Methods for a blind analysis of isobar data collected by the STAR collaboration

J. Adam[6], L. Adamczyk[2], J. R. Adams[39], J. K. Adkins[30], G. Agakishiev[28], M. M. Aggarwal[40], Z. Ahammed[59], I. Alekseev[3,35], D. M. Anderson[53], A. Aparin[28], E. C. Aschenauer[6], M. U. Ashraf[11], F. G. Atetalla[29], A. Attri[40], G. S. Averichev[28], V. Bairathi[22], K. Barish[10], A. J. Bassill[10], A. Behera[51], R. Bellwied[20], A. Bhasin[27], J. Bielcik[14], J. Bielcikova[38], L. C. Bland[6], I. G. Bordyuzhin[3], J. D. Brandenburg[48,6], A. V. Brandin[35], J. Butterworth[44], H. Caines[62], M. Calderón de la Barca Sánchez[8], D. Cebra[8], I. Chakaberia[29,6], P. Chaloupka[14], B. K. Chan[9], F-H. Chang[37], Z. Chang[6], N. Chankova-Bunzarova[28], A. Chatterjee[11], D. Chen[10], J. H. Chen[18], X. Chen[47], J. Cheng[55], M. Cherney[13], M. Chevalier[10], S. Choudhury[18], W. Christie[6], H. J. Crawford[7], M. Csanád[16], S. Das[11], M. Daugherity[1], T. G. Dedovich[28], I. M. Deppner[19], A. A. Derevschikov[42], L. Didenko[6], X. Dong[31], J. L. Drachenberg[1], J. C. Dunlop[6], T. Edmonds[43], N. Elsey[61], J. Engelage[7], G. Eppley[44], R. Esha[51], S. Esumi[56], O. Evdokimov[12], J. Ewigleben[32], O. Eyser[6], R. Fatemi[30], S. Fazio[6], P. Federic[38], J. Fedorisin[28], C. J. Feng[37], Y. Feng[43], P. Filip[28], E. Finch[50], Y. Fisyak[6], A. Francisco[62], L. Fulek[2], C. A. Gagliardi[53], T. Galatyuk[15], F. Geurts[44], A. Gibson[58], K. Gopal[23], D. Grosnick[58], W. Guryn[6], A. I. Hamad[29], A. Hamed[5], J. W. Harris[62], W. He[18], X. He[26], S. Heppelmann[8], S. Heppelmann[41], N. Herrmann[19], E. Hoffman[20], L. Holub[14], Y. Hong[31], S. Horvat[62], Y. Hu[18], B. Huang[12], H. Z. Huang[9], S. L. Huang[51], T. Huang[37], X. Huang[55], T. J. Humanic[39], P. Huo[51], G. Igo[9], D. Isenhower[1], W. W. Jacobs[25], C. Jena[23], A. Jentsch[6], Y. JI[47], J. Jia[6,51], K. Jiang[47], S. Jowzaee[61], X. Ju[47], E. G. Judd[7], S. Kabana[29], M. L. Kabir[10], S. Kagamaster[32], D. Kalinkin[25], K. Kang[55], D. Kapukchyan[10], K. Kauder[6], H. W. Ke[6], D. Keane[29], A. Kechechyan[28], M. Kelsey[31], Y. V. Khyzhniak[35], D. P. Kikoła [60], C. Kim[10], D. Kincses[16], T. A. Kinghorn[8], I. Kisel[17], A. Kiselev[6], A. Kisiel[60], M. Kocan[14], L. Kochenda[35], L. K. Kosarzewski[14], L. Kramarik[14], P. Kravtsov[35], K. Krueger[4], N. Kulathunga Mudiyanselage[20], L. Kumar[40], R. Kunnawalkam Elayavalli[61], J. H. Kwasizur[25], R. Lacey[51], S. Lan[11], J. M. Landgraf[6], J. Lauret[6], A. Lebedev[6], R. Lednicky[28], J. H. Lee[6], Y. H. Leung[31], C. Li[47], W. Li[49], W. Li[44], X. Li[47], Y. Li[55], Y. Liang[29], R. Licenik[38], T. Lin[53], Y. Lin[11], M. A. Lisa[39], F. Liu[11], H. Liu[25], P. Liu[51],

P. Liu[49], T. Liu[62], X. Liu[39], Y. Liu[53], Z. Liu[47], T. Ljubicic[6], W. J. Llope[61], M. Lomnitz[31], R. S. Longacre[6], N. S. Lukow[52], S. Luo[12], X. Luo[11], G. L. Ma[49], L. Ma[18], R. Ma[6], Y. G. Ma[49], N. Magdy[12], R. Majka[62], D. Mallick[36], S. Margetis[29], C. Markert[54], H. S. Matis[31], O. Matonoha[14], J. A. Mazer[45], K. Meehan[8], J. C. Mei[48], N. G. Minaev[42], S. Mioduszewski[53], B. Mohanty[36], M. M. Mondal[36], I. Mooney[61], Z. Moravcova[14], D. A. Morozov[42], M. Nagy[16], J. D. Nam[52], Md. Nasim[22], K. Nayak[11], D. Neff[9], J. M. Nelson[7], D. B. Nemes[62], M. Nie[48], G. Nigmatkulov[35], T. Niida[56], L. V. Nogach[42], T. Nonaka[11], G. Odyniec[31], A. Ogawa[6], S. Oh[62], V. A. Okorokov[35], B. S. Page[6], R. Pak[6], A. Pandav[36], Y. Panebratsev[28], B. Pawlik[2], D. Pawlowska[60], H. Pei[11], C. Perkins[7], L. Pinsky[20], R. L. Pintér[16], J. Pluta[60], J. Porter[31], M. Posik[52], N. K. Pruthi[40], M. Przybycien[2], J. Putschke[61], H. Qiu[26], A. Quintero[52], S. K. Radhakrishnan[29], S. Ramachandran[30], R. L. Ray[54], R. Reed[32], H. G. Ritter[31], J. B. Roberts[44], O. V. Rogachevskiy[28], J. L. Romero[8], L. Ruan[6], J. Rusnak[38], O. Rusnakova[14], N. R. Sahoo[48], H. Sako[56], S. Salur[45], J. Sandweiss[62], S. Sato[56], W. B. Schmidke[6], N. Schmitz[33], B. R. Schweid[51], F. Seck[15], J. Seger[13], M. Sergeeva[9], R. Seto[10], P. Seyboth[33], N. Shah[24], E. Shahaliev[28], P. V. Shanmuganathan[6], M. Shao[47], F. Shen[48], W. Q. Shen[49], S. S. Shi[11], Q. Y. Shou[49], E. P. Sichtermann[31], R. Sikora[2], M. Simko[38], J. Singh[40], S. Singha[26], N. Smirnov[62], W. Solyst[25], P. Sorensen[6], H. M. Spinka[4], B. Srivastava[43], T. D. S. Stanislaus[58], M. Stefaniak[60], D. J. Stewart[62], M. Strikhanov[35], B. Stringfellow[43], A. A. P. Suaide[46], M. Sumbera[38], B. Summa[41], X. M. Sun[11], Y. Sun[47], Y. Sun[21], B. Surrow[52], D. N. Svirida[3], P. Szymanski[60], A. H. Tang[6], Z. Tang[47], A. Taranenko[35], T. Tarnowsky[34], J. H. Thomas[31], A. R. Timmins[20], D. Tlusty[13], M. Tokarev[28], C. A. Tomkiel[32], S. Trentalange[9], R. E. Tribble[53], P. Tribedy[6], S. K. Tripathy[16], O. D. Tsai[9], Z. Tu[6], T. Ullrich[6], D. G. Underwood[4], I. Upsal[48,6], G. Van Buren[6], J. Vanek[38], A. N. Vasiliev[42], I. Vassiliev[17], F. Videbæk[6], S. Vokal[28], S. A. Voloshin[61], F. Wang[43], G. Wang[9], J. S. Wang[21], P. Wang[47], Y. Wang[11], Y. Wang[55], Z. Wang[48], J. C. Webb[6], P. C. Weidenkaff[19], L. Wen[9], G. D. Westfall[34], H. Wieman[31], S. W. Wissink[25], R. Witt[57], Y. Wu[10], Z. G. Xiao[55], G. Xie[31], W. Xie[43], H. Xu[21], N. Xu[31], Q. H. Xu[48], Y. F. Xu[49], Y. Xu[48], Z. Xu[6], Z. Xu[9], C. Yang[48], Q. Yang[48], S. Yang[6], Y. Yang[37], Z. Yang[11], Z. Ye[44], Z. Ye[12], L. Yi[48], K. Yip[6], H. Zbroszczyk[60], W. Zha[47], D. Zhang[11], S. Zhang[47], S. Zhang[49], X. P. Zhang[55], Y. Zhang[47], Z. J. Zhang[37], Z. Zhang[6], J. Zhao[43],

C. Zhong[49], C. Zhou[49], X. Zhu[55], Z. Zhu[48], M. Zurek[31], M. Zyzak[17]

[1] *Abilene Christian University, Abilene, Texas 79699*

[2] *AGH University of Science and Technology, FPACS, Cracow 30-059, Poland*

[3] *Alikhanov Institute for Theoretical and Experimental Physics NRC "Kurchatov Institute", Moscow 117218, Russia*

[4] *Argonne National Laboratory, Argonne, Illinois 60439*

[5] *American University of Cairo, New Cairo 11835, New Cairo, Egypt*

[6] *Brookhaven National Laboratory, Upton, New York 11973*

[7] *University of California, Berkeley, California 94720*

[8] *University of California, Davis, California 95616*

[9] *University of California, Los Angeles, California 90095*

[10] *University of California, Riverside, California 92521*

[11] *Central China Normal University, Wuhan, Hubei 430079*

[12] *University of Illinois at Chicago, Chicago, Illinois 60607*

[13] *Creighton University, Omaha, Nebraska 68178*

[14] *Czech Technical University in Prague, FNSPE, Prague 115 19, Czech Republic*

[15] *Technische Universität Darmstadt, Darmstadt 64289, Germany*

[16] *ELTE Eötvös Loránd University, Budapest, Hungary H-1117*

[17] *Frankfurt Institute for Advanced Studies FIAS, Frankfurt 60438, Germany*

[18] *Fudan University, Shanghai, 200433*

[19] *University of Heidelberg, Heidelberg 69120, Germany*

[20] *University of Houston, Houston, Texas 77204*

[21] *Huzhou University, Huzhou, Zhejiang 313000*

[22] *Indian Institute of Science Education and Research (IISER), Berhampur 760010 , India*

[23] *Indian Institute of Science Education and Research (IISER) Tirupati, Tirupati 517507, India*

[24] *Indian Institute Technology, Patna, Bihar 801106, India*

[25] *Indiana University, Bloomington, Indiana 47408*

[26] *Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou, Gansu 730000*

[27] *University of Jammu, Jammu 180001, India*

[28] *Joint Institute for Nuclear Research, Dubna 141 980, Russia*

[29] *Kent State University, Kent, Ohio 44242*

[30] *University of Kentucky, Lexington, Kentucky 40506-0055*

[31] *Lawrence Berkeley National Laboratory, Berkeley, California 94720*

[32] *Lehigh University, Bethlehem, Pennsylvania 18015*

[33] *Max-Planck-Institut für Physik, Munich 80805, Germany*

[34] *Michigan State University, East Lansing, Michigan 48824*

[35] *National Research Nuclear University MEPhI, Moscow 115409, Russia*

[36] *National Institute of Science Education and Research, HBNI, Jatni 752050, India*

[37] *National Cheng Kung University, Tainan 70101*

[38] *Nuclear Physics Institute of the CAS, Rez 250 68, Czech Republic*

[39] *Ohio State University, Columbus, Ohio 43210*

[40] *Panjab University, Chandigarh 160014, India*

[41] *Pennsylvania State University, University Park, Pennsylvania 16802*

[42] *NRC "Kurchatov Institute", Institute of High Energy Physics, Protvino 142281, Russia*

[43] *Purdue University, West Lafayette, Indiana 47907*

[44] *Rice University, Houston, Texas 77251*

[45] *Rutgers University, Piscataway, New Jersey 08854*

[46] *Universidade de São Paulo, São Paulo, Brazil 05314-970*

[47] *University of Science and Technology of China, Hefei, Anhui 230026*

[48] *Shandong University, Qingdao, Shandong 266237*

[49] *Shanghai Institute of Applied Physics, Chinese Academy of Sciences, Shanghai 201800*

[50] *Southern Connecticut State University, New Haven, Connecticut 06515*

[51] *State University of New York, Stony Brook, New York 11794*

[52] *Temple University, Philadelphia, Pennsylvania 19122*

[53] *Texas A&M University, College Station, Texas 77843*

[54] *University of Texas, Austin, Texas 78712*

[55] *Tsinghua University, Beijing 100084*

[56] *University of Tsukuba, Tsukuba, Ibaraki 305-8571, Japan*

[57] *United States Naval Academy, Annapolis, Maryland 21402*

[58] *Valparaiso University, Valparaiso, Indiana 46383*

[59] *Variable Energy Cyclotron Centre, Kolkata 700064, India*

[60] *Warsaw University of Technology, Warsaw 00-661, Poland*

[61] *Wayne State University, Detroit, Michigan 48201*

[62] *Yale University, New Haven, Connecticut 06520*

Abstract

In 2018, the STAR collaboration collected data from ${}^{96}_{44}Ru + {}^{96}_{44}Ru$ and ${}^{96}_{40}Zr + {}^{96}_{40}Zr$ at $\sqrt{s_{NN}} = 200$ GeV to search for the presence of the chiral magnetic effect in collisions of nuclei. The isobar collision species alternated frequently between ${}^{96}_{44}Ru + {}^{96}_{44}Ru$ and ${}^{96}_{40}Zr + {}^{96}_{40}Zr$. In order to conduct blind analyses of studies related to the chiral magnetic effect in these isobar data, STAR developed a three-step blind analysis procedure. Analysts are initially provided a "reference sample" of data, comprised of a mix of events from the two species, the order of which respects time-dependent changes in run conditions. After tuning analysis codes and performing time-dependent quality assurance on the reference sample, analysts are provided a species-blind sample suitable for calculating efficiencies and corrections for individual $\approx$ 30-minute data-taking runs. For this sample, species-specific information is disguised, but individual output files contain data from a single isobar species. Only run-by-run corrections and code alteration subsequent to these corrections are allowed at this stage. Following these modifications, the "frozen" code is passed over the fully un-blind data, completing the blind analysis. As a check of the feasibility of the blind analysis procedure, analysts completed a "mock data challenge," analyzing data from $Au + Au$ collisions at $\sqrt{s_{NN}} = 27$ GeV, collected in 2018. The $Au + Au$ data were prepared in the same manner intended for the isobar blind data. The details of the blind analysis procedure and results from the mock data challenge are presented.

Keywords:

1. Introduction

For more than a decade, the STAR Collaboration has been searching for evidence of chiral magnetic effects (CME) [1, 2, 3]. CME [4, 5] refers to the induction of an electric current ($\vec{J}_e$) by the magnetic field ($\vec{B}$) in a chiral system: $\vec{J}_e \propto \mu_5 \vec{B}$. A chiral system bears a nonzero μ_5, which characterizes the imbalance of right-handed and left-handed fermions in the system. The discovery of CME in high-energy heavy-ion collisions would confirm the simultaneous existence of ultra-strong magnetic fields, chiral symmetry restoration, and topological charge changing transitions in these collisions. On average, $\vec{B}$ is perpendicular to the reaction plane ($\Psi_{\rm RP}$) that contains the impact parameter and the beam momenta. CME, therefore, will manifest a charge transport across the reaction plane.

A set of observables common to CME searches are the charge-separation fluctuations perpendicular to $\Psi_{\rm RP}$, e.g. with a three-point correlator [6], $\gamma \equiv \langle \cos(\phi_\alpha + \phi_\beta - 2\Psi_{\rm RP}) \rangle$, where averaging is done over all particles in an event and over all events. To draw firm conclusions on the presence of CME, an effective way is needed to disentangle the signal and background contributions, the latter of which are intertwined with collective flow. Collisions of isobaric nuclei, e.g. ${}^{96}_{44}Ru + {}^{96}_{44}Ru$ and ${}^{96}_{40}Zr + {}^{96}_{40}Zr$, present an opportunity to vary the initial magnetic field while keeping background conditions approximately the same [7]. Ruthenium-96 and Zirconium-96 each have 96 nucleons but with different numbers of protons, 44 and 40 for Ru and Zr, respectively. Monte Carlo Glauber simulations indicate ${}^{96}_{44}Ru + {}^{96}_{44}Ru$ and ${}^{96}_{40}Zr + {}^{96}_{40}Zr$ collisions at the same beam energy are almost identical in terms of particle production [8]. The ratio of the multiplicity distributions from the two collision systems is consistent with unity almost everywhere, except in $0-5\%$ most central collisions, where the slightly larger charge radius of Ru ($R_0 = 5.085$ fm) plays a role against that of Zr ($R_0 = 5.02$ fm). CME analyses can focus on the centrality range of $20-60\%$, where the background difference due to the multiplicity are negligible. A theoretical calculation using the HIJING model [8] indicates the relative difference in the square of the initial magnetic field between ${}^{96}_{44}Ru + {}^{96}_{44}Ru$ and ${}^{96}_{40}Zr + {}^{96}_{40}Zr$ collisions approaches $15-18\%$ for peripheral events and $\approx 13\%$ for central events. These estimates translate into a relative difference in the CME signal observable between the two isobars of 3%, assuming an 80% background from elliptic flow, requiring a minimum of 1.2×10^9 events to pass the various analysis selection criteria to achieve a result of 5σ significance. Due to the small dif-

ference in the CME signal observables, of critical importance to the analysis is control of systematic uncertainties, in particular those related to detector acceptance and efficiency, which may vary in a time-dependent way.

In 2018, STAR collected data from isobar collisions, ${}^{96}_{44}Ru + {}^{96}_{44}Ru$ and ${}^{96}_{40}Zr + {}^{96}_{40}Zr$, at $\sqrt{s_{NN}} = 200$ GeV. For the first time, the STAR Collaboration has implemented blind analyses of these data in studies related to CME. While blind analyses are not uncommon in particle physics, e.g. Ref. [9], the typical methods were not found to be suitable for the specific needs of STAR CME analyses. What follows is the description of the blind analysis procedure for the 2018 isobar collision data. The procedure was accepted by the STAR Collaboration prior to CME data-taking. While primarily relevant for CME-related studies, the opportunity for a blind analysis was open to all STAR analyses of 2018 isobar data. In identifying as a "STAR blind analysis" for the 2018 isobar running, analysts adhere to the following procedure. Subsequent STAR publications clearly identify as a blind analysis or an "un-blind" analysis according to the accepted procedure. The following procedure takes advantage of frequent switching of the isobar collision species during 2018 RHIC running to interleave isobar data samples from each species in a way that respects the time-variation of data running conditions. STAR collected 6.3 billion isobar events, evenly split between the two species, during the two months of isobar running. The RHIC isobar stores or "fills" typically lasted 20 hours, with STAR collecting data during 30-minute "runs" of the data acquisition system. Accelerator operators adjusted the beam optics throughout the 20 hour fills to maintain nearly constant collision rates, with the same target rate for the two isobar species.

2. Blinding Techniques

2.1. General principle

Blind analyses often rely on a "reference sample" and an inability to differentiate two or more samples or a particular sample from the reference (see Ref. [9] for a brief overview of blind analyses in particle physics). The reference sample is often used either to tune an analysis without predetermined bias or to provide a reference for evaluating the significance of a result, e.g. eliminating placebo effects or genetic conditions that may bias the result of medical studies.

2.2. Considerations

While many possibilities exist, the blinding method for a particular analysis should be well-matched to the specific needs of that analysis. Among the specific considerations for analysis of the 2018 STAR isobar data are the following:

- The un-blind data should not be accessible by physics analysts prior to analysis tuning.
- Accounting for time-dependent detector fluctuations is a critical component of analysis quality assurance (Q/A).
- Accounting for run-by-run anomalies is a critical component of final analysis Q/A.
- Randomizing variables within an event may severely compromise the quality of analysis Q/A and associated corrections. For example, randomizing the sign of charged particle tracks would prevent charge-dependent efficiency corrections; and randomizing particle azimuthal angle would destroy correlations from secondary decays. Because of these considerations, such methods are not retained as part of this procedure.
- To ensure the isobar species have statistically comparable behaviors in terms of luminosity, event trigger composition, energy, vertex distribution, occupancy of tracks, etc., the 2018 RHIC run involved frequent switching of the isobar collision species.
- With this consideration in mind, it is feasible to interleave or "mix" events from the two collision species in a given output data file as an efficient method to disguise the collision species.
- Certain STAR experts, recused from blind physics analyses, may require isobar information during RHIC running to ensure data of sufficient quality to achieve target physics goals.
- Calibration experts, who are recused from conducting blind physics analyses, may need access to un-blind data to ensure sufficiently robust calibrations and corrections to achieve the desired physics goals.

- Runs of quality suitable for inclusion in physics analyses, e.g. not exhibiting large detector inefficiencies, must be selected prior to the mixing of events from different species.

For the blind analysis of isobar data collected in 2018, STAR adopted a three-step blinding procedure. For the first step, analysts are provided output data files that mix events from the two isobar collision species, while respecting the time-dependence of run conditions. Analysts use this data sample to perform time-dependent Q/A of the data and to tune analysis codes. At the conclusion of these studies, analysts commit their code to a repository. In the second step, analysts are provided an "unmixed-blind," sample suitable for calculating corrections that vary according to individual $\approx$ 30-minute data-taking runs. The run identification number are disguised, but the output data files do not mix events from different runs. Only these "run-by-run" corrections (e.g. for changing detector efficiencies) and code alterations subsequent to these corrections are allowed during this step. At the conclusion of these studies, the final codes are committed to the repository, so that differences may be evaluated. After the analysis codes are verified, the final data analysis pass is completed using these final codes and the fully un-blind data released.

2.3. Initial procedure

Initial implementation of the analysis blinding procedure began prior to and during the 2018 RHIC run. To the extent possible, information pertaining to the isobar species was restricted during the run. Access to raw data for purposes of Q/A during the run was restricted to identified experts, approximately 5% of the collaboration, recused from blind physics analyses. To the extent possible, all raw data samples were limited in size below the level needed for sensitivity to a CME signal, e.g. less than 10000 events. Un-blind experts produced species-blind performance plots to evaluate data quality for the run in-progress.

Prior to the software production of the blind data, it was necessary to set detector calibrations and determine an appropriate list of quality data-taking runs. Due to the importance of robust calibrations to the physics analyses, these calibrations were performed by the relevant experts using un-blind data. These calibration experts were recused from participation in blinded physics analyses. Additionally, a committee was designated to determine data-taking runs of sufficient quality for inclusion in physics analyses.

Members of this run selection committee were also recused from participation in blinded physics analyses. Production of the blind data commenced after calibrations and the designation of good runs.

No physics analysis groups are provided with un-blinded data prior to completion of the un-blinding procedure.

2.4. Blind data production

In the blind production of data, the following information encoded in the data stream (DST) are obfuscated: the identification numbers for the event, its particular data-taking run, and RHIC fill; the event timestamp; the event collision species; and the hit rates for the east and west STAR zero-degree calorimeters (ZDC) [10] and beam-beam counters (BBC) [11], as well as their coincidence and background rates. All output data files are assigned a generic name and pseudo-run-number that monotonically increases with time. The exact start time of a data production is not known to ensure, e.g. that a particular pseudo-run-number is not trivially related to a particular isobar species. The mixing procedure and exact algorithm to re-assign pseudo-run numbers are encrypted and only known by two experts, who are recused from performing blind physics analyses. The reference sample, species-separated samples, and fully unblind samples are provided in a three-step process.

2.5. Step-1: "The Reference"

Analysts are initially provided output files composed of events from a mix of the two isobar species. The mixing procedure is not a priori known. As much as possible, the order of events respects temporal changes in running conditions. Events showing peculiar discrepancies from the initial Q/A are excluded from the sample, and events from the two species are only combined if the detector performance, e.g. acceptance, was similar for the two events. Events are randomly rejected at the level of $\sim$ 10%, so that "counters," e.g. trigger words, cannot be used to determine the species. Analysis code and time-dependent Q/A are tuned on this reference sample, committed to the analysis code repository, and kept unchanged at this stage. Among other aspects, this step enables extraction of time-dependent spectra for Q/A, detection of time-dependent anomalies, detection of secondary decays and measurement of peak widths relevant to momentum resolution.

2.6. Step-2: "The run by run Q/A sample"

After analysis of the reference data, analysts are provided an "unmixed-blind sample" comprised of files that obscure the true run number (and, hence, the isobar species) but do not mix events across different runs. The pseudo-run-number uniquely maps to one true run number and one (unknown) isobar species. The data are provided in such a way that a mix of files from each species appear in the same directory. As, in the first step, a fraction of events from each run is rejected to ensure that simple counting of events could not decipher the species. This sample enables species-blind run-by-run Q/A. Only run-by-run corrections and code alteration directly resulting from these corrections are allowed at this stage. The number of events provided per file is tuned so that statistics are sufficient for robust corrections but insufficient for deciphering the isobar species.

2.7. Step-3: Full Un-blinding

Once Q/A is complete and analyses of the run-by-run Q/A data are final, full un-blinding proceeds. At this stage, physics results are produced with the previously tuned, vetted, and fixed analysis codes. In this data production, all information is un-blinded and restored to the data files.

3. Implementation and Timeline for Blinded Analyses

No STAR physics analyses had access to species information prior to un-blinding. The timelines for un-blinding are estimated by the blind analysts, who present regular updates to their respective physics working groups (PWG) to document progress and to inform adjustments to the timeline. Decisions to un-blind are based upon a review of thoroughly documented analysis procedures, codes, and analysis reports–including estimates of measurement uncertainty–by the relevant PWG. In addition, for blind analyses of the isobar data, so-called "godparent committees" or "GPCs," are set early and follow analyses closely throughout their development. The GPCs serve an important role in verifying that analyses are ready to proceed to the next stages of the blinding procedure. After the step-1 data are available, blind-data analysts estimate a timeline for completing the necessary analyses for advancing to step-2. Based on this input from the analysts, management approves a date for the beginning of the second step. Analysts present regular updates to document progress. Regardless of progress, un-blinding occurs no earlier than the original estimate unless all blind analyses are deemed

ready to proceed by STAR Management. Based upon the progress reports, un-blinding may be delayed to ensure the quality of the final results. An analogous timeline procedure is done for the full un-blinding. Prior to the first un-blinding step, analysts prepare detailed notes documenting the procedures, cuts, corrections, systematic uncertainties, and criteria for any future run-by-run cuts and corrections. Prior to the second un-blinding step, analysts ensure that the documentation is updated and complete, including the run-by-run portion of analyses. Prior to each un-blinding step, analysts provide analysis codes for vetting and Q/A by the GPC in addition to the standard vetting within the physics working groups.

When the GPC is satisfied that an analysis is ready for un-blinding, analysts present the status of their analyses to the physics working group conveners and the physics analysis coordinator. As the un-blinding date approaches, analysts discuss with STAR management any need for delays to un-blinding to ensure the quality of results. If an unresolved disagreement exists between analysts, the decision to un-blind or extend the date lies with STAR management. After physics results are produced with un-blinded data, a review is conducted to verify that the frozen analysis code was used to produce the results.

While un-blinded data are not accessible to physics analyses until the blinding timeline is completed, management uses discretion in applying blinding to any calibration analysis. To ensure the integrity of calibrations, e.g. those of the beamline and TPC [12], STAR calibration experts may require access to un-blind data. Without robust calibrations, the physics analyses may not be able to achieve the required precision for deciphering a CME signal. Therefore, the relevant experts are allowed access to the un-blind data for these tasks. Furthermore, access to un-blind data is restricted to these experts alone and the experts recuse themselves from participation in any blind physics analysis.

4. Mock data challenge

As the recommended analysis blinding procedure represents a substantial departure from that typical for STAR analyses, testing feasibility is critical. Toward this end, a "mock data challenge" was conducted utilizing data from $Au + Au$ collisions at $\sqrt{s_{NN}} = 27$ GeV, also collected in 2018. Additionally, this exercise served as an opportunity for the software and computing team to develop, tune, and test the machinery necessary for producing the blind

data samples. "Blinded" samples of these data were provided to analysts, utilizing the same techniques intended for blinding the isobar data. One sample was provided with output data files containing events from a mixture of data-taking runs, simulating the first stage of blinding, where data files contain a mix of isobar species. Another sample was provided using output files containing events from single data-taking runs but still blinding certain variables that in the isobar data sample could be used to identify the isobar species. For completeness, a final un-mixed sample was provided with no information obscured, simulating the fully "un-blind" phase of the analysis. Analysts used the two mock blind-data samples to perform quality control studies and appropriately tune analysis codes, selection cuts, and corrections. Once completed, the analysts then ran the same codes over the un-blind data sample to verify that the analysis was feasible with the given data structures and that results were appropriately consistent within the statistical differences between the samples. Example quality assurance plots for the three different samples are shown in Fig. 1. Note that the different samples did not contain identical sets of events.

5. After Un-blinding

After un-blinding, only changes to correct "mistakes," defined for this purpose as errors in arithmetic or unintended departures from the approved and documented analysis procedures, are allowed. If such a correction is made, the analysis results with the error will also be provided with a detailed explanation of the specific correction applied and why it was needed. On a case-by-case basis, the collaboration considers announcing the result from a blind analysis simultaneously with the submission of the corresponding paper to the journal and the preprint arXiv. Regardless, only one set of "final" results from the blind analysis will be released, e.g. there will be no set of "preliminary" results prior to the "final" results. All STAR publications of 2018 results state explicitly whether the analysis followed the approved STAR blinding procedure.

6. Conclusion

The STAR Collaboration has developed a procedure to carry out blind analyses of isobar collision data, collected in 2018. The procedure described in this manuscript was accepted by the STAR Council in January 2018, prior

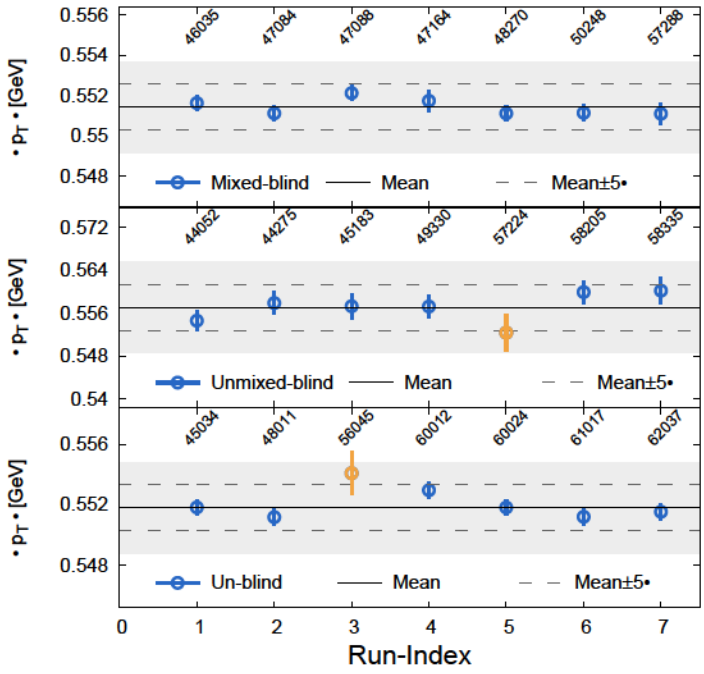

Figure 1: Mean transverse momentum of charged particle tracks associated with the primary collision vertex for three data samples: (top) "mixed-blind" output files containing events from a mix of data-taking runs, (middle) "unmixed-blind" output files containing events not mixed across data-taking runs, and (bottom) "un-blind" output files containing events not mixed across data-taking runs. The observable is shown as a function of an arbitrary data-taking run index with the pseudo-run-number shown in the top labels of the panel. Note that for the "mixed-blind" sample this index simply indicates events from the same output file.

to the isobar collision runs. The initial step in the procedure is an analysis of blinded data samples that interleave events from the two collision species, while the second step involves analysis of blinded data samples that do not mix events from the two collision species, followed by complete un-blinding of the data. Prior to commencing with analysis of the isobar data, a mock data challenge was successfully conducted to demonstrate the feasibility of the procedure both from an analysis standpoint and a computational standpoint. Analyses of the blind data are underway, following the procedure outlined in this manuscript.

References

[1] B. I. Abelev, et al., Azimuthal Charged-Particle Correlations and Possible Local Strong Parity Violation, Phys. Rev. Lett. 103 (2009) 251601. arXiv:0909.1739, doi:10.1103/PhysRevLett.103.251601.

[2] B. I. Abelev, et al., Observation of charge-dependent azimuthal correlations and possible local strong parity violation in heavy ion collisions, Phys. Rev. C81 (2010) 054908. arXiv:0909.1717, doi:10.1103/PhysRevC.81.054908.

[3] G. Wang, Search for Chiral Magnetic Effects in High-Energy Nuclear Collisions, Nucl. Phys. A904-905 (2013) 248c–255c. arXiv:1210.5498, doi:10.1016/j.nuclphysa.2013.01.069.

[4] D. Kharzeev, Parity violation in hot QCD: Why it can happen, and how to look for it, Phys. Lett. B633 (2006) 260–264. arXiv:hep-ph/0406125, doi:10.1016/j.physletb.2005.11.075.

[5] D. E. Kharzeev, L. D. McLerran, H. J. Warringa, The Effects of topological charge change in heavy ion collisions: 'Event by event P and CP violation', Nucl. Phys. A803 (2008) 227–253. arXiv:0711.0950, doi:10.1016/j.nuclphysa.2008.02.298.

[6] S. A. Voloshin, Parity violation in hot QCD: How to detect it, Phys. Rev. C70 (2004) 057901. arXiv:hep-ph/0406311, doi:10.1103/PhysRevC.70.057901.

[7] S. A. Voloshin, Testing the Chiral Magnetic Effect with Central U+U collisions, Phys. Rev. Lett. 105 (2010) 172301. arXiv:1006.1020, doi:10.1103/PhysRevLett.105.172301.

[8] W.-T. Deng, X.-G. Huang, G.-L. Ma, G. Wang, Test the chiral magnetic effect with isobaric collisions, Phys. Rev. C94 (2016) 041901. arXiv:1607.04697, doi:10.1103/PhysRevC.94.041901.

[9] J. R. Klein, A. Roodman, Blind analysis in nuclear and particle physics, Ann. Rev. Nucl. Part. Sci. 55 (2005) 141–163. doi:10.1146/annurev.nucl.55.090704.151521.

[10] F. S. Bieser, et al., The STAR trigger, Nucl. Instrum. Meth. A499 (2003) 766–777. doi:10.1016/S0168-9002(02)01974-5.

[11] J. Kiryluk, Relative luminosity measurement in STAR and implications for spin asymmetry determinations, AIP Conf. Proc. 675 (1) (2003) 424–428. doi:10.1063/1.1607171.

[12] M. Anderson, et al., The Star time projection chamber: A Unique tool for studying high multiplicity events at RHIC, Nucl. Instrum. Meth. A499 (2003) 659–678. arXiv:nucl-ex/0301015, doi:10.1016/S0168-9002(02)01964-2.