Simulation-Based Detection for Composite Hypothesis Testing with Application to Time-Delay Signals

Ella Hadar[1] **Ilana Nisky**[2] **Tirza Routtenberg**[1]

[1]Electrical and Computer Engineering, Ben-Gurion University of the Negev, Israel

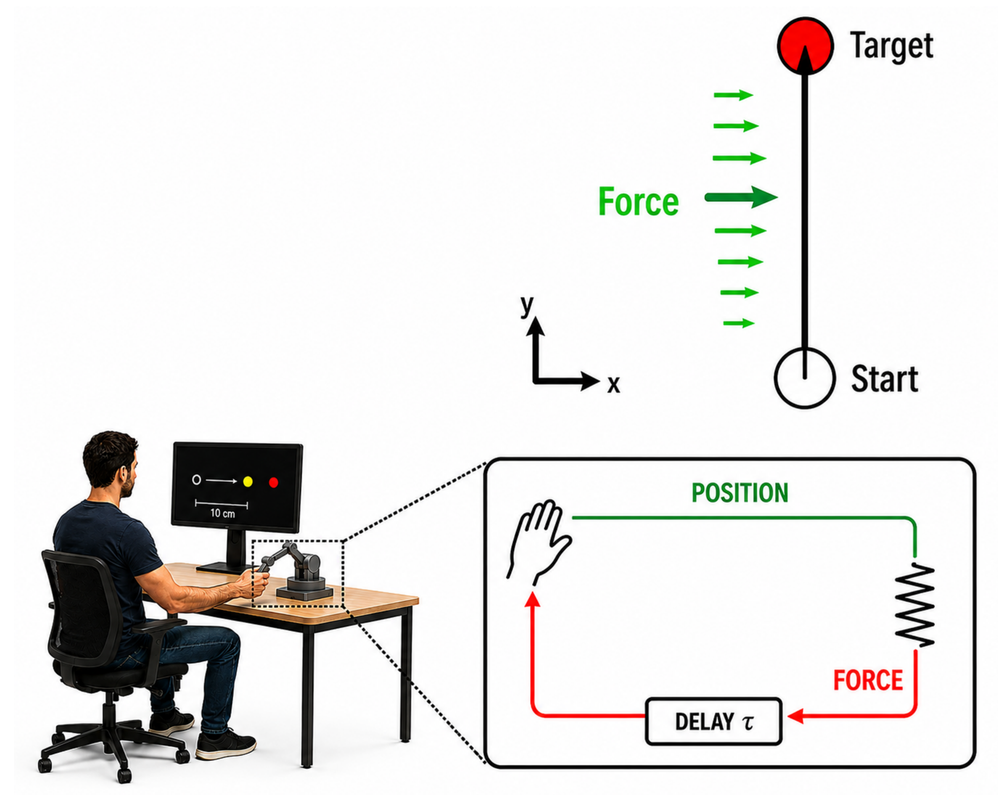
[2]Biomedical Engineering, Ben-Gurion University of the Negev, Israel

1. Motivation

- Many detection problems involve composite hypotheses with unknown deterministic nuisance parameters.
- In finite-sample, low-SNR regimes, estimating the parameters is unstable and computationally demanding.
- In human-in-the-loop robotic systems, human response delays directly affect closed-loop stability and safety [1].
- Selecting the model that best captures human behavior is essential for adaptive control design $\Rightarrow$ a composite hypothesis testing problem with unknown delays.
- Classical tests may be computationally challenging and less reliable in finite-sample, low-SNR regimes.

Main Idea

Replace test-time likelihood optimization with offline simulation-trained detection.



2. Composite Hypothesis Testing

We formulate the problem as a binary composite hypothesis testing problem:

$$\begin{cases} H_1 : x \sim f_1(x;\theta_1), \\ H_2 : x \sim f_2(x;\theta_2), \end{cases}$$

The parameters θ_1 and θ_2 are deterministic nuisance parameters: they are fixed but unknown at test time.

Goal

Given observation block x, decide whether the data is generated under H_1 or H_2.

3. Likelihood-Based Tests

- **Oracle Likelihood Ratio Test (Oracle LRT)**

$$\Lambda_{\mathrm{LRT}}(x) = \log \frac{f_2(x;\theta_2)}{f_1(x;\theta_1)} \underset{H_1}{\overset{H_2}{\gtrless}} \gamma.$$

$\Rightarrow$ Uses the true (unknown) parameters, i.e. unattainable. Serves as a performance benchmark.

- **Generalized Likelihood Ratio Test (GLRT)**

$$\Lambda_{\mathrm{GLRT}}(x) = \log \frac{\max_{\theta_2} f_2(x;\theta_2)}{\max_{\theta_1} f_1(x;\theta_1)} \underset{H_1}{\overset{H_2}{\gtrless}} \gamma.$$

$\Rightarrow$ Replaces unknown parameters by maximum likelihood estimators $\Rightarrow$ computationally expensive orintractable.

- **Bayesian LRT (Bayes LRT)**

$$\Lambda_{\mathrm{Bayes}}(x) = \log \frac{\int f_2(x;\theta_2)\pi_2(\theta_2)d\theta_2}{\int f_1(x;\theta_1)\pi_1(\theta_1)d\theta_1} \underset{H_1}{\overset{H_2}{\gtrless}} \gamma.$$

$\Rightarrow$ Marginalizes over chosen design distributions, mixture-based Bayesian benchmark, but the setting is non-Bayesian!

Challenge

GLRT requires maximization over unknown parameters, which can be computationally expensive and unstable in finite-sample, low-SNR regimes [2]. Bayesian LRT depends on chosen design priors and may require numerical integration. These limitations motivate an offline simulation-based alternative.

4. Simulation-Based Inference

- Simulation-Based Inference (SBI) enables hypothesis testing when likelihood-based methods are computationally expensive.
- Likelihood-Free Inference (LFI) learns inference procedures directly from simulated data without explicit likelihood evaluation.

Key takeaway

Motivated by SBI and LFI [3,4], we leverage simulated data to learn a decision rule directly in composite hypothesis testing, rather than estimating the nuisance parameters or approximating the posterior.

5. Proposed Method: SBD

Under each hypothesis H_j, $j = 1,2$, we sample nuisance parameters from a design distribution and generate labeled training data:

$$\theta_j^{(\ell)} \sim \pi_j(\theta_j), \qquad y^{(\ell)} \sim f_j(y;\theta_j^{(\ell)}), \qquad j \in \{1,2\}.$$

This yields a labeled training set:

$$\mathcal{T} = \left\{ y_n^{(\ell)},\ c^{(\ell)} \right\}_{\ell=1}^{2L}, \qquad c^{(\ell)} \in \{1,2\}.$$

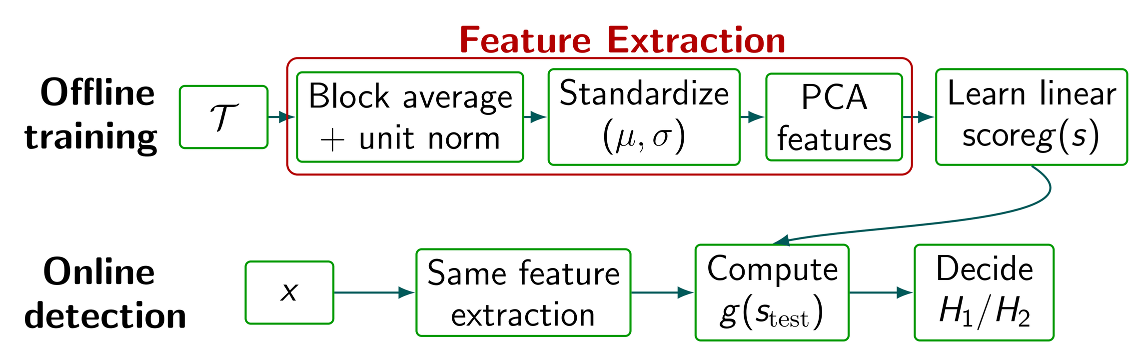
Each simulated observation block is mapped into low-dimensional features used for detector learning.

Key Idea

SBD learns the decision rule $g(\cdot)$ from simulated labeled blocks, avoiding online likelihood maximization.

6. Feature Pipeline & Decision Rule

Using the labeled training set $\mathcal{T}$, each block is converted into stable low-dimensional features used to learn the decision rule.



Feature Rationale

Normalization stabilizes the features, while PCA reduces dimensionality and classifier complexity.

Decision Rule

For the test block x, compute the test statistic $g(s_{\mathrm{test}})$, i.e., a linear decision score, and apply the threshold rule:

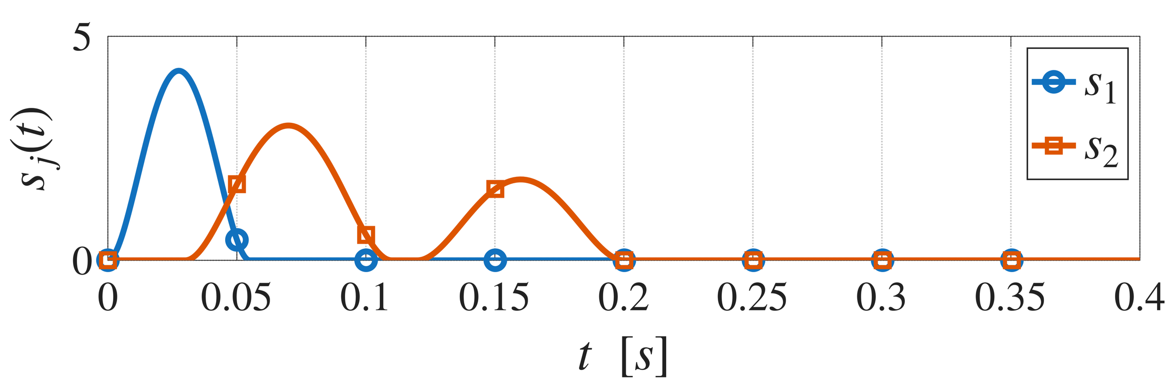
$$g(s_{\mathrm{test}}) \underset{H_1}{\overset{H_2}{\gtrless}} \gamma_{\mathrm{SBD}}.$$

7. Time-Delay Signal Application

We consider binary detection of delayed waveform structures:

$$H_j: \quad x = s_j(\tau_j) + w, \qquad w \sim \mathcal{N}(0,\eta^2 I), \qquad j \in \{1,2\},$$

- τ_j is an unknown deterministic delay.
- H_1 is a single-pulse minimum-jerk profile
- H_2 is a two-pulse profile



8. Experimental Setup

We evaluate the proposed SBD on the time-delay signal detection problem and compare its performance with GLRT, Bayes LRT, and Oracle LRT. The SNR is defined with respect to the clean waveform energy.

Each test block contains:

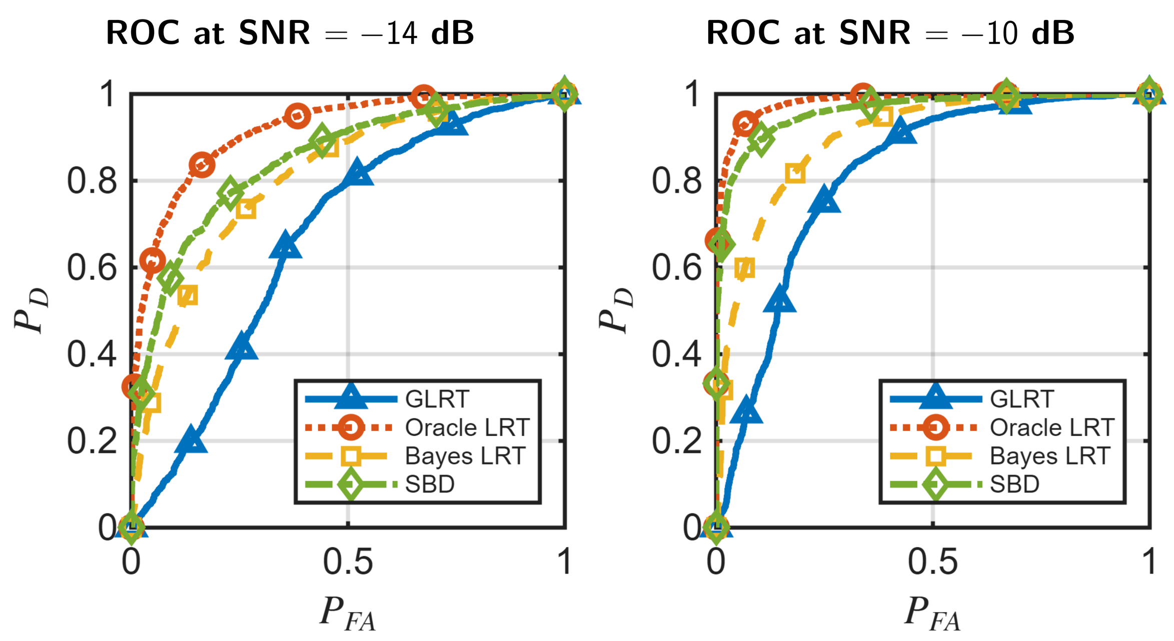
- Each block contains $N = 6$ repeated noisy observations.
- $L = 10^4$ simulated labeled blocks per hypothesis used for offline training.
- True delay parameter: $\tau = 0.7\,\mathrm{ms}$.

Evaluation objective

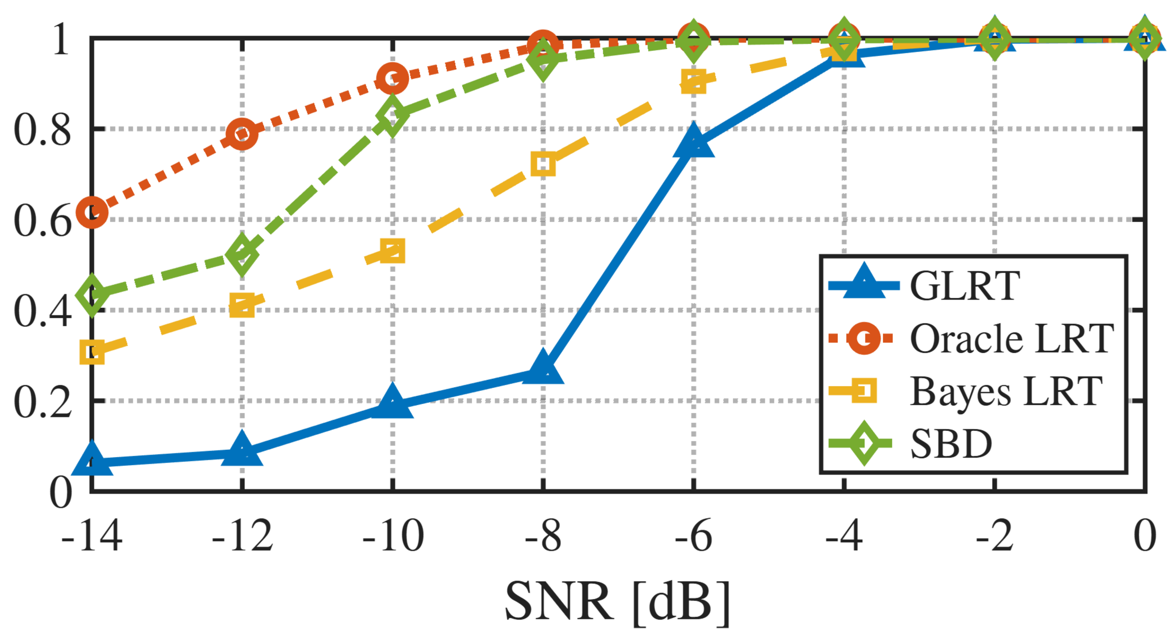
Assess whether the proposed SBD can outperform classical likelihood-based detectors and approach Oracle performance while avoiding online likelihood maximization and Bayesian marginalization.

9. Results

ROC Performance in Low-SNR Regimes



Detection Probability Across SNR at $P_{FA} = 0.05$



Main findings

- SBD achieves higher detection performance than GLRT and Bayes LRT in challenging low-SNR regimes.
- As SNR increases, SBD approaches the oracle-LRT benchmark.

10. Main Takeaways & Future Work

Main Takeaways

- Offline simulations replace online likelihood optimization.
- SBD provides a likelihood-free alternative.
- Offline storage is traded for lower real-time complexity.

Future Work

- Theoretical analysis of SBD.
- Robustness to model misspecification.
- Adaptive robotic control.

References

[1] G. Avraham, F. Mawase, A. Karniel, L. Shmuelof, O. Donchin, F. A. Mussa-Ivaldi, and I. Nisky, "Representing delayed force feedback as a combination of current and delayed states," *Journal of Neurophysiology*, vol. 118, no. 5, pp. 2677–2687, 2017.

[2] O. Zeitouni, J. Ziv, and N. Merhav, "When is the generalized likelihood ratio test optimal?" *IEEE Transactions on Information Theory*, vol. 38, no. 5, pp. 1597–1602, 1992.

[3] K. Cranmer, J. Brehmer, and G. Louppe, "The frontier of simulation-based inference," *Proceedings of the National Academy of Sciences*, vol. 117, no. 48, pp. 30 055–30 062, 2020.

[4] K. Cranmer, J. Pavez, and G. Louppe, "Approximating likelihood ratios with calibrated discriminative classifiers," *arXiv preprint arXiv:1506.02169*, 2015.

hadarell@post.bgu.ac.il