Politecnico di Torino

POLITECNICO DI TORINO

Master of Science in Physics of Complex Systems

Department of Applied Science and Technology

Inference of the size of nonlinear network systems from perceptible dynamics

Supervisors
Prof. Lorenzo Zino
Prof. Maurizio Porfiri

Candidate
Francesca Bianca Brovia

Academic Year 2025–2026

Abstract

Network dynamical systems are widely present in science and engineering. The primary property of a network dynamical system is its size, which, for scalar dynamics, corresponds to the number of nodes. For linear network systems, recent studies have developed reliable tools for inferring the size of the system from perceptible dynamics (measurements of one or some of the network nodes) across multiple experiments. In this work, these tools are extended to nonlinear network systems by putting forward a model-agnostic approach that combines clustering techniques, detection matrices, and spectral analysis. The theoretical premise of the algorithm is that, under mild assumptions, the variation between the dynamics of some nodes across multiple measurements can be used to bound the variation between the dynamics of all nodes across the same measurements. By applying clustering techniques on perceptible dynamics, nearby measurements are identified in regions where the variational dynamics can be approximated as linear, allowing the detection matrix to be reliably constructed. The spectrum of this detection matrix is then analyzed to infer its rank, which corresponds to the size of the nonlinear network system. The approach is demonstrated through numerical experiments on different nonlinear network systems, including various hypergraph structures. Whether nonlinearity comes from individual dynamics of the nodes or the interactions among them, it is rarely a feature that one can dismiss. My work paves the way to infer the size of a nonlinear network system when governing equations are unknown and only limited data are accessible.

Contents

Chapter 1

Introduction

1.1 Motivation

Two of the most fascinating problems in the history of science suggest that insights about nature's organization can be drawn from limited empirical observations. The first comes from Einstein's celebrated work on Brownian motion.[1] By studying trajectories of colloidal particles, Einstein was able to provide strong evidence for the atomic theory of matter and to estimate Avogadro's number. What appeared as random fluctuations turned out to reveal a universal constant, encapsulating how macroscopic data can encode hidden microscopic information. The second example comes from the seminal work of Maxwell[2], Boltzmann[3], and other pioneers of statistical mechanics, where ergodicity plays an important role[4]. By tracking a single particle in a real gas in time, one can infer the statistical properties of the entire ensemble, such as the Maxwell–Boltzmann distribution. This demonstrates how microscopic observations can translate into the macroscopic laws governing a system.

These examples point to a broader, unifying question: to what extent can system-level dynamics be inferred from partial, noisy, or indirect observations? Tackling such questions lies at the heart of understanding network dynamical systems and the hidden structures that govern them. Network dynamical systems are a fundamental mathematical framework for modeling a wide variety of natural and technological phenomena that involve multiple interacting units, each one characterized by internal dynamics and interactions with others. Real-world internal dynamics and interactions are seldom linear. For instance, nonlinear network systems arise in neuroscience, where interactions beyond simple pairwise coupling shape brain functional networks[5–7]; in social systems, where behaviors are governed by higher-order social ties, such as collaborative networks[8–13]; and

in biological dynamics, where nonlinear and higher-order processes influence the organization of microbiomes[14–17]. These multidisciplinary examples motivate the need for rigorous tools to investigate nonlinear network systems.

1.2 Background and Literature Review

One of the most fundamental properties of a network dynamical system is its size. In the case of scalar dynamics, this corresponds to the number of nodes, while in the more general case of vector dynamics it corresponds to the number of degrees of freedom. Inferring the size of a dynamical system from partial observations is a fundamental challenge in many scientific fields. In many real-world systems, only a small subset of variables can be measured, either due to experimental limitations or because the internal structure of the system is not directly accessible. As a consequence, developing methods that allow one to infer global properties of the system from limited observations has become an important research direction in data-driven modeling and system identification. Questions about the inference of the size of a system from limited measurements are pervasive in the natural sciences[18–22]. For instance, is it possible to determine the number of interacting species by observing the population trajectory of just a single species over time? Or, is it possible to infer how many planets are in a system by observing the orbital motions of just a few of them? Several studies have recently attempted to answer these questions using different data-driven methods, including: i) the assembly of representative matrices from time-series measurements, whose rank reflects the size of the system[19,20,23]; ii) probing the system with specific input signals to observe its response[18]; and iii) estimating the diffusion coefficients of agents and relating them to the size of the collective[21].

Among these methods, the detection matrix has emerged as a simple and reliable tool[23], which is built using time-series from multiple measurements taken on one or multiple nodes. For linear, deterministic dynamics, the rank of this matrix is equal to the size of the system under mild assumptions; this claim holds whether or not the dynamics are autonomous or non-autonomous[19]. In the case of stochastic dynamics, such as Markov chains, one can construct a detection matrix that captures the evolution of probability distributions. This approach allows for the identification of the correct number of states in a Markov chain, even when some states are indistinguishable to the observer[24].

The detection matrix is conceptually related to the control-theoretic notion of observability[25]. Observability is a fundamental concept in control theory, as it connects a system's outputs to its internal states, determining whether those states

can be reconstructed from available measurements. Observability plays a central role in system identification and state reconstruction problems, as it determines whether the internal state of a system can be uniquely determined from its outputs. In practical applications, this property is crucial when direct access to the full system state is not available and only indirect measurements can be obtained. For linear systems, this condition holds if and only if the rank of the so-called observability matrix is equal to the dimension of the system. Porfiri[19] demonstrated that, under mild assumptions, the rank of the detection matrix coincides with the rank of the observability matrix. To date, the application of the detection matrix to the nonlinear network systems remains limited and not fully understood. When it was first introduced by Haehne *et al.*[23], the method was successfully tested on periodic and chaotic oscillators through linearization in the phase space. The idea is to focus on a specific point in the phase space and examine the variation of nearby trajectories, which should be well described by linear dynamics. Although numerical results are promising, a theoretical guarantee is still lacking.

In this thesis, I aim to address this gap by developing an algorithm that integrates concepts from nonlinear system theory and control theory to support the reliable use of the detection matrix for nonlinear network systems. I seek to infer the size of a nonlinear network system, in which nonlinearities may arise from internal dynamics or interactions. I am particularly keen towards hypergraph-based interactions, extending dyadic models that have dominated the literature on the inference of system size[18–20,23,24,26,27]. Hypergraphs are a generalization of standard graphs that allow edges to connect more than two nodes simultaneously. In many real-world systems, interactions naturally involve more than two entities simultaneously, making pairwise network representations insufficient. Hypergraph models provide a natural mathematical framework to describe such higher-order interactions. In a standard graph, an edge encodes an interaction between exactly two nodes. In contrast, a hyperedge in a hypergraph can link a set of nodes, capturing interactions that involve multiple participants at once. This framework is well suited to describe collective phenomena where outcomes depend on the joint state of several entities rather than pairwise relationships alone. Hypergraph-based models have recently attracted significant attention in the statistical physics, network science, and control communities[13,28–38], as they capture the complexity of real-world systems far more accurately than their dyadic counterparts. For example, in ecological systems, interactions often involve three or more species competing for shared resources or territory, while in social systems, collaboration networks frequently extend beyond pairwise relationships —such as in multi-author scientific publications, where a single paper may connect several researchers at once.

1.3 Contributions of this Thesis

To achieve this objective, I establish a theoretical result that allows for bounding the distance between the state of a nonlinear system to a reference state using information only about the system output. Building on this result, I develop an algorithm that combines clustering techniques, detection matrices, and spectral analysis to infer the size of a nonlinear network system. For network systems with dense interactions, numerical experiments indicate that observing time-series from a single node may suffice to estimate the global network size precisely. Sparser networks require knowledge of more nodes, as one might expect. These results suggest that the amount of information required for reliable size identification depends strongly on the structure and density of the network interactions. A key advantage of the proposed method lies in its model-agnostic nature, as it requires no assumptions about the underlying dynamics governing the system and does not rely on strong assumptions about the form of the underlying interactions, making it broadly applicable to real-world settings where a mathematical model may not have been yet developed.

In summary, the main contributions of this thesis can be summarized as follows:

- Theoretical result. I establish a theoretical result that provides a bound on the distance between the state of a nonlinear dynamical system and a reference state using information solely derived from the system output. This result establishes a formal connection between partial observations and global properties of the system, offering a theoretical foundation for inferring the size of nonlinear network dynamical systems from limited measurements. More broadly, it contributes to the understanding of how output observations can be related to the internal structure of complex networked dynamics.

- Algorithmic framework. Building upon this theoretical result, I develop a model-agnostic algorithm for estimating the size of nonlinear network systems from time-series data. The proposed method combines clustering techniques, detection matrices, and spectral analysis in order to extract structural information from observed trajectories. In particular, the algorithm exploits local dynamical information obtained from measurements collected at a subset of nodes to reconstruct the dimension of the underlying network system. Importantly, the method does not require explicit knowledge of the governing dynamical equations, making it applicable in settings where the system model is unknown or only partially specified.

- Numerical validation and robustness analysis. Finally, I validate the proposed approach through a series of numerical experiments on both standard graphs and hypergraphs. These experiments illustrate the effectiveness of the method in estimating the size of nonlinear network systems under different interaction structures. In particular, the results indicate that for networks with dense interactions, time-series measurements from a single node may be sufficient to accurately estimate the global network size, while sparser networks typically require observations from a larger subset of nodes. These findings highlight the relationship between network structure and the amount of information required for reliable system size identification.

The rest of the thesis is organized as follows. Section 2 introduces the theoretical foundations required for the subsequent development of my method, which is detailed in Section 3. Section 4 presents numerical results on standard graphs and hypergraphs. Finally, conclusions are drawn in Section 5.

Chapter 2

Theory

2.1 Notation

I use the following notation. I denote real numbers as $\mathbb{R}$ and, given a positive integer N, I denote as $\mathbb{R}^N$ the N-dimensional Euclidean vector space. The 2-norm of vector $u \in \mathbb{R}^N$ is $\|u\| \coloneqq \sqrt{u^\top u}$, with $\top$ indicating vector and matrix transposition. Given a real matrix M, $\|M\|$ denotes the induced 2-norm, $\|M\| \coloneqq \sqrt{\rho(M^\top M)}$, where $\rho(M^\top M)$ is the spectral radius (the largest eigenvalue in magnitude). Given an m-dimensional, time-dependent vector $v(t)$, $v\colon [t_0, T] \to \mathbb{R}^m$, its L^2 norm on $\Omega \coloneqq [t_0, T]$

$$\|v\|_{L^2} \coloneqq \sqrt{\int_{t_0}^{T} \|v(\tau)\|^2 \,\mathrm{d}\tau}, \tag{2.1}$$

where t_0 and T denote the beginning and end of an observation throughout the thesis.

2.2 Linearization and observability

I deal with nonlinear dynamical systems in the following form:

$$\dot{x}(t) = F(x(t)), \tag{2.2a}$$
$$y(t) = h(x(t)), \tag{2.2b}$$

where $x(t) \in \mathbb{R}^N$ is the state and $y(t) \in \mathbb{R}^n$ is the output, $\forall t \in \Omega$. The nonlinear vector fields $F\colon \mathbb{R}^N \to \mathbb{R}^N$ and $h\colon \mathbb{R}^N \to \mathbb{R}^n$ are assumed to be differentiable

functions on their domain[1].

A widely used technique to analyze the dynamics of a nonlinear system is to linearize about a nominal trajectory $x^\star(t)$, thereby mapping the dynamics onto a linear time-varying (LTV) system[39]. I define the variation from the nominal trajectory as $\delta_x(t) \coloneqq x(t) - x^\star(t)$, and the variation of the output as $\delta_y(t) \coloneqq y(t) - y^\star(t)$, where $y^\star(t) = h(x^\star(t))$.

Nonlinear system (2.2) is locally approximated about the nominal trajectory $x^*(t)$ by LTV system

$$\dot{\delta}_x(t) = A(t)\delta_x(t), \tag{2.3a}$$
$$\delta_y(t) = C(t)\delta_x(t), \tag{2.3b}$$

where $A : \Omega :\to R^{N\times N}$ is the Jacobian matrix, $A(t) \coloneqq \left.\frac{\partial F(x)}{\partial x}\right|_{x^\star(t)}$, and $C : \Omega \to R^{n\times N}$ the output matrix, $C(t) \coloneqq \left.\frac{\partial h(x)}{\partial x}\right|_{x^\star(t)}$. In practical applications, I typically measure some of the network dynamical systems, so that function h in (2.2) is linear and the corresponding output matrix C is composed of zeros and ones —the latter corresponding to what it is called perceptible dynamics.

The solution of linear system (2.3) is

$$\delta_x(t) = \Phi(t,t_0)\delta_x(t_0), \tag{2.4}$$

where the state transition matrix $\Phi : \Omega \times \Omega \to R^{N\times N}$ is computed from

$$\frac{\partial}{\partial t}\Phi(t,t_0) = A(t)\Phi(t,t_0), \tag{2.5a}$$
$$\Phi(t_0,t_0) = I, \tag{2.5b}$$

with I being the identity matrix. The solution of (2.5) is written in the form of Neumann series[40], which corresponds to the Dyson expansion in quantum physics and to the Peano–Baker series in control theory[41],

$$\Phi(t,t_0) = I + \int_{t_0}^{t} A(\tau)\,\mathrm{d}\tau + \int_{t_0}^{t}\int_{t_0}^{\tau_1} A(\tau_1)A(\tau_2)\,\mathrm{d}\tau_2\,\mathrm{d}\tau_1 + \ldots \tag{2.6}$$

Consequently, the output can be expressed as

$$\delta_y(t) = C(t)\Phi(t,t_0)\delta_x(t_0). \tag{2.7}$$

[1]In principle, global differentiability can be relaxed in favor of local differentiability in the neighborhood of a nominal trajectory.

Upon linearization, the system can be locally approximated by an LTV model, which allows to apply classical concepts from linear systems theory. For completeness, a brief theoretical background is provided in Appendix A.1. A system is said to be observable if one can uniquely reconstruct the initial state from the output over that window[41]. Observability can be analyzed through the observability Gramian

$$W(t_0,T) = \int_{t_0}^{T} \Phi^\top(t,t_0)\, C^\top(t)\, C(t)\, \Phi(t,t_0)\, \mathrm{d}t. \tag{2.8}$$

The system is said to be observable in Ω if and only if the symmetric matrix $W(t_0,T)$ is positive definite[41]; this condition ensures that no nontrivial initial state can generate identically zero output.

A slightly weaker notion of observability is structural observability[42]. For linear systems, structural observability depends on the sparsity pattern of the state transition matrix rather than on the specific numerical values of its entries[43] —the locations of zero and nonzero (fixed but otherwise arbitrary) entries. A system is said to be structurally observable if it is observable for almost all numerical assignments of the nonzero entries of the transition matrix. In particular, a fully connected system —one whose transition matrix has all entries nonzero— is structurally observable[44]. Intuitively, this is because information can propagate between any state variable and any potential sensor node. In what follows, I assume that the systems under consideration are observable. This assumption is reasonable, as the networks I study are connected, their coefficients are drawn randomly —as opposed to specific value that might hinder observability[42]— and self-loops are present —avoiding dynamics unfolding in low-dimensional space[45].

2.3 Bounding the internal state from the output

I consider nonlinear system (2.2), along with its linearized form (2.3), and I make the following assumptions.

Assumption 1. *Matrix function A is piecewise continuous in Ω, so that*

$$\alpha := \operatorname*{ess\,sup}_{\tau\in[t_0,T]} \|A(\tau)\| \tag{2.9}$$

is well defined and Peano–Baker series (2.6) *satisfies* [2]

$$\|\Phi(t_1,t_2)\| \le e^{\alpha(t_1-t_2)}, \quad \forall t_1 \ge t_2. \tag{2.10}$$

[2]This claim can be proved by computing the norm of the right-hand side of (2.6) , applying the triangle inequality, and the norm submultiplicativity.

Assumption 2. *Output matrix $C(t)$ is piecewise continuous in Ω.*

Assumption 3. *Observability Gramian matrix* (2.8) *is positive definite.*

Theorem 1. *Consider nonlinear system* (2.2)*, whose linearization about a nominal trajectory yields LTV system* (2.3)*. Under Assumptions 1–3, the following bound holds*

$$\|\delta_x(t)\| \leq \frac{e^{\alpha(t-t_0)}}{\sqrt{\lambda_{\min}}}\|\delta_y\|_{L^2}, \tag{2.11}$$

where $\lambda_{\min}$ is the smallest eigenvalue of $W(t_0,T)$.

Proof. Recalling (2.4) and (2.7), I connect the L^2 norm of output δ_{y_i} to initial condition $\delta_x(t_0)$ through the observability Gramian,

$$\begin{aligned}
\|\delta_y\|_{L^2}^2 &= \int_{t_0}^{T} \left(C(\tau)\Phi(\tau,t_0)\delta_x(t_0)\right)^\top C(\tau)\Phi(\tau,t_0)\delta_x(t_0)\,\mathrm{d}\tau \\
&= \delta_x(t_0)^\top \left(\int_{t_0}^{T} \Phi(\tau,t_0)^\top C(\tau)^\top C(\tau)\Phi(\tau,t_0)\,\mathrm{d}\tau\right)\delta_x(t_0) \\
&= \delta_x(t_0)^\top W(t_0,T)\delta_x(t_0).
\end{aligned} \tag{2.12}$$

Since the Gramian is symmetric and positive definite, all the eigenvalues are positive and Rayleigh's quotient is bounded by the smallest eigenvalue[41], that is,

$$\frac{z^\top W(t_0,t)z}{\|z\|^2} \geq \lambda_{\min}, \quad \forall z \in \mathbb{R}^N. \tag{2.13}$$

Combining (2.12) and (2.13) with $z = \delta_x(t_0)$, I bound the initial condition of the entire state using the L^2 norm of the output,

$$\|\delta_x(t_0)\| \leq \frac{\|\delta_y\|_{L^2}}{\sqrt{\lambda_{\min}}}. \tag{2.14}$$

Next, I relate the initial condition of the state with the state at any time in Ω. Specifically, by applying Cauchy-Schwartz inequality and using the submultiplicativity of the norm[41], I establish

$$\frac{\mathrm{d}\|\delta_x(t)\|}{\mathrm{d}t} \leq \|A(t)\|\|\delta_x(t)\|, \tag{2.15}$$

which, upon integration from t_0 to t, yields

$$\|\delta_x(t)\| \leq \|\delta_x(t_0)\| + \int_{t_0}^{t} \|A(\tau)\|\|\delta_x(\tau)\|\,\mathrm{d}\tau. \tag{2.16}$$

By using Grönwall-Bellman inequality[41] and Assumption 1, I obtain

$$\|\delta_x(t)\| \leq \|\delta_x(t_0)\| e^{\alpha(t-t_0)}. \tag{2.17}$$

To prove the claim, I bound the norm of the initial condition in (2.17) using (2.14). □

The bound in (2.11) depends on the smallest eigenvalue of the observability Gramian, $\lambda_{\min}$, and the intrinsic system dynamics, encapsulated by α. The smaller $\lambda_{\min}$ (the larger α), the looser the bound.

I comment that the linearized dynamics in (2.3) and the associated transition matrix are linked to the variational equations commonly used in dynamical system theory[46,47], especially their asymptotics, which are used to estimate Lyapunov exponents for chaotic systems[48]. Here, instead, I leverage the transient information encoded in the transition matrix to assess the system's observability and infer its size. For completeness, the full proof of the state reconstruction bound is provided in Appendix A.6.

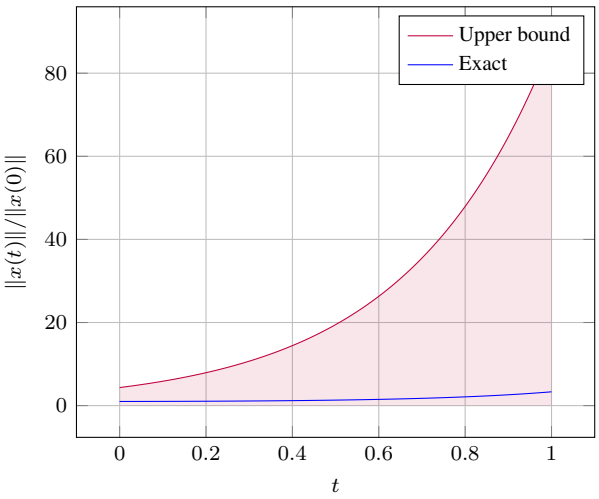


Figure 2.1. Dynamics of Example 1. The red-shaded region represents the upper bound derived from Theorem 1, while the blue curve depicts the exact analytical solution of the system.

Example 1. *I consider LTV system*

$$\dot{x}(t) = A(t)x(t), \qquad y(t) = Cx(t), \tag{2.18}$$

with

$$A = \begin{bmatrix} t+1 & -1 \\ -1 & t+1 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 \end{bmatrix}. \tag{2.19}$$

The state-transition matrix is given by

$$\Phi(t,s) = e^{\frac{t^2-s^2}{2}+t-s} \begin{bmatrix} \cosh(t-s) & -\sinh(t-s) \\ -\sinh(t-s) & \cosh(t-s) \end{bmatrix}. \tag{2.20}$$

The system is observable over $[0,1]$*, as Gramian*

$$W(0,1) = \begin{bmatrix} 9.66 & -6.17 \\ -6.17 & 4.15 \end{bmatrix}, \tag{2.21}$$

is positive definite with $\lambda_{\min} = 0.144$*. I can calculate bound in Theorem 1, by noting that* $\alpha = 3$*; see Fig. 2.1 for a comparison between the evolution of the state norm and the theoretical bound.*

Chapter 3

Methodology

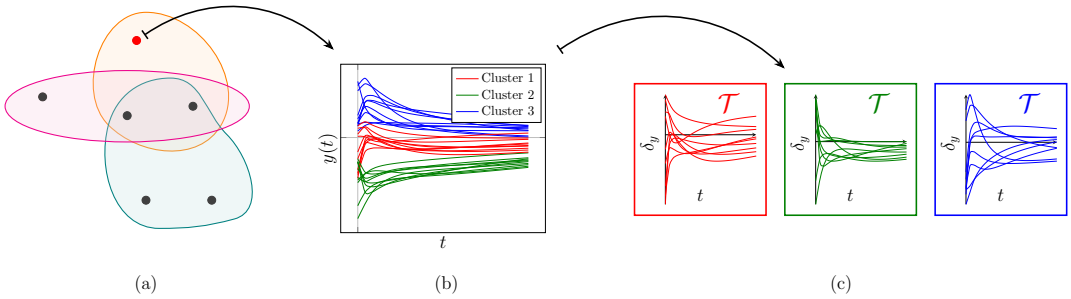


Figure 3.1. Schematic illustration of the proposed approach. (a) Network of $N = 6$ nodes, where one node (red point) is accessible. (b) Time-series of the perceptible dynamics, from different initial conditions. (c) Trajectories are clustered in three groups; different colors are associated with different clusters. A detection matrix is assembled for each cluster to solve the inference problem.

Here, I describe my methodology to infer the size of a nonlinear network system from perceptible dynamics (Algorithm 1), illustrated in Fig. 3.1. I assume I have access to M independent measurements, which may come from multiple experiments in which the researcher can initiate the dynamics from different initial conditions, or a single one in which they partition a longer observation into shorter segments[26,27]. Each measurement consists of the discrete-time recording of the output of a subset of n nodes at S sampling times, all contiguous in time, which I denote as $\{y^{(i,\ell)}(t_j)\}$, for nodes $i = 1,\dots,n$, experiments $\ell = 1,\dots,M$, and sampling times $t_1,\dots,t_S$. In the rest of the thesis, it is assumed that the output of the generic i-th node at time t_j is a scalar quantity, $y^{(i,\ell)}(t_j) \in \mathbb{R}$. This allows a simple notation without loss in generality. In fact, one can deal with vectorial nodal outputs by simply considering each entry as a distinct duplicate of the same

node, ultimately obtaining an estimate of the number of degrees of freedom in the system; see, for instance, Celli and Porfiri [26] for linear systems. A more detailed discussion of the scalar output variable is provided in Appendix A.2.

The first step of the proposed approach consists in the partition of the M measurements into K clusters of nearby dynamics. Based on Theorem 1, one can argue that two measurements with close perceptible dynamics will also have nearby evolution of the internal state of the network system. Hence, measurements within a cluster can be viewed as variations around a nominal trajectory, described by a linearized model that can be studied using a detection matrix. In particular, I apply the standard K-means algorithm to minimize the objective function

$$\sum_{k=1}^{K} \sum_{\ell \in \mathscr{Y}_k} \sum_{i=1}^{n} \sum_{j=1}^{S} |y^{(i,\ell)}(t_j) \mu_{k,j}^{(i)}|^2$$

over the set of clusters $\mathscr{Y} = \{\mathscr{Y}_1, \ldots, \mathscr{Y}_K\}$, where $\mu_{k,j}^{(i)} \in \mathbb{R}$ is the mean of $y^{(i,\ell)}(t_j)$'s over the measurements $\ell \in \mathscr{Y}_k$ [49] (see Line 1 of Algorithm 1). For further implementation details on the clustering procedure, see Appendix A.3. A relatively small value of K is preferred to avoid artificially shrinking the intra-cluster distance. When dealing with a broad range of initial conditions and a larger ensemble of measurements, one may consider increasing K to refine the local characterization of the dynamics. If a cluster contains less than $S+1$ measurements, I discard it from the analysis, so that I work with the $\hat{K}$ clusters that have at least $S+1$ measurements, $\hat{\mathscr{Y}}$ (see Line 2 of Algorithm 1). A more detailed discussion on the rationale for the choice of S and on the minimum admissible value of this parameter is provided in Appendix A.5.

The second step is to construct a detection matrix for each cluster. Given the ultimate goal of aggregating inferences from all the detection matrices into a single estimate, I proceed as follows. If the cluster contains more than $S+1$ measurements, I prune the cluster to $S+1$ measurements by eliminating the measurements that are the furthest away from the mean over the length of the measurement. In particular, for each trajectory $\ell \in \mathscr{Y}_k$, I compute $d^{(\ell)} = \sum_{i=1}^{n} \sum_{j=1}^{S} |y^{(i,\ell)}(t_j) - \mu_{k,j}^{(i)}|^2$ and eliminate trajectories with the largest values of $d^{(\ell)}$ (see Lines 4–5 of Algorithm 1). For notational convenience, I assume that trajectories are ordered so that the closest to the mean are the first $S+1$.

The third step is to identify the trajectory that is the closest to the mean in each of the retained and pruned clusters, that is, the trajectory with the smallest $d^{(\ell)}$. Such a trajectory is regarded as the nominal trajectory for cluster $\mathscr{Y}_k$, denoted as $\hat{y}_k$, and it is used to define the variational dynamics within the cluster at each timestep t_j, $j = 1, \ldots, S$, denoted as $\delta y^{(\ell)}(t_j) \in \mathbb{R}^n$, which is a vector whose generic i-th

entry is equal to the difference $y^{(i,\ell)}(t_j) - \hat{y}_k^{(i)}(t_j)$ (see Line 6 of Algorithm 1). I do not use the centroid itself as the nominal trajectory because, in nonlinear systems, superposition principle does not hold. Consequently, cluster centroids generally do not satisfy the system dynamics and cannot serve as valid nominal trajectories.

The fourth step entails the assembly of the detection matrix[19,23] of cluster $\mathscr{Y}_k$ (see Line 7 of Algorithm 1)

$$\mathscr{T}_k = \begin{bmatrix} \delta y^{(1)}(t_1) & \delta y^{(2)}(t_1) & \cdots & \delta y^{(S)}(t_1) \\ \delta y^{(1)}(t_2) & \delta y^{(2)}(t_2) & \cdots & \delta y^{(S)}(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ \delta y^{(1)}(t_S) & \delta y^{(2)}(t_S) & \cdots & \delta y^{(S)}(t_S) \end{bmatrix}. \tag{3.1}$$

Columns encode the temporal evolution of distinct trajectories, while rows represent time across trajectories.

Then, I estimate the rank of the detection matrix (see Line 8 of Algorithm 1), which for observable LTVs, is equal to the size of the system[19]. To estimate the rank, I follow the method proposed in Haehne et al.[23], which identifies the rank as the index of the largest gap in the ordered singular value spectrum, computed on the logarithmic scale as $\Delta_s = \log(\sigma_s) - \log(\sigma_{s+1})$, for $s \in \{1, ..., S-1\}$. The spectral gap allows for separating large singular values associated with system dynamics from small ones those due to the linear approximation and noise (see Lines 8–9 of Algorithm 1).

Finally, the size of the nonlinear network system is estimated as the average rank over the different clusters $\hat{N}$ (see Line 11 of Algorithm 1).

Algorithm 1 Network size estimation

Require: Nonlinear system (black box), n accessible nodes, M measurements, S recordings per measurement

Ensure: Estimated system size $\hat{N}$

1: Apply K-means clustering at each time-step to partition the set of measurements into K clusters.
2: Remove clusters with less than $S+1$ elements, keeping $\hat{K}$ clusters, termed $\{\hat{\mathscr{Y}}_1, \ldots, \hat{\mathscr{Y}}_{\hat{K}}\}$
3: **for** each cluster $k = 1, \ldots, \hat{K}$ **do**
4: Compute $d^{(\ell)} = \sum_{i=1}^{n} \sum_{j=1}^{S} \left| y^{(i,\ell)}(t_j) - \mu_{k,j}^{(i)} \right|^2$.
5: Select the $S+1$ nearest trajectories to the cluster mean
6: Select the nominal trajectory $\hat{y}_k$ as the trajectory that is closest to the cluster mean.
7: Construct the detection matrix $\mathscr{T}_k$ using (3.1).
8: Compute singular value decomposition
9: Identify largest spectral gap and draw the size estimate for the cluster, $\hat{N}_k$
10: **end for**
11: Compute estimate $\hat{N} = \frac{1}{\hat{K}} \sum_{k=1}^{\hat{K}} \hat{N}_k$

Chapter 4

Results

In this chapter, I demonstrate the proposed method by performing a series of numerical experiments on different classes of nonlinear network systems. The aim is threefold. First, I verify the accuracy of the estimated system size with respect to the true dimension. Second, I evaluate the robustness of the method against parameter changes. Third, I demonstrate the applicability of the approach across different model systems, including hypergraph-based systems.

4.1 Model systems

A hypergraph $H = (V,E)$ is a generalization of a standard graph where $V = \{1,\ldots,N\}$ is the set of nodes and E is the set of hyperedges, each of which can connect any number of vertices. A k-uniform hypergraph is a special case where each and every hyperedge contains exactly k vertices. In a compact form, hyperedges can be represented by a family of adjacency tensors $A^{(p)}$, $p = 2,3,\ldots$, which determine which p-tuple of nodes are connected. In other words, the entry $a^{(p)}_{ijk\ldots\ell}$ of $A^{(p)}$ is non-zero if the p-tuple $\{i,j,k,\ldots,\ell\}$ is connected by an hyperedge, and its magnitude determines the strength of such a connection. Standard graphs can be interpreted as 2-uniform hypergraphs, where the adjacency tensors reduce to the standard adjacency matrix. Assuming scalar dynamics for each node, the dynamics can be written as [38]

$$\begin{aligned}\dot{x}_i(t) = & f_i(x_i(t)) + \sum_{j=1}^{N} a^{(2)}_{ij}\, h^{(2)}(x_i(t),x_j(t)) \\ & + \sum_{j,k=1}^{N} a^{(3)}_{ijk}\, h^{(3)}(x_i(t),x_j(t),x_k(t)) + \ldots, \end{aligned} \tag{4.1}$$

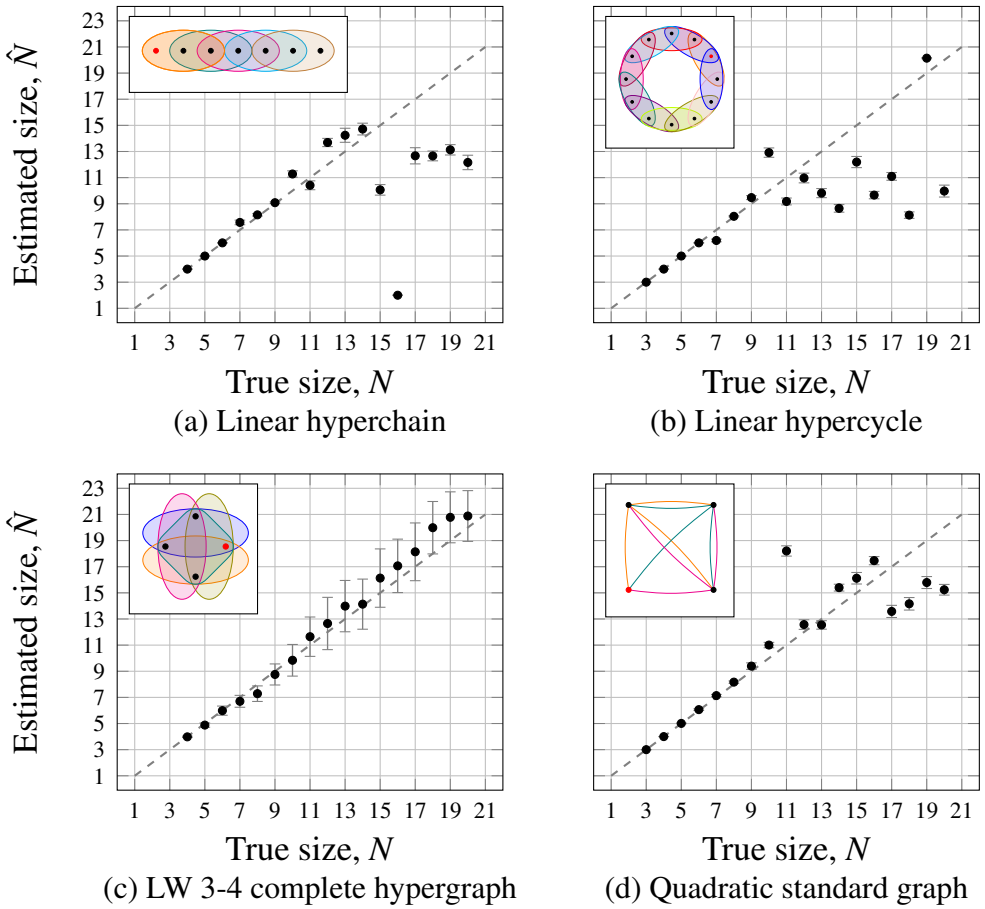


Figure 4.1. Inference of network size from measurements at single node. For each model (a–d), a marker represents the mean estimate across 100 independent realizations; vertical error bars indicate one standard deviation. The bisectrix (dashed gray) serves as a visual reference of perfect inference. In the inferences, it is used $M = 1{,}000$ trajectories, $K = 2$ clusters, and (a,b) $S = 25$ or (c,d), $S = 30$ time-steps. Measured nodes are marked in red.

for $i = 1, \ldots, N$, where f_i describes the intrinsic dynamics of node i. Interaction function $h^{(p)}$ captures the functional form through which multiple nodes belonging to a hyperedge collectively influence each other. I consider four representative instances of (4.1) that illustrate different structural and dynamical features.

Example 2 (Linear hyperchain). *In the linear hyperchain, a system of N nodes evolves according to linear internal dynamics* ($f_i(x_i) = a_i x_i$) *and nearest-neighbor interactions. I consider a third-order hyperchain, where each node interacts with the node before and after, except for the first node that interacts with the second*

and third and the last one that interacts with the second and third last. Hence, the dynamics of each node depend only on the states of its neighbors along a chain, yielding the following equations:

$$\begin{cases} \dot{x}_1(t) = a_1 x_1(t) + a^{(2)}_{1,2,3} x_2(t) x_3(t), \\ \dot{x}_i(t) = a_i x_i(t) + a^{(2)}_{i-1,i,i+1} x_{i-1}(t) x_{i+1}(t), \quad i = 2, \dots, N-1, \\ \dot{x}_N(t) = a_N x_N(t) + a^{(2)}_{N,N-1,N-2} x_{N-2}(t) x_{N-1}(t). \end{cases} \tag{4.2}$$

This simple topology allows to test the performance of the algorithm in a setting where the interaction pattern is minimal and local. Note that, while the internal dynamics is linear, nonlinearities arise from the coupling terms.

Example 3 (Linear hypercycle). *A closely related model to the hyperchain is the hypercycle, obtained by imposing periodic boundary conditions to the linear hyperchain. In this case, the last node is coupled again to the first one, resulting in a cyclic structure. The hypercycle provides a slightly more complex configuration than the hyperchain, as it introduces long-range coupling through the closure of the loop, while still maintaining a regular and homogeneous topology. The evolution of the system over time is described by the following ordinary differential equation:*

$$\dot{x}_i(t) = a_i x_i(t) + a^{(2)}_{i-1,i,i+1} x_{i+1}(t) x_{i-1}(t), \qquad i = 1, \dots, N, \tag{4.3}$$

where cyclic boundary conditions are applied, that is, $N \equiv 0$ and $N+1 \equiv 1$.

Example 4 (Linear weighted 3-4 complete hypergraph). *In this example, I focus on interactions of order three and four, corresponding to three-body and four-body terms, respectively,*

$$\dot{x}_i(t) = a_i x_i(t) + \sum_{j,k=1}^{N} a^{(2)}_{ijk} x_j(t) x_k(t) + \sum_{j,k,l=1}^{N} a^{(3)}_{ijkl} x_j(t) x_k(t) x_l(t). \tag{4.4}$$

This model captures more intricate dependencies than the previous ones, since higher-order terms regulate how groups of three nodes collectively influence one another. It is particularly suitable for testing the capacity of the method to handle higher-order interactions.

Example 5 (Quadratic standard graph). *Finally, I investigate a scenario where the nodes interact over a standard graph —only interactions of order two are present but internal dynamics and coupling terms feature quadratic terms. I re-write the*

dynamics using coefficients a_{ij} and b_{ij}, which are the matrix entries of matrices A and B, to capture quadratic and linear couplings between node i and node j, respectively (the terms a_{ii} and b_{ii} capture the corresponding terms in the internal dynamics). Matrix A weights the local quadratic terms x_j^2 and is chosen to be diagonal with strictly negative entries, thereby inducing dissipative nonlinear self-dynamics at each node. Matrix B encapsulates the linear coupling among nodes and is symmetric and negative definite, ensuring the stability of linear dynamics. Ultimately, I write the following equation set:

$$\dot{x}_i(t) = \sum_{j=1}^{N} a_{ij} x_j(t)^2 + \sum_{j=1}^{N} b_{ij} x_j(t), \qquad i = 1, \ldots, N. \tag{4.5}$$

This example is representative of nonlinear network systems based on standard dyadic interactions between the nodes.

Together, these four classes of models span a spectrum of structural complexity, from simple chains to fully connected higher-order hypergraphs, providing a comprehensive benchmark for the proposed method.

4.2 Inference from a single node

To gather statistical insight, I work with 100 realizations of the model systems. For the hyperchain (Example 2), a_i and a_{ij} are sampled uniformly at random in $\{-1, -0.9, -0.8\}$ and $\{-5.5, -5.0, -4.5, 4.5, 5.0, 5.5\}$, respectively. For the hypercycle (Example 3), a_i and $a_{i-1,i,i+1}$ are sampled uniformly at random in $\{-1.8, -1.7\}$ and $\{-4.5, -4.0, -3.5, 3.5, 4.0, 4.5\}$, respectively. For the linear weighted (LW) 3-4 complete hypergraph (Example 4), a_i is sampled uniformly at random in $\{-0.05, -0.04, -0.03, -0.02\}$, while a_{ijk} and a_{ijkl} are sampled uniformly at random in $\{-0.005, -0.002, 0.005, 0.002\}$. For the quadratic standard graph (Example 5), A is a diagonal matrix with each entry sampled uniformly at random in $\{-1.5, -0.5\}$ and $B = -(M^\top M + \mu I)$, with the each entry of matrix M sampled from a normally distributed random variable with 0 mean and unit variance, and $\mu = 0.5$, in order to have a negative definite matrix. The initial conditions are sampled uniformly at random in $[0.5 - 10^{-5}, 0.5 + 10^{-5}]$. Without loss of generality, node 1 is the accessible one. Each trajectory is simulated for a time interval of unit length and sampled with a uniform spacing, unless differently specified.

In agreement with my intuition, measurements from a single node are sufficient to accurately infer the size of a nonlinear network system of a small size (Fig. 4.1).

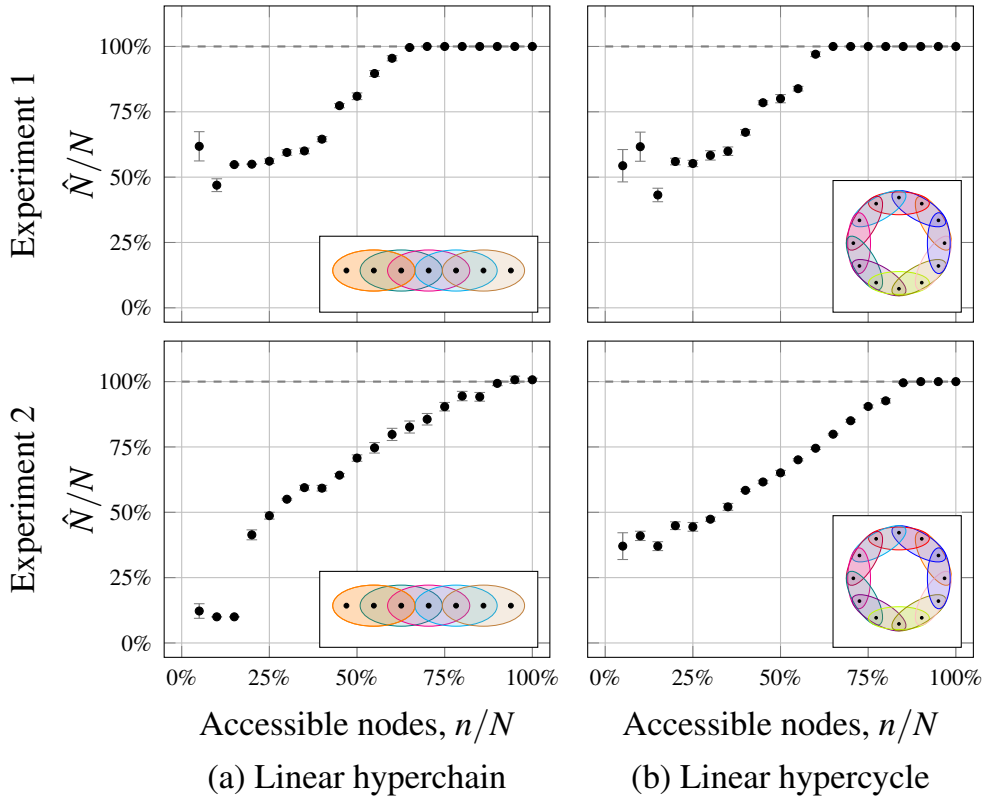
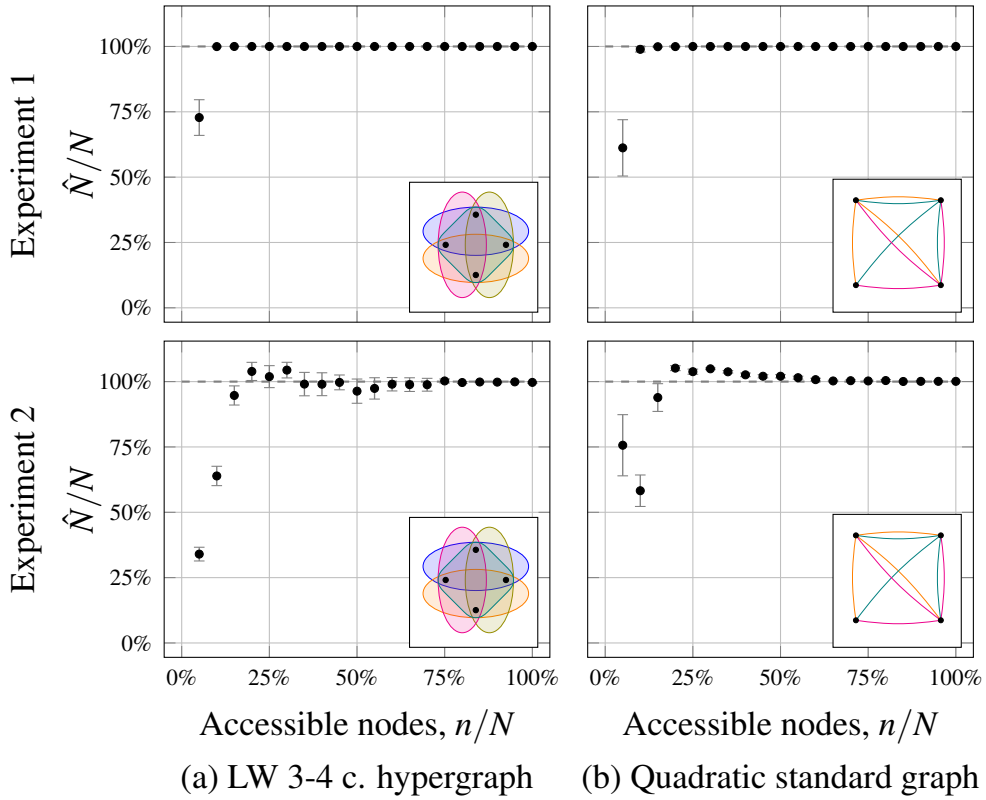


Figure 4.2. Inference of network size from measurements at multiple nodes. Plot of $\hat{N}/N$ versus the fraction of selected nodes n/N; markers show averages over 100 independent realizations and error bars denote one standard deviation. The number of nodes is $N = 20$. In the inferences, it is used $M = 1{,}000$ trajectories, $K = 2$ clusters, and $S = 25$ time-steps. Two sets of experiments are performed with different ranges of initial conditions: for Experiment 1 (first row), initial states are uniformly sampled in $[0.5 - 10^{-5}, 0.5 + 10^{-5}]$, for Experiment 2 (second row) they are uniformly sampled in $[0.5 - 10^{-2}, 0.5 + 10^{-2}]$. Insets show the corresponding interaction topologies.

Across multiple model systems, I demonstrate the possibility of inferring the system size for networks with less than 10 nodes. Whether the nonlinearity is due to the interaction, as in a hypergraph-based model, or in the internal dynamics, as in the standard graph, my methodology yields accurate predictions of the system size. As the network size increases past 10 nodes, consistent recovery from a single-node measurement is only possible for complete hypergraphs. In this case, any node contains a significant footprint of the whole network, thereby enriching the information contained in the detection matrix.



(a) LW 3-4 c. hypergraph (b) Quadratic standard graph

Figure 4.3. Inference of network size from measurements at multiple nodes. Plot of $\hat{N}/N$ versus the fraction of selected nodes n/N; markers show averages over 100 independent realizations and error bars denote one standard deviation. The number of nodes is $N = 20$. In the inferences, it is used $M = 1{,}000$ trajectories, $K = 2$ clusters, and 30 time-steps. Two sets of experiments are performed with different ranges of initial conditions: for Experiment 1 (first row), initial states are uniformly sampled in $[0.5 - 10^{-5}, 0.5 + 10^{-5}]$, for Experiment 2 (second row) they are uniformly sampled in $[0.5 - 10^{-2}, 0.5 + 10^{-2}]$. Insets show the corresponding interaction topologies.

4.3 Inference from multiple nodes

The setup for the simulation is equivalent to the one presented above, with the main difference being that more than one node are accessible. Without loss of generality, I assume that the first n/N nodes are accessible. Also different from the analysis above, I consider two sets of initial conditions: sampling at random in $[0.5 - 10^{-5}, 0.5 + 10^{-5}]$ (Experiment 1) and sampling at random in $[0.5 - 10^{-2}, 0.5 + 10^{-2}]$ (Experiment 2).

Measuring more than one node opens the door for accurate inferences of larger

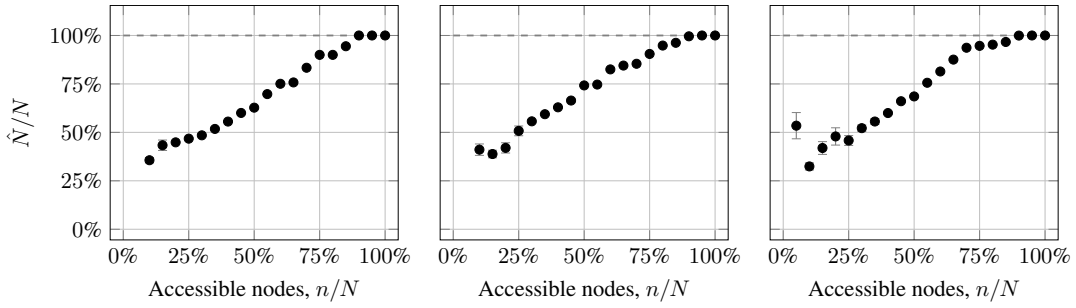
systems (Figs. 4.2, 4.3). In agreement with my intuition, size recovery improves in a nearly monotonic fashion as the number of measured nodes increases. The more information is collected from the network, the closer is my estimate to the true system size. Across model systems, measurements of about a half of the network nodes seem to be sufficient to exactly determine the system size. Yet, for dense networks, such as complete hypergraphs or quadratic standard graphs, even 10% of the nodes are sufficient to achieve exact size recovery. In contrast, for sparse topologies such as chains or cycles, a larger number of measured nodes is required before the system size can be reliably identified. These results highlight how the structural properties of a network influence the amount of information that shall be sampled in order to achieve accurate inference.

These conclusions are fairly robust to changes in the range of the initial conditions (Figs. 4.2, 4.3). Widening the initial conditions by three orders of magnitude yields some reductions in the accuracy of the inference. Specifically, for the linear hyperchain and the linear hypercycle, one needs to increase the fraction of accessible nodes to 80% to attain perfect inference. With respect to the LW 3-4 complete hypergraph and the quadratic standard graph, performance deteriorates marginally. In these cases, 15% of the network nodes are needed for a perfect inference, offering promise for the application of the methodology beyond synthetic data.

4.4 Robustness analysis

To demonstrate the dependence of my findings with respect to the parameters of the algorithms, I perform additional simulations on the most challenging case examined above (Fig. 4.2a), namely, the linear hyperchain in Experiment 2. I individually vary (doubling) the number of time-steps S, the number of trajectories M, and the number of clusters K compared to the case in Fig. 4.2. Overall, I do not unveil a critical effect of S and M (Fig. 4.4a,b), warning prudence against embarking in higher computational costs with longer time-series or more measurements. Increasing the number of clusters K to four seems to have a more beneficial effect, whereby I attain nearly perfect inference by measuring about 70% of the network, instead of the 80% mark required when using two clusters (Fig. 4.4c). This evidence is in line with my intuition that a small number of clusters may be inadequate to describe dynamics with wider initial conditions.

I also test the performance of my algorithm on a larger network. Specifically, I consider a network with $N = 100$ nodes, and I repeat the analysis performed in Section 4.3 for the quadratic standard graph (Example 5). My results (Fig. 4.5),



(a) S=50 (b) M=2,000 (c) K=4

Figure 4.4. Inference of system size from multiple perceptible dynamics for hyperchain with varying parameters. Plot of $\hat{N}/N$ versus the fraction of selected nodes n/N; markers show averages over 100 independent realizations and error bars denote one standard deviation. Unless differently stated in the sub-caption, $N = 20$, $S = 25$, $M = 1{,}000$, and $K = 2$. Initial states are uniformly sampled in $[0.5 - 10^{-2}, 0.5 + 10^{-2}]$.

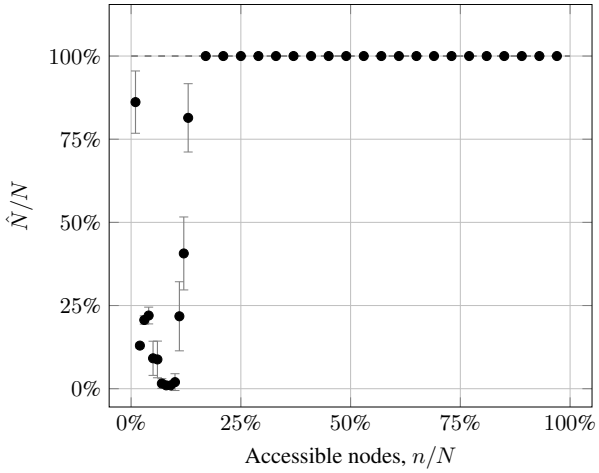


Figure 4.5. Inference of system size from multiple perceptible dynamics for quadratic standard graph with $N = 100$. Plot of $\hat{N}/N$ versus the fraction of selected nodes n/N; markers show averages over 100 independent realizations and error bars denote one standard deviation. In the inferences, we use $M = 4{,}000$ trajectories, $K = 2$ clusters, and $S = 110$. Each trajectory is simulated for a time interval of [0.0,0.75]. Initial states are uniformly sampled in $[0.5 - 10^{-5}, 0.5 + 10^{-5}]$.

are consistent with my findings for smaller networks in Fig. 4.3d, suggesting that the algorithm can also be applied to larger networks. In fact, my numerical simulations (Fig. 4.5) reveal that it is enough to access 17% of the nodes to obtain the

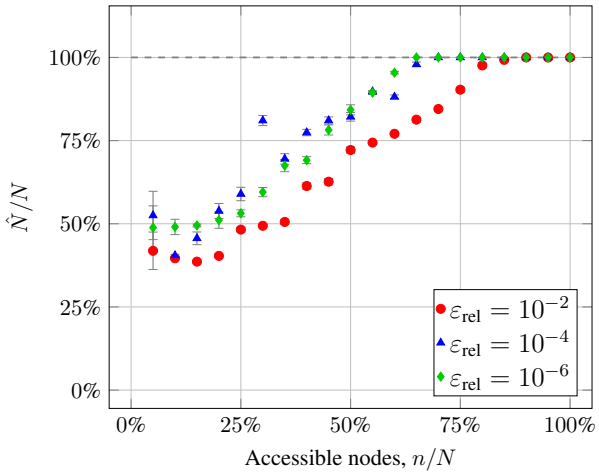


Figure 4.6. Inference of system size from multiple nodes for hyperchain with different colors representing different levels of noise, ε_{rel}. Plot of $\hat{N}/N$ versus the fraction of selected nodes n/N; markers show averages over 100 independent realizations and error bars denote one standard deviation. The number of nodes is $N = 20$. In the inferences, it is used $M = 1{,}000$ trajectories, $K = 2$ clusters, and $S = 25$. Initial states are uniformly sampled in $[0.5 - 10^{-5}, 0.5 + 10^{-5}]$.

correct estimate of the network size N.

Finally, I test the robustness of the method with respect to the presence of noise in the measurements. Specifically, I consider the scenario of the linear hyperchain in Experiment 1 (Fig. 4.3a, top row), and simulate an additive Gaussian noise on the measurement. Specifically, each measurement $y^{(i,\ell)}(t_j)$ is defined as the sum of the value of the output of node i at time t_j in experiment ℓ and the realization of a Gaussian with 0 mean and different standard deviations, each realization independent of the others. For the analysis, I consider three different values of standard deviation, defined in relative terms with respect to the average value of the state, denoted as ε_{rel}. Results (Fig. 4.6), suggest that my method is robust with respect to small-to-medium levels of noise. For noise with relative standard deviation $\varepsilon_{\text{rel}} < 1\%$ with respect to the value of the state, the results obtained are almost indistinguishable from those obtained in the absence of noise, with a visible but slight worsening of the inference quality when $\varepsilon_{\text{rel}} = 1\%$. Importantly, the algorithm does not seem to tend to overestimate the network size.

Chapter 5

Conclusion

Recent years have seen a surge of methodologies to estimate the size of a network dynamical system from perceptible dynamics[18–24,26,27]. Most of the research has focused on linear network systems, hindering application to real-world systems where nonlinearities arise from internal dynamics and coupling. The latter case is one that is becoming more and more timely, as hypergraphs are gaining traction[5,13,22,28–38]. To address this methodological gap, I put forward a novel methodology that leverages time-series from accessible nodes across multiple experiments.

My approach unfolds in a sequence of algorithmic steps, which start from clustering analysis of available trajectories to inform the identification of close-by dynamics for the entire state. For each cluster, I construct a detection matrix, encoding the evolution of the variational dynamics about the cluster mean trajectory. By computing the rank of all the detection matrices, I estimate the size of the system. Knowledge of a single network node is sufficient to determine the size of dense networks, where the dynamics of each node contains a footprint of the entire network. For sparse networks, instead, one may need to rely on a more complete knowledge of the system dynamics, unless the network contains less than ten nodes.

The main limitations of the work are its reliance on the observability of the nonlinear network systems —a property that is not universal to all network systems. For example, structured networks without self-loops may fail to display observability[50,51]. Beyond the relaxation of the assumption of observability, several extensions are promising. First, an exact quantification of the role of network density on the algorithm success is lacking. While I numerically document

a difference between sparse and dense networks, a theoretical framework to examine such a difference is still missing. Second, my validation focuses on non-divergent dynamics and to a limited set of parameters. Performing an extensive numerical parametric study as well as further simulations on other scenarios — including vectorial dynamics— is an important direction of future research. Third, I presently rely on the mean rank of the detection matrix across clusters; it is tenable that further insight could be gathered by examining the variations across clusters, especially when dealing with experimental trajectories that are best described by several clusters. Fourth, the algorithm assumes deterministic dynamics, while numerical simulations suggest that the proposed method is robust with respect to small-to-medium levels of noise. Formally extending the approach to account for stochasticity in the form of added noise or network dynamics is an important area of research.

Another promising direction for future research lies in the integration of Koopman operator theory into the proposed framework. Koopman analysis provides an operator-theoretic perspective on nonlinear dynamical systems by describing their evolution through a linear operator acting on a space of observable functions of the state. Although the underlying system may be nonlinear, the dynamics of these observables evolve linearly in an infinite-dimensional function space. This perspective is particularly appealing for the analysis of complex network dynamics, since it enables the use of linear spectral methods to study nonlinear systems. In practice, the main challenge lies in identifying a suitable set of observable functions that span an approximately invariant subspace of the Koopman operator, allowing the construction of finite-dimensional linear representations of the dynamics. In the context of the present work, Koopman-based approaches may provide an alternative strategy for inferring the size of nonlinear network systems by constructing linear representations of the observed dynamics in a lifted observable space. Combining the detection-matrix methodology developed in this thesis with Koopman-based embeddings could therefore represent an interesting direction for future investigation. More broadly, the integration of Koopman operator techniques with data-driven system identification methods may offer new tools for uncovering hidden structure in nonlinear network dynamics and for improving the robustness of network size estimation from partial observations.

In conclusion, this work opens a pathway toward model-agnostic inference of hidden dynamics in nonlinear networked systems. By relying solely on observable time-series data and avoiding strong assumptions on the governing equations, the proposed methodology provides a flexible framework for investigating complex networked dynamics. As such, it offers promising perspectives for the analysis of large-scale interconnected systems arising in physics, biology, engineering, and

the social sciences. Moreover, the framework may serve as a basis for future developments aimed at understanding the structure and dimensionality of complex nonlinear networks from limited observations.

Chapter 6

Data Availability

The data that support the findings of this study and the code used for the simulations is available: `https://github.com/dynamicalsystemslaboratory/dynamics-size-inference` The repository contains the scripts used to reproduce the numerical experiments presented in this thesis. In particular, a separate code is provided for each graph example discussed in the text, including the linear hyperchain, linear hypercycle, LW 3-4 complete hypergraph, and quadratic standard graphs.

The scripts whose name contains the label "1" correspond to the inference procedure based on time-series observations from a single node, while those labeled "2" implement the inference from multiple observed nodes. In addition, the repository includes a dedicated script for the hyperchain example with additive Gaussian noise, which is used to assess the robustness of the proposed method under noisy observations.

Chapter 7

Related Publication and Author Contributions

Part of the results presented in this thesis has been developed within the following scientific publication:

> F. B. Brovia, R. Succar, L. Zino, and M. Porfiri, "Inference of the size of nonlinear network systems from perceptible dynamics", *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 2026.

The published article is available at:
`https://doi.org/10.1063/5.0318179`
The contributions of the authors to the published work are summarized below.

- **Francesca Bianca Brovia:** Conceptualization, data curation, formal analysis, investigation, methodology, software implementation, validation, visualization, and writing of the original draft.
- **Rayan Succar:** Investigation, methodological support, and writing — review and editing.
- **Lorenzo Zino:** Conceptualization, investigation, methodology, supervision, and writing — review and editing.
- **Maurizio Porfiri:** Conceptualization, funding acquisition, investigation, methodology, resources, supervision, and writing — review and editing.

Appendix A

Additional Numerical Experiments and Methodological Clarifications

This appendix provides additional material that complements the results presented in the main text. In particular, I report further numerical experiments and clarify several aspects of the mathematical formulation of the proposed algorithm. These additions address specific points raised during the review process.

First, I clarify the concept of observability of a dynamical system, an essential notion for understanding the thesis. Second, I discuss the notation used for the observed measurements and the assumption that the measured output associated at each node at each time is a scalar quantity. I then provide additional details on the clustering procedure employed in the algorithm, with particular focus on the use of the k-means method. Finally, I present supplementary numerical simulations performed with different parameter values to verify that the performance of the proposed inference procedure does not depend on a specific choice of model parameters. At the end of Appendix A the full proof of the theoretical result is presented.

A.1 Observability concept

Observability is a fundamental concept in control theory that determines whether the internal state of a dynamical system can be reconstructed from its external outputs over time. A system is said to be observable if its current state can be uniquely determined from the history of output measurements.

Consider the linear time-invariant system

$$\dot{x}(t) = Ax(t), \qquad y(t) = Cx(t), \tag{A.1}$$

where $x(t) \in \mathbb{R}^N$ is the state vector and $y(t) \in \mathbb{R}^p$ is the measured output.
The observability matrix is defined as

$$\mathscr{O} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{N-1} \end{bmatrix}. \tag{A.2}$$

The system is observable if and only if

$$\operatorname{rank}(\mathscr{O}) = N, \tag{A.3}$$

which ensures that the entire internal state vector $x(t)$ can be reconstructed from the output measurements.

An equivalent characterization of observability can be obtained through the observability Gramian. For a time interval $[t_0, T]$, the observability Gramian is defined as

$$W(t_0, T) = \int_{t_0}^{T} \Phi^\top(t, t_0) C^\top C \Phi(t, t_0)\, \mathrm{d}t\,, \tag{A.4}$$

where $\Phi(t, t_0)$ is the state transition matrix of the system. The system is observable on $[t_0, T]$ if and only if the Gramian $W(t_0, T)$ is positive definite.

In many practical scenarios the system matrices A and C are unknown and only time-series measurements of the outputs are available. In such cases, classical observability analysis cannot be directly applied. Instead, data-driven approaches can be used to infer properties of the underlying system.

In the methodology considered in this work, nonlinear network dynamics are locally approximated through linearization around a nominal trajectory. Under suitable observability conditions, variations in the measured outputs encode information about the variations of the internal state.

To exploit this property, measurements collected from multiple experiments or initial conditions are arranged into a detection matrix

$$T_{(k,M)} = \begin{bmatrix} y^{(1)}(t_1) & \dots & y^{(M)}(t_1) \\ \vdots & \ddots & \vdots \\ y^{(1)}(t_k) & \dots & y^{(M)}(t_k) \end{bmatrix}, \tag{A.5}$$

where M denotes the number of trajectories and k the number of sampled time points.

When the system is observable, the rank of this matrix reflects the dimension of the underlying dynamical system. Therefore, estimating the rank of the detection matrix provides a data-driven mechanism to infer hidden states or the size of a networked dynamical system from empirical time-series data.

A.2 Scalar Observation Variable

In the formulation of the algorithm, the measurements are denoted by $y^{(i,\ell)}(t_j)$.

More precisely, I assume that I have M independent measurements, which may correspond, for instance, to repeated experiments performed on the same system. Each measurement consists of the observed dynamics of n nodes sampled at S time instants t_j. To make this structure explicit, the measurements can be written as $y^{(i,\ell)}(t_j)$ where $\ell = 1,\ldots,M$ denotes the measurement index, $i = 1,\ldots,n$ identifies the observed node, and $j = 1,\ldots,S$ indicates the sampling time.

In the implementation adopted in this work, I assume that the observation associated with each node at each time step is a scalar quantity. This assumption simplifies the notation and the construction of the detection matrix used by the algorithm. In many practical applications, this assumption is also natural, since each sensor typically measures a single physical quantity.

It is worth noting, however, that the proposed method is not fundamentally restricted to scalar observations. If the measurement at each node were a vector quantity, its components could be stacked to form an augmented observation vector, and the detection matrix could be constructed accordingly. While this extension does not introduce conceptual difficulties, it leads to heavier notation and would obscure the main ideas of the method. For this reason, in the present work I focus on the scalar measurement case.

A.3 Clustering Procedure

An essential step of the proposed inference algorithm is the identification of clusters of similar observed dynamics. These clusters are used to locally approximate the nonlinear dynamics through linearization, allowing the construction of detection matrices within each cluster.

To perform this task, I employ the k-means clustering algorithm. Clustering is performed with respect to the M measurements, each described by the scalar

observations $y^{(i,\ell)}(t_j)$, where $\ell = 1,\dots,M$ indexes the measurements, $i = 1,\dots,n$ denotes the node index, and $j = 1,\dots,S$ represents the sampling time.

The clustering procedure minimizes the objective function

$$\sum_{k=1}^{K} \sum_{\ell \in \mathscr{Y}_k} \sum_{i=1}^{n} \sum_{j=1}^{S} \left| y^{(i,\ell)}(t_j) - \mu_{k,j}^{(i)} \right|^2, \tag{A.6}$$

over the set of clusters $\mathscr{Y} = \{\mathscr{Y}_1,\dots,\mathscr{Y}_K\}$, where $\mu_{k,j}^{(i)} \in \mathbb{R}$ denotes the mean of the observations $y^{(i,\ell)}(t_j)$ over the measurements $\ell \in \mathscr{Y}_k$.

In this way, measurements that exhibit similar temporal evolution are grouped within the same cluster. Each cluster can then be interpreted as a collection of trajectories evolving in a neighborhood of a nominal trajectory, which allows the nonlinear dynamics to be locally approximated through linearization. The resulting linear models are then used to construct the detection matrices employed in the subsequent inference step.

Another potential concern regarding the use of k-means is its sensitivity to initialization. Indeed, different initial choices of cluster centroids may in principle lead to different clustering outcomes. In my numerical experiments, however, I observed that the clustering results are quite robust provided that the dynamics is simulated for a sufficiently long time and that the number of time samples S is sufficiently large. Under these conditions, the trajectories exhibit clear dynamical patterns, which makes the clustering procedure relatively insensitive to the initial conditions of the k-means algorithm.

A.4 Additional Simulations with Modified Parameters

In the numerical experiments reported in the main text, the model parameters are randomly drawn from predefined sets. As noted during the review, these ranges may appear somewhat arbitrary. The main purpose of these choices is to ensure heterogeneity in the model parameters while avoiding parameter values that would lead to unstable or divergent dynamics.

To further verify that the performance of the proposed method does not rely on a specific choice of parameters, I performed additional simulations using different parameter ranges. In particular, I considered a hyperchain system in which the self-loop coefficients were selected from the set {-1.1, -1.0, -0.9, -0.8, -0.7}, while the interaction coefficients were drawn from the sets {-8.0 + 0.5k | k=0,1,...,12}

$\cup \{2.0 + 0.5k \mid k=0,1,\ldots,12\}$. These values differ from those used in the simulations reported in the main text. Nevertheless, the inference results remain comparable with the original experiments, confirming that the proposed algorithm correctly identifies the network size even when the model parameters are varied (see Fig. A.1).

A systematic exploration of the parameter space would require a large campaign of numerical simulations and is beyond the scope of the present work, whose main goal is methodological. However, these additional experiments provide further evidence of the robustness of the proposed approach.

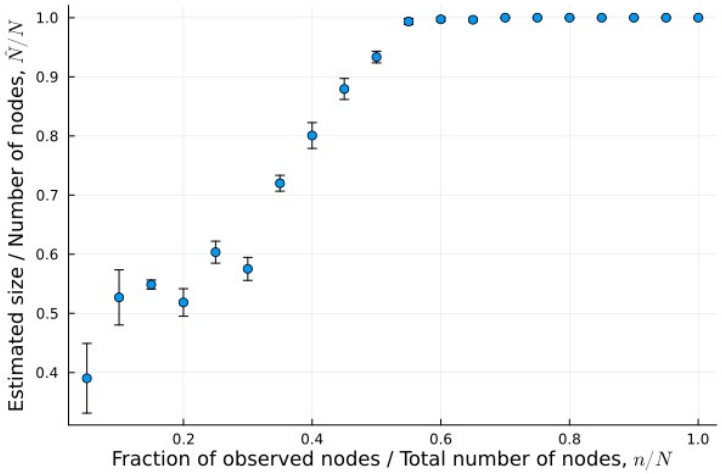


Figure A.1. Inference of network size from measurements at multiple nodes for a hyperchain with modified parameters.

A.5 Dependence on the time-series length

The proposed algorithm requires the time-series length S to satisfy $S > N$ in order to construct a valid detection matrix. In main experiments I considered $N = 20$, which implies that S cannot be smaller than 20.

In the main text I already reported results obtained with a moderate value $S = 25$. Increasing the time-series length to $S = 50$ did not lead to significant improvements in the inference results, while it increases the computational cost of the simulations.

As an additional verification, I performed a simulation on the hyperchain dynamics with $S = 21$, which is the smallest admissible value satisfying $S > N$. The

corresponding results are shown in Fig.A.2 and exhibit no substantial differences with respect to those obtained for $S = 25$ (see in the main text).

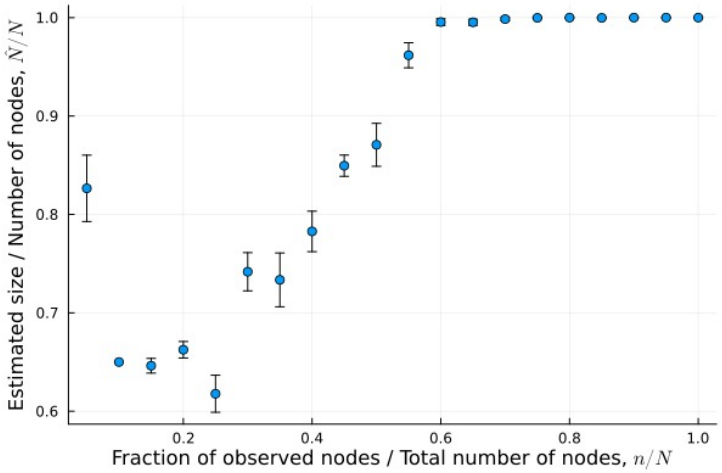


Figure A.2. Inference of network size from measurements at multiple nodes for a hyperchain with S=21.

A.6 Proof of the state reconstruction bound

In the following, I provide a detailed explanation of each step in the proof of the theorem. Consider the nonlinear dynamical system:

$$\dot{x}(t) = F(x,t), \qquad y(t) = h(x(t)) \tag{A.7}$$

and fix a nominal trajectory $x^\star$ on $[t_0, T]$. Let $\delta_x := x - x^\star$, $\delta_y := y - y^\star$ be the displacement.

The linearization along $x^\star$ is

$$\dot{\delta}_x(t) = A(t)\,\delta_x(t), \qquad \delta_y(t) = C(t)\,\delta_x(t), \tag{A.8}$$

with $A(t) = \partial F/\partial x|_{x^\star}$, $C(t) = \partial h/\partial x|_{x^\star}$ and state transition matrix $\Phi(\tau, t_0)$.
The solution of the linearized system is $\delta_x(t) = \Phi(t,t_0)\delta_x(t_0)$ and as a result the output is $\delta_y(t) = C(t)\Phi(t,t_0)\delta_x(t_0)$.
Define the observability Gramian on $[t_0, T]$ as

$$W(t_0,T) := \int_{t_0}^{T} \Phi(\tau,t_0)^\top C(\tau)^\top C(\tau)\,\Phi(\tau,t_0)\,\mathrm{d}\tau\,. \tag{A.9}$$

The output energy over the finite horizon $[t_0, T]$ is defined as

$$\|\delta_y\|_{L^2}^2 \;=\; \int_{t_0}^{T} \|\delta_y(\tau)\|^2 \,\mathrm{d}\tau, \tag{A.10}$$

which connects the output signal δ_y to the input displacement $\delta_x(t_0)$ through the observability Gramian

$$\begin{aligned}
\int_{t_0}^{T} \|\delta_y(\tau)\|^2 \,\mathrm{d}\tau &= \int_{t_0}^{T} \langle C(\tau)\delta_x(\tau), C(\tau)\delta_x(\tau)\rangle \,\mathrm{d}\tau \\
&= \int_{t_0}^{T} \delta_x(\tau)^\top C(\tau)^\top C(\tau)\, \delta_x(\tau) \,\mathrm{d}\tau \\
&= \int_{t_0}^{T} \big(\Phi(\tau,t_0)\delta_x(t_0)\big)^\top C(\tau)^\top C(\tau) \big(\Phi(\tau,t_0)\delta_x(t_0)\big) \,\mathrm{d}\tau \\
&= \delta_x(t_0)^\top \left[\int_{t_0}^{T} \Phi(\tau,t_0)^\top C(\tau)^\top C(\tau)\, \Phi(\tau,t_0) \,\mathrm{d}\tau\right] \delta_x(t_0) \\
&= \delta_x(t_0)^\top W(t_0,T)\, \delta_x(t_0).
\end{aligned} \tag{A.11}$$

Since W is symmetric, it admits the spectral decomposition

$$W = Q\Lambda Q^\top, \tag{A.12}$$

where Q is an orthogonal matrix whose columns are the eigenvectors of W, and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix of eigenvalues with $\lambda_{\min} \le \lambda_i \le \lambda_{\max}$. Then, for any $z \in \mathbb{R}^n$,

$$\lambda_{\min}\|z\|^2 \;\le\; z^\top W z \;\le\; \lambda_{\max}\|z\|^2. \tag{A.13}$$

Note that if W is positive semidefinite, its eigenvalues satisfy the strict inequality $\lambda_{\min} > 0$ under the assumption that the system is observable on $[t_0, T]$, which ensures that no nontrivial initial condition can produce identically zero output.

With $z = \delta_x(t_0)$ and the identity above,

$$\lambda_{\min}\, \|\delta_x(t_0)\|^2 \;\le\; \|\delta_y\|_{L^2}^2 \;\le\; \lambda_{\max}\, \|\delta_x(t_0)\|^2, \tag{A.14}$$

hence

$$\frac{1}{\sqrt{\lambda_{\max}(W)}}\, \|\delta_y\|_{L^2} \;\le\; \|\delta_x(t_0)\| \;\le\; \frac{1}{\sqrt{\lambda_{\min}(W_o)}}\, \|\delta_y\|_{L^2}\,. \tag{A.15}$$

I recall the classical Cauchy–Schwarz inequality: for all $x, y \in \mathbb{R}^n$

$$|x^\top y| \;\le\; \|x\|\,\|y\|. \tag{A.16}$$

From $\dot{\delta}_x(t) = A(t)\delta_x(t)$, and for $\delta_x(t) \neq 0$, the Cauchy–Schwarz inequality yields

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\|\delta_x(t)\| &= \frac{\delta_x(t)^\top \dot{\delta}_x(t)}{\|\delta_x(t)\|} \\ &\le \frac{\|\delta_x(t)\|\,\|\dot{\delta}_x(t)\|}{\|\delta_x(t)\|} \\ &= \|A(t)\delta_x(t)\| \\ &\le \|A(t)\|\,\|\delta_x(t)\|. \end{aligned} \tag{A.17}$$

The use of the Cauchy–Schwarz inequality in the derivation

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\delta_x(t)\| \;\le\; \|A(t)\|\,\|\delta_x(t)\| \tag{A.18}$$

only requires that $A(t)$ is well defined at each $t \in [t_0, T]$, so that its induced norm $\|A(t)\|$ exists.

To proceed, I need an upper bound on the growth of $\|\delta_x(t)\|$. This is provided by the classical Grönwall–Bellman inequality.

Let $\phi(t)$ and $\nu(t)$ be continuous functions defined for $t \ge t_0$, with $\nu(t) \ge 0$, and let ψ be a constant. If

$$\phi(t) \;\le\; \psi + \int_{t_0}^{t} \nu(\sigma)\phi(\sigma)\,\mathrm{d}\sigma, \qquad t \ge t_0, \tag{A.19}$$

then

$$\phi(t) \;\le\; \psi \exp\!\left(\int_{t_0}^{t} \nu(\sigma)\,\mathrm{d}\sigma\right), \qquad t \ge t_0. \tag{A.20}$$

In the present framework, I set $\phi(t) = \|\delta_x(t)\|$, $\nu(t) = \|A(t)\|_2$ and $\psi = \|\delta_x(t_0)\|$. Since by assumption $\|A(\cdot)\| \in L^\infty([t_0, T])$, the conditions of the lemma are satisfied.

Applying Grönwall's inequality to the differential estimate obtained via Cauchy–Schwarz, I obtain

$$\|\delta_x(t)\| \;\le\; \exp\!\left(\int_{t_0}^{t} \|A(\sigma)\|_2\,\mathrm{d}\sigma\right)\|\delta_x(t_0)\|, \qquad t \in [t_0, T]. \tag{A.21}$$

To conclude, I combine the observability inequality with the Grönwall estimate.

For every $t \in [t_0, T]$, I obtain the bound

$$\|\delta_x(t)\| \;\leq\; \frac{\|\Phi(t,t_0)\|}{\sqrt{\lambda_{\min}(W(t_0,T))}}\,\|\delta_y\|_{L^2} \;\leq\; \frac{\exp\!\left(\int_{t_0}^{t}\|A(\sigma)\|\,\mathrm{d}\sigma\right)}{\sqrt{\lambda_{\min}(W(t_0,T))}}\,\|\delta_y\|_{L^2}. \tag{A.22}$$

From Assumption 1:

$$\alpha = \operatorname*{ess\,sup}_{t\in[t_0,T]} \|A(t)\|. \tag{A.23}$$

Hence

$$\int_{t_0}^{t} \|A(\sigma)\|\,\mathrm{d}\sigma \leq \alpha(t-t_0), \tag{A.24}$$

which implies

$$\exp\!\left(\int_{t_0}^{t} \|A(\sigma)\|\,\mathrm{d}\sigma\right) \leq e^{\alpha(t-t_0)}. \tag{A.25}$$

Therefore, the following bound holds

$$\|\delta_x(t)\| \leq \frac{e^{\alpha(t-t_0)}}{\sqrt{\lambda_{\min}(W(t_0,T))}}\|\delta_y\|_{L^2}. \tag{A.26}$$

This coincides with the estimate stated in Theorem 1.

Bibliography

[1] A. Einstein. Über die von della molekularkinetischen theorie della wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen della Physik*, 322:549–560, 1905.

[2] J. C. Maxwell. On the dynamical theory of gases. *Philosophical Transactions of the Royal Society of London*, pages 49–88, 1867.

[3] L. Boltzmann. Weitere studien über das wärmegleichgewicht unter gasmolekülen. In *Kinetische Theorie II: Irreversible Prozesse Einführung und Originaltexte*, pages 115–225. Springer, 1970.

[4] M. Baldovin, R. Marino, and A. Vulpiani. Ergodic observables in non-ergodic systems: The example of the harmonic chain. *Physica A: Statistical Mechanics and its Applications*, 630:129273, 2023.

[5] A. Santoro, F. Battiston, M. Lucas, G. Petri, and E. Amico. Higher-order connectomics of human brain function reveals local topological signatures of task decoding, individual identification, and behavior. *Nature Communications*, 15:10244, 2024.

[6] G. Petri, P. Expert, F. Turkheimer, R. Carhart-Harris, D. Nutt, P. J. Hellyer, and F. Vaccarino. Homological scaffolds of brain functional networks. *Journal of The Royal Society Interface*, 11:20140873, 2014.

[7] C. J. Honey, R. Kötter, M. Breakspear, and O. Sporns. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences*, 104:10240–10245, 2007.

[8] A. Patania, G. Petri, and F. Vaccarino. The shape of collaborations. *EPJ Data Science*, 6:18, 2017.

[9] M. Ye, L. Zino, Z. Mlakar, J. W. Bolderdijk, H. Risselada, B. M. Fennis, and M. Cao. Collective patterns of social diffusion are shaped by individual inertia and trend-seeking. *Nature Communications*, 12:5698, 2021.

[10] C. Castellano, S. Fortunato, and V. Loreto. Statistical physics of social dynamics. *Reviews of Modern Physics*, 81:591–646, 2009.

[11] D. Centola and M. Macy. Complex contagions and the weakness of long ties. *American Journal of Sociology*, 113:702–734, 2007.

[12] G. Cencetti, F. Battiston, B. Lepri, and M. Karsai. Temporal properties of higher-order interactions in social networks. *Scientific Reports*, 11:7028, 2021.

[13] A. Antelmi, G. Cordasco, C. Spagnuolo, and P. Szufel. Social influence maximization in hypergraphs. *Entropy*, 23:796, 2021.

[14] S. Arya, A. B. George, and J. P. O'Dwyer. Sparsity of higher-order landscape interactions enables learning and prediction for microbiomes. *Proceedings of the National Academy of Sciences*, 120:e2307313120, 2023.

[15] J. Friedman, L. M. Higgins, and J. Gore. Community structure follows simple assembly rules in microbial microcosms. *Nature Ecology & Evolution*, 1:0109, 2017.

[16] J. M. Levine, J. Bascompte, P. B. Adler, and S. Allesina. Beyond pairwise mechanisms of species coexistence in complex communities. *Nature*, 546: 56–64, 2017.

[17] S. Cui, Q. Zhao, G. Zhang, H. Jardon-Kojakhmetov, and M. Cao. Analysis of higher-order lotka-volterra models: Application of s-tensors and the polynomial complementarity problem. *IEEE Transactions on Automatic Control*, pages 1–16, 2025.

[18] M. Tyloo and R. Delabays. System size identification from sinusoidal probing in diffusive complex networks. *Journal of Physics: Complexity*, 2: 025016, 2021.

[19] M. Porfiri. Validity and limitations of the detection matrix to determine hidden units and network size from perceptible dynamics. *Physical Review Letters*, 124:168301, 2020.

[20] X. Tang, W. Huo, Y. Yuan, X. Li, L. Shi, H. Ding, and J. Kurths. Dynamical network size estimation from local observations. *New Journal of Physics*, 22:093031, 2020.

[21] P. De Lellis and M. Porfiri. Inferring the size of a collective of self-propelled vicsek particles from the random motion of a single unit. *Communications Physics*, 5:86, 2022.

[22] L. Neuhäuser, M. Scholkemper, F. Tudisco, and M. T. Schaub. Learning the effective order of a hypergraph dynamical system. *Science Advances*, 10: eadh4053, 2024.

[23] H. Haehne, J. Casadiego, J. Peinke, and M. Timme. Detecting hidden units and network size from perceptible dynamics. *Physical Review Letters*, 122: 158301, 2019.

[24] R. Succar, A. Boldini, and M. Porfiri. Detecting hidden states in stochastic dynamical systems. *Physical Review Research*, 6:013149, 2024.

[25] R. E. Kalman. On the general theory of control systems. In *Proceedings of the First International Conference on Automatic Control*, pages 481–493, Moscow, USSR, 1960. IFAC.

[26] P. Celli and M. Porfiri. The detection matrix as a model-agnostic tool to estimate the number of degrees of freedom in mechanical systems and engineering structures. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 32:033106, 2022.

[27] G. Börner, H. Haehne, J. Casadiego, and M. Timme. Revealing system dimension from single-variable time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 33, 2023.

[28] S. Boccaletti, P. De Lellis, C. Del Genio, K. Alfaro-Bittner, R. Criado, S. Jalan, and M. Romance. The structure and dynamics of networks with higher order interactions. *Physics Reports*, 1018:1–64, 2023.

[29] L. Gallo, R. Muolo, L. V. Gambuzza, V. Latora, M. Frasca, and T. Carletti. Synchronization induced by directed higher-order interactions. *Communications Physics*, 5:263, 2022.

[30] R. Lambiotte, M. Rosvall, and I. Scholtes. From networks to optimal higher-order models of complex systems. *Nature Physics*, 15:313–320, 2019.

[31] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, and G. Petri. Networks beyond pairwise interactions: Structure and dynamics. *Physics Reports*, 874:1–92, 2020.

[32] F. Battiston, E. Amico, A. Barrat, G. Bianconi, G. Ferraz de Arruda, B. Franceschiello, I. Iacopini, S. Kéfi, V. Latora, Y. Moreno, M. M. Murray, T. P. Peixoto, F. Vaccarino, and G. Petri. The physics of higher-order interactions in complex systems. *Nature Physics*, 17:1093–1098, 2021.

[33] S. Majhi, M. Perc, and D. Ghosh. Dynamics on higher-order networks: a review. *Journal of The Royal Society Interface*, 19:20220043, 2022.

[34] J. Pickard, A. Surana, A. Bloch, and I. Rajapakse. Observability of hypergraphs. In *Proceedings of the 62nd IEEE Conference on Decision and Control*, pages 2445–2451, Marina Bay Sands, Singapore, 2023.

[35] X.-J. Xu, S. He, and L.-J. Zhang. Dynamics of the threshold model on hypergraphs. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 32: 023125, 2022.

[36] D. Schlager, K. Clauß, and C. Kuehn. Stability analysis of multiplayer games on adaptive simplicial complexes. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 32:053128, 2022.

[37] C. Zhang, H. Yang, S. Cui, B. Jiang, and M. Cao. Global and local observability of hypergraphs. *arXiv preprint arXiv:2402.00078*, 2024.

[38] R. Delabays, G. De Pasquale, F. Dörfler, and Y. Zhang. Hypergraph reconstruction from dynamics. *Nature Communications*, 16:2691, 2025.

[39] A. Isidori. *Nonlinear Control Systems*. Springer, London, UK, 3rd edition, 1999.

[40] J. Mathews and R. L. Walker. *Mathematical Methods of Physics*. W.A. Benjamin, San Francisco CA, US, 2nd edition, 1970.

[41] W. J. Rugh. *Lynear System Theory*. Pearson, London, UK, 2nd edition, 1996.

[42] F. Lo Iudice, F. Sorrentino, and F. Garofalo. On node controllability and observability in complex dynamical networks. *IEEE Control Systems Letters*, 3:847–852, 2019.

[43] G. Reissig, C. Hartung, and F. Svaricek. Strong structural controllability and observability of linear time-varying systems. *IEEE Transactions on Automatic Control*, 59:3087–3092, 2014.

[44] E. D. Sontag. *Mathematical Control Theory*. Springer, New York NY, US, 2nd edition, 1998.

[45] V. Thibeault, A. Allard, and P. Desrosiers. The low-rank hypothesis of complex systems. *Nature Physics*, 20:294–302, 2024.

[46] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn. Lyapunov characteristic exponents for smooth dynamical systems and for hamiltonian systems; a method for computing all of them. part 1: Theory. *Meccanica*, 15: 9–20, 1980.

[47] J. P. Eckmann and D. Ruelle. Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics*, 57:617–656, 1985.

[48] L.-S. Young. Mathematical theory of lyapunov exponents. *Journal of Physics A: Mathematical and Theoretical*, 46:254001, 2013.

[49] J. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pages 281–297, University of California, Berkeley, US, 1967.

[50] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási. Observability of complex systems. *Proceedings of the National Academy of Sciences*, 110:2460–2465, 2013.

[51] A. Braunstein, L. Dall'Asta, G. Semerjian, and L. Zdeborová. Network dismantling. *Proceedings of the National Academy of Sciences*, 113:12368–12373, 2016.