PROCEEDINGS OF SCIENCE

Towards a more robust intermittency analysis technique in heavy ion collisions: achievements and challenges.

Nikolaos Davis[a,*]

[a] *Institute of Nuclear Physics PAN ,*
ul. Radzikowskiego 152, Kraków, Poland

E-mail: nikolaos.davis@ifj.edu.pl

The search for experimental signatures of the critical point (CP) of strongly interacting matter is one of the main objectives of numerous heavy ion collision experiments today, notably at the CERN SPS (NA49, NA61/SHINE) and RHIC (STAR Collaboration). In particular, critical fluctuations of the order parameter of the chiral phase transition, the net baryon density n_B, have been studied as a promising observable connected to the approach of a system of colliding nuclei to the CP. Experimentally, critical fluctuations can be probed through proton multiplicity intermittency analysis of the second scaled factorial moments (SSFMs) in transverse momentum space. Proton intermittency analyses on NA49 [1] as well as NA61/SHINE SPS data [2] have been published, providing some evidence of critical fluctuations. However, a more careful study of intermittency as applied in experimental sets with their limitations in system size, available event statistics and detector effects indicates that factorial moments exhibit much greater uncertainties than previously estimated, rendering published results inconclusive [3]. We attempt to shed light on factorial moment behaviour via the Critical Monte Carlo (CMC) simulation in the proton channel; by implementing the model in realistic conditions of heavy-ion reactions, we perform a systematic study of factorial moments for various cases of power-law exponents and possible critical component strength in the data. Using novel statistical techniques to overcome the limitations of traditional intermittency analysis, notably to handle intricate data point correlations, we show that the availability and proper use of a large number of events is a decisive factor in performing a robust intermittency analysis.

Proceedings of the Corfu Summer Institute 2024 "School and Workshops on Elementary Particle Physics and Gravity" (CORFU2024) 12 - 26 May, and 25 August - 27 September, 2024
Corfu, Greece

*Speaker

PoS(CORFU2024)097

1. Introduction

A key question in the study of QCD is to determine the structure of the QCD Phase Diagram as a function of temperature T and baryochemical potential μ_B (equivalently: nuclear density, n_B); that is, to determine the various states of strongly interacting matter, the Equations of State (EoS) that govern them, and the location, nature, and type of the phase transition boundaries that separate them. Fig. 1 shows a hypothetical sketch of the QCD Phase Diagram in $(T - \mu_B)$, where the phases of hadronic (nuclear) matter and quark-gluon plasma are separated by different types of phase transitions, namely a smooth cross-over at high T and low μ_B, and a 1$^{\text{st}}$ order transition at low T and high μ_B, ending on a critical point (CP), in the vicinity of which a 2$^{\text{nd}}$ order transition is expected to occur. Both the cross-over and the 1$^{\text{st}}$ order transition line are evidenced by Lattice QCD and effective models, respectively, and therefore there is ample evidence to support the existence of a critical point (CP) as an end point of the 1$^{\text{st}}$ order transition line.

A characteristic feature of a second order phase transition (expected to occur at the CP) is the divergence of the correlation length, leading to a scale-invariant system effectively described by a universality class. Of particular interest are local fluctuations of the order parameter of the QCD chiral phase transition, the chiral condensate $\sigma(\mathbf{x}) = \langle \bar{q}(\mathbf{x}) q(\mathbf{x}) \rangle$. At finite baryochemical potential, the critical fluctuations of the chiral condensate are transferred to the net-baryon density [4]. For a critical system, we expect fluctuations of the order parameter to be self-similar [5], obeying power-laws with critical exponents determined by the 3D Ising universality class [6–8].

Candidates for the role of order parameter include the chiral σ-condensate as reconstructed through $\pi^+\pi^-$ (dipion) pairs, as well as local fluctuations of the net baryon density n_B, and its proxy, the proton density in transverse momentum space. The choice of dipions offers the advantage of being the main decay channel of the σ-condensate, leading to an abundance of pions, and thus high-multiplicity events to analyse. However, this comes at the price of sifting through a large combinatorial background of unrelated $\pi^+\pi^-$ pairs, which restricts the analysis to a window of invariant mass close to the production threshold of dipions[1]. For that reason, intermittency analysis efforts were shifted to the study of net baryon density fluctuations, of which the net proton density is a proxy.

Such fluctuations correspond to a power-law scaling of the proton density-density correlation function, which can be detected in transverse momentum space within the framework of an intermittency analysis [8, 10–12] of proton scaled factorial moments (SFMs). A detailed analysis can be found in Ref. [1], where we study various heavy nuclei collision datasets recorded in the NA49 experiment at maximum energy (158A GeV/c, $\sqrt{s_{NN}} \approx 17$ GeV) of the SPS (CERN).

2. Methodology

2.1 The method of intermittency analysis

As mentioned in Section 1, critical fluctuations of the chiral phase transition order parameter follow a power-law form at the vicinity of the CP; specifically, for the idealized case of an infinite

[1]A detailed dipion intermittency analysis of NA49 experimental data can be found in [9].

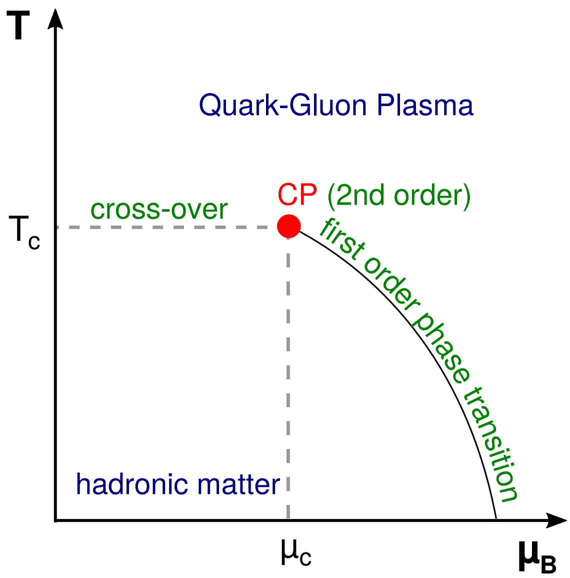


Figure 1: Hypothetical sketch of the phase diagram of strongly interacting matter with critical point, drawn as a function of baryochemical potential μ_B and temperature T. A crossover transition is predicted at low μ_B and high T; whereas, a 1st order transition is predicted at low T and high μ_B. A critical point (CP) is therefore hypothesized to exist as an end point of the 1st order transition line.

size system belonging to the 3D-Ising universality class, we obtain the following forms of density-density correlations in momentum space, for the σ-condensate [6] and the net baryon density n_B [8], respectively:

$$\langle n_\sigma(\mathbf{k})\, n_\sigma(\mathbf{k}')\rangle \sim |\mathbf{k}-\mathbf{k}'|^{-4/3} \tag{1a}$$

$$\langle n_B(\mathbf{k})\, n_B(\mathbf{k}')\rangle \sim |\mathbf{k}-\mathbf{k}'|^{-5/3} \tag{1b}$$

where $\mathbf{k}-\mathbf{k}'$ is the momentum transfer.

In order to probe a set of particle momenta for presence of power-law density-density correlations, we use the method of intermittency analysis of the Second Scaled Factorial Moments (SSFM) $F_2(M)$, pioneered by Białas and others [8, 10–12] as a method to detect non-trivial dynamical fluctuations in high energy nuclear collisions. The method consists of partitioning an analysis window in transverse momentum space into a number of equal size bins (Fig.2 *left*), then examining how $F_2(M)$ of particle transverse momenta scale with the number M^2 of 2D bins:

$$F_2(M) \equiv \left\langle \frac{1}{M^2}\sum_{i=1}^{M^2} n_i(n_i-1)\right\rangle \Big/ \left\langle \frac{1}{M^2}\sum_{i=1}^{M^2} n_i\right\rangle^2 \tag{2}$$

where n_i is the number of particles in the i-th bin, and $\langle\ldots\rangle$ denotes average over events.

For a pure critical system, $F_2(M)$ is predicted to follow a power-law [6, 8]:

$$F_2(M) \sim M^{2\cdot\phi_{2,cr}} \quad , \quad \phi_{2,cr}^{(\sigma)} = 2/3 \quad , \quad \phi_{2,cr}^{(p)} = 5/6 \tag{3}$$

where the exponent ϕ_2 is called the intermittency index.

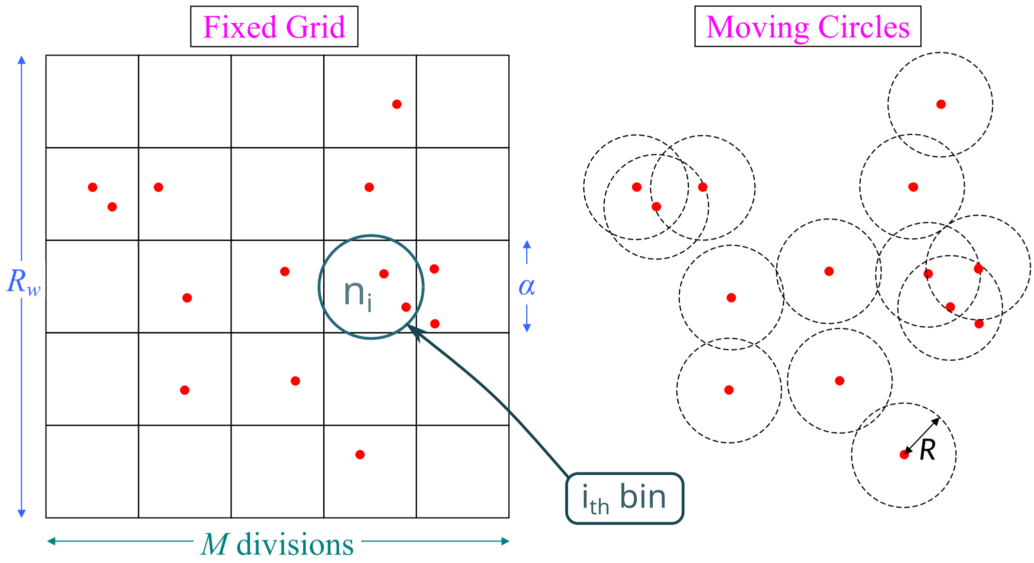


Figure 2: *Left:* Counting particle pairs in a transverse momentum space partitioning of $M \times M$ equal size bins. *Right:* By centering circles around each point in the set, it is possible to account for all pairs of points within a given distance R; demanding equal area for bins and circles allows for a correspondence of scales: $\pi R^2 = \alpha^2$. Image adapted from [13]

For a noisy system, mixed event moments must be subtracted from the data moments in order to recover the critical component [1]. This is a non-trivial operation, starting with notionally partitioning all pairs in eq.(2) into critical/background pairs, plus a cross-term:

$$\langle n(n-1)\rangle = \underbrace{\langle n_c(n_c-1)\rangle}_{\text{critical}} + \underbrace{\langle n_b(n_b-1)\rangle}_{\text{background}} + \underbrace{2\langle n_b n_c\rangle}_{\text{cross term}} \tag{4}$$

Rearranging (4), and normalizing to the mean particle multiplicity, we obtain:

$$\underbrace{\Delta F_2(M)}_{\text{correlator}} = \underbrace{F_2^{(d)}(M)}_{\text{data}} - \lambda(M)^2 \cdot \underbrace{F_2^{(b)}(M)}_{\text{background}} - 2 \cdot \underbrace{\lambda(M)}_{\text{ratio } \frac{\langle n\rangle_b}{\langle n\rangle_d}} \cdot (1-\lambda(M))\, f_{bc} \tag{5}$$

where $\lambda(M) \equiv \langle n\rangle_b / \langle n\rangle_d$ is defined as the ratio of background to total (data) multiplicity in bins of size M. λ is in general a function of bin size, but in practice it converges to a constant value at the limit of $M \to \infty$, unless the 1-particle distribution is singular. The cross-term factor, f_{bc}, cannot in general be factored out into background and critical contributions, due to correlations between the two sets. However, Critical Monte Carlo [8] simulations show that the cross-term can be safely neglected [1] in two limit cases:

1. When $\lambda(M) \sim 0$, i.e. for an almost pure system, when background can altogether be ignored;
2. When $\lambda(M) \lesssim 1$, i.e. when the background is dominant, and background moments can be approximated by mixed event moments $F_2^{(b)}(M) \sim F_2^{\text{mix}}(M)$.

The latter (dominant background) has proved to be the case in virtually all experimental systems we have studied so far. We can thus simplify eq.(5), and define an effective correlator $\Delta F_2(M)$ as simply the difference of data and mixed event moments:

$$\Delta F_2(M) = F_2^{(d)}(M) - F_2^{(m)}(M) \tag{6}$$

Intermittent behavior, if present, will then be revealed in $\Delta F_2(M)$,

$$\Delta F_2(M) \sim \left(M^2\right)^{\varphi_2},\; M \gg 1 \tag{7}$$

and we obtain the predictions of eq.(3) for the intermittency index ϕ_2, for the σ-condensate and proton density, respectively.

The usual methodology of computing $F_2(M)$ on a grid is computationally intensive, and for the case of proton intermittency where proton multiplicity per event is low, introduces artifacts. Due to the arbitrary positioning of grid lines, pairs within the same scale may be split across borders, skewing pair statistics. In the past, this problem was corrected for by introducing a lattice average over slightly displaced grids [1]; however, a computationally faster alternative to the lattice average has been developed [13] using the correlation integral $C(R)$ [14], defined as:

$$C(R) = \frac{2}{\langle N_{mul}\,(N_{mul}-1)\rangle_{ev}} \left\langle \sum_{\substack{i,j\\ i<j}} \Theta\left(|x_i - x_j| \leq R\right) \right\rangle_{ev} \tag{8}$$

where R is a given length scale, N_{mul} is the event multiplicity, and Θ is the step function, counting the number of pairs of particles in an event within a distance R from each other. $C(R)$ can be calculated by (notionally) placing circles of radius R around each of the points in the set (Fig. 2 *right*), and counting the points falling within them. We can then match the circle radius R to the bin side α, and thus the number of divisions M per dimension, by demanding equal area for bins and circles: $\pi R_M^2 = \alpha^2$. Thus, we arrive at a correspondence between $F_2(M)$ and $C(R_M)$,

$$F_2(M) = \frac{\langle N_{mul}\,(N_{mul}-1)\rangle_{ev}}{\langle N_{mul}\rangle_{ev}^2} M^2\, C(R_M) \tag{9}$$

that we can use to compute $F_2(M)$ efficiently, at the same time eliminating grid artifacts. We have adopted the correlation integral technique throughout our latest intermittency analyses, both for experimental and for Monte Carlo simulated data.

2.2 Handling of statistical and systematic uncertainties

SSFMs statistical errors are estimated via the bootstrap method [15, 16], which is a well-established statistical technique for obtaining unbiased error estimates of statistical quantities. In applying the bootstrap method to intermittency analysis, the original set S of events is first randomly sampled, with replacement, i.e., a number of events equal to that of the original set are selected uniformly at random; thus, a new bootstrap set S_B is created, in which some events are omitted and others duplicated. By repeating this process, a large number of bootstrap samples S_{B_i}, $i = 1, \ldots, N_B$, $N_B \gtrsim 1000$, are created from the original set. Subsequently the quantity of interest, in particular the moments $\Delta F_2(M)$, are calculated for each bootstrap sample in the same manner as for the original; the resulting values can be used to obtain the bootstrap statistical distribution of $\Delta F_2(M)$, as well as its standard error, confidence intervals, or any other measure of variance desirable.

Bootstrap estimation of uncertainties has certain advantages over error propagation: it is straightforward to calculate, only requiring calculation of the original statistics (the SSFMs), in contrast to error propagation, which requires calculating higher moments [15]. It is relatively cheap computationally, as only the weights of each original event need to be calculated and stored in advance; the SSFMs of bootstrap samples can then be computed in one pass, along with those of the original. Finally, it allows us to naturally and effortlessly calculate the correlation matrix of $\Delta F_2(M)$ between different bins M.

It must, however, be emphasized that the bootstrap cannot help us estimate or correct for the systematic uncertainties that may be present in the original sample. As verified by Monte Carlo simulations, as well as theoretical analysis, bootstrap estimates of the magnitude of variance and covariance of SSFMs can be trusted, but the centroids (average, median) estimated by bootstrap will certainly be biased towards the original sample. This is especially important to bear in mind when attempting to fit SSFMs with a power-law model, as in eq.(7): $\Delta F_2(M)$ values for different M are not independent, as the same events are used to calculate all $\Delta F_2(M)$, and this invalidates a simple least-squares fit. We will address this issue, as well as methods to deal with it, in Section 5.

3. Results

3.1 NA49 proton intermittency analysis

As mentioned in Section.1, pion intermittency analysis has the drawback of a large combinatorial background of non-critical $\pi^+\pi^-$ pairs that need to be subtracted in order to reveal the critical dipion contribution originating from σ meson decays. For that reason, analysis efforts have been shifted to proton intermittency analysis, which probes the density-density correlations of the proton density, a proxy to the net-baryon density [4]. There are solid theoretical predictions for the expected ϕ_2 value of critical proton transverse momenta SSFMs [8]. This mode of analysis has the advantage of directly probing the proton density, and thus not requiring the subtraction of a combinatorial background. On the other hand, it has its own shortcomings: proton per event multiplicity is low in medium sized nuclei collisions (typically, of the order of $\sim 2-5$ protons per event in systems such as C+C and Si+Si, and ~ 10 in Pb+Pb). Therefore, an exceptionally large number of events (events statistics), and good proton identification are required in order for a proton intermittency analysis to be conclusive. At minimum of the order of $\sim 100K$ events, and ideally more than $\sim 1M$, are needed in order to confidently establish a trend in $\Delta F_2(M)$; proton purity (the percentage of actual protons in the candidate protons selected from the data) should ideally be in excess of 90%.

A proton intermittency analysis was performed on a number of NA49 collision data sets of different sizes (C+C, Si+Si, Pb+Pb), at the maximum collision energy (158A GeV/c, corresponding to $\sqrt{s_{NN}} \approx 17$ GeV) of the Super Proton Synchrotron (SPS), CERN [1]. For the purposes of the analysis, the most central (12% C+C, 12% Si+Si, 10% Pb+Pb) collisions were selected, as determined by the energy deposited in the Projectile Spectator Detector (PSD) downstream

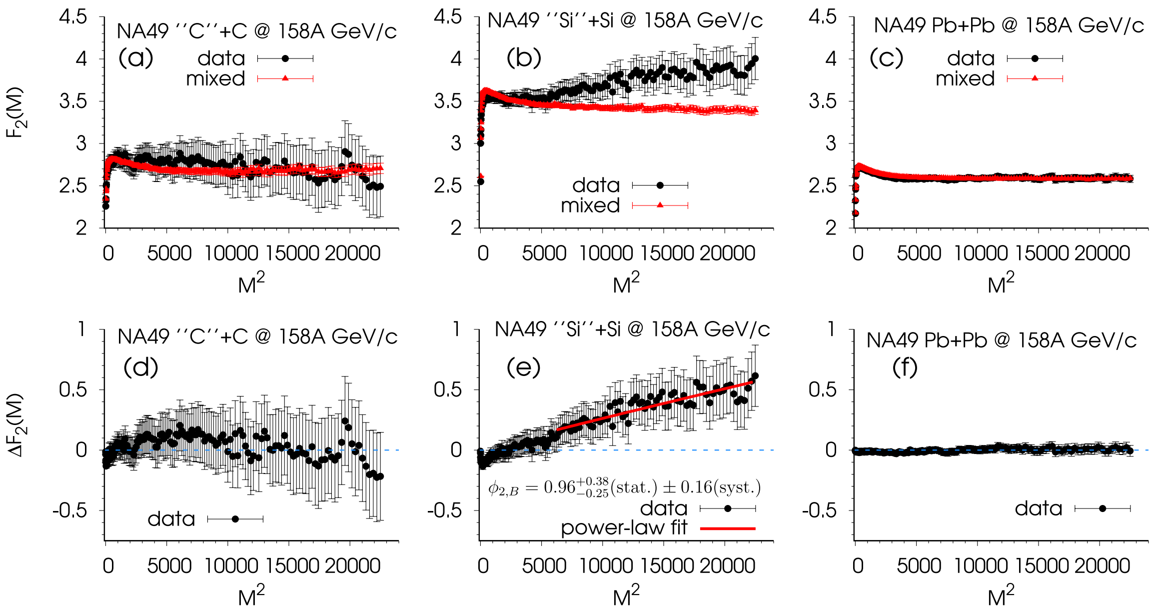


Figure 3: Proton SSFMs $F_2(M)$ *(top row)* and the correlator $\Delta F_2(M)$ *(bottom row)* for NA49 (a,d) C+C , (b,e) Si+Si , and (c,f) Pb+Pb most central (12%, 12%, 10%) collisions at 158A GeV/c ($\sqrt{s_{NN}} \approx 17$ GeV) [1]. Error bars are calculated through the bootstrap method [15].

calorimeter [17]. The event statistics amounted to 148K events for C+C, 166K events for Si+Si, and 330K events for Pb+Pb. The standard event and track selection cuts of the NA49 experiment were applied. Proton identification used the measurements of particle energy loss dE/dx in the gas of the time projection chambers; tracks were accepted as candidate protons when the estimated probability of being a proton exceeded 80% for the C+C and Si+Si systems and 90% for Pb+Pb collisions. Finally, a window of analysis was selected in transverse momentum space ($-1.5 \leq p_{x,y} \leq 1.5$ GeV/c), and candidate protons in the mid-rapidity region ($|y_{CM}| \leq 0.75$) were projected in transverse space, where their SSFMs were calculated.

Fig.3 shows the SSFMs $F_2(M)$ and the correlator $\Delta F_2(M)$ for the analyzed NA49 systems. No intermittency effect is observed for the C+C (a,d) and Pb+Pb (c,f) systems; original data and mixed event moments overlap, and the correlator fluctuates around zero. In contrast, a significant intermittency effect is observed in the Si+Si system (b,e), as evidenced by the scaling of the corresponding $\Delta F_2(M)$. The fitted power-law value for the intermittency index ϕ_2 is compatible with the theoretical prediction, eq.(3), however the statistical uncertainties are large.

3.2 NA61/SHINE proton intermittency analysis

Motivated by the positive, if ambiguous, NA49 proton intermittency Si+Si result, the search for the critical point through intermittency analysis has continued within the framework of the NA61/SHINE experiment [18], a fixed target, high-energy collision experiment at the SPS, CERN, and the direct continuation of NA49.

The NA49 intermittency result suggests an experimental intermittency scan of medium-sized nuclei as the best candidates for the detection of the critical point. Preliminary analysis of a number of medium-sized SHINE systems (Be+Be, Ar+Sc at 150A GeV/c, corresponding to $\sqrt{s_{NN}}$ =

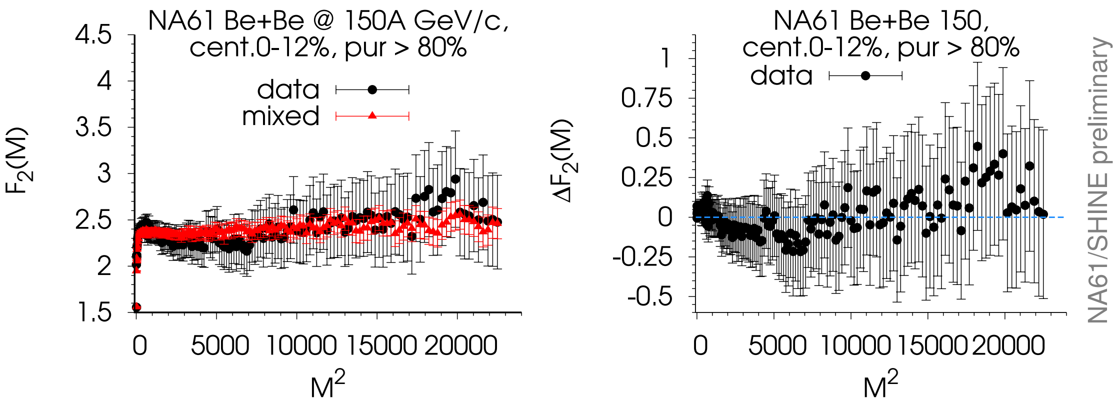


Figure 4: Proton SSFMs $F_2(M)$ *(left)* and the correlator $\Delta F_2(M)$ *(right)* for NA61/SHINE Be+Be 12% most central collisions at 150A GeV/c ($\sqrt{s_{NN}} \approx 16.8$ GeV) [19]. Error bars are calculated through the bootstrap method [15].

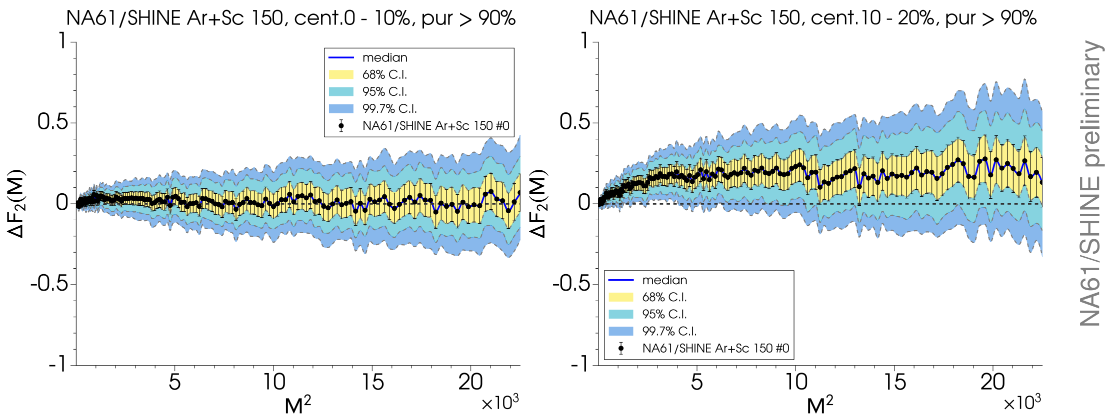


Figure 5: Proton SSFMs correlator $\Delta F_2(M)$ for NA61/SHINE Ar+Sc 0-10% *(left)*, and 10-20% *(right)* most central collisions at 150A GeV/c ($\sqrt{s_{NN}} \approx 16.8$ GeV) [2]. Error bars are calculated through the bootstrap method [15]. Colored bands correspond to 1-(*yellow*), 2-(*light blue*), and 3-σ (*dark blue*) confidence intervals, respectively.

16.8 GeV) close in nuclear size to the NA49 Si+Si system was performed, to which end great effort was exerted in applying proper experimental cuts, the pre-selection of events and proton identification and selection.

Fig.4 shows the SSFMs $F_2(M)$ and the correlator $\Delta F_2(M)$ for the analyzed Be+Be NA61/SHINE system [19]. No intermittency effect is observed, as $F_2(M)$ of data and mixed events overlap, and therefore the correlator $\Delta F_2(M)$ fluctuates around zero. It should be noted, however, that the average proton multiplicity per event in the Be+Be system was ~ 1.5 in the mid-rapidity range, excluding events with a zero proton multiplicity; that is far too low for proton pair correlations to be firmly established, making it unlikely for an intermittency analysis to be able to detect a weak critical component, even if present, given the event statistics available ($\sim$ 160K events).

Following Be+Be analysis, focus was shifted to studying the Ar+Sc system at 150A GeV/c, the closest in system size and collision energy to the NA49 Si+Si system. In this case, a full scan in collision centrality was performed, in the 0-20% most central range, in 10% intervals; the decision was due to experimental evidence, as well as theoretical understanding [3], that changes in collision peripherality influence the freeze-out conditions (μ_B, T) in a mild manner.

The first indication of intermittency in mid-central Ar+Sc collisions at 150A GeV/c was presented at the CPOD2018 international conference [20]. In 2019, an extended event statistics set, approved by the NA61/SHINE Collaboration, was subjected to careful analysis. Event statistics were of the order of $\sim$ 400K events per 10% centrality interval. Fig.5 shows the results of the SHINE Ar+Sc intermittency centrality scan, in the form of the correlator $\Delta F_2(M)$ of proton SSFMs, for each of the 10%-wide collision centrality ranges. Centrality dependence is evident in the scaling of factorial moments, with the 0-10% (most central) collisions showing no evidence of intermittency, whereas the more peripheral collisions (10-20%) exhibit a mild intermittent effect. Fig.5 also illustrates the magnitude and form of $\Delta F_2(M)$ statistical uncertainties via confidence intervals (colored bands) corresponding roughly to 1,2 and 3-σ variation around the experimental values (black points), as calculated via the statistical bootstrap.

It must be emphasized that such plots do not provide the full picture as to the $\Delta F_2(M)$ variation and uncertainties, due to the presence of M-bin correlations: the errors of points with different M values, especially neighboring ones, are correlated, since the same set of events was used to calculate all the $F_2(M)$. This is the reason why confidence intervals for the intermittency index ϕ_2 cannot be properly obtained through simple power-law fitting, not even by fitting different bootstrap samples independently. The full $\Delta F_2(M)$ covariance matrix has to be estimated and taken into account, which is subject to potential biases. The solution, as we will detail in the following sections, is to avoid fitting and use model-weighting over a broad collection of parametrized models.

4. The Critical Monte Carlo simulation

It becomes clear from the above review of experimental intermittency analysis that a better understanding is desirable of the way criticality in transverse momentum space is expressed through SSFMs, as well as the interplay of critical and non-critical protons and the effect a (dominant) background has on the strength and functional form of scaling of factorial moments. Monte Carlo simulations provide a royal road towards such insight.

For this purpose, we use a modified version of the Critical Monte-Carlo (CMC) event generator [3, 6, 8] suited for simulating protons in transverse momentum space; simulated protons are produced by sampling a truncated Lévy walk process to exhibit density-density correlations mimicking those originating from a fireball freezing out at the QCD critical point. The power-law exponent is adjustable within a range of values; for example, it can be chosen to describe correlations characterizing a critical system in the 3d-Ising universality class. The associated intermittency index range currently achievable is $\phi_2 \in [0.1, 1]$; the value $\phi_2 = 5/6$ corresponds to a fractal mass dimension of $d_F = 1/3$ for the 2-dimensional Lévy walk. Furthermore, the algorithm can be parametrized to reproduce any empirical 2-dimensional average per event proton $p_{x,y}$ distribution for the random walk cluster centers, and a (possibly) Poissonian per-event proton multiplicity distribution, the characteristics of which can be plugged in; finally, truncated Lévy walk bounds can be fine-tuned in order to produce critical density-density correlations within the desired scales.

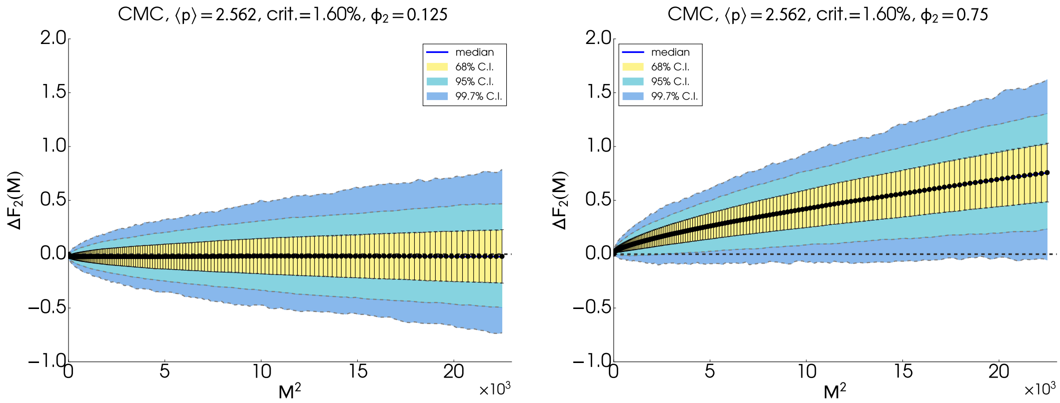


Figure 6: Correlator $\Delta F_2(M)$ for two sets of $\sim$ 400K CMC simulated SHINE-like non-central Ar+Sc collisions at 150A GeV/c ($\sqrt{s_{NN}} \approx 16.8$ GeV). For both sets, the percentage of critical to total simulated protons has been set to 1.60%. Critical exponent (intermittency index) is set to $\phi_2 = 0.125$ *(left)* and $\phi_2 = 0.75$ *(right)* respectively. The black points correspond to the average $\Delta F_2(M)$ of $\sim$ 8K independent iterations (samples) of the simulation. Colored bands correspond to 1-(*yellow*), 2-(*light blue*), and 3-σ (*dark blue*) confidence intervals of independent sample variation, respectively.

A number of uncorrelated proton momenta drawn from a plugged-in one-particle p_T distribution replace the critical protons with an adjustable probability per particle. These simulate the effect of non-critical background contamination on the critical signal, with the desired background level λ, eq.(5).

Additionally, an "afterburner" can be applied to CMC events in order to simulate detector effects for better comparison with experimental data. To this purpose, the 2-dimensional simulated proton momenta are assigned a rapidity distribution, extracted from experimental data, thus allowing us to apply all appropriate experimental cuts at the momentum level. In particular, rapidity, pair quality and acceptance cuts are applied, in the same manner as for SHINE experimental data. We also apply gaussian smearing of simulated proton momenta with an adjustable radius, in order to simulate the limited track momentum resolution in the detector.

Fig. 6 shows two examples of the correlator $\Delta F_2(M)$ obtained by the CMC simulation, for SHINE-like parametrized Ar+Sc collisions at 150A GeV/c. Two different values of the intermittency index exponent ϕ_2 have been used, one low *(left)* and one close to the critical prediction *(right)*. In each case, the factorial moments are calculated on sets of $\sim$ 400K events, similar to SHINE available statistics, and simulation is repeated independently for $\sim$ 8K iterations, keeping the same simulation parameters throughout. Independent simulations allow us to estimate the variability of the resulting $\Delta F_2(M)$ at this level of event statistics; the average $\Delta F_2(M)$ gives us the overall trend of the model. We observe qualitative behavior of the correlator similar to that of the experimental Ar+Sc SHINE set, Fig. 5. We also note a very similar internal spread of values as obtained for Ar+Sc SHINE via the bootstrap method, with the crucial difference that now the different samples are statistically independent.

Critical Monte Carlo simulations give us an intuitive illustration of the way a critical effect is diluted in experimental data, and of the limitations in detecting critical scaling in systems with a dominant background component. By directly comparing simulated and experimental $\Delta F_2(M)$

behaviour, we can gauge the range of plausible values for the critical exponent ϕ_2 as well as the critical percentage of protons in experimental data. Additionally, CMC simulations provide an empirical justification for replacing the full correlator $\Delta F_2(M)$ expression, eq.(5) with the "dominant background" approximation, eq.(6).

5. Challenges in intermittency analysis, and possible solutions

5.1 Challenges in intermittency analysis

In performing intermittency analysis, in the form it is traditionally pursued, one faces a number of practical and methodological challenges that need to be dealt with in order to obtain statistically significant results. We list the most important among them below:

1. Particle species, especially protons, cannot be perfectly identified experimentally; candidates will always contain a small percentage of impurities. In the case of protons, the main contamination comes from K^+ tracks misidentified as protons. In NA61/SHINE [18], we identify particles through their energy loss dE/dx in the Time Projection Chambers (TPCs), as a function of their momentum. Therefore, good particle identification requires quality decomposition of the total dE/dx spectra of tracks into a sum of gaussians of all particle species (p, K, π, e) in selected momentum space slices. Ideally, we accept candidate protons if they have an estimated probability of $> 90\%$ of being a proton; however, there is a delicate balance between achieving a high enough proton purity, and keeping the total multiplicity of accepted protons large enough for the demands of an intermittency analysis.

2. Experimental momentum resolution sets a limit to how small a bin size (large M) we can probe. Empirically, we have settled for a maximum of $M = 150$ one-dimensional bins, which corresponds to a Δp_T of 20 MeV/c. Experimental momentum resolution is of the order of ~ 5 MeV/c, well below our minimum bin size.

3. A finite (small) number of usable events is available for analysis; the "infinite statistics" behaviour of $\Delta F_2(M)$ must be extracted from these. The best way to guard against finite statistics artifacts is to investigate as many CMC simulated data sets as possible, with an event statistics comparable to that of the experimental data.

4. Proton multiplicity for medium-sized systems is low (typically $\sim 2-3$ protons per event, in the window of analysis) – and the demand for high proton purity lowers it still more. There is really no satisfactory solution to this issue, other than trying to increase event statistics. Failing that, we must turn to Monte Carlo simulation to gain insight about the statistical significance of our experimental data.

5. M-bins are correlated due to the fact that the same events are used to calculate all $F_2(M)$. This biases the fit algorithm for the intermittency index ϕ_2, and makes confidence interval estimation hard. Note that bin correlations persist even when $F_2(M)$ is calculated through the correlation integral, eq.(8,9), as can be evidenced by calculating the correlation matrix between different M-values.

PoS(CORFU2024)097

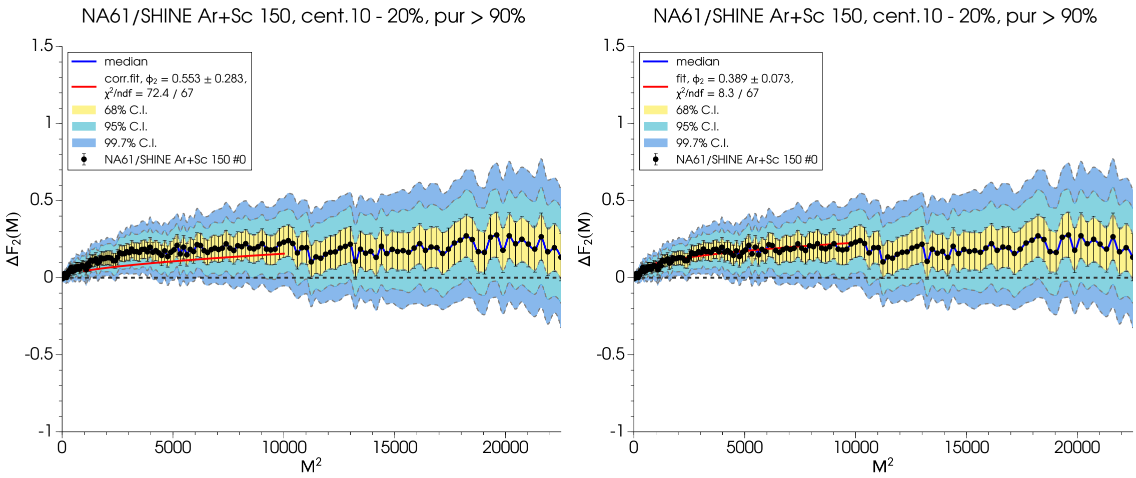


Figure 7: Correlated (*left*) and uncorrelated fit (*right*) for the NA61/SHINE Ar+Sc 10-20% most central system at 150*A* GeV/*c*.

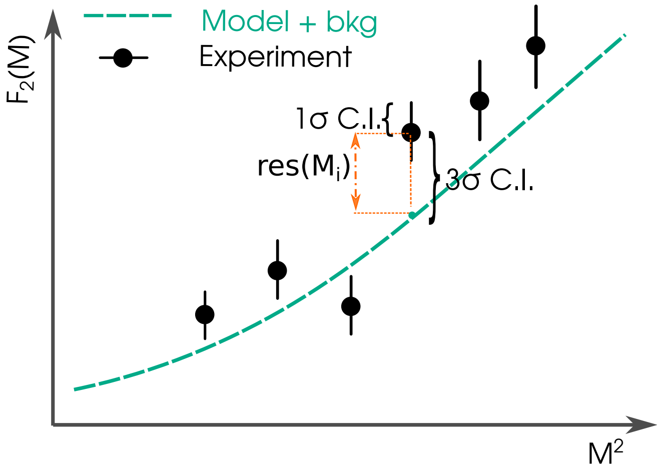
5.2 The statistical bootstrap and correlated fits

The method of the statistical bootstrap (see Section 2.2) can help us up to a point to resolve issues #3-5, in providing unbiased estimators for the magnitude of statistical uncertainties of SSFMs. As mentioned, however, it is not sufficient for taking into account M-bin correlations and systematic errors in SSFMs. Replication of events means bootstrap sets are not independent of the original: magnitude of variance and covariance estimates can be trusted, but central values will be biased to the original sample.

Correlated fits for obtaining ϕ_2 can be performed, using M-correlation matrix estimated via the bootstrap; however, these are known to be unstable: [21, 22]. Fig. 7 shows the results of a correlated (*left*) vs an uncorrelated (*right*) fit for the NA61/SHINE Ar+Sc 10-20% most central system at 150*A* GeV/*c*. The correlated best fit line unintuitively passes below all experimental points, in contrast to the uncorrelated fit, which passes through them; the two methods also give considerably different values for ϕ_2 and χ^2 of fit. Another possible approach, the method of independent bins, attempts to eliminate bin correlations by randomly partitioning the events into non-overlapping sets, and using a different set to calculate $F_2(M)$ for different M (see, for example, [23] for an application of this technique). This method, however, decimates the already low event statistics by spreading it over too few bins, thus inflating per bin uncertainties. A way to handle bin correlations without decimating the statistics will be presented in Section 5.4.

5.3 Performing a scan of models

The proposed solution to the problem of M-bin correlation is to avoid fitting for ϕ_2 altogether; rather, one should attempt to build a large number of Monte Carlo models, corresponding to different values of power-law scaling and percentage of critical protons, then weigh (test) these models against the experimental data. Fig. 8 *(left)* illustrates the main idea: $F_2(M)$ experimental values are compared to the average $F_2(M)$ of a given model via a goodness-of-fit function, such as the sum of their residual squares, $\chi^2 = \sum_i res^2(M_i)$; subsequently, a p-value is calculated for the



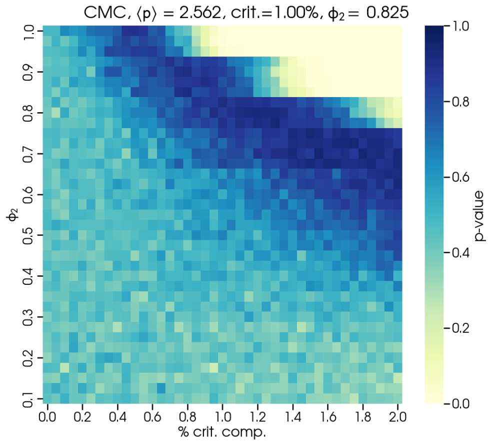


Figure 8: *Left*: A sketch of the comparison between model and experimental $F_2(M)$. A goodness-of-fit function can be defined based on the residuals between model and data. *Right*: Aggregating p-values for a grid of different models gives us a map of the likelihood of each model.

model by comparing the experimentally obtained χ^2 value to the χ^2-distribution of model samples. By creating a grid of models spanning different values of ϕ_2 and critical percentage of protons, Fig. 8 *(right)*, we can obtain a map of p-values, showing which models are more likely to fit the experimental data.

5.4 Handling bin correlations through Principal Component Analysis

As already mentioned, great caution must be exercised in defining a goodness-of-fit function for the correlator $\Delta F_2(M)$, because neighboring M-bins are strongly correlated. This is true even when constructing CMC samples of independent events, since the events in a single sample are used to calculate multiple M-scales. Additionally, when we attempt to study the joint distribution of $\Delta F_2(M)$ for all M, we are faced with a multi-dimensional structure of vast complexity (typically, $M \in [1, 150]$), which cannot possibly be probed with a few thousand samples (each independent sample being a point in this multi-dimensional M-space). It is thus necessary to reduce the effective dimensionality of this space, and untangle correlations between bins/dimensions.

This can be achieved through the well-established statistical tool of *Principal Component Analysis* (PCA). PCA works by taking a set of (in general) correlated multi-dimensional observations – seen as a cloud of points in a high-dimensional space – and identifying the principal axes going through it, i.e. the directions along which variance of points is maximal. The first principal axis accounts for the largest variance present in the set; subsequent axes account for progressively less and less variance, in a hierarchical structure. If we then rotate the original axes (in our case, the $F_2(M)$) to the principal axes directions, we can define new quantities, the Principal Components (PCs), which are by construction statistically independent linear combinations of the original M-bins. Finally, we keep only the first few significant PCs, by applying an appropriately chosen metric to determine the number of PCs that optimizes reconstruction of the original set of points; ideally, we want to keep only the main features of the original distribution, discarding the noise.

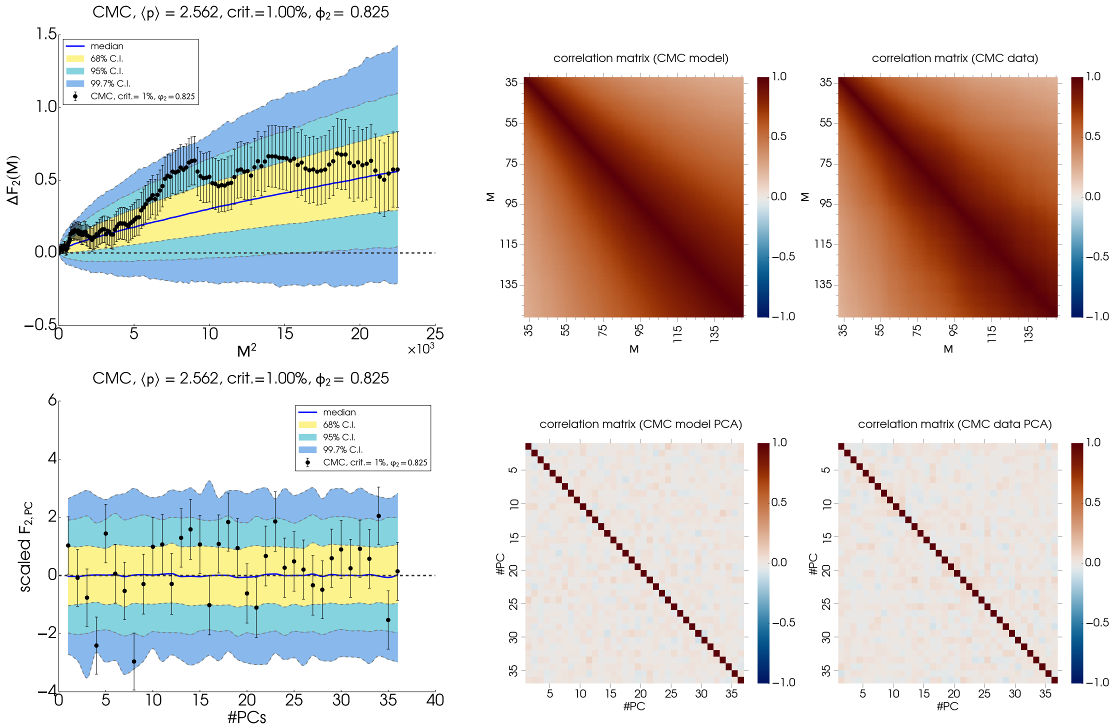


Figure 9: *(Top Row) Left:* Correlator $\Delta F_2(M)$ of a synthetic CMC data set (black points) versus $\sim$ 8K independent samples of CMC-generated events (colored bands); *Middle/Right:* The correlation matrix between M-bins, for CMC model and data samples, respectively. *(Bottom Row) Left:* The correlator $\Delta F_2(M)$ for the same CMC data sets, transformed to the PC coordinates; *Middle/Right:* The correlation matrix between PCs, for CMC model and data samples, respectively. Colored bands correspond to 1-(*yellow*), 2-(*light blue*), and 3-σ (*dark blue*) confidence intervals of independent sample variation, respectively.

Fig. 9 illustrates the action of PCA: in the top left plot, we show the $\Delta F_2(M)$ of a synthetic data set based on the Critical Monte Carlo [3, 6, 8] adapted to the SHINE detector conditions (black points). This set corresponds to the conditions of SHINE non-central (10-20%) Ar+Sc collisions at 150A GeV/c; the actual percentage of CMC-generated critical protons (critical component) was set at 1%. On the same plot, we show the $\Delta F_2(M)$ of $\sim$ 8K samples of CMC-generated events (colored bands) with parameters adjusted as to closely approximate the synthetic set. The top middle and top right plots show the correlation matrices of $\Delta F_2(M)$ values between different bins M for CMC model and data sets[2], respectively. We notice the two correlation matrices are qualitatively very similar, and both exhibit strong correlations around the main diagonal.

The bottom row of Fig. 9 shows the transformed $\Delta F_2(M)$ of CMC model and data sets, and their correlation matrices, in the rotated coordinates of the PCs. We use the correlation matrix of the $\Delta F_2(M)$ of the independent CMC samples in order to determine the PC directions, then we apply the transformation consistently to both CMC model and synthetic data sets. It must be noted that a different subset of CMC samples is used in the construction of the rotation, and a different one in the illustration of the effect, thus averting overfitting. Nevertheless, we clearly see the decoupling

[2]The correlation matrix for the CMC synthetic data set is approximated through the bootstrap, in the way outlined in Section 2.2

effect the PCA rotation has on the $\Delta F_2(M)$ values: the rotated correlation matrices of both CMC model and synthetic data set are diagonal to an excellent approximation, and therefore we can define χ^2 of samples simply as the sum of squares of their PC values (in the rotated coordinates, zero corresponds by construction to the average model behavior, illustrated by the dark blue line on the bottom left plot, Fig. 9).

5.5 CMC model scan, and the role of event statistics

Armed with the PCA analysis technique, we can systematically scan a wide range of CMC models and compare them against any experimental data set – in particular, the (itself CMC-generated) synthetic data set presented in Section 5.4. First, we match the overall characteristics of the models (proton multiplicity and momenta distributions, total number of events) to that of the experimental data set. Subsequently, we could form a grid of models in the parameters of critical component percentage and intermittency index, such as shown in Fig. 8 *(Right)*, with the parameters ranging in some reasonable intervals. However, in order to better illustrate the behaviour of the experimental set vs the models, it is actually better to focus on a single row of the 2-dimensional parameter grid, and plot the p-values as a function of varying the critical component percentage. We choose for this purpose a row roughly corresponding to the critically predicted exponent $\phi_2 = 5/6$, and vary the percentage of critical protons from 0-2% in steps of 0.05%. For each critical percentage choice, we simulate $\sim$ 10K independent CMC samples, and for each sample we calculate the correlator $\Delta F_2(M)$, using the sets with zero critical component as a substitute for the mixed events. We extract from these simulations the parameters of the PCA rotation. Finally, we perform the PCA rotation consistently for both model CMC and data moments, and compare the χ^2 of experimental to Monte Carlo samples per PC dimension. Our investigation on the reconstruction of the original CMC distributions indicates that $\sim$ 35 PCs should be kept in the analysis.

The results are shown in Fig. 10 *(top row)*. We fix $\phi_2 = 0.825$, and plot the p-values along this row of $\sim$ 1600 pseudo-experimental samples[3] (CMC simulations) of a model with no signal (Fig. 10 *(top left)*, critical component 0%, $\phi_2 = 0.1$), and a model with 1% signal (Fig. 10 *(top right)*, critical component 1%, $\phi_2 = 0.825$). In both scenarios, and all simulations, the total number of events per simulation is $\sim$ 400K, the same as in NA61/SHINE Ar+Sc collisions. Each CMC sample corresponds to a p-value curve in the plot as a function of critical component percentage; peaks of the curves correspond to more likely values of critical component percentage; the width of the peak (ignoring minor fluctuations) indicates the resolution, i.e. the power of our scan to tell models apart. Narrower peaks have more resolution power, as they can exclude most models outside the peak's width. We observe that for the usual NA61/SHINE event statistics the curves are wide, allowing us only to exclude a very strong signal ($\sim$ 2%) in the no-signal scenario, and not being at all able to exclude zero signal in the 1% signal scenario.

The picture is very different when we repeat the scan with $\sim$ 1600 pseudo-experimental samples of $\sim$ 4M events each, i.e. 10 times the usual Ar+Sc NA61/SHINE event statistics, Fig. 10 *(bottom row)*. In both scenarios (no-signal & 1% signal), the curves are now tightly clustered and narrowly peaked around the true critical component percentage value. We get almost twice the resolution of

[3]In the plot, only a small random subset of these $\sim$ 1600 curves is shown so as not to obfuscate the structure.

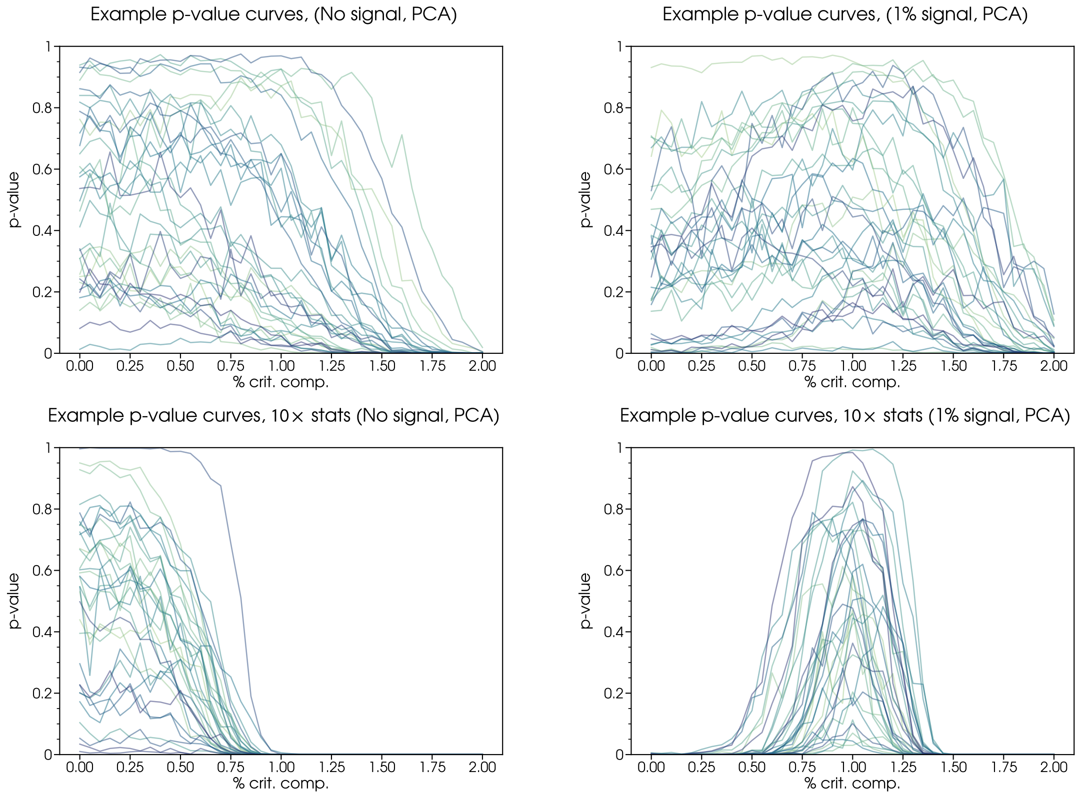


Figure 10: *(Top Row)* p-value curves for a collection of CMC pseudo-experimental samples corresponding to a no-signal model *(left)*, and a 1% signal model *(right)*, compared against a range of CMC models with $\phi_2 = 5/6$ and varying critical component percentages. All comparisons are performed in the Principal Component (PC) space; each experimental sample consists of $\sim$ 400K events; *(Bottom Row)* The same comparisons for the same models as in *Top Row*, calculated for experimental samples with 10× the event statistics ($\sim$ 4M events).

the low statistics scenario, being able to exclude as weak as 1% signals in the no-signal case, and anything outside the $\sim 0.5\% - 1.5\%$ signal range for the 1% signal case.

6. Concluding remarks

We have presented a review of the current status of experimentally-oriented intermittency analysis, focusing on proton intermittency. Intermittency analysis of proton density is a promising strategy for detecting the Critical Point of strongly interacting matter. However, it poses certain challenges in the context of an actual heavy-ion collision experiment, which is always constrained in terms of available event statistics, particle multiplicity, and proton identification. Furthermore, it is evident that large uncertainties in the determination of factorial moments, and especially strong bin correlations between different scales cannot be handled by the conventional analysis methodology.

New techniques have been developed, and are constantly being perfected, that allow us to handle statistical and systematic uncertainties without sacrificing event statistics. This is achieved through building Monte Carlo models of both critical behaviour and the non-critical background of a heavy-ion collision system near the critical point. These models are then compared against the experimental data via a scan in parameter space. At the same time, rotating from the original M-bins to appropriately determined principal components eliminates bin correlations, and ensures that the results of model comparison remain valid.

Critical Monte Carlo (CMC) investigation of Principal Component analysis performance indicates that large factorial moment fluctuations are inherent to proton intermittency analysis. Therefore, the effect of event statistics (number of analyzed events) on model parameter scan resolution is decisive. The event statistics currently available in NA61/SHINE recorded data sets, particularly Ar+Sc at $\sqrt{s_{NN}} \approx 16.8$ GeV, (on the order of $\sim$ 500K events) does not allow us to distinguish between a no-signal scenario and one with reasonable ($\leq 2\%$) levels of critical component. On the other hand, we could plausibly discern as weak as a $\sim 1\%$ critical component from no signal with about 10 times the currently available statistics, i.e. $\sim$ 5M events. Such an event statistics is well within the reach of e.g. the NA61/SHINE experiment, given its upgraded detector status.

Further exploration of refined models with critical and non-critical components is certainly desirable, in order to properly assess the requirements of an intermittency analysis of experimental data, and design proper protocols for robust analysis.

References

[1] NA49 collaboration, *Critical fluctuations of the proton density in A+A collisions at 158A GeV*, *Eur. Phys. J. C* **75** (2015) 587 [1208.5292].

[2] NA61/SHINE collaboration, *Searching for the Critical Point of Strongly Interacting Matter in Nucleus–Nucleus Collisions at CERN SPS*, *Acta Phys. Polon. Supp.* **13** (2020) 637 [2002.06636].

[3] N.G. Antoniou et al., *Decoding the QCD critical behaviour in $A + A$ collisions*, *Nucl. Phys. A* **1003** (2020) 122018 [2004.12133].

[4] Y. Hatta and M.A. Stephanov, *Proton number fluctuation as a signal of the QCD critical endpoint*, *Phys. Rev. Lett.* **91** (2003) 102003 [hep-ph/0302002].

[5] T. Vicsek, *Fractal Growth Phenomena*, World Scientific (1989).

[6] N.G. Antoniou, Y.F. Contoyiannis, F.K. Diakonos, A.I. Karanikas and C.N. Ktorides, *Pion production from a critical QCD phase*, *Nucl. Phys. A* **693** (2001) 799 [hep-ph/0012164].

[7] N.G. Antoniou, Y.F. Contoyiannis, F.K. Diakonos and G. Mavromanolakis, *Critical QCD in nuclear collisions*, *Nucl. Phys. A* **761** (2005) 149 [hep-ph/0505185].

[8] N.G. Antoniou, F.K. Diakonos, A.S. Kapoyannis and K.S. Kousouris, *Critical opalescence in baryonic QCD matter*, *Phys. Rev. Lett.* **97** (2006) 032002 [hep-ph/0602051].

[9] NA49 collaboration, *Search for the QCD critical point in nuclear collisions at* 158*A GeV at the CERN Super Proton Synchrotron (SPS)*, *Phys. Rev. C* **81** (2010) 064907.

PoS(CORFU2024)097

[10] A. Bialas and R.B. Peschanski, *Moments of Rapidity Distributions as a Measure of Short Range Fluctuations in High-Energy Collisions*, *Nucl. Phys. B* **273** (1986) 703.

[11] A. Bialas and R.B. Peschanski, *Intermittency in Multiparticle Production at High-Energy*, *Nucl. Phys. B* **308** (1988) 857.

[12] A. Bialas and R.C. Hwa, *Intermittency parameters as a possible signal for quark - gluon plasma formation*, *Phys. Lett. B* **253** (1991) 436.

[13] F.K. Diakonos and A.S. Kapoyannis, *Correlation integral vs. second order factorial moments and an efficient computational technique*, *Eur. Phys. J. C* **82** (2022) 200 [2109.12571].

[14] P. Grassberger and I. Procaccia, *Measuring the strangeness of strange attractors*, *PHYSICA D* **9** (1983) 189.

[15] W.J. Metzger, "Estimating the uncertainties of factorial moments." HEN-455 (unpublished), (2004).

[16] B. Efron, *Bootstrap Methods: Another Look at the Jackknife*, *The Annals of Statistics* **7** (1979) 1 .

[17] NA49 collaboration, *The NA49 large acceptance hadron detector*, *Nucl. Instrum. Meth. A* **430** (1999) 210.

[18] NA61 collaboration, *NA61/SHINE facility at the CERN SPS: beams and detector system*, *JINST* **9** (2014) P06005 [1401.4699].

[19] NA61/SHINE collaboration, *Search for the critical point of strongly interacting matter through power-law fluctuations of the proton density in NA61/SHINE*, *PoS* **CPOD2017** (2018) 054.

[20] NA61/SHINE collaboration, *Recent results from proton intermittency analysis in nucleus-nucleus collisions from NA61/SHINE at CERN SPS*, *PoS* **CORFU2018** (2019) 154.

[21] B. Wosiek, *Intermittency analysis of correlated data*, *Acta Phys. Polon. B* **21** (1990) 1021.

[22] C. Michael, *Fitting correlated data*, *Phys. Rev. D* **49** (1994) 2616 [hep-lat/9310026].

[23] NA61/SHINE collaboration, *Search for the critical point of strongly-interacting matter in $^{40}Ar + ^{45}Sc$ collisions at 150A GeV/c using scaled factorial moments of protons*, *Eur. Phys. J. C* **83** (2023) 881 [2305.07557].