# LEARNING REQUIREMENTS FOR STEALTH ATTACKS

*Ke Sun**, Iñaki Esnaola**†, Antonia M. Tulino*§‡, H. Vincent Poor*†

* Dept. of Automatic Control and Systems Engineering, University of Sheffield, Sheffield S1 3JD, UK
§ Nokia Bell Labs, Holmdel, NJ 07733, USA
†Dept. of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA
‡University degli Studi di Napoli Federico II, 80138 Naples, Italy

## ABSTRACT

The learning data requirements are analyzed for the construction of stealth attacks in state estimation. In particular, the training data set is used to compute a sample covariance matrix that results in a random matrix with a Wishart distribution. The ergodic attack performance is defined as the average attack performance obtained by taking the expectation with respect to the distribution of the training data set. The impact of the training data size on the ergodic attack performance is characterized by proposing an upper bound for the performance. Simulations on the IEEE 30-Bus test system show that the proposed bound is tight in practical settings.

***Index Terms***— stealth attacks, data injection attacks, random matrix theory, information theory

## 1. INTRODUCTION

Data injection attacks [1] are one of the main threats that the smart grid faces. Attack constructions that exploit the sparsity of the data injection vector have been proposed [2] as practical constructions that can disrupt the state estimation performed by the operator. Distributed attack construction and detection strategies are studied in [3, 4, 5, 6] where it is shown that the bad data detection procedures put in place by the operator can be defeated by several attackers that control a subset of the sensing infrastructure in the grid. Modelling the state variables as a random process, attack constructions that exploit the statistical knowledge of the state variables are proposed in [7, 8]. The addition of probabilistic structure to the state variables opens the door to the definition of information theoretic attacks for which the damage and probability of detection are characterized in terms of information measures [9]. In [10] the assumption of perfect knowledge of the statistics of the state variables is relaxed by considering a training data set to learn the statistics. Therein, it is numerically shown that the performance of the attack when imperfect knowledge of the statistics is available changes significantly with respect to the case with perfect knowledge. In this paper, we analytically characterize the impact of the training data size and the correlation between state variables over the attack performance.

Ke Sun acknowledges the support of China Scholarship Council (CSC) and the support from the Department of Automatic Control and Systems Engineering for travelling. H.Vincent Poor was supported in part by the U.S. National Science Foundation under Grants DMS-1736417 and ECCS-1824710.

## 2. SYSTEM MODEL

### 2.1. State Estimation and Bad Data Detection

The measurement model for state estimation with linearized dynamics is given by

$$Y^M = \mathbf{H}X^N + Z^M, \tag{1}$$

where $Y^M \in \mathbb{R}^M$ is a vector of random variables describing the measurements; $X^N \in \mathbb{R}^N$ is a vector of random variables describing the state variables; $\mathbf{H} \in \mathbb{R}^{M\times N}$ is the linearized Jacobian measurement matrix which is determined by the power network topology and the admittances of the branches; and $Z^M \in \mathbb{R}^M$ is the additive white Gaussian noise (AWGN) with distribution $\mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_M)$ where $\sigma^2$ is the variance of the error introduced by the sensors [11], [12, Chapter 15]. The vector of the state variables is assumed to follow a multivariate Gaussian distribution given by $X^N \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{XX})$, where $\mathbf{\Sigma}_{XX}$ is the positive-definite covariance matrix of the distribution of the state variables. The Gaussian assumption for the vector of the state variables is also adopted by [7] and [8]. As a result of the linear measurement model in (1), the vector of measurements also follows a multivariate Gaussian distribution denoted by $Y^M \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{YY})$, where $\mathbf{\Sigma}_{YY} = \mathbf{H}\mathbf{\Sigma}_{XX}\mathbf{H}^{\mathrm{T}} + \sigma^2\mathbf{I}_M$.

Data injection attacks corrupt the measurements available to the operator by adding an attack vector to the measurements. The resulting vector of compromised measurements is given by

$$Y_A^M = \mathbf{H}X^N + Z^M + A^M, \tag{2}$$

where $A^M \in \mathbb{R}^M$ is the attack vector and $Y_A^M \in \mathbb{R}^M$ is the vector containing the compromised measurements

[1]. Following the approach in [9] we adopt a Gaussian framework for the construction of the attack vector, i.e. $A^M \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{AA})$, where $\mathbf{\Sigma}_{AA}$ is the covariance matrix of the attack distribution. The rationale for choosing a Gaussian distribution for the attack vector follows from the fact that for the attack model in (2) the additive attack distribution that minimizes the mutual information between the vector of state variables and the compromised measurements is Gaussian [13]. Because of the Gaussianity of the attack distribution, the vector of compromised measurements is distributed as $Y_A^M \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{Y_AY_A})$, where $\mathbf{\Sigma}_{Y_AY_A} = \mathbf{H}\mathbf{\Sigma}_{XX}\mathbf{H}^{\mathrm{T}} + \sigma^2\mathbf{I}_M + \mathbf{\Sigma}_{AA}$.

The operator of the power system makes use of the acquired measurements to detect the attack. The detection problem is cast as a hypothesis testing problem with hypotheses

$$\mathcal{H}_0 : Y^M \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{YY}), \quad \text{versus} \tag{3}$$

$$\mathcal{H}_1 : Y^M \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{Y_AY_A}). \tag{4}$$

The null hypothesis $\mathcal{H}_0$ describes the case in which the power system is not compromised, while the alternative hypothesis $\mathcal{H}_1$ describes the case in which the power system is under attack. The Neyman-Pearson lemma [14] states that for a fixed probability of Type I error, the likelihood ratio test (LRT) achieves the minimum Type II error when compared with any other test with an equal or smaller Type I error. Consequently, the LRT is chosen to decide between $\mathcal{H}_0$ and $\mathcal{H}_1$ based on the available measurements. The LRT between $\mathcal{H}_0$ and $\mathcal{H}_1$ takes following form:

$$L(\mathbf{y}) \triangleq \frac{f_{Y_A^M}(\mathbf{y})}{f_{Y^M}(\mathbf{y})} \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \tau, \tag{5}$$

where $\mathbf{y} \in \mathbb{R}^M$ is a realization of the vector of random variables modelling the measurements, $f_{Y_A^M}$ and $f_{Y^M}$ denote the probability density functions (p.d.f.'s) of $Y_A^M$ and $Y^M$, respectively, and $\tau$ is the decision threshold set by the operator to meet the false alarm constraint.

### 2.2. Information-Theoretic Attacks

The purpose of the attacker is to disrupt the normal state estimation procedure by minimizing the information that the operator acquires about the state variables, while guaranteeing that the probability of attack detection is small enough, and therefore, remain concealed in the system. To that end, the attacker aims to minimize the mutual information between the vector of state variables and the vector of compromised measurements denoted by $I(X^N; Y_A^M)$. On the other hand, we assess the performance of attack detection by the LRT via the Chernoff-Stein lemma [15], which characterizes the asymptotic exponent of the probability of detection when the number of observations of measurement vectors grows to infinity. In our setting, the Chernoff-Stein lemma states that for any LRT and $\epsilon \in (0, 1/2)$, it holds that

$$\lim_{n\to\infty} \frac{1}{n} \log \beta_n^\epsilon = -D(P_{Y_A^M} \| P_{Y^M}), \tag{6}$$

where $D(\cdot\|\cdot)$ is the Kullback-Leibler (KL) divergence, $\beta_n^\epsilon$ is the minimum Type II error such that the Type I error $\alpha$ satisfies $\alpha < \epsilon$, and $n$ is the number of $M$-dimensional measurement vectors that are available for the LRT. Therefore, for the attacker, minimizing the asymptotic detection probability is equivalent to minimizing $D(P_{Y_A^M} \| P_{Y^M})$, where $P_{Y_A^M}$ and $P_{Y^M}$ denote the probability distributions of $Y_A^M$ and $Y^M$, respectively.

A stealthy attack construction that combines these two information measures in one cost function is proposed in [10]. Interestingly, the resulting cost function boils down to the effective secrecy proposed in [16] which can be written as

$$I(X^N; Y_A^M) + D(P_{Y_A^M} \| P_{Y^M}) = D(P_{X^N Y_A^M} \| P_{X^N} P_{Y^M}), \tag{7}$$

where $P_{X^N Y_A^M}$ is the joint distribution of $X^N$ and $Y_A^M$. The resulting attack construction problem is equivalent to solving the following optimization problem:

$$\min_{A^M} D(P_{X^N Y_A^M} \| P_{X^N} P_{Y^M}). \tag{8}$$

Under the attack Gaussianity assumption the cost function in (7) is a function of the attack covariance matrix $\mathbf{\Sigma}_{AA}$. Let us define the cost function for the Gaussian case as

$$f(\mathbf{\Sigma}_{AA}) \overset{\Delta}{=} \frac{1}{2}\left[\mathrm{tr}(\mathbf{\Sigma}_{YY}^{-1}\mathbf{\Sigma}_{AA}) - \log|\mathbf{\Sigma}_{AA} + \sigma^2\mathbf{I}_M| - \log|\mathbf{\Sigma}_{YY}^{-1}|\right]. \tag{9}$$

It is shown in [9] that (8) is a convex optimization problem and that the covariance matrix for the optimal Gaussian attack is $\mathbf{\Sigma}_{AA}^\star \triangleq \mathbf{H}\mathbf{\Sigma}_{XX}\mathbf{H}^{\mathrm{T}}$.

## 3. LEARNING ATTACK CONSTRUCTION

The stealth attack construction proposed above requires perfect knowledge of the covariance matrix of the state variables and the linearized Jacobian measurement matrix. In the following we study the performance of the attack when the second order statistics are not perfectly known by the attacker but the linearized Jacobian measurement matrix is known. We model the partial knowledge by assuming that the attacker has access to a sample covariance matrix of the state variables. Specifically, the training data consisting of $K$ state variable realizations $\{X_i^N\}_{i=1}^K$ is available to the attacker. That being the case the attacker computes the unbiased estimate of the covariance matrix of the state variables given by

$$\mathbf{S}_{XX} = \frac{1}{K-1} \sum_{i=1}^{K} X_i^N (X_i^N)^{\mathrm{T}}. \tag{10}$$

The stealth attack constructed using the sample covariance matrix follows a multivariate Gaussian distribution given by

$$\tilde{A}^M \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{\tilde{A}\tilde{A}}), \tag{11}$$

where $\mathbf{\Sigma}_{\tilde{A}\tilde{A}} = \mathbf{H}\mathbf{S}_{XX}\mathbf{H}^{\mathrm{T}}$.

Since the sample covariance matrix in (10) is a random matrix with central Wishart distribution given by

$$\mathbf{S}_{XX} \sim \frac{1}{K-1} W_N(K-1, \mathbf{\Sigma}_{XX}), \tag{12}$$

the ergodic counterpart of the cost function in (7) is defined in terms of the conditional KL divergence given by

$$\mathbb{E}_{\mathbf{S}_{XX}}\left[D\left(P_{X^N Y_{\tilde{A}}^M|\mathbf{S}_{XX}} \| P_{X^N} P_{Y^M}\right)\right]. \tag{13}$$

The ergodic cost function characterizes the expected performance of the attack averaged over the realizations of training data. Note that the performance using the sample covariance matrix is suboptimal [10] and that the ergodic performance converges asymptotically to the optimal attack construction when the size of the training data set increases.

## 4. UPPER BOUND FOR ERGODIC ATTACK PERFORMANCE

In this section, we analytically characterize the ergodic attack performance defined in (13) by providing an upper bound using random matrix theory tools. Before introducing the upper bound, some auxiliary results on the expected value of the extreme eigenvalues of Wishart random matrices are presented below.

### 4.1. Auxiliary Results in Random Matrix Theory

**Lemma 1.** *Let $\mathbf{Z}_L$ be an $(K-1)\times L$ matrix whose entries are independent standard normal random variables, then*

$$\mathrm{var}\left(s_{max}(\mathbf{Z}_L)\right) \leq 1, \tag{14}$$

*where* $\mathrm{var}\left(\cdot\right)$ *denotes the variance and $s_{max}(\mathbf{Z}_L)$ is the maximum singular value of $\mathbf{Z}_L$.*

*Proof.* Note that $s_{max}(\mathbf{Z}_L)$ is a 1-Lipschitz function of matrix $\mathbf{Z}_L$, the maximum singular value of $\mathbf{Z}_L$ is concentrated around the mean [17, Proposition 5.34] given by $\mathbb{E}[s_{max}(\mathbf{Z}_L)]$. Then for $t \geq 0$, it holds that

$$\mathbb{P}[|s_{max}(\mathbf{Z}_L) - \mathbb{E}[s_{max}(\mathbf{Z}_L)]| > t] \leq 2\exp\{-t^2/2\} \tag{15}$$

$$\leq \exp\{1 - t^2/2\}. \tag{16}$$

Therefore $s_{max}(\mathbf{Z}_L)$ is a sub-gaussian random variable with variance proxy $\sigma_p^2 \leq 1$. The lemma follows from the fact that $\mathrm{var}\left(s_{max}(\mathbf{Z}_L)\right) \leq \sigma_p^2$. □

**Lemma 2.** *Let $\mathbf{W}_L$ denote a central Wishart matrix distributed as $\frac{1}{K-1}W_L(K-1, \mathbf{I}_L)$, then the non-asymptotic expected value of the extreme eigenvalues of $\mathbf{W}_L$ is bounded by*

$$\left(1 - \sqrt{L/(K-1)}\right)^2 \leq \mathbb{E}[\lambda_{min}(\mathbf{W}_L)] \tag{17}$$

*and*

$$\mathbb{E}[\lambda_{max}(\mathbf{W}_L)] \leq \left(1 + \sqrt{L/(K-1)}\right)^2 + 1/(K-1), \tag{18}$$

*where $\lambda_{min}(\mathbf{W}_L)$ and $\lambda_{max}(\mathbf{W}_L)$ denote the minimum eigenvalue and maximum eigenvalue of $\mathbf{W}_L$, respectively.*

*Proof.* Note that [17, Theorem 5.32]

$$\sqrt{K-1} - \sqrt{L} \leq \mathbb{E}[s_{min}(\mathbf{Z}_L)] \tag{19}$$

and

$$\sqrt{K-1} + \sqrt{L} \geq \mathbb{E}[s_{max}(\mathbf{Z}_L)], \tag{20}$$

where $s_{min}(\mathbf{Z}_L)$ is the minimum singular value of $\mathbf{Z}_L$. Given the fact that $\mathbf{W}_L = \frac{1}{K-1}\mathbf{Z}_L^{\mathrm{T}}\mathbf{Z}_L$, then it holds that

$$\mathbb{E}[\lambda_{min}(\mathbf{W}_L)] = \frac{\mathbb{E}\left[s_{min}(\mathbf{Z}_L)^2\right]}{K-1} \geq \frac{\mathbb{E}\left[s_{min}(\mathbf{Z}_L)\right]^2}{K-1} \tag{21}$$

and

$$\mathbb{E}[\lambda_{max}(\mathbf{W}_L)] = \frac{\mathbb{E}\left[s_{max}(\mathbf{Z}_L)^2\right]}{K-1} \leq \frac{\mathbb{E}[s_{max}(\mathbf{Z}_L)]^2 + 1}{K-1}, \tag{22}$$

where (22) follows from Lemma 1. Combining (19) with (21), and (20) with (22), respectively, yields the lemma. □

### 4.2. Main Result

The ergodic attack performance is given by

$$\begin{aligned}
&\mathbb{E}\left[f(\mathbf{\Sigma}_{\tilde{A}\tilde{A}})\right] \\
&= \frac{1}{2}\mathbb{E}\left[\mathrm{tr}(\mathbf{\Sigma}_{YY}^{-1}\mathbf{\Sigma}_{\tilde{A}\tilde{A}}) - \log|\mathbf{\Sigma}_{\tilde{A}\tilde{A}} + \sigma^2\mathbf{I}_M| - \log|\mathbf{\Sigma}_{YY}^{-1}|\right] \\
&= \frac{1}{2}\Big(\mathrm{tr}\big(\mathbf{\Sigma}_{YY}^{-1}\mathbf{\Sigma}_{AA}^{\star}\big) - \log\big|\mathbf{\Sigma}_{YY}^{-1}\big| - \mathbb{E}\big[\log|\mathbf{\Sigma}_{\tilde{A}\tilde{A}} + \sigma^2\mathbf{I}_M|\big]\Big).
\end{aligned} \tag{23}$$

The assessment of the ergodic attack performance boils down to evaluating the last term in (23). Closed form expressions for this term are provided in [18] for the same case considered in this paper. However, the resulting expressions are involved and are only computable for small dimensional settings. For systems with a large number of dimensions the expressions are computationally prohibitive. To circumvent this challenge we propose a lower bound on the term that yields an upper bound on the ergodic attack performance. Before presenting the main result we provide the following auxiliary convex optimization result.

**Lemma 3.** *Let $\mathbf{B} = \mathrm{diag}(b_1, \ldots, b_p)$ denote a positive definite diagonal matrix. Then*

$$\mathbb{E}\left[\log\left|\mathbf{B} + \mathbf{W}_p^{-1}\right|\right] \geq \sum_{i=1}^{p} \log\left(b_i + 1/x_i^{\star}\right), \tag{24}$$

*where $x_i^{\star}$ is the solution to the convex optimization problem given by*

$$\min_{\{x_i\}_{i=1}^p} \sum_{i=1}^{p} \log\left(b_i + 1/x_i\right) \tag{25}$$

$$s.t. \quad \sum_{i=1}^{p} x_i = p \tag{26}$$

$$\max\left(x_i\right) \leq \left(1 + \sqrt{p/(K-1)}\right)^2 + 1/(K-1) \tag{27}$$

$$\min\left(x_i\right) \geq \left(1 - \sqrt{p/(K-1)}\right)^2. \tag{28}$$

*Proof.* Note that

$$\mathbb{E}\left[\log\left|\mathbf{B}+\mathbf{W}_p^{-1}\right|\right] = \sum_{i=1}^{p}\mathbb{E}\left[\log\left(b_i+\frac{1}{\lambda_i(\mathbf{W}_p)}\right)\right] \quad (29)$$

$$\geq \sum_{i=1}^{p}\log\left(b_i+\frac{1}{\mathbb{E}[\lambda_i(\mathbf{W}_p)]}\right) \quad (30)$$

where in (29), $\lambda_i(\mathbf{W}_p)$ is the $i$-th eigenvalue of $\mathbf{W}_p$ in decreasing order; (30) follows from Jensen's inequality due to the convexity of $\log\left(b_i+\frac{1}{x}\right)$ for $x>0$. Constraint (26) follows from the fact that $\mathbb{E}[\text{trace}(\mathbf{W}_p)] = p$, and constraints (27) and (28) follow from Lemma 2. This completes the proof. □

The following theorem provides a lower bound for the last term in (23), and therefore, it enables us to characterize the ergodic attack performance.

**Theorem 1.** *Let $\mathbf{\Sigma}_{\tilde{A}\tilde{A}} = \mathbf{H}\mathbf{S}_{XX}\mathbf{H}^T$ and denote by $\mathbf{\Lambda}_p = \text{diag}(\lambda_1,\ldots,\lambda_p)$ the diagonal matrix containing the nonzero eigenvalues in decreasing order. Then*

$$\mathbb{E}\left[\log|\mathbf{\Sigma}_{\tilde{A}\tilde{A}}+\sigma^2\mathbf{I}_M|\right] \geq \left(\sum_{i=0}^{p-1}\psi(K-1-i)\right) - p\log(K-1) + \sum_{i=1}^{p}\log\left(\frac{\lambda_i}{\sigma^2}+\frac{1}{\lambda_i^\star}\right) + 2M\log\sigma, \quad (31)$$

*where $\psi(\cdot)$ is the Euler digamma function, $p = \text{rank}(\mathbf{H}\mathbf{\Sigma}_{XX}\mathbf{H}^T)$, and $\{\lambda_i^\star\}_{i=1}^p$ is the solution to the optimization problem given by (25) - (28) with $b_i = \frac{\lambda_i}{\sigma^2}$, for $i=1,\ldots,p$.*

*Proof.* We proceed by noticing that

$$\mathbb{E}\left[\log|\mathbf{\Sigma}_{\tilde{A}\tilde{A}}+\sigma^2\mathbf{I}_M|\right]$$

$$= \mathbb{E}\left[\log\left|\frac{1}{(K-1)\sigma^2}\mathbf{Z}_M^{\mathrm{T}}\mathbf{\Lambda}\mathbf{Z}_M+\mathbf{I}_M\right|\right] + 2M\log\sigma \quad (32)$$

$$= \mathbb{E}\left[\log\left|\frac{\mathbf{\Lambda}_p}{\sigma^2}\frac{\mathbf{Z}_p^{\mathrm{T}}\mathbf{Z}_p}{K-1}+\mathbf{I}_M\right|\right] + 2M\log\sigma \quad (33)$$

$$= \mathbb{E}\left[\log\left|\frac{\mathbf{Z}_p^{\mathrm{T}}\mathbf{Z}_p}{K-1}\right| + \log\left|\frac{\mathbf{\Lambda}_p}{\sigma^2}+\left(\frac{\mathbf{Z}_p^{\mathrm{T}}\mathbf{Z}_p}{K-1}\right)^{-1}\right|\right] + 2M\log\sigma \quad (34)$$

$$\geq \left(\sum_{i=0}^{p-1}\psi(K-1-i)\right) - p\log(K-1) + \sum_{i=1}^{p}\log\left(\frac{\lambda_i}{\sigma^2}+\frac{1}{\lambda_i^\star}\right) + 2M\log\sigma, \quad (35)$$

where in (32), $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues of $\mathbf{H}\mathbf{\Sigma}_{XX}\mathbf{H}^{\mathrm{T}}$ in decreasing order; (33) follows from the fact that $p = \text{rank}(\mathbf{H}\mathbf{\Sigma}_{XX}\mathbf{H}^{\mathrm{T}})$; (35) follows from substituting (2.12) in [19] and Lemma 3 into (34). This completes the proof. □

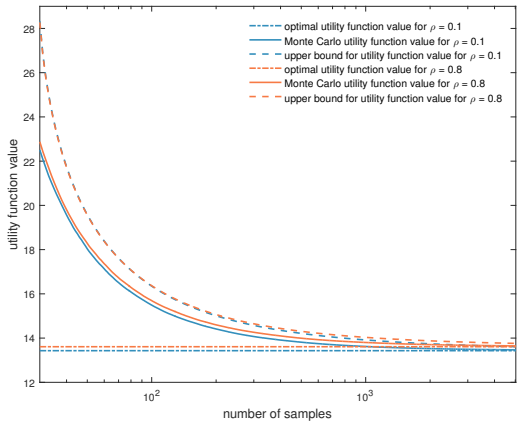


**Fig. 1**. Performance of the upper bound in Theorem 2 as a function of number of sample for $\rho = 0.1$ and $\rho = 0.8$ when SNR $= 20$ dB.

**Theorem 2.** *The ergodic attack performance given in (23) is upper bounded by*

$$\mathbb{E}\left[f(\mathbf{\Sigma}_{\tilde{A}\tilde{A}})\right] \leq \frac{1}{2}\Bigg(tr\left(\mathbf{\Sigma}_{YY}^{-1}\mathbf{\Sigma}_{AA}^{\star}\right) - \log\left|\mathbf{\Sigma}_{YY}^{-1}\right| - 2M\log\sigma - \left(\sum_{i=0}^{p-1}\psi(K-1-i)\right) + p\log(K-1) - \sum_{i=1}^{p}\log\left(\frac{\lambda_i}{\sigma^2}+\frac{1}{\lambda_i^\star}\right)\Bigg). \quad (36)$$

*Proof.* The proof follows immediately from combing Theorem 1 with (23). □

## 5. NUMERICAL RESULTS

The numerical results are obtained on the IEEE 30-Bus test system where the Jacobian matrix $\mathbf{H}$ is obtained using MATPOWER [20]. For the construction of the stealth attack the covariance matrix of the state variables is chosen to be a Toeplitz matrix with exponential decay parameter $\rho$ as in [8]. Specifically, the Toeplitz matrix of dimension $N \times N$ with exponential decay parameter $\rho$ is given by $\mathbf{\Sigma}_{XX} = [s_{ij} = \rho^{|i-j|}; i,j = 1,2,\ldots,N]$. We define the Signal-to-Noise Ratio (SNR) as

$$\text{SNR} = 10\log_{10}\left(\frac{\text{tr}(\mathbf{H}\mathbf{\Sigma}_{XX}\mathbf{H}^{\mathrm{T}})}{M\sigma^2}\right). \quad (37)$$

Fig.1 depicts the upper bound in Theorem 2 as a function of number of samples for $\rho = 0.1$ and $\rho = 0.8$ when SNR $=$ 20 dB. Interestingly, the upper bound in Theorem 2 is tight for large values of the training data set size for all values of the exponential decay parameter determining the correlation.

## 6. REFERENCES


[1] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proc. ACM Conf. on Computer and Communications Security*, Chicago, IL, USA, Nov. 2009, pp. 21–32.

[2] T. T. Kim and H. V. Poor, "Strategic protection against data injection attacks on power grids," *IEEE Trans. Smart Grid*, vol. 2, no. 2, pp. 326–333, Jun. 2011.

[3] S. Cui, Z. Han, S. Kar, T. T. Kim, H. V. Poor, and A. Tajer, "Coordinated data-injection attack and detection in the smart grid: A detailed look at enriching detection solutions," *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 106–115, Sep. 2012.

[4] M. Ozay, I. Esnaola, F. T. Y. Vural, S. R. Kulkarni, and H. V. Poor, "Sparse attack construction and state estimation in the smart grid: Centralized and distributed models," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1306–1318, Jul. 2013.

[5] U. A. Khan and A. M. Stanković, "Secure distributed estimation in cyber-physical systems," in *Proc. IEEE Int. Conf. on Acoust., Speech and Signal Process.*, Vancouver, Canada, May 2013, pp. 5209–5213.

[6] A. Tajer, S. Kar, H. V. Poor, and S. Cui, "Distributed joint cyber attack detection and state recovery in smart grids," in *Proc. IEEE Int. Conf. on Smart Grid Commun.*, Brussels, Belgium, Oct. 2011, pp. 202–207.

[7] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 645–658, Dec. 2011.

[8] I. Esnaola, S. M. Perlaza, H. V. Poor, and O. Kosut, "Maximum distortion attacks in electricity grids," *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 2007–2015, Jul. 2016.

[9] K. Sun, I. Esnaola, S.M. Perlaza, and H.V. Poor, "Stealth attacks on the smart grid," *arXiv preprint arXiv:1808.04184*, 2018.

[10] K. Sun, I. Esnaola, S.M. Perlaza, and H.V. Poor, "Information-theoretic attacks in the smart grid," in *Proc. IEEE Int. Conf. on Smart Grid Commum.*, Dresden, Germany, Oct. 2017, pp. 455–460.

[11] A. Abur and A. G. Expósito, *Power System State Estimation: Theory and Implementation*, CRC Press, Mar. 2004.

[12] J. J. Grainger and W. D. Stevenson, *Power System Analysis*, McGraw-Hill, 1994.

[13] I. Shomorony and A. S. Avestimehr, "Worst-case additive noise in wireless networks," *IEEE Trans. Inf. Theory*, vol. 59, no. 6, pp. 3833–3847, Jun. 2013.

[14] H. V. Poor, *An Introduction to Signal Detection and Estimation*, Springer, New York, 1994.

[15] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Nov. 2012.

[16] J. Hou and G. Kramer, "Effective secrecy: Reliability, confusion and stealth," in *Proc. IEEE Int. Symp. on Information Theory*, Honolulu, HI, USA, Jun. 2014, pp. 601–605.

[17] R. Vershynin, "Introduction to the non-asymptotic analysis of random matrices," in *Compressed Sensing: Theory and Applications*, Y. Eldar and G. Kutyniok, Eds., chapter 5, pp. 210–268. Cambridge University Press, Cambridge, UK, 2012.

[18] G. Alfano, A. M. Tulino, A. Lozano, and S. Verdú, "Capacity of MIMO channels with one-sided correlation," in *Proc. IEEE Int. Symp. on Spread Spectrum Techniques and Applications*, Sydney, Australia, Aug 2004.

[19] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*, Now Publishers Inc, 2004.

[20] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.