Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

QuadNet: a database for complex network studies

Wendy Yadira Eras Herrera

wendy.herrera@ufop.edu.br

Universidade Federal de Ouro Preto

Leandro Freitas

Universidade Federal de Minas Gerais

Luis Antonio Aguirre

Universidade Federal de Minas Gerais

Research Article

Keywords: Nonlinear oscillators, synchronization, oscillator networks, complex systems, state estimation, Unscented Kalman filter

Posted Date: June 15th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9866484/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

QuadNet: a database for complex network studies

Wendy Y. Eras-Herrera [*], Leandro Freitas[†] and Luis Antonio Aguirre[†]

Departamento de Engenharia Elétrica, Instituto de Ciências Exatas e Aplicadas, Universidade Federal de Ouro Preto, João Monlevade, Minas Gerais, Brazil.

[†] Departamento de Engenharia Eletrônica, Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627, Belo Horizonte, 31270-901, Minas Gerais, Brazil.

*Corresponding author(s). E-mail(s): wendy.herrera@ufop.edu.br;
Contributing authors: leandrofabreu@ufmg.br; aguirre@ufmg.br;

Abstract

This work presents a benchmark dataset that can be used in studies related to networks, synchronization, state estimation, among others. Ten similar coupled quadrature oscillators are connected using the following topologies: ring, small world and scale-free. Several values of the coupling strength are used so that there are datasets with different synchronization behaviors, namely: asynchronous, phase-synchronized, and (almost) fully synchronized. The dataset is publicly available at (Freitas et al., 2025) where the simple case of two coupled oscillators is also offered. In order to illustrate a potential use this work addresses the state estimation problem where the network full state is estimated from a limited set of measurements using the Unscented Kalman Filter. Practical aspects as what, where and how much to measure is investigated using Monte Carlo simulations. Results show that to consider the node degree and the formation of clusters in deciding where to measure has a positive effect in estimation performance.

Keywords: Nonlinear oscillators; synchronization; oscillator networks; complex systems; state estimation; Unscented Kalman filter.

1 Introduction

Networks have become a paradigm for modeling complex systems and complex behavior in almost all fields of knowledge, from electrical power grids to biological systems (Milo et al., 2002; Feliks et al., 2013; Montanari et al., 2019; He et al., 2024; Li et al., 2024; Monteiro, 2026). Those systems are often modeled as connected dynamical systems

$$\dot{x}_i = f_i(x_i) + K \sum_{j=1}^{N} A_{ij}\, h(x_i, x_j), \tag{1}$$

where $f_i : \mathbb{R}^n \to \mathbb{R}^n$ determines the dynamics of the isolated system, the adjacency matrix $A = [A_{ij}] \in \mathbb{R}^{N \times N}$ defines the topology of connections, $K \in \mathbb{R}$ the coupling strength between the nodes, and $h : \mathbb{R}^n \to \mathbb{R}^n$ the coupling function between each pair of nodes. In this context, several important questions have been raised concerning challenges such as estimation and control (Liu et al., 2011, 2013; Haber, 2021; Montanari et al., 2022; Skardal and Arenas, 2023; Chen et al., 2025; Wu et al., 2026). Because not all nodes in a practical network are available for measurement, what to measure and where to measure naturally emerges as important questions (Montanari et al., 2022; Montanari and Aguirre, 2020). In particular, for the estimation of the network state, it is clear that the dynamics of the node, $f_i(x_i)$, the strength of the coupling K, the topology of the connection A_{ij}, and the coupling function $h(x_i, x_j)$ play a significant role in the expected performance (Montanari and Aguirre, 2020; Haber, 2021), although there is a lack of understanding of how each aspect influences the performance of the estimation.

In order to address such questions in a more systematic way and to enable a more direct comparison of algorithms and techniques, it is convenient to have a benchmark dataset. The purpose of this paper is to describe a public benchmark database (Freitas et al., 2025) and to illustrate its use to address problems related to state estimation. The database is based on connected quadrature oscillators (Bittencourt et al.,

2023, 2024; Dias et al., 2024; Eras-Herrera et al., 2025), which can be implemented electronically for real-world validation.

1.1 Nomenclature

The main acronyms used in the work are listed below.

Asyn: asynchronous or not synchronized;

ClS: cluster synchronization;

CS: complete synchronization;

GS: global synchronization;

MAPE: mean absolute percentual error;

MC: nodes with more connections;

PC: nodes with fewer connections;

PS: phase synchronization;

UKF: unscented Kalman filter.

The work is organized as follows. Section 2 presents the database used and how it was built from numerical simulations of electronic circuits. Section 3 proposes the challenge and presents the adopted approach to the state estimation problem, with a formal analysis of the observability of the electronic oscillator. Section 4 presents the results for two coupled oscillators. Section 5 presents the results for the small-world network, where the most influential aspects for state estimation are discussed. Section 6 synthesizes all findings and derives the main conclusions.

2 The Database

The freely available database (Freitas et al., 2025) currently consists of simulated data and will soon include experimental data. Here, details of the experimental oscillators and simulated equations are provided.

2.1 The quadrature oscillator

The equations that describe the dynamics of the quadrature oscillator are (Bittencourt et al., 2023):

$$\begin{aligned}
\dot{v}_{\text{sen}} &= \left(\frac{1}{\tau_2}\right) v_{\cos} + \left(\frac{1}{\tau_1} - \frac{1}{\tau_2} + \frac{1}{C_1 R_a}\right) v_1 + u_1, \\
\dot{v}_{\cos} &= \left(-\frac{1}{\tau_3}\right) v_{\text{sen}} - \frac{1}{C_3} I_l + u_2, \\
\dot{v}_1 &= \left(\frac{1}{\tau_2}\right) v_{\cos} - \frac{1}{\tau_2} v_1,
\end{aligned} \tag{2}$$

where $\tau_1 = R_1 C_1$, $\tau_2 = R_2 C_2$, $\tau_3 = R_3 C_3$ and the *coupling terms* are given by:

$$u_1 = \frac{R_f}{R R_a C_1} \sum_j v_{\text{sen}}^j \quad \text{e} \quad u_2 = \frac{R_f}{R R_a C_1} \sum_j v_{\cos}^j. \tag{3}$$

It is noticed that although the choice of a single coupling variable is very important (Braga and Aguirre, 2024), in this work for the sake of electronic implementation, the coupling is always accomplished through two variables as indicated in (2) and (3). The limiting current I_l is:

$$I_l = \begin{cases} \frac{1}{R_5} v_{\cos} + \frac{V_{cc}}{R_4}, & \text{if } v_{\cos} < -\frac{R_5}{R_4} V_{cc} - \left(1 + \frac{R_5}{R_4}\right) V_d, \\ \frac{1}{R_5} v_{\cos} - \frac{V_{cc}}{R_4}, & \text{if } v_{\cos} > \frac{R_5}{R_4} V_{cc} + \left(1 + \frac{R_5}{R_4}\right) V_d, \\ 0, & \text{otherwise}, \end{cases} \tag{4}$$

where the parameters are given in Table 1. Time dependence is omitted to simplify notation.

In this work, the state vector at time k for the ith oscillator is indicated by

$$[x_{j,k}^i]^T \coloneqq \begin{bmatrix} v_{\text{sen}}^i & v_{\cos}^i & v_1^i \end{bmatrix}^T, \tag{5}$$

where $x_{j,k}^i$ indicates the jth state variable of the ith oscillator at the kth time instant.

2.2 The database used

The database was obtained using different network topologies in different dynamical regimes as described in this section, where all data are freely available online (Freitas et al., 2025). In order to quantify the dynamical regime (i.e. level of synchronization) of the network, the Kuramoto order parameter r was used (Kuramoto, 1975):

$$re^{j\psi(t)} = \frac{1}{N}\sum_{k=1}^{N} e^{j\theta^k(t)}, \tag{6}$$

where θ^k is the phase of the kth oscillator. The values $r \approx 1$ indicate identical synchronization, and very small values of r indicate lack of synchronization.

The simulated data are provided for two coupled oscillators and ten oscillators connected in ring, small world, and scale-free configurations (see Tab. 2). In all cases small differences in the main frequency were used, coupling is always bidirectional with used-defined coupling force. As that force is increased, cluster and close to global identical synchronization is observed.

The networks were simulated using a fourth-order Runge-Kutta numerical integration with an integration step and a sampling time $h = 0.005\,\mathrm{s}$. The initial conditions were taken randomly around the origin, the simulation time is 120 s, and the first 60 s were discarded.

Table 1: Circuit parameters. Source: Freitas et al. (2025).

Component	Value	Unit
R_f	$[1, 100]$	kΩ
R, R_a, R_b	100	kΩ
R_1	10	kΩ
R_2	$[10, 100]$	kΩ
R_3	12	kΩ
R_4	22	kΩ
R_5	4.7	kΩ
C_1, C_2, C_3	2.2	μF

All data and other information such as the coupling strength (R_f), pair-wise order parameter r, parameter values, and oscillator frequencies are available online at (Freitas et al., 2025) and can be used in a number of problems related to synchronization (Braga and Aguirre, 2024), control (Zhang, 2023), or estimation (de Paula et al., 2026). Having a benchmark to test and compare methods is methodologically important. In the reminder of this work the use of the data set is illustrated in the context of nonlinear state estimation.

In this work, the datasets `exp0001` and `exp0003` were used, respectively corresponding to two oscillators (Fig. 1a) and Small World network (Fig. 1b) with ten non-identical oscillators, with natural frequencies ranging from 3.80 to 4.00 Hz (Two oscillators network) and from 3.80 to 3.85 Hz (Small World network).

3 State Estimation

3.1 The Challenge

From a set of data, the challenge is to estimate the full state of the network using the smallest number of measurements. Performance will be assessed in view of topology, dynamical regime, and what measurement is used.

Table 2: A sample of datasets in *QuadNet Estimation Challenge*. Source: (Freitas et al., 2025).

·	Label	Nodes	Topology	Regime
×	`exp001a`	2	two osc.	no sync.
×	`exp001b`	2	two osc.	phase sync.
×	`exp001c`	2	two osc.	complete sync.
	`exp002a`	10	ring	no sync.
	`exp002b`	10	ring	cluster sync.
	`exp002c`	10	ring	global sync.
×	`exp003a`	10	small world	no sync.
×	`exp003b`	10	small world	global sync.
×	`exp003c`	10	small world	cluster sync.
	`exp004a`	10	scale-free	no sync.
	`exp004b`	10	scale-free	cluster sync.
	`exp004c`	10	scale-free	cluster sync.



(a) Two Oscillators (b) Small World (SW)

Figure 1: Networks topologies considered.

The problem is stated as follows. It is assumed that the system is described by

$$x_k = f(x_{k-1}, u_{k-1}, w_{k-1}, k-1), \tag{7}$$

$$y_k = h(x_k, \nu_k, k), \tag{8}$$

where $x_k \in \mathbb{R}^n$ is the state vector, $f : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{N} \to \mathbb{R}^n$ is the system model and $h : \mathbb{R}^n \times \mathbb{R}^r \times \mathbb{N} \to \mathbb{R}^m$ is the measuring function. The inputs are $u_k \in \mathbb{R}^p$ and the measured vector is $y_k \in \mathbb{R}^m$. It is assumed that the process $w_k \in \mathbb{R}^q$ and the measurement noise $\nu_k \in \mathbb{R}^r$ are white random variables and mutually independent of the covariance matrices Q_{k-1} and R_k, respectively.

From a Bayesian perspective, the aim of the state estimation problem is to find a posterior probability density function $J(x_k) = \rho(x_k|y_{1:k})$ from the system model, noisy measurements $y_{1:k} = \{y_1, \ldots, y_k\}$, the driving signal, and the initial estimate. The problem is solved in two steps: *propagation* and *assimilation* (Bar-Shalom Y., Li X. R., Kirubarajan T., 2001).

The solution of the state estimation problem for nonlinear systems in the form of (7) and (8) considers that the marginal posterior PDF $p(x_k|y_{1:k})$ cannot be completely characterized by its mean $\hat{x}_k$ and covariance P_k^{xx}, because it is not Gaussian in general (Daum F., 2005). To avoid this problem, suboptimal algorithms are used, such as the

Extended Kalman Filter (EKF) (Maybeck, P. S., 1979) and the Unscented Kalman Filter (UKF) (Julier, S. J. and Uhlmann, J. K. and Durrant-Whyte, H. F., 2000), where a small number of points, called sigma points, are chosen in a deterministic way to numerically approximate the mean and covariance of the PDF (Ristic, B., Arulampalam, S., Gordon, N., 2004). In this work, state estimation is achieved using the UKF algorithm.

3.2 Dynamical Observability

Observability is the property that enables the full and unambiguous reconstruction of the state from the measurements. This concept is due to Kalman who proposed criteria to classify a system either as observable or non observable. This is called structural observability. A common scenario in practice occurs when a system is observable from two different sets of measurements; however, from one set the ease with which one is able to estimate the state is greater. Hence, *dynamical observability* aims at quantifying practical observability in the context of *observable systems*. Several observability coefficients have been proposed to quantify *dynamical observability* and rank measured variables in terms of how effective they are in discriminating among two similar states. For a general discussion of this, see (Aguirre et al., 2018) and the references therein.

In short, one computes the nonlinear observability matrix using Lie derivatives and evaluates some measure of numerical conditioning averaged along a trajectory. For the quadrature oscillator, we get: $\delta_{\sin} = 1.09 \times 10^{-7}$, $\delta_{\cos} = 9.63 \times 10^{-8}$, and $\delta_{v1} = 1.33 \times 10^{-7}$. In view of such results, the following remarks are in order.

Remark 1. *The oscillator is observable from any of the measured variables as the observability matrix never vanishes along the trajectory.*

Remark 2. *Other electronic oscillators, such as Chua's circuit, also have rather small absolute values of the observability coefficients. This is a consequence of poor overall conditioning due to the wide range of values in which the parameters are found, e.g.*

from μF to $k\Omega$. Nevertheless, the overall average condition number varies in the range $2.4k \leq \kappa \leq 2.9k$ which is considered workable. Hence, the oscillator is sufficiently observable from any variable to enable state estimation.

Remark 3. *The observability coefficients are similar, suggesting that there is no a priori reason to prefer one variable to the other.*

3.3 Performance Metric

The performance of the estimators was evaluated using the Mean Absolute Percentage Error (MAPE):

$$\text{MAPE}_j^i = \sum_{k=k_\text{i}}^{k_\text{f}} \frac{|x_{j,k}^i - \hat{x}_{j,k}^i|}{|\max(x_j^i)|} \times 100, \tag{9}$$

where k_i and k_f are the first and last discrete instants considered. k_i it must be greater than the transient period, so the transient period is not used. Therefore, $x_{j,k}^i$ indicates the j-th state variable of the i-th oscillator at time k. The hat indicates estimated quantities.

4 Choice of Observables

Here, the challenge is to estimate the state of a system of two non-identical coupled oscillators from measurements taken only from one of them and using the UKF approach, the process model (2)-(4) is the nonlinear system (7) while the observation model (8) is linear with output matrix $C_k \in \mathbb{R}^{m \times n}$ given by

$$C_k = \begin{bmatrix} I_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} \end{bmatrix}, \tag{10}$$

in which only the states of oscillator 1 are measured.

The noise of the process Q_k was chosen empirically, and the measurement noise R_k was tuned from the information provided by the instrumentation specifications. It is assumed that the process noise, $w_k \sim \mathcal{N}(0, Q_k)$, is a zero-mean white noise with constant covariance $Q_k = 10^{-4} I_{n\times n}$ and the observation noise, $v_k \sim \mathcal{N}(0,\, R_k)$, is a zero-mean white noise with constant covariance $R_k = 10^{-3} I_{n\times n}$.

The initial state vector was $x_{0|0} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T$ and the covariance matrix $P^{xx}_{0|0} = \text{diag}(10 \cdot I_{n\times n})$. It is important to note that, in the dataset used in the simulation, the transient period corresponding to the first 5 s was discarded.

4.1 State estimation for two coupled oscillator

To perform the simulation, the process model (2)-(4) and the observation model (10) are considered, along with the values of all parameters available in the file `exp001a.mat` of the dataset `exp001`.

Table 3 shows the MAPE index in the following regimes: asynchronous (Asyn) with $R_f = 10\,\text{k}\Omega$, phase synchronization (PS) with $R_f = 20\,\text{k}\Omega$ and complete synchronization (CS) with $R_f = 70\,\text{k}\Omega$. As expected, the estimation error is greater for the oscillator 2 which is not measured.

To evaluate the performance of state estimation, the number of measurements used was varied. All measurements were performed on oscillator 1. Table 3 shows the MAPE index (9) of the state estimates provided by the UKF approach. In general, a greater error is observed in the state estimates $\hat{x}^i_j$ $j = 4, 5, 6$ when only the oscillator 1 is measured. Specifically, for Asyn, measuring two variables instead of three results in an error increase of approximately 1% to 3%, with the pair $y = [v^1_{\text{cos}}\ v^1_1]^T$ resulting in the smallest errors. When measuring only one variable, errors increase, as expected. In this case, the smallest errors occurred when measuring $y = v^1_{\text{cos}}$. In this case, the errors increased only by about 3% compared to the case where all the oscillator variables 1

Table 3: Index MAPE (9). Measures are taken only from oscillator 1 (first column).

Measurement vector y	MAPE Asyn	MAPE PS	MAPE CS
Data file	`exp001a`	`exp001b`	`exp001c`
$y = [v^1_{\text{sen}}\ v^1_{\text{cos}}\ v^1_1]^T$	0.0929	0.6176	0.5113
	0.2174	0.6150	0.5020
	0.1265	1.0754	0.9191
	10.8083	5.6573	4.1581
	11.0726	5.6810	4.2053
	10.6658	5.4731	4.0869
$y = [v^1_{\text{sen}}\ v^1_{\text{cos}}]^T$	0.9441	0.6781	0.5545
	0.2221	0.6234	0.5061
	0.1942	0.7180	0.6028
	13.3435	6.8536	4.1985
	13.7467	6.9431	4.2453
	13.2501	6.7040	4.1184
$y = [v^1_{\text{sen}}\ v^1_1]^T$	0.8153	0.6604	0.5354
	0.5419	0.7343	0.5619
	0.2041	1.1570	0.9817
	12.7985	6.1395	4.1152
	13.1512	6.2003	4.1804
	12.7475	6.0196	4.0860
$y = [v^1_{\text{cos}}\ v^1_1]^T$	0.2793	0.8722	0.7513
	0.2621	0.7634	0.6294
	0.1304	1.1080	0.9587
	11.8421	5.6201	4.2808
	12.1266	5.6884	4.2852
	11.6828	5.4407	4.1036
$y = v^1_{\text{sen}}$	0.1640	0.7327	0.5956
	0.6062	0.7584	0.5779
	0.3643	0.8713	0.7148
	17.5945	*6.3668*	**4.2725**
	18.1897	*6.4744*	**4.3359**
	17.6465	*6.2842*	**4.2477**
$y = v^1_{\text{cos}}$	0.3700	1.0868	0.9246
	0.2756	0.8408	0.7004
	0.2466	0.6193	0.5066
	13.7024	*6.8532*	**4.1443**
	14.0570	*6.8860*	**4.1561**
	13.6030	*6.6514*	**4.0246**
$y = v^1_1$	0.8776	1.0426	0.8110
	1.0502	1.0562	0.8990
	0.3569	1.2880	1.1009
	32.8389	*6.2008*	**4.3930**
	34.1314	*6.2909*	**4.4613**
	33.0916	*6.1314*	**4.3150**

were measured. Electronically, the signal v^1_{cos} is the one most affected by nonlinearity (voltage clamping) (Bittencourt et al., 2023).

The quality of the estimates of oscillator 2 state variables improves when the regime becomes phase-synchronized, and improves even further in the case of full synchronization. This can be understood as follows: these dynamic regimes occur for the most strongly coupled oscillators, and therefore, the information flow is more intense, facilitating estimation. Furthermore, the dynamic behavior of the oscillators becomes more similar as the coupling is stronger. In PS, the oscillation frequencies are equal, but there are fluctuations in the phase difference; whereas in CS, both the frequencies and phases of the oscillators are essentially equal.

Another aspect that deserves attention is noting that the dependence on what is measured is much less relevant in phase synchronization and full synchronization regimes. This should not be considered a general result, as is argued next. The study of dynamical observability in Sec. 3.2 was carried out *for one* oscillator, hence the dimension of the state space is three. In this section, we investigate a *different* system, which is the result of coupling two oscillators, hence yielding a six-dimensional state space and a new observability analysis would have to be accomplished. In that respect, it is not contradictory that measuring $v_{\cos}$ is slightly better than measuring $v_{\sin}$ and that measuring v_1 is significantly worse (Tab. 3). As synchronization improves from Asyn to PS, and further to CS, the six-dimensional state space becomes increasingly similar to the original three dimensional space. In fact, this remark has recently been used to propose a measure of synchronization (Castro and Aguirre, 2025). As a consequence, in PS and CS one may assume that the results in Sec. 3.2 become valid. In fact, the MAPE values are quite similar regardless of the measured variable (see italic and underlined values in Tab. 3).

The quality of the estimates of the state of oscillator 1 worsens as the coupling strength of the oscillators increases. Since the coupling is bidirectional, oscillator 1 is increasingly influenced by oscillator 2 as the coupling strength increases, but this

alone does not explain the worsening of the estimation. This behavior deserves a more detailed study.

5 Sensor Allocation in a Small-World Network

The nonlinear network model (see Eq. 7) was used together with the linear observation model $y_k = C_k x_k + \nu_k$, with:

$$C_k = \text{diag}\left(\left[\begin{array}{ccc} C^0_{3\times 3} & \cdots & C^9_{3\times 3} \end{array}\right]\right) \in \mathbb{R}^{m\times n}, \tag{11}$$

where $C^i_{3\times 3}$ contains the three voltages of the i-th oscillator in the main diagonal.

In each experiment, only the voltages of five oscillators, referred to as "measured oscillators," were used. Twenty Monte Carlo simulations were performed in which the measured oscillators were chosen randomly, following the rules defined for each experiment, as detailed in the sequence.

5.1 State estimation

The following synchronization regimes will be used in the sequel: asynchronous (Asyn) with $R_f = 1\,\text{k}\Omega$, cluster synchronization (ClS) with $R_f = 5\,\text{k}\Omega$ and global synchronization (GS) with $R_f = 9\,\text{k}\Omega$. In ClS the oscillators within a cluster are synchronized with high value of the order parameter (6), but are not synchronized with oscillators outside the cluster. In GS all the oscillators in the network are well synchronized to each other.

5.1.1 `ExpRand`: random choice of measured oscillators.

Figures 2 and Table 4 show the performance of the state estimation in 20 Monte Carlo simulations in which the measured oscillators were randomly chosen. This was done in three dynamical regimes: not synchronized ($R_f = 1$ kΩ), cluster synchronization

($R_f = 5$ kΩ) and global synchronization ($R_f = 9$ kΩ). The variance of the distributions is rather large.

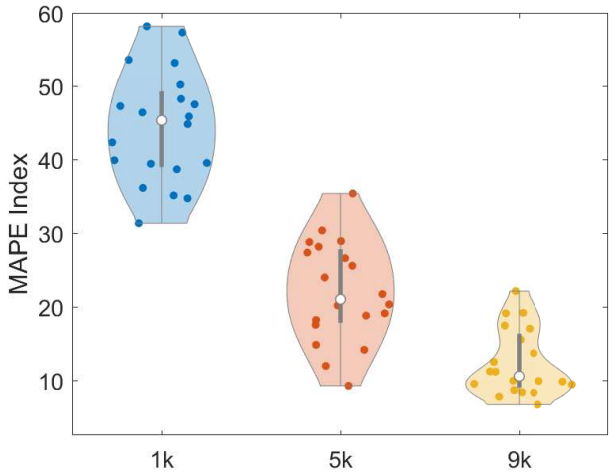


Figure 2: Estimation performance for a random choice of the measured oscillators along 20 Monte Carlo Simulations. From the left: not synchronized, cluster synchronization and generalized synchronization for the small world network.

Table 4: Data corresponding to Figure 2. Index MAPE (9) – Values from the boxplot (first quartile (Q1), second quartile (Q2=median), third quartile (Q3) and the Interquartile Range (IQR) of the experiment `ExpRand` for the Small World network.

	Asyn ($R_f = 1$ kΩ)	ClS ($R_f = 5$ kΩ)	GS ($R_f = 9$ kΩ)
Q3	49.20	27.77	16.30
Q2 (median)	45.33	21.04	10.58
Q1	39.04	17.89	9.08
IQR=Q3-Q1	10.16	9.88	7.21

5.1.2 `ExpFreq`: choice based on the diversity of natural frequencies.

To evaluate whether measuring oscillators that configure a wider range of natural frequencies improves state estimation, two experiments were defined: `ExpFreqC1`, less diverse; and `ExpFreqC2`, which covers a wider range of natural frequencies.

In group `ExpFreqC1`, oscillators exceeding 0.5% of the average natural frequency of the network were always excluded from the measurement. Group `ExpFreqC2` always includes nodes with the most diverse natural frequencies, which are the oscillators $(0, 8, 9)$. The remaining measured oscillators, to complete the five, are chosen randomly in the 20 Monte Carlo runs (see Tab. 5).

Table 5: Natural frequency diversity based experiment.

Name	Rules
`ExpFreqC1`	do not include osc. 0, 8 & 9
`ExpFreqC2`	measure 0, 8 & 9 plus other two chosen randomly

To simplify the captions of the figures, `ExpFreqC1` and `ExpFreqC2` are denoted as C1 and C2, respectively, in ClS (left) and GS (right). The graph scales have been fixed in all figures to facilitate comparison.

From Fig. 3 and Tab. 9, it can be seen that the estimation performance for GS is always better than for ClS. This result is true for all experiments in this work and is in agreement with previous investigations (Eras-Herrera et al., 2025).

In ClS, estimation for C1 is better than for C2, with less dispersion and a lower median. The opposite seems to be the case in GS. This suggests that measuring nodes with greater variance of natural frequencies with respect to the population average does not generally lead to improved state estimation.

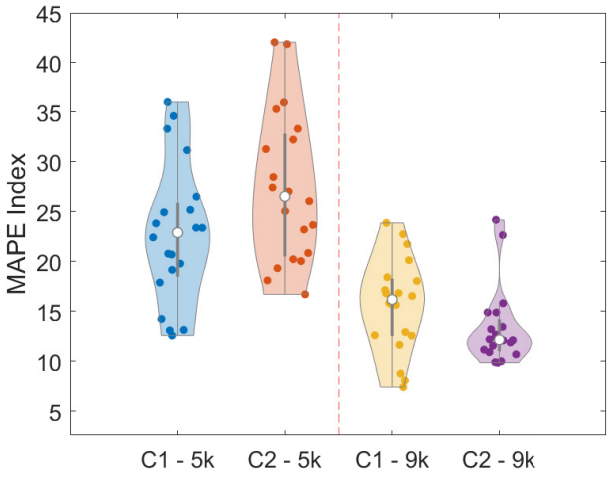


Figure 3: Estimation performance in `ExpFreq`. ClS (left of the red dotted line), GS (right). For the difference between C1 and C2 see Tab. 5.

Table 6: Topology based experiment.

Name	Rules
`ExpTopoC1`	nodes with more connections (MC) (0, 6, 8) and plus other two chosen randomly
`ExpTopoC2`	nodes with fewer connections (PC) (1, 2, 3) and plus other two chosen randomly

5.1.3 `ExpTopo`: choice based on network topology.

To evaluate the impact of a choice based solely on topology, the experiment was subdivided into two scenarios. In the Small World network, the influence of the number of connections (degree) of the measured nodes was investigated, where `ExpTopoC1` denotes the estimation based on nodes with more connections (MC), which always include nodes $(0, 6, 8)$. The experiment `ExpTopoC2` always includes nodes with fewer connections (PC), which are $(1, 2, 3)$. To complete the five, the remaining measured nodes are randomly selected from among the intermediate nodes, excluding the PC and MC nodes (see Tab. 6).

For simplicity, in the figure captions, `ExpTopoC1` and `ExpTopoC2` are denoted, respectively, as MC (more connections) and PC (fewer connections) for the Small World network, in *cluster* synchronization (left side) and global synchronization (right side).

Figure 4 presents the result using the `ExpTopo` experiment on the Small World network. It can be observed that MC is marginally better than PC in the cluster synchronization regime and PC is better in the global synchronization regime (Table 10). The result suggests that, in strongly synchronized networks, it is interesting to measure nodes with fewer connections.

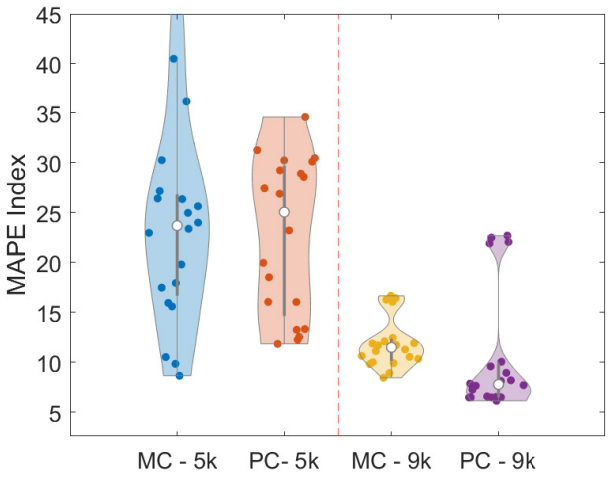


Figure 4: Estimation performance in `ExpTopo`. ClS (left of the red dotted line), GS (right). For the difference between MC and PC see Tab. 10.

5.1.4 `Exp<Freq/Topo>-Clus`: cluster information in addition to `ExpFreq` and `ExpTopo`.

In this experiment, the goal is to ensure the measurement of at least one oscillator from each cluster. This determination of which clusters are present is done *a posteriori*, after the simulation. The considered clusters are presented in Tab. 7 and were

obtained based on the matrix of order parameters calculated pairwise available in Freitas et al. (2025). The objective is to verify the effect of ensuring the measurement of the maximum possible clusters, in addition to the criteria based on the natural frequency (`ExpFreq`) and topology (`ExpTopo`).

Table 7: Oscillator clusters for the Small World network. Source: (Freitas et al., 2025).

#	ClS	GS
1	7, 9, 3, 0 6, 8, 4, 5	3, 1, 2
2	1	6, 7, 0, 9
3	2	8, 4, 5
4		
5		

`ExpFreq-ClusC1`: For the Small World network, oscillators $(0, 8, 9)$ were excluded (Tab. 5). Note that only three clusters were identified for this network (Tab. 7). Therefore, oscillators 1 and 2 were measured, and three more were randomly selected from $\{3, 4, 5, 6, 7\}$, as shown in Tab. 8.

`ExpFreq-ClusC2`: For the Small World network, oscillators $(0, 8, 9)$, which are in the same cluster, were measured, plus oscillators 1 and 2, resulting in the measurement of at least one from each of the three clusters (Tab. 8).

The experiments `ExpTopo-ClusC1` and `ExpTopo-ClusC2` follow the same logic, which is summarized in Tables 6 and 8, and the results are shown in Tab. 10.

Experiments `ExpFreq-Clus`. In Fig. 5 and Tab. 9, it can be observed that C2 is better than C1 except for the *cluster* regime for the Small World network. The same was verified in Fig. 3. The results in Tab. 9 suggest that, when the measurement of at least one oscillator in each cluster is guaranteed, a greater diversity of natural frequencies improves the estimation in both dynamic regimes, even if the gain is marginal.

The `ExpTopo-Clus` experiments, in the Small World network, are presented in Fig. 6 and Tab. 10. It is observed that PC is better than MC in both dynamic regimes,

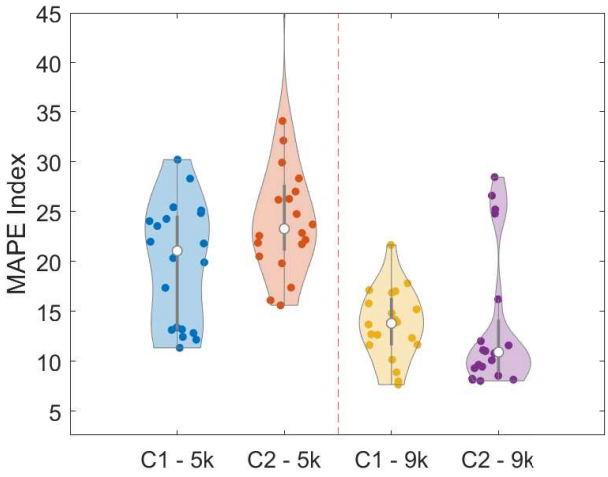


Figure 5: Estimation performance in `ExpFreq-Clus`. ClS (left of the red dotted line), GS (right). For the difference between C1 and C2 see Tab. 9.

Table 8: `Exp<Freq/Topo>-Clus` experiments. The brackets indicate that there was a random selection among the elements of the respective group to complete the measured oscillators.

Name	ClS	GS
`ExpFreq-ClusC1`	{3, 4, 5, 6, 7}, 1, 2	{1, 2, 3}, {6, 7}, {4, 5}
`ExpFreq-ClusC2`	0, 8, 9, 1, 2	{1, 2, 3}, 0, 9, 8
`ExpTopo-ClusC1`	0, 6, 8, {4, 5, 7, 9}	0, 6, 8, {7, 9}, {4, 5}
`ExpTopo-ClusC2`	1, 2, 3, {4, 5, 7, 9}	1, 2, 3, {7, 9}, {4, 5}

which is considered counter-intuitive, as well as the fact that (see Tab. 10) in the *cluster* regime the gain when measuring at least one oscillator from each cluster is significant when considering nodes with fewer connections.

5.2 Size effect analysis of the allocation rules

In this section a statistical analysis of the results is provided. The main goal is to quantify the effect size of each sensor allocation rule studied in the presented Small

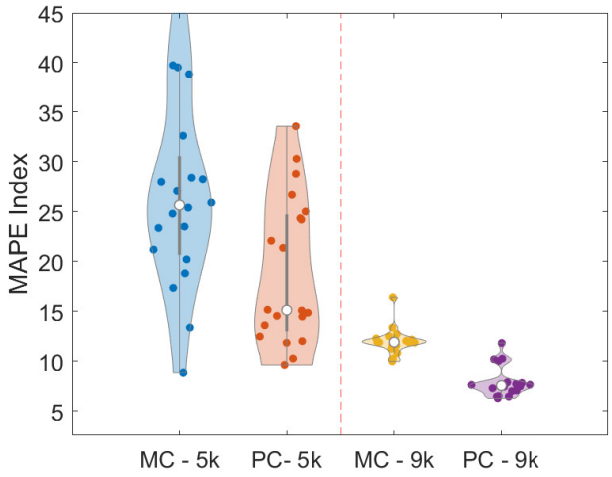


Figure 6: Estimation performance in `ExpTopo-Clus`. ClS (left of the red dotted line), GS (right). For the difference between MC and PC see Tab. 10.

Table 9: Index MAPE (9). See caption of Table 4 for details. Experiments `ExpFreq` and `ExpFreq-Clus` for the Small World network.

ClS	`ExpFreqC1`	`ExpFreqC2`	`ExpFreq-ClusC1`	`ExpFreq-ClusC2`
Q3	25.83	32.77	24.55	27.66
Q2 (median)	22.89	26.52	21.07	23.28
Q1	18.51	20.53	13.16	21.11
IQR=Q3-Q1	7.32	12.24	11.39	6.55
GS				
Q3	18.22	14.15	16.32	14.10
Q2 (median)	16.17	12.14	13.79	10.90
Q1	12.57	11.03	11.63	8.92
IQR=Q3-Q1	5.64	3.12	4.69	5.19

World network (Fig. 1b). To compare two allocation rules, a Wilcoxon test (Wilcoxon, 1945) is performed to investigate if both performance results (MAPE) come from a different probability distribution, with a significance level of 80% (p-value $p = 0.20$) chosen for convenience. In this study, the Wilcoxon test is adequate because it does not assume any particular distribution of the MAPE indices, that are positive and probably non-Gaussian.

Table 10: Index MAPE (9). See caption of Table 4 for details. Experiments `ExpTopo` and `ExpTopo-Clus` for the Small World network.

ClS	ExpTopoC1	ExpTopoC2	ExpTopo-ClusC1	ExpTopo-ClusC2
Q3	26.78	29.66	30.50	24.67
Q2 (median)	23.68	25.05	25.65	**15.12**
Q1	16.70	14.67	20.69	13.03
IQR=Q3-Q1	10.08	14.99	9.80	11.64
GS				
Q3	12.26	9.80	12.22	7.86
Q2 (median)	11.48	7.77	11.90	**7.56**
Q1	10.16	6.49	11.59	7.02
IQR=Q3-Q1	2.09	3.31	0.62	0.83

The allocation rules that provided a significant improvement, measured by the Wilcoxon test, are further analyzed to quantify the size effect, assessed by the biserial correlation coefficient (Pearson, 1910), γ, given by

$$\gamma = \frac{Z}{\sqrt{N}}, \tag{12}$$

where Z is the Z-statistic and N, the sample size. As a general rule, we can interpret $0.1 < \gamma < 0.3$ as small, $0.3 < \gamma < 0.5$ medium, and $\gamma \geq 0.5$ large effect size (Cohen, 1977).

To illustrate the use and interpretation of statistical tests, Table 11 compares the results of uniformly random sensor allocation among all nodes in the Small World network, with different coupling strength values (R_f), previously presented in Table 4 and Figure 2. The small p-values in Table 11 indicate that the results are significantly different from each other with high confidence. The effect size value γ (12) shows that the increase from $R_f = 1$ kΩ to $R_f = 5$ kΩ, in the first line with $\gamma = 0.8765$, has a larger effect on performance than the increase from $R_f = 5$ kΩ to $R_f = 9$ kΩ, with $\gamma = 0.8515$. It should be noted that, in both cases, the increase in coupling strength are equal, of +4 kΩ, but the effect on improving performance are not, as shown in the

first line of Table 11. In the following, these statistics are used to compare effect sizes on each allocation rule adopted in sensor allocation in the network.

Table 11: Effect size of MAPE (9) from randomized.

Experiment	p**-value**	**Effect size** (12)	**Interpretation**
$R_f = 1$ kΩ vs $R_f = 5$ kΩ	1.9073×10^{-6}	0.8765	Large effect
$R_f = 5$ kΩ vs $R_f = 9$ kΩ	9.5367×10^{-6}	0.8515	Large effect

To assess the improvement of a sensor allocation rule, each one is compared to the baseline experiment, with no rule, that is, the completely random sensor allocation through the network in 20 Monte Carlo runs. The individual result of this baseline experiment was presented in Table 4 and Figure 2. The allocation rules analyzed were:

- `+diversity`: increasing the natural frequency diversity (Sec. 5.1.2);
- `-diversity`: reducing the natural frequency diversity (Sec. 5.1.2);
- `+degree`: measuring nodes with more connections in the network (Sec. 5.1.3);
- `-degree`: measuring nodes with less connections in the network (Sec. 5.1.3);
- `+diversity-clus`: enhancing the natural frequency diversity, but ensuring the measurement of at least one node per cluster (Sec. 5.1.4);
- `-diversity-clus`: reducing the natural frequency diversity but ensuring the measurement of at least one node per cluster (Sec. 5.1.4);
- `+degree-clus`: measuring nodes with more connections in the network but ensuring the measurement of at least one node per cluster (Sec. 5.1.4);
- `-degree-clus`: measuring nodes with fewer connections in the network but ensuring the measurement of at least one node per cluster (Sec. 5.1.4).

Table 12 shows the results of experiments for which $p < 0.20$ at the confidence level of 80%; and the effect size γ (12) was medium or large. The mean improvement of each rule compared to the baseline is colored green (red) for those that reduced (increased) the average MAPE index.

In terms of the diversity of natural frequencies, the results in Table 12 show that this rule usually degrades the estimation performance by increasing the diversity (with $R_f = 5\,\mathrm{k}\Omega$) or decreasing it (with $R_f = 9\,\mathrm{k}\Omega$). Hence, the recommendation is not to use diversity as a single rule.

As for the node degree, the results suggest that to measure nodes with smaller degree – either with the added rule of measuring at least one oscillator of each cluster or not – is advantageous.

Table 12: Effect size of MAPE (9) comparing different sensor allocation rules with the completely randomized (no rule) in small-world networks.

Experiment	R_f	**Improvement**	p**-value**	**Effect size** (12)	**Interpretation**
`+diversity`	5 kΩ	-24 %	$\mathbf{1.0689 \times 10^{-2}}$	**0.5593**	**large effect**
`-diversity-clus`	5 kΩ	+10 %	1.2309×10^{-1}	0.3506	medium effect
`+degree-clus`	5 kΩ	-20 %	1.5365×10^{-1}	0.3256	medium effect
`-degree-clus`	5 kΩ	+14 %	7.5851×10^{-2}	0.4007	medium effect
`-diversity`	9 kΩ	-26 %	$\mathbf{7.2956 \times 10^{-3}}$	**0.5843**	**large effect**
`-degree`	9 kΩ	+16 %	1.6496×10^{-1}	0.3172	medium effect
`-degree-clus`	9 kΩ	+36 %	$\mathbf{6.2943 \times 10^{-5}}$	**0.8014**	**large effect**

Among all experiments in the small-world network, the results indicate that measuring nodes with lower degree in the network can provide benefits on state estimation of the entire network, for weak or strongly connected networks. When the measured nodes come from different synchronized clusters, the effect is enhanced from a medium to a large effect on the MAPE performance metric of the estimated states.

6 Conclusion

This work presents a simulation dataset that can be used as a benchmark in state estimation studies for networked systems. The data were generated by numerical simulation of networked quadrature oscillators using different network topologies in different dynamical regimes. In this work, the challenge is to estimate the full state

of the network using a limited number of measurements from a set of data available online (Freitas et al., 2025).

For the case of two coupled oscillators, it was found that the estimation error is greater for the oscillator that is not measured. The MAPE index for that oscillator varies from approximately 4% in synchronized regimes to 33% in non-synchronized ones. These results, combined with the analysis of the observability of the electronic oscillators, suggest that the dependence on what is measured is much less relevant in PS and CS regimes.

In the state estimation problem of ten non-identical oscillators connected in a Small World network, in the ClS and GS dynamic regimes, based on the measurements of five oscillators using the UKF, the aim was to investigate which factors are most relevant for selecting the oscillators to be measured and to quantify the impact of the resulting choices on the quality of state estimation.

For the factor based on natural frequency diversity, the results show that the quality of state estimation worsens when measuring a more diverse group of sensor nodes for low coupling strength (5 kΩ), while the inverse effect is observed for high coupling strength (9 kΩ), where the quality of state estimation worsens when measuring a low diversity group of sensor nodes. However, by ensuring the measurement of every cluster in each case of the `diversity-clus` experiments improves the quality of state estimation when measuring less diverse oscillators in the weak coupling (5 kΩ) case. No clear evidence was found that suggests that it is advantageous to take into account the diversity of frequencies in the choice of measured oscillators.

For the factor based on the network topology, the results show that taking into account the node degree of the measured oscillators is relevant. In particular, choosing nodes with smaller degrees tends to worsen the estimation performance, whereas choosing nodes with larger degrees improves the estimation. In addition, ensuring the

measurement of every cluster in each case of the `degree-clus` experiments further improves the estimation.

Some topics for future research include investigating the influence of the coupling strength; investigating whether the results hold across different network scales and topologies, including fully-connected, random, ring, and scale-free configurations; and investigating the influence of network motifs (Milo et al., 2002) on state estimation to determine their applicability in large-scale networks.

References

Aguirre, L.A., Portes, L.L., Letellier, C.: Structural, dynamical and symbolic observability: From dynamical systems to networks. Plos ONE **13**(e0206180) (2018) https://doi.org/10.1371/journal.pone.0206180

Braga, P.A.S., Aguirre, L.A.: Selecting the coupling variable to synchronize nonlinear oscillators. Nonlinear Dynamics **112**(17), 15177–15191 (2024) https://doi.org/10.1007/s11071-024-09795-2

Bar-Shalom Y., Li X. R., Kirubarajan T.: Estimation with Applications to Tracking and Navigation, Wiley-interscience edn. John Wiley & Sons, New York (2001)

Bittencourt, V.H.S., Silva Junior, W.C., Dias, A.C.B., Aguirre, L.A., Freitas, L.: Influência da topologia na emergência de sincronismo explosivo em redes de osciladores. In: Anais do Congresso Brasileiro de Automática, p. 3598. SBA, Rio de Janeiro (2024)

Bittencourt, V.H.S., Silva Junior, W.C., Montanari, A.N., Freitas, L.: Projeto, construção e modelagem não linear do oscilador de quadratura. In: Anais do Simpósio Brasileiro de Automação Inteligente. SBA, Manaus (2023)

Castro, H.C.d., Aguirre, L.A.: The latent variable subspace in synchronization problems. Journal of Control, Automation and Electrical Systems **36**(4), 648–662 (2025) https://doi.org/10.1007/s40313-025-01179-0

Cohen, J.: Statistical Power Analysis for the Behavioral Sciences, Revised edn. Elsevier, Amsterdam (1977). https://doi.org/10.1016/C2013-0-10517-X . https://linkinghub.elsevier.com/retrieve/pii/C2013010517X

Chen, W., Zhu, W., Fan, Z., Yang, J., Wang, X., Yu, H., Du, S.: Implementation and optimization of amplitude and phase measurement based on kalman filtering. In: Minasian, R., Chai, L. (eds.) Sixteenth International Conference on Signal Processing Systems (ICSPS 2024), vol. 13559, p. 1355916. SPIE, Kunming (2025). https://doi.org/10.1117/12.3061540 . https://doi.org/10.1117/12.3061540

Daum F.: Nonlinear filters: Beyond the Kalman filter. IEEE Aerospace and Electronics Systems Magazine **20**(8), 57–69 (2005)

Paula, A.A., Batista, A.P., Oliveira, E.S., Rodrigues, G.G.: An iterative data assimilation scheme for the unscented Kalman filter. Journal of Control, Automation and Electrical Systems (2026) https://doi.org/10.1007/s40313-026-01267-9

Dias, A.C.B., Silva Junior, W.C., Bittencourt, V.H.S., Freitas, L., Aguirre, L.A.: Verificação experimental de estados de quimera em redes de osciladores de quadratura. In: Anais do Congresso Brasileiro de Automática, p. 3004. SBA, Rio de Janeiro (2024)

Eras-Herrera, W.Y., Freitas, L., Aguirre, L.A.: QuadNet: Uma base de dados para a estimação de estados em sistemas conectados em rede. In: Anais do Simpósio Brasileiro de Automação Inteligente. SBA, São João del-Rei (2025)

Feliks, Y., Groth, A., Robertson, A.W., Ghil, M.: Oscillatory climate modes in the

Indian Monsoon, North Atlantic, and Tropical Pacific. Journal of Climate **26**, 9528–9544 (2013)

Freitas, L., Herrera, W.Y.E., Aguirre, L.A.: QuadNet Estimation Challenge (2025). https://github.com/leandrofabreu/quadNetEstimationChallenge

Haber, A.: Joint sensor node selection and state estimation for nonlinear networks and systems. IEEE Transactions on Network Science and Engineering **8**, 1722–1732 (2021) https://doi.org/10.1109/TNSE.2021.3069890

He, X., Häberle, V., Dörfler, F.: Complex-frequency synchronization of converter-based power systems. IEEE Transactions on Control of Network Systems, 1–12 (2024) https://doi.org/10.1109/TCNS.2024.3420983

Julier, S. J. and Uhlmann, J. K. and Durrant-Whyte, H. F.: A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators. IEEE Transactions on Automatic Control **45**(3), 477–482 (2000)

Kuramoto, Y.: Self-entrainment of a Population of Coupled Non-linear Oscillators. In: Araki, H. (ed.) Lecture Notes in Physics Vol. 39, pp. 420–422. Springer, Berlin, Heidelberg (1975)

Liu, Y.Y., Slotine, J.J., Barabási, A.L.: Controllability of complex networks. Nature **473**, 167–173 (2011) https://doi.org/10.1038/nature10011

Liu, Y.Y., Slotine, J.J., Barabási, A.L.: Observability of complex systems. Proceedings of the National Academy of Sciences of the United States of America **110**(7), 2460–2465 (2013) https://doi.org/10.1073/pnas.1215508110

Li, R., Sun, C., Dong, M., Wang, M., Gao, Q., Liu, X.: The controllability analysis of brain networks during rhythmic propagation. IEEE Transactions on Network

Science and Engineering **11**(4), 3812–3823 (2024) https://doi.org/10.1109/TNSE.2024.3386949

Montanari, A.N., Aguirre, L.A.: Observability of network systems: A critical review of recent results. Journal of Control, Automation and Electrical Systems **31**(1), 1348–1374 (2020) https://doi.org/10.1007/s40313-020-00633-5

Maybeck, P. S.: Stochastic Models, Estimation and Control, Academic press, inc. edn. Academic Press, San Diego (1979)

Montanari, A.N., Duan, C., Aguirre, L.A., Motter, A.E.: Functional observability and target state estimation in large-scale networks. Proceedings of the National Academy of Sciences: PNAS **119**(1), 2113750119 (2022) https://doi.org/10.1073/pnas.2113750119

Montanari, A.N., Freitas, L., Proverbio, D., Gonçalves, J.: Functional observability and subspace reconstruction in nonlinear systems. Physical Review Research **4**, 043195 (2022) https://doi.org/10.1103/PhysRevResearch.4.043195

Montanari, A.N., Freitas, L., Torres, L.A.B., Aguirre, L.A.: Phase synchronization analysis of bridge oscillators between clustered networks. Nonlinear Dynamics **97**, 2399–2411 (2019) https://doi.org/10.1007/s11071-019-05135-x

Monteiro, L.H.A.: Complex Networks in Contemporary Science and Technology. Multidisciplinary Digital Publishing Institute (MDPI) (2026). https://doi.org/10.3390/e28040456

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., Alon, U.: Network motifs: Simple building blocks of complex networks. Science **298**, 824–827 (2002) https://doi.org/10.1126/science.298.5594.824

Pearson, K.: On a new method of determining correlation, when one variable is given by alternative and the other by multiple categories. Biometrika **7**, 248 (1910) https://doi.org/10.2307/2345385

Ristic, B., Arulampalam, S., Gordon, N.: Beyond the Kalman Filter: Particle Filters for Tracking Applications. Artech House, Artech House, Boston, MA, USA. (2004)

Skardal, P.S., Arenas, A.: Dual control of coupled oscillator networks. IEEE Control Systems Society **2**, 146–154 (2023) https://doi.org/2210.01103

Wilcoxon, F.: Individual comparisons by ranking methods. Biometrics Bulletin **1**, 80 (1945) https://doi.org/10.2307/3001968

Wu, Y., Zhou, X., Jiang, G., Zheng, Y., Wang, X.: Target layer state estimation in multilayer complex dynamical networks considering multidimensional node dynamics. International Journal of Robust and Nonlinear Control **36**, 5137–5147 (2026) https://doi.org/10.1002/rnc.70475

Zhang, M.: Almost output synchronization for multi-agent systems with non-identical agents under time-varying topologies. Journal of Control, Automation and Electrical Systems **34**, 29–39 (2023) https://doi.org/10.1007/s40313-022-00957-4