More with less: *sparse* kernel methods with dictionary learning

Expressive, *regularized* and *interpretable* models for **statistical anomaly detection**

Gaia Grosso[1,2,3], Demba Ba[1,2], Phil Harris[1,3]

[1]NSF Institute for Artificial Intelligence and Fundamental Interaction (IAIFI)
[2]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA [3]MIT Laboratory for Nuclear Science, Cambridge, MA

Harvard University
Massachusetts Institute of Technology

EUROPEAN AI FOR FUNDAMENTAL PHYSICS CONFERENCE
EuCAIFCon 2024

GOAL

Signal-agnostic statistical detection of new physical processes

Maximum-likelihood-ratio goodness-of-fit test:

$$t(\mathcal{D}) = 2 \max_{\boldsymbol{\theta}} \log \frac{\mathcal{L}(\mathcal{D}|\mathrm{H}_{\boldsymbol{\theta}})}{\mathcal{L}(\mathcal{D}|\mathrm{H}_{\mathbf{0}})}$$

$$= -2 \min_{\boldsymbol{\theta}} L_{\mathrm{LR}}[f_{\boldsymbol{\theta}}]$$

$$n(x|\mathrm{H}_{\boldsymbol{\theta}}) = n(x|\mathrm{H}_{\mathbf{0}}) \, \exp[f_{\boldsymbol{\theta}}(x)]$$

Loss function:

$$L_{\mathrm{LR}}[f_{\boldsymbol{\theta}}] = \sum_{x \in \mathcal{R}} w_{\mathbf{0}}(x) \left(\exp[f_{\boldsymbol{\theta}}(x)] - 1\right) - \sum_{x \in \mathcal{D}} f_{\boldsymbol{\theta}}(x)$$

PROBLEM

How to design $f_{\boldsymbol{\theta}}(x)$ to capture *rare* and *unexpected* subtle perturbations on top of the known physics?

More with less: *sparse* kernel methods with dictionary learning

Expressive, *regularized* and *interpretable* models for **statistical anomaly detection**

Gaia Grosso[1,2,3], Demba Ba[1,2], Phil Harris[1,3]

[1]NSF Institute for Artificial Intelligence and Fundamental Interaction (IAIFI)
[2]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA [3]MIT Laboratory for Nuclear Science, Cambridge, MA

HARVARD UNIVERSITY

Massachusetts Institute of Technology

EUROPEAN AI FOR FUNDAMENTAL PHYSICS CONFERENCE
EuCAIFCon 2024

SOLUTION
Sparse linear combination of Gaussian Kernels (SGK)

$$f_{\mu,w}(x) = \sum_{i=1}^{M} w_i\, k(x;\, \mu_i, \sigma_i)$$

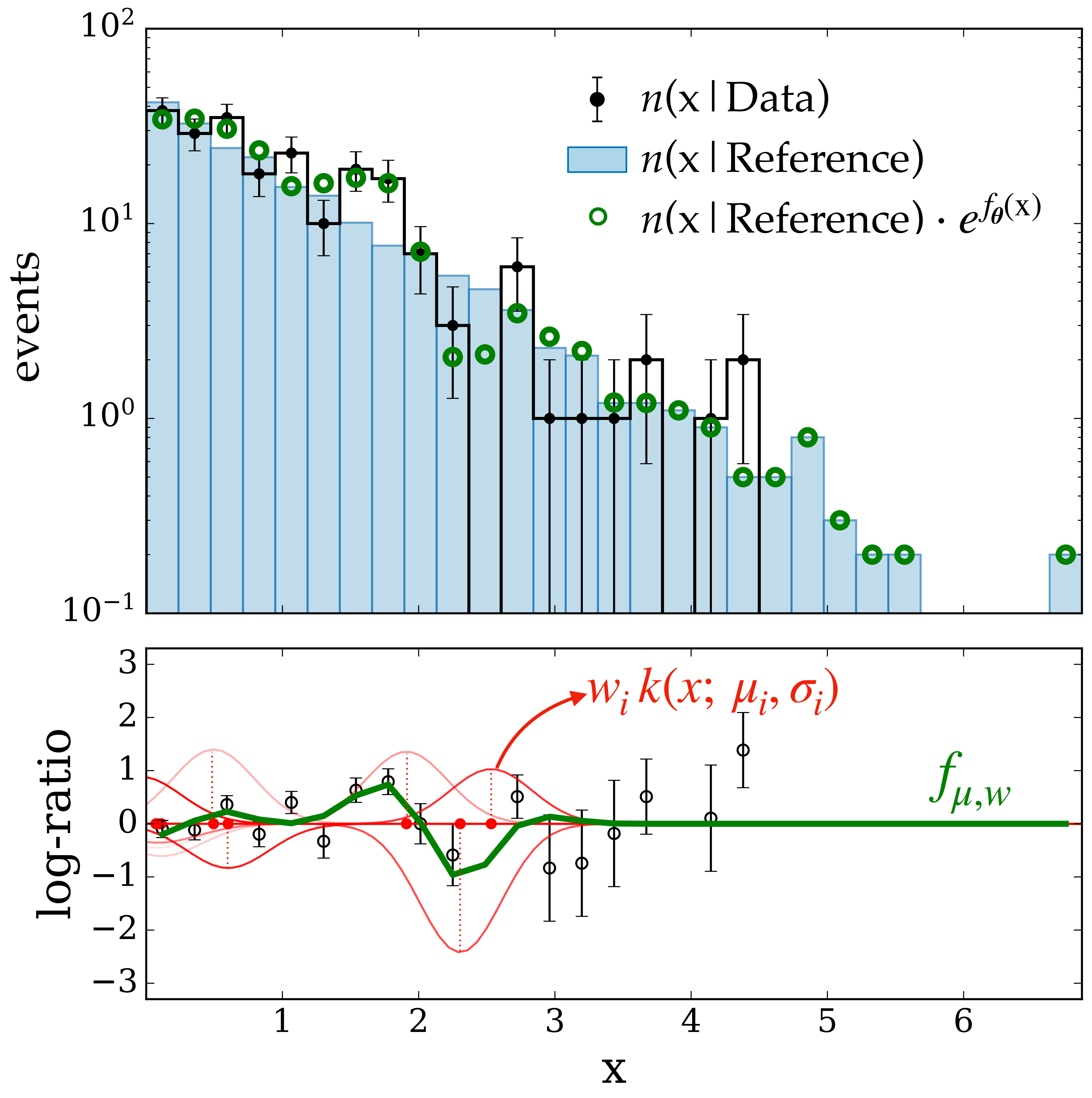
Local interpretability
Active kernels highlight anomalous regions

$$k(x;\, \mu_i, \sigma_i) = \mathrm{A}\exp\left[-\frac{||x-\mu_i||^2}{2\sigma_i^2}\right]$$

Sparse model (M ≪ N)
competition between data points to attract the kernels

Adaptive model (learnable μ)
directing *attention* to anomalous features

Smooth model ($\sigma^2 = \sigma^2_{\mathrm{exp}} + \sigma^2_X$)
Physics constraints (e.g. experimental resolution).
What is the scale of New Physics?



More with less: *sparse* kernel methods with dictionary learning

Expressive, *regularized* and *interpretable* models for **statistical anomaly detection**

Gaia Grosso[1,2,3], Demba Ba[1,2], Phil Harris[1,3]

[1]NSF Institute for Artificial Intelligence and Fundamental Interaction (IAIFI)
[2]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA [3]MIT Laboratory for Nuclear Science, Cambridge, MA

HARVARD UNIVERSITY

Massachusetts Institute of Technology

EUROPEAN AI FOR FUNDAMENTAL PHYSICS CONFERENCE
EuCAIFCon 2024

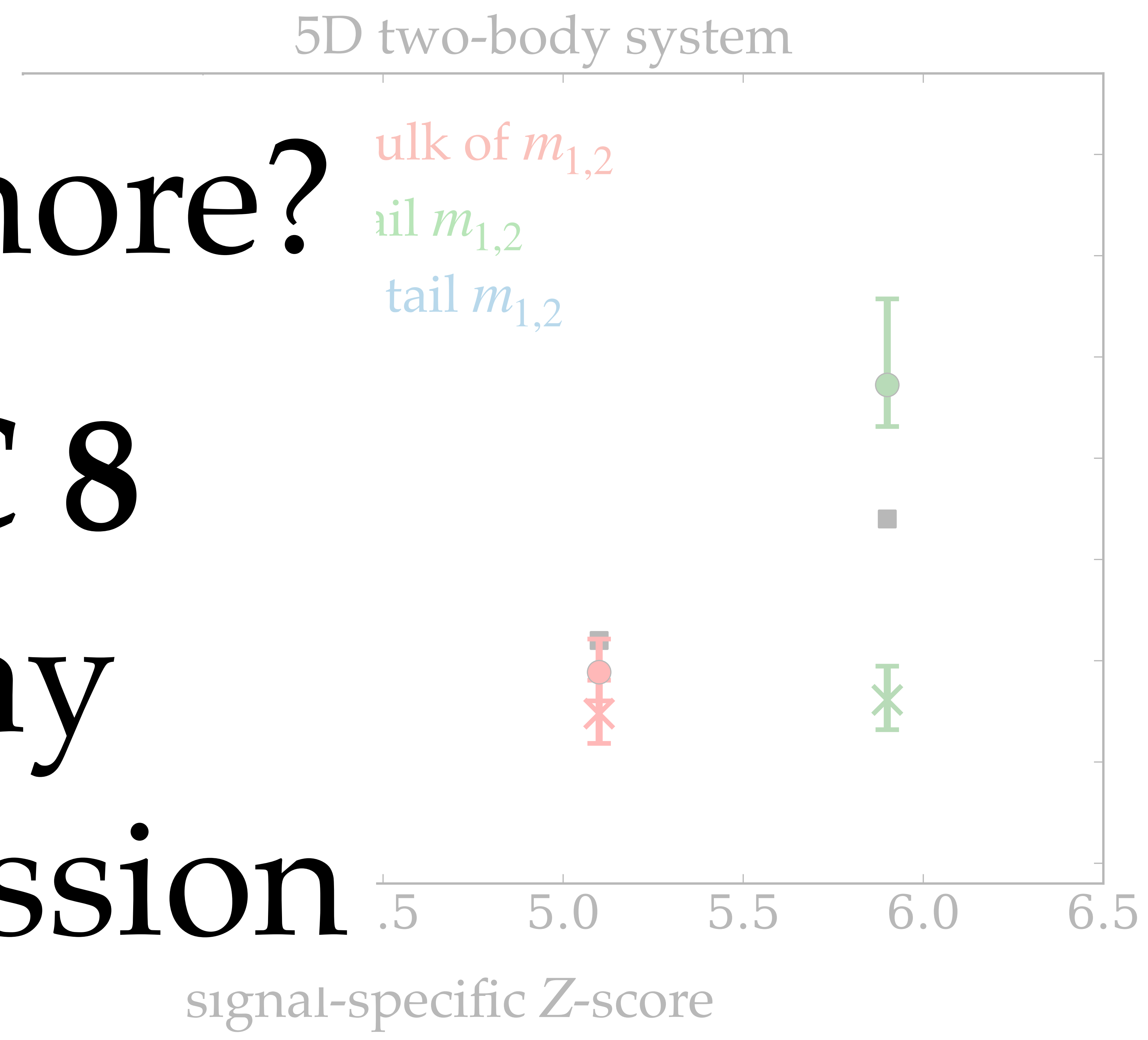
RESULTS

	Model	#par		time	Ref
■	NN	96			
✕	GK*	10k	(M=10k, rando		
○	SGK*	600	(M=100, learn		

$^*\sigma = q_{50\%}$: median of pair-wise dist

more with l

Same or improved sensitivity t



Want to know more?

Drop by **LOC 8** on Wednesday for the poster session

IMPLICATIONS

Resource efficient represent

→ Interpretability

→ Data compression?

[1] "Learning multivariate new physics" Eur. Phys. J. C 81, 89 (2021)
[2] "Learning new physics efficiently with nonparametric methods" Eur. Phys. J. C, 82(10)