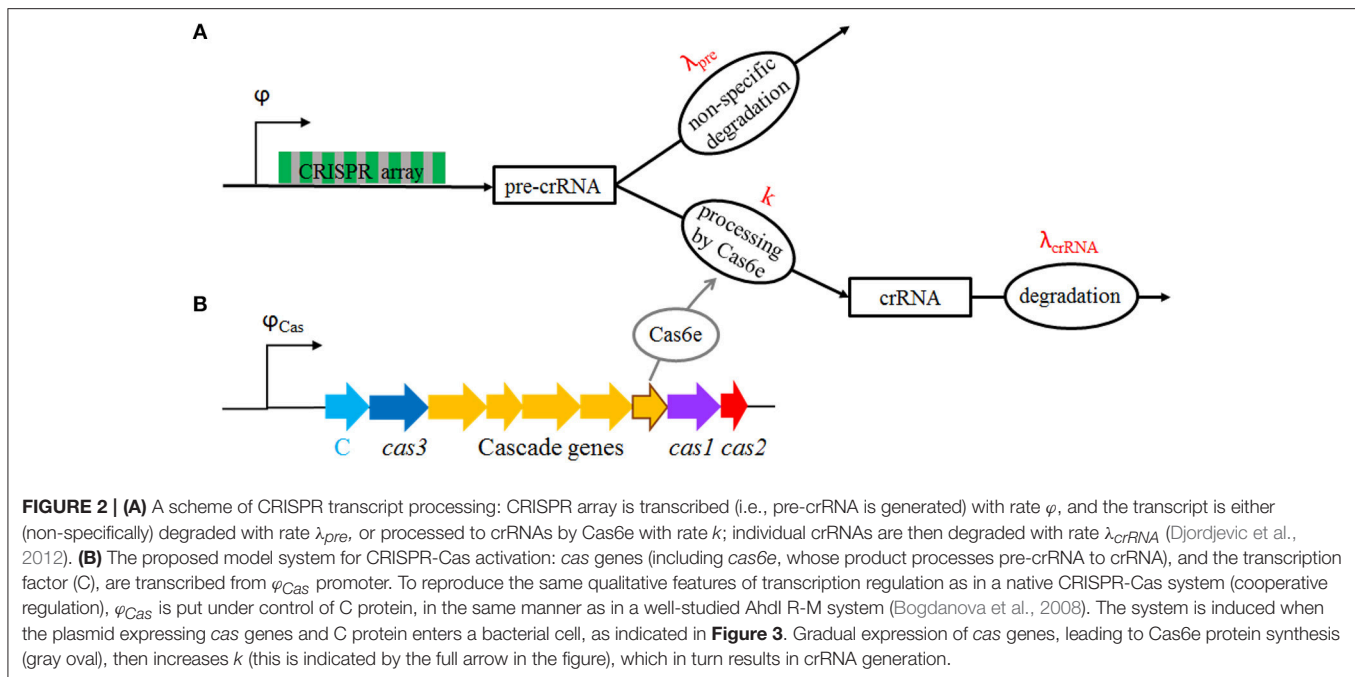

FIGURE 1 | A scheme of a Type I-E CRISPR-Cas system from *E coli* (Al-Attar et al., 2011; Makarova et al., 2011, 2015). The *cas* genes and the CRISPR array are indicated. R and S within the CRISPR array correspond, respectively, to repeats and spacers; note that the spacer sequences differ from each other, and are labeled by consecutive numbers (1, 2, 3...). IGLB and L correspond to intergenic regions where promoters for the Cascade complex genes (*cse1,2, cas7,5,6e*) and the Cas1,2 adaptation proteins (IGLB) and the CRISPR array (L) are located. The two promoters within IGLB and L are indicated by arrows. One of the Cas proteins (Cas6e) is responsible for processing pre-crRNA to crRNA. The effector Cascade complex is composed of proteins encoded by genes marked with yellow color. It binds crRNA, which recognizes invading DNA. Once recognized, foreign DNA is destroyed by the product of *cas3* (Brouns et al., 2008).

FIGURE 2 | (A) A scheme of CRISPR transcript processing: CRISPR array is transcribed (i.e., pre-crRNA is generated) with rate φ, and the transcript is either (non-specifically) degraded with rate λ_{pre}, or processed to crRNAs by Cas6e with rate k; individual crRNAs are then degraded with rate λ_{crRNA} (Djordjevic et al., 2012). **(B)** The proposed model system for CRISPR-Cas activation: *cas* genes (including *cas6e*, whose product processes pre-crRNA to crRNA), and the transcription factor (C), are transcribed from φ_{Cas} promoter. To reproduce the same qualitative features of transcription regulation as in a native CRISPR-Cas system (cooperative regulation), φ_{Cas} is put under control of C protein, in the same manner as in a well-studied AhdI R-M system (Bogdanova et al., 2008). The system is induced when the plasmid expressing *cas* genes and C protein enters a bacterial cell, as indicated in **Figure 3**. Gradual expression of *cas* genes, leading to Cas6e protein synthesis (gray oval), then increases k (this is indicated by the full arrow in the figure), which in turn results in crRNA generation.

FIGURE 3 | A scheme of the basic setup: pre-crRNA is transcribed in a cell from CRISPR array with rate φ. Transcription of *cas* genes occurs from a plasmid, which enters the cell, inducing crRNA generation. A transcription regulator (C) is transcribed together with *cas* genes, regulating transcription of the φ_{Cas} promoter in a same qualitative manner as exhibited in native CRISPR-Cas system. To achieve this, C protein controls transcription of φ_{Cas} promoter in the same way as in a well-studied R-M system. This leads to a gradual synthesis of Cas6e and C protein, leading to system activation, as schematically shown in **Figure 2B**.

Cas6e (which is responsible for pre-crRNA processing) generates highly abundant crRNAs from pre-crRNA which is present at low abundance (Pougach et al., 2010). We previously showed that a simple quantitative model—whose relevant kinetic scheme is shown in **Figure 2A**—explains this observation (Djordjevic et al., 2012), so that a small decrease in pre-crRNA abundance leads to a much larger (around two orders of magnitude) increase in crRNA abundance. Interestingly, the main mechanism responsible for this strong amplification is fast non-specific degradation of pre-crRNA (see **Figure 2**) by unidentified nuclease(s). In particular, when cas genes expression increases, processing of pre-crRNA by Cas6e is favored and diverts the entire pre-crRNA molecule away from the path of non-specific degradation. Therefore, the fast non-specific degradation of pre-crRNA should be considered as a second major regulatory feature of CRISPR-Cas expression.

The modeling described in Djordjevic et al. (2012) took into account only the transcript processing step, i.e., it was assumed that there is an infinitely abrupt (stepwise) increase of pre-crRNA

to crRNA processing rate, and pre-crRNA generation rate. This is, however, a clear idealization of the induction mechanism, as transcription regulation of *cas* genes and CRISPR array promoters is neglected. That is, in reality, pre-crRNA processing rate can be increased only gradually, as it takes time to synthesize the needed Cas proteins. The rate of Cas proteins synthesis is in turn directly related to the transcription control of the *cas* gene promoter in the IGLB region (see **Figure 1**). Similarly, the rate by which pre-crRNA is synthesized is determined by the transcription control of the CRISPR array promoter (L region).

Consequently, a more realistic model of CRISPR-Cas expression dynamics has to take into account both the regulation of CRISPR array and Cas protein synthesis, and CRISPR transcript processing. However, a major obstacle in achieving such model is that signaling which leads to the system induction, and detailed mechanism of CRISPR-Cas transcription regulation, is still unclear. We here propose a model system for CRISPR-Cas induction by assuming that activation of crRNA production

is put under transcriptional control exhibited in a restriction-modification (R-M) immune system (Pingoud et al., 2014). As argued below, such model system would have qualitative features of transcription regulation expected for a CRISPR-Cas, and will keep the same transcript processing mechanism as that described for native system. On the other hand, this model system allows bypassing the currently unknown signaling that leads to CRISPR-Cas activation, and can be readily analyzed *in silico*, since transcription regulation of a well-studied R-M system (AhdI, see Bogdanova et al., 2008)—for which we previously showed that it can be reliably modeled (see below)—is used as a proxy for transcription regulation of CRISPR-Cas system.

Through this approach, we expect to:

i. Obtain quantitatively more realistic model of CRISPR-Cas induction dynamics, in which the transcription regulation, i.e., the gradual synthesis of relevant enzymes and transcription regulators is explicitly taken into account.
ii. Qualitatively understand the main features of CRISPR-Cas induction, in particular the roles of cooperative transcription regulation, and of fast non-specific degradation of pre-crRNA.
iii. Propose an experimental setup for CRISPR-Cas induction that mimics the main qualitative features of the native system.

The setup of the model will be explicitly considered in the next subsection.

RESULTS

In silico Experiment Setup
The Model System
We start from a CRISPR transcript processing scheme, which is shown in **Figure 2**. According to this scheme, pre-crRNA is generated with rate φ, and subsequently either non-specifically degraded (due to activity of an unspecified nuclease) with rate λ_{pre}, or is processed by Cas6e to crRNAs with rate k. crRNAs are subsequently degraded with rate λ_{crRNA}. All the parameters in the scheme are experimentally determined in (Djordjevic et al., 2012) (for Type I-E CRISPR-Cas in *E. coli*) and explicitly stated in Methods. In particular, the main feature of the transcript processing is a large (non-specific) pre-crRNA degradation rate (with $\lambda_{pre} \sim 1$ 1/min), which is much larger than crRNA degradation rate (with $\lambda_{crRNA} \sim 1/100$ 1/min). In the experiments, crRNA production is artificially activated, by overexpressing Cas6e from a plasmid, which increases pre-crRNA processing rate (k) for between one and two orders of magnitude (between $10\lambda_{pre}$ and $100\lambda_{pre}$). While the repression of the *cas* promoter in IGLB region (see **Figure 1**) is very strong, with very small amount of Cas6e synthesized when the system is uninduced, the repression of the CRISPR array promoter is significantly weaker, with rather strong basal rate of pre-crRNA generation ($\varphi \sim 10$ 1/min) (Pougach et al., 2010; Pul et al., 2010; Westra et al., 2010; Djordjevic et al., 2012).

As indicated in the Introduction, we previously modeled the transcript processing mechanism (Djordjevic et al., 2012), where we took that k is increased abruptly, i.e., as a step function at $t = 0$. This neglects the transcription regulation of *cas*

and CRISPR array promoters. Such abrupt increase of k will provide a baseline for our predictions, which will now take into account that Cas6e (the enzyme which processes pre-crRNA to crRNA) is synthesized gradually. While in the experiments crRNA generation is activated by overexpressing Cas6e from a plasmid (see e.g., Pougach et al., 2010), it is likely that in the native system the expression of CRISPR array is activated as well (Pul et al., 2010). Consequently, we will also take into account a gradual synthesis of the regulator [in our case, a C-protein (Tao et al., 1991; Bogdanova et al., 2008)], which can activate CRISPR array transcription by increasing the basal rate φ to a higher value.

To include transcription regulation of the *cas* promoter, i.e., the gradual synthesis of Cas6e and C transcriptional regulator, we here propose the model system whose setup is schematically shown in **Figures 2**, **3**. This setup includes a CRISPR array which is expressed from a promoter with basal transcription activity φ (**Figure 3**). The second component is a vector (plasmid, virus) which expresses *cas* genes and the control protein C that are jointly transcribed from a promoter with transcription activity φ_{Cas}. While Cas3 is not directly relevant for the problem considered here (dynamics of crRNA generation), as it does not take part in crRNA biogenesis, it is necessary for CRISPR interference (Hille and Charpentier, 2016). We therefore include it in the setup to allow expression of all *cas* genes, i.e. to have a fully functional CRISPR-Cas system.

As detailed below, φ_{Cas} is regulated by C. To mimic the qualitative features of transcription regulation in native CRISPR-Cas system, we employ the transcription regulation found in some R-M systems, as explained in the next subsection. The system is activated when the vector enters a bacterial cell lacking its own *cas* genes, which leads to a gradual synthesis of Cas proteins (including Cas6e), therefore increasing the processing rate k, which in turn leads to crRNA generation (see **Figure 2B**—the full arrow) by pre-crRNA processing. Gradual increase of pre-crRNA generation rate can be also considered through this model, through activation of CRISPR array promoter by gradually synthesized C.

Note that the setup above, where *cas* genes are introduced in a cell on a vector, allows bypassing the unknown signaling step in CRISPR-Cas induction. That is, the vector entering the cell marks the start of the system activation (setting zero time in the dynamics simulations), and mimics the signaling which starts synthesis of the transcription activator. Therefore, the key regulatory features which characterize the downstream steps (CRISPR array transcription and transcript processing) can be studied both *in silico* (which will be done here), and also potentially experimentally. In terms of experimental implementation, introducing *cas* genes in a cell on a virus also allows synchronizing the cell population, which is an approach previously implemented to visualize R-M protein kinetics (Mruk and Blumenthal, 2008).

Putting CRISPR-Cas under Transcription Control of an R-M System
As discussed above, *cas* promoter will be put under transcription control exhibited by R-M systems. Below, the main elements

necessary for modeling the system transcription regulation are introduced.

R-M systems are often mobile, and can spread from one bacterial host to the other (Mruk and Kobayashi, 2013). When a plasmid carrying R-M system genes enters a naive bacterial host, the host genome is initially unmethylated, and can consequently be cut by the restriction enzyme. It is, therefore, evident that expression of the restriction enzyme and methyltransferase must be tightly regulated in order to ensure that bacterial genome is protected by the methyltransferase ("antidote"), before it is cut by the restriction enzyme. This tight regulation is often achieved through a dedicated control (C) proteins (Tao et al., 1991; Vijesurier et al., 2000).

We here concentrate on the AhdI R-M system, whose transcription control by C protein has been well-studied (Bogdanova et al., 2008). The activation of AhdI by C protein is reminiscent of CRISPR-Cas activation, as strong cooperative interactions are involved in both cases. In particular, C proteins bound at promoter-proximal and promoter-distal operators interact with high binding cooperativity, so that configuration in which only one operator is occupied cannot be observed in the absence of RNA polymerase (RNAP). At lower C protein concentrations, RNAP can outcompete C protein bound at promoter-proximal operator, leading to transcriptionally active configuration (Bogdanova et al., 2009). Moreover, another feature exhibited in AhdI transcription control, i.e., autoregulation by C protein, is also likely found in CRISPR-Cas transcription regulation. That is, LeuO that activates CRISPR-Cas expression (Westra et al., 2010) also regulates its own transcription. In particular, similarly to transcription regulation of cas genes, leuO is repressed by H-NS, while this repression is abolished by LeuO (Chen et al., 2001). At high concentrations, C protein is bound at both promoter-proximal and promoter-distal position, leading to the promoter repression—see **Figure 5** in (Bogdanova et al., 2009) and the scheme of the transcription configurations shown in **Figure 5** (framed in the figure). Negative autoregulation is also exhibited by LeuO, as it inhibits transcription activation of its gene by BglJ-RcsB (Stratmann et al., 2012). Therefore, putting cas genes under transcription control found in AhdI mimics the main qualitative features of CRISPR-Cas transcription regulation, namely, gradual synthesis of Cas proteins, cooperativity in transcription regulation, and putative autoregulation.

Another advantage of this setup is that we previously showed that biophysical modeling can be used to:(i) explain *in vitro* measurements of the *wild type* and mutant R-M system transcription control (Bogdanova et al., 2008), (ii) explain *in vivo* measurements of the system dynamics (Morozova et al., 2015), (iii) effectively perturb the main R-M system features and relate these perturbations with the system dynamics (Rodic et al., 2017). Consequently, transcription control of a well-studied AhdI R-M system, whose transcription regulation can be reliably modeled (Bogdanova et al., 2008), will serve as a proxy for the transcription control of a much less understood CRISPR-Cas system.

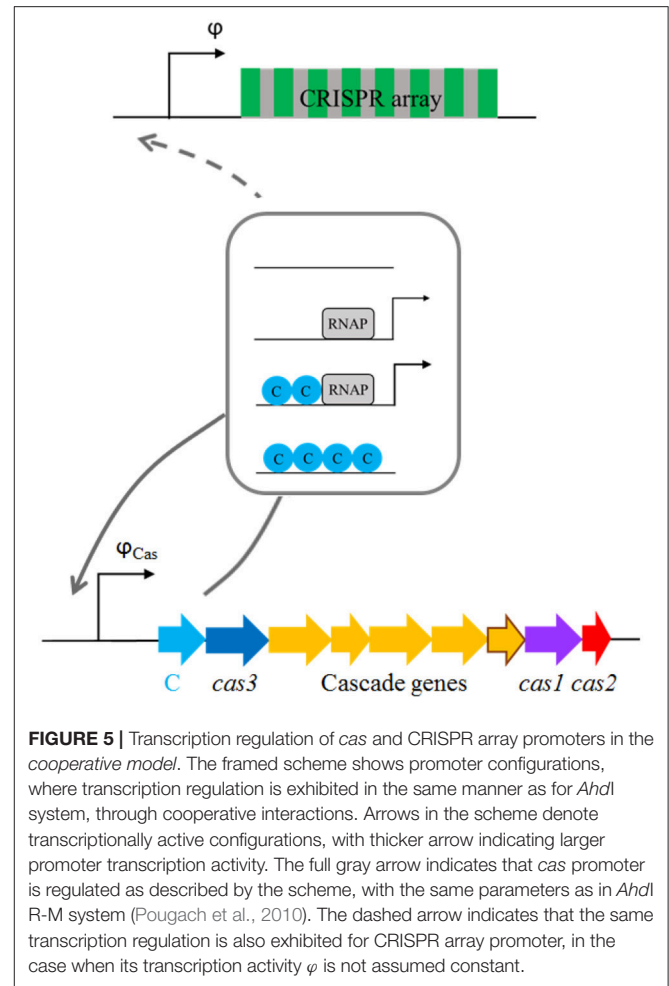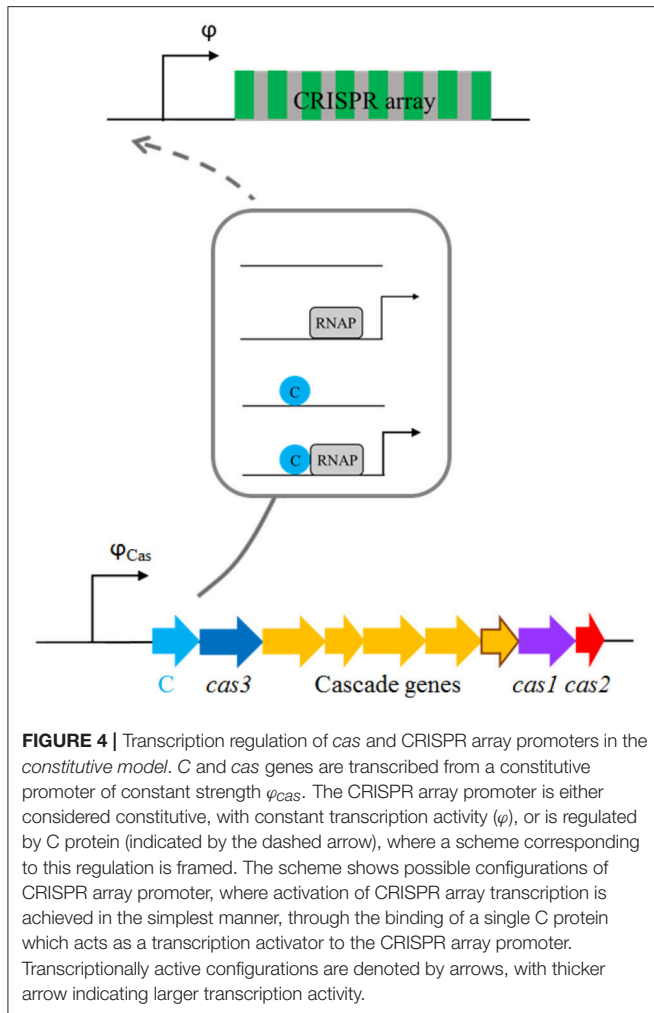## *In silico* Analysis of the Main System Features

The baseline for our predictions will be provided by a model in which the increase of pre-crRNA to crRNA processing rate k is infinitely abrupt—we will call this the baseline model. Comparing the baseline model with predictions that take into account the system transcription regulation (as schematically shown in **Figures 2**, **3**), allows analyzing how gradual synthesis of Cas6e affects kinetics of crRNA generation.

While in the native CRISPR-Cas both cas genes and CRISPR array promoters are repressed by global regulators, the repression of cas genes was found to be much stronger (Pul et al., 2010; Westra et al., 2010)—consequently, when the system is (experimentally) artificially induced, this is commonly done by expressing only cas genes (Pougach et al., 2010; Semenova et al., 2016; Musharova et al., 2017). However, in the native system, it is likely that expression of both CRISPR array and cas genes is activated when the appropriate induction signal(s) is received (Pul et al., 2010). We will therefore investigate the system dynamics when only cas genes are activated (i.e., only pre-crRNA processing rate is gradually increased), and when cas genes and CRISPR array promoter transcription are jointly (and gradually) increased. Consequently, in both of the models introduced below (constitutive and cooperative), we will consider two options. First, when only transcription of cas genes is activated, while transcription activity of CRISPR array remains constant. Second, we will consider the case when the transcription activity of CRISPR array is increased as well.

We further introduce two models of cas gene and CRISPR array transcription regulation:

i *The constitutive model* (**Figure 4**). In this model cas genes are expressed from a constitutive promoter, so that they are transcribed with the constant rate once the plasmid enters a cell. In the case when we consider that the system is activated by only increasing pre-crRNA processing rate, the transcription activity φ is kept constant. When CRISPR array transcription rate is increased as well, increasing φ is exhibited in the simplest manner, by binding of a single C protein activator. Note that, in accordance with its name, no cooperativity is exhibited for transcription regulation described by this model.

ii *The cooperative model* (**Figure 5**). In this model, C protein regulates the transcription of cas genes, and its own transcription, in the same manner as in AhdI R-M system. As noted above, such transcription regulation is characterized by strong cooperative interactions. CRISPR array transcription rate is either kept constant, or in the case when it is increased, we take that it is exhibited in the same way as for cas promoter transcription (the dashed arrow in **Figure 5**).

Studying of the two models allows one to assess how the cooperative transcription regulation (which also characterizes the native CRISPR-Cas system) compares to the activation in which no cooperativity is exhibited, and therefore allows us to assess the role of this key system feature. Also, considering the two models when φ is first kept constant, and then increased together with k, allows assessing significance of CRISPR array transcription

FIGURE 4 | Transcription regulation of *cas* and CRISPR array promoters in the *constitutive model*. *C* and *cas* genes are transcribed from a constitutive promoter of constant strength φ_{cas}. The CRISPR array promoter is either considered constitutive, with constant transcription activity (φ), or is regulated by C protein (indicated by the dashed arrow), where a scheme corresponding to this regulation is framed. The scheme shows possible configurations of CRISPR array promoter, where activation of CRISPR array transcription is achieved in the simplest manner, through the binding of a single C protein which acts as a transcription activator to the CRISPR array promoter. Transcriptionally active configurations are denoted by arrows, with thicker arrow indicating larger transcription activity.

FIGURE 5 | Transcription regulation of *cas* and CRISPR array promoters in the *cooperative model*. The framed scheme shows promoter configurations, where transcription regulation is exhibited in the same manner as for *Ahd*I system, through cooperative interactions. Arrows in the scheme denote transcriptionally active configurations, with thicker arrow indicating larger promoter transcription activity. The full gray arrow indicates that *cas* promoter is regulated as described by the scheme, with the same parameters as in *Ahd*I R-M system (Pougach et al., 2010). The dashed arrow indicates that the same transcription regulation is also exhibited for CRISPR array promoter, in the case when its transcription activity φ is not assumed constant.

control. To allow a direct comparison of models dynamics, the overall strength of φ_{Cas} is adjusted so that the same value of maximal pre-crRNA processing rate is achieved. Similarly, when the transcription rate of CRISPR array is increased, the interaction parameters are adjusted so that the same equilibrium increase of φ is achieved in both models (see Methods).

Modeling Results
Kinetics of Pre-crRNA and crRNA Production

We first consider the situation in which crRNA generation is activated by expressing Cas proteins, such that the processing rate k is gradually increased, while the CRISPR array transcription activity remains constant. In this case, we compare the system dynamics for: (i) *baseline model*, in which the processing rate k is increased as a step function, which corresponds to the limit of infinitely fast system induction, (ii) constitutive model (see **Figure 4**), and (iii) cooperative model (see **Figure 5**).

In constitutive and cooperative models, the gradual synthesis of Cas6e leads to gradual change of transcript processing rate k (k^* is a processing constant):

$$k(t) = [Cas6e](t) \cdot k^* \qquad (1)$$

Figure 6 illustrates how the processing rate (k) changes with time, when the baseline, constitutive, and cooperative models of *cas* gene expression are assumed. For the constitutive model (the dash-dotted curve), the processing rate uniformly increases and reaches an equilibrium value, for all values of k_{eq} considered in three panels of **Figure 6**. On the other hand, for cooperative model (the dashed curve) and at higher values of k_{eq} (**Figures 6B,C**), we see a rapid increase of k at initial times, followed by a fast return to the equilibrium value due to repression at higher C protein concentrations.

In **Figure 7**, we address how different k dynamics (shown in **Figure 6**), affects pre-crRNA and crRNA generation. Specifically, φ is held constant at its initial value (10 1/min), while k changes according to the baseline, constitutive, or cooperative models until reaching the same equilibrium value of $10\lambda_{pre}$, $100\lambda_{pre}$, and $1,000\lambda_{pre}$ (left, central, and right columns of **Figure 7**, respectively). The model of abrupt Cas6e expression serves as a baseline for assessing the dynamics in the other two models (constitutive and cooperative), in which Cas6e is realistically (gradually) expressed.

In **Figures 7A–D**, we see that cooperative model leads to the steepest transition from ON to OFF state (in the case
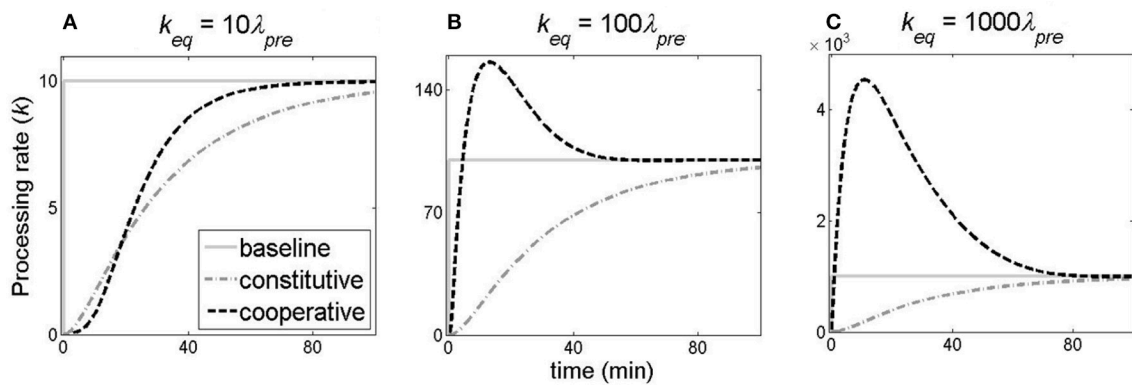
FIGURE 6 | Comparing the dynamics of the pre-crRNA processing rate change. The change of the processing rate k with time is shown for: the baseline model (solid gray curve), the constitutive model (dash-dotted gray curve) and the cooperative model (dashed black curve). **(A–C)** correspond to different k_{eq} values ($k_{eq} = 10\lambda_{pre}$, $100\lambda_{pre}$, $1,000\lambda_{pre}$, respectively). CRISPR transcription activity is constant (10 1/min).
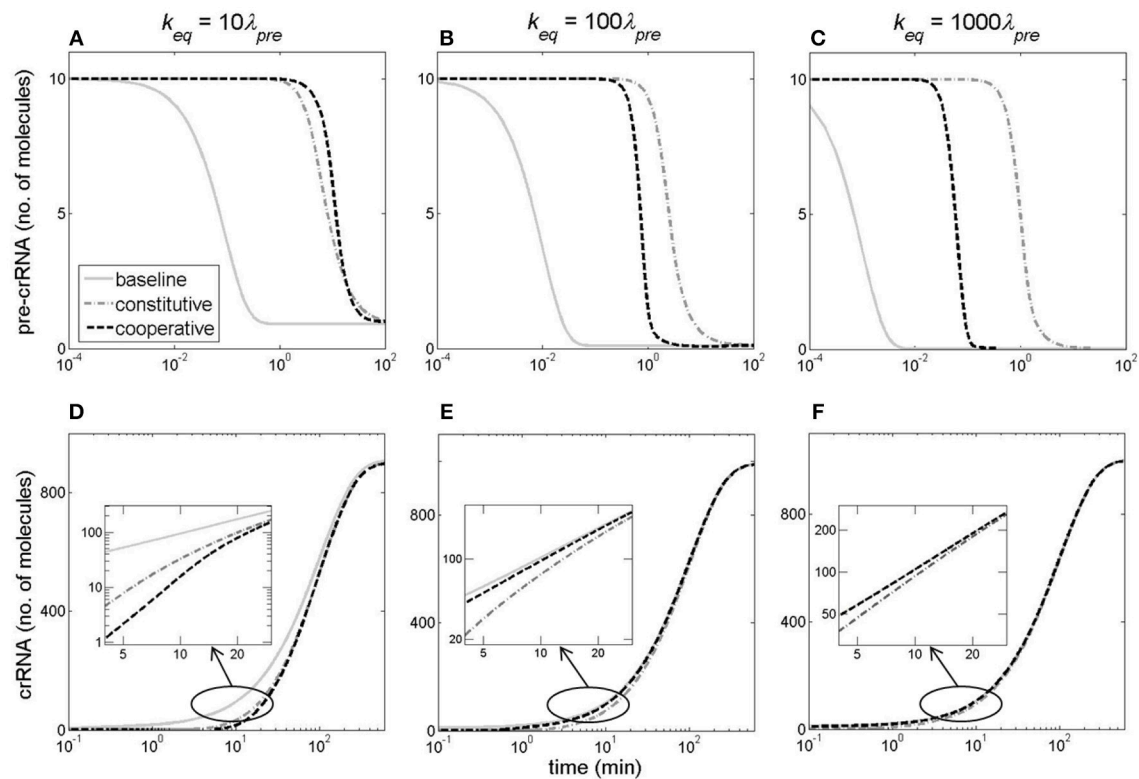

FIGURE 7 | Kinetics of pre-crRNA and crRNA generation. The columns correspond to k_{eq} values of $10\lambda_{pre}$ **(A,D)**, $100\lambda_{pre}$ **(B,E)**, and $1,000\lambda_{pre}$ **(C,F)**, which are reached through the baseline model (the gray solid curve), the constitutive model (the gray dash-dotted curve) or cooperative model (the black dashed curve). The upper **(A–C)** and the lower **(D–F)** rows correspond, respectively, to pre-crRNA and crRNA kinetics. CRISPR array promoter transcription activity is kept constant at 10 1/min.

of pre-crRNA), and from OFF to ON state (in the case of crRNA). Furthermore, we can distinguish between two different regimes in **Figure 7**. At lower k_{eq} (left column in **Figure 7**), there is a noticeably slower accumulation of crRNA at early times in both cooperative and constitutive models compared to the

baseline model of infinitely abrupt processing rate (k) increase (**Figure 7D**). On the other hand, at higher k_{eq} ($k_{eq} \geq 100$ 1/min, the central and right columns in **Figure 7**), the dynamics of crRNA accumulation for cooperative model becomes faster compared to constitutive model dynamics at early times, and

approaches the limit of infinitely abrupt k increase (see the inserts in **Figures 7E,F**). The faster kinetics of crRNA increase in cooperative model is due to the fast increase of k at early times in this model (**Figures 6B,C**).

Effects of *cas* Genes Regulation

From **Figure 7**, we observe that transcripts reach their steady-state levels quite late, i.e., >100 min post-induction. Such late time is, however, not relevant for cell response to phage infection, since infected *E. coli* lyse ~20 min post-infection, while shut-off of essential cell functions happens earlier (Kruger and Schroeder, 1981). Therefore, in **Figure 8** we estimate pre-crRNA and crRNA levels for all three models at 20 min post-induction, as the maximal value of pre-crRNA processing rate k_{eq} is changed from very low to high values (>100λ_{pre}, characteristic for artificial Cas6e induction), while keeping the level of CRISPR array transcription constant ($\varphi = 10$ 1/min).

The following features emerge from **Figure 8**:

i. A switch-like system behavior for both pre-crRNA and crRNA curves in the cooperative model, while the constitutive and baseline models yield much more gradual responses to changes in k_{eq}. For crRNA, the cooperative model leads to a rapid transition from the OFF state (with essentially no crRNA generated at 20 min), to the ON state (with high abundance of crRNA), and reciprocal situation for pre-crRNA. Consequently, for small amounts of synthesized Cas6e (i.e., small k_{eq} values), which can be caused by leaks in *cas* promoter activity, the system remains in OFF state. On the other hand, once the system is activated when the processing rate (directly related to the amount of Cas6e available) reaches a certain threshold ($k_{eq} \gtrsim 50$), a large amount of crRNA is generated at early times, which should allow protection from foreign DNA invasion. The significance of this behavior is considered in Discussion.

ii. An interesting cross-over behavior in the cooperative model, where at low k_{eq} values crRNA amounts are low, while at high k_{eq} values the synthesized crRNA amounts become larger than in the constitutive model, and approach the baseline model curve. Therefore, at high k-values (~100 1/min), which are encountered in experiments, (Pougach et al., 2010; Djordjevic et al., 2012) the model of cooperative *cas* gene expression leads to accumulation of protective crRNA amounts close to those achievable in the limit of infinitely abrupt k increase. Consequently, the high cooperativity in transcription regulation, characteristic for native CRISPR-Cas system regulation, leads to a highly efficient crRNA generation at high transcript processing rates.

iii. Sufficient crRNA levels are generated to protect host cell against bacteriophage infection, at early times post-induction, even at relatively low values of pre-crRNA processing rate. That is, k_{eq} somewhat larger than 11/min leads to ~10 crRNAs which already corresponds to the amount that negatively affects phage development (Pougach et al., 2010); moreover, a small additional k_{eq} increase leads to a large increase in generated crRNAs in the cooperative model, due to the rapid transition from OFF to ON state.

iv. A saturation in generated crRNA amounts at early times post-induction. That is, for k_{eq}~100 1/min the amount of generated crRNAs at 20 min stops significantly increasing with further increase in k_{eq}. This saturation can be relieved (leading to increase in the amount of generated crRNA), if CRISPR array transcription activity is increased, which is further analyzed below.

Perturbing Pre-crRNA Degradation Rate

We next perturb the second key feature of CRISPR-Cas regulation—fast non-specific degradation of pre-crRNA. The consequence of pre-crRNA degradation rate λ_{pre} decrease at constant φ was next investigated for all three models. The
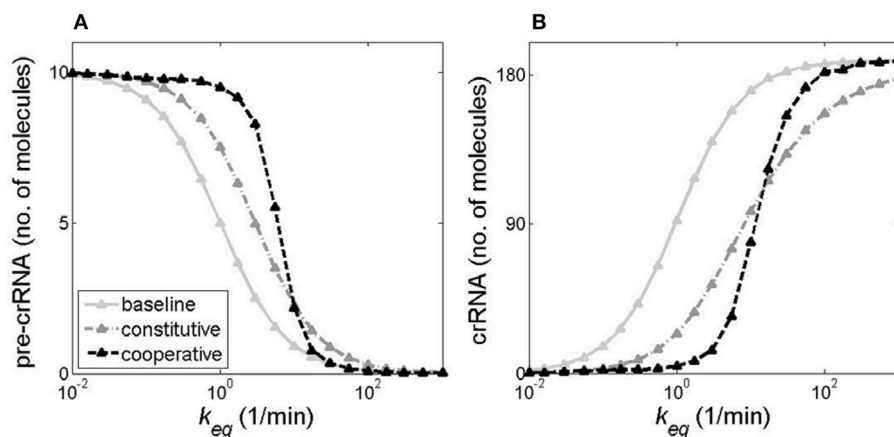

FIGURE 8 | Pre-crRNA and crRNA amounts early post-induction for different models of *cas* gene transcription regulation. The figure shows **(A)** pre-crRNA and **(B)** crRNA amounts 20 min post-induction (i.e., 20 min after introduction of the vector expressing *cas* genes), as a function of the maximal (equilibrium) value of the transcript processing rate k. CRISPR promoter transcription activity is kept constant ($\varphi = 10$ 1/min). The gray solid, the gray dash-dotted, and the black dashed curves correspond, respectively, to baseline, constitutive, and cooperative models of *cas* regulation.

decrease was followed at different k_{eq} values (i.e., at different levels of Cas6e activity), where φ is held constant.

The effects of λ_{pre} decrease are similar for all three models, so in **Figure 9** we show the results only for the cooperative model. For all k_{eq} values we see that abolishing the fast decay of pre-crRNA (decreasing λ_{pre}), significantly decreases the time delay of the onset of crRNA generation. This effect is most pronounced at high k_{eq} values (**Figure 9C**). Also, perturbing the degradation rate deforms crRNA dynamics curve with respect to the standard Hill (sigmoidal) shape that is exhibited at high λ_{pre} such as $\lambda_{pre} = 1/50$. Furthermore, analogously to **Figure 8**, in Figure S1 (Supplementary Material), we show how crRNA amount at 20 min after induction depends on pre-crRNA degradation rate λ_{pre}. One can clearly observe that as λ_{pre} decreases, the amount of generated crRNA early post-induction significantly increases, consistently with the decrease of the time delay of onset of crRNA generation observed in **Figure 9**.

Relieving crRNA Production Saturation by Increasing Pre-crRNA Generation

In addition to *cas* genes, CRISPR array promoter is also repressed (though more weakly) by global transcription regulators (Pul et al., 2010; Westra et al., 2010). Consequently, crRNA generation can be also augmented by increasing CRISPR array transcription activity. Therefore, we next assess how joint increase of k (achieved by activating *cas* gene transcription) and φ (achieved by increasing CRISPR array transcription) affects generated crRNA amount 20 min post-induction for all three regulatory models.

As can be seen from **Figure 10**, increasing φ robustly relieves crRNA saturation (see also discussion of **Figure 8**). Moreover, one can see that a relatively modest, factor of two increase of φ (from 10 1/min to 20 1/min) can abolish the need of a significant, order of magnitude, k increase to produce the same amount or crRNA. As above, we observe a switch-like behavior for the cooperative model (compare **Figure 10C** with **Figures 10A,B**),

with cooperative model curves exhibiting the steepest transition from OFF to ON state for all φ values.

Regulation of CRISPR Array Transcription Activity

We next consider how different models of regulation of CRISPR array transcription affect crRNA dynamics. For all three models, the transcription activity φ is increased by an order of magnitude (from $\varphi = 10$ 1/min to $\varphi = 100$ 1/min), for different k_{eq} values ($k_{eq} = \lambda_{pre}$, $10\lambda_{pre}$, and $100\lambda_{pre}$), see Figure S2 (Supplementary Material). We obtain that the cooperative model leads to a more controlled (attenuated) pre-crRNA dynamics, which is due to the presence of repressing mechanism at high C protein amounts (see Figure S3). For crRNA dynamics, we observe that the cooperative model exhibits the steepest transition from OFF to ON state. Moreover, this model leads to the largest delay in crRNA generation. Consequently, in addition to pre-crRNA degradation rate, the cooperative transcription regulation also contributes to the delay between the activating signal and the onset of crRNA generation.

We previously (**Figure 9**) perturbed pre-crRNA degradation rate while keeping the transcription rate φ constant. Finally, we now also decrease λ_{pre} under the conditions when both *cas* genes and CRISPR array transcription is activated according to all three models (see Figure S4). The results are qualitatively similar to **Figure 9** (where φ is constant), i.e., decreasing λ_{pre} diminishes the switch-like system response and/or decreases the time-delay in the onset of pre-crRNA generation.

DISCUSSION AND SUMMARY

One of the most prominent problems in understanding CRISPR-Cas function is assessing dynamics of the system activation, i.e., understanding the roles of the key features of CRISPR-Cas regulation. Addressing this problem is complicated by the fact that exact conditions for system activation remain unclear. In fact, for Type I-E CRISPR-Cas system in *E. coli*, even bacteriophage infection itself is not sufficient

FIGURE 9 | The effect of perturbing pre-crRNA degradation rate on the dynamics of crRNA generation. The pre-crRNA processing rate increases to its equilibrium value through the cooperative model, while φ is held constant (at 10 1/min). Different curves correspond to different λ_{pre} values: 1 (solid black), 1/10 (dashed black), 1/50 (solid gray), and 1/100 1/min (dashed gray). **(A–C)**, correspond to different k_{eq} values indicated at the top of each panel.

FIGURE 10 | Relieving saturation in generated crRNA amounts through joint k and φ increase. crRNA amount as a function of k_{eq} at 20 min post-induction, obtained for **(A)** baseline, **(B)** constitutive, and **(C)** cooperative models. Curves marked with diamonds, filled squares, circles, and filled triangles, correspond, respectively, to φ of 10, 20, 50, and 100 1/min.

to induce the system. We here proposed a synthetic setup which allows inducing CRISPR-Cas with qualitative features that correspond to native system regulation, while bypassing currently unclear conditions under which the system is activated. This setup involves putting *cas* genes and/or CRISPR array under transcription control found in a well-studied R-M system, which exhibits cooperative transcription regulation that is also characteristic of CRISPR-Cas regulation (Bouffartigues et al., 2007; Westra et al., 2010). A major advantage of the setup is that it can be readily experimentally implemented, e.g., by introducing *cas* genes and the regulator (C protein) in a cell on a virus. This would allow synchronizing the cell population, and experimentally observing the system dynamics, where such measurements could be directly compared with the predictions provided here. Another advantage is that major parameters in the setup have been inferred from experimental data, as both CRISPR transcript processing, and AhdI transcription regulation, have been experimentally well-studied (Bogdanova et al., 2008; Pougach et al., 2010; Djordjevic et al., 2012).

Consequently, this setup allows us to directly (*in silico*) address how the system regulation contributes to its dynamical response. In particular, previous experimental and computational work point to cooperative regulation of *cas* gene and CRISPR array transcription, and fast non-specific degradation of pre-crRNA, as two main system regulatory features (Pougach et al., 2010; Pul et al., 2010; Westra et al., 2010; Djordjevic et al., 2012). We therefore investigated two alternative regulatory architectures, one with constitutive, and the other with cooperative *cas* gene regulation. The dynamics corresponding to these two architectures was then compared with the baseline model, in which pre-crRNA processing rate is increased infinitely abruptly. We assessed the dynamics in the case when only *cas* genes are activated (i.e., only pre-crRNA processing rate is gradually increased), and when *cas* genes and CRISPR array promoter transcription is jointly increased. We focused on early system dynamics (within the first 20 min post-induction), as this period is most relevant for defending the cell against invading viruses. Finally, we also perturbed the high pre-crRNA non-specific

degradation rate, under different system conditions described above, and assessed what effect such perturbation has on system dynamics.

The main result of the analysis is that the system regulation leads to a clear switch-like behavior, characterized by an initial delay of crRNA synthesis, followed by a steep transition from OFF to ON state. Unexpectedly, it is not only the cooperative transcription regulation, but also fast non-specific pre-crRNA degradation, which leads to such dynamics. That is, decreasing the high pre-crRNA degradation rate effectively abolishes the delay in crRNA generation, and deforms the crRNA kinetics from the standard sigmoidal (Hill) shape (Hill, 2013) typical for switch-like system response (**Figure 9**). Interestingly, we also found that, when pre-crRNA processing rate and CRISPR array transcription rate are jointly (and gradually) increased, as likely exhibited in the native system, the system is more robust to perturbations in the degradation rate (Figure S4).

The cooperative transcription regulation leads to an interesting cross-over behavior in the early system dynamics. At low pre-crRNA processing rates, cooperative regulation leads to much smaller crRNA amounts at early times compared to constitutive expression. On the other hand, at higher processing rates, there is a large increase in synthesized crRNA amounts, which approach the limit of infinitely abrupt system induction. Interestingly, when the system is artificially activated by overexpressing *cas* genes, pre-crRNA processing rates correspond to the regime of the highly enhanced crRNA production (Djordjevic et al., 2012). While the parameters of the native system induction are unclear, it is tempting to hypothesize that they may also reach this cross-over, allowing the system to generate crRNAs with the rate close to the limit of infinitely fast induction at times when they are needed.

The rapid transition of the system from OFF to ON state is straightforward to interpret in terms of its function in immune response. When a potential signal indicating infection is received by the cell, CRISPR-Cas has a very short time to generate sufficient crRNA amounts to protect the cell, as bacteriophages are typically highly efficient in shutting-down essential cell functions. Thus, there is a question whether enough crRNA can

be generated in a model which accounts for gradual synthesis of proteins that process pre-crRNA and/or are responsible for gradual CRISPR array activation. We robustly obtained that enough crRNA can be generated at early times, even when the system is activated by only increasing the pre-crRNA processing rate. Moreover, a much smaller increase of the processing rate is needed to achieve certain crRNA amount, if CRISPR array transcription is activated as well. Therefore, these results may explain the relatively inefficient repression of CRISPR array promoter, since even a small increase of CRISPR array transcription rate efficiently increases generated crRNA amounts. In fact, the need to rapidly produce large amounts of crRNAs may be a major constraint on system dynamics.

In distinction to the rapid transition of the system from "OFF" to "ON" state, interpretation of the delay in crRNA generation, which comes as a model prediction, is less straightforward. One possibility is that such delay is related with primed adaptation in CRISPR-Cas, which relies on a pre-existing (priming) spacer that enables a biased uptake of new spacers—therefore serving to minimize infection by phage escape mutants that would otherwise evade the interference (Sternberg et al., 2016). In particular, it has been found that priming is facilitated by slow or delayed CRISPR interference, leading to a steady-state flux of substrates from which new spacers can be acquired (Kunne et al., 2016; Severinov et al., 2016; Musharova et al., 2017). Such delay in CRISPR interference can clearly be achieved by a delay in crRNA generation that is predicted in our work.

It has been proposed that Type I-E CRISPR-Cas in *E. coli* may have functions other than immunity. For example, it was found by bioinformatics analysis that the system is changing very slowly, in distinction to rapid diversification of CRISPR arrays in other species, indicating that the system is not taking an active role in defense against immediate viral threats (Touchon et al., 2011). In this respect, it may be useful to view the dynamical properties inferred here in a more general terms, namely of a capability of expressing a large number of molecules in a narrow time interval, with a specific time-delay with respect to reception of an external signal. It is clear that such highly efficient, and temporally specific response, may be highly desirable for multiple cellular functions. It would be very interesting to find out how functions of *E. coli* Type I-E CRISPR-Cas, yet to be discovered in the future, would fit within the dynamical properties inferred here.

METHODS

We start from a previously introduced model of CRISPR transcript processing by Cas proteins (Djordjevic et al., 2012). In this model (see **Figure 2A**), a short-living transcript [pre-crRNA] is synthesized with a promoter transcription activity φ, and further, either quickly degraded with a degradation rate λ_{pre}, or processed (cut) into shorter, long-living RNAs [crRNA] with a processing rate k. Processed transcripts are degraded with a rate λ_{crRNA}. In the equations below, we assume that the processing rate depends linearly on the substrate (pre-crRNA) amount, since the amount of pre-crRNA is small [<10 molecules per cell (Pougach et al., 2010)], so that the corresponding kinetic

equations are:

$$\frac{d[\text{pre} - \text{crRNA}]}{dt} = \varphi - (\lambda_{\text{pre}} + k) \cdot [\text{pre} - \text{crRNA}] \qquad (2)$$

$$\frac{d[\text{crRNA}]}{dt} = k \cdot [\text{pre} - \text{crRNA}] - \lambda_{\text{crRNA}} \cdot [\text{crRNA}]$$

$$(3)$$

The equations above are further solved deterministically, as both CRISPR array and *cas* genes are expressed from promoters with strong basal transcription. Furthermore, the small pre-crRNA amount is due to fast non-specific degradation, i.e., due to the transcript processing step. With respect to this, note that there is an access of enzyme (Cas6e) over substrate (pre-crRNA) (Djordjevic et al., 2012), so the equations describing the transcript processing are linear. Therefore, their deterministic solution accurately describes the mean of the stochastic simulations.

In the previous study (Djordjevic et al., 2012), we considered a model in which transcription regulation is neglected, so that k and φ increase in an idealized manner, i.e., infinitely abruptly. We now introduce models where the relevant enzymes and transcription regulators are synthesized in a realistic (i.e., gradual) manner. Specifically, k in Equation now explicitly depends on time, and is proportional to the enzyme (the processing protein, Cas6e) concentration, i.e., $k = [Cas6e] \cdot k^*$, where k^* is processing constant. We here consider that this processing rate k can change with time in the following ways:

1) Infinitely abruptly, from 0 to its equilibrium value, k_{eq} at $t = 0$, which we refer to as the baseline model.
2) Gradually, with [Cas6e](t), where Cas6e is expressed from a constitutive promoter (promoter with constant transcription activity), see **Figure 4**.
3) Also gradually with [Cas6e](t), where Cas6e is now expressed from an AhdI-like regulated promoter (see **Figure 5**).

As noted above, we either keep the CRISPR array transcription rate φ constant (which allows us investigating the dynamics in response to changing only pre-crRNA processing rate), or allow φ to change:

1) Infinitely abruptly (the baseline model), so that at $t = 0$ it increases from its starting value (10 1/min) to the equilibrium value.
2) Gradually, through the simplest activation mechanism, where a single C protein activates transcription from the CRISPR array promoter (the dashed arrow in **Figure 4**).
3) Also gradually with C(t), where the same transcription regulation as in AhdI RM system is exhibited (the dashed arrow in **Figure 5**).

In constructing Cas6e and CRISPR expression models, we refer to our existing model of AhdI restriction-modification (RM) system control (Bogdanova et al., 2008), which describes expression of the control protein (C) and the restriction endonuclease (R)—C and R are co-transcribed in AhdI RM system. We here use a thermodynamical model of CR operon transcription regulation, and a dynamical model of transcript and protein expression.

For $t = 0$ we take the moment when plasmid carrying C and *cas* genes enters the naïve host. Thus, all initial conditions are set to zero, except for $[pre\text{-}crRNA](t = 0) = \varphi/\lambda_{pre} = 10\,(1/\min)$ (Djordjevic et al., 2012), as extracted from the Equation in equilibrium. Note that while C and *cas* genes enter the cell on a plasmid, CRISPR array is expressed within the cell, with the transcription rate φ.

Constitutive Model of *cas* Gene and CRISPR Array Expression

We assume that C and *cas* genes are co-transcribed from a constitutive (unregulated) *cas* promoter (see above and **Figure 4**). C and *cas* transcript and protein concentrations change with time:

$$\frac{d[c - cas](t)}{dt} = \varphi_{Cas} - \lambda_{Cas} \cdot [c - cas](t) \tag{4}$$

$$\frac{dC(t)}{dt} = k_C \cdot [c - cas](t) - \lambda_C \cdot C(t) \tag{5}$$

$$\frac{d[Cas6e](t)}{dt} = k_{Cas6e} \cdot [c - cas](t) - \lambda_{Cas6e} \cdot [Cas6e](t). \tag{6}$$

Note that all the notation (including in the equation above), is introduced in **Table 1**. The first terms on the right-hand side represent transcript/protein synthesis by transcription/translation, while the second terms represent transcript/protein decay by degradation. The parameter values are as in AhdI RM system model (with Cas6e now replacing R in AhdI system), and are also provided in the table at the end of the methods. Since C and Cas6e protein degradation rates are taken to be the same, it follows:

$$[Cas6e](t) = \frac{k_{Cas6e}}{k_C} C(t), \tag{7}$$

So that the differential equation for Cas6e dynamics can be omitted. We set the value of φ_{Cas} to one (see the next subsection) so that the equilibrium processing rate is the same for the constitutive and the cooperative models (see e.g., **Figure 6**), which allows a direct comparison of the dynamics in these two models. Consequently, we set k^* so that $k_{eq} = [Cas6e]_{eq} \cdot k^* = 10\,(1/\min)$. Regarding CRISPR array transcription φ, we keep it constant, in the case when we consider the system activation by overexpression of *cas* genes. In the case when we also consider activation of CRISPR transcription, we introduce a simple model of CRISPR expression regulation (the dashed arrow in **Figure 4**), where CRISPR promoter, apart from being unoccupied, can be found in the following three configurations, which are represented by the reactions shown below: (i) RNAP alone bound to the promoter (8), (ii) a C monomer alone bound to its binding site (9), and (iii) RNAP recruited by a C monomer bound to its binding site, acting as a transcription activator —note that these configurations correspond to the second, third and fourth line in the framed part of **Figure 4**, respectively.

$$DNA + RNAP \underset{K_{1A}}{\xrightarrow{\hspace{1cm}}} RNAP - DNA \tag{8}$$

$$DNA + C \underset{K_{2A}}{\xrightarrow{\hspace{1cm}}} C - DNA \tag{9}$$

$$C - DNA + RNAP \underset{K_{3A}}{\xrightarrow{\hspace{1cm}}} C - DNA - RNAP \tag{10}$$
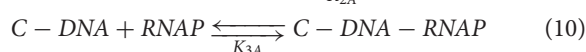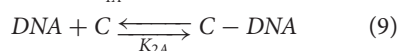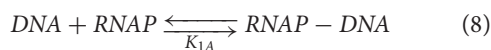
TABLE 1 | Notations used in model equations.

Variables	Description	
φ_{Cas}	Transcription activity of *cas* promoter	
φ	Transcription activity of CRISPR promoter	
[c-cas]	Concentration of *cas* operon transcript	
[pre-crRNA]	Concentration of unprocessed CRISPR array transcript	
[crRNA]	Concentration of processed CRISPR array transcript	
C	Concentration of control protein	
[Cas6e]	Concentration of processing protein	
KINETIC MODEL CONSTANTS		
k^*	CRISPR transcript processing constant	0.02
k_C	Translation constant for control protein	0.60
k_{Cas6e}	Translation constant for processing protein	3.00
λ_{Cas}	Rate of *cas* transcript decay	0.20
λ_{pre}	Rate of unprocessed CRISPR transcript decay	1.00
λ_{crRNA}	Rate of processed CRISPR transcript decay	0.01
λ_C	Rate of control protein decay	0.033
λ_{Cas6e}	Rate of Cas6e processing protein decay	0.033
TRANSCRIPTION REGULATION MODELS CONSTANTS		
α	Proportionality constants	1.663
γ		110
α'		110
a	Constants which absorb the relevant	1.60×10^{-1}
p	equilibrium dissociation constants and RNA	9.25×10^{-1}
q	polymerase concentration	1.41×10^{-5}
d		1.00×10^{-1}
e		Adjusted
f		2.00×10^{2}
a'		1.00×10^{-1}
p'		Adjusted
q'		2.50×10^{-5}
K_D		6.50×10^{2}

The equilibrium dissociation constants of the above reactions are given by:

$$K_{1A} = [DNA]\,[RNAP]\,/\,[RNAP - DNA] \tag{11}$$

$$K_{2A} = [DNA]\,[C]\,/\,[C - DNA] \tag{12}$$

$$K_{3A} = [C - DNA]\,[RNAP]\,/\,[C - DNA - RNAP]. \tag{13}$$

Using the Shea-Ackers based approach, i.e. assuming that the transcription activity is proportional to the equilibrium promoter occupancy by RNAP, we derive the expression for CRISPR promoter transcriptional activity:

$$\varphi = \gamma \frac{Z_{RNAP} + Z_{C-RNAP}}{1 + Z_{RNAP} + Z_C + Z_{C-RNAP}} \tag{14}$$

where γ is a proportionality constant, while configuration statistical weights correspond to: $Z_{RNAP} = [RNAP - DNA]\,/\,[DNA]$ — RNAP alone bound to the promoter, $Z_C = [C - DNA]\,/\,[DNA]$ –C monomer alone bound to its binding site, $Z_{C-RNAP} = [C - DNA - RNAP]\,/\,[DNA]$ — RNAP

recruited to the promoter by a bound C monomer. We can obtain φ dependence on C concentration:

$$\varphi(C) = \gamma \frac{d + def[C]}{1 + d + e[C] + def[C]} \tag{15}$$

If we introduce parameters expressed in terms of the equilibrium binding constants and RNAP concentration:

$$d = [RNAP]/K_{1A} \tag{16}$$

$$e = 1/K_{2A} \tag{17}$$

$$f = K_{1A}/K_{3A}. \tag{18}$$

To estimate the parameters, we use a condition:

$$\varphi(0) = 10 \frac{1}{min} \tag{19}$$

which corresponds to the value in Djordjevic et al. (2012), and:

$$\varphi(C_{eq}) = 100 \frac{1}{min} \tag{20}$$

Another (evident) condition is that the fraction, which appears on the right-hand side of the Equation (15), has to be smaller than 1. By adjusting the parameters to satisfy the conditions (19) and (20), we obtain $d < 1/9$, which allows setting the values of d and γ. Further, we notice that $e = 99/([C]_{eq} \cdot (f - 100))$ and, having fixed the value of f, we can adjust e with respect to $[C]_{eq}$.

The unprocessed [pre-crRNA] and processed [crRNA] transcript amounts change with time according to the Equations (2) and (3), where φ is given by .

Cooperative Model of *cas* and CRISPR Expression

As opposed to the constitutive *cas* operon expression, we here assume that the *cas* promoter is regulated by C as in the wild type AhdI RM system (Bogdanova et al., 2008), through cooperative interactions (see **Figure 5**). The following set of reactions describes the transcriptional regulation of the *cas* promoter by the C protein (note the promoter configurations shown in **Figure 5**):

$$C + C \underset{K_1}{\xleftrightarrow{\hspace{1cm}}} D \tag{21}$$

$$DNA + RNAP \underset{K_2}{\xleftrightarrow{\hspace{1cm}}} RNAP - DNA \tag{22}$$

$$D + DNA \underset{K_3}{\xleftrightarrow{\hspace{1cm}}} D - DNA \tag{23}$$

$$D - DNA + D \underset{K_4}{\xleftrightarrow{\hspace{1cm}}} T - DNA \tag{24}$$

$$D - DNA + RNAP \underset{K_5}{\xleftrightarrow{\hspace{1cm}}} D - DNA - RNAP \tag{25}$$
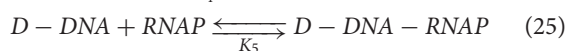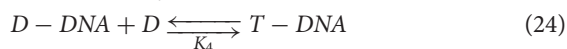
where C and D stand for C protein monomers and dimers, respectively.

The reactions (21)–(25) represent:

- (21) C monomers dimerization;
- (22) *RNAP* binding to the *cas* promoter forming *RNAP-DNA* complex;

- (23) *D* binding to the distal binding site forming *D-DNA* complex;
- (24) *D* recruitment to the proximal binding site forming *T-DNA* complex;
- (25) *RNAP* recruitment to the *cas* promoter forming *D-DNA-RNAP* complex.

In equilibrium the above reactions lead to the following equations of the equilibrium dissociation constants:

$$K_1 = \frac{[C]^2}{[D]} \tag{26}$$

$$K_2 = \frac{[DNA][RNAP]}{[RNAP - DNA]} \tag{27}$$

$$K_3 = \frac{[D][DNA]}{[D - DNA]} \tag{28}$$

$$K_4 = \frac{[D][D - DNA]}{[T - DNA]} \tag{29}$$

$$K_5 = \frac{[RNAP][D - DNA]}{[D - DNA - RNAP]} \tag{30}$$

Taking into account the aforementioned Shea-Ackers assumption we obtain:

$$\varphi_{Cas} = \alpha \frac{Z_{RNAP} + Z_{D-RNAP}}{1 + Z_{RNAP} + Z_{D-RNAP} + Z_T}, \tag{31}$$

α is a proportionality constant, $Z_{RNAP} = [RNAP-DNA]/[DNA]$, $Z_{D-RNAP} = [D - DNA - RNAP]/[DNA]$ and $Z_T = [T - DNA]/[DNA]$ denote the statistical weights of only RNAP bound to the promoter, RNAP recruited to the promoter by a C dimer bound to the distal binding site, and a C tetramer repressing transcription, respectively.

By using Equations (26)–(30), the Equation (31) can be rewritten in terms of C monomer concentration (following the notation in Bogdanova et al., 2008; Rodic et al., 2017):

$$\varphi_{Cas}(C) = \alpha \frac{a + b[C]^2}{1 + a + b[C]^2 + c[C]^4} \tag{32}$$

which can be expressed, by using the redefined parameters, in the following form:

$$\varphi_{Cas}(C) = \alpha \frac{a + ap[C]^2}{1 + a + ap[C]^2 + p^2q[C]^4}. \tag{33}$$

We set α so that the equilibrium value of *cas* transcription activity corresponds to one (adapted from Bogdanova et al., 2008). Parameters a, p, and q depend on the equilibrium dissociation constants and RNAP concentration and are given by:

$$a = [RNAP]/K_2 \tag{34}$$

$$p = \frac{K_2}{K_1 K_3 K_5} \tag{35}$$

$$q = \frac{1}{K_1^2 K_3 K_4 p^2} = \frac{K_3 K_5^2}{K_2^2 K_4} \tag{36}$$

While their values are deduced from the already determined a, b, and c, that correspond to the best fit to the AhdI experimentally measured transcription activity vs. C (Bogdanova et al., 2008).

Regarding the dynamics, note that C and Cas6e transcript and protein amounts change with time according to the Equations (4)–(6), where φ_{Cas} is given by .

Similarly as for the constitutive model, we keep φ constant, in the case when we consider inducing the system through increasing pre-crRNA processing rate. When we also consider regulation of CRISPR array transcription, we assume that CRISPR promoter is regulated by C in the same way as *cas* promoter. Thus, following the same procedure we obtain for the CRISPR promoter transcription activity:

$$\varphi = \alpha' \frac{a' + a'p'[C]^2}{1 + a' + a'p'[C]^2 + p'^2 q'[C]^4} \tag{37}$$

where constants α', a', p', and q' are determined by imposing the same constraints on φ as above (-). Specifically, these constraints lead to the condition $a' < \frac{1}{9}$, which allows setting parameters a' and α'. Further, from Equation (20) we express p' in terms of q' and get $q' < \frac{1}{400*99}$ (deduced from the real roots criterion of quadratic equation), based on which we set q', and subsequently obtain the relation for adjusting p' with respect to k_{eq} (i.e., C_{eq}). Again, the unprocessed [pre-crRNA] and processed [crRNA] transcript amounts change with time according to the Equations (2) and (3), where φ is replaced with (37).

Changing Pre-crRNA Processing Rate

From Equation (1) we have that

$$k_{eq} = [Cas6e]_{eq} \cdot k^*, \tag{38}$$

where we adjust the equilibrium value of k in the constitutive and the cooperative case by varying the concentration of Cas6e in equilibrium. The equilibrium Cas6e concentration can be derived from the steady-state conditions for Equations and :

$$[Cas6e]_{eq} = \frac{k_{Cas6e}}{\lambda_{Cas}\lambda_{Cas6e}} \varphi_{Cas}(C_{eq}). \tag{39}$$

In the model of constitutive C and Cas6e expression, the equilibrium concentration of Cas6e is adjusted through the change of φ_{Cas} (being constant with time). In the case of cooperative C and Cas6e expression, $[Cas6e]_{eq}$ is adjusted through the change of α in Equation (33), i.e., through the change of overall *cas* promoter strength, taking into account that $[C]_{eq}$ is proportional to $[Cas6e]_{eq}$ according to (7).

Joint Change of *k* and φ

We here investigate how the joint change of k and φ, which corresponds to the joint increase of *cas6e* and CRISPR array gene expression, affects the dynamics of [pre-crRNA] and [crRNA] transcripts. We start from the baseline model of infinitely abrupt increase of k and φ. We then compare the baseline model to the more realistic case of constitutive and the cooperative models. We take φ change from the initial value of 10 1/min to 100 1/min in equilibrium, while k_{eq} takes on values λ_{pre}, $10\lambda_{pre}$, and

$100\lambda_{pre}$. Note that the change in k_{eq}, implies joint change of φ_{Cas} in Equation (4) and e in Equation (15) in the constitutive case; in the cooperative case it implies joint change of α and p in Equation (33) and p' in Equation (37), which ensures the same functional dependency $\varphi(t)$, for different values of k_{eq}.

Perturbing Pre-crRNA Degradation Rate λ_{pre}

The pre-crRNA degradation rate λ_{pre} is perturbed (decreased) in the following two cases:

i. With the transcription rate φ (10 1/min) held constant. The equilibrium value of k is then adjusted by varying φ_{Cas} in the constitutive, and α in the cooperative model.
ii. When both φ and the processing rate k reach the equilibrium value (100 1/min) gradually, with the effect of the change assessed in all three models (baseline, constitutive and cooperative). k_{eq} reaches the value 100 1/min through the change of φ_{Cas} in the constitutive, and α and p in the cooperative model, while φ increases from $\varphi(0) = 10$ 1/min to $\varphi(C_{eq}) = 1,001/min$ through adjusting the parameters e in the constitutive, and p' in the cooperative model.

Note that changing λ_{pre} affects the initial amount of pre-crRNA (which is an initial condition for the differential equations) according to the relation $[pre - crRNA]_{eq}(t = 0) = \varphi(t = 0)/\lambda_{pre}$ (see Equation 2), which follows from the steady-state condition for pre-crRNA when the system is not activated.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2017.02139/full#supplementary-material

REFERENCES

Al-Attar, S., Westra, E. R., van der Oost, J., and Brouns, S. J. (2011). Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes. *Biol. Chem.* 392, 277–289. doi: 10.1515/bc.2011.042

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., et al. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709–1712. doi: 10.1126/science.1138140

Bogdanova, E., Djordjevic, M., Papapanagiotou, I., Heyduk, T., Kneale, G., and Severinov, K. (2008). Transcription regulation of the type II restriction-modification system AhdI. *Nucleic Acids Res.* 36, 1429–1442. doi: 10.1093/nar/gkm1116

Bogdanova, E., Zakharova, M., Streeter, S., Taylor, J., Heyduk, T., Kneale, G., et al. (2009). Transcription regulation of restriction-modification system Esp1396I. *Nucleic Acids Res.* 37, 3354–3366. doi: 10.1093/nar/gkp210

Bouffartigues, E., Buckle, M., Badaut, C., Travers, A., and Rimsky, S. (2007). H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. *Nat. Struct. Mol. Biol.* 14, 441–448. doi: 10.1038/nsmb1233

Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J., Snijders, A. P., et al. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321, 960–964. doi: 10.1126/science.1159689

Chen, C. C., Fang, M., Majumder, A., and Wu, H. Y. (2001). A 72-base pair AT-rich DNA sequence element functions as a bacterial gene silencer. *J. Biol. Chem.* 276, 9478–9485. doi: 10.1074/jbc.M010501200

Djordjevic, M., Djordjevic, M., and Severinov, K. (2012). CRISPR transcript processing: a mechanism for generating a large number of small interfering RNAs. *Biol. Direct* 7:24. doi: 10.1186/1745-6150-7-24

Hill, T. L. (2013). *Cooperativity Theory in Biochemistry: Steady-State and Equilibrium Systems*. New York, NY: Springer Science & Business Media.

Hille, F., and Charpentier, E. (2016). CRISPR-Cas: biology, mechanisms and relevance. *Proc. Natl. Acad. Sci. U.S.A.* 371:20150496. doi: 10.1098/rstb.2015.0496

Kruger, D. H., and Schroeder, C. (1981). Bacteriophage T3 and bacteriophage T7 virus-host cell interactions. *Microbiol. Rev.* 45, 9–51.

Kunne, T., Kieper, S. N., Bannenberg, J. W., Vogel, A. I., Miellet, W. R., Klein, M., et al. (2016). Cas3-derived target DNA degradation fragments fuel primed CRISPR adaptation. *Mol. Cell* 63, 852–864. doi: 10.1016/j.molcel.2016.07.011

Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I., and Koonin, E. V. (2006). A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct* 1:7. doi: 10.1186/1745-6150-1-7

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., et al. (2011). Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* 9, 467–477. doi: 10.1038/nrmicro2577

Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., et al. (2015). An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* 13, 722–736. doi: 10.1038/nrmicro3569

Medina-Aparicio, L., Rebollar-Flores, J., Gallego-Hernandez, A., Vazquez, A., Olvera, L., Gutierrez-Rios, R., et al. (2011). The CRISPR/Cas immune system is an operon regulated by LeuO, H-NS, and leucine-responsive regulatory protein in *Salmonella enterica* serovar Typhi. *J. Bacteriol.* 193, 2396–2407. doi: 10.1128/JB.01480-10

Mojica, F. J., and Diez-Villasenor, C. (2010). The on-off switch of CRISPR immunity against phages in *Escherichia coli*. *Mol. Microbiol.* 77, 1341–1345. doi: 10.1111/j.1365-2958.2010.07326.x

Morozova, N., Sabantsev, A., Bogdanova, E., Fedorova, Y., Maikova, A., Vedyaykin, A., et al. (2015). Temporal dynamics of methyltransferase and restriction endonuclease accumulation in individual cells after introducing a restriction-modification system. *Nucleic Acids Res.* 44, 790–800. doi: 10.1093/nar/gkv1490

Mruk, I., and Blumenthal, R. M. (2008). Real-time kinetics of restriction–modification gene expression after entry into a new host cell. *Nucleic Acids Res.* 36, 2581–2593. doi: 10.1093/nar/gkn097

Mruk, I., and Kobayashi, I. (2013). To be or not to be: regulation of restriction–modification systems and other toxin–antitoxin systems. *Nucleic Acids Res.* 42, 70–86. doi: 10.1093/nar/gkt711

Musharova, O., Klimuk, E., Datsenko, K. A., Metlitskaya, A., Logacheva, M., Semenova, E., et al. (2017). Spacer-length DNA intermediates are associated with Cas1 in cells undergoing primed CRISPR adaptation. *Nucleic Acids Res.* 45, 3297–3307. doi: 10.1093/nar/gkx097

Patterson, A. G., Yevstigneyeva, M. S., and Fineran, P. C. (2017). Regulation of CRISPR–Cas adaptive immune systems. *Curr. Opin. Microbiol.* 37, 1–7. doi: 10.1016/j.mib.2017.02.004

Pingoud, A., Wilson, G. G., and Wende, W. (2014). Type II restriction endonucleases—a historical perspective and more. *Nucleic Acids Res.* 42, 7489–7527. doi: 10.1093/nar/gku447

Pougach, K., Semenova, E., Bogdanova, E., Datsenko, K. A., Djordjevic, M., Wanner, B. L., et al. (2010). Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. *Mol. Microbiol.* 77, 1367–1379. doi: 10.1111/j.1365-2958.2010.07265.x

Pul, U., Wurm, R., Arslan, Z., Geissen, R., Hofmann, N., and Wagner, R. (2010). Identification and characterization of E. coli CRISPR-cas promoters and their silencing by H-NS. *Mol. Microbiol.* 75, 1495–1512. doi: 10.1111/j.1365-2958.2010.07073.x

Ratner, H. K., Sampson, T. R., and Weiss, D. S. (2015). I can see CRISPR now, even when phage are gone: a view on alternative CRISPR-Cas functions from the prokaryotic envelope. *Curr. Opin. Infect. Dis.* 28:267. doi: 10.1097/QCO.0000000000000154

Richter, C., Chang, J. T., and Fineran, P. C. (2012). Function and regulation of clustered regularly interspaced short palindromic repeats (CRISPR) / CRISPR associated (Cas) systems. *Viruses* 4, 2291–2311. doi: 10.3390/v4102291

Rodic, A., Blagojevic, B., Zdobnov, E., Djordjevic, M., and Djordjevic, M. (2017). Understanding key features of bacterial restriction-modification systems through quantitative modeling. *BMC Syst. Biol.* 11, 1–15. doi: 10.1186/s12918-016-0377-x

Semenova, E., Savitskaya, E., Musharova, O., Strotskaya, A., Vorontsova, D., Datsenko, K. A., et al. (2016). Highly efficient primed spacer acquisition from targets destroyed by the *Escherichia coli* type IE CRISPR-Cas interfering complex. *Proc. Natl. Acad. Sci. U.S.A.* 113, 7626–7631. doi: 10.1073/pnas.1602639113

Severinov, K., Ispolatov, I., and Semenova, E. (2016). The influence of copy-number of targeted extrachromosomal genetic elements on the outcome of CRISPR-Cas defense. *Proc. Natl. Acad. Sci. U.S.A.* 3:45. doi: 10.3389/fmolb.2016.00045

Sternberg, S. H., Richter, H., Charpentier, E., and Qimron, U. (2016). Adaptation in CRISPR-Cas Systems. *Mol. Cell* 61, 797–808. doi: 10.1016/j.molcel.2016.01.030

Stratmann, T., Pul, U., Wurm, R., Wagner, R., and Schnetz, K. (2012). RcsB-BglJ activates the *Escherichia coli* leuO gene, encoding an H-NS antagonist and pleiotropic regulator of virulence determinants. *Mol. Microbiol.* 83, 1109–1123. doi: 10.1111/j.1365-2958.2012.07993.x

Tao, T., Bourne, J. C., and Blumenthal, R. M. (1991). A family of regulatory genes associated with type II restriction-modification systems. *J. Bacteriol.* 173, 1367–1375. doi: 10.1128/jb.173.4.1367-1375.1991

Touchon, M., Charpentier, S., Clermont, O., Rocha, E. P., Denamur, E., and Branger, C. (2011). CRISPR distribution within the *Escherichia coli* species is not suggestive of immunity-associated diversifying selection. *J. Bacteriol.* 193, 2460–2467. doi: 10.1128/JB.01307-10

Vijesurier, R. M., Carlock, L., Blumenthal, R. M., and Dunbar, J. C. (2000). Role and mechanism of action of C · *Pvu*II, a regulatory protein

conserved among restriction-modification systems. *J. Bacteriol.* 182, 477–487. doi: 10.1128/JB.182.2.477-487.2000

Westra, E. R., Pul, U., Heidrich, N., Jore, M. M., Lundgren, M., Stratmann, T., et al. (2010). H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Mol. Microbiol.* 77, 1380–1393. doi: 10.1111/j.1365-2958.2010. 07315.x

Yang, C. D., Chen, Y. H., Huang, H. Y., Huang, H. D., and Tseng, C. P. (2014). CRP represses the CRISPR/Cas system in *Escherichia coli*: evidence that endogenous CRISPR spacers impede phage P1 replication. *Mol. Microbiol.* 92, 1072–1091. doi: 10.1111/mmi.12614

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Calculating hard probe radiative energy loss beyond soft-gluon approximation: how valid is the approximation?

Bojana Blagojevic,[1] Magdalena Djordjevic,[1] and Marko Djordjevic[2]

[1]*Institute of Physics Belgrade, University of Belgrade, Belgrade, Serbia*
[2]*Faculty of Biology, Institute of Physiology and Biochemistry, University of Belgrade, Belgrade, Serbia*
(Dated: April 23, 2018)

The soft-gluon approximation, which implies that radiated gluon carries away a small fraction of initial parton's energy, is a commonly used assumption in calculating radiative energy loss of high momentum partons traversing QGP created at RHIC and LHC. While soft-gluon approximation is convenient, different theoretical approaches reported significant radiative energy loss of high $p_\perp$ partons, thereby questioning its validity. To address this issue, we relaxed the soft-gluon approximation within DGLV formalism. The obtained analytical expressions are quite distinct compared to the soft-gluon case. However, numerical results for the first order in opacity fractional energy loss lead to small differences in predictions for the two cases. The difference in the predicted number of radiated gluons is also small. Moreover, the effect on these two variables has an opposite sign, which when combined results in almost overlapping suppression predictions. Therefore, our results imply that, contrary to the commonly held doubts, the soft-gluon approximation in practice works surprisingly well in DGLV formalism. Finally, we also discuss generalizing this relaxation in the dynamical QCD medium, which suggests a more general applicability of the conclusions obtained here.

PACS numbers: 12.38.Mh; 24.85.+p; 25.75.-q

I. INTRODUCTION

One of the main assumptions in the radiative energy loss calculations of energetic parton (in the further text referred to as jet) traversing the QGP medium, is the soft-gluon approximation which assumes that radiated gluon carries away a small portion of initial jet energy, i.e. $x = \omega/E \ll 1$, where E is the energy of initial jet and ω is the radiated gluon energy.

Such assumption was widely used in various energy loss models: *i)* in multiple soft scattering based ASW model [1–3]; *ii)* BDMPS [4, 5] and BDMPS-Z [6, 7]; *iii)* in opacity expansion based GLV model [8, 9] and *iv)* in multi-gluon evolution based HT approach [10, 11], etc. These various energy loss models predict a significant medium induced radiative energy loss, questioning the validity of the soft-gluon approximation. To address this issue, a finite x (or large x limit) was introduced in some of these models [12, 13] or their extensions [14]. However, introduction of finite x lead to different conclusions on the importance of relaxing the soft-gluon approximation, which was assessed from relatively small [14], but noticeable, to moderately large [13].

The soft-gluon approximation was also used in the development of our dynamical energy loss formalism [15–17], specifically in its radiative energy loss component. This formalism was comprehensively tested against angular averaged nuclear modification factor R_{AA} [18, 19] data, where we obtained robust agreement for wide range of probes [17, 20], centralities [21] and beam energies [20, 22], including clear predictions for future experiments [23, 24]. This might strongly suggest that our energy loss formalism can well explain the jet-medium interactions in QGP, making this formalism suitable for

the tomography of QCD medium.

However, the soft-gluon approximation obviously breaks-down for: *i)* intermediate momentum ranges ($5 < p_\perp < 10$ GeV) where the experimental data are most abundant and with the smallest error-bars, and *ii)* gluon energy loss, since due to the color factor of 9/4 gluons lose significantly more energy compared to quark jets, therefore questioning the reliability of our formalism in such cases. Due to this, and for precise predictions, it became necessary to relax the soft-gluon approximation, and consequently test its validity in dynamical energy loss formalism.

This paper presents our first step toward this goal. Since the dynamical energy loss is computationally very demanding, we will, in this study, start with relaxing this approximation on its simpler predecessor, i.e. DGLV [25] formalism. Within this, we will concentrate on gluon jets, since, due to their color factor, the soft-gluon approximation has the largest impact for this type of partons. For the gluon jets, we perform the radiative energy loss calculation, to the first order in the number of scattering centers (opacity), where we consider that the radiation of one gluon is induced by one collisional interaction with the medium.

Our calculation is done within the pQCD approach for a finite size, optically thin QCD medium and since it is technically demanding – it will be divided in several steps: *i)* First, the calculation will be done in the simplest case of massless gluons in the system of static scattering centers [26] within GLV, *ii)* Then it will be extended towards the gluons with the effective mass [27], which presents expansion of DGLV [25] toward larger loss of jet energy via radiated gluon, and *iii)* Finally, we will discuss the impact of finite x on the radiative energy loss, when

dynamical medium [17] (i.e. a recoil with the medium constituents) is accounted.

In that manner we will assess the validity of the soft-gluon assumption for gluon jets, and this will also provide an insight into whether or not a finite x has to be implemented in quark-jet radiative energy loss calculations within our formalism. Namely, if the relaxation of the soft-gluon approximation only slightly modifies gluon-jet radiative energy loss, then even smaller modification would be expected in quark-jet case, thus making this relaxation redundant. Otherwise, if the effect of a finite x appears to be a significant in gluon-jet case, then the relaxation in quark-jet case may also be required, and would represent an important future task.

Secondly, as stated above, the relaxation of the soft-gluon approximation is needed in order to extend the applicability of our model [17] towards intermediate momentum region. Thus, the another benefit of this relaxation would be to extend the $p_\perp$ range in which our predictions are valid.

The sections are organized as follows: In section II, we provide the theoretical framework. In section III, we outline the computation of the zeroth order in opacity gluon-jet radiative energy loss in static QCD medium, beyond soft-gluon approximation, in the cases of both massless and massive gluons. For $x \ll 1$ the results from [9, 25] are reproduced.

Section IV contains concise description of relaxing the soft-gluon approximation in calculating the first order in opacity radiative energy loss for massless gluon jet in static QCD medium. In a limit of very small x result from [9] is recovered.

In section V we explain the computation of the first order in opacity gluon-jet energy loss in static QCD medium, with effective gluon mass [27] included, and beyond soft-gluon approximation. This presents an extension of the calculations from [25] toward finite x, so that results from [25] can be recovered in $x \ll 1$ limit. The detailed calculations corresponding to sections III - V are presented in the Appendices C-J.

In section VI we outline the numerical estimates based on our beyond soft-gluon calculations for gluon jet and the comparison with our previous results from [25], i.e. the results with soft-gluon approximation. Particularly, we investigate the effect of finite x on gluon-jet fractional radiative energy loss, number of radiated gluons, differential radiative energy loss, single gluon radiation spectrum and gluon suppression [28]. Conclusions and outlook are presented in section VII.

II. THEORETICAL FRAMEWORK

In this work, we concentrate on relaxing soft-gluon approximation in calculating the first order in opacity radiative energy loss of high $p_\perp$ eikonal gluon jets within (GLV) DGLV [25] formalism. That is, we assume that high $p_\perp$ gluon jet is produced inside a "thin" finite QGP

medium at some initial point $(t_0, z_0, \mathbf{x}_0)$, and that the medium is composed of static scattering centers [26]. Therefore, we model the interactions in QGP assuming a static (Debye) colored-screened Yukawa potential, whose Fourier and color structure acquires the following form ([9, 25, 26]):

$$V_n = V(q_n)e^{iq_n x_n} = 2\pi\delta(q_n^0)v(\vec{\mathbf{q}}_n)e^{-i\vec{\mathbf{q}}_n\vec{\mathbf{x}}_n}\times$$
$$\times T_{a_n}(R)\otimes T_{a_n}(n), \qquad (1)$$

$$v(\vec{\mathbf{q}}_n) = \frac{4\pi\alpha_s}{\vec{\mathbf{q}}_n^2 + \mu^2}, \qquad (2)$$

where x_n denotes time-space coordinate of the nth scattering center, μ is Debye screening mass, $\alpha_s = g_s^2/4\pi$ is strong coupling constant, while $T_{a_n}(R)$ and $T_{a_n}(n)$ denote the generators in $SU(N_c = 3)$ color representation of gluon jet and target (scattering center), respectively.

For consistency with [9, 25], we use the same notation for 4D vectors (e.g. momenta), which is described in detail in Appendix A and proceed throughout using Light-cone coordinates. The same Appendix contains algebra manipulation and identities for $SU(N_c)$ generators, as well as the Feynman rules, used in these calculations.

The approximations that we assume throughout the paper are stated in Appendix B.

The small transverse momentum transfer elastic cross section for interaction between gluon jet and target parton in GW approach [8, 26] is given by:

$$\frac{d\sigma_{el}}{d^2\mathbf{q}_1} = \frac{C_2(G)C_2(T)}{d_G}\frac{|v(0,\mathbf{q}_1)|^2}{(2\pi)^2}, \qquad (3)$$

where $\mathbf{q}_1$ corresponds to transverse momentum of exchanged gluon, $C_2(G)$ represents Casimir operator in adjoint representation (G) of gluons $SU(N_c = 3)$ with dimension $d_G = 8$, whereas $C_2(T)$ denotes Casimir operator in target (T) representation.

Since this formalism assumes optically "thin" plasma, the final results are expanded in powers of opacity, which is defined as the mean number of collisions in the medium: $L/\lambda = N\sigma_{el}/A_\perp$ [9], where L is the thickness of the QCD medium, λ is a mean free path, while N denotes the number of scatterers (targets) in transverse area $A_\perp$. Note that, we restrict our calculations to the first order in opacity, which is shown to be the dominant term ([29, 30]).

III. ZEROTH ORDER RADIATIVE ENERGY LOSS

To gradually introduce technically involving beyond soft-gluon calculations, we first concentrate on massless gluons traversing static QCD medium.

We start with M_0 Feynman diagram, which corresponds to the source J that produces off-shell gluon with momentum $p + k$, that further, without interactions with

QCD medium, radiates on-shell gluon with momentum k and emerges with momentum p. We will further refer to these two outgoing gluons as the radiated (k) and the final (p) gluon. Note that, both in this and consecutive sections that involve interactions with one and two scattering centers, we consistently assume that initial jet propagates along the longitudinal z axis. The detailed calculation of M_0 for finite x in massless case is presented in Appendix C, with all assumptions listed in Appendix B.

We also assume that gluons are transversely polarized particles and although we work in covariant gauge, we can choose any polarization vector for the external on-shell gluons [14], so in accordance with [9, 14, 25] we choose $n^\mu = [0, 2, \mathbf{0}]$ (i.e. $\epsilon(k) \cdot k = 0$, $\epsilon(k) \cdot n = 0$ and $\epsilon(p) \cdot p = 0$, $\epsilon(p) \cdot n = 0$). Likewise, we assume that the source has also the same polarization as real gluons [14] (i.e. $\epsilon(p+k) \cdot (p+k) = 0$, $\epsilon(p+k) \cdot n = 0$). Thus, for massless gluon's momenta we have:

$$p + k = [E^+, E^-, \mathbf{0}], \quad k = [xE^+, \frac{\mathbf{k}^2}{xE^+}, \mathbf{k}],$$
$$p = [(1-x)E^+, \frac{\mathbf{p}^2}{(1-x)E^+}, \mathbf{p}], \tag{4}$$

where $E^+ = p^0 + k^0 + p_z + k_z$, $E^- = p^0 + k^0 - p_z - k_z$ and due to 4-momentum conservation:

$$\mathbf{p} + \mathbf{k} = 0. \tag{5}$$

The polarization vectors read:

$$\epsilon_i(k) = [0, \frac{2\boldsymbol{\epsilon}_i \cdot \mathbf{k}}{xE^+}, \boldsymbol{\epsilon}_i], \qquad \epsilon_i(p) = [0, \frac{2\boldsymbol{\epsilon}_i \cdot \mathbf{p}}{(1-x)E^+}, \boldsymbol{\epsilon}_i],$$
$$\epsilon_i(p+k) = [0, 0, \boldsymbol{\epsilon}_i], \tag{6}$$

where $i = 1, 2$, and we also make use of Eq. (5). So, the amplitude that gluon jet, produced at x_0 inside QCD medium, radiates a gluon of color c without final state interactions reads:

$$M_0 = J_a(p+k)e^{i(p+k)x_0}(-2ig_s)(1 - x + x^2)\frac{\boldsymbol{\epsilon} \cdot \mathbf{k}}{\mathbf{k}^2}(T^c)_{da}. \tag{7}$$

The radiation spectrum is obtained when Eq. (7) is substituted in:

$$d^3N_g^{(0)}d^3N_J \approx \mathrm{Tr}\langle |M_0|^2 \rangle \frac{d^3\vec{\mathbf{p}}}{(2\pi)^3 2p^0} \frac{d^3\vec{\mathbf{k}}}{(2\pi)^3 2\omega}, \tag{8}$$

where $\omega = k_0$, and where d^3N_J reads:

$$d^3N_J = d_G|J(p+k)|^2\frac{d^3\vec{\mathbf{p}}_J}{(2\pi)^3 2E_J}. \tag{9}$$

Here $E_J = E = p_0 + k_0$ and $\vec{\mathbf{p}}_J$ denotes energy and 3D momentum of the initial gluon jet, respectively. The jet part can be decoupled by using the equality:

$$\frac{d^3\vec{\mathbf{p}}}{(2\pi)^3 2p^0}\frac{d^3\vec{\mathbf{k}}}{(2\pi)^3 2\omega} = \frac{d^3\vec{\mathbf{p}}_J}{(2\pi)^3 2E_J}\frac{dxd^2\mathbf{k}}{(2\pi)^3 2x(1-x)}, \tag{10}$$

which is obtained by substituting $p_z, k_z \to p_z^J, xE$. Finally, energy spectrum acquires the form:

$$\frac{xd^3N_g^{(0)}}{dxd\mathbf{k}^2} = \frac{\alpha_s}{\pi}\frac{C_2(G)}{\mathbf{k}^2}\frac{(1 - x + x^2)^2}{1 - x}, \tag{11}$$

which recovers the well-known Altarrelli-Parisi [31] result.

We now briefly concentrate on generating result in finite temperature QCD medium, since in [27], it was shown that gluons in finite temperature QGP can be approximated as a massive transverse plasmons with mass $m_g = \mu/\sqrt{2}$, where μ is the Debye mass. In this case, M_0 amplitude becomes:

$$M_0 = J_a(p+k)e^{i(p+k)x_0}(-2ig_s)(1 - x + x^2)\times$$
$$\times \frac{\boldsymbol{\epsilon} \cdot \mathbf{k}}{\mathbf{k}^2 + m_g^2(1 - x + x^2)}(T^c)_{da}, \tag{12}$$

leading to:

$$\frac{xd^3N_g^{(0)}}{dxd\mathbf{k}^2} = \frac{\alpha_s}{\pi}\frac{C_2(G)\,\mathbf{k}^2}{(\mathbf{k}^2 + m_g^2(1 - x + x^2))^2}\times$$
$$\times \frac{(1 - x + x^2)^2}{1 - x}. \tag{13}$$

IV. FIRST ORDER RADIATIVE ENERGY LOSS IN MASSLESS CASE

In accordance with [25], we compute the first order in opacity radiative energy loss of gluon jet for finite x starting from the expression:

$$d^3N_g^{(1)}d^3N_J = \left(\frac{1}{d_T}\mathrm{Tr}\langle |M_1|^2 \rangle + \frac{2}{d_T}Re\,\mathrm{Tr}\langle M_2M_0^* \rangle\right)\frac{d^3\vec{\mathbf{p}}}{(2\pi)^3 2p^0}\frac{d^3\vec{\mathbf{k}}}{(2\pi)^3 2\omega}, \tag{14}$$

where M_0 corresponds to the diagram without final state interactions with QCD medium, introduced in previous

section, M_1 is the sum of all diagrams with one scattering center, M_2 is the sum of all diagrams with two scattering centers in the contact-limit case, while d_T denotes the dimension of the target color representation (for pure gluon medium $d_T = 8$). In obtaining the expression for differential energy loss, we again incorporate Eqs. (9, 10) in Eq. (14).

The assumption that initial jet propagates along z-axis, takes the following form in the two cases stated below:

1. One interaction with QCD medium (M_1):

$$p + k - q_1 = [E^+ - q_{1z}, E^- + q_{1z}, \mathbf{0}], \qquad (15)$$

where $p + k - q_1$ corresponds to the initial jet, while k and p retain the same expressions as in Eq. (4), with the distinction that now $\mathbf{p} \neq -\mathbf{k}$, since due to 4-momentum conservation, the following relation holds:

$$\mathbf{q}_1 = \mathbf{p} + \mathbf{k}; \qquad (16)$$

The rest of the notation is the same as in Eq. (4).

2. Two interactions with QCD medium (M_2):

$$p + k - q_1 - q_2 = [E^+ - q_{1z} - q_{2z}, E^- + q_{1z} + q_{2z}, \mathbf{0}], \qquad (17)$$

where $p + k - q_1 - q_2$ corresponds to the initial jet and $q_i = [q_{iz}, -q_{iz}, \mathbf{q}_i]$ to exchanged gluons, $i = 1, 2$ with $q_i^0 = 0$, while p, k retain the same expressions as in Eq. (4). Also, due to 4-momentum conservation, the following relation between gluon transverse momenta holds:

$$\mathbf{p} + \mathbf{k} = \mathbf{q}_1 + \mathbf{q}_2, \qquad (18)$$

which in the contact-limit case (when $\mathbf{q}_1 + \mathbf{q}_2 = 0$) reduces to $\mathbf{p} + \mathbf{k} = 0$.

Note that Eq. (16) has to be satisfied for M_1 diagrams in order to claim that initial jet propagates along z-axis, i.e. for M_1 diagrams $\mathbf{p} + \mathbf{k}$ is different from 0. This is an important distinction between the calculations presented in our study, and the calculations done within SCET formalism (see e.g. [14]), where $\mathbf{p} + \mathbf{k} = 0$ was used in calculation of *both* M_1 and M_2 diagrams, though the assumption that initial jet propagates along z-axis was used in that study as well.

The transverse polarization vectors $\epsilon_i(k)$ and $\epsilon_i(p)$ for both: M_1 and M_2 amplitudes are given by the same expression as in the previous section (with an addition that in M_1 case: $\mathbf{p} \neq -\mathbf{k}$, as discussed above), while ϵ for initial jets consistently has the same form as in Eq. (6), i.e. $\epsilon_i(p + k - q_1) = [0, 0, \boldsymbol{\epsilon}_i]$ for M_1 amplitudes, and $\epsilon_i(p + k - q_1 - q_2) = [0, 0, \boldsymbol{\epsilon}_i]$ for M_2 amplitudes.

The detailed calculation of the remaining 10 Feynman diagrams, under the approximations stated in Appendix B, contributing to the first order in opacity radiative energy loss, is given in Appendices D-H, whereas thorough derivation of the single gluon radiation spectrum beyond soft-gluon approximation in massless case is given in Appendix I and reads (energy loss expression can be straightforwardly extracted by using $dE^{(1)}/dx \equiv \omega dN_g^{(1)}/dx \approx xE dN_g^{(1)}/dx$):

$$\frac{dN_g^{(1)}}{dx} = \frac{C_2(G)\alpha_s}{\pi} \frac{L}{\lambda} \frac{(1 - x + x^2)^2}{x(1-x)} \int \frac{d^2\mathbf{q}_1}{\pi} \frac{\mu^2}{(\mathbf{q}_1^2 + \mu^2)^2} \int d\mathbf{k}^2 \times$$
$$\times \left\{ \frac{(\mathbf{k} - \mathbf{q}_1)^2}{(\frac{4x(1-x)E}{L})^2 + (\mathbf{k} - \mathbf{q}_1)^4} \left(2 - \frac{\mathbf{k} \cdot (\mathbf{k} - \mathbf{q}_1)}{\mathbf{k}^2} - \frac{(\mathbf{k} - \mathbf{q}_1) \cdot (\mathbf{k} - x\mathbf{q}_1)}{(\mathbf{k} - x\mathbf{q}_1)^2} \right) + \right.$$
$$\left. + \frac{\mathbf{k}^2}{(\frac{4x(1-x)E}{L})^2 + \mathbf{k}^4} \left(1 - \frac{\mathbf{k} \cdot (\mathbf{k} - x\mathbf{q}_1)}{(\mathbf{k} - x\mathbf{q}_1)^2} \right) + \left(\frac{1}{(\mathbf{k} - x\mathbf{q}_1)^2} - \frac{1}{\mathbf{k}^2} \right) \right\}, \qquad (19)$$

where we assumed a simple exponential distribution $\frac{2}{L} e^{\frac{-2\Delta z}{L}}$ of longitudinal distance between the gluon-jet production site and target rescattering site, emerging as $(\frac{4x(1-x)E}{L})^2$ in the denominators of the integrand. Note that, Eq. (19) reduces to massless case of Eq. (11) from [25] in the $x \to 0$ limit, as expected.

It is straightforward to show that our result is symmetric under the exchange of radiated (k) and final (p) gluon, as expected beyond soft-gluon approximation, and due to inability to distinguish between these two identical gluons.

V. GLUON RADIATIVE ENERGY LOSS IN FINITE TEMPERATURE QCD MEDIUM

Next, we note that in ultra-relativistic heavy ion collisions, finite temperature QCD medium is created, that modifies the gluon self energies, and can consequently significantly influence the radiative energy loss results. It is therefore essential to include finite temperature effects in gluon radiative energy loss calculations beyond soft-gluon approximation, which is the main goal of this section. To address this issue, we note that in [27], it was shown that gluons can be approximated as massive transverse plasmons with effective mass m_g (for gluons with the hard momenta $k \gtrsim T$) equal to its asymptotic value. The assumption of initial jet propagating along z-axis, for massive case, leads to the following form of momenta, in the three cases stated below:

1. No interaction with QCD medium (M_0):

$$p + k = [E^+, E^-, \mathbf{0}], \quad k = [xE^+, \frac{\mathbf{k}^2 + m_g^2}{xE^+}, \mathbf{k}],$$

$$p = [(1-x)E^+, \frac{\mathbf{p}^2 + m_g^2}{(1-x)E^+}, \mathbf{p}], \tag{20}$$

where Eq. (5) holds;

2. One interaction with QCD medium (M_1):

k and p retain the same expressions as in Eq. (20), with addition that (as in the previous section) Eq. (16) holds due to conservation of 4-momentum, while initial jet has the momentum of a same form as in Eq. (15).

3. Two interactions with QCD medium (M_2):

p, k have the same expressions as in Eq. (20). Also, due to 4-momentum conservation Eq. (18) holds and in the contact-limit case reduces to $\mathbf{p} + \mathbf{k} = 0$, while initial jet momentum retains the same form as in Eq. (17).

The transverse polarization vectors remain the same as in the massless case.

We retain all approximation from the previous section, which are reviewed in Appendix B, and recalculate the same 11 diagrams from Appendices C-H, also beyond soft-gluon approximation. The overview of all intermediate results is contained in Appendix J. Thus, Eq. (19) in the massive case acquires more complex form given by:

$$\frac{dN_g^{(1)}}{dx} = \frac{C_2(G)\alpha_s}{\pi} \frac{L}{\lambda} \frac{(1-x+x^2)^2}{x(1-x)} \int \frac{d^2\mathbf{q}_1}{\pi} \frac{\mu^2}{(\mathbf{q}_1^2 + \mu^2)^2} \int d\mathbf{k}^2 \times$$

$$\times \left\{ \frac{(\mathbf{k} - \mathbf{q}_1)^2 + \chi}{(\frac{4x(1-x)E}{L})^2 + ((\mathbf{k} - \mathbf{q}_1)^2 + \chi)^2} \left(2\frac{(\mathbf{k} - \mathbf{q}_1)^2}{(\mathbf{k} - \mathbf{q}_1)^2 + \chi} - \frac{\mathbf{k} \cdot (\mathbf{k} - \mathbf{q}_1)}{\mathbf{k}^2 + \chi} - \frac{(\mathbf{k} - \mathbf{q}_1) \cdot (\mathbf{k} - x\mathbf{q}_1)}{(\mathbf{k} - x\mathbf{q}_1)^2 + \chi} \right) + \right.$$

$$\left. + \frac{\mathbf{k}^2 + \chi}{(\frac{4x(1-x)E}{L})^2 + (\mathbf{k}^2 + \chi)^2} \left(\frac{\mathbf{k}^2}{\mathbf{k}^2 + \chi} - \frac{\mathbf{k} \cdot (\mathbf{k} - x\mathbf{q}_1)}{(\mathbf{k} - x\mathbf{q}_1)^2 + \chi} \right) + \left(\frac{(\mathbf{k} - x\mathbf{q}_1)^2}{((\mathbf{k} - x\mathbf{q}_1)^2 + \chi)^2} - \frac{\mathbf{k}^2}{(\mathbf{k}^2 + \chi)^2} \right) \right\}, \tag{21}$$

where $\chi = m_g^2(1 - x + x^2)$. It can easily be verified that, in the soft-gluon limit, we recover Eq. (11) from [25] (note that for gluon jet $M \equiv m_g$, so that the term $M^2 x^2$ from [25] should be neglected), and that in the massless limit Eq. (21) reduces to Eq. (19).

To our knowledge, this result presents the first introduction of effective gluon mass beyond-soft-gluon-approximation radiative energy loss. Additionally, we again verified that single gluon radiation spectrum is symmetric to substitution of p and k gluons, as necessary (see the previous section and Appendix J). Furthermore, note that the analytical form of Eq. (21) is quite different from the corresponding expression with the soft-gluon approximation (Eq. (11) from [25]). In the next section, we will evaluate the extent of numerical differences to which these two different analytical expressions lead.

In particular, we are interested in what is the effect of finite x on gluon fractional radiative energy loss ($\frac{\Delta E^{(1)}}{E}$), number of radiated gluons ($N_g^{(1)}$) and on the suppres-

sion (R_{AA}). We accordingly note that $\frac{dE^{(1)}}{dx} \equiv \omega \frac{dN_g^{(1)}}{dx} \approx xE \frac{dN_g^{(1)}}{dx}$ from which we can further straightforwardly numerically evaluate $\frac{\Delta E^{(1)}}{E}$, as well as the number of radiated gluons ($N_g^{(1)}$).

VI. NUMERICAL RESULTS

We next assess how the relaxation of soft-gluon approximation modifies gluon-jet energy loss to the 1st order in opacity. We consequently compare the predictions based on the results derived in this paper, with the one obtained in the soft-gluon limit from [25] (applied to gluons) - the comparison is done for gluons with effective mass $m_g = \mu/\sqrt{2}$, where $\mu = \sqrt{4\pi\alpha_s(1 + n_f/6)}T$ and $n_f = 3$ is the number of the effective light-quark flavor. For all figures, we assume constant strong coupling [32]
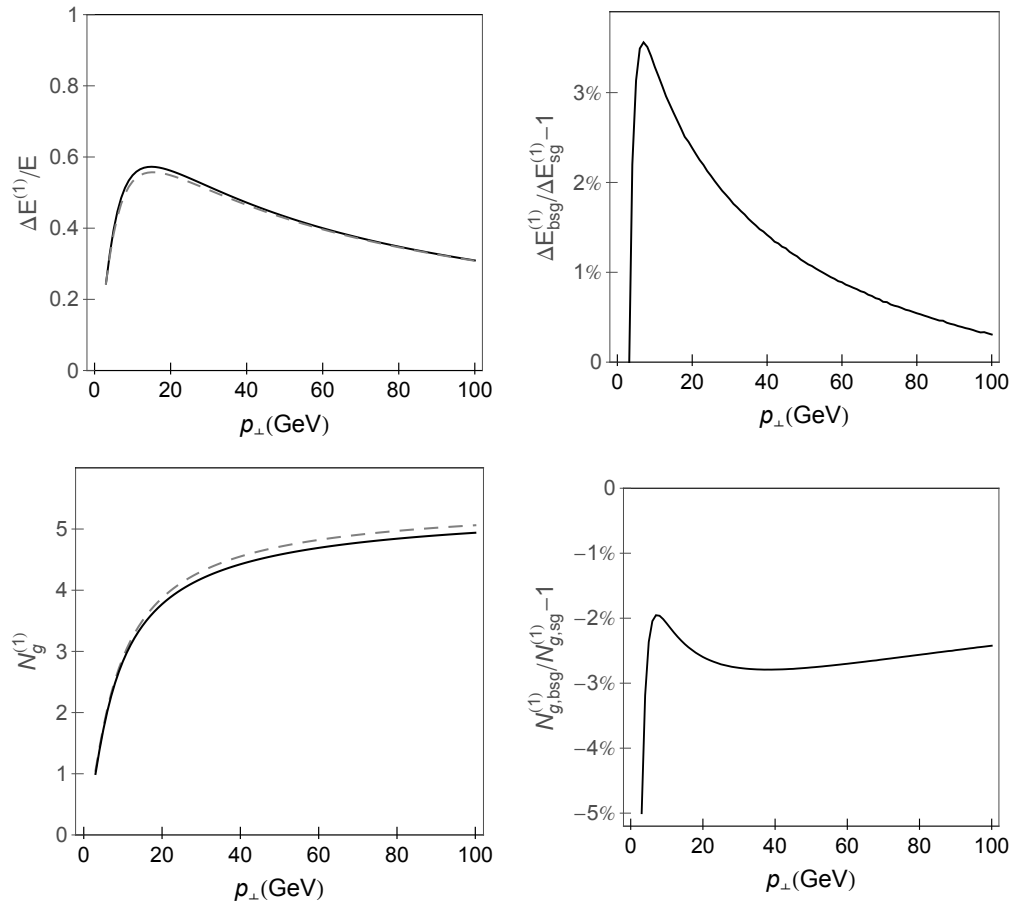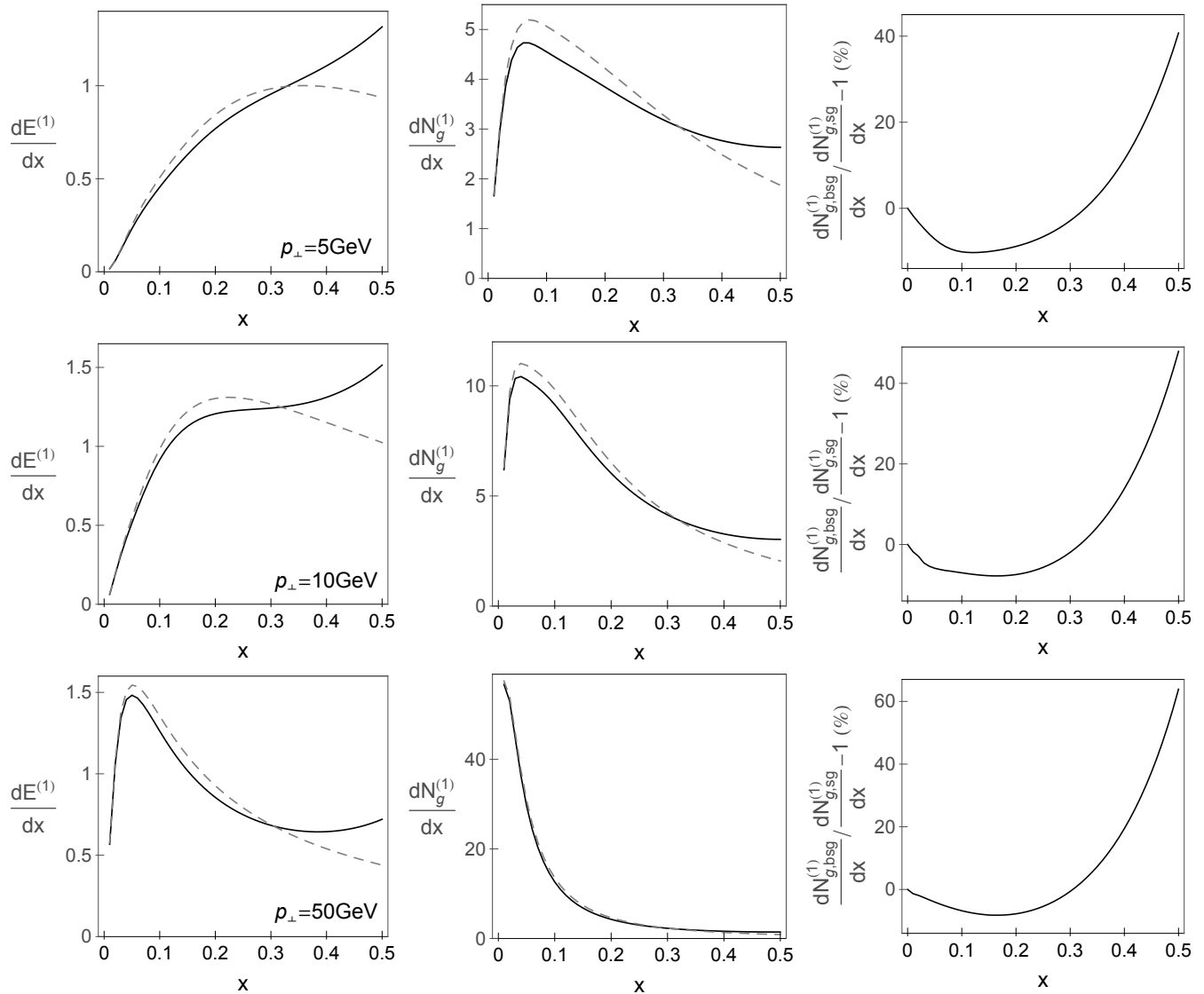
FIG. 1: The effect of relaxing the soft-gluon approximation on integrated variables to the 1st order in opacity of DGLV formalism, as a function of $p_\perp$. The top left panel compares gluon's fractional radiative energy loss without (the solid curve) and with (the dashed curve) soft-gluon approximation. The top right panel represents relative change of the radiative energy loss when the soft-gluon approximation is relaxed with respect to the soft-gluon limit. The bottom left panel compares number of radiated gluons without (the solid curve) and with (the dashed curve) soft-gluon approximation, whereas the bottom right panel provides a percentage of radiated gluon number change when soft-gluon approximation is relaxed.

$\alpha_s = \frac{g_s^2}{4\pi} = 0.3$, and use $L = 5$ fm, $\lambda = 1$ fm, $T = 300$ MeV, to mimic standard LHC conditions.

The top left panel of Fig. 1 shows comparison of the fractional radiative energy loss $\frac{\Delta E^{(1)}}{E}$, for calculations beyond the soft-gluon approximation, and with the soft-gluon approximation, as a function of initial jet transverse momentum ($p_\perp$). More specifically, the curve corresponding to beyond soft-gluon approximation (bsg) case is obtained from Eq. (21) multiplied by xE and integrated over x, while the curve corresponding to soft-gluon approximation (sg) case is obtained by numerically integrating Eq. (11) from [25]. These two curves almost overlap, even converge towards one another at higher $p_\perp$. Note that, the upper limit of x integration is equal to $1/2$ instead of 1, in order to avoid double counting. The upper limits of integration for $|\mathbf{k}|$ and $|\mathbf{q_1}|$, determined kinematically, are $2x(1-x)E$ and $\sqrt{4ET}$, respectively [25].

The bottom left panel of Fig. 1 presents comparison of number of radiated gluons in bsg and sg cases. These two curves also nearly overlap, with a slight disagreement at

higher $p_\perp$.

Quantitative assessment of relaxing the soft-gluon approximation on these two variables can be observed from the two right panels of Fig. 1. We see that finite values of x slightly increase fractional radiative energy loss by maximum of $\sim 3\%$ up to $p_\perp \approx 10$ GeV compared to sg case. Afterwards, the difference between bsg and sg $\frac{\Delta E^{(1)}}{E}$ steeply decreases towards 0%. Additionally, finite x also decreases number of radiated gluons for a small amount (up to 5%) compared to sg case for very low transverse momenta. Further the relative difference reaches a peak of -2% also at $p_\perp \approx 10$ GeV, and for higher transverse momenta remains nearly constant somewhat below -2%. Consequently, the overall conclusion from Fig. 1 is that the effect on both variables is small and with opposite signs.

The effect of finite x value is further assessed on the differential radiative energy loss ($\frac{dE^{(1)}}{dx}$), and on single gluon radiation spectrum ($\frac{dN_g^{(1)}}{dx}$) and it's relative change. These effects are shown as a function of x on Fig. 2, for

FIG. 2: The effect of relaxing the soft-gluon approximation on differential variables to the 1st order in opacity of DGLV formalism, as a function of x. The comparison of: *i)* differential gluon radiative energy loss $(dE^{(1)}/dx)$; *ii)* single gluon radiation (spectrum) distribution in momentum fraction $(dN_g^{(1)}/dx)$ between *bsg* (the solid curve) and *sg* (the dashed curve) case, for different values of initial jet transverse momenta (5 GeV, 10 GeV, 50 GeV, as indicated in panels) is shown in the first and second column, respectively. The relative change of the single gluon radiation spectrum with respect to soft-gluon limit is shown in the third column.

different values of initial jet transverse momentum $p_\perp$; *bsg* curves for $\frac{dE^{(1)}}{dx}$ are obtained by numerically integrating Eq. (21) multiplied by xE over $|\mathbf{k}|$ and $\mathbf{q}_1$, whereas *sg* curves correspond to Eq. (11) in [25]. From Fig. 2, we observe a small difference between *bsg* and *sg* results for $x \lesssim 0.3$ (roughly up to 0.4), i.e. for smaller x, as expected. We also recognize $x \approx 0.3$ as a "cross-over" value, below which differential radiative energy loss and single gluon radiation spectrum are somewhat lower in *bsg* compared to *sg* case, and above which the opposite is true. At high value of x, i.e. $0.4 < x \leq 0.5$, the differences between our *bsg* differential radiative energy loss

and previously obtained *sg* [25] ascend to notable values ($\sim 50\%$) and increases with increasing $p_\perp$.

To investigate the effect of relaxing the soft-gluon approximation on the single gluon radiation spectrum in more detail, the third column is added in Fig. 2 (see also Fig. 3), showing relative change of $\frac{dN_g^{(1)}}{dx}$. This quantitative estimation (difference smaller than 10% for $x \lesssim 0.4$) is in agreement with the previous discussion. In particular, at higher x values, there is a notably larger spectra in *bsg* compared to *sg* case, and this difference enhances (up to 60% at $p_\perp = 50$ GeV) with increasing $p_\perp$. Never-
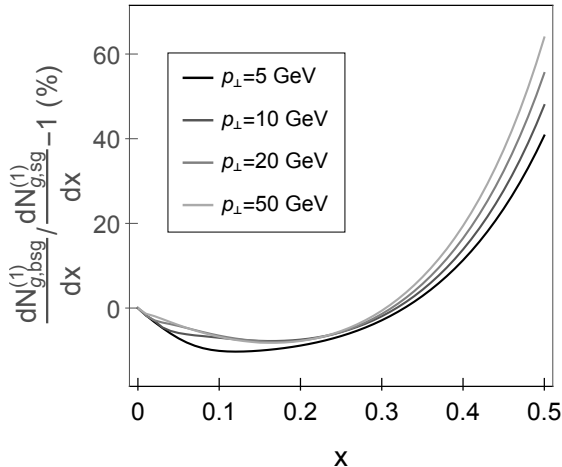
FIG. 3: The effect of relaxing the soft-gluon approximation on $dN_g^{(1)}/dx$ for different $p_\perp$ values. The relative change of the single gluon radiation spectrum with respect to soft-gluon case, calculated to the 1st order in opacity of DGLV formalism, for different values of initial $p_\perp$ (as indicated in the legend) is depicted as a function of x. The curves fade as transverse momentum increases.

theless, for both variables: ($\frac{dE^{(1)}}{dx}$ and $\frac{dN_g^{(1)}}{dx}$) bsg and sg cases lead to similar results for $x \lesssim 0.4$.

The effect of relaxing the soft-gluon approximation on single gluon radiation spectrum for different transverse momentum values of initial gluon jet is further addressed in Fig. 3. We observe that a notable, that is, tenfold increase of $p_\perp$ leads to a modest increase (less than 25%) of $\frac{dN_g^{(1)}}{dx}$ in bsg compared to sg case. Note that the same dependence is obtained for $\frac{dE_{bsg}^{(1)}}{dx}/\frac{dE_{sg}^{(1)}}{dx} - 1$ (since $\frac{dE^{(1)}}{dx} \sim x\frac{dN_g^{(1)}}{dx}$, so that x cancels when taking the relative ratio). Therefore, we conclude that the relaxation of the soft-gluon approximation has nearly the same effect on $\frac{dN_g^{(1)}}{dx}$ and $\frac{dE^{(1)}}{dx}$ (across the whole x region) independently on $p_\perp$ of the initial jet.

Although we showed that relaxing the soft-gluon approximation has small numerical impact on both integrated ($\frac{\Delta E^{(1)}}{E}$, $N_g^{(1)}$, across the whole x region) and differential ($\frac{dE^{(1)}}{dx}$, $\frac{dN_g^{(1)}}{dx}$, up to $x \approx 0.4$) variables, the difference between bsg and sg cases can go up to 10% (and with different signs), and moreover can be quite large for $x > 0.4$. This, therefore, leads to a question, how the relaxation of the soft-gluon approximation affects predictions for measured observables, such as the angular averaged nuclear modification factor R_{AA} [18, 19]. Comparing R_{AA} with and without soft-gluon approximation allows assessing how adequate is this approximation in obtaining reliable numerical predictions.

To that end, we next concentrate on generating the predictions for bare gluon R_{AA}, based only on radiative energy loss, with and without soft-gluon approximation.

R_{AA} is defined as the ratio of the quenched $A + A$ spectrum to the $p + p$ spectrum, scaled by the number of binary collisions N_{bin}:

$$R_{AA}(p_\perp) = \frac{dN_{AA}/dp_\perp}{N_{bin}dN_{pp}/dp_\perp}. \qquad (22)$$

In order to obtain gluon quenched spectra, we use generic pQCD convolution [33]:

$$\frac{E_f d^3\sigma(g)}{dp_f^3} = \frac{E_i d^3\sigma(g)}{dp_i^3} \otimes P(E_i \to E_f), \qquad (23)$$

where $\frac{E_i d^3\sigma(g)}{dp_i^3}$ denotes initial gluon spectrum, which is computed according to [34, 35], while $P(E_i \to E_f)$ denotes radiative energy loss probability, which includes multi-gluon [29] and path-length [33] fluctuations. The path-length distributions for $0 - 5\%$ most central collisions are implemented as described in [19]. Note that we omitted fragmentation and decay functions, because we are considering the parton's quenching, as we are primarily interested in how the relaxation of the soft-gluon approximation in energy loss affects R_{AA}. Thereupon, we will also investigate how the initial gluon distribution influences R_{AA}.

Therefore, the left panel of Fig. 4 compares R_{AA} predictions with and without soft-gluon approximation accounted, while the percentage change arising from relaxing the approximation is given by the right panel of Fig. 4 as a function of the final $p_\perp$. We observe that this relaxation barely modifies R_{AA}, in particular the relative change drops to somewhat less than -1% at $p_\perp \approx 10$ GeV and further rises to the constant value of 2%, with increasing $p_\perp$. This very good agreement (with even smaller differences compared to previously studied variables) between bsg and sg R_{AA} raises questions of: $i)$ why relaxing the soft-gluon approximation has negligible effect on R_{AA} and $ii)$ why the large discrepancy observed in Figs.(2, 3) for high x values does not lead to larger difference in R_{AA}?

Regarding $i)$ above, we argue that this pattern is expected, as it is well-known that in suppression calculations both $\frac{\Delta E^{(1)}}{E}$ and $N_g^{(1)}$ non-trivially affect the R_{AA}. Namely, by comparing the two right panels of Fig. 1 with the right panel of Fig. 4 we observe that relaxing the soft-gluon approximation has opposite effects on $\frac{\Delta E^{(1)}}{E}$ and $N_g^{(1)}$, while their interplay is responsible for the negligible effect on R_{AA} - i.e. the effect on R_{AA} is qualitatively a superposition of the effects on $\frac{\Delta E^{(1)}}{E}$ and $N_g^{(1)}$.

To answer $ii)$ above, it is convenient to recall that suppression of gluon jet (see Eq. (23)) depends not only on the energy loss probability, but also on the initial gluon distribution. In order to intuitively interpret the role of the initial gluon distribution, we refer to a descriptive Fig. 5, which represents its dependence on initial transverse momentum. The concept considered is the following: Some parent gluon with unknown initial momentum traverses QGP, looses its energy by gluon bremsstrahlung, and emerges with final
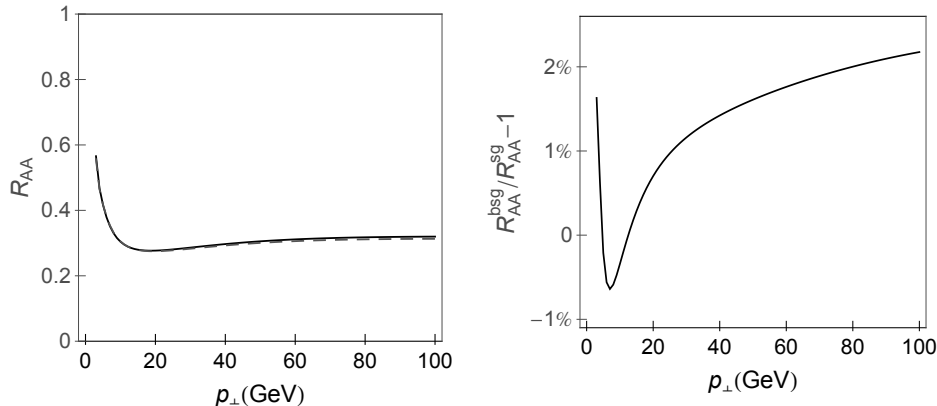
FIG. 4: The effect of relaxing the soft-gluon approximation on gluon nuclear modification factor R_{AA} versus $p_\perp$. The suppression of gluon jet beyond soft-gluon approximation (the solid curve) is compared to soft-gluon R_{AA} (the dashed curve) as a function of transverse momentum in the left panel. The right panel quantifies the effect and expresses it in percentage.
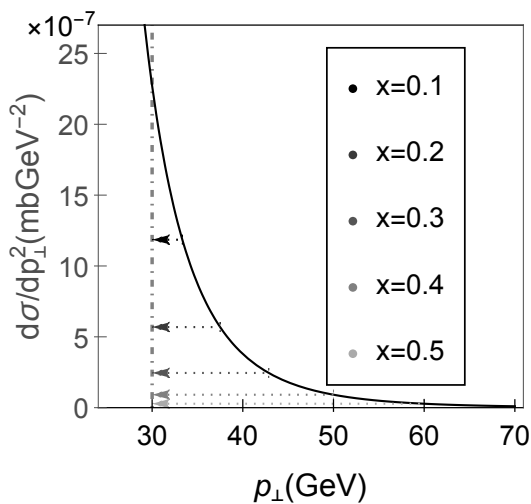

FIG. 5: The role of initial gluon distribution in constraining relevant x region. The solid black curve represents initial gluon distribution as a function of $p_\perp$ at the LHC [34, 35]. The dot-dashed gray line marks the final gluon transverse momentum, while dotted arrows link parent gluons, that lost momentum fraction equal to x, with their corresponding initial transverse momenta. The arrows fade as x increases, as indicated in the legend.

momentum $p_\perp = 30$ GeV. This final gluon can descend from the parent gluon with any $p_\perp$ higher than its own, but we restrict ourselves to 5 different initial momenta, corresponding to different fractional momentum loss $x \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$. For instance, $x = 0.5$ corresponds to initial gluon momentum of $30/0.5$ GeV$= 60$ GeV, i.e. to the parent gluon that lost half of its momentum etc. The question is which of these 5 gluons is the most likely to be the parent one, and how is this proba-

bility correlated with x? From Fig. 5 we infer that, due to the exponentially decreasing initial gluon momentum distribution, the initial gluon corresponding to $x = 0.1$ has the highest probability to be the parent one, and as x increases the probability sharply decreases (i.e. for $x \gtrsim 0.4$ it diminishes for 2 orders of magnitude compared to the $x = 0.1$ case). This makes only the region below the crossover noted above ($x \lesssim 0.3$) relevant for generating the predictions, and also explains why the large inconsistency between bsg and sg $\frac{dE^{(1)}}{dx}$ (or equivalently $\frac{dN_g^{(1)}}{dx}$) curves from Figs.(2, 3) at higher x does not affect R_{AA}. Additionally, the effect of relaxing the soft-gluon approximation on $\frac{dN_g^{(1)}}{dx}$ and $\frac{dE^{(1)}}{dx}$ is practically insensitive to initial transverse momentum (see Fig. 3), which is the reason why finite x affects equivalently gluon R_{AA} regardless of it's transverse momentum, as observed in Fig. 4.

Finally, we also recalculated our finite x results, when running coupling $\alpha_s(Q^2)$, as defined in [36], instead of constant value $\alpha_s = 0.3$, is introduced in radiative energy loss formula. The obtained predictions lead to the same conclusions as obtained above (and are consequently omitted), which supports the generality of the obtained results.

VII. CONCLUSIONS AND OUTLOOK

The main theoretical goal of this paper was to investigate what effect relaxing of the soft-gluon approximation has on radiative energy loss, and consequently on suppression, which depends only on initial distribution and energy loss of high-momentum parton in QGP. Particularly we chose high $p_\perp$ gluon, as due to the color factor of $9/4$ compared to the quarks, this assumption affects gluons the most. To this end, we analytically calculated all Feynman diagrams contributing to the first order in opac-

ity radiative energy loss beyond soft-gluon approximation, first within GLV [9] (massless case), and later within DGLV [25] (massive case), formalism, and numerically predicted: fractional and differential energy loss, number of radiated gluons, single gluon radiation spectrum and gluon's suppression. Unexpectedly we obtained that, although the analytic results significantly differ from the corresponding soft-gluon results, the numerical predictions are nearly indistinguishable, i.e. within few percents. We then explained that, due to exponentially decreasing initial gluon distribution, only $x \lesssim 0.3$ region effectively contributes to the integrated variable predictions. We also showed that negligible suppression change is due to an interplay between the finite x effects on *i)* fractional energy loss and *ii)* number of radiated gluons, that have opposite sign. The presented comparisons are done under the assumption of fixed strong coupling constant, but also tested with running coupling leading to the same conclusions. Since we showed that gluon quenching in QCD medium composed of static scattering centers is not affected by the soft-gluon assumption, quark radiative energy loss is even less likely to be notably altered, though this remains to be further tested.

This, to our knowledge, presents the first opportunity to assess the effect of relaxing the soft-gluon approximation on radiative energy loss within DGLV formalism. Some other radiative energy loss formalisms, which also imply static scatterers, generated their results on a finite x. However, contrary to the conclusions derived for these formalisms, where significant difference in the radiative energy loss was obtained, we found that relaxing soft-gluon approximation brings negligible change to the results. Consequently, this implies that, within DGLV formalism, there is no need to go beyond the soft-gluon approximation.

Based on the results of this paper, we also expect that the soft-gluon approximation can be reliably applied to the dynamical energy loss formalism, as implicitly suggested by the previous robust agreement [17, 20–22] of our theoretical predictions with a comprehensive set of experimental data. In particular, the effective cross section $v(\mathbf{q})$ (which corresponds to interaction between the jet and exchanged gluon) [37] does not depend on x, so introduction of finite x will not affect this term. We also expect that the rest of the energy loss expression (i.e. $f(\mathbf{k}, \mathbf{q}, x)$, which corresponds to interaction between the jet and radiated gluon [37]) will be modified in the similar manner as in the static case, since when $x \to 0$, these two expression coincide. However, relaxing the soft-gluon approximation in dynamical energy loss model is out of the scope of this paper, and this claim still remains to be rigorously tested in the future.

Appendix A: NOTATIONS AND USEFUL FORMULAS

In this paper we used the following notation for vectors, in consistency with both [9, 25]:

- $\vec{\mathbf{p}}$ denotes momentum 3D vector

- $\mathbf{p}$ denotes transverse momentum 2D vector

- p_z denotes component of momentum vector along the initial jet

- $p = (p^0, p_z, \mathbf{p}) = [p^+, p^-, \mathbf{p}]$ denotes momentum 4D vector in Minkowski and Light Cone coordinates, respectively, where $p^+ = p^0 + p_z$ and $p^- = p^0 - p^z$.

For simplicity, we here consider QCD medium consisting of static partons and model the interactions of the gluon jet with the medium via static color-screened Yukawa potential, whose Fourier and color structure acquires the following form ([9, 26]):

$$V_n = V(q_n)e^{iq_nx_n} = 2\pi\delta(q_n^0)v(\vec{\mathbf{q}}_n)e^{-i\vec{\mathbf{q}}_n\vec{\mathbf{x}}_n} \times$$
$$\times T_{a_n}(R) \otimes T_{a_n}(n), \quad (A1)$$

$$v(\vec{\mathbf{q}}_n) = \frac{4\pi\alpha_s}{\vec{\mathbf{q}}_n^2 + \mu^2}, \quad (A2)$$

where x_n denotes space-time coordinate of the nth scatterer (target), $T_{a_n}(R)$ and $T_{a_n}(n)$ denote generators in $SU(N_c = 3)$ color representation of jet and target, respectively, while μ is Debye screening mass and $\alpha_s = g_s^2/4\pi$ is strong coupling constant. In the following lines we will briefly display the identities and algebra that $SU(N_c = 3)$ generators meet:

$$\text{Tr}(T^a(n)) = 0, \quad (A3)$$

$$\text{Tr}(T^a(i)T^b(j)) = \delta_{ij}\delta^{ab}\frac{C_2(i)d_i}{d_G}, \quad (A4)$$

where $d_G = 8$ is the dimension of the adjoint representation (G). We assume that all target partons are in the same d_T dimensional representation (T) with Casimir operator $C_2(T)$, while the gluon jet is in the adjoint representation (G), with Casimir operator $C_2(G)$.

In $SU(N_c = 3)$ color algebra, the following identities hold as well:

$$[T^a, T^b] = if^{abc}T^c, \quad (A5)$$

while in the adjoint representation we have:

$$(T^b)_{ab} = if^{abc}, \quad (A6)$$

$$T^a(G)T^a(G) = C_2(G)I, \quad (A7)$$

where I denotes identity matrix of dimension d_G and the $SU(N_c = 3)$ structure constants f^{abc} are completely antisymmetric to indices permutations, which we frequently apply. In the adjoint representation the following equalities also stand:

$$C(G) = C_2(G) = N_c = 3, \tag{A8}$$
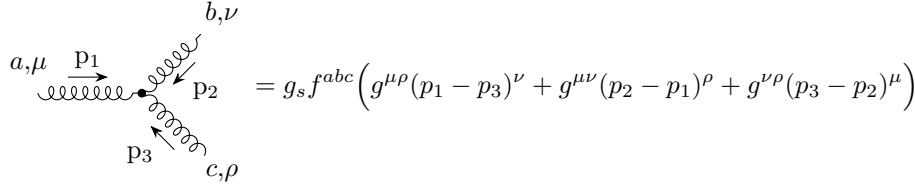
$$\text{Tr}(T^a(G)T^a(G)) = d_G C_2(G). \tag{A9}$$

And finally, in our computations we frequently make use of the fact that trace is invariant under cyclic permutations and that generators are Hermitian matrices.

Since our extensive calculations are done in pQCD at finite temperature and include only gluon interactions, below we list the necessary Feynman rules in covariant gauge that we employ:

- massless gluon propagator in Feynman gauge:

$$a,\mu \xrightarrow{\quad p \quad} b,\nu \quad = \frac{-i\delta_{ab}g_{\mu\nu}}{p^2 + i\epsilon}, \tag{A10}$$

- 3-gluon vertex:

$$= g_s f^{abc}\left(g^{\mu\rho}(p_1 - p_3)^\nu + g^{\mu\nu}(p_2 - p_1)^\rho + g^{\nu\rho}(p_3 - p_2)^\mu\right). \tag{A11}$$

Note that, all diagrams in this paper are obtained by using [38]. Since only physical transverse gluon states must be accounted, summing is done according to Eq. (57) from [31]:

$$\sum_{pol} \epsilon^i(k)\epsilon^j(k) = \delta^{ij} - \frac{k^i k^j}{\vec{\mathbf{k}}^2}, \tag{A12}$$

where $i,j = 1,2,3$.

Appendix B: ASSUMPTIONS

Throughout the paper we assume that initial gluon jet propagates along the z-axis, i.e. has transverse momentum equal to zero, while radiated gluon carries away a finite rate x of initial gluon longitudinal momentum and energy, and final gluon emerges with momentum p. Therefore, instead of assuming soft-gluon approximation ($x \ll 1$), as it was done in [9, 25], we allow x to acquire finite non-zero values, thus relaxing the soft-gluon approximation.

Since we are calculating radiative energy loss within the (GLV) DGLV formalism apart from abandoning the soft-gluon approximation, the following assumptions remain:

- *The soft-rescattering approximation.* Consistently with [9, 25] we assume that partons energies and longitudinal momenta are high compare to their transverse momenta, which disables the radiated

and the final gluon to digress much from the initial longitudinal direction (the eikonal approximation).

$$E^+ \sim (1-x)E^+ \sim xE^+ \gg |\mathbf{p}|, |\mathbf{k}|, |\mathbf{q}_i|, \tag{B1}$$

- *The first order approximation.* The gluon-jet radiative energy loss is calculated up to the first order in opacity expansion, as argued in [9, 29, 30].

- *Scattering centers distribution and ensemble average.* We consider that all scattering centers x_i are distributed with the same transversely homogeneous density:

$$\rho(\vec{\mathbf{x}}) = \frac{N}{A_\perp}\bar{\rho}(z), \tag{B2}$$

where $\int dz\bar{\rho}(z) = 1$ and also that impact parameter (i.e. relative transverse coordinate) $\mathbf{b} = \mathbf{x}_i - \mathbf{x}_0$ alters within a large transverse area $A_\perp$ compared to the interaction area $\frac{1}{\mu^2}$. Therefore, the ensemble average over the scattering center locations reduces to an impact parameter average:

$$\langle ... \rangle = \int \frac{d^2\mathbf{b}}{A_\perp} ..., \tag{B3}$$

which in our case is mainly used in the following form:

$$\left\langle e^{-i(\mathbf{q}_i + \mathbf{q}_j)\mathbf{b}} \right\rangle = \frac{(2\pi)^2}{A_\perp}\delta^2(\mathbf{q}_i + \mathbf{q}_j). \tag{B4}$$

We also assume that the energy of initial hard probe is large compared to the potential screening scale:

$$E^+, (1-x)E^+, xE^+ \gg \mu. \tag{B5}$$

Next, we assume that the distance between the source J and the scattering centers is large relative to the inter-action length:

$$z_i - z_0 \gg \frac{1}{\mu}, \tag{B6}$$

then, that source current varies slowly with momentum:

$$J(p+k-q) \approx J(p+k), \tag{B7}$$

and that the source current can be written explicitly in terms of polarization vector:

$$J_a^\mu(p+k-q) \equiv J_a(p+k-q)\epsilon^\mu(p+k-q)$$
$$\approx J_a(p+k)\epsilon^\mu(p+k-q). \tag{B8}$$

In the following sections first we assume that gluons are massless (GLV) in order to make the comprehensive derivations more straightforward and easier to follow, but later we recalculate all the results with gluon mass [27] included (DGLV) (Appendix J).

Appendix C: Gluon jet M_0

First we calculate gluon-jet radiation amplitude to emit a gluon, carrying a finite fraction x of initial jet energy, with momentum, polarization and color (k, $\epsilon(k)$, c) and without interactions with the medium M_0.

We assume that initial gluon $(p+k)$ propagates along z-axis. By using M_0 amplitude as an example, we will implement the aforementioned assumptions in order to acquire momentum and polarization expressions. Thus, the initial gluon 4-momentum reads:

$$p+k = (p^0+k^0, p_z+k_z, \mathbf{0}),$$
$$p+k = [E^+, E^-, \mathbf{0}], \tag{C1}$$

where $E^+ = p^0 + k^0 + p_z + k_z$ and $E^- = p^0 + k^0 - p_z - k_z$.

Assuming massless (real) gluons for simplicity, the momentum vectors of the radiated (k) and the final (p) gluons acquire the following form:

$$k^2 = 0 \Rightarrow k = [xE^+, \frac{\mathbf{k}^2}{xE^+}, \mathbf{k}], \tag{C2}$$

$$p^2 = 0 \Rightarrow p = [(1-x)E^+, \frac{\mathbf{p}^2}{(1-x)E^+}, \mathbf{p}]. \tag{C3}$$

We also assume that gluons are transversely polarized particles. Although we work in covariant gauge, we can choose any polarization vector for the external on-shell

$$\mathrm{M_0}$$
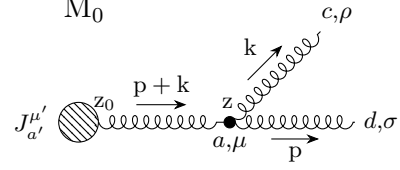
FIG. 6: Zeroth order diagram that includes no interaction with the QCD medium, and contributes to gluon radiation amplitude to the first order in opacity L/λ. The dashed circle represents the source J, which at longitudinal coordinate z_0 produces an off-shell gluon jet, propagating along z-axis. z denotes longitudinal coordinate at which the gluon is radiated. Latin indices denote color charges, while Greek indices denote components of 4-vectors. k denotes 4-momentum of the radiated gluon carrying the color c, and p denotes 4-momentum of the final gluon jet carrying the color d.

gluons, so in accordance with [9, 14, 25] we choose $n^\mu = [0, 2, \mathbf{0}]$, as stated above:

$$\epsilon(k) \cdot k = 0, \qquad \epsilon(k) \cdot n = 0, \qquad \epsilon(k)^2 = -1,$$
$$\epsilon(p) \cdot p = 0, \qquad \epsilon(p) \cdot n = 0, \qquad \epsilon(p)^2 = -1, \tag{C4}$$

while we assume that the source has also the physical polarizations as the real gluons [14]:

$$\epsilon(p+k) \cdot (p+k) = 0, \qquad \epsilon(p+k) \cdot n = 0,$$
$$\epsilon(p+k)^2 = -1. \tag{C5}$$

Using Eqs. (C2- C5) we can now obtain the following expressions for the gluon polarizations:

$$\epsilon_i(k) = [0, \frac{2\boldsymbol{\epsilon}_i \cdot \mathbf{k}}{xE^+}, \boldsymbol{\epsilon}_i], \qquad \epsilon_i(p) = [0, \frac{2\boldsymbol{\epsilon}_i \cdot \mathbf{p}}{(1-x)E^+}, \boldsymbol{\epsilon}_i],$$
$$\epsilon_i(p+k) = [0, 0, \boldsymbol{\epsilon}_i], \tag{C6}$$

where $i = 1, 2$ counts for polarization vectors. Note that the 4-momentum is conserved, which leads to the relation:

$$\mathbf{p} + \mathbf{k} = 0, \tag{C7}$$

that we implement in Eqs. (C3, C6) in order to ensure that everything is expressed in terms of $\mathbf{k}$. Also, $E^+ \approx 2E$, $E^- = \frac{\mathbf{k}^2}{x(1-x)E^+}$, where $E = p^0 + k^0$ is the energy of initial jet.

Using the notation from Fig. 6 we obtain:

$$M_0 = \epsilon_\sigma^*(p)\epsilon_\rho^*(k)g_s f^{acd}\Big(g^{\mu\sigma}(2p+k)^\rho + g^{\mu\rho}(-p-2k)^\sigma + g^{\rho\sigma}(-p+k)^\mu\Big)\frac{-i\delta_{aa'}g_{\mu\mu'}}{(p+k)^2+i\epsilon}iJ_{a'}(p+k)e^{i(p+k)x_0}\times$$

$$\times \epsilon^{\mu'}(p+k) \approx J_a(p+k)e^{i(p+k)x_0}(-2g_s)(1-x+x^2)\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2}f^{acd}$$

$$= J_a(p+k)e^{i(p+k)x_0}(-2ig_s)(1-x+x^2)\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2}(T^c)_{da}. \tag{C8}$$

Eq. (C8) after summation by using Eq. (A12) gives:

$$\langle|M_0|^2\rangle = |J(p+k)|^2(4g_s^2)\frac{C_2(G)d_G}{\mathbf{k}^2}(1-x+x^2)^2. \tag{C9}$$

Next we substitute the Eq. (C9) in:

$$d^3N_g^{(0)}d^3N_J \approx \text{Tr}\langle|M_0|^2\rangle\frac{d^3\vec{\mathbf{p}}}{(2\pi)^3 2p^0}\frac{d^3\vec{\mathbf{k}}}{(2\pi)^3 2\omega}, \tag{C10}$$

Note that, contrary to the soft-gluon approximation [25], where:

$$d^3N_J \approx d_G|J(p+k)|^2\frac{d^3\vec{\mathbf{p}}}{(2\pi)^3 2p^0}, \tag{C11}$$

now p, denoting the momentum of the final gluon jet, is not approximately equal to the momentum of initial gluon jet (i.e. the radiated gluon can carry away a substantial amount of the initial jet energy and longitudinal momentum). Thus instead of Eq. (C11) throughout this paper we use the general one:

$$d^3N_J = d_G|J(p+k)|^2\frac{d^3\vec{\mathbf{p}}_J}{(2\pi)^3 2E_J}, \tag{C12}$$

where $E_J = E$ and $\vec{\mathbf{p}}_J$ denotes energy and 3D momentum of the initial gluon jet, respectively. Knowing that the substitution of variables $(p_z, k_z \to p_z^J, xE)$ gives:

$$\frac{d^3\vec{\mathbf{p}}}{(2\pi)^3 2p^0}\frac{d^3\vec{\mathbf{k}}}{(2\pi)^3 2\omega} = \frac{d^3\vec{\mathbf{p}}_J}{(2\pi)^3 2E_J}\frac{dxd^2\mathbf{k}}{(2\pi)^3 2x(1-x)}, \tag{C13}$$

and by substituting Eqs. (C9, C12, C13) in Eq. (C10), for radiation spectrum we now obtain:

$$\frac{xd^3N_g^{(0)}}{dxd\mathbf{k}^2} = \frac{\alpha_s}{\pi}\frac{C_2(G)}{\mathbf{k}^2}\frac{(1-x+x^2)^2}{1-x}, \tag{C14}$$

which recovers well-known Altarelli-Parisi result [31] and for $x \ll 1$ reduces to the massless soft-gluon limit

of Eq. (9) from [25]. We have also checked that, the same result can be obtained by directly implementing polarization vectors (Eq. (C6)) in Eq. (C8), instead of using Eq. (A12) when averaging.

Appendix D: Diagrams $M_{1,1,0}$, $M_{1,0,0}$, $M_{1,0,1}$

In this section we provide a detailed calculations of Feynman amplitudes, corresponding to gluon-jet interaction with one scattering center, which are depicted in Fig. 7. Again for consistency, we assume that initial jet $(p+k-q)$ propagates along z-axis. Throughout this section momentum and polarization vector for initial gluon read:

$$p+k-q_1 = [E^+ - q_{1z}, E^- + q_{1z}, \mathbf{0}], \tag{D1}$$

$$\epsilon_i(p+k-q_1) = [0, 0, \boldsymbol{\epsilon}_i], \tag{D2}$$

where $q_1 = [q_{1z}, -q_{1z}, \mathbf{q}_1]$, with $q_1^0 = 0$, denotes the momentum of exchanged gluon, while p, k and corresponding polarization vectors retain the same expression as in Eqs. (C2, C3, C6), with the distinction that the following relation between gluon transverse momenta, due to 4-momentum conservation, holds:

$$\mathbf{q}_1 = \mathbf{p} + \mathbf{k}. \tag{D3}$$

1. Computation of $M_{1,1,0}$ diagram

We chose to start with thorough derivation of the expression for $M_{1,1,0}$ amplitude, simply because it has no counterpart regarding the symmetry under $(p \leftrightarrow k, x \leftrightarrow (1-x), c \leftrightarrow d)$ substitutions, and it provides all necessary steps for calculating the remaining two amplitudes from this chapter, apart from having one less singularity compared to the amplitudes $M_{1,0,0}$ and $M_{1,0,1}$. Thus, using the notation from the left diagram of Fig. 7, we write:
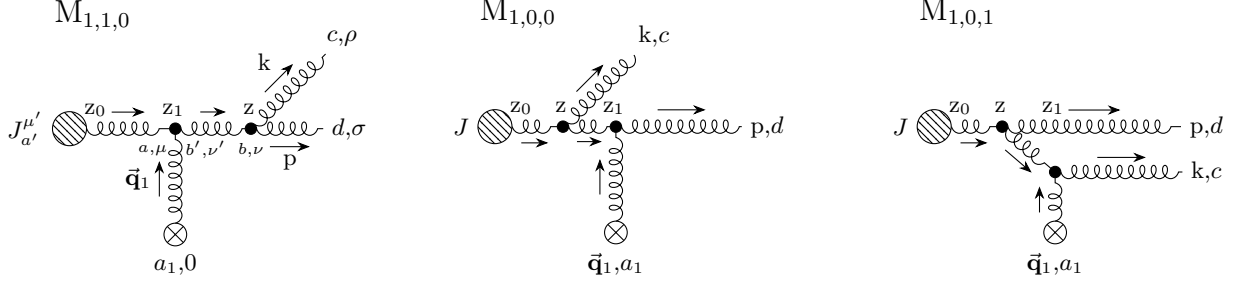
FIG. 7: Three diagrams, corresponding to interaction with one static scattering center, that contribute to gluon-jet radiation amplitude to the first order in opacity L/λ. z_1 denotes longitudinal coordinate of the interactions with one scattering center. Crossed circle represents scatterer that exchanges 3D momentum $\vec{q}_1$ with the jet. Note that, all three diagrams assume equivalently ordered Latin and Greek indices as indicated by the first diagram. Remaining labeling is the same as in Fig. 6.

$$
\begin{aligned}
M_{1,1,0} =& \int \frac{d^4 q_1}{(2\pi)^4} \epsilon_\sigma^*(p)\epsilon_\rho^*(k) g_s f^{bcd}\Big(g^{\nu\sigma}(2p+k)^\rho + g^{\nu\rho}(-p-2k)^\sigma + g^{\rho\sigma}(-p+k)^\nu\Big)\frac{(-i)\delta_{bb'}g_{\nu\nu'}}{(p+k)^2+i\epsilon} \times \\
& \times f^{ab'a_1}\Big(g^{\mu 0}(p+k-2q_1)^{\nu'} + g^{\mu\nu'}(-2p-2k+q_1)^0 + g^{\nu'0}(p+k+q_1)^\mu\Big)T_{a_1}V(q_1)e^{iq_1 x_1} \times \\
& \times \frac{(-i)\delta_{aa'}g_{\mu\mu'}}{(p+k-q_1)^2+i\epsilon}iJ_{a'}(p+k-q_1)\epsilon^{\mu'}(p+k-q_1)e^{i(p+k-q_1)x_0} \\
\approx& J_a(p+k)e^{i(p+k)x_0}f^{bcd}f^{a_1ab}T_{a_1}(-i)(1-x+x^2)\int \frac{d^2\mathbf{q}_1}{(2\pi)^2}e^{-i\mathbf{q}_1\cdot(\mathbf{x}_1-\mathbf{x}_0)}(2g_s)\frac{(1-x)\,\boldsymbol{\epsilon}\cdot\mathbf{k}-x\,\boldsymbol{\epsilon}\cdot\mathbf{p}}{x(1-x)} \times \\
& \times E^+ \int \frac{dq_{1z}}{2\pi}\frac{v(q_{1z},\mathbf{q}_1)e^{-iq_{1z}(z_1-z_0)}}{((p+k-q_1)^2+i\epsilon)((p+k)^2+i\epsilon)},
\end{aligned}
\tag{D4}
$$

where we used the equation:

$$
(p+k)^2 = \frac{((1-x)\mathbf{k}-x\mathbf{p})^2}{x(1-x)},
\tag{D5}
$$

and assumed that J varies slowly with momentum q_1, i.e. Eq. (B7). The longitudinal momentum transfer integral:

$$
I_1(p,k,\mathbf{q}_1,z_1-z_0) \equiv \int \frac{dq_{1z}}{2\pi}\frac{v(q_{1z},\mathbf{q}_1)e^{-iq_{1z}(z_1-z_0)}}{(p+k-q_1)^2+i\epsilon}
\tag{D6}
$$

has to be performed in the lower half-plane of the complex plain, since $z_1 > z_0$. In order to determine the pole arising from potential, we rewrite Eq. (A2) in a more appropriate form:

$$
v(\vec{q}_n) = \frac{4\pi\alpha_s}{(q_{nz}+i\mu_{n\perp})(q_{nz}-i\mu_{n\perp})},
\tag{D7}
$$

where $\mu_{n\perp}^2 = \mu^2 + \mathbf{q}_n^2$, with n denoting the corresponding scattering center.

Aside from the pole originating from Eq. (D7) ($q_{1z} = -i\mu_{1\perp}$), there is also a singularity emerging from the gluon propagator:

$$
\begin{aligned}
\bar{q}_1 =& -\frac{\mathbf{k}^2}{xE^+} - \frac{\mathbf{p}^2}{(1-x)E^+} - i\epsilon \\
=& -\frac{\mathbf{k}^2}{2\omega} - \frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1)^2}{2\omega} - i\epsilon,
\end{aligned}
\tag{D8}
$$

where $\omega = k_0 \approx \frac{xE^+}{2}$. The residue around the pole at $\bar{q}_1$ is computed as (the negative sign is due to the clock-wise orientation of the closed contour in the complex plain):

$$
\begin{aligned}
Res(\bar{q}_1) \approx& -v(-\frac{\mathbf{k}^2}{xE^+} - \frac{\mathbf{p}^2}{(1-x)E^+},\mathbf{q}_1)\frac{i}{E^+}e^{i(\frac{\mathbf{k}^2}{xE^+}+\frac{\mathbf{p}^2}{(1-x)E^+})(z_1-z_0)} \\
=& -v(-\frac{\mathbf{k}^2}{2\omega} - \frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1)^2}{2\omega},\mathbf{q}_1)\frac{i}{E^+}e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1)^2)(z_1-z_0)}.
\end{aligned}
\tag{D9}
$$

The pole originating from the potential ($q_{1z} = -i\mu_{1\perp}$) does not contribute to the longitudinal integral, since residue around that pole is exponentially suppressed due to Eq. (B6), i.e. $\mu(z_1 - z_0) = \mu\lambda \gg 1$ (and $\mu \sim \mu_{1\perp}$):

$$Res(-i\mu_{1\perp}) \approx -i\frac{4\pi\alpha_s}{(-2i\mu_{1\perp})E^+(-i\mu_{1\perp})}e^{-\mu_{1\perp}(z_1-z_0)} \to 0, \tag{D10}$$

where we assumed that $E^+ \gg \mu$ and soft-rescattering approximation.

This makes only $\bar{q}_1$ singularity relevant for calculating longitudinal integral. Therefore I_1 coincides with Eq. (D9), i.e.:

$$I_1(p,k,\mathbf{q}_1,z_1-z_0) \approx -v(-\frac{\mathbf{k}^2}{xE^+} - \frac{\mathbf{p}^2}{(1-x)E^+}, \mathbf{q}_1)\frac{i}{E^+}e^{i(\frac{\mathbf{k}^2}{xE^+} + \frac{\mathbf{p}^2}{(1-x)E^+})(z_1-z_0)} \approx -v(0,\mathbf{q}_1)\frac{i}{E^+}e^{i(\frac{\mathbf{k}^2}{xE^+} + \frac{\mathbf{p}^2}{(1-x)E^+})(z_1-z_0)}$$

$$= -v(0,\mathbf{q}_1)\frac{i}{E^+}e^{\frac{i}{2\omega}(\mathbf{k}^2 + \frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1)^2)(z_1-z_0)}, \tag{D11}$$

where we used eikonal approximation (i.e. for a finite x: $\frac{\mathbf{k}^2}{(xE^+)^2} \ll 1$ and $\frac{\mathbf{p}^2}{((1-x)E^+)^2} \ll 1$). Finally, $M_{1,1,0}$ amplitude reads:

$$M_{1,1,0} = J_a(p+k)e^{i(p+k)x_0}(-i)(1-x+x^2)f^{bcd}f^{a_1ab}T_{a_1}\int\frac{d^2\mathbf{q}_1}{(2\pi)^2}v(0,\mathbf{q}_1)e^{-i\mathbf{q}_1\cdot\mathbf{b}_1}(-2ig_s)\frac{\boldsymbol{\epsilon}\cdot((1-x)\mathbf{k}-x\mathbf{p})}{((1-x)\mathbf{k}-x\mathbf{p})^2} \times$$

$$\times e^{i(\frac{\mathbf{k}^2}{xE^+} + \frac{\mathbf{p}^2}{(1-x)E^+})(z_1-z_0)}$$

$$= J_a(p+k)e^{i(p+k)x_0}(-i)(1-x+x^2)(T^cT^{a_1})_{da}T_{a_1}\int\frac{d^2\mathbf{q}_1}{(2\pi)^2}v(0,\mathbf{q}_1)e^{-i\mathbf{q}_1\cdot\mathbf{b}_1}(-2ig_s)\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2} \times$$

$$\times e^{\frac{i}{2\omega}(\mathbf{k}^2 + \frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1)^2)(z_1-z_0)}, \tag{D12}$$

where we denoted $\mathbf{b}_1 \equiv \mathbf{x}_1 - \mathbf{x}_0$. In this subsection, we constantly make use of Eq. (D3) in the following form:

$$\mathbf{p}^2 = (\mathbf{k} - \mathbf{q}_1)^2, \tag{D13}$$

and also manipulate with $SU(N_c = 3)$ structure constants by using Eqs. (A5, A6). Note from Fig. 7 that, as expected, $M_{1,1,0}$ is symmetric under the substitutions: $(p \leftrightarrow k, x \leftrightarrow (1-x), c \leftrightarrow d)$, where the symmetry can be straightforwardly verified by implementing these substitutions in the first two lines of Eq. (D12).

2. Computation of $M_{1,0,0}$ and $M_{1,0,1}$ diagrams

Applying the same procedure as in the previous subsection, we proceed with calculating $M_{1,0,0}$. Note that the order of the color and Dirac indices denoting vertices is the same for all three diagrams in Fig. 7, and are therefore omitted in the last two diagrams.

$$M_{1,0,0} = \int\frac{d^4q_1}{(2\pi)^4}\epsilon_\sigma^*(p)f^{bda_1}\Big(g^{\nu 0}(p-2q_1)^\sigma + g^{\nu\sigma}(-2p+q_1)^0 + g^{\sigma 0}(p+q_1)^\nu\Big)T_{a_1}V(q_1)e^{iq_1x_1}\frac{(-i)\delta_{bb'}g_{\nu\nu'}}{(p-q_1)^2+i\epsilon} \times$$

$$\times g_sf^{acb'}\Big(g^{\mu\nu'}(2p+k-2q_1)^\rho + g^{\mu\rho}(-p-2k+q_1)^{\nu'} + g^{\rho\nu'}(-p+k+q_1)^\mu\Big)\epsilon_\rho^*(k)\frac{(-i)\delta_{aa'}g_{\mu\mu'}}{(p+k-q_1)^2+i\epsilon} \times$$

$$\times iJ_{a'}(p+k-q_1)\epsilon^{\mu'}(p+k-q_1)e^{i(p+k-q_1)x_0}$$

$$\approx J_a(p+k)e^{i(p+k)x_0}f^{bda_1}f^{acb}T_{a_1}(-i)(1-x+x^2)E^+\int\frac{d^2\mathbf{q}_1}{(2\pi)^2}e^{-i\mathbf{q}_1\cdot\mathbf{b}_1}(2g_s)\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{x}I_2, \tag{D14}$$

where:

$$I_2(p,k,\mathbf{q}_1,z_1-z_0) \equiv \int \frac{dq_{1z}}{2\pi} \frac{v(q_{1z},\mathbf{q}_1)e^{-iq_{1z}(z_1-z_0)}}{(p+k-q_1)^2+i\epsilon} \times$$
$$\times \frac{1}{(p-q_1)^2+i\epsilon}. \tag{D15}$$

In order to calculate the previous integral, due to $z_1 > z_0$

we again have to close the contour below the real axis. Similarly as in $M_{1,1,0}$ amplitude, again only the poles originating from the propagators contribute to the integral: $(-\frac{\mathbf{k}^2}{xE^+} - \frac{\mathbf{p}^2}{(1-x)E^+} - i\epsilon)$ and $(\frac{\mathbf{k}^2-\mathbf{p}^2}{(1-x)E^+} - i\epsilon)$, while $(-i\mu_{1\perp})$ is exponentially suppressed (due to $\mu(z_1-z_0) \gg 1$). Therefore we obtain:

$$I_2(p,k,\mathbf{q}_1,z_1-z_0) \approx \frac{ix}{E^+\mathbf{k}^2}v(0,\mathbf{q}_1)\Big(e^{i(\frac{\mathbf{k}^2}{xE^+}+\frac{\mathbf{p}^2}{(1-x)E^+})(z_1-z_0)} - e^{i\frac{(\mathbf{p}^2-\mathbf{k}^2)}{(1-x)E^+}(z_1-z_0)}\Big)$$
$$\approx \frac{ix}{E^+\mathbf{k}^2}v(0,\mathbf{q}_1)\Big(e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1)^2)(z_1-z_0)} - e^{\frac{i}{2\omega}\frac{x}{1-x}((\mathbf{k}-\mathbf{q}_1)^2-\mathbf{k}^2)(z_1-z_0)}\Big), \tag{D16}$$

leading to:

$$M_{1,0,0} = J_a(p+k)e^{i(p+k)x_0}(-i)(1-x+x^2)f^{bda_1}f^{acb}T_{a_1}\int \frac{d^2\mathbf{q}_1}{(2\pi)^2}v(0,\mathbf{q}_1)e^{-i\mathbf{q}_1\cdot\mathbf{b}_1}(2ig_s)\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2}\times$$
$$\times\Big(e^{i(\frac{\mathbf{k}^2}{xE^+}+\frac{\mathbf{p}^2}{(1-x)E^+})(z_1-z_0)} - e^{i\frac{(\mathbf{p}^2-\mathbf{k}^2)}{(1-x)E^+}(z_1-z_0)}\Big)$$
$$= J_a(p+k)e^{i(p+k)x_0}(-i)(1-x+x^2)(T^{a_1}T^c)_{da}T_{a_1}\int \frac{d^2\mathbf{q}_1}{(2\pi)^2}v(0,\mathbf{q}_1)e^{-i\mathbf{q}_1\cdot\mathbf{b}_1}(2ig_s)\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2}\times$$
$$\times\Big(e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1)^2)(z_1-z_0)} - e^{-\frac{i}{2\omega}\frac{x}{1-x}(\mathbf{k}^2-(\mathbf{k}-\mathbf{q}_1)^2)(z_1-z_0)}\Big). \tag{D17}$$

By applying similar procedure for $M_{1,0,1}$ we obtain:

$$M_{1,0,1} = J_a(p+k)e^{i(p+k)x_0}(-i)(1-x+x^2)f^{bca_1}f^{adb}T_{a_1}\int \frac{d^2\mathbf{q}_1}{(2\pi)^2}v(0,\mathbf{q}_1)e^{-i\mathbf{q}_1\cdot\mathbf{b}_1}(2ig_s)\frac{\boldsymbol{\epsilon}\cdot\mathbf{p}}{\mathbf{p}^2}\times$$
$$\times\Big(e^{i(\frac{\mathbf{k}^2}{xE^+}+\frac{\mathbf{p}^2}{(1-x)E^+})(z_1-z_0)} - e^{i\frac{(\mathbf{k}^2-\mathbf{p}^2)}{xE^+}(z_1-z_0)}\Big)$$
$$= J_a(p+k)e^{i(p+k)x_0}(-i)(1-x+x^2)[T^c,T^{a_1}]_{da}T_{a_1}\int \frac{d^2\mathbf{q}_1}{(2\pi)^2}v(0,\mathbf{q}_1)e^{-i\mathbf{q}_1\cdot\mathbf{b}_1}(2ig_s)\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2}\times$$
$$\times\Big(e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1)^2)(z_1-z_0)} - e^{\frac{i}{2\omega}(\mathbf{k}^2-(\mathbf{k}-\mathbf{q}_1)^2)(z_1-z_0)}\Big). \tag{D18}$$

Notice from Fig. 7 that $M_{1,0,1}$ and $M_{1,0,0}$ are symmetric under the following substitutions: $(p \leftrightarrow k, x \leftrightarrow (1-x), c \leftrightarrow d)$; it can be straightforwardly verified that Eqs. (D17, D18) are symmetric under these substitutions.

Appendix E: Diagram $M_{2,2,0}$

Next we concentrate on the diagrams containing two interactions with the static scattering centers, since they also contribute to the gluon radiative energy loss to the first order in opacity, when multiplied by M_0^*. There are seven such diagrams, that we gather into four groups,

each of which contains two (or one) diagrams symmetric under $(p \leftrightarrow k, x \leftrightarrow (1-x), c \leftrightarrow d)$ substitutions.

For consistency the initial gluon jet (with momentum $p + k - q_1 - q_2$) propagates along z-axis, i.e.:

$$p + k - q_1 - q_2 = [E^+ - q_{1z} - q_{2z}, E^- + q_{1z} + q_{2z}, \mathbf{0}], \tag{E1}$$

$$\epsilon_i(p + k - q_1 - q_2) = [0, 0, \boldsymbol{\epsilon}_i], \tag{E2}$$

where $q_i = [q_{iz}, -q_{iz}, \mathbf{q}_i]$, $i = 1, 2$ with $q_i^0 = 0$ denote momenta of exchanged gluons, while p, k and corresponding polarizations retain the same expressions as in Eqs. (C2, C3, C6), with distinction that, due to 4-momentum conservation, the following relation between

gluon transverse momenta holds:

$$\mathbf{p} + \mathbf{k} = \mathbf{q}_1 + \mathbf{q}_2. \tag{E3}$$

Again, from seven diagrams we chose one model diagram $M_{2,2,0}$, based on the same reason as in Appendix D, for thorough derivation of the final amplitude expression. From Fig. 8, where gluon jet after two consecutive interactions with scattering centers radiates a gluon with momentum k, we observe that there are two limiting cases that we consider.

Using the notation from Fig. 8 we write:

$$
\begin{aligned}
M_{2,2,0} &= \int \frac{d^4 q_1}{(2\pi)^4} \frac{d^4 q_2}{(2\pi)^4} \epsilon_\sigma^*(p) \epsilon_\rho^*(k) g_s f^{ecd} \Big(g^{\xi\sigma}(2p+k)^\rho + g^{\xi\rho}(-p-2k)^\sigma + g^{\rho\sigma}(-p+k)^\xi \Big) \frac{-i\delta_{ee'} g_{\xi\xi'}}{(p+k)^2 + i\epsilon} \times \\
&\quad \times f^{be'a_2} \Big(g^{\nu 0}(p+k-2q_2)^{\xi'} + g^{\nu\xi'}(-2p-2k+q_2)^0 + g^{\xi'0}(p+k+q_2)^\nu \Big) T_{a_2} V(q_2) e^{iq_2 x_2} \frac{-i\delta_{bb'} g_{\nu\nu'}}{(p+k-q_2)^2 + i\epsilon} \times \\
&\quad \times f^{ab'a_1} \Big(g^{\mu 0}(p+k-2q_1-q_2)^{\nu'} + g^{\mu\nu'}(-2p-2k+q_1+2q_2)^0 + g^{\nu'0}(p+k+q_1-q_2)^\mu \Big) T_{a_1} V(q_1) e^{iq_1 x_1} \times \\
&\quad \times \frac{-i\delta_{aa'} g_{\mu\mu'}}{(p+k-q_1-q_2)^2 + i\epsilon} i J_{a'}(p+k-q_1-q_2) \epsilon^{\mu'}(p+k-q_1-q_2) e^{i(p+k-q_1-q_2)x_0} \\
&\approx i J_a(p+k) e^{i(p+k)x_0} f^{ecd} f^{bea_2} f^{aba_1} T_{a_2} T_{a_1} (1-x+x^2)(-i) \int \frac{d^2\mathbf{q}_1}{(2\pi)^2}(-i) \int \frac{d^2\mathbf{q}_2}{(2\pi)^2}(2ig_s) \frac{\boldsymbol{\epsilon} \cdot ((1-x)\mathbf{k} - x\mathbf{p})}{((1-x)\mathbf{k} - x\mathbf{p})^2} \times \\
&\quad \times e^{-i\mathbf{q}_1 \cdot \mathbf{b}_1} e^{-i\mathbf{q}_2 \cdot \mathbf{b}_2} (E^+)^2 \int \frac{dq_{1z}}{2\pi} \frac{dq_{2z}}{2\pi} \frac{v(q_{1z}, \mathbf{q}_1) v(q_{2z}, \mathbf{q}_2) e^{-iq_{1z}(z_1 - z_0)} e^{-iq_{2z}(z_2 - z_0)}}{((p+k-q_1-q_2)^2 + i\epsilon)((p+k-q_2)^2 + i\epsilon)},
\end{aligned} \tag{E4}
$$

where $\mathbf{b}_i \equiv \mathbf{x}_i - \mathbf{x}_0$, $i = 1, 2$ denote transverse impact parameters. We used Eq. (D5) and assumed that J varies slowly with momentum q_i, i.e. $J(p + k - q_1 - q_2) \approx J(p + k)$.

Regarding the longitudinal q_{1z} integral, we introduce a new variable: $q_z = q_{1z} + q_{2z}$ throughout this, and the following sections involving Feynman amplitudes which include interactions with two scattering centers. Therefore, we rewrite the exponent in the following manner: $e^{-iq_{1z}(z_1 - z_0)} e^{-iq_{2z}(z_2 - z_0)} = e^{-iq_z(z_1 - z_0)} e^{-iq_{2z}(z_2 - z_1)}$. Rewriting q_{1z} longitudinal integral in terms of q_z, i.e. changing the variables, we obtain:

$$I_2(p, k, \mathbf{q}_1, \vec{q}_2, z_1 - z_0) = \int \frac{dq_z}{2\pi} \frac{v(q_z - q_{2z}, \mathbf{q}_1) e^{-iq_z(z_1 - z_0)}}{(p+k-q_1-q_2)^2 + i\epsilon}. \tag{E5}$$

Again, due to $z_1 > z_0$, the contour must be closed in the

lower half-plane of complex q_z plain, so additional minus sign arises from the negative orientation of the contour and also we neglect the pole at $q_z = -i\mu_{1\perp} + q_{2z}$, since it is exponentially suppressed due to Eq. (B6). Thus, only one pole, originating from the gluon propagator, contributes to the first longitudinal integral:

$$
\begin{aligned}
\bar{q} &= -\frac{\mathbf{k}^2}{xE^+} - \frac{\mathbf{p}^2}{(1-x)E^+} - i\epsilon \\
&= -\frac{\mathbf{k}^2}{2\omega} - \frac{x}{1-x} \frac{(\mathbf{k} - \mathbf{q}_1 - \mathbf{q}_2)^2}{2\omega} - i\epsilon, \tag{E6}
\end{aligned}
$$

where we used, as well as throughout the Appendices F-H the relation between transverse momenta Eq. (E3). The residue at Eq. (E6) then gives:
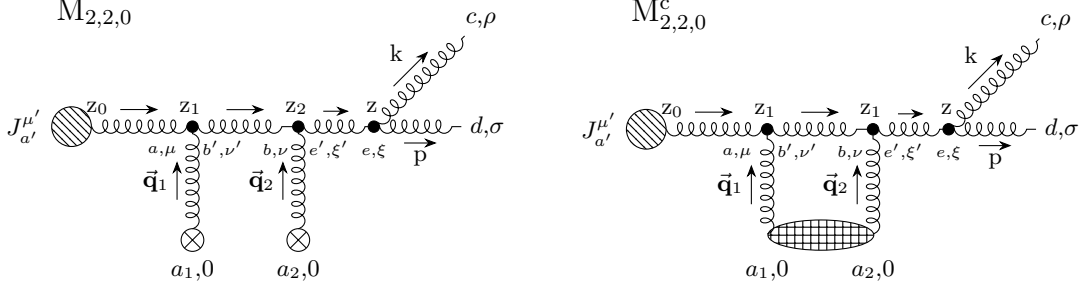
FIG. 8: Feynman diagram $M_{2,2,0}$ and its contribution to the first order in opacity gluon-jet radiative energy loss: contact-limit $M_{2,2,0}^c$. z_i, where $i = 1, 2$, denotes longitudinal coordinate of the interactions with the consecutive scattering centers (or in the contact limit $z_1 = z_2$). Crossed circles represent scatterers that exchange 3D momentum $\vec{\mathbf{q}}_i$ with the jet, which in contact-limit case merge into one gridded ellipse. Note that, all the following figures assume equivalently ordered Latin and Greek indices as in this figure. Remaining labeling is the same as in Figs.(6, 7).

$$I_2(p,k,\mathbf{q}_1,\vec{\mathbf{q}}_2,z_1 - z_0) \approx -v(-q_{2z} - \frac{\mathbf{k}^2}{2\omega} - \frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega}, \mathbf{q}_1)\frac{i}{E^+}e^{\frac{i}{2\omega}(\mathbf{k}^2 + \frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2)(z_1-z_0)}. \qquad \text{(E7)}$$

Next we need to solve the remaining q_{2z} longitudinal momentum transfer integral:

$$I_3(p,k,\mathbf{q}_1,\mathbf{q}_2,z_2 - z_1) = \int \frac{dq_{2z}}{2\pi}\frac{v(q_{2z},\mathbf{q}_2)e^{-iq_{2z}(z_2-z_1)}}{(p+k-q_2)^2 + i\epsilon} \times$$
$$\times v(-q_{2z} - \frac{\mathbf{k}^2}{2\omega} - \frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega}, \mathbf{q}_1). \qquad \text{(E8)}$$

Luckily, we are interested only in two limiting cases:

- The limit of well-separated scattering centers $z_2 - z_1 \gg 1/\mu$, where poles originating from Yukawa potentials are exponentially suppressed,

- The contact limit $z_1 = z_2$, where these poles contribute to the final results.

In the case of two distinct scatterers ($z_1 \neq z_2$) and in the limit of well-separated scattering centers there is only one pole that contributes to the residue (the singularities originating from Yukawa potential once again are exponential suppressed):

$$\bar{q}_{2z} = -\frac{\mathbf{k}^2}{xE^+} - \frac{\mathbf{p}^2}{(1-x)E^+} + \frac{\mathbf{q}_1^2}{E^+} - i\epsilon$$
$$= -\frac{\mathbf{k}^2}{2\omega} - \frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega} + \frac{\mathbf{q}_1^2}{E^+} - i\epsilon. \qquad \text{(E9)}$$

Since $z_2 > z_1$ again we close the contour below the real q_{2z} axis and thus obtain:

$$I_3(p,k,\mathbf{q}_1,\mathbf{q}_2,z_2 - z_1) \approx -v(0,\mathbf{q}_1)v(0,\mathbf{q}_2)\frac{i}{E^+} \times$$
$$\times e^{\frac{i}{2\omega}(\mathbf{k}^2 + \frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2 - x\mathbf{q}_1^2)(z_2-z_1)}. \qquad \text{(E10)}$$

In the special case of contact limit, i.e. when $z_1 = z_2$, instead of Eq. (E8) we need to calculate the following q_{2z} integral:

$$I_3^c(p,k,\mathbf{q}_1,\mathbf{q}_2,0) = \int \frac{dq_{2z}}{2\pi}\frac{v(q_{2z},\mathbf{q}_2)}{(p+k-q_2)^2 + i\epsilon} \times$$
$$\times v(-q_{2z} - \frac{\mathbf{k}^2}{2\omega} - \frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega}, \mathbf{q}_1). \qquad \text{(E11)}$$

Now, the contributions from Yukawa singularities ($q_{2z} = -i\mu_{1\perp}$, $q_{2z} = -i\mu_{2\perp}$) are not negligible and need to be included together with Eq. (E9). By choosing the same integration contour we obtain:

$$I_3^c(p,k,\mathbf{q}_1,\mathbf{q}_2,0) \approx \frac{-i}{E^+}\left(v(0,\mathbf{q}_1)v(0,\mathbf{q}_2) - \frac{(4\pi\alpha_s)^2}{2}\frac{1}{\mu_{2\perp}^2 - \mu_{1\perp}^2}\left(\frac{1}{\mu_{1\perp}^2} - \frac{1}{\mu_{2\perp}^2}\right)\right) = -v(0,\mathbf{q}_1)v(0,\mathbf{q}_2)\frac{i}{2E^+}, \qquad \text{(E12)}$$

which is exactly $\frac{1}{2}$ of the strength of Eq. (E10). Note

that, in previous calculations we applied soft-rescattering

approximation and also assumed $E^+ \gg \mu_{i\perp}$, $i = 1, 2$.

Finally, contact limit of this amplitude reads:

$$
\begin{aligned}
M_{2,2,0}^c = & -iJ_a(p+k)e^{i(p+k)x_0}f^{ecd}f^{bea_2}f^{aba_1}T_{a_2}T_{a_1}(1-x+x^2)(-i)\int\frac{d^2\mathbf{q}_1}{(2\pi)^2}(-i)\int\frac{d^2\mathbf{q}_2}{(2\pi)^2}v(0,\mathbf{q}_1)v(0,\mathbf{q}_2)e^{-i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{b}_1}\times \\
& \times\frac{1}{2}(2ig_s)\frac{\boldsymbol{\epsilon}\cdot((1-x)\mathbf{k}-x\mathbf{p})}{((1-x)\mathbf{k}-x\mathbf{p})^2}e^{i(\frac{\mathbf{k}^2}{xE^+}+\frac{\mathbf{p}^2}{(1-x)E^+})(z_1-z_0)} \\
= & -J_a(p+k)e^{i(p+k)x_0}(T^cT^{a_2}T^{a_1})_{da}T_{a_2}T_{a_1}(1-x+x^2)(-i)\int\frac{d^2\mathbf{q}_1}{(2\pi)^2}(-i)\int\frac{d^2\mathbf{q}_2}{(2\pi)^2}v(0,\mathbf{q}_1)v(0,\mathbf{q}_2)e^{-i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{b}_1}\times \\
& \times\frac{1}{2}(2ig_s)\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x(\mathbf{q}_1+\mathbf{q}_2))}{(\mathbf{k}-x(\mathbf{q}_1+\mathbf{q}_2))^2}e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2)(z_1-z_0)},
\end{aligned} \tag{E13}
$$

where we applied Eq. (E3) and manipulated with $SU(N_c = 3)$ structure constants by using Eqs. (A5, A6). Also we assumed that $\mathbf{x}_1 = \mathbf{x}_2$, since diagrams with two different centers will not contribute to the final result due to Eq. (A3,A4).

Note from Fig. 8 that $M_{2,2,0}$ is symmetric under the substitutions: $(p \leftrightarrow k, x \leftrightarrow (1-x), c \leftrightarrow d)$, which can be straightforwardly verified by implementing these substitutions in the first two lines of Eq. (E13).

medium (the first row of Fig. 9).

Appendix F: Diagrams $M_{2,0,3}$ and $M_{2,0,0}$

Next we consider $M_{2,0,3}$ diagram, where the radiated gluon suffers two consecutive interactions with the QCD

Note that the order of the color and Dirac indices denoting vertices is the same for all the remaining diagrams containing two interactions with the scatterers as in Fig. 8, and therefore omitted onward.

$$
\begin{aligned}
M_{2,0,3} = & \int\frac{d^4q_1}{(2\pi)^4}\frac{d^4q_2}{(2\pi)^4}\epsilon_\rho^*(k)f^{eca_2}\Big(g^{\xi 0}(k-2q_2)^\rho+g^{\xi\rho}(-2k+q_2)^0+g^{\rho 0}(k+q_2)^\xi\Big)T_{a_2}V(q_2)e^{iq_2x_2}\frac{-i\delta_{ee'}g_{\xi\xi'}}{(k-q_2)^2+i\epsilon}\times \\
& \times f^{be'a_1}\Big(g^{\nu 0}(k-2q_1-q_2)^{\xi'}+g^{\nu\xi'}(-2k+q_1+2q_2)^0+g^{\xi' 0}(k+q_1-q_2)^\nu\Big)T_{a_1}V(q_1)e^{iq_1x_1}\frac{-i\delta_{bb'}g_{\nu\nu'}}{(k-q_1-q_2)^2+i\epsilon}\times \\
& \times\epsilon_\sigma^*(p)g_sf^{adb'}\Big(g^{\mu\nu'}(p+2k-2q_1-2q_2)^\sigma+g^{\mu\sigma}(-2p-k+q_1+q_2)^{\nu'}+g^{\sigma\nu'}(p-k+q_1+q_2)^\mu\Big)\times \\
& \times\frac{-i\delta_{aa'}g_{\mu\mu'}}{(p+k-q_1-q_2)^2+i\epsilon}iJ_{a'}(p+k-q_1-q_2)\epsilon^{\mu'}(p+k-q_1-q_2)e^{i(p+k-q_1-q_2)x_0} \\
\approx & \ iJ_a(p+k)e^{i(p+k)x_0}f^{eca_2}f^{bea_1}f^{adb}T_{a_2}T_{a_1}\frac{(1-x+x^2)}{1-x}(-i)\int\frac{d^2\mathbf{q}_1}{(2\pi)^2}(-i)\int\frac{d^2\mathbf{q}_2}{(2\pi)^2}(2ig_s)\boldsymbol{\epsilon}\cdot\mathbf{p}\ e^{-i\mathbf{q}_1\cdot\mathbf{b}_1}e^{-i\mathbf{q}_2\cdot\mathbf{b}_2}\times \\
& \times\int\frac{dq_{1z}}{2\pi}\frac{dq_{2z}}{2\pi}\frac{E^+k^+v(q_{1z},\mathbf{q}_1)v(q_{2z},\mathbf{q}_2)e^{-iq_{1z}(z_1-z_0)}e^{-iq_{2z}(z_2-z_0)}}{((p+k-q_1-q_2)^2+i\epsilon)((k-q_1-q_2)^2+i\epsilon)((k-q_2)^2+i\epsilon)}.
\end{aligned} \tag{F1}
$$

Next, again by changing the variables $q_{1z} \to q_z = q_{1z} + q_{2z}$, we define the following integral:

$$
I_2(p,k,\mathbf{q}_1,\vec{q}_2,z_1-z_0) = \int\frac{dq_z}{2\pi}\frac{v(q_z-q_{2z},\mathbf{q}_1)e^{-iq_z(z_1-z_0)}}{((p+k-q_1-q_2)^2+i\epsilon)((k-q_1-q_2)^2+i\epsilon)}. \tag{F2}
$$

$$\mathrm{M}_{2,0,3}$$

$$\mathrm{M}^{c}_{2,0,3}$$

$$\mathrm{M}_{2,0,0}$$

$$\mathrm{M}^{c}_{2,0,0}$$
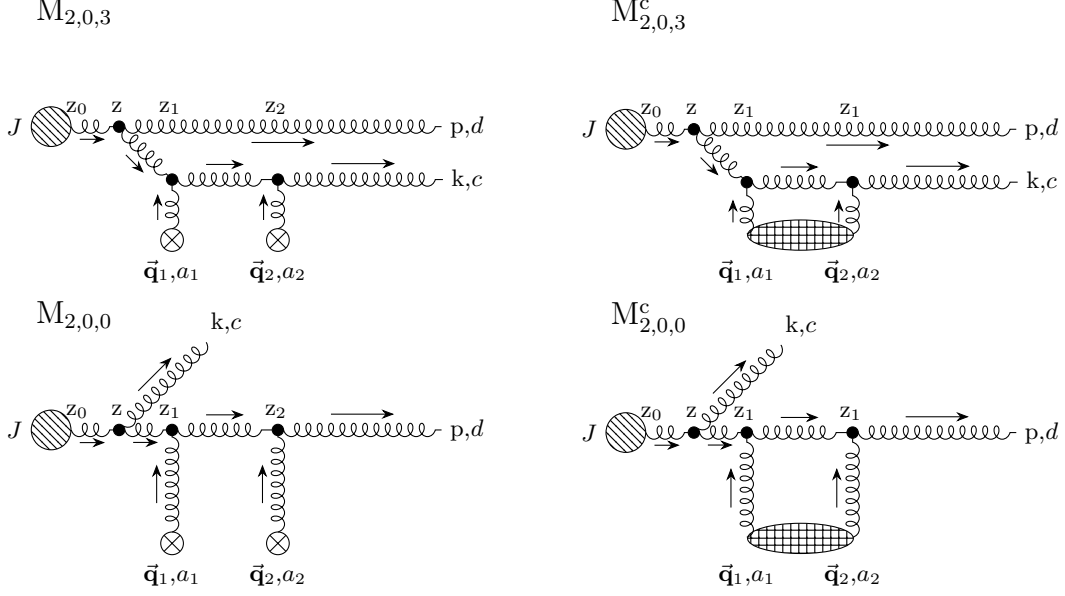
FIG. 9: Feynman diagrams $M_{2,0,3}$ and $M_{2,0,0}$ in well-separated (left column) and in contact-limit case ($z_1 = z_2$), which contributes to the first order in opacity of gluon-jet radiative energy loss: $M^{c}_{2,0,3}$ and $M^{c}_{2,0,0}$ (right column). Remaining labeling is the same as in Fig. 8.

Again, as explained in the previous section, we close the contour in lower half-plane, and since $\mu(z_1 - z_0) \gg 1$ the pole at $q_z = -i\mu_{1\perp} + q_{2z}$ is again exponentially suppressed. Therefore the remaining q_z singularities originating from gluon propagators are:

$$\bar{q}_1 = -\frac{\mathbf{k}^2}{xE^+} - \frac{\mathbf{p}^2}{(1-x)E^+} - i\epsilon = -\frac{\mathbf{k}^2}{2\omega} - \frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega} - i\epsilon,$$

$$\bar{q}_2 = -\frac{\mathbf{k}^2}{xE^+} + \frac{\mathbf{p}^2}{xE^+} - i\epsilon = -\frac{\mathbf{k}^2}{2\omega} + \frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega} - i\epsilon. \tag{F3}$$

After performing the integration, i.e. summing the residues at these two poles, I_2 now reads:

$$I_2(p, k, \mathbf{q}_1, \vec{\mathbf{q}}_2, z_1 - z_0) \approx v(-q_{2z}, \mathbf{q}_1)\frac{i(1-x)}{E^+(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}\left(e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2)(z_1-z_0)} - e^{\frac{i}{2\omega}(\mathbf{k}^2-(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2)(z_1-z_0)}\right). \tag{F4}$$

The remaining integral over q_{2z} is:

$$I_3(p, k, \mathbf{q}_1, \mathbf{q}_2, z_2 - z_1) = \int \frac{dq_{2z}}{2\pi} \frac{v(q_{2z}, \mathbf{q}_2)e^{-iq_{2z}(z_2-z_1)}}{(k-q_2)^2 + i\epsilon} \times$$
$$\times v(-q_{2z}, \mathbf{q}_1), \tag{F5}$$

and since we are interested only in the contact-limit case (i.e. $z_1 = z_2$), we need to calculate:

$$I^c_3(p, k, \mathbf{q}_1, \mathbf{q}_2, 0) = \int \frac{dq_{2z}}{2\pi} \frac{v(q_{2z}, \mathbf{q}_2)v(-q_{2z}, \mathbf{q}_1)}{(k-q_2)^2 + i\epsilon}, \tag{F6}$$

which gives:

$$I_3^c(p, k, \mathbf{q}_1, \mathbf{q}_2, 0) \approx -v(0, \mathbf{q}_1)v(0, \mathbf{q}_2)\frac{i}{2xE^+}, \quad \text{(F7)}$$

which can readily be shown to represent exactly $\frac{1}{2}$ of the

strength of the well-separated limit Eq. (F5), as for $M_{2,2,0}$ amplitude. The contact limit of this amplitude reduces to:

$$M_{2,0,3}^c = iJ_a(p+k)e^{i(p+k)x_0}f^{eca_2}f^{bea_1}f^{adb}T_{a_2}T_{a_1}(1-x+x^2)(-i)\int\frac{d^2\mathbf{q}_1}{(2\pi)^2}(-i)\int\frac{d^2\mathbf{q}_2}{(2\pi)^2}v(0,\mathbf{q}_1)v(0,\mathbf{q}_2)e^{-i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{b}_1}\times$$
$$\times\frac{1}{2}(2ig_s)\frac{\boldsymbol{\epsilon}\cdot\mathbf{p}}{\mathbf{p}^2}\left(e^{i(\frac{\mathbf{k}^2}{xE^+}+\frac{\mathbf{p}^2}{(1-x)E^+})(z_1-z_0)}-e^{i\frac{(\mathbf{k}^2-\mathbf{p}^2)}{xE^+}(z_1-z_0)}\right)$$
$$= J_a(p+k)e^{i(p+k)x_0}[[T^c,T^{a_2}],T^{a_1}]_{da}T_{a_2}T_{a_1}(1-x+x^2)(-i)\int\frac{d^2\mathbf{q}_1}{(2\pi)^2}(-i)\int\frac{d^2\mathbf{q}_2}{(2\pi)^2}v(0,\mathbf{q}_1)v(0,\mathbf{q}_2)e^{-i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{b}_1}\times$$
$$\times\frac{1}{2}(2ig_s)\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)}{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}\left(e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2)(z_1-z_0)}-e^{\frac{i}{2\omega}(\mathbf{k}^2-(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2)(z_1-z_0)}\right). \quad \text{(F8)}$$

Proceeding in the same manner, for $M_{2,0,0}^c$ amplitude (the second row of Fig. 9) we obtain:

$$M_{2,0,0}^c = iJ_a(p+k)e^{i(p+k)x_0}f^{eda_2}f^{bea_1}f^{acb}T_{a_2}T_{a_1}(1-x+x^2)(-i)\int\frac{d^2\mathbf{q}_1}{(2\pi)^2}(-i)\int\frac{d^2\mathbf{q}_2}{(2\pi)^2}v(0,\mathbf{q}_1)v(0,\mathbf{q}_2)e^{-i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{b}_1}\times$$
$$\times\frac{1}{2}(2ig_s)\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2}\left(e^{i(\frac{\mathbf{k}^2}{xE^+}+\frac{\mathbf{p}^2}{(1-x)E^+})(z_1-z_0)}-e^{i\frac{(\mathbf{p}^2-\mathbf{k}^2)}{(1-x)E^+}(z_1-z_0)}\right)$$
$$= J_a(p+k)e^{i(p+k)x_0}(T^{a_2}T^{a_1}T^c)_{da}T_{a_2}T_{a_1}(1-x+x^2)(-i)\int\frac{d^2\mathbf{q}_1}{(2\pi)^2}(-i)\int\frac{d^2\mathbf{q}_2}{(2\pi)^2}v(0,\mathbf{q}_1)v(0,\mathbf{q}_2)e^{-i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{b}_1}\times$$
$$\times\frac{1}{2}(2ig_s)\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2}\left(e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2)(z_1-z_0)}-e^{\frac{i}{2\omega}\frac{x}{1-x}((\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2-\mathbf{k}^2)(z_1-z_0)}\right). \quad \text{(F9)}$$

From Fig. 9 we infer that $M_{2,0,3}$ and $M_{2,0,0}$ are symmetric under the following substitutions: $(p \leftrightarrow k, x \leftrightarrow (1-x), c \leftrightarrow d)$, which can be straightforwardly verified by implementing these substitutions in Eqs. (F8, F9).

We provide only the contact-limit case diagrams $M_{2,0,1}^c$ and $M_{2,0,2}^c$ (Fig. 10), since, in the end, only they are used in calculating radiative energy loss to the first order in opacity.

Appendix G: Diagrams $M_{2,0,1}$ and $M_{2,0,2}$

Here we consider the case when both initial gluon jet and radiated gluon interact with one scattering center.

$$M^c_{2,0,1} \qquad\qquad\qquad M^c_{2,0,2}$$
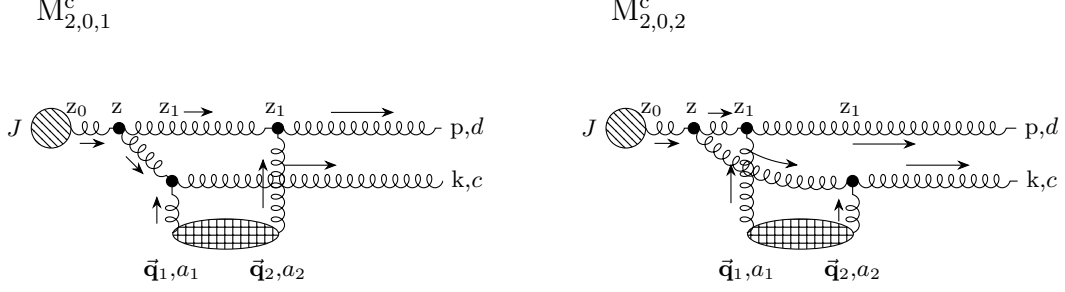
FIG. 10: Topologically indistinct Feynman diagrams $M^c_{2,0,1}$ and $M^c_{2,0,2}$ in contact limit ($z_1 = z_2$), which contribute to the first order in opacity of gluon-jet radiative energy loss. Remaining labeling is the same as in Fig. 8.

$$
M_{2,0,1} = \int \frac{d^4q_1}{(2\pi)^4}\frac{d^4q_2}{(2\pi)^4}\epsilon^*_\sigma(p)f^{eda_2}\Big(g^{\xi 0}(p-2q_2)^\sigma + g^{\xi\sigma}(-2p+q_2)^0 + g^{\sigma 0}(p+q_2)^\xi\Big)T_{a_2}V(q_2)e^{iq_2 x_2}\times
$$
$$
\times\frac{-i\delta_{ee'}g_{\xi\xi'}}{(p-q_2)^2+i\epsilon}g_s f^{ae'b'}\Big(g^{\mu\nu'}(p+2k-2q_1-q_2)^{\xi'} + g^{\mu\xi'}(-2p-k+q_1+2q_2)^{\nu'} + +g^{\xi'\nu'}(p-k+q_1-q_2)^\mu\Big)\times
$$
$$
\times\epsilon^*_\rho(k)f^{bca_1}\Big(g^{\nu 0}(k-2q_1)^\rho + g^{\nu\rho}(-2k+q_1)^0 + g^{\rho 0}(k+q_1)^\nu\Big)T_{a_1}V(q_1)e^{iq_1 x_1}\frac{-i\delta_{bb'}g_{\nu\nu'}}{(k-q_1)^2+i\epsilon}\times
$$
$$
\times\frac{-i\delta_{aa'}g_{\mu\mu'}}{(p+k-q_1-q_2)^2+i\epsilon}iJ_{a'}(p+k-q_1-q_2)\epsilon^{\mu'}(p+k-q_1-q_2)e^{i(p+k-q_1-q_2)x_0}
$$
$$
\approx -iJ_a(p+k)e^{i(p+k)x_0}f^{eda_2}f^{aeb}f^{bca_1}T_{a_2}T_{a_1}(1-x+x^2)(-i)\int\frac{d^2\mathbf{q}_1}{(2\pi)^2}(-i)\int\frac{d^2\mathbf{q}_2}{(2\pi)^2}(2ig_s)\,\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)\,e^{-i\mathbf{q}_1\cdot\mathbf{b}_1}\times
$$
$$
\times e^{-i\mathbf{q}_2\cdot\mathbf{b}_2}(E^+)^2\int\frac{dq_{1z}}{2\pi}\frac{dq_{2z}}{2\pi}\frac{v(q_{1z},\mathbf{q}_1)v(q_{2z},\mathbf{q}_2)e^{-iq_{1z}(z_1-z_0)}e^{-iq_{2z}(z_2-z_0)}}{((p+k-q_1-q_2)^2+i\epsilon)((k-q_1)^2+i\epsilon)((p-q_2)^2+i\epsilon)}. \tag{G1}
$$

Again, by changing the variables $q_{1z}\to q_z = q_{1z}+q_{2z}$, we define the following integral:

$$
I_2(p,k,\mathbf{q}_1,\vec{q}_2,z_1-z_0) = \int\frac{dq_z}{2\pi}\frac{v(q_z-q_{2z},\mathbf{q}_1)e^{-iq_z(z_1-z_0)}}{(p+k-q_1-q_2)^2+i\epsilon}\times
$$
$$
\times\frac{1}{(k-q_1)^2+i\epsilon}. \tag{G2}
$$

Since $z_1 > z_0$ we must close the contour in lower half-plane, and since $\mu(z_1-z_0)\gg 1$ again we neglect the

pole at $q_z = -i\mu_{1\perp}+q_{2z}$. Therefore the remaining q_z singularities originating from gluon propagators are:

$$
\bar{q}_1 = -\frac{\mathbf{k}^2}{2\omega} - \frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega} - i\epsilon,
$$
$$
\bar{q}_2 = -\frac{\mathbf{k}^2}{2\omega} + \frac{(\mathbf{k}-\mathbf{q}_1)^2}{2\omega} + q_{2z} - i\epsilon. \tag{G3}
$$

Summing the residues gives:

$$
I_2(p,k,\mathbf{q}_1,\vec{q}_2,z_1-z_0) \approx \frac{ie^{i\frac{\mathbf{k}^2}{2\omega}(z_1-z_0)}}{E^+k^+(q_{2z}+\frac{(\mathbf{k}-\mathbf{q}_1)^2}{2\omega}+\frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega})}\times
$$
$$
\times\left(v(-q_{2z}-\frac{\mathbf{k}^2}{2\omega}-\frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega},\mathbf{q}_1)e^{i\frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega}(z_1-z_0)} - v(\frac{(\mathbf{k}-\mathbf{q}_1)^2}{2\omega}-\frac{\mathbf{k}^2}{2\omega},\mathbf{q}_1)e^{-i(q_{2z}+\frac{(\mathbf{k}-\mathbf{q}_1)^2}{2\omega})(z_1-z_0)}\right). \tag{G4}
$$

The remaining q_{2z} integral is:

$$I_3(p,k,\mathbf{q}_1,\mathbf{q}_2,z_2-z_0,z_2-z_1) = \int \frac{dq_{2z}}{2\pi} \frac{1}{q_{2z} + \frac{(\mathbf{k}-\mathbf{q}_1)^2}{2\omega} + \frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega}} \frac{v(q_{2z},\mathbf{q}_2)}{(p-q_2)^2+i\epsilon} \times$$

$$\times \left(e^{-iq_{2z}(z_2-z_1)} e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2)(z_1-z_0)} v\left(-q_{2z} - \frac{\mathbf{k}^2}{2\omega} - \frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega},\mathbf{q}_1\right) - \right.$$

$$\left. - e^{-iq_{2z}(z_2-z_0)} e^{-\frac{i}{2\omega}((\mathbf{k}-\mathbf{q}_1)^2-\mathbf{k}^2)(z_1-z_0)} v\left(\frac{(\mathbf{k}-\mathbf{q}_1)^2}{2\omega} - \frac{\mathbf{k}^2}{2\omega},\mathbf{q}_1\right) \right), \tag{G5}$$

where the singularity on q_{2z} real axis: $q_{2z} = -\frac{(\mathbf{k}-\mathbf{q}_1)^2}{2\omega} - \frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega} \equiv -a$, $(a > 0)$ has to be avoided by taking Cauchy principal value of I_3 according to the Fig. 11, i.e.:

$$I_3 \equiv I_{PV} = I_B - I_C - I_D, \tag{G6}$$

where $I_B = -2\pi i \sum_i Res(I_3(\bar{q}_i))$, with i counting the poles in the lower-half plane. Additionally $I_C = 0$, and it's straightforward to show, that after the following substitution $q_{2z} = -a + re^{i\varphi}$, where $r \to 0$, also $I_D = 0$. Therefore, principal value of I_3 reduces to I_B,

i.e. $-2\pi i \sum_i Res(I_3(\bar{q}_i))$.

In the well-separated case Eq. (G5) poles originating from Yukawa potentials ($q_{2z} = -\frac{\mathbf{k}^2}{2\omega} - \frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega} - i\mu_{1\perp}$ and $q_{2z} = -i\mu_{2\perp}$) are again exponentially suppressed ($e^{-\mu_{i\perp}(z_2-z_{0,1})} \to 0$, $i = 1, 2$) and therefore can be neglected, so only the pole from the propagator survives $q_{2z} = \frac{x}{1-x}(\frac{(\mathbf{k}-\mathbf{q}_1)^2}{2\omega} - \frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega}) - i\epsilon$. However, since we are interested only in the contact-limit case (i.e. $z_1 = z_2$), instead of Eq. (G5) we need to calculate the principal value of the following integral:

$$I_3^c(p,k,\mathbf{q}_1,\mathbf{q}_2,z_1-z_0) = \int \frac{dq_{2z}}{2\pi} \frac{1}{q_{2z} + \frac{(\mathbf{k}-\mathbf{q}_1)^2}{2\omega} + \frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega}} \times$$

$$\times \left(\frac{e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2)(z_1-z_0)}}{(p-q_2)^2+i\epsilon} v\left(-q_{2z} - \frac{\mathbf{k}^2}{2\omega} - \frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega},\mathbf{q}_1\right)v(q_{2z},\mathbf{q}_2) - \right.$$

$$\left. - \frac{e^{-iq_{2z}(z_1-z_0)}e^{-\frac{i}{2\omega}((\mathbf{k}-\mathbf{q}_1)^2-\mathbf{k}^2)(z_1-z_0)}}{(p-q_2)^2+i\epsilon} v\left(\frac{(\mathbf{k}-\mathbf{q}_1)^2}{2\omega} - \frac{\mathbf{k}^2}{2\omega},\mathbf{q}_1\right)v(q_{2z},\mathbf{q}_2) \right), \tag{G7}$$

which again reduces to the sum of residua, with $-a$ effectively not being a pole (Fig. 11). Particularly, for the second term in the bracket of Eq. (G7), only the propagator pole survives, while for the first term in the bracket all three poles have to be accounted, although

residues at poles from potentials sum to the order of $\mathcal{O}(\frac{(\mathbf{k}-\mathbf{q}_1)^2}{x(1-x)E^+(\mu_{1\perp}+\mu_{2\perp})})$, and thus can be neglected compared to the remaining residue.

Finally, in the contact-limit case we obtain:

$$M_{2,0,1}^c \approx -iJ_a(p+k)e^{i(p+k)x_0} f^{eda_2} f^{aeb} f^{bca_1} T_{a_2} T_{a_1}(1-x+x^2)(-i)\int \frac{d^2\mathbf{q}_1}{(2\pi)^2}(-i)\int \frac{d^2\mathbf{q}_2}{(2\pi)^2} v(0,\mathbf{q}_1)v(0,\mathbf{q}_2)e^{-i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{b}_1} \times$$

$$\times (2ig_s)\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2}\left(e^{i(\frac{\mathbf{k}^2}{xE^+}+\frac{\mathbf{p}^2}{(1-x)E^+})(z_1-z_0)} - e^{i(\frac{\mathbf{k}^2}{xE^+}+\frac{\mathbf{p}^2}{(1-x)E^+}-\frac{(\mathbf{k}-\mathbf{q}_1)^2}{x(1-x)E^+})(z_1-z_0)}\right)$$

$$= J_a(p+k)e^{i(p+k)x_0}(T^{a_2}[T^c,T^{a_1}])_{da} T_{a_2} T_{a_1}(1-x+x^2)(-i)\int \frac{d^2\mathbf{q}_1}{(2\pi)^2}(-i)\int \frac{d^2\mathbf{q}_2}{(2\pi)^2} v(0,\mathbf{q}_1)v(0,\mathbf{q}_2)e^{-i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{b}_1} \times$$

$$\times (2ig_s)\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2}\left(e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2)(z_1-z_0)} - e^{\frac{i}{2\omega}(\mathbf{k}^2-\frac{(\mathbf{k}-\mathbf{q}_1)^2}{1-x}+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2)(z_1-z_0)}\right). \tag{G8}$$

Notice that, contrary to the previous three amplitudes that also included two scattering centers, in Eq. (G8) no
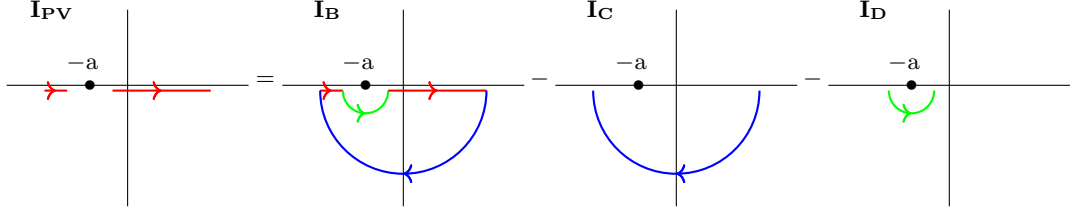
FIG. 11: Illustration of calculating Cauchy principal value (I_{PV}) in the case when singularity (-a) on the real (horizontal) axis arises.

factor $\frac{1}{2}$ when comparing to well-separated limit appears.

Proceeding in the same manner, for $M_{2,0,2}^c$ we obtain:

$$
M_{2,0,2}^c \approx iJ_a(p+k)e^{i(p+k)x_0}f^{eca_2}f^{abe}f^{bda_1}T_{a_2}T_{a_1}(1-x+x^2)(-i)\int\frac{d^2\mathbf{q}_1}{(2\pi)^2}(-i)\int\frac{d^2\mathbf{q}_2}{(2\pi)^2}v(0,\mathbf{q}_1)v(0,\mathbf{q}_2)e^{-i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{b}_1}\times
$$
$$
\times(2ig_s)\frac{\boldsymbol{\epsilon}\cdot(\mathbf{p}-\mathbf{q}_1)}{(\mathbf{p}-\mathbf{q}_1)^2}\left(e^{i(\frac{\mathbf{k}^2}{xE^+}+\frac{\mathbf{p}^2}{(1-x)E^+})(z_1-z_0)}-e^{i(\frac{\mathbf{k}^2}{xE^+}+\frac{\mathbf{p}^2}{(1-x)E^+}-\frac{(\mathbf{p}-\mathbf{q}_1)^2}{x(1-x)E^+})(z_1-z_0)}\right)
$$
$$
=J_a(p+k)e^{i(p+k)x_0}(T^{a_1}[T^c,T^{a_2}])_{da}T_{a_2}T_{a_1}(1-x+x^2)(-i)\int\frac{d^2\mathbf{q}_1}{(2\pi)^2}(-i)\int\frac{d^2\mathbf{q}_2}{(2\pi)^2}v(0,\mathbf{q}_1)v(0,\mathbf{q}_2)e^{-i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{b}_1}\times
$$
$$
\times(2ig_s)\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_2)}{(\mathbf{k}-\mathbf{q}_2)^2}\left(e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2)(z_1-z_0)}-e^{\frac{i}{2\omega}(\mathbf{k}^2-\frac{(\mathbf{k}-\mathbf{q}_2)^2}{1-x}+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2)(z_1-z_0)}\right). \tag{G9}
$$

As for $M_{2,0,1}^c$ amplitude, no factor of $\frac{1}{2}$ appears. From well-separated analogon of Fig. 10 we could infer that $M_{2,0,1}$ and $M_{2,0,2}$ are symmetric under the following substitutions: $(p\leftrightarrow k, x\leftrightarrow(1-x), c\leftrightarrow d)$, which can readily be verified by implementing these substitutions in the first two lines of either of the two Eqs. (G8, G9) and by using structure constant asymmetry. Note that, in Eq. (G9) we applied Eq. (E3). Also, since in contact-limit case these two diagrams are topologically indistinct, we need to either omit one of them in order to avoid over counting, or to include both, but multiply each by a factor $\frac{1}{2}$ (we will do the latter).

the case when one interaction with the scattering center located at $\vec{\mathbf{x}}_1$ occurs before and the other interaction at the same place occurs after the gluon has been radiated.

In order to avoid redundant derivations (i.e. repetition of the above calculations) we briefly outline our derivation of Feynman amplitudes for only contact-limit case.

Appendix H: Diagrams $M_{2,1,0}$ and $M_{2,1,1}$

The contact-limit case of the remaining two diagrams is presented in Fig. 12. These diagrams correspond to

In the light of time-ordered perturbation theory from [8, 39] these two diagrams are identically equal to zero, since $\int_{t_1}^{t_1}dt... = 0$, but for the consistency we will provide a brief verification of this argument.

$M^c_{2,1,0}$ k,c $M^c_{2,1,1}$
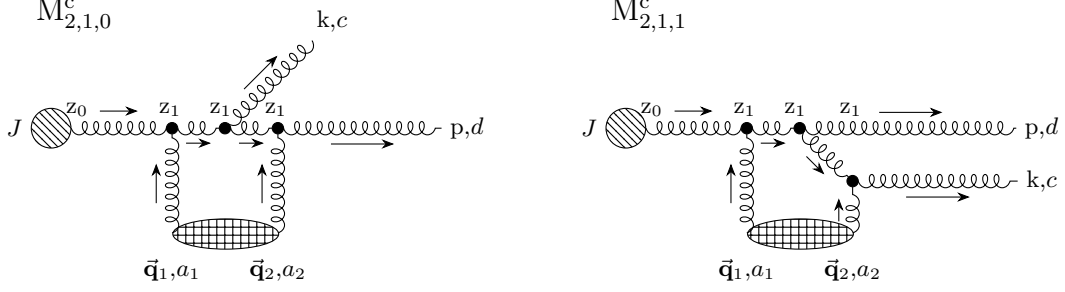
FIG. 12: Feynman diagrams $M^c_{2,1,0}$ and $M^c_{2,1,1}$ in contact limit ($z_1 = z_2$), which have negligible contribution to the first order in opacity gluon-jet radiative energy loss. Remaining labeling is the same as in Fig. 8.

$$
M_{2,1,0} = \int \frac{d^4q_1}{(2\pi)^4} \frac{d^4q_2}{(2\pi)^4} \epsilon^*_\sigma(p) f^{eda_2} \Big(g^{\xi 0}(p-2q_2)^\sigma + g^{\xi\sigma}(-2p+q_2)^0 + g^{\sigma 0}(p+q_2)^\xi \Big) T_{a_2} V(q_2) e^{iq_2 x_2} \frac{-i\delta_{ee'} g_{\xi\xi'}}{(p-q_2)^2+i\epsilon} \times
$$

$$
\times \epsilon^*_\rho(k) g_s f^{bce'} \Big(g^{\nu\xi'}(2p+k-2q_2)^\rho + g^{\nu\rho}(-p-2k+q_2)^{\xi'} + g^{\rho\xi'}(-p+k+q_2)^\nu \Big) \frac{-i\delta_{bb'} g_{\nu\nu'}}{(p+k-q_2)^2+i\epsilon} \times
$$

$$
\times f^{ab'a_1} \Big(g^{\mu 0}(p+k-2q_1-q_2)^{\nu'} + g^{\mu\nu'}(-2p-2k+q_1+2q_2)^0 + g^{\nu' 0}(p+k+q_1-q_2)^\mu \Big) T_{a_1} V(q_1) e^{iq_1 x_1} \times
$$

$$
\times \frac{-i\delta_{aa'} g_{\mu\mu'}}{(p+k-q_1-q_2)^2+i\epsilon} i J_{a'}(p+k-q_1-q_2) \epsilon^{\mu'}(p+k-q_1-q_2) e^{i(p+k-q_1-q_2)x_0}
$$

$$
\approx i J_a(p+k) e^{i(p+k)x_0} f^{eda_2} f^{bce} f^{aba_1} T_{a_2} T_{a_1} \frac{(1-x+x^2)}{x} (-i) \int \frac{d^2\mathbf{q}_1}{(2\pi)^2} (-i) \int \frac{d^2\mathbf{q}_2}{(2\pi)^2} (2ig_s) \, \boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1) \, e^{-i\mathbf{q}_1\cdot\mathbf{b}_1} \times
$$

$$
\times e^{-i\mathbf{q}_2\cdot\mathbf{b}_2} (E^+)^2 \int \frac{dq_{2z}}{2\pi} \frac{v(q_{2z},\mathbf{q}_2) e^{-iq_{2z}(z_2-z_1)}}{((p+k-q_2)^2+i\epsilon)((p-q_2)^2+i\epsilon)} \int \frac{dq_z}{2\pi} \frac{v(q_z-q_{2z},\mathbf{q}_1) e^{-iq_z(z_1-z_0)}}{(p+k-q_1-q_2)^2+i\epsilon}
$$

$$
\approx J_a(p+k) e^{i(p+k)x_0} f^{eda_2} f^{bce} f^{aba_1} T_{a_2} T_{a_1} (1-x+x^2)(-i) \int \frac{d^2\mathbf{q}_1}{(2\pi)^2} (-i) \int \frac{d^2\mathbf{q}_2}{(2\pi)^2} (2ig_s) \, \boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1) \, e^{-i\mathbf{q}_1\cdot\mathbf{b}_1} \times
$$

$$
\times e^{-i\mathbf{q}_2\cdot\mathbf{b}_2} (E^+)^2 \int \frac{dq_{2z}}{2\pi} \frac{v(q_{2z},\mathbf{q}_2) e^{-iq_{2z}(z_2-z_1)}}{((p+k-q_2)^2+i\epsilon)((p-q_2)^2+i\epsilon)} \frac{1}{k^+} v\Big(-\frac{\mathbf{k}^2}{2\omega} - \frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega} - q_{2z}, \mathbf{q}_1\Big) \times
$$

$$
\times e^{\frac{i}{2\omega}(\mathbf{k}^2 + \frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2)(z_1-z_0)}. \tag{H1}
$$

In the contact-limit case there are four q_{2z} poles of the above integral in the lower half-plane: $-\frac{\mathbf{k}^2}{2\omega} -$ $\frac{x}{1-x}\frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega} + \frac{x\mathbf{q}_1^2}{2\omega} - i\epsilon$, $\frac{x}{1-x}\big(\frac{(\mathbf{k}-\mathbf{q}_1)^2}{2\omega} - \frac{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2}{2\omega}\big) - i\epsilon$, $-i\mu_{1\perp}$ and $-i\mu_{2\perp}$, which give:

$$
M^c_{2,1,0} = i J_a(p+k) e^{i(p+k)x_0} (T^{a_2} T^c T^{a_1})_{da} T_{a_2} T_{a_1} (1-x+x^2)(-i) \int \frac{d^2\mathbf{q}_1}{(2\pi)^2} (-i) \int \frac{d^2\mathbf{q}_2}{(2\pi)^2} v(0,\mathbf{q}_1) v(0,\mathbf{q}_2) e^{-i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{b}_1} \times
$$

$$
\times (ig_s) \frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2} e^{\frac{i}{2\omega}(\mathbf{k}^2 + \frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2)(z_1-z_0)} \frac{\mu_{1\perp}^2 + \mu_{1\perp}\mu_{2\perp} + \mu_{2\perp}^2}{\mu_{1\perp}\mu_{2\perp}} \frac{(\mathbf{k}-x\mathbf{q}_1)^2}{x(1-x)E^+(\mu_{1\perp}+\mu_{2\perp})}, \tag{H2}
$$

where the residues at first two poles (i.e. originating from the gluon propagators) cancel each other exactly, leading to the result Eq. (H2) that is suppressed by a factor of $\mathcal{O}(\frac{(\mathbf{k}-x\mathbf{q}_1)^2}{x(1-x)E^+(\mu_{1\perp}+\mu_{2\perp})})$ compared to the all previous amplitudes (note that x is finite), as in the case of soft-gluon

approximation [9, 25].

The same conclusion applies to $M^c_{2,1,1}$ amplitude, which can be straightforwardly verified by repeating the analogous procedure as for $M^c_{2,1,0}$, and by the fact that these two amplitudes are symmetric (see Fig. 12) to the

exchange $(p \leftrightarrow k, x \leftrightarrow (1-x), c \leftrightarrow d)$.

Appendix I: Calculation of radiative energy loss

In this section we provide concise outline of calculating the first order in opacity radiative energy loss. We start with the equation:

$$
d^3 N_g^{(1)} d^3 N_J = \left(\frac{1}{d_T} \operatorname{Tr} \langle |M_1|^2 \rangle + \frac{2}{d_T} Re \operatorname{Tr} \langle M_2 M_0^* \rangle \right) \times
$$
$$
\times \frac{d^3 \vec{\mathbf{p}}}{(2\pi)^3 2p^0} \frac{d^3 \vec{\mathbf{k}}}{(2\pi)^3 2\omega}, \qquad (I1)
$$

where M_1 is sum of all diagrams with one scattering center from Appendix D, M_2 is sum of all diagrams with two scattering centers in the contact limit from Appendices E,F,G and M_0^* is obtained from Appendix C.

The final results from Appendix D yield:

$$
M_1 = M_{1,1,0} + M_{1,0,0} + M_{1,0,1} = J_a(p+k) e^{i(p+k)x_0} (1 - x + x^2) T_{a_1}(-i) \int \frac{d^2 \mathbf{q}_1}{(2\pi)^2} v(0, \mathbf{q}_1) e^{-i\mathbf{q}_1 \cdot \mathbf{b}_1} (2ig_s) \times
$$
$$
\times \left\{ \left(\frac{\boldsymbol{\epsilon} \cdot (\mathbf{k} - \mathbf{q}_1)}{(\mathbf{k} - \mathbf{q}_1)^2} [T^c, T^{a_1}]_{da} - \frac{\boldsymbol{\epsilon} \cdot (\mathbf{k} - x\mathbf{q}_1)}{(\mathbf{k} - x\mathbf{q}_1)^2} (T^c T^{a_1})_{da} + \frac{\boldsymbol{\epsilon} \cdot \mathbf{k}}{\mathbf{k}^2} (T^{a_1} T^c)_{da} \right) e^{\frac{i}{2\omega}(\mathbf{k}^2 + \frac{x}{1-x}(\mathbf{k} - \mathbf{q}_1)^2)(z_1 - z_0)} -
$$
$$
- \frac{\boldsymbol{\epsilon} \cdot (\mathbf{k} - \mathbf{q}_1)}{(\mathbf{k} - \mathbf{q}_1)^2} [T^c, T^{a_1}]_{da} e^{\frac{i}{2\omega}(\mathbf{k}^2 - (\mathbf{k} - \mathbf{q}_1)^2)(z_1 - z_0)} - \frac{\boldsymbol{\epsilon} \cdot \mathbf{k}}{\mathbf{k}^2} (T^{a_1} T^c)_{da} e^{\frac{i}{2\omega} \frac{x}{1-x}((\mathbf{k} - \mathbf{q}_1)^2 - \mathbf{k}^2)(z_1 - z_0)} \right\}, \qquad (I2)
$$

leading to:

$$
\frac{1}{d_T} \operatorname{Tr} \langle |M_1|^2 \rangle = N |J(p+k)|^2 (4g_s^2) \frac{1}{A_\perp} (1 - x + x^2)^2 \int \frac{d^2 \mathbf{q}_1}{(2\pi)^2} |v(0, \mathbf{q}_1)|^2 \frac{C_2(T)}{d_G} \left\{ \left(\frac{\boldsymbol{\epsilon} \cdot (\mathbf{k} - x\mathbf{q}_1)}{(\mathbf{k} - x\mathbf{q}_1)^2} \right)^2 \operatorname{Tr}((T^c)^2 (T^{a_1})^2) + \right.
$$
$$
+ 2\alpha \left(2 \frac{\boldsymbol{\epsilon} \cdot (\mathbf{k} - \mathbf{q}_1)}{(\mathbf{k} - \mathbf{q}_1)^2} - \frac{\boldsymbol{\epsilon} \cdot \mathbf{k}}{\mathbf{k}^2} - \frac{\boldsymbol{\epsilon} \cdot (\mathbf{k} - x\mathbf{q}_1)}{(\mathbf{k} - x\mathbf{q}_1)^2} \right) \frac{\boldsymbol{\epsilon} \cdot (\mathbf{k} - \mathbf{q}_1)}{(\mathbf{k} - \mathbf{q}_1)^2} - \alpha \frac{\boldsymbol{\epsilon} \cdot \mathbf{k}}{\mathbf{k}^2} \frac{\boldsymbol{\epsilon} \cdot (\mathbf{k} - \mathbf{q}_1)}{(\mathbf{k} - \mathbf{q}_1)^2} 2 \cos \left(\frac{\mathbf{k}^2 - (\mathbf{k} - \mathbf{q}_1)^2}{x(1-x)E^+} (z_1 - z_0) \right) +
$$
$$
+ 2 \left(\frac{\boldsymbol{\epsilon} \cdot \mathbf{k}}{\mathbf{k}^2} \operatorname{Tr}((T^c)^2 (T^{a_1})^2) - \frac{\boldsymbol{\epsilon} \cdot (\mathbf{k} - x\mathbf{q}_1)}{(\mathbf{k} - x\mathbf{q}_1)^2} \operatorname{Tr}(T^c T^{a_1} T^c T^{a_1}) \right) \frac{\boldsymbol{\epsilon} \cdot \mathbf{k}}{\mathbf{k}^2} -
$$
$$
- 2\alpha \left(\frac{\boldsymbol{\epsilon} \cdot (\mathbf{k} - \mathbf{q}_1)}{(\mathbf{k} - \mathbf{q}_1)^2} - \frac{1}{2} \frac{\boldsymbol{\epsilon} \cdot \mathbf{k}}{\mathbf{k}^2} - \frac{1}{2} \frac{\boldsymbol{\epsilon} \cdot (\mathbf{k} - x\mathbf{q}_1)}{(\mathbf{k} - x\mathbf{q}_1)^2} \right) \frac{\boldsymbol{\epsilon} \cdot (\mathbf{k} - \mathbf{q}_1)}{(\mathbf{k} - \mathbf{q}_1)^2} 2 \cos \left(\frac{(\mathbf{k} - \mathbf{q}_1)^2}{x(1-x)E^+} (z_1 - z_0) \right) +
$$
$$
+ \left. \left(\alpha \frac{\boldsymbol{\epsilon} \cdot (\mathbf{k} - \mathbf{q}_1)}{(\mathbf{k} - \mathbf{q}_1)^2} - \frac{\boldsymbol{\epsilon} \cdot \mathbf{k}}{\mathbf{k}^2} \operatorname{Tr}((T^c)^2 (T^{a_1})^2) + \frac{\boldsymbol{\epsilon} \cdot (\mathbf{k} - x\mathbf{q}_1)}{(\mathbf{k} - x\mathbf{q}_1)^2} \operatorname{Tr}(T^c T^{a_1} T^c T^{a_1}) \right) \frac{\boldsymbol{\epsilon} \cdot \mathbf{k}}{\mathbf{k}^2} 2 \cos \left(\frac{\mathbf{k}^2}{x(1-x)E^+} (z_1 - z_0) \right) \right\}, \qquad (I3)
$$

where the number of scattering centers N comes from summation over scattering centers Eqs. (B2, B3), then $\alpha \equiv \operatorname{Tr}((T^c)^2 (T^{a_1})^2 - T^c T^{a_1} T^c T^{a_1})$, and we also used the definition of commutator, the fact that trace is invariant under the cyclic permutations, Eq. (A4) (with $i = j$ and $d_i = d_T$) and the relation $E^+ \approx 2E$. We verified that this result is also symmetric under the substitutions: $(p \leftrightarrow k, x \leftrightarrow (1-x), c \leftrightarrow d)$ when written in terms of structure constants.

Next, we summarize contact limits of all diagrams that contain two scattering centers from Appendices E-H and then take their ensemble average according to Eqs. (B2, B3, B4) in order to obtain M_2:

$$M_2 = M_{2,2,0}^c + M_{2,0,3}^c + M_{2,0,0}^c + \frac{1}{2}(M_{2,0,1}^c + M_{2,0,2}^c) = \frac{1}{2}NJ_a(p+k)e^{i(p+k)x_0}(-2ig_s)\frac{1}{A_\perp}(1-x+x^2)T_{a_2}T_{a_1}\times$$

$$\times \int \frac{d^2\mathbf{q}_1}{(2\pi)^2}|v(0,\mathbf{q}_1)|^2\Big\{\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2}\Big(e^{\frac{i}{2\omega}\frac{\mathbf{k}^2}{1-x}(z_1-z_0)}\big([[T^c,T^{a_2}],T^{a_1}]_{da} + [T^{a_2}T^{a_1},T^c]_{da}\big) - [[T^c,T^{a_2}],T^{a_1}]_{da}-$$

$$- (T^{a_2}T^{a_1}T^c)_{da}\Big) + \frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2}\Big(e^{\frac{i}{2\omega}\frac{\mathbf{k}^2}{1-x}(z_1-z_0)} - e^{\frac{i}{2\omega}\frac{\mathbf{k}^2-(\mathbf{k}-\mathbf{q}_1)^2}{1-x}(z_1-z_0)}\Big)\big((T^{a_2}[T^c,T^{a_1}])_{da} + (T^{a_1}[T^c,T^{a_2}])_{da}\big)\Big\}.$$

$$(I4)$$

Then, by multiplying the previous expression by M_0^*, we obtain:

$$\frac{2}{d_T}Re\,\mathrm{Tr}\,\langle M_2 M_0^*\rangle = N|J(p+k)|^2(4g_s^2)\frac{1}{A_\perp}(1-x+x^2)^2\int\frac{d^2\mathbf{q}_1}{(2\pi)^2}|v(0,\mathbf{q}_1)|^2\frac{C_2(T)}{d_G}\times$$

$$\times\Big\{(\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2})^2\Big(2\alpha\cos\big(\frac{\mathbf{k}^2}{x(1-x)E^+}(z_1-z_0)\big) - 2\alpha - \mathrm{Tr}((T^c)^2(T^{a_1})^2)\Big)-$$

$$- 2\alpha\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2}\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2}\Big(\cos\big(\frac{\mathbf{k}^2}{x(1-x)E^+}(z_1-z_0)\big) - \cos\big(\frac{\mathbf{k}^2-(\mathbf{k}-\mathbf{q}_1)^2}{x(1-x)E^+}(z_1-z_0)\big)\Big)\Big\},\quad (I5)$$

which can easily be verified to be symmetric to the exchange $(p\leftrightarrow k, x\leftrightarrow(1-x), c\leftrightarrow d)$, when written in terms of structure constants. By summing the expressions Eqs. (I3, I5) we obtain:

$$\frac{1}{d_T}\mathrm{Tr}\,\langle|M_1|^2\rangle + \frac{2}{d_T}Re\,\mathrm{Tr}\,\langle M_2 M_0^*\rangle = Nd_G|J(p+k)|^2(4g_s^2)\frac{C_2(T)}{d_G}C_2^2(G)\frac{1}{A_\perp}(1-x+x^2)^2\int\frac{d^2\mathbf{q}_1}{(2\pi)^2}|v(0,\mathbf{q}_1)|^2\times$$

$$\times\Big\{\Big(1-\cos\big(\frac{\mathbf{k}^2}{x(1-x)E^+}(z_1-z_0)\big)\Big)\Big(\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2} - \frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2}\Big)\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2} + \Big((\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2})^2 - (\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2})^2\Big)+$$

$$+ \Big(1-\cos\big(\frac{(\mathbf{k}-\mathbf{q}_1)^2}{x(1-x)E^+}(z_1-z_0)\big)\Big)\Big(2\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2} - \frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2} - \frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2}\Big)\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2}\Big\},\quad (I6)$$

which in the soft-gluon approximation coincides with massless limit of Eq. (82) from [25] and where we used the following equalities that are valid in adjoint representation: $\mathrm{Tr}(T^cT^{a_1}T^cT^{a_1}) = \frac{1}{2}C_2^2(G)d_G = \alpha = \frac{1}{2}\mathrm{Tr}((T^c)^2(T^{a_1})^2)$, which follow from Eqs. (A4-A9) and the commutator definition.

Since we are considering optically "thin" QCD plasma, it would be convenient to expand energy loss in powers of opacity, which is defined by the mean number of collisions in QCD medium [9]:

$$\bar{n} = \frac{L}{\lambda} = \frac{N\sigma_{el}}{A_\perp}, \quad (I7)$$

where the small transverse momentum transfer elastic cross section between the jet and the target partons is

taken from GW model (Eq. (6) from [9]), which in our case reads:

$$\frac{d\sigma_{el}}{d^2\mathbf{q}_1} = \frac{C_2(G)C_2(T)}{d_G}\frac{|v(0,\mathbf{q}_1)|^2}{(2\pi)^2}. \quad (I8)$$

Combining the Eqs. (I7, I8) we obtain:

$$\frac{L}{\lambda} = \frac{N}{A_\perp}\frac{C_2(G)C_2(T)}{4\pi d_G}\frac{(4\pi\alpha_s)^2}{\mu^2}. \quad (I9)$$

Next we incorporate Eq. (I9) in Eq. (I6), substitute obtained expression in Eq. (I1), keeping in mind that $\vec{\mathbf{p}}$ is 3D momentum of a final jet, and that we need to apply Eqs. (C12, C13). The single gluon radiation spectrum in the first order in opacity then becomes:

$$\frac{dN_g^{(1)}}{dx} = \frac{C_2(G)\alpha_s}{\pi}\frac{L}{\lambda}\frac{(1-x+x^2)^2}{x(1-x)}\int\frac{d^2\mathbf{q}_1}{\pi}\frac{\mu^2}{(\mathbf{q}_1^2+\mu^2)^2}\int\frac{d^2\mathbf{k}}{\pi}\Bigg\{-\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2}\Big(\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2}+\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2}-2\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2}\Big)\times$$

$$\times\int dz_1(1-\cos(\frac{(\mathbf{k}-\mathbf{q}_1)^2}{x(1-x)E^+}(z_1-z_0)))\frac{2}{L}e^{-\frac{2(z_1-z_0)}{L}}+\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2}\Big(\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2}-\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2}\Big)\times$$

$$\times\int dz_1(1-\cos(\frac{\mathbf{k}^2}{x(1-x)E^+}(z_1-z_0)))\frac{2}{L}e^{-\frac{2(z_1-z_0)}{L}}+\Big((\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2})^2-(\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2})^2\Big)\int dz_1\frac{2}{L}e^{-\frac{2(z_1-z_0)}{L}}\Bigg\},$$

$$(I10)$$

and the differential radiative energy loss $\frac{dE^{(1)}}{dx} \equiv \omega\frac{d^3N_g^{(1)}}{dx} \approx xE\frac{d^3N_g^{(1)}}{dx}$ acquires the form:

$$\frac{dE^{(1)}}{dx} = \frac{C_2(G)\alpha_s}{\pi}\frac{L}{\lambda}E\frac{(1-x+x^2)^2}{1-x}\int\frac{d^2\mathbf{q}_1}{\pi}\frac{\mu^2}{(\mathbf{q}_1^2+\mu^2)^2}\int\frac{d^2\mathbf{k}}{\pi}\Bigg\{-\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2}\Big(\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2}+\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2}-2\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2}\Big)\times$$

$$\times\int dz_1(1-\cos(\frac{(\mathbf{k}-\mathbf{q}_1)^2}{x(1-x)E^+}(z_1-z_0)))\frac{2}{L}e^{-\frac{2(z_1-z_0)}{L}}+\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2}\Big(\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2}-\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2}\Big)\times$$

$$\times\int dz_1(1-\cos(\frac{\mathbf{k}^2}{x(1-x)E^+}(z_1-z_0)))\frac{2}{L}e^{-\frac{2(z_1-z_0)}{L}}+\Big((\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2})^2-(\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2})^2\Big)\int dz_1\frac{2}{L}e^{-\frac{2(z_1-z_0)}{L}}\Bigg\},$$

$$(I11)$$

where we also assumed a simple exponential distribution $\frac{2}{L}e^{-2\frac{z_1-z_0}{L}}$ between the scattering centers (as in [25]). So we finally obtain:

$$\frac{dN_g^{(1)}}{dx} = \frac{C_2(G)\alpha_s}{\pi}\frac{L}{\lambda}\frac{(1-x+x^2)^2}{x(1-x)}\int\frac{d^2\mathbf{q}_1}{\pi}\frac{\mu^2}{(\mathbf{q}_1^2+\mu^2)^2}\int d\mathbf{k}^2\times$$

$$\times\Bigg\{\frac{(\mathbf{k}-\mathbf{q}_1)^2}{(\frac{4x(1-x)E}{L})^2+(\mathbf{k}-\mathbf{q}_1)^4}\Big(2-\frac{\mathbf{k}\cdot(\mathbf{k}-\mathbf{q}_1)}{\mathbf{k}^2}-\frac{(\mathbf{k}-\mathbf{q}_1)\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2}\Big)+$$

$$+\frac{\mathbf{k}^2}{(\frac{4x(1-x)E}{L})^2+\mathbf{k}^4}\Big(1-\frac{\mathbf{k}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2}\Big)+\Big(\frac{1}{(\mathbf{k}-x\mathbf{q}_1)^2}-\frac{1}{\mathbf{k}^2}\Big)\Bigg\},$$

$$(I12)$$

which is symmetric to the exchange of p and k gluons, and:

$$\frac{dE^{(1)}}{dx} = \frac{C_2(G)\alpha_s}{\pi}\frac{L}{\lambda}E\frac{(1-x+x^2)^2}{1-x}\int\frac{d^2\mathbf{q}_1}{\pi}\frac{\mu^2}{(\mathbf{q}_1^2+\mu^2)^2}\int d\mathbf{k}^2\times$$

$$\times\Bigg\{\frac{(\mathbf{k}-\mathbf{q}_1)^2}{(\frac{4x(1-x)E}{L})^2+(\mathbf{k}-\mathbf{q}_1)^4}\Big(2-\frac{\mathbf{k}\cdot(\mathbf{k}-\mathbf{q}_1)}{\mathbf{k}^2}-\frac{(\mathbf{k}-\mathbf{q}_1)\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2}\Big)+$$

$$+\frac{\mathbf{k}^2}{(\frac{4x(1-x)E}{L})^2+\mathbf{k}^4}\Big(1-\frac{\mathbf{k}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2}\Big)+\Big(\frac{1}{(\mathbf{k}-x\mathbf{q}_1)^2}-\frac{1}{\mathbf{k}^2}\Big)\Bigg\},$$

$$(I13)$$

which in soft-gluon approximation reduces to massless limit of Eq. (84) from [25].

Appendix J: Diagrams and radiative energy loss in finite T QCD medium

Next we recalculate the results from Appendices C-H when the gluon mass $m_g = \frac{\mu}{\sqrt{2}}$ is included, i.e. gluon propagator has the following form [27]:

- gluon propagator with mass m_g in Feynman gauge:

$$
\begin{array}{c} a,\mu \qquad \xrightarrow{p} \qquad b,\nu \end{array} = \frac{-i\delta_{ab}P_{\mu\nu}}{p^2 - m_g^2 + i\epsilon}, \qquad (J1)
$$

where $P_{\mu\nu}$, given by Eq. (12) from [27] (specifically $P_{\mu\nu} = g_{\mu\nu} - \frac{p_\mu p_\nu n^2 + n_\mu n_\nu p^2 - n_\mu p_\nu (np) - n_\nu p_\mu (np)}{n^2 p^2 - (np)^2}$), represents the transverse projector. Note that, since the transverse projectors act directly or indirectly on transverse polarization vectors one may immediately replace $P_{\mu\nu}$ with $g_{\mu\nu}$ in gluon propagators, in order to facilitate the calculations.

This observation is obvious for off-shell gluon propagator, whereas the derivation for the remaining internal gluon lines is straightforward.

Consistently throughout this section, initial jet propagates long z-axis, 4-momentum is conserved and minus Light cone coordinate of p and k momenta acquire an additional term $+m_g^2$ in the numerator compared to massless case (Appendices C-H), due to relations $k^2 = p^2 = m_g^2$, while the polarizations remain the same.

We provide only the final expressions for all 11 Feynman diagrams beyond soft-gluon approximation, when the gluon mass is included, since its derivation is similar to the case of massless gluons and in order to avoid unnecessary repetition (Appendices C-H).

Thus, for M_0 we obtain:

$$
M_0 = J_a(p+k)e^{i(p+k)x_0}(-2ig_s)(1-x+x^2)\times
$$
$$
\times \frac{\boldsymbol{\epsilon} \cdot \mathbf{k}}{\mathbf{k}^2 + m_g^2(1-x+x^2)}(T^c)_{da}. \qquad (J2)
$$

The expression for $M_{1,1,0}$ now reads:

$$
M_{1,1,0} = J_a(p+k)e^{i(p+k)x_0}(-i)(1-x+x^2)(T^cT^{a_1})_{da}T_{a_1}\int \frac{d^2\mathbf{q}_1}{(2\pi)^2}v(0,\mathbf{q}_1)e^{-i\mathbf{q}_1\cdot\mathbf{b}_1}\times
$$
$$
\times(-2ig_s)\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2+m_g^2(1-x+x^2)}e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1)^2+\frac{m_g^2(1-x+x^2)}{1-x})(z_1-z_0)}, \qquad (J3)
$$

which differs from Eq. (D12) in the term $\chi \equiv m_g^2(1-x+x^2)$, which now appears in the denominator and in exponent, accompanying the squared transverse momentum.

Further on, we will use the shorthand notation χ.

Similarly, for $M_{1,0,0}$ and $M_{1,0,1}$ we obtain, respectively:

$$
M_{1,0,0} = J_a(p+k)e^{i(p+k)x_0}(-i)(1-x+x^2)(T^{a_1}T^c)_{da}T_{a_1}\int \frac{d^2\mathbf{q}_1}{(2\pi)^2}v(0,\mathbf{q}_1)e^{-i\mathbf{q}_1\cdot\mathbf{b}_1}\times
$$
$$
\times(2ig_s)\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2+\chi}\left(e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1)^2+\frac{\chi}{1-x})(z_1-z_0)} - e^{-\frac{i}{2\omega}\frac{x}{1-x}(\mathbf{k}^2-(\mathbf{k}-\mathbf{q}_1)^2)(z_1-z_0)}\right), \qquad (J4)
$$

$$
M_{1,0,1} = J_a(p+k)e^{i(p+k)x_0}(-i)(1-x+x^2)[T^c,T^{a_1}]_{da}T_{a_1}\int \frac{d^2\mathbf{q}_1}{(2\pi)^2}v(0,\mathbf{q}_1)e^{-i\mathbf{q}_1\cdot\mathbf{b}_1}\times
$$
$$
\times(2ig_s)\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2+\chi}\left(e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1)^2+\frac{\chi}{1-x})(z_1-z_0)} - e^{\frac{i}{2\omega}(\mathbf{k}^2-(\mathbf{k}-\mathbf{q}_1)^2)(z_1-z_0)}\right). \qquad (J5)
$$

Proceeding in the similar manner, we obtain the following expressions for contact-limit diagrams which include

interactions with two scattering centers:

$$M_{2,2,0}^c = -J_a(p+k)e^{i(p+k)x_0}(T^c T^{a_2} T^{a_1})_{da} T_{a_2} T_{a_1}(1-x+x^2)(-i)\int \frac{d^2\mathbf{q}_1}{(2\pi)^2}(-i)\int \frac{d^2\mathbf{q}_2}{(2\pi)^2} v(0,\mathbf{q}_1)v(0,\mathbf{q}_2)e^{-i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{b}_1} \times$$
$$\times \frac{1}{2}(2ig_s)\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x(\mathbf{q}_1+\mathbf{q}_2))}{(\mathbf{k}-x(\mathbf{q}_1+\mathbf{q}_2))^2+\chi}e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2+\frac{\chi}{1-x})(z_1-z_0)}, \tag{J6}$$

$$M_{2,0,3}^c = J_a(p+k)e^{i(p+k)x_0}[[T^c,T^{a_2}],T^{a_1}]_{da} T_{a_2} T_{a_1}(1-x+x^2)(-i)\int \frac{d^2\mathbf{q}_1}{(2\pi)^2}(-i)\int \frac{d^2\mathbf{q}_2}{(2\pi)^2} v(0,\mathbf{q}_1)v(0,\mathbf{q}_2)e^{-i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{b}_1} \times$$
$$\times \frac{1}{2}(2ig_s)\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)}{(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2+\chi}\left(e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2+\frac{\chi}{1-x})(z_1-z_0)} - e^{\frac{i}{2\omega}(\mathbf{k}^2-(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2)(z_1-z_0)}\right), \tag{J7}$$

$$M_{2,0,0}^c = J_a(p+k)e^{i(p+k)x_0}(T^{a_2} T^{a_1} T^c)_{da} T_{a_2} T_{a_1}(1-x+x^2)(-i)\int \frac{d^2\mathbf{q}_1}{(2\pi)^2}(-i)\int \frac{d^2\mathbf{q}_2}{(2\pi)^2} v(0,\mathbf{q}_1)v(0,\mathbf{q}_2)e^{-i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{b}_1} \times$$
$$\times \frac{1}{2}(2ig_s)\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2+\chi}\left(e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2+\frac{\chi}{1-x})(z_1-z_0)} - e^{\frac{i}{2\omega}\frac{x}{1-x}((\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2-\mathbf{k}^2)(z_1-z_0)}\right), \tag{J8}$$

$$M_{2,0,1}^c = J_a(p+k)e^{i(p+k)x_0}(T^{a_2}[T^c,T^{a_1}])_{da} T_{a_2} T_{a_1}(1-x+x^2)(-i)\int \frac{d^2\mathbf{q}_1}{(2\pi)^2}(-i)\int \frac{d^2\mathbf{q}_2}{(2\pi)^2} v(0,\mathbf{q}_1)v(0,\mathbf{q}_2)e^{-i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{b}_1} \times$$
$$\times (2ig_s)\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2+\chi}\left(e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2+\frac{\chi}{1-x})(z_1-z_0)} - e^{\frac{i}{2\omega}(\mathbf{k}^2-\frac{(\mathbf{k}-\mathbf{q}_1)^2}{1-x}+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2)(z_1-z_0)}\right), \tag{J9}$$

$$M_{2,0,2}^c = J_a(p+k)e^{i(p+k)x_0}(T^{a_1}[T^c,T^{a_2}])_{da} T_{a_2} T_{a_1}(1-x+x^2)(-i)\int \frac{d^2\mathbf{q}_1}{(2\pi)^2}(-i)\int \frac{d^2\mathbf{q}_2}{(2\pi)^2} v(0,\mathbf{q}_1)v(0,\mathbf{q}_2)e^{-i(\mathbf{q}_1+\mathbf{q}_2)\cdot\mathbf{b}_1} \times$$
$$\times (2ig_s)\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_2)}{(\mathbf{k}-\mathbf{q}_2)^2+\chi}\left(e^{\frac{i}{2\omega}(\mathbf{k}^2+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2+\frac{\chi}{1-x})(z_1-z_0)} - e^{\frac{i}{2\omega}(\mathbf{k}^2-\frac{(\mathbf{k}-\mathbf{q}_2)^2}{1-x}+\frac{x}{1-x}(\mathbf{k}-\mathbf{q}_1-\mathbf{q}_2)^2)(z_1-z_0)}\right). \tag{J10}$$

The amplitudes $M_{2,1,0}^c$ and $M_{2,1,1}^c$ are omitted as they are suppressed compared to the remaining amplitudes.

After adding Eqs. (J3, J4, J5), we obtain:

$$\frac{1}{d_T}\operatorname{Tr}\left\langle |M_1|^2\right\rangle = N|J(p+k)|^2(4g_s^2)\frac{1}{A_\perp}(1-x+x^2)^2\int\frac{d^2\mathbf{q}_1}{(2\pi)^2}|v(0,\mathbf{q}_1)|^2\frac{C_2(T)}{d_G}\left\{\left(\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2+\chi}\right)^2\operatorname{Tr}((T^c)^2(T^{a_1})^2)+\right.$$

$$+2\alpha\left(2\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2+\chi}-\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2+\chi}-\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2+\chi}\right)\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2+\chi}-\alpha\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2+\chi}\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2+\chi}2\cos\left(\frac{\mathbf{k}^2-(\mathbf{k}-\mathbf{q}_1)^2}{x(1-x)E^+}(z_1-z_0)\right)$$

$$+2\left(\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2+\chi}\operatorname{Tr}((T^c)^2(T^{a_1})^2)-\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2+\chi}\operatorname{Tr}(T^cT^{a_1}T^cT^{a_1})\right)\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2+\chi}-$$

$$-2\alpha\left(\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2+\chi}-\frac{1}{2}\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2+\chi}-\frac{1}{2}\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2+\chi}\right)\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2+\chi}2\cos\left(\frac{(\mathbf{k}-\mathbf{q}_1)^2+\chi}{x(1-x)E^+}(z_1-z_0)\right)+$$

$$+\left.\left(\alpha\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2+\chi}-\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2+\chi}\operatorname{Tr}((T^c)^2(T^{a_1})^2)+\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2+\chi}\operatorname{Tr}(T^cT^{a_1}T^cT^{a_1})\right)\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2+\chi}2\cos\left(\frac{\mathbf{k}^2+\chi}{x(1-x)E^+}(z_1-z_0)\right)\right\}.$$

$$(J11)$$

Likewise, after adding Eqs. (J6)-(J10), we obtain:

$$\frac{2}{d_T}Re\operatorname{Tr}\left\langle M_2M_0^*\right\rangle = N|J(p+k)|^2(4g_s^2)\frac{1}{A_\perp}(1-x+x^2)^2\int\frac{d^2\mathbf{q}_1}{(2\pi)^2}|v(0,\mathbf{q}_1)|^2\frac{C_2(T)}{d_G}\times$$

$$\times\left\{\left(\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2+\chi}\right)^2\left(2\alpha\cos\left(\frac{\mathbf{k}^2+\chi}{x(1-x)E^+}(z_1-z_0)\right)-2\alpha-\operatorname{Tr}((T^c)^2(T^{a_1})^2)\right)-\right.$$

$$\left.-2\alpha\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2+\chi}\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2+\chi}\left(\cos\left(\frac{\mathbf{k}^2+\chi}{x(1-x)E^+}(z_1-z_0)\right)-\cos\left(\frac{\mathbf{k}^2-(\mathbf{k}-\mathbf{q}_1)^2}{x(1-x)E^+}(z_1-z_0)\right)\right)\right\},$$

$$(J12)$$

leading to:

$$\frac{1}{d_T}\operatorname{Tr}\left\langle |M_1|^2\right\rangle + \frac{2}{d_T}Re\operatorname{Tr}\left\langle M_2M_0^*\right\rangle = Nd_G|J(p+k)|^2(4g_s^2)\frac{C_2(T)}{d_G}C_2^2(G)\frac{1}{A_\perp}(1-x+x^2)^2\int\frac{d^2\mathbf{q}_1}{(2\pi)^2}|v(0,\mathbf{q}_1)|^2\times$$

$$\times\left\{\left(1-\cos\left(\frac{\mathbf{k}^2+\chi}{x(1-x)E^+}(z_1-z_0)\right)\right)\left(\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2+\chi}-\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2+\chi}\right)\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2+\chi}+\left(\left(\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2+\chi}\right)^2-\left(\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2+\chi}\right)^2\right)+\right.$$

$$\left.+\left(1-\cos\left(\frac{(\mathbf{k}-\mathbf{q}_1)^2+\chi}{x(1-x)E^+}(z_1-z_0)\right)\right)\left(2\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2+\chi}-\frac{\boldsymbol{\epsilon}\cdot\mathbf{k}}{\mathbf{k}^2+\chi}-\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2+\chi}\right)\frac{\boldsymbol{\epsilon}\cdot(\mathbf{k}-\mathbf{q}_1)}{(\mathbf{k}-\mathbf{q}_1)^2+\chi}\right\}.$$

$$(J13)$$

In the soft-gluon approximation the previous expression coincides with Eq. (82) from [25] (note that contrary to the cited paper, we here consider gluon jet, so that M no longer denotes heavy quark mass, but instead $M\equiv m_g$ and therefore the term M^2x^2 is also negligible).

If we further apply the same procedure as in Appendix I, and again assume the simple exponential distribution $\frac{2}{L}e^{-2\frac{z_1-z_0}{L}}$ between the scattering centers, we obtain:

$$\frac{dN_g^{(1)}}{dx} = \frac{C_2(G)\alpha_s}{\pi}\frac{L}{\lambda}\frac{(1-x+x^2)^2}{x(1-x)}\int\frac{d^2\mathbf{q}_1}{\pi}\frac{\mu^2}{(\mathbf{q}_1^2+\mu^2)^2}\int d\mathbf{k}^2\times$$

$$\times\left\{\frac{(\mathbf{k}-\mathbf{q}_1)^2+\chi}{(\frac{4x(1-x)E}{L})^2+((\mathbf{k}-\mathbf{q}_1)^2+\chi)^2}\left(2\frac{(\mathbf{k}-\mathbf{q}_1)^2}{(\mathbf{k}-\mathbf{q}_1)^2+\chi}-\frac{\mathbf{k}\cdot(\mathbf{k}-\mathbf{q}_1)}{\mathbf{k}^2+\chi}-\frac{(\mathbf{k}-\mathbf{q}_1)\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2+\chi}\right)+\right.$$

$$\left.+\frac{\mathbf{k}^2+\chi}{(\frac{4x(1-x)E}{L})^2+(\mathbf{k}^2+\chi)^2}\left(\frac{\mathbf{k}^2}{\mathbf{k}^2+\chi}-\frac{\mathbf{k}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2+\chi}\right)+\left(\frac{(\mathbf{k}-x\mathbf{q}_1)^2}{((\mathbf{k}-x\mathbf{q}_1)^2+\chi)^2}-\frac{\mathbf{k}^2}{(\mathbf{k}^2+\chi)^2}\right)\right\}, \quad \text{(J14)}$$

which is symmetric to the exchange of p and k gluons, and which for $m_g \to 0$ coincides with Eq. (I12). Also:

$$\frac{dE^{(1)}}{dx} = \frac{C_2(G)\alpha_s}{\pi}\frac{L}{\lambda}E\frac{(1-x+x^2)^2}{1-x}\int\frac{d^2\mathbf{q}_1}{\pi}\frac{\mu^2}{(\mathbf{q}_1^2+\mu^2)^2}\int d\mathbf{k}^2\times$$

$$\times\left\{\frac{(\mathbf{k}-\mathbf{q}_1)^2+\chi}{(\frac{4x(1-x)E}{L})^2+((\mathbf{k}-\mathbf{q}_1)^2+\chi)^2}\left(2\frac{(\mathbf{k}-\mathbf{q}_1)^2}{(\mathbf{k}-\mathbf{q}_1)^2+\chi}-\frac{\mathbf{k}\cdot(\mathbf{k}-\mathbf{q}_1)}{\mathbf{k}^2+\chi}-\frac{(\mathbf{k}-\mathbf{q}_1)\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2+\chi}\right)+\right.$$

$$\left.+\frac{\mathbf{k}^2+\chi}{(\frac{4x(1-x)E}{L})^2+(\mathbf{k}^2+\chi)^2}\left(\frac{\mathbf{k}^2}{\mathbf{k}^2+\chi}-\frac{\mathbf{k}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2+\chi}\right)+\left(\frac{(\mathbf{k}-x\mathbf{q}_1)^2}{((\mathbf{k}-x\mathbf{q}_1)^2+\chi)^2}-\frac{\mathbf{k}^2}{(\mathbf{k}^2+\chi)^2}\right)\right\}, \quad \text{(J15)}$$

which, in soft-gluon approximation, reduces to Eq. (84) from [25], and which for $m_g \to 0$ coincides with our massless beyond soft-gluon approximation expression Eq. (I13).

Further, we display the beyond soft-gluon approxima-

tion expressions needed for numerical evaluation of the corresponding variables. So, the number of radiated gluons to the first order in opacity for gluons with effective mass m_g and for finite x reads:

$$N_g^{(1)} = \frac{C_2(G)\alpha_s}{\pi}\frac{L}{\lambda}\int_0^{\frac{1}{2}}dx\frac{(1-x+x^2)^2}{x(1-x)}\int\frac{d^2\mathbf{q}_1}{\pi}\frac{\mu^2}{(\mathbf{q}_1^2+\mu^2)^2}\int d\mathbf{k}^2\times$$

$$\times\left\{\frac{(\mathbf{k}-\mathbf{q}_1)^2+\chi}{(\frac{4x(1-x)E}{L})^2+((\mathbf{k}-\mathbf{q}_1)^2+\chi)^2}\left(2\frac{(\mathbf{k}-\mathbf{q}_1)^2}{(\mathbf{k}-\mathbf{q}_1)^2+\chi}-\frac{\mathbf{k}\cdot(\mathbf{k}-\mathbf{q}_1)}{\mathbf{k}^2+\chi}-\frac{(\mathbf{k}-\mathbf{q}_1)\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2+\chi}\right)+\right.$$

$$\left.+\frac{\mathbf{k}^2+\chi}{(\frac{4x(1-x)E}{L})^2+(\mathbf{k}^2+\chi)^2}\left(\frac{\mathbf{k}^2}{\mathbf{k}^2+\chi}-\frac{\mathbf{k}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2+\chi}\right)+\left(\frac{(\mathbf{k}-x\mathbf{q}_1)^2}{((\mathbf{k}-x\mathbf{q}_1)^2+\chi)^2}-\frac{\mathbf{k}^2}{(\mathbf{k}^2+\chi)^2}\right)\right\}. \quad \text{(J16)}$$

Similarly, the fractional radiative energy loss is obtained after numerically integrating the following expression:

$$\frac{\Delta E^{(1)}}{E} = \frac{C_2(G)\alpha_s}{\pi}\frac{L}{\lambda}\int_0^{\frac{1}{2}}dx\frac{(1-x+x^2)^2}{1-x}\int\frac{d^2\mathbf{q}_1}{\pi}\frac{\mu^2}{(\mathbf{q}_1^2+\mu^2)^2}\int d\mathbf{k}^2\times$$

$$\times\left\{\frac{(\mathbf{k}-\mathbf{q}_1)^2+\chi}{(\frac{4x(1-x)E}{L})^2+((\mathbf{k}-\mathbf{q}_1)^2+\chi)^2}\left(2\frac{(\mathbf{k}-\mathbf{q}_1)^2}{(\mathbf{k}-\mathbf{q}_1)^2+\chi}-\frac{\mathbf{k}\cdot(\mathbf{k}-\mathbf{q}_1)}{\mathbf{k}^2+\chi}-\frac{(\mathbf{k}-\mathbf{q}_1)\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2+\chi}\right)+\right.$$

$$\left.+\frac{\mathbf{k}^2+\chi}{(\frac{4x(1-x)E}{L})^2+(\mathbf{k}^2+\chi)^2}\left(\frac{\mathbf{k}^2}{\mathbf{k}^2+\chi}-\frac{\mathbf{k}\cdot(\mathbf{k}-x\mathbf{q}_1)}{(\mathbf{k}-x\mathbf{q}_1)^2+\chi}\right)+\left(\frac{(\mathbf{k}-x\mathbf{q}_1)^2}{((\mathbf{k}-x\mathbf{q}_1)^2+\chi)^2}-\frac{\mathbf{k}^2}{(\mathbf{k}^2+\chi)^2}\right)\right\}. \quad \text{(J17)}$$

[1] U.A. Wiedemann, Nucl. Phys. B **588**, 303 (2000).

[2] C.A. Salgado and U.A. Wiedemann, Phys. Rev. D **68**, 014008 (2003).

[3] N. Armesto, C.A. Salgado and U.A. Wiedemann, Phys. Rev. D **69**, 114003 (2004).

[4] R. Baier, Y.L. Dokshitzer, A.H. Mueller, S. Peigne and D. Schiff, Nucl. Phys. B **483**, 291 (1997).

[5] R. Baier, Y.L. Dokshitzer, A.H. Mueller, S. Peigne and D. Schiff, Nucl. Phys. B **484**, 265 (1997).

[6] B.G. Zakharov, JETP Lett. **63**, 952 (1996).

[7] B.G. Zakharov, JETP Lett. **65**, 615 (1997).

[8] M. Gyulassy, P. Levai and I. Vitev, Nucl. Phys. B **571**, 197 (2000).

[9] M. Gyulassy, P. Levai and I. Vitev, Nucl. Phys. B **594**, 371 (2001).

[10] X.N. Wang and X.F. Guo, Nucl. Phys. A **696**, 788 (2001).

[11] A. Majumder and M. Van Leeuwen, Prog. Part. Nucl. Phys. A **66**, 41 (2011).

[12] L. Apolinario, N. Armesto and C.A. Salgado, Phys. Lett. B **718**, 160-168 (2012).

[13] B.W. Zhang and X.N. Wang, Nucl.Phys. A **720**, 429-451 (2003).

[14] G. Ovanesyan and I. Vitev, JHEP **1106**, 080 (2011); Phys. Lett. B **706**, 371 (2012).

[15] M. Djordjevic and U. Heinz, Phys. Rev. Lett. **101**, 022302 (2008).

[16] M. Djordjevic, Phys. Rev. C **80**, 064909 (2009).

[17] M. Djordjevic and M. Djordjevic, Phys. Lett. B **734**, 286 (2014).

[18] D. d'Enterria and B. Betz, Lect. Notes Phys. **785** , 285 (2010).

[19] A. Dainese, Eur. Phys. J. C **33**, 495 (2004).

[20] M. Djordjevic and M. Djordjevic, Phys. Rev. C **90**, 034910 (2014).

[21] M. Djordjevic, M. Djordjevic and B. Blagojevic, Phys. Lett. B **737**, 298 (2014).

[22] M. Djordjevic, Phys. Rev. Lett. **112**, 042302 (2014).

[23] M. Djordjevic and M. Djordjevic, Phys. Rev. C **92**, no. 2, 024918 (2015).

[24] M. Djordjevic, B. Blagojevic and L. Zivkovic, Phys. Rev. C **94**, no. 4, 044908 (2016).

[25] M. Djordjevic and M. Gyulassy, Nucl. Phys. A **733**, 265-298 (2004).

[26] M. Gyulassy and X.N. Wang, Nucl. Phys. B **420**, 583 (1994).

[27] M. Djordjevic and M. Gyulassy, Phys. Rev. C **68**, 034914 (2003).

[28] J.D. Bjorken: FERMILAB-PUB-82-059-THY, 287-292 (1982).

[29] M. Gyulassy, P. Levai and I. Vitev, Phys. Lett. B **538**, 282 (2002).

[30] S. Wicks (2008), arXiv:0804.4704.

[31] G. Altarelli and G. Parisi, Nucl. Phys. B **126**, 298 (1977).

[32] B. Betz and M. Gyulassy, Phys. Rev. C **86**, 024903 (2012).

[33] S. Wicks, W. Horowitz, M. Djordjevic and M. Gyulassy, Nucl. Phys. A **784**, 426 (2007).

[34] Z.B. Kang, I. Vitev and H. Xing, Phys. Lett. B **718**, 482 (2012).

[35] R. Sharma, I. Vitev and B.W. Zhang, Phys. Rev. C **80**, 054902 (2009).

[36] R.D. Field, *Applications of Perturbative QCD* (Perseus Books, Cambridge, Massachussetts, 1995).

[37] M. Djordjevic and M. Djordjevic, Phys. Lett. B **709**, 229-233 (2012).

[38] J. Ellis, *TikZ-Feynman: Feynman diagrams with TikZ*, arXiv:hep-ph/1601.05437 (2016).

[39] M. Gyulassy, P.Levai and I. Vitev, Nucl. Phys. B **661**, 637c (1999).

Systems Biology of Bacterial Immune Systems: Regulation of Restriction-Modification and CRISPR-Cas Systems

Andjela Rodic, Bojana Blagojevic, and Marko Djordjevic

Contents

Abstract Restriction-modification (R-M) and CRISPR-Cas are bacterial immune systems which defend their prokaryotic hosts from invasive DNA. Understanding how these systems are regulated is necessary for both biotechnology applications, and for understanding how they modulate horizontal gene transfer (including acquisition of virulence factors). We here review results on modeling these systems which point to common general principles underlying their architecture and dynamical response, with particular emphasis on modeling methods. We show that the modeling predictions are in a good agreement with both in vitro measurements

A. Rodic
Institute of Physiology and Biochemistry, Faculty of Biology, University of Belgrade, Belgrade, Serbia

Multidisciplinary PhD program in Biophysics, University of Belgrade, Belgrade, Serbia

M. Djordjevic (✉)
Institute of Physiology and Biochemistry, Faculty of Biology, University of Belgrade, Belgrade, Serbia
e-mail: dmarko@bio.bg.ac.rs

B. Blagojevic
Institute of Physics, University of Belgrade, Belgrade, Serbia

of promoter transcription activity and the first in vivo measurements of gene expression dynamics in R-M systems. Modeling induction of CRISPR-Cas systems is challenging, as signaling which leads to their activation is currently unknown. However, based on similarities between transcription regulation in CRISPR-Cas and some R-M systems, we argue that transcription regulation of much simpler (and better studied) R-M systems can be used as a proxy for CRISPR-Cas transcription regulation, allowing to in silico assess CRISPR-Cas dynamical properties. Based on the obtained results, we propose that mechanistically otherwise different bacterial immune systems, presumably due to a common function, share the same unifying principles governing their expression dynamics.

1   Introduction

Two types of prokaryotic "immune systems," known as restriction-modification (R-M) and CRISPR-Cas (*C*lustered, *r*egularly *i*nterspaced *s*hort *p*alindromic *r*epeats-*C*RISPR-*a*ssociated *p*roteins) systems, resemble the mammalian immune system in their ability to actively and with high selectivity combat infectious elements (foreign DNA) (Goldberg and Marraffini 2015). Apart from their immune function, these systems significantly influence evolution and ecology of prokaryotes in a number of ways and have a range of applications in biotechnology (Ershova et al. 2015; Hille and Charpentier 2016).

In type II R-M systems, which are often found on plasmids, separate genes code for two main system components: a restriction enzyme, which cuts specific DNA sequences, and a methyltransferase, which methylates the same sequences and thereby protects them from cutting (Nagornykh et al. 2008). It is widely considered that the main condition for safely and efficiently establishing an R-M system in a naïve host cell, is a delayed beginning of expression of restriction enzyme with respect to methyltransferase. This delay provides enough time for a methyltransferase to protect a host genome, so that restriction enzyme later targets only invasive DNA. Apart from this constraint on their dynamics imposed by their function, we propose other potentially common R-M system dynamical properties, and ask if these can be achieved by a wide variety of R-M systems architectures and regulatory features (Rodic et al. 2017b). These hypotheses are tested by analyzing dynamical properties of different R-M systems, predicted by biophysical models including thermodynamically modeled transcription regulation and dynamically modeled transcript and protein expression.

Unlike R-M systems, which are considered rudimentary for their lack of ability to memorize past infections, CRISPR-Cas are advanced, adaptive prokaryotic immune systems, which store partial DNA sequences of former infectors as spacers flanked by direct repeats in a so-called CRISPR array (Hille and Charpentier 2016). Another constitutive part of a CRISPR-Cas system are genes coding for Cas proteins. In Type

I-E CRISPR-Cas system in *E. coli*, which is a model system for studying CRISPR-Cas regulation, CRISPR array is transcribed as a long pre-crRNA molecule which is further cut by Cas6e protein into small crRNAs, containing separate spacers. These crRNAs guide Cascade complexes constituted of Cas proteins to complementary foreign DNA, which is consequently destroyed. Somewhat surprisingly, while CRISPR-Cas is extensively used for designing various biotechnological tools, its native function and regulation in bacterial cells are not well understood. In particular, CRISPR-Cas is silenced in *E. coli* cells under standard conditions, which hinders observing its expression dynamics (Pul et al. 2010). However, transcription regulation of this system involves general features similar to those found in certain R-M systems, which can be used to predict the main features of CRISPR-Cas expression dynamics (Rodic et al. 2017a).

In this chapter, we aim to explain how a thermodynamic model of a given promoter regulation is formulated, by briefly describing a theoretical basis of thermodynamic modeling and showing how this approach is applied on examples of R-M systems, AhdI and EcoRV. Further, thermodynamic modeling of transcription is used as an input for dynamic modeling, predicting appropriate protein expression in a cell in time, which is discussed on the example of Esp1396I R-M system, for which protein expression dynamics were experimentally measured. We also show how measures for dynamical properties of interest were defined to compare expression dynamics of different R-M systems and to propose unifying principles that characterize their regulatory dynamics. To in silico predict the main qualitative properties of CRISPR-Cas dynamics, and to understand the significance of few characteristic regulatory features found in CRISPR-Cas, we introduce the idea of using a synthetic setup where R-M system transcription regulation with similar features is used as a proxy for not–well understood CRISPR-Cas transcription regulation. Based on the obtained results, we propose that regulatory dynamics of CRISPR-Cas and R-M systems may be governed by similar design principles imposed by their immune function.

2   Thermodynamic Modeling of Transcription Regulation

Thermodynamic modeling approach of gene transcription control is based on principles of statistical mechanics. As an input it takes levels of transcription factors, and patterns and affinities of their binding sites, while as an output it provides predictions of promoter transcription activity (Dresch et al. 2013).

As regulation of transcription initiation, which is a rate-limiting step in gene transcription, involves binding of protein molecules (RNA polymerase, transcription factors) to DNA (promoter region), let us start with a simple scenario in which one molecule of protein, present in some copy number in a cell, binds to one binding site on DNA. From a thermodynamics point of view, the cell interior can be approximated by a system exchanging energy with a much larger heat reservoir (its surroundings) (Phillips et al. 2012). Protein molecules in this system, among

which energy is distributed, are approximated by noninteracting particles randomly moving in space confined to the cell volume. These particles can be arranged in a number of different ways, and every unique arrangement of particles corresponds to a particular *microstate* of the system. The probability of finding different microstates is given by the Boltzmann distribution, which we derive below.

2.1 Derivation of the Boltzmann Distribution

Consider a system (s) in contact with a thermal reservoir (r), which together constitute an isolated system with fixed total energy $E = E^{(s)} + E^{(r)}$. According to the second law of thermodynamics, such an isolated system evolves toward such partition of energy between the system and the reservoir, which corresponds to the largest number of microstates of the whole system (Phillips et al. 2012). Therefore, the probability that the system has energy $E_i^{(s)}$ is proportional to the number of the corresponding microstates of the overall system, $\Omega(E, E_i^{(s)}) = \Omega^{(s)}(E_i^{(s)}) \times \Omega^{(r)}(E - E_i^{(s)})$. System degeneracy is directly related to its entropy $S = k_B \ln(\Omega)$, where k_B is the Boltzmann constant, so the probability that the system has energy $E_i^{(s)}$ reads:

$$
\begin{aligned}
P\left(E_i^{(s)}\right) &\propto \exp\left(S^{(s)}\left(E_i^{(s)}\right)/k_B\right) \cdot \exp\left(S^{(r)}\left(E - E_i^{(s)}\right)/k_B\right) \\
&\approx \exp\left(S^{(s)}\left(E_i^{(s)}\right)/k_B\right) \cdot \exp\left(\left(S^{(r)}(E) - \frac{dS^{(r)}}{dE} \cdot E_i^{(s)}\right)/k_B\right) \\
&\propto \exp\left(S^{(s)}\left(E_i^{(s)}\right)/k_B\right) \cdot \exp\left(-E_i^{(s)}/(k_B \cdot T)\right),
\end{aligned}
\tag{1}
$$

where in the second step, the reservoir entropy is expended about $S^{(r)}(E)$ (note that this approximation is valid when a reservoir is much bigger than a system, so $E_i^{(s)} << E$), while in the third step the thermodynamic definition of temperature $(\partial S/\partial E)_{V,N} = 1/T$ is used. The first term in Eq. (1) gives the number of microstates of a system with energy $E_i^{(s)}$ (i.e., $\Omega^{(s)}(E_i^{(s)})$), while the second term is called the Boltzmann factor, and represents the unnormalized probability of selecting one particular system microstate at energy $E_i^{(s)}$, i.e. it represents a statistical weight of that microstate (Sneppen and Zocchi 2005).

2.2 Statistical Weights from Statistical Mechanics

In the problem of binding of a protein to its binding site considered above, all of the microstates can be grouped in one of the two system *macrostates*: the one in which the DNA binding site is occupied by the protein, or the one in which it is empty, where binding sites in these two states are characterized by the energies $\varepsilon_i^{(bs)}$ (so that i corresponds to *bound* or *unbound*). Thereby, the energy of the system $(E_i^{(s)})$ is a sum of the binding site energy and the kinetic energies of all unbound

protein molecules. Since the probability of finding different microstates is given by the Boltzmann distribution, the weight associated with the macrostate with energy $E_i^{(s)}$ is proportional to the corresponding number of the system microstates ($\Omega^{(s)}$), multiplied by the Boltzmann factor (the numerator in the equation below):

$$P\left(E_i^{(s)}\right) = \frac{\Omega^{(s)}\left(E_i^{(s)}\right) \cdot \mathrm{e}^{-E_i^{(s)}/(k_B \cdot T)}}{\sum_i \left(\Omega^{(s)}\left(E_i^{(s)}\right) \cdot \mathrm{e}^{-E_i^{(s)}/(k_B \cdot T)}\right)}. \tag{2}$$

In the denominator of Eq. (2) is the so-called *partition function,* which represents a sum of statistical weights of all possible system microstates.

To determine $\Omega^{(s)}$ from Eq. (2), i.e. to count in how many ways protein molecules can be arranged, one needs to know how many states are available to one freely moving protein molecule with kinetic energy $\varepsilon_k = p^2/(2m)$ in a cell. According to the uncertainty principle from quantum mechanics, this question amounts to counting discrete cells of the size h (Planck's constant) in the phase-space containing three dimensions of particle position (r) and three dimensions of its momentum (p) (Stowe 2007; Sneppen and Zocchi 2005).

Therefore, the statistical weight of the system macrostate with binding site energy $\varepsilon_{bound}^{(bs)}$, where the protein binding site is occupied, is obtained by summing through all possible arrangements (permutations) of $N-1$ indistinguishable protein molecules (because 1 is bound) in a cell phase-space, with that sum weighted by a corresponding Boltzmann factor (Phillips et al. 2012; Sneppen and Zocchi 2005):

$$Z_{ON} = \frac{1}{(N-1)!}\left(\int_V \int \frac{\mathrm{d}^3 r \cdot \mathrm{d}^3 p}{h^3} \mathrm{e}^{-p^2/(2mk_B T)}\right)^{N-1} \mathrm{e}^{-\varepsilon_{bound}^{(bs)}/(k_B T)}$$
$$\propto k^{N-1}\rho^{-(N-1)}\mathrm{e}^{-\varepsilon_{bound}^{(bs)}/(k_B T)}, \tag{3}$$

where $k = (2mk_B T\pi/h^2)^{3/2}$ and $\rho = N/V$ (V is cell volume). Equivalently, a statistical weight of a system macrostate in which all protein molecules are free in a cell (with binding site energy $\varepsilon_{unbound}^{(bs)}$) reads:

$$Z_{OFF} = \frac{1}{N!}\left(\int_V \int \frac{\mathrm{d}^3 r \cdot \mathrm{d}^3 p}{h^3} \mathrm{e}^{-p^2/(2mk_B T)}\right)^{N} \mathrm{e}^{-\varepsilon_{unbound}^{(bs)}/(k_B T)}$$
$$\propto k^{N}\rho^{-N}\mathrm{e}^{-\varepsilon_{unbound}^{(bs)}/(k_B T)}. \tag{4}$$

Taking into account that the total statistical weight (partition function) of this system is $Z = Z_{ON} + Z_{OFF}$, one can express the ratio of probabilities of finding a binding site in occupied and unoccupied state:

$$\frac{P_{ON}}{P_{OFF}} = \frac{Z_{ON}}{Z} \cdot \left(\frac{Z_{OFF}}{Z}\right)^{-1} = \frac{\rho}{k}\mathrm{e}^{-\Delta\varepsilon/(k_B T)}, \tag{5}$$

where $\Delta\varepsilon = \varepsilon_{bound}^{(bs)} - \varepsilon_{unbound}^{(bs)}$ is the binding energy. Conveniently, statistical weights are expressed in terms of $\Delta\varepsilon$ (i.e., normalized with Z_{OFF}). One should have in mind that binding of a protein to DNA induces significant conformational changes in both molecules, so $\Delta\varepsilon$ in the above equations corresponds to the (Gibbs) free energy of binding (often written as ΔG, which we will adopt below).

2.3 Statistical Weights from Equilibrium Biochemical Reactions

Binding of a protein present in a cell in concentration [P], to a binding site of concentration [BS] is, alternatively, described by the following chemical reaction:

$$[P] + [BS] \xrightleftharpoons[K_d]{} [P \sim BS], \tag{6}$$

characterized by the equilibrium dissociation constant $K_d = [P]\cdot[BS]/[P \sim BS]$. The ratio of probabilities of finding a binding site occupied and unoccupied is then

$$\frac{P_{ON}}{P_{OFF}} = \frac{[P \sim BS]}{[BS]_{tot}} \cdot \left(\frac{[BS]}{[BS]_{tot}}\right)^{-1} = \frac{[P]}{K_d}, \tag{7}$$

where $[BS]_{tot} = [BS] + [P \sim BS]$ is a total binding site concentration. Equation (7) is equivalent to Eq. (5) obtained using statistical mechanics, where $[P] = \rho$ and $K_d = k\cdot\exp(\Delta\varepsilon/(k_BT))$ (Sneppen and Zocchi 2005).

   If a protein from the analyzed example is RNA polymerase (RNAP) binding to a promoter site, the promoter transcription activity can be approximated through a classical assumption that the transcription activity is proportional to equilibrium binding probability of RNAP to the promoter (Shea and Ackers 1985). Transcription from promoters with more complex regulation, including combinatorial binding of multiple transcription factors which results in more than two promoter configurations, can also be modeled in this way, as in the following example.

2.4 Modeling Transcription Regulation of AhdI R-M System

Thermodynamic modeling approach introduced above was applied in modeling transcription regulation of the R-M system AhdI, which belongs to a large group of R-M systems coding for an additional, control protein (C) which regulates transcription of system genes (Bogdanova et al. 2008). In this system, an operon containing control protein and restriction endonuclease genes (*c* and *res*), and a gene coding for methyltransferase (*met*) are oriented convergently and transcribed from
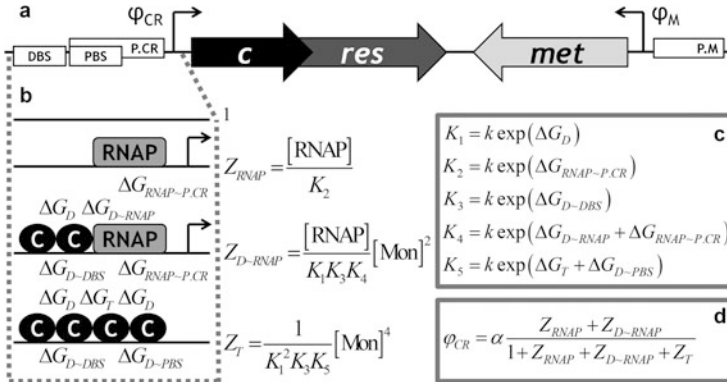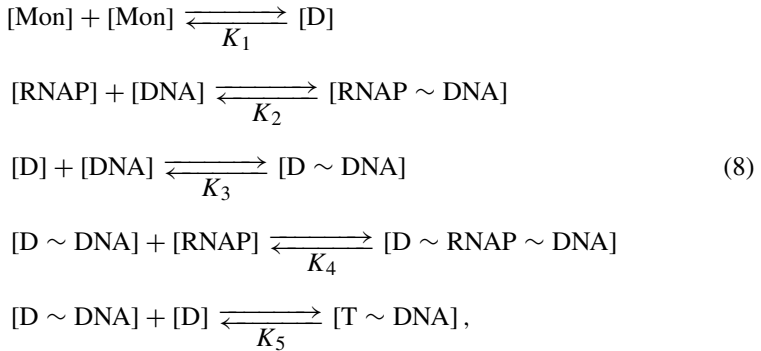
Fig. 1 Thermodynamic modeling of P.CR transcription regulation in AhdI R-M system. (**a**) Gene organization in AhdI system. P.CR, DBS, PBS, and P.M mark relative positions of the P.CR promoter, the distal and the proximal C protein binding site, and the P.M promoter, respectively. (**b**) Allowed P.CR configurations with their statistical weights denoted on the right, expressed in terms of the equilibrium dissociation constants (K) of reactions (8). Protein–DNA (below each configuration) and protein–protein (above interacting proteins) binding free energies (ΔG, in units of k_BT) are related to the appropriate equilibrium dissociation constants following the equations in (**c**). (**d**) P.CR transcription activity (φ_{CR}) is proportional to the fraction of statistical weights corresponding to transcriptionally active configurations (those containing an arrow in **b**)

the promoters denoted as P.CR and P.M, respectively (Fig. 1a). Methyltransferase methylates the P.M promoter, thereby repressing transcription of its own gene. On the other hand, transcription of the operon genes is regulated by binding of C protein dimers to the distal (DBS) and the proximal binding sites (PBS) in the P.CR promoter region.

Prior experiments of in vitro transcription from a wild type P.CR showed that transcription from this promoter is virtually inactive in the absence of C protein, and that it becomes first activated and then repressed with increasing C protein concentration (Bogdanova et al. 2008). This suggests that RNAP is presumably recruited to the promoter through a protein-protein contact with a bound C protein which, therefore, acts as a transcription activator. However, in the electrophoretic mobility shift assay experiments, only free DNA and complexes comprised of C protein tetramers bound to DNA were revealed in the whole range of varying C protein concentrations (Bogdanova et al. 2008; McGeehan et al. 2006). Furthermore, it was shown that DBS has a few orders of magnitude larger binding affinity than PBS, indicating that binding of C dimers to DNA is highly cooperative, i.e., a C dimer bound to DBS immediately recruits a second C dimer to PBS. As a bound C tetramer prevents RNAP from binding to the P.CR and thereby represses transcription of *c* and *res* genes, this raises a question of how transcription from the P.CR is activated. Therefore, quantitative modeling was used to test the proposed mechanism: that RNAP can passively outcompete a second C dimer from binding to PBS, which results in activation of transcription from the P.CR (Bogdanova et al. 2008).

The proposed thermodynamic model of the P.CR transcription regulation takes into account the following chemical reactions, characterized by the appropriate equilibrium dissociation constants (K):

$$[\text{Mon}] + [\text{Mon}] \underset{K_1}{\rightleftharpoons} [\text{D}]$$

$$[\text{RNAP}] + [\text{DNA}] \underset{K_2}{\rightleftharpoons} [\text{RNAP} \sim \text{DNA}]$$

$$[\text{D}] + [\text{DNA}] \underset{K_3}{\rightleftharpoons} [\text{D} \sim \text{DNA}] \tag{8}$$

$$[\text{D} \sim \text{DNA}] + [\text{RNAP}] \underset{K_4}{\rightleftharpoons} [\text{D} \sim \text{RNAP} \sim \text{DNA}]$$

$$[\text{D} \sim \text{DNA}] + [\text{D}] \underset{K_5}{\rightleftharpoons} [\text{T} \sim \text{DNA}],$$

where [RNAP], [Mon], [D] and [DNA] stand for concentrations of RNA polymerase, C protein monomers and dimers, and DNA containing the P.CR promoter region, while [RNAP $\sim$ DNA], [D $\sim$ DNA], [D $\sim$ RNAP $\sim$ DNA] and [T $\sim$ DNA] denote concentrations of established complexes of, respectively, RNAP bound to the P.CR, a C dimer bound to DBS, RNAP recruited to the promoter by a bound C dimer, and a bound C tetramer. This system of reactions describes establishing of the allowed P.CR equilibrium configurations characterized by the following statistical weights (Fig. 1b):

- 1—empty promoter;
- $Z_{RNAP} = [\text{RNAP} \sim \text{DNA}]/[\text{DNA}]$—only RNAP bound to the promoter, which corresponds to basal transcription of the operon genes;
- $Z_{D\text{-}RNAP} = [\text{D} \sim \text{RNAP} \sim \text{DNA}]/[\text{DNA}]$—RNAP recruited to the promoter by a C dimer bound to DBS, resulting in transcription activation;
- $Z_T = [\text{T} \sim \text{DNA}]/[\text{DNA}]$—a second C dimer recruited to PBS by a C dimer bound to DBS, with obtained C tetramer repressing transcription.

Note that the configuration representing only a C dimer bound to PBS was not taken into account, as such a configuration was not observed in the experiments and has a very low probability due to a large cooperativity in C dimers binding. One should also note that this modeling approach involves the rapid equilibrium assumption applied to the binding reactions, which is justified by the fact that association and dissociation processes between a protein and a DNA molecule, or two protein molecules, are much faster compared to transcription, translation and protein/RNA degradation processes (Phillips et al. 2012). Consequently, the model considers only the frequency of different promoter configurations in equilibrium and cannot distinguish between different sequences of binding events leading to a given configuration—e.g., whether protein A binds to DNA first and prevents binding of protein B, or it displaces protein B when it is already bound to DNA.

The measured value of C protein dimerization constant (K_1) is by an order of magnitude larger than the range of C protein concentrations used in experiments, indicating that C protein is present in a cell in the form of monomers. Therefore, statistical weights of the corresponding configurations are expressed in terms of C monomer and RNAP concentrations and, either appropriate equilibrium dissociation constants (Fig. 1b), or binding free energies (Fig. 1c). According to the assumption introduced above, transcription activity of the P.CR is proportional to the fraction of statistical weights that correspond to bound RNAP (Fig. 1d). Absorbing all constants into few parameters (x, y, and z), P.CR transcription activity is obtained as a function of C protein monomer concentration:

$$\varphi_{CR}(\text{Mon}) = \alpha \frac{x + y[\text{Mon}]^2}{1 + x + y[\text{Mon}]^2 + z[\text{Mon}]^4}, \tag{9}$$

where α is a proportionality constant with units transcript amount over time. Equation (9) was fitted to the experimentally measured data, obtained for a wild type system (Fig. 2a), but also for systems in which mutations were introduced in the DNA sequences of DBS or/and PBS (Fig. 2b–d), which corresponds to changing $\Delta G_{D\text{-}DBS}$ or/and $\Delta G_{D\text{-}PBS}$ (see Fig. 1b) (Bogdanova et al. 2008). Fig. 2 shows that the proposed model, with only three free parameters (x, y, and z; α was given the value 1), is in very good agreement with the data for both the wild type and the mutated systems. Furthermore, when fitted to the mutants data, parameter values change as expected with respect to the wild-type case—e.g., decreasing the affinity of DBS strongly negatively affects parameters y and z, while it has no effect on parameter x (compare the Eq. (9) with statistical weights in Fig. 1b and c). All of the above indicates that the modeling can realistically explain in vitro measured transcription activities and, accordingly, that the proposed model appropriately describes the P.CR transcription regulation in AhdI system.
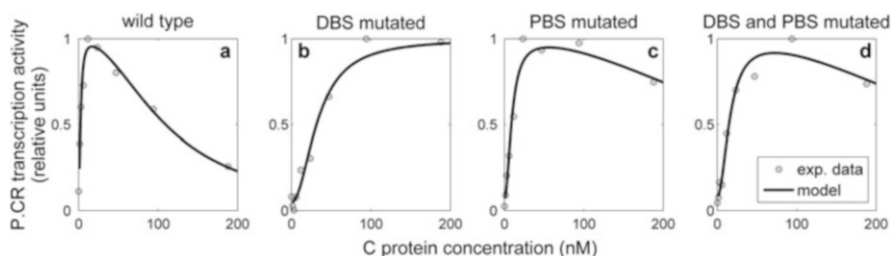

Fig. 2 Fitting experimentally measured dependence of P.CR transcription activity on C protein concentration in wild type and experimentally mutated systems, with a thermodynamic model of this promoter transcription regulation. Transcription activity was measured in arbitrary units and the values (grey circles) were normalized. Solid curves represent the fitted model Eq. (9). (**a**) Wild type system, (**b**) DBS affinity decreased, (**c**) PBS affinity decreased, (**d**) Decreased affinity of both DBS and PBS (Bogdanova et al. 2008)

3 Dynamic Modeling of Protein Expression

Dynamic modeling is the most common approach to model molecular networks and can be used to predict how protein amounts of interest—e.g. those of restriction enzyme and methyltransferase—change with time. State variables of the model represent concentrations (or numbers of molecules) of all mRNA and protein species in the system. These quantities dynamically depend on the combination of all processes that increase or decrease the corresponding amounts, characterized by appropriately defined rates (Le Novère 2015).

Experimentally observing dynamics of protein expression in a cell is, however, challenging due to a prerequirement for a synchronized cell population. Consequently, such measurements have been conducted on R-M systems in only two cases: for PvuII system, by introducing the system in a cell on a phage vector (Mruk and Blumenthal 2008), and for Esp1396I system, by monitoring fluorescently labeled R-M system proteins at the level of single cells (Morozova et al. 2016). In the latter case, experimental measurements were compared with predictions of a biophysical model of Esp1396I R-M system expression during its establishment in a newly transformed host (Morozova et al. 2016).

Similarly to AhdI system, Esp1396I system contains *c* and *res* genes in an operon, expressed from a promoter controlled by cooperative binding of two C dimers (see Fig. 1a and b). In contrast to an autoregulated *m* gene in AhdI system, in Esp1396I system, P.M is under control of C protein, where binding of one C dimer to its single binding site in this promoter region represses transcription of *m* gene (Bogdanova et al. 2009). P.CR and P.M regulation was thermodynamically modeled as explained above, to obtain relations for their transcription activities (φ_i) as functions of C protein concentration, which were further used as an input for a dynamic model describing how appropriate transcript (m_i) and protein (p_i) amounts change with time, for all three system components ($i = C, Res, Met$ denoting C protein, restriction enzyme, and methyltransferase, respectively):

$$\frac{dm_i(t)}{dt} = \varphi_i - \lambda_i^m \cdot m_i, \qquad \frac{dp_i(t)}{dt} = \kappa_i \cdot m_i - \lambda_i^p \cdot p_i \qquad (10)$$

Equation (10) takes into account that transcript and protein amounts are increased by transcription of the corresponding genes and translation of their transcripts (with translation constants κ_i), respectively, while these amounts are decreased with decay constants λ_i^m and λ_i^p, which account for both degradation and dilution of molecules due to cell division.

The proposed model of Esp1396I expression is minimal, in a sense that it takes into account only the experimentally established regulatory mechanisms, and that all model parameters are considered time-independent. Estimating the parameters by fitting this model to the data (Fig. 3a and b), is a difficult task due to the relatively large parameter space. This task is simplified by the fact that the parameters related to restriction enzyme expression can be estimated separately from those
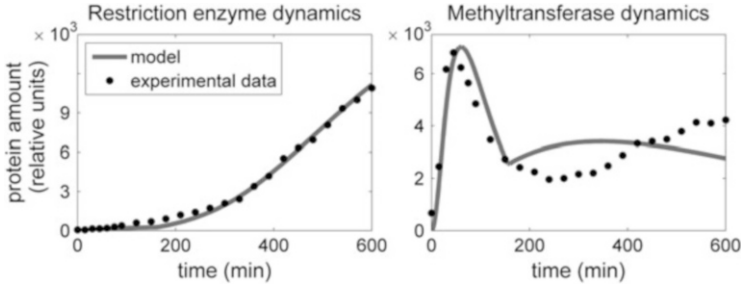
Fig. 3 Fitting experimentally measured data of single cell Esp1396I R-M system expression dynamics with a biophysical model. The zero time point corresponds to the plasmid entry in a naïve cell. (**a**) Restriction enzyme expression dynamics, (**b**) Methyltransferase expression dynamics (Morozova et al. 2016)

describing methytransferase expression, as methyltransferase does not control *c* and *res* expression. The observed good agreement of the model with the data is also supported by a subsequent experimental confirmation of very large restriction enzyme stability, which is consistent with inferred parameter values. Moreover, this minimal model can explain the main qualitative features of expression dynamics observed for Esp1396I system and proposed for R-M systems in general (Fig. 3a and b): a delayed beginning of restriction enzyme synthesis and high expression of methyltransferase early upon transforming a naïve cell. Improved quantitative agreement of the model with the data can likely be achieved by involving the dependence of at least some parameter values with time, imposed by changing conditions in a cell population or a desynchronization of cell and plasmid division. Specifically, during the first ∼160 min cells in the culture divided with different (faster) rate compared to the rest of the experiment (Morozova et al. 2016), which is taken into account through decay parameters in the model, as previously explained. Therefore, it is plausible to assume that population dynamics also has significant effect on some other parameters of the model, which may be a subject of future modeling.

4 Modeling Expression of EcoRV R-M System

In contrast to AhdI and Esp1396I systems presented above, in EcoRV R-M system P.CR and P.M are oriented divergently and partially overlap causing mutually exclusive binding of RNAP to these promoters (Fig. 4a), which represents the most distinctive regulatory feature of EcoRV system (Semenova et al. 2005). Consequently, P.CR and P.M control is strongly coupled, making transcription regulation of this system more complex compared to AhdI system. Furthermore,
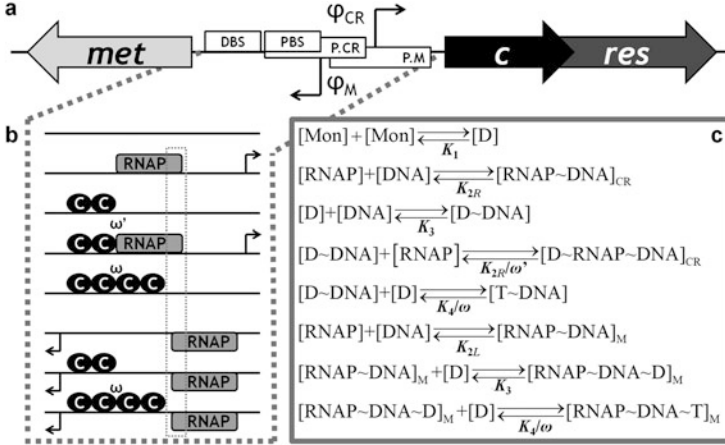
Fig. 4 Thermodynamic modeling of EcoRV R-M system transcription regulation. (**a**) Scheme of gene organization in EcoRV. Relative positions of operon and *met* promoters (P.CR and P.M) and distal and proximal C dimer binding sites (DBS and PBS) are denoted. (**b**) Allowed configurations of a DNA fragment separating *met* and *c* genes, with those transcriptionally active denoted with an arrow. Overlapping of P.CR and P.M is emphasized by framing their common fragment. (**c**) Chemical reactions in the model, with their equilibrium dissociation constants (K). Unlike in modeling AhdI transcription regulation (Fig. 1), cooperativities in binding of a second C dimer to PBS ($\omega \equiv \exp(-\Delta G_T)$) and of RNAP to P.CR ($\omega' \equiv \exp\left(-\Delta G^{CR}_{D \sim RNAP}\right)$) are here introduced as separate parameters, to enable perturbation of ω alone (see below in the text)

all characteristic AhdI regulatory features are absent from EcoRV. Namely: (1) no cooperativity in C dimers binding to DBS and PBS was experimentally found for EcoRV system, (2) *c* transcript is not leaderless in EcoRV, contrary to AhdI system whose leaderless *c* transcript is translated less efficiently than *res* and *met* transcripts, and (3) the equilibrium dissociation constant for a reaction of C protein dimerization is significantly lower than in AhdI system, leading to mostly C dimers in solution (Semenova et al. 2005).

To thermodynamically model EcoRV transcription regulation, one first needs to determine the allowed configurations of a DNA region separating the two divergent genes (Fig. 4b). Transcription regulation of the P.CR by C protein is similar to that found in AhdI system, except that now an additional configuration, consisting of only one C dimer bound to DBS, has to be included due to the absence of cooperativity. Regarding the P.M regulation, contrary to AhdI where it was C-independent, in EcoRV it is indirectly influenced by C protein, as it dictates when RNAP can bind to P.M due to overlapping promoters. From the equilibrium chemical reactions (Fig. 4c), which describe establishing of the allowed configurations, statistical weights can be determined and further used to obtain the

equations for P.CR and P.M transcription activities:

$$\varphi_{CR}(\text{Mon}) = \alpha \frac{\left(1 + \omega' \frac{[\text{Mon}]^2}{K_1 K_3}\right)}{u\left(1 + \frac{[\text{Mon}]^2}{K_1 K_3} + \omega \frac{[\text{Mon}]^4}{5 K_1^2 K_3^2}\right) + \left(1 + \omega' \frac{[\text{Mon}]^2}{K_1 K_3}\right)}, \qquad (11)$$

$$\varphi_{M}(\text{Mon}) = \alpha \frac{u\left(1 + \frac{[\text{Mon}]^2}{K_1 K_3} + \omega \frac{[\text{Mon}]^4}{5 K_1^2 K_3^2}\right)}{u\left(1 + \frac{[\text{Mon}]^2}{K_1 K_3} + \omega \frac{[\text{Mon}]^4}{5 K_1^2 K_3^2}\right) + \left(1 + \omega' \frac{[\text{Mon}]^2}{K_1 K_3}\right)}, \qquad (12)$$

relying, again, on the assumption that promoter transcription activity is proportional to its equilibrium occupancy by RNAP. In deriving the above Eqs. (11) and (12), the following information from the experiments was used: a C dimer binds to DBS with approximately five times higher affinity compared to PBS, setting $K_4/K_3 = 5$, and the P.CR is considerably weaker than the P.M ($K_{2R} >> K_{2L}$, $u = K_{2R}/K_{2L}$) (Semenova et al. 2005). The thermodynamic model of EcoRV transcription regulation (Eqs. (11) and (12)) is incorporated in an appropriate dynamic model of transcript and protein expression, of the form given by Eq. (10). Furthermore, to estimate the model parameters, and since EcoRV expression dynamics has not been experimentally measured, it is useful to reduce their number by rescaling the appropriate variables. A detailed explanation of parameter estimation in the case of EcoRV is available in (Rodic et al. 2017b). Overall, this presents to our knowledge the first model of a divergent R-M system, which provides an opportunity to assess the effect of regulatory features found in such a system on its expression dynamics, by in silico introducing AhdI features in EcoRV system (see below).

5 Inferring Effects of R-M Systems Regulatory Features on Their Dynamical Properties

As all R-M systems share the same function, namely, efficiently destroying foreign DNA without harming the host cell, it is reasonable to hypothesize that their expression dynamics, constrained by their function, should exhibit some universal properties, regardless of the underlying regulation. Specifically, the following common dynamical properties of R-M system establishment in a naïve host cell have been proposed (Rodic et al. 2017b): (1) a time delay in expression of restriction enzyme with respect to methyltransferase, which provides time for genome protection, (2) a fast transition of restriction enzyme expression from the OFF to the ON state, to ensure rapid cell protection from incoming foreign DNA, and (3) a stable steady-state of the toxic molecule (restriction enzyme), as
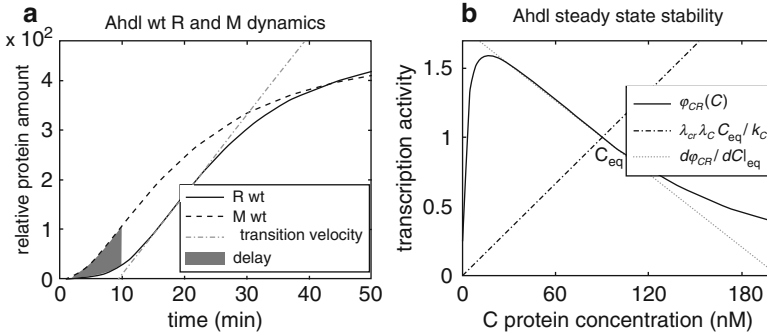
Fig. 5 Dynamical property observables. (**a**) Predicted restriction enzyme (R) and methyltransferase (M) expression dynamics upon system entry in a naïve bacterial host. Measures of R expression time delay and transition velocity are graphically represented, (**b**) Dependence of AhdI P.CR transcription activity on C protein concentration is provided by the full line, whose intersection with the dash-dotted line determines the equilibrium C protein concentration. Slope of the transcription activity curve at this equilibrium concentration (the dotted line) is related with the steady state stability (Rodic et al. 2017b)

fluctuations in restriction enzyme amount not matched by appropriate fluctuations in methyltransferase amount could lead to host cell death.

To quantify these properties, corresponding *dynamical property observables* were defined, which are graphically represented on the example of predicted AhdI wild type dynamics in Fig. 5 (Bogdanova et al. 2008; Rodic et al. 2017b). As a measure of the time delay, the ratio of the shaded areas in a perturbed and in a wild type system, spanning the first 10 min postinduction was introduced (Fig. 5a). A maximal slope of the sigmoidal restriction enzyme expression curve (the dash-dotted line in Fig. 5a) measures the transition velocity from the OFF (low restriction enzyme amount) to the ON (high restriction enzyme amount) state. Finally, the third dynamical property observable (Ω^2) related with the slopes of the dash-dotted and dotted curves shown in Fig. 5b quantifies stability of the restriction enzyme steady-state (Rodic et al. 2017b).

From Fig. 5 it can be readily inferred that AhdI exhibits all the listed dynamical properties. Moreover, perturbing characteristic AhdI regulatory features—i.e., large cooperativity in C dimers binding, high dissociation constant for C dimerization and low translation rate for the leaderless *c* transcript—abolishes these properties, leading to, presumably, less optimal AhdI expression dynamics (Rodic et al. 2017b). Thus, the requirement for the proposed dynamical properties might explain the existence of these characteristic AhdI regulatory features.

Despite missing all regulatory features inherent to AhdI system, and having a unique feature not present in AhdI (overlapping of P.CR and P.M), wild type EcoRV system also meets the same three dynamical properties (see the darkest R and M curves in Fig. 6a), arguing in favor of universality of these properties in different R-M systems. Therefore, the question emerges: why are AhdI regulatory features,
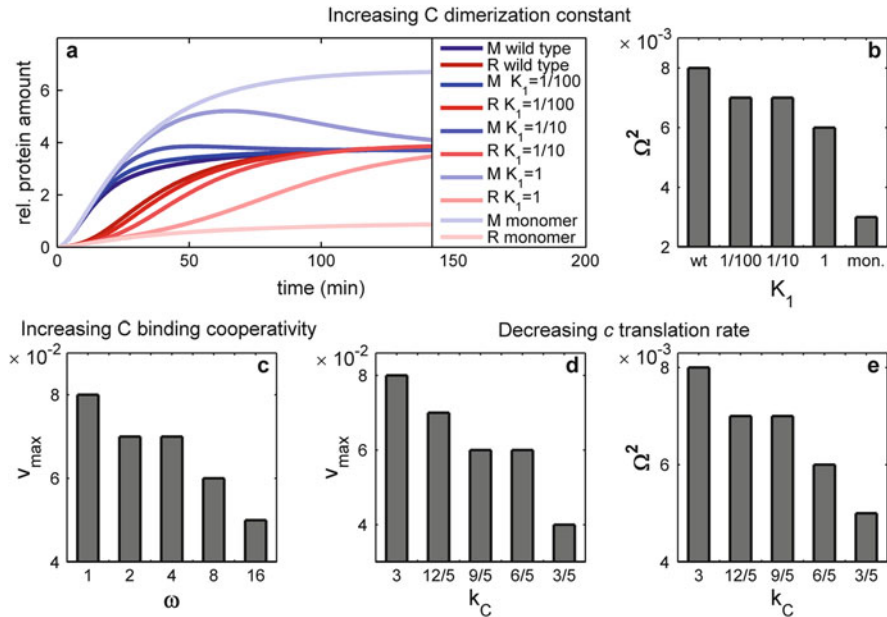
Fig. 6 In silico introducing AhdI-characteristic regulatory features in EcoRV system. Effect of increasing C dimerization constant K_1 on (**a**) dynamics of restriction enzyme (R) and methyltransferase (M) and (**b**) the steady-state stability. Effect of increasing cooperativity ω in C binding on (**c**) R OFF-ON transition velocity. Effect of decreasing the c translation constant k_C on (**d**) transition velocity and (**e**) steady-state stability (Rodic et al. 2017b)

apparently successful in optimizing this R-M system immune function, absent from EcoRV? This question can be addressed by in silico introducing characteristic AhdI regulatory features in EcoRV and observing their effect on the system dynamics (Rodic et al. 2017b).

To that end, the equilibrium dissociation constant of C dimerization (K_1 in Fig. 4c) was gradually increased toward an AhdI-characteristic value, which corresponds to a transition from the case where the solution contains mostly C dimers to the case where it contains mostly C monomers (Fig. 6a and b). This perturbation clearly has an adverse effect on two dynamical properties: on the OFF-ON transition velocity (note in Fig. 6a that the slope of R curves decreases as the dimerization constant is increased) and on the steady-state stability (Fig. 6b). Transition velocity of restriction enzyme expression is also decreased by introducing higher C binding cooperativity (increasing ω), as can be seen from Fig. 6c, and by decreasing c translation constant k_C (Fig. 6d). Less efficient c translation additionally leads to a less stable steady-state (Fig. 6e). Apparently, none of these three perturbations affect an extent of the time delay in restriction enzyme expression.

Overall, the observation that perturbing of the characteristic regulatory features abolishes one or more of the proposed dynamical properties for both AhdI and

EcoRV R-M systems (that are characterized by two different architectures, convergent and divergent), provides a possible unifying principle behind seemingly different designs of these systems. Particularly, specific combinations of different regulatory features found in these two systems appear to be optimized to meet the same dynamical properties, related with their successful establishment in a new host cell. Moreover, it seems that some regulatory features are specifically found together in the same system because of their complementary effects on system expression dynamics. Namely, high binding cooperativity in AhdI system is accompanied by a large C dimer dissociation constant, which may be a consequence of the opposite effects these features have on the steady-state level of restriction enzyme, thereby balancing the amount of this toxic molecule, while at the same time providing more optimal dynamical properties of system expression (Rodic et al. 2017b). In line with this proposal is the absence of both features in EcoRV system, where their separate introduction leads to abolishing of some of the dynamical properties and disrupting the steady-state ratio of the amounts of methyltransferase and restriction enzyme (see for example Fig. 6a). Furthermore, Esp1396I system with convergent gene organization and C regulated transcription similar as in AhdI system, exhibits both lower cooperativity in C binding and lower dissociation constant of dimerization compared to AhdI (Bogdanova et al. 2009). It would be interesting to see if other R-M systems, with different regulatory features, such as control by antisense RNAs or by translational coupling (Nagornykh et al. 2008), are similarly constrained by the proposed dynamical principles, and what are the roles of their regulatory features in achieving these principles.

6 Assessing the Significance of CRISPR-Cas Regulatory Features

Protection of a host bacterium by a CRISPR-Cas system requires its activation upon infection by foreign DNA, or upon setting the right environmental conditions (Ratner et al. 2015). However, expression dynamics of a native CRISPR-Cas have not been observed experimentally because this system (Type I-E) is silent in cells under standard conditions, even in the presence of an infecting phage, and signaling resulting in system activation is unknown (Pul et al. 2010). To date, experimental research on CRISPR-Cas transcription control in *E. coli* and *S. enterica* revealed some elements and regulatory features involved in system activation (Pul et al. 2010; Westra et al. 2010; Medina-Aparicio et al. 2011), specifically: (1) it is known that both CRISPR array and (Cascade) *cas* genes promoters are repressed by highly cooperative binding of global regulators, such as H-NS and LRP, (2) repressors can be outcompeted in binding by some global activators (e.g., LeuO), when present at elevated amounts. Therefore, highly cooperative repression, which is abolished by transcription activators, can be considered as one of the major Type I-E CRISPR-Cas regulatory features (Rodic et al. 2017a).
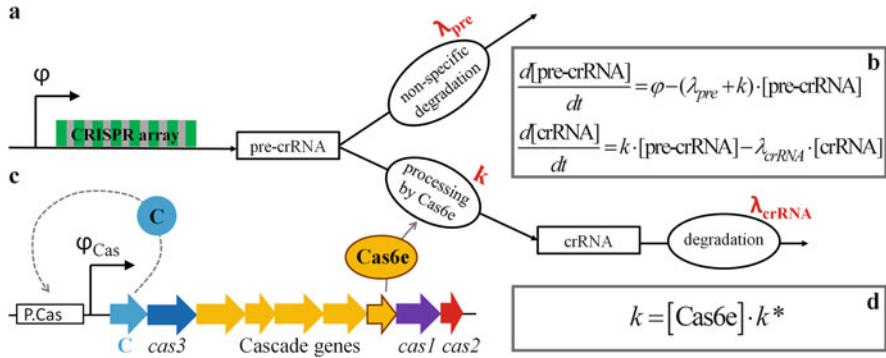
Fig. 7 Scheme of the proposed setup for CRISPR-Cas activation. (**a**) Pre-crRNA processing model scheme. Notation: φ—CRISPR array transcription rate (assumed constant in modeling), λ_{pre}, k and λ_{crRNA}—rates of the processes specified in the ovals (Djordjevic et al. 2012); (**b**) Equations which correspond to the scheme under **a** and which describe time dependence of pre-crRNA and crRNA amounts; (**c**) A schematic diagram of plasmid encoded *c* and *cas* genes under control of the Cas promoter (P.Cas), with a transcription rate φ_{Cas}. AhdI-like regulation of P.Cas by C protein, as denoted by a dashed arrow, is included in the cooperative model; (**d**) Dependence of a processing rate k on Cas6e amount is considered linear, in line with an experimentally determined very low amount of its substrate (pre-crRNA; k^*—processing constant) (Rodic et al. 2017a)

Furthermore, another key regulatory feature of CRISPR-Cas expression is the fast nonspecific pre-crRNA degradation by an unidentified endonuclease (Djordjevic et al. 2012). Particularly, it was shown by modeling CRISPR transcript processing upon artificial overexpression of Cas proteins (for a scheme of a model and corresponding dynamic equations see Fig. 7a and b), that the main mechanism responsible for a large increase in crRNA amount from a small decrease in substrate (pre-crRNA) amount is the fast substrate degradation. Processing of pre-crRNA by an elevated amount of Cas proteins diverts the whole molecule from the path of nonspecific degradation, thereby amplifying the equilibrium crRNA amount for a few orders of magnitude (Djordjevic et al. 2012).

However, to more realistically predict CRISPR-Cas expression dynamics and to understand the significance of the established main regulatory features of this system, an appropriate mathematical description of gradual expression of a processing Cas6e protein upon system induction is also needed, as the pre-crRNA processing rate (k) directly depends on the level of Cas6e (Fig. 7d). As a detailed mechanism of Cas promoter (P.Cas) control is unknown, one possible approach involves replacing transcription control (while keeping the CRISPR-Cas pre-crRNA processing regulation) of a native P.Cas with that of a qualitatively and mechanistically similar, but better explored system, and in silico analyzing expression dynamics of the obtained construct (Rodic et al. 2017a). Such an approach would allow assessing the significance of CRISPR-Cas regulatory features, with a minimum of guessing (if a system used as a proxy was already well-studied so that all its major parameters are fixed). At the same time, in silico analysis would provide

a much simpler and faster way of fulfilling the task of interest, in comparison to a complementary experimental approach, which would require creating an artificial gene circuit, extensive mutations of the system regulatory features and measuring in vivo expression dynamics of pre-crRNA and crRNA.

Similarities in transcription regulation can be noted between a well-studied AhdI R-M system, for which a biophysical model was already constructed and evaluated (see Fig. 1) (Bogdanova et al. 2008), and Type I-E CRISPR-Cas. Strong cooperative interactions are involved in both systems, in particular, in binding of C dimers to the P.CR region and in binding of H-NS molecules to the P.Cas region, thereby repressing transcription (Pul et al. 2010). This repression by H-NS can be abolished by transcriptional activator LeuO (Westra et al. 2010; Medina-Aparicio et al. 2011). Consequently, autoregulation (both positive and negative), similar to that exhibited by C protein in AhdI system, is likely found in CRISPR-Cas regulation, as LeuO activates, and also indirectly represses its own promoter (Chen et al. 2001; Stratmann et al. 2012). Thus, the main features of CRISPR-Cas transcription regulation, namely, gradual synthesis of Cas proteins, cooperativity in transcription regulation, and putative autoregulation, can be qualitatively mimicked by putting *cas* genes under transcription control found in AhdI. More precisely, the proposed setup for CRISPR-Cas activation analyzed in silico includes *cas* genes put under control of the P.CR promoter from AhdI (see Fig. 1b), which are introduced in a cell on a plasmid, marking the start of CRISPR-Cas activation (Fig. 7c); dynamics of crRNA generation due to the Cas6e processing activity is monitored.

To understand the effect of cooperative transcription regulation of *cas* genes, three (sub)models of *cas* genes regulation in the proposed setup are analyzed: (1) a baseline model, in which P.Cas transcription activity acquires its equilibrium value infinitely fast, (2) a constitutive model, with constant P.Cas transcription activity, and (3) a cooperative model, where P.Cas is cooperatively regulated by C protein in the same manner as the AhdI P.CR promoter. Figure 8 shows how the amount of crRNA, determined 20 min after the start of system activation, depends on the level of the processing rate k reached in equilibrium, in all three models. The crRNA
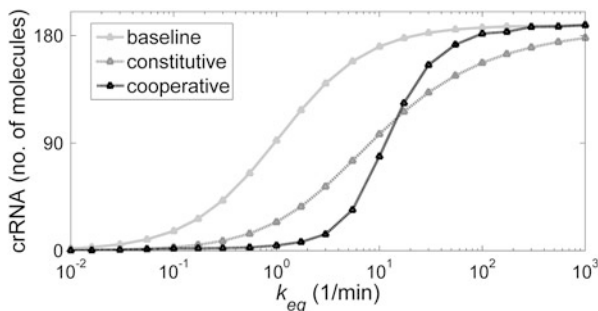

Fig. 8 Dependence of crRNA amount 20 min postinduction on the equilibrium value of a pre-crRNA processing rate k_{eq} in the three models of CRISPR-Cas activation (Rodic et al. 2017a)

amount at 20 min postinduction was specifically chosen, as this period is most relevant for a successful cell defense against an incoming virus.

The most prominent characteristic of the cooperative model is its switch-like behavior (Fig. 8), in contrast to much more gradual responses of the constitutive and the baseline models. This feature enables keeping the system in the OFF state in the case of possible leaks in P.Cas activity (corresponding to low k_{eq} values) and, on the other side, rapidly generating a sufficient amount of crRNAs once the induction signal is received, to efficiently combat viral infection. Furthermore, taking into account that the amount of crRNAs sufficient to negatively affect phage development was determined to be $\sim$10 molecules (Pougach et al. 2010), the models predict that enough crRNAs is generated even at low (somewhat larger than 1 1/min) k_{eq} values, corresponding to the activated system regime. Therefore, CRISPR-Cas system expression regulation is probably mainly constrained by a requirement of rapidly producing a large amount of crRNAs (Rodic et al. 2017a).

In line with its switch-like behavior, at low k_{eq} values the cooperative model produces less crRNAs than the constitutive one. However, at high k_{eq} values an interesting cross-over behavior is observed, where the cooperative model approaches the limit of infinitely fast crRNA production (given by the baseline model). Even more crRNAs can be generated by jointly activating transcription of both *cas* genes and a CRISPR array, which relieves the saturation obtained when increasing only *cas* expression (Djordjevic et al. 2012; Rodic et al. 2017a). As k values around 100 1/min were encountered in experiments with artificial overexpression of *cas* genes (Pougach et al. 2010; Djordjevic et al. 2012), it is tempting to speculate that such high rates of pre-crRNA processing may also be reached in the native system, providing highly efficient protection to a bacterium.

Effect of abolishing the second major CRISPR-Cas regulatory feature, i.e., of decreasing the pre-crRNA nonspecific degradation rate (λ_{pre}) in the cooperative model, can be deduced from Fig. 9. Namely, the crRNA dynamics curve is gradually deformed with respect to the standard Hill (sigmoidal) shape, indicating that fast
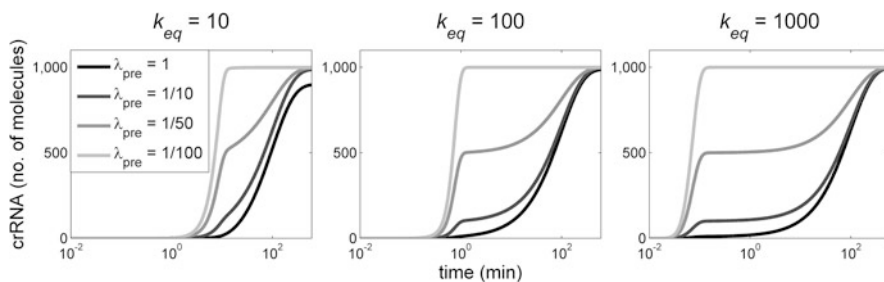

Fig. 9 Perturbing the fast nonspecific degradation of pre-crRNA. Effect of decreasing the pre-crRNA degradation rate λ_{pre} on crRNA expression dynamics at three different k_{eq} values can be seen in the figures under (**a**, **b**, and **c**). Native CRISPR-Cas is characterized by $\lambda_{pre} = 1$ 1/min. Zero time point corresponds to the start of system activation, i.e., to the moment of entrance of a plasmid carrying *cas* genes in a cell (Rodic et al. 2017a)

nonspecific pre-crRNA degradation is, together with cooperative transcription regulation, responsible for the system switch-like behavior. Another effect of decreasing λ_{pre}, which comes as a model prediction, is a decrease in the time delay of the onset of crRNA generation, most pronounced at high k_{eq} values. It has been shown previously that slow or delayed CRISPR interference (targeting of foreign DNA by Cascade-crRNA complexes) facilitates the *primed adaptation* (Künne et al. 2016; Musharova et al. 2017), i.e., the acquisition of invasive DNA fragments similar to already possessed spacers, to be incorporated as new spacers in the CRISPR array. This mechanism serves to minimize infection by phages with mutated genome sequences, which would otherwise evade the interference (Sternberg et al. 2016). The required delay could be achieved by a delay in crRNA generation, resulting from fast pre-crRNA degradation, as predicted by the model (Rodic et al. 2017a). Consequently, both main dynamical features, i.e., rapid transition of the system from OFF to ON state and the delay in the expression of the effector molecules (restriction enzyme and crRNAs), are observed in both R-M systems and mechanistically more complicated CRISPR-Cas systems.

7 Summary and Conclusion

Seemingly very different architectures and regulatory properties of AhdI and EcoRV systems can be explained by few common design principles, ensuring that expression dynamics of every R-M system is optimized to serve its purpose—namely, safe and efficient host cell protection from foreign DNA. Other R-M systems, representative of different regulatory mechanisms, should be investigated to test if unifying design principles can be defined at the level of the whole group of these prokaryotic immune systems. Moreover, having the same immune function, CRISPR-Cas systems may also obey similar design principles, as it was theoretically predicted by using a synthetic system to bypass the unknown CRISPR-Cas transcription control, while keeping the same transcript processing mechanism as in native CRISPR-Cas. Thereby, thermodynamic modeling proved to be an appropriate approach in describing R-M system transcription regulation, while in combination with a dynamic model of protein expression, it can be used to describe the main qualitative properties of R-M system dynamics of establishment in a single, naïve host cell. Further experimental and theoretical studies of CRISPR-Cas regulation may allow to more accurately understand its dynamics, in line with what is already done for R-M systems. Overall, the studies on bacterial immune systems summarized here underline a major goal of systems biology which is to discover common design principles in mechanistically otherwise different biological systems.

References

Bogdanova E, Djordjevic M, Papapanagiotou I et al (2008) Transcription regulation of the type II restriction-modification system AhdI. Nucleic Acids Res 36:1429–1442

Bogdanova E, Zakharova M, Streeter S et al (2009) Transcription regulation of restriction-modification system Esp1396I. Nucleic Acids Res 37:3354–3366

Chen C-C, Fang M, Majumder A et al (2001) A 72-base pair AT-rich DNA sequence element functions as a bacterial gene silencer. J Biol Chem 276:9478–9485

Djordjevic M, Djordjevic M, Severinov K (2012) CRISPR transcript processing: a mechanism for generating a large number of small interfering RNAs. Biol Direct 7:24–34

Dresch JM, Richards M, Ay A (2013) A primer on thermodynamic-based models for deciphering transcriptional regulatory logic. BBA-Gene Regul Mech 1829:946–953

Ershova A, Rusinov I, Spirin S et al (2015) Role of restriction-modification systems in prokaryotic evolution and ecology. Biochemistry-Moscow 80:1373–1386

Goldberg GW, Marraffini LA (2015) Resistance and tolerance to foreign elements by prokaryotic immune systems – curating the genome. Nat Rev Immunol 15:717–724

Hille F, Charpentier E (2016) CRISPR-Cas: biology, mechanisms and relevance. Philos T Roy Soc B 371:20150496

Künne T, Kieper SN, Bannenberg JW et al (2016) Cas3-derived target DNA degradation fragments fuel primed CRISPR adaptation. Mol Cell 63:852–864

Le Novère N (2015) Quantitative and logic modelling of molecular and gene networks. Nat Rev Genet 16:146–158

McGeehan J, Papapanagiotou I, Streeter S et al (2006) Cooperative binding of the C.AhdI controller protein to the C/R promoter and its role in endonuclease gene expression. J Mol Biol 358:523–531

Medina-Aparicio L, Rebollar-Flores J, Gallego-Hernández A et al (2011) The CRISPR/Cas immune system is an operon regulated by LeuO, H-NS, and leucine-responsive regulatory protein in Salmonella enterica serovar Typhi. J Bacteriol 193:2396–2407

Morozova N, Sabantsev A, Bogdanova E et al (2016) Temporal dynamics of methyltransferase and restriction endonuclease accumulation in individual cells after introducing a restriction-modification system. Nucleic Acids Res 44:790–800

Mruk I, Blumenthal RM (2008) Real-time kinetics of restriction–modification gene expression after entry into a new host cell. Nucleic Acids Res 36:2581–2593

Musharova O, Klimuk E, Datsenko KA et al (2017) Spacer-length DNA intermediates are associated with Cas1 in cells undergoing primed CRISPR adaptation. Nucleic Acids Res 45:3297–3307

Nagornykh M, Bogdanova E, Protsenko A et al (2008) Regulation of gene expression in a type II restriction-modification system. Russ J Genet 44:523–532

Phillips R, Kondev J, Theriot J et al (2012) Physical biology of the cell. Garland Science, New York

Pougach K, Semenova E, Bogdanova E et al (2010) Transcription, processing and function of CRISPR cassettes in Escherichia coli. Mol Microbiol 77:1367–1379

Pul Ü, Wurm R, Arslan Z et al (2010) Identification and characterization of E. coli CRISPR-cas promoters and their silencing by H-NS. Mol Microbiol 75:1495–1512

Ratner HK, Sampson TR, Weiss DS (2015) I can see CRISPR now, even when phage are gone: a view on alternative CRISPR-Cas functions from the prokaryotic envelope. Curr Opin Infect Dis 28:267–274

Rodic A, Blagojevic B, Djordjevic M et al (2017a) Features of CRISPR-Cas regulation key to highly efficient and temporally-specific crRNA production. Front Microbiol 8:2139

Rodic A, Blagojevic B, Zdobnov E et al (2017b) Understanding key features of bacterial restriction-modification systems through quantitative modeling. BMC Syst Biol 11:377–391

Semenova E, Minakhin L, Bogdanova E et al (2005) Transcription regulation of the EcoRV restriction–modification system. Nucleic Acids Res 33:6942–6951

Shea MA, Ackers GK (1985) The OR control system of bacteriophage lambda: a physical-chemical model for gene regulation. J Mol Biol 181:211–230

Sneppen K, Zocchi G (2005) Physics in molecular biology. Cambridge University Press, Cambridge

Sternberg SH, Richter H, Charpentier E et al (2016) Adaptation in CRISPR-Cas systems. Mol Cell 61:797–808

Stowe K (2007) An introduction to thermodynamics and statistical mechanics. Cambridge University Press, New York

Stratmann T, Pul Ü, Wurm R et al (2012) RcsB-BglJ activates the Escherichia coli leuO gene, encoding an H-NS antagonist and pleiotropic regulator of virulence determinants. Mol Microbiol 83:1109–1123

Westra ER, Pul Ü, Heidrich N et al (2010) H-NS-mediated repression of CRISPR-based immunity in Escherichia coli K12 can be relieved by the transcription activator LeuO. Mol Microbiol 77:1380–1393

Modeling and bioinformatics of bacterial immune systems: understanding regulation of CRISPR/Cas and restriction-modification systems

Jelena Guzina[1], Anđela Rodić[1], Bojana Blagojević[2] and Marko Đorđević[1*]

[1]University of Belgrade – Faculty of Biology, Studentski trg 3, 11000 Belgrade, Serbia
[2]Institute of Physics Belgrade, University of Belgrade, Pregrevica 118, 11080 Belgrade, Serbia

Summary. Bacterial immune systems protect bacterial cells from foreign DNA, such as viruses and plasmids. They also critically affect bacterial pathogenicity by reducing the flow of genes between bacteria. Two such major systems are restriction-modification and the recently discovered CRISPR/Cas systems. Here we review our work on understanding gene expression regulation in these systems, which takes a systems biology approach, combining modeling, bioinformatics and data analysis from quantitative experiments. Specifically, we address the following: (i) modeling gene expression regulation during restriction-modification system establishment in a naïve bacterial host, (ii) modeling the dynamics of CRISPR/Cas activation, in particular, how the features characterizing system transcription regulation and transcript processing affect the dynamics, (iii) predictions of transcription start sites for alternative σ factors that have been poorly studied up-to-now, but are important as CRISPR/Cas likely responds to bacterial cell envelope stress, (iv) our preliminary results on predictions of different CRISPR/Cas components, in particular, small RNAs associated with the systems, which likely have a key role in their regulation.

Keywords: Bacterial defense systems, bioinformatics, biophysical modeling, CRISPR/Cas, restriction-modification systems.

Introduction

Bacteria are continuously exposed to foreign nucleic acids, such as phage DNA, plasmids or other mobile genetic elements. In order to protect genome integrity, cells are equipped with immune systems that target invasive extra-chromosomal elements for degradation (Shabbir et al. 2016), whereby the immune response reduces the rate of horizontal gene transfer (HGT), thus also affecting related aspects of cell functioning (e.g. virulence) (Vasu and Nagaraja 2013; Hatoum-Aslan and Marraffini 2014). Analogous to eukaryotic modes of defense, bacterial immune systems can be recognized as innate or adaptive, where restriction-modification and CRISPR/Cas (*C*lustered *R*egularly *I*nterspaced *S*hort *P*alindromic *R*epeats/*C*RISPR-*as*sociated proteins), respectively, are two major representatives of such systems (Goldberg and Marraffini 2015).

Restriction-modification (RM) systems are considered innate since they target invasive elements without prior immunization with fragments of foreign genetic material. Two major components of RM systems are the enzymes restriction endonuclease (R) and methyltransferase (M) (Fig. 1A), which are frequently encoded on mobile genomic loci (e.g. plasmids), so that these systems easily propagate through bacterial populations (Fig. 1B) (Heitman 1993; Kobayashi et al. 1999). Once an RM system enters the cell, tight regulation of its expression becomes essential for ensuring safe and efficient establishment in the naïve bacterial host. Precisely, R that represents the effector component of a RM system, cuts short specific DNA sequences, irrespective of their location, so that self-targeting can easily arise. To evade autoimmunity, R has to be expressed with a delay with respect to M, as methylation of the genomic sites recognized by R protects them from cleavage (Fig. 1A) (Wilson 1991).

Unlike RM systems, CRISPR/Cas provides adaptive component to bacterial immunity, which arises as a consequence of its dynamical structure (Barrangou et al. 2007; van der Oost et al. 2009). A major system component is the CRISPR array,
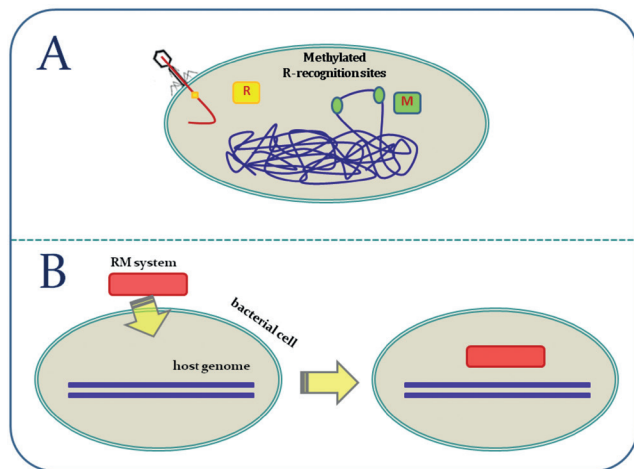
Fig. 1. RM system functioning (A) and establishment (B) in a host bacterial cell. A. Restriction-endonuclease (shown as a yellow rectangle) cuts the DNA at R-specific recognition sites (shown in yellow); Methyltransferase (shown as a green rectangle) methylates R-recognition sites on the host genome, thus protecting these sites (shown in green) from cleavage. **B.** RM systems are usually found on mobile genetic elements (e.g. plasmids), which enables them to efficiently propagate throughout bacterial populations. RM system, entering the bacterial cell (red rectangle), is shown.

which is characterized by a series of tandem repeats separated with unique spacer sequences (Fig. 2) (Al-Attar et al. 2011). The spacers are derived from previously encountered foreign genetic material, so that small interfering RNAs (crRNAs), which are generated upon array expression, target invasive elements based on complementarity; this makes the basic mechanism that confers resistance against foreign DNA/RNA (Bolotin et al. 2005). In addition to CRISPR array, the system also includes Cas proteins with mainly nucleolytic activity, which act as effectors during array immunization with new spacers, crRNA processing/expression and target degradation. CRISPR/Cas components typically remain silent under standard physiological conditions (Pul et al. 2010), whereby sudden activation leads to the production of large crRNA amounts, thus enabling efficient target eradication.

Despite the fact that RM and CRISPR/Cas systems markedly differ mechanistically, they likely embody the same design principles as a consequence of the general characteris-

Fig. 2. A typical organization of CRISPR/Cas locus in *E. coli*. CRISPR array is schematically presented with successive blue diamonds (direct repeats) and yellow rectangles (spacers); the upstream *cas* genes, characteristic of Type I CRISPR/Cas systems, are indicated with rightwards-oriented pentagons. Intergenic regions that contain promoters transcribing *cas* genes (IGLB), and CRISPR array (L) are also shown.

tics that shape the immune response. Namely, the induction of the CRISPR/Cas system probably faces similar dynamical constraints as the establishment of an RM system in a naïve bacterial host, as both require a rapid transition of the "toxic" (auto-immunogenic) molecule – R or crRNA – from "OFF" to "ON" state (Djordjevic 2013) to enable efficient target eradication. In addition to rapid transition, the expression of "toxic" immune molecules is also characterized by an initial delay, so that crRNAs in CRISPR/Cas are not expressed before the virus genome enters the cell, and M (the antidote) in RM systems has enough time to act.

Common design principles that impose similar dynamical constraints on RM and CRISPR/Cas activity are linked to the equivalent regulatory expression patterns in these systems. To understand the underlying transcription regulation, it is necessary to map transcription start sites (TSS) associated with different components of RM and CRISPR/Cas systems. This, however, is non-trivial since: (i) promoter elements of house-keeping σ factors are highly degenerate, so that a search usually results in a large fraction of false positives (Djordjevic 2014); (ii) information on the specificity of alternative σ factors (related to stress response) is largely missing, which is relevant since CRISPR/Cas is likely induced by cell-envelope stress (Ratner et al. 2015),which, in turn, is connected to Group IV (ECF) σ factors (Raivio and Silhavy 2001; Ratner et al. 2015).

In addition, an important aspect of CRISPR/Cas regulation are small RNAs associated with CRISPR/Cas (tracrRNAs) encoded outside the array, which have an essential role in CRISPR-transcript processing (Deltcheva et al. 2011), and possibly other system functions. Consequently, in this review we briefly present our work on:

1. modeling gene expression regulation during RM system establishment in a naïve bacterial host;
2. modeling dynamics of CRISPR/Cas activation, in particular how key features that characterize systems transcription regulation and transcript processing affect its dynamics;
3. predictions of bacterial TSS, particularly those related to alternative σ factors, which are poorly studied to date, but highly relevant as CRISPR/Cas likely responds to bacterial cell-envelope stress;
4. our preliminary results on predictions of different CRISPR/Cas components, in particular small RNAs associated with the system, which likely have a key role in its regulation.

Modeling *in vivo* expression of restriction-modification systems

Certain dynamical constraints imposed by their immune function have been proposed for RM systems in general. However, RM system dynamics have been observed in live cells in only two cases, as such experimental measure-

ments are complicated by a requirement for synchronous populations of cells transformed with RM system genes (Mruk and Blumenthal 2008; Morozova et al. 2016). In an earlier attempt, Mruk and Blumenthal synchronously introduced the PvuII system genes placed on an M13 phage into naïve cells by phage infection (Mruk and Blumenthal 2008). Our collaborators, on the other hand, conducted the first single-cell measurements of RM system dynamics for the Esp1396I system: they fused sequences encoding fluorescent proteins to the R and M genes and monitored the dynamics of the appearance of fluorescent signals in individual cells, transformed with a plasmid carrying the modified Esp1396I system (Morozova et al. 2016). To check if the regulatory features found in this particular system allow establishing observed dynamics, and if they can provide the proposed dynamical constraints, we constructed a quantitative model of the Esp1396I system regulation, which we will briefly outline below.

Among type II RM systems, whose main characteristic is that R and M are encoded by separate genes, a large group contains a third gene encoding a control (C) protein, which is typically transcribed as a part of the operon with the R gene; the example for such a gene arrangement is the RM system Esp1396I represented in Fig. 3A. C proteins regulate transcription by binding in the form of dimers to their binding sites, partially overlapping with a promoter (Nagornykh et al. 2008). The transcription of Esp1396I system genes was thermodynamically modeled by considering all allowed configurations of the system promoters and determining their statistical weights (Figs. 3B and 3C). The most frequently
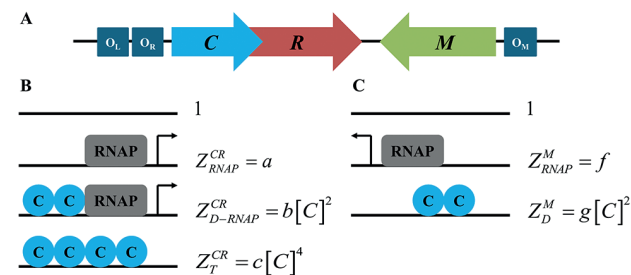

Fig. 3. Modeling transcription regulation in the Esp1396I RM system. A. Esp1396I gene organization scheme. The convergently oriented genes encoding R and M in the Esp1396I system are represented by the red and the green arrows, respectively, while the blue arrow represents the C gene, partially overlapping with the R gene. The dark blue boxes denoted by O_L, O_R and O_M represent operator sequences in the CR and the M promoter, which bind C dimers. **B and C.** The allowed configurations of RNAP (grey rectangle) and C protein (blue circle) molecules on the CR and the M promoter are illustrated, respectively, in B and C, where the transcriptionally active configurations contain an arrow. The corresponding statistical weights (Z) of the configurations, indicated on their right, depend on constant RNAP concentration and protein-protein and protein-DNA interaction energies (absorbed into parameters a, b, c, f and g) and variable C protein concentration.

observed regulation mechanism of the weak C and R operon (CR) promoter (also found in the Esp1396I system, Fig. 3B) involves highly cooperative binding of two C dimers to the left and the right operator sequences (O_L and O_R in Fig. 3A), where a C dimer bound to the high affinity left binding site can recruit either RNA polymerase (RNAP) to the promoter (thus activating transcription; the corresponding configuration has a statistical weight Z_{D-RNAP}^{CR}), or a second C dimer to the low affinity right binding site (establishing a tetramer that represses transcription; configuration Z_T^{CR}) (Bogdanova et al. 2008; Nagornykh et al. 2008). In the Esp1396I RM system, transcription of the M gene is also under the control of the C protein (Fig. 3C), whose binding to a single binding site (for a dimer; O_M in Fig. 3A) partially overlapping with the strong M promoter, excludes RNAP binding to the promoter and represses transcription of the M gene (configuration Z_D^M) (Bogdanova et al. 2009). For both the CR and the M promoter, configurations corresponding to basal transcription (configurations Z_{RNAP}^{CR} and Z_{RNAP}^M in Fig. 3, respectively) and empty promoters (statistical weight 1) were also assumed (Bogdanova et al. 2009). According to the classical Shea-Ackers assumption, which states that promoter transcription activity is proportional to the equilibrium probability of RNAP binding (Shea and Ackers 1985), the transcription activities of the CR(φ_{CR})and the M promoters(φ_M)are proportional to the probability of establishing their transcriptionally active configurations (for the statistical weights, see Fig. 3 caption):

$$\varphi_{CR} \propto \frac{a+b[C]^2}{1+a+b[C]^2+c[C]^4} \quad , \quad \varphi_M \propto \frac{f}{1+f+g[C]^2} \quad . \quad (1.1)$$

Transcripts (with concentration m_i, where $i = R, M, C$ denotes corresponding system components) synthesized from these promoters are degraded with a rate λ_i^m, while proteins (p_i) are generated by transcript translation with a rate k_i and are further degraded with a rate λ_i^p, as described by the following dynamical model equations:

$$\frac{dm_i(t)}{dt} = \varphi_i - \lambda_i^m \cdot m_i \quad , \quad \frac{dp_i(t)}{dt} = k_i \cdot m_i - \lambda_i^p \cdot p_i \quad . \quad (1.2)$$

It should be noted that the decay terms (λ) in equations include not only degradation of the transcripts and the proteins, but also their dilution due to cell division, which occurred with two very different rates during the first (0-160 min) and second time intervals (after 160 min) of the experiment. Consequently, the cell population dynamics are in part taken into account in the model through the decay terms. However, there are likely significant additional population dynamics effects that should, in principle, be included in the model, e.g. those related to possible changes in the cell metabolism and different plasmid and cell division

rates. Namely, our model describing the inherent RM system regulation and assuming constant parameters throughout the experiment (apart from different λ in the two time intervals) can successfully explain the main proposed qualitative features of system dynamics (Fig. 4), i.e. a large accumulation of M early upon plasmid entry into a naïve cell and a delay in the expression of R with respect to M, necessary for complete host genome protection. However, our model cannot completely quantitatively reproduce the system dynamics, i.e. there is a quantitative disagreement between the experimental data and the model predictions for M dynamics in the second time interval (after 160 min), likely arising from the additional population effects that we discussed above.

Design principles behind RM systems

The features of RM systems can be explained in terms of a few simple dynamical constraints that ensure safe and efficient RM system establishment. To this end, we proposed that all RM systems should exhibit the same simple dynamical properties: firstly, in every RM system there should be a significant expression of M prior to R, to avoid autoimmunity (Rodic et al. 2017). Once the host genome is protected (i.e. methylated), R should be rapidly generated, to "immunize" the host cell against virus infection, as fast as possible. Additionally, fluctuations of the toxic molecule R should be minimized, so as to evade that, due to large fluctuations, the toxic molecule amount is not matched by the antidote (M). Consequently, the following three dynamical properties are relevant to characterize RM system dynamics: (i) the time delay of R expression with respect to M; (ii) the transition

velocity of the system from "OFF" to "ON" state; (iii) the stability of R steady state levels.

To quantify these dynamical properties, we referred to the predicted system dynamics and the stability of R steady-state levels in the wild type (wt) AhdI system (Fig. 5). Accordingly, we introduced the following dynamical property observables (Rodic et al. 2017): (i) the ratio of the shaded areas in the perturbed and in the wt system for the first 10 min post-system entry as a measure of the time delay (Fig. 5A); (ii) the maximal slope of the sigmoidal R expression curve as a measure of the transition velocity from "OFF" (low R value) to "ON" (high R value) state (Fig. 5A); (iii) a measure of the stability of R steady-state levels (Fig. 5B) as derived in Bogdanova et al. (2008) – note that greater steady-state stability leads to smaller R fluctuations.

We here employed the biophysical model of wt AhdI transcription regulation that we previously developed and which was verified by the *in vitro* experimental measurements of the AhdI transcription activity dependence on C protein concentration (Bogdanova et al. 2008), and also the dynamical model of transcript and protein expression, which was also verified by *in vivo* measurements (see above and Morozova et al. (2016)). The described methodology, which involves a combination of thermodynamic and kinetic modeling, has been successfully applied to various systems in molecular biology (Munro et al. 2016). While there are few studies concerned with modeling some aspects of RM systems expression regulation (Williams et al. 2013), to our knowledge our work is the first to employ this modeling approach to systematically understand the relation between RM system regulation and its dynamics.

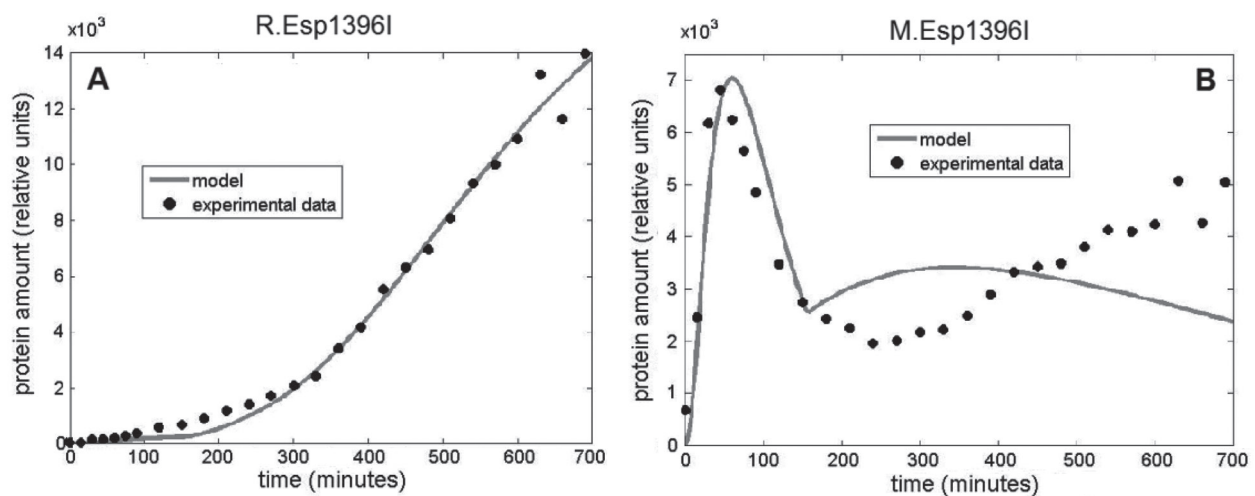In order to explain the (three) AhdI features, we per-

Fig. 4. Predicted Esp1396I RM system expression dynamics vs. experimental data. The change of R and M protein amounts in time is presented, respectively, in **A and B**. Circles correspond to the experimentally measured concentrations of protein fusions, while full lines correspond to the best fit of the model (described by the system of equations 1.1 and 1.2) to the data, obtained by varying parameters in biologically reasonable ranges. Time is set to zero at the point of the first available measurement. Adapted from Morozova et al. (2016).
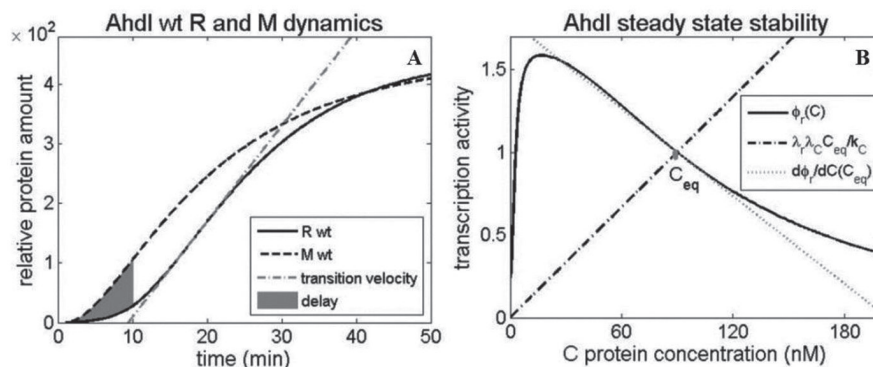
Fig. 5. Quantifying RM system dynamical properties A. R and M expression dynamics for AhdI RM system (Bogdanova et al. 2008). The shaded area presents a measure of a time delay between M (the dashed curve) and R (the solid curve) expression. The maximal slope of the sigmoidal R expression curve (dash-dotted line) is taken as a measure of the transition velocity from OFF to ON state. **B. Stability of the steady-state level**. The steady-state (C_{eq}) is obtained at the intersection of the CR promoter transcription activity (the solid curve) and the dash-dotted line, whose slope depends on the transcript and the protein decay rates and the protein translation rate (see Supplements in (Bogdanova et al. 2008) and (Rodic et al. 2017)). Larger difference in the slopes of the dash-dotted line and the solid curve at their intersection point leads to a more stable steady state. Adapted from Rodic et al. (2017).

turbed them *in silico*, one by one, to observe how this affects the dynamical property observables (Rodic et al. 2017). Firstly, we gradually increased the (initially low) C transcript translation initiation rate k_C towards the value characteristic of R and M transcripts. In Fig. 6A we observe a reduction in the delay between R and M expression, and a decreasing of the R steady-state level as the main effect of this perturbation. This finding can be intuitively explained by the fact that in increasing the translation initiation rate, C is generated faster, which hastens the formation of the activating and repressing complexes on the CR promoter. The effect on the other two observables is negligible. Consequently, this perturbation has a significantly adverse effect on one of the three dynamical properties (the delay between R and M expression), decreasing the ability of the system to protect the host genome from the cleavage.

Next, we gradually lowered the C subunit dissociation constant of dimerization K_1 from the very high value characteristic to the AhdI system, which corresponds to mostly C monomers in the solution, to low values, which correspond to predominantly C dimers in the solution, as shown in Fig. 6B (Bogdanova et al. 2008; Rodic et al. 2017). The three main effects of this perturbation are significant decreases in the time delay, in the transition velocity and in the steady-state levels of R. The stability of R steady-state levels is not significantly affected. Consequently, this perturbation has a significantly adverse effect on two dynamical properties, greatly reducing the ability of the system to protect the host genome from cleavage, and increasing the time window needed for the system to become protected from foreign DNA infection.

Finally, we gradually decreased only the extremely high cooperativity ω in C dimers binding to the CR promoter, which is shown in Fig. 6C (Rodic et al. 2017). We observe that this perturbation affects only the late R dynamics (see the left panel of Fig. 6C), since only efficiency in forming the repressor complex, whose probability is proportional to C^4, is affected, which becomes important only later on, when enough C is generated. Namely, this perturbation significantly decreases the stability of the steady state (see the right panel of Fig. 6C), thus having a significantly adverse effect on one dynamical property but not affecting the others. Also, contrary to the previous two perturbations, it significantly increases the steady-state levels of R, so that exhibiting different perturbations allows a balancing of the amount of the toxic molecule R in the cell.

To summarize, all three AhdI control features, in general, have the same effect on the dynamical properties, i.e. perturbing them makes at least one dynamical property much less optimal, while not notably affecting the other properties. This, together with the fact that decreasing the binding cooperativity ω has the opposite effect on the R steady-state levels from the other two perturbations (which facilitates controlling the toxic molecule R level) can explain the unusually large binding cooperativity in AhdI (Semenova et al. 2005; Bogdanova et al. 2009).

Dynamics of CRISPR/Cas system expression

Despite being intensively used in biotechnology for developing powerful genetic tools, the adaptive prokaryotic immune system CRISPR/Cas still appears to be underexplored when it comes to understanding the mechanism of its natural induction in a cell. In fact, the dynamics of CRISPR/Cas expression upon foreign DNA invasion have not been observed experimentally *in vivo*. What crucially hinders observing these dynamics is that CRISPR/Cas of Type I-E, which is the model system for CRISPR/Cas induction and regulation (most extensively studied in *E. coli*), is silent under
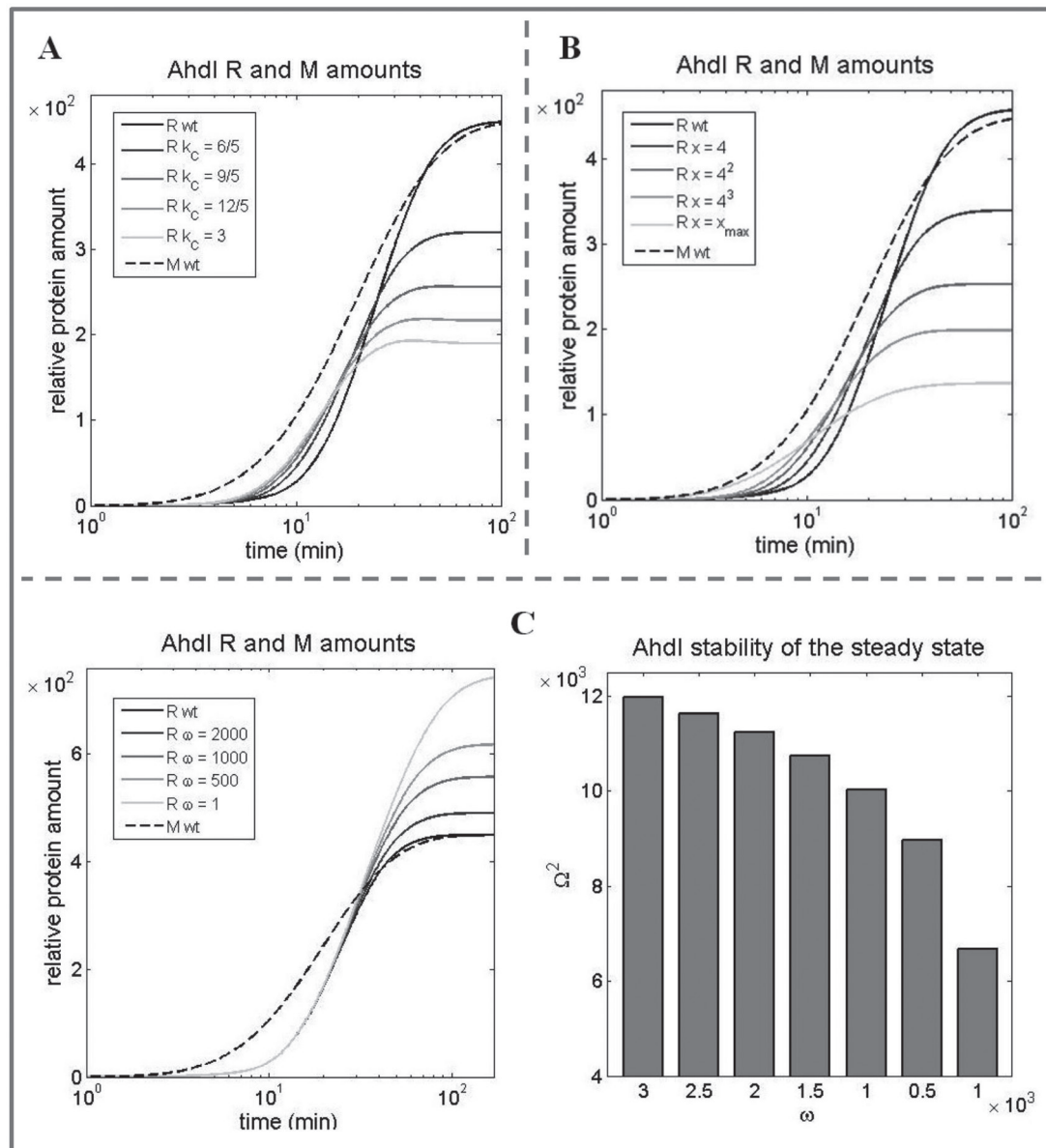
Fig. 6. Perturbing AhdI control features. A. Increasing C transcript translation initiation rate k_C. The effect of gradual k_C increase (from wt putative 3/5 1/min towards 3 1/min, which corresponds to the R and M (Bogdanova et al. 2008) is assessed on the protein expression dynamics, with R (solid) curves fading as k_C increases. The dashed curve corresponds to M expression, which is not affected by any of the three perturbations. **B.** Decreasing dissociation constant of C dimerization K_I. The effect of gradual K_I decrease from the high value, corresponding to only monomers in the solution, to the low value, corresponding to only dimers in the solution, is assessed on the protein expression dynamics, with R (solid) curves fading as K_I decreases. The relative protein amounts are derived from *in vitro* wt transcription activity measurements (Bogdanova et al. 2008). x denotes the ratio of K_I decrease. **C.** Decreasing cooperativity ω of C dimers binding to CR promoter in AhdI. The effect of gradual decrease of extremely high ω, inherent to the wt AhdI system (Bogdanova et al. 2008), to ω corresponding to no binding cooperativity is assessed on the protein expression dynamics (the left figure), with R (solid) curves fading as ω decreases. The stability of R steady-state levels (the right figure). Adapted from Rodic et al. (2017).

normal growth conditions, even in the presence of bacteriophage infection, and the induction mechanism is only partially known (Westra et al. 2010). However, the dynamical properties of CRISPR/Cas induction can be understood by examining how the system regulatory features contribute to the expression dynamics, which can be efficiently performed using quantitative modeling.

Our group previously dynamically modeled pre-crRNA processing into crRNAs upon CasE (processing) protein overexpression (Djordjevic et al. 2012). The proposed model (schematically represented in Fig. 7) takes into account that pre-crRNA is synthesized by transcription of the CRISPR array and then either nonspecifically degraded by an unidentified endonuclease or processed by CasE into crRNAs, which are further relatively slowly degraded. The model predicts that the system feature crucial for enabling the experimentally measured, very large (~2 orders of magnitude) amplification of crRNAs from a small decrease in pre-crRNA concentration upon CasE overexpression, is the rapid, nonspecific degradation of pre-crRNA. Therefore, the unidentified endonuclease is probably an essential component for achieving the fast system transition from "OFF" to "ON" state.

However, CasE proteins, which process pre-crRNA and which determine how the processing rate (k in the Fig. 7) depends on time, are gradually synthesized when the induction signal is received. Therefore, to model CRISPR/Cas system induction, in addition to the transcript processing, transcription regulation of the *cas* promoter also has to be incorporated in the model. As the mechanism of transcription induction is not known, to address this problem, we noted clear qualitative similarities in transcription regulation of CRISPR/Cas and RM systems. In particular, while the *cas* promoter is repressed by very cooperative binding of global regulators (such as H-NS proteins), which can be displaced from the promoter by some transcription activa-

tors (such as LeuO) (Westra et al. 2010), in the RM systems described above RNAP itself acts as an activator, displacing the recruited C dimer from the repressor position (see Fig. 3B) (Bogdanova et al. 2008, 2009). Therefore, our main idea is to consider a synthetic gene circuit where transcript processing, which is exhibited in the CRISPR/Cas system (Fig. 7), is put under the transcription control of an RM system that was previously studied in detail. Specifically, we assume that *cas* (including *casE*) genes are transcribed together with a gene encoding the C protein from a promoter regulated by the cooperative binding of C dimers, as described above (Fig. 3B). In this way, transcription control of a well-studied RM system serves as a proxy for the transcription control of a much less understood CRISPR/Cas system and can be thermodynamically modeled as described above.

In our future work, we plan to compare the behavior of the model described above with that of a setup in which *cas* genes are constitutively expressed, which we will use to explore: (i) how the cooperative *cas* promoter regulation (see above) is related to the expected sharp switch-like behavior of the system; (ii) how the dynamics of crRNA generation in the cooperative model compares to the limit of infinitely fast (abrupt) system induction (Djordjevic et al. 2012), and (iii) how the fast nonspecific degradation of pre-crRNA (which is the main feature of CRISPR transcript processing) affects the system dynamics.

Predicting CRISPR/Cas system components

As previously mentioned, CRISPR/Cas systems are the focus of current intensive research; however, efforts are predominantly invested into the development of promising biotechnology applications that revolutionize the concepts of programmable genome editing and gene expression regulation (Singh et al. 2017). Consequently, insights into the
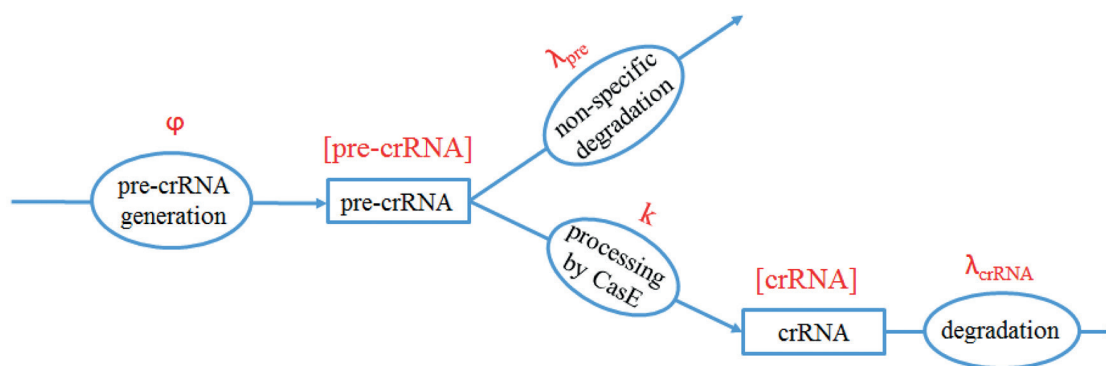

Fig. 7. The model scheme of pre-crRNA processing in CRISPR/Cas system. Notation used: φ – CRISPR promoter transcription activity, λ_{pre} – rate of (nonspecific) pre-crRNA degradation, k – rate of pre-crRNA processing to crRNAs by CasE, λ_{crRNA} – rate of crRNA degradation; square brackets denote concentrations of appropriate RNAs. Adapted from Djordjevic et al. (2012).

mechanisms that control the functioning of native CRISPR/Cas systems remain insufficiently explored. On the other hand, understanding native CRISPR/Cas function is crucial for the advancement of applied CRISPR/Cas research, which depends equally on the diversity of engineered CRISPR-based constructs and the capacity to control these constructs with sufficient precision.

An attractive avenue to improve the knowledge about native CRISPR/Cas systems, which could also lead to more powerful biotech applications, is investigating small CRISPR-associated RNAs. These RNA molecules (tracrRNAs), encoded outside the CRISPR array, are increasingly recognized as carriers of important regulatory and effector roles in the system. Namely, tracrRNAs are indispensable in Type II CRISPR/Cas systems for processing CRISPR array transcripts into mature crRNAs and subsequent targeting of the invasive genetic elements for degradation (in a complex with crRNA and Cas9 nuclease) (Deltcheva et al. 2011). At the same time, the underlying mechanism of action of this effector complex forms the basis for the Cas9:sgRNA paradigm that is extensively exploited for current CRISPR-based biotechnology applications (Hille and Charpentier 2016).

Despite their central role in CRISPR/Cas immunity and immense potential for translational research, small CRISPR-associated RNAs are largely unexplored, since their experimental discovery is complicated by (under standard conditions) a silent CRISPR/Cas system and still limited RNA-seq data in bacteria. An efficient alternative for the systematic identification and analysis of these small RNAs across different bacterial genomes is a bioinformatics-based approach, where the availability of sequenced genomic loci that encode CRISPR/Cas systems is the only prerequisite for computational analysis.

In general, small non-coding RNAs in bacteria are characterized by variable length, a low level of conservation and often indistinguishable secondary structure and nucleotide composition, so that *ab initio* detection, which is based on mining transcription signals (TSS and terminators) associated with small RNA expression units represents the most reliable search procedure (Sridhar and Gunasekaran 2013). However, a major shortfall of such an approach is that TSSs are often predicted with poor accuracy in bacterial genomes (Djordjevic 2014); for example, a standard supervised (information-theory based) search of the housekeeping (RpoD) promoter elements is associated with high rates of false positives.

Namely, due to considerable degeneracy of RpoD promoter elements, accurately aligning the -35 element to the -10 element is highly non-trivial, which was evidenced by our finding that the available -35 element alignments show a significant discrepancy with the biochemical data on σ^{70}-DNA interactions (Djordjevic 2011). In line with this, many implementations of the information-theory method use only the -10 element as the predictor of promoter specificity,

which negatively affects the search accuracy. To address this problem, we performed systematic *de novo* MLSA (*Multiple Local Sequence Alignment*) alignment of RpoD promoter elements in *E. coli*, based on a Gibbs search (for more details on methods see Djordjevic 2011), which provided improved -35 element characterization, along with the identification of the -15 element, a previously unrecognized determinant of RpoD specificity (Djordjevic 2011). As illustrated in Fig. 8, employing this new alignment for a weight matrix-based TSS search resulted in false-positive reduction by 50% (Nikolic et al. 2017), which clearly advocates the implementation of the new alignment within small CRISPR-associated RNA search procedure.
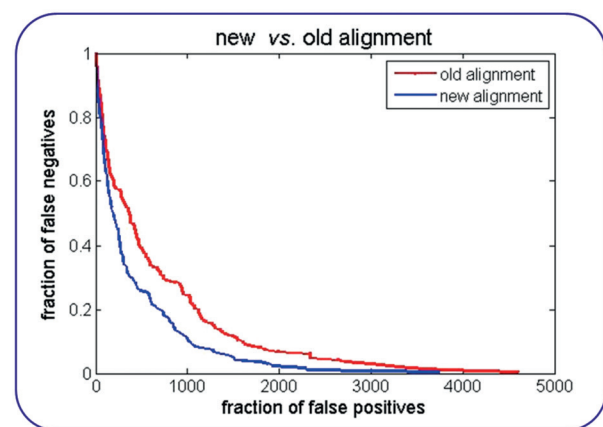

Fig. 8. DET (Detection Error Tradeoff) curve for the old and the new alignment of *E. coli* RpoD promoters. Fraction of false negatives is shown on the *y*-axis, and the fraction of false positives on the *x*-axis. DET-curve for the old alignment is colored red, and for the new alignment in blue. Adapted from Nikolic et al. (2017).

Compared to TSS, a terminator search is characterized by substantially higher accuracy, so that adaptation of the standard algorithm for Rho-independent terminator prediction in bacteria (Ermolaeva et al. 2000) can be used to detect small CRISPR-associated RNAs. Actually, for both TSSs and terminators, the search parameters can be trained against experimentally determined tracrRNAs across Type II CRISPR/Cas systems, where distinguishing true predictions (small RNAs) can be further aided by querying the predicted expression units for complementarity to the array direct repeats. Finally, secondary evidence for these *ab initio* predictions can be obtained through conservation analysis across related bacterial strains and mining available RNA-seq data. This, altogether, will be the core approach in our future research, which will focus on the systematic identification of small associated RNAs across diverse (Type II) CRISPR/Cas systems, with the goal of acquiring deeper insight into the functioning of native CRISPR/Cas systems.

The proposed procedure for small CRISPR-associated RNA detection is based on predicting housekeeping (RpoD)

promoter elements; however, CRISPR/Cas induction is also (likely) related to the activity of alternative (ECF) σ factors, that takeover bacterial transcription in response to cell-envelope stress (Ratner et al. 2015). However, ECF promoter prediction is far more challenging, as the binding specificity in this highly versatile group of alternative σ factors (Staron et al. 2009) was largely unknown. Consequently, to address this problem we firstly systematically explored protein and DNA interaction motifs that are involved in transcription initiation by alternative σ factors, as described in the next section.

Transcription by ECF σ factors

Distinct from housekeeping (RpoD) σ factors that globally control bacterial transcription under standard growth conditions, alternative σ factors transcribe more specialized regulons in response to signals related with stress, metabolic changes or development. Among these, ECF σ factors are the most abundant and diverse, yet the underlying mechanisms of ECF transcription initiation are largely unexplored (Helmann 2002). Signaling cascades that activate ECF-specific transcriptional response are mainly triggered at the level of the cell membrane (Brooks and Buchanan 2008), which, on the other hand, is related to the invasion of foreign genetic elements into the bacterial cell. Consequently, equivalent signaling cascades are likely connected with CRISPR/Cas and ECF induction, so the analysis of ECF transcriptional mechanisms might further elucidate the regulatory mechanisms behind CRISPR/Cas activity.

Structurally, ECF σ factors are the simplest in the entire σ70 family, and, at the same time, characterized by the most versatile protein sequences (including DNA-binding domains). Accordingly, promoter specificity in this group is also highly diverse, as evidenced by the very limited capacity for ECF promoter cross-recognition (Rhodius et al. 2013). Clearly, inferring specificity for unexplored group members through comparative analysis against a number of experimentally characterized representatives is not applicable in the ECF σ group. However, it is this approach that underlies the current paradigm on ECF functioning, which assumes interaction with rigid promoters characterized by obligatory and well-conserved -35 and -10 elements (Staron et al. 2009; Feklistov et al. 2014).

The paradigm on ECF functioning is completely opposite to the mix-and-match mechanism of promoter recognition, which was well established in the housekeeping (RpoD) σ70 group (Hook-Barnard and Hinton 2007). Namely, the mix-and-match paradigm allows a flexible promoter element structure as long as the threshold transcription activity is accomplished through mutual complementation of promoter element interaction energies with the σ factor. The most extreme, and altogether best known example of this mechanism is -35 element absence in RpoD promoters, which is

accommodated through σ factor interactions with a strong -10 element extension (also recognized as dsDNA).

Contrary to current considerations, we identified this ultimate example of promoter element complementation in ECF promoter sequences, recognized by the outlier group members (phage 7-11 and phiEco32 σ factors), during our systematic computational analysis of ECF promoter specificity, where we employed an extensive comparison of protein and DNA sequences through pairwise and multiple, global and local alignments (Fig. 9), for details see Methods in (Guzina and Djordjevic 2016). The presence of the classical mix-and-match trademark in phage ECF promoters is the first example of promoter recognition flexibility in the group, which we further corroborated by identifying a (putatively interacting) conserved protein motif, immediately C-terminal from the domain σ2 boundary, through multiple global alignment of ECF protein sequences (Guzina and Djordjevic 2016).
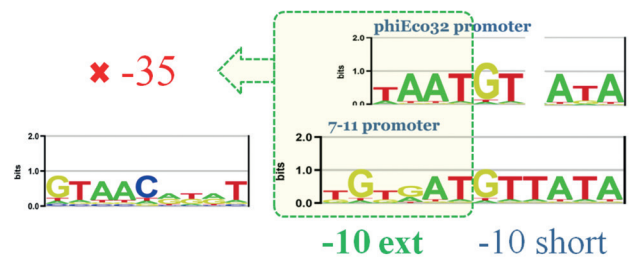

Fig. 9. Alignment of phage 7-11 and phiEco32 ECF promoters. Sequence-logo for 7-11 ECF promoters, with the presence of both -35 elements and long -10 element extensions is shown in the lower part of the figure; the logo for phiEco32 ECF promoters, where the presence of -10 element extension is followed by the absence of the -35 element, is shown in the upper part of the figure. Adapted from Guzina and Djordjevic (2016).

The coexistence of the conserved protein-DNA motifs was inferred in the bacterial ECF02 subgroup (containing experimentally well-characterized σE from *E. coli*) through multiple global and local alignments (Guzina and Djordjevic 2016). Interestingly, this novel σ-promoter interaction, whose partial conservation was also found in σW of *B. subtilis* (another canonical ECF member belonging to the ECF01 subgroup), appears further away from the domain σ2/-10 element boundary (Guzina and Djordjevic 2016). At the same time, protein-DNA interactions in the spacer with inversed polarity (i.e. closer to the domain σ4/-35 element boundary) are present in the ECF32 subgroup, which indicates that ECF σ factors display even greater flexibility during promoter recognition compared to the RpoD group. In fact, the observed flexibility in ECF promoter recognition aligns very well with the common biophysical mechanism of transcription initiation in the σ70 family, which is characterized by two major steps – closed and open complex formation (Djordjevic and

Bundschuh 2008). In the first step, σ^{70} factors interact with dsDNA promoter elements, while the second step depends on σ^{70} interactions with ssDNA elements. The interplay between these different energetic contributions determines the transcriptional output on the promoter, whose kinetic profile, in the framework of the mix-and-match mechanism, is indicated by the mutual complementation of the promoter elements, affecting the (most) relevant initiation step(s) for a given σ factor group.

In line with this, a biophysics-based correlation analysis we performed on a larger number of (*E. coli*) σ^{E} promoters (for more details on the analysis see ref. Guzina and Djordjevic 2017) revealed strong complementation between dsDNA elements, indicating that an efficient bacterial response to stress-related stimuli essentially depends on a high dsDNA-binding affinity of ECF σ factors for their promoters (Guzina and Djordjevic 2017). Correlations found between newly discovered spacer and canonical σ^{E} (-35 and -10) elements further corroborate the observed kinetic profile of ECF transcription initiation, which could, in turn, provide an alternative regulatory avenue for shaping the dynamics of CRISPR/Cas induction, where rapid expression of effector components (crRNA and Cas) appears as the main underlying signature. In our future research, we will use this detailed analysis of ECF σ factor specificity to develop methods for the accurate detection of TSS associated with these σ factors, which will, in turn, allow more accurate prediction of important CRISPR/Cas components, and consequently a better insight into the native system function.

Conclusion

Here we have reviewed our research on the modeling and bioinformatics of CRISPR/Cas and RM systems. We argue that the results presented to date show that combining experiments with modeling and bioinformatics is an optimal approach to understand the function of these exciting systems. Moreover, such an approach provides a better understanding of the common principles in design of these seemingly mechanistically quite different systems – understanding the principles that unify different biological systems is a major goal of systems biology. We believe that our current results provide a good starting point for understanding the regulation of diverse CRISPR/Cas and RM systems, including newly discovered CRISPR/Cas types. Regarding CRISPR/Cas, this can lead to new and improved biotechnology applications for a system that has already revolutionized the biotechnology field.

Acknowledgments

References

Al-Attar S, Westra ER, van der Oost J, Brouns SJ. 2011. Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes. Biological Chemistry. 392(4):277-289.

Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science. 315(5819):1709-1712.

Bogdanova E, Djordjevic M, Papapanagiotou I, Heyduk T, Kneale G, Severinov K. 2008. Transcription regulation of the type II restriction-modification system AhdI. Nucleic Acids Research. 36(5):1429-1442.

Bogdanova E, Zakharova M, Streeter S, Taylor J, Heyduk T, Kneale G, Severinov K. 2009. Transcription regulation of restriction-modification system Esp1396I. Nucleic Acids Research. 37(10):3354-3366.

Bolotin A, Quinquis B, Sorokin A, Enrlich SD. 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology. 151(Pt 8):2551-2561.

Brooks BE, Buchanan SK. 2008. Signaling mechanisms for activation of extracytoplasmic function (ECF) sigma factors. Biochimica et Biophysica Acta. 1778(9):1930-1945.

Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao J, Pirzada ZA, Eckert MR, Vogel J, Charpentier E. 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature. 471(7340):602-607.

Djordjevic M. 2011. Redefining *Escherichia coli* σ^{70} promoter elements: – 15 motif as a complement of the – 10 motif. Journal of Bacteriology. 193(22):6305-6314.

Djordjevic M. 2013. Modeling bacterial immune systems: strategies for expression of toxic - but useful - molecules. Biosystems. 112(2):139-144.

Djordjevic M. 2014. Integrating sequence analysis with biophysical modelling for accurate transcription start site prediction. Journal of Integrative Bioinformatics. 11(2):240.

Djordjevic M, Bundschuh R. 2008. Formation of the open complex by bacterial RNA polymerase--a quantitative model. Biophysical Journal. 94(11):4233-4248.

Djordjevic M, Djordjevic M, Severinov K. 2012. CRISPR transcript processing: a mechanism for generating a large number of small interfering RNAs. Biology Direct. 7(1):24.

Ermolaeva MD, Khalak HG, White O, Smith HO, Salzberg SL. 2000. Prediction of transcription terminators in bacterial genomes. Journal of Molecular Biology. 301(1):27-33.

Feklistov A, Sharon BD, Darst SA, Gross SA. 2014. Bacterial sigma factors: a historical, structural, and genomic perspective. Annual Review of Microbiology. 68:357-376.

Goldberg GW, Marraffini LA. 2015. Resistance and tolerance to foreign elements by prokaryotic immune systems - curating the genome. Nature Reviews. Immunology. 15(11):717-724.

Guzina J, Djordjevic M. 2016. Promoter recognition by ECF sigma factors: analyzing DNA and protein interaction motifs. Journal of Bacteriology. 198(14):1927-1938.

Guzina J, Djordjevic M. 2017. Mix-and-matching as a promoter recognition mechanism by ECF sigma factors. BMC Evolutionary Biology. 17(Suppl 1):12.

Hatoum-Aslan A, Marraffini LA. 2014. Impact of CRISPR immunity on the emergence and virulence of bacterial pathogens. Current Opinion in Microbiology. 17:82-90.

Heitman J. 1993. On the origins, structures and functions of restriction-modification enzymes. Genetic Engeneering (N Y). 15:57-108.

Helmann JD. 2002. The extracytoplasmic function (ECF) sigma factors. Advances in Microbial Physiology. 46:47-110.

Hille F, Charpentier E. 2016. CRISPR-Cas: biology, mechanisms and relevance. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences. 371(1707).

Hook-Barnard IG, Hinton DM. 2007. Transcription initiation by mix and match elements: flexibility for polymerase binding to bacterial promoters. Gene Regulation and Systems Biology. 1:275-293.

Kobayashi I, Nobusato A, Kobayashi-Takahashi N, Uchiyama I. 1999. Shaping the genome – restriction–modification systems as mobile genetic elements. Current Opinion in Genetics and Development. 9(6):649-656.

Morozova N, Sabantsev A, Bogdanova E, Fedorova Y, Maikova A, Vedyaykin A, Rodic A, Djordjevic M, Khodorkovskii M, Severinov K. 2016. Temporal dynamics of methyltransferase and restriction endonuclease accumulation in individual cells after introducing a restriction-modification system. Nucleic Acids Research. 44(2):790-800.

Mruk I, Blumenthal RM. 2008. Real-time kinetics of restriction-modification gene expression after entry into a new host cell. Nucleic Acids Research. 36(8):2581-2593.

Munro PD, Ackers GK, Shearwin KE. 2016. Aspects of protein–DNA interactions: a review of quantitative thermodynamic theory for modelling synthetic circuits utilising LacI and CI repressors, IPTG and the reporter gene *lacZ*. Biophysical Reviews. 8(4):331-345.

Nagornykh MO, Bogdanova ES, Protsenko AS, Zakharova MV, Solonin AS, Severinov KV. 2008. [Regulation of gene expression in type II restriction-modification system]. Genetika 44(5):606-615.

Nikolic M, Stankovic T, Djordjevic M. 2017. Contribution of bacterial promoter elements to transcription start site detection accuracy. Journal of Bioinformatics and Computational Biology 15(2):1650038.

Pul U, Wurm R, Arslan Z, Geissen R, Hofmann N, Wagner R. 2010. Identification and characterization of E. coli CRISPR-cas promoters and their silencing by H-NS. Molecular Microbiology. 75(6):1495-1512.

Raivio TL, Silhavy TJ. 2001. Periplasmic stress and ECF sigma factors. Annual Reviews of Microbiology. 55:591-624.

Ratner HK, Sampson TR, Weiss DS. 2015. I can see CRISPR now, even when phage are gone: a view on alternative CRISPR-Cas functions from the prokaryotic envelope. Current Opinion in Infectious Diseases. 28(3):267-274.

Rhodius VA, Segall-Shapiro TH, Sharon BD, Ghodasara A, Orlova E, Tabakh H, Brukhardt DH, Clancy K, Peterson TC, Gross CA, et al. 2013. Design of orthogonal genetic switches based on a crosstalk map of sigmas, anti-sigmas, and promoters. Molecular Systems Biology. 9:702.

Rodic A, Blagojevic B, Zdobnov E, Djordjevic M, Djordjevic M. 2017. Understanding key features of bacterial restriction-modification systems through quantitative modeling. BMC Systems Biology 11(Supplement 1):2.

Semenova E, Minakhin L, Bogdanova E, Nagornykh M, Vasilov A, Heyduk T, Solonin A, Zakharova M, Severinov K. 2005. Transcription regulation of the EcoRV restriction-modification system. Nucleic Acids Research. 33(21):6942-6951.

Shabbir MA, Hao H, Shabbir MZ, Wu Q, Sattar A, Yuan Z. 2016. Bacteria vs. Bacteriophages: Parallel Evolution of Immune Arsenals. Frontiers in Microbiology. 7:1292.

Shea MA, Ackers GK. 1985. The OR control system of bacteriophage lambda. A physical-chemical model for gene regulation. Journal of Molecular Biology. 181(2):211-230.

Singh V, Braddick D, Dhar PK. 2017. Exploring the potential of genome editing CRISPR-Cas9 technology. Gene. 599:1-18.

Sridhar J, Gunasekaran P. 2013. Computational small RNA prediction in bacteria. Bioinformatics and Biology Insights. 7:83-95.

Staron A, Sofia HJ, Dietrich S, Ulrich LE, Liesegang H, Mascher T. 2009. The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) sigma factor protein family. Molecular Microbiology. 74(3):557-581.

van der Oost J, Jore MM, Westra ER, Lundgren M, Brouns SJ. 2009. CRISPR-based adaptive and heritable immunity in prokaryotes. Trends in Biochemical Sciences. 34(8):401-407.

Vasu K, Nagaraja V. 2013. Diverse functions of restriction-modification systems in addition to cellular defense. Microbiology and Molecular Biology Reviews. 77(1):53-72.

Westra E., Pul U, Heidrich N, Jore MM, Lundgren M, Stratmann T, Wurm R, Raine A, Mescher M, Van Heerevald L, et al. 2010. H-NS-mediated repression of CRISPR-based immunity in Escherichia coli K12 can be relieved by the transcription activator LeuO. Molecular Microbiology. 77(6):1380-1393.

Williams K, Savageau AM, Blumenthal RM. 2013. A bistable hysteretic switch in an activator–repressor regulated restriction–modification system. Nucleic Acids Research. 41(12):6045-6057.

Wilson GG. 1991. Organization of restriction-modification systems. Nucleic Acids Research. 19(10):2539-2566.

Energy loss in jet suppression - what effects matter?

To cite this article: Bojana Blagojevic and Magdalena Djordjevic 2015 *J. Phys.: Conf. Ser.* **612** 012006

View the article online for updates and enhancements.

Related content

Energy loss in jet suppression - what effects matter?

Bojana Blagojevic and Magdalena Djordjevic

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

E-mail: `bojanab@ipb.ac.rs`

Abstract. Jet suppression is considered to be an excellent probe of QCD matter created in ultra-relativistic heavy ion collisions. Our theoretical predictions of jet suppression, based on our recently developed dynamical energy loss formalism, show a robust agreement with various experimental data for different probes, experiments (RHIC and LHC) and centrality regions. Our dynamical energy loss formalism includes the following key ingredients: dynamical scattering centers, collisional energy loss, finite magnetic mass and running coupling. Although all these ingredients are theoretically justified, it is currently unclear how they individually contribute to accurate suppression predictions. Natural question rises: is there one effect which is crucial for the agreement, or is the agreement a joint effect of several smaller improvements. To answer this question, we study how the above mentioned key effects affect the suppression calculations. Our results show that each energy loss effect is important and that a robust agreement between theoretical predictions and experimental data is a cumulative effect of all improvements.

1. Introduction

Suppression of high transverse momentum observables [1] is considered to be an excellent probe for mapping the properties of QCD matter created in ultra-relativistic heavy ion collisions at RHIC and LHC. Therefore comparison of available suppression experimental data with the theoretical predictions [2–4] tests different theoretical models and provides the insight into underlying QGP physics. For generating these predictions, we developed dynamical energy loss formalism which includes the following energy loss effects: *i)* dynamical scattering centers, *ii)* QCD medium of a finite size [5, 6], *iii)* both radiative [5, 6] and collisional [7] energy losses, *iv)* finite magnetic mass effects [8] and *v)* running coupling [9]. We further incorporated this energy loss formalism into a numerical procedure [9] in order to obtain suppression predictions. In the numerical procedure, accurate energy loss calculations are considered to be crucial for obtaining reliable suppression predictions.

We have shown that the suppression predictions obtained from this dynamical energy loss formalism are in a very good agreement with the available experimental data for both RHIC and LHC experiments, light and heavy flavor probes and different centrality ranges [9–11].

We here address the importance of different energy loss effects in the suppression calculations for D mesons (as a clear energy loss probe) in central 200 GeV Au+Au collisions at RHIC, because fragmentation function does not modify bare charm quark suppression [10, 12]. Our approach is to systematically include different energy loss effects. In particular, we first investigate the importance of including collisional energy loss and thus necessity of abolishing static in favor of dynamical approximation. Next we address the importance of including finite magnetic mass in the suppression calculations and finally the running coupling.

2. Theoretical and computational formalism

In this section, we give a brief description of our dynamical energy loss formalism [9] with regression on how each effect, when added, altered energy loss expression, while in Section 4 we take the reverse approach - the historical approach, starting from a static approximation and moving to systematically include all the effects.

In order to obtain quenched spectra we use generic pQCD convolution given by Eq.(1) from [9]. The initial charm quark spectrum is computed according to [13] and energy loss probability includes both radiative and collisional energy losses in a finite size dynamical QCD medium, multi-gluon [14] and path length [15, 16] fluctuations. In our calculations we do not use the fragmentation function of charm quark into D meson, as explained in Section 1.

The radiative energy loss in a finite size dynamical QCD medium is given by Eq.(2.12) from [5], while the finite magnetic mass and running coupling are introduced according to [8] and [9], respectively. For the finite magnetic mass case we use the following range of magnetic to electric mass ratio: $0.4 < \mu_M/\mu_E < 0.6$, according to non-perturbative approaches [17–21], otherwise, $\mu_M = 0$ is used. Also when the running coupling is not included, in our calculations we use $\alpha_S = \frac{g^2}{4\pi} = 0.3$ and Debye mass $\mu_E = gT$, $(g = 2)$. Collisional energy loss is calculated in accordance with Eq.(14) from [7]. Transition from the static [22] to the dynamical approximation in terms of radiative energy loss is explained in [6].

In our calculations for the charm quark mass we use $M_c = 1.2$ GeV, for 0-5% central 200 GeV Au+Au collisions we assume an average medium temperature of T=225 MeV [10] and for the number of effective light quark flavors we use $n_f = 2.5$.

3. Comparison with experimental data

As we mentioned in Section 1, our dynamical energy loss formalism [9] leads to a very good agreement with suppression experimental data for diverse probes at both RHIC [10] and LHC [9] and for different centrality regions [11]. The suppression is expressed by the nuclear modification factor R_{AA} [4], which quantifies the QCD medium effects on the yield of high-p_T particles. Fig. 1, which shows comparison of the D meson R_{AA} predictions with corresponding R_{AA} measured at the LHC and comparison of the single electron R_{AA} predictions with non-photonic single electron R_{AA} measured at RHIC, reflects the above mentioned agreement.
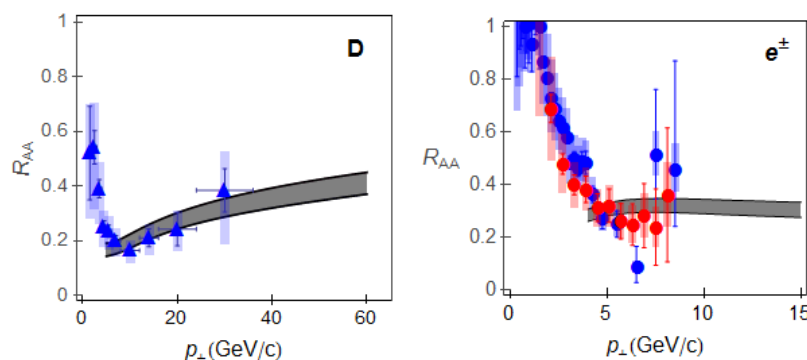

Figure 1. Theory vs. experimental data for D meson and single electron suppressions as a function of transverse momentum. Left panel shows comparison of D meson R_{AA} predictions with experimentally measured R_{AA} (triangle) in most central 2.76 TeV Pb+Pb collisions at the LHC. Right panel shows comparison of single electron R_{AA} predictions with non-photonic single electron R_{AA} (circle) measured in most central 200 GeV Au+Au collisions at RHIC. Left (right) panel is adapted from [9] ([10]).

4. Results and discussion

We start from the static approximation [22, 23] and use a constant value of the strong coupling constant and of Debye mass (as mentioned above), and no finite magnetic mass effects (μ_M=0); note that these values are used in Figs. 2 and 3, while the importance of finite magnetic mass is considered in Fig. 3. Previously, the static approximation was widely used, which assumed that collisional energy loss can be neglected compared to radiative. Left panel of Fig. 2 shows that static approximation has to be abolished, because collisional energy loss suppression is comparable or even larger than static radiative one. Therefore, central panel of Fig. 2 addresses the significance of including dynamical effects by comparing static with dynamical radiative energy loss R_{AA}. We observe a significant suppression increase in the dynamical approximation, so we conclude that dynamical effects are important. Right panel of Fig. 2 investigates whether collisional energy loss is still relevant in dynamical approximation, by comparing radiative with collisional contribution to R_{AA} in the dynamical QCD medium. We conclude that even in dynamical approximation, both radiative and collisional contributions are important, so we further include both radiative and collisional (total) energy losses in dynamical QCD medium.
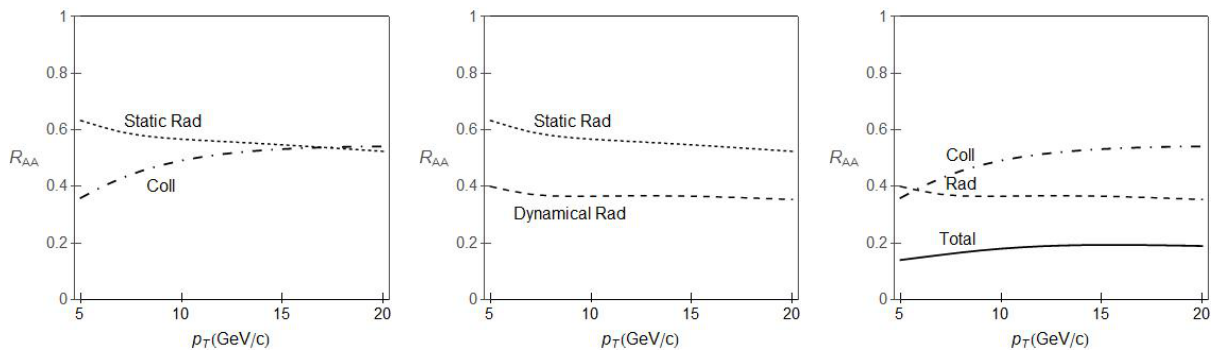

Figure 2. Static vs. dynamical approximation. D meson suppression predictions are shown as a function of transverse momentum. Left panel shows comparison of static radiative (dotted curve) with dynamical collisional (dot-dashed curve) contribution to R_{AA}. Central panel shows comparison of static radiative (dotted curve) with dynamical radiative (dashed curve) contribution to R_{AA}. Right panel shows radiative (dashed curve), collisional (dot-dashed curve) and radiative + collisional (solid curve) contribution to R_{AA} in dynamical QCD medium. Debye mass is $\mu_E = gT$, coupling constant is $\alpha_S = 0.3$ and no finite magnetic mass effects are included ($\mu_M = 0$). Adapted from [24].

Next we consider how inclusion of finite magnetic mass in radiative energy loss calculations [8] affects the R_{AA} predictions, as indicated in Section 2. By comparing R_{AA} with and without finite magnetic mass (Fig. 3), we observe significant suppression decrease due to finite magnetic mass effects. Hence, we conclude that finite magnetic mass effects are important.

Also, the importance of taking into account running coupling [9] is considered in Fig.7 from [24], where we observe suppression increase due to running coupling only at lower jet energies. Consequently running coupling is also important.

5. Conclusions

Since dynamical energy loss formalism led to a robust agreement with the suppression data for different experiments, probes and centrality ranges [9–11], we wanted to determine whether the agreement was a consequence of one dominant effect or a joint effect of several smaller
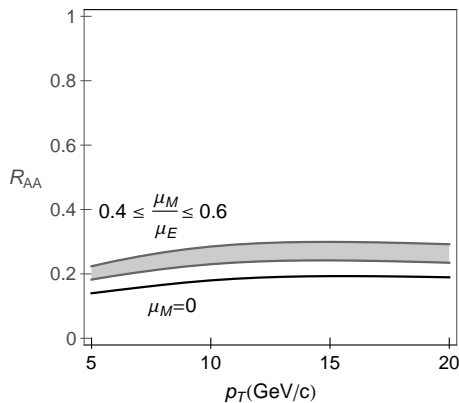
Figure 3. Magnetic mass effects on R_{AA}**.** D meson suppression predictions are shown, as a function of transverse momentum, for radiative and collisional energy loss in dynamical QCD medium, with (band) and without (solid curve) magnetic mass. Debye mass is $\mu_E = gT$ and coupling constant is $\alpha_S = 0.3$. The upper (lower) boundary of the band corresponds to $\mu_M/\mu_E = 0.6$ ($\mu_M/\mu_E = 0.4$). Adapted from [24].

improvements introduced to energy loss calculations. In order to examine the importance of each effect we followed first a historical approach starting from the static approximation and gradually introduced different energy loss effects in D meson suppression calculations (as a clear energy loss probe) until reaching dynamical energy loss formalism [9]. The conclusion is that each energy loss effect is important and that a robust agreement is a cumulative effect of all these improvements. Therefore, in order to obtain reliable suppression predictions we need to accurately account for all the relevant energy loss ingredients.

Acknowledgments

References
[1] Bjorken J D 1982 FERMILAB-PUB-82-059-THY pp 287-92
[2] Brambilla N et al 2004 *Preprint* hep-ph/0412158
[3] Gyulassy M 2002 *Lect. Notes Phys.* **583** 37
[4] d'Enterria D and Betz B 2010 *Lect. Notes Phys.* **785** 285
[5] Djordjevic M 2009 *Phys. Rev.* C **80** 064909
[6] Djordjevic M and Heinz U 2008 *Phys. Rev. Lett.* **101** 022302
[7] Djordjevic M 2006 *Phys. Rev.* C **74** 064907
[8] Djordjevic M and Djordjevic M 2012 *Phys. Lett.* B **709** 229
[9] Djordjevic M and Djordjevic M 2014 *Phys. Lett.* B **734** 286
[10] Djordjevic M and Djordjevic M 2014 *Phys. Rev.* C **90** 034910
[11] Djordjevic M, Djordjevic M and Blagojevic B 2014 *Phys. Lett.* B **737** 298
[12] Djordjevic M 2014 *Phys. Rev. Lett.* **112** 4 042302
[13] Kang Z B, Vitev I and Xing H 2012 *Phys. Lett.* B **718** 482-7
[14] Gyulassy M, Levai P and Vitev I 2002 *Phys. Lett.* B **538** 282
[15] Wicks S, Horowitz W, Djordjevic M and Gyulassy M 2007 *Nucl. Phys.* A **784** 426
[16] Dainese A 2004 *Eur. Phys. J.* C **33** 495
[17] Maezawa Yu, Aoki S, Ejiri S, Hatsuda T, Ishii N, Kanaya K, Ukita N and Umeda T 2010 *Phys. Rev.* D **81** 091501
[18] Maezawa Yu, Aoki S, Ejiri S, Hatsuda T, Ishii N, Kanaya K, Ukita N and Umeda T 2008 *Proc. of Science Lattice 2008 (Williamsburg)* p 194 (*Preprint* hep-lat/0811.0426)
[19] Nakamura A, Saito T and Sakai S 2004 *Phys. Rev.* D **69** 014506
[20] Hart A, Laine M and Philipsen O 2000 *Nucl. Phys.* B **586** 443
[21] Bak D, Karch A and Yaffe L G 2007 *J. High Energy Phys.* JHEP0708(2007)049
[22] Djordjevic M and Gyulassy M 2004 *Nucl. Phys.* A **733** 265-98
[23] Gyulassy M and Wang X N 1994 *Nucl. Phys.* B **420** 583
    Wang X N, Gyulassy M and Plumer M 1995 *Phys. Rev.* D **51** 3436
[24] Blagojevic B and Djordjevic M 2014 *Preprint* nucl-th/1411.1649

Modeling jet-medium interactions at RHIC and LHC - which energy loss effect is crucial?

To cite this article: B Blagojevic and M Djordjevic 2016 *J. Phys.: Conf. Ser.* **668** 012044

View the article online for updates and enhancements.

Modeling jet-medium interactions at RHIC and LHC - which energy loss effect is crucial?

B Blagojevic and M Djordjevic

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

E-mail: bojanab@ipb.ac.rs

Abstract. High momentum hadrons' suppression is considered to be excellent probe of QCD matter created in ultra-relativistic heavy ion collisions. Here we apply our recently developed dynamical energy loss formalism, which includes the following effects: dynamical scattering centers, QCD medium of a finite size, both radiative and collisional energy losses, running coupling and finite magnetic mass, and which we further incorporate into numerical procedure, to generate angular averaged R_{AA} predictions and to compare them with experimental R_{AA} data, by using no free parameters. A robust agreement of our predictions and experimentally measured R_{AA} for different energies, probes and all available centrality regions, raised the question whether this agreement is consequence of a single effect or of a superposition of all these effects. We obtained that, although the inclusion of dynamical scattering center has the largest relative importance, all the other effects are also important, since they lead to fine improvements of the agreement. Therefore, the robust agreement is a cumulative effect of all these features, with dynamical effect being crucial for accurate R_{AA} predictions.

1. Introduction

High momentum light and heavy flavor suppression [1] is considered to be excellent tool for studying QCD matter created in ultra-relativistic heavy ion collisions at RHIC and LHC. An abundance of suppression data, that has become available at RICH and LHC since recently, and its comparison with theoretical R_{AA} predictions [2–4], allows testing our understanding of QGP matter. In order to generate these predictions, we recently developed dynamical energy loss formalism, which we further integrated into numerical procedure as described in [5]. This formalism includes the following energy loss effects: *i)* dynamical scattering centers, *ii)* QCD medium of a finite size [6,7], *iii)* both radiative [6,7] and collisional [8] energy losses, *iv)* running coupling [5] and *v)* finite magnetic mass [9]. Also, note that, accurate energy loss calculation is considered to be the main ingredient responsible for obtaining reliable R_{AA} predictions.

In our previous papers [5, 10, 11], we demonstrated a robust agreement between our R_{AA} predictions, obtained as explained in previous paragraph, and R_{AA} data for both RHIC and LHC experiments, diverse set of probes and all available centrality ranges.

Here we address the relative importance of different energy loss effects in obtaining accurate angular averaged R_{AA} predictions for D mesons (as the clearest energy loss probe), for which it was previously shown that fragmentation function does not alter bare charm quark R_{AA} [10,12]. High momentum D meson R_{AA} data, obtained recently at LHC [13], serve as a baseline for testing the models. We concentrate on central 200 GeV Au+Au collisions at RHIC and 2.76 TeV Pb+Pb collisions at LHC. Our approach is to systematically include the effects in energy

loss calculations [14], i.e. we first compare the relative importance of radiative and collisional contribution to R_{AA} predictions, next we address the importance of including the dynamical scattering centers, then the running coupling and finally the finite magnetic mass. Note that only the main results are presented here; for the full account on the results, please see [14].

2. Theoretical and computational formalism

In this section, we concisely describe computational formalism, our dynamical energy loss formalism [5] and how each effect, when introduced, changed energy loss expressions.

For obtaining quenched spectra we apply generic pQCD convolution given by Eq.(2) from [14] ([15]). The initial charm quark spectrum is calculated in accordance with [16], while energy loss probability includes both radiative and collisional energy losses in a finite size dynamical QCD medium, multi-gluon [17] and path length fluctuations [15, 18].

The expression for radiative energy loss in a finite size dynamical QCD medium is given by Eq.(2.12) from [6], while the transition from static to dynamical scattering centers is explained in [7]. The collisional energy loss is calculated according to Eq.(14) from [8]. The running coupling is introduced in accordance with [5], while for constant coupling we use $\alpha_S = 0.3$ ($\alpha_S = 0.25$) in RHIC (LHC) case. Debye screening mass is $\mu_E = gT$ ($g = 2$). The finite magnetic mass is introduced as in [9], and its range ($0.4 < \mu_M/\mu_E < 0.6$) is set according to many non-perturbative approaches [19–23], otherwise $\mu_M = 0$ is used.

We model the medium by assuming an effective temperature of 221 MeV at RHIC [24] and 304 MeV at LHC [25]. No medium evolution is accounted. The validity of this assumption is discussed in [14]. For charm quark mass we use $M_c = 1.2$ GeV, and for the number of effective light quark flavors we use $n_f = 2.5$ ($n_f = 3$) in RHIC (LHC) case.
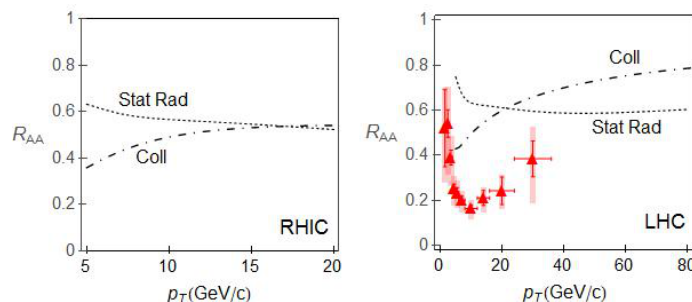
3. Results and discussion

Figure 1. Necessity of abolishing static approximation. D meson R_{AA} predictions, as a function of transverse momentum, are shown for only static radiative (dotted curve) and for only dynamical collisional (dot-dashed curve) contribution in a finite size QCD medium. Left (right) panel corresponds to RHIC (LHC) case. Right panel also displays D meson R_{AA} data in $0 - 7.5\%$ central 2.76 TeV Pb+Pb collisions at LHC (red triangles) [13]. Debye mass is $\mu_E = gT$, coupling constant is $\alpha_S = 0.3$ ($\alpha_S = 0.25$) for RHIC (LHC) and finite magnetic mass is not included ($\mu_M = 0$). Adapted from [14].

In this section we apply historically-driven approach, starting from static approximation [26, 27] and gradually adding energy loss effects. We display only the main results of our study carried out in [14]. Finite size QCD medium is assumed throughout the paper. The constant coupling and constant Debye mass (as mentioned above), and no finite magnetic mass are considered in Fig. 1 and Fig. 2. Static approximation, which assumes that the medium is composed of static

scattering centers, was firstly commonly used. It entails also, that collisional energy loss can be neglected compared to radiative one. However, Fig. 1 clearly shows that static approximation has to be abolished in favor of dynamical scatering centers' approximation, since collisional R_{AA} is comparable with radiative one. Further, we compute these both energy losses within the
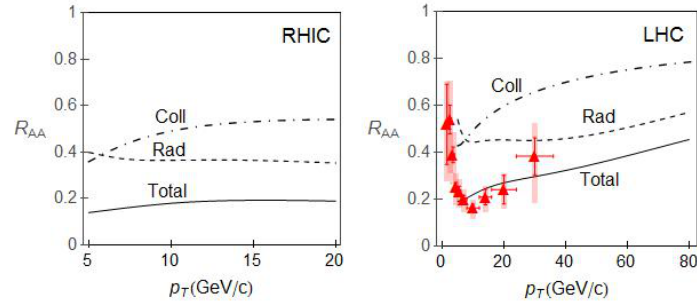

Figure 2. Dynamical approximation as the main effect. D meson R_{AA} predictions, as a function of transverse momentum, are shown for radiative (dashed curve), collisional (dot-dashed curve) and radiative + collisional (solid curve) energy losses in a finite size dynamical QCD medium. Left (right) panel corresponds to RHIC (LHC) case. Right panel also displays D meson R_{AA} data in $0 - 7.5\%$ central 2.76 TeV Pb+Pb collisions at LHC (red triangles) [13]. Debye mass is $\mu_E = gT$, coupling constant is $\alpha_S = 0.3$ ($\alpha_S = 0.25$) for RHIC (LHC) and finite magnetic mass is not included ($\mu_M = 0$). Adapted from [14].

same dynamical framework (Fig. 2) and we draw three conclusions: 1) dynamical radiative R_{AA} alone is not sufficient to explain qualitatively nor quantitatively the LHC experimental data; 2) radiative and collisional R_{AA} are still both important; 3) the total R_{AA} is in rough agreement with experimental data. Therefore, the inclusion of dynamical scattering centers is the main effect responsible for obtaining accurate R_{AA} predictions. Finally, we address the importance of
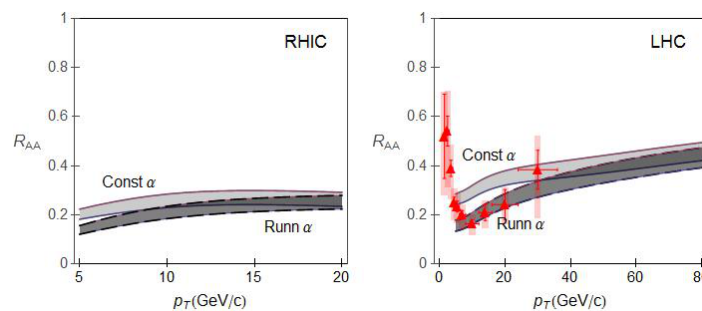

Figure 3. Our dynamical energy loss formalism. D meson R_{AA} predictions, as a function of transverse momentum, are shown for the constant coupling $\alpha_S = 0.3$ ($\alpha_S = 0.25$) for RHIC (LHC) (light gray band) and for the running coupling (dark gray band). In both cases radiative + collisional contributions in a finite size dynamical QCD medium are accounted. Upper (lower) boundary of each band corresponds to $\mu_M/\mu_E = 0.6$ ($\mu_M/\mu_E = 0.4$). Left (right) panel corresponds to RHIC (LHC) case. Right panel also displays D meson R_{AA} data in $0 - 7.5\%$ central 2.76 TeV Pb+Pb collisions at LHC (red triangles) [13]. Adapted from [14].

including the running coupling (leads to a significant R_{AA} decrease at lower jet momenta) and

finite magnetic mass (leads to a significant R_{AA} increase) [14]. From Fig. 3 we see that these two effects (although taken alone worsens the agreement [14]) taken together lead to quantitatively and qualitatively better agreement with the LHC R_{AA} data, compared to the case when these effects are omitted. This illustrates possible synergy in including these two effects.

4. Conclusions

A robust agreement of angular averaged R_{AA} predictions, based on our dynamical energy loss formalism, with R_{AA} data, for different energies, probes and centrality ranges, initiated the question: whether this agreement is a consequence of a one dominant energy loss effect or a joint effect of several smaller improvements [14]. With the LHC suppression data serving as a baseline, we here showed that (for the clearest energy loss probe: D meson R_{AA}), inclusion of dynamical scattering centers has the largest relative importance in obtaining accurate R_{AA} predictions. Furthermore, we found that all other considered effects are also important and responsible for the finer agreement with the data. So the good agreement is a result of a superposition of all these effects. Therefore, detailed study of partons' energy loss, as well as, inclusion of all relevant medium effects is necessary to correctly model the jet-medium interactions.

References

[1] Bjorken J D 1982 FERMILAB-PUB-82-059-THY pp 287-92
[2] Brambilla N *et al* 2014 *Eur. Phys. J.* C **74** 2981
[3] Gyulassy M 2002 *Lect. Notes Phys.* **583** 37
[4] d'Enterria D and Betz B 2010 *Lect. Notes Phys.* **785** 285
[5] Djordjevic M and Djordjevic M 2014 *Phys. Lett.* B **734** 286
[6] Djordjevic M 2009 *Phys. Rev.* C **80** 064909
[7] Djordjevic M and Heinz U 2008 *Phys. Rev. Lett.* **101** 022302
[8] Djordjevic M 2006 *Phys. Rev.* C **74** 064907
[9] Djordjevic M and Djordjevic M 2012 *Phys. Lett.* B **709** 229
[10] Djordjevic M and Djordjevic M 2014 *Phys. Rev.* C **90** 034910
[11] Djordjevic M, Djordjevic M and Blagojevic B 2014 *Phys. Lett.* B **737** 298
[12] Djordjevic M 2014 *Phys. Rev. Lett.* **112** 042302
[13] Grelli A 2013 *Nucl. Phys.* A **904-905** 635c
     Abelev B *et al* 2012 *J. High Energy Phys.* JHEP1209(2012)112
[14] Blagojevic B and Djordjevic M 2015 *J. Phys.* G **42** 075105
[15] Wicks S, Horowitz W, Djordjevic M and Gyulassy M 2007 *Nucl. Phys.* A **784** 426
[16] Kang Z B, Vitev I and Xing H 2012 *Phys. Lett.* B **718** 482-7
[17] Gyulassy M, Levai P and Vitev I 2002 *Phys. Lett.* B **538** 282
[18] Dainese A 2004 *Eur. Phys. J.* C **33** 495
[19] Maezawa Yu, Aoki S, Ejiri S, Hatsuda T, Ishii N, Kanaya K, Ukita N and Umeda T 2010 *Phys. Rev.* D **81** 091501
[20] Maezawa Yu, Aoki S, Ejiri S, Hatsuda T, Ishii N, Kanaya K, Ukita N and Umeda T 2008 *Proc. of Science Lattice* 2008 (Williamsburg) p 194 (*Preprint* hep-lat/0811.0426)
[21] Nakamura A, Saito T and Sakai S 2004 *Phys. Rev.* D **69** 014506
[22] Hart A, Laine M and Philipsen O 2000 *Nucl. Phys.* B **586** 443
[23] Bak D, Karch A and Yaffe L G 2007 *J. High Energy Phys.* JHEP0708(2007)049
[24] Adare A *et al* 2010 *Phys. Rev. Lett.* **104** 132301
[25] Wilde M 2013 *Nucl. Phys.* A **904-905** 573c
[26] Djordjevic M and Gyulassy M 2004 *Nucl. Phys.* A **733** 265-98
[27] Gyulassy M and Wang X N 1994 *Nucl. Phys.* B **420** 583
     Wang X N, Gyulassy M and Plumer M 1995 *Phys. Rev.* D **51** 3436

На основу члана 161 Закона о општем управном поступку и службене евиденције издаје се

УВЕРЕЊЕ

Благојевић (Боја) Бојана, бр. индекса 2013/8031, рођена 24.08.1984. године, Приједор, Република Српска, Босна и Херцеговина, уписана школске 2017/2018. године, у статусу: самофинансирање; тип студија: докторске академске студије; студијски програм: Физика.

Према Статуту факултета студије трају (број година): три.
Рок за завршетак студија: у двоструком трајању студија.

Ово се уверење може употребити за регулисање војне обавезе, издавање визе, права на дечији додатак, породичне пензије, инвалидског додатка, добијања здравствене књижице, легитимације за повлашћену вожњу и стипендије.

Овлашћено лице факултета

РЕПУБЛИКА СРБИЈА

ФИЗИЧКИ ФАКУЛТЕТ
УНИВЕРЗИТЕТА У БЕОГРАДУ

ДИПЛОМА

О СТЕЧЕНОМ ВИСОКОМ ОБРАЗОВАЊУ

Благојевић (Боја) Бојана

РОЂЕН-А 24.08.1984. ГОДИНЕ У ПРИЈЕДОРУ, ПРИЈЕДОР

БИХ , УПИСАН-А 2003/2004. ГОДИНЕ,

А ДАНА 01. ФЕБРУАРА 2013. ГОДИНЕ, ЗАВРШИО-ЛА ЈЕ СТУДИЈЕ НА

ФИЗИЧКОМ ФАКУЛТЕТУ УНИВЕРЗИТЕТА У БЕОГРАДУ, НА

СТУДИЈСКОЈ ГРУПИ ФИЗИКА

СА ОПШТИМ УСПЕХОМ 10 (ДЕСЕТ) У ТОКУ СТУДИЈА И

ОЦЕНОМ 10 (ДЕСЕТ) НА ДИПЛОМСКОМ ИСПИТУ.

НА ОСНОВУ ТОГА ИЗДАЈЕ МУ-ЈОЈ СЕ ОВА ДИПЛОМА О СТЕЧЕНОМ ВИСОКОМ

ОБРАЗОВАЊУ И СТРУЧНОМ НАЗИВУ

ДИПЛОМИРАНИ ФИЗИЧАР

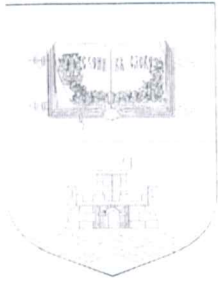РЕДНИ БРОЈ ИЗ ЕВИДЕНЦИЈЕ О ИЗДАТИМ ДИПЛОМАМА 242013

У БЕОГРАДУ, 07. ФЕБРУАРА 2013. ГОДИНЕ
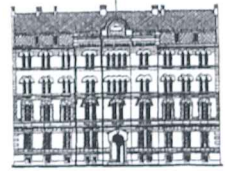
ДЕКАН   РЕКТОР

ПРОФ. ДР Јаблан Дојчиловић   ПРОФ. ДР Владимир Бумбаширевић

УНИВЕРЗИТЕТ У БЕОГРАДУ

Физички факултет

ДОДАТАК ДИПЛОМИ
диплома број 242013

1. Информације које идентификују имаоца квалификације

1.1 Презиме: Благојевић

1.2 Име: Бојана

1.3 Име једног од родитеља: Боја

1.4 Датум рођења (дд.мм.гггг.): 24. 08.1984.

1.5 Место рођења: Приједор

1.6 Држава рођења: Босна и Херцеговина

1.7 ЈМБГ 2408984165001

1.8 Број досијеа: 2030002

1.9 На овај факултет је први пут уписан 2003/2004. школске године, на прву годину студија

2. Информације које идентификују квалификацију

2.1 Стечени стручни назив квалификације: **Дипломирани физичар**

2.2 Главне области студирања за квалификацију: **Природно-математичке науке**

2.3 Назив установе која даје квалификацију: **Универзитет у Београду**

2.4 Врста установе која даје квалификацију: **универзитет**

2.5 Оснивач установе која даје квалификацију: **Република Србија**

2.6 Назив установе која управља студијама: **Физички факултет, Универзитета у Београду**

2.7 Оснивач установе која управља студијама: **Република Србија**

2.8 Језик студија: **српски**

3. Информације о нивоу квалификације

3.1 Ниво квалификације: Основне студије VII/1 степен

3.2 Званична дужина трајања студија (година, семестара): 4 године, 8 семестара

3.3 **Услови уписа:** У прву годину основних студија на Физичком факултету може се уписати лице које има средње образовање у четворогодишњем трајању

А: Упис после завршене средње школе

3.3.1а Назив и седиште школе: Гимназија Јован Дучић, Добој
Завршила 2003. године.

Б: Прелаз са другог факултета

3.3.1б Назив и седиште факултета: --

В: Упис после завршене више школе

3.3.1в Назив и седиште више школе: --
Завршио _____ године.

4.1 Начин студирања: **редован студент (са пуним обимом ангажовања)**

4.2 Захтеви програма: Студент који положи све испите утврђене наставним планом основних студија и одбрани дипломски рад стиче високо образовање и стручни назив ДИПЛОМИРАНИ ФИЗИЧАР

4.3 Појединости оствареног програма:

ши-фра	предмет	недељни фонд часова			испит	Наставник код кога је положен испит / Одговорни наставник
		преда-вања	вежбе	остало	оцена	
1б	Физичка механика и термофизика	120	240		10	Драгомир Крпић
2б	Математика IБ	120	120		10	Енес Удовичић
3б	Обрада резултат мерења	45	30		10	Иван Аничин
4б	Основи хемије	30	15		10	Зоран Миодраговић
5б	Енглески језик	60	60		10	Радмила Драговић
6б	Математика II Б	120	120		10	Мирослав Павловић
7б	Електромагнетизам	60	105		10	Наташа Недељковић
8б	Математичка физика I	60	45		10	Иванка Милошевић
9б	Таласи и оптика	60	105		10	Наташа Недељковић
10б	Теоријска механика	60	60		10	Сунчица Елезовић
11б	Филозофија физике	60	30		10	Радомир Ђорђевић
12б	Квантна механика	90	60		10	Милан Дамјановић
13б	Електрдинамика	60	60		10	Воја Радовановић
14б	Статистичка физика	60	60		10	Сава Милошевић
15б	Математичка физика II	60	60		10	Иванка Милошевић
16б	Физичка електроника	60	75		10	Александар Стаматовић
17б	Физика атома	60	75		10	Јагош Пурић
18б	Нуклеарна физика	60	75		10	Иван Аничин
19б	Физика молекула	60	75		10	Драгољуб Белић
20б	Квантна електродинамика	60	60		10	Маја Бурић
21т	Квантна статистичка физика	60	60		10	Сава Милошевић
22т	Симетрије у физици	45	30		10	Иванка Милошевић
23т	Теорија кондензованог стања	60	30		10	Зоран Радовић
24т	Физичка кинетика	45	30		10	Слободан Вуковић
25т	Теорија елементарних честица	45	30		10	Воја Радовановић
26т	Дипломски рад		90		10	Марија Вранеш Милосављевић

Тема дипломског рада: ПРОДУКЦИЈА МАСЕ И РАСПАДИ СУПЕРСИМЕТРИЧНИХ ЧЕСТИЦА У ОКВИРУ CMSSM МОДЕЛА НА LHC-У.
Наставник - ментор дипломског рада: др Марија Вранеш-Милосављевић
Датум одбране дипломског рада (дд.мм.гггг.): 01. 02. 2013. године

4.4 Начин оцењивања:

Оцена	Значење оцене	Проценат од укупног броја поена (%)
10	Изузетан	95-100
9	Одличан	85-94
8	Врло добар	75-84
7	Добар	65-74
6	Довољан	55-64
5	Недовољан	≤ 54

4.5 Укупна оцена квалификације:
Просечна оцена: 10,00 (десет и 00/100)

5   Информације о функцији квалификације

5.1   Приступ даљем студирању:Ималац дипломе има право да настави даље образовање и усавршавање у складу са Законом о високом образовању

6   Додатне информације

6.1   --

6.2 Извори додатних информација:
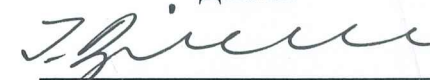О установи: www.ff.bg.ac.rs ; www.mps.sr.gov.rs
О програму: www.ff.bg.ac.rs

7.   Оверавање додатка:

7.1 Овај додатак дипломи односи се на следећа оригинална документа :
Диплома факултета број: 242013 издата дана (дд.мм.гггг.): 07. 02. 2013. године

Декан: Ректор:

_____   _____
Проф. др Јаблан Дојчиловић   **Проф. др Владимир Бумбаширевић**

Универзитет у Београду ФИЗИЧКИ ФАКУЛТЕТ
Студентски трг 12, 11000 Београд
Поштански фах 44
Тел. 011 7158 151, 3281 375
ПИБ 100039173, Мат. бр. 07048190

University of Belgrade FACULTY OF PHYSICS
Studentski trg 12, 11000 Belgrade
Postal Box 44
Phone +381 11 7158 151, Fax +381 11 3282 619
www.ff.bg.ac.rs, dekanat@ff.bg.ac.rs

ПОТВРДА

Овим се потврђује да је БОЈАНА БЛАГОЈЕВИЋ, дипломирани физичар, пријавила Наставно-научном већу Физичког факултета докторску дисертацију под називом: „ТЕОРИЈСКА ПРЕДВИЂАЊА ГУБИТАКА ЕНЕРГИЈЕ ВИСОКО ЕНЕРГИЈСКИХ ЧЕСТИЦА У КВАРК-ГЛУОНСКОЈ ПЛАЗМИ“.

Наставно-научно веће Физичког факултета на својој седници одржаној 27. априла 2016. године усвојило је предложену тему и за ментора одредило др Магдалену Ђорђевић, научног сарадника Института за физику.

На седници Већа научних области природно математичких наука Универзитета у Београду одржаној 23. маја 2016. године кандидату је усвојена тема и одобрен рад на дисертацији.

Београд, 26.9.2018.

ДЕКАН ФИЗИЧКОГ ФАКУЛТЕТА

Проф. др Јаблан Дојчиловић