Research Training Group Physics of the Heaviest Particles at the LHC

RWTH AACHEN UNIVERSITY

Collaborative Research Center TRR 257

P H

Particle Physics Phenomenology after the Higgs Discovery

Institute for Theoretical Particle Physics and Cosmology

RWTH AACHEN UNIVERSITY

2411.00085

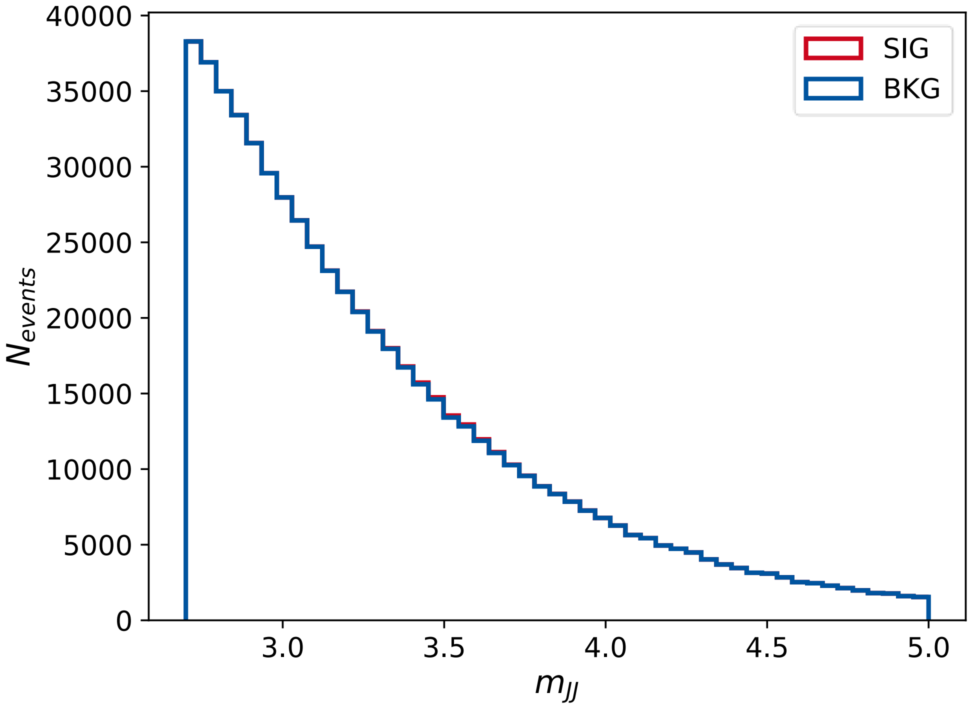
Resonant Searches as Cut and Count Experiments

Marie Hein

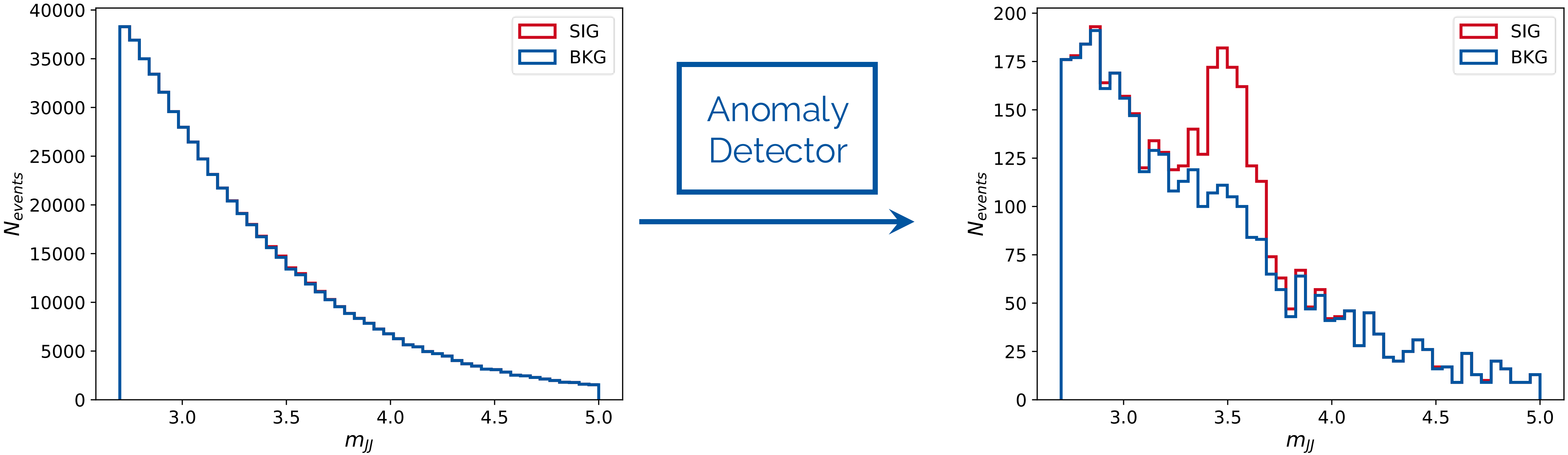
with Ranit Das, Thorben Finke, Gregor Kasieczka, Michael Krämer, Alexander Mück and David Shih

ML4Jets 2024

Anomaly Detection



Anomaly Detection



Weak Supervision

“Classification without labels: Learning from mixed samples in high energy physics” [1709.02949], E. Metodiev, B. Nachman, J. Thaler

- Optimal classifier

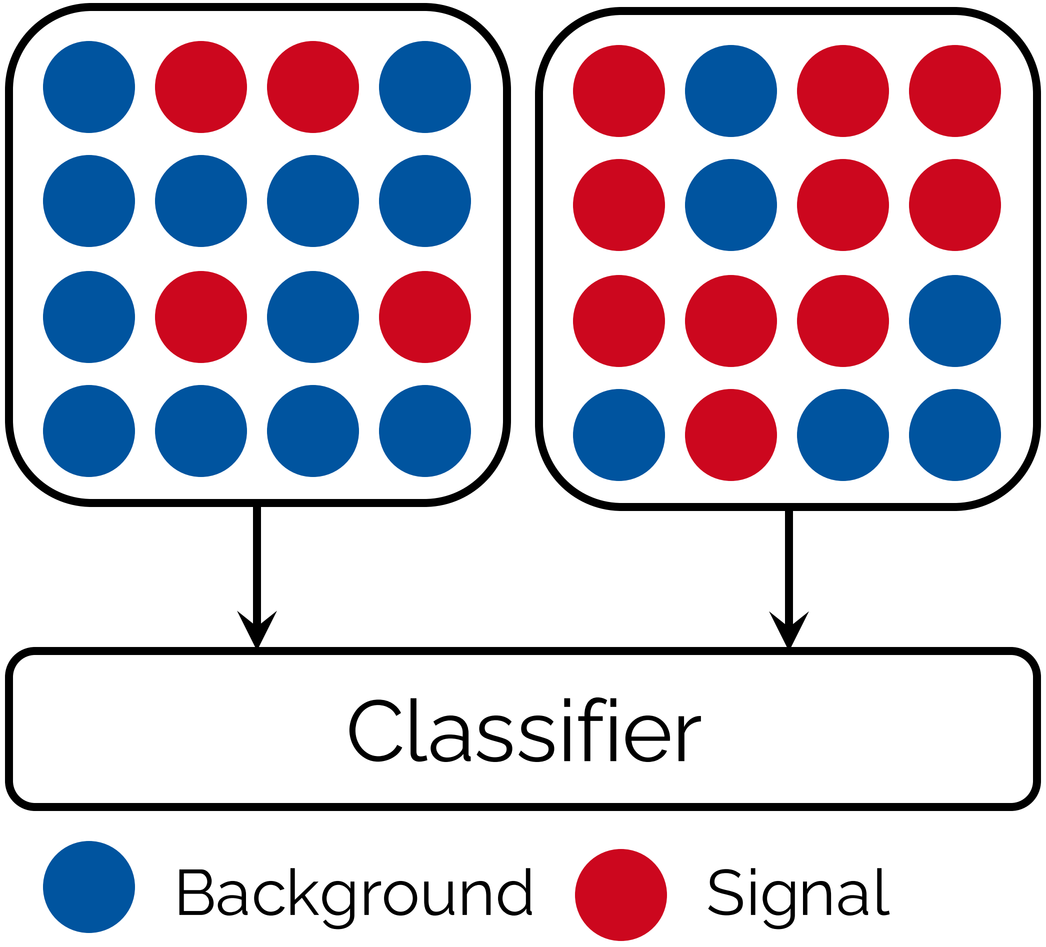
$$R_{\text{optimal}}(x) = \frac{p_S(x)}{p_B(x)}$$

- For mixed datasets with signal fractions f_i

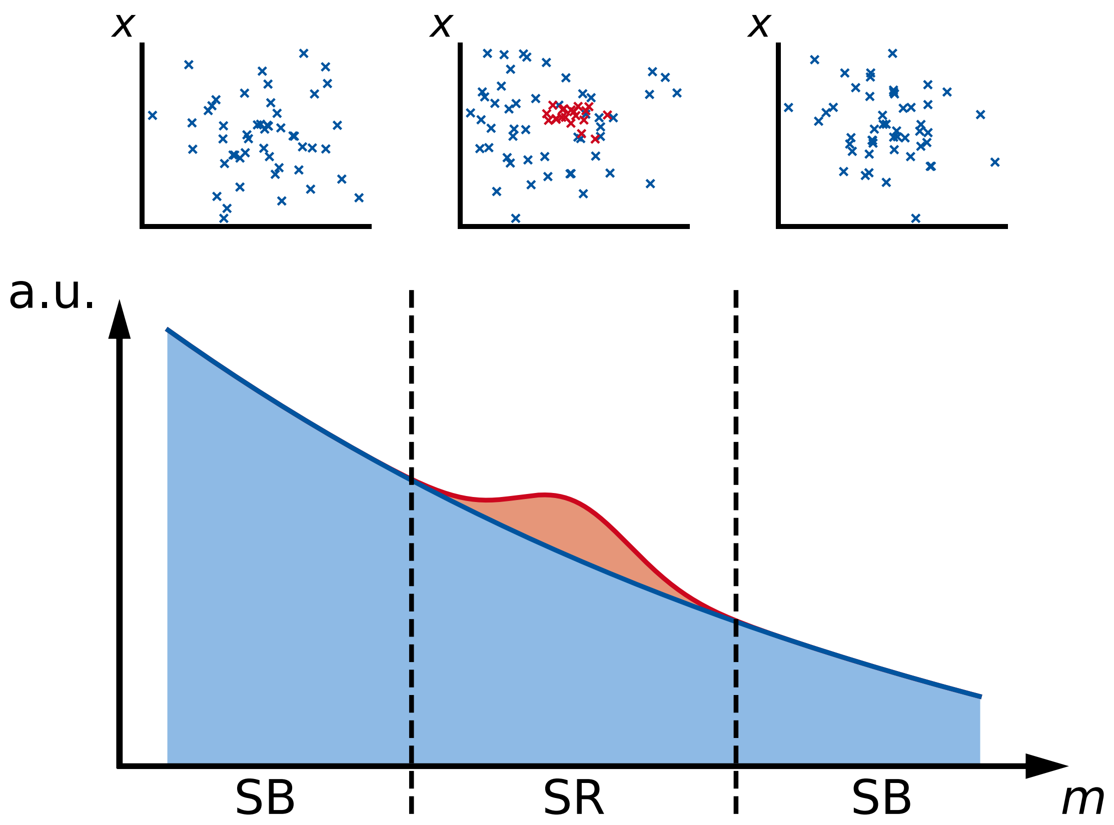
$$R_{\text{mixed}}(x) = \frac{f_1 R_{\text{optimal}}(x) + (1 - f_1)}{f_2 R_{\text{optimal}}(x) + (1 - f_2)}$$

→Monotonically increasing function of $R_{\text{optimal}}(x)$ as long as $f_1 > f_2$

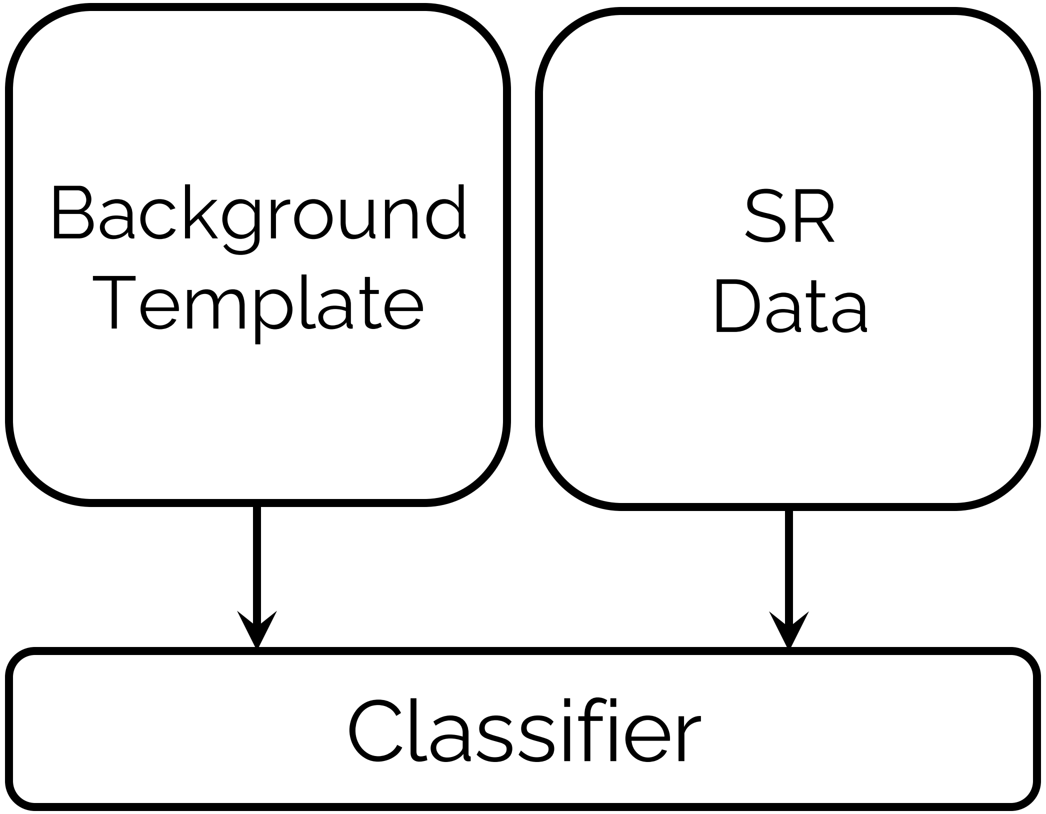
→Same decision boundaries



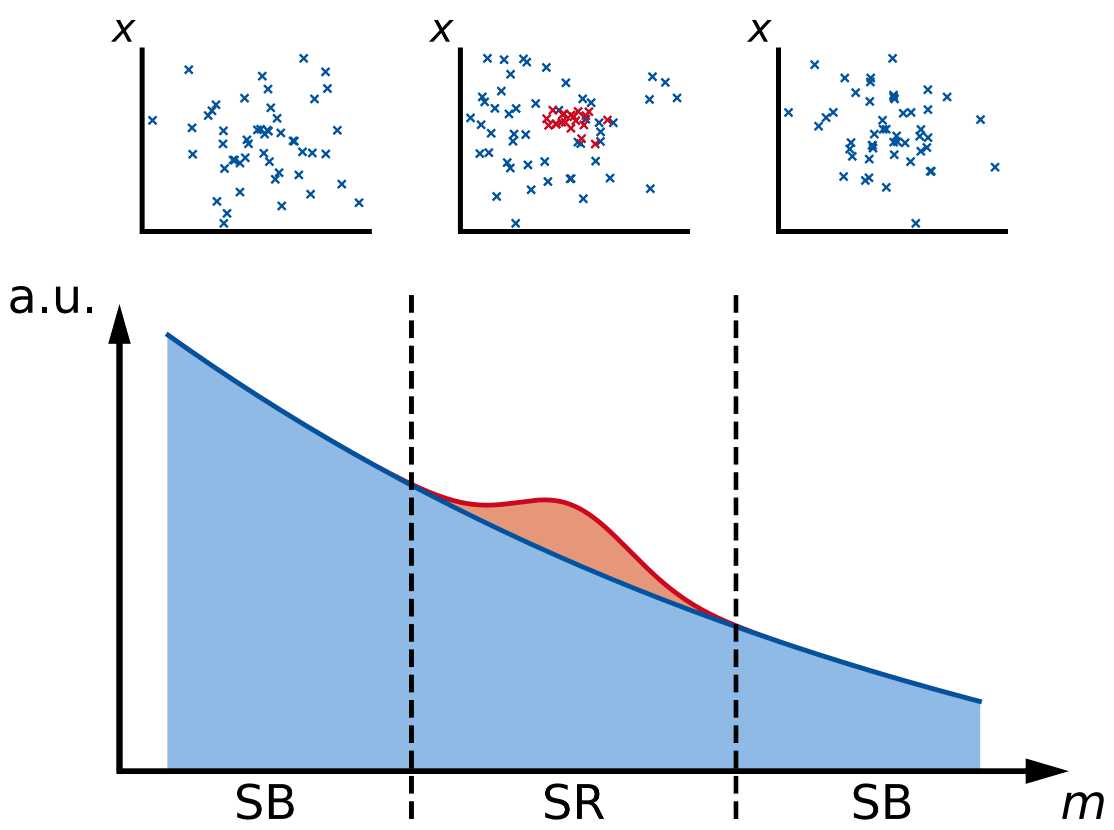
Application to resonance searches



Recreated from [2109.00546]



Application to resonance searches



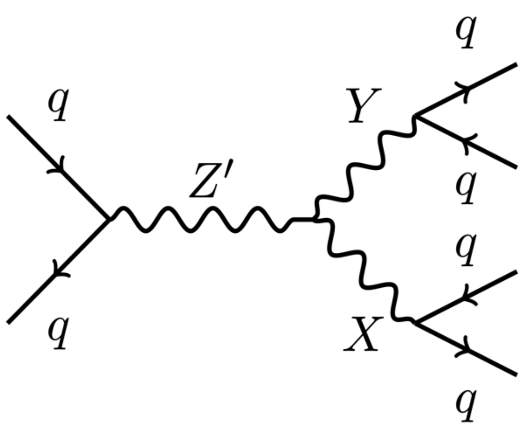
Recreated from [2109.00546]

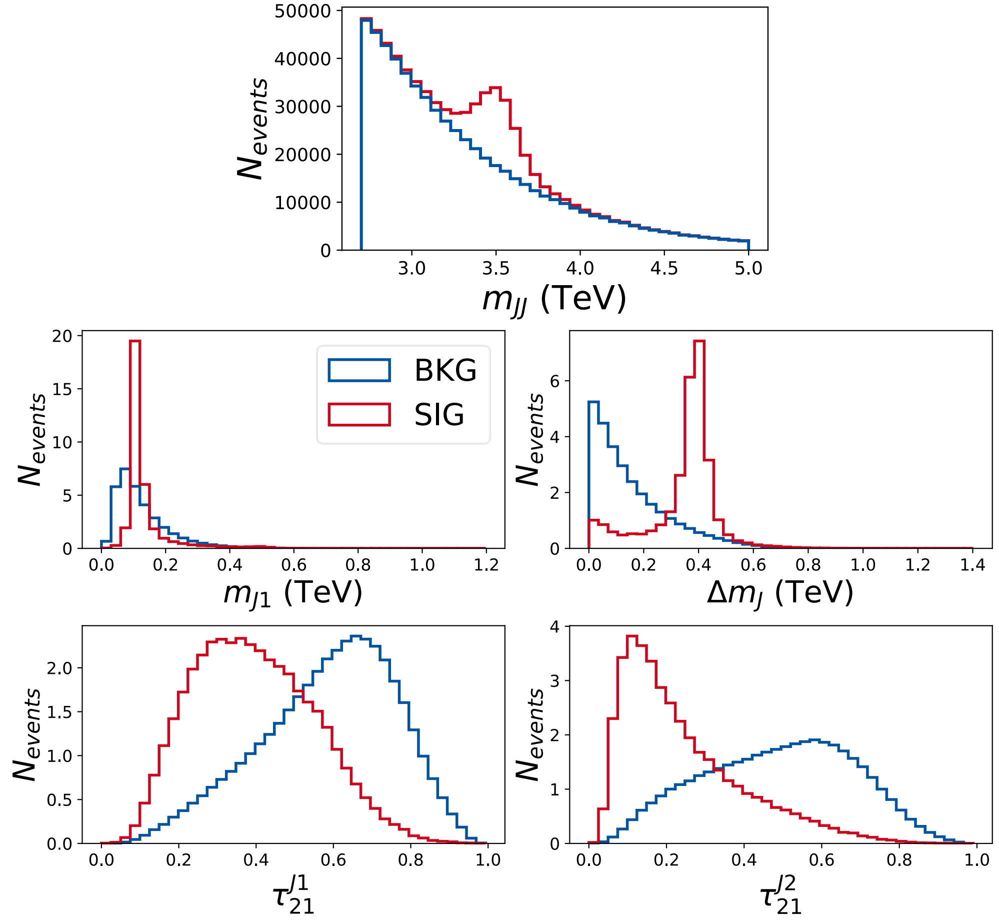
1. Idealized Anomaly Detector: SR background
2. CWoLa Hunting: SB data
3. CATHODE:
   a. Train DE on SB data to learn $p_{\mathrm{bkg}}(x|m)$
   b. Sample into SR

LHCO R&D dataset

"The LHC Olympics 2020: A Community Challenge for Anomaly Detection in High Energy Physics" [2101.08320], G. Kasieczka, B. Nachman, D. Shih et. al.

- Benchmark dataset for anomaly detection
- QCD dijet background (1M events)
- Signal (0 or 1000 events)



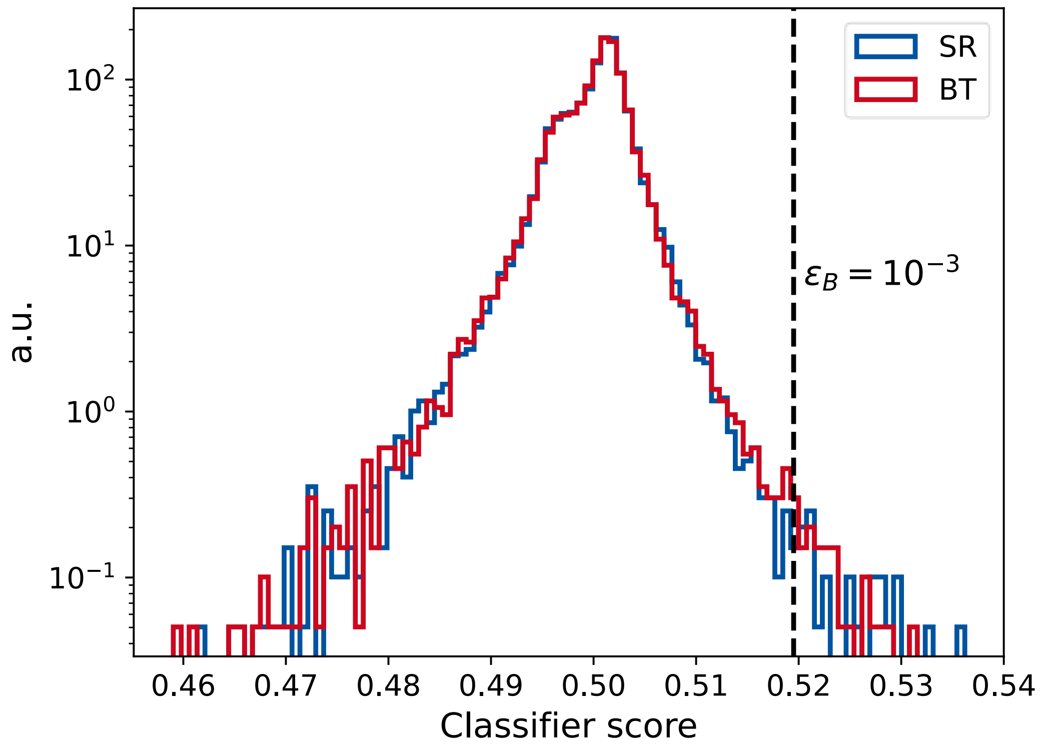


- Use 9 windows with centers at

$$m_{JJ,n} = 3.5\ \text{TeV} + (5 - n) \cdot 0.1\ \text{TeV}$$

Idealized Anomaly Detector

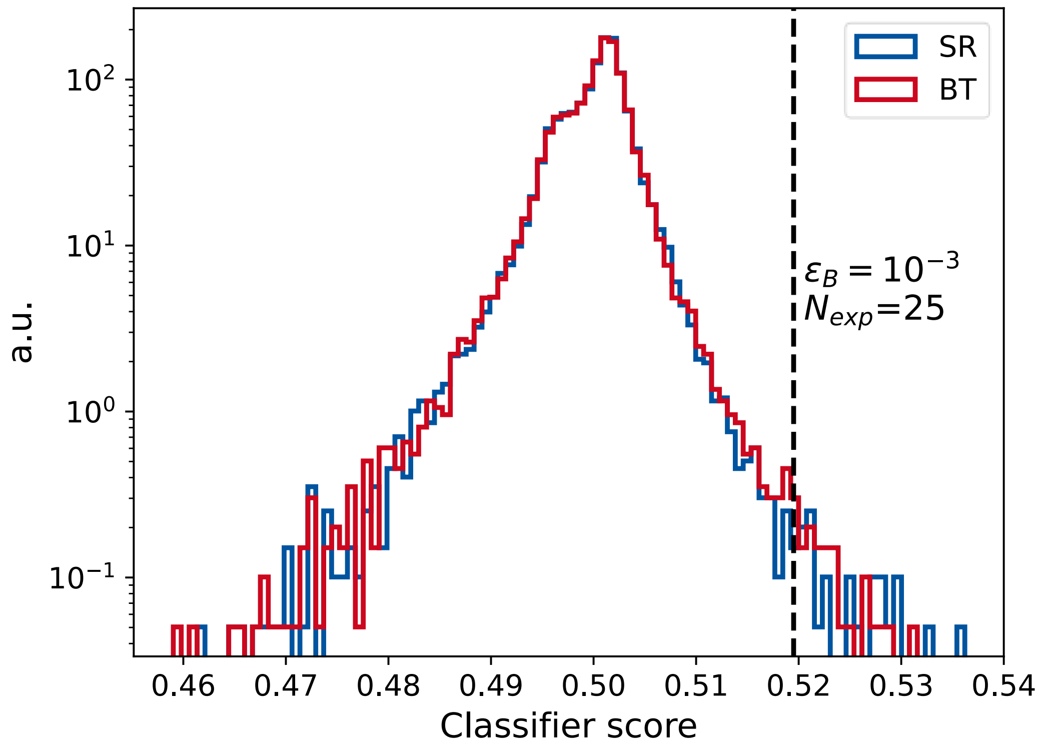
Without signal



Idealized Anomaly Detector

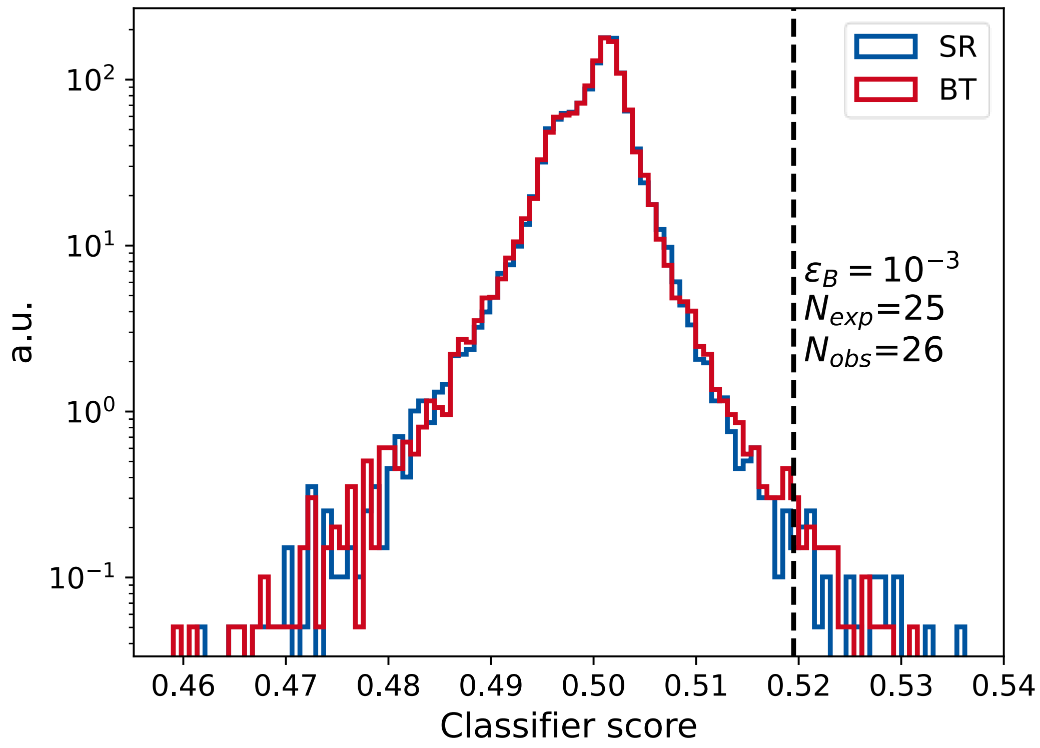
Without signal

$$N_{\mathrm{exp}} = \epsilon_B N_{SR}$$



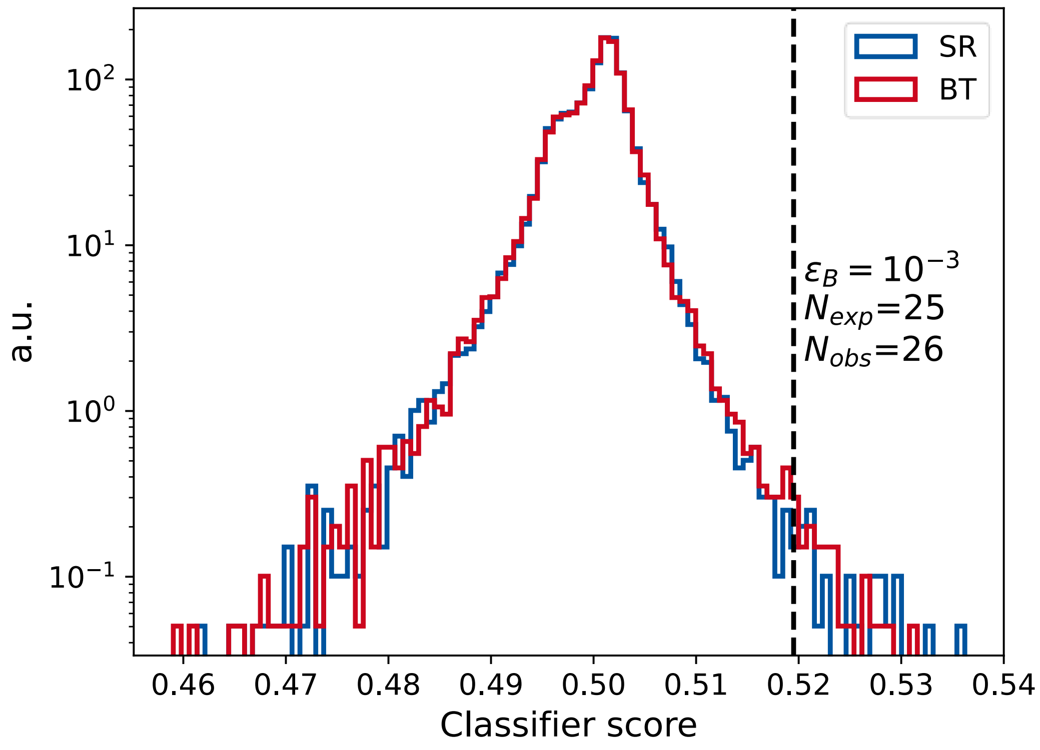
Idealized Anomaly Detector

Without signal

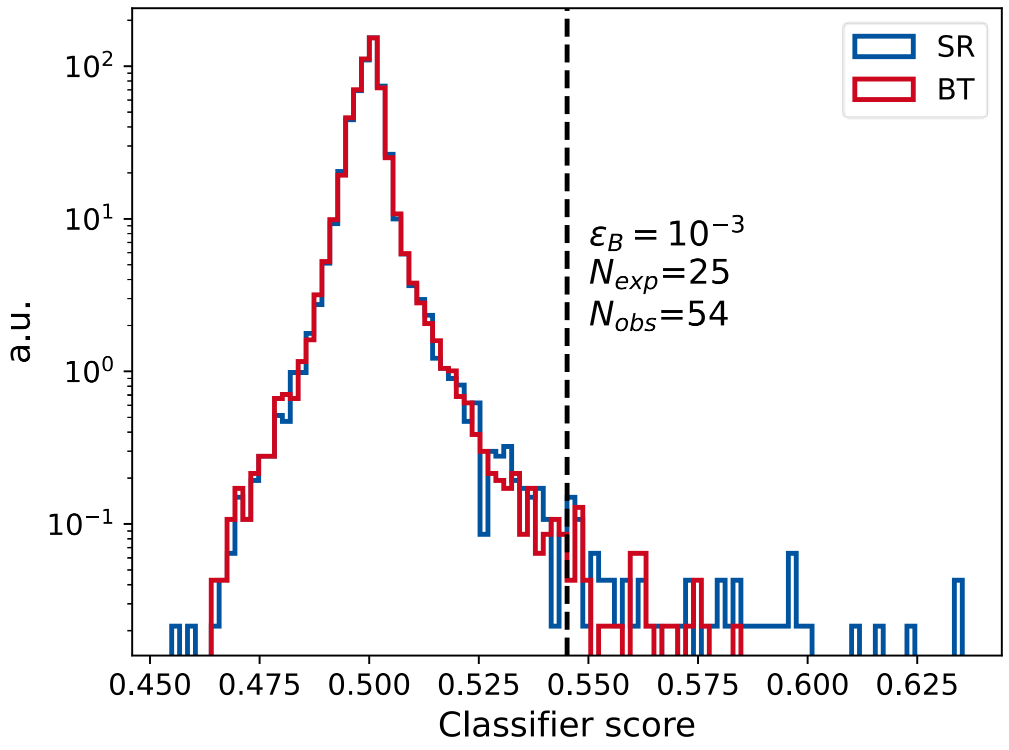


Idealized Anomaly Detector

Without signal



With signal



Calculating the Significance

$$\mathcal{S} = \frac{N_{\mathrm{obs}} - N_{\mathrm{exp}}}{\sqrt{N_{\mathrm{exp}}^2 \left(N_{\mathrm{exp}}^{-1} + \sigma_{\mathrm{exp}}^2\right)}}$$

- ϵ_B — Working point (determined on samples)
- $N_{\mathrm{obs}}(\epsilon_B)$ — Observed number of events
- $N_{\mathrm{exp}}(\epsilon_B)$ — Expected number of events (background-only assumption)
- $\sigma_{\mathrm{exp}}(\epsilon_B)$ — Relative error on expectation (from determination of working point)

Calculating the Significance

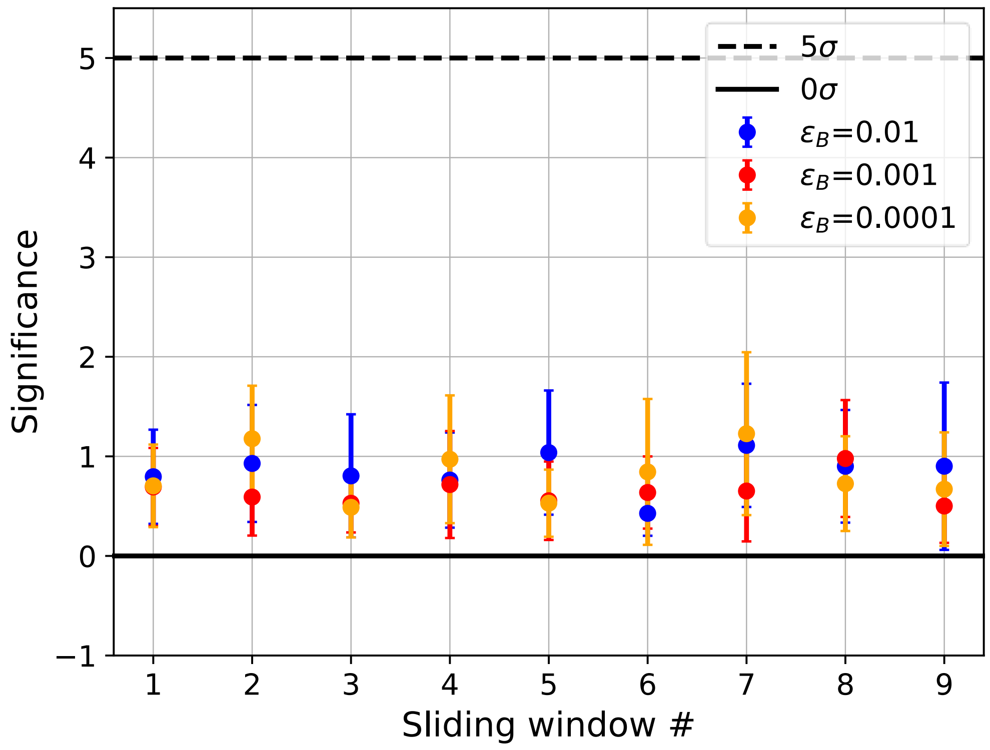
"Asymptotic formulae for likelihood-based tests of new physics" [1007.1727], G. Cowan, K. Cranmer, E. Gross, and O. Vitells

$$\mathcal{S} = \left[2\left(N_{\text{obs}} \ln \frac{N_{\text{obs}}\left(N_{\text{exp}}^{-1} + \sigma_{\text{exp}}^2\right)}{1 + N_{\text{obs}}\sigma_{\text{exp}}^2} - \frac{1}{\sigma_{\text{exp}}^2} \ln \frac{1 + N_{\text{obs}}\sigma_{\text{exp}}^2}{1 + N_{\text{exp}}\sigma_{\text{exp}}^2}\right)\right]^{1/2}$$

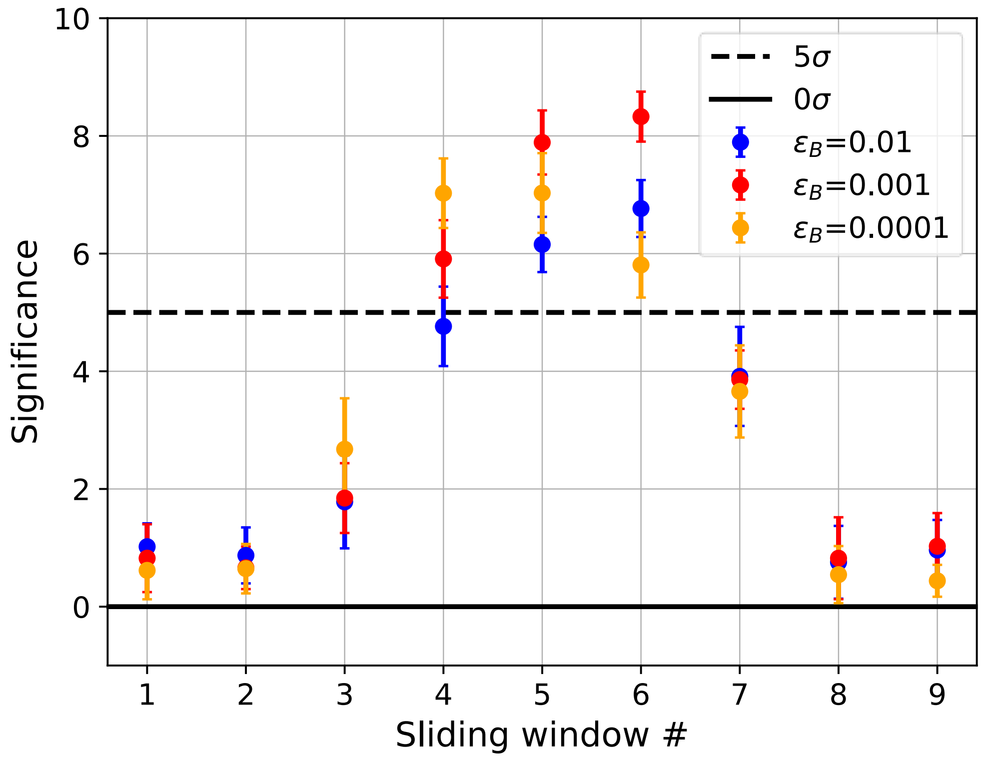
- ϵ_B — Working point (determined on samples)
- $N_{\text{obs}}(\epsilon_B)$ — Observed number of events
- $N_{\text{exp}}(\epsilon_B)$ — Expected number of events (background-only assumption)
- $\sigma_{\text{exp}}(\epsilon_B)$ — Relative error on expectation (from determination of working point)

Idealized Anomaly Detector

Without signal

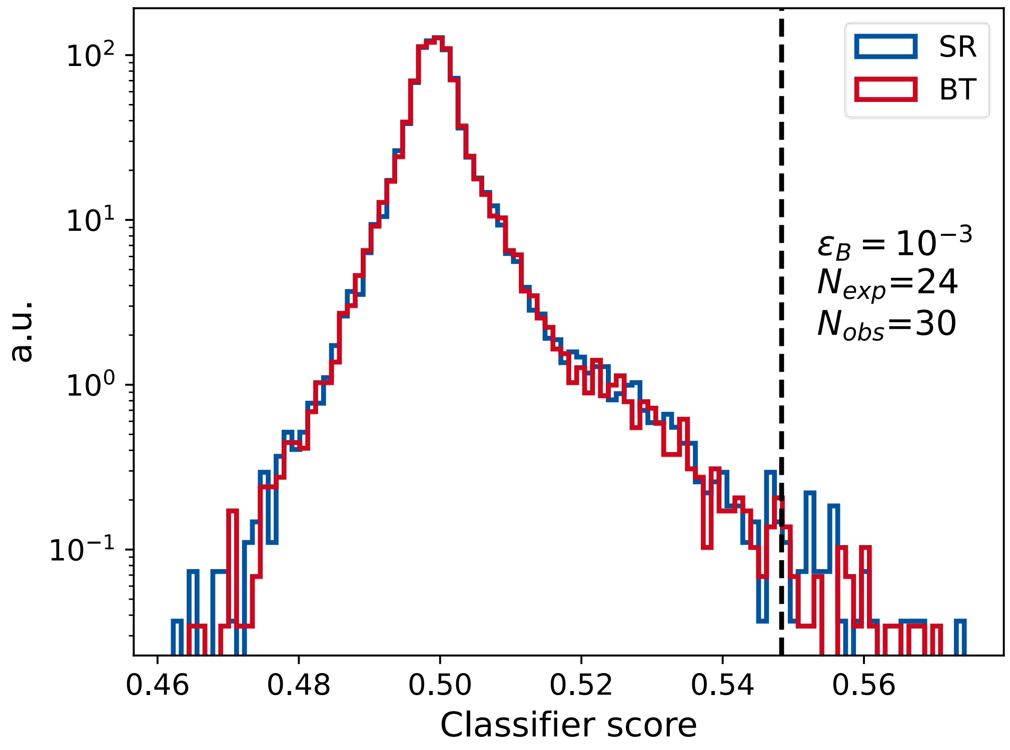


With signal

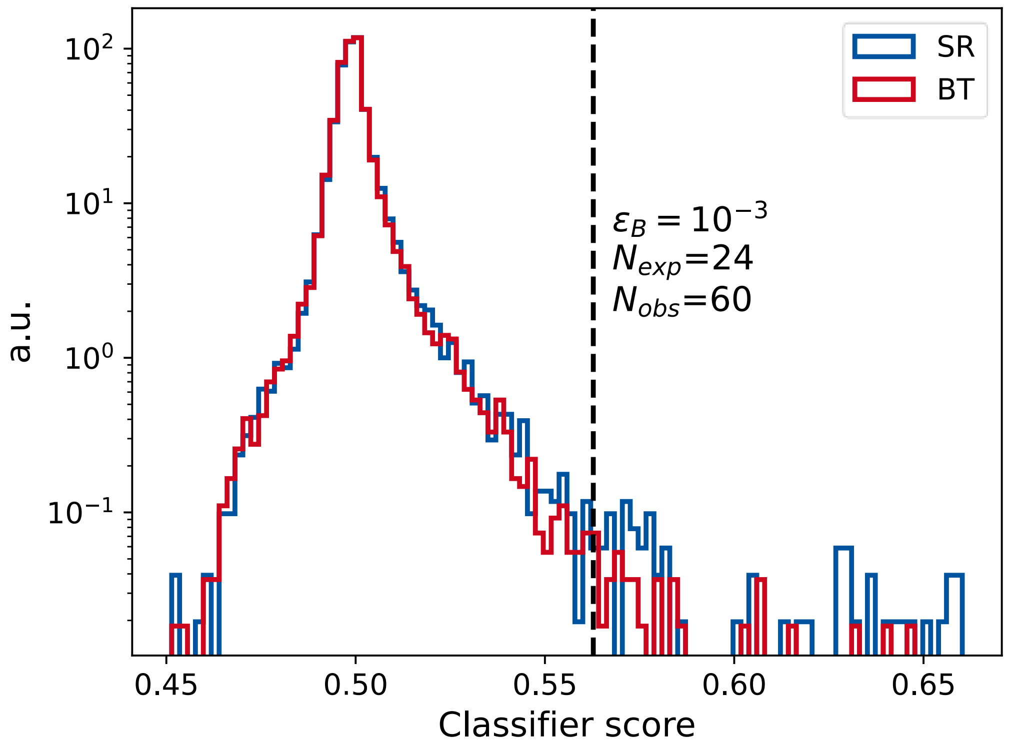


CWoLa Hunting

Without signal

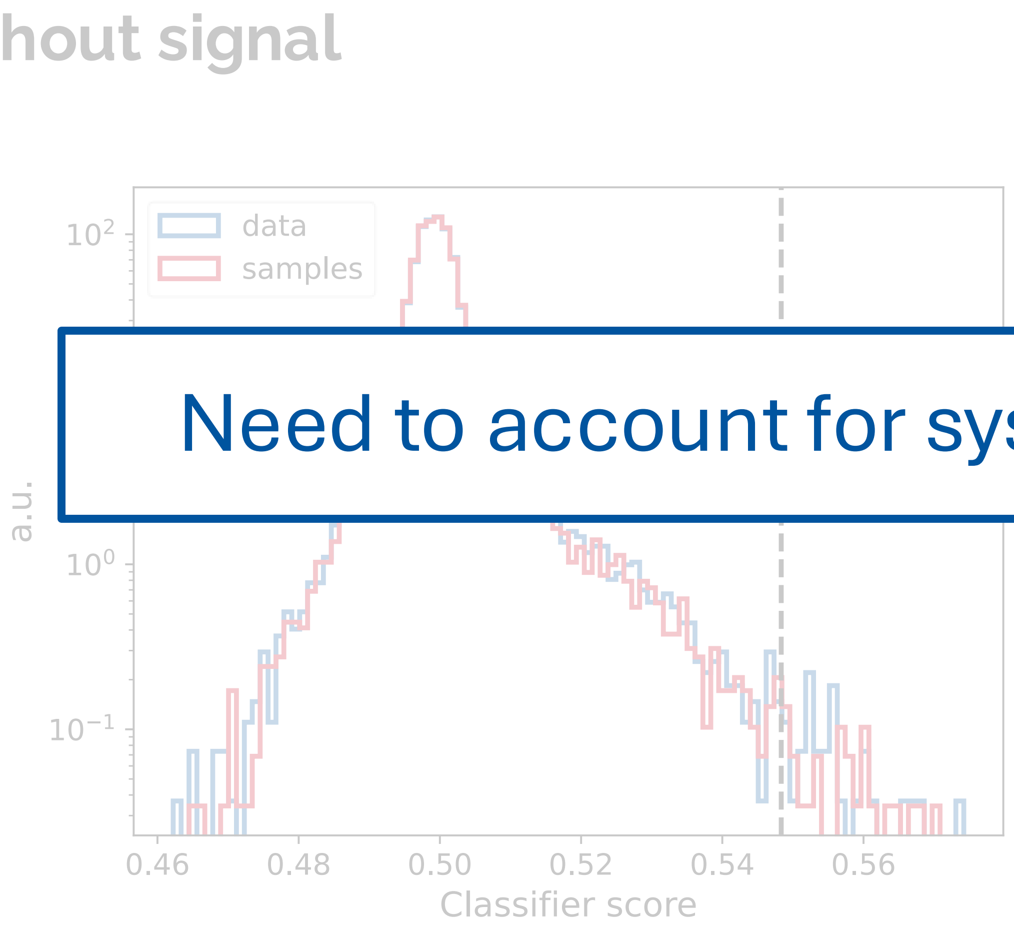


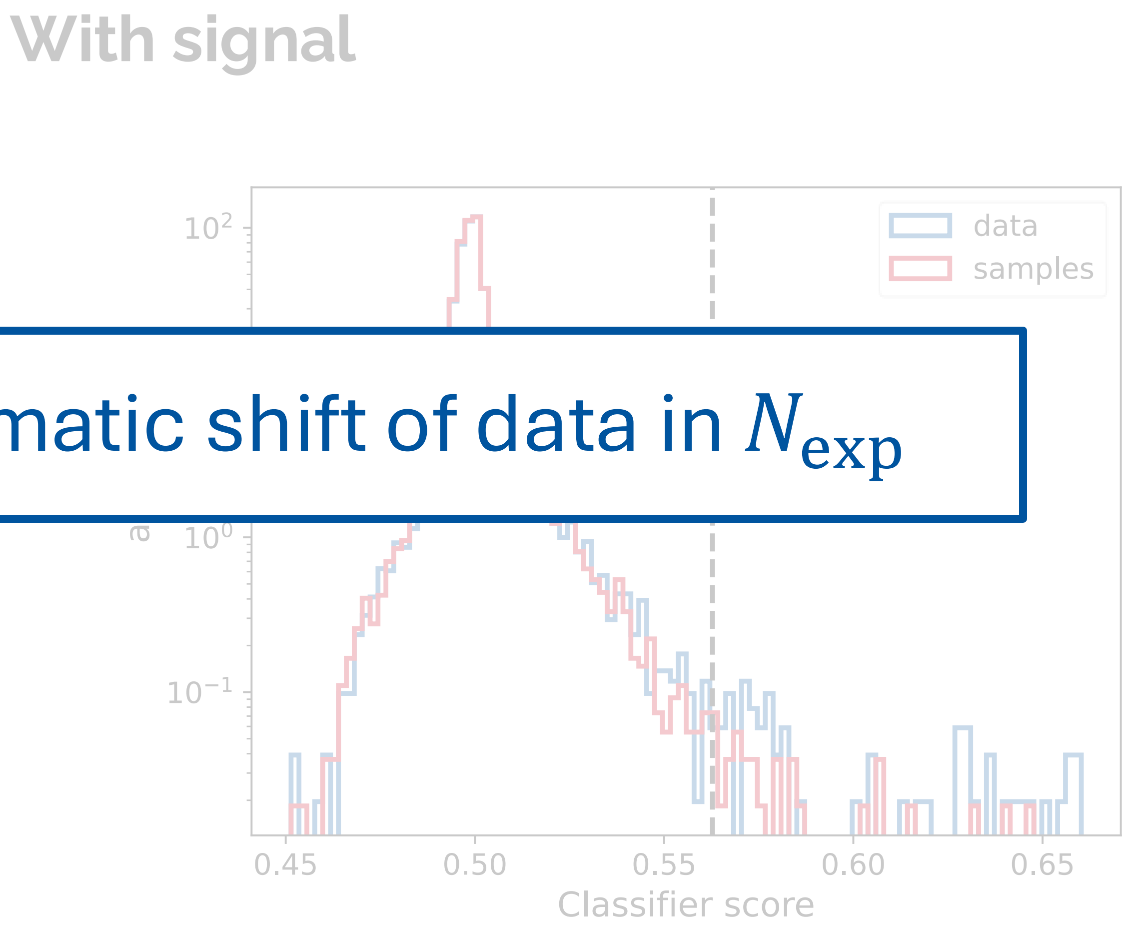
With signal



CWoLa

Institute for Theoretical Particle Physics and Cosmology
RWTH AACHEN UNIVERSITY





Need to account for systematic shift of data in N_{exp}

Systematic shift

Define a relative systematic shift per window n

$$\delta_{\text{sys},\, n} = \frac{N_{\text{obs,n}} - \epsilon_B N_{\text{SR,n}}}{\epsilon_B N_{\text{SR,n}}}$$

and its average over all windows δ_{sys}.

Systematic shift

Define a relative systematic shift per window n

$$\delta_{\mathrm{sys},\,n} = \frac{N_{\mathrm{obs,n}} - \epsilon_B N_{\mathrm{SR,n}}}{\epsilon_B N_{\mathrm{SR,n}}}$$

and its average over all windows δ_{sys}.

	IAD	CWoLa & CATHODE
N_{exp}	$\epsilon_B N_{\mathrm{SR}}$	$\epsilon_B N_{\mathrm{SR}}(1 + \delta_{\mathrm{sys}})$
σ_{exp}	σ_{ϵ_B}	$\sqrt{\sigma_{\epsilon_B}^2 + \sigma_{\delta_{\mathrm{sys}}}^2}$

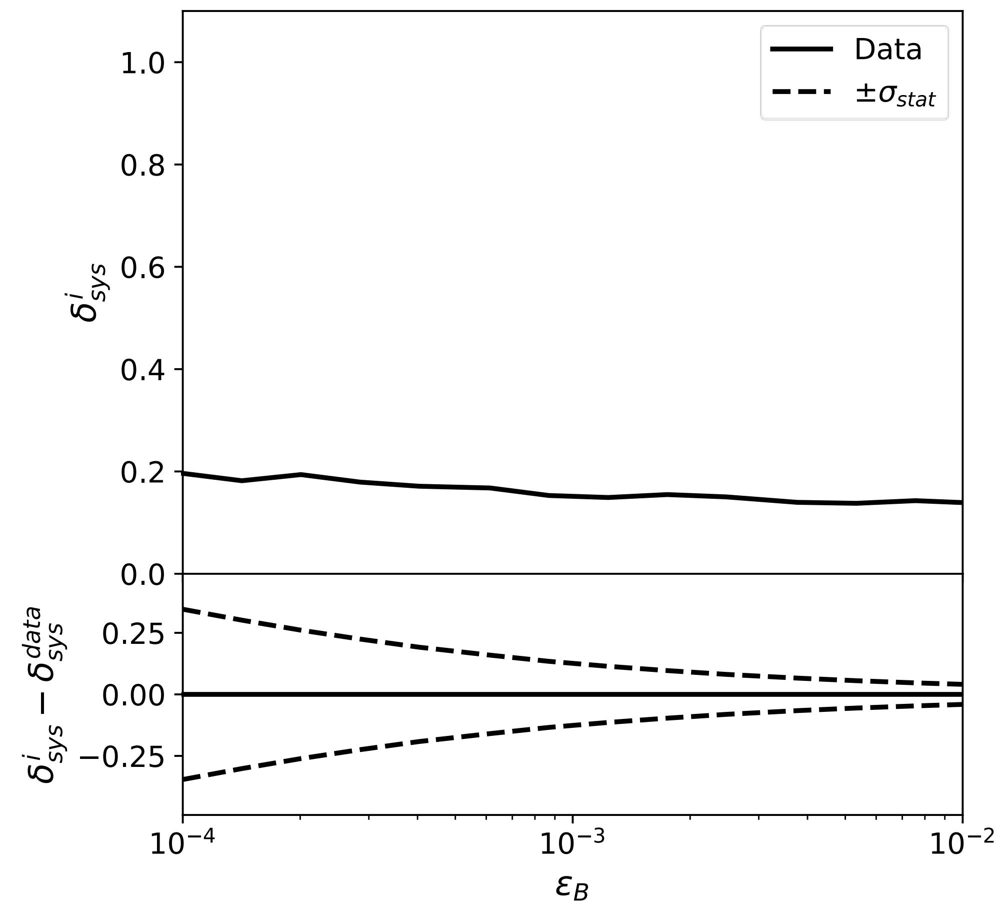
Systematic shift

Define a relative systematic shift per window n

$$\delta_{\text{sys},\,n} = \frac{N_{\text{obs,n}} - \epsilon_B N_{\text{SR,n}}}{\epsilon_B N_{\text{SR,n}}}$$

and its average over all windows δ_{sys}.

	IAD	CWoLa & CATHODE
N_{exp}	$\epsilon_B N_{\text{SR}}$	$\epsilon_B N_{\text{SR}}(1+\delta_{\text{sys}})$
σ_{exp}	σ_{ϵ_B}	$\sqrt{\sigma_{\epsilon_B}^2 + \sigma_{\delta_{\text{sys}}}^2}$



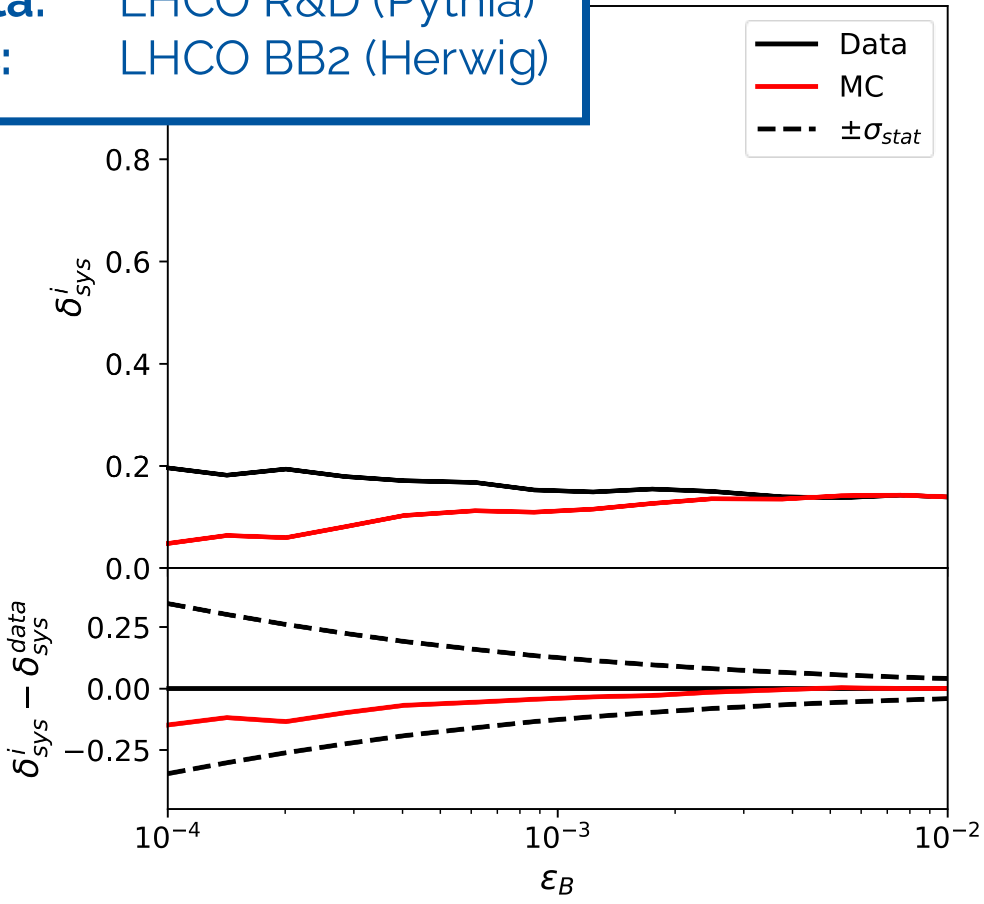
Define a relative systematic shift per window n

$$\delta_{\text{sys},\,n} = \frac{N_{\text{obs,n}} - \epsilon_B N_{\text{SR,n}}}{\epsilon_B N_{\text{SR,n}}}$$

and its average over all windows δ_{sys}.

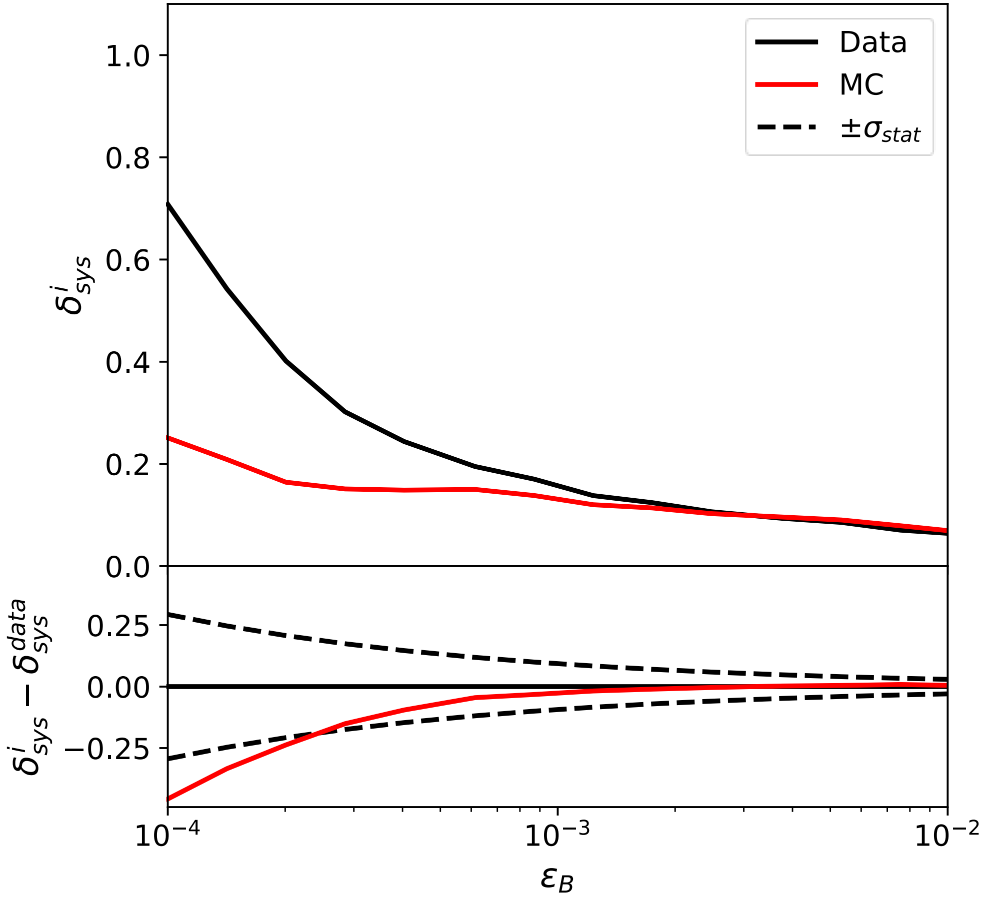
	IAD	CWoLa & CATHODE
N_{exp}	$\epsilon_B N_{\text{SR}}$	$\epsilon_B N_{\text{SR}}(1+\delta_{\text{sys}})$
σ_{exp}	σ_{ϵ_B}	$\sqrt{\sigma_{\epsilon_B}^2 + \sigma_{\delta_{\text{sys}}}^2}$

Data: LHCO R&D (Pythia)
MC: LHCO BB2 (Herwig)

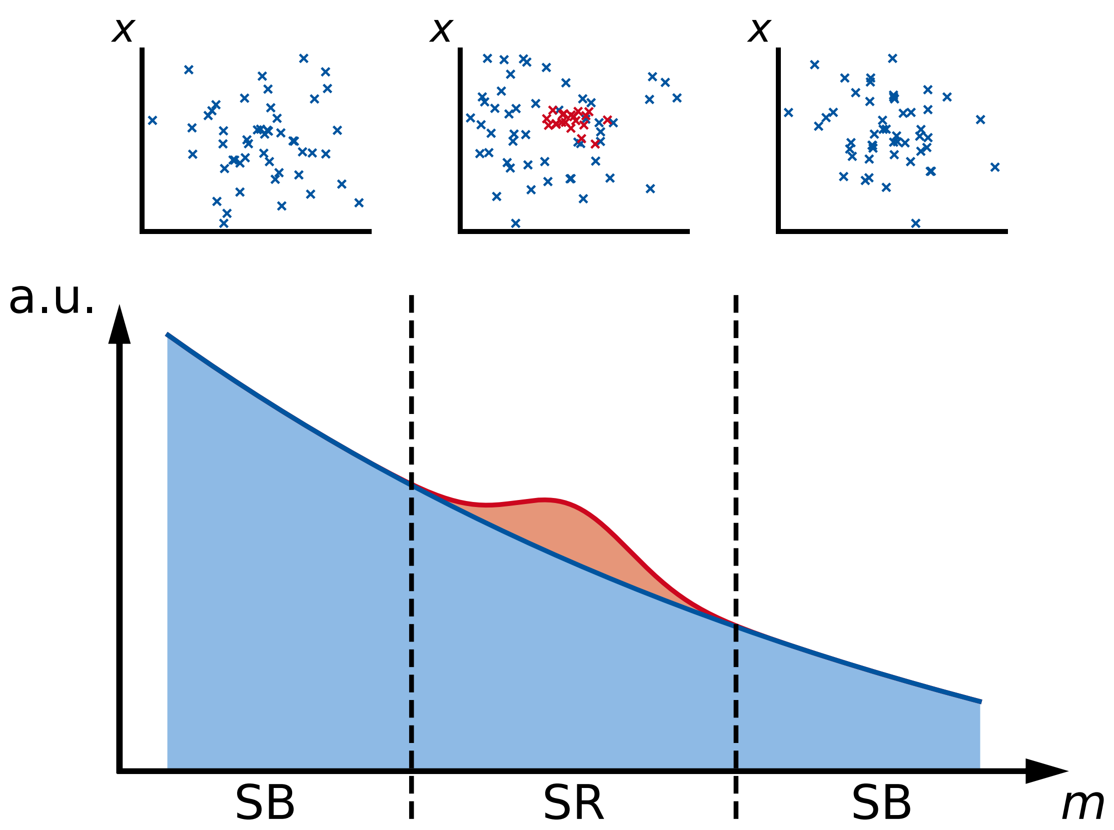


Systematic shift for CATHODE

- For CATHODE, large deviation of MC value from data value visible

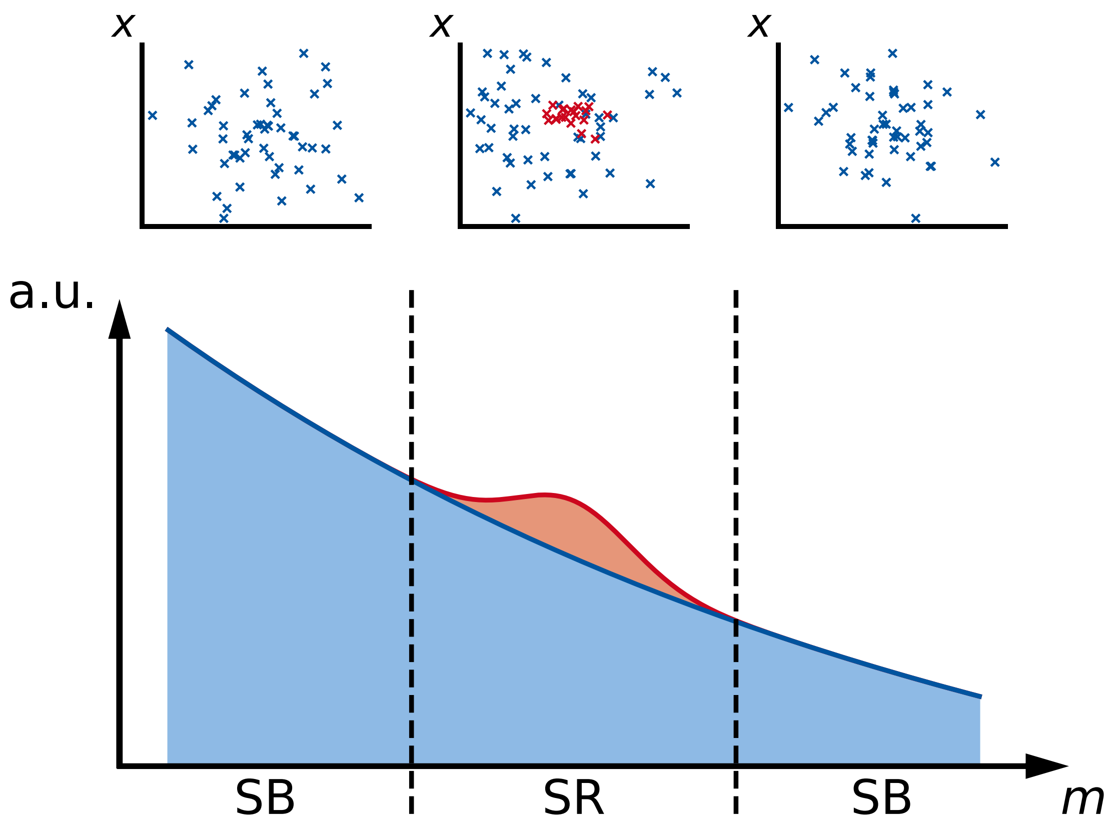


Data-driven determination of δ_{sys}



Recreated from [2109.00546]

Data-driven determination of δ_{sys}

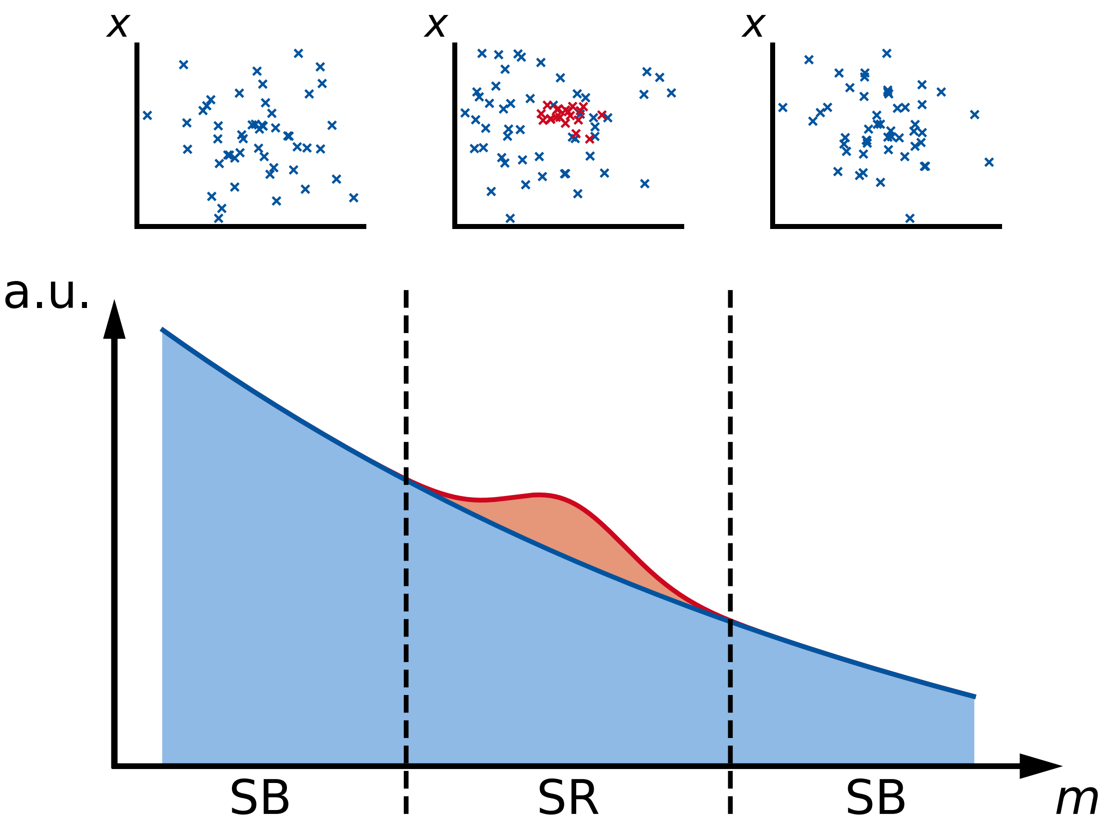


Recreated from [2109.00546]

Normal CATHODE procedure:

- Train DE on SB
- Sample DE in SR
- Train SR samples vs SR data classifier

Data-driven determination of δ_{sys}



Recreated from [2109.00546]

Normal CATHODE procedure:

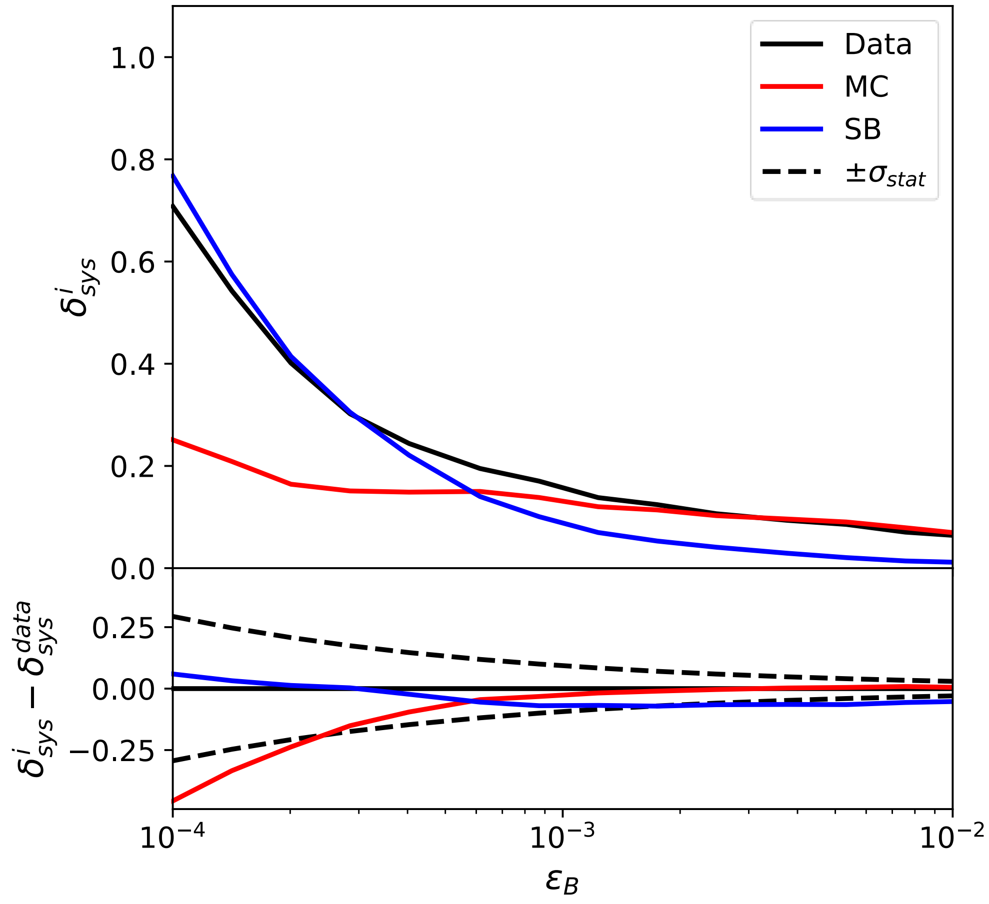
- Train DE on SB
- Sample DE in SR
- Train SR samples vs SR data classifier

Determine δ_{sys} by

- Train DE on SB
- Sample DE in SB
- Train SB samples vs SB data classifier

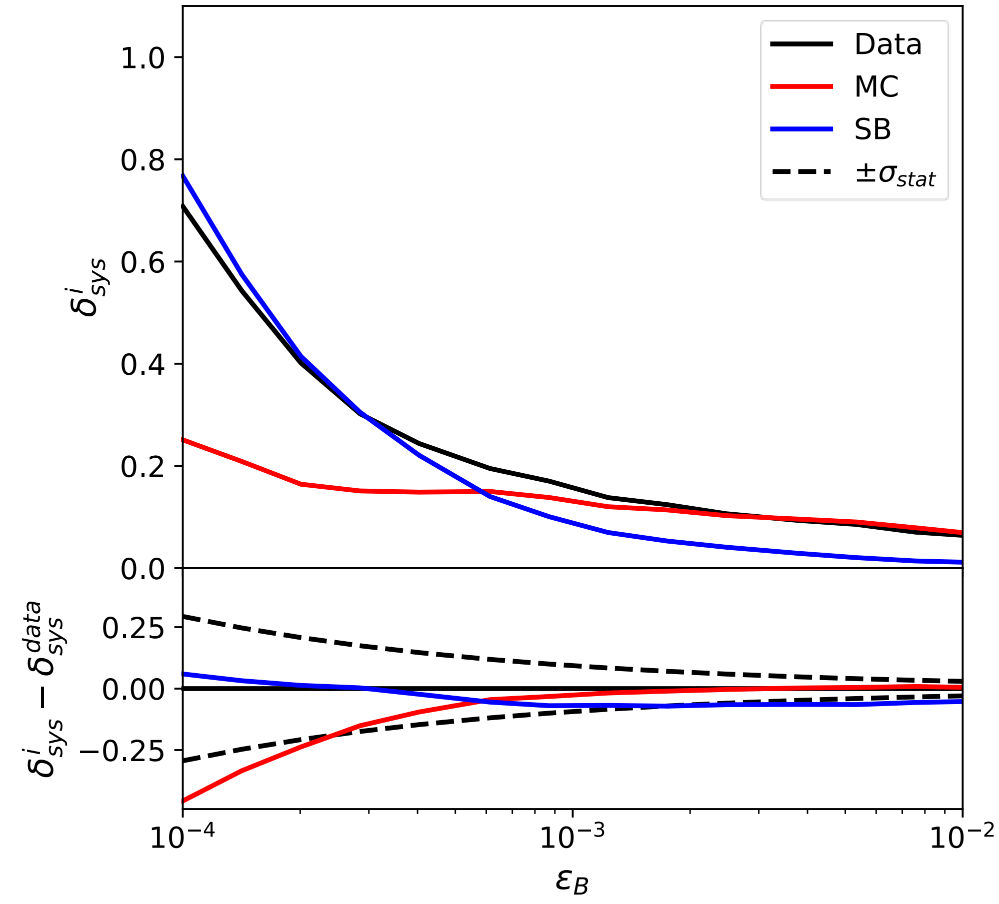
Systematic shift for CATHODE

- For CATHODE, large deviation of MC value from data value visible
- δ_{sys} from SB fits other part of the spectrum



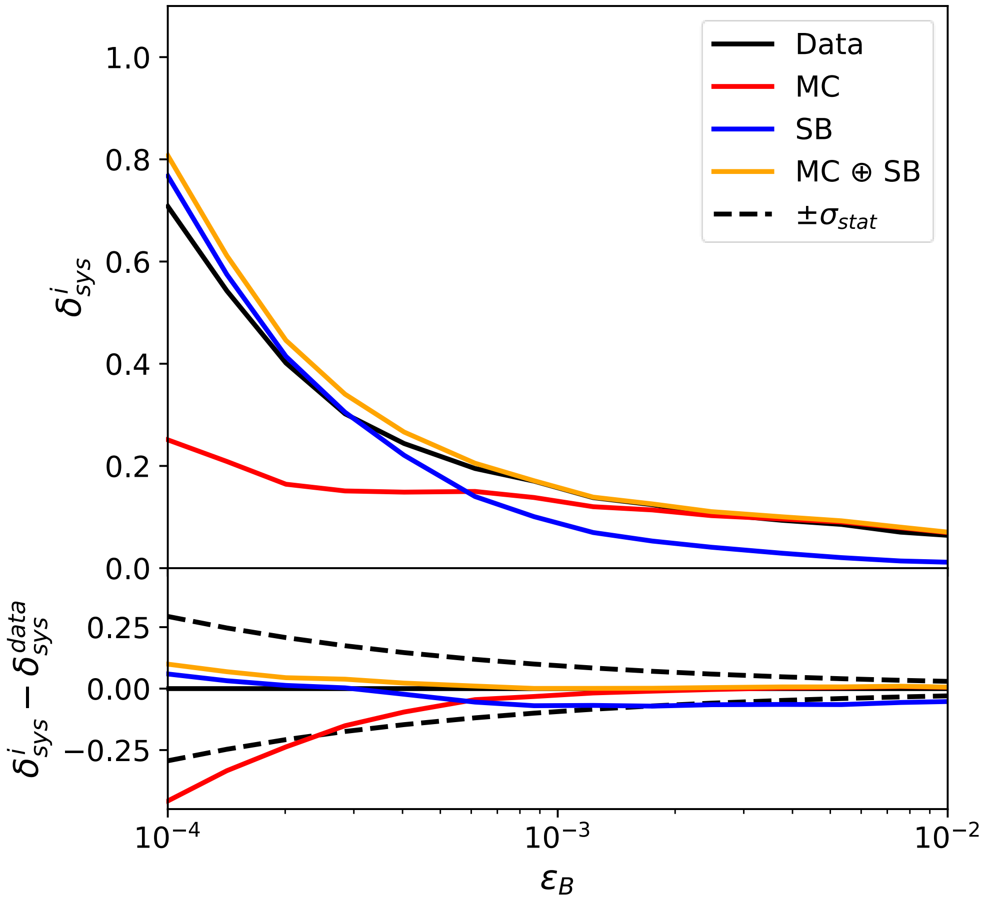
Systematic shift for CATHODE

- For CATHODE, large deviation of MC value from data value visible
  - ➢ Estimates interpolation error
- δ_{sys} from SB fits other part of the spectrum
  - ➢ Estimates difficulty of DE on tails



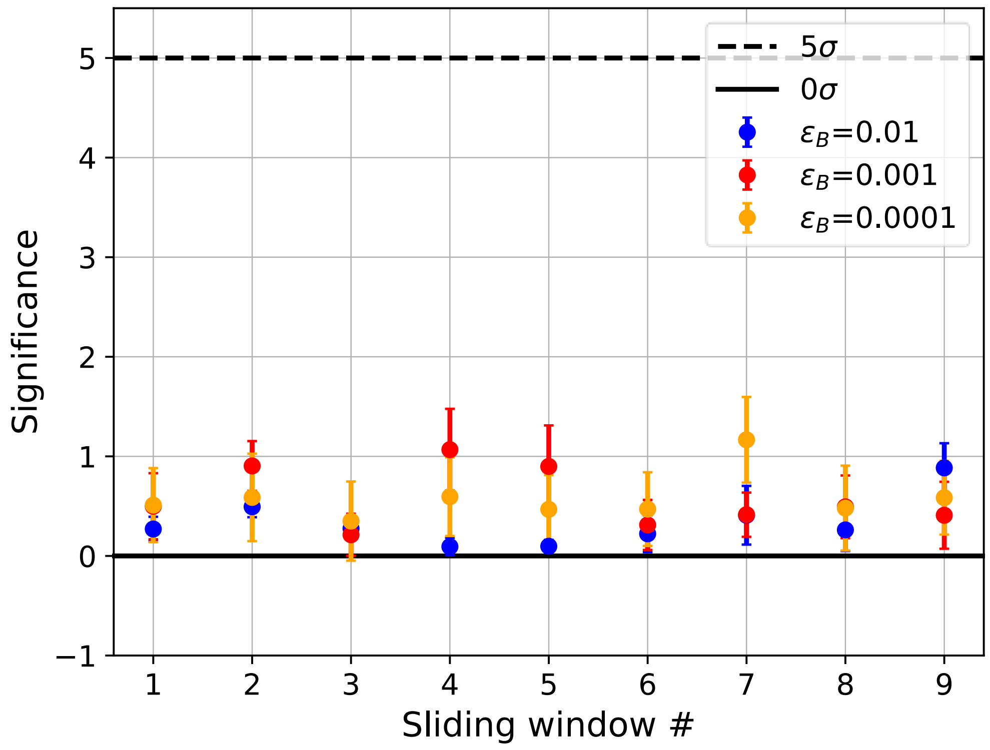
Systematic shift for CATHODE

- For CATHODE, large deviation of MC value from data value visible
  - ➢ Estimates interpolation error
- δ_{sys} from SB fits other part of the spectrum
  - ➢ Estimates difficulty of DE on tails
- Quadratic addition of both error sources
  - ➢ Fits whole spectrum

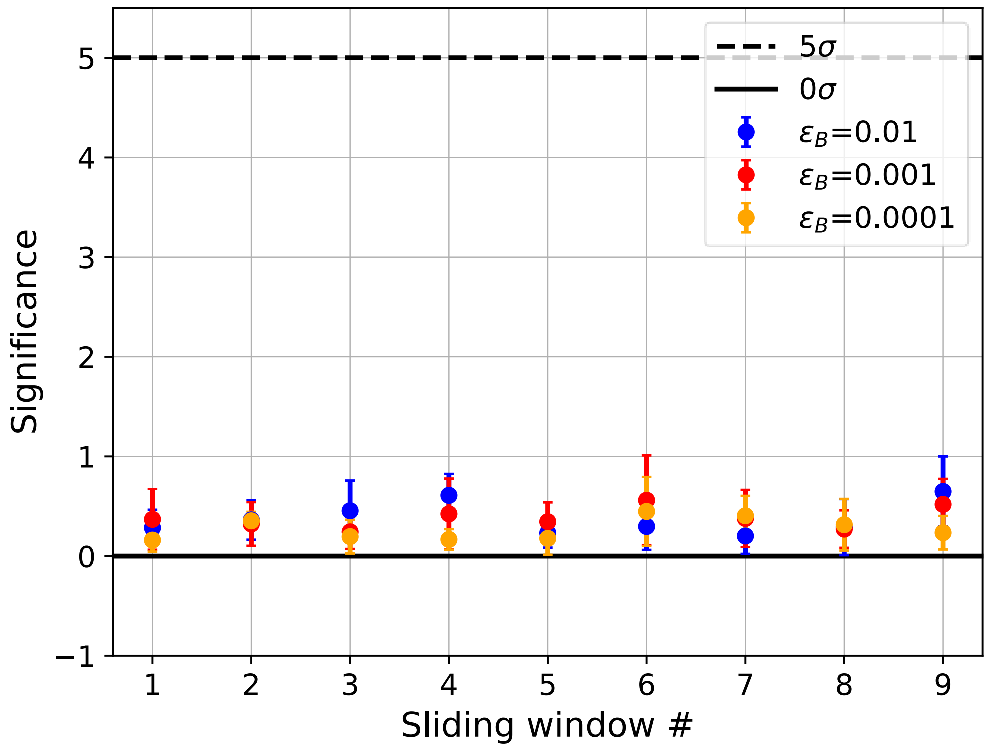


Background-only test

CWoLa Hunting

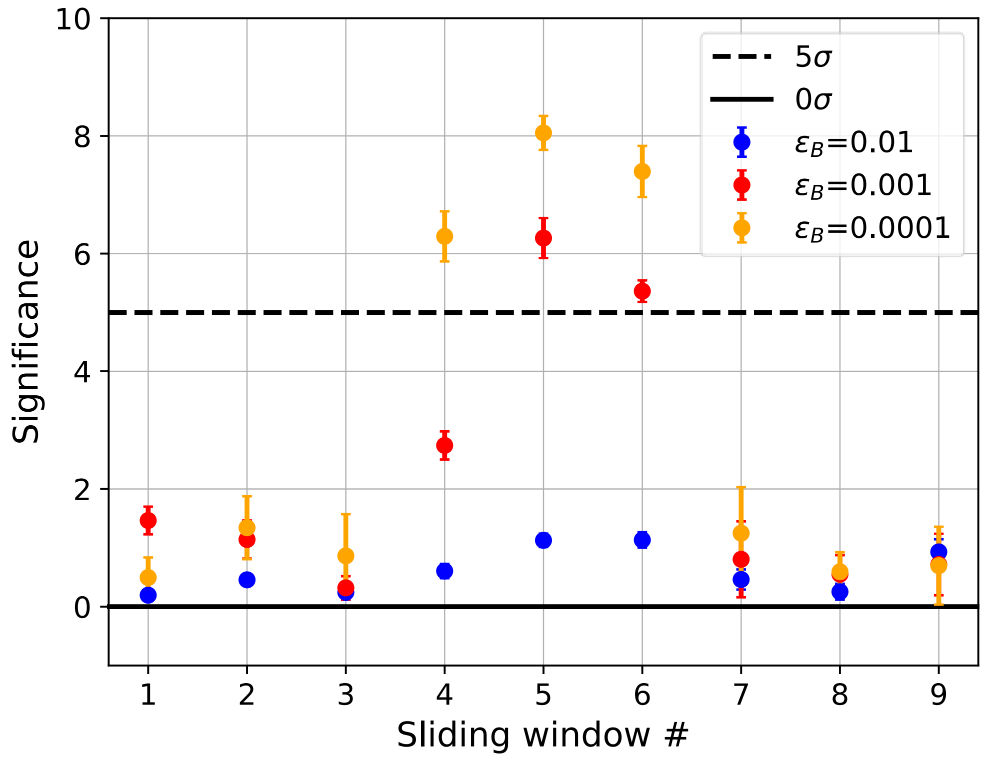


CATHODE

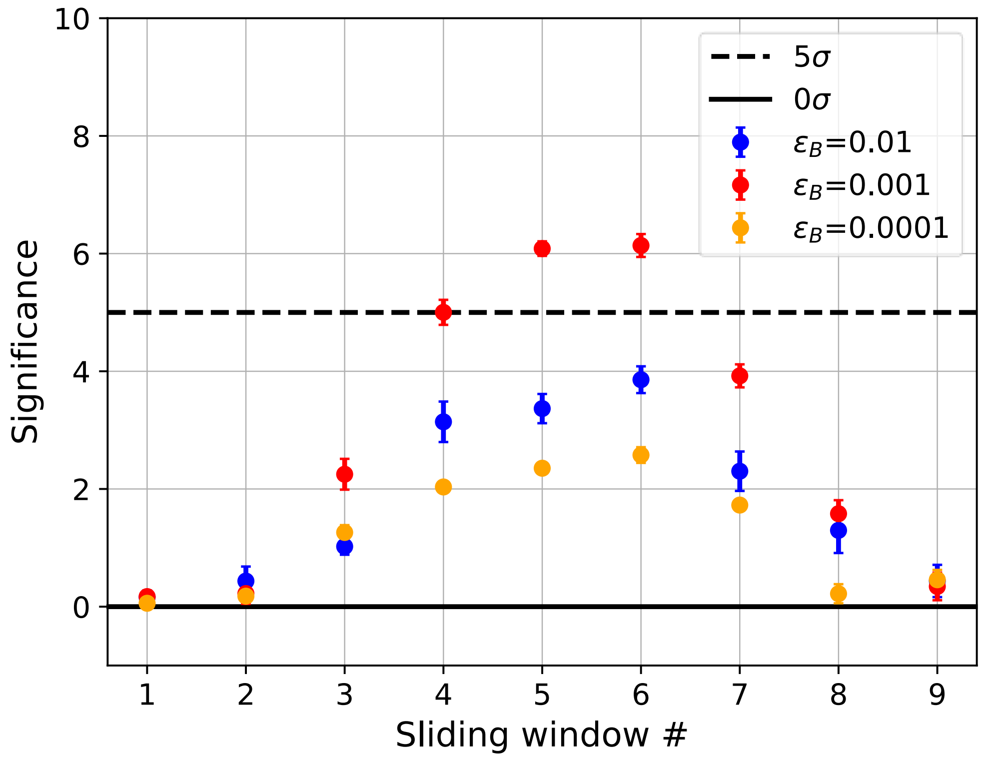


Adding signal

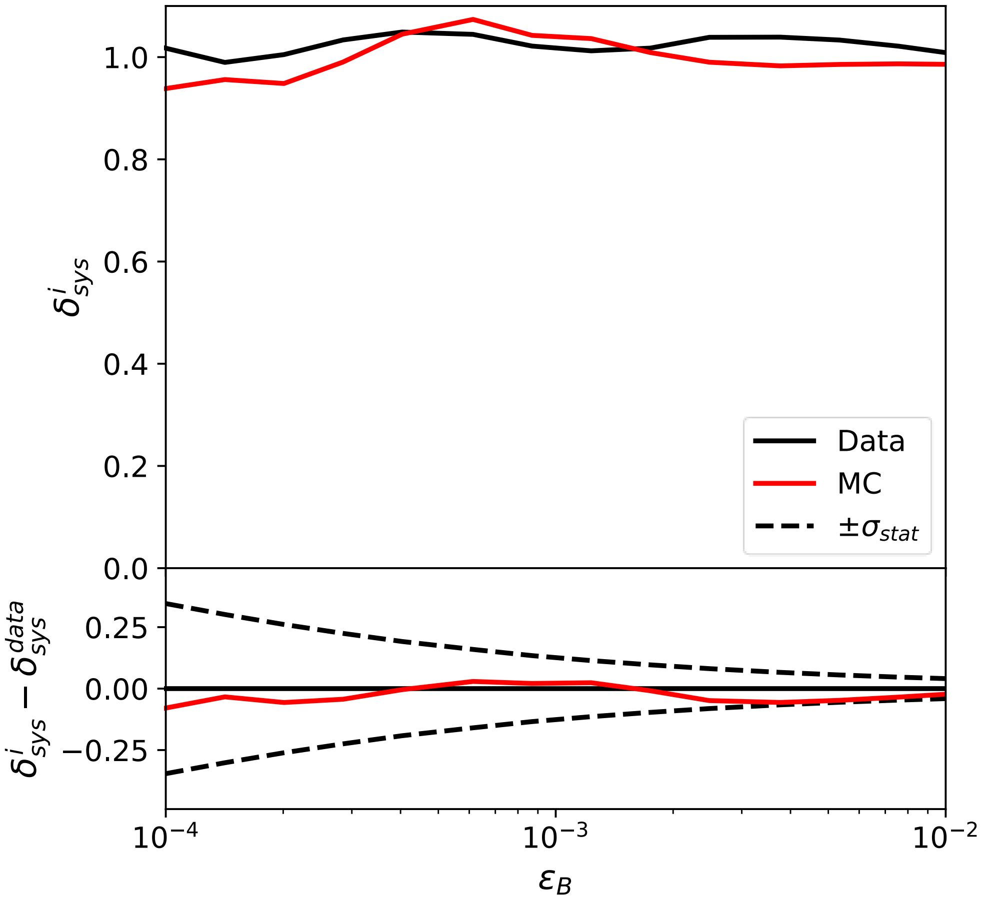
CWoLa Hunting



CATHODE



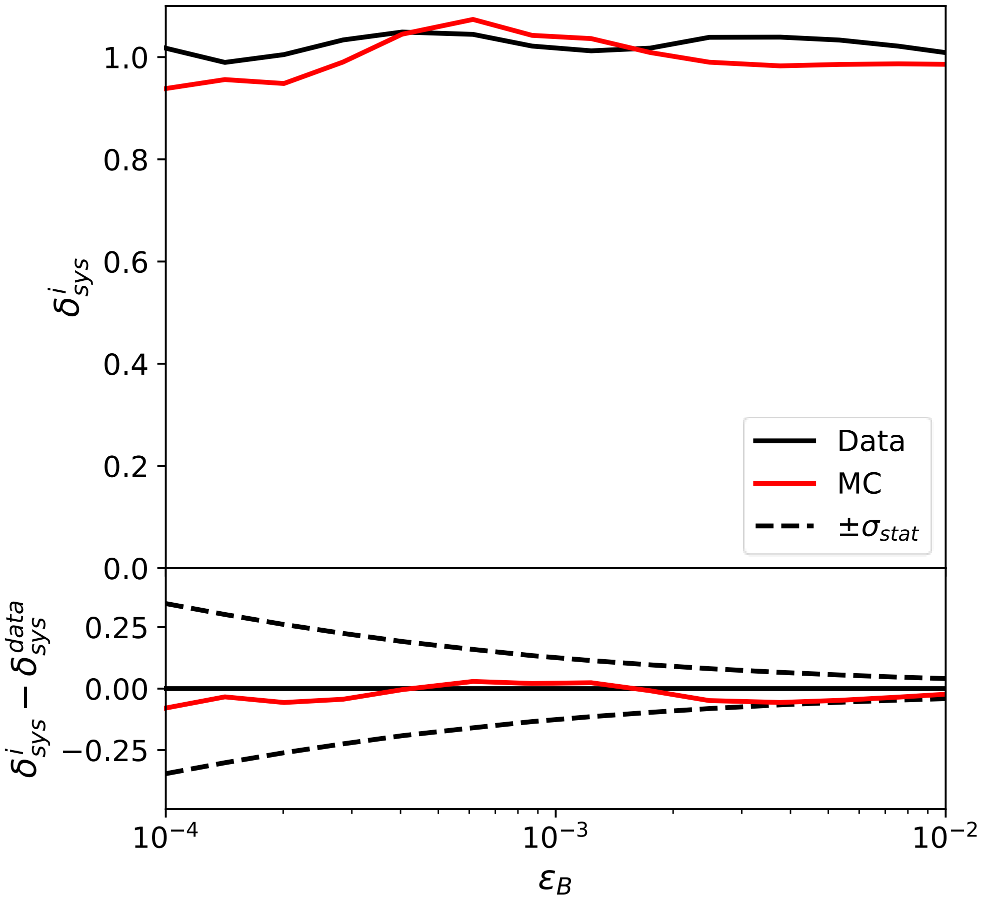
Breaking the Assumptions: CWoLa



Add a feature correlated with m_{JJ}:

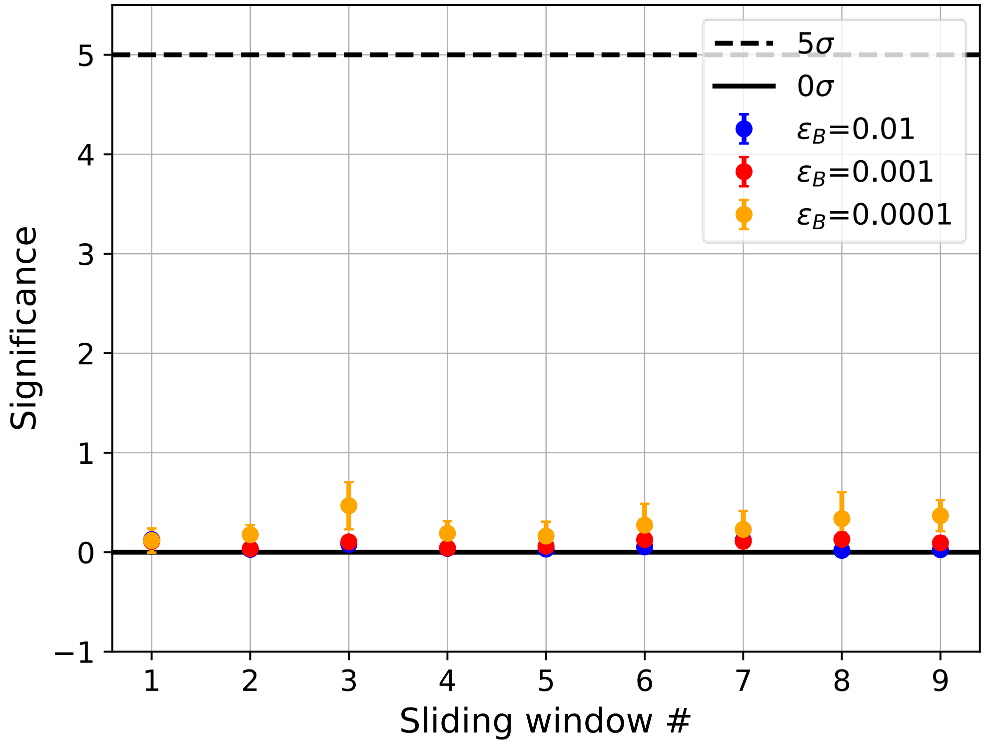
$$\Delta R = \sqrt{(\phi_{J1} - \phi_{J2})^2 + (\eta_{J1} - \eta_{J2})^2}$$

Breaking the Assumptions: CWoLa



Add a feature correlated with m_{JJ}:

$$\Delta R = \sqrt{(\phi_{J1} - \phi_{J2})^2 + (\eta_{J1} - \eta_{J2})^2}$$



Conclusion

- Alternative approach to bump hunts with direct background estimation instead of fits
  - Robust
    - No false discoveries observed
    - Even when method assumptions are violated
  - Simple
    - Statistical procedure
    - Systematic bias evaluation
  - Powerful
    - Significant deviations observed

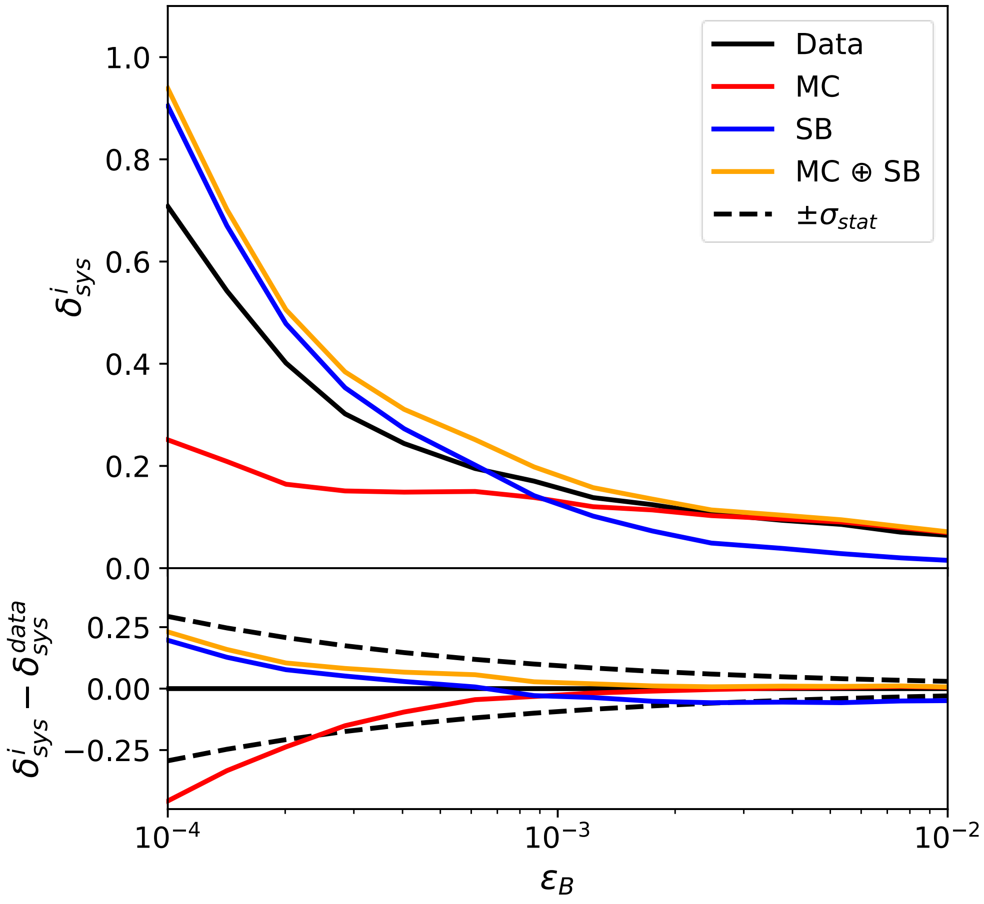
Accurate and robust methods for direct background estimation in resonant anomaly detection

R. Das, T. Finke, **MH**, G. Kasieczka, M. Krämer, A. Mück, D. Shih

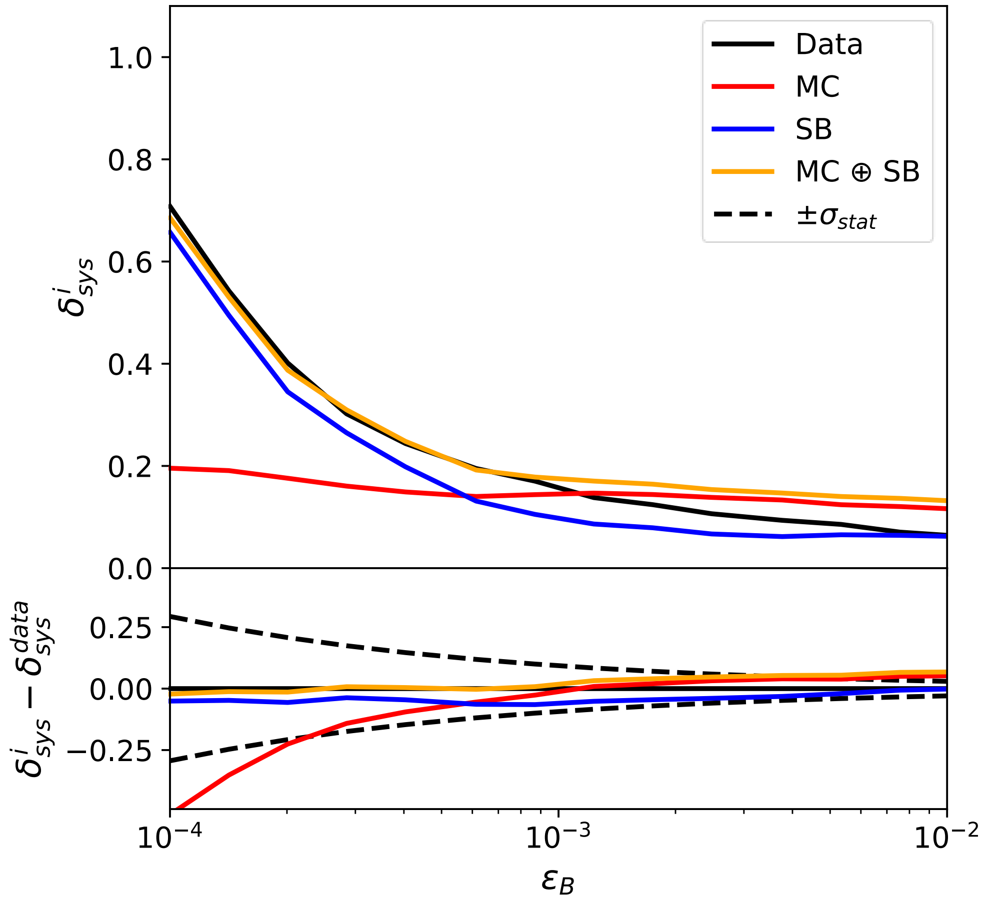
Backup

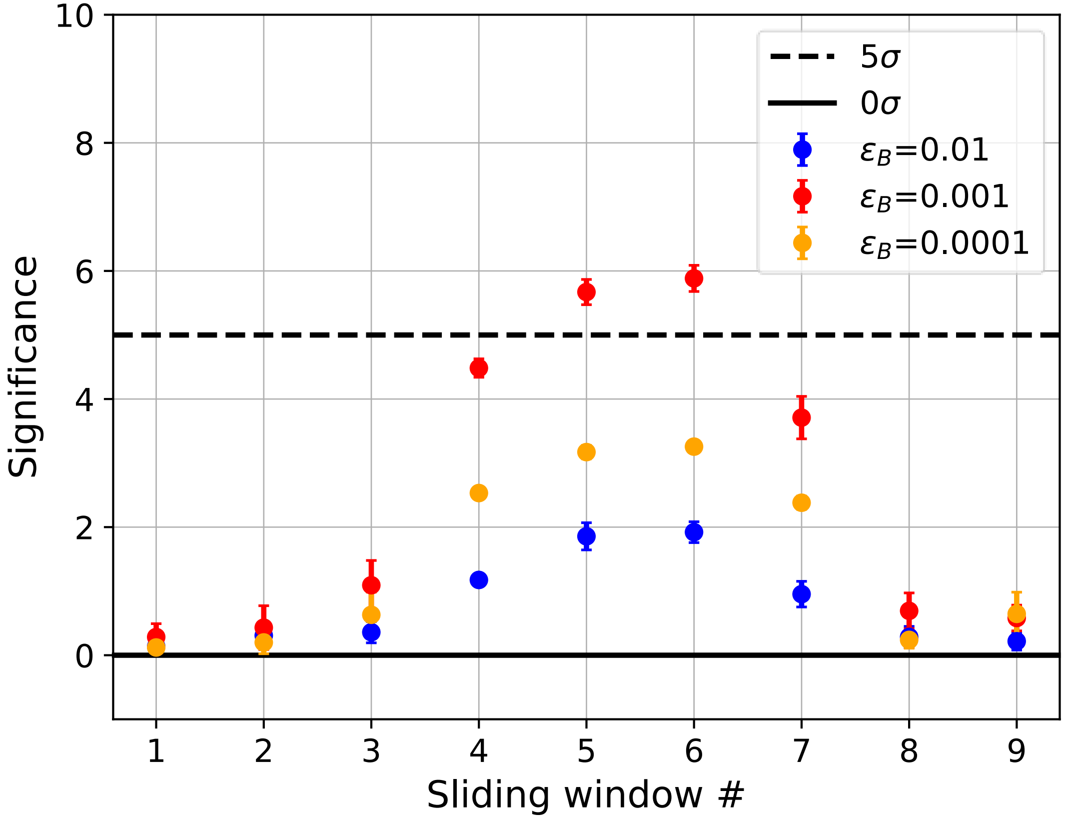
CATHODE with signal

- Determination of δ_{sys} on SB is affected by the presence of signal
  - →Obtain larger δ_{sys} with signal than without
  - →Dampening of significances
- To mitigate this we use the whole sideband with the statistics present in the SR
  - →Dilution of signal



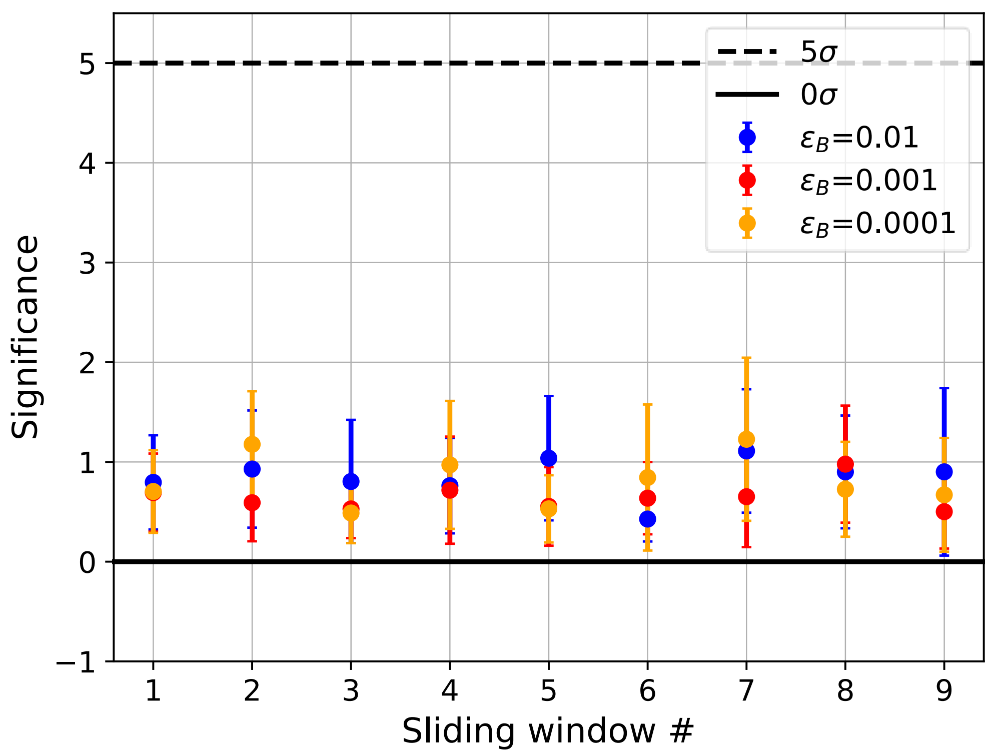
CATHODE with ΔR



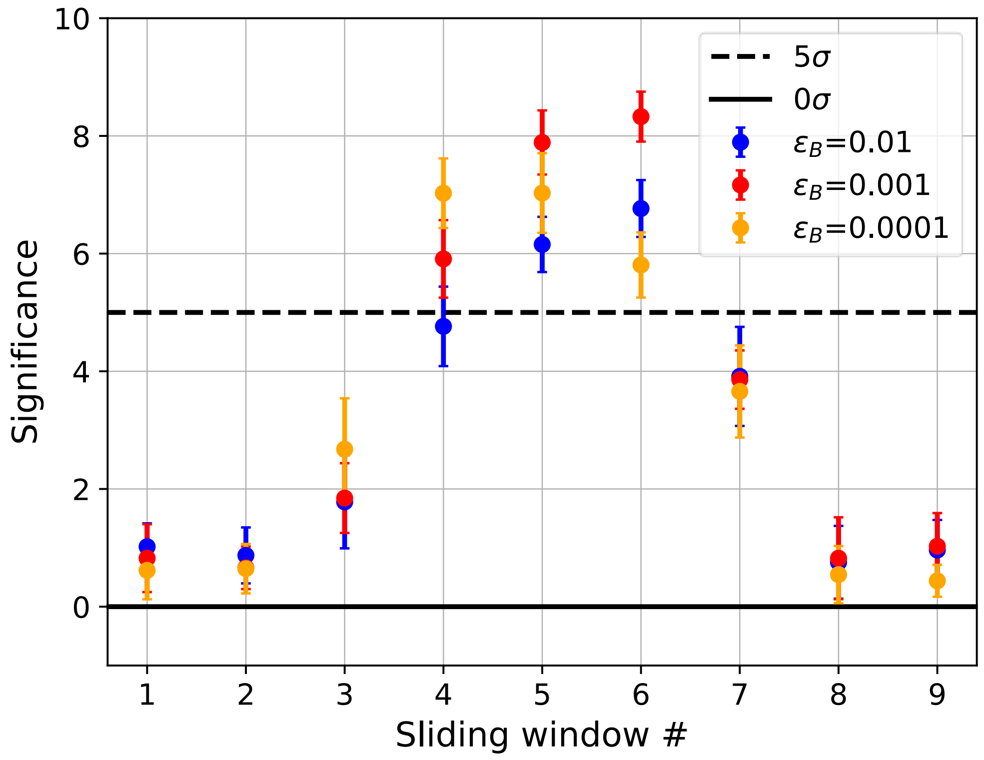


IAD Significances

Without Signal

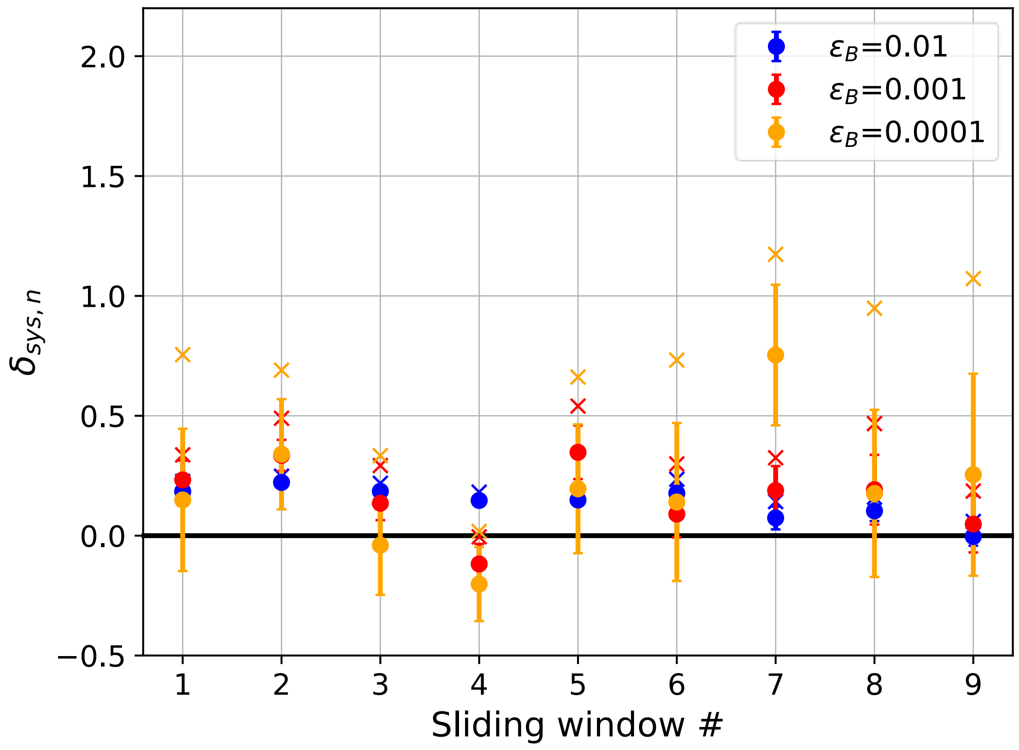


With Signal



δ_{sys} for different window numbers

CWoLa



CATHODE