Signal Selection for Oscillation Monitoring with Guarantees on Data Recovery under Corruption

Kaustav Chatterjee, *Student Member, IEEE,* Nilanjan Ray Chaudhuri, *Senior Member, IEEE* and George Stefopoulos, *Senior Member, IEEE*

***Abstract*—Insights are developed into grouping PMU signals for guaranteeing data recovery under sparse corruption. Analytical relations are derived to express the denseness of the subspace spanned by a measurement window in terms of the modal observabilities of its constituent signals. It is shown that grouping signals by minimizing variation in phase angles and amplitudes of observabilities for each poorly-damped mode minimizes the numerical-rank of the measurement window, enhances denseness of the subspace, and helps in attaining the sufficiency condition guaranteeing exact recovery using Robust Principal Component Analysis-based signal reconstruction methods. These insights are structured into lemmas and propositions for signal selection and are validated on synthetic data from IEEE test systems, as well as field PMU data from a US utility.**

***Index Terms*—Commpressive Sensing, Robust PCA, Oscillation Monitoring, PMU Signal Recovery, Denseness, RIC.**

I. Introduction

MONITORING low-frequency oscillations is important in ensuring stability and security of large power networks. If not adequately damped, these oscillations can grow and eventually cause large-scale outages – as observed in the North American blackout of August 1996. Ever since, major utilities around the world have invested extensively in building their networks of Phasor Measurement Units (PMUs). This has enabled monitoring of the oscillatory modes owing to time-synchronization and higher reporting rates of these devices [1].

However, due to variation in modal energy among the measurements, the locations of the PMUs (and the choice of signals from installed PMUs) are critical in monitoring these oscillations. Several methods have been proposed in literature, which use relative modal observabilities [2] and participation ratios [3] obtained from the linearized system model in deciding the signals of interest. There also exist measurement-driven approaches that derive metrics from the power spectral density (PSD) of a time-series data for the selection of candidate signals [4]. The objective of these approaches is to choose the signals that are rich in information of the critical modes.

The accuracy of oscillation monitoring algorithms however, is contingent upon the control center receiving error-free PMU measurements. Missing data samples, spurious outliers, and malicious injections from cyber attacks can lead to erroneous estimates of damping ratios, modal frequencies, and mode-shapes. Ideally, bad data detection and correction can be done using power system state estimators in control centers, but this requires full observability of the system with PMUs and an accurate knowledge of system topology – a review of literature in this area can be found in [5]–[9]. However, at present most systems have a limited number of PMUs, which may not ensure full observability of the system.

On the other hand, several measurement-driven data-preprocessing approaches have been proposed in [10]–[17] that neither require full observability with PMUs, nor need the network topology information. These works exploit the spatio-temporal correlation in measurements to solve low-rank matrix completion problems for recovering clean measurements from corrupted PMU data. However, as one understands, this requires careful selection of signals to ensure correlation and low-rankness of the measurement window. In this context, we ask the following research question – *can we intelligently locate the PMUs and/or group signals from installed PMUs such that the resulting combinations guarantee data recovery from corruption while capturing every information necessary to estimate the critical modes.*

We argue that a group of signals best suited for estimation of modes might not be the ideal choice for corruption-resilient signal recovery. Implying that, a set of signals selected purely based on the maximum relative modal observability criterion might fail in the accurate recovery of the individual signals corrupted by anomalies. This can affect the overall accuracy of the modal estimation. To solve this, we make propositions about grouping candidate signals to increase denseness of the subspace spanned by the measurement window, thereby, enhancing the chances of exact data recovery under corruption.

We make the following contributions– first, we show that the denseness of a subspace derived from a measurement window can be bounded by denseness of the observability submatrix obtained from the small-signal model corresponding to the poorly-damped modes in the system – under the weak assumption that other modes are sufficiently damped in the window and/or sufficiently unobservable. Second, using these deductions, we draw insights into grouping of signals to enhance the denseness of the observability submatrix to meet the sufficiency condition for guaranteed exact data recovery [18]. Further, we quantify the perturbation in the denseness values for small variations in the magnitudes of modal observabilities within a signal group. And finally, we extend these insights onto recommendations for grouping signals directly from

K. Chatterjee and N. R. Chaudhuri are with The School of Electrical Engineering and Computer Science, The Pennsylvania State University, University Park, PA 16802, USA (e-mail: kuc760@psu.edu; nuc88@psu.edu)

G. Stefopoulos is with New York Power Authority, White Plains, NY 10601 USA (e-mail: George. Stefopoulos@nypa.gov).

Financial support from the NSF grant under award CNS 1739206 is gratefully acknowledged.

PMU data in absence of a small-signal model. We demonstrate the applicability of our proposed method on synthetic data obtained from IEEE 4-machine and 16-machine systems, as well as field PMU measurements obtained from New York Power Authority.

Notations: Matrices are denoted by bold uppercase, vectors by bold lowercase, and scalars by normal font. Superscript H denotes Hermitian of a matrix. For any matrix $\mathbf{Y}$, range($\mathbf{Y}$) denotes the subspace spanned by its columns, and basis($\mathbf{Y}$) denotes the orthonormal basis matrix whose columns span the same subspace as range($\mathbf{Y}$).

II. Preliminaries

A. Small-signal Model and Modal Observability

The linearized state-space model of a power system can be described as,

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) \tag{1}$$

where, $\mathbf{x}(t) \in \Re^{m\times 1}$, $\mathbf{y}(t) \in \Re^{n\times 1}$, and $\mathbf{u}(t) \in \Re^{p\times 1}$ are respectively the vectors of state, output, and input variables capturing the perturbations from their respective equilibria.

Assuming $\mathbf{A}$ is diagonalizable, consider the transformation $\mathbf{P}^{-1}\mathbf{x}(t) = \tilde{\mathbf{x}}(t)$ where, $\mathbf{P}$ is the matrix of right eigenvectors of $\mathbf{A}$. The equations in (1) can then be re-written as,

$$\begin{aligned} \dot{\tilde{\mathbf{x}}}(t) &= \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\tilde{\mathbf{x}}(t) + \mathbf{P}^{-1}\mathbf{B}\mathbf{u}(t) = \mathbf{\Lambda}\tilde{\mathbf{x}}(t) + \tilde{\mathbf{B}}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{P}\tilde{\mathbf{x}}(t) = \mathbf{\Psi}\tilde{\mathbf{x}}(t) \end{aligned} \tag{2}$$

where, $\mathbf{\Lambda} \in \Re^{m\times m}$ is the diagonal matrix of the eigenvalues of $\mathbf{A}$. Henceforth, in this paper, we shall refer to each complex-conjugate eigenvalue pair of $\mathbf{\Lambda}$ as a mode, and $\mathbf{\Psi} \in \Re^{n\times m}$ as the matrix of relative modal observabilities [19] mapping the extent to which each mode is visible in the output variables measured by the PMUs.

B. Corruption Model

Owing to the chances of corruption − missing data, spurious outliers, and malicious injections mentioned before, let the measurements observed at the control center be different from $\mathbf{y}(t)$. We model this as,

$$\mathbf{z}(t) = \mathbf{y}(t) + \mathbf{e}(t) \tag{3}$$

where, $\mathbf{e}(t)$ is the vector of additive signal corruptions. In this paper, we consider that the corruptions are limited to only a fraction of all signals. This is a realistic assumption for wide-area monitoring applications, since PMUs are spread across wide geographies and thus, coordinated corruption in a large number of channels is improbable. Therefore, $\mathbf{e}(t)$ is a sparse vector with most entries as zeros. In this context, note that a vector is called $s-$sparse if it has at most s non-zero entries.

C. Robust Signal Recovery

The objective of the data pre-processor is to recover the actual measurements $\mathbf{y}(t)$ from the corrupted observations $\mathbf{z}(t)$. To that end, we shall use robust principal component analysis (R-PCA)-based anomaly correction approach from [15]–[17]. Building on the theory of compressive sensing [20]–[22], the algorithms in [15]–[17] solve sparse optimization problems to estimate the signal corruption $\hat{\mathbf{e}}(t)$, which can then be subtracted from $\mathbf{z}(t)$ to recover $\hat{\mathbf{y}}(t)$. Henceforth, we shall refer to $\hat{\mathbf{y}}(t)$ as the recovered signal vector at time t. The recovery framework is summarised below.

Let $\mathbf{Y} \in \Re^{n\times N}$ be the matrix representing the time-window of recovered samples till the latest instant.

$$\mathbf{Y} = \begin{bmatrix} \hat{\mathbf{y}}(t - N\tau) & \dots & \hat{\mathbf{y}}(t-\tau) \end{bmatrix} \tag{4}$$

where, τ is the reporting interval between two successive PMU samples. In a large interconnected system, the matrix $\mathbf{Y}$ is low-rank, and thus, can be spanned by a few basis vectors. Let these bases be obtained from singular value decomposition of the window. Hence, $\mathbf{Y} \approx \hat{\mathbf{U}}\hat{\mathbf{\Sigma}}\hat{\mathbf{V}}^H$, where columns of $\hat{\mathbf{U}}$ are the singular vectors corresponding to r dominant singular values.

The basic idea behind the recovery is to decompose $\mathbf{z}(t)$ into vectors $\hat{\mathbf{y}}(t)$ and $\hat{e}(t)$, such that $\hat{\mathbf{y}}(t)$ is in the low-rank subspace spanned by $\hat{\mathbf{U}}$ and $\hat{e}(t)$ is a sparse vector of estimated signal corruptions. To do so, $\mathbf{z}(t)$ is projected onto the space orthogonal to the span of $\hat{\mathbf{U}}$ as shown,

$$\boldsymbol{\gamma}(t) = \mathbf{\Phi}\mathbf{z}(t) = \mathbf{\Phi}(\mathbf{y}(t) + \mathbf{e}(t)) = \mathbf{\Phi}\mathbf{e}(t) + \boldsymbol{\nu}(t) \tag{5}$$

where, $\mathbf{\Phi} = \mathbf{I} - \hat{\mathbf{U}}\hat{\mathbf{U}}^H$. Ideally, $\mathbf{y}(t)$ belongs to the span of $\hat{\mathbf{U}}$, and therefore, the projection ensures that the contribution of $\mathbf{y}(t)$ is nullified. However, because of noise in the measurements and due to approximations in limiting the dimension of $\hat{\mathbf{U}}$ to r columns, the term $\mathbf{\Phi}\mathbf{y}(t)$ is not exactly zero, but negligibly small as captured in $\boldsymbol{\nu}(t)$. The estimation of the sparse corruption vector $\mathbf{e(t)}$ from $\mathbf{z}(t)$ can then be posed as an optimization problem [23] shown in (6) that ensures maximal sparsity of the corruption estimate.

$$\min_{\mathbf{e}(t)} \ \left\|\mathbf{e}(t)\right\|_0 \quad \text{s.t.} \ \left\|\boldsymbol{\gamma}(t) - \mathbf{\Phi}\mathbf{e}(t)\right\|_2 \leq \eta(t) \tag{6}$$

However, minimizing l_0-norm in eqn. (6) is non-convex and a hard combinatorial problem. Through convex relaxation, this is posed as an l_1-norm minimization problem [23].

$$\min_{\mathbf{e}(t)} \ \left\|\mathbf{e}(t)\right\|_1 \quad \text{s.t.} \ \left\|\boldsymbol{\gamma}(t) - \mathbf{\Phi}\mathbf{e}(t)\right\|_2 \leq \eta(t) \tag{7}$$

The solution $\hat{\mathbf{e}}(t)$ to eqn. (7), is an estimate of corruption. The clean signal can then be recovered as, $\hat{\mathbf{y}}(t) = \mathbf{z}(t) - \hat{\mathbf{e}}(t)$. The thresholding term $\eta(t)$ in (7) is updated in every time-step as, $\eta(t) = ||\mathbf{\Phi}\hat{\mathbf{y}}(t-\tau)||_2$. Initialization of the thresholding term can done by setting $\eta(0)$ to a small value estimated by projecting a sample from the archived measurements onto $\mathbf{\Phi}$. In our problem, assuming the first sample is uncorrupted, $\eta(t)$ is initialized by $||\mathbf{\Phi}y(0)||_2$ − as described in [16].

D. Guarantees for Exact Signal Recovery

Although the relaxation ensures a solution in polynomial time, it does not guarantee a sparse solution, implying that the recovery using eqn. (7) may not be accurate. It is only for some special structures of the projection matrix $\mathbf{\Phi}$ that the solutions of eqn. (7) and eqn. (6) coincide. We call this the $l_0 - l_1$ equivalence. This equivalence is achieved and the formulation in eqn. (7) is guaranteed to return a $s-$sparse solution if the $s-$restricted isometry constant $\delta_s(\mathbf{\Phi})$ is below a desired threshold [22].

Candés in [22] proved that for $\delta_{2s}(\mathbf{\Phi}) < \sqrt{2}-1$, the l_0 and l_1 problems are equivalent. However, this is conservative and several other upper bounds [18], [24], [25] have subsequently been derived, which have relaxed the requirement on $\delta_{2s}(\mathbf{\Phi})$. Notable of these, is the proposition from Cai et.al in [18] that guarantees $s-$sparse recovery for $\delta_s(\mathbf{\Phi}) < 0.307$. In this paper, we shall refer to this bound from [18] as the *sufficiency condition* for guaranteeing exact recovery of corrupted signals.

In [26], authors relate $\delta_s(\mathbf{\Phi})$ to the denseness of $\hat{\mathbf{U}}$. Denseness coefficient κ_s [26] for any matrix $\mathbf{Y}$ is defined as,

$$\kappa_s(\mathbf{Y}) = \kappa_s(\text{range}(\mathbf{Y})) = \max_{|T|\leq s} \left\| (\mathbf{I}_T)^H \, \text{basis}(\mathbf{Y}) \right\|_2 \tag{8}$$

where, $\mathbf{I}_T$ is the submatrix of the identity matrix $\mathbf{I}$ containing the columns with indices in set T.

Maximum value that κ_s can attain is 1. *Lower the value of κ_s, higher is the denseness of the range space.* For any $r-$rank $\mathbf{Y}$, the minimum value $\kappa_1(\mathbf{Y})$ can achieve is $\sqrt{r/n}$. The minimum value is attained when $\mathbf{Y}$ is spanned by basis vectors whose entries all have magnitude $\sqrt{1/n}$.

As derived in [26], for a basis matrix $\hat{\mathbf{U}}$,

$$\delta_s(\mathbf{\Phi}) = \delta_s(\mathbf{I} - \hat{\mathbf{U}}\hat{\mathbf{U}}^H) = \kappa_s(\hat{\mathbf{U}})^2 \tag{9}$$

Therefore, attaining the sufficiency condition,

$$\delta_s(\mathbf{\Phi}) < 0.307 \implies \kappa_s(\mathbf{Y}) = \kappa_s(\hat{\mathbf{U}}) < 0.554 = \kappa^* \tag{10}$$

Going forward, we ask how can one ensure that $\kappa_s(\hat{\mathbf{U}}) < \kappa^*$. Intuitively, different choices of signals (both the size of the set and the individual constituents) would alter the singular vectors spanning $\mathbf{Y}$, thereby changing the denseness coefficient. Thus motivated, next in this paper, we develop analytical insights on to grouping signals to minimize $\kappa_s(\mathbf{Y})$.

III. Insights into Signal Selection

Lemma 1. $\kappa_s(\mathbf{Y}) \leq \kappa_s(\mathbf{\Psi}) = \max_{|T|\leq s} \left\| (\mathbf{I}_T)^H \, \mathbf{\Psi}\mathbf{\Psi}^\dagger \right\|_2$

Proof. Referring to eqns (2) and (4),

$$\mathbf{Y} = \mathbf{\Psi} \begin{bmatrix} \tilde{\mathbf{x}}(t-N\tau) & \dots & \tilde{\mathbf{x}}(t-\tau) \end{bmatrix} \triangleq \mathbf{\Psi}\tilde{\mathbf{X}} \tag{11}$$

Therefore, $\text{range}(\mathbf{Y}) \subseteq \text{range}(\mathbf{\Psi}) \implies \kappa_s(\mathbf{Y}) \leq \kappa_s(\mathbf{\Psi})$. Equality is attained when $\tilde{\mathbf{X}}$ is full row rank.

Next, let $\hat{\mathbf{U}}_{\mathbf{\Psi}}$ be the matrix of singular vectors spanning the range of $\mathbf{\Psi}$. Therefore, following the definition of denseness,

$$\kappa_s(\mathbf{\Psi}) = \max_{|T|\leq s} \left\| (\mathbf{I}_T)^H \, \hat{\mathbf{U}}_{\mathbf{\Psi}} \right\|_2 = \max_{|T|\leq s} \left\| (\mathbf{I}_T)^H \, \hat{\mathbf{U}}_{\mathbf{\Psi}} \hat{\mathbf{U}}_{\mathbf{\Psi}}^H \right\|_2 \tag{12}$$

Moreover, the range-spaces of $\hat{\mathbf{U}}_{\mathbf{\Psi}}$ and $\mathbf{\Psi}$ being the same, the orthogonal projection matrices onto this space can be equated as follows: $\hat{\mathbf{U}}_{\mathbf{\Psi}}\hat{\mathbf{U}}_{\mathbf{\Psi}}^H = \mathbf{\Psi}\mathbf{\Psi}^\dagger$, where $\mathbf{\Psi}^\dagger = (\mathbf{\Psi}^H\mathbf{\Psi})^{-1}\mathbf{\Psi}^H$. Substituting this in eqn. (12) we get,

$$\kappa_s(\mathbf{\Psi}) = \max_{|T|\leq s} \left\| (\mathbf{I}_T)^H \, \hat{\mathbf{U}}_{\mathbf{\Psi}} \hat{\mathbf{U}}_{\mathbf{\Psi}}^H \right\|_2 = \max_{|T|\leq s} \left\| (\mathbf{I}_T)^H \, \mathbf{\Psi}\mathbf{\Psi}^\dagger \right\|_2 \quad \square$$

Corollary: Consider a data window $\mathbf{Y}$ exhibiting oscillatory response due to k poorly-damped modes. Under the weak assumption that the rest of the modes are sufficiently damped and/or sufficiently unobservable in the data window, we can write $\kappa_s(\mathbf{Y}) \leq \kappa_s(\hat{\mathbf{\Psi}})$ where, $\hat{\mathbf{\Psi}} = \begin{bmatrix} \hat{\mathbf{\Psi}}_\mathbf{1} & \hat{\mathbf{\Psi}}_\mathbf{2} & \cdots & \hat{\mathbf{\Psi}}_\mathbf{k} \end{bmatrix}$ is the submatrix of $\mathbf{\Psi}$ containing the complex conjugate column-pairs corresponding to the poorly-damped modes of interest, implying each $\hat{\mathbf{\Psi}}_j = \begin{bmatrix} \boldsymbol{\psi}_j & \overline{\boldsymbol{\psi}}_j \end{bmatrix}$ for $j = 1, 2, \dots, k$.

Remarks: Assuming that only k poorly-damped modes are observable in the signals, the data matrix may be approximated as $\mathbf{Y} \approx \hat{\mathbf{\Psi}}\tilde{\mathbf{X}}$. Therefore, it can be inferred that, the numerical-rank$(\mathbf{Y}) \leq$ rank$(\hat{\mathbf{\Psi}}) \leq 2k$.

A. Selection of Signals with Same Phase Relationship

As discussed, for a $r-$rank $\hat{\mathbf{\Psi}}$, the minimum value that $\kappa_1(\hat{\mathbf{\Psi}})$ can achieve is $\sqrt{r/n}$. Therefore, to achieve a low $\kappa_1(\hat{\mathbf{\Psi}})$ it is reasonable to select signals in a way that $\hat{\mathbf{\Psi}}$ has rank 1. To that end, following lemmas are presented.

Lemma 2. For a unimodal case, if the complex entries in $\boldsymbol{\psi}_1$ have same phase angle then, rank$(\hat{\mathbf{\Psi}})$ is 1.

Proof. Let $\boldsymbol{\psi}_1 = \begin{bmatrix} |\psi_{11}|\underline{/\theta} & \dots & |\psi_{1i}|\underline{/\theta} & \dots & |\psi_{1n}|\underline{/\theta} \end{bmatrix}^T$ and therefore, $\overline{\boldsymbol{\psi}}_1 = \begin{bmatrix} |\psi_{11}|\underline{/-\theta} & \dots & |\psi_{1i}|\underline{/-\theta} & \dots & |\psi_{1n}|\underline{/-\theta} \end{bmatrix}^T$.

Next, we perform Gram-Schmidt orthonormalization to find the bases spanning the columns of $\hat{\mathbf{\Psi}}_1$. Let the orthonormal bases be $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$, such that, $\hat{\mathbf{u}}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|_2} = \frac{\boldsymbol{\psi}_1}{\|\boldsymbol{\psi}_1\|_2}$, and $\hat{\mathbf{u}}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|_2}$ where, $\mathbf{u}_2 = \overline{\boldsymbol{\psi}}_1 - \frac{\langle \boldsymbol{\psi}_1, \overline{\boldsymbol{\psi}}_1 \rangle}{\langle \boldsymbol{\psi}_1, \boldsymbol{\psi}_1 \rangle}\boldsymbol{\psi}_1$. Since $\langle \boldsymbol{\psi}_1, \overline{\boldsymbol{\psi}}_1 \rangle = \|\boldsymbol{\psi}_1\|_2^2 \underline{/-2\theta}$, we can express any i^{th} entry of $\mathbf{u}_2$ as,

$$u_{2i} = |\psi_{1i}|\underline{/-\theta} - \frac{\|\boldsymbol{\psi}_1\|_2^2 \underline{/-2\theta}}{\|\boldsymbol{\psi}_1\|_2^2} |\psi_{1i}|\underline{/\theta} = 0 \tag{13}$$

This implies $\hat{\mathbf{u}}_2 = \mathbf{0}$, and therefore, rank$(\hat{\mathbf{\Psi}})$ is 1. $\square$

Corollary: The rank 1 property ensures that $\kappa_1(\hat{\mathbf{\Psi}}) = \kappa_1(\boldsymbol{\psi}_1)$. This simplifies the problem as the analysis now reduces to calculating the denseness of a column vector.

Lemma 3. For a unimodal case, the minimum value of $\kappa_1(\hat{\mathbf{\Psi}})$ is attained when signals are selected from a coherent group with minimum variance in the magnitudes of relative modal observabilities.

Proof. Signals selected from a single coherent group oscillate in unison, and hence, it can be inferred that their modal observabilities have same phase. Therefore, following Lemma 2, rank$(\hat{\mathbf{\Psi}})$ is 1.

Next, let $\{\mathbf{e}_1, \dots \mathbf{e}_i, \dots \mathbf{e}_n\}$ be the set of standard basis vectors. The denseness κ_1 can then be calculated as,

$$\begin{aligned} \kappa_1(\hat{\mathbf{\Psi}}) = \kappa_1(\boldsymbol{\psi}_1) = \max_i \left\| \mathbf{e}_i^H \, \hat{\mathbf{u}}_1 \right\|_2 &= \|\hat{\mathbf{u}}_1\|_\infty = \frac{\|\boldsymbol{\psi}_1\|_\infty}{\|\boldsymbol{\psi}_1\|_2} \\ &= \frac{|\psi_{1i}|_{max}}{\sqrt{\sum_{i=1}^n |\psi_{1i}|^2}} \geq \frac{1}{\sqrt{n}} \end{aligned} \tag{14}$$

The minimum value $\sqrt{1/n}$ is attained when $|\psi_{1i}|$-s are equal.

Since this is an idealistic scenario, next we investigate how denseness changes with variation in the magnitudes of modal observabilities, especially in presence of a large-magnitude outlier in the observability vector. For this, we divide the signal set into two groups– (1) the signal with index i_{max} having the highest magnitude of observability $|\psi_{1i}|_{max}$, and (2) the remaining $n-1$ signals with observability magnitudes $|\psi_{1i}|$-s dispersed with mean μ and standard deviation σ.

The mean and variance in the observability magnitudes for the group of $n-1$ signals can be computed as

$$\mu = \frac{1}{n-1} \sum_{i=1,\ i\neq i_{max}}^{n} |\psi_{1i}| \tag{15}$$

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1,\ i\neq i_{max}}^{n} (|\psi_{1i}| - \mu)^2$$

$$\implies (n-1)\sigma^2 = \sum_{i=1,\ i\neq i_{max}}^{n} |\psi_{1i}|^2 + \sum_{i=1,\ i\neq i_{max}}^{n} \mu^2 - 2\mu \sum_{i=1,\ i\neq i_{max}}^{n} |\psi_{1i}| \tag{16}$$

Substituting eqn. (15) in eqn. (16)

$$\implies (n-1)\sigma^2 = \sum_{i=1,\ i\neq i_{max}}^{n} |\psi_{1i}|^2 + (n-1)\mu^2 - 2(n-1)\mu^2$$

$$\implies \sum_{i=1,\ i\neq i_{max}}^{n} |\psi_{1i}|^2 = (n-1)(\mu^2+\sigma^2) \tag{17}$$

As $|\psi_{1i}|_{max} \geq |\psi_{1i}|\ \forall\ i$, we can write $|\psi_{1i}|_{max} = \mu + \rho\sigma$ for some $\rho \geq 1$. The parameter ρ controls the extent to which the signal with the highest observability magnitude deviates from the remaining. Therefore,

$$\sum_{i=1}^{n} |\psi_{1i}|^2 = |\psi_{1i}|_{max}^2 + \sum_{i=1,\ i\neq i_{max}}^{n} |\psi_{1i}|^2 = (\mu+\rho\sigma)^2 + (n-1)(\mu^2+\sigma^2) \tag{18}$$

Substituting this in eqn. (14),

$$\kappa_1(\boldsymbol{\psi}_1) = \frac{|\psi_{1i}|_{max}}{\sqrt{\sum_{i=1}^{n} |\psi_{1i}|^2}} = \frac{\mu+\rho\sigma}{\sqrt{(\mu+\rho\sigma)^2 + (n-1)(\mu^2+\sigma^2)}}$$

$$\implies \kappa_1(\hat{\boldsymbol{\Psi}}) = \kappa_1(\boldsymbol{\psi}_1) = \frac{1}{\sqrt{1 + \frac{(n-1)(\mu^2+\sigma^2)}{(\mu+\rho\sigma)^2}}} \tag{19}$$

From eqn. (19) it is evident that $\kappa_1(\hat{\boldsymbol{\Psi}})$ is an increasing function of ρ and a decreasing function of n. Therefore, one obvious way to decrease $\kappa_1(\hat{\boldsymbol{\Psi}})$ is to increase the number of signals in the selected set. However, this might lead to other challenges– first, it may not be easy to obtain a large number of signal variables with observabilities in same phase, and second, even if obtained, it is difficult to ensure minimal variation in their observability magnitudes. Moreover, if n is increased such that a signal with large obervability magnitude (outlier) compared to the rest appears in the set, it would lead to a high ρ and in turn a high $\kappa_1(\hat{\boldsymbol{\Psi}})$. This is demonstrated in the example below.

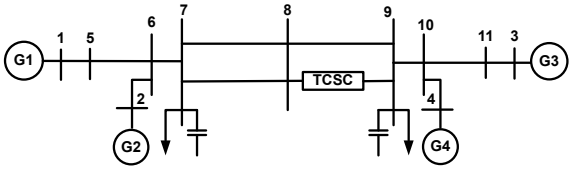

Fig. 1: 2−area 4−machine test system with TCSC

Example 1: Consider the fundamental-frequency phasor model of the 4−machine, 2−area system from [19] as shown in Fig. 1. Under nominal loading condition, it has one poorly-damped inter-area mode with frequency 0.628 Hz and damping ratio 0.012. We consider real power flows in lines $4-10$, $3-11$, and $11-10$ as a signal set. The relative observabilities of the mode in these signals are listed in Table I. Following eqn. (19), $\kappa_1(\hat{\boldsymbol{\Psi}})$ is calculated as 0.6053 and from nonlinear time-domain simulations we obtain $\kappa_1(\mathbf{Y}) = 0.6051$. Although the denseness values are reasonable, they are still above κ^*.

TABLE I: SIGNALS AND RELATIVE MODAL OBSERVABILITIES

Signals	P_{4-10}	P_{3-11}	P_{11-10}	P_{10-9}
$\lvert\psi_{1i}\rvert$	0.5922	0.6935	0.6935	1.2621
$\angle\psi_{1i}$	$\angle 90°$	$\angle 91°$	$\angle 91°$	$\angle 90°$

If we add one more signal− real power flow in line $10-9$ to the set, which has a high $|\psi_{1i}|$ relative to the rest (and thus, ideal for modal estimation), $\kappa_1(\hat{\boldsymbol{\Psi}})$ and $\kappa_1(\mathbf{Y})$ increase to 0.7405 and 0.7404, respectively. Thus, instead of aiding to resilience the added signal worsens the chances of data recovery. □

From this, we conclude that for guaranteed signal recovery, it is not enough to increase the number of signals – care should be taken to minimize variance in observability magnitudes and avoid outliers. This could mean sacrificing the best signal (having the highest $|\psi_{1i}|$ among all signals) for another signal with relatively lower $|\psi_{1i}|$. *This is in conflict with our conventional wisdom of signal selection for modal estimation.*

Next, considering the recommendations above, let signals be grouped with variance in $|\psi_{1i}|$ minimized, such that we can assume $\frac{\mu}{\sigma} >> 1$ and ρ in the neighborhood of 1. Under these conditions, we can approximate eqn. (19) as follows.

$$\kappa_1(\hat{\boldsymbol{\Psi}}) = \kappa_1(\boldsymbol{\psi}_1) = \frac{\rho + \frac{\mu}{\sigma}}{\sqrt{(\rho+\frac{\mu}{\sigma})^2 + (n-1)(1+\frac{\mu^2}{\sigma^2})}} \approx \frac{\rho+\frac{\mu}{\sigma}}{\sqrt{n}\frac{\mu}{\sigma}} = \frac{1}{\sqrt{n}}\left(1+\frac{\rho\sigma}{\mu}\right) \tag{20}$$

Eqn. (20) shows that for sets sufficiently dense, κ_1 increases linearly with ρ and σ from its minimum (and ideal) value $\frac{1}{\sqrt{n}}$.

Next, we validate our proposition using the following example from the system in Fig. 1.

Example 2: First, consider signals P_{6-7}, P_{7-8}, P_{8-9}, P_{9-10}, and $\theta_7-\theta_9$ as a set. Refer to Table II for the ψ_{1i} values. Clearly, $\theta_7-\theta_9$ appears as an outlier in the set considering the $|\psi_{1i}|$ values. Therefore, the set is expected to have a high $\kappa_1(\hat{\boldsymbol{\Psi}})$. Note that, the phase angles $\angle\psi_{1i}$-s are not exactly equal but the difference is small enough to ensure $\mathbf{Y}$ has only 1 dominant singular value with the corresponding singular vector capturing 95% variance in data (largest singular value: 24.04, second largest: 0.65), and thus, for all practical purposes this can be analyzed as a rank 1 case.

TABLE II: SIGNALS AND RELATIVE MODAL OBSERVABILITIES

Signals	P_{6-7}	P_{7-8}	P_{8-9}	P_{9-10}	$\theta_7-\theta_9$
$\lvert\psi_{1i}\rvert$	0.4554	0.4997	0.5064	1.2621	5.0733
$\angle\psi_{1i}$	$\angle -74°$	$\angle -83°$	$\angle -83°$	$\angle -90°$	$\angle -84°$

Using eqn. (19), we obtain $\kappa_1(\hat{\boldsymbol{\Psi}}) = 0.9580$, which is significantly higher than κ^*. Therefore, when measurement

$\theta_7 - \theta_9$ is corrupted by maliciously injecting a signal with negative damping at $t = 23$ s as shown in Fig. 2 – the signal recovery is not guaranteed. Figure 2(a) shows that the recovery using the framework in eqn. (7) failed.

This implies modal estimation on the reconstructed signals would result in erroneous estimates of damping ratio and frequency for the oscillation. This is seen in the damping ratio estimation plots of Fig. 3(a). In this example, we use multi-channel Prony analysis [27] on the signal set in Table II with a moving window of 150 samples for estimation of damping.

TABLE III: SIGNALS AND RELATIVE MODAL OBSERVABILITIES

Signals	P_{6-7}	P_{7-8}	P_{8-9}	P_{10-4}	$\theta_4 - \theta_{11}$
$\lvert\psi_{1i}\rvert$	0.4554	0.4997	0.5064	0.5922	0.4500
$\angle\psi_{1i}$	$\angle -74°$	$\angle -83°$	$\angle -83°$	$\angle -90°$	$\angle -84°$

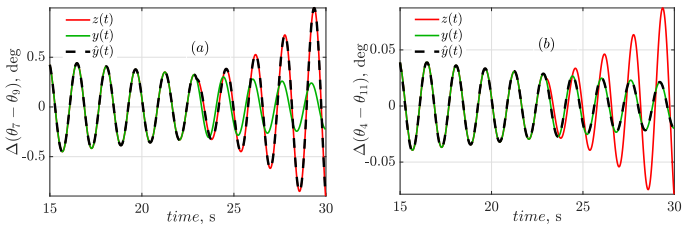

Fig. 2: (a) Incorrect $1-$sparse recovery of $\Delta(\theta_7 - \theta_9)$ from signals in Table II and (b) exact $1-$sparse recovery of $\Delta(\theta_4 - \theta_{11})$ from signals in Table III.

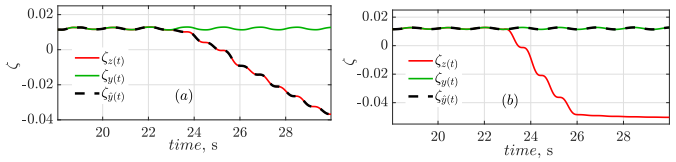

Fig. 3: (a) Damping ratio estimation using multi-channel Prony from incorrect $1-$sparse recovery of $\Delta(\theta_7 - \theta_9)$ from signals in Table II and (b) exact $1-$sparse recovery of $\Delta(\theta_4 - \theta_{11})$ from signals in Table III.

TABLE IV: FROM LINEARIZED MODEL OF THE SYSTEM IN FIG. 1

Damp. ratio ζ : 0.0121	Modal freq. f (Hz) : 0.628

Next, adhering to our recommendations on minimizing variance in $|\psi_{1i}|$-s we replace signals P_{9-10} and $\theta_7 - \theta_9$ by P_{10-4} and $\theta_4 - \theta_{11}$. The relative modal observabilities for the new signal set is listed in Table III. For this set, $\sigma = 0.025$, $\mu = 0.447$ and $\rho = 4.51$, and therefore, from eqn. (19), $\kappa_1(\hat{\mathbf{\Psi}}) = 0.5262$. This is less than κ^* and therefore, guarantees exact recovery under all $1-$sparse corruptions. To verify that, we corrupt signal $\theta_4 - \theta_{11}$ and as shown in Fig. 2(b), the reconstruction is accurate. And therefore, the damping ratio estimation from the recovered signals (see Fig. 3(b)) matches that of the linear model in Table IV.

Also, in this case, $\frac{\mu}{\sigma} = 18.83 >> \rho$, and therefore, we can use the expression in eqn. (20) as an approximate estimate of denseness. Substituting the values of ρ, σ, μ and n in eqn. (20), we get $\frac{1}{\sqrt{n}}\left(1 + \frac{\rho\sigma}{\mu}\right) = 0.554$, which is a reasonable approximation of $\kappa_1(\hat{\mathbf{\Psi}})$ for this case. □

Next, we demonstrate the robustness of the proposed approach to measurement noise. White Gaussian noise is added to all five signals in each of the two groups in Tables II and III

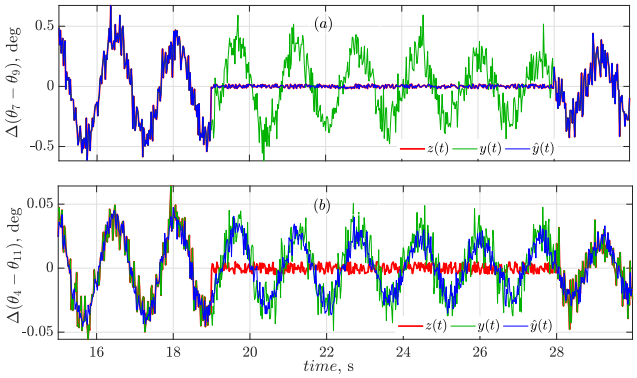

Fig. 4: (a) Incorrect $1-$sparse recovery of $\Delta(\theta_7 - \theta_9)$ from noisy signals in Table II and (b) correct $1-$sparse recovery of $\Delta(\theta_4 - \theta_{11})$ from noisy signals in Table III under missing data attack. SNR in simulation is 40 dB.

to achieve a desired signal-to-noise ratio (SNR). It is assumed that there happens a missing data attack in signals $\theta_7 - \theta_9$ and $\theta_4 - \theta_{11}$ for a period of 9 s (between $t = 19$ and 28 s, see Fig. 4). Following the recovery framework in Section II-C as before, these signals are then reconstructed from the remaining signals in their respective groups. Signal reconstruction for an SNR of 40 dB is shown in Fig. 4. It is seen that the recovery of $\theta_7 - \theta_9$ from the signals in Table II has failed and the recovery of $\theta_4 - \theta_{11}$ from signals in Table III has succeeded as before. This reinforces our claim.

To study the effects of the quality of data recovery on oscillation monitoring, we next perform modal estimation for both the signal sets using multi-channel Prony analysis [27] on the 15 s measurement window shown in Fig. 4. This is repeated for SNRs 50 and 60 dBs, and in each case 100 simulations are performed with uncorrelated noise being added each time. Statistical dispersion of the estimated damping ratio ζ and oscillation frequency f for each of the three noise levels and the respective signal sets are listed in Tables V and VI.

TABLE V: MODAL ESTIMATION USING MULTI-CHANNEL PRONY ON NOISY SIGNALS IN TABLE II WITH MISSING DATA ATTACK ON $\Delta(\theta_7 - \theta_9)$ AND AFTER RECOVERY FOR DIFFERENT NOISE LEVELS

SNR (dB)	With Corruption		After Recovery	
	mean ζ $\pm$ std. dev.	mean f (Hz) $\pm$ std. dev.	mean ζ $\pm$ std. dev.	mean f (Hz) $\pm$ std. dev.
60	0.0630 $\pm$ 0.0012	0.627 $\pm$ 0.0047	0.0610 $\pm$ 0.0017	0.625 $\pm$ 0.0086
50	0.0677 $\pm$ 0.0011	0.624 $\pm$ 0.0080	0.0626 $\pm$ 0.0036	0.625 $\pm$ 0.0053
40	0.0640 $\pm$ 0.0007	0.625 $\pm$ 0.0051	0.0621 $\pm$ 0.0019	0.625 $\pm$ 0.0073

TABLE VI: MODAL ESTIMATION USING MULTI-CHANNEL PRONY ON NOISY SIGNALS IN TABLE III WITH MISSING DATA ATTACK ON $\Delta(\theta_4 - \theta_{11})$ AND AFTER RECOVERY FOR DIFFERENT NOISE LEVELS

SNR (dB)	With Corruption		After Recovery	
	mean ζ $\pm$ std. dev.	mean f (Hz) $\pm$ std. dev.	mean ζ $\pm$ std. dev.	mean f (Hz) $\pm$ std. dev.
60	0.0277 $\pm$ 0.0017	0.633 $\pm$ 0.0068	0.0127 $\pm$ 0.0011	0.628 $\pm$ 0.0032
50	0.0274 $\pm$ 0.0020	0.637 $\pm$ 0.0055	0.0128 $\pm$ 0.0014	0.628 $\pm$ 0.0022
40	0.0260 $\pm$ 0.0027	0.632 $\pm$ 0.0088	0.0121 $\pm$ 0.0016	0.629 $\pm$ 0.0062

As we can see from Table V, an incorrect signal recovery during an attack can lead to erroneous estimation of system modes. In this particular case, during a missing data attack the estimated values of damping ratio from the recovered signals in Table II could lead to a false judgement that the inter-area mode is better damped. In contrast, if the signals in Table III are used for oscillation monitoring, the damping ratio and oscillation frequency can be estimated accurately from the recovered signals. This underscores the merit of our proposition.

B. Generic Signal Selection for Unimodal Case

So far, we considered a group of signals with same phase-relationship in observabilities that lead to a rank-1 $\hat{\boldsymbol{\Psi}}$ matrix. We now study a generic case of selecting signals with varying observability magnitudes and angles. This would mean the rank of $\hat{\boldsymbol{\Psi}}$ would now be 2, and therefore, the minimum attainable value for denseness would be $\sqrt{2/n}$.

We ask under what conditions can we achieve $\kappa_1(\hat{\boldsymbol{\Psi}})$ close to $\sqrt{2/n}$. This is challenging, because the denseness calculation now involves two orthonormal vectors, which although derived from complex conjugate column-pairs, do not have any structural similarities. However, with some special conditions imposed on signal selection, we show that denseness of $\hat{\boldsymbol{\Psi}}$ can be expressed as a multiple of the denseness of its column $\boldsymbol{\psi}_1$. This allows us to build on our previously gained insights on minimizing denseness of a column vector to minimize $\kappa(\hat{\boldsymbol{\Psi}})$. This is presented in the following lemma.

Lemma 4. If the entries in $\boldsymbol{\psi}_1$ are such that $\langle \boldsymbol{\psi}_1, \overline{\boldsymbol{\psi}}_1 \rangle = 0$, then $\kappa_1(\hat{\boldsymbol{\Psi}}) = \sqrt{2}\ \kappa_1(\boldsymbol{\psi}_1)$.

Proof. Let $\boldsymbol{\psi}_1 = \big[\ |\psi_{11}|\underline{/\theta_1}\ \ldots\ |\psi_{1i}|\underline{/\theta_i}\ \ldots\ |\psi_{1n}|\underline{/\theta_n}\ \big]^T$ and $\overline{\boldsymbol{\psi}}_1 = \big[\ |\psi_{11}|\underline{/-\theta_1}\ \ldots\ |\psi_{1i}|\underline{/-\theta_i}\ \ldots\ |\psi_{1n}|\underline{/-\theta_n}\ \big]^T$.

Following Gram-Schmidt orthonormalization as before, $\hat{\mathbf{u}}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|_2} = \frac{\boldsymbol{\psi}_1}{\|\boldsymbol{\psi}_1\|_2}$, and $\hat{\mathbf{u}}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|_2}$ where, $\mathbf{u}_2 = \overline{\boldsymbol{\psi}}_1 - \frac{\langle \boldsymbol{\psi}_1, \overline{\boldsymbol{\psi}}_1 \rangle}{\langle \boldsymbol{\psi}_1, \boldsymbol{\psi}_1 \rangle}\boldsymbol{\psi}_1$. If the complex inner product $\langle \boldsymbol{\psi}_1, \overline{\boldsymbol{\psi}}_1 \rangle = \sum_{i=1}^{n} |\psi_{1i}|^2\ \underline{/-2\theta_i} = 0$, then $\big[\ \hat{\mathbf{u}}_1\ \ \hat{\mathbf{u}}_2\ \big] = \Big[\ \frac{\boldsymbol{\psi}_1}{\|\boldsymbol{\psi}_1\|_2}\ \ \frac{\overline{\boldsymbol{\psi}}_1}{\|\boldsymbol{\psi}_1\|_2}\ \Big]$. So,

$$\kappa_1(\hat{\boldsymbol{\Psi}}) = \kappa_1\big(\big[\ \hat{\mathbf{u}}_1\ \ \hat{\mathbf{u}}_2\ \big]\big) = \frac{1}{\|\boldsymbol{\psi}_1\|_2} \max_i \Big\| \mathbf{e}_i^* \Big[\ \boldsymbol{\psi}_1\ \ \overline{\boldsymbol{\psi}}_1\ \Big] \Big\|_2$$

$$= \frac{\max_i \sqrt{|\psi_{1i}|^2 + |\psi_{1i}|^2}}{\|\boldsymbol{\psi}_1\|_2} = \frac{\sqrt{2}|\psi_{1i}|_{max}}{\sqrt{\sum_{i=1}^{n} |\psi_{1i}|^2}} = \sqrt{2}\ \kappa_1(\boldsymbol{\psi}_1) \quad \square$$

Lemma 4 presents an useful insight towards achieving the minimum value of $\kappa_1(\hat{\boldsymbol{\Psi}})$ for the generic signal selection case. In addition to minimizing the variance in $|\psi_{1i}|$ s as prescribed before, if we can ensure the inner product $\langle \boldsymbol{\psi}_1, \overline{\boldsymbol{\psi}}_1 \rangle = 0$, then

$$\kappa_1(\hat{\boldsymbol{\Psi}}) = \sqrt{\frac{2}{n}}\big(1 + \frac{\rho\sigma}{\mu}\big) \tag{21}$$

This follows from our previous results. Since the variance in observability magnitudes is small (small σ, and ρ in neighborhood of 1), $\kappa_1(\boldsymbol{\psi}_1)$ can be approximated to $\frac{1}{\sqrt{n}}\big(1 + \frac{\rho\sigma}{\mu}\big)$. Next, with the condition imposed on the complex inner product, we arrive at the expression in eqn. (21) using Lemma 4. Note, that this is also consistent with our understanding that the minimum value $\sqrt{2/n}$ is attained for $\sigma = 0$.

Although Lemma 4 provides useful insights on attaining lower values for denseness, finding signals to achieve $\langle \boldsymbol{\psi}_1, \overline{\boldsymbol{\psi}}_1 \rangle = 0$ in practice can be quite challenging. Therefore, instead of making it exactly zero, we look for the implications of small but non-zero inner products on the denseness $\kappa_1(\hat{\boldsymbol{\Psi}})$. This is summarized in the lemma below.

Lemma 5. If signals are selected such that the variation in observability magnitudes is small ($\frac{\mu}{\sigma} >> \rho$) and the absolute value of the inner product $|\langle \boldsymbol{\psi}_1, \overline{\boldsymbol{\psi}}_1 \rangle| \leq \epsilon$ for some small ϵ, then

$$\kappa_1(\hat{\boldsymbol{\Psi}}) \leq \sqrt{\frac{2}{n}}\big(1 + \frac{\rho\sigma}{\mu}\big)\big(1 + \frac{\epsilon}{2\ \|\boldsymbol{\psi}_1\|_2^2}\big)$$

Proof. Let $\langle \boldsymbol{\psi}_1, \overline{\boldsymbol{\psi}}_1 \rangle = \epsilon\ \underline{/\phi}$, then following Gram-Schmidt orthonormalization as before,

$$\hat{\mathbf{u}}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|_2} = \frac{\boldsymbol{\psi}_1}{\|\boldsymbol{\psi}_1\|_2} \quad \text{and} \quad \hat{\mathbf{u}}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|_2},$$

$$\text{where, } \mathbf{u}_2 = \overline{\boldsymbol{\psi}}_1 - \frac{\langle \boldsymbol{\psi}_1, \overline{\boldsymbol{\psi}}_1 \rangle}{\langle \boldsymbol{\psi}_1, \boldsymbol{\psi}_1 \rangle}\boldsymbol{\psi}_1 = \overline{\boldsymbol{\psi}}_1 - \frac{\epsilon\ \underline{/\phi}}{\|\boldsymbol{\psi}_1\|_2^2}\boldsymbol{\psi}_1$$

Therefore, any i^{th} entry of $\mathbf{u}_2$ can be expressed as,

$$u_{2i} = |\psi_{1i}|\underline{/-\theta_i} - \frac{\epsilon\ |\psi_{1i}|}{\|\boldsymbol{\psi}_1\|_2}\underline{/\theta_i + \phi}$$

We calculate $|u_{2i}|^2$, and ϵ being small, neglect the term $\frac{\epsilon^2}{\|\boldsymbol{\psi}_1\|_2^4}$ hereafter. Therefore,

$$|u_{2i}|^2 \approx |\psi_{1i}|^2\big\{1 - \frac{2\ \epsilon}{\|\boldsymbol{\psi}_1\|_2^2}\cos(2\theta_i + \phi)\big\} \tag{22}$$

The norm $\|\mathbf{u}_2\|_2$ can then be calculated as,

$$\|\mathbf{u}_2\|_2^2 = \sum_{i=1}^{n} |u_{2i}|^2 = \sum_{i=1}^{n} |\psi_{1i}|^2 + \frac{2\ \epsilon}{\|\boldsymbol{\psi}_1\|_2^2}\sum_{i=1}^{n} |\psi_{1i}|^2\cos(2\theta_i + \phi) \tag{23}$$

$$\text{Now, } \langle \boldsymbol{\psi}_1, \overline{\boldsymbol{\psi}}_1 \rangle = \epsilon\ \underline{/\phi} \implies \sum_{i=1}^{n} |\psi_{1i}|^2\ \underline{/-2\theta_i} = \epsilon\ \underline{/\phi} \tag{24}$$

If each phasor $|\psi_{1i}|^2\underline{/-2\theta_i}$ in the summation is rotated by angle ϕ counter-clockwise, the resultant phasor also gets rotated in same amount.

$$\text{This implies, } \sum_{i=1}^{n} |\psi_{1i}|^2\ \underline{/-2\theta_i - \phi} = \epsilon\ \underline{/\phi - \phi} = \epsilon \tag{25}$$

Substituting the real part of (25) in eqn. (23),

$$\|\mathbf{u}_2\|_2^2 = \|\boldsymbol{\psi}_1\|_2^2 + \frac{2\ \epsilon^2}{\|\boldsymbol{\psi}_1\|_2^2} = \|\boldsymbol{\psi}_1\|_2^2\big(1 + \frac{2\ \epsilon^2}{\|\boldsymbol{\psi}_1\|_2^4}\big) \approx \|\boldsymbol{\psi}_1\|_2^2 \tag{26}$$

Therefore, following the definition of denseness,

$$\kappa_1(\hat{\boldsymbol{\Psi}}) = \kappa_1\big(\big[\ \hat{\mathbf{u}}_1\ \ \hat{\mathbf{u}}_2\ \big]\big) = \frac{1}{\|\boldsymbol{\psi}_1\|_2}\max_i \Big\| \big[\ u_{1i}\ \ u_{2i}\ \big] \Big\|_2$$

$$= \frac{1}{\|\boldsymbol{\psi}_1\|_2}\max_i \Big\{|\psi_{1i}|^2 + |\psi_{1i}|^2\big\{1 - \frac{2\ \epsilon}{\|\boldsymbol{\psi}_1\|_2^2}\cos(2\theta_i + \phi)\big\}\Big\}^{\frac{1}{2}}$$

$$\leq \frac{\sqrt{2}\ |\psi_{1i}|_{max}}{\|\boldsymbol{\psi}_1\|_2}\max_i \Big\{1 - \frac{\epsilon}{\|\boldsymbol{\psi}_1\|_2^2}\cos(2\theta_i + \phi)\Big\}^{\frac{1}{2}}$$

$$\leq \frac{\sqrt{2}\ |\psi_{1i}|_{max}}{\|\boldsymbol{\psi}_1\|_2}\Big\{1 + \frac{\epsilon}{\|\boldsymbol{\psi}_1\|_2^2})\Big\}^{\frac{1}{2}} = \sqrt{2}\kappa_1(\boldsymbol{\psi}_1)\Big\{1 + \frac{\epsilon}{\|\boldsymbol{\psi}_1\|_2^2})\Big\}^{\frac{1}{2}} \tag{27}$$

Finally for small ϵ, we can say $\frac{\epsilon}{\|\boldsymbol{\psi}_1\|_2^2} << 1$, and therefore, we apply binomial expansion to write,

$$\kappa_1(\hat{\boldsymbol{\Psi}}) \leq \sqrt{\frac{2}{n}}\big(1+\frac{\rho\sigma}{\mu}\big)\big(1+\frac{\epsilon}{2\ \|\boldsymbol{\psi}_1\|_2^2}\big) \quad \square \tag{28}$$

This is demonstrated in the example below.

Example 3: In addition to two power flow signals P_{3-11} and P_{11-10} as before, consider two frequency difference signals $f_1 - f_3$ and $f_2 - f_3$, and the angle difference signal $\theta_3 - \theta_{11}$ from the system in Fig. 1. The relative modal observabilities are listed in Table VII. Firstly, owing to difference in phases, the rank of $\hat{\boldsymbol{\Psi}}$ is 2. This is also reflected in the two dominant singular values of $\mathbf{Y}$. Following the definition of denseness, the actual values of $\kappa_1(\hat{\boldsymbol{\Psi}})$ and $\kappa_1(\mathbf{Y})$ for this case is calculated as 0.7385 and 0.7377, respectively. Next, we study the validity of our approximations stated in Lemma 4.

TABLE VII: SIGNALS AND RELATIVE MODAL OBSERVABILITIES

Signals	$\theta_3 - \theta_{11}$	P_{3-11}	P_{11-10}	$f_1 - f_3$	$f_2 - f_3$
$\lvert\psi_{1i}\rvert$	0.6340	0.6935	0.6935	0.8231	0.7501
$\angle\psi_{1i}$	$\angle 84.1°$	$\angle 91°$	$\angle 91°$	$\angle 3.5°$	$\angle 3.8°$

In this case, $|\langle\boldsymbol{\psi}_1, \overline{\boldsymbol{\psi}}_1\rangle| = \epsilon = 0.138$, also $\frac{\epsilon^2}{\|\boldsymbol{\psi}_1\|_2^4} = 0.0028 << 1$ and thus, can be neglected. From the variation in magnitudes of $|\psi_{1i}|$s, we obtain $\mu = 0.693$, $\sigma = 0.041$ and $\rho = 3.19$. Substituting these in eqn. (28) we get, $\kappa_1(\hat{\boldsymbol{\Psi}}) < 0.7711$. This validates our proposition on the upper bound on denseness due to a small non-zero perturbation in inner product. $\square$

It can be seen from Table VII that there are two distinct groups with phases approximately in quadrature. Also, the signals are selected such that the vector norms of $|\psi_{1i}|$ in each group are nearly equal. These two conditions together achieve a small inner product, and therefore, a reasonably accurate estimate of $\kappa_1(\hat{\boldsymbol{\Psi}})$.

C. Selection of Signals with Multiple Modes

Next, we extend these insights on minimizing κ_s to signal sets with multiple poorly-damped modes. Building on the notion that a signal with multiple modes can be decomposed into its constituent frequencies, we seek to enhance denseness for each of these mono-frequency components. For a set of candidate signals, ensuring minimal variation in observability phase angle and magnitudes for each mode results in a $\mathbf{Y}$ with numerical-rank approximately 1, since the basis vector corresponding to the largest singular value captures most of the variance in data. As we show next, this can be used to contain $\kappa_s(\mathbf{Y})$ below κ^*.

Example 4: We consider the positive-sequence fundamental-frequency model of the IEEE 5−area, 16−machine New England − New York test system [28] shown in Fig. 5. Eigen-analysis of the system under nominal loading indicates presence of 4 poorly-damped modes with frequencies− 0.51 Hz, 0.39 Hz, 0.62 Hz and 0.79 Hz, and settling times− 28.8 s, 25.7 s, 18.1 s and 16.1 s, respectively. We assume that all generator buses are equipped with PMUs.

We study two signal selection scenarios to compare the denseness of the respective signal subspaces and its implication in recovering clean signals corrupted with malicious injections. We first consider voltage magnitude signals at the terminals of generators− 1, 6, 7, and 8. The relative observabilities of each of the 4 modes in these signals is listed in Table VIII. The denseness κ_1 of the observability submatrix for each of the 4 modes is listed in Table IX. Clearly, none of the $\kappa_1(\hat{\boldsymbol{\Psi}}_i)$-s are below the desired κ^*, and therefore, recovery of any of these signals for neither of the modes can be guaranteed. Moreover, $\kappa_1(\hat{\boldsymbol{\Psi}}_4)$ is very high and nearly approximates 1. This can be attributed to the fact that the relative observability of 0.79 Hz mode in $|V_1|$ is significantly higher as compared to remaining 3 signals.

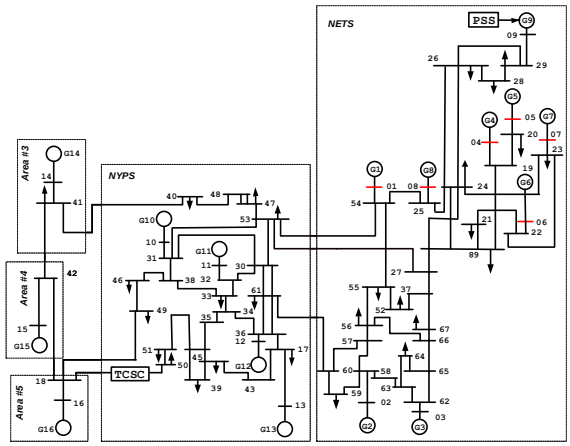

Fig. 5: IEEE 5−area 16-machine New England - New York test system.

TABLE VIII: SIGNALS AND RELATIVE MODAL OBSERVABILITIES

Signals	$\lvert\psi_{1i}\rvert\angle\psi_{1i}$ (0.51 Hz)	$\lvert\psi_{2i}\rvert\angle\psi_{2i}$ (0.39 Hz)	$\lvert\psi_{3i}\rvert\angle\psi_{3i}$ (0.62 Hz)	$\lvert\psi_{4i}\rvert\angle\psi_{4i}$ (0.79 Hz)
$\lvert V_1\rvert$	0.0013 $\angle -58.8°$	0.0032 $\angle 98.4°$	0.0048 $\angle -159.7°$	0.0025 $\angle 101.5°$
$\lvert V_8\rvert$	0.0079 $\angle -84.6°$	0.0044 $\angle -74.6°$	0.0095 $\angle -168.7°$	0.0004 $\angle -142.6°$
$\lvert V_6\rvert$	0.0075 $\angle -85.4°$	0.0046 $\angle -79.7°$	0.0089 $\angle -169.2°$	0.0008 $\angle -124.5°$
$\lvert V_7\rvert$	0.0075 $\angle -85.3°$	0.0046 $\angle -80.2°$	0.0089 $\angle -169.1°$	0.0008 $\angle -122.4°$
$\lvert V_4\rvert$	0.0063 $\angle -85.5°$	0.0039 $\angle -80.9°$	0.0074 $\angle -168.4°$	0.0007 $\angle -125.7°$
$\lvert V_5\rvert$	0.0078 $\angle -85.0°$	0.0045 $\angle -76.9°$	0.0093 $\angle -169.1°$	0.0006 $\angle -130.9°$

Also, observe that for the 0.51 Hz mode, having the largest settling time of the three, the observability magnitude in the signal set varies between 0.0013 and 0.0078. This negatively impacts the overall denseness[1] of the signal set. $\kappa_1(\hat{\boldsymbol{\Psi}})$ is calculated as 0.8116. To study the implications of such a high value, we corrupt the signal $|V_8|$ as shown in Fig. 6 (b). All 4 clean signals without corruption are shown in Fig. 6 (a). It is evident from Fig. 6 (b) that the recovery of $|V_8|$ is inaccurate, as the reconstructed signal traces the corruption.

TABLE IX: MODE-WISE DENSENESS[2] OF $|V_1|, |V_8|, |V_6|$, and $|V_7|$

Modes	0.51 Hz	0.39 Hz	0.62 Hz	0.79 Hz
$\kappa_1(\hat{\boldsymbol{\Psi}}_\mathbf{i})$	0.5959	0.5470	0.5765	0.9968

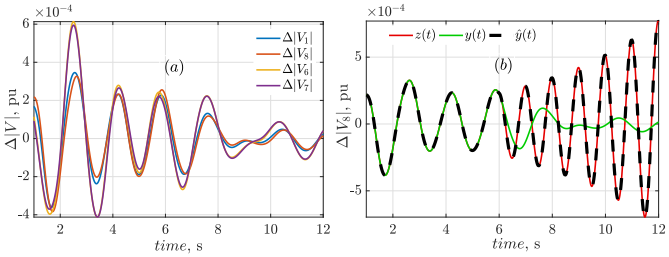

Fig. 6: Plots for (a) signals $\Delta|V_1|, \Delta|V_8|$, $\Delta|V_4|$, and $\Delta|V_5|$ without corruption, and (b) incorrect $1-$sparse recovery of $\Delta|V_8|$ after corruption.

Next, we replace signals $|V_1|$ and $|V_8|$ by signals $|V_4|$ and $|V_5|$. The observabilities are listed in Table VIII and as shown in Table X, mode-wise denseness $\kappa_1(\mathbf{\Psi}_i) < \kappa^* \ \forall \ i$. The overall denseness[1] $\kappa_1(\mathbf{\Psi})$ is calculated as 0.5359, which is also less than κ^*, i.e. signal recovery is guaranteed. To validate this claim, we corrupt $|V_5|$ as in Fig. 7 (b). All 4 clean signals without corruption are shown in Fig. 7 (a). It can be seen from Fig. 7 (b) that the recovery of $|V_5|$ is exact.

TABLE X: MODE-WISE DENSENESS[2] OF $|V_4|, |V_5|, |V_6|$, and $|V_7|$

Modes	0.51 Hz	0.39 Hz	0.62 Hz	0.79 Hz
$\kappa_1(\mathbf{\Psi}_i)$	0.5328	0.5229	0.5377	0.5519
$\frac{1}{\sqrt{n}}\left(1+\frac{\rho_i \sigma_i}{\mu_i}\right)$	0.5468	0.5322	0.5541	0.5761

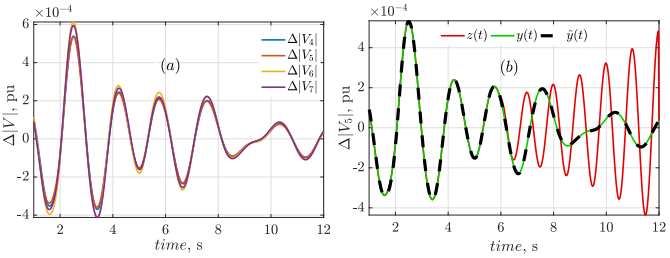

Fig. 7: Plots for (a) signals $\Delta|V_6|, \Delta|V_7|$, $\Delta|V_4|$, and $\Delta|V_5|$ without corruption, and (b) exact $1-$sparse recovery of $\Delta|V_5|$ after corruption.

TABLE XI: MODAL ESTIMATION FROM SIGNALS $|V_8|$ and $|V_5|$ WITH CORRUPTION AND AFTER RECOVERY

Mode-wise Damp. ratio & Frequency	Linear Model	Prony on $\Delta\|V_8\|$		Prony on $\Delta\|V_5\|$	
		$z(t)$	$\hat{y}(t)$	$z(t)$	$\hat{y}(t)$
ζ_1	0.063	-0.751	-0.760	-0.889	0.060
f_1 (Hz)	0.391	0.466	0.466	0.330	0.393
ζ_2	0.043	-0.994	-0.994	-0.565	0.045
f_2 (Hz)	0.508	0.515	0.515	0.581	0.512
ζ_3	0.056	-0.302	-0.302	-0.982	0.055
f_3 (Hz)	0.622	0.660	0.660	0.612	0.622
ζ_4	0.050	-0.999	-0.999	-0.131	0.061
f_4 (Hz)	0.792	0.787	0.787	0.747	0.799

Table XI lists the results of modal estimation on the signals $|V_5|$ and $|V_8|$ – both when corrupted and after recovery (as discussed above), and compares those with that obtained from the linearized model. We can see from the table that for V_5 the damping ratios and frequencies estimated from the recovered signal closely match those of the linear model for all 4 poorly-damped modes. This validates our proposition that modal estimation from the signal set in Fig. 7 is robust to corruption. On the other hand, the modal estimation from the reconstructed $|V_8|$ is inaccurate leading to an impression that the modes are unstable.

IV. Signal Selection from Field PMU Data

Finally, we extend the insights derived before for grouping signals directly from field PMU data in absence of a small-signal model. This is entirely data-driven and can find potential application during system operation. To that end, we consider detrended bus voltage magnitude signals $|V_1|$ to $|V_{40}|$, from 40 different PMU locations in New York Power Authority (NYPA). Denseness calculation on the entire data set yields $\kappa_1(\mathbf{Y}) = 0.549$ – just enough to guarantee recovery of 1 in 40 signals. This ratio being abysmally low, our objective is to selectively group signals from this set using our proposed recommendations to improve corruption resilience.

We consider a 50 s moving window of archived data (assume previously recovered, and therefore, trusted) and perform spectral decomposition on each of the 40 signals. We observe two frequencies within the signal set– 0.25 Hz and 0.06 Hz. Next, with $|V_1|$ as reference we compute the output-to-output transfer function, and the relative modeshapes for all 40 signals at the mentioned frequencies– 0.25 Hz and 0.06 Hz, as described in [29]. Since the output-to-output transfer functions were computed with respect to a fixed reference, modeshapes thus obtained are equivalent to relative modal observabilities. Next, going by our propositions, we group signals having similar phase and magnitude of relative modal observabilities for *each* of the two frequency components. To do so, we use $k-$means clustering [30] on the signal set using the 2×1 feature vector of complex-valued relative observabilites.

TABLE XII: SIGNAL SETS I AND II

Signal Set I	$\|V_6\|$, $\|V_{15}\|$, $\|V_{19}\|$, $\|V_{20}\|$, $\|V_{21}\|$, $\|V_{26}\|$, $\|V_{27}\|$, $\|V_{28}\|$
Signal Set II	$\|V_4\|$, $\|V_7\|$, $\|V_{11}\|$, $\|V_{13}\|$, $\|V_{16}\|$, $\|V_{17}\|$, $\|V_{22}\|$, $\|V_{24}\|$

Two such clusters– signal sets I and II, obtained for $k = 10$ are listed in Table XII. Figures 8 (a) and 9 (a) show the time-domain plots of all 8 signals in each cluster, along with power spectral density (PSD) plots of one representative signal from each cluster in Figs 8 (b) and 9 (b). Clearly, signals in set I exhibit unimodal oscillations of 0.25 Hz, while those in set II indicate presence of both 0.25 Hz and 0.06 Hz oscillations.

The denseness coefficients[1] κ_1 and κ_2 corresponding to $1-$sparse and $2-$sparse recoveries for these signal sets are listed in Table XIII. It can be seen that for both the signal sets, $\kappa_2(\mathbf{Y}) < \kappa_2(\hat{\mathbf{\Psi}})$. This validates our claim in lemma 1. Further, in both cases, $\kappa_2(\hat{\mathbf{\Psi}}) < \kappa^*$, which guarantees exact recovery upto 2 of 8 signals in each group. This is a significant improvement from 1 in 40.

TABLE XIII: DENSENESS VALUES FOR SIGNAL SETS I AND II

Group	$\kappa_1(\mathbf{Y})$	$\kappa_1(\hat{\mathbf{\Psi}})$	$\kappa_2(\mathbf{Y})$	$\kappa_2(\hat{\mathbf{\Psi}})$
Signal Set I	0.3806	0.4109	0.5382	0.5530
Signal Set II	0.3752	0.3804	0.5275	0.5363

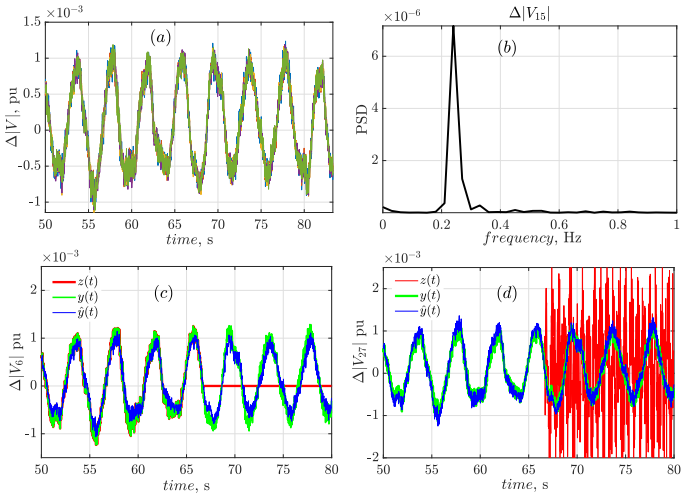

Fig. 8: Plots for (a) all signals in set I prior to corruption with (b) PSD, and (c)-(d) correct $2-$sparse recovery after corruption of $|V_6|$ and $|V_{27}|$.

TABLE XIV: SET II: FREQUENCY-WISE DENSENESS[2] AND BOUNDS

Freq.	$\kappa_1(\hat{\mathbf{\Psi}}_\mathbf{i})$	$\frac{1}{\sqrt{n}}\left(1+\frac{\rho_i\sigma_i}{\mu_i}\right)$	$\kappa_2(\hat{\mathbf{\Psi}}_\mathbf{i})$	$\sqrt{\frac{2}{n}}\left(1+\frac{\rho_i\sigma_i}{\mu_i}\right)$
0.25 Hz	0.3794	0.3842	0.5259	0.5433
0.06 Hz	0.3884	0.3953	0.5438	0.5590

[1] computed on singular vectors capturing 95% or more variance in data

[2] Ideally, $\kappa_s(\hat{\mathbf{\Psi}}) \geq \kappa_s(\hat{\mathbf{\Psi}}_\mathbf{j})\ \forall\ j \leq k$, when computed on all k singular vectors. Here, mode-wise variance minimization results in 1^{st} singular vector capturing 97% variance. $\kappa_s(\hat{\mathbf{\Psi}})$ thus computed, appears slightly less than some $\kappa_s(\hat{\mathbf{\Psi}}_\mathbf{j})$-s due to numerical approximations in truncation.

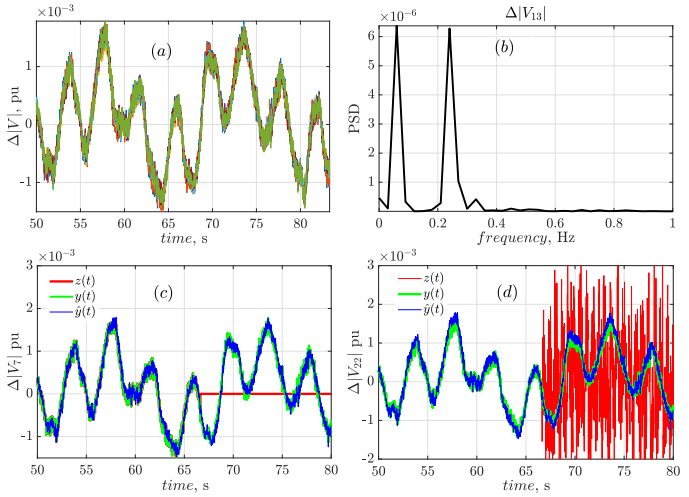

Fig. 9: Plots for (a) all signals in set II prior to corruption with (b) PSD, and (c)-(d) correct $2-$sparse recovery after corruption of $|V_7|$ and $|V_{22}|$.

Table XIV shows the denseness coefficients[2] of set II for each constituent frequency component, obtained corresponding to each complex-conjugate column pair $\hat{\mathbf{\Psi}}_\mathbf{j}$ of $\mathbf{\Psi} = \left[\hat{\mathbf{\Psi}}_\mathbf{1}\hat{\mathbf{\Psi}}_\mathbf{2}\right]$. As is seen, for each frequency component, the sufficiency condition for $2-$sparse recovery is attained. Moreover, it is seen that the bound on κ_1 derived in Section III, and extended to κ_2 using the identity $\kappa_s < \sqrt{s}\kappa_1$ [26], is sufficiently tight.

To validate our claim on resilience, we next corrupt signals $|V_6|$ and $|V_{27}|$ from set I, and $|V_7|$ and $|V_{22}|$ from set II.

TABLE XV:
ESTIMATION OF 0.25 Hz OSCILLATION FROM $\Delta|V_6|$ AND $\Delta|V_{27}|$ – WITH CORRUPTION AND AFTER RECOVERY USING SIGNAL SET I

Parameters	Prony on $\Delta\lvert V_6\rvert$			Prony on $\Delta\lvert V_{27}\rvert$		
	$y(t)$	$z(t)$	$\hat{y}(t)$	$y(t)$	$z(t)$	$\hat{y}(t)$
Damp. Ratio	0.026	0.069	0.026	0.023	0.078	0.025
Freq. (Hz)	0.250	0.216	0.251	0.251	0.211	0.251

TABLE XVI:
ESTIMATION OF 0.25 Hz OSCILLATION FROM $\Delta|V_7|$ AND $\Delta|V_{22}|$ – WITH CORRUPTION AND AFTER RECOVERY USING SIGNAL SET II

Parameters	Prony on $\Delta\lvert V_7\rvert$			Prony on $\Delta\lvert V_{22}\rvert$		
	$y(t)$	$z(t)$	$\hat{y}(t)$	$y(t)$	$z(t)$	$\hat{y}(t)$
Damp. ratio	0.025	0.113	0.027	0.024	0.104	0.023
Freq. (Hz)	0.251	0.214	0.251	0.251	0.212	0.251

In signals $|V_6|$ and $|V_7|$, we perform missing data attack as before, while in signals $|V_{22}|$ and $|V_{27}|$ we corrupt the actual signal value with random spurious outliers. Corruption is introduced in 200 consecutive samples starting at $t = 68$ s, as shown in Figs 8 (c)-(d) and 9 (c)-(d). Recovery is performed independently for each set with subspace derived from its constituent signals. It can be observed that the recovery is exact as the reconstructed signal tracks the original. The results of modal estimation using Prony analysis on a 30 s window (see Figs 8 and 9) of original, corrupted, and reconstructed signals are listed in Tables XV and XVI. It can be seen that the 0.25 Hz mode estimated from the recovered signals closely matches that of the original.

Further, to underscore the importance of our propositions in grouping signals, we arbitrarily form signal set III taking 4 signals each from sets I and II, as shown in Table XVII. As before, we corrupt signals $|V_7|$ and $|V_{22}|$. For this group, the denseness values obtained are: $\kappa_1(\mathbf{Y}) = 0.709$ and $\kappa_2(\mathbf{Y}) = 0.995$, each of which is greater than κ^*. Therefore, recovery of signals is not guaranteed. This is evident from the erroneous reconstruction plots in Figs 10 (a)-(b). The damping ratio and frequency of oscillation estimated from these reconstructed signals are: 0.092 and 0.210 Hz respectively, which are different from those estimated from the original signals in Table XVI.

TABLE XVII: SIGNAL SET III

Signal Set III	$\lvert V_{15}\rvert$, $\lvert V_{20}\rvert$, $\lvert V_{21}\rvert$, $\lvert V_{27}\rvert$, $\lvert V_4\rvert$, $\lvert V_{11}\rvert$, $\lvert V_7\rvert$, $\lvert V_{22}\rvert$

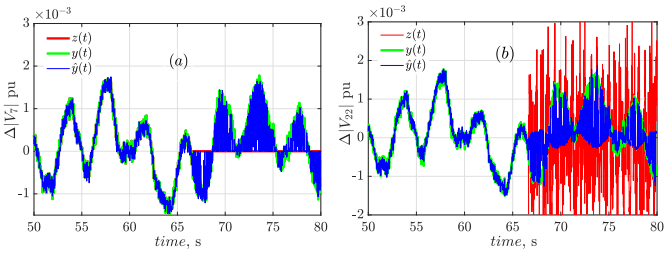

Fig. 10: Incorrect recovery of 2 corrupted signals from signal set III.

V. Conclusion

Insights were derived and recommendations were made for grouping signals to enhance denseness of a set – both

considering availability of small-signal model, and in absence thereof. It was shown that denseness of a set increased when signals were grouped with observabilities in same phase and variation in magnitudes minimized, for each poorly-damped mode. Further, for a signal set sufficiently dense, the denseness coefficient increases linearly with increase in relative magnitude of the largest observability. This implies signals that are conventionally preferred for oscillation monitoring might have to be sacrificed for guaranteed recovery of the signal set.

ACKNOWLEDGMENTS

The authors would like to thank New York Power Authority for providing the PMU data.

REFERENCES

[1] F. Aminifar, M. Fotuhi-Firuzabad, A. Safdarian, A. Davoudi, and M. Shahidehpour, "Synchrophasor Measurement Technology in Power Systems: Panorama and State-of-the-Art," *IEEE Access*, vol. 2, pp. 1607–1628, 2014.

[2] D. J. Trudnowski, "Estimating Electromechanical Mode Shape From Synchrophasor Measurements," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1188–1195, 2008.

[3] T. Huang, M. Wu, and L. Xie, "Prioritization of PMU Location and Signal Selection for Monitoring Critical Power System Oscillations," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 3919–3929, 2018.

[4] V. S. Perić, X. Bombois, and L. Vanfretti, "Optimal Signal Selection for Power System Ambient Mode Estimation Using a Prediction Error Criterion," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2621–2633, Jul. 2016.

[5] Y. Gu, T. Liu, D. Wang, X. Guan, and Z. Xu, "Bad Data Detection Method for Smart Grids based on Distributed State Estimation," in *2013 IEEE International Conference on Communications (ICC)*, Jun. 2013, pp. 4483–4487.

[6] A. S. Dobakhshari and A. M. Ranjbar, "A Wide-Area Scheme for Power System Fault Location Incorporating Bad Data Detection," *IEEE Trans. Power Del.*, vol. 30, no. 2, pp. 800–808, Apr. 2015.

[7] 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 Processing Magazine*, vol. 29, no. 5, pp. 106–115, Sep. 2012.

[8] T. A, S. Sihag, and K. Alnajjar, "Non-linear state recovery in power system under bad data and cyber attacks," *Journal of Modern Power Systems and Clean Energy*, vol. 7, no. 5, pp. 1071–1080, Sep. 2019.

[9] A. Ashok, M. Govindarasu, and V. Ajjarapu, "Online Detection of Stealthy False Data Injection Attacks in Power System State Estimation," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1636–1646, May 2018.

[10] Y. Hao, M. Wang, J. H. Chow, E. Farantatos, and M. Patel, "Modelless Data Quality Improvement of Streaming Synchrophasor Measurements by Exploiting the Low-Rank Hankel Structure," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6966–6977, Nov. 2018.

[11] P. Gao, R. Wang, M. Wang, and J. H. Chow, "Low-Rank Matrix Recovery From Noisy, Quantized, and Erroneous Measurements," *IEEE Trans. Signal Process.*, vol. 66, pp. 2918–2932, Jun. 2018.

[12] P. Gao, M. Wang, J. H. Chow, M. Berger, and L. M. Seversky, "Missing Data Recovery for High-Dimensional Signals With Nonlinear Low-Dimensional Structures," *IEEE Trans. Signal Process.*, vol. 65, no. 20, pp. 5421–5436, Oct. 2017.

[13] M. Liao, D. Shi, Z. Yu, Z. Yi, Z. Wang, and Y. Xiang, "An Alternating Direction Method of Multipliers Based Approach for PMU Data Recovery," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4554–4565, 2019.

[14] K. Mahapatra and N. R. Chaudhuri, "Malicious Corruption-Resilient Wide-Area Oscillation Monitoring Using Principal Component Pursuit," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1813–1825, 2019.

[15] ——, "Malicious Corruption-Resilient Wide-Area Oscillation Monitoring using Online Robust PCA," in *2018 IEEE Power Energy Society General Meeting (PESGM)*, Aug. 2018, pp. 1–5.

[16] ——, "Online Robust PCA for Malicious Attack-Resilience in Wide-Area Mode Metering Application," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 2598–2610, Jul. 2019.

[17] K. Chatterjee, K. Mahapatra, and N. R. Chaudhuri, "Robust Recovery of PMU Signals with Outlier Characterization and Stochastic Subspace Selection," *IEEE Trans. Smart Grid*, pp. 1–1, 2019.

[18] T. T. Cai, L. Wang, and G. Xu, "New bounds for restricted isometry constants," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4388–4394, Sep. 2010.

[19] P. Kundur, *Power System Stability and Control*. McGraw-Hill, 1994.

[20] E. J. Candes and M. B. Wakin, "An Introduction To Compressive Sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[21] E. J. Candes and B. Recht, "Exact low-rank matrix completion via convex optimization," in *2008 46th Annual Allerton Conference on Communication, Control, and Computing*, Sep. 2008, pp. 806–812.

[22] E. J. Candes, "The Restricted Isometry Property and Its Implications for Compressed Sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9, pp. 589–592, May 2008.

[23] H. Guo, C. Qiu, and N. Vaswani, "An Online Algorithm for Separating Sparse and Low-Dimensional Signal Sequences From Their Sum," *IEEE Trans. Signal Process.*, vol. 62, no. 16, pp. 4284–4297, Aug. 2014.

[24] T. T. Cai, L. Wang, and G. Xu, "Shifting inequality and recovery of sparse signals," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1300–1308, Mar. 2010.

[25] S. Foucart and M. Lai, "Sparsest solutions of underdetermined linear systems via l_q-minimization for $0 < q \leq 1$," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 395–407, 2009.

[26] C. Qiu, N. Vaswani, B. Lois, and L. Hogben, "Recursive Robust PCA or Recursive Sparse Recovery in Large but Structured Noise," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 5007–5039, Aug. 2014.

[27] D. J. Trudnowski, J. M. Johnson, and J. F. Hauer, "Making Prony Analysis More Accurate using Multiple Signals," *IEEE Trans. Power Syst.*, vol. 14, no. 1, pp. 226–231, Feb. 1999.

[28] N. R. Chaudhuri, "Wide-area Monitoring and Control of Future Smart Grids," PhD thesis, Imperial College, London, U.K., 2011. [Online]. Available: http://hdl.handle.net/10044/1/7026.

[29] L. Dosiek, N. Zhou, J. W. Pierre, Z. Huang, and D. J. Trudnowski, "Mode Shape Estimation Algorithms under Ambient Conditions: A Comparative Review," *IEEE Trans. on Power Syst.*, vol. 28, no. 2, pp. 779–787, May 2013.

[30] S. Lloyd, "Least squares quantization in pcm," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.

Kaustav Chatterjee (S'19) received the B.E. degrees in electrical engineering from Jadavpur University, India, in 2015, and the M.Tech. degree in electrical engineering from the Indian Institute of Technology Bombay, India, in 2018. He is currently pursuing the Ph.D. degree in electrical engineering with The Pennsylvania State University, USA. His research interests lie in the intersection of signal processing and system theory for real-time monitoring of large power systems and online estimation of system dynamics.

Nilanjan Ray Chaudhuri (S'08-M'09-SM'16) received the Ph.D. degree in power systems from Imperial College London, London, UK in 2011. From 2005 to 2007, he worked in General Electric (GE) John F. Welch Technology Center. He came back to GE and worked in GE Global Research Center, NY, USA as a Lead Engineer during 2011 to 2014. Presently, he is an Assistant Professor with the School of Electrical Engineering and Computer Science at Penn State, University Park, PA. He was an Assistant Professor with North Dakota State University, Fargo, ND, USA during 2014-2016. He is a member of the IEEE and IEEE PES. Dr. Ray Chaudhuri is the lead author of the book Multi-terminal Direct Current Grids: Modeling, Analysis, and Control (Wiley/IEEE Press, 2014). He served as an Associate Editor of the IEEE TRANSACTIONS ON POWER DELIVERY (2013 – 2019) and IEEE PES LETTERS (2016 - present). Dr. Ray Chaudhuri was the recipient of the National Science Foundation Early Faculty CAREER Award in 2016 and Joel and Ruth Spira Excellence in Teaching Award in 2019.

George Stefopoulos (S'98-M'08-SM'17) is the Director of the Advanced Grid Innovation Laboratory for Energy (AGILe) at the New York Power Authority. He has been with the New York Power Authority since 2009 holding positions as Research and Technology Development Engineer and Smart Grid Solution Architect. His research areas of interest include modeling and simulation of electric power systems, real-time simulation and hardware-in-the-loop testing, power system control and operation, synchrophasor applications, and computer application in electric power engineering. George is the author or co-author of over 60 journal and conference research papers as well as many technical reports. He is a senior member of the Institute of Electrical and Electronics Engineers (IEEE) and an associate member of Institute of Engineering Technology (IET). George received his Diploma in Electrical and Computer Engineering from the National Technical University of Athens, Greece in 2001 and he also holds Master's and Ph.D. degrees from the Georgia Institute of Technology (2002 and 2009 respectively). He also obtained a MBA degree in Executive Management from Pace University of New York in 2015.