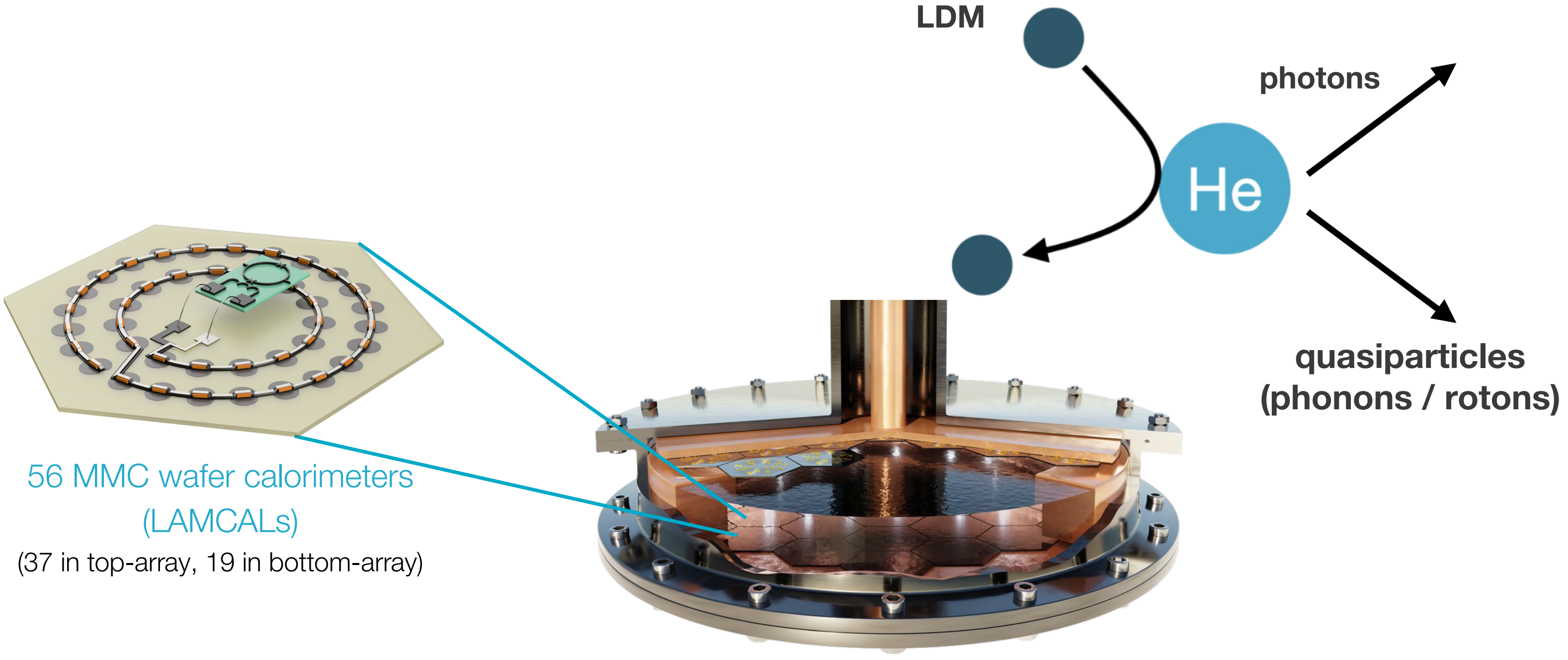
From Detector Likelihood to Geometry-Aware Reconstruction

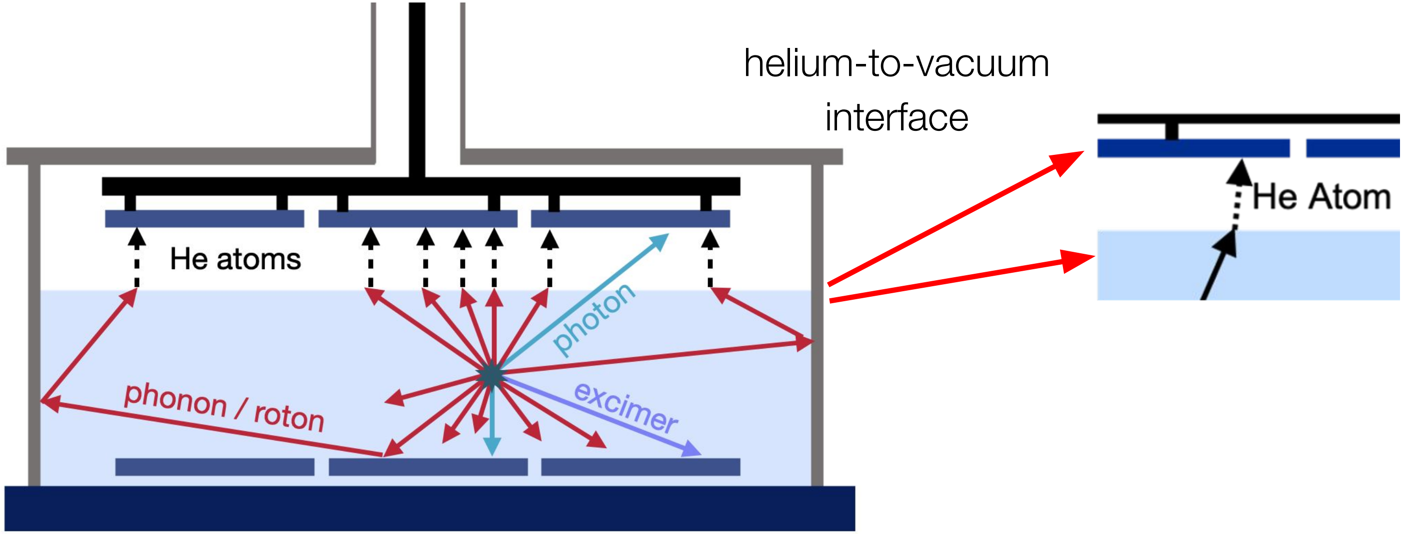
Dowling Wong
Institute of Experimental Particle Physics (ETP), KIT
NPML 2026

Why Low-Energy Reconstruction Matters



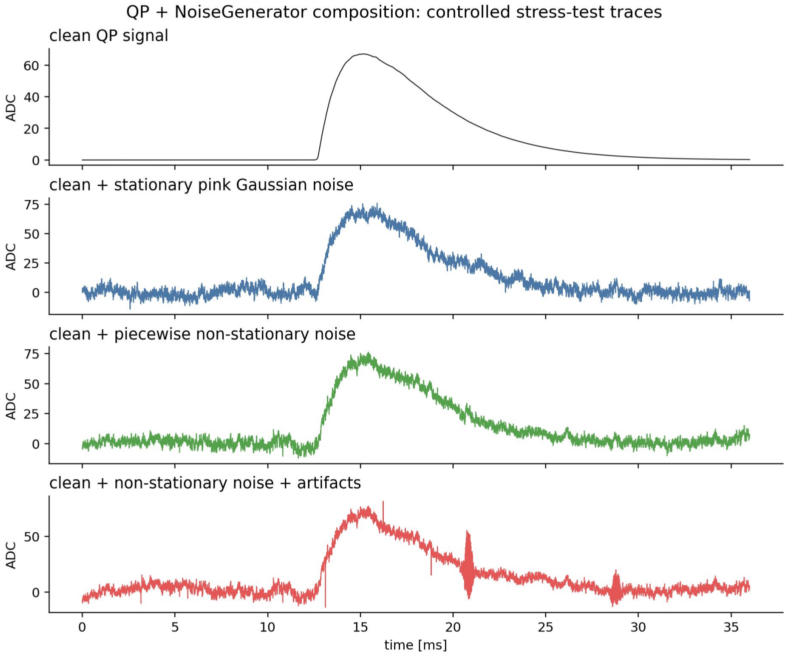
DELight is a direct detection dark matter experiment using Superfluid helium and LAMCAL to probe the thus far uncharted low mass dark matter parameter space.

Quasiparticle signals encode the event information



To reach lower dark matter masses, low-energy interaction of quasiparticle becomes the signature signal for the experiment, which carries information with various position, channel sharing, arriving time and pulse shape.

Statistical model of the detector response



A latent physical event generates an expected detector response s(z), while readout electronics and thermal fluctuations contribute stochastic noise n, yielding the observed trace x.

$$x = s(z) + n$$

$z = (E, t_0, r, \phi, c, \alpha, b, \eta)$: event parameters

$n \sim N(0,\Sigma)$: detector fluctuations

The likelihood determines the loss

Given a signal hypothesis ŝ, what is the likelihood of observing x?

With first-order approximation Gaussian noise:

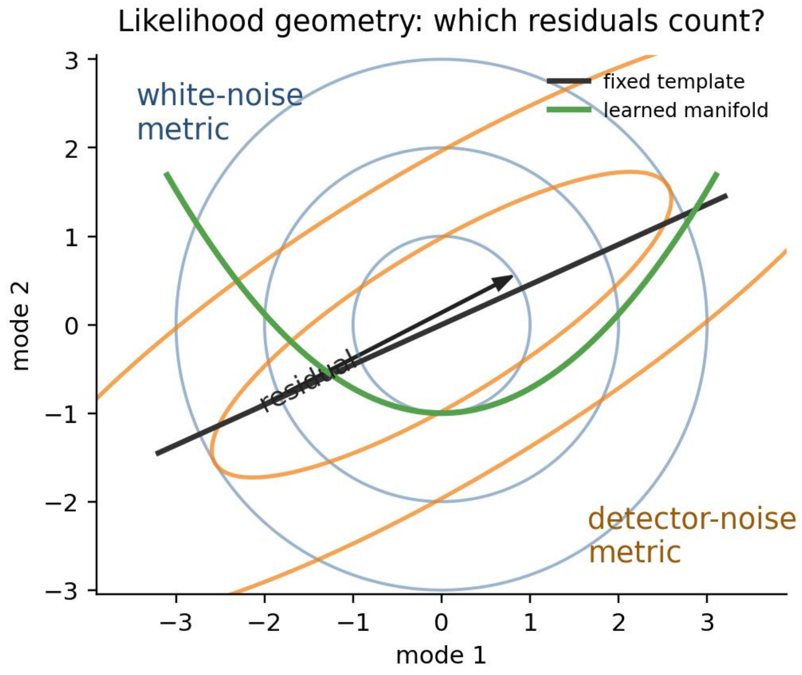
$$n \sim \mathcal{N}(0, \Sigma)$$

Then likelihood becomes:

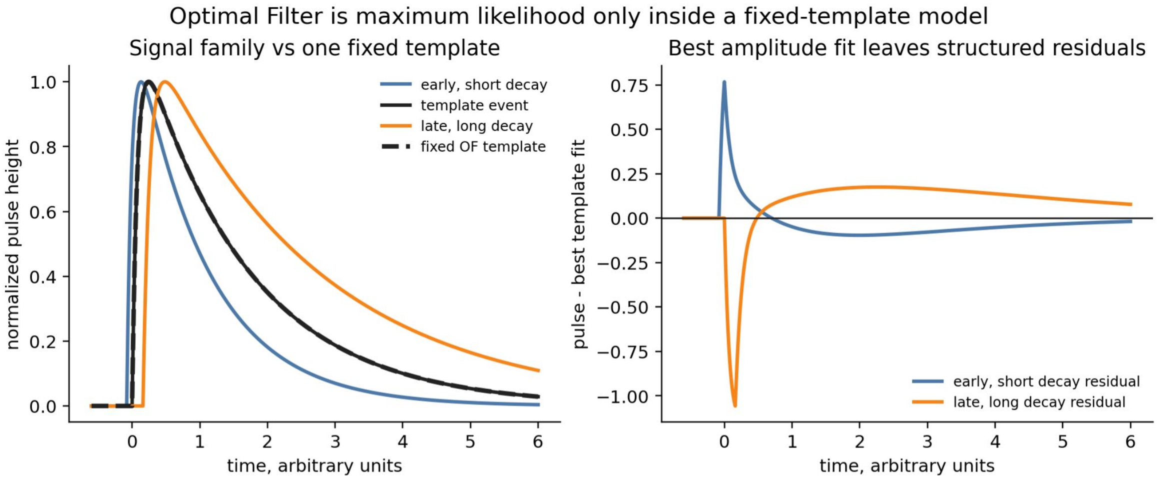
$$-\log p(x \mid \hat{s}) = \frac{1}{2}(x - \hat{s})^\top \Sigma^{-1}(x - \hat{s}) + \text{const}$$

Therefore maximizing likelihood is equivalent to minimizing the Mahalanobis distance

$$\|x - \hat{s}\|^2_{\Sigma^{-1}} = (x - \hat{s})^\top \Sigma^{-1}(x - \hat{s}) \, .$$



Optimal Filter as Local Maximum Likelihood



Instead of an arbitrary reconstruction ŝ, assume $\hat{s} = As_0$

Likelihood becomes:

$$\hat{A} = \underset{A}{\operatorname{argmin}}(x - As_0)^T \Sigma^{-1} (x - As_0).$$

Maximizing Gaussian likelihood:

$$\frac{\partial}{\partial A}(x - As_0)^T \Sigma^{-1} (x - As_0) = 0,$$

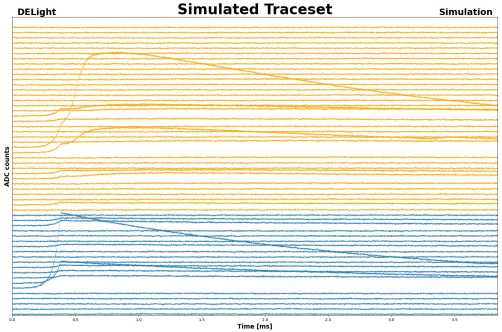
Yields maximum likelihood estimator under a fixed-template signal model, Optimal Filter:

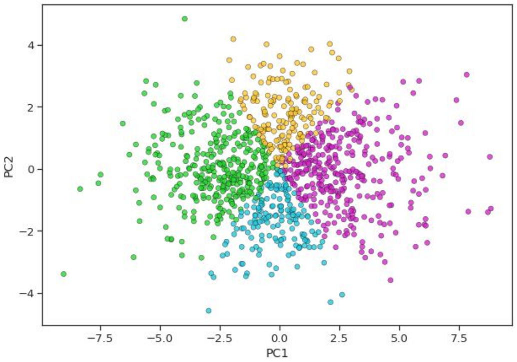
$$\hat{A} = \frac{s_0^\top \Sigma^{-1} x}{s_0^\top \Sigma^{-1} s_0}$$

PCA Learns Euclidean Signal Geometry

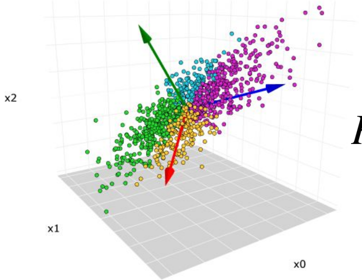
Real detector signals do not lie on a single template;
they form a low-dimensional manifold generated by varying physical parameters.

Signal reconstruction becomes projection onto a **learned signal manifold**.





Learn principal signal modes



$$\hat{x} = Pc$$

$$P = [p_1, \ldots, p_k]$$

P = learned signal basis
c = latent coordinates

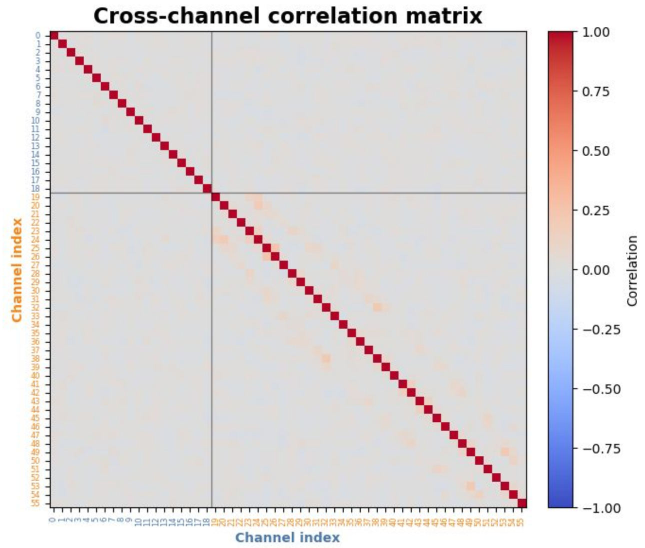
Projection for reconstruction

EMPCA Learns Noise-Weighted Geometry

PCA assumes isotropic noise.
Detector noise is colored and correlated.

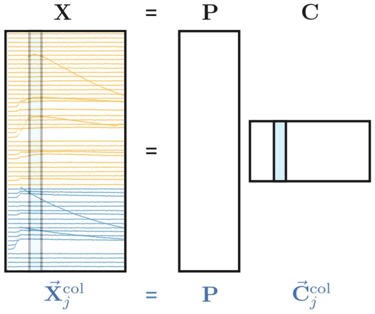
$$\min_{P,C}\sum_i (x_i - Pc_i)^\top \Sigma^{-1}(x_i - Pc_i)$$

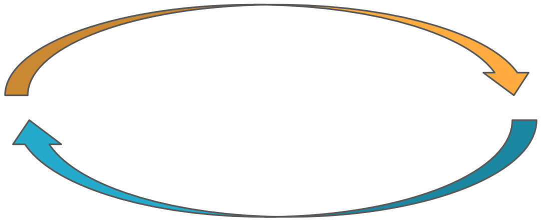
Under detector geometry,
Σ^{-1} encodes colored noise and cross-channel correlations

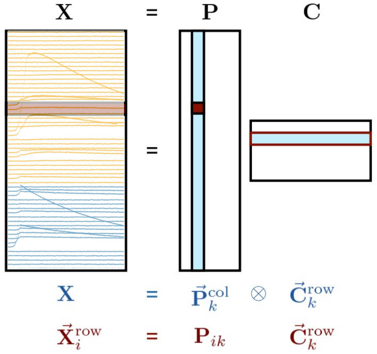


Σ^{-1} = inverse detector covariance

Expectation: Solving coefficients

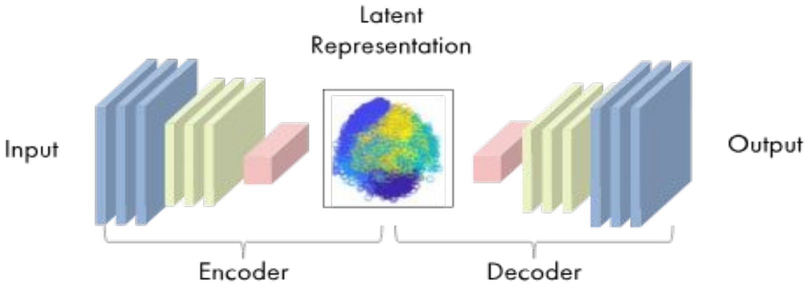


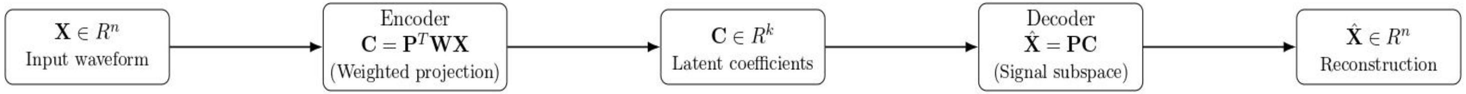




Maximization: refine basis

Autoencoders Learn Nonlinear Geometry





$$L = (x - W_d W_e x)^T \Sigma^{-1} (x - W_d W_e x) \qquad \hat{x} = Pc = W_d W_e x.$$

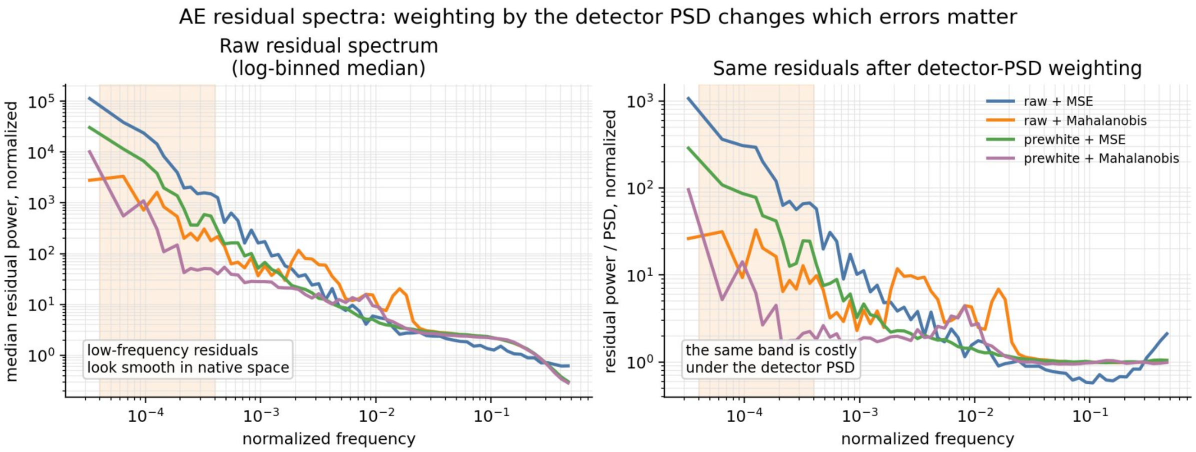
EMPCA is mathematically equivalent to a noise-weighted linear autoencoder under the same likelihood geometry.

Validated by:

- subspace equivalence: principal angle = 0.0046°
- residual equivalence: KS statistic = 9.18×10^{-4}

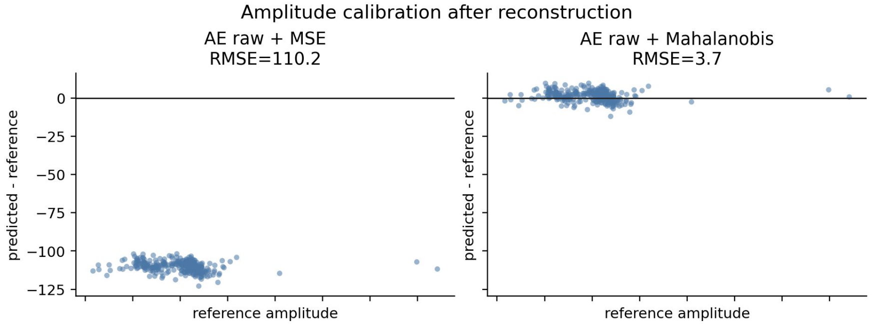
Autoencoders naturally extend this framework to nonlinear signal manifolds and non-Gaussian noise.

Detector Geometry Defines Which Errors Matter



Isotropic PCA: $\min_{P,C} |X - PC|_F^2$

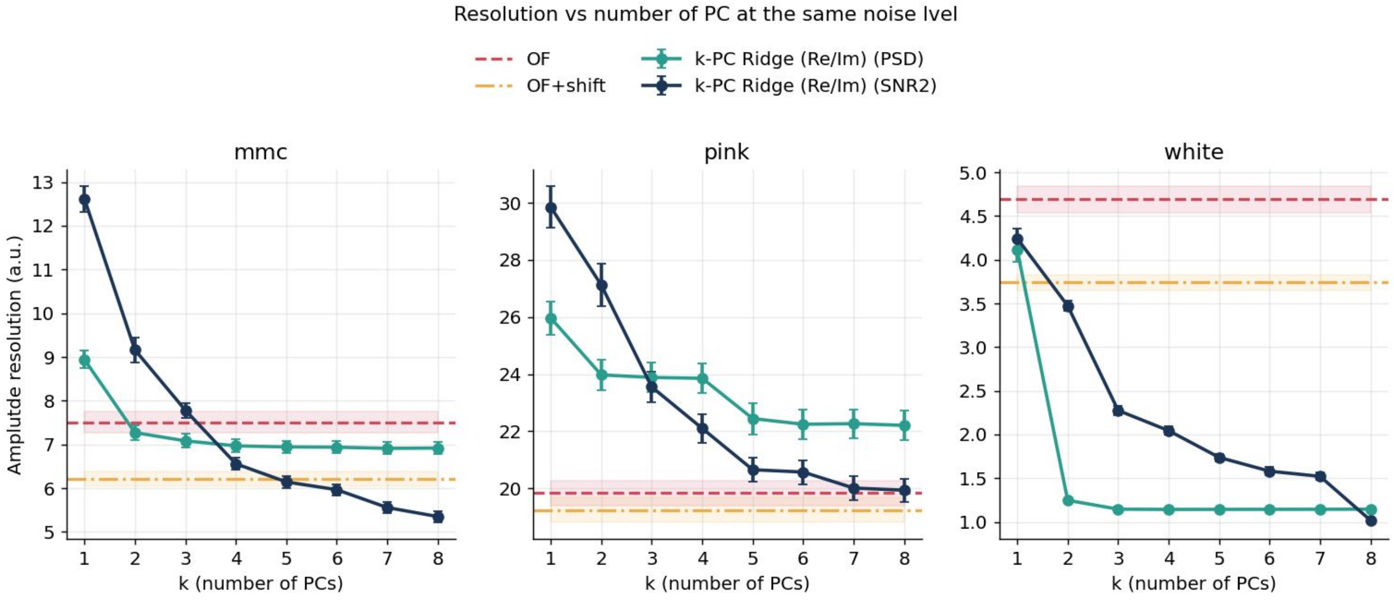
Noise-aware EMPCA: $min_{P,C} |X - PC|^2_{\Sigma^{-1}}$



- MSE assumes all frequencies are equally important.
- Mahalanobis weights residuals by detector covariance Σ^{-1}.
- The same residual spectrum can yield different reconstruction quality.

Exploratory methodology under development

Learning the Signal Manifold Improves Reconstruction

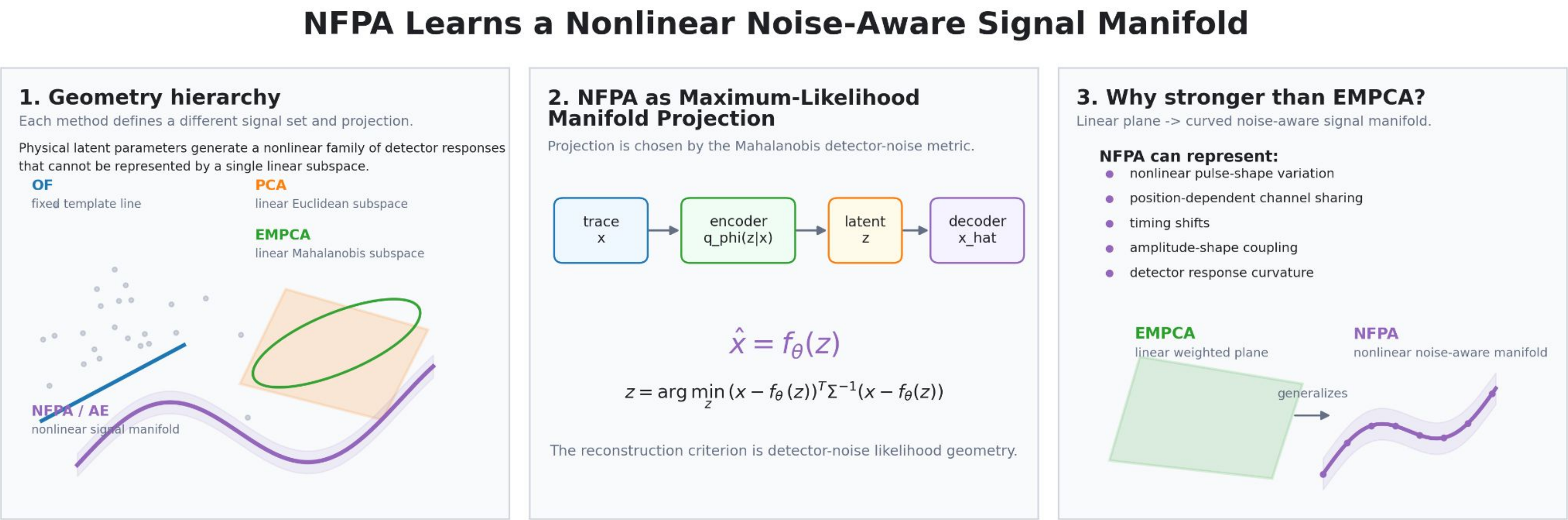


σ = width of reconstructed amplitude distribution

More PCs
↓
Richer signal manifold
↓
Better reconstruction

- If geometry defines which residuals matter, then learning a better approximation of the signal manifold should improve reconstruction under that geometry.
  - OF relies on a fixed signal template.
  - EMPCA learns a low-dimensional signal manifold.

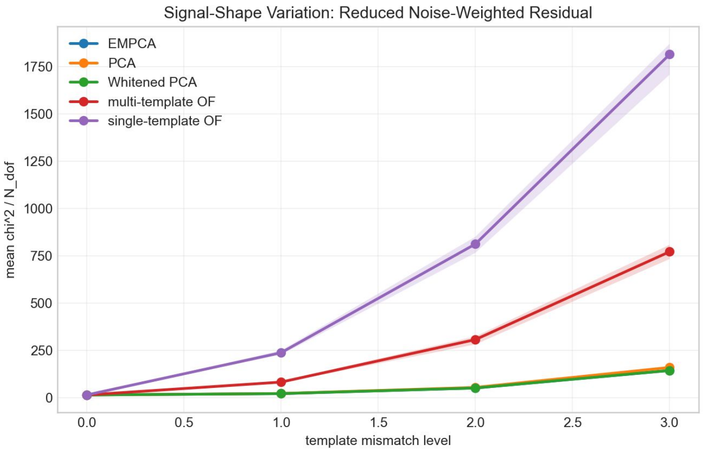
Architecture Should Respect Learned Geometry

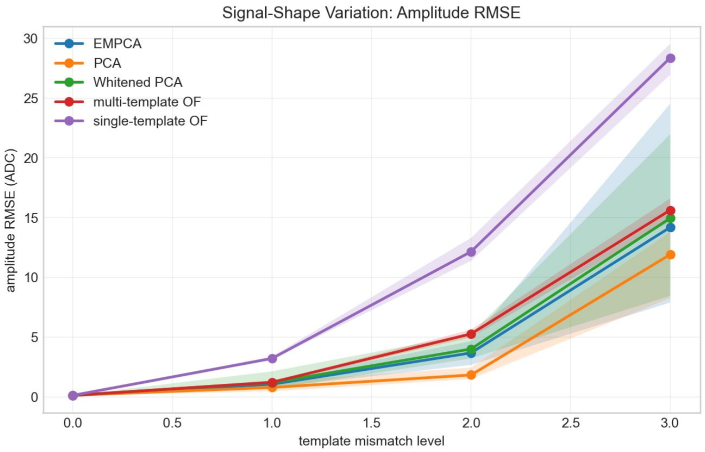


	EMPCA	NFPA
Reconstruction	$\hat{x} = \sum_{i=1}^{k} c_i p_i$	$\hat{X} = \sum_{i=1}^{k} c_i \left(a_i b_i^T \right)$
Signal space basis	$P = [p_1, \dots, p_k]$	$a_i = channel\ mode$ $b_i = time\ mode$

Channel Structure
×
Temporal Structure

Geometry Is Necessary but Not Sufficient





Mahalanobis likelihood answers: $\min_{P,C} \sum_i (x_i - Pc_i)^\top \Sigma^{-1} (x_i - Pc_i)$

Exploratory methodology under development

- How should residuals be weighted?

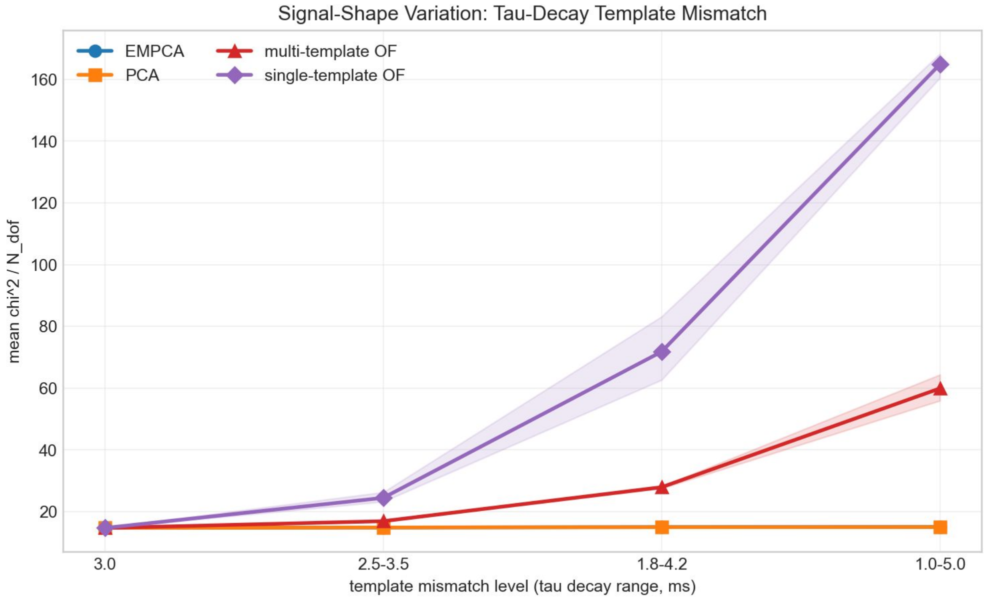
It does not define:

- coverage
- identifiability
- inductive bias

$$z = (E, t_0, r, \phi, c, \alpha, b, \eta)$$

Coverage determines what can be learned

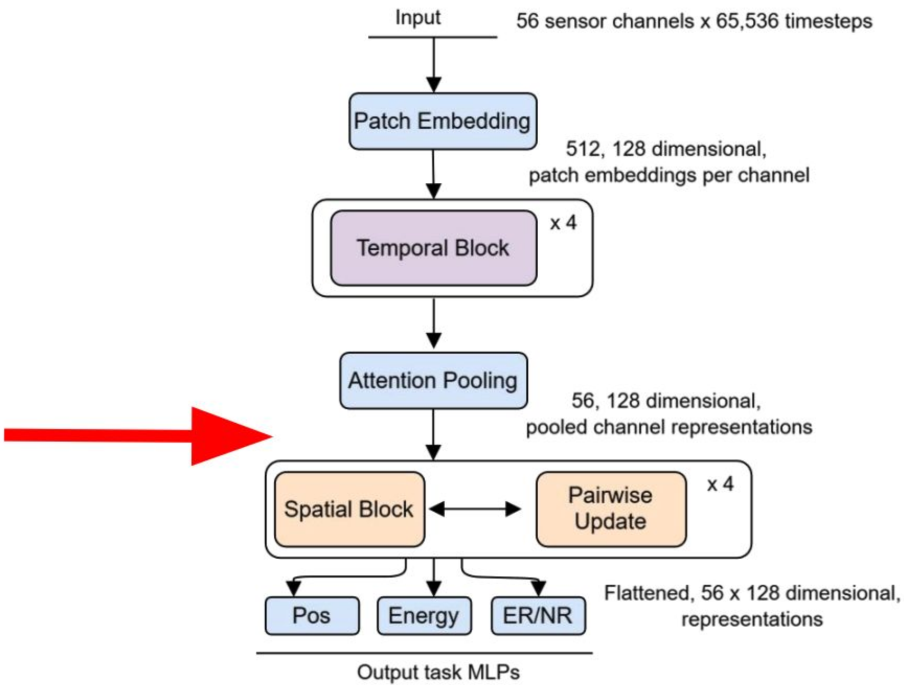
- Recovery depends on:
  - coverage: did training data excite the relevant latent factors?
  - inductive bias: does the architecture prefer the right solution among likelihood-compatible reconstructions?



- Coverage determines learnability
- OF assumes fixed signal templates.
- Signals vary across latent parameters
- More calibration data → better manifold approximation
- Coverage is a data problem

Exploratory methodology under development

Physics-Informed Attention for Global Reconstruction



Once geometry and coverage are fixed, architecture determines which reconstruction is learned.

Input
↓
Temporal Block
↓
Attention Pooling
↓
Spatial Block

time representation
↓
channel representation

The likelihood defines the geometry, coverage determines what can be learned, and the architecture selects which likelihood-compatible solution is learned.

Detector Geometry as an Attention Prior

-Pairwise Geometry Representation

Δx, Δy, Δz

same detector plane

channel adjacency

noise covariance

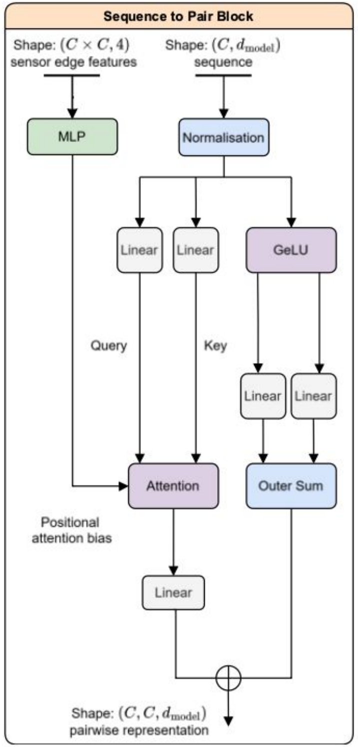
↓

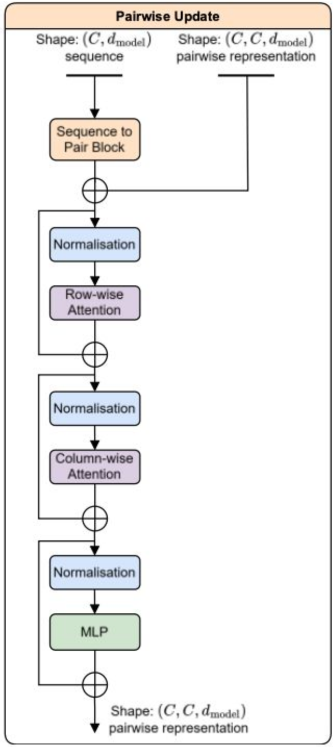
pair representation Zij

↓

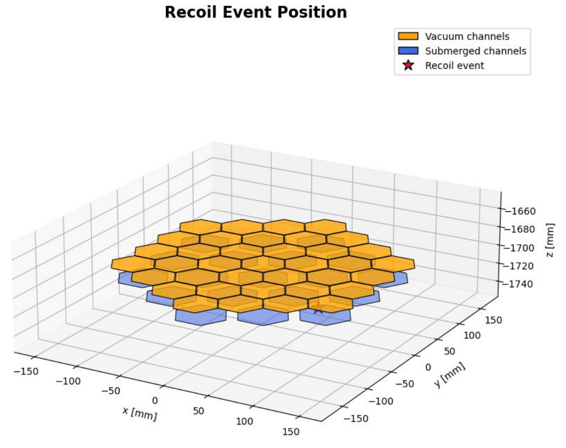
attention bias

OF amplitudes injected as local maximum-likelihood statistics



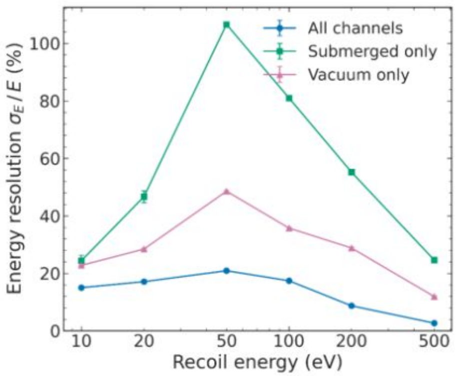


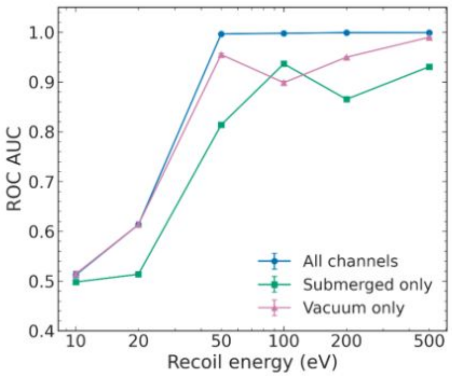
Does the Model Learn Detector Physics?

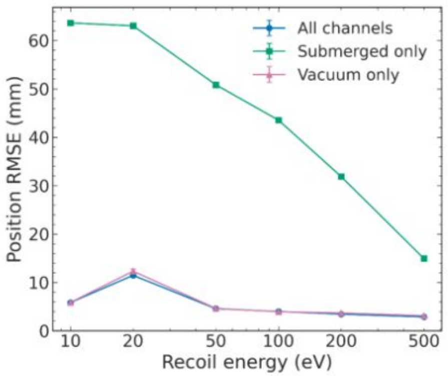


Hypothesis:

If the model learned detector structure, removing informative sensors should cause predictable degradation.



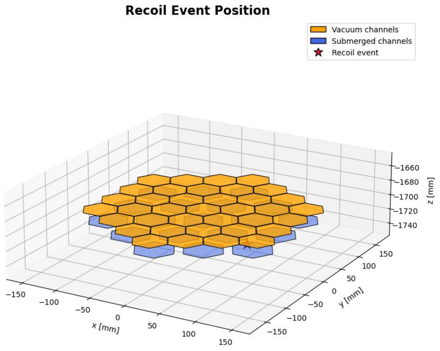


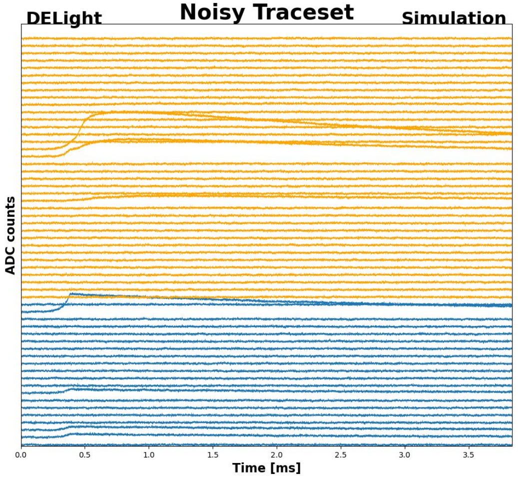


Observation:
Top-channel masking causes the largest degradation.

Conclusion:
The model learned the detector's signal-partitioning structure.

Conclusion: reconstruction is not only a loss-function problem





Likelihood → Geometry → Coverage → Information → Architecture → Inductive bias

The loss determines the geometry. Coverage determines the information. Architecture determines the solution.

Thank you for your attention

Questions, comments, and feedback are very welcome.

Interested in geometry-aware reconstruction, or related collaborations?

dowling.wong@kit.edu