Exploring the string shoving model in PYTHIA8 for collective behaviors in pp collisions

Junlee Kim¹ · Eun-Joo Kim¹ · SuJeong Ji² · SangHoon Lim²

Received: 18 June 2021 / Revised: 8 July 2021 / Accepted: 26 July 2021 / Published online: 19 August 2021
© The Korean Physical Society 2021

Abstract
Two-particle correlations have been used to study property of a produced medium called quark–gluon plasma (QGP) in relativistic heavy-ion collisions, and a long-range correlation among produced particles has been considered as a strong evidence of strongly interacting QGP. Strikingly, clear long-range correlations have been also observed in small collision systems like pp and p+A at a wide range of collision energies. A model introducing a repulsive force between flux tubes called string shoving can qualitatively describe the long-range near-side correlations in high multiplicity pp collisions. This model implemented in the PYTHIA8 Monte Carlo event generator allows us to perform detailed study. In this work, a full data analysis to calculate associated yields and flow coefficients is performed with PYTHIA8 events, and the results are compared with experimental results. We find that the string shoving model overestimates long-range near-side correlations particularly in low multiplicity events. We also introduce a method to subtract such overestimated correlations in low multiplicity to have a more relevant comparison with the data.

Keywords String shoving · Collectivity · pp Collision

1 Introduction
A strong momentum correlation of produced particles from relativistic heavy-ion collisions is thought to be originated from an initial geometric anisotropy and hydrodynamic evolution Busza et al. [1]. Two-particle correlations in $\Delta \phi$ (azimuthal angle difference) and $\Delta \eta$ (pseudorapidity difference) have been measured to study azimuthal correlations among produced particles, and a clear correlation has been observed in long-range ($|\Delta \eta| > 2$) where correlations from resonance decay and jet fragmentation are expected to be small. Interestingly, such long-range correlation has been also observed in small collision systems like pp, p+Au, and p+Pb collisions over a wide range of collision energy at Relativistic Heavy Ion Collider (RHIC) and the Large Hadron Collider (LHC) Nagle and Zajc [2]. Extensive studies have been performed to quantify the collective behavior in small collision systems, and a significant elliptic flow has been measured even in pp collisions Aad et al. [3] and Khachatryan et al. [4]. Many theoretical approaches have been introduced to explain the data. One is to apply hydrodynamic evolution developed to describe the flow in heavy-ion collisions Weller and Romatschke [5], and the other is a model of initial-state correlations among gluons Dumitru et al. [6]. The models considering hydrodynamic behavior generally provide a better description of the data, but it is not yet conclusive.

Recently, another model called string shoving model was proposed to describe the long-range correlation in high multiplicity pp collisions Bierlich et al. [7]. This model introduces a repulsive force between flux tubes so that the flux tube expands both longitudinally and transversely. In a high multiplicity pp event, where many partonic interactions can occur, one string may overlap with many other strings. The repulsion among these overlapping strings right after the collision results in correlation between particles even in a large $\Delta \eta$ range. In the comparison with the experimental data shown in Bierlich et al. [8], this model implemented in PYTHIA8 event generator Sjöstrand et al. [9] shows a qualitative agreement in long-range correlation functions. For

---

1 Division of Science Education, Jeonbuk National University, Jeonju 54896, South Korea
2 Department of Physics, Pusan National University, Busan 46241, South Korea
more detailed quantitative comparisons, we perform a full analysis of Pythia8 events with string shoving and calculate quantities which have been measured in experiments. In the following sections, we describe the analysis procedure and show results using Pythia8 event generator with different configurations. The comparison with experimental results and discussion will be followed.

2 Study with Pythia8

The string shoving model is quantitatively studied by analyzing charged hadrons from millions of non-diffractive Pythia8 pp events at $\sqrt{s} = 13$ TeV. The event multiplicity is defined with charged hadrons satisfying $|\eta| < 2.4$ and $p_T > 0.4$ GeV/c, which is equivalent definition to the experimental results. To calculate two-particle correlations, we follow the formula for per-trigger yields Khachatryan et al. [10], which is defined as,

$$\frac{1}{\text{N}_{\text{trig}}} \, \frac{d^2 \text{N}_{\text{pair}}}{d\Delta \eta d\Delta \varphi} = B(0, 0) \, S(\Delta \eta, \Delta \varphi),$$

where $N_{\text{trig}}$ and $N_{\text{pair}}$ are the number of trigger particles and the number of pairs, respectively. $S(\Delta \eta, \Delta \varphi)$ and $B(\Delta \eta, \Delta \varphi)$ are per-trigger yields in same and mixed events:

$$S(\Delta \eta, \Delta \varphi) = \frac{1}{\text{N}_{\text{trig}}} \, \frac{d^2 \text{N}_{\text{same}}}{d\Delta \eta d\Delta \varphi},$$

$$B(\Delta \eta, \Delta \varphi) = \frac{1}{\text{N}_{\text{trig}}} \, \frac{d^2 \text{N}_{\text{mixed}}}{d\Delta \eta d\Delta \varphi},$$

and $B(0, 0)$ is the per-trigger yield at $(\Delta \eta, \Delta \varphi) = (0, 0)$, where the pair acceptance is maximum.

Figures 1, 2, 3 and 4 show two-particle $\Delta \eta–\Delta \varphi$ correlation functions with charged hadrons for $1 < p_T < 2$ GeV/c.

![Fig. 1](image1.png)  
**Fig. 1** Two-particle correlation functions for charged hadrons in $|\eta| < 2.4$ and $1 < p_T < 2$ GeV/c from low (left) and high (right) multiplicity Pythia8 pp events at $\sqrt{s} = 13$ TeV with the default configuration. $N_{\text{ch}}$ is calculated with charged hadrons in $|\eta| < 2.4$ and $p_T > 0.4$ GeV/c.

![Fig. 2](image2.png)  
**Fig. 2** Two-particle correlation functions for charged hadrons in $|\eta| < 2.4$ and $1 < p_T < 2$ GeV/c from low (left) and high (right) multiplicity Pythia8 pp events at $\sqrt{s} = 13$ TeV with the string shoving model. $N_{\text{ch}}$ is calculated with charged hadrons in $|\eta| < 2.4$ and $p_T > 0.4$ GeV/c.
Exploring the string shoving model in \textsc{Pythia8} for collective behaviors in pp collisions

Fig. 3 Two-particle correlation functions for charged hadrons in $|\eta| < 2.4$ and $1 < p_T < 2$ GeV/c from low (left) and high (right) multiplicity \textsc{Pythia8} pp events at $\sqrt{s} = 13$ TeV with the default configuration. Initial charged particles produced directly from partons are used for the correlations, and $N_{ch}$ is calculated with final charged hadrons in $|\eta| < 2.4$ and $p_T > 0.4$ GeV/c.

Fig. 4 Two-particle correlation functions for charged hadrons in $|\eta| < 2.4$ and $1 < p_T < 2$ GeV/c from low (left) and high (right) multiplicity \textsc{Pythia8} pp events at $\sqrt{s} = 13$ TeV with the string shoving model. Initial charged particles produced directly from partons are used for the correlations, and $N_{ch}$ is calculated with final charged hadrons in $|\eta| < 2.4$ and $p_T > 0.4$ GeV/c.

In low (left) and high (right) multiplicity \textsc{Pythia8} pp events at $\sqrt{s} = 13$ TeV with different configurations. In \textsc{Pythia8} events with the default configuration (Fig. 1), there is no long-range near-side correlation in two different multiplicity bins as expected. On the other hand, the correlation functions with the string shoving option shown in Fig. 2 clearly exhibit a long-range near-side correlation called “ridge”. It is also worth mentioning that a ridge structure is more prominent than one with final charged hadrons, and this implies that a smearing effect in decay processes decreases the correlation strength among particles.

One-dimensional $\Delta \varphi$ correlation functions are constructed by averaging the two-dimensional correlation functions at a certain $\Delta \eta$ range:

\[
\frac{1}{N_{\text{trig}}} \frac{dN_{\text{pair}}}{d\Delta \varphi} = \int d\Delta \eta \frac{1}{N_{\text{trig}}} \frac{d^2N_{\text{pair}}}{d\Delta \eta d\Delta \varphi}.
\]

For long-range correlations, the $\Delta \eta$ range is determined as $2 < |\Delta \eta| < 4$, where non-flow effects from di-jets and particle decay are expected to be not significant. The correlated part is estimated by implementing the Zero-Yield-At-Minimum (ZYAM) procedure Ajitanand et al. [11]. The
minimum yield ($C_{ZYAM}$) at the $\Delta \varphi = \Delta \varphi_{\text{min}}$ for each $\Delta \varphi$ correlation function is obtained by fitting the $\Delta \varphi$ correlation function with a Fourier series up to the third term and is subtracted from the $\Delta \varphi$ correlation function. This procedure results in the $\Delta \varphi$ correlation function at $\Delta \varphi_{\text{min}}$ to be zero.

Figure 5 shows one-dimensional $\Delta \varphi$ correlation functions at long range ($2 < |\Delta \eta| < 4$) after applying the ZYAM procedure for different $p_T$ ranges in low (upper) and high (lower) multiplicity events. PYTHIA8 events with the default configuration show no near-side structure in different $p_T$ and multiplicity ranges. PYTHIA8 events with the string shoving option, on the other hand, clearly show near-side peaks. The CMS results Khachatryan et al. [10] are compared with the correlation functions from PYTHIA8 events, and the multiplicity ranges for PYTHIA8 events are selected $\sim 15\%$ higher than the multiplicity ranges for CMS results to consider a track reconstruction efficiency. The string shoving model overestimates the near-side correlation for $1 < p_T < 2 \text{ GeV}/c$ in low multiplicity events. In the high multiplicity range, the correlation function of final charged hadrons for $1 < p_T < 2 \text{ GeV}/c$ in PYTHIA8 events with the string shoving nicely agrees with the CMS result. However, the model underestimates the near-side correlation for higher $p_T$ ranges.

3 Results

To have more quantitative comparisons, associated yields of long-range ($2 < |\Delta \eta| < 4$) near-side correlation functions are calculated by integrating one-dimensional correlation functions over the region $|\Delta \varphi| < |\Delta \varphi_{\text{min}}|$

$$y_{\text{assoc}} = \int_{|\Delta \varphi|<|\Delta \varphi_{\text{min}}|} d\Delta \varphi \left( \frac{1}{N_{\text{orig}}} \frac{dN_{\text{pair}}}{d\Delta \varphi} - C_{\text{ZYAM}} \right),$$

where $\Delta \varphi_{\text{min}}$ is the $\Delta \varphi$ value from the ZYAM method. The left panel of Fig. 6 shows the associated yield of long-range near-side correlation functions as a function of $N_{\text{ch}}$ for charged hadrons in $1 < p_T < 2 \text{ GeV}/c$, and the multiplicity $N_{\text{ch}}$ is calculated with charged hadrons in $|\eta| < 2.4$ and $p_T > 0.4 \text{ GeV}/c$. Note that associated yields with initial charged hadrons directly produced from partons are also calculated in the same $N_{\text{ch}}$ range. In case of the string shoving model (magenta lines), significant amount of associated yields are observed at the entire $N_{\text{ch}}$ range. In case of the string shoving model (magenta lines), significant amount of associated yields are observed at the entire $N_{\text{ch}}$ range, and it increases with charged hadron multiplicity. The larger associated yields of initial hadrons

![Fig. 5 One-dimensional long-range ($2 < |\Delta \eta| < 4$) $\Delta \varphi$ correlations in low and high multiplicity events for different $p_T$ bins. The CMS results (closed circle) Khachatryan et al. [10] are compared with correlation functions using initial (dashed) and final (solid) charged hadrons from PYTHIA8 events with (magenta) and without (orange) the string shoving. $\Delta \varphi_{\text{min}}$ values for correlation functions with final charged hadrons are indicated.](image-url)
than those of final hadrons may indicate that the long-range correlation built from the string shoving smears out through particle decay processes.

The CMS results Khachatryan et al. [10] are compared with the string shoving model. Note that the CMS results are as a function of the number of reconstructed charged particles ($N_{\text{ch}}$), the CMS results are shifted along the $x$-axis ($1.15 \times N_{\text{ch}}$) based on the reported track reconstruction efficiency. The associated yields of final hadrons from the string shoving model agree with the CMS result at high multiplicity events ($N_{\text{ch}} > 120$), but it overestimates associated yields at lower multiplicity events ($N_{\text{ch}} < 50$) indicating there is a long-range correlation from jets in the string shoving model. For a more detailed comparison, associated yields as a function of $p_T$ in high multiplicity events are shown in the right panel of Fig. 6. Although the associated yields for charged hadrons in $1 < p_T < 2 \text{ GeV/c}$, shown in the left panel of Fig. 6, are comparable between the string shoving model and the CMS result in high multiplicity events, they show a different $p_T$ dependence. The associated yields from the string shoving model is higher than the CMS result around $p_T = 1 \text{ GeV/c}$, and it drops more rapidly for $p_T > 2 \text{ GeV/c}$.

For further investigation on the non-zero associated yields from the string shoving model in low multiplicity events ($N_{\text{ch}} \lesssim 50$), we generated Pythia8 events with the string shoving configuration but turning off the multi-parton interaction option, so charged hadrons are mostly produced from hard scattered partons. In these events where the number of strings is expected to be small, a clear long-range near-side correlation is still observed. Figure 7 shows associated yields of long-range near-side correlations as a function of $p_T$ for different event classes. Solid lines are from normal low ($N_{\text{ch}} < 20$) and high ($N_{\text{ch}} \geq 120$) multiplicity Pythia8 events with the string shoving configuration, and the dashed line is from inclusive events without multi-parton interaction. The associated yields in normal low multiplicity events and events without multi-parton interaction are consistent. This indicates that the string shoving model introduces a long-range correlation for particles associated with jets, which is not observed in experimental results at low multiplicity events ($N_{\text{ch}} < 40$).

We perform a further analysis to subtract associated yields from jets to have a better comparison with the data. One simple assumption is that there are two sources of particles, jets and bulk, and there is no correlation between one particle from jets and the other particle from bulk. Another assumption is that the associated yields from jets can be estimated at low multiplicity events where jet contribution is dominant. In such case, long-range two-particle correlation functions after the ZYAM procedure can be expressed as

![Fig. 6](image1.png)

![Fig. 7](image2.png)
\[ \frac{1}{N_{\text{trig}}} \frac{dN_{\text{pair}}}{d\Delta \varphi} = \frac{1}{N_{\text{trig-j}} + N_{\text{trig-b}}} \left( \frac{dN_{\text{pair-j}}}{d\Delta \varphi} + \frac{dN_{\text{pair-b}}}{d\Delta \varphi} \right). \]  

where \( N_{\text{trig-j}} (N_{\text{trig-b}}) \) is the number of trigger particles from jets (bulk), and \( N_{\text{pair-j}} (N_{\text{pair-b}}) \) is the number of pairs between particles from jet (bulk). The number of pairs from jets \( (N_{\text{pair-j}}) \) in a certain high multiplicity bin can be defined as

\[ N_{\text{pair-j}} = N_{\text{pair-j}}^{\text{HM}} = \frac{N_{\text{pair-j}}^{\text{env}}}{N_{\text{pair-j}}^{\text{nLM}}}. \]

where \( N_{\text{env}} \) is the number of events in a certain multiplicity range. Subtracted associated yields are calculated as

\[ Y_{\text{assoc,sub}} = \int_{|\Delta \varphi| < |\Delta \varphi_{\text{min}}|} d\Delta \varphi \left( \frac{1}{N_{\text{trig}}} \frac{dN_{\text{HM}}}{d\Delta \varphi} \right), \]

where \( N_{\text{HM}} \) is the number of trigger particle in high multiplicity events containing trigger particles both from jets and bulk.

Figure 8 shows results from the subtraction method. Associated yields in \( 20 \leq N_{\text{ch}} < 30 \) are used for the subtraction based on the distribution of associated yields as a function of \( N_{\text{ch}} \) from the string shoving model. The slope of the increase of associated yields with \( N_{\text{ch}} \) changes in \( 20 \leq N_{\text{ch}} < 30 \). The jet contribution would be dominating at low multiplicity, so we assume that the associated yields in \( 20 \leq N_{\text{ch}} < 30 \) are maximum yields from jets. In addition, no additional multiplicity dependence is considered for yields from jets, as no strong multiplicity dependence is observed in near-side jet yields calculated using a method in Khachatryan et al. [12].

The left panel shows associated yields for charged hadrons in \( 1 < p_T < 2 \text{ GeV/c} \) as a function of \( N_{\text{ch}} \), and the dashed line is a result from the subtraction procedure. The subtraction procedure shows a larger impact in lower multiplicity, and the subtracted yields shows a similar multiplicity dependence with the CMS results. In case of associated yields as a function of \( p_T \) shown in the right panel, there is no significant change in \( p_T \) dependence for the high multiplicity bin.

In experiments, long-range two-particle correlations are also quantified with flow coefficients. A template fit method developed by the ATLAS collaboration Aad et al. [3] to subtract non-flow effects is applied to Pythia8 events to obtain flow coefficients. In the template fit method, correlation functions from low multiplicity events are used to estimate non-flow contributions:

\[ Y_{\text{template}}(\Delta \varphi) = Y_{\text{ridge}}(\Delta \varphi) + FY_{\text{nLM}}(\Delta \varphi), \]

\[ Y_{\text{ridge}}(\Delta \varphi) = G \left( 1 + \sum_{n=2}^{\infty} 2v_{n,n} \cos(n\Delta \varphi) \right), \]

where \( F \) is a scale factor for correlation functions in low multiplicity events, and \( v_{n,n} \) is a coefficient related to the \( n \)-th order flow coefficient. The coefficient \( G \) is fixed by requiring that integrals of \( Y_{\text{template}}(\Delta \varphi) \) and \( Y_{\text{HM}}(\Delta \varphi) \) are equal. Based on the assumption of factorization of flow coefficients, \( v_{n,n} \) from two-particle correlations between trigger particles of \( p_T^a \) and associated particles of \( p_T^b \) is \( v_{n,n}(p_T^a, p_T^b) = v_n(p_T^a) v_n(p_T^b) \).

More details on the template fit method can be found in Aad et al. [3].

Figure 9 shows the example of the template fits to one-dimensional correlation functions of charged hadrons in \( 0.5 < p_T < 5.0 \text{ GeV/c} \) from high multiplicity Pythia8 pp events with the default configuration (left) and the string shoving option (right). To compare with the ATLAS results, slightly different selections are used for multiplicity \((|\eta| < 2.5 \text{ and } p_T > 0.4 \text{ GeV/c})\) and long range \((2 < |\Delta \varphi| < 5)\). In case of the default configuration, the correlation function from high multiplicity events is well.

![Figure 8](image-url)  
**Fig. 8** Associated yields of long-range near-side correlation functions as a function of \( N_{\text{ch}} \) for charged hadrons in \( 1 < p_T < 2 \text{ GeV/c} \) (left) and as a function of \( p_T \) for different multiplicity bins (right). Associated yields in \( 20 \leq N_{\text{ch}} < 30 \) are used for the subtraction procedure.
Exploring the string shoving model in PyTHIA8 for collective behaviors in $pp$ collisions

The string shoving model in PyTHIA8 is described by the scaled correlation function from low multiplicity events. Hence, the obtained $v_{2,2}$ is very small. With the string shoving model, a clear near-side structure is seen in the correlation function from high multiplicity events ($Y_{HM}(\Delta \phi)$). However, there is also a near-side peak in the correlation function in low multiplicity events ($Y_{LM}(\Delta \phi)$), the scaled $Y_{LM}(\Delta \phi)$ take a part of the near-side correlation in $Y_{HM}(\Delta \phi)$. Therefore, one can expect that the obtained $v_{n,n}$ from the template fit method will be smaller than the true $v_{n,n}$ at the same multiplicity range.

Figure 10 shows extracted $v_{2,2}$ from the template fit method in PyTHIA8 events with the default configuration (orange lines) and the string shoving model (magenta lines) as a function of charged hadron multiplicity $N_{ch}$ for charged hadrons in $0.5 < p_T < 5.0 \text{ GeV}/c$ (left) and as a function of $p_T$ for high multiplicity events (right). The $v_{2,2}$ from PyTHIA8 events with the default configuration are close to zero, and this demonstrates the non-flow subtraction method as the previous study in Lim et al. [13]. The results from PyTHIA8 events with the string shoving model show non-zero $v_{2,2}$ values which can be expected from long-range near-side correlation in the correlation functions. In the comparison of the $v_{2,2}$ with ATLAS results Aad et al. [3], the $v_{2,2}$ from the string shoving model is significantly lower than ATLAS results. Based on the similar correlation functions in high multiplicity events shown in Fig. 5 and the over-subtraction of non-flow contribution shown in Fig. 9, the lower $v_{2,2}$ from the string shoving model is mainly due to the non-flow subtraction procedure with $Y_{LM}(\Delta \phi)$ showing a near-side peak.

![Fig. 9](image1.png)

**Fig. 9** Template fits to one-dimensional correlation functions in high multiplicity ($100 \leq N_{ch} < 110$) PyTHIA8 events with the default configuration (left) and the string shoving model (right). The correlation function in low multiplicity ($0 \leq N_{ch} < 20$) events are used for the non-flow subtraction procedure. The solid points indicate the correlation functions in high multiplicity, and the open points and curves represent different components of the template shifted along the y-axis by $G$ or $FY_{LM}(0)$.

![Fig. 10](image2.png)

**Fig. 10** $v_{2,2}$ obtained from the template fit method in PyTHIA8 events with the default configuration and the string shoving model, as a function of charged hadron multiplicity (left) and $p_T$ (right). Results from PyTHIA8 events are compared with ATLAS results Aad et al. [3].
4 Summary

In summary, a quantitative study of long-range two-particle correlations in PYTHIA8 events with the string shoving model has been performed. Clear long-range near-side correlations are observed in the string shoving model, whereas no such structure is seen in PYTHIA8 events with the default configuration. In the comparison with experimental data, the string shoving model qualitatively describes long-range correlations in high multiplicity events. However, one difference is that long-range near-side correlations exist even at low multiplicity region in the string shoving model. Due to this fact, the model overestimates associated yields of long-range near-side correlation functions, and flow coefficients ($v_2, v_3$) obtained from the template fit method are lower than the experimental results. A method to subtract jet correlations in associated yields is tested, and the subtracted yields of charged hadrons for $1 < p_T < 2$ GeV/$c$ shows a similar multiplicity dependence with the experimental results.

Acknowledgements  S.J. Ji and S.H. Lim acknowledge support from a 2-year Research Grant of Pusan National University.

References

1. W. Busza, K. Rajagopal, W. van der Schee, Ann. Rev. Nucl. Part. Sci. 68, 339 (2018)
2. J.L. Nagle, W.A. Zajc, Ann. Rev. Nucl. Part. Sci. 68, 211 (2018)
3. G. Aad et al., ATLAS. Phys. Rev. Lett. 116, 172301 (2016)
4. V. Khachatryan et al., CMS. Phys. Lett. B 765, 193 (2017a)
5. R.D. Weller, P. Romatschke, Phys. Lett. B 774, 351 (2017)
6. A. Dumitru, K. Dusling, F. Gelis, J. Jalilian-Marian, T. Lappi, R. Venugopalan, Phys. Lett. B 697, 21 (2011)
7. C. Bierlich, G. Gustafson, and L. Lönnblad (2016), arXiv:1612.05132
8. C. Bierlich, G. Gustafson, L. Lönnblad, Phys. Lett. B 779, 58 (2018)
9. T. Sjöstrand, S. Mrenna, P.Z. Skands, Comput. Phys. Commun. 178, 852 (2008)
10. V. Khachatryan et al., CMS. Phys. Rev. Lett. 116, 172302 (2016)
11. N.N. Ajitanand, J.M. Alexander, P. Chung, W.G. Holzmann, M. Issah, R.A. Lacey, A. Shevel, A. Taranenko, P. Danielewicz, Phys. Rev. C 72, 011902 (2005)
12. V. Khachatryan et al., CMS. Phys. Lett. B 765, 193 (2017b)
13. S.H. Lim, Q. Hu, R. Belmont, K.K. Hill, J.L. Nagle, D.V. Perepelitsa, Phys. Rev. C 100, 024908 (2019)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.