

# Generating Hidden Anomalies by Statistical Graph Signal Processing with Applications to Cyber Security in Electrical Networks



COLUMBIA  
UNIVERSITY

Gal Morgenstern, Jip Kim, James Anderson, Gil Zussman, and Tirza Routtenberg

G. Morgenstern and T. Routtenberg are with the Ben-Gurion University of the Negev, Israel. J. Kim is with KENTECH, South Korea.

J. Anderson and G. Zussman are with Columbia University, New York. Email: galmo@post.bgu.ac.il.

## Introduction

Let  $\mathcal{G}(\mathcal{V}, \mathcal{E})$  be a connected, undirected graph with  $N$  nodes. The graph Laplacian matrix  $\mathbf{L}$  captures the interactions between the graph nodes with entries defined by

$$L_{k,l} = \begin{cases} \sum_{m \in \mathcal{N}_k} \omega_{k,m} & k = l \\ -\omega_{k,l} & (k,l) \in \mathcal{E} \\ 0 & \text{otherwise,} \end{cases} \quad (1)$$

where  $\mathcal{N}_k$  is the set of nodes connected to node  $k$ . Denote the eigenvalue decomposition of  $\mathbf{L}$  as

$$\mathbf{L} = \mathbf{U} \text{diag}(\lambda_1, \dots, \lambda_n) \mathbf{U}^T$$

where the eigenvalues are ordered such that

$$0 = \lambda_1 < \lambda_2 \leq \dots \leq \lambda_n.$$

We define graph signals as the mapping  $\mathbf{x} : \mathcal{V} \rightarrow \mathbb{R}^N$  and consider graph filters of the

form:

$$f(\mathbf{L}) = \mathbf{U} f(\boldsymbol{\Lambda}) \mathbf{U}^T,$$

where  $f(\boldsymbol{\Lambda})$  defines the graph filter frequency response with diagonal entries given by  $f(\lambda_i)$ . The graph Fourier transform of the graph signal  $\mathbf{x}$  is

$$\tilde{\mathbf{x}} \triangleq \mathbf{U}^T \mathbf{x}.$$

Consider the hypothesis testing

$$\begin{cases} \mathcal{H}_0 : \mathbf{z} = \mathbf{x} + \text{noise} \\ \mathcal{H}_1 : \mathbf{z} = \mathbf{x} + \mathbf{c} + \text{noise,} \end{cases}$$

where  $\mathbf{z} \in \mathbb{R}^N$  represents the measured graph signal and  $\mathbf{c}$  represents the hidden anomaly. We can then solve the problem of finding anomalies over the graph using the test statistic

$$T^f(\mathbf{z}) = \|f(\mathbf{L})\mathbf{z}\|^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \gamma,$$

where  $f(\mathbf{L})$  is a graph filter. By setting the frequency response to

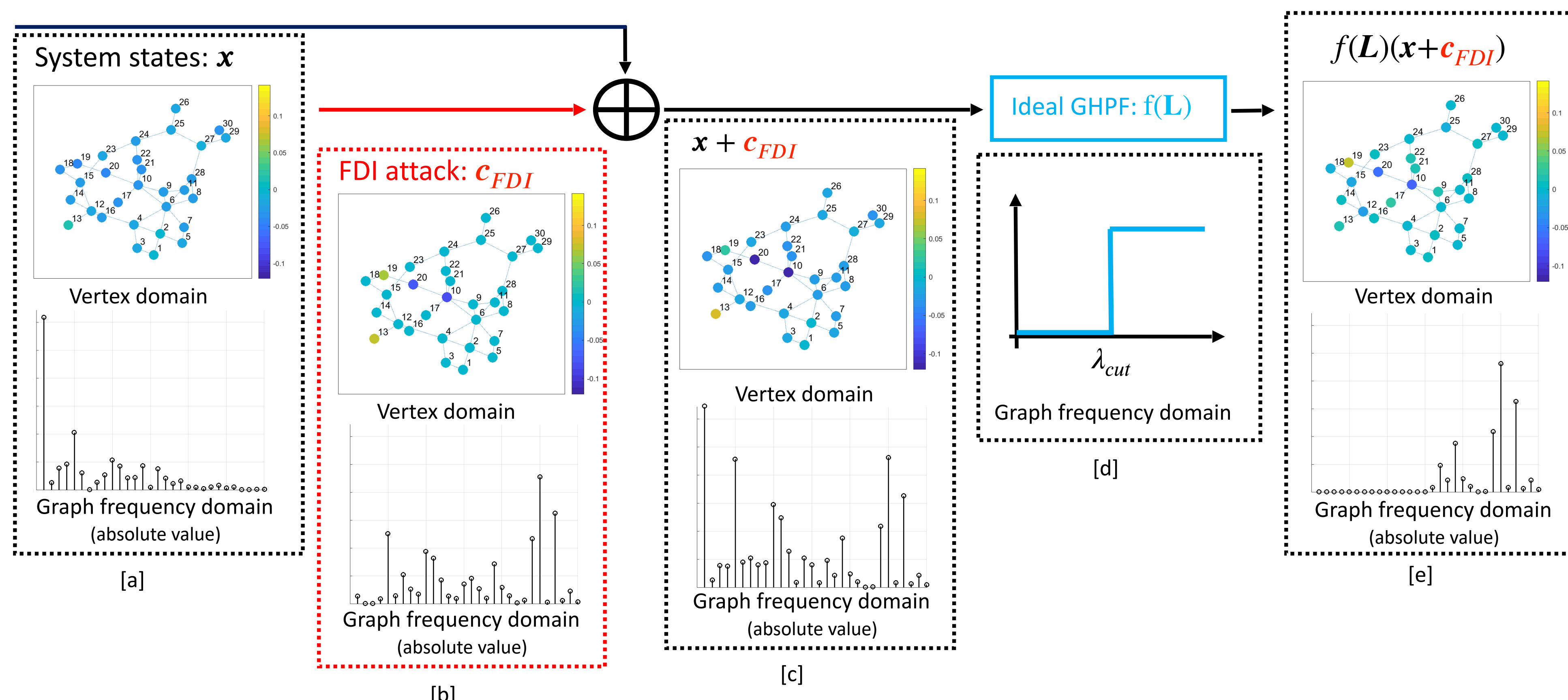
$$f^{TV}(\lambda_i) = \sqrt{\lambda_i}, \quad i = 1, \dots, N,$$

we obtain that the test statistic can be rewritten as:

$$T^f(\mathbf{z}) = \mathbf{z}^T \mathbf{L} \mathbf{z}$$

which is the well known Graph Tikhonov regularization. However, although graph signal processing (GSP)-based detectors can effectively detect ordinary anomalies, they may not be sufficient for detecting smooth graph signals with a low graph TV. To address this issue, we present a new type of attack, called graph false data injection (GFDI), and propose a defense mechanism against it.

## GSP-based detection: Schematic Diagram



## Strategic protection

Our protection scheme identifies a minimal set of state variables ( $\mathcal{D}$ ), such that if secured, it would disable the possibility of generating a GFDI attack.

The proposed protection scheme is

$$\hat{\mathcal{D}} = \arg \min_{\mathcal{D} \subseteq \mathcal{V}} |\mathcal{D}| \text{ s.t. } (\hat{\mathbf{c}})^T \mathbf{L} \hat{\mathbf{c}} > \delta.$$

The number of possible  $\mathcal{D}$ 's grows exponentially with the system size. Therefore, we propose a low-complexity greedy algorithm:

### Algorithm 1: GSP-based Protection

**Input** :  $\mathbf{L}, k, \tau, \delta$

**Output**:  $\mathcal{D}$

- 1  $\mathcal{D} \leftarrow \emptyset$ ;
- 2 **repeat**
- 3     Derive  $\mathcal{S}$  from  $\mathcal{D}$  by including power injections in the buses in  $\mathcal{D}$  and power flows entering the same buses;
- 4     Get  $\hat{\mathbf{i}}$  and  $\hat{\mathbf{c}}$  from GFDI solution;
- 5     Add  $\hat{\mathbf{i}}$  to  $\mathcal{D}$ ;
- 6 **until**  $(\hat{\mathbf{c}})^T \mathbf{L} \hat{\mathbf{c}} > \delta$ ;
- 7 **return**  $\mathcal{D}$ ;

## GFDI attacks (Hidden anomalies)

The GFDI attack is a smooth graph signal characterized by a low graph (TV), defined as  $TV^G(\mathbf{c}) = \mathbf{c}^T \mathbf{L} \mathbf{c}$ , under practical constraints:

1. achieving a certain impact:  $\|\mathbf{c}\|_\infty \geq \tau$
2. being sparse:  $|\mathbf{c}|_0 \leq k$
3. not having access to a subset of protected measurements  $\mathcal{S}$ , i.e.  $\mathbf{H}^S \mathbf{c} = \mathbf{0}$

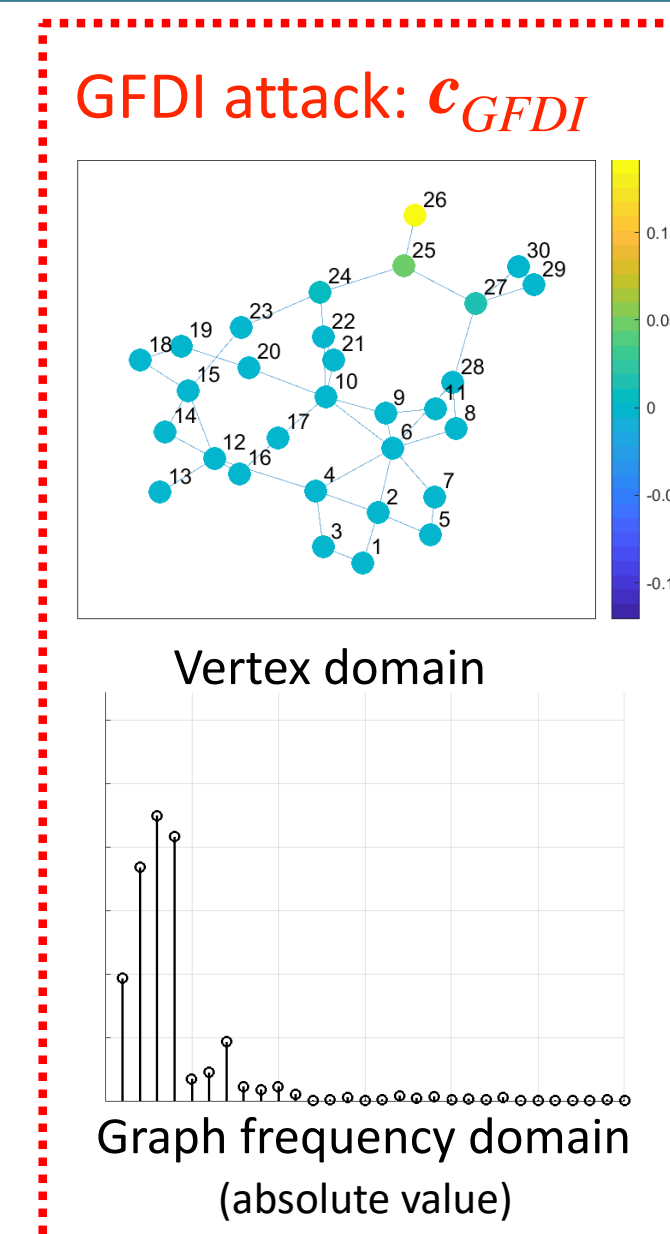
The resulting optimization problem is

$$\begin{aligned} \min_{\mathbf{c} \in \mathbb{R}^N} \quad & \mathbf{c}^T \mathbf{L} \mathbf{c} \\ \text{s.t.} \quad & \begin{cases} \|\mathbf{c}\|_\infty \geq \tau \\ \|\mathbf{c}\|_0 \leq k \\ \mathbf{H}^S \mathbf{c} = \mathbf{0}. \end{cases} \end{aligned}$$

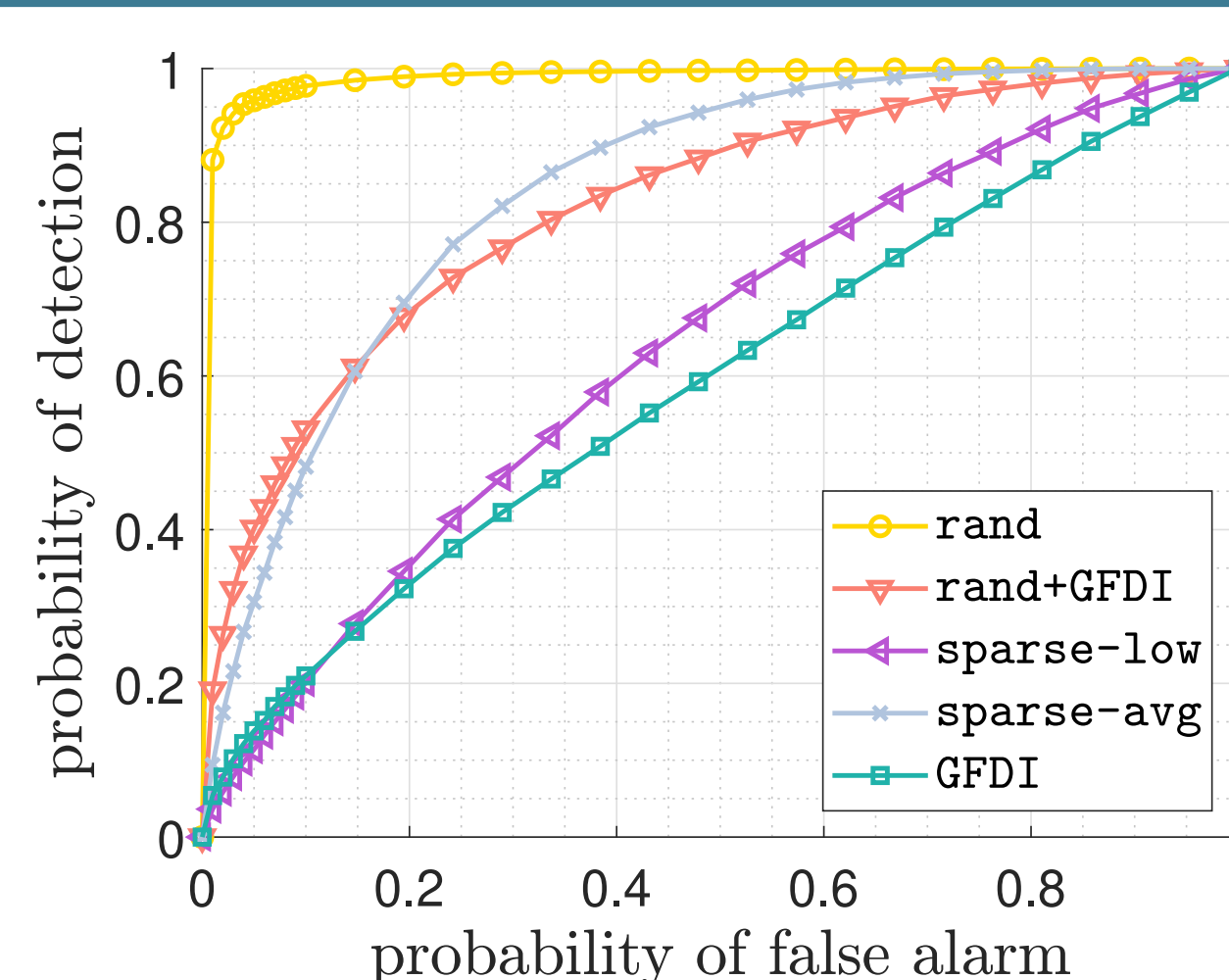
By using the  $\ell_1$  convex relaxation, the problem can be solved by solving  $N$  convex optimization problems where for case  $i$   $\|\mathbf{c}\|_\infty \geq \tau$  is replaced by  $c_i = \tau$ . The convex optimization problems are solved by using the projected gradient descent method and the alternating method of multipliers.

## GFDI attack: Schematic Diagram

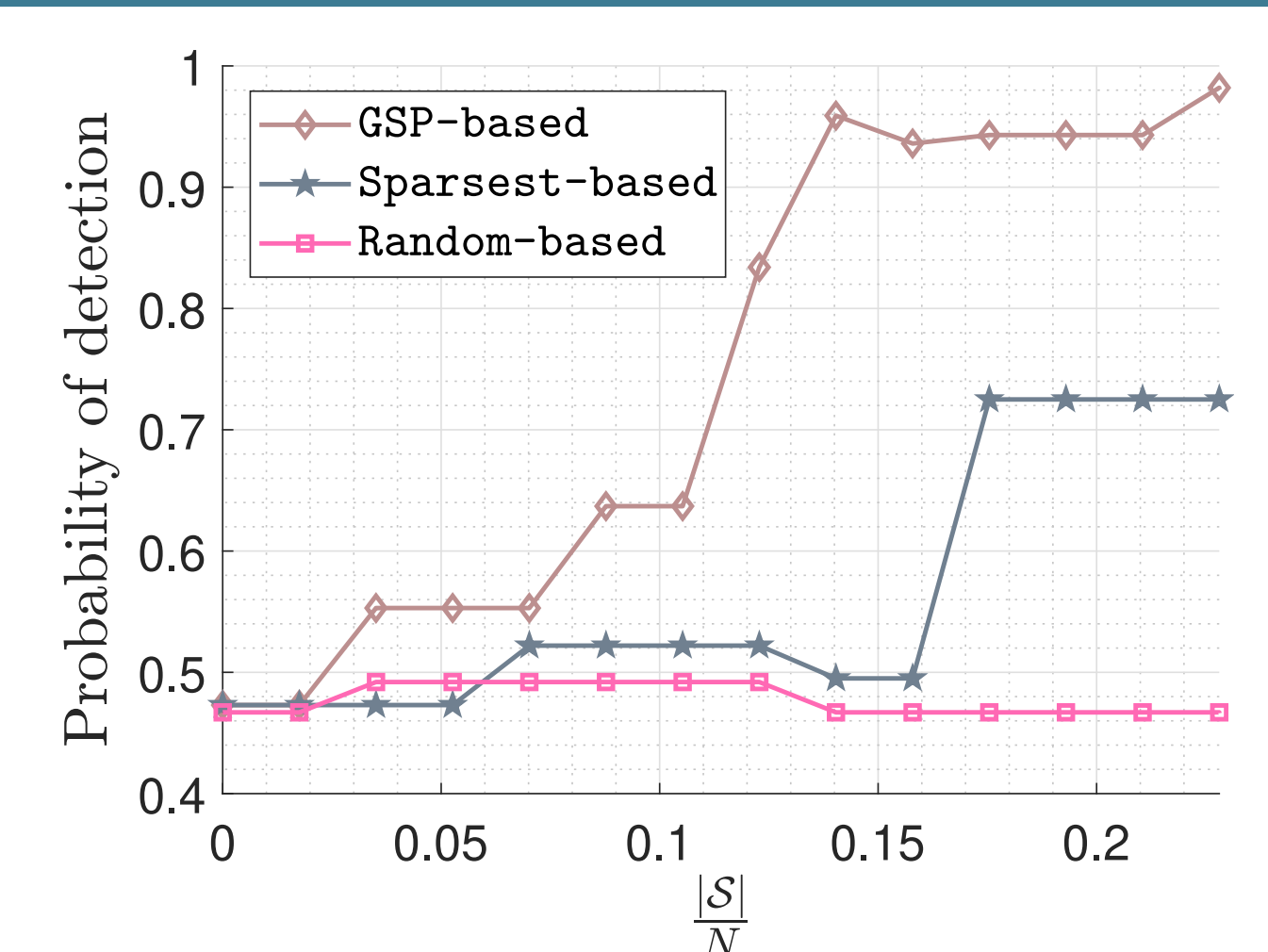
The GFDI attack is located at the lower graph frequencies. Therefore, in contrast to the FDI attack in [b], when added to the system states, the output will not obtain abnormal energy.



## GFDI attack: Simulations



## Strategic protection: Simulations



## Acknowledgements

This work was supported in part by the Next Generation Internet (NGI) program, the Jabotinsky Scholarship from the Israel Ministry of Technology and Science, and the Israel Ministry of National Infrastructure, Energy.