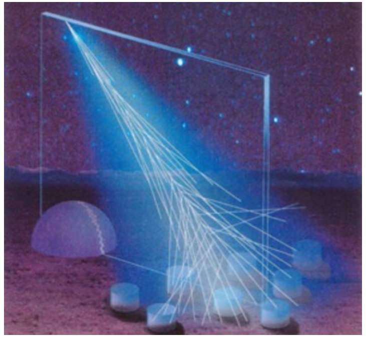
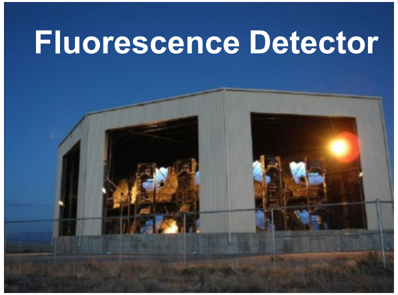
Bayesian hierarchical cross-calibration of hybrid air-shower detectors

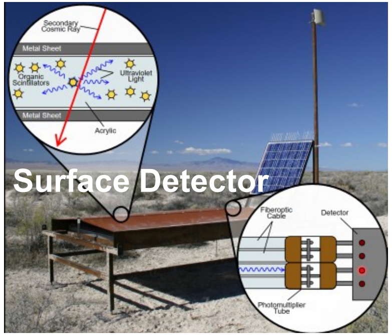
Anton Prosekin, Anatoli Fedynitch, Kozo Fujisue
Academia Sinica, Taipei, Taiwan

TeVPA 2026,
Aug 30 - Sept 04, 2026

Hybrid air shower detectors







Surface Detectors (SD):

- Operate continuously (~100% duty cycle)
- Large event statistics but higher uncertainty in energy estimation

Fluorescence Detectors (FD):

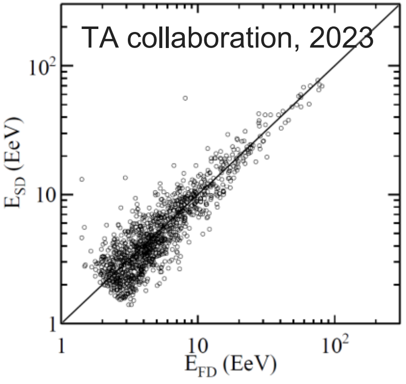
- Operate only on clear, moonless nights (~10–15% duty cycle)
- Provide nearly calorimetric energy measurement with smaller uncertainty

Hybrid Detection:

- Events observed simultaneously by SD and FD
- Used to calibrate the SD energy scale

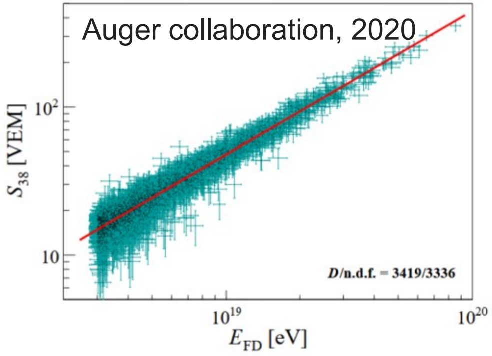
SD-FD calibration approaches

Least square fit



$$E_{FD} = E_{SD}/1.27$$

Loglikelihood minimization

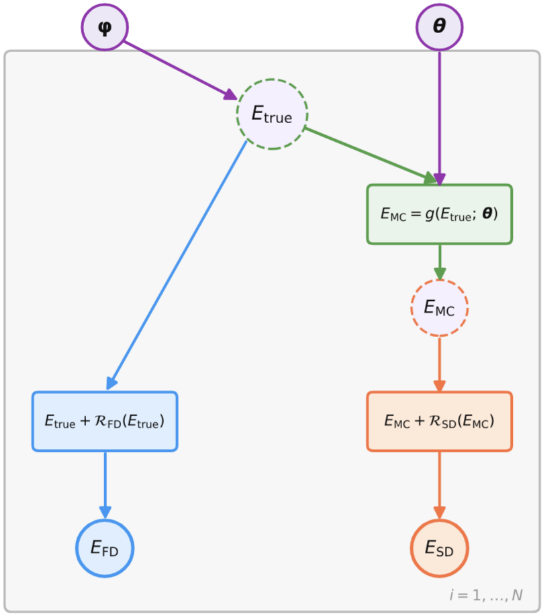


$$E_{FD} = A\ {S_{38}}^{B}$$

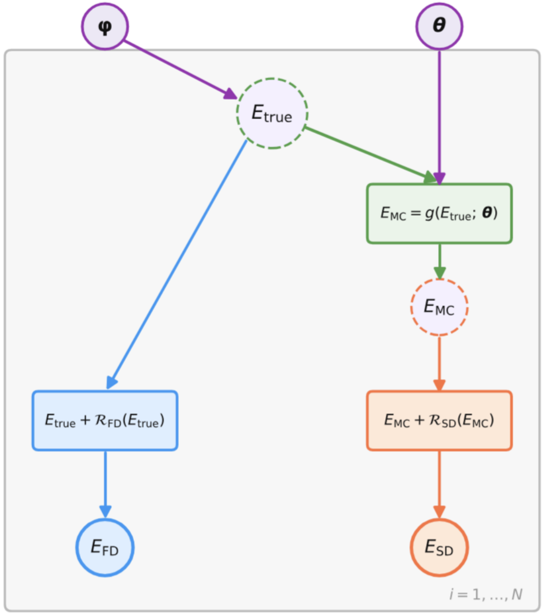
$$A = (1.86 \pm 0.03) \times 10^{17}\ \text{eV}$$

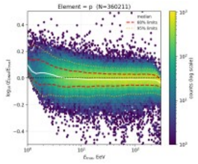
$$B = 1.031 \pm 0.004$$

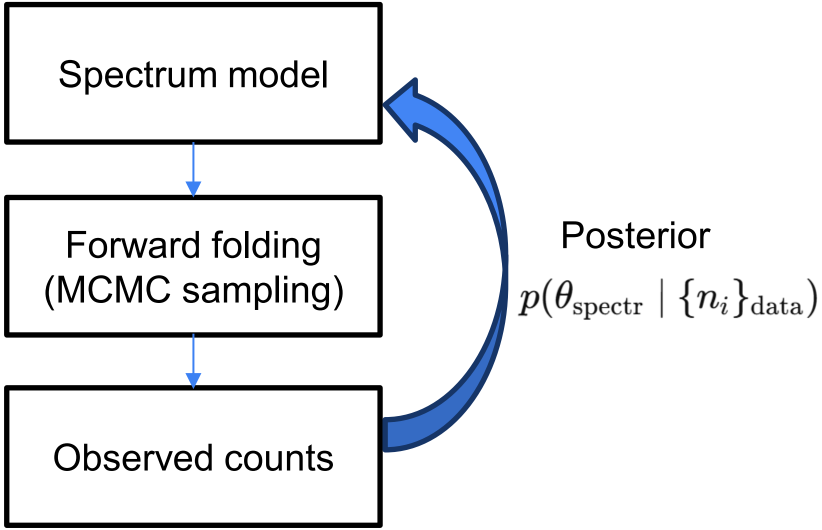
Hierarchical MCMC



Motivation: Calibration and forward folding



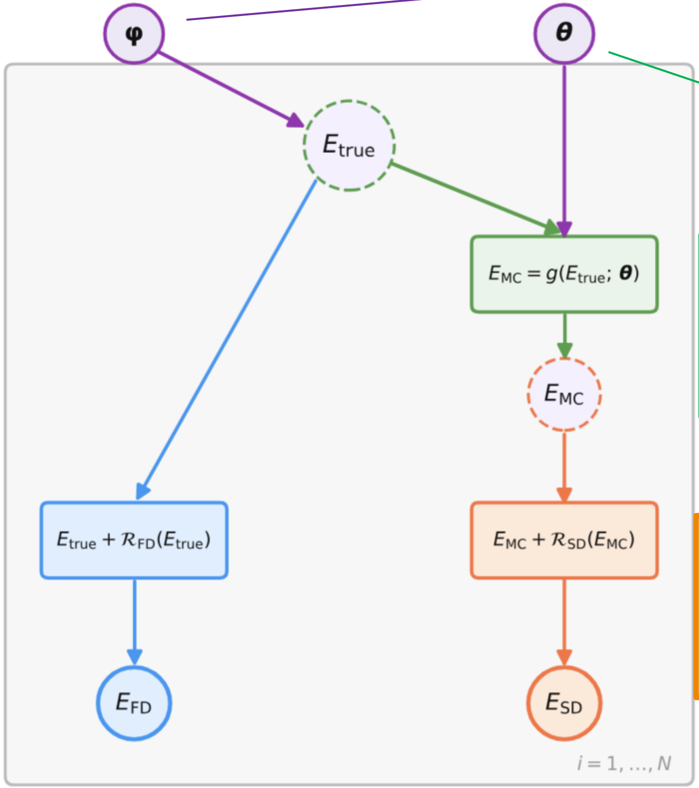




Kozo Fujisue, ICRC 2025

- MCMC calibration yields posterior distributions of calibration parameters
- Enables propagation of these uncertainties into forward folding

Bayesian model



Hierarchical prior

$$\lg E_{\text{true},i} \sim \text{SkewNormal}(\xi,\ \omega,\ \alpha)$$

$$z_i \sim \text{SkewNormal}(0,\ 1,\ \alpha)$$

$$\lg E_{\text{true},i} = \xi + \omega\, z_i$$

Calibration function

$$\lg E_{\text{MC},i} = g(\lg E_{\text{true},i};\ a, b, c)$$

$$g(\lg E;\ a, b, c) = a\,(\lg E)^2 + (1 + b)\ \lg E + c$$

FD residual

$$\lg E_{\text{FD},i} \sim \mathcal{N}(\lg E_{\text{true},i},\ \sigma_{\text{FD}})$$

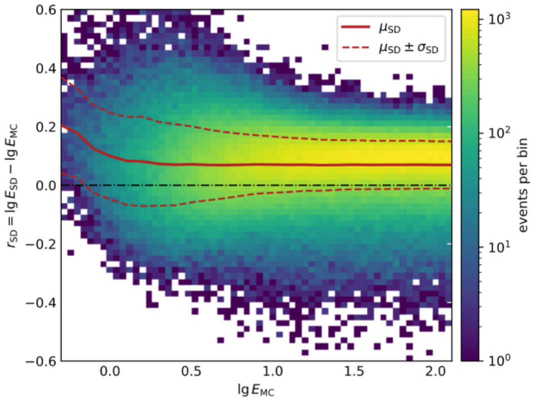
SD reconstruction residual model

$$\lg E_{\text{SD},i} \sim \lg E_{\text{MC},i} + \mathcal{R}_{\text{SD}}(\lg E_{\text{MC},i})$$

$$R_{\text{SD}}(x \mid \lg E_{\text{MC}}) = \mathcal{N}(x \mid \mu_{\text{SD}}(\lg E_{\text{MC}}), \sigma_{\text{SD}}(\lg E_{\text{MC}}))$$

SD reconstruction residuals

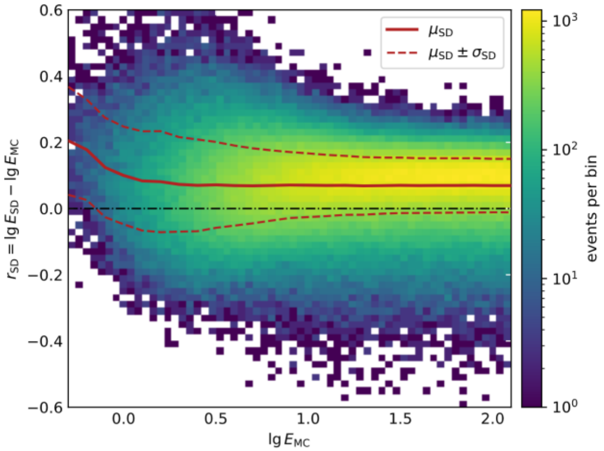
- Reconstruct sample of MC events
- Take only events that pass trigger and quality cuts → allows to include cuts
- Histogram residuals:

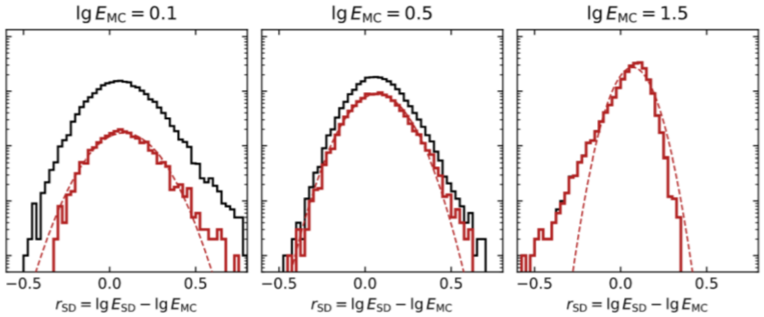


- Stan Hamiltonian MCMC requires analytical representation of the model
- Parametrize the SD residual $\mathbf{r = lg(E_{SD} / E_{true})}$
  - fit the conditional $\mathbf{p(r \mid lg\ E_{true})}$ in true energy bins
  - fit the energy dependence of distributions parameters with spline

SD reconstruction residuals

- Residual distributions are non-Gaussian with long tails
- We tested Gaussian mixture, Student-t mixture and skewed Student-t
- The calibration sensitive only to first two moments
    - Gaussian with **the moments** as μ(E) and σ(E) is enough for calibration with N=1500 events
    - Note: It is not direct Gaussian fit! The fit finds μ(E) and σ(E) different from the moments





Toy simulations

- Model reconstruction residuals after trigger and quality cuts as **non-Gaussian** with long tails
- Sample 400k events and build residual **Gaussian distribution for Stan**
- Sample another pool of events and get a true energies for each event

$$\lg E_{\mathrm{true},j} = g^{-1}(\lg E_{\mathrm{MC},j}; a, b, c) \qquad g(x; a, b, c) = ax^2 + (1+b)x + c$$

- For each hybrid data sample of 1500 events draw from (importance sampling):

$$\lg E_{\mathrm{true},i} \sim \mathrm{SkewNormal}(\xi,\ \omega,\ \alpha)$$

- Generate for each event observed:

$$\lg E_{\mathrm{FD},i} \sim \mathcal{N}(\lg E_{\mathrm{true},i},\ \sigma_{\mathrm{FD}})$$

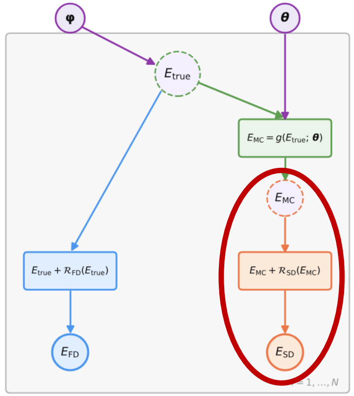
$$\lg E_{\mathrm{SD},i} \sim \lg E_{\mathrm{MC},i} + \mathcal{R}_{\mathrm{SD}}(\lg E_{\mathrm{MC},i})$$

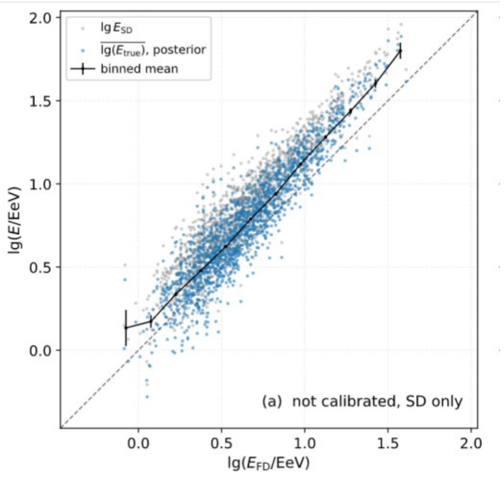
- Test “round trip” inference to get parameters of the generated dataset
- Test hidden true energies

Parameters
$a_{\mathrm{true}} = 0.03$
$b_{\mathrm{true}} = 0.03$
$c_{\mathrm{true}} = 0.10$
$\sigma_{\mathrm{FD}} = 0.043$
$\xi = 0.30$
$\omega = 0.45$
$\alpha = 3.0$
$N_{\mathrm{events}} = 1500$

E_{true} inference: MC scale

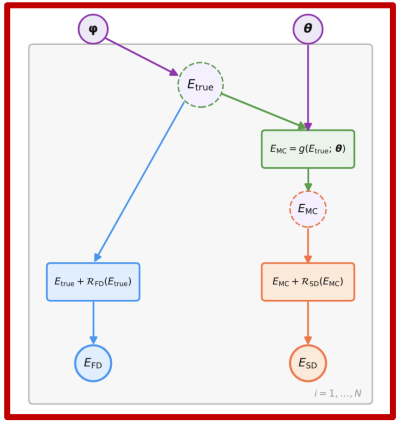
- Use only SD residuals
  - Shift from SD measurements
  - True energies E_{true} on MC energy scale
  - SD residuals can account only for MC bias and variance

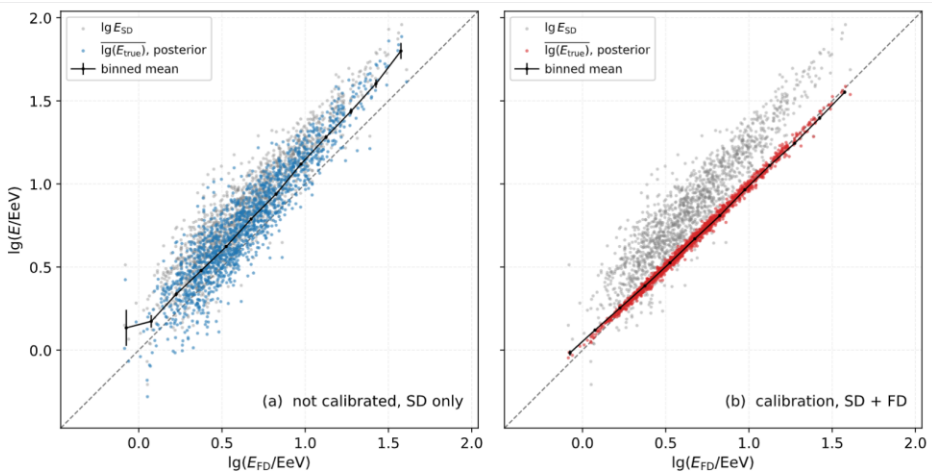




E_{true} inference: full calibration

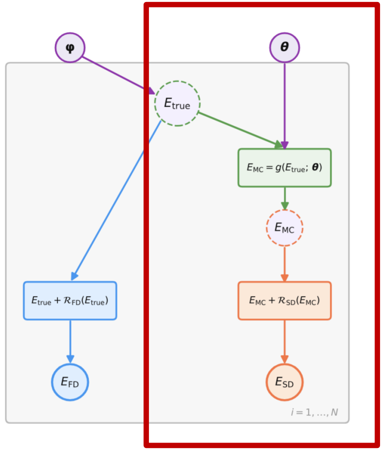
- Full calibration that use both SD and FD:
  - True energies E_{true} on FD energy scale
  - Variance is narrower than SD and FD separately
  - Distribution the same as for thrown energies

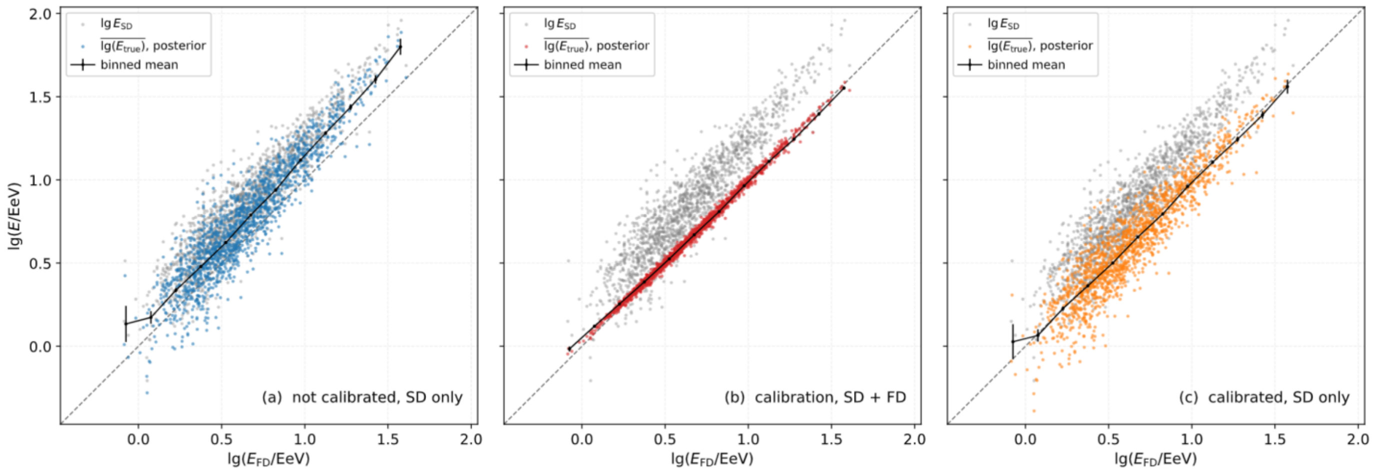




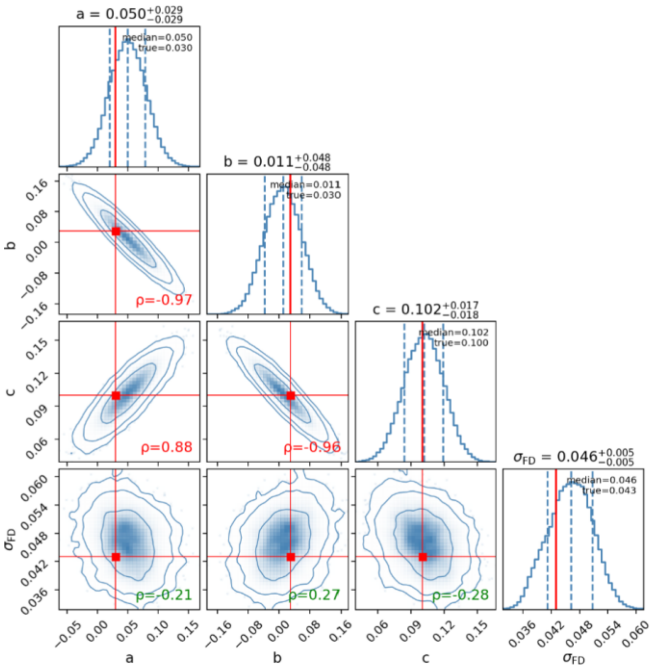
E_{true} inference: calibrated SD energies

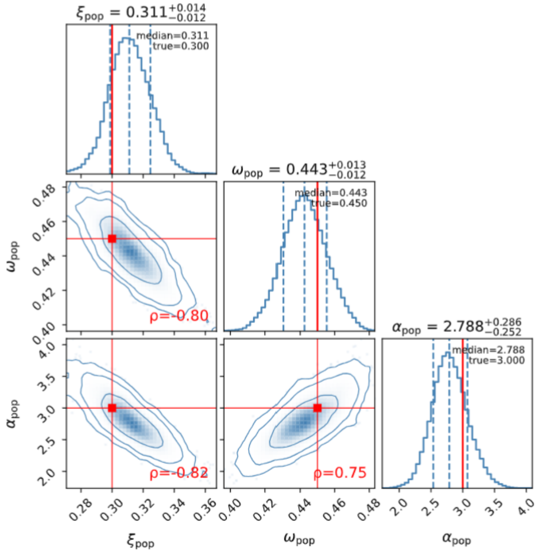
- Use calibrated SD:
  - True energies E_{true} on FD energy scale
  - Variance reflects SD's variance (wider than in case SD+FD)





Posterior and recovery of parameters



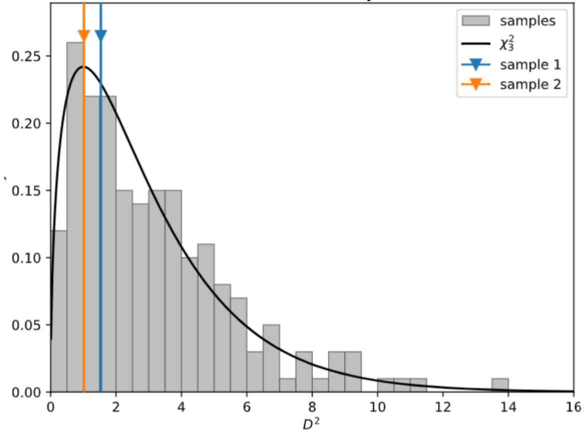


- Recovery inside 68% of posterior but not at the centre
- Position of posterior relative to true values depends on fluctuations of the sample
- Strong correlation because of the polynomial form
- Position of posterior doesn’t sensitive to reparameterization of the calibration function

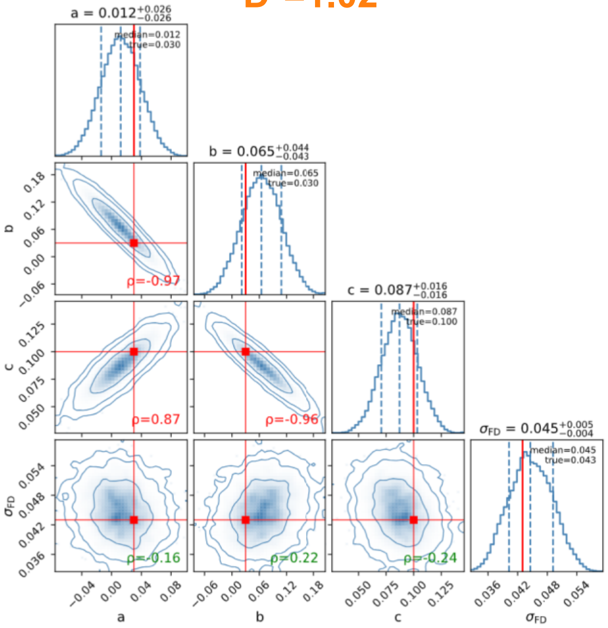
Recovery for different samples

- Generate 200 toy samples and test statistical properties
- Mahalanobis distance: $D^2 = (\boldsymbol{\theta}_0 - \bar{\boldsymbol{\theta}})^{\mathsf{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\theta}_0 - \bar{\boldsymbol{\theta}}), \qquad \boldsymbol{\theta} = (a, b, c)$
- Corner plots might look worse in spite of smaller fluctuations

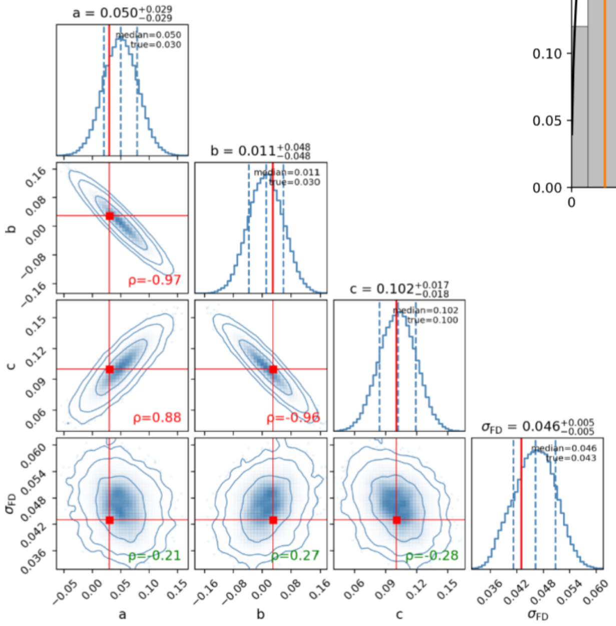
Mahalanobis distance
for 200 samples



D^2=1.02



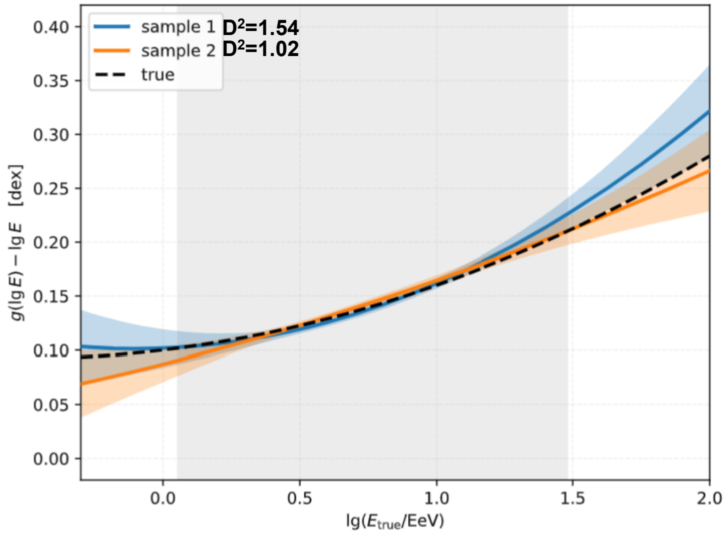
D^2=1.54



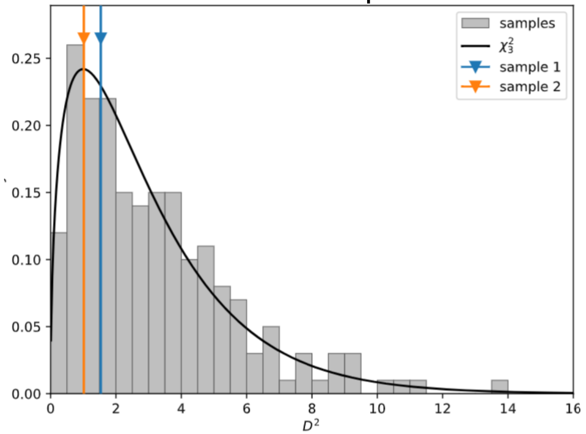
Calibration function inference

- Mahalanobis distance: $D^2 = (\boldsymbol{\theta}_0 - \bar{\boldsymbol{\theta}})^\mathsf{T}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}_0 - \bar{\boldsymbol{\theta}}), \qquad \boldsymbol{\theta} = (a, b, c)$

Calibration function

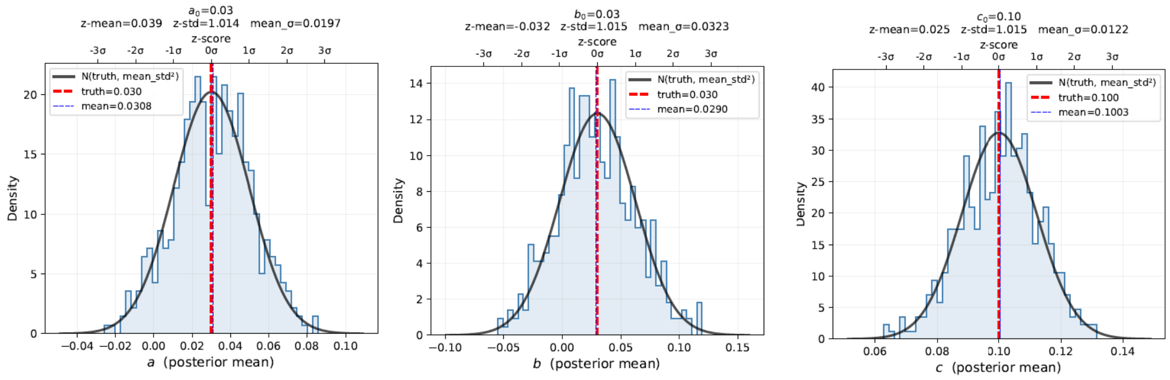


Mahalanobis distance
for 200 samples



- Mahalanobis distance defines which posterior contains true function in narrower band
  - Provide upper limit for the minimum width of posterior band to contain true function everywhere
- Posterior are wider outside data region but still contain true function in extrapolated region

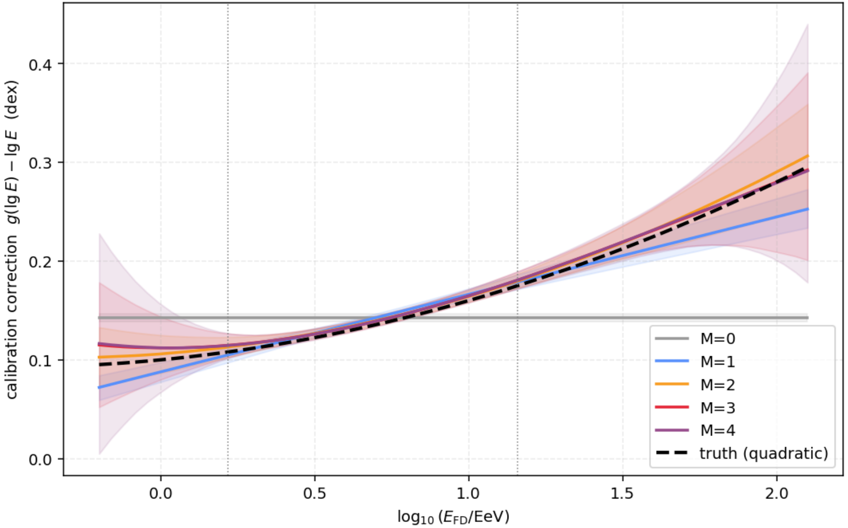
Posterior means and uncertainties



- Check means of 200 samples
- Pull means consistent with zero indicate no bias
- Pull widths consistent with 1:
  - indicate correctly estimated posterior uncertainties and nominal coverage
- the 68% credible intervals contain the true values in approximately 68% of the toy samples

Calibration function complexity

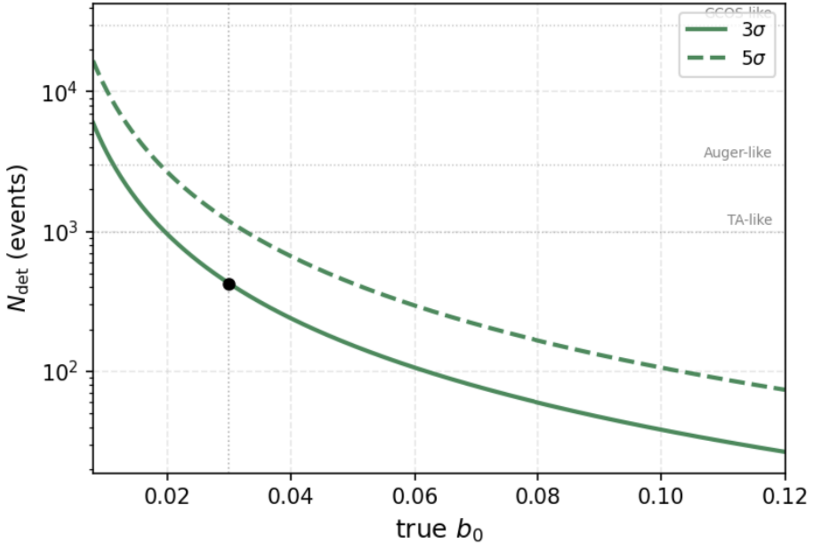
- Stan allows to compare models via leave-one-out (LOO) method
- Generate toy data with a quadratic calibration (a=b=0.03, c=0.10)
- Compare metric improvements for Bernstein polynomials of degree M = 0…5 (M=2 is the quadratic truth)

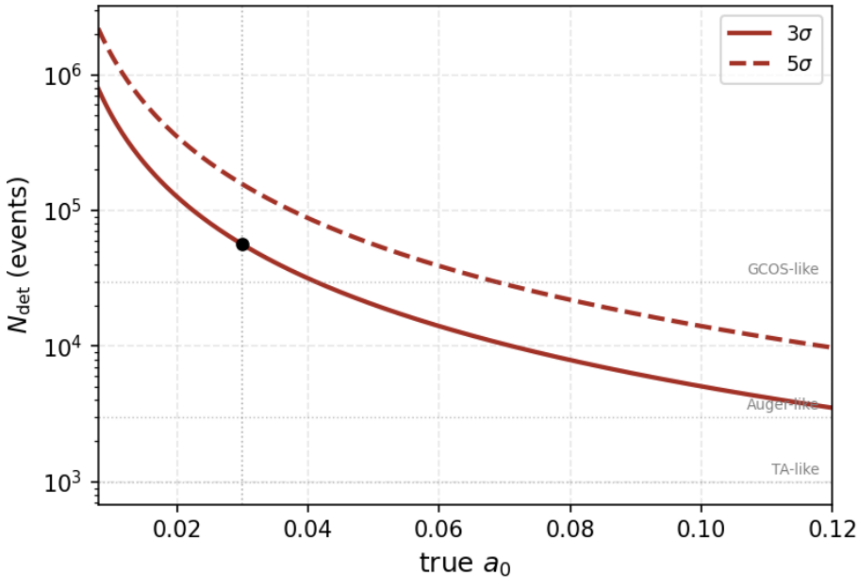


Even though the truth is quadratic, LOO stops at linear at N≈1000: “a” is below the LOO noise floor

Number of events vs complexity

- Detectability: does adding a term improve out-of-sample predictive fit? (nested model comparison)
- Metric: LOO cross-validation score ΔELPD (significance z = ΔELPD/SE)
- Number of events needed to detect a term at z=3σ and 5σ vs its true amplitude using z ~ $\sqrt{N}$

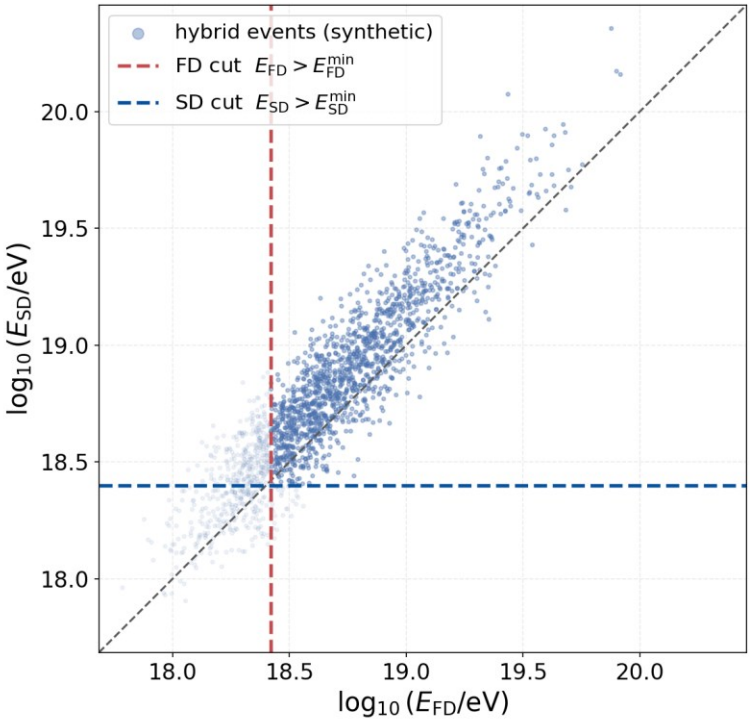




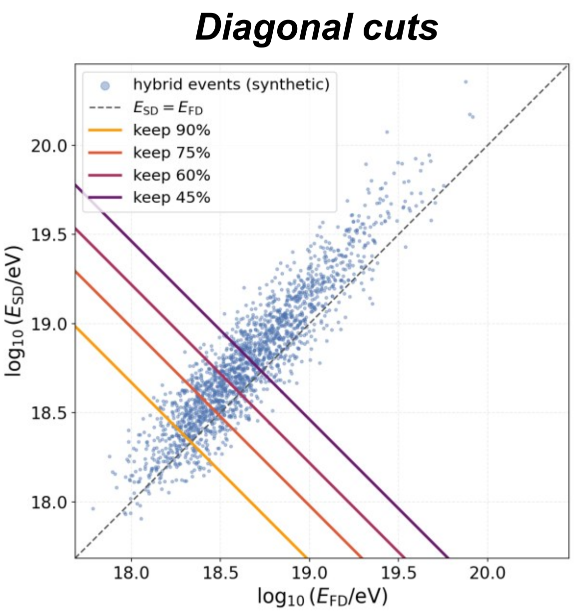
- Linear model is detected at 400 events
- Quadratic model is detected 60k events

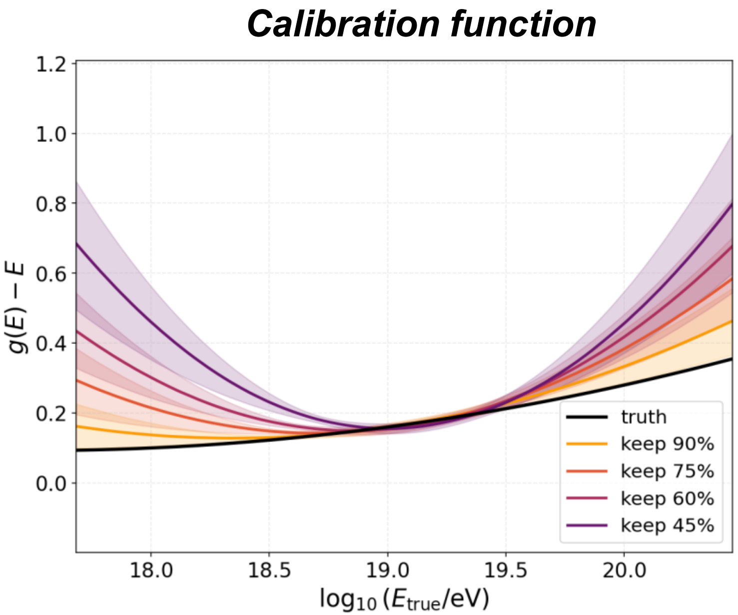
Cuts in the FD-SD energy plane

- SD residuals account for trigger and quality cuts but not for **energy cuts**
- Hybrid data sets might have hard low energy cuts
    - FD cuts
    - SD cuts
    - Diagonal cuts
- Upward fluctuation pass the cut
- Downward fluctuations removed



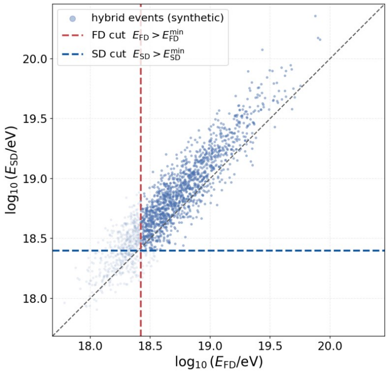
Test: diagonal cuts





- Synthetic data, known truth for calibration function
    - 45° cut $\lg E_{FD} + \lg E_{SD} > t$, four depths
- For quadratic function lower energy bias influence high energies as well
- Ignore the cut → bias grows with depth

Truncated likelihood



- Fix: normalize conditional distribution to account for lost events
  - Example for FD cut modelled by Θ function

$$p_{\rm cut}(x_{\rm FD} \mid \lg E) = \frac{\Theta(x_{\rm FD} - x_{\rm cut})\, \mathcal{N}(x_{\rm FD} \mid \lg E,\ \sigma_{\rm FD})}{Z}$$

  - Normalization factor:

$$Z = \int \Theta(x - x_{\rm cut})\, \mathcal{N}(x \mid \lg E, \sigma_{\rm FD})\, dx = \int_{x_{\rm cut}}^{\infty} \mathcal{N}(x \mid \lg E, \sigma_{\rm FD})\, dx = \Phi\left(\frac{\lg E - x_{\rm cut}}{\sigma_{\rm FD}}\right)$$

- Θ implements the cut; it automatically sets the limits of the normalizing integral
- Per event: add $(-\log Z)$ to the log-likelihood in Stan
- Z depends on the parameters → cannot be dropped as a constant

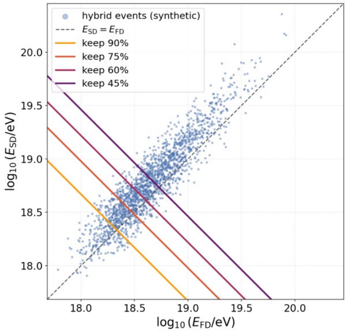
Truncated likelihood: diagonal cut

- Cut $w_{\mathrm{FD}}\, x_{\mathrm{FD}} + w_{\mathrm{SD}}\, x_{\mathrm{SD}} > t$ (FD, SD independent Gaussians given $\lg E_{\mathrm{true}}$)

$$Z = \iint \Theta(w_{\mathrm{FD}} x + w_{\mathrm{SD}} y - t)\, \mathcal{N}(x \mid \cdot)\, \mathcal{N}(y \mid \cdot)\, dx\, dy = \Phi\left(\frac{\mu_Z - t}{\sigma_Z}\right)$$

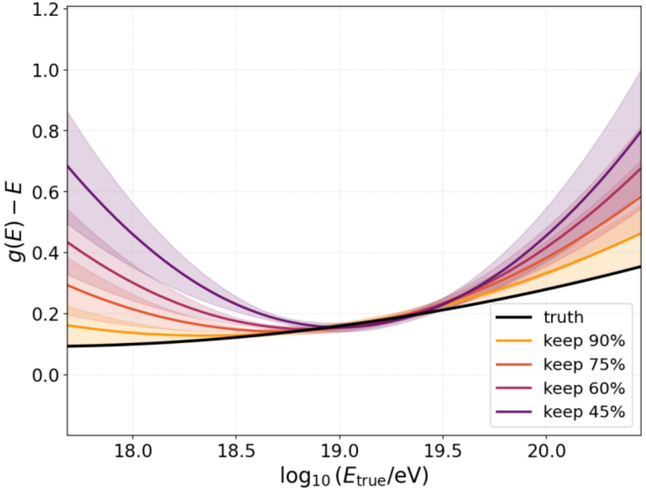
$$\mu_Z = w_{\mathrm{FD}}\ \lg E + w_{\mathrm{SD}}\, g(\lg E), \qquad \sigma_Z^2 = w_{\mathrm{FD}}^2 \sigma_{\mathrm{FD}}^2 + w_{\mathrm{SD}}^2 \sigma_{\mathrm{SD}}^2$$

- Linear cut on jointly Gaussian observables → 2-D integral collapses to one Φ
- Per event add: $-\log \Phi\ \left((\mu_Z - t)/\sigma_Z\right)$
- μ_Z carries $g(\mathrm{lgE}; \mathrm{a}, \mathrm{b}, \mathrm{c})$ → the term is parameter-dependent

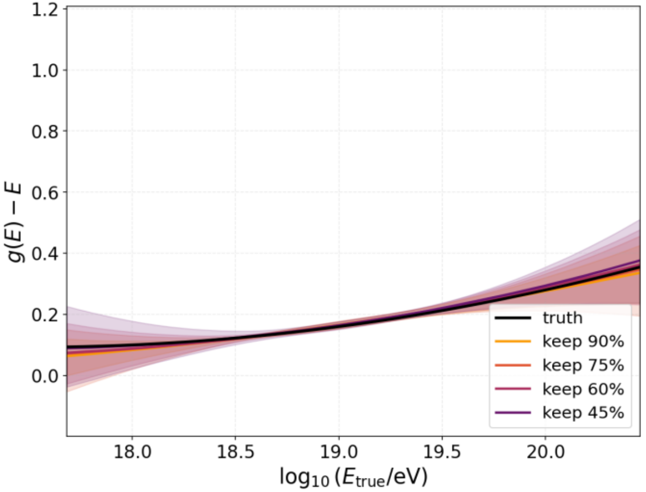


Calibration function fit

No account for cuts



With truncated likelihood



Conclusions

- Calibration of the response function is required to correctly perform forward folding on FD energy scale
- Bayesian MCMC calibration allows a natural extension of the MCMC forward folding method and enables uncertainty propagation
- The model takes into account different types of cuts that significantly influence calibration
- Bayesian methods allows to choose model complexity