# 1 Introduction and Summary

Here we present a review of a method for treating highly correlated sets of nuisance parameters. In short, one can transform the nuisance parameter basis into one where the parameters are uncorrelated. Then, we find whih new parameters have the largest effect on ones reconstructed signal, and identify the minimum number that will be necessary to span the space of shapes presentable by the full set of nuisance parameters in your reconstructed space.

# 2 De-correlating systematic uncertainties

Consider a set of $n$ nuisance parameters $\vec{x}$, each with mean $\mu$. A $n \times n$ matrix, $\Sigma$ defines the positive-definite covariance between these nuisance parameters. If we assume the distribution is jointly normally distributed, then the probability density function defining the odds of sampling a set of nuisance parameters, $\vec{y}$, is given by

$$P(\vec{y}) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left[-\tfrac{1}{2}(\vec{y} - \vec{\mu})^T \Sigma^{-1} (\vec{y} - \vec{\mu})\right] \tag{1}$$

where $|\Sigma|$ is the determinant of $\Sigma$. We can first consider the transformation

$$\vec{p} \equiv \vec{y} - \vec{\mu} \tag{2}$$

and consider instead the *log* of the likelihood. This allows us to drop off the prefactors, since we ultimately will only care about the differences in log likelihoods. What we then have is

$$\ln P \equiv LLH \sim -\tfrac{1}{2}\vec{p}^T \Sigma^{-1} \vec{p} \tag{3}$$

We can now, by construction, define a transformation $U$, satisfying $UU^T = I$, and a diagonal matrix $A$ such that

$$A = U\Sigma^{-1}U^T \qquad\qquad LLH \sim -\tfrac{1}{2}(U\vec{p})^T U\Sigma^{-1}U^T(U\vec{p}) \tag{4}$$

and we recover

$$LLH \sim -\tfrac{1}{2}\vec{p}'^T A \vec{p}', \tag{5}$$

where the primed vectors refer to the vectors transformed by $U$. The left side of equation (4) can be re-expressed as

$$A = U\Sigma^{-1}U^T \qquad\qquad \rightarrow \qquad\qquad A' = U^T\Sigma U, \tag{6}$$

or, since $A$ is diagonal,

$$\Sigma U = U \begin{bmatrix} \lambda_! & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}. \tag{7}$$

If we re-write the transformation matrix $U$ as a vector of its column vectors $\eta_i$, $U = \begin{bmatrix} \eta_1 & \dots & \eta_n \end{bmatrix}$ then Equation can be expressed as

$$\Sigma\eta_i = \lambda_i\eta_i. \tag{8}$$

This tells us that the column vectors of the transformation matrix $U$ are the right eigenvectors of the covariance matrix $\Sigma$. And each of the corresponding eigenvalues of $\Sigma$ are the entries along the diagonals in $A'$.

So, for a given covariance matrix $\Sigma$, we solve for its eigenvalues and eigenvectors. The eigenvalues each represent a gaussian prior width for new, uncorrelated, parameters in the new basis. The matrix of the eigenvectors allows for transformations from the correlated to the uncorrelated basis. Similarly, the inverse of that eigenvector matrix facilitates transformations back from the uncorrelated basis to the correlated basis.

# 3  Ranking

Moving forwards, we need a function that takes a set of correlated nuisance parameters $\vec{x}$ and calculates a binned expectation. Generally this is done in three steps

1. Generate a Monte Carlo sample

2. Use the Monte Carlo sample to calculate a binned central expectation for your sample

3. For each parameter $i$, and for each event $j$, calculate the 1D gradient for the weight about the central expectation, $\left.\frac{\partial w_j}{\partial x_i}\right|_{central}$

Then each parameter could in principle be perturbed independently, a new weight calculated for each simulated event, and a new overall expectation calculated. Since the nuisance parameters are correlated, this isn't a good idea.

Using the methods of Section 2, we solve for the eigenvalues and vectors of the covariance matrix to determine the transformation to the un-correlated basis and the widths of the uncorrelated priors. For each uncorrelated parameter, we transform its $1\sigma$ deviation from its central value into the correlated basis, and use this set of nuisance parameters to calculate an expectation from the Monte Carlo gradients.

We can then use a likelihood function, such as Equation9, to calculate the impact of the uncorrelated parameter.

$$-\Delta LLH = \tfrac{1}{2} \sum_{i=0}^{bins} \left( \frac{p_i - \mu_i}{\sigma_i} \right)^2, \tag{9}$$

summing over the $i$ bins, where $p_i$ is the perturbed value in a bin, $\mu$ is its mean, and $\sigma$ is the expected variance in the bin. For a Poissonian error, $\sigma = \sqrt{\mu}$.

# 4  Trimming

# 5  Daemonflux

Daemonflux REF is combines the uncertainties on both the overall cosmic ray flux and the hadronic production rates in the cosmic ray air showers that follow when they impact our atmosphere. It is comprised of six "global spline fit" parameters and seventeen more that relate to pion and kaon production at various energies. These parameters are highly correlated, as shown in Figure 5.

Write about the gradient extraciton.

We proceed with calculating a new basis as in the previous sections, and calculate the impact of the new orthogonal parameters. This is done for the cascades sample developed by the diffuse

Daemon Correlation
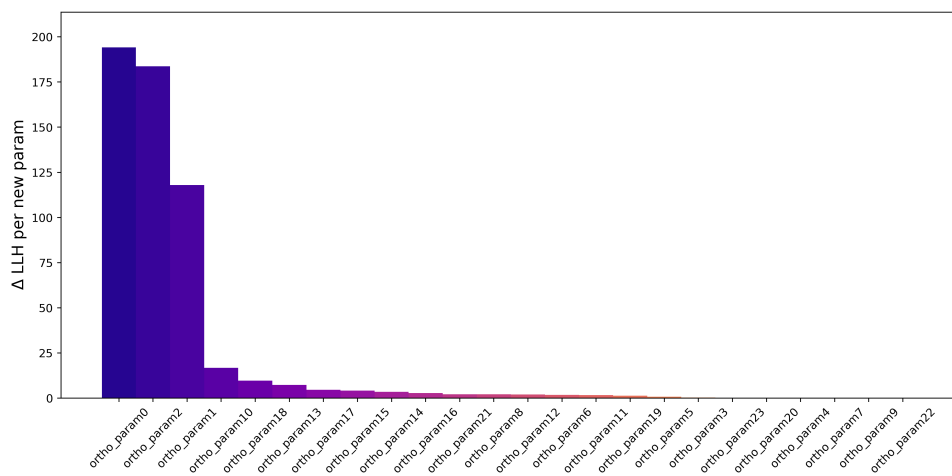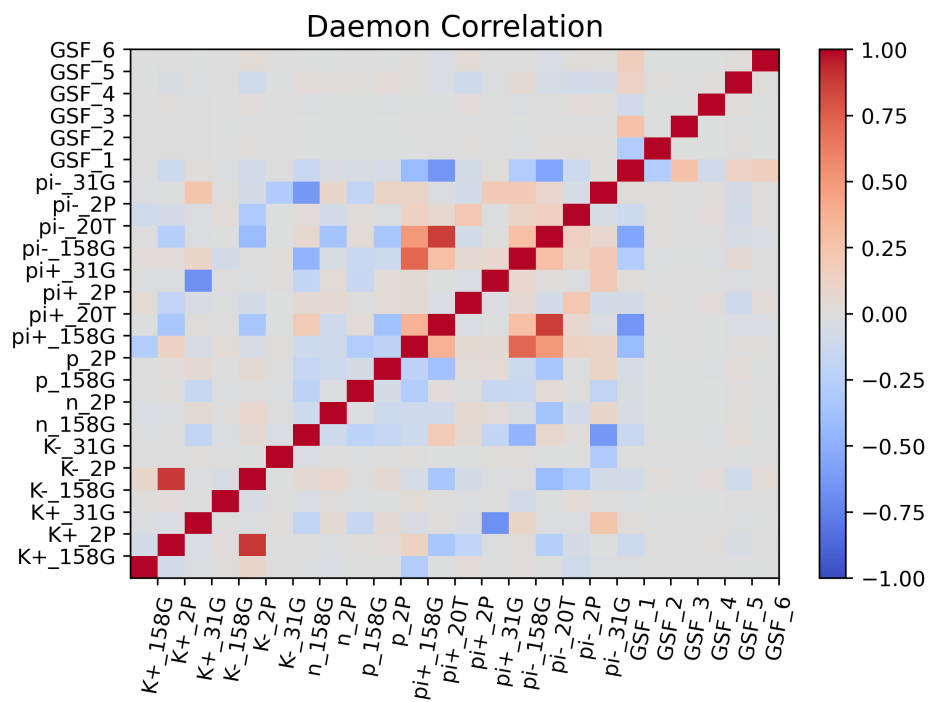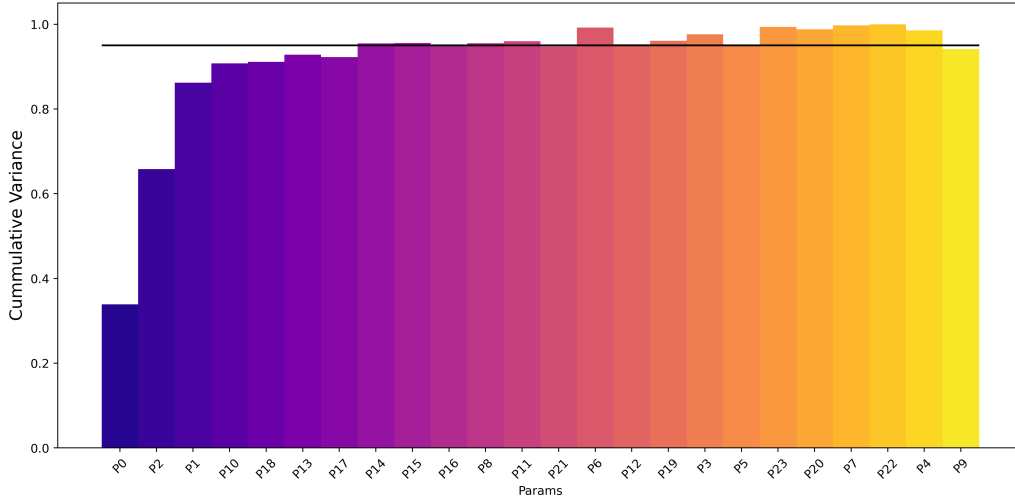
group, and used in ongoing MEOWS analyses[1]. These are shown in Figure 5. As is evident, the first few parameters have a much greater impact than those that follow.

## 5.1 Variance Testing

We first allow only the parameter with the biggest pull to vary, or "ortho_param0". We then sample this value 10,000 times and calculate the $\Delta LLH$ at each sample. We note the threshold value of $\Delta LLH$ below which 90% of the samples $\Delta LLH$'s lie. This process is then repeated for the two parameters with the largest pulls, and the new threshold noted. This is repeated until all 23 parameters are being varied at each step. The $\Delta LLH$ thresholds, normalized to the one varrying the most, are shown in Figure 5.1, where a horizontal line has been drawn at 0.95. From this figure, we gather that seven or eight parameters are needed to account for 95% of the variance in the whole ensemble.

## 5.2 Shape Testing

It is also important to ensure the full configuration of shapes achievable, in reconstructed space, with a 23 parameter ensemble is also possible to achieve with a reduced number of parameters. To test this, we sample all 23 parameters randomly. Then, by allowing only the $m$ strongest pulling parameters to vary, we carry out a fit to the fully shuffled expectation while leaving the last $23 - m$ parameters fixed at their central values. The distributions of best-fit $\Delta LLH$, are shown in Figure 5.2. From this, most shapes appear to be realizable with as few as five parameters, although with some occasional deviations. Once seven parameters are allowed to vary, no discrepancies are observed between fits and expectations generated with a full 23-parameter sampling.

---

[1]MEOWS-C-2023