Limits on reconstruction of dynamics in networks

Jiajing Guan, Tyrus Berry, and Timothy Sauer

Limits on reconstruction
of dynamics in networks

Jiajing Guan, Tyrus Berry, Timothy Sauer
George Mason University
Fairfax, VA 22030

An observability condition number is defined for physical systems modeled by network dynamics. Assuming the dynamical equations of the network are known and a noisy trajectory is observed at a subset of the nodes, we calculate the expected distance to the nearest correct trajectory as a function of the observation noise level, and discuss how it varies over the unobserved nodes of the network. When the condition number is sufficiently large, reconstructing the trajectory from observations from the subset will be infeasible. This knowledge can be used to choose an optimal subset from which to observe a network.

I. INTRODUCTION

The study of network dynamics is increasingly useful in modeling physical processes. Networks present a fascinating departure from generic dynamical systems due to the constraints imposed on direct communication between nodes, resulting in complicated dynamics and nontrivial bifurcation structures [1–4]. Modeling by networks has become an important topic in almost every area of physical and biological science, including distributed mechanical processes, weather and climate, and metabolic, genomic and neural networks.

A crucial aspect of studying distributed systems is the difficulty of finding generic observables that facilitate reconstruction of the entire collective dynamics of the network. The theory of observability was pioneered for linear dynamics by Kalman [5]. For nonlinear dynamics, the theory of attractor reconstruction [6, 7] provides hope that for generic observables of sufficiently high dimension, the dynamics can be reconstructed. Although observations at single or even multiple nodes of a network may not be provably generic, the results of Joly [8] show that some aspects of reconstructibility may be present by observing even a single node in a strongly connected network, i.e. a network for which every node is downstream from every other node. Observability in both linear and nonlinear networks is a topic of intense recent interest [9–14] and has close connections to controllability [15–17].

However, observability in theory does not guarantee a satisfactory reconstruction from data collected from a sparsely–connected network, or far from target nodes, even in the case where the equations of motion are known. To date, even in this more tractable scenario, surprisingly little in the way of general practical requirements have been developed for inferring information from measurements. A critical obstruction is the presence of noise in the observations, and how this is magnified in efforts to reconstruct the dynamics. In this article, we offer a definition of observability condition number for reconstruction of network trajectories, demonstrate its asymptotic properties for limiting cases such as full observability, and exhibit its behavior for some relevant examples. The main conclusion is that for practical use of network reconstruction techniques, theoretical observability may be only a first step, and that a condition number measuring error magnification may fundamentally govern the limits of reconstructibility. In short, if the observational noise level is σ times the macroscopic level of the dynamics, then the error magnification must be less than $1/\sigma$ to allow accuracy in the reconstructed dynamics. Our aim is to quantify this magnification for each specific observing subset and target node of the network.

II. OBSERVABILITY CONDITION NUMBER

In the following, we denote by S a subset of observed nodes or variables of a dynamical network. Let X be a network node whose trajectory needs to be reconstructed. We consider an ergodic trajectory of a compact attractor which is observed with noise, and consider the trajectory reconstruction error at one node X of the network.

In this scenario, we conjecture that there is a constant κ depending on S, X, and the dynamics, such that in the low noise limit, the expected error of reconstructing a length-N trajectory satisfies

$$\frac{\text{reconstruction error per step at X}}{\text{observation error per step at S}} \sim \frac{\kappa}{\sqrt{N}}. \tag{1}$$

asymptotically as $N \to \infty$. To be more precise, consider the observation noise to be mean zero with variance σ^2, and let $e = \{e_1, \ldots, e_N\}$ denote the trajectory error at X with component $e_i = z_i - x_i$ in terms of the exact trajectory x_i and the reconstructed trajectory z_i. Since the standard deviation of the trajectory errors e_i on each step is the expected value $\mathbb{E}\left[||e||_2^2/N\right]^{1/2}$, the asymptotic relation (1) translates to the existence of the limit

$$\kappa = \lim_{\substack{N \to \infty \\ \sigma \to 0}} \frac{\sqrt{N}\,\mathbb{E}\left[||e||_2^2/N\right]^{1/2}}{\sigma} = \lim_{\substack{N \to \infty \\ \sigma \to 0}} \frac{\mathbb{E}\left[||e||_2^2\right]^{1/2}}{\sigma}. \tag{2}$$

We call $\kappa = \kappa_{S,X}$ the *observability condition number* of node X observed by S. This is a single constant that encapsulates the ability to reconstruct the dynamics at X from the subset S.

The study of condition number as a measure of controllability and observability is classical, beginning with Friedland [18] in the context of linear systems. Here we consider the nonlinear case, and append the asymptotics expressed in (2). In addition, we describe a direct computational approach to approximating the condition number: Small observational noise is added to a length N trajectory of the dynamical system and a variational data assimilation technique is used to reconstruct the nearest exact trajectory. The ratio of output (reconstruction) error to input (observation) error is $\kappa_{S,X}/\sqrt{N}$.

The fundamental importance of the existence of a universal quantity $\kappa_{S,X}$, independent of trajectory length, is that it allows us to compare the varying fidelity of various observation sets S at node X. This has direct implications for

sensor placement in general systems, such as measurements in a metabolic network or the location of electrodes in a neural assembly.

We begin by establishing the formula (2) for the case of completely observed linear dynamics, where κ exists and equals 1. By "completely observed", we mean that the subset S of observed variables includes all variables. Consider first the scalar case and assume the dynamics $f(x) = ax$. Let $\{x_1, \ldots, x_N\}$ be a trajectory under f, so that $x_t = f^{(t-1)}(x_1)$. Let $y_t = x_t + \delta_t$ be the (completely observed) trajectory observed with i.i.d. observation noise δ_t of mean 0 and covariance $\Sigma(\delta) = \sigma^2 \mathbb{I}_{N \times N}$ for some $\sigma > 0$. We search for a trajectory $\{z_1, \ldots, z_N\}$ where $z_t = f^{(t-1)}(z_1)$, that minimizes the sum squared error

$$\sum_{t=1}^{N}(z_t - y_t)^2 = \sum_{t=1}^{N}(a^{t-1}z_1 - a^{t-1}x_1 - \delta_i)^2. \tag{3}$$

In the sense of least squares, the $\{z_t\}$ trajectory is the one closest to the observations. Setting the derivative with respect to z_t to zero and solving yields

$$0 = z_1 \sum_{t=1}^{N} a^{2(t-1)} - x_1 \sum_{t=1}^{N} a^{2(t-1)} - \sum_{t=1}^{N} a^{t-1}\delta_t$$

$$z_1 = x_1 + \frac{\sum_{t=1}^{N} a^{t-1}\delta_t}{\sum_{t=1}^{N} a^{2(t-1)}}. \tag{4}$$

The square of the numerator of (2) is the expected squared error of the reconstructed trajectory $\{z_1, \ldots, z_N\}$ compared with the original trajectory $\{x_1, \ldots, x_N\}$, or using $z_t - x_t = a^{t-1}(z_1 - x_1)$ and (4),

$$\mathbb{E}\left[\sum_{t=1}^{N}(z_t - x_t)^2\right] = \mathbb{E}(z_1 - x_1)^2 \sum_{t=1}^{N} a^{2(t-1)}$$

$$= \mathbb{E}\left[\left(\frac{\sum_{t=1}^{N} a^{t-1}\delta_t}{\sum_{t=1}^{N} a^{2(t-1)}}\right)^2\right] \sum_{t=1}^{N} a^{2(t-1)}$$

$$= \frac{\sum_{t=1}^{N} a^{2(t-1)}\mathbb{E}(\delta_t^2)}{\sum_{t=1}^{N} a^{2(t-1)}} = \sigma^2$$

where we used the fact that the noises δ_t are uncorrelated. Dividing by the observation noise level σ, we conclude that $\kappa = 1$ for the completely observed case.

The scalar case can be extended to linear dynamics $f(x) = Ax$ for a symmetric matrix by diagonalizing A and applying the above argument componentwise. We expect the result to extend to nonlinear dynamics as well under appropriate technical conditions, whose details will be presented elsewhere.

For partial observations, such as observing at a proper subset S of nodes of a network, $\kappa_{S,X}$ will be substantially greater than 1, which is the focus of this article. As an illustrative example, consider the undirected network of nine nodes illustrated in Fig. 1, where the update equations at node j follow the nonlinear discrete dynamics

$$x_{t+1}^j = a \cos x_t^j + b y_t^j + c \sum_{k=1}^{9} A_{jk} x_t^k, \quad y_{t+1}^j = x_t^j \tag{5}$$

where $a = 2.2, b = 0.4$ for $j = 1, \ldots, 9$ and $A = \{A_{jk}\}$ is the (symmetric) adjacency matrix of the network. The discrete dynamical map used here at each node is a variant of the classical Hénon map [19] that is suitable for distributed dynamics.

We describe an algorithm to compute κ as in (2) from a general network. Generate an exact trajectory $\{x_1, \ldots, x_N\}$, which is observed by a function $h(x)$ plus Gaussian observational noise with variance σ^2 at each point of the trajectory to get $\{y_1, \ldots, y_N\}$, where $y_t = h(x_t) + \epsilon_t$. In the examples to follow, h will be the projection from the current state x, the set of all nodes, to the observation subset S. We apply what amounts to a variational data assimilation method, described in the remainder of the paragraph, to the inexact y_t observations to find an exact trajectory $\{z_1, \ldots, z_N\}$ that minimizes the least squares difference between the y_t and z_t trajectories, analogous to (3). To accomplish this, we applied a Gauss-Newton iteration [20] that enforces z_t being a trajectory while minimizing the observation difference. More precisely, we search for z_t minimizing

$$R^2 = \frac{1}{q^2} \sum_{t=1}^{N-1}(f(z_t) - z_{t+1})^2 + \frac{1}{r^2} \sum_{t=1}^{N-1}(h(z_t) - y_t)^2 \tag{6}$$
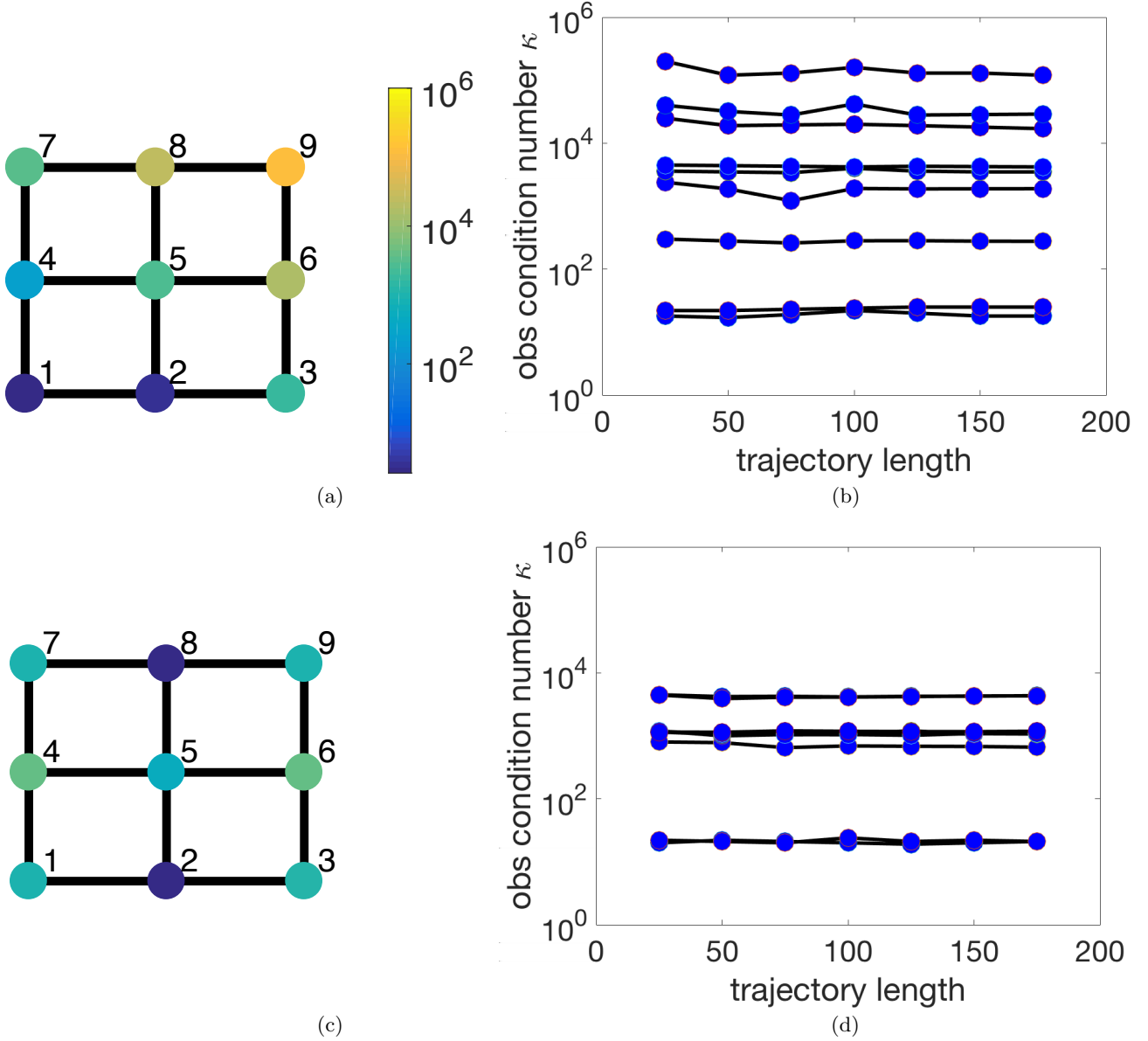
FIG. 1. Undirected network of nine nodes with dynamics (5) and adjacency matrix A as shown. (a) Node X is shaded (color online) according to $\kappa_{S,X}$ where $S = \{1, 2\}$. (b) Estimates of $\kappa_{S,X}$ in (a) as a function of trajectory length.(c) Same as (a), but $S = \{2, 8\}$. (d) Estimates of $\kappa_{S,X}$ in (c).

where q and r are weights that specify the trajectory noise and observation noise tolerances, respectively. We use $q << r$ to ensure that the z_t trajectory is effectively exact, at least relative to the observation errors. At the conclusion of the Gauss-Newton iteration for z_t, we compute the errors $e_t = z_t - x_t$ and the approximation (2) to $\kappa_{S,X}$.

III.   EXAMPLES

The results of this algorithm applied to the network in Fig. 1(a), observed with the x-coordinates at the set $S = \{1, 2\}$, are shown in Fig. 1(b). The nine traces correspond to each of the nine network nodes X. The two observed nodes are at the bottom, and remaining traces show various levels of $\kappa_{S,X}$. In this example, the asymptotic $N \to \infty$ limit in (2) is reached for relatively short trajectory lengths. The nodes in Fig. 1(a) are colored according to the respective observability condition numbers.
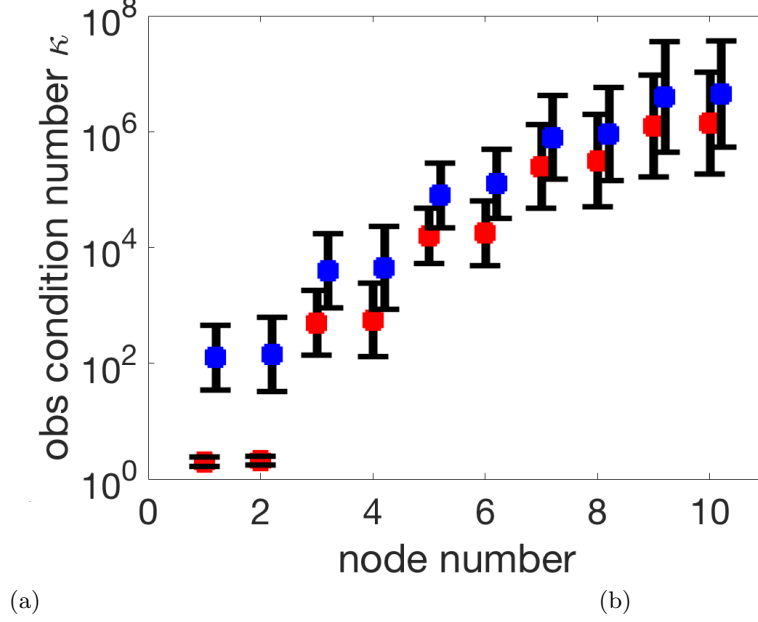
FIG. 2. (a) A network of 10 nodes with dynamics from (5). (b) Plot of $\kappa_{S,X}$ for nodes $1 - 10$ when observed at x-coordinates of subset $S = \{1, 2\}$. The means and standard deviations for the x^j (left, red online) and y^j (right, blue online) variables are shown. The colors correspond to the color bar in Fig. f1.

The fact that $\kappa_{S,X}$ can be arbitrarily large is illustrated by undirected networks such as Fig. 2. The equations (5) are used, and the observing set is $S = \{1, 2\}$. As expected, the resulting condition numbers grow with the distance from the observing set. However, calculating large $\kappa_{S,X}$ is delicate, as we discuss below.

As an example of a discretely-sampled continuous dynamical system we built a directed network of FitzHugh-Nagumo neurons [21, 22] as shown in Fig. 3(a)

$$v'_j = -w_j + d_j v_j - v_j^3/3 + I + g \sum_{k=1}^{9} A_{jk} v_k$$
$$w'_j = a_j - b_j w_j + c_j v_j \tag{7}$$

where the parameters were varied by about 5% from $a = 0.42, b = 0.8, c = 0.08, d = 0.01$ and $I = -0.025$ among the nodes. The system was observed at nodes $1 - 4$ at a step size $\Delta t = 0.1$, and small observation noise was added at each step. Fig. 3(b) shows the observability condition number calculated at the remaining four nodes.

A. Erdős-Rényi and scale-free networks

We tested the computation of $\kappa_{S,X}$ in two network constructions to compare the effects of degree distribution on the observability condition number. For each construction, we used measures of degree and centrality in our exploration. The example in Fig. 4(a) is a scale-free network of 20 nodes. After sorting the nodes in descending order by degree, $\kappa_{S,X}$ was computed for S equal to the first (largest) four, second four, etc and denoted by the round markers. The same analysis carried out using closeness centrality instead of degree is plotted in diamonds. Note that the more sparsely-connected observer sets lead to much increased mean $\kappa_{S,X}$.

The same analysis, but for an Erdős-Rényi network of the same size, is shown in Fig. 4(b). Note that values of $\kappa_{S,X}$ vary much less with the choice of the subset S. This comparison appears to show that according to the observability condition number, there is an advantage to concentrating observers at hub-like nodes, which is quite intuitive.
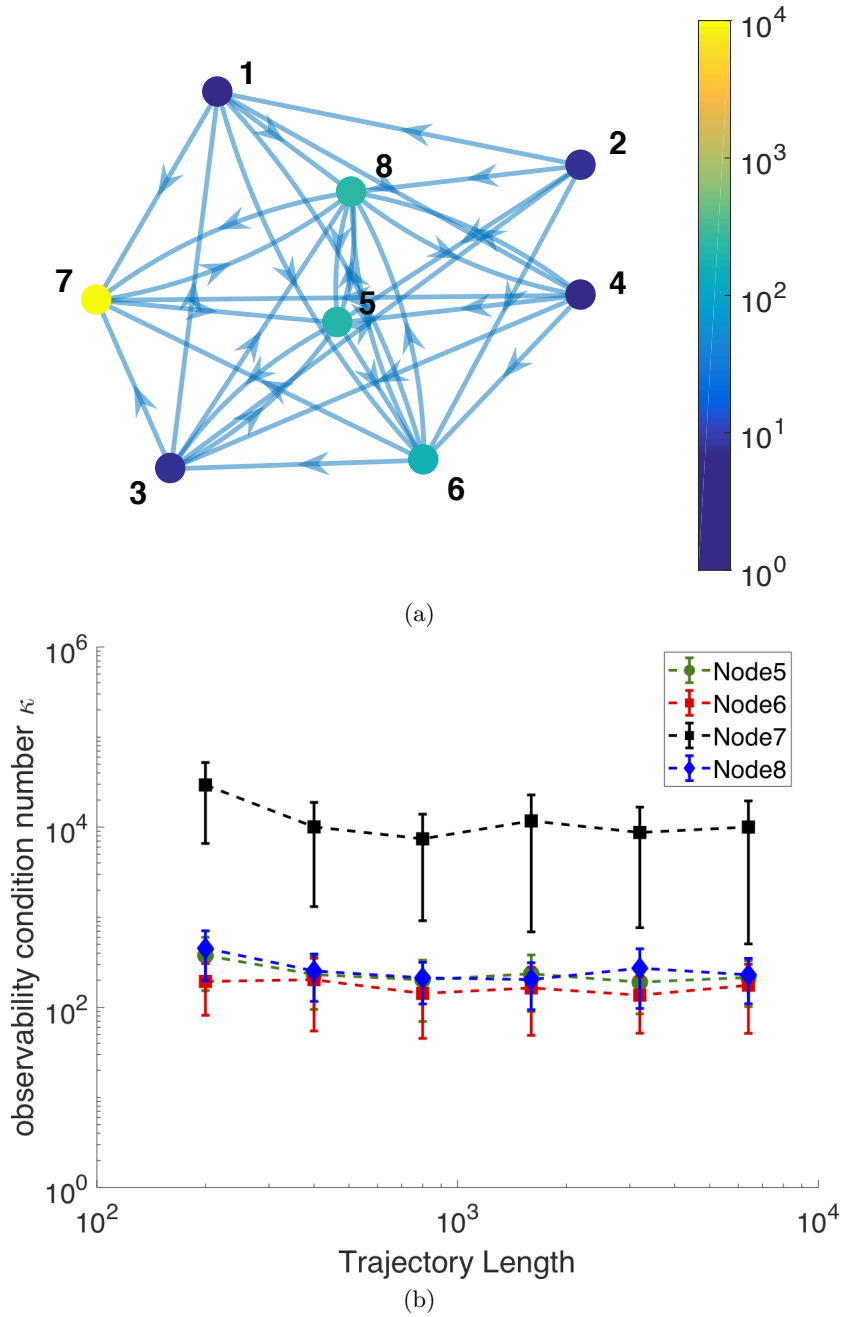
(a)

(b)

FIG. 3. (a) $\kappa_{S,X}$ for FitzHugh-Nagumo network. (b) Estimates of $\kappa_{S,X}$ versus trajectory length. Although not obvious from the adjacency matrix, node 7 is more difficult to observe from the set $S = \{1, 2, 3, 4\}$ than the remaining nodes.

B. Regulatory network

The application of observability condition number to the regulatory network in [23] shows how alternative choices of observation subsets can be usefully compared as a component of experimental design. A representation of the model of 21 differential equations, representing a network that regulates circadian rhythm, is shown in Fig. 5(a), where a directed edge from one variable to another denotes the former is in the right-hand side of the latter's equation. The Supplementary Material contains a Matlab implementation of the equations.

This system was also studied in [24], and we adopt a set of parameter values used there, resulting in periodic dynamics. Fig. 5(b) shows a sample trajectory of all 21 variables undergoing periodic behavior.

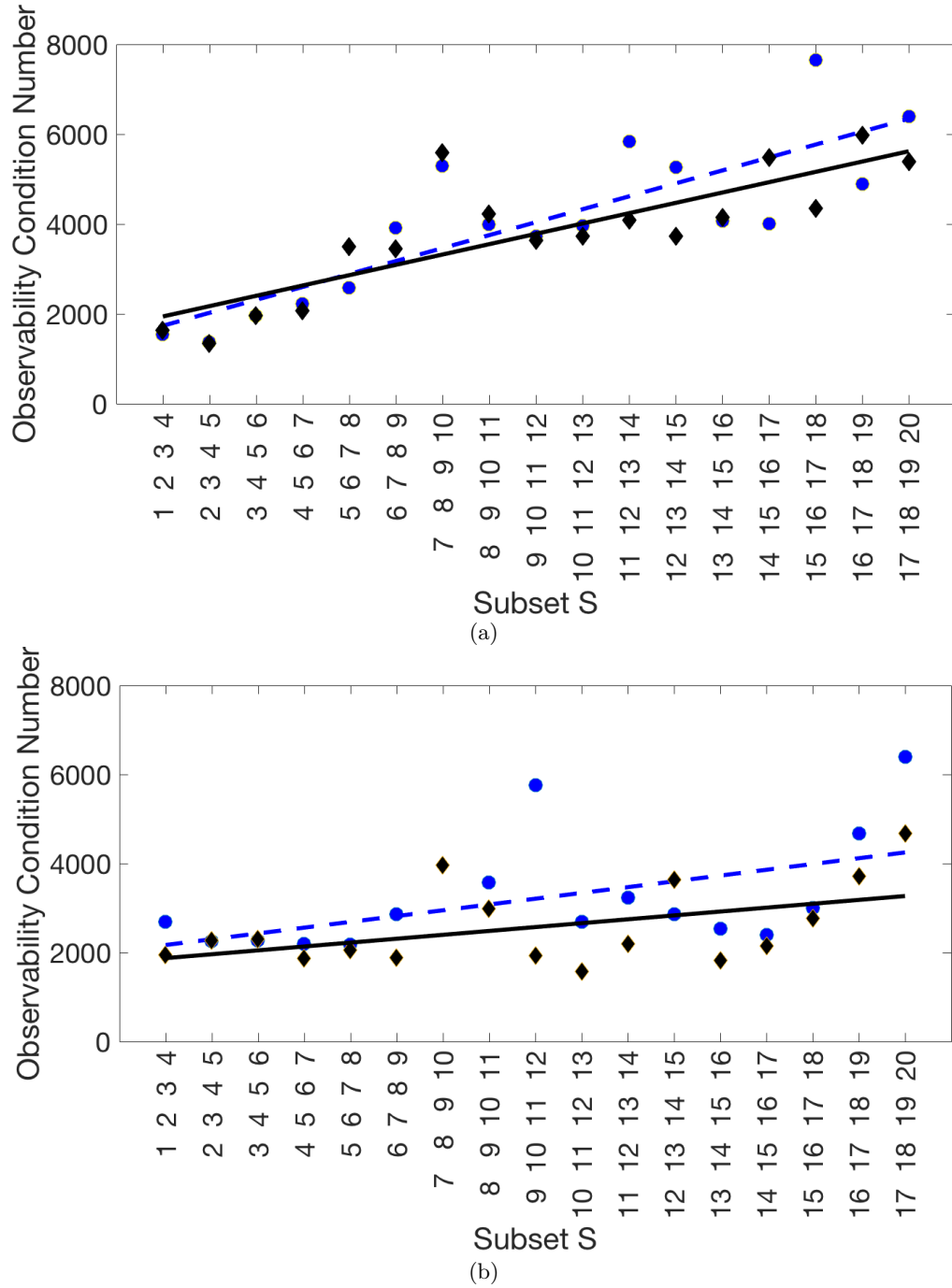Fig. 6(a) displays the observability condition number for 100 random choices of three-node subsets S, computed

FIG. 4. (a) Mean observability condition number of scale-free network, as a function of four-member observation subsets S. Round markers (dashed lines) denote nodes sorted by descending degree; diamond markers (solid lines) denote nodes sorted by descending closeness centrality. (b) Same for Erdős-Renyi network.

along the periodic trajectory of Fig. 5(b). For each three-node subset, a simulation was used to compute $\kappa_{S,X}$ fir each network node X, and the mean was calculated over all nodes X. This calculation was averaged over 10 realizations and the mean and standard error are plotted in the figure. On the horizontal axis, the subsets S are shown, sorted by the value of $\kappa_{S,X}$.

Upon carrying out this numerical exploration, it became obvious that there were two distinct classes of subsets. In particular, Fig. 6(a) shows a clear discontinuity; there is a noticeable gap between the values of $\kappa_{S,X}$ for the leftmost 35 subsets compared to the remainder . Closer analysis (see Supplemental Material, Fig. 1 for details) shows that

each group in the leftmost set contains node 8 or 16, which correspond to reactants Rorc and RORc, respectively. From this striking plot, we can conclude at least that Rorc and RORc are key observables in the system.

Fig. 6(b) shows the sorted observability condition numbers for 100 subsets of six nodes. The same separation is apparent as in the three-node subset case; the subsets to the left of the gap turn out to contain either node 8 or 16 (see Supplementary Material, Fig. 2), similar to Fig. 6(a).

Interestingly, a group of six nodes were isolated in [24] to satisfy the hypotheses of a theorem guaranteeing good control and observability properties. The theorem refers to a "reduced feedback vertex set," which is a set of nodes such that deleting any one causes the network to lack the strong connectedness property. The subset chosen in [24] contains the reactants PER1, PER2, CRY1, CRY2, RORc and CLK/BMAL1, which correspond to nodes numbered 9, 10, 11, 12, 16, and 21. This special subset has node 16 in common with our results, and is included (non-randomly) in Fig. 6(b). However, it is 37th from the left in the plot (see Supplementary Material), meaning that it is not particularly distinguished as an observing subset from the point of view of observability condition number. This underlines another advantage of our approach compared to the otherwise very interesting theorem in [24], in that $\kappa_{S,X}$ can be computed for an arbitrary subset of interest, without satisfying the very specific hypotheses in that result.

The emergence of the importance of nodes 8 and 16 is not obvious from any other technique known to us. This striking result shows how the concept can be directly useful in the choice of observers in a relevant regulatory network.

IV. DISCUSSION

To conclude, we note that special care must be taken to carry out the Gauss-Newton iteration

$$z^{k+1} = z^k - (DR^T DR)^{\dagger} (DR)^T R(z^k),$$

that minimizes the sum (6). where $R = [R_1, \ldots, R_n]$ and $R_1(z)^2 + \ldots + R_n(z)^2$ It is a central tenet of uncertainty quantification that the accuracy of the solution of the iteration will be dependent on the condition number (see [20], for example) of the Gauss-Newton problem, in this case the (conventional) condition number C of the matrix $(DR)^T DR$, where DR denotes the Jacobian. Since the errors added are of size $\sigma << 1$, the residuals R_i can be expected to be on the same order. Thus one may run out of correct significant digits if $C/\sigma > \epsilon_{mach}^{-1} \approx 10^{16}$ for double precision computations.

In Fig. 7 we explore this issue. For simplicity, we consider a discrete stochastic map that multiplies by a random 2×2 matrix at each step. We observe at both phase variables, so that the system is completely observable and $\kappa_{S,X} = 1$. Each marker represents a calculation of $\kappa_{S,X}$ where S is both variables, i.e. completely observed, and X is one of the variables. As shown above, in this case $\kappa_{S,X} = 1$. The markers correspond to trajectories of length between 20 and 240, going from bottom to top, with input noise σ. The vertical axis denotes the condition number C of $(DR)^T DR$. The color of the marker corresponds to the observability condition number $\kappa_{S,X}$. The dashed line is drawn at $C/\sigma = 10^{16}$. Note that as the trajectories become longer, C becomes larger and when the dashed line is passed, $\kappa_{S,X}$ is incorrectly determined (in some cases by a factor of more than 1000), due to lack of significant digits caused by ill-conditioning. To avoid this difficulty, the length of trajectories must be limited to the safe area below the dashed line. Alternatively, computations beyond double precision could be used.

In this article, we have introduced the concept of observability condition number $\kappa_{X,S}$ that has a consistent asymptotic definition in the limit of long ergodic trajectories and the limit of small noise. We have shown that the definition in relatively straightforward to compute in multidimensional systems. This settles a fundamental, long-standing problem in network dynamics, namely where to locate a minimal set of sensors to measure remote dynamics. Computation of $\kappa_{S,X}$ allows a direct comparison of all options. In particular, an exhaustive enumeration among subsets S to find the minimum mean or maximum over the network is feasible for moderate-sized networks, and establishes a guiding principle for large networks where exhaustive search may not be feasible.

ACKNOWLEDGMENTS

[1] M. Golubitsky and I. Stewart, Bulletin of the American Mathematical Society **43**, 305 (2006).
[2] T. D. Sauer, Physical Review Letters **93**, 198701 (2004).

[3] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Physics Reports **424**, 175 (2006).

[4] M. Newman, A.-L. Barabasi, and D. J. Watts, *The structure and dynamics of networks* (Princeton University Press, 2011).

[5] R. Kalman, IRE Transactions on Automatic Control **4**, 110 (1959).

[6] F. Takens, in *Dynamical Systems and Turbulence, Warwick*, Lecture Notes in Mathematics, Vol. 898, edited by D. Rand and L.-S. Young (Springer Berlin / Heidelberg, 1981) pp. 366–381.

[7] T. Sauer, J. Yorke, and M. Casdagli, Journal of Statistical Physics **65**, 579 (1991).

[8] R. Joly, Nonlinearity **25**, 657 (2012).

[9] C. Letellier, L. A. Aguirre, and J. Maquet, Physical Review E **71**, 066213 (2005).

[10] C. Letellier and L. A. Aguirre, Physical Review E **72**, 056202 (2005).

[11] C. Letellier and L. A. Aguirre, Physical Review E **79**, 066210 (2009).

[12] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, Proceedings of the National Academy of Sciences **110**, 2460 (2013).

[13] I. Sendiña-Nadal, S. Boccaletti, and C. Letellier, Physical Review E **94**, 042205 (2016).

[14] A. J. Whalen, S. N. Brennan, T. D. Sauer, and S. J. Schiff, Physical Review X **5**, 011005 (2015).

[15] C.-T. Lin, IEEE Transactions on Automatic Control **19**, 201 (1974).

[16] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, Nature **473**, 167 (2011).

[17] B. Wang, L. Gao, Y. Gao, Y. Deng, and Y. Wang, Scientific Reports **4**, 5399 (2014).

[18] B. Friedland, Journal of Dynamic Systems, Measurement, and Control **97**, 444 (1975).

[19] M. Hénon, in *The Theory of Chaotic Attractors* (Springer, 1976) pp. 94–102.

[20] T. Sauer, *Numerical Analysis*, 3rd ed. (Pearson Education, 2018).

[21] R. FitzHugh, Biophysical Journal **1**, 445 (1961).

[22] J. Nagumo, S. Arimoto, and S. Yoshizawa, Proceedings of the IRE **50**, 2061 (1962).

[23] H. P. Mirsky, A. C. Liu, D. K. Welsh, S. A. Kay, and F. J. Doyle, Proceedings of the National Academy of Sciences **106**, 11107 (2009).

[24] B. Fiedler, A. Mochizuki, G. Kurosawa, and D. Saito, Journal of Dynamics and Differential Equations **25**, 563 (2013).

[25] A. Mochizuki, B. Fiedler, G. Kurosawa, and D. Saito, Journal of Theoretical Biology **335**, 130 (2013).
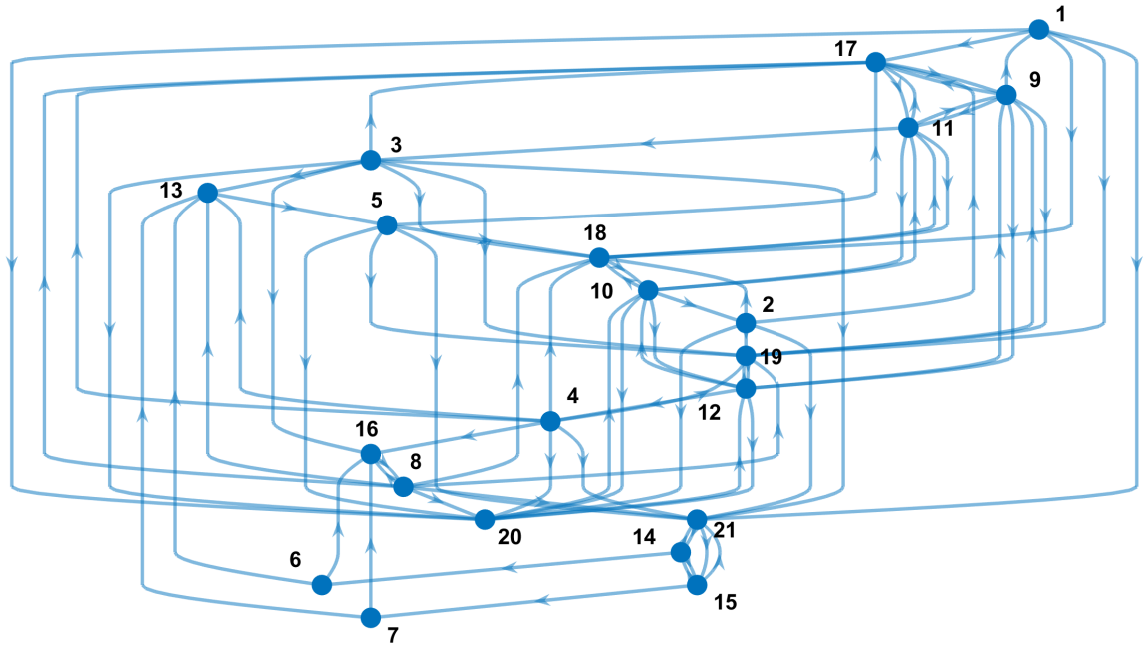
(a)
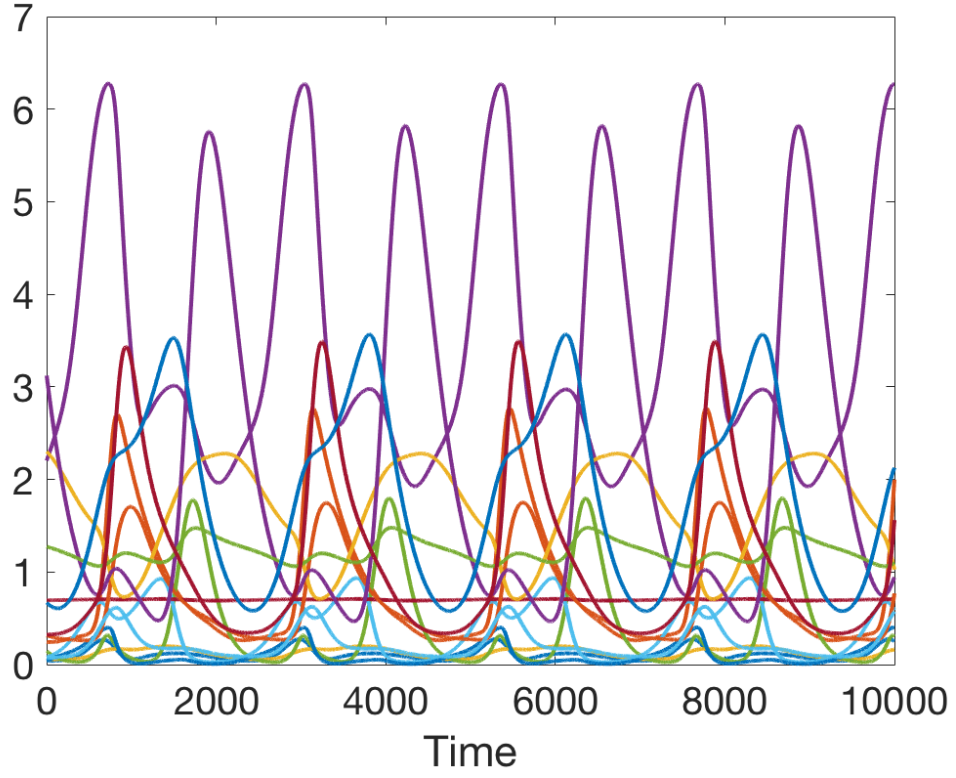
(b)

FIG. 5. (a) Network diagram represents 21-variable differential equation of circadian rhythm model described in [23]. (b) The 21 traces show a periodic trajectory of the 21 variables of the regulatory network in (a).
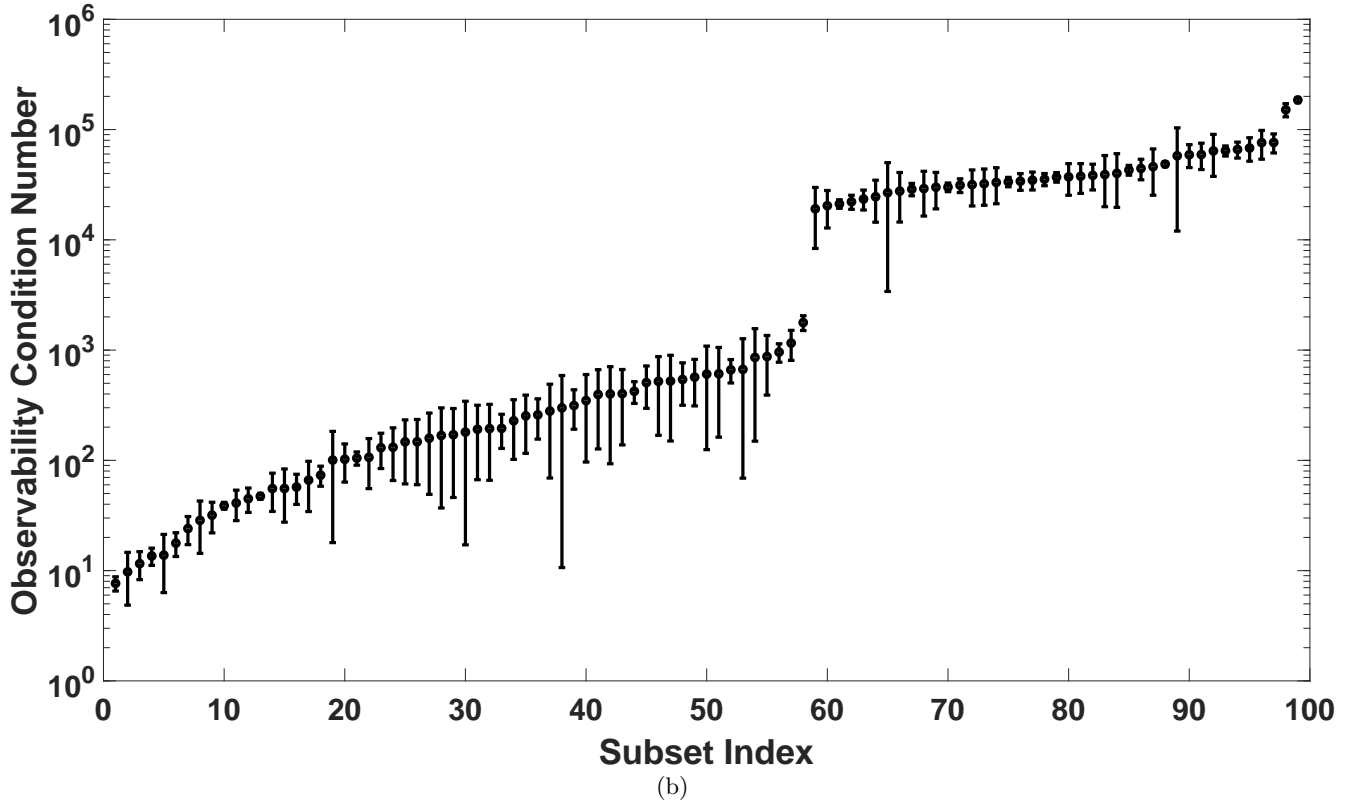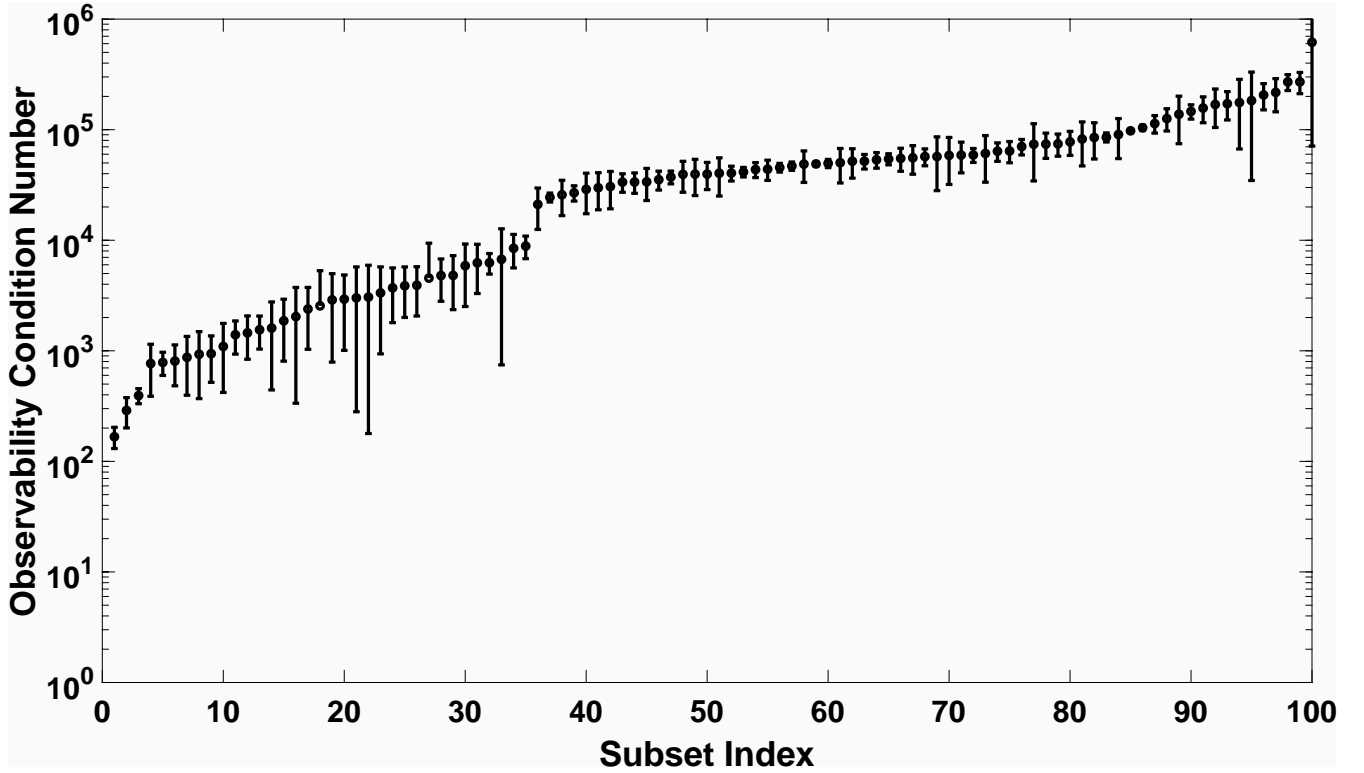
(a)

(b)

FIG. 6. Mean observability condition number $\kappa_{S,X}$ for 21-node regulatory network of [23], for the trajectory in Fig. 5(b). (a) A total of 100 subsets S of 3 nodes, chosen randomly, are compared in the plot. The subsets are sorted from left to right by increasing $\kappa_{S,X}$. There is a noticeable gap after the first 35 subsets. All subsets to the left of the gap contain either node 8 or 16; corresponding to reactants RORc and Rorc, respectively. None of the subsets to the right of the gap contain either RORc or Rorc. (b) Average observability condition number $\kappa_{S,X}$ for 6-node subsets S along the horizontal axis. See Supplementary Material for identification of the 3-node and 6-node subsets, respectively.
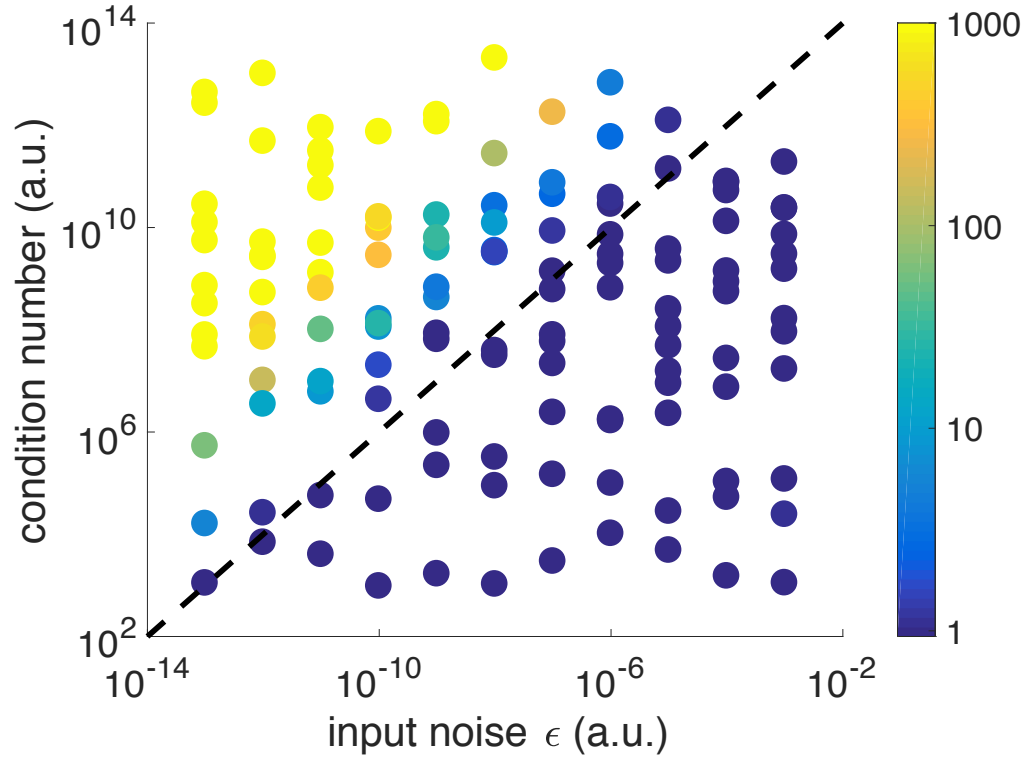
FIG. 7. Markers are shaded by estimate of $\kappa_{S,X}$ and plotted versus input noise and the condition number C of the Gauss-Newton iteration, for completely observed network of random 2×2 matrices. Trajectory lengths vary from 20 (lower points) to 240 (upper points). Calculations below the dashed line $C = 10^{16}\sigma$ are reliable.