A Short Review on Jet Identification

Sevil Salur\textsuperscript{1,a} for the STAR Collaboration

Lawrence Berkeley National Laboratory, 1 Cyclotron Road MS-70R0319, Berkeley, CA 94720

Received: date / Revised version: date

Abstract. Jets can be used to probe the physical properties of the high energy density matter created in collisions at the Relativistic Heavy Ion Collider (RHIC). Measurements of strong suppression of inclusive hadron distributions and di-hadron correlations at high \( p_T \) have already provided evidence for partonic energy loss. However, these measurements suffer from well-known geometric biases due to the competition of energy loss and fragmentation. These biases can be avoided if the jets are reconstructed independently of their fragmentation details - quenched or unquenched. In this paper, we discuss modern jet reconstruction algorithms (cone and sequential recombination) and their corresponding background subtraction techniques required by the high multiplicities of heavy ion collisions. We review recent results from the STAR experiment at RHIC on direct jet reconstruction in central Au+Au collisions at \( \sqrt{s_{NN}} = 200 \text{ GeV} \).

PACS. 21.65.Qr Relativistic Heavy Ion Collisions – 13.87.Ce Jet Production

1 Introduction

A highly collimated “spray” of particles also known as “jets” is produced in high energy collisions. Hadronic jets are the experimental signatures of quarks and gluons (partons) and they reflect the underlying parton kinematics \cite{1,2,3,4}. Cross section measurements of jets are performed at many hadronic and leptonic colliders to check in detail perturbative QCD (pQCD) calculations, to help determine parton distribution functions and to look for new physics. The inclusive jet cross section at Tevatron is measured very precisely over 20 orders of magnitude and is found to be in a very good agreement with the NLO pQCD calculations using CTEQ 6.1 parton distribution functions \cite{5,6}. The robustness of the theoretical calculations on jet cross sections in \( p + \bar{p} \) collisions motivates the use of jets as direct probes of partonic energy loss in dense matter generated in ultra-relativistic heavy ion collisions at RHIC and in near future at the LHC \cite{7,8}.

Due to the limited pseudo-rapidity (\( \eta \)) and azimuthal \( \phi \) coverage of the electromagnetic calorimeters of RHIC experiments during the first 6 years of RHIC operation, the background determination needed for direct jet reconstruction was not possible. Instead other observables such as the strong suppression of inclusive hadron distributions and di-hadron correlations at high \( p_T \) were measured. However, such measurements suffer from geometric biases due to the competition of energy loss and fragmentation - the leading particle spectrum is dominated by relatively low energy jets that happen to lose little energy in the medium and fragment into higher \( p_T \) particles \cite{9}.

These biases can be removed if the reconstructed partonic kinematics is independent of whether the fragmentation is modified by the medium or not. An unbiased jet reconstruction measurement in heavy ion collisions would give access to the full spectrum of fragmentation topologies without geometric biases, enabling full exploration of quenching dynamics. In addition, fully reconstructed jets allow the measurement of qualitatively new observables such as jet shapes, fragmentation functions, and energy flow. Since 2006, the STAR barrel electromagnetic calorimeter (BEMC) with full azimuthal coverage and unit pseudorapidity acceptance is operational. This enables the study of the underlying event background required for full jet reconstruction in heavy ion collisions for the first time at RHIC. The experimental results discussed in this article were presented for the first time during the Hard Probes 2008 meeting \cite{10} for the first direct measurement of jets and \cite{11} for the accompanying measurement of jet fragmentation studies in heavy ion collisions.

2 Jet Reconstruction Algorithms

During the last 20 years, various jet reconstruction algorithms were developed to combine measured particles into jets. For a detailed overview of jet algorithms in high energy collisions, see \cite{12,13,14,15} and references therein. The primary requirement for jet algorithms is low sensitivity to hadronization, radiation and splitting. They should also be defined equally at hadron and parton level. This was suggested by Weinberg 30 years ago \cite{4}. According to him “Quark and gluon jets can be compared to detector jets, if jet algorithms respect collinear and infrared safety.”
Full jet reconstruction in heavy ion collisions is a new frontier. As in leptonic and hadronic collisions, in heavy ion collisions the chosen jet reconstruction method should also be theoretically and experimentally consistent. As the experimental energy determination degrades the resolution somewhat with the detectors, the algorithm should also aim to minimize resolution effects from unrelated sources such as the underlying event. The expected increase in Large Hadron Collider (LHC) luminosities (20 to 200 collisions in a detector) requires that the traditional jet algorithms have to be improved with underlying event subtraction techniques for \( p + p \) collisions to resolve events for pile up. These improved techniques can also be used in heavy ion environments (Au+Au or Pb+Pb) where the background subtraction is required due to large multiplicities of produced particles [14, 16].

The algorithm should also be detector independent i.e., allow the combination of particles detected in various detectors. The results discussed in this article are from jets that are reconstructed by combining the neutral energy from the BEMC and charged particles from the Time Projection Chamber (TPC) of the STAR experiment. STAR’s TPC and BEMC detectors cover full azimuth (0 < \( \phi \) < 2\( \pi \)) and mid rapidity (\(-1 < \eta < 1\)) of the events. The Figure 1 shows an example of a reconstructed di-jet event in the STAR experiment.

![Fig. 1](image)

**Fig. 1.** 21 GeV di-jet reconstructed from a single event with a combined transverse momentum and energy for charged and neutral particles per grid cell in the \( \eta \) and \( \phi \) plane from 0-20% most central Au+Au collisions [11].

Corrections for double-counting of energy due to hadronic energy deposition in the BEMC and to electrons are applied. Two kinds of jet reconstruction algorithms are utilized: seeded cone (leading order high seed cone (LOHSC)) and sequential recombination (\( k_T \) and Cambridge/Aachen). In the following, we briefly discuss these two algorithms and the corresponding underlying event subtractions.

### 2.1 Cone Algorithms

The cone algorithms have been used as a primary tool to identify jets at hadron colliders since the early 1980s. This algorithm is based on the picture that a jet consists of a large amount of hadronic energy in a small angular region. Therefore, the main method is to combine particles in \( \eta - \phi \) space with their neighbors within a cone of radius \( R (R = \sqrt{\Delta\phi^2 + \Delta\eta^2}) \). This is illustrated schematically in Figure 2.

![Fig. 2](image)

**Fig. 2.** The transverse momentum of tracks and/or deposition of energy in calorimeter towers are represented as the solid squares in \( \eta \) and \( \phi \) space. Red squares defined as seeds are the ones above a given threshold. Jet cones are solid circles around the seeds. All the energy of the particles is added for the given cone around the seed particles to estimate the jet energy. The background in the jet energy is estimated with the average of the total energy in cones without seeds (shown in dashed circles) and is subtracted on an event-by-event basis.

To optimize the search and effectiveness of jet finding, these algorithms use splitting, merging, and iteration steps in the events of leptonic and hadronic collisions. However, to avoid instabilities in cone-finding due to large heavy ion background, we use a simple seeded cone without iteration or split-merging steps, with cone radius \( R = 0.4 \) and minimum seed of 4.6 GeV. The choice of the relatively small cone size is to suppress the underlying heavy ion background [17,18]. In \( p + p \) collisions \( \sim 80\% \) of the jet energy is observed to be within \( R \sim 0.3 \) for 50 GeV jets in the Tevatron data [6]. However, broadening of the jet fragmentation due to quenching in the medium formed in heavy ion collisions may reduce the fraction of the measured energy in a given cone size and needs further exploration. To reduce the heavy ion background, the minimum accepted transverse momentum of charged particles, and the transverse energy of the calorimeter cells (\( p_T^{cut} \)) is varied between 0.1 to 2 GeV. This threshold cut does not remove all the background contamination on the jet energy and additional subtraction is needed. As presented in Figure 2 schematically, the residual background is corrected based on the out-of-cone energy for the same \( p_T^{cut} \), averaged over the STAR acceptance but measured on an event-by-event basis, and scaled to the cone area.
A recently developed seedless infrared-safe cone algorithm (SISCone) [19] resolves most of the ambiguities associated with the various cone algorithms. The SISCone algorithm is already used in $p + p$ collisions at $\sqrt{s} = 200$ GeV and the first results can be found in [20].

### 2.2 Sequential Recombination Algorithms

The sequential recombination algorithms have been used extensively in the Tevatron as they are collinear and infrared safe [6, 21, 22]. In these types of algorithms, arbitrarily shaped jets are allowed to follow the energy flow resulting in less bias on the reconstructed jet shape than with cone algorithms [14]. Figure 3 represents a schematic comparison of the jet areas for cone and $k_T$ type algorithms.

![Fig. 3. A schematic comparison of the jet area for cone and $k_T$ type algorithms.](image)

The sequential recombination algorithms combine objects in relative to the closeness of their $p_T$. Particles are merged into a new cluster via successive pair-wise recombination. Algorithmic details can be found in [6, 21] and references therein. The FastJet code package for sequential recombination algorithms was used for the STAR analyses in $p + p$ and $Au + Au$ collisions [10, 11, 20, 23, 24]. This package includes $k_T$, Cambridge/Aachen (CAMB), anti-$k_T$, and an interface to external jet finders such as Seedless Infrared Safe Cone (SisCone) via a plugin mechanism [19]. For infrared and collinear safe algorithms an active area ($A_j$) of each jet is estimated by filling an event with many very soft particles and then counting how many are clustered into a given jet. If the underlying event is distributed uniformly in $\eta$ and $\phi$ then this noise density can be subtracted from the measured jet energy on an event-by-event basis to correct for the background energy underlying the jet. In simulations, this correction is observed to recover most of the jet energy when they are reconstructed in pile up and heavy ion backgrounds [14]. The $k_T$, Cambridge/Aachen and anti-$k_T$ algorithms are all based on the same sequential recombination algorithm, but they differ in the distance measure that is used to group particles into jets.

Figure 4 shows the jet-area from $k_T$ algorithm from FastJet code for 0-10% central $Au + Au$ events (MB-Trig), in PYTHIA isolated jet events embedded in real central $Au + Au$ events (PyEmbed) and in PYTHIA isolated jet events (PyTrue).

![Fig. 4. Jet area from $k_T$ algorithm is reconstructed utilizing the FastJet code package [14] for real jets in minimum bias triggered 0-10% central $Au + Au$ collision (MB-Trig), in PYTHIA isolated jet events embedded in real central $Au + Au$ events (PyEmbed) and in PYTHIA isolated jet events (PyTrue).](image)

### 3 Jet Reconstruction Analysis

#### 3.1 Event Selection and Terminology

This analysis utilizes events of $Au+Au$ collisions at $\sqrt{s_{NN}} = 200$ GeV recorded by STAR. Using the multiplicity measurement performed with the STAR TPC, only the most central (0-10%) $Au + Au$ collisions are selected. Two event sets were analyzed, based on fast on-line trigger configurations:

(i) MB-Trig, (minimum-bias trigger) utilizing the coincidence between the two calorimeters at beam rapidity (Zero Degree Calorimeters) with signals in each greater than $\sim 40\%$ of the most probable amplitude for a single neutron, and

(ii) HT-Trig (high-tower trigger) that satisfies the MB-Trig conditions and the additional requirement of 2-tower EMC clusters having at least a 7.5 GeV energy deposition.

Three million MB-Trig events, corresponding to 300 thousand 0-10% most central events are used for this study. A total of 80 million MB-Trig events were recorded by STAR during the 2007 $Au + Au$ run, but only the event set used have been fully reconstructed off-line. The HT-Trig is designed to enhance the recorded rate of high $p_T$ photons and electrons. It may also serve to enhance the
In order to assess jet reconstruction energy resolution, background subtraction, efficiency and acceptance, Monte-Carlo model studies based on PYTHIA 8.107 [25] are performed. PYTHIA events with high $E_T$ jets are generated in three different configurations:

(i) PyTrue: PYTHIA isolated jets including all particles except neutrinos. Jets are reconstructed using the PYTHIA internal jet finder, PyCell, for the cone algorithm, and FastJet for the sequential recombination algorithm.

(ii) PyDet: PYTHIA isolated jets (parameterized detector response level) reconstructed using the jet algorithms that are also applied to the real data.

(iii) PyEmbed: PyDet that are embedded in a background of real Au+Au 0-10% central events, with jets reconstructed with the jet algorithms that are also applied to the real data.

In all the real events and simulations, only the highest energy jet per event is selected as the reconstructed jet.

### 3.2 Energy Resolution

The energy resolution for jet reconstruction with various algorithms has been studied with isolated jets simulated with PYTHIA [25, 26, 27]. Figure 5 shows the event by event comparison of PyTrue, PyDet and PyEmbed from LOHSC algorithm. See [10] for comparison of the energy resolution with the $k_T$ and Cambridge/Aachen algorithms. Applied cuts and jet energy are specified in the figures. A shift of median due to un-measured particles, primarily neutrons and $K_0^S$, and the applied $p_T$ cut (hence loss of jet energy), is observed for the $\Delta E = E_{PyDet} - E_{PyTrue}$ histogram. Background effects are simulated using PYTHIA jets that are embedded in real Au+Au events.

The distribution in Figure 5 is convoluted with the true jet spectrum to produce the observed jet spectrum. The effect of the heavy ion background on the jet energy can be seen in the $\Delta E = E_{PyEmbed} - E_{PyTrue}$ distribution. A positive $\Delta E$ in this distribution can distort the measured inclusive jet spectrum substantially, increasing the apparent yield at high $E_T$ and resulting in a harder spectrum. A correction to the spectrum must be applied to account for this effect.

The influence of energy resolution on jet spectrum can be observed with PyDet, PyEmbed and PyTrue distributions shown in Figure 6 for the LOHSC algorithm for $p_T^{cut} = 0.1$ GeV. A large difference between PyEmbed and PyDet is observed as expected from the tail at positive $\Delta E$ due to large background in the red distribution in Figure 5. When the $p_T$ threshold is increased, the background fluctuations are reduced and the enhancement in the spectrum relative to the case without background is reduced to a negligible level [10]. Jet reconstruction in 0-10% most central Au+Au collisions is similar to that of $p+p$ collisions with a larger $p_T^{cut}$ GeV threshold requirement. However, a reduction in the measured jet energy ($p_T^{cut}$ dependent bias) is introduced. Similar effects, though smaller in magnitude, are also observed for the $k_T$ and Cambridge/Aachen algorithms.

### 3.3 Jet Spectra Corrections and Comparisons

The correction factors for the jet spectrum are estimated using PYTHIA simulated jets embedded in real Au+Au collisions. The $p_T^{cut}$ = 0.1 GeV on track momentum and calorimeter cell energy is applied for PyDet and PyEmbed. Note the lower threshold on generated jet energy $E_T^{PyTrue} > 5$ GeV, which affects the reconstructed spectrum up to $E_T = 20$ GeV.
the ratio distribution is used as a multiplicative correction to the inclusive spectrum. Table 1 shows the inclusive jet spectrum correction factors for various $p_T^{cut}$ values. For the sequential clustering algorithms, the correction factors are closer to unity.

The corrected inclusive jet spectrum for the LOHSC algorithm for the $p_T^{cut} = 1$ GeV is presented in Figure 7. The $p_T^{cut} = 1$ GeV is selected as it corresponds to correction factors close to unity. At this $p_T^{cut}$ the competing effects of energy loss due to momentum threshold cut and the kick up in the jet spectrum due to the positive tail of energy resolution cancel each other. The solid triangles are for the MB-Trig data set and are corrected for resolution, acceptance and efficiency.

| $p_T^{cut}$ (GeV) | LOHSC | kT | CAMB |
|-------------------|-------|----|------|
| 0.1               | 0.2-10| 1-4| 2-6  |
| 1                 | 0.2-1 | 0.7-1| 1-2  |
| 2                 | 0.2-0.3| 0.5-1| 0.5-1 |

In order to assess the biases in the jet spectrum reconstructed in central Au+Au collisions, we compare to the spectrum measured in p+p collisions. To account for nuclear geometric effects we scale the p+p spectrum by $N_{\text{Binary}}$, the number of binary nucleon+nucleon collisions equivalent to a central Au+Au collisions, as calculated by a Glauber model [28]. The cross section for hard processes is expected to scale with $N_{\text{Binary}}$ if no nuclear effects are present. In the case of jet reconstruction, $N_{\text{Binary}}$ scaling is expected if the reconstruction is unbiased, i.e. the jet energy is recovered independent of the particular mode of fragmentation, even in the presence of strong jet quenching. The $N_{\text{Binary}}$ scaled jet spectrum from $p+p$ collisions is shown in solid squares [29]. The yellow band represents the systematic uncertainty of the $p+p$ jet measurement. Heavy ion jet spectrum is observed to agree with $N_{\text{Binary}}$ scaled p+p measurement within the $\sim 50$% systematic uncertainty of the normalization.

The open circles in Figure 7 show the uncorrected jet spectrum from HT-Trig data which is substantially lower than the corrected MB-Trig spectrum. The correction factors of energy resolution, efficiency and acceptance for the HT-Trig are expected to be small. A large trigger bias due to the additional 7.5 GeV energy deposition in EMCAL requirement for HT-Trig relative to MB-Trig is seen to persist at least to 30 GeV. Further statistics of MB-Trig data is needed to assess the bias at high $p_T$.

Figures 8,9,10 show the comparison of inclusive jet spectra from the MB-Trig Au+Au data and the $N_{\text{Binary}}$ scaled p+p for the $p_T = 0.1$ GeV threshold cut for kT, Cambridge/Aachen and LOHSC algorithms. While the agreement between $N_{\text{Binary}}$ scaled p+p and MB-Trig measurement is good for $p_T^{cut} = 0.1$ GeV, it is also seen to be poorer with the larger $p_T$ threshold cut [11]. This suggests that $p_T^{cut}$ introduces biases which are not fully corrected.
4 Summary

The full reconstruction of jets in 0-10% most central heavy ion collisions at RHIC energies is presented. Systematics of the underlying heavy ion background subtraction is studied utilizing various algorithms and consideration of the jet area. The $N_{\text{Binary}}$ scaling is observed for the least-biased cuts with the given $\sim 50\%$ systematic uncertainty of the $p+p$ jet spectrum measurement. An unbiased jet reconstruction in heavy ion collisions with MB-Trig data appears to be feasible. However, spectrum corrections are currently based on model calculations using PYTHIA fragmentation. This aspect, together with background subtraction techniques, spectrum variations due to cuts and reconstruction algorithms, must be investigated further in order to assess the systematic uncertainties of this measurement.

The first heavy ion run at LHC is expected in late 2009. The heavy ion background is predicted to be larger at the LHC than at RHIC, but there will be copious production of very energetic jets, well above background [30]. The large kinematic reach at the LHC may provide sufficient lever-arm to map out the QCD evolution of jet quenching [31]. The comparison of full jet measurements in the different physical systems generated at RHIC and the LHC will provide unique and crucial insights into our understanding of jet quenching and the nature of hot QCD matter.

References

1. S. D. Drell and D. J. Levy, T. M. Yan, Phys. Rev. 187 2159 (1969).
2. N. Cabibbo, G. Parisi and M. Testa, Lett. Nuovo Cimento 4, 35 (1970).
3. J. D. Bjorken and S. D. Brodsky, Phys. Rev. D 1, 1416 (1970).
4. G. Sterman and S. Weinberg, Phys. Rev. Lett. 39, 1436 (1977).
5. A. Abulencia et al., Phys. Rev. D 75 092006 (2007); 119901 (2007) ; A. Abulencia et al., Phys. Rev. Lett. 96 122001 (2006).
6. http://www-cdf.fnal.gov/physics/new/qcd/QCD.html
7. J. Adams, et. al. Phys. Rev. Lett. 91 072304 (2003); J. Adams, et. al. Phys. Rev. Lett. 97 162301 (2006).
8. S. S. Adler, et. al. Phys. Rev. Lett. 96 202301 (2006).
9. R. Baier, Y. L. Dokshitzer, S. Peigne and D. Schiff, Phys. Lett. 345B, 277 (1995); R. Baier, Y. L. Dokshitzer, A. H. Mueller, S. Peigne, D. Schiff, Nucl. Phys. B, 483, 291-320 (1997).
10. S. Salur for STAR Collaboration, nuc-ex/0809.1609.
11. J. Putschke for STAR Collaboration, nuc-ex/0809.1419.
12. G. C. Blazey et al. FERMILAB-CONF-00-092-E, hep-ex/0005012.
13. G. P. Salam, hep-ph 0801.0070.
14. M. Cacciari, G. P. Salam, G. Soyez, JHEP 0804:005 (2008).
15. B. Flaugher and K. Meier. FERMILAB-CONF-90-248-E, Dec, (1990).
16. M. Cacciari, G. Salam Phys.Lett. B659, 119-126, (2008); M. Cacciari, G. Salam hep-ph/0512210.
17. S. L. Blyth et al., J. Phys. G 34 271 (2007).
18. J. Putschke for the ALICE Collaboration, Workshop on "Parton fragmentation processes: in the vacuum and in the medium" Feb 2008.
19. G. P. Salam and G. Soyez, [http://projects.hepforge.org/siscone](http://projects.hepforge.org/siscone)
20. E. Bruna for STAR Collaboration, see Hot Quarks 2008 proceedings.
21. N. Brown, W. J. Stirling Z.Phys. C 53:629-636 (1992).
22. S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock and B. R. Webber, Phys. Lett. B 269, 432 (1991); S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, Nucl. Phys. B 406, 187 (1993); S. D. Ellis and D. E. Soper, Phys. Rev. D 48, 3160 (1993).
23. FastJet, M. Cacciari, G. P. Salam and G. Soyez, [http://www.lpthe.jussieu.fr~salam/fastjet](http://www.lpthe.jussieu.fr/~salam/fastjet).
24. M. Cacciari, G. P. Salam, G. Soyez, JHEP 0804:063 (2008).
25. T. Sjostrand, S. Mrenna, P. Skands, Comput. Phys. Commun. 178, 852-867 (2008); T. Sjostrand, S. Mrenna, P. Skands hep-ph:0710.3820; T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna, E. Norrbin Comput. Phys. Commun. 135, 238-259 (2001).
26. J. Adams, et. al. Phys. Lett. B 637, 161 (2006).
27. B. I. Abelev, et. al. Phys. Rev. C 75, 049901 (2007).
28. R. J. Glauber, Phys. Rev. 100,242 (1955).
29. B. I. Abelev, et. al. Phys. Rev. Lett. 97 252001 (2006).
30. P.M. Jacobs and M. van Leeuwen, Nucl. Phys. A774 237-246 (2006).
31. J. Casalderrey-Solano and X. N. Wang, Phys.Rev.C 77, 024902 (2008).