Application of Deep Learning to Jet Charge Discrimination

Meisam Ghasemi Bostanabad* and Mojtaba Mohammadi Najafabadi†
Department of Particles and Accelerators, Institute for research fundamental science

The Large Hadron Collider (LHC) generates an enormous volume of data every year. A significant challenge in this vast dataset is the analysis of hadronic jets, which are clusters of particles resulting from quantum chromodynamics processes. Determining the charge of the parton that initiates a light-quark jet can be highly valuable for testing components of the Standard Model and identifying potential signals of physics beyond the Standard Model. An overview of a study of using classical and quantum machine learning models to classify jet charges for up and anti-up quark jets is described here. In addition to better efficiency and rejection rate than the traditional methods, Jet charge offers a wide range of potential applications in both measurements and exploratory searches.

INTRODUCTION

The Large Hadron Collider (LHC) is the world's most powerful and complex particle accelerator, designed to probe the fundamental laws of nature. Its primary goals include testing the predictions of the Standard Model of particle physics, such as the existence of the Higgs boson, while also searching for signs of new physics beyond the Standard Model, such as supersymmetry and extra dimensions. However, one of the greatest challenges in analyzing the enormous volumes of data generated by the LHC is the study of hadronic jets — collimated sprays of particles resulting from quantum chromodynamic (QCD) processes. These jets are produced in high-energy collisions and their detailed characterization is essential. Accurate identification of the origin of these jets, whether from quarks or gluons [3], is critical for testing the Standard Model's predictions and for uncovering any potential signals of new, unknown phenomena. Improving our ability to use electric charge to discriminate between jets originating from different particles will enhance our understanding of both well-established theories and potential discoveries at the LHC.

The concept of the jet charge observable, first introduced by Field and Feynman [2], has been explored in experimental studies such as measurement of top quark charge at the LHC [3], and identifying the charge of heavy bosons W'/Z' [4]. Previous studies of jet charge [] involve measuring variants of a momentum-weighted jet charge, typically defined as a summation of the particle tracks within the jet, weighted by their transverse momentum (p_T).

$$\mathcal{Q}_j = \frac{1}{(p_T^{\text{jet}})^\kappa} \sum_{i \in \text{Tr}} Q_i (p_T^i)^\kappa \tag{1}$$

where the summation runs over all particles tracks recorded for the jet, Q_i and p_T^i represent the charge of the object and the magnitude of its transverse momentum relative to the beam axis, and p_T^{jet} is the total transverse momentum of the jet. The parameter κ is a tuning parameter defined the weighting with values ranging from 0 to 1. When κ approaches 0, the jet charge becomes highly influenced by soft tracks, which are often not detectable. Conversely, as κ increases towards infinity, the jet charge becomes primarily determined by the charge of the leading track [6, 8]. Fig. 1 presents the distributions of $\mathcal{Q}_j$ for $u, \bar{u}$ jets at two values of κ. Although there is a large overlap between the positive and negative jet charge distributions, $\mathcal{Q}_j$ remains useful for distinguishing the charge of the originating parton.

ANALYSIS VARIABLES

By utilizing various deep learning models such as convolutional or graph neural network, we can capture the complex features of jet charge distributions, which are often challenging to interpret using traditional methods. This discrimination helps in reducing background noise in

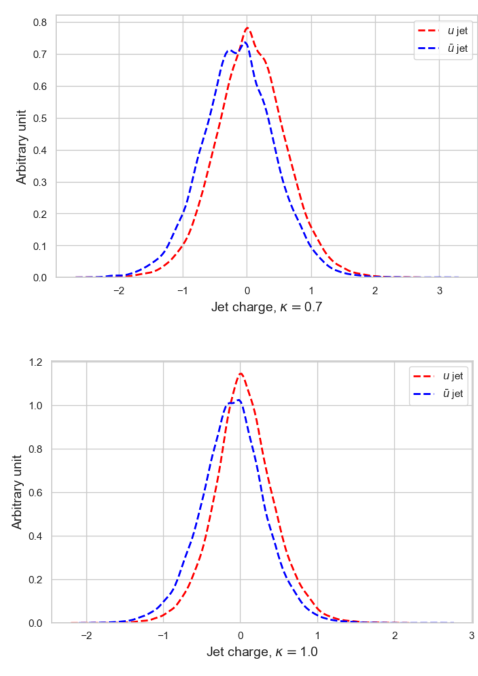

FIG. 1: Distributions of $\mathcal{Q}_j$ for $u, \bar{u}$ jets obtained from $pp \to ug$ or $pp \to \bar{u}g$ events with $\kappa = 0.7, 1$.

measurements, improving the accuracy of particle identification, and providing valuable insights into the behavior of quarks within jets.

The training dataset comprises 18000 simulated up quark and anti-up quark jet events within $pp \to ug$ and $pp \to \bar{u}g$ processes, simulated by `MADGRAPH` []. For this study, jets are clustered using anti-k_t algorithm with radius parameter $R = 0.4$ []. The `PYTHIA` [] generator is used to collect information for every jet in an event such as p_T, η, ϕ, momentum weighted charge for κ of 0, 0.3, 0.5, 0.7, and 1. In addition to jet variables which are necessary for deep neural network, we provide a list of information for jet's components to train convolutional neural network, including the particle's p_T, η, ϕ, and charge. In addition to the variables for jets and subjets, we also used some features from traditional jet charge calculation searches []:

$$
\begin{aligned}
Q_{1,\kappa} &= \sum_{i \in Tr} q_i z_i^{\kappa} \\
Q_{3,\kappa} &= \frac{\sum_{i \in Tr} q_i \left|\Delta\eta_i\right|^{\kappa}}{\sum_{i \in Tr} \left|\Delta\eta_i\right|^{\kappa}}
\end{aligned}
\tag{2}
$$

where $z_i = p_{T_i}/p_{T_{jet}}$ is the fraction of the jet momentum, and summation is over every particle in a jet, which provides much more information than just the particles with leading p_T. The last variables used for model training are charge ratio as the ratio of the sum of positive charges to the negative ones, and total jet charge as the sum of track's charge.

RESULT

Another sample of interest for validating jet charge is hadronically decaying W bosons coming from top decays. In a semi-leptonic $t\bar{t}$ sample, the leptonically decaying W can be used to determine the two charges of the jets from the hadronically decaying W. The distributions of these charges can then be

Deep Neural Network

A deep neural network (DNN) is a machine learning model inspired by the human brain, consisting of multiple layers of interconnected artificial neurons. Each neuron in the network performs a computation based on its inputs, weights, and biases. The relationship between the inputs and the output of a neuron can be expressed as $z = \sum_{i=1}^{n} w_i x_i + b$, where w_i are the weights associated with each input x_i, and b is the bias. The output a of a neuron is obtained by applying a non-linear activation function σ, such that $a = \sigma(z)$. Weights and biases are crucial as they determine how the input data is transformed as it propagates through the network. Learning these parameters effectively allows the model to recognize patterns and make accurate predictions.

The training of a DNN involves forward and backward propagation. During forward propagation, the input data is passed through the network layer by layer, and each layer's output becomes the input for the next layer, ultimately generating a prediction. This process can be expressed as $\mathbf{a}^{(l+1)} = \sigma(\mathbf{W}^{(l)}\mathbf{a}^{(l)} + \mathbf{b}^{(l)})$, where $\mathbf{W}^{(l)}$ represents the weights, $\mathbf{b}^{(l)}$ represents the biases of layer l, and $\mathbf{a}^{(l)}$ is the activation. In the backward propagation, the goal is to minimize the error between the predicted output and the actual output using an error function, such as mean squared error or cross-entropy. Gradients of the error function with respect to each parameter are calculated using the chain rule of differentiation. These gradients are then used in an optimization algorithm, such as gradient descent, to iteratively update the weights and biases, following the equation $\mathbf{W}^{(l)} \leftarrow \mathbf{W}^{(l)} - \eta \frac{\partial L}{\partial \mathbf{W}^{(l)}}$, where η is the learning rate, and L is the loss function. This process allows the network to improve its performance through multiple iterations and effectively learn the complex relationships within the data.

Another sample of interest for validating jet charge is hadronically decaying W bosons coming from top decays. In a semi-leptonic $t\bar{t}$ sample, the leptonically decaying

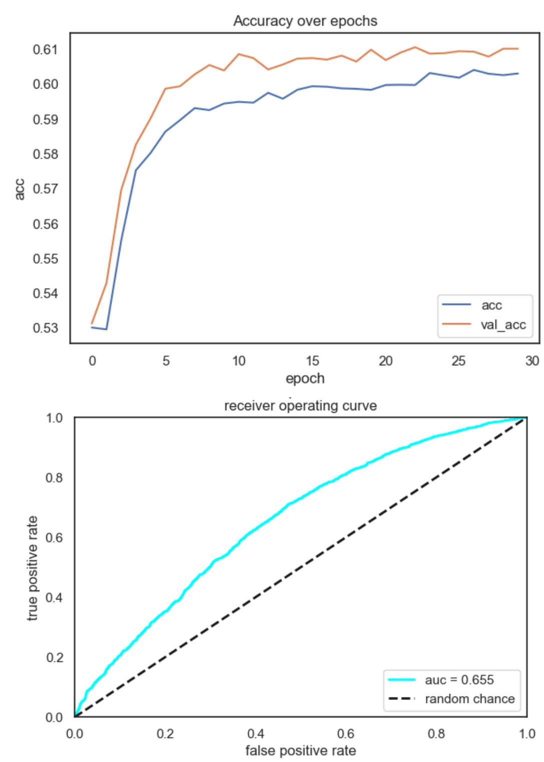

FIG. 2: Top: final state composition in dijet production. Bottom: Sum of the two jet charges in dijet events, for various κ. The growth with dijet invariant mass reflects the larger fraction of valence quark PDFs at large x and corresponding decrease in gg final states.

W can be used to determine the two charges of the jets from the hadronically decaying W. The distributions of these charges can then be compared to expectations, an example comparison is shown in Fig. 3. Validating this simulation on data would establish weighted jet charge as a trustworthy tool, which could then be used for new physics applications. Perhaps it could even be employed within the context of W decays to help with top-tagging or W polarization measurements.

Other ML Classifiers

Next, we consider the effects of pile-up and contamination on jet charge. One might worry that at high luminosity jet charge would be diluted by pile-up events, as up to $\mathcal{O}(100)$ proton-proton collisions can take place in the same bunch crossing. However, the products of these interactions tend to be soft, and are thus assigned little weight as long as κ is not too small. Further, charged particles can be traced to their collision vertex allowing most contamination to be removed. Finally, jet grooming techniques like trimming [11] can be applied to further reduce contamination. We present a comparison of effects of contamination and techniques to mitigate it in Fig. 4.

Convolutional Neural Network

Having demonstrated the practicality of jet charge for new physics searches and proposed ways to validate it on standard model data, we now turn to the feasibility of systematically improvable jet charge calculations. While Monte-Carlo programs like PYTHIA often provide an excellent approximation to full quantum chromodynamics, they are only valid to leading-order in perturbation theory including the resummation of leading Sudakov double logarithms [9].

A precise calculation of jet charge is challenging because it is not an infrared-safe quantity. Jet charge is sensitive to hadronization and cannot be calculated without knowledge of the fragmentation functions $D_j^h(x,\mu)$. These functions give the average probability that a hadron h will be produced by a parton j with the hadron having a fraction z of the parton's energy. Fragmentation functions, like parton distribution functions, are non-perturbative objects with perturbative evolution equations which simplify in moment space. The Mellin moments are defined by

which evolve through local renormalization group equations, just like the moments of parton distribution functions.

We first consider the average value of the jet charge

where $z = E_h/E_{\text{jet}}$ is the fraction of the jet's energy the hadron carries. For narrow jets $z \sim p_T^h/p_T^{\text{jet}}$.

To connect to the fragmentation functions, we first observe that for $\kappa > 0$ the the charge is dominated by collinear and not soft radiation. Thus the contributions of the hard and soft sectors of phase space, while contributing to the formation of the jet, should have a suppressed effect on $\mathcal{Q}_\kappa^i$. We can therefore use the fragmenting jet functions introduced in Refs. [13, 14] to write

Here $\mathcal{J}_i(E,R,\mu)$ is a jet function and $\mathcal{J}_{ij}(E,R,x,\mu)$ a set of calculable coefficients which depend on the jet definition and flavor i of the hard parton originating the jet. The hard and soft contributions conveniently canceled in this ratio. Therefore

with $\widetilde{\mathcal{J}}_{ij}$ related to $\mathcal{J}_{ij}$ by a Mellin-transform as in Eq. (**??**). By charge conjugation $\sum_h Q_h \widetilde{D}_g^h(\kappa,\mu) = 0$, so in particular $\langle \mathcal{Q}_\kappa^g \rangle = 0$. We have checked that the μ-dependence of $\mathcal{J}_{ij}/\mathcal{J}_i$ exactly compensates for the μ-dependance of the fragmentation functions at order α_s.

We have written both $\mathcal{J}_i(E,R,\mu)$ and $\mathcal{J}_{ij}(E,R,x,\mu)$ as if they depend on the energy E and size R of the jet, however, these functions only give a valid description to leading power of a single scale corresponding to the transverse size of the jet. Here we use the e^+e^- version

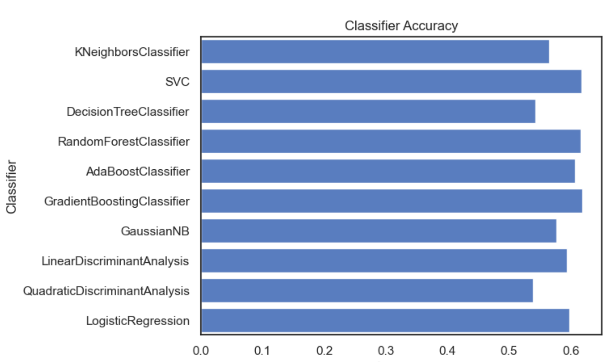

FIG. 3: Sum of jet charges of the two non b-jets in semi-leptonic $t\bar{t}$ events with a positively (solid) or negatively (dashed) charged lepton.

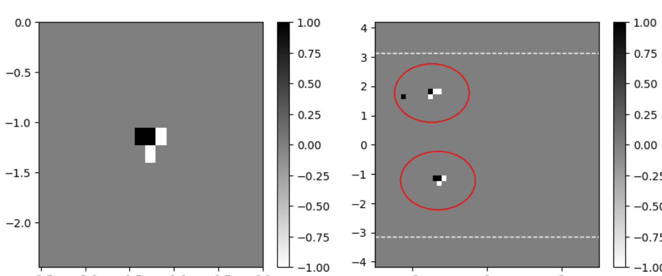

FIG. 4: Comparison of W' vs. Z' discrimination subject to contamination from initial state radiation (ISR), multiple interactions (MI), and pile-up events. We also show the result with and without jet trimming ($R_{\text{sub}} = 0.2$, $f_{\text{cut}} = 0.02$).

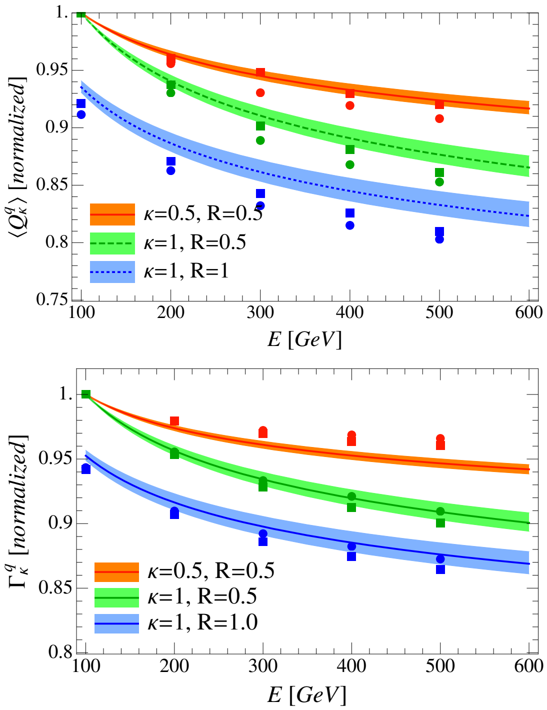

FIG. 5: Comparison of theory prediction (bands) for the average (top) and width (bottom) of the jet charge distribution to PYTHIA (squares and circles for d and u quarks) for e^+e^- collisions. Bands show uncertainty from varying the factorization scale by a factor of 2. Normalizing to 1 at $E = 100$ GeV and $R = 0.5$ removes the dependence on the nonperturbative input and quark flavor.

of anti-k_T jets of size R, for which the natural scale is $\mu_j = 2E\tan(R/2)$ [15]. We can therefore calculate the average jet charge by evaluating the Mellin-moments of fragmentation functions at the scale μ_j and multiplying by the jet functions.

Since only one linear combination of fragmentation functions appears in Eq.(**??**), the theoretical prediction is not significantly limited by the large uncertainty on $D_j^h(\kappa,\mu)$. One can simply measure $D_j^h(\kappa,\mu)$ by observing the average jet charge for each flavor at one value for μ and then using the theoretical calculation to predict it at other values. In the absence of data, we simulate such a comparison using PYTHIA. The result is shown in Figure 5 for various values of κ and R, and normalized at a reference point. Already we can see a clear agreement between the theory and PYTHIA.

To calculate other properties of the jet charge distribution requires correlations among hadrons. For example, we can consider the width of the jet charge, $(\Gamma_\kappa^i)^2 = \langle \mathcal{Q}_\kappa^i \rangle^2 - \langle (\mathcal{Q}_\kappa^i)^2 \rangle$. This depends on the moment

$$\big\langle (\mathcal{Q}_\kappa^i)^2 \big\rangle = \sum_n \sum_{h_1,\dots,h_n} \int \mathrm{d}z_1 \cdots \mathrm{d}z_n \, (Q_1 z_1^\kappa + \cdots + Q_n z_n^\kappa)^2 \times \frac{1}{\sigma_{\rm jet}} \frac{\mathrm{d}^n \sigma_{h_1\cdots h_n \in \rm jet}}{\mathrm{d}z_1 \cdots \mathrm{d}z_n}\,, \tag{3}$$

where the sum runs over all hadronic final states. After integrating over most of the z_i and including a factor of $\frac{1}{2}$ for identical hadrons, this simplifies to

$$\big\langle (\mathcal{Q}_\kappa^i)^2 \big\rangle = \int \mathrm{d}z\, z^{2\kappa} \sum_h Q_h^2 \frac{1}{\sigma_{\rm jet}} \frac{\mathrm{d}\sigma_{h\in \rm jet}}{\mathrm{d}z} + \int \mathrm{d}z_1\, \mathrm{d}z_2\, z_1^\kappa z_2^\kappa \sum_{h_1,h_2} Q_{h_1} Q_{h_2} \frac{1}{\sigma_{\rm jet}} \frac{\mathrm{d}\sigma_{h_1 h_2 \in \rm jet}}{\mathrm{d}z_1\, \mathrm{d}z_2}\,. \tag{4}$$

The first term on the right hand side can be expressed in terms of products of fragmentation functions and jet functions as for $\langle \mathcal{Q}_\kappa^i \rangle$. The second term can be expressed in terms of something we call a dihadron fragmenting jet function, $\mathcal{G}_i^{h_1 h_2}$. Its matching onto (dihadron) fragmentation functions is given by

$$\begin{aligned} &\mathcal{G}_i^{h_1 h_2}(E,R,z_1,z_2,\mu) \\ &= \sum_j \int \frac{\mathrm{d}u}{u^2}\, \mathcal{J}_{ij}(E,R,u,\mu) D_j^{h_1 h_2}\Big(\frac{z_1}{u},\frac{z_2}{u},\mu\Big) \\ &+ \sum_{j,k} \int \frac{\mathrm{d}u}{u}\frac{\mathrm{d}v}{v}\, \mathcal{J}_{ijk}(E,R,u,v,\mu) D_j^{h_1}\Big(\frac{z_1}{u},\mu\Big) D_k^{h_2}\Big(\frac{z_2}{v},\mu\Big)\,, \end{aligned} \tag{5}$$

The second term is due to a perturbative parton splitting before hadronization and only starts at 1-loop order,

$$\mathcal{J}_{ijk}^{(1)}(E,R,u,v,\mu) = \mathcal{J}_{ij}^{(1)}(E,R,u,\mu)\delta(1-u-v)\delta_{k,a(ij)}\,, \tag{6}$$

where $\delta_{k,a(ij)}$ indicates that the flavor k is completely fixed by ij. E.g. $a(qq) = g$, $a(gq) = \bar{q}$. We then find

$$\big\langle (\mathcal{Q}_\kappa^q)^2 \big\rangle = \frac{1}{16\pi^3} \sum_j \frac{\widetilde{\mathcal{J}}_{qj}(E,R,2\kappa,\mu)}{J_q(E,R,\mu)} \Big[\sum_h Q_h^2 \tilde{D}_j^h(2\kappa,\mu) + \sum_{h_1,h_2} Q_{h_1} Q_{h_2} \tilde{D}_j^{h_1 h_2}(\kappa,\kappa,\mu) \Big]\,. \tag{7}$$

(For a gluon jet, which we do not consider here, there is a contribution from the last line of Eq. (5) corresponding to a perturbative $g \to q\bar{q}$ splitting.) We have checked that this equation is renormalizat-group invariant at order α_s.

Unfortunately, the dihadron fragmentation functions are even more poorly known than the regular fragmentation functions. However, we can use the same trick as for the average jet charge to calculate the E and R dependence of the width, given measurements at some reference scale. As with the average jet charge, we can now calculate the width by fitting one parameter for each flavor, corresponding to the term in brackets in Eq. (7), and predicting the E and R dependence. Results compared to PYTHIA for the width are shown in Fig. 5 and show good agreement. The gluon mixing contribution is not included in these figures since it requires additional matching; a discussion of the effect of gluon mixing can be found in Ref. [16].

To go beyond the average and the width, for example to the 3rd or higher moments, multi-hadron fragmentation functions would be needed. From a practical point of view, such functions are nearly impossible to measure with any precision. However, we have found that the discriminating power of jet charge is nearly as strong using Gaussians based on the average and width as it is with the full differential jet charge distribution. It follows that accurate calculations of the phenomenologically relevant part of jet charge distributions are achievable with the formalism we have introduced in this paper. The full fragmenting jet functions, both for the single hadron and dihadron case, and the evolution kernels, are now known at 1-loop order. To see whether higher precision is required, and to explore the importance of power corrections, requires some LHC data to compare with. The calculations and issues discussed here are expanded on in Ref. [16].

As we have shown, the weighted jet charge, and its moments, are measurable and testable already at the LHC. With potential to uniquely determine quantum number of certain new physics particles, should they show up, it is important to verify jet charge on standard model processes. Thus jet charge holds promise as a measurable, calculable and useful observable.

We thank G. Kane, E. Kuflik, S. Rappoccio, G. Stavenga, and M. Strassler for helpful discussions. DK is supported by the Simons foundation and by an LHC-TI travel grant. TL is partially supported by NASA Theory Program grant NNX10AD85G and by the National Science Foundation under Grant No. PHYS-1066293 and the hospitality of the Aspen Center for Physics. MDS is supported by DOE grant DE-SC003916. WW is supported by DOE grant DE-FG02-90ER40546.

* mghasemi@ipm.ir
† mojtaba@ipm.ir

[1] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, (2008), arXiv:0810.0409 [hep-ph].
[2] D. E. Kaplan, K. Rehermann, M. D. Schwartz, and B. Tweedie, Phys.Rev.Lett. **101**, 142001 (2008), arXiv:0806.0848 [hep-ph].
[3] J. Gallicchio and M. D. Schwartz, Phys.Rev.Lett. **107**, 172001 (2011), arXiv:1106.3076 [hep-ph].
[4] A. Altheimer *et al.*, J. Phys. **G39**, 063001 (2012).
[5] R. Field and R. Feynman, Nucl. Phys. **B136**, 1 (1978).
[6] J. Berge *et al.*, Nucl. Phys. **B184**, 13 (1981).
[7] J. Albanese *et al.* (EMC), Phys. Lett. **B144**, 302 (1984).
[8] D. Decamp *et al.* (ALEPH), Phys. Lett. **B259**, 377 (1991).
[9] A note on simulation: Our Monte-Carlo events are generated in MADGRAPH 5 [17], showering and hadronization are modeled using PYTHIA 8 [18], and jets are clustered with FASTJET [19] with anti-kT [20] jets of $R = 0.5$. Unless stated otherwise, we assume a 14 TeV LHC.
[10] G. Aad *et al.* (ATLAS), (2012), arXiv:1208.0949.
[11] D. Krohn, J. Thaler, and L.-T. Wang, JHEP **1002**, 084 (2010).
[12] Note that in addition Monte-Carlo programs differ in their formulation of the parton shower and in their treatment of hadronization. Comparing PYTHIA to HERWIG++ [21], we find agreement for the mean jet charge at the $\mathcal{O}(10\%)$ level.
[13] M. Procura and I. W. Stewart, Phys. Rev. **D81**, 074009 (2010).
[14] M. Procura and W. J. Waalewijn, Phys. Rev. **D85**, 114041 (2012).
[15] S. D. Ellis *et al.*, JHEP **1011**, 101 (2010).
[16] W. J. Waalewijn, Phys.Rev. **D86**, 094030 (2012), arXiv:1209.3019 [hep-ph].
[17] J. Alwall *et al.*, JHEP **1106**, 128 (2011), arXiv:1106.0522.
[18] T. Sjostrand, S. Mrenna, and P. Z. Skands, Comput. Phys. Commun. **178**, 852 (2008).
[19] M. Cacciari, G. P. Salam, and G. Soyez, Eur. Phys. J. **C72**, 1896 (2012).
[20] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **0804**, 063 (2008).
[21] S. Gieseke *et al.*, (2011), arXiv:1102.1672.