

Many Wrongs Make a Right

Leveraging Biased Simulations Towards Unbiased Parameter Inference

Sean Benevedes

Pheno 2026

May 12th, 2026

Problem statement

- We have some target data which is a mixture of signal and background $s_{\text{target}}(x)$ and $b_{\text{target}}(x)$; want to estimate signal fraction κ
- We don't have the functions $s_{\text{target}}(x)$ and $b_{\text{target}}(x)$, but we have simulated signal and background samples

$$p_{\text{target}}(x) = \kappa_{\text{target}} s_{\text{target}}(x) + (1 - \kappa_{\text{target}}) b_{\text{target}}(x)$$

→ Need **simulation-based inference**

What if our simulations are poor?

- Despite heroic efforts, we know that in some circumstances our models are untrustworthy
 - Limited perturbative precision
 - Details of hadronization
- If our simulations are poor, inference for κ is generically biased!

$$s_{\text{sim}}(x) \neq s_{\text{target}}(x), \quad b_{\text{sim}}(x) \neq b_{\text{target}}(x)$$

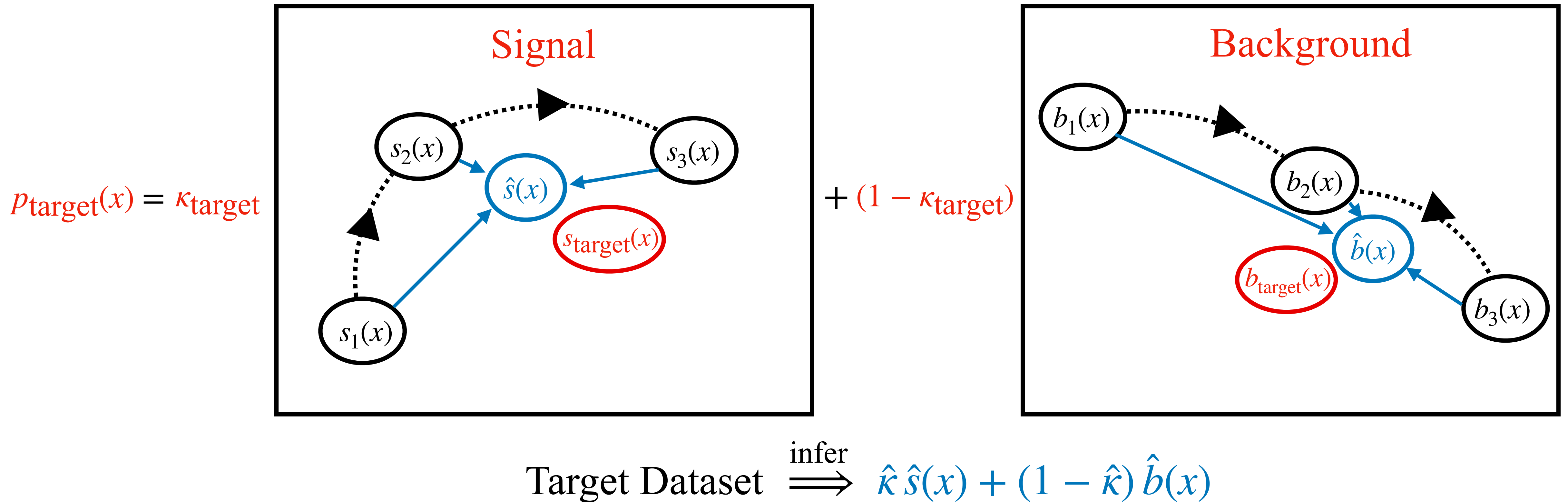
Assumption that lets us make progress: Our signal and background simulations are misspecified **individually**, but we have several of each!

(E.g. different event generators, different showers, or even just different nuisance parameter settings)

Previous work

- Idea of combining simulations to model data better than any individual simulation not new, e.g. template morphing (Read 1999) and moment morphing (Baak, Gadatsch, Harrington, Verkerke [1410.7388](#))
 - These assume simulations are generated via different values of a set of nuisance parameters
 - They then interpolate in nuisance parameter space → only expect success if there is some value of nuisance parameters for which your simulation is correct

Template-Adapted Mixture Models (TAMMs)



TAMM Details

How do we put this into practice? Need to make some choices

1. Feature representation: Unbinned or binned
2. Model components \mathfrak{s}_k and \mathfrak{b}_k : raw simulation or preprocessed
3. Model form: Exponential or linear
4. Statistical framework: Frequentist or Bayesian

Exponential: $s_{\text{exp}}(x) = c_s e^{w_k \ln \mathfrak{s}_k(x)}, \quad b_{\text{exp}}(x) = c_b e^{v_k \ln \mathfrak{b}_k(x)}$

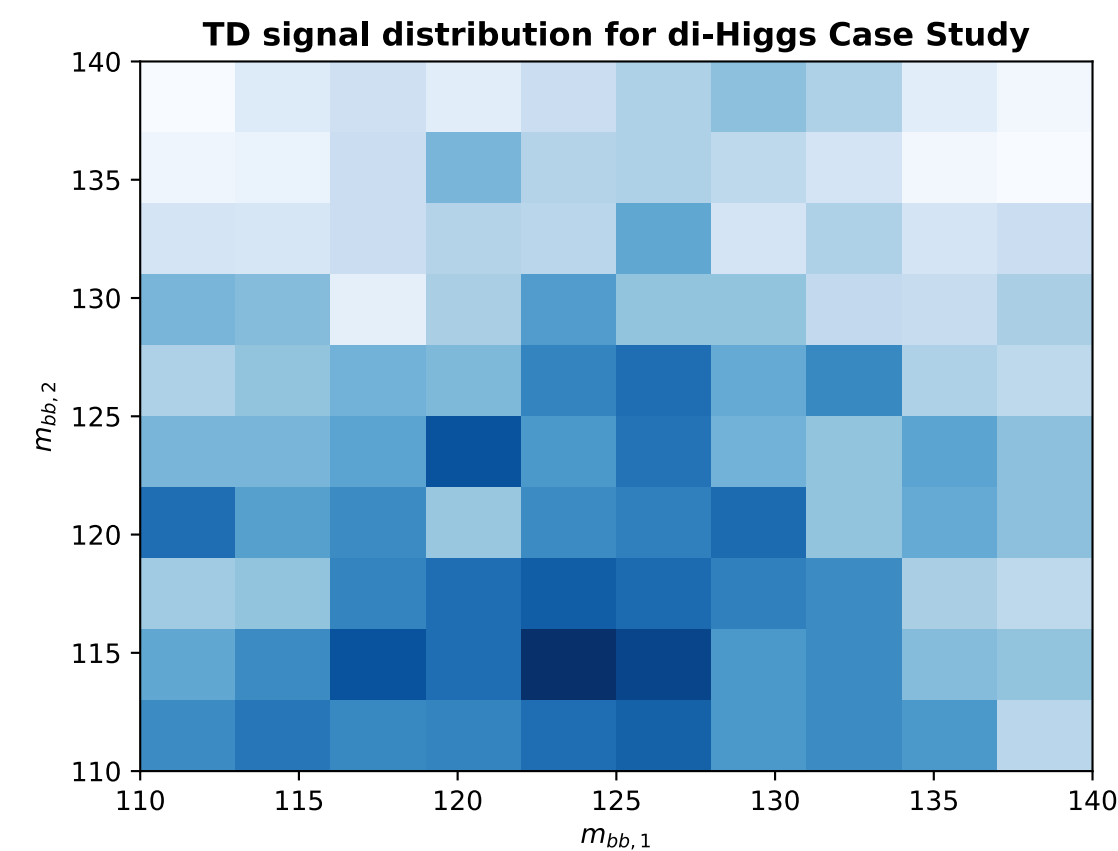
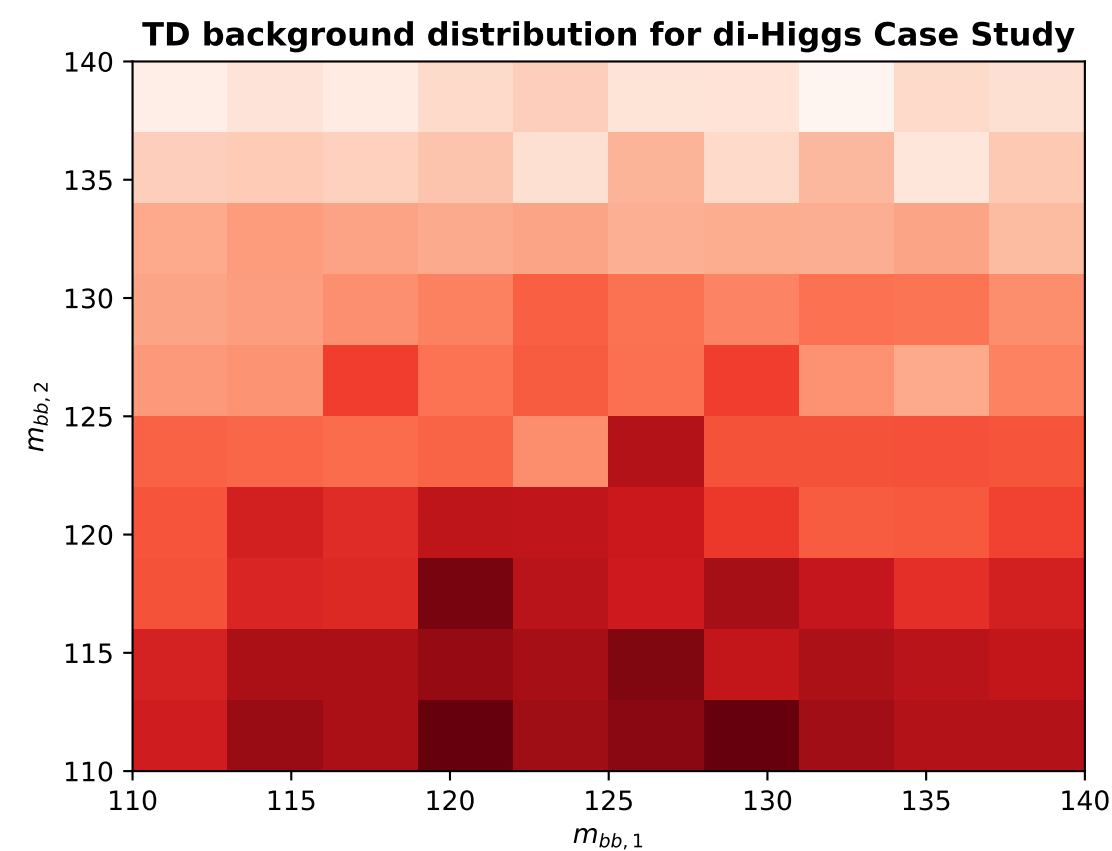
Linear: $s_{\text{lin}}(x) = w_k \mathfrak{s}_k(x), \quad b_{\text{lin}}(x) = v_k \mathfrak{b}_k(x)$

TAMM Details

	Bayesian Topic Modeling	Frequentist Neural Estimation
Representation	Binned	Unbinned
Components	Topic models	Raw simulation densities
Model	Linear	Exponential
Inference	Bayesian	Frequentist

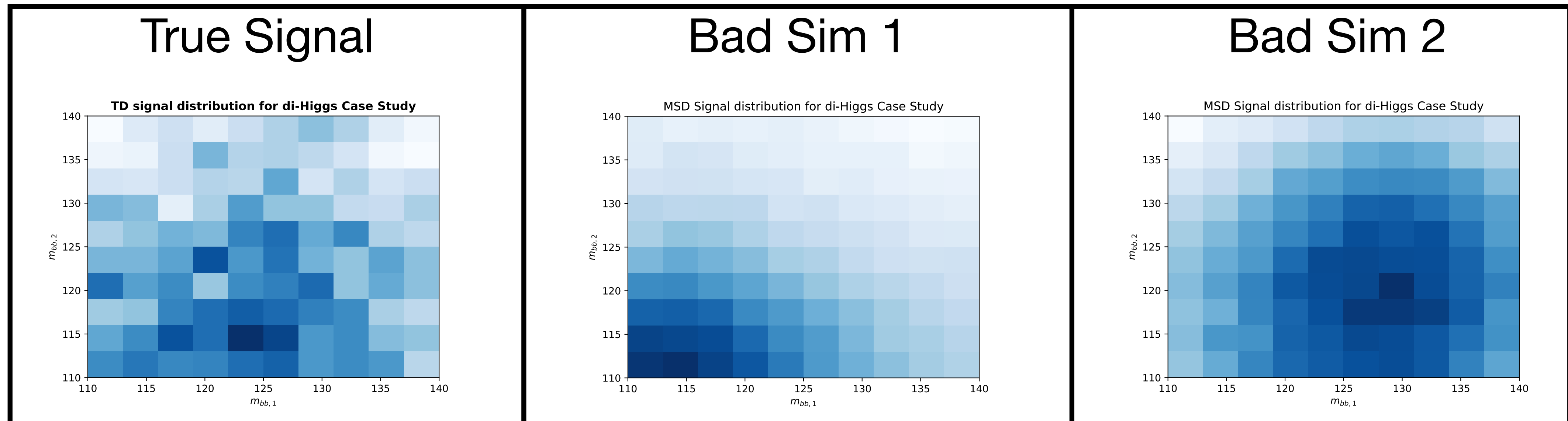
Physics case study: simulated di-Higgs

- We take the signal to be di-Higgs to four b-jets, with nonresonant QCD background. Observed data consists of 5k signal and 50k background events
 - We generate our events in MadGraph, shower them in Pythia, and perform fast detector simulation in Delphes
- Our features are the two dijet masses obtained after pairing the b-jets which minimize squared distance from the Higgs boson mass

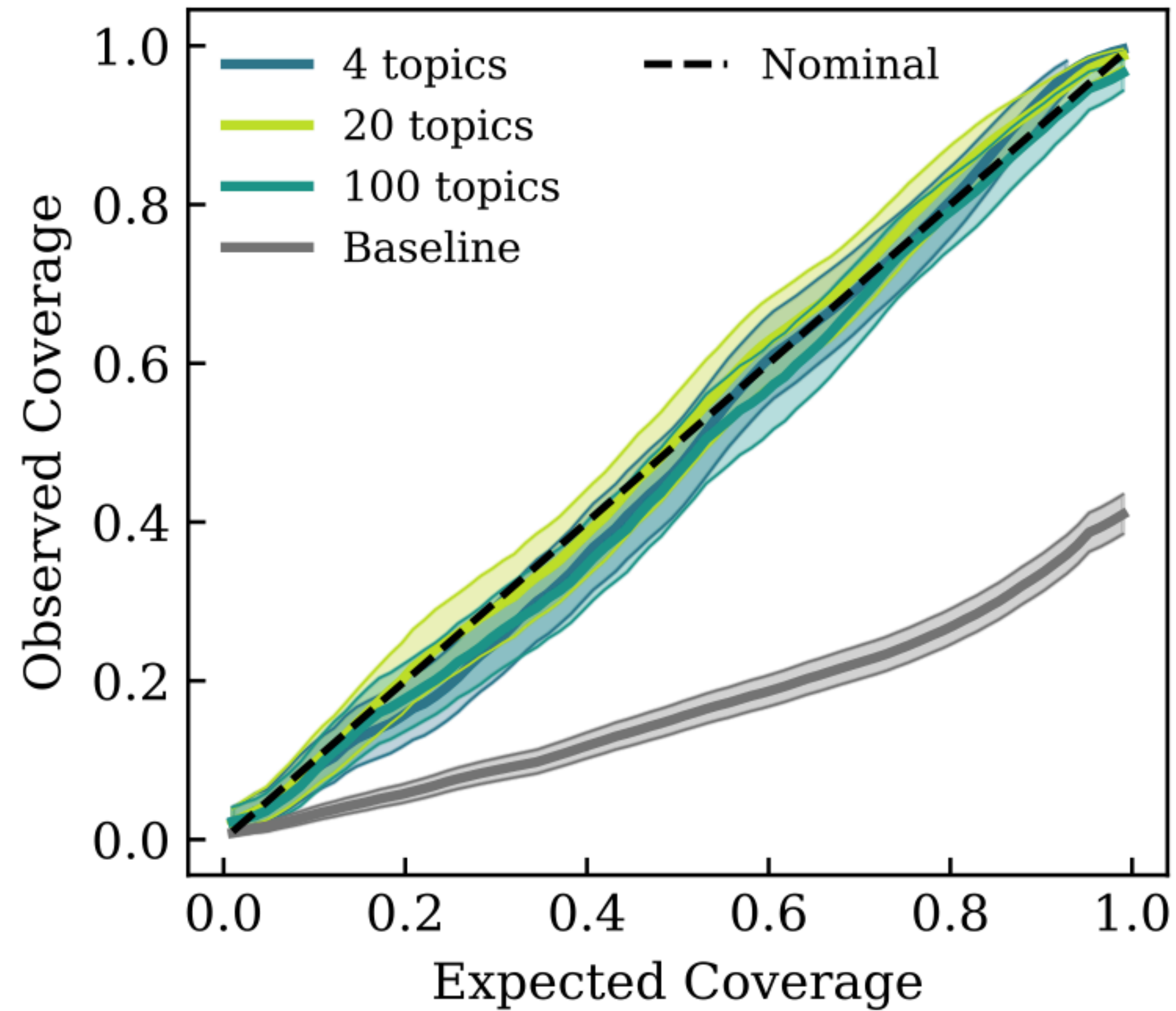


Physics case study: simulated di-Higgs

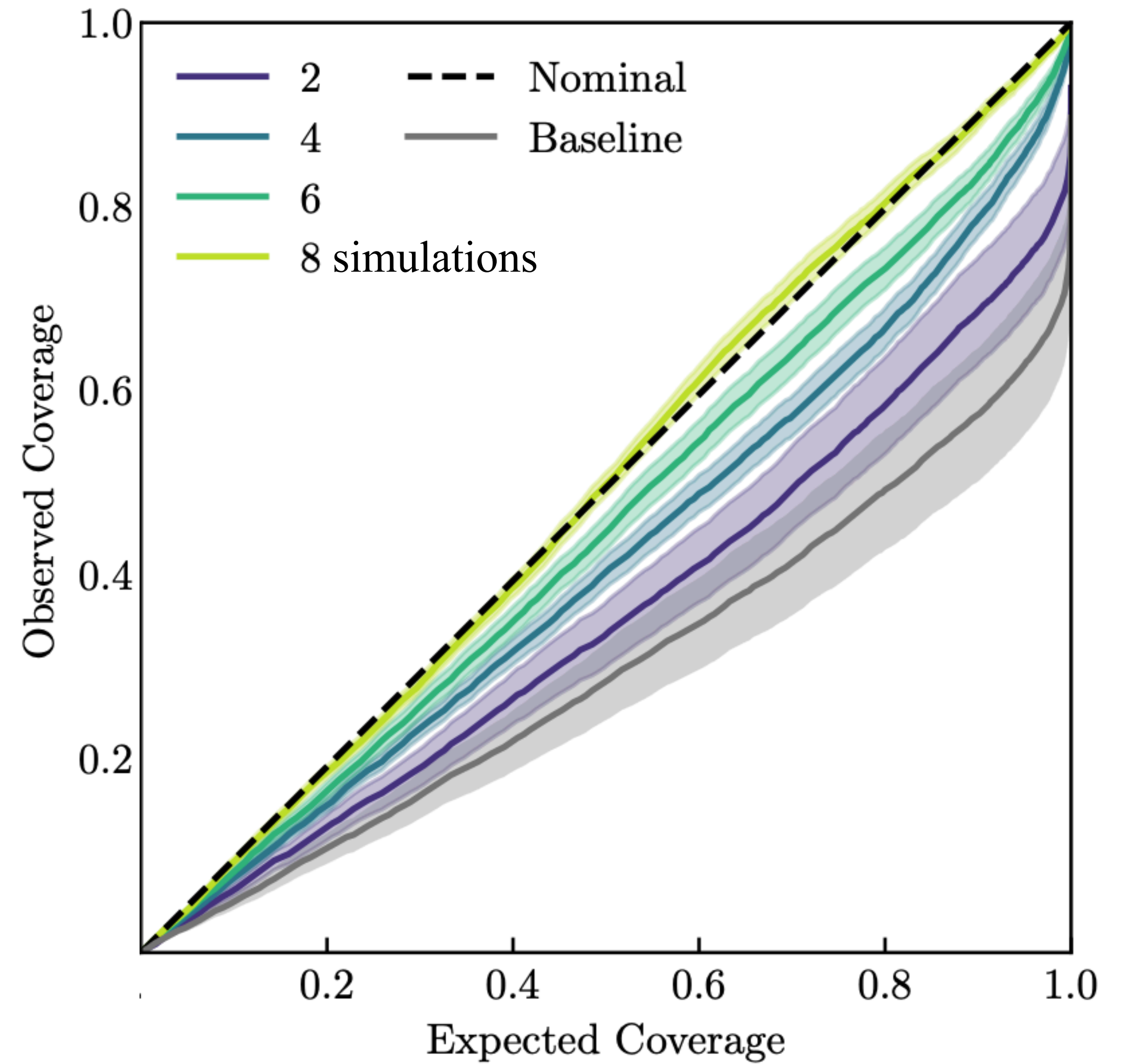
- We produce misspecified simulations (both of background and signal) by modifying the parameters of the Jet Energy Scale function in the CMS card of Delphes
- We perform a coverage study on the inferred signal fraction $\hat{\kappa}$
- All jets are available on Zenodo!



Bayesian Topic Modeling



Frequentist Neural Estimation



TAMM Summary

- TAMMs achieve **nominal coverage** on semi-realistic signal fraction extraction **with misspecified simulations!**
- Bayesian Topic Modeling and Frequentist Neural Estimation are both effective, and are complementary
 - Bayesian Topic Modeling performs well with low-dimensional feature spaces, and can harness information from a large number of misspecified simulations
 - Frequentist Neural Estimation should, in principle, scale well to high-dimensional feature spaces, and achieves nominal coverage using a small number of misspecified simulations
- All jets are available [on Zenodo](#), and try our code [on GitHub](#)!