



Application of Bayesian Gaussian Process for Optimization of String Fusion Model Parameters

Vladimir Kovalenko

Saint Petersburg State University

from 27 May 2018 to 2 June 2018
Valday, Russia
Europe/Moscow timezone

QUARKS-2018
XXth International Seminar on High Energy Physics

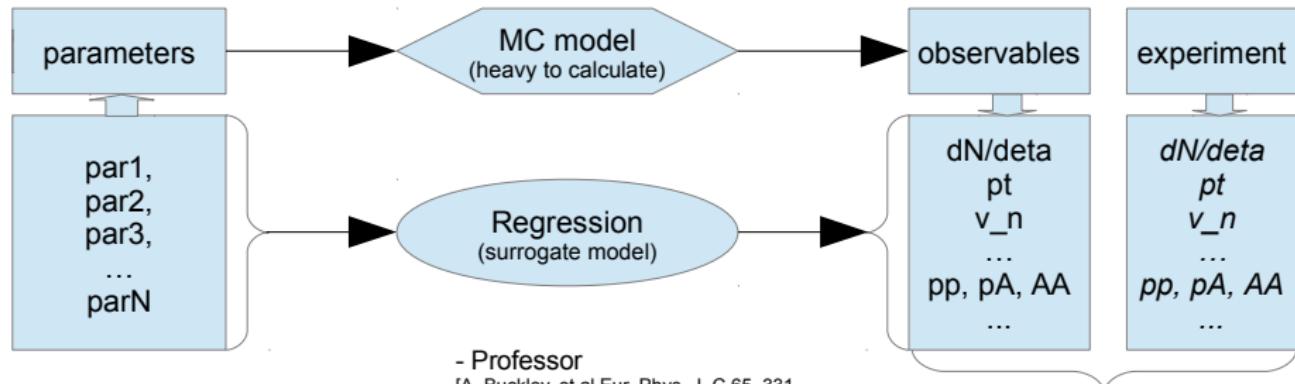
This work is supported by the Russian Science Foundation under grant 17-72-20045.

Outline

- Formulation of the problem for Bayesian Gaussian process optimization
- Description of the Non-Glauber Monte Carlo model with string fusion
- Results of Bayesian Optimization for string parameters
- Conclusions and Outlook

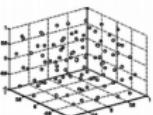
Parameter dependence approximation

- Monte Carlo models generally use many parameters, but can calculate many observables to compare with experiment



Sampling:

- grid (brute force)
- Latin hypercube (maximize minimum distance)



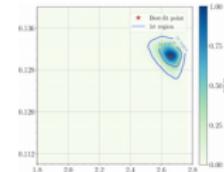
- Professor
[A. Buckley, et al Eur. Phys. J. C 65, 331 (2010)], (new version 2.2.2 more advance)

- Gaussian Processes
[C. E. Rasmussen C.K. I. Williams
The MIT Press, 2006]
<http://www.gaussianprocess.org/gpml/>

Example for heavy ion:
[Jonah E. Bernhard et al, Phys. Rev. C 94, 024907 (2016)]

Parameter estimation
+ error (Posterior)

Bayesian
parameter
estimation



4

Parameter dependence approximation Gaussian processes regression

Stochastic process is Gaussian if and only if for every finite set of indices

$\mathbf{X}_{t_1, \dots, t_k} = (\mathbf{X}_{t_1}, \dots, \mathbf{X}_{t_k})$ is a multivariate Gaussian random variable.

Gaussian processes can be completely defined by covariance functions. Examples:

- Constant : $K_C(x, x') = C$
- Linear: $K_L(x, x') = x^T x'$
- Gaussian noise: $K_{GN}(x, x') = \sigma^2 \delta_{x,x'}$
- Squared exponential: $K_{SE}(x, x') = \exp\left(-\frac{\|d\|^2}{2\ell^2}\right)$

In our analysis we will use sum of Squared exponential and Gaussian noise:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{|\mathbf{x}_i - \mathbf{x}_j|^2}{2\ell^2}\right) + \sigma_n^2 \delta_{ij}$$

Software we used:

GPML Matlab Code, v. 4.1, in GNU Octave, v. 4.0.0 <http://www.gaussianprocess.org/gpml/>
Scikit-learn Gaussian Process Regressor (python 2.7.12, Sklearn 0.18.1)
http://scikit-learn.org/stable/modules/generated/sklearn.gaussian_process.GaussianProcessRegressor.html

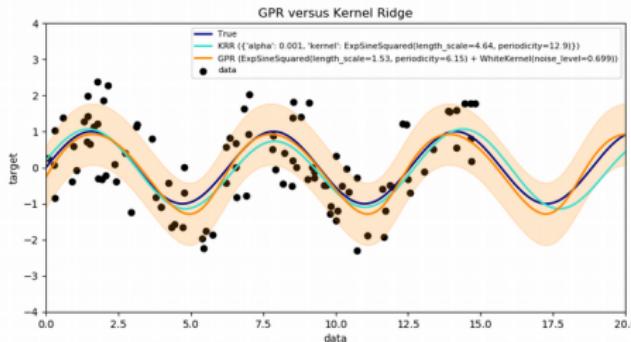
Parameter dependence approximation Gaussian processes regression

Advantages of Gaussian processes:

- The prediction interpolates the observations (at least for regular kernels).
- The prediction is probabilistic (Gaussian), one can compute empirical confidence intervals - can be used for online fitting, adaptive refitting

Disadvantages of Gaussian processes:

- They are not sparse, i.e., they use the whole samples/features information to perform the prediction.
- They lose efficiency in high dimensional spaces – namely when the number of features exceeds a few dozens.



Bayesian parameters optimization

Bayes' theorem

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{marginal likelihood}}, \quad p(x|y, X) = \frac{p(y|X, x)p(x)}{p(y|X)}$$

Prior – we take uniform (if the model good in constraining the data, the posterior will slightly depend on prior)

$$P(\mathbf{x}) \propto \begin{cases} 1 & \text{if } \min(x_i) \leq x_i \leq \max(x_i) \text{ for all } i \\ 0 & \text{else.} \end{cases}$$

Marginal likelihood – normalization constant

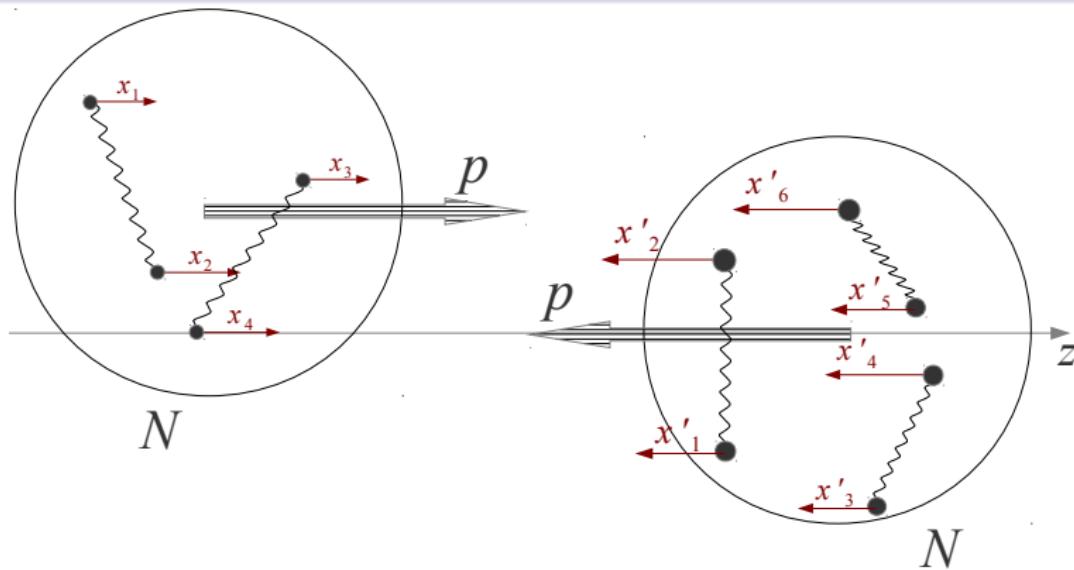
Likelihood (determined from the Gaussian Processes)

$$\ln P(x|y, X) = -\frac{1}{2} \frac{(y - y_{\text{exp}})^2}{\sigma_{\text{exp}}^2} + \text{const}$$

$$\text{Posterior } P(x|y, X) \propto e^{-\frac{1}{2} \frac{(y - y_{\text{exp}})^2}{\sigma_{\text{exp}}^2}}$$

For several observables we take independent Gaussian Process for each one and then combine in the Likelihood

Model description: Color dipoles inside a nucleon



$$\sum_i x_i p = p$$

$$\sum_i x_i = 1$$

$$\sum_i x'_i p' = p$$

$$\sum_i x'_i = 1$$

p-p interaction: parton distributions

- Inclusive momentum distributions are taken from [1,2]:

$$f_u(x) = f_{\bar{u}}(x) = C_{u,n} x^{-\frac{1}{2}} (1-x)^{\frac{1}{2}+n},$$

$$f_d(x) = f_{\bar{d}}(x) = C_{d,n} x^{-\frac{1}{2}} (1-x)^{\frac{3}{2}+n},$$

$$f_{ud}(x) = C_{ud,n} x^{\frac{3}{2}} (1-x)^{-\frac{3}{2}+n},$$

$$f_{uu}(x) = C_{uu,n} x^{\frac{5}{2}} (1-x)^{-\frac{3}{2}+n}.$$

- At $n > 1$ the sea quarks and antiquarks have the same distribution as the valence quarks.
- Poisson distribution for the number of quark-antiquark (diquark) pairs (n) is assumed with some parameter λ

[1] A.B. Kaidalov, O.I.Piskunova. Zeitschrift fur Physik C 30(1):145-150, 1986

[2] G.H. Arakelyan, A.Capella, A.B.Kaidalov, and Yu.M.Shabelski. Eur.Phys.J (C), 26(1):81-90, 2002

p-p interaction: parton distributions

- Corresponding exclusive distribution of the momentum fractions:

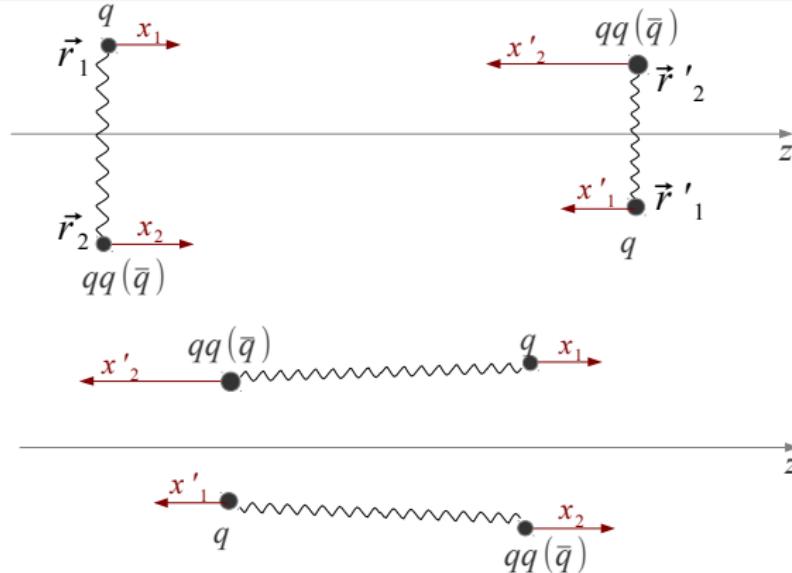
$$\rho(x_1, \dots, x_N) = c \cdot \prod_{j=1}^{N-1} x_j^{-\frac{1}{2}} \cdot x_N^{\alpha_N} \cdot \delta\left(\sum_{i=1}^N x_i - 1\right)$$

- Valence quark is labelled by N-1, the diquark by N, and the other refers to sea quarks and antiquarks.
- $N=2^*n$

10 Distribution in the impact parameter plane

- Exclusive distribution in the impact parameter plane is constructed from the following suppositions:
 - 1 Centre of mass is fixed: $\sum_{j=1}^N \vec{r}_j \cdot x_j = 0$.
 - 2 Inclusive distribution of each parton is the 2-dimentional Gaussian distribution.
 - 3 Normalization condition $\langle r^2 \rangle = \langle \frac{1}{N} \sum_{j=1}^N r_j^2 \rangle = r_0^2$.
- The parameter r_0^2 is connected with the mean square radius of the proton by the formula: $\langle r_N^2 \rangle = \frac{3}{2} r_0^2$.

Monte Carlo model: Color dipoles



Interaction probability amplitude [4, 5]:

$$(1) \quad f = \frac{\alpha_s^2}{2} \ln^2 \frac{|\vec{r}_1 - \vec{r}'_1| |\vec{r}_2 - \vec{r}'_2|}{|\vec{r}_1 - \vec{r}'_2| |\vec{r}_2 - \vec{r}'_1|}$$

Two dipoles interact more probably, if the ends are close to each other, and (others equal) if they are wide.

[4] G. Gustafson, Acta Phys. Polon. B40, 1981 (2009)

[5] C. Flensburg, G. Gustafson, and L. Lonnblad, Eur. Phys. J. (C) 60, 233 (2009)

Confinement radius

- With confinement taken into we obtain [4, 5]:

$$f = \frac{\alpha_s^2}{2} \left[K_0\left(\frac{|\vec{r}_1 - \vec{r}_1'|}{r_{max}}\right) + K_0\left(\frac{|\vec{r}_2 - \vec{r}_2'|}{r_{max}}\right) - K_0\left(\frac{|\vec{r}_1 - \vec{r}_2'|}{r_{max}}\right) - K_0\left(\frac{|\vec{r}_2 - \vec{r}_1'|}{r_{max}}\right) \right]^2 \quad (2)$$

where K_0 is modified Bessel function.

- At $r \rightarrow 0$ $K_0(r/r_{max}) \approx -\ln(r/(2r_{max}))$ and we return back to the formula (1).
- At $r \rightarrow \infty$: $K_0(r/r_{max}) \approx \sqrt{\frac{\pi r_{max}}{2r}} e^{-r/r_{max}}$
- and amplitude decrease exponentially.
- The total probability of the inelastic interaction of two protons in the eikonal approximation:

$$p = 1 - e^{-\sum_{i,j} f_{ij}}$$

[4] G. Gustafson, Acta Phys. Polon. B40, 1981 (2009)

[5] C. Flensburg, G. Gustafson, and L. Lonnblad, Eur. Phys. J. (C) 60, 233 (2009)

Calculation of multiplicity

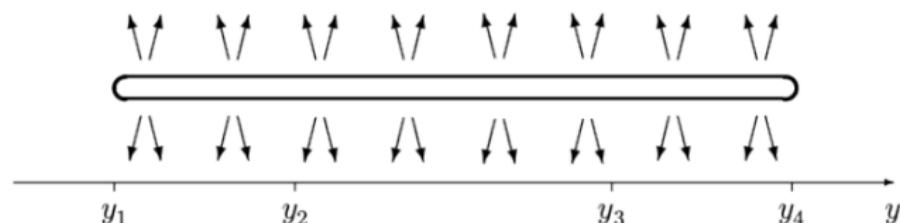
- Multiplicity is calculated in the framework of colour strings, stretched between colliding partons; x_i determine rapidity ends of strings.



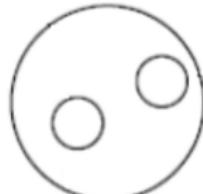
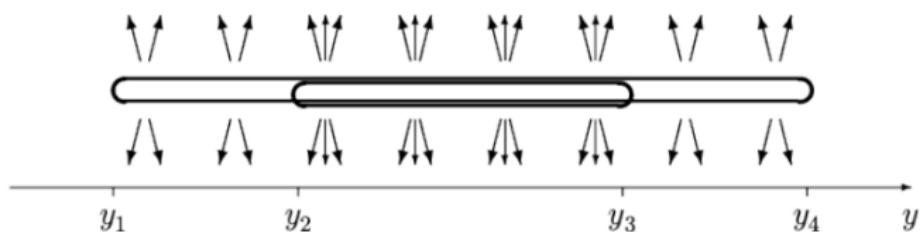
y_{\min} and y_{\max} are calculated supposing that a string fragments into only two particles with masses 0.15 GeV (for pion) and 0.94 GeV for proton and transverse momentum of 0.3 GeV (and higher at LHC)

- dN/dy from one string is supposed to be constant μ_0 .
- String fusion effects considered

- Uniform and independent distribution of particles on rapidity from y_{\min} to y_{\max}



- Can study string overlaps:

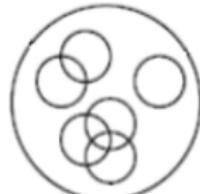


Multi-parton interactions

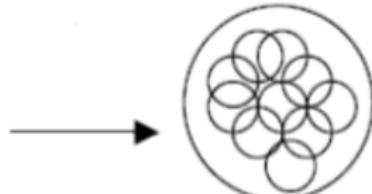


heavy ions

-->>> $\text{sqrt}(s)$ increases -->>>



-->>>

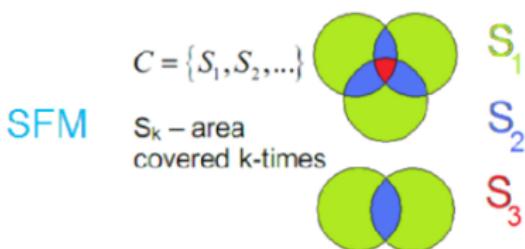


-->>>

$$Q^2(n) = \left(\sum_{i=1}^n \vec{Q}_i(1) \right)^2 = \sum_{i=1}^n Q_i^2(1) + \sum_{i \neq j} \vec{Q}_i(1) \cdot \vec{Q}_j(1)$$

$$\langle Q^2(n) \rangle = n Q^2(1)$$

overlaps



$$\langle \mu \rangle_k = \mu_0 \sqrt{k} \frac{S_k}{\sigma_0} \quad \langle p_t^2 \rangle_k = p_0^2 \sqrt{k} \quad \langle p_t \rangle_k = p_0 \sqrt[4]{k}$$

S_k – area, where k strings are overlapping, σ_0 single string transverse area, μ_0 and p_0 – mean multiplicity and transverse momentum from one string

String fusion mechanism predicts (agrees with experiment):

- decrease of multiplicity
- increase of p_T
- growth of p_T with multiplicity in pp, pA and AA collisions
- growth of strange particle yields

Key parameter – transverse radius of the string r_{str} : larger string

area – bigger overlapping

$r_{str} = 0$ - no fusion;

M. A. Braun, C. Pajares, Nucl. Phys. B 390 (1993) 542.

M. A. Braun, R. S. Kolevatov, C. Pajares, V. V. Vechernin, Eur. Phys. J. C 32 (2004) 535.

N.S. Amelin, N. Armesto, C. Pajares, D. Sousa, Eur.Phys.J.C22:149-163 (2001), arXiv:hep-ph/0103060

G. Ferreiro and C Pajares J. Phys. G: Nucl. Part. Phys. 23 1961 (1997)

- We use usual Woods-Saxon form of nuclear distribution:

$$\rho(r) = \rho_0 \frac{1}{1 + \exp\left(\frac{r - R}{\alpha}\right)}$$

- All partons from each nucleon are considered together
- A nucleus is participating in the collision if at least one of its partons collides with other from the proton.
- **Every parton can interact with other one only once** – this provides energy conservation in the initial state

- We use usual Woods-Saxon form of nuclear distribution:

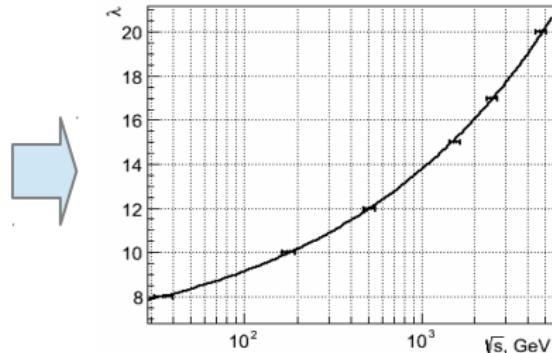
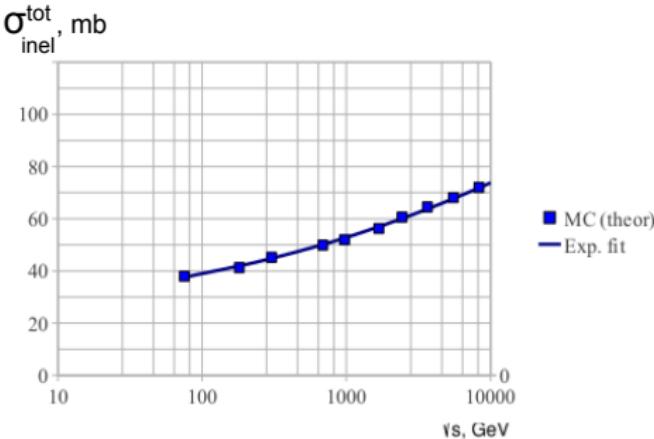
$$\rho(r) = \rho_0 \frac{1}{1 + \exp\left(\frac{r - R}{\alpha}\right)}$$

- All partons from each nucleon are considered together
- A nucleus is participating in the collision if at least one of its partons collides with other from the proton.
- **Every parton can interact with other one only once** – this provides energy conservation in the initial state

p-p interaction: parameter fixing

Strategy for parameters fixing:

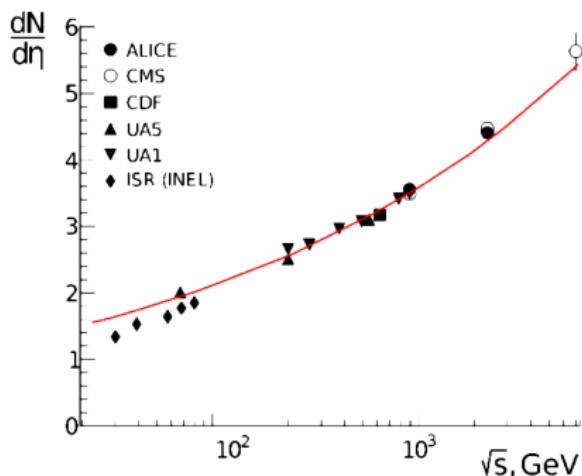
- Correspondence of mean number of dipoles λ and energy is obtained using data on total inelastic cross section
- Performed for each parameters combination and tabulated



p-p interaction: parameter fixing

Strategy for parameters fixing:

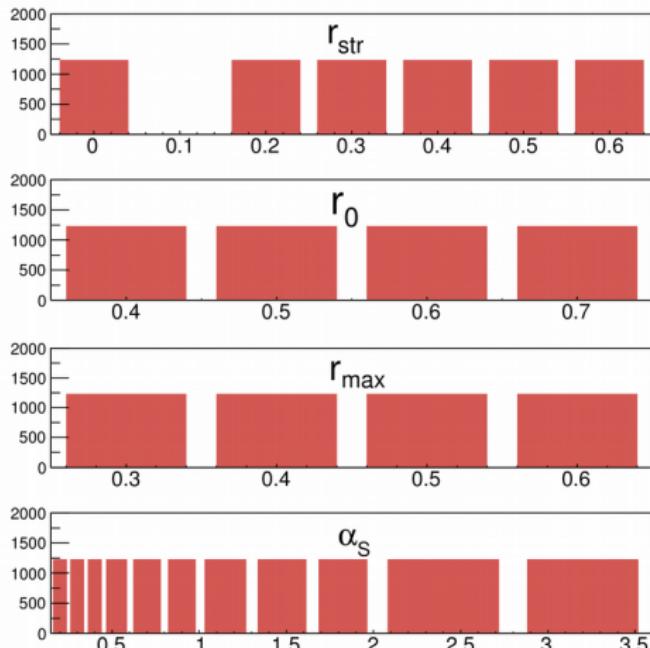
- Mean multiplicity per rapidity from one string μ_0 is fixed once at intermediate LHC energy (2.36 TeV)
- Data on energy dependence of multiplicity in pp collisions is used to constrain the rest of parameters
- p-Pb at 5.02 GeV minimum bias:
 $\langle dN/d\eta \rangle = 16.81 \pm 0.71$ [10]
- Look at PbPb collisions



Primary parameter distribution

Initial range of parameters:

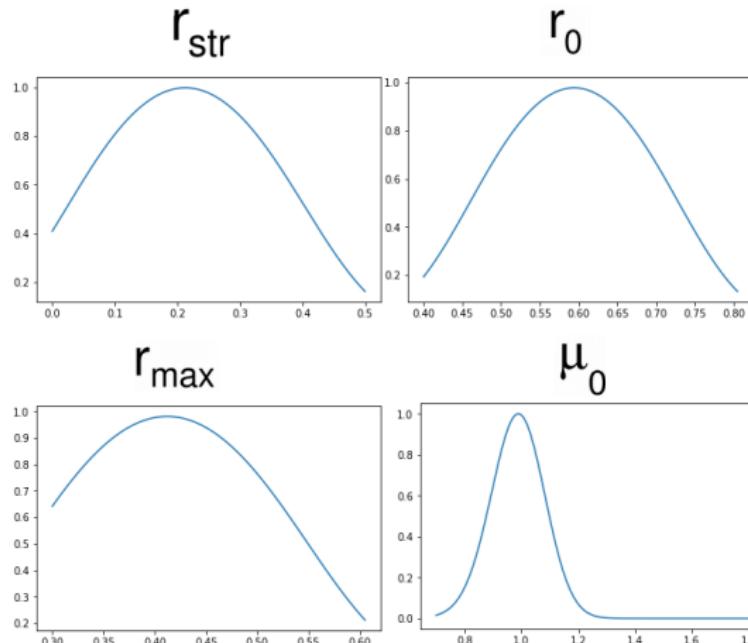
- r_0 : 0.4 – 0.7 fm
- r_{\max}/r_0 : 0.3 – 0.6
- α_S : 0.2 – 2.8
- r_{str} : 0 (no fusion); 0.2-0.6 fm
- Energy range: 53 – 7000 GeV



Posterior distributions after accounting of pp multiplicity

pp multiplicity in the model

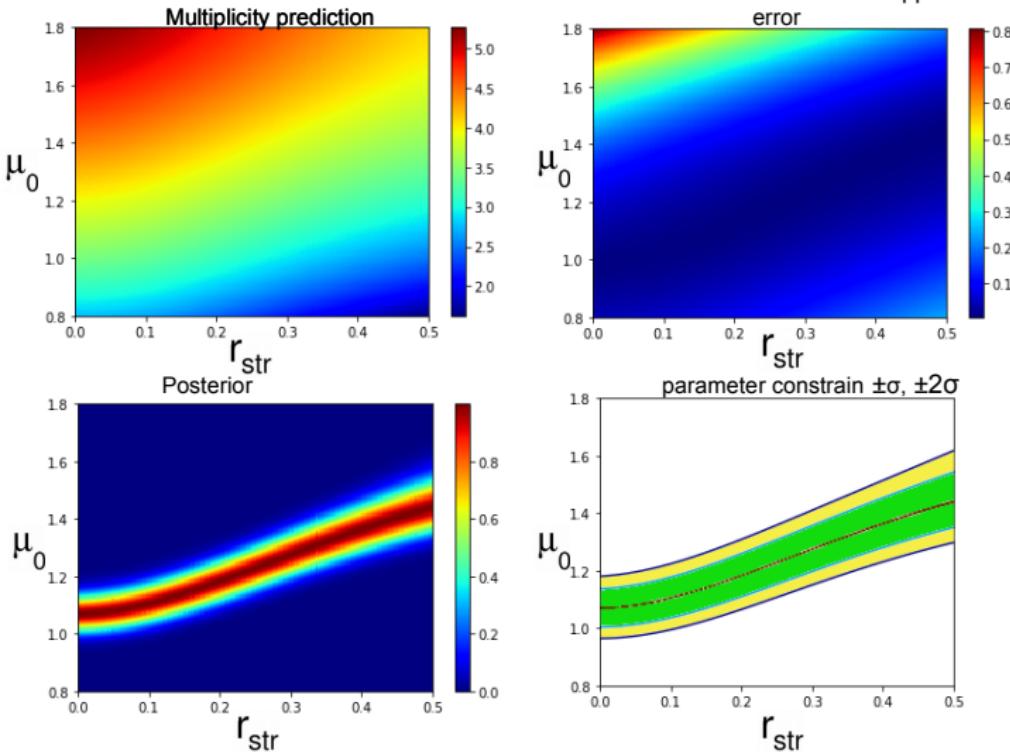
- Not much sensitive to string fusion r_{str}
- r_0 around 0.6 is favoured
- r_{max} and α_s : not well restricted
- μ_0 : peaks around 1.0



Posterior parameter estimation from energy dependence of pp multiplicity

- Focus on string related parameters: r_{str} and μ_0 :

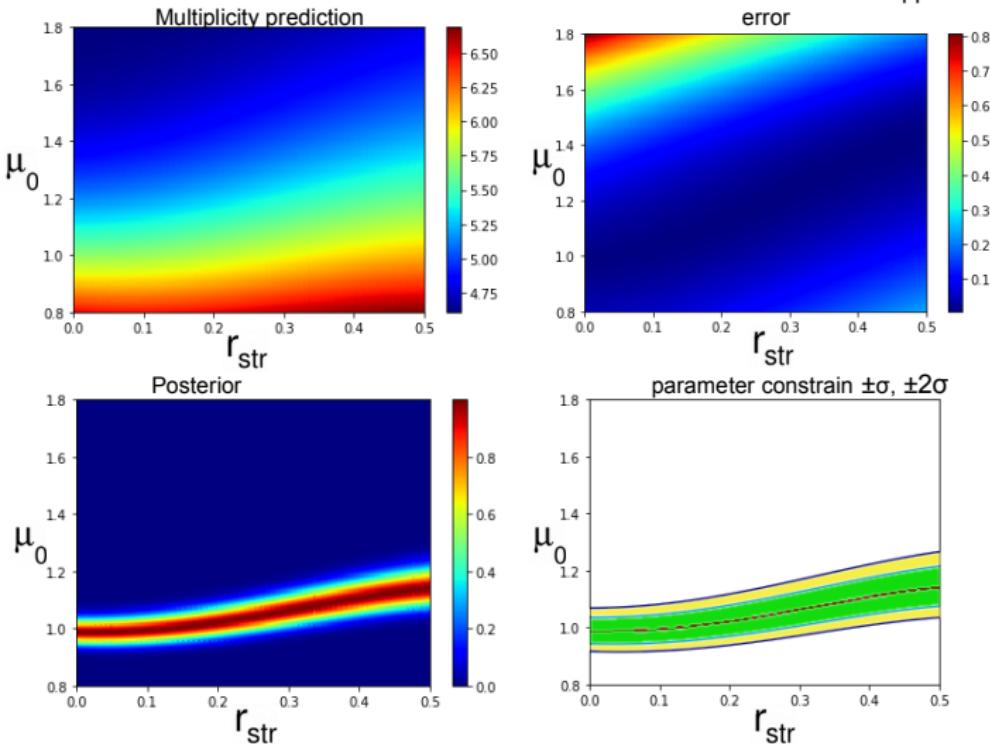
Experimental data:
 $\mu_{\text{pp}} = 3.61 \pm 0.17$ at 0.9 TeV



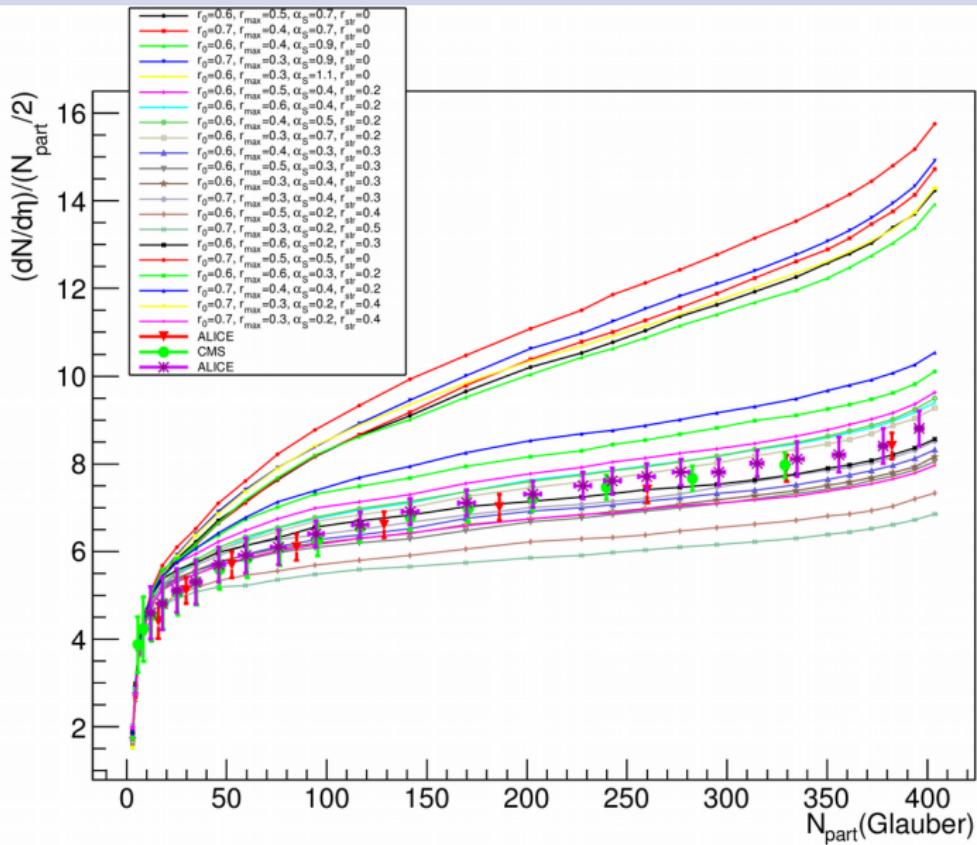
Posterior parameter estimation from energy dependence of pp multiplicity

- Focus on string related parameters: r_{str} and μ_0 :

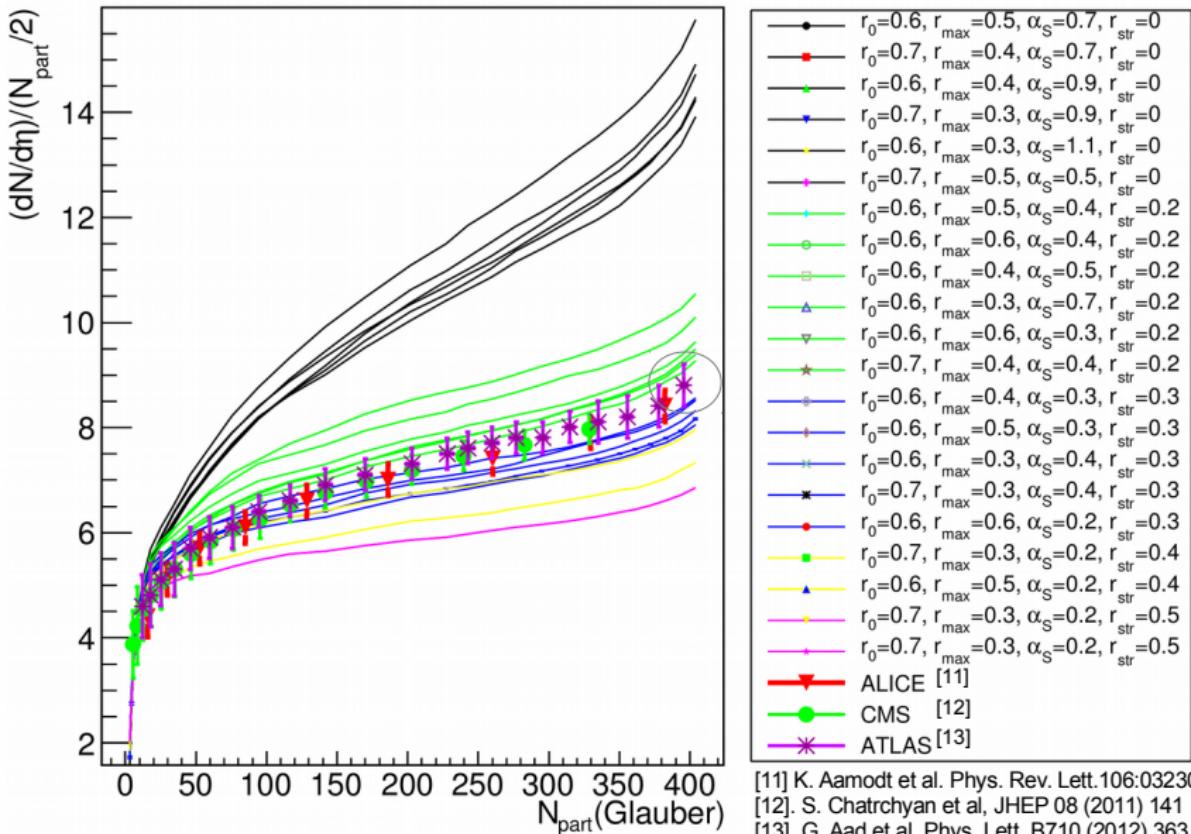
Experimental data:
 $\mu_{\text{pp}} = 5.74 \pm 0.15$ at 7.0 TeV



Model predictions for PbPb collisions at 2.76 TeV



Results: PbPb collisions at 2.76 TeV

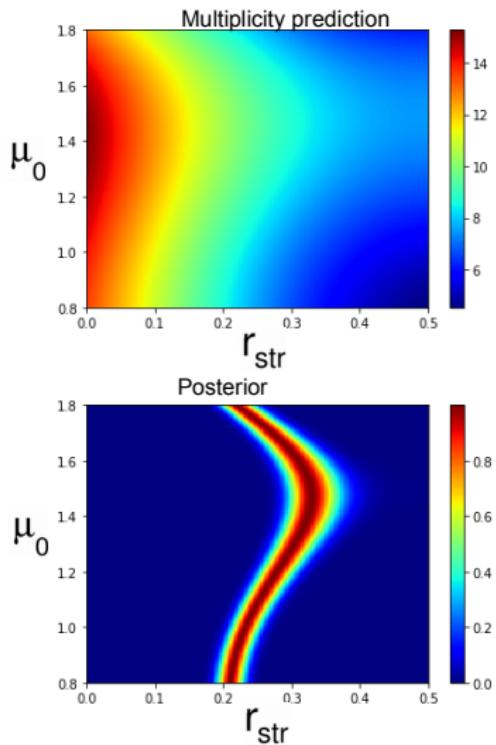


[11] K. Aamodt et al. Phys. Rev. Lett. 106:032301, 2011

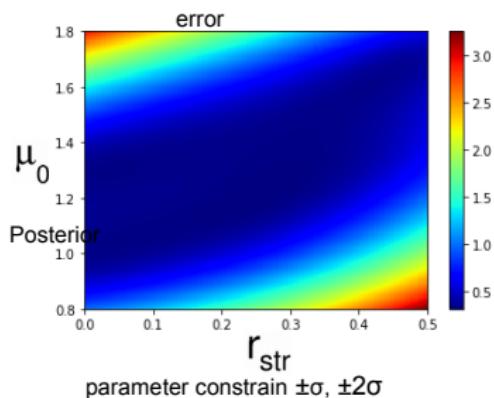
[12] S. Chatrchyan et al. JHEP 08 (2011) 141

[13] G. Aad et al. Phys. Lett. B710 (2012) 363

Posterior parameter estimation from central PbPb collisions at 2.76 TeV

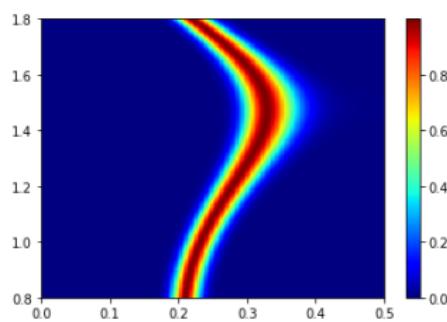


Experimental data:
 $\mu_{\text{PbPbcentral}}/\text{Npart}^2 = 8.4 \pm 0.4$

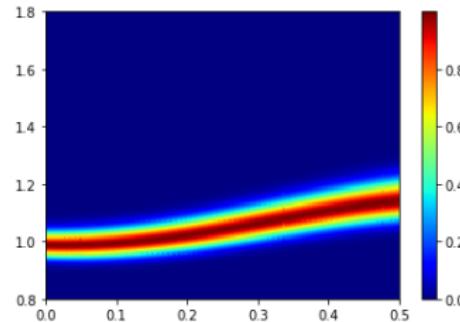


Combined results on Posterior and parameter estimation

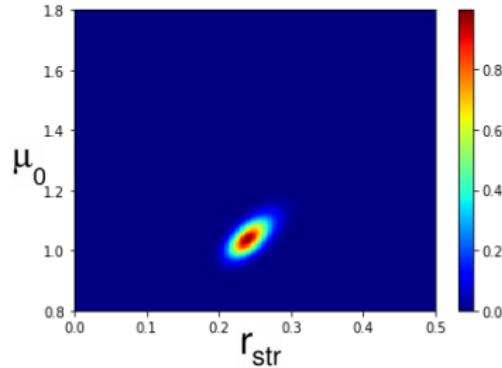
PbPb collisions at 2.76 TeV plus energy dependence of multiplicity in pp collisions



+

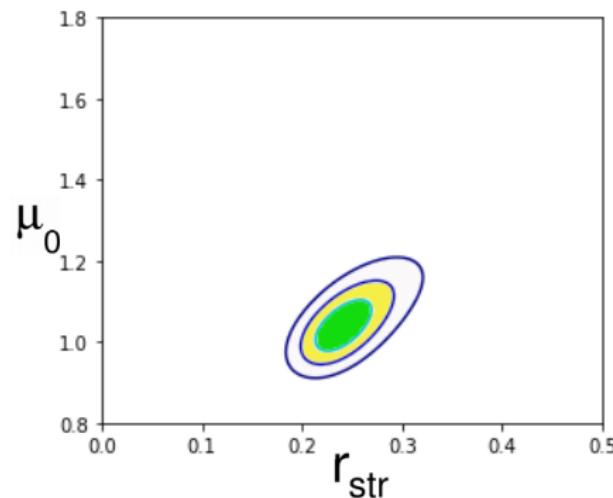
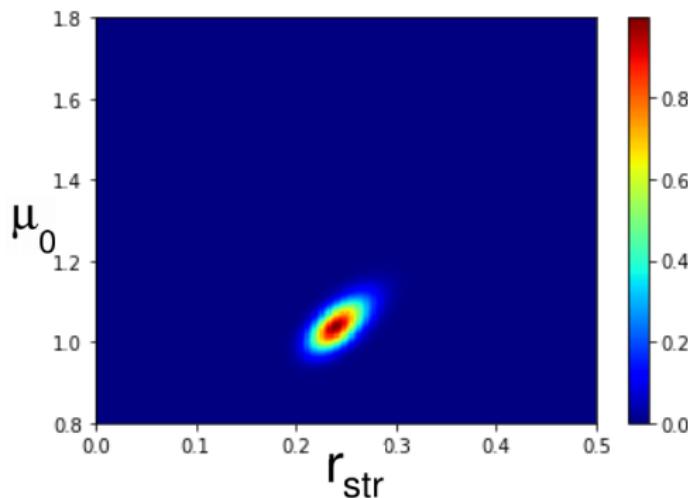


=



Combined results on Posterior and parameter estimation

PbPb collisions at 2.76 TeV plus energy dependence of multiplicity in pp collisions



Estimation

$$r_{\text{str}} = 0.25 \pm 0.03 \text{ fm}$$

$$\mu_0 = 1.1 \pm 0.03$$

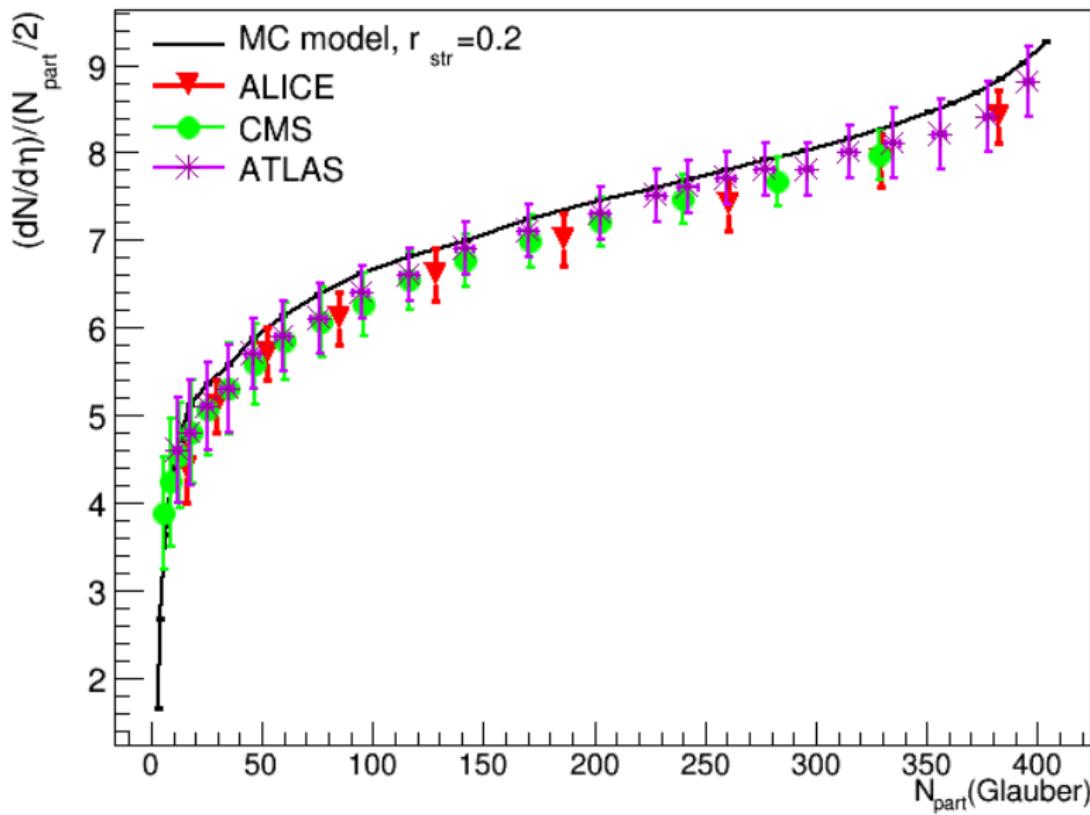
Conclusions

- Bayesian Gaussian process optimization has been applied for the parameter tuning of the non-Glauber Monte-Carlo with string fusion
- In the model the inelastic cross section and multiplicity are described wide energy range and for different colliding systems.
- Multiplicity per rapidity from one single string μ_0 is constrained by the energy dependence of the multiplicity in pp collisions
- The transverse radius of string r_{str} (string fusion parameter) is constrained by multiplicity in central PbPb collisions.

Outlook

- Improvement of the parameter estimation by considering more data
- Extension of the energy range
- Application of the Principal Component Analysis to observables

Backup



References

- V.Kovalenko. Modelling of exclusive parton distributions and long-range rapidity correlations for pp collisions at the LHC energy
accepted at Phys. Atom. Nucl. Vol. 93, N 10 (2013)
arXiv:1211.6209 [hep-ph]
- V.Kovalenko, V.Vechernin. Model of pp and AA collisions for the description of long-range correlations
PoS (Baldin ISHEPP XXI) 077
arXiv:1212.2590 [nucl-th]

- We have to introduce a new parameter – r_{max}
- Confinement effects can be taken into account by the replacement of the Coulomb propagator $\Delta(\vec{r}) = \int \frac{d^2 \vec{k}}{(2\pi)^2} \frac{e^{i\vec{k}\cdot\vec{r}}}{k^2}$, by the Yukawa one: $\frac{1}{k^2 + M^2}$, where $M = 1/r_{max}$ is the confinement specific scale.
- As a result, we get for the probability amplitude the following:

$$f = \frac{\alpha_s^2}{2} \left[K_0 \left(\frac{|\vec{r}_1 - \vec{r}_3|}{r_{max}} \right) + K_0 \left(\frac{|\vec{r}_2 - \vec{r}_4|}{r_{max}} \right) - K_0 \left(\frac{|\vec{r}_1 - \vec{r}_4|}{r_{max}} \right) - K_0 \left(\frac{|\vec{r}_2 - \vec{r}_3|}{r_{max}} \right) \right]^2 \quad (4)$$

- Squared ratio of the quark and hadron radiuses should be about $\frac{1}{10}$. It leads $r_{max} \simeq 0.2 - 0.3 fm$.

p-p interaction: color dipoles

- The probability amplitude for the collision of two dipoles with coordinates $(\mathbf{r}_1, \mathbf{r}_2), (\mathbf{r}_3, \mathbf{r}_4)$ [3,4]:

$$f = \frac{\alpha_s^2}{2} \ln^2 \frac{|\vec{r}_1 - \vec{r}_3| \cdot |\vec{r}_2 - \vec{r}_4|}{|\vec{r}_1 - \vec{r}_4| \cdot |\vec{r}_2 - \vec{r}_3|}$$

- Confinement is taken into account by introduction of some cut off at $r_{max} \approx 0.2 - 0.3 \text{ fm}$. It leads:

$$f = \frac{\alpha_s^2}{2} \left[K_0 \left(\frac{|\vec{r}_1 - \vec{r}_3|}{r_{max}} \right) + K_0 \left(\frac{|\vec{r}_2 - \vec{r}_4|}{r_{max}} \right) - K_0 \left(\frac{|\vec{r}_1 - \vec{r}_4|}{r_{max}} \right) - K_0 \left(\frac{|\vec{r}_2 - \vec{r}_3|}{r_{max}} \right) \right]^2$$

- The total probability of the inelastic interaction of two protons in the eikonal approximation:

$$p = 1 - e^{-\sum_{i,j} f_{ij}}$$

[3] G. Gustafson, Acta Phys. Polon. B40, 1981 (2009)

[4] C. Flensburg, G. Gustafson, and L. Lonnblad, Eur. Phys. J. (C) 60, 233 (2009)

p-p interaction: string fusion

The interaction of colour strings in transverse plane is carried out in the framework of local string fusion model [5] with the introduction of the lattice in the impact parameter plane. The finite rapidity length of strings is taken into account [6-8].

$$\langle \mu \rangle_k = \mu_0 \sqrt{k} \frac{S_k}{\sigma_0} \quad \langle p_t^2 \rangle_k = p_0^2 \sqrt{k} \quad \langle p_t \rangle_k = p_0 \sqrt[4]{k}$$

S_k – area, where k strings are overlapping, σ_0 single string transverse area, μ_0 and p_0 – mean multiplicity and transverse momentum from one string

[5] Braun, M.A. and Pajares, C. Eur. Phys. J. (C), 16, 349, 2000

[6] V. Vechernin and R. Kolevatov, Physics of Atomic Nuclei 70, 1797 (2007)

[7] V. Vechernin and R. Kolevatov, Physics of Atomic Nuclei 70, 1809 (2007)

[8] Vechernin, V. V. and Kolevatov, R. S., Simple cellular model of long-range multiplicity and pt correlations in high-energy nuclear collisions 2003 <http://arxiv.org/abs/hep-ph/0304295v1>

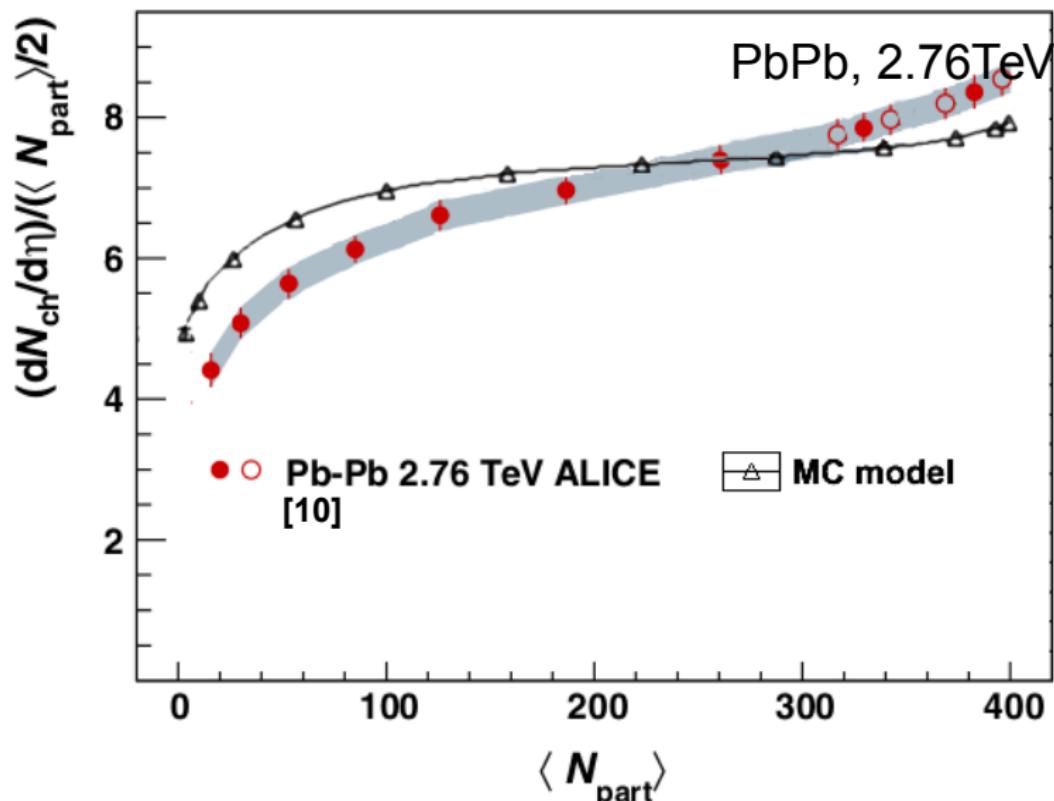
string fusion mechanism versions

	"overlaps" (local fusion)	"clusters" (global fusion)
SFM	<input type="checkbox"/> <p>$C = \{S_1, S_2, \dots\}$</p> <p>S_k – area covered k-times</p>	<input checked="" type="checkbox"/> <p>$C = \left\{ S_i^{cl}, S_2^{cl}, \dots \right\}$</p> <p>$N_1^{str} = 3$</p> <p>$S_1^{cl}$</p> $k_i^{cl} = \frac{N_i^{str} \cdot \sigma_0}{S_i^{cl}}$ <p>$N_2^{str} = 2$</p> <p>S_2^{cl}</p>
cellular analog of SFM	<input type="checkbox"/> <p>$C = \left\{ N_{ij}^{str} \right\}$</p> <p>$k_{ij} = N_{ij}^{str}$ – "occupation" numbers</p>	<input checked="" type="checkbox"/> <p>$C = \left\{ S_1^{cl}, S_2^{cl}, \dots \right\}$</p> <p>$N_1^{str} = 5$</p> <p>$S_1^{cl} = 3\sigma_0$</p> <p>$N_2^{str} = 4$</p> <p>$S_2^{cl} = 2\sigma_0$</p> <p>$k_1^{cl} = 5/3$</p> <p>$k_2^{cl} = 2$</p>

- Nucleus-Nucleus collision is a sequence of nucleons collisions
- Nucleons are distributed according to Woods-Saxon:

$$\rho(r) = \rho_0 \frac{1}{1 + \exp\left(\frac{r - R}{\alpha}\right)}$$

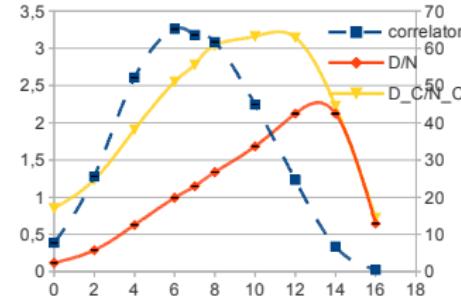
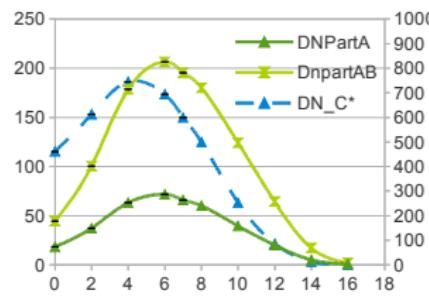
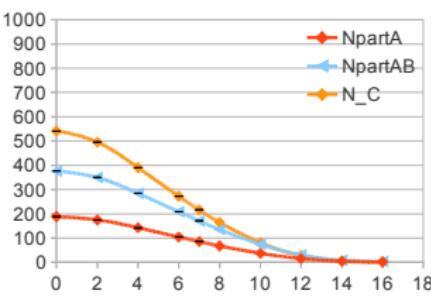
- Trajectories of nucleons are linear
- Each nucleus can collide several times with the same inelastic cross section: $\sigma_{inel}^{nn} = \text{const}$
corresponding to proton-proton inelastic cross section
- Energy loss due to particle production is not considered

AA interaction:
charged multiplicity

AA interactions

Compare with Glauber's model

Number of participant, number of binary collisions, their variations and scaled variations and correlator for $\sigma_{NN}^{inel} = 34\text{mb}$, calculated in the *model of this work*:



The same for the *Glauber's model* ($\sigma_{NN} = 34\text{mb}$):

