BUILDING BETTER CONE JET ALGORITHMS

S.D. Ellis,[1] J. Huston,[2] and M. Tönnesmann[3]

[1]Seattle, WA 98195 USA
[2]E. Lansing, MI 48824 USA
[3]Föhringer Ring 6, 80805 München, Germany
(Dated: August 16, 2003)

The search for physics beyond the Standard Model is being greatly enhanced by improved theoretical tools and ideas at the same time that vast amounts of new high energy data are becoming availiable. To make the most of this situation it is essential that we simultaneously improve our phenomenological tools, such as jet algorithms, to more reliably bridge the gap between theory and experiment. We present recent results on the development of better cone jet algorithms.

I. INTRODUCTION

A common facet of essentially all physics studies at current and future accelerators is the important role of hadronic jets in the characterization of the final states. These jets are intended to serve as surrogates for the underlying partons (quarks and gluons) in terms of which, along with leptons, high energy scattering events are most easily analyzed. The explicit connection between the sprays of final state particles observed in the detectors and the underlying partons is typically supplied by a jet algorithm. These algorithms are employed to map final states, both in QCD perturbation theory (*i.e.*, in terms of a relatively small number of energetic partons) and in the data (*i.e.*, in terms of the observed hadrons and/or calorimeter towers), onto jets. Thus, in principle, we can connect the observed final states, in all of their complexity, with the short-distance perturbative final states, which are easier to interpret and to analyze theoretically. For example, one would like to be able to use jets to identify and reconstruct the presence of W's, top quarks and Higgs bosons from their hadronic decays into quarks (and gluons). Given the large challenge inherent in such a procedure, it is no surprise that an important component of the preparations[1] for Run II at the Tevatron, and for future data taking at the LHC, has been the study of ways in which to improve jet algorithms. It is this topic that we will address in the following discussion.

In order to understand how jet algorithms work it is useful to have in mind a simple (essentially classical) space-time picture of how hard-scattering events evolve. Think of time zero as flagged by the basic hard-scattering process that, via the exchange of large transverse momentum, transforms 2 energetic but longitudinally moving (initial state) partons into 2, or more, energetic (final state) partons with substantial transverse momentum. This picture serves to characterize the short distance ($\ll 1$ fermi) configuration accurately described by perturbative QCD. At next-to-leading order (NLO) in perturbative QCD there are at most 3 such large transverse momentum final state partons, with at most 2 in a single jet. As the final state evolves in time and space there is further (softer and largely collinear) radiation or showering. At distance scales of order a fermi or larger (non-perturbative) confinement physics or hadronization is relevant. The partons with nonzero QCD charge organize themselves into color-singlet hadrons. To accurately perform the task of connecting fixed order perturbative final states to the experimentally observed final states the jet algorithms must be robust under the impact of both higher order perturbative and non-perturbative physics. In the context of a theoretical calculation the jet algorithm must be insensitive to the corners of phase space where the perturbation theory is potentially divergent, *e.g.*, insensitive to emission of soft and/or collinear gluons. Likewise the resulting jets should not depend on the details of the long distance process by which the partons associate themselves into color singlet hadrons. This includes both the inherently nonperturbative "hadronization" process and the essentially perturbative, but all orders showering process. All of these processes tend to smear out the local structure of the low order perturbative description. On the experimental side we desire insensitivity to the effects introduced by the detectors themselves. We also need to limit the size of the required numerical analysis in order to ensure timely analysis with finite computer resources. The past implementation of jet algorithms has invariably involved compromises on at least some of these issues, as we will discuss. The (optimistic) quantitative goal is a precision of order 1% in the comparison between theory and experiment. In this note we will ignore the uncertainties due to higher order perturbative corrections and uncertainties in the parton distributions functions and focus on the issues that arise from the algorithms themselves.

II. CONE JET HISTORY

We begin by reviewing cone jet algorithms, which have formed the basis of jet studies at hadron colliders. First employed at the $Sp\bar{p}S$ at CERN[2], the idea of the cone jet was explicitly spelled out in the Snowmass Algorithm[3], which was developed by a collaboration of theorists and experimentalists. The Snowmass Algorithm led to the jet algorithms used by the CDF and DØ collaborations during Run I at the Tevatron, although neither Collaboration followed the Snowmass prescription precisely. The intuitive picture is that jets should be composed of hadrons or partons that are, in some sense, nearby each other. The fundamental cone jet idea is that nearness be defined in a simple geometric fashion: jets are composed of hadrons or partons whose 3-momenta lie within a cone defined by a circle in the angular variables (η, ϕ), where $\eta = \ln(\cot\theta/2)$ is the pseudorapidity and ϕ is the azimuthal angle. This idea of being nearby in angle can be contrasted with an algorithm based on being nearby in transverse momentum as illustrated by the so-called k_T Algorithm[4] that has been widely used at e^+e^- and ep colliders, and is now being studied at the Tevatron[5]. Intuitively we also expect the jets to be aligned with the most energetic particles in the final state and correspondingly we expect to be able to identify a unique set of jets on an event-by-event basis. This uniqueness expectation is realized in the Snowmass Algorithm by identifying jets with "stable" cones. A stable cone has the property that the geometric center of the cone coincides with the E_T weighted centroid of the particles in the cone. We imagine defining a final state in terms of a list of final state partons (theory) or hadrons and/or calorimeter towers (experiment) each labeled by an index k, a direction (η_k, ϕ_k) and a scalar "transverse energy" (assuming that we start with massless 4-vectors), $E_{T,k} = E\sin\theta = |\vec{p}_{T,k}|$. According to the Snowmass Algorithm the constituents of a jet (J) of cone radius R and the corresponding jet properties are defined by the following set of equations

$$
\begin{aligned}
k \in J : (\phi_k - \phi_J)^2 + (\eta_k - \eta_J)^2 &\leq R^2, \\
\phi_J = \sum_{k \in J} \frac{E_{T,k}\phi_k}{E_{T,J}}, \quad \eta_J &= \sum_{k \in J} \frac{E_{T,k}\eta_k}{E_{T,J}}, \\
E_{T,J} = \sum_{k \in J} E_{T,k}. &
\end{aligned}
\tag{1}
$$

The first line specifies that the jet consists of all partons, hadrons or towers that lie within the distance R of the jet cone center, (η_J, ϕ_J), while the second line specifies that the E_T weighted centroid of this list of particles coincides with the geometric jet cone center (*i.e.*, the second line is the stability condition). The third line defines the Snowmass total scalar E_T for the jet. Typically the value $R = 0.7$ has been employed as it leads to fairly stable results both theoretically and experimentally.

It is important to recognize that jet algorithms involve two distinct steps. The first step is to identify the "constituents" of the jet, *i.e.*, the list of calorimeter towers or hadrons or partons that comprise the stable cone that is the jet. The second step involves constructing the kinematic properties that will characterize the jet, *i.e.*, determine into which bin the jet will be placed. In the Snowmass Algorithm the E_T weighted variables defined in Eq. 1 are used both to identify and bin the jet, *i.e.*, in terms of η_J and $E_{T,J}$. The simplicity and intuitive appeal of the cone algorithm justify its wide spread usage at hadron colliders and it has proved to be a very useful tool. However, its practical implementation has involved several compromises that now constitute limits to its precision, as we will attempt to explain.

In a theoretical calculation[6] the Snowmass Algorithm can be applied quite literally by integrating only over the portions of multi-parton phase space corresponding to parton configurations that satisfy the stability conditions in Eq. 1. In the experimental case it was imagined that the Snowmass Algorithm would be applied to each event via an iterative process to search the entire final state, *i.e.*, the entire particle/tower list mentioned above, for sets of final state particles/towers that satisfy the stability constraint. In practice[1] the experimental implementation of the cone algorithm has been more complicated, involving at least 3 detailed steps without direct analogues in the theoretical implementation. First, experimentalists have employed various short cuts to minimize the (computer) search time. In particular, Run I algorithms made use of both "seeds" and "pre-clustering". The iterative search for stable cones was initiated only at locations provided by preclusters assembled from contiguous seed towers, *i.e.*, calorimeter towers with deposited energy exceeding a predefined limit (typically ˜1 GeV). Preclusters are constrained to extend no more than $2R$ in either η or ϕ. Starting only at such a seeded location in (η, ϕ), a list is constructed of the particles (towers) within a distance R of the seed. Then the E_T weighted centroid for the particles in the list is found (calculated as in Eq. 1). If the calculated centroid is consistent with the initial cone center, a stable cone has been identified. If not, the calculated centroid is used as the center of a new cone with a new list of particles inside and the calculation of the centroid is repeated. This process is iterated (as envisioned in the Snowmass Algorithm), with the cone

center migrating with each repetition, until a stable cone is identified or until the cone centroid has migrated out of the fiducial volume of the detector. Note, however, that since the search starts only at the seeds, this process may not identify all possible stable cones and thus may not find all jets. Of particular concern are the final states that arise from two energetic partons of nearly equal energy that are separated in angle by nearly $2R$. This configuration will be identified as a single jet in the theoretical perturbative analysis with the center of the jet approximately mid-way between the partons. The smeared version of this configuration in the actual (or simulated) final state, on the other hand, will likely exhibit a local minimum of energy (a saddlepoint) at the midpoint between the partons (assuming the angular smearing due to showering and hadronization is not large compared to R), with seeds only at the location of the partons. As discussed in more detail below, the experimental iterative search initiated (*i.e.*, seeded) only at the location of the two partons will likely find two jets centered on the partons, rather than the higher energy solution containing the remnants of both partons found by the theoretical analysis.

When all of the stable cones in an event have been identified, there will typically be some overlap between cones. This situation raises the next issue that has no obvious, nontrivial analogue in the perturbative calculation (at least at low order) and that was not envisioned in the Snowmass Algorithm. It is addressed with a splitting/merging routine in the experimental jet algorithms. In practice the splitting/merging routine involves the definition of a parameter f_{merge}, typically with values in the range $0.5 \leq f_{merge} \leq 0.75$, such that, if the overlapping transverse energy fraction (the transverse energy in the overlap region divided by the smaller of the total energies in the two overlapping cones) is greater than f_{merge}, the two cones are merged to make a single jet. If this constraint is not met, the calorimeter towers/hadrons in the overlap region are individually assigned to the cone whose center is closer. This situation yields 2 final jets. As the initial seeding/preclustering process and the final merging/splitting process were not foreseen in the Snowmass Algorithm, it is perhaps not surprising that the detailed specification of these steps were somewhat different for CDF and DØ in Run I (*e.g.*, CDF used $f_{merge} = 0.75$, while DØ used $f_{merge} = 0.5$). Note that the trivial application of the merging routine to the NLO perturbative analysis explains why the theoretical algorithm is defined to always keep the jet containing two partons out to a separation of $2R$, whenever such a solution exists. In such a configuration the 2 possible jets that are composed 1 parton each exhibit 100% overlap with the 2-parton jet. Thus by the usual merging routines, the single parton jets should always be merged with the 2-parton jet.

The final point that distinguishes the theoretical algorithm from the experimental algorithms is specific to the JetClu algorithm employed by the CDF Collaboration. The JetClu algorithm exhibits a not widely documented and troubling feature called "ratcheting". All seed towers initially included in a "seed cone" (the first cone around a seed) remain associated with that cone (*i.e.*, on its list) even as the center of the cone migrates during the search for a stable solution to Eq. 1. This remains true, by construction, even if the center moves further than R away. This ratcheting feature leads to several results not mimicked in the theoretical algorithm, or in other experiments. It leads to final "stable cones" that are often not cone-shaped (although this issue already arises from splitting/merging). Ratcheting also implies that all of the initial seed towers are included in (at least one of) the final jets. More importantly it means that cones tend to become stable after only a short migration, but in a fashion dependent on the details of the showering/hadronization process (and the initial pre-clustering process). This property argues against using JetClu for precision comparisons with theory.

The essential challenge in the use of jet algorithms is to understand the quantitative implications of the differences between the algorithms as applied by different experiments and as applied in theoretical calculations. This is the only way to control the uncertainties. It is this goal that is the primary concern of this paper. Clearly differences that lead to substantial uncertainties ($> 1\%$) should be eliminated, or well enough understood that they can be corrected for. Consider first the better known issues raised above. As suggested earlier the use of seeds in the experimental algorithms means that certain configurations kept by the theoretical algorithm are likely to be missed by the experimental one. In particular the symmetric configuration of two energetic partons nearly $2R$ apart can satisfy the strict application of Eq. 1 to form a single jet in a perturbative calculation, with no energy at the center of the jet. Yet the same configuration, including the effects of showering and hadronization, will likely lead to two well separated seeds (with only limited E_T in between) and thus two jets in the experimental algorithm, which may not be subsequently merged. When this point was recognized, the theoretical parameter R_{sep}[7] was introduced in the perturbative calculation to mimic this feature of the experimental analysis. With this parameter included two partons with angular separation

$$\Delta = \sqrt{(\phi_k - \phi_J)^2 + (\eta_k - \eta_J)^2} > R_{sep} \cdot R \leq 2R \qquad (2)$$

are not allowed to comprise a single jet, even if they satisfy Eq. 1. Comparison of the theory to data[7] suggested that the value $R_{sep} = 1.3$ provided the best description of the data at NLO. In completely independent analyses[8] the CDF and DØ Collaborations also found that two initially independent jets (selected, for example, from different events) had to placed within $\Delta /R \simeq 1.3$ of each other in a simulated event in order to be identified as a single jet by the experimental algorithms. Thus the theoretical parameter R_{sep} is presumably

also providing some level of simulation of the splitting/merging procedure. One could also imagine explicitly including seeds in the perturbative calculations at higher orders, *e.g.*, requiring the presence of a soft gluon between the two partons just mentioned. However, this explicit seed definition introduces a highly undesirable (logarithmic) dependence on the seed E_T cut (the minimum E_T required to be treated as a seed cell)[9]. The kinematic configurations corresponding to this scenario are suggested in Fig. 1, where the two energetic partons
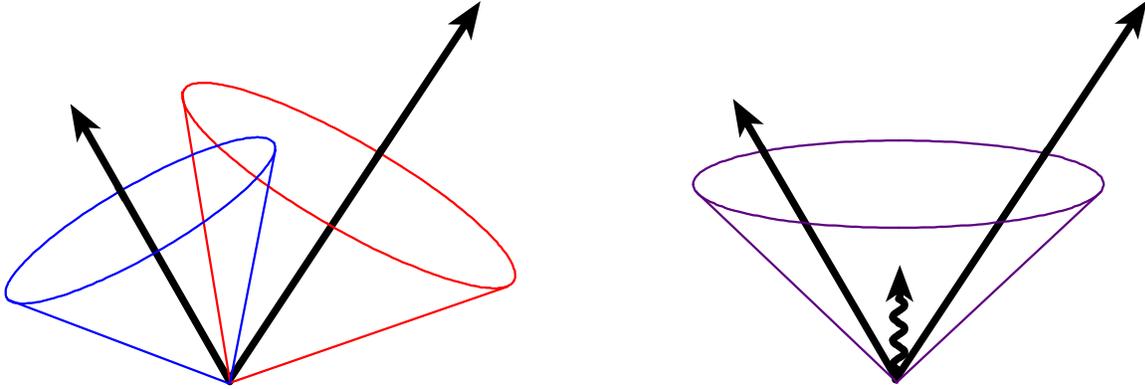
FIG. 1: Two partons in two cones or in one cone with a (soft) seed present.

could be in either 2 cones (LHS) or a single cone seeded by a soft gluon (RHS). Of course, the most desirable approach is redefine the experimental algorithms in order to remove the need for the R_{sep} parameter as will be described in more detail below.

Concerning the issue of ratcheting in the JetClu algorithm, we note that there is no reasonable way to simulate the role of ratcheting in the theoretical calculations, since its role depends in detail on the level of secondary (soft) radiation present in the event. As suggested above, this point should be interpreted as a strong argument against the continued use of JetClu. For reliable comparisons of theory and experiment it is essential that we eliminate such differences between the theoretical and experimental algorithms.

III. RUN II IMPROVEMENTS

With the advent of a new, higher energy and higher luminosity, Run II at the Tevatron, theorists and experimentalists have been offered another opportunity to construct algorithms whose application to perturbation theory and to data analysis will be more in parallel. This will in turn allow greater precision in comparing theory and experiment. In particular the Run II Workshop Proceedings[1] recommend that the CDF and DØ Collaborations employ, as much as possible, identical jet algorithms, including the use of 4-vector kinematics (4-vector instead of scalar summation and $y = 0.5 * \ln\left[(E + p_z)/(E - p_z)\right]$ instead of $\eta = \ln\left(\cot\theta/2\right)$) and a common splitting/merging procedure. The Workshop Proceedings also provide algorithms developed specifically to address the issue of seeds, in particular the Midpoint Algorithm and the Seedless Algorithm. With the Midpoint Algorithm one proceeds initially as in the original algorithms, *i.e.*, starting with seeds. Using the stable cones thus identified, one explicitly tests for the possibility of a new stable cone centered (in η, ϕ or y, ϕ) between two stable cones already found by placing a new seed midway between all pairs of stable cones separated by less than $2R$. The Midpoint algorithm is included in the analysis discussed below.

The central innovation of the Seedless Algorithm is to place an initial trial cone at every point on a randomly located but regular lattice in (y, ϕ). For maximum coverage the lattice should be approximately as fine-grained as the detector. It is not so much that this algorithm lacks seeds, but rather that the algorithm puts seed cones "everywhere" of interest. The Seedless Algorithm can be streamlined by imposing the constraint that a given trial cone is removed from the analysis if the center of the cone migrates outside of its original lattice cell during the iteration process. The streamlined version still samples every lattice cell for stable cone locations, but is less computationally intensive. Our experience with the streamlined version of this algorithm suggests that there can be problems finding stable cones with centers located very close to the cell boundaries of the lattice. This technical difficulty is easily addressed by enlarging the distance that a trial cone must migrate before being discarded. For example, if this distance is 60% of the lattice cell width instead of the default value of 50%, the problem essentially disappears with only a tiny impact on the required time for analysis. While this algorithm is less familiar than the Run I algorithms and apparently is more computer intensive than

the Midpoint Algorithm, we still advocate its use. It is inherently more similar in structure to the theoretical algorithm.

IV. NEW ISSUES

The rest of this paper is devoted primarily to an analysis of the Midpoint algorithm, especially as compared to the JetClu algorithm of CDF and the theoretical algorithm, with an eye towards identifying issues that could limit the precision of the comparison with theory. In particular, we have found that there exist previously unrecognized final state configurations that are likely to be missed in the data, compared to the theoretical result. We describe the application of representative jet algorithms to data sets that were generated with the HERWIG 6.1 Monte Carlo[10] and then processed through a CDF detector simulation. We include results from the JetClu algorithm both with and without the ratcheting feature described above. Following the recommendation of the Run II Workshop, we use 4-vector kinematics for the Midpoint Algorithm and place the cone at the midpoint in (y, ϕ), where y is the true rapidity. In the JetClu Algorithm the value $f_{merge} = 0.75$ was used (as in the Run I analyses), while for the Midpoint and Seedless Algorithms the value $f_{merge} = 0.5$ was used, as suggested in the Workshop Proceedings[1]. Finally, for completeness, we also include in our analysis a sample k_T Algorithm.

Starting with a sample of 250,000 events, which were generated with HERWIG 6.1 and processed through a CDF detector simulation and which were required to have at least 1 initial parton with $E_T > 200$ GeV, we applied the various algorithms to find jets with $R = 0.7$ in the central region ($|y| < 1$). We then identified the corresponding jets from each algorithm by finding jet centers differing by $\Delta R < 0.1$. The plots in Fig. 2 indicate the average difference in E_T for these jets as a function of the jet E_T. From these results we can draw several conclusions. First, the k_T Algorithm identifies jets with E_T values similar to those found by JetClu, finding slightly more energetic jets at small E_T and somewhat less energetic jets at large E_T. We will not discuss this algorithm further here except to note that DØ has applied it in a study of Run I data[12] and in that analysis the k_T Algorithm jets seems to exhibit somewhat *larger* E_T than expected from NLO perturbation theory. The cone algorithms, including the JetClu Algorithm without ratcheting, which is labeled JetCluNR, identify jets with systematically *smaller* E_T values (by approximately 0.5% to 1 %) than those identified by the JetClu Algorithm (with ratcheting). Due to the rapid falloff of the jet cross section with the variable E_T, this systematic displacement in E_T results in a corresponding systematic reduction of approximately 5% in the jet cross section at a given E_T value. This sort of systematic variation falls well within our definition of an important effect. We believe that this systematic shortfall can be understood as resulting from the smearing effects of perturbative showering and non-perturbative hadronization, which have not previously been considered.

V. A SIMPLE MODEL

To provide insight into the issues raised by Fig. 2 we now discuss a simple, but informative analytic picture. It will serve to illustrate the impact of showering and hadronization on the operation of jet algorithms. We consider the scalar potential $V(\overrightarrow{r})$ defined as a function of the 2-dimensional variable $\overrightarrow{r} = (y, \phi)$ by the integral over the transverse energy distribution of either the partons or the hadrons/calorimeter towers in the final state with the indicated weight function,

$$V(\overrightarrow{r}) = -\frac{1}{2} \int d^2\rho \times \left(R^2 - (\overrightarrow{\rho} - \overrightarrow{r})^2\right) \times \Theta\left(R^2 - (\overrightarrow{\rho} - \overrightarrow{r})^2\right) \times E_T(\overrightarrow{\rho}) \tag{3}$$

$$= -\frac{1}{2} \sum_i E_{T,i} \times \left(R^2 - (\overrightarrow{\rho_i} - \overrightarrow{r})^2\right) \times \Theta\left(R^2 - (\overrightarrow{\rho_i} - \overrightarrow{r})^2\right).$$

The second expression arises from replacing the continuous energy distribution $E_T(\overrightarrow{\rho})$ with a discrete set, $i = 1$ to N, of delta functions, representing the contributions of either a discrete configuration of partons or a set of calorimeter towers (and hadrons). Each parton direction or the location of the center of each calorimeter tower is defined in y, ϕ by the 2-D vector $\overrightarrow{\rho_i} = (y_i, \phi_i)$, while the parton/calorimeter cell has a transverse energy (or E_T) content given by $E_{T,i}$. This function is clearly related to the energy in a cone of size R containing the towers whose centers lie within a circle of radius R around the point $\overrightarrow{r}$ as defined by the Θ function. More importantly it carries information about the locations of "stable" cones. The points of equality between the E_T weighted centroid and the geometric center of the cone (*i.e.*, the stable cones) correspond precisely to the
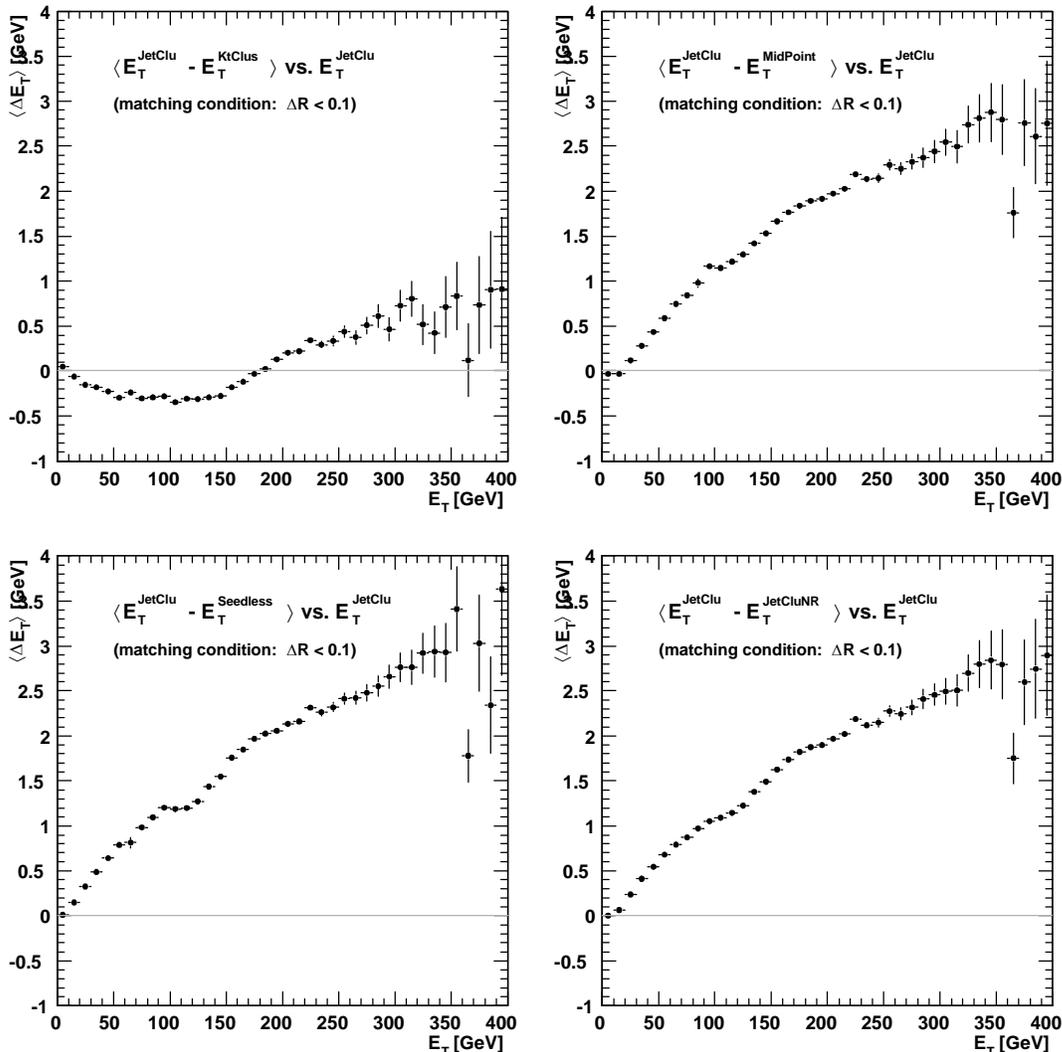
FIG. 2: Comparison of E_T in matched jets identified by various jet algorithms.

minima of V, as in familiar mechanics problems. The corresponding force or gradient of this function has the form (note that the delta function arising from the derivative of the theta function cannot contribute as it is multiplied by a factor equal to its argument)

$$\overrightarrow{F}\left(\overrightarrow{r}\right) = -\overrightarrow{\nabla}V\left(\overrightarrow{r}\right) = \sum_i E_{T,i} \times \left(\overrightarrow{\rho_i} - \overrightarrow{r}\right) \times \Theta\left(R^2 - \left(\overrightarrow{\rho_i} - \overrightarrow{r}\right)^2\right). \tag{4}$$

This expression vanishes at points where the weighted centroid coincides with the geometric center, *i.e.*, at points of stability (and at maxima of V, points of extreme instability). One can think of this "force" as driving the flow of cone centers during the iteration process described earlier. The corresponding expression for the energy in the cone centered at $\overrightarrow{r}$ is

$$E_C\left(\overrightarrow{r}\right) = \sum_i E_{T,i} \times \Theta\left(R^2 - \left(\overrightarrow{\rho_i} - \overrightarrow{r}\right)^2\right). \tag{5}$$

Since the primary underlying issue is under what conditions two energetic partons are associated with the same final jet , we can develop our understanding of the above equations by considering a scenario (containing all of the interesting effects) involving 2 partons separated by $|\overrightarrow{\rho_2} - \overrightarrow{\rho_1}| = d$. It is sufficient to specify the
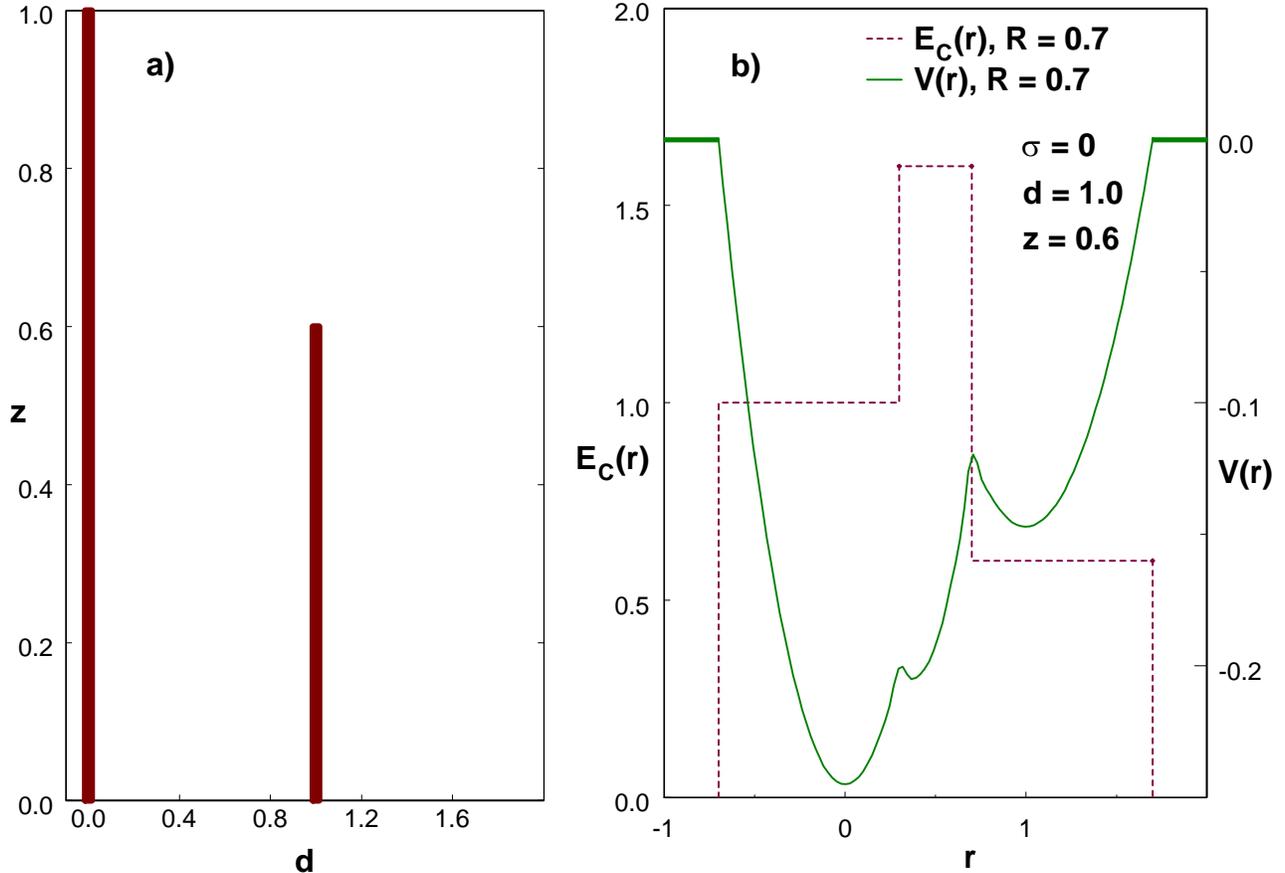
FIG. 3: 2-Parton distribution: a) transverse energy distribution; b) distributions $V(r)$ and $E_C(r)$ in the perturbative limit of no smearing.

energies of the 2 partons simply by their ratio, $z = E_2/E_1 \leq 1$. Now we can study what sorts of 2 parton configurations yield stable cones in the 2-D phase space specified by $0 \leq z \leq 1$, $0 \leq d \leq 2R$ (beyond $2R$ the 2 partons are surely in different cones). As a specific example consider the case $\rho_1 = 0$, $\rho_2 = d = 1.0$ and $z = 0.6$ with $R = 0.7$ (a typical experimental value). The underlying (partonic) energy distribution is illustrated in Fig. 3a, representing a delta function at $\rho = 0$ (with scaled weight 1) and another at $\rho = 1.0$ (with scaled weight 0.7). This simple distribution leads to the functions $V(r)$ and $E_C(r)$ indicated in Fig. 3b, where the 1-D variable r is the distance along the direction between the partons, $r = \overrightarrow{r} \cdot (\overrightarrow{\rho_2} - \overrightarrow{\rho_1})/d$. In going from the true energy distribution to the cone distribution $E_C(r)$ the energy is effectively smeared over a range given by R, each region having sharp boundaries. In the potential $V(r)$ the distribution is further shaped by the quadratic factor $R^2 - (\rho_i - r)^2$. We see that $V(r)$ exhibits 3 local minima corresponding to the expected stable cones around the two original partons ($r_1 = 0$ and $r_2 = 1$), plus a third stable cone in the middle ($r_3 = zd/(1+z) = 0.375$ in the current case). This middle cone includes the contributions from both partons as indicated by the magnitude (1.6) of the middle peak in the function $E_C(r)$. Note further that the middle cone is found at a location where there is initially no energy in Fig. 3a, and thus no seeds. One naively expects that such a configuration is not identified as a stable cone by the experimental implementations of the cone algorithm that use seeds simply because they do not look for it. Note also that, since both partons are entirely within the center cone($|r_1 - r_3| < R, |r_2 - r_3| < R$), the overlap fractions are unity and the usual merging/splitting routine will lead to a single jet containing all of the initial energy ($1 + z$). This is precisely how this configuration was treated in the NLO perturbative analysis of the Snowmass Algorithm[6] (*i.e.*, only the leading jet, the middle cone, was kept).
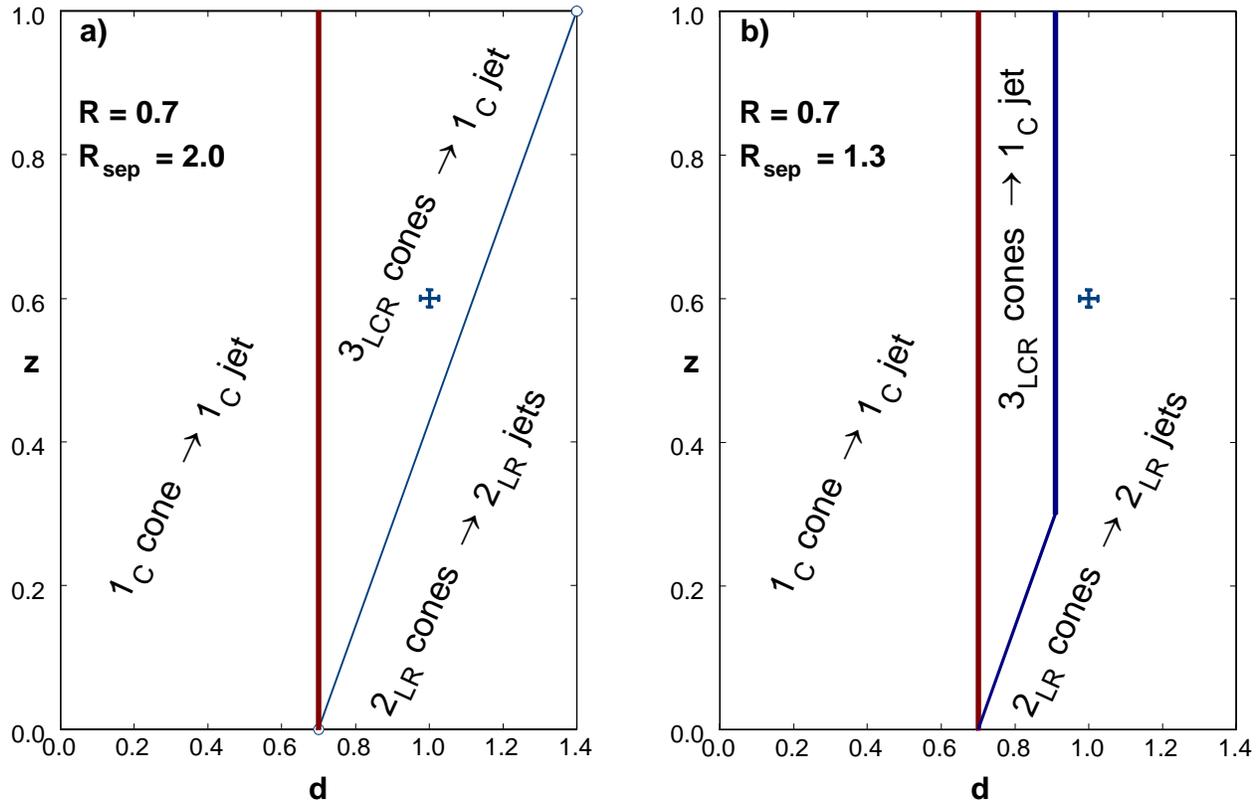
FIG. 4: Perturbation theory cone/jet structure: a) $R_{sep} = 2$; b) $R_{sep} = 1.3$. The subscripts, LCR, refer to whether the cone/jet is found at the position of the Left parton, Right parton or at a intermediate Central position. The cross indicates the configuration explored in subsequent figures.

Similar reasoning leads to Fig. 4a, which indicates the various 2 parton configurations found by the perturbative stable cone algorithm. In the 2-D domain of (d, z) there are three distinct regions. For $d < R$ and all values of z the potential $V(r)$ has just a single minimum in the central region, as indicated by the "C" subscript ($r_C = zd/(1 + z) < R$). Thus this region yields a single stable cone and a single jet containing both partons. In the triangular region $R < d < (1 + z)R$ the potential function has three minima (as above) and one finds 3 stable cones at the locations of the two partons ("L" and "R", $r_L = \rho_1 = 0$, $r_R = \rho_2 = d$) and at an intermediate central location indicated by the "C" subscript ($r_C = zd/(1 + z) < zR$). As in the specific example described above, both partons lie within this central cone ($|r_L - r_C| = r_C < zR < R$, $|r_R - r_C| = d(1 - z)/(1 + z) < (1 - z)R < R$) and by the usual merging rules (the overlap is 100%) all three cones are merged to a single central jet (at r_C), again with all of the energy. For $d > (1 + z)R > R$ the potential exhibits two well separated minima at $r_L = 0$ and $r_R = d > R$ with no overlap. Thus in this region we find 2 stable cones and 2 jets, each containing one parton, of scaled energies 1 and z. Thus, except in the far right region of the graph, the 2 partons are (in perturbation theory) always merged to form a single jet. In particular, the parton configuration of Fig. 3a, which is indicated by the cross symbol in Fig. 4a, yields 3 stable cones and a single final jet encompassing both partons.

As suggested above we expect that the use of seeds in experimental algorithms will play a role in the middle, 3_{LCR}, region, but not elsewhere. While the potential $V(r)$ in perturbation theory has three minima in this region, we may not find all three of the corresponding stable cones if we only look for stable cones near the seeds. Looking back at Fig. 3a we expect to have "seeds" near ρ_1 and ρ_2 but presumably not in the central region between the two partons where there is no energy. Trial cones near ρ_1 will migrate to the L stable

position under iteration while a similar flow will occur for seeds near ρ_2 to the R stable cone. With no seeds in between there may be no trial cones flowing to the C stable cone. The *ad hoc* parameter R_{sep} was introduced to provide a (crude) simulation of this situation (see further discussion below) in the NLO calculations[7]. The parameter R_{sep} is defined such that stable cones containing 2 partons are not allowed for partons separated by $d > R_{sep} \times R$. As a result cones are no longer merged in this kinematic region. Thus we are modeling the effect of seeds by the replacement 3_{LCR} cones $\longrightarrow 2_{LR}$ cones in this region. In the present language this situation is illustrated in Fig. 4b corresponding to $R_{sep} = 1.3$, $R \times R_{sep} = 0.91$. This specific value for R_{sep} was chosen[7] to yield reasonable agreement with the Run I data. The conversion of much the 3 cones $\to 1$ jet region ($d > 0.91$, $z > (d - R)/R$) to 2 cones $\to$ 2 jets has the impact of lowering the average E_T of the leading jet and hence the jet cross section at a fixed $E_{T,J}$. Parton configurations that naively produced jets with energy characterized by $1 + z$ now correspond to jets of maximum energy 1. However, the large z, large d part of phase space is far from the collinear and soft singularities of the perturbation theory and does not make a major contribution to the cross section. Thus the resulting reduction in the jet cross section in going from $R_{sep} = 2.0$ to $R_{sep} = 1.3$ is only approximately 5.2% at $E_T = 100$ GeV and 3.7% at $E_T = 400$ GeV. This level of precision was not essential for Run I, but is relevant to the present discussion. This magnitude of change in the jet cross section corresponds approximately to a 1% decrease in the average jet E_T consistent with the level of mismatch exhibited in Fig. 2. Note that with this value of R_{sep} the specific parton configuration in Fig. 3a, which is indicated by the cross symbol in Fig. 4b, will now yield 2 jets (and not 1 merged jet) in the theoretical calculation.

As mentioned earlier, it has been recommended that this issue with seeds in the 3_{LCR} region be addressed in Run II via the Midpoint and Seedless Algorithms. The expectation was that both would succeed in identifying the missing stable cone at r_C. However, as indicated in Fig. 2, neither of these two algorithms reproduces the results of JetClu. Further, they both identify jets that are similar to JetClu *without* ratcheting. Thus we expect that there is more to this story.

As suggested earlier, a major difference between the perturbative level, with a small number of partons, and the experimental level of multiple hadrons is the smearing that results from perturbative showering and nonperturbative hadronization. For the present discussion the primary impact is that the starting energy distribution will be smeared out in the variable $\vec{r}$. We can simulate this effect in our simple model using Gaussian smearing, *i.e.*, we replace the delta functions in Eq. 3 with 2-D Gaussians of width $\sigma (\propto e^{-\left((\vec{r} - \vec{\rho_i})^2/\sigma^2\right)})$. (Since this corresponds to smearing in an angular variable, we would expect σ to be a decreasing function of E_T, *i.e.*, more energetic jets are narrower. We also note that this naive picture does not include the expected color coherence in the products of the showering/hadronization process, nor its stochastic nature.) The first impact of the simulated smearing is that some of the energy initially associated with the partons now lies outside of the cones centered on the partons. This effect, typically referred to as "splashout" in the literature, is (exponentially) small in this model for $\sigma < R$. Here we will focus on less well known but phenomenologically more relevant effects of the smearing. The distributions corresponding to Fig. 3b with $d = 1.0$ and $z = 0.6$, but now with $\sigma = 0.10$ (instead of $\sigma = 0$), are exhibited in Fig. 5a. With the initial energy distribution smeared by σ, the potential $V(r)$ is now even more smeared and, in fact, we see that the middle stable cone (the minimum in the middle of Fig. 3b) has been washed out by the increased smearing. Thus the cone algorithm applied to data (where such smearing is present) may not find the middle cone that is present in perturbation theory, not only due to the use of seeds but also due to this new variety of smearing correction, which renders this cone unstable. Note that this conclusion obtains even though there is a clear maximum in $E_C(r)$ at the location where the middle stable cone used to be. Since, as a result of this smearing correction, the middle cone is not stable, this problem is *not* addressed by either the Midpoint Algorithm or the Seedless Algorithm. Both algorithms look in the correct place, but they look for stable cones, which may not present in real data. This point is presumably part of the explanation for why both of these algorithms disagree with the JetClu results in Fig. 2.

Our studies also suggest a further impact of the smearing of showering/hadronization that was previously unappreciated. This new effect is illustrated in Fig. 5b, which shows $V(r)$, still for $z = 0.6$ and $d = 1.0$, but now for $\sigma = 0.25$. With the increased smearing the second stable cone, corresponding to the second parton, has now also been washed out, *i.e.*, the right hand local minimum has also disappeared. While the effect of the smearing is intuitively reasonable, *i.e.*, the distributions become increasingly smoother, the resulting loss of stable cones was previously unappreciated. To help make this scenario more visual Figs. 6 to 8 exhibit the 2-D structure inherent in Eq. 3 in terms of contour plots of the potential $V(\vec{r})$. The darker shades indicate lower values of potential, while the white dots indicate the location of the initial partons. In each figure the white circle(s) indicates the location of the final jet(s). The smooth evolution of the potential as the smearing increases is clear. With no smearing symmetrical minima around each parton are apparent along with the deeper minimum in the center. This figure also makes the point that, even when the central minimum is present, it may serve as an attractor for only a small area in the (y, ϕ) plane. With some smearing as in Fig. 10 the central minimum has evolved into just a lobe on the left-hand minimum and we find just two final jets.
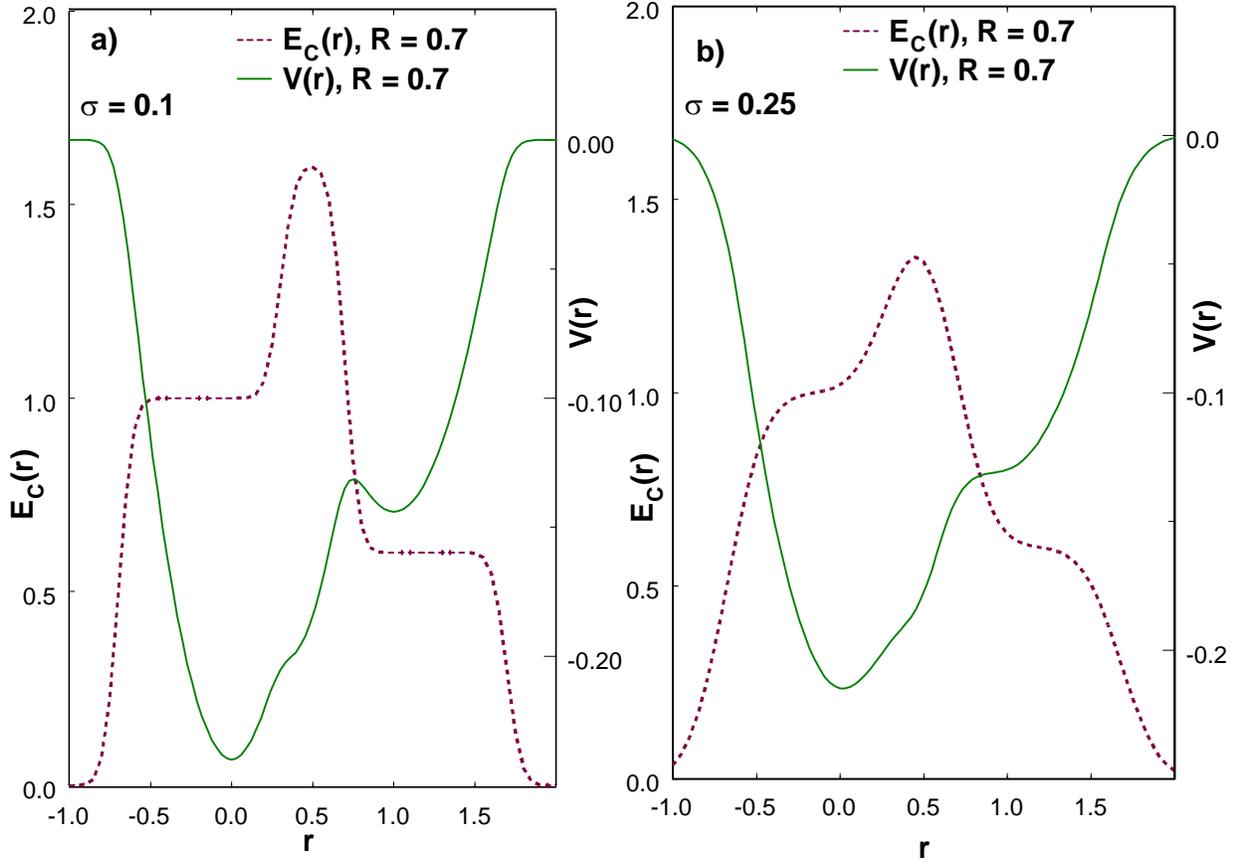
FIG. 5: The distributions $V(r)$ and $E_C(r)$ *with* $d = 1.0$ and $z = 0.6$ for smearing width $\sigma =$ a) 0.1; b) 0.25.
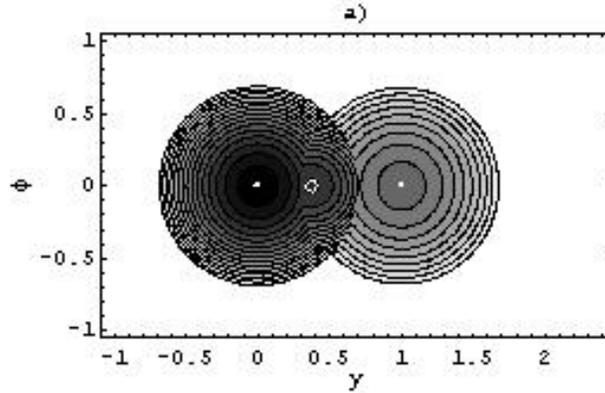
FIG. 6: Contour plot of the potential $V(r)$ in 2-D for $z = 0.6$ and $d = 1$ in the perturbative limit ($\sigma = 0$).

With further smearing as in Fig. 11 the lobe extends to the right-hand parton and only one minimum and one jet remain.

This same, large smearing, situation is exhibited in the case of Monte Carlo "data" by the lego plot in Fig. 9 indicating the jets found by the Midpoint Algorithm in a specific Monte Carlo event. The Midpoint Algorithm does not identify the energetic towers (shaded in black) to the right of the (more lightly shaded) energetic identified central jet as either part of that jet or as a separate jet, *i.e.*, these obviously relevant towers are not
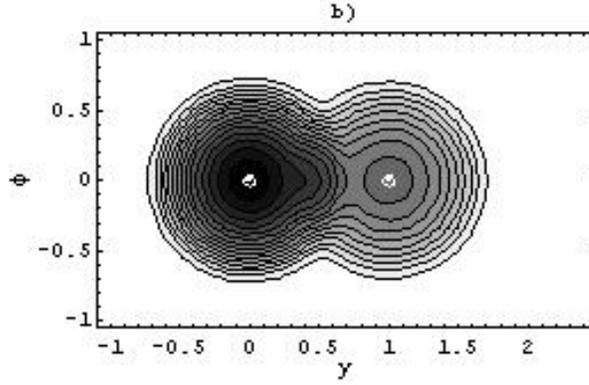
FIG. 7: Contour plot of the potential $V(r)$ in 2-D with $z = 0.6$ and $d = 1$ for $\sigma = 0.1$.
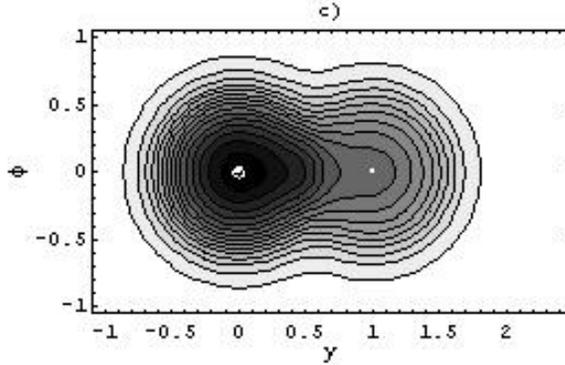
FIG. 8: Contour plot of the potential $V(r)$ in 2-D with $z = 0.6$ and $d = 1$ for $\sigma = 0.25$.

found to be in a stable cone. The iteration of any cone containing these towers invariably migrates to the nearby higher E_T towers.

We can further describe the impact of these smearing effects by looking again at the stable cone content in the (d, z) plane as we did in Fig. 4. The resulting structure is indicated in Figs. 10 and 11. The special case $(d = 1.0, z = 0.6)$ that we have already discussed is again marked in the figures by the cross. Comparing to the perturbative result in Fig. 4b (the three regions of the previous graph with $R_{sep} = 1.3$ are indicated by the dashed-dotted line as a reference) we note several features. The small separation region, $d < R$, remains unchanged. There is still a single stable cone intermediate in position between the 2 partons, which will yield a single central jet with both partons in the jet. The bulk of the region for $R < d < R_{sep} * R$ has changed from yielding 3 stable cones to yielding just a single central stable cone. This will yield a single central jet with both partons included, This conclusion is presumably independent of the seed issue as trial cones seeded near ρ_1 and ρ_2 will now flow to the central stable cone and not be stopped by the now absent LR stable cones as was expected in the previous discussion without smearing. Thus the jet structure in this region is largely unchanged from the unsmeared case perturbative including the R_{sep} parameter. On the fine details, $e.g.$, the specific location of the boundary, depends on the value of σ. Note that this effect of smearing in this region is consistent with our use of the R_{sep} parameter ($i.e.$, we assumed that the central jet was identified in this region even if seeds were used). The bulk of the region of larger separation, $d > R_{sep} * R$, now yields two stable cones, each one containing a single parton and located essentially at that parton, as indicated by the "LR" subscript. The example, $d = 1$, $z = 0.6$, studied above falls in this region for the smaller smearing. For the smaller smearing case of Fig. 10 there is still a discernible region where 3 stable cones are identified. Depending on the role of seeds this may or may not yield a central jet. Based on the discussion above, we expect the use of seeds in the experimental algorithms to mean that in most of this region only two jets, "L" and "R" for each parton, are found and there is no central combined jet (as suggested by the R_{sep} boundary). There is also a small region where only the left-hand and central cones are still stable, yielding essentially the original central jet with both partons. Even with seeds invoked, a trial cone seeded by the right-hand parton
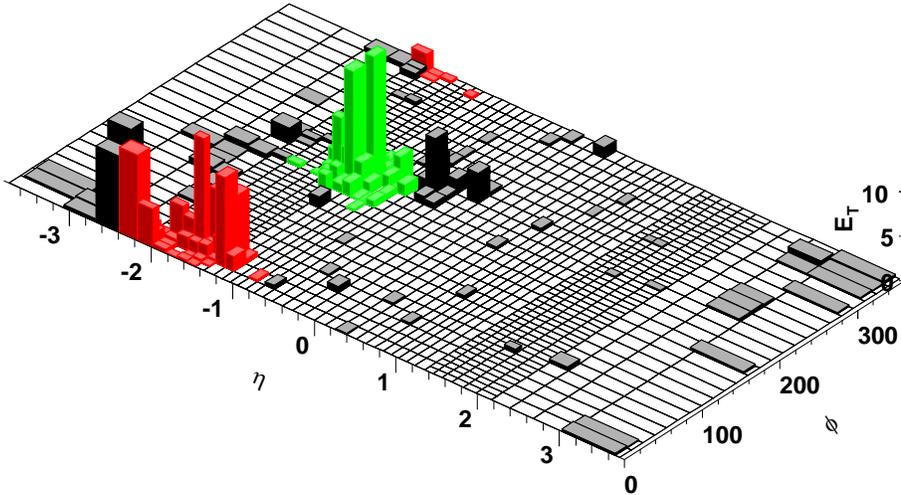
FIG. 9: Result of applying the Midpoint Algorithm to a specific Monte Carlo event in the CDF detector.

should flow to this central stable cone. Both of these regions (2_{LC} and 3_{LCR}) have essentially disappeared with the larger smearing case of Fig. 11. The other new configuration is a single stable cone near the location of the left-hand parton, with both the central and right-hand cones washed out by the smearing. The dashed boundary between the 1_{C} and 1_{L} regions is meant to indicate that this is a "soft" transition with the location of the single stable cone moving smoothly leftward as d increases. As discussed earlier, our special case (the cross) falls into this region for the larger smearing (recall that there is only the L stable cone in Fig. 8). We expect that an effective smearing intermediate between $\sigma = 0.1$ and $\sigma = 0.25$ is the most "realistic" value for this parameter in the sense of most closely matching the effects of "real" showering and hadronization for the purposes of the present discussion. Thus Figs. 10 and 11 suggest that the *ad hoc* parameter $R_{sep} = 1.3$ affords a reasonable description of how the cone algorithm works in practice and offering at least a partial explanation of why it produced perturbative results similar to the data. Based on the current analysis, however, its role is more to simulate the effects of smearing and less the effect of seeds. The efficacy of the parameter R_{sep} is perhaps best highlighted by reinterpreting the stable cones of Figs. 10 and 11 in terms of identified jets, assuming seeds act as described above ($3_{\text{LCR}} \longrightarrow 2_{\text{LR}}$, $2_{\text{LC}} \longrightarrow 1_{\text{C}}$), as illustrated in Fig. 12. The point is to notice the similarity between Fig. 12 and the perturbative result, with R_{sep}, in Fig. 4. This similarity is made even stronger when we note that, in terms of the leading jet, 1_{L} is essentially equivalent to 2_{LR} at least for small z values where the R jet has much lower energy. It is worthwhile mentioning that "merging" plays essentially no role here. In the 2_{LR} region the two cones are too well separated to be merged, even if $f_{merge} = 0.5$. This will not be the case in the next section.

VI. RATCHETING

Since our goal is to understand the detailed differences between theoretical and experimental applications of cone jet algorithms, we must say something about the role of the previously mentioned ratcheting effect in the CDF JETCLU jet algorithm. This feature of the algorithm ensures that as the trial cone migrates away from the initial seed position, any calorimeter tower included in the initial trial cone stays with trial "cone" ("once in a cone, always in a cone"), which is therefore no longer cone shaped in the (y, ϕ) plane. Thus with ratcheting turned on the "history" of a migrating trial cone matters, an effect with no analogue in the perturbative calculations. Ratcheting means that the lists of towers to be summed over in Eqs. 1 and 3 grow with migration and a tower k no longer needs to be within R of the cone center to be included. In the current analysis we can define the integral (the potential) over the smeared E_T profiles to take this ratcheting into account, with two distinct potentials depending on whether we are "seeded" (by assumption) by the left-most
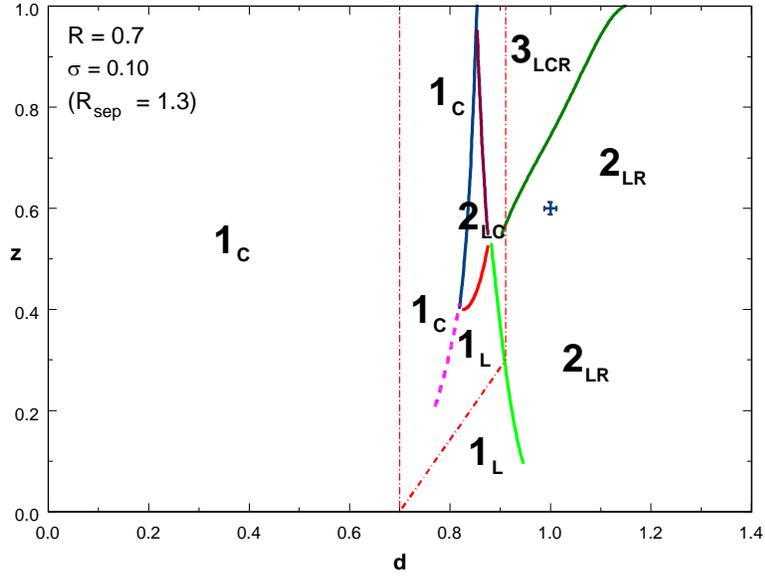
FIG. 10: Stable cone structure found by the cone jet algorithm with underlying partons including gaussian smearing with width $\sigma = 0.1$. The dashed-dotted lines indicated the unsmeared, perturbative result with $R_{sep} = 1.3$.
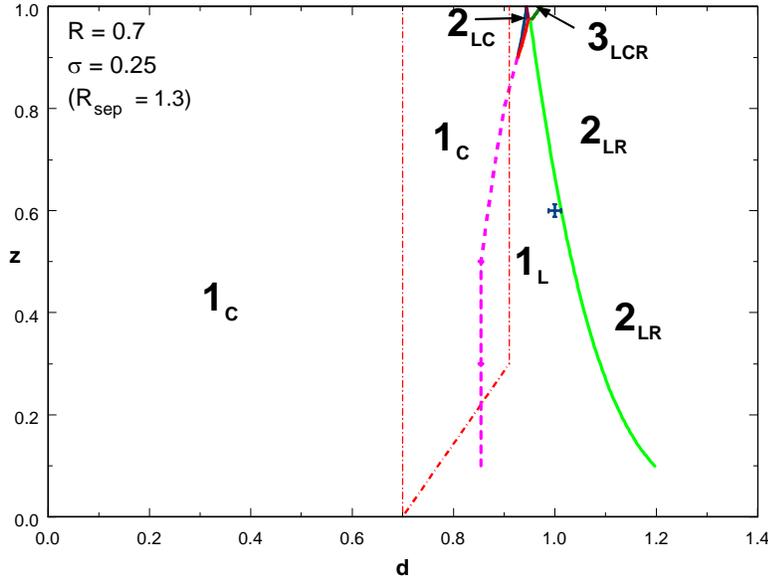
FIG. 11: Stable cone structure found by the cone jet algorithm with underlying partons including gaussian smearing with width $\sigma = 0.25$. The dashed-dotted lines indicated the unsmeared, perturbative result with $R_{sep} = 1.3$.
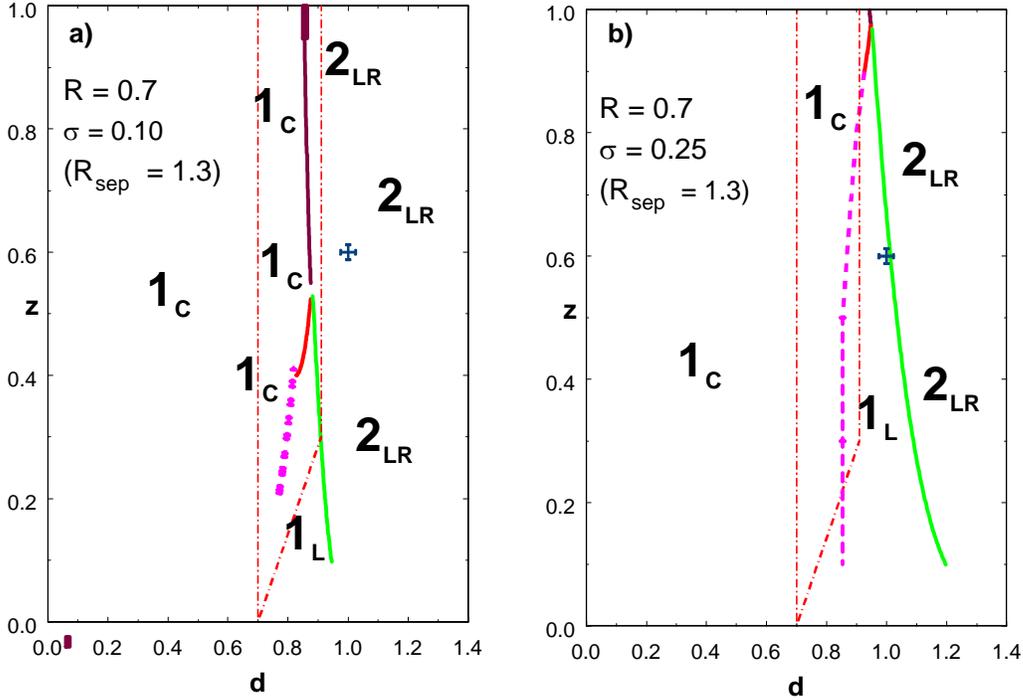
FIG. 12: Identified jets structure in (d, z) plane including smearing and seeds: a) $\sigma = 0.1$, b) $\sigma = 0.25$.

or right-most parton. The stable cones are then the sum of the stable cones found from both potentials. In general, the central stable cones identified by the two potentials (if present) are nearly coincident. On the other hand, the right-hand potential will never exhibit a left-hand stable cone. Due to ratcheting, the trial cone cannot migrate that far. Likewise the left-hand potential will never exhibit a right-hand stable cone. With enough smearing present the left-hand potential (based on the larger energy left-hand parton) will tend to have only a left-hand stable cone with its central stable cone washed out by the smearing. In contrast the right-hand potential (based on the lower energy right-hand parton) will still exhibit a central stable cone even when the right-hand stable cone is lost due to smearing. The analogues of the stable cone structures above, but now with ratcheting simulated as just described, are exhibited in Figs. 13 and 14. We see from these figures that the impact of ratcheting is to greatly expand the 2_{LC} and 3_{LCR} regions at the expense of the 1_L and 2_{LR} regions. As suggested above, the effect of ratcheting is to keep the central stable cone even when smearing is present. Of particular importance is the appearance of a large region below the line $z = (d - R)/R$ that is now of type 2_{LC}, instead of 1_L (or even 2_{LR} as it was in the perturbative calculation), where the central stable cone arises from the ratcheted cone seeded by the right-hand parton. Thus this central stable cone, due to ratcheting, will include the energy from the right-hand parton. Further these two stable cones are close enough to be merged into a single jet, 2_{LC}(stable cones)$\longrightarrow 1_C$ (jet). This implies that in this region, where the cross section is large, we will identify a merged jet (encompassing both partons) using the experimental algorithm with seeds and ratcheting on (realistic) data with smearing. Note that this is just the region corresponding to Fig. 9 where the difference between 2_{LC} with ratcheting and 1_L without ratcheting explains why the dark towers in the LEGO were not assigned to any jet. This point is further illustrated in Fig. 15, which shows the identified jet structure and is the analogue of Fig. 12. As before we have assumed that the impact of seeds is to ensure that we miss the central cone in the configuration 3_{LCR}, i.e., $3_{LCR} \longrightarrow 2_{LR}$. These figures are to be compared to each other and the perturbative result in Fig. 4. Again we see that the R_{sep} parameter provides a reasonable simulation of seed and smearing effects. However, in the region below the line $z = (d - R)/R$ mentioned above, the result with ratcheting is different in that this region now has a leading C type jet (with both partons) while the perturbative result and the non-ratcheted result has leading L type jets (with just the left-hand parton). Since this merging of the two partons has the effect of increasing the E_T of the primary
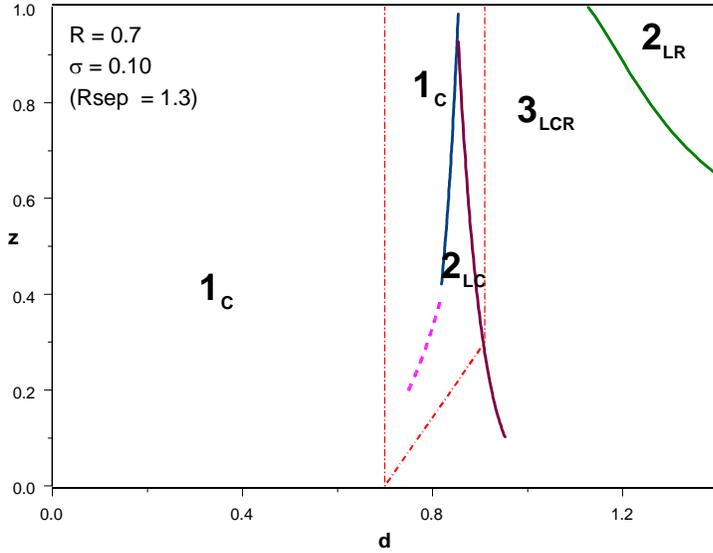
FIG. 13: Stable cone structure found by the cone jet algorithm with underlying partons including gaussian smearing with width $\sigma = 0.1$ and ratcheting. The dashed-dotted lines indicated the unsmeared, perturbative result with $R_{sep} = 1.3$.

jet in the event, we expect the MidPoint Algorithm (without ratcheting) to yield a smaller E_T on average than JETCLU, just as we saw in Fig. 2. The current analysis also suggests a similar difference from the NLO result with R_{sep}, i.e., the effects described here suggest that the inclusive jet cross section defined by JETCLU should be larger by 3-5% than that defined by NLO perturbation theory with $R_{sep} = 1.3$.

VII.   POSSIBLE IMPROVEMENT

In summary, we have found that the impact of smearing due to showering and hadronization is expected to be much more important than simply the leaking of energy out of the cone (splashout). Certain stable cone configurations, present at the perturbative level, can disappear from the analysis of real data due to the effects of showering and hadronization. This situation leads to corrections to the final jet yields that are relevant to our goal of 1% precision in the mapping between perturbation theory and experiment. Compared to the perturbative analysis of the 2-parton configuration, both the central stable cone and the stable cone centered on the lower energy parton can be washed out by smearing. Further, this situation is not addressed by either the Midpoint Algorithm or the Seedless Algorithm, both of which assume that the central stable cone is present. One possibility for addressing the missing middle cone would be to eliminate the stability requirement for the added midpoint cone in the Midpoint Algorithm. However, if there is enough smearing to eliminate also the second (lower energy) cone, even this scenario will not help, as we do not find two cones to put a third cone between. Another approach, which we will explore here, is to attempt to reduce the impact of smearing. Recall that nothing precludes us from employing two different cone sizes. We can use one during the search for the stable cones and the second during the calculation of the jet properties. Since the effective smearing in the potential $V(\vec{r})$ that describes the location of the stable cones is a combination of that due to the physical effects of showering and hadronization and that due to the size of the cone, using a smaller value of cone radius, e.g., $R' = R/2$, during the search for stable cones will reduce the smearing effects. We can still use $R = 0.7$ in the jet construction phase. With this choice of R' we find the stable cone structure illustrated in Fig. 16. We can understand this figure by noting that in first approximation the boundaries in these graphs are function of d/R. This when we decrease R by a factor of 2, the boundaries move to the left by a corresponding amount. The major change with this "fixed" algorithm is that the boundary between 1 stable cone and 2 stable cones occurs at much smaller d. There is also a new region labeled $2'_{LR'}$ separated from the usual 2_{LR} region by a "soft" boundary indicated by the dashed line (similar in spirit to the $1_C \longrightarrow 1_L$ boundary), that is intended to indicate that in this region the 2 stable cones are actually both rather central, i.e., not at the parton locations.
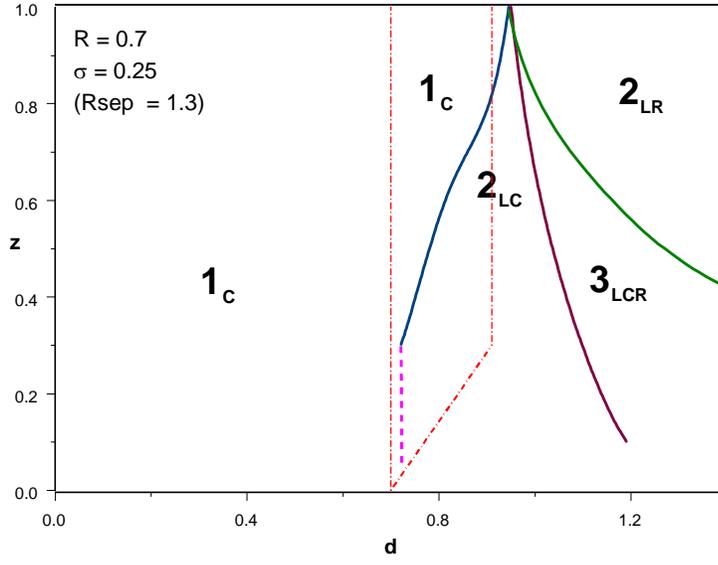
FIG. 14: Stable cone structure found by the cone jet algorithm with underlying partons including gaussian smearing with width $\sigma = 0.25$ and ratcheting. The dashed-dotted lines indicated the unsmeared, perturbative result with $R_{sep} = 1.3$.
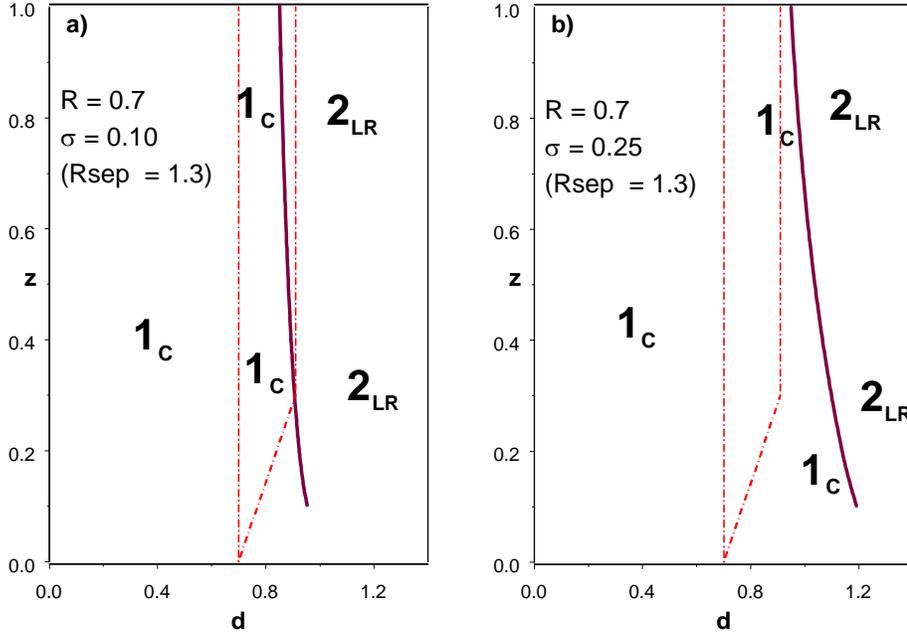
FIG. 15: Identified jets structure in (d, z) plane including smearing, seeds and *ratcheting*: a) $\sigma = 0.1$, b) $\sigma = 0.25$.
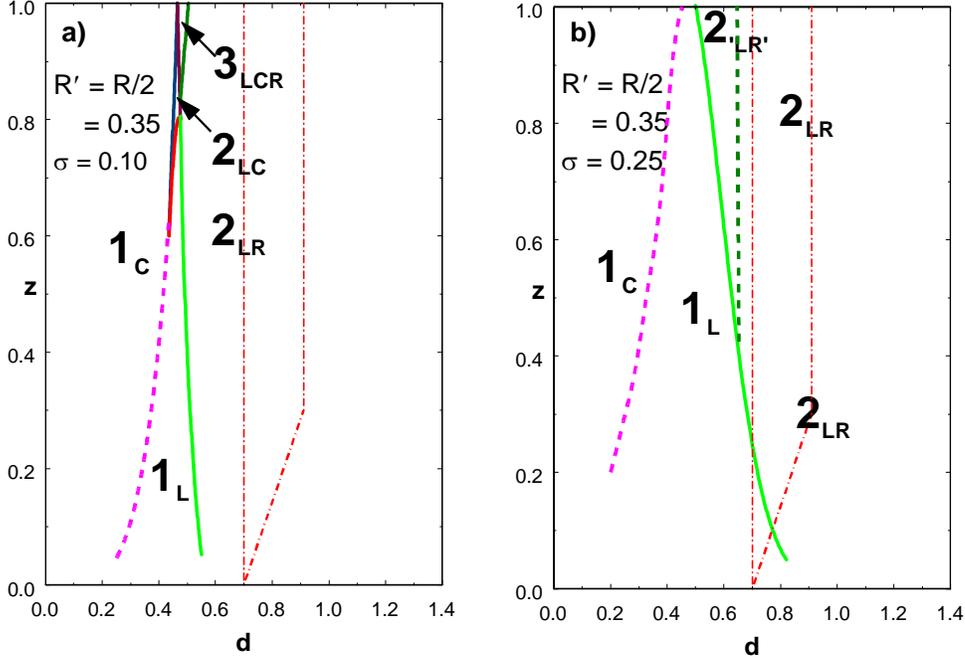
FIG. 16: Stable cone structure found using $R' = R/2$ for the two values of the smearing parameter: a) $\sigma = 0.1$, b) $\sigma = 0.25$.

This changes smoothly as d increases.

To convert these stable cones into jets we use their locations for cones of size $R = 0.7$ to determine what energy is in the jet. Unlike the situations discussed above, the merging of the cones to make a single jet is important for this analysis. Two stable cones of radius R' can be considerably closer together than those of size R. Including the splitting/merging algorithm with $f_{merge} = 0.5$ we find the jet structure exhibited in Fig. 17. To interpret this figure it is important to note for $d < R$, a jet of the type 1_L is essentially identical to one labeled 1_C in the sense that, once we have switched to the larger cone radius, essentially all of the energy from both partons is included in that jet. For practical purposes all of the single jet region corresponds to jets with both partons. The essential change from the unfixed case of Fig. 12 occurs in the region $R < d < R_{sep*R}$, $z < (d - R)/R$. Here the change from 1_L jets to 1_C jets is essential. The 1_L jets in the unfixed case contain only the left-hand parton, while the 1_C jets in the fixed case include both partons. This leads to both higher energy jets, similar to those identified by JETCLU, and ensures that there are few instances of the situation illustrated in Fig. 9 where some energetic calorimeter towers are not assigned to any jet. The improved agreement between the JetClu results and those of the Midpoint Algorithm with the last "fix" (using the smaller R' value during discovering but still requiring cones to be stable) is indicated in Fig. 18. Clearly most of the differences between the jets found by the JetClu and Midpoint Algorithms are removed in the fixed version of the latter. The small R "fix" suggested for the Midpoint Algorithm can also be employed for the Seedless Algorithm but, like the Midpoint Algorithm, it will still miss the middle (now unstable) cone.

To further illustrate the improvement offered by this fix and to confirm that the difference between JETCLU with ratcheting and the MidPoint algorithm arises primarily from the $R < d < R_{sep*R}$, $z < (d - R)/R$ region of the (d, z) plane we have performed the following exercise. Using the set of Monte Carlo events noted earlier, we have identified the jets found with JETCLU and the various versions of the MidPoint algorithm. Then we removed the calorimeter towers and particles contained in those jets from event and performed a second pass search for jets with the same algorithms. The point is that clusters of towers that are potential jets but not identified as such in the first pass, e.g., the cluster of towers in Fig. 9, may still be found as jets in this second pass when the stability of the corresponding cone is not compromised by the nearby identified jet. Then we can ask how the second pass jets are related to the nearest first pass jets in terms of their location (as pairs) in the (d, z) plane. A scatter plot of this result is shown in Fig. (I want to put Matthias's plot z_vs_d_plt_NEW_6
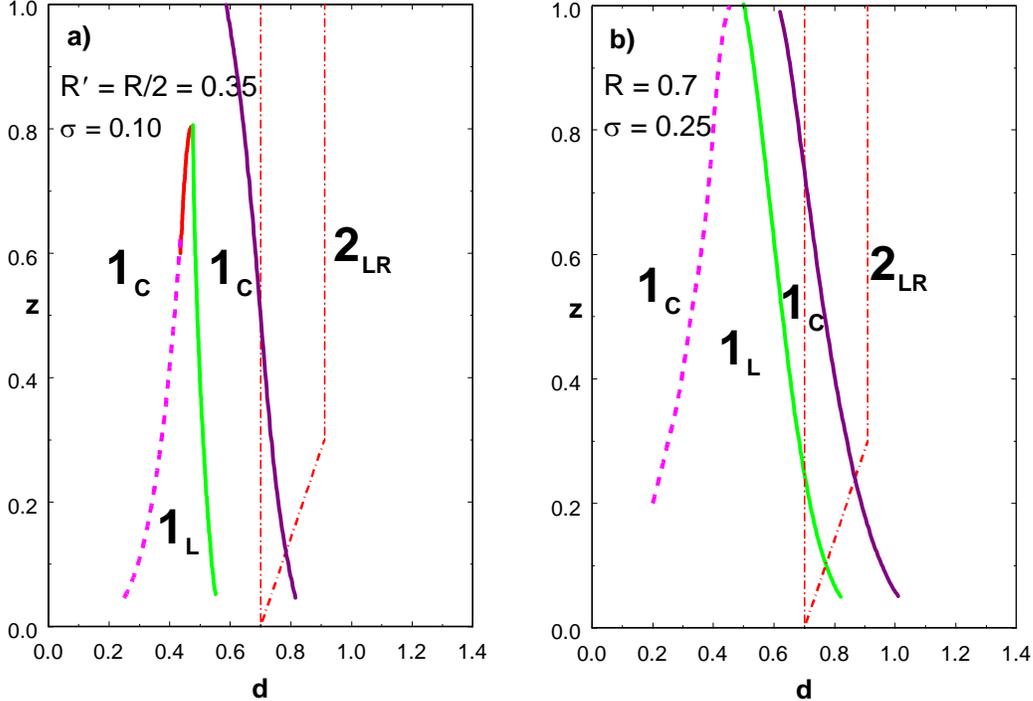
FIG. 17: Jet structure, for jets of size $R = 0.7$, found for stable cones defined by cone radius $R' = 0.35$, including smearing effects: a) $\sigma = 0.1$, b) $\sigma = 0.25$. Splitting/merging uses $f_{merge} = 0.5$.

from his webpage here but my Tex program won't accept the postscript and I have not succeeded in turning it into an acceptable EPS file - can you give me a native encapsulated version?). It is clear from these scatter plots that the difference between the JETCLU and MidPoint algorithms arise from just the configurations we have been discussing. Further, when the fix described above is applied, the second pass jets in this region largely vanish. The corresponding calorimeter towers are now included in the first pass jets just as in the JETCLU analysis.

VIII.   CONCLUSION

In this paper we have discussed various reasons that cone jet algorithms can yield differing results when applied in the context of perturbation theory and to real experimental data. The underlying goal is to identify the techniques necessary to provide precision (˜1%) comparisons between theory and experiment. We have analyzed the results of applying a range of jet algorithms, especially the JETCLU algorithm used by CDF in Run I and the MidPoint algorithm proposed for use in Run II, to a large set of Monte Carlo events. We have typically used the JETCLU result as the reference result. We have also used a simple analytic model to simulate how the algorithms work when applied to 2-parton perturbative configurations, allowing a direct connection to perturbation theory. We have focussed on the impact of including smearing in this simple model to simulate the effect of showering and hadronization in the real world. This model allows us to characterize the more complex correlations in real data to the behavior in a simple 2-D plane representing the relative energy of the partons z and the angular separation d. In this (d, z) plane the differences of interest arise in two regions as identified in Fig. In the triangular region at large z and d (Region I), the use of seeds and the effects of smearing render experimental cone jet algorithms unable to identify the 2-partons in a jet configuration characteristic of the NLO perturbative calculation. The issues with seeds has been, in fact, long appreciated and this issue was the reason for the introduction of the R_{sep} parameter in the perturbative calculations. The current studies help to
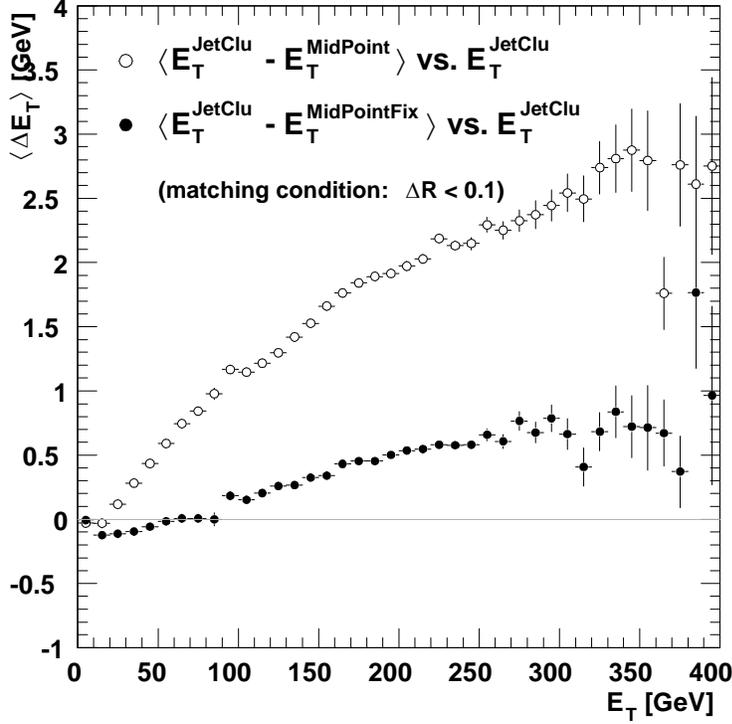
FIG. 18: Comparison of E_T in matched jets with and without the fix described in the text.
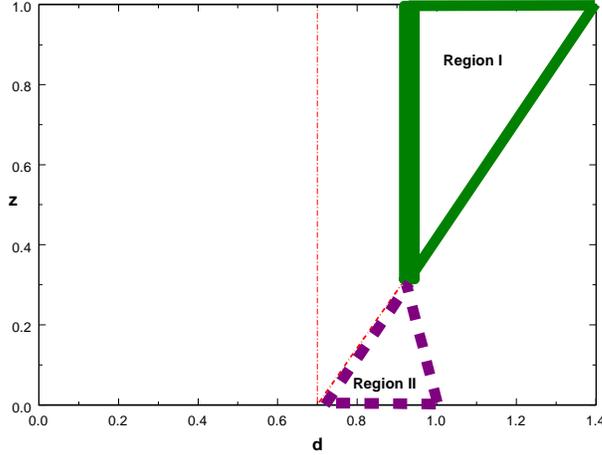
verify the efficacy of this parameter in simulating the real effects of both seeds and smearing. In any case the QCD amplitude is not large in this corner of phase space. The other region of interest corresponds to small z but larger d ($> R$) (Region II). In this region the perturbative calculation predicts 2 identified jets. On the experimental side, seeds play little role but smearing is quite important, often leading to the remnants of the lower energy parton not being identified as part of any jet by the MidPoint Algorithm. While this has little impact on the leading jet, it is troubling that this information is lost from our analysis of the event. It also turns out that JETCLU includes a feature called ratcheting that ensures that in the region of the (d, z) plane JETCLU does not find just a single jet with 1 parton (like the MidPoint algorithm), nor two distinct jets, each with a single parton (like perturbation theory) but rather both partons included in a single. Thus in this region

JETCLU jets have large energies than either the jets fond by the MidPoint algorithm or perturbation theory. By using a finer cone size to search for stable cones, the MidPoint algorithm can be brought into reasonable agreement with JETCLU. On the other hand, this is a mixed blessing. The recommendation of the current work is rather that JETCLU, with its dependence on the history of the search for stable cones, not be employed in future jet physics studies where comparison to theoretical predictions is desired. The MidPoint algorithm without the fix outlined above provides a better match to the theoretical studies. Even in this case the role of smearing in removing central stable cones remains an issue for precision comparisons of theory and experiment, and seems to require the continued use of the R_{sep} parameter.

---

[1] G. C. Blazey, *et al.*, *Run II Jet Physics: Proceedings of the Run II QCD and Weak Boson Physics Workshop*, hep-ex/0005012; P. Aurenche, *et al.*, *The QCD and Standard Model Working Group: Summary Report from Les Houches*, hep-ph/0005114.

[2] G. Arnison, *et al.* (UA1 Collaboration), Phys. Lett. **123B**, 115 (1983); M. Banner, *et al.* (UA2 Collaboration), Phys. Lett. **118B**, 203 (1982).

[3] J. E. Huth, *et al.*, in *Proceedings of Research Directions for the Decade: Snowmass 1990*, July 1990, edited by E. L. Berger (World Scientific, Singapore, 1992) p. 134.

[4] S.D. Ellis and D. Soper, Phys. Rev. **D48**, 3160 (1993); S. Catani, Yu.L. Dokshitzer and B.R. Webber, Phys. Lett. **B285**, 291 (1992); S. Catani, Yu.L. Dokshitzer, M.H. Seymour and B.R. Webber, Nucl. Phys. **B406**, 187 (1993).

[5] V.M. Abazov, *et al.* (DØ Collaboration), Phys.Lett. **B525,** 211 (2002); Phys.Rev. **D65,** 052008 (2002).

[6] S.D. Ellis, Z. Kunszt and D. Soper, Phys. Rev Lett. **69**, 1496 (1992); Phys. Rev Lett. **64**, 2121 (1990); Phys. Rev. **D40**, 2188 (1989); Phys. Rev. Lett. **62**, 726 (1989).

[7] S.D. Ellis, Z. Kunszt and D. Soper, Phys. Rev Lett. **69**, 3615 (1992); S.D. Ellis in Proceedings of the *28th Rencontres de Moriond: QCD and High Energy Hadronic Interactions*, March 1993, p. 235; B. Abbott, *et al.*, Fermilab Pub-97-242-E (1997).

[8] F. Abe *et al.* (CDF Collaboration), Phys. Rev. **D45**, 1448 (1992); B. Abbott, *et al.* (DØ Collaboration), FERMILAB-Pub-97/242-E.

[9] M.H. Seymour, Nucl. Phys. **B513**, 269 (1998).

[10] G. Marchesini, B.R. Webber, G. Abbiendi, I.G. Knowles, M.H. Seymour, L. Stanco, Computer Phys. Commun. **67,** 465 (1992) ; G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M.H. Seymour, B.R. Webber, JHEP **01,** 010 (2001) [hep-ph/9912396].

[11] F. Abe, *et al.* (CDF Collaboration), Phys. Rev. **D45**, 1448 (1992).

[12] V.M. Abazov, *et al.* (DØ Collaboration), hep-ex/0106032.