

PARTIAL CORRELATION ANALYSIS IN
ULTRA-RELATIVISTIC NUCLEAR COLLISIONS* **WOJCIECH BRONIOWSKI^{a,b}, ADAM OLSZEWSKI^a^aInstitute of Physics, Jan Kochanowski University, 25-406 Kielce, Poland^bThe H. Niewodniczański Institute of Nuclear Physics Polish Academy of Sciences
Radzikowskiego 152, 31-342 Kraków, Poland

(Received January 8, 2019)

We show that the method of partial covariance is a very efficient way to introduce constraints (such as the centrality selection) in data analysis in ultra-relativistic nuclear collisions. The technique eliminates spurious event-by-event fluctuations of physical quantities due to fluctuations of control variables. Moreover, in the commonly used superposition approach to particle production, the method can be used to impose constraints on the initial sources rather than on the finally produced particles, thus separating out the trivial fluctuations from statistical hadronization or emission from sources and focusing strictly on the initial-state physics. As illustration, we use simulated data from hydrodynamics started on the wounded-quark event-by-event initial conditions, followed with statistical hadronization, to show the practicality of the approach in analyzing the forward-backward multiplicity fluctuations. We mention generalizations to the case with several constraints and other observables, such as the transverse momentum or eccentricity correlations.

DOI:10.5506/APhysPolBSupp.12.405

This paper is based on [1] where more details can be found. The technique of *partial covariance* is widely used in other areas of science in situations where one can distinguish the *physical* variables and the *control* (spurious, nuisance) variables in multivariate statistical samples (see, e.g., [2, 3]). Such a separation occurs in typical setups in ultra-relativistic nuclear collisions, where response of certain detectors is used to determine *centrality*, the quantile from a given measured quantity, which plays the role of a control variable, whereas other quantities correspond to physical variables.

* Presented at the XIII Workshop on Particle Correlations and Femtoscopy, Kraków, Poland, May 22–26, 2018.

** Supported by the National Science Centre, Poland (NCN) grant 2015/19/B/ST2/00937.

The problem of eliminating fluctuations of centrality, spuriously correlating to physical quantities, has a long history with numerous methods, *e.g.*, [4–8], developed precisely for this purpose. We argue that the advocated partial covariance method is particularly simple, bringing up a general understanding of centrality as a control variable, whose interpretation depends on the experimental arrangement.

With partial correlations, the constraints on the control variables emerge from the relationship of the partial covariance, defined as

$$\text{cov}(X_i, X_j \bullet Z) \equiv \text{cov}(X_i, X_j) - \text{cov}(X_i, Z_a) [\text{cov}(Z)]_{ab}^{-1} \text{cov}(Z_b, X_j), \quad (1)$$

to the *conditional covariance* $\text{cov}(X_i, X_j|Z)$, where i, j label the physical variables X and a, b the control variables Z . The subtraction in Eq. (1) projects out the components spuriously correlated via the control variables, one is thus left with physical correlations only. The conditional covariance is defined by first fixing the values of Z to obtain the covariance of X_i and X_j , and then averaging over the control variable(s) Z . Note that the prescription of using very narrow centrality bins and then averaging over them, advocated in [9–11], precisely conforms to this recipe. It has been shown [10, 12, 13] that $\text{cov}(X_i, X_j \bullet Z) = \text{cov}(X_i, X_j|Z)$ iff $\langle X|Z \rangle = \alpha + BZ$, *i.e.*, iff the expectation value of X at Z fixed is an affine function of Z , with α a constant vector and B a constant matrix. This feature is well-satisfied when the centrality classes are sufficiently narrow, as is typically the case, hence the equality between conditional and partial correlations holds to a very good approximation.

To understand in simple terms this relation, let us consider the case with a single physical variable X and a single constraint Z . In panel (a) of Fig. 1, we show a data sample (represented with an oval) which is cut into narrow stripes with fixed values of Z . By assumption, the expectation value of X

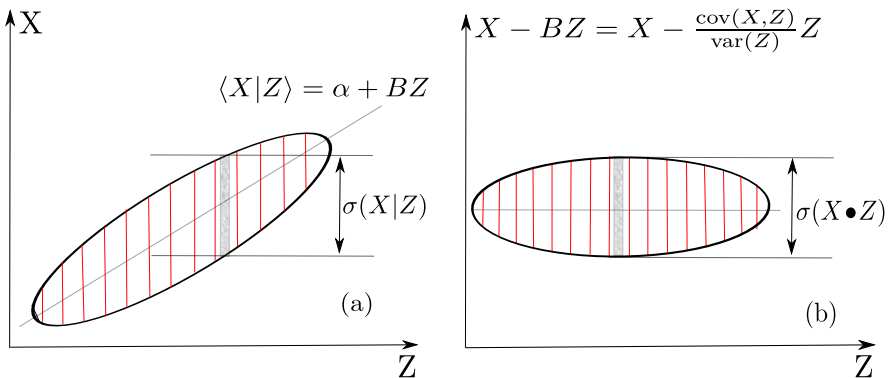


Fig. 1. Graphical illustration of the proof of equality of the partial and conditional variance.

in each stripe aligns along a straight line. The typical width at a fixed Y is indicated with $\sigma(X|Y)$. Note that it is much narrower than the whole span of the sample in X , which is also due to the extension of the sample in Y , which is correlated with X . Now, averaging the affine relation over Z yields $B = \text{cov}(X, Z)/\text{var}(Z)$. We may thus lower each stripe from panel (a) of Fig. 1 by BZ , with the result shown in panel (b), where the data oval is “straightened”. The variance of this representation of the sample is $\text{var}(X - BZ) = \text{var}(X) - \text{cov}(X, Z)^2/\text{var}(Z)$, which is nothing but the partial variance of X , which completes the proof in this simple case. The derivation from right to left proceeds analogously.

The construction of Fig. 1 indicates the two equivalent ways to evaluate the covariance of the physical variables: slicing into narrow bins and computing the conditional covariance, or computing the partial covariance on the whole sample. We note that the partial covariance method is simpler, as it avoids possible problems encountered for low multiplicity samples, where narrow bins may be poorly populated, or even empty.

We also wish to remark here that the meaning of centrality in nuclear collisions is by no means universal, but is strictly related to the response of the chosen “control” detectors, be it multiplicity in the central or peripheral bins, multiplicity of a sub-event sample, multiplicity of spectators in peripheral detectors, or the transverse energy from calorimeters. Definition (1) allows one to combine several control variables simultaneously in a straightforward way.

A novel development presented in [1] was the combination of the technique of partial covariance with the superposition approach [14] to particle production in ultra-relativistic nuclear collisions, based on the assumption that particles are emitted from independent sources, such as wounded nucleons [15] or quarks [16, 17]. The partial covariance method makes it possible to impose constraints *at the level of these initial sources*, which we find non-trivial. We have tested our formalism by analyzing the forward–backward (FB) multiplicity correlations, defined via the correlation function

$$C(X_1, X_2) = \frac{\text{cov}(X_1, X_2)}{\langle X_1 \rangle \langle X_2 \rangle}. \quad (2)$$

The FB correlation of the numbers of sources, with the constraint imposed at the number of sources in the mid-rapidity bin C , is given by the formula [1]

$$\begin{aligned} C(S_F, S_B \bullet S_C) &= C(N_F, N_B) - \frac{C(N_F, N_C)C(N_B, N_C)}{\text{var}(N_C) - \omega(m)\langle N_C \rangle} \\ &\simeq C(S_F, S_B|S_C), \end{aligned} \quad (3)$$

where $\omega(m)$ is the scaled variance of the overlaid distribution (m is a random number of particles emitted from a single source). In the case of the Poisson distribution, $\omega(m) = 1$. Note that except for $\omega(m)$, the quantities on the right-hand side of Eq. (3) involve measurable particle multiplicities only.

To test formula (3) in a practical application, we have taken a sample of simulated events which uses the wounded quark event-by-event initial conditions [18], the 3+1D viscous hydrodynamics [19], and THERMINATOR for statistical hadronization [20, 21]. The longitudinal profile in spatial rapidity is assumed to have the phenomenologically successful form of “triangles” [22–24], where each source has the emission profile peaked in the direction of its motion. For our test, on the one hand, we have computed the right-hand side with the generated hadrons, on the other hand, we have evaluated the left-hand side directly from the known correlation of sources. In the adopted Bzdak–Teaney (BT) model [25], it has the form of

$$C(S_F, S_B \bullet S_C) = \frac{\text{var}(Q_A - Q_B)}{\langle Q_A + Q_B \rangle^2} \frac{y_1 y_2}{y_B^2}, \quad (4)$$

where $Q_{A,B}$ denotes the number of wounded quarks in the colliding nucleus A or B and y_B stands for the rapidity of the beam.

As demonstrated in Fig. 2, we are able to recover the FB partial correlations in the initial condition to a very reasonable accuracy. The left panel shows the partial covariance from the BT model, whereas the right panel presents its estimate extracted with multiplicities of positively charged pions from the simulated events (the use of same charge pions reduces the correlations from resonance decays). We note a remarkably similar shape

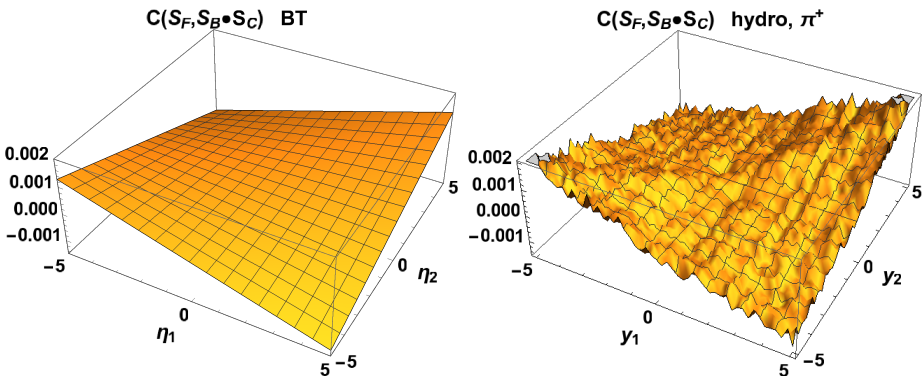


Fig. 2. Forward–backward partial correlations for the source multiplicities with the constraint fixing the number of sources in the mid-rapidity bin, evaluated directly in the BT model [25] (a), and extracted from the simulated data for positively charged pions (b), plotted against the rapidities of the pions, $y_{1,2}$.

and magnitude of the two results. Some discrepancy may be attributed to breaking of the independence of sources, as well as from the mixing of neighboring rapidity bins, here coming from cascades of resonance decays.

Needless to say, it would be very interesting to apply the proposed method to real data and to infer information on correlation between multiplicities of the initial sources. Generalizations to multiple simultaneous constraints (such as from response of various detectors controlling centrality) have been discussed in [1]. The method can also be straightforwardly extended to FB correlations of other observables, such as the transverse momentum [26] or the harmonic flow coefficients.

We thank Piotr Bożek for providing a sample from hydrodynamic simulations in the wounded quark model.

REFERENCES

- [1] A. Olszewski, W. Broniowski, *Phys. Rev. C* **96**, 054903 (2017).
- [2] H. Cramer, *Mathematical Methods of Statistics*, Princeton Mathematical Series, no. 9. Princeton University Press, Princeton 1946.
- [3] W. Krzanowski, *Principles of Multivariate Analysis*, Oxford Statistical Science Series. Oxford University Press, Oxford 2000.
- [4] M. Gaździcki, S. Mrówczyński, *Z. Phys. C* **54**, 127 (1992).
- [5] M.I. Gorenstein, M. Gaździcki, *Phys. Rev. C* **84**, 014904 (2011).
- [6] R.S. Bhalerao, J.-Y. Ollitrault, S. Pal, D. Teaney, *Phys. Rev. Lett.* **114**, 152301 (2015).
- [7] W. Broniowski, A. Olszewski, *Phys. Rev. C* **95**, 064910 (2017).
- [8] R. Rogly, G. Giacalone, J.-Y. Ollitrault, *Phys. Rev. C* **99**, 034902 (2019).
- [9] B. Abelev *et al.*, *Phys. Rev. Lett.* **103**, 172301 (2009).
- [10] A. Bzdak, *Phys. Rev. C* **85**, 051901 (2012).
- [11] S. De *et al.*, *Phys. Rev. C* **88**, 044903 (2013).
- [12] A.J. Lawrance, *Am. Stat.* **30**, 146 (1976).
- [13] K. Baba, R. Shibata, M. Sibuya, *Aust. NZ J. Stat.* **46**, 657 (2004).
- [14] A. Olszewski, W. Broniowski, *Phys. Rev. C* **88**, 044913 (2013).
- [15] A. Białas, M. Błeszyński, W. Czyż, *Nucl. Phys. B* **111**, 461 (1976).
- [16] A. Białas, W. Czyż, W. Furmański, *Acta Phys. Pol. B* **8**, 585 (1977).
- [17] V.V. Anisovich, Yu.M. Shabelski, V.M. Shekhter, *Nucl. Phys. B* **133**, 477 (1978).
- [18] P. Bożek, W. Broniowski, M. Rybczyński, *Phys. Rev. C* **94**, 014902 (2016).
- [19] P. Bożek, *Phys. Rev. C* **81**, 034909 (2010).

- [20] A. Kisiel, T. Tałuć, W. Broniowski, W. Florkowski, *Comput. Phys. Commun.* **174**, 669 (2006).
- [21] M. Chojnacki, A. Kisiel, W. Florkowski, W. Broniowski, *Comput. Phys. Commun.* **183**, 746 (2012).
- [22] A. Białas, W. Czyż, *Acta Phys. Pol. B* **36**, 905 (2005).
- [23] A. Adil, M. Gyulassy, T. Hirano, *Phys. Rev. D* **73**, 074006 (2006).
- [24] P. Bożek, I. Wyskiel, *Phys. Rev. C* **81**, 054902 (2010).
- [25] A. Bzdak, D. Teaney, *Phys. Rev. C* **87**, 024906 (2013).
- [26] A. Olszewski, Ph.D. Thesis, to be published.