1 Reviewing and updating Run 1 analysis

We will start with the starting point from the Run 1 note. Namely, we define the per-jet measurement errors in terms of terms that are fully correlated between EMTopo and HI collections and terms that are completely uncorrelated. For now, we assume the correlated contribution is the same for both collections.

$$\begin{aligned} \Delta p_{\mathrm{T}}^{\mathrm{EM}} &\equiv \Delta_{\mathrm{c}} + \Delta_{\mathrm{EM}} \\ \Delta p_{\mathrm{T}}^{\mathrm{HI}} &\equiv \Delta_{\mathrm{c}} + \Delta_{\mathrm{HI}} \end{aligned} \tag{1}$$

Then, the jet energy resolution, defined as the standard deviation of the jet energy from the true can be written,

$$\begin{aligned} R_{\mathrm{EM}} &\equiv Var[\Delta p_{\mathrm{T}}^{\mathrm{EM}}] = Var[\Delta_{\mathrm{c}} + \Delta_{\mathrm{EM}}] = Var[\Delta_{\mathrm{c}}] + Var[\Delta_{\mathrm{EM}}] \\ R_{\mathrm{HI}} &\equiv Var[\Delta p_{\mathrm{T}}^{\mathrm{HI}}] = Var[\Delta_{\mathrm{c}} + \Delta_{\mathrm{HI}}] = Var[\Delta_{\mathrm{c}}] + Var[\Delta_{\mathrm{HI}}]. \end{aligned} \tag{2}$$

Following the convention of the note, we will simplify the notation: $Var[\Delta_{\mathrm{c}}] \equiv s_{\mathrm{c}}^2$, $Var[\Delta_{\mathrm{EM}}] \equiv s_{\mathrm{EM}}^2$, $Var[\Delta_{\mathrm{HI}}] \equiv s_{\mathrm{HI}}^2$ with the results that

$$\begin{aligned} R_{\mathrm{EM}} &\equiv s_{\mathrm{c}}^2 + s_{\mathrm{EM}}^2 \\ R_{\mathrm{HI}} &\equiv s_{\mathrm{c}}^2 + s_{\mathrm{HI}}^2 \end{aligned} \tag{3}$$

We can use the MC to evaluate R_{EM} and R_{HI} and we can evaluate the difference between the result which we define as A:

$$A \equiv R_{\mathrm{EM}} - R_{\mathrm{HI}} = s_{\mathrm{c}}^2 + s_{\mathrm{EM}}^2 - s_{\mathrm{c}}^2 - s_{\mathrm{HI}}^2 = s_{\mathrm{EM}}^2 - s_{\mathrm{HI}}^2. \tag{4}$$

We can also evaluate a quantity that is differently sensitive to the relative resolution of the two jet collections. Namely we can evaluate the variance of the difference between $\Delta p_{\mathrm{T}}^{\mathrm{EM}}$ and $\Delta p_{\mathrm{T}}^{\mathrm{HI}}$:

$$B_{MC} \equiv Var[(\Delta p_{\mathrm{T}}^{\mathrm{EM}} - \Delta p_{\mathrm{T}}^{\mathrm{HI}})] = s_{\mathrm{EM}}^2 + s_{\mathrm{HI}}^2. \tag{5}$$

However, we can also evaluate the same quantity but using data. The difference between these two values we will take as an estimate of the (squared) uncertainty on B:

$$\delta^2 B \equiv (B_{MC} - B_{data})^2. \tag{6}$$

The Jet-EtMiss group has provided an estimate for the uncertainty on R_{EM}, which we will write $\delta^2 R_{\mathrm{EM}}$. The question we want to answer is "given $\delta^2 R_{\mathrm{EM}}$ and $\delta^2 B$, what limits can we put on $\delta^2 R_{\mathrm{HI}}$?

Using Eq. 3, we can write

$$\begin{aligned} \delta^2 R_{\mathrm{EM}} &\equiv \delta^2 s_{\mathrm{c}}^2 + \delta^2 s_{\mathrm{EM}}^2 \\ \delta^2 R_{\mathrm{HI}} &\equiv \delta^2 s_{\mathrm{c}}^2 + \delta^2 s_{\mathrm{HI}}^2 \end{aligned} \tag{7}$$

Then, given a value of $\delta^2 R_{\mathrm{EM}}$, $\delta^2 s_{\mathrm{c}}^2$ can only vary over the range $0 \leq s_{\mathrm{c}}^2 \leq \delta^2 R_{\mathrm{EM}}$. Suppose we define $\delta^2 s_{\mathrm{c}}^2 \equiv f \delta^2 R_{\mathrm{EM}}$. Then, $\delta^2 s_{\mathrm{EM}}^2 = (1-f)\delta^2 R_{\mathrm{EM}}$. We do not know *a priori* the value of f, but $\delta^2 B$ constrains f since $\delta^2 B = \delta^2 s_{\mathrm{EM}}^2 + \delta^2 s_{\mathrm{HI}}^2$. Since both terms are positive definite, $\delta^2 s_{\mathrm{EM}}^2 \leq \delta^2 B$ or, more explicitly,

$$(1-f)\delta^2 R_{\mathrm{EM}} \leq \delta^2 B \rightarrow f \geq 1 - \frac{\delta^2 B}{\delta^2 R_{\mathrm{EM}}}. \tag{8}$$

Let's consider the extreme values of f: $1 - \frac{\delta^2 B}{\delta^2 R_{\mathrm{EM}}}$ and 1. For $f = 1 - \frac{\delta^2 B}{\delta^2 R_{\mathrm{EM}}}$, $\delta^2 s_{\mathrm{HI}}^2 = 0$ and then $\delta^2 R_{\mathrm{HI}} = \delta^2 s_{\mathrm{c}}^2 = f\delta^2 R_{\mathrm{EM}} = \delta^2 R_{\mathrm{EM}} - \delta^2 B$. In contrast, for $f = 1$, $\delta^2\ s_{\mathrm{EM}}^2 = 0$ so $\delta^2 B = \delta^2 s_{\mathrm{HI}}^2$ and $\delta^2 R_{\mathrm{EM}} = \delta^2 s_{\mathrm{c}}^2$. As result, $\delta^2 R_{\mathrm{HI}} = \delta^2 R_{\mathrm{EM}} + \delta^2 B$. Thus, the uncertainly on the HI jet energy resolution can span the range $[\delta^2 R_{\mathrm{EM}} - \delta^2 B, \delta^2 R_{\mathrm{EM}} + \delta^2 B]$ depending on the value of f. Since we don't know f, the upper value provides a conservative estimate for the uncertainty on the HI jet energy resolution.

The difference between the above result and equation 20 in the Run 1 note results from the neglect of a covariance term in Eq. 20 in the note. That term reflects the fact that the uncertainty in the EMTopo jet energy resolution can be correlated with the systematic uncertainty on A. See e.g. Eq. 4. In the above analysis, that covariance is explicitly accounted for in the decomposition in Eq. 3. To evaluate the covariance using the above formulation, we can write the *uncertainty covariance* between R_{EM} ($\delta^2\sigma_{HI}$ in the note) and A using Eq. 4

$$Cov[R_{\mathrm{EM}}, A] = Var[R_{\mathrm{EM}}] - Cov(R_{\mathrm{EM}}, R_{\mathrm{HI}}) = \delta^2 R_{\mathrm{EM}} - Cov(R_{\mathrm{EM}}, R_{\mathrm{HI}}). \tag{9}$$

Here all of the variances and covariances are related to uncertainties on the parameters not their statistical distributions. Using Eq. 4 and assuming no uncertainty covariance between s_{c}^2, s_{EM}^2, and s_{HI}^2, then $Cov(R_{\mathrm{EM}}, R_{\mathrm{HI}}) = \delta^2 s_{\mathrm{c}}^2$ with thew result that

$$Cov[R_{\mathrm{EM}}, A] = \delta^2 R_{\mathrm{EM}} - \delta^2 s_{\mathrm{c}}^2 = \delta^2 R_{\mathrm{EM}}\,(1 - f)\,. \tag{10}$$

Repeating the procedure of the note, we re-write Eq. 4

$$R_{\mathrm{HI}} = R_{\mathrm{EM}} - A. \tag{11}$$

This is equivalent to equation 19 in the Run 1 note, but without taking the square root. Then, following the usual propagation of errors including the covariance term,

$$\begin{aligned} \delta^2 R_{\mathrm{HI}} =& \delta^2 R_{\mathrm{EM}} + \delta^2 A - 2Cov[R_{\mathrm{EM}}, A] \\ =& \Delta^2 R_{\mathrm{EM}}\,[1 - 2(1-f)] + \delta^2 B \\ =& \Delta^2 R_{\mathrm{EM}}\,(2f - 1) + \delta^2 B. \end{aligned} \tag{12}$$

Now, if we evaluate Eq. 12 at the minimum and maximum values of f per the above analysis we obtain estimates for $\delta^2 R_{\mathrm{HI}}$ of $\delta^2 R_{\mathrm{EM}} - \delta^2 B$ and $\delta^2 R_{\mathrm{EM}} + \delta^2 B$, respectively. These results are consistent with the above analysis and yield $\delta^2 R_{\mathrm{EM}} + \delta^2 B$ as a conservative estimate for $\delta^2 R_{\mathrm{HI}}$.

2 Impact of GSC calibration

The application of the GSC to the pp jets introduces a complication because it likely violates the assumptions in Eq. 2 that the correlated contribution to the per-jet measurement errors are the same in the EMTopo and HI jets. The GSC calibration factors may introduce a scale difference between the correlated contributions to the two jet collections. For the following we will assume

$$\begin{aligned} \Delta p_{\mathrm{T}}^{\mathrm{EM}}\big|_{\mathrm{GSC}} &\equiv \Delta_{\mathrm{c}} + \Delta_{\mathrm{EM}} \\ \Delta p_{\mathrm{T}}^{\mathrm{HI}} &\equiv \lambda\Delta_{\mathrm{c}} + \Delta_{\mathrm{HI}} \end{aligned} \tag{13}$$

Following the above,

$$\begin{aligned} R_{\mathrm{EM}} &\equiv s_{\mathrm{c}}^2 + s_{\mathrm{EM}}^2, \\ R_{\mathrm{HI}} &\equiv \lambda^2 s_{\mathrm{c}}^2 + s_{\mathrm{HI}}^2, \end{aligned} \tag{14}$$

and

$$A = s_c^2 + s_{EM}^2 - \lambda^2 s_c^2 - s_{HI}^2 = (1-\lambda^2) s_{EM}^2 - s_{HI}^2, \tag{15}$$

$$B = s_c^2 + s_{EM}^2 + \lambda^2 s_c^2 + s_{HI}^2 - 2\lambda s_c^2 = (1-\lambda)^2 s_c^2 + s_{EM}^2 + s_{HI}^2. \tag{16}$$

For the moment, we will assume that λ can be independently estimated without any uncertainty. We will come back below and discuss how to estimate λ and evaluate the effects of non-zero uncertainty on its value.

From Eq. 16, we can write the uncertainty on B,

$$\delta^2 B = (1-\lambda)^4 \delta^2 s_c^2 + \delta^2 s_{EM}^2 + \delta^2 s_{HI}^2. \tag{17}$$

Now, we follow the above approach and define $\delta^2 s_c^2 \equiv f \delta^2 R_{EM}$ and $\delta^2 s_{EM}^2 = (1-f)\delta^2 R_{EM}$. Then

$$\delta^2 B = \delta^2 R_{EM} \left(1 + f\left[(1-\lambda)^4 - 1\right]\right) + \delta^2 s_{HI}^2. \tag{18}$$

Similar to the situation above, f is bounded by the requirement that $\delta^2 s_{HI}^2 \geq 0$, so

$$1 + f\left[(1-\lambda)^4 - 1\right] \leq \frac{\delta^2 B}{\delta^2 R_{EM}} \rightarrow f \geq \left[\frac{1}{1-(1-\lambda)^4}\right]\left(1 - \frac{\delta^2 B}{\delta^2 R_{EM}}\right). \tag{19}$$

Depending on the value of λ, this "minimum" value of f may be larger than unity which means that there is an inconsistency in the formulation of the result. In other words, the uncertainty on B limits the possibile values of λ. More specifically,

$$\frac{1}{1-(1-\lambda)^4} \leq 1 - \frac{\delta^2 B}{\delta^2 R_{EM}} \tag{20}$$

or

$$(1-\lambda)^4 \leq \frac{\frac{\delta^2 B}{\delta^2 R_{EM}}}{1 - \frac{\delta^2 B}{\delta^2 R_{EM}}}. \tag{21}$$

Assuming that this limited is satisfied, at the lower limit of f, $\delta^2 s_{HI}^2 = 0$ and $\delta^2 R_{HI} = \lambda^4 \delta^2 s_c^2$ or

$$\delta^2 R_{HI}\big|_{f\ \min} = \left[\frac{\lambda^4}{1-(1-\lambda)^4}\right]\left(\delta^2 R_{EM} - \delta^2 B\right). \tag{22}$$

At the upper limit, $\delta^2 s_{EM}^2 = 0$ and $\delta^2 B = \delta^2 s_{HI}^2$. Then, $\delta^2 R_{HI} = \lambda^4 \delta^2 s_c^2 + \delta^2 B$ or

$$\delta^2 R_{HI}\big|_{f\ \max} = \lambda^4 \delta^2 R_{EM} + \delta^2 B. \tag{23}$$

This value should represent the maximum uncertainty on $\delta^2 R_{HI}$. Even though Eq. 22 appears to produce larger values in the case of small B, those values are not allowed by the restriction that $f \leq 1$.

The above analysis depends on the parameter λ which we can estimate by studying the correlation between Δp_T^{EM} and Δp_T^{HI}. If we measure in the MC samples, the statistical covariance $Cov(\Delta p_T^{EM}, \Delta p_T^{HI})$, then in the formulation here, the covariance reduces to $Cov(\Delta p_T^{EM}, \Delta p_T^{HI}) = \lambda Var \Delta_c = \lambda s_c^2$. We could also measure the covariance when not applying the GSC calibration to the EMTopo jets for which

$$\Delta p_T^{EM}\big|_{\text{No GSC}} \equiv \lambda \Delta_c + \Delta_{EM}' \tag{24}$$

and, for which the covariance would reduce to $Cov(\Delta p_T^{EM}\big|_{\text{No GSC}}, \Delta p_T^{HI}) = \lambda^2 s_c^2$. From these two covariances, an estimate for λ could be obtained:

$$\lambda = \frac{Cov(\Delta p_T^{EM}\big|_{\text{No GSC}}, \Delta p_T^{HI})}{Cov(\Delta p_T^{EM}, \Delta p_T^{HI})}. \tag{25}$$