Nuclear Instruments and Methods in Physics Research A 858 (2017) 1–11

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

Mathematical properties of numerical inversion for jet calibrations

Aviv Cukierman[a,b], Benjamin Nachman[a,b,c,*]

[a] Physics Department, Stanford University, Stanford, CA 94305, USA
[b] SLAC National Accelerator Laboratory, Stanford University, Menlo Park, CA 94025, USA
[c] Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94704, USA

A B S T R A C T

Numerical inversion is a general detector calibration technique that is independent of the underlying spectrum. This procedure is formalized and important statistical properties are presented, using high energy jets at the Large Hadron Collider as an example setting. In particular, numerical inversion is inherently biased and common approximations to the calibrated jet energy tend to over-estimate the resolution. Analytic approximations to the closure and calibrated resolutions are demonstrated to effectively predict the full forms under realistic conditions. Finally, extensions of numerical inversion are presented which can reduce the inherent biases. These methods will be increasingly important to consider with degraded resolution at low jet energies due to a much higher instantaneous luminosity in the near future.

1. Introduction

At a proton-proton collider like the Large Hadron Collider (LHC), quarks and gluons are produced copiously. These partons fragment to produce collimated streams of colorless particles that leave their energy in the calorimeters of the ATLAS and CMS detectors.[1] The energy depositions are organized using jet clustering algorithms to stand as experimental proxies for the initiating quarks and gluons. The most widely used clustering scheme in ATLAS and CMS is the anti-k_t algorithm [3] with radius parameter R=0.4. Even though the inputs to jet clustering (topological clusters for ATLAS [4,5] and particle flow objects for CMS [6,7] are themselves calibrated, the average reconstructed jet energy is not the same as the true jet energy, because of various detector effects. To account for this, calibrations are applied to each reconstructed jet.

2. Numerical inversion

The jet calibration procedures of ATLAS [8] and CMS [9,10] involve several steps to correct for multiple nearly simultaneous pp collisions (pileup), the non-linear detector response, the η-dependence of the jet response, flavor-dependence of the jet response, and residual data/simulation differences in the jet response. The simulation-based corrections to correct for the calorimeter non-linearities in transverse energy E_{T} and pseudorapidity η are accounted for using *numerical inversion*.

The purpose of this note is to formally document numerical inversion and describe (with proof) some of its properties. In what follows, X will be a random variable representing the particle-jet E_{T} and Y will be a random variable representing the reconstructed jet E_{T}. Define.[2]

$$f_{\mathrm{me}}(x) = \mathbb{E}[Y|X = x] \tag{1}$$

$$R_{\mathrm{me}}(x) = \mathbb{E}\left[\frac{Y}{x}\middle| X = x\right] = \frac{f_{\mathrm{me}}(x)}{x}. \tag{2}$$

Where the subscript indicates that we are taking the mean of the stated distribution and '$\mathbb{E}$' stands for *expected value* (= average). In practice, sometimes the core of the distribution of $Y|X = x$ is fit with a Gaussian and so the effective measure of central tendency is the mode of the distribution. Therefore in analogy to Eqs. (1) and (2), we define

$$f_{\mathrm{mo}}(x) = \mathrm{mode}[Y|X = x] \tag{3}$$

$$R_{\mathrm{mo}}(x) = \mathrm{mode}\left[\frac{Y}{x}\middle| X = x\right] = \frac{f_{\mathrm{mo}}(x)}{x}. \tag{4}$$

We will often drop the subscript of f and R for brevity in the text, when

* Corresponding author at: Physics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94704, USA.
E-mail address: bnachman@cern.ch (B. Nachman).

[1] Jets have been calibrated in previous experiments, such as the Tevatron CDF [1] and D0 [2] experiments, but the methods were significantly different and so this note focuses on the general purpose LHC experiments.

[2] Capital letters represent random variables and lower case letters represent realizations of those random variables, i.e. X=x means the random variable X takes on the (non-random) value x.

http://dx.doi.org/10.1016/j.nima.2017.03.038
Received 30 October 2016; Received in revised form 6 March 2017; Accepted 20 March 2017
Available online 21 March 2017
0168-9002/

it is clear which definition we are referring to. If not specified, f and R will refer to a definition using a generic definition of central tendency. For all sensible notions of central tendency, we still have that $R(x) = \frac{f(x)}{x}$.

We will often think of $Y|X = x \sim \mathcal{N}(f(x), \sigma(x))$, where this notation means 'Y given X=x is normally distributed with mean $f(x)$ and standard deviation $\sigma(x)$'; however, in this note, we will remain general unless stated otherwise. The function $R(x)$ is called the *response function*. Formally, numerical inversion is the following procedure:

1. Compute $f(x)$, $R(x)$.
2. Let $\widetilde{R}(y) = R(f^{-1}(y))$.
3. Apply a jet-by-jet correction: $Y \mapsto Y/\widetilde{R}(Y)$.

The intuition for step 2 is that for a given value y drawn from the distribution $Y|X = x$, $f^{-1}(y)$ is an estimate for x and then $R(f^{-1}(y))$ is an estimate for the response at the value of x that gives rise to Y. Let $p(x)$ be the prior probability density function of E_T. Then we note that we do not want to use $\mathbb{E}[X|Y]$ instead of $f^{-1}(Y)$ because the former depends on $p(x)$, whereas f (and thus f^{-1}) does not depend on $p(x)$, by construction.

We can see now our first result, which will be useful for the rest of this note:

The correction derived from numerical inversion is $Y \mapsto Z = f^{-1}(Y)$.

Proof.

$$\widetilde{R}(Y) = R(f^{-1}(Y)) = \frac{f(f^{-1}(Y))}{f^{-1}(Y)} = \frac{Y}{f^{-1}(Y)} \to Z = \frac{Y}{\widetilde{R}(Y)} = f^{-1}(Y) \tag{5}$$

□

2.1. Closure

One important property of numerical inversion is the concept of *closure*, which quantifies whether the new distribution $f^{-1}(Y|X = x)$ obtained after numerical inversion is centered at x, using the same notion of central tendency as in the definition of f. In particular, define the closure as

$$C_{\rm me}(x) \equiv \mathbb{E}\left[\frac{Z}{x}\middle|X = x\right] = \mathbb{E}\left[\frac{f^{-1}(Y)}{x}\middle|X = x\right], \tag{6}$$

and $C_{\rm mo}$ is defined in an analogous way. The symbol C will denote the closure for a generic notion of central tendency. We say that numerical inversion has *achieved closure* or simply *closes* if, for all x,

$$C = 1. \tag{7}$$

2.2. Assumptions and definitions

The general results presented in the following sections are based on three assumptions listed below. These requirements should be satisfied by real detectors using calorimeters and trackers to reconstruct jets, given that the detector-level reconstruction is of sufficiently high quality.

1. $f^{-1}(y)$ exists for all y in the support of Y, and f^{-1} is single-valued. These may seem like obvious statements, but are not vacuous, even for a real detector. For example, pileup corrections can result in non-zero probability that $Y < 0$, so the function f must be computed for all possible values of Y, even if the transverse energy is negative. At the high-luminosity LHC (HL-LHC), the level of pileup will be so high that the jet energy resolution may be effectively infinite at low transverse energies (no correlation between particle-level and detector-level jet energy). In that case, f^{-1} may not be single valued and numerical inversion cannot be strictly applied as described in Section 2.
2. $f(x)$ is monotonically increasing: $f'(x) > 0$ for all x. This condition should trivially hold for any reasonable detector: detector-level jets resulting from particle-level jets with a higher E_T should on average have a higher E_T than those originating from a lower E_T particle-level jet. Note that this is only true for a fixed η. Detector technologies depend significantly on η and therefore the η-dependence of f (for a fixed x) need not be monotonic. We note also that Assumption 1 implies that $f'(x) \geq 0$ or $f'(x) \leq 0$ for all x; so Assumption 2 is equivalent to the additional assumptions that $f'(x) \neq 0$ for any x, and that $f'(x) > 0$ (as opposed to $f'(x) < 0$).
3. f is twice-differentiable. The first derivative of f has already been assumed to exist in Assumption 2, and the second derivative will also be required to exist for some of the later results. In practice we expect f to be differentiable out to any desired order.

We note that as long as the above three assumptions hold, the theorems stated in the remainder of this paper are valid. In particular, this implies that x could be any calibrated quantity that satisfies the above constraints, e.g. the jet transverse momentum p_T or the jet mass m. We focus on the case of calibrating the E_T for sake of concreteness.

We have separated the results in this paper into "Proofs" and "Derivations". The "Proofs" require only the three assumptions stated above, and in particular do not assume anything about the shape of the underlying distributions, e.g. that the distributions $Y|X = x$ are Gaussian or approximately Gaussian. The "Derivations" are useful approximations that apply in the toy model described in Appendix I; we expect them to apply in a wide variety of cases relevant to LHC jet physics. In particular, we expect these approximations to hold in cases with properties similar to the toy model presented here - e.g., good approximation of f by its truncated Taylor series about each point and approximately Gaussian underlying distributions of $Y|X = x$.[3]

Finally, in the rest of this paper, we write $\rho_{Y|X}(y|x)$ to represent the probability distribution of Y given X=x, and $\rho_{Z|X}(z|x)$ to be the probability distribution of Z given X=x. A standard fact about the probability distribution from changing variables is that

$$\rho_{Z|X}(z|x) = f'(z)\rho_{Y|X}(f(z)|x). \tag{8}$$

To ease the notation, we will often use $\rho_Y(y)$ and $\rho_Z(z)$ interchangeably with $\rho_{Y|X}(y|x)$ and $\rho_{Z|X}(z|x)$, respectively, when it is clear (as is usually the case) that we are conditioning on some true value x.[4]

3. Results

In the subsequent sections, we will derive properties about the closure C for three different definitions of the central tendancy: mean (Section 3.1), mode (Section 3.2), and median (Section 3.3).

3.1. Mean

In the following section only, for brevity, we will let f be $f_{\rm me}$ and C be $C_{\rm me}$.

3.1.1. Closure

We can write the closure Eq. (6) as

[3] Note that we do not require that $Y|X = x$ is exactly Gaussian, only that it is approximately Gaussian, which is true for a wide range of energies and jet reconstruction algorithms at ATLAS and CMS. In particular, there are non-negligible (but still often small) asymmetries at low and high E_T at ATLAS and CMS [8–10]. In any case, even if $Y|X = x$ *is* Gaussian, $Z|X = x$ is in general *not* Gaussian, for non-linear response functions; see Appendix A.

[4] In practice it is necessary to condition on a small range of X, e.g. $X \in [x, (1+\epsilon)x]$. If ϵ is large then there can be complications from the changing of $f(x)$ over the specified range and from the shape of the prior distribution of X over the specified range. These challenges can be solved by generating large enough Monte Carlo datasets. We therefore assume that $\epsilon \ll 1$ and consider complications from finite ϵ beyond the scope of this paper.