Angular and energy resolution in Neutrino Telescopes

Jordan Seneca

April 2020

This document is intended to clarify the relationship between the distribution of *resolutions* and the performance of their underlying energy and direction estimators. We will give a short introduction on basic statistical concept to base my arguments, then present the case of energy and direction estimation.

Introduction

An estimator is a rule which gives the estimated value, or estimate, $\hat{\Theta}$, for a true parameter Θ of a random observable X. The estimator can be among other things, a reconstruction algorithm or a function, which inputs data and outputs $\hat{\Theta}$. The mean squared error, variance, and bias of the estimator are some of the most important quantities that help determining the performance of the estimator.

The resolution Ø of a variable for a detector is a chosen parameter which measures the performance of the estimator, and which roughly communicates how close we can expect $\hat{\Theta}$ to be to Θ. It is important to understand the relationship between the chosen Ø parameter and the estimator in order to correctly interpret the meaning of the resolution with respect to the performance of the detector.

Basic properties of estimators

The *error* is the difference between an estimate and the true value,

$$\mathrm{Err}(\hat{\Theta}) \equiv \hat{\Theta} - \Theta, \tag{1}$$

and the *mean squared error* is, as the name suggests, the mean of the squared errors,

$$\mathrm{MSE}(\hat{\Theta}) \equiv \left\langle \left(\hat{\Theta} - \Theta\right)^2 \right\rangle. \tag{2}$$

The *deviation* is the difference between an estimate and the expected value of the estimator,

$$\mathrm{Dev}(\hat{\Theta}) \equiv \hat{\Theta} - \langle\hat{\Theta}\rangle. \tag{3}$$

The *variance* of an estimator is the expected square of its deviations,

$$\mathrm{Var}(X) \equiv \langle(\hat{\Theta} - \Theta)^2\rangle. \tag{4}$$

The *standard deviation* is simply the squareroot of the variance,

$$\sigma \equiv \sqrt{\mathrm{Var}(X)} = \sqrt{\langle(\hat{\Theta} - \Theta)^2\rangle}. \tag{5}$$

The *bias* is the difference between the expected estimate and the true value,

$$B(\hat{\Theta}) \equiv \langle \hat{\Theta} \rangle - \Theta. \tag{6}$$

The mean squared error can also be expressed in terms of the variance and the bias,

$$\mathrm{MSE}(\hat{\Theta}) = \mathrm{Var}(\hat{\Theta}) + B(\hat{\Theta})^2, \tag{7}$$

for which reason the variance and bias will take the main focus from now on. Fig.1 is an example of a biased estimator with a Gaussian distribution.

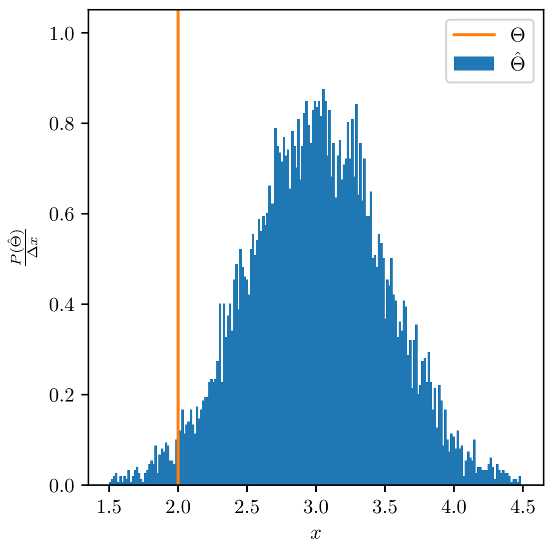


Figure 1: Example of the estimation of a parameter Θ, with $B(\hat{\Theta}) = 1$, $\mathrm{Var}(\hat{\Theta}) = 0.25$, and $\mathrm{MSE}(\hat{\Theta}) = 1.25$

Fig.2 is an example of a 2-dimensional biased estimator with a 2-dimensional Gaussian distribution.

Energy resolution

We will now compare typical estimators for the energy and direction of neutrino Cherenkov telescope events. These estimators are often probability density function based maximum likelihood estimators and are often called *reconstruction algorithms*, please take *estimate* to mean *reconstruction* here.

An observable of interest in neutrino events is its energy, E_ν. We consider the neutrino energy estimator $\hat{E}_\nu(\mathrm{data})$. Let's imagine that we receive a monoenergetic flux of neutrinos with energy $E_\nu = 40$ TeV. A distribution of the energy reconstruction could look like Fig.3. The performance of this estimator can be calculated by

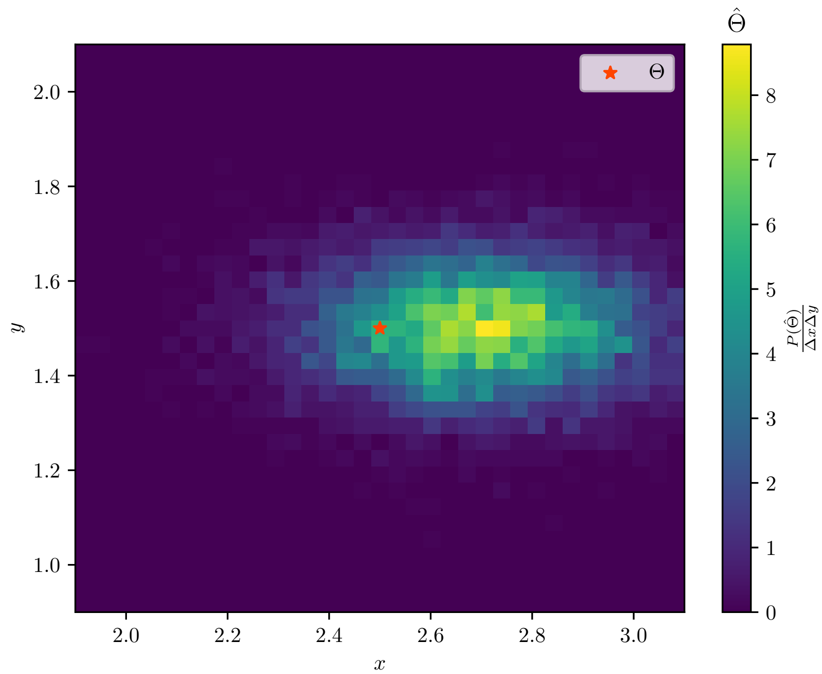


Figure 2: Example of the estimation of a parameter $\Theta = \begin{bmatrix}\Theta_x \\ \Theta_y\end{bmatrix}$, with $B(\hat{\Theta}) = \begin{bmatrix}0.2 \\ 0\end{bmatrix}$, $\mathrm{Var}(\hat{\Theta}) = \begin{bmatrix}0.04 \\ 0.01\end{bmatrix}$, and $\mathrm{MSE}(\hat{\Theta}) = \begin{bmatrix}0.08 \\ 0.01\end{bmatrix}$

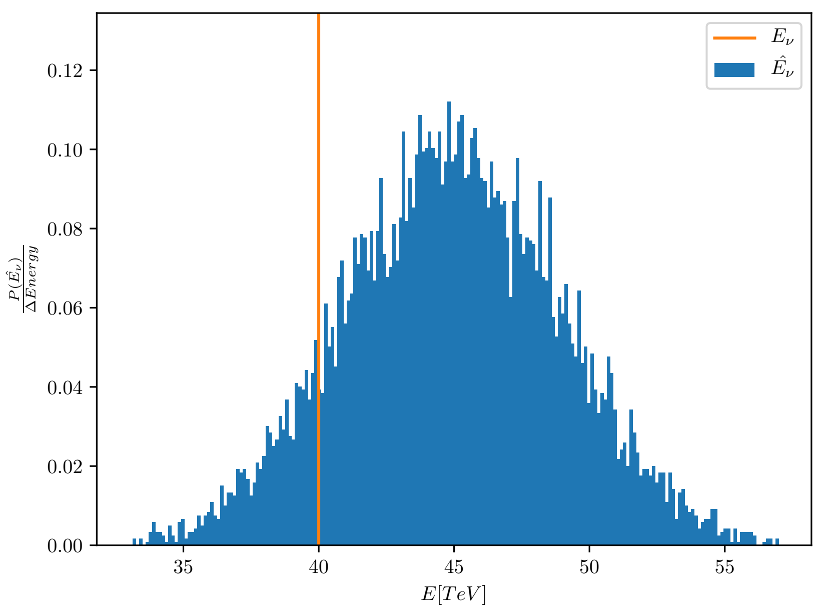


Figure 3: Example distribution of $\hat{E}_\nu$, with $B(\hat{E}_\nu) = 5$ TeV, $\mathrm{Var}(\hat{E}_\nu) = 16$ TeV^2, and $\mathrm{MSE}(\hat{E}_\nu) = 41$ TeV^2

direct calculation on a large sample of measurements using Eqn.6, Eqn.2, and Eqn.4,

$$\text{Bias}(\hat{E}_\nu) \simeq 5 \text{ TeV}$$
$$\text{MSE}(\hat{E}_\nu) \simeq 41 \text{ TeV}^2$$
$$\text{Var}(\hat{E}_\nu) \simeq 16 \text{ TeV}^2$$

However, we can also study the relationship between the performance of the estimator and the properties of its distribution, notably its mean μ and standard deviation σ:

$$\mu_{\hat{E}_\nu} = \langle \hat{E}_\nu \rangle = \text{Bias}(\hat{E}_\nu) + E_\nu \tag{8}$$
$$\sigma^2_{\hat{E}_\nu} = \langle \left(\hat{E}_\nu - \langle \hat{E}_\nu \rangle\right)^2 \rangle = \text{Var}(\hat{E}_\nu) \tag{9}$$

Clearly, neutrino telescopes deal with a large range of energies, so a normalization is needed to look at the performance over the whole energy range in single distribution. We therefore define a new quantity, the relative energy difference $\hat{E}_\nu / E_\nu$. Its distribution function could look like Fig.4 Again, we can find the relation between the properties of this distribution and the performance of the estimator.

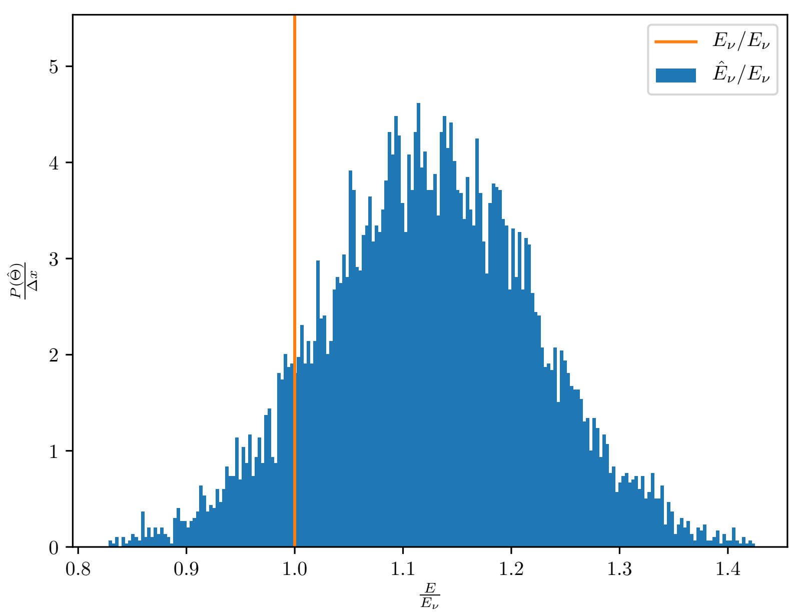


Figure 4: Example distribution of $\hat{E}_\nu / E_\nu$, with $B(\hat{E}_\nu) = 5$ TeV, $\text{Var}(\hat{E}_\nu) = 16 \text{ TeV}^2$, and $\text{MSE}(\hat{E}_\nu) = 41 \text{ TeV}^2$

$$\mu_{\hat{E}_\nu / E_\nu} = \frac{\langle \hat{E}_\nu \rangle}{E_\nu} = \frac{\text{Bias}(\hat{E}_\nu)}{E_\nu} + 1 \tag{10}$$
$$\sigma^2_{\hat{E}_\nu / E_\nu} = \frac{1}{{E_\nu}^2} \langle \left(\hat{E}_\nu - \langle \hat{E}_\nu \rangle\right)^2 \rangle = \frac{\text{Var}(\hat{E}_\nu)}{{E_\nu}^2} \tag{11}$$

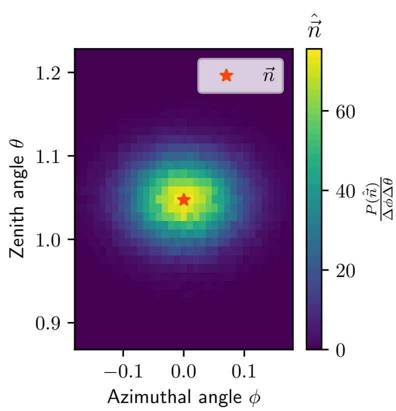
If the distribution is well behaved, i.e. Gaussian or otherwise single peaked and symmetric, the mean and standard deviation of $\hat{E}_\nu / E_\nu$ can be used as average and resolution respectively. When the distribution is non Gaussian, the median and FWHM, or other quantities can be preferred to measure the average and resolution of the estimator.

Angular resolution

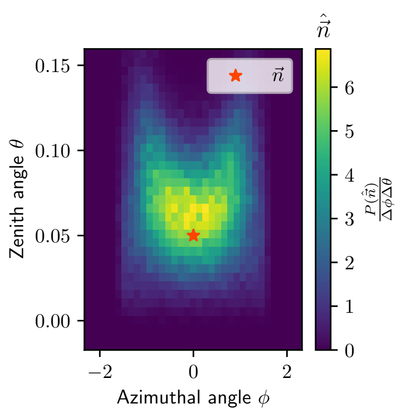
Another observable of interest for neutrino telescopes is the incoming direction of a neutrino, $\vec{n} = \begin{bmatrix} 1 \\ \theta \\ \phi \end{bmatrix}$, with estimate $\hat{\vec{n}} = \begin{bmatrix} 1 \\ \hat{\theta} \\ \hat{\phi} \end{bmatrix}$

We are interested in knowing the relationship between the true and estimated direction.

Let's imagine that we observe a flux of neutrinos from a given source, all with the same incoming direction. We must first choose a coordinate system in which to work. Spherical coordinates seem a natural choice, so let's plot a reasonable expectation for an estimator located right above the horizon, at a zenith angle of $\theta = \pi/3$, shown in Fig. **??**. If we repeat this experiment at a higher altitude, say $\theta = 0.05$, the distribution of the estimates looks like Fig. **??**. Clearly, spherical coordinates are not very useful since geometrical effects are introduced which depend on the location of the sky and not on the performance of the estimator. We would like to find variables



(a) Example distribution of $\hat{\vec{n}}$ in spherical coordinates with $\vec{n} = \begin{bmatrix} \pi/3 \\ 0 \end{bmatrix}$, with $B(\hat{\vec{n}}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\mathrm{Var}(\hat{\vec{n}}) = \mathrm{MSE}(\hat{\vec{n}}) = \begin{bmatrix} 3\text{ deg} \\ 2\text{ deg} \end{bmatrix}$



(b) Example distribution of $\hat{\vec{n}}$ in spherical coordinates with $\vec{n} = \begin{bmatrix} 0.05 \\ 0 \end{bmatrix}$, with $B(\hat{\vec{n}}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\mathrm{Var}(\hat{\vec{n}}) = \mathrm{MSE}(\hat{\vec{n}}) = \begin{bmatrix} 3\text{ deg} \\ 1.5\text{ deg} \end{bmatrix}$

to measure the performance of the estimator which remain invariant under rotations. We define the angle difference α along the zenith angle, and the orthogonal angle difference β,

$$\alpha \equiv \sin^{-1} |p_\alpha - \vec{n}| \tag{12}$$

$$\beta \equiv \sin^{-1} \left|p_\beta - \hat{\vec{n}}\right| \tag{13}$$

where we define

$$p_\alpha, p_\beta : \begin{cases} p_{\alpha\phi} & = \vec{n}_\phi \\ (p_\alpha - \vec{n}) \cdot (p_\beta - \hat{\vec{n}}) & = 0 \end{cases} \tag{14}$$

where p_α and p_β are two points along the line of sight such that $p_\alpha \cdot (p_\alpha - \vec{n}) = 0$ and $p_\beta \cdot (p_\beta - \hat{\vec{n}}) = 0$, see Fig. 6.[1] This yields

[1] If it seems surprising to you that $\theta_0 \neq \hat{\theta}$, imagine the same situation close to the zenith, here it is clear that $\hat{\vec{n}}$ is pointing at a different θ.

$$p_\alpha = \cos(\theta_0 - \theta) \begin{bmatrix} \sin\theta_0 \cos\phi \\ \sin\theta_0 \sin\phi \\ \cos\theta_0 \end{bmatrix} = \cos(\theta_0 - \theta)\vec{n}_0 \tag{15}$$

$$p_\beta = (\sin\theta_0 \sin\hat{\theta} \cos\Delta\phi + \cos\theta_0 \cos\hat{\theta}) \begin{bmatrix} \sin\theta_0 \cos\phi \\ \sin\theta_0 \sin\phi \\ \cos\theta_0 \end{bmatrix} = (\sin\theta_0 \sin\hat{\theta} \cos\Delta\phi + \cos\theta_0 \cos\hat{\theta})\vec{n}_0 \tag{16}$$

Now, again looking at the spread of $\hat{\vec{n}}$ in α and β in Fig. **??**, we discover a difference between the distribution at low and high θ. This time however, the effect is due to a real change in performance of the detector which was introduced in the sample.

Unfortunately, the expression for the zenith angle θ_0 of $\vec{n}_0$ is a very complicated function of θ, $\hat{\theta}$ and $\Delta\phi$ obtained by solving

$$\frac{\cos\theta}{\cos\theta_0} = \cos\Delta\phi \tan\hat{\theta} \sin(\theta_0 + \theta) + \cos(\theta_0 - \theta), \tag{17}$$

so it is not very useful to study the exact relationship between α, β and the spherical coordinates. However it is clear from Eqn.17, Eqn.15, and Eqn.16 that the two angle differences are functions $\alpha(\theta, \phi, \Delta\theta, \Delta\phi)$ and $\beta(\theta, \phi, \Delta\theta, \Delta\phi)$ where $\Delta\theta \equiv \hat{\theta} - \theta$ and $\Delta\phi \equiv \hat{\phi} - \phi$ are the errors of the zenith angle and azimuthal angle, respectively. In general, due to heterogeneities of the detector, we cannot expect the distributions of α and β to be constant nor proportional. In the context of neutrino telescopes however, the distributions of α and β are independent on the location of the source in the sky, which is evidenced by performing the test above for various sources. This assumption lets us assert that σ_α and σ_β are independent of θ and ϕ.

A typical reconstruction algorithm will perform as shown in Fig. **??**, albeit with non-Gaussian artifacts, but the main defining features are present.

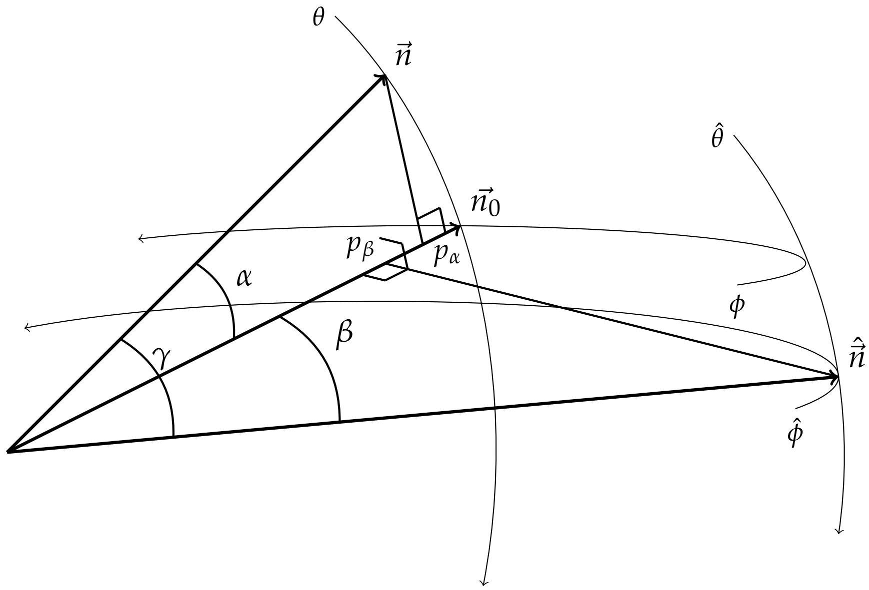
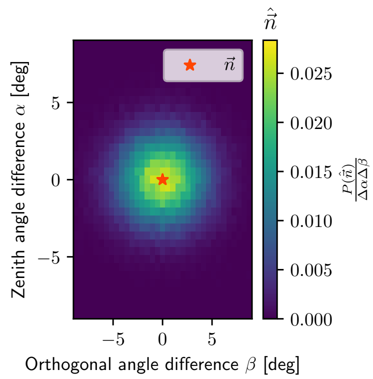
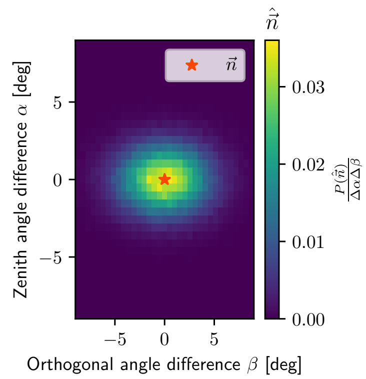


Figure 6: Relation between α, β, $\vec{n}$, $\hat{\vec{n}}$ and $\vec{n}_0$ in the spherical coordinate system.



(a) Example distribution of $\hat{\vec{n}}$ in α and β with $\vec{n} = \begin{bmatrix} \pi/3 \\ 0 \end{bmatrix}$, with $B(\hat{\vec{n}}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\text{Var}(\hat{\vec{n}}) = \text{MSE}(\hat{\vec{n}}) = \begin{bmatrix} 3 \text{ deg} \\ 2.0 \text{ deg} \end{bmatrix}$



(b) Example distribution of $\hat{\vec{n}}$ in α and β with $\vec{n} = \begin{bmatrix} 0.05 \\ 0 \end{bmatrix}$, with $B(\hat{\vec{n}}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\text{Var}(\hat{\vec{n}}) = \text{MSE}(\hat{\vec{n}}) = \begin{bmatrix} 3 \text{ deg} \\ 1.5 \text{ deg} \end{bmatrix}$

- $\hat{\vec{n}}$ is *unbiased*, $\langle \alpha \rangle = \langle \beta \rangle = 0$
- The distribution of $\hat{\vec{n}}$ is symmetrical about $\vec{n}$, $\text{Var}(\alpha) \simeq \text{Var}(\beta)$

With these assumptions in mind, the performance of the estimator can safely be measured with a single parameter. We use the following parametrization for a *total* angular difference (from now on simply *angular difference*) between the true and estimated neutrino direction,

$$\gamma \equiv \cos^{-1}\left(\hat{\vec{n}} \cdot \vec{n}\right) \tag{18}$$

$$= \cos^{-1}\left(\cos\Delta\phi\cos\Delta\theta + \cos\theta\cos\hat{\theta}(1-\cos\Delta\phi)\right) \tag{19}$$

$$= \cos^{-1}\left(\frac{1-\cos^2\beta(\tan^2\alpha - 1)}{2}\right) \tag{20}$$

We see that γ is related to α and β only. Knowing that β and α are independent of the location in the sky, γ must be too. γ is also related to the *errors* of θ and ϕ with an added θ and $\hat{\theta}$ dependence as evidenced in Fig. **??**. Let's now take a look at the distribution of γ in Fig. 8 and connect its properties to the performance of the estimator. The distribution vanished at the true value, which can seem odd

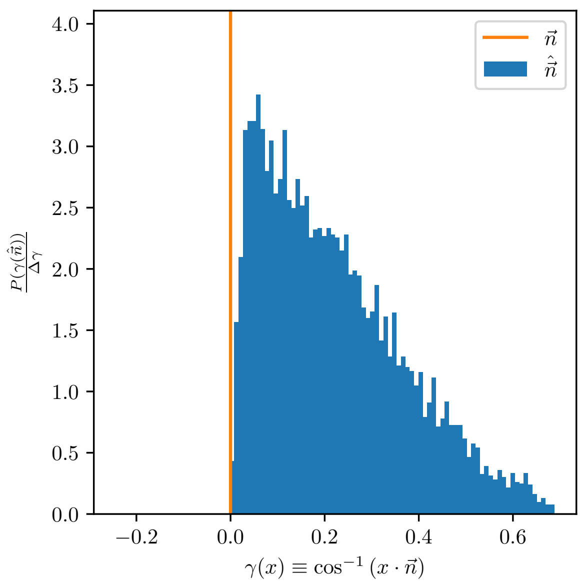


Figure 8: Example distribution for the angular difference γ, with $B(\hat{\vec{n}}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\mathrm{Var}(\hat{\vec{n}}) = \mathrm{MSE}(\hat{\vec{n}}) = \begin{bmatrix} 10\text{ deg} \\ 10\text{ deg} \end{bmatrix}$

since is is clear from Fig 6 that this is where most of the estimates are. This is because the histogram has been plotted in equal bins of γ, rather than the more natural equal bins of solid angle. We account for this effect in Fig. 9 We can now recognize the symmetrical profile of γ, which in this idealized case is the half-Gaussian distribution. Finally, we identify the properties of this distribution to the properties of the estimator.

$$\mu_{\gamma\text{corr}} = \frac{\sigma_\alpha\sqrt{2}}{\sqrt{\pi}} = \frac{\sigma_\beta\sqrt{2}}{\sqrt{\pi}} \tag{21}$$

$$\sigma^2_{\gamma\text{corr}} = \sigma^2_\alpha\left(1-\frac{2}{\pi}\right) = \sigma^2_\beta\left(1-\frac{2}{\pi}\right) \tag{22}$$

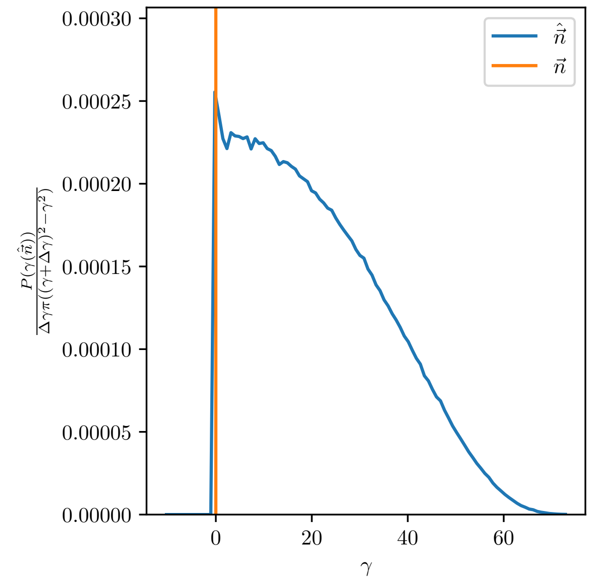


Figure 9: Example corrected distribution for the angular difference γ, with $B(\hat{\vec{n}}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\mathrm{Var}(\hat{\vec{n}}) = \mathrm{MSE}(\hat{\vec{n}}) = \begin{bmatrix} 10\text{ deg} \\ 10\text{ deg} \end{bmatrix}$

Notice how the different parameters chosen for the resolution result in very different interpretations of their distributions. Typically, the angular difference will not follow a half-Gaussian distribution, where, again, the median or other variables can be preferred to give information about the estimator's variance.

Conclusion

The distributions of quantities related to energy and direction estimation have non-trivial relationships to the properties of the estimators. In the case of the relative energy difference, the mean and spread of the distribution inform the bias and standard deviation of the energy estimator respectively. In the case of angular difference, the mean and spread of the distribution inform the standard deviation and the variance of the direction estimator respectively. Due to finite statistics, every of these quantities have associated errors which can also be estimated, all of which need to be assessed thoroughly in any analysis of reconstruction to provide the full picture of how well the detector is performing.