SIAM J. MATRIX ANAL. APPL.
Vol. 44, No. 4, pp. 1799–1821

IDENTIFIABILITY IN CONTINUOUS LYAPUNOV MODELS*

PHILIPP DETTLING[†], ROSER HOMS[†], CARLOS AMÉNDOLA[§], MATHIAS DRTON[†‡], AND NIELS RICHARD HANSEN[¶]

Abstract. The recently introduced graphical continuous Lyapunov models provide a new approach to statistical modeling of correlated multivariate data. The models view each observation as a one-time cross-sectional snapshot of a multivariate dynamic process in equilibrium. The covariance matrix for the data is obtained by solving a continuous Lyapunov equation that is parametrized by the drift matrix of the dynamic process. In this context, different statistical models postulate different sparsity patterns in the drift matrix, and it becomes a crucial problem to clarify whether a given sparsity assumption allows one to uniquely recover the drift matrix parameters from the covariance matrix of the data. We study this identifiability problem by representing sparsity patterns by directed graphs. Our main result proves that the drift matrix is globally identifiable if and only if the graph for the sparsity pattern is simple (i.e., does not contain directed 2-cycles). Moreover, we present a necessary condition for generic identifiability and provide a computational classification of small graphs with up to 5 nodes.

Key words. identifiability, Lyapunov equation, graphical modeling

MSC codes. 62R01, 62H22, 15A24

DOI. 10.1137/22M1520311

1. Introduction.

Recent work of Fitch (2019) and Varando and Hansen (2020) introduces Lyapunov models as a new paradigm of probabilistic graphical modeling (Maathuis et al., 2019). When capturing cause-effect relations among observations, standard graphical models directly postulate noisy functional relations among the considered random variables (Pearl, 2009; Peters et al., 2017; Spirtes et al., 2000). In contrast, the new Lyapunov models introduce a temporal perspective that simplifies, in particular, modeling of feedback loops. Suppose the data at hand are collected by observing a p-dimensional random vector. Lyapunov models assume the random vector to arise as a one-time cross-sectional observation of a p-dimensional dynamic process in equilibrium. When working in continuous time, the natural model for the process is an Ornstein–Uhlenbeck process $X(t)$ that is given by the stochastic differential equation

$$\mathrm{d}X(t) = M(X(t) - a)\,\mathrm{d}t + D\,\mathrm{d}W(t), \tag{1.1}$$

where $W(t)$ is a Wiener process, and $a \in \mathbb{R}^p$ and $M, D \in \mathbb{R}^{p\times p}$ are nonsingular parameter matrices. In this context, the matrix M is a drift matrix that quantifies

*Received by the editors September 7, 2022; accepted for publication (in revised form) by Z. Drmac July 10, 2023; published electronically December 4, 2023.
https://doi.org/10.1137/22M1520311

Funding: The work of the authors was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme grant 883818. The work of the first author was also supported by the Hanns-Seidel Foundation.

[†]TUM School of Computation, Information and Technology, and Department of Mathematics, Technical University of Munich, 85748 Garching bei München, Germany (philipp.dettling@tum.de, roser.homs@tum.de).

[‡]Munich Data Science Institute, Technical University of Munich, 85748 Garching bei München, Germany (mathias.drton@tum.de).

[§]Institute of Mathematics, Technical University of Berlin, 10623 Berlin, Germany (amendola@math.tu-berlin.de).

[¶]Department of Mathematical Sciences, University of Copenhagen, Copenhagen 2100, Denmark (niels.r.hansen@math.ku.dk).

temporal cause-effect relations among the variables, and $C = DD^\top$ is a positive definite volatility matrix.

If M is a stable matrix, i.e., all its eigenvalues have negative real part, then $X(t)$ has a stationary Gaussian distribution $N(a,\Sigma)$, where the covariance matrix Σ is the unique solution (Risken, 1996) to the continuous Lyapunov equation

$$M\Sigma + \Sigma M^\top + C = 0. \tag{1.2}$$

A graphical continuous Lyapunov model as defined by Fitch (2019) and Varando and Hansen (2020) refines this setup by assuming that the drift matrix $M = (m_{ij})$ exhibits a specific zero pattern that is given by a directed graph. A similar perspective was presented by Young et al. (2019) for discrete time autoregressive models, which leads to an equilibrium covariance matrix solving the *discrete* Lyapunov equation. Based on an estimated covariance matrix $\hat{\Sigma}$, both Fitch (2019) and Varando and Hansen (2020) develop estimation techniques for M using (1.1). They consider a setting where the data are comprised of a sample of independent and identically distributed random observation vectors that are obtained from several independent copies of the multivariate Ornstein–Uhlenbeck process in equilibrium. An application is shown in both works using the data set of Sachs et al. (2005). These data contain measurements of expression of different proteins in human immune system cells that are harvested and subjected to flow cytometry (and thus "destroyed"). The estimation methods of Fitch (2019) and Varando and Hansen (2020) are applied to obtain estimates of the protein signaling network, recovering substantial parts of the version accepted among biologists. Naturally, statisticians then aim to provide theoretical guarantees for the methods applied. Under the assumption of parameter identifiability, a consistency result for the estimation method of Fitch (2019) is derived in Dettling et al. (2022). Parameter identifiability means that given a fixed support of the matrix M and positive definite matrices C,Σ, we are able to uniquely recover the entries in M. This central question for statistical theory for Lyapunov models is the object of study in this work.

Organization and results of the paper. In section 2 we introduce graphical continuous Lyapunov models and motivate the question of identifiability with the help of the directed 3-cycle as a running example. In section 3 we formally introduce the notions of generic and global identifiability and make some preliminary observations. In section 4, we explain the structure of the matrix $A(\Sigma)$ that arises from (half-)vectorization of the Lyapunov equation. We also highlight how the rank of a submatrix of $A(\Sigma)$ determines generic and global identifiability of a model. Exploiting block structure in the relevant submatrix of $A(\Sigma)$, we prove global identifiability for all directed acyclic graphs (DAGs) in section 5. Our proof also yields that the models given by DAGs are closed algebraic subsets of PD_p, and that the models associated to complete DAGs are equal to PD_p (Corollary 5.4). In section 6, we turn to cyclic graphs for which the relevant matrices no longer exhibit block structure. We demonstrate that for small graphs the approach studying factorizations of determinants can still be implemented using sum of squares (SOS) methods to certify that the relevant polynomials are positive on PD_p. In section 7 we present our main result (Theorem 7.1), which proves that global model identifiability holds if the underlying graph is simple (i.e., does not contain any 2-cycle). If C is diagonal—the case of primary practical interest—then the requirement that the graph be simple is also necessary for global identifiability. Moreover, we are able to show that for all $C \in \mathrm{PD}_p$, all simple graphs yield models $\mathcal{M}_{G,C}$ that are closed algebraic subsets of PD_p. We discuss further the diagonal hypothesis on C in Appendix A. In section 8, we turn to the weaker notion of generic identifiability, for which we develop a necessary criterion and computationally classify all nonsimple graphs with up to 5 nodes. The paper concludes in section 9. Some details on the structure of the matrix $A(\Sigma)$ and the factorization of its minors are deferred to Appendix B.

The code we used for our computations is available at the repository website https://mathrepo.mis.mpg.de/LyapunovIdentifiability.

2. Preliminaries.

A graphical continuous Lyapunov model as defined in Dettling et al. (2022) considers the setup that the drift matrix $M = (m_{ij})$ exhibits a specific zero pattern that is given by a directed graph G on the set of nodes $[p] = \{1, \ldots, p\}$, with $m_{ji} = 0$ whenever $i \to j$ is not an edge in G. In this setting our graphs will always include self-loops $i \to i$.

Example 2.1. The directed 3-cycle G with vertex set $V = \{1,2,3\}$ and edge set $E = \{1 \to 1, 2 \to 2, 3 \to 3, 1 \to 2, 2 \to 3, 3 \to 1\}$, which is displayed in Figure 1, encodes drift matrices of the form

$$M = \begin{pmatrix} m_{11} & 0 & m_{13} \\ m_{21} & m_{22} & 0 \\ 0 & m_{32} & m_{33} \end{pmatrix}.$$

The Lyapunov equation from (1.2) is a symmetric matrix equation providing $p(p+1)/2$ constraints. In contrast, the drift matrix M is a $p \times p$ matrix that need not be symmetric. Hence, without any assumptions on its structure, M is never uniquely determined by the covariance matrix Σ of the observations. For graphical Lyapunov models, this leads to a key identifiability question: For which sparsity patterns can the drift matrix M be recovered from the positive definite covariance matrix Σ? Our treatment of this question will assume that the volatility matrix C is a known positive definite matrix. While some of our results hold for all positive definite C, others require the assumption that C is diagonal. This is a sensible assumption as it corresponds to the setting of uncorrelated noise. A special case is the assumption $C = 2I_p$ that covers the natural setting of homoscedastic noise.

Remark 2.2. Evidently, if a matrix Σ solves the Lyapunov equation (1.2) for a pair (M, C) then Σ also solves the equation given by $(\gamma M, \gamma C)$ for any $\gamma \in \mathbb{R}$. An implication of this fact is that our results on recovery of M for fixed C also address the setting of models in which $C = \gamma C'$, with C' known and positive definite but $\gamma > 0$ an unknown parameter. In this latter setting, one can only hope to recover M up to a scalar multiple and this is possible if and only if M can be recovered uniquely in the setting where we fix $C = C'$.

Before proceeding to illustrate the identifiability problem for Example 2.1, we give a formal definition of graphical continuous Lyapunov models as sets of covariance matrices. We write PD_p for the cone of $p \times p$ positive definite matrices.

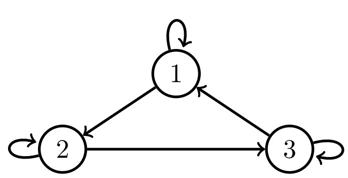

FIG. 1. *The directed* 3*-cycle.*

DEFINITION 2.3. *Let $G = (V, E)$ be a directed graph with vertex set $V = [p]$ and an edge set E that includes all self-loops $i \to i$, $i \in [p]$. We write $\mathbb{R}^E$ for the space of matrices $M = (m_{ij}) \in \mathbb{R}^{p\times p}$ with $m_{ji} = 0$ whenever $i \to j \notin E$. Given a choice of $C \in \mathrm{PD}_p$, the graphical continuous Lyapunov model of G is the set of covariance matrices*

$$\mathcal{M}_{G,C} = \left\{\Sigma \in \mathrm{PD}_p : M\Sigma + \Sigma M^\top = -C \text{ for some } M \in \mathbb{R}^E\right\}.$$

Remark 2.4. Let $\mathrm{Stab}(E) \subseteq \mathbb{R}^E$ be the subset of stable matrices, which is always nonempty and open, since E is assumed to always include all self-loops; see Definition 2.3. In particular, this implies that $\dim(\mathrm{Stab}(E)) = |E|$. When C is positive definite, the Lyapunov equation from (1.2) has a positive definite solution Σ if and only if M is stable (Bhaya et al., 2003, Theorem 1.1). Hence, the definition of the model $\mathcal{M}_{G,C}$ remains unchanged if we replace the requirement $M \in \mathbb{R}^E$ by $M \in \mathrm{Stab}(E)$.

The identifiability question we pose asks whether a covariance matrix Σ in the model $\mathcal{M}_{G,C}$ may simultaneously solve the Lyapunov equation for more than one choice of a matrix $M \in \mathbb{R}^E$. In other words, we study the injectivity of the (rational) parametrization map

$$\begin{aligned} \phi_{G,C} : \mathrm{Stab}(E) &\to \mathrm{PD}_p \\ M &\mapsto \Sigma(M, C), \end{aligned} \tag{2.1}$$

where $\Sigma(M, C)$ is the unique matrix Σ that solves the Lyapunov equation given by the stable matrix M and positive definite C. See (4.1) for details on this uniqueness.

By vectorization, the Lyapunov equation (1.2) is transformed into the linear equation system

$$A(\Sigma)\mathrm{vec}(M) = -\mathrm{vech}(C), \tag{2.2}$$

where $\mathrm{vech}(C)$ is the half-vectorization of a fixed symmetric matrix $C \in \mathrm{PD}_p$, and $A(\Sigma)$ is a $p(p + 1)/2 \times p^2$ matrix depending on Σ whose form will be discussed in section 4.

Example 2.5. In the case of $p = 3$ variables the matrix $A(\Sigma)$ equals

$$\begin{array}{c} \\ (1,1) \\ (1,2) \\ (1,3) \\ (2,2) \\ (2,3) \\ (3,3) \end{array}
\begin{array}{c}
\begin{array}{ccccccccc} 1\to1 & 1\to2 & 1\to3 & 2\to1 & 2\to2 & 2\to3 & 3\to1 & 3\to2 & 3\to3 \end{array} \\
\begin{pmatrix}
2\Sigma_{11} & 0 & 0 & 2\Sigma_{12} & 0 & 0 & 2\Sigma_{13} & 0 & 0 \\
\Sigma_{12} & \Sigma_{11} & 0 & \Sigma_{22} & \Sigma_{12} & 0 & \Sigma_{23} & \Sigma_{13} & 0 \\
\Sigma_{13} & 0 & \Sigma_{11} & \Sigma_{23} & 0 & \Sigma_{12} & \Sigma_{33} & 0 & \Sigma_{13} \\
0 & 2\Sigma_{12} & 0 & 0 & 2\Sigma_{22} & 0 & 0 & 2\Sigma_{23} & 0 \\
0 & \Sigma_{13} & \Sigma_{12} & 0 & \Sigma_{23} & \Sigma_{22} & 0 & \Sigma_{33} & \Sigma_{23} \\
0 & 0 & 2\Sigma_{13} & 0 & 0 & 2\Sigma_{23} & 0 & 0 & 2\Sigma_{33}
\end{pmatrix}
\end{array},$$

where the column index $i \to j$ corresponds to entry m_{ji} of the drift matrix $M = (m_{ij})$.

Given a graph G with $p(p + 1)/2$ edges, unique solvability of (2.2) for $M \in \mathbb{R}^E$ is equivalent to a certain maximal square submatrix of $A(\Sigma)$ being invertible. This submatrix is formed by all columns of $A(\Sigma)$ corresponding to edges of the graph. Observe that two columns indexed by $i \to j$ and $k \to l$ have the same zero pattern whenever $j = l$. This motivates ordering the columns of $A(\Sigma)_{\cdot,E}$ increasingly with

$$i \to j < k \to l \quad \text{if } j < l \text{ or } j = l, i < k. \tag{2.3}$$

Moreover, note that for simple graphs there is a natural pairing between pairs (i,j) with $i \leq j$ and edges between i and j. In this case, we will order rows accordingly with their corresponding pair (i,j).

Example 2.6. Consider the 3-cycle G from Example 2.1. The submatrix of $A(\Sigma)$ in Example 2.5 associated to G is

$$A(\Sigma)_{\cdot,E} = \begin{array}{c} \\ (1,1) \\ (1,3) \\ (1,2) \\ (2,2) \\ (2,3) \\ (3,3) \end{array} \begin{array}{c} \begin{array}{cccccc} 1\to 1 & 3\to 1 & 1\to 2 & 2\to 2 & 2\to 3 & 3\to 3 \end{array} \\ \begin{pmatrix} 2\Sigma_{11} & 2\Sigma_{13} & 0 & 0 & 0 & 0 \\ \Sigma_{13} & \Sigma_{33} & 0 & 0 & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{12} & \Sigma_{23} & \Sigma_{11} & \Sigma_{12} & 0 & 0 \\ 0 & 0 & 2\Sigma_{12} & 2\Sigma_{22} & 0 & 0 \\ 0 & 0 & \Sigma_{13} & \Sigma_{23} & \Sigma_{22} & \Sigma_{23} \\ 0 & 0 & 0 & 0 & 2\Sigma_{23} & 2\Sigma_{33} \end{pmatrix} \end{array}.$$

To show invertibility of $A(\Sigma)_{\cdot,E}$, we may inspect its determinant, which factorizes as

$$2^3 \cdot \det(\Sigma) \cdot (\Sigma_{11}\Sigma_{22}\Sigma_{33} - \Sigma_{12}\Sigma_{13}\Sigma_{23}). \tag{2.4}$$

All displayed factors are positive when Σ is positive definite. Indeed, $\det(\Sigma) > 0$ and the fact that $\det(\Sigma_{ij,ij}) = \Sigma_{ii}\Sigma_{jj} - \Sigma_{ij}^2 > 0$ for all $i \neq j$ implies that $\Sigma_{11}^2\Sigma_{22}^2\Sigma_{33}^2 > \Sigma_{12}^2\Sigma_{13}^2\Sigma_{23}^2$, which clarifies that the last factor is also positive. Alternatively, we can show this using the identity

$$\begin{aligned}(\Sigma_{11}\Sigma_{22}\Sigma_{33})^2 &- (\Sigma_{12}\Sigma_{13}\Sigma_{23})^2 \\ &= (\Sigma_{13}\Sigma_{23})^2 \det(\Sigma_{12,12}) + \Sigma_{11}\Sigma_{22}\Sigma_{23}^2 \det(\Sigma_{13,13}) + \Sigma_{11}^2\Sigma_{22}\Sigma_{33} \det(\Sigma_{23,23}) > 0.\end{aligned}$$

We conclude that when G is the 3-cycle, then for *all* covariance matrices $\Sigma \in \mathcal{M}_{G,C} \subseteq \mathrm{PD}_3$ there is a unique matrix $M \in \mathbb{R}^E$ such that $\Sigma = \phi_{G,C}(M)$. We will refer to this property as the 3-cycle defining a *globally identifiable* model. Note that our argument also shows that $\mathcal{M}_{G,C} = \mathrm{PD}_3$.

This small example already reveals some of the subtleties arising when analyzing identifiability of continuous Lyapunov models. The problem can be reduced to determining whether a particular submatrix that is sparsely populated with covariances has full rank (see Lemmas 4.3 and 6.4) but the resulting matrices have involved graph-dependent structures. The choice of ordering in (2.3) is especially insightful for DAGs. After sorting the nodes such that if $i \to j$, then $i \leq j$, any DAG yields a block upper-triangular matrix, as in Example 5.2, from which identifiability for all associated models follows (Theorem 5.3). For cyclic graphs, however, the polynomials that appear while factoring determinants, as in (2.4), quickly increase in complexity, and it is not easy to determine whether they are nonzero. In our main result (Theorem 7.1) we thus consider alternative spectral arguments that use the stability of the drift matrix M in order to derive identifiability.

3. Notions of identifiability. We begin by recalling the concept of fibers that is useful in defining the different notions of identifiability we study in subsequent sections. Let $C \in \mathrm{PD}_p$, and let $\mathcal{M}_{G,C}$ be the graphical continuous Lyapunov model associated to a directed graph $G = (V,E)$ with vertex set $V = [p]$ and edge set E. Let $\phi_{G,C}$ be the parametrization from (2.1). The *fiber* of a matrix $M_0 \in \mathrm{Stab}(E)$ is the set

$$\mathcal{F}_{G,C}(M_0) = \{M \in \mathrm{Stab}(E) : \phi_{G,C}(M) = \phi_{G,C}(M_0)\}. \tag{3.1}$$

In other words, a fiber comprises all drift matrices $M \in \mathbb{R}^E$ whose Lyapunov equation (for the fixed matrix $C \in \mathrm{PD}_p$) is solved by a given covariance matrix Σ.

We will consider three natural notions of identifiability.

Definition 3.1. *Let $\mathcal{M}_{G,C}$ be the graphical continuous Lyapunov model given by a directed graph $G = (V, E)$ with $V = [p]$ and $C \in \mathrm{PD}_p$. The model $\mathcal{M}_{G,C}$ is*

(i) *globally identifiable if $\mathcal{F}_{G,C}(M_0) = \{M_0\}$ for all $M_0 \in \mathrm{Stab}(E)$;*

(ii) *generically identifiable if $\mathcal{F}_{G,C}(M_0) = \{M_0\}$ for almost all $M_0 \in \mathrm{Stab}(E)$, i.e., the matrices with $\mathcal{F}_{G,C}(M_0) \neq \{M_0\}$ form a Lebesgue null set in $\mathbb{R}^E$;*

(iii) *nonidentifiable if $|\mathcal{F}_{G,C}(M_0)| = \infty$ for all $M_0 \in \mathrm{Stab}(E)$.*

Remark 3.2. The generic properties we prove in this paper are derived by showing that they hold outside a strict subset of $\mathrm{Stab}(E)$ that is described by polynomials in the entries of the drift matrix; see, e.g., Lemma 4.3. Hence, in a generically identifiable model the exception set is not merely a set of Lebesgue measure zero, but also a lower-dimensional algebraic subset of $\mathrm{Stab}(E)$.

Remark 3.3. Characterizing identifiability is also a key problem for standard directed graphical models; see Drton (2018) and Sullivant (2018, Chap. 16) for a discussion of the different notions of identifiability in this context. For standard graphical models, necessary and sufficient conditions for global identifiability have been obtained (Drton et al., 2011). However, many models of interest are not globally identifiable, and much work has also gone into criteria for generic identifiability (Brito and Pearl, 2006; Drton and Weihs, 2016; Foygel et al., 2012; Kumor et al., 2019).

The 3-cycle from Example 2.6 is an example of global identifiability. Under global identifiability, no two distinct stable matrices may define the same covariance matrix in the model given by the graph. Unfortunately, this is not always the case.

Example 3.4. Consider the 2-cycle $G = (V, E)$ with $V = \{1, 2\}$ and $E = \{1 \to 1, 2 \to 2, 1 \to 2, 2 \to 1\}$. Then $\phi_{G,C}$ maps the 4-dimensional parameter space $\mathrm{Stab}(E)$ to the 3-dimensional PD_2-cone. Hence, when computing any fiber we have to solve a linear system that is underdetermined, with three equations in four unknowns. Therefore, $\mathcal{M}_{G,C}$ is nonidentifiable, no matter the choice of $C \in \mathrm{PD}_2$.

The example just given generalizes as follows.

Lemma 3.5. *Let $G = (V, E)$ be a directed graph with vertex set $V = [p]$, and let $C \in \mathrm{PD}_p$. If $|E| > \dim(\mathcal{M}_{G,C})$, i.e., the number of free parameters in* $\mathrm{Stab}(E)$ *is greater than the dimension of the model, then $\mathcal{M}_{G,C}$ is nonidentifiable. In particular, all graphs with $|E| > p(p + 1)/2$ give nonidentifiable models.*

Proof. By the Hurwitz criterion, the set of sparse stable matrices $\mathrm{Stab}(E)$ is semialgebraic; see Horn and Johnson (1991, Theorem 2.3.3). As its dimension is $\dim(\mathrm{Stab}(E)) = |E| > \dim(\mathcal{M}_{G,C})$, it follows that the rational map $\phi_{G,C}$ defined on $\mathrm{Stab}(E)$ is generically infinite-to-one; see, e.g., Barber et al., (2022, Lemma 2.5). Apply Lemma 4.3 below to conclude that all fibers are infinite. ∎

A straightforward but very useful fact when studying global identifiability is that if a graph $G = (V, E)$ yields a globally identifiable model, then so does every subgraph $H = (V, E')$, $E' \subseteq E$, that is obtained by removing edges of the form $i \to j$ with $i \neq j$. We record this fact as follows.

Proposition 3.6. *Let $\mathcal{M}_{G,C}$ be a globally identifiable model given by a directed graph $G = (V, E)$ with $V = [p]$ and $C \in \mathrm{PD}_p$. Let $E' \subset E$ be a subset of the edges. Then the model $\mathcal{M}_{H,C}$ defined by the subgraph $H = (V, E')$ is globally identifiable.*

Proof. It holds that $\operatorname{Stab}(E') \subseteq \operatorname{Stab}(E)$. Therefore, for every matrix $M_0 \in \operatorname{Stab}(E')$, we have $\mathcal{F}_{H,C}(M_0) \subseteq \mathcal{F}_{G,C}(M_0) = \{M_0\}$, where the last equality is due to the assumed global identifiability of $\mathcal{M}_{G,C}$. ∎

In the case where C is diagonal, further conclusions can be made.

PROPOSITION 3.7. *Let $G=(V,E)$ be a directed graph with $V=[p]$. Let $C \in \mathrm{PD}_p$ be diagonal, and let I_p be the $p \times p$ identity matrix. Then the models for C versus I_p are isomorphic, and so are their fibers:*

(i) $\mathcal{M}_{G,C} = C^{1/2}\mathcal{M}_{G,I_p}C^{1/2}$, *and*

(ii) $\mathcal{F}_{G,C}(M) = \mathcal{F}_{G,I_p}(C^{1/2}MC^{-1/2})$ *for all* $M \in \operatorname{Stab}(E)$.

In particular, $\mathcal{M}_{G,C}$ is globally/generically identifiable if and only if $\mathcal{M}_{G,I_p}$ is globally/generically identifiable.

Proof. Since C is diagonal, the similarity transformation $\tau_1 : M \mapsto C^{-1/2}MC^{1/2}$ is an automorphism of $\mathbb{R}^E$, with $\tau_1(\operatorname{Stab}(E)) = \operatorname{Stab}(E)$. Define a second linear map $\tau_2 : \Sigma \mapsto C^{-1/2}\Sigma C^{-1/2}$, an automorphism of the space of symmetric matrices with $\tau_2(\mathrm{PD}_p) = \mathrm{PD}_p$. Now

$$
\begin{aligned}
&M\Sigma + \Sigma M^\top + C = 0 \iff \\
&\quad (C^{-1/2}MC^{1/2})(C^{-1/2}\Sigma C^{-1/2}) + (C^{-1/2}\Sigma C^{-1/2})(C^{-1/2}MC^{1/2})^\top + I_p = 0.
\end{aligned}
$$

Thus, $\mathcal{M}_{G,I_p} = \tau_2(\mathcal{M}_{G,C})$ and $\mathcal{F}_{G,I_p}(M) = \mathcal{F}_{G,C}(\tau_1^{-1}(M))$. ∎

In Proposition 3.6 only edges are removed when forming a subgraph. When C is diagonal we may strengthen the result to subgraphs in which we also remove vertices; compare Drton et al. (2011, Lemma 1) in the context of standard graphical models.

PROPOSITION 3.8. *Let $G = (V,E)$ be a directed graph with $V = [p]$, and let $H = (V',E')$ be a subgraph with $V' \subseteq V$ and $E' \subseteq E$. If the model $\mathcal{M}_{G,C}$ is globally identifiable for a diagonal matrix $C \in \mathrm{PD}_p$, then $\mathcal{M}_{H,C'}$ is globally identifiable for all diagonal matrices $C' \in \mathrm{PD}_{p'}$, where $p' = |V'|$.*

Proof. By Proposition 3.6, it suffices to prove that removing an isolated vertex from G preserves global identifiability of the model for C diagonal. By Proposition 3.7, we may assume that $C = I_p$ and $C' = I_{p-1}$, where p is an isolated node of G. Let $M \in \operatorname{Stab}(E)$, and let $M_{[p-1],[p-1]}$ be the submatrix comprising the first $p-1$ rows and columns. Since p is isolated, the pth row and column of M is zero with the exception of the diagonal entry m_{pp}. It is not difficult to see that $\Sigma = \phi_{G,I_p}(M)$ also has its pth row and column equal to zero except for the diagonal entry which equals $\Sigma_{pp} = -1/(2m_{pp})$. Hence, the entry m_{pp} is always uniquely determined by Σ, and we conclude that the cardinality of the fiber $\mathcal{F}_{G,I_p}(M)$ is equal to the cardinality of $\mathcal{F}_{H,I_{p-1}}(M_{[p-1],[p-1]})$. Since every matrix in $\operatorname{Stab}(E')$ is a submatrix $M_{[p-1],[p-1]}$ of a matrix $M \in \operatorname{Stab}(E)$, the model $\mathcal{M}_{H,I_{p-1}}$ is globally identifiable. ∎

Combining Proposition 3.8 with Example 3.4, we obtain that the graph of a globally identifiable model cannot contain any 2-cycles.

DEFINITION 3.9. *A directed graph $G = (V,E)$ is simple if it is free of 2-cycles, i.e., there do not exist 2 distinct nodes $i,j \in V$ such that $i \to j \in E$ and $j \to i \in E$. Otherwise, we call G nonsimple.*

PROPOSITION 3.10. *If a directed graph $G = (V,E)$, $V = [p]$, defines a globally identifiable model $\mathcal{M}_{G,C}$ when $C \in \mathrm{PD}_p$ is diagonal, then G must be simple.*

Remark 3.11. Propositions 3.8 and 3.10 may fail for nondiagonal $C \in \mathrm{PD}_p$. See Appendix A for an example.

Unfortunately, similar subgraph arguments cannot be made for generic instead of global identifiability. Indeed, generic identifiability may be lost but also restored when removing an edge. Example 8.4 illustrates this phenomenon.

4. Rank conditions. In this section, we discuss solving the Lyapunov equation (1.2) for the generally nonsymmetric drift matrix M given the symmetric matrices Σ and C. We will proceed by vectorizing the Lyapunov equation, and we will state necessary and sufficient conditions for identifiability based on the ranks of submatrices of the coefficient matrix $A(\Sigma)$ of the vectorized Lyapunov equation.

First, recall that when the matrices M and C are given, the continuous Lyapunov equation from (1.2) is uniquely solvable for the symmetric matrix Σ if and only if no two eigenvalues of M add up to zero. This well-known fact can be shown by vectorizing the equation to

$$(I_p \otimes M + M \otimes I_p)\mathrm{vec}(\Sigma) = -\mathrm{vec}(C), \tag{4.1}$$

where $\otimes$ is the Kronecker product and $\mathrm{vec}(\cdot)$ is the columnwise vectorization of a matrix; see, e.g., Bernstein (2011). The coefficient matrix $I_p \otimes M + M \otimes I_p$ is a Kronecker sum, and it follows that its eigenvalues are the pairwise sums of the eigenvalues of M. If we now additionally assume that C is positive definite, then Lyapunov's theorem (Horn and Johnson, 1991, Theorem 2.2.1) yields that the Lyapunov equation from (1.2) has a unique positive definite solution Σ if and only if M is a stable matrix.

However, solving for M given two symmetric (and in our context positive definite) matrices Σ and C is a more difficult question. In general, it is not possible to have a unique solution for M due to the dimensionality problems mentioned in Lemma 3.5. The graphical perspective of the Lyapunov models motivates considering sparse matrices M and asking the solvability question in a new light, as we illustrated in Example 2.6.

LEMMA 4.1. *Vectorizing the Lyapunov equation* (1.2)*, we obtain the system*

$$((\Sigma \otimes I_p) + (I_p \otimes \Sigma)K_p)\mathrm{vec}(M) = -\mathrm{vec}(C), \tag{4.2}$$

where K_p is the $p \times p$ commutation matrix.

The commutation matrix K_p is the symmetric permutation matrix that transforms the vectorization of a $p \times p$ matrix to the vectorization of its transpose (Magnus and Neudecker, 1999, p. 54).

Proof of Lemma 4.1. It holds that

$$\begin{aligned}\mathrm{vec}(M\Sigma + \Sigma M^\top) &= \mathrm{vec}(M\Sigma) + \mathrm{vec}(\Sigma M^\top)\\ &= (\Sigma^\top \otimes I_p)\mathrm{vec}(M) + (I_p \otimes \Sigma)\mathrm{vec}(M^\top) = ((\Sigma \otimes I_p) + (I_p \otimes \Sigma)K_p)\mathrm{vec}(M). \qquad \square\end{aligned}$$

The Lyapunov equation (1.2) is symmetric and therefore $p(p-1)/2$ equations of the equation system (4.2) are redundant.

DEFINITION 4.2. *Given a $p \times p$ symmetric matrix Σ, we define the $p(p+1)/2 \times p^2$ matrix $A(\Sigma)$ by selecting the rows of*

$$(\Sigma \otimes I_p) + (I_p \otimes \Sigma)K_p$$

indexed by pairs (k,l) with $k \le l$.

Let $\mathrm{vech}(C) = (C_{kl} : k \le l)$ be the half-vectorization of the symmetric matrix C. Then we can write the Lyapunov equation as

$$A(\Sigma)\mathrm{vec}(M) = -\mathrm{vech}(C).$$

As noted, we index the rows of $A(\Sigma)$ by pairs (k,l) with $k \le l$. To index the columns of $A(\Sigma)$ we will use the potential edges $i \to j$, where we recall that the edge $i \to j$ corresponds to the entry m_{ji} of the matrix M.

Example 2.6 displayed $A(\Sigma)$ for the case of $p = 3$. In general, we have

$$A(\Sigma)_{(k,l),i\to j} = \begin{cases} 0 & \text{if} \quad j \ne k, l, \\ \Sigma_{li} & \text{if} \quad j = k,\, k \ne l, \\ \Sigma_{ki} & \text{if} \quad j = l,\, l \ne k, \\ 2\Sigma_{ji} & \text{if} \quad j = k = l. \end{cases} \tag{4.3}$$

Any specific graphical continuous Lyapunov model assumes that M has nonzero entries only for pairs (j,i) for which the underlying graph contains the edge $i \to j$. We are thus led to select a subset of columns of the coefficient matrix $A(\Sigma)$ when studying solvability of the Lyapunov equation. By the next lemma, generic and global identifiability of a graphical continuous Lyapunov model are equivalent to rank conditions on the relevant submatrix of $A(\Sigma)$.

LEMMA 4.3. *Let $G = (V,E)$ be a directed graph with $V = [p]$, and let $C \in \mathrm{PD}_p$. Let $A(\Sigma)_{\cdot,E}$ be the submatrix of $A(\Sigma)$ obtained by selecting the columns indexed by the edges in E. Then the model $\mathcal{M}_{G,C}$ is*

(i) *globally identifiable if and only if $A(\Sigma)_{\cdot,E}$ has full column rank $|E|$ for all $\Sigma \in \mathcal{M}_{G,C}$;*

(ii) *generically identifiable if and only if there exists a matrix $\Sigma \in \mathcal{M}_{G,C}$ such that $A(\Sigma)_{\cdot,E}$ has full column rank $|E|$.*

If $\mathcal{M}_{G,C}$ is not generically identifiable, then it is nonidentifiable.

Proof. Let $M_0 \in \mathrm{Stab}(E)$, and let $\Sigma_0 = \phi_{G,C}(M_0)$ be the associated covariance matrix. The fiber $\mathcal{F}_{G,C}(M_0)$ is the set of all matrices $M \in \mathbb{R}^E$ with

$$A(\Sigma_0)_{\cdot,E}\,\mathrm{vec}(M)_E = -\mathrm{vech}(C), \tag{4.4}$$

where $\mathrm{vec}(M)_E$ is the subvector of $\mathrm{vec}(M)$ that comprises the entries indexed by (j,i) with $i \to j \in E$. Hence, $\mathcal{F}_{G,C}(M_0) = \{M_0\}$ precisely when $A(\Sigma_0)_{\cdot,E}$ has full column rank such that (4.4) has a unique solution. Claim (i) is now evident.

To prove (ii), note that $A(\Sigma)_{\cdot,E}$ has full column rank if and only if the vector of all maximal minors of $A(\Sigma)_{\cdot,E}$ is nonzero. By (4.1), the map $\phi_{G,C}$ is a rational map. Consequently, the map taking $M \in \mathrm{Stab}(E)$ to the maximal minors of $A(\phi_{G,C}(M))_{\cdot,E}$ is rational as well. Now a rational map is nonzero outside a measure zero set if and only if there exists a single point where it is nonzero. Consequently, the existence of $\Sigma \in \mathcal{M}_{G,C}$ with $A(\Sigma)_{\cdot,E}$ of full column rank implies generic identifiability of $\mathcal{M}_{G,C}$.

Finally, if $\mathcal{M}_{G,C}$ is not generically identifiable, then the column rank of $A(\Sigma_0)_{\cdot,E}$ is strictly smaller than $|E|$ for all $\Sigma_0 = \phi_{G,C}(M_0) \in \mathcal{M}_{G,C}$. The fiber $\mathcal{F}_{G,C}(M_0) \subseteq \mathrm{Stab}(E)$ is then the affine subspace of solutions to (4.4) of dimension ≥ 1. Hence, $|\mathcal{F}_{G,C}(M_0)| = \infty$ for all $M_0 \in \mathrm{Stab}(E)$, and $\mathcal{M}_{G,C}$ is nonidentifiable. ∎

5. Directed acyclic graphs. In this section, we prove that all models that are given by *directed acyclic graphs* (DAGs) are globally identifiable. In our setting, a

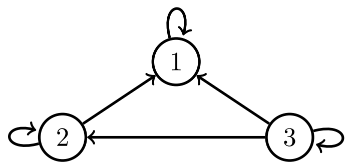

FIG. 2. *The complete DAG G^* on 3 nodes.*

DAG is a directed graph that does not contain any directed cycles other than the always present self-loops $i \to i$, $i \in [p]$. This case is special in that we are able to make a simple argument based on block structure in the coefficient matrix $A(\Sigma)$.

By Proposition 3.6, in order to prove global identifiability for all DAGs it suffices to treat DAGs that are complete in the sense of the following definition.

DEFINITION 5.1. *A directed simple graph $G = (V, E)$ with $V = [p]$ is* complete *if there is an edge between every pair of distinct nodes.*

A simple graph that also contains all self-loops $i \to i$, $i \in [p]$, is complete if and only if $|E| = p(p+1)/2$. Because vertex relabelling has no impact on identifiability, we can furthermore restrict attention to a single topological ordering. In other words, it suffices to consider the single complete DAG G^* whose edge set comprises all edges $i \to j$ with $i \geq j$.

Example 5.2. Consider the case of $p = 3$ nodes, for which the complete DAG $G^* = (V, E^*)$ is shown in Figure 2. The graph encodes the drift matrix

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ 0 & m_{22} & m_{23} \\ 0 & 0 & m_{33} \end{pmatrix},$$

and the submatrix $A(\Sigma)_{\cdot, E^*}$ is equal to

$$\begin{array}{c} \\ (1,1) \\ (1,2) \\ (1,3) \\ (2,2) \\ (2,3) \\ (3,3) \end{array} \begin{array}{c} \begin{array}{cccccc} 1 \to 1 & 2 \to 1 & 3 \to 1 & 2 \to 2 & 3 \to 2 & 3 \to 3 \end{array} \\ \begin{pmatrix} 2\Sigma_{11} & 2\Sigma_{12} & 2\Sigma_{13} & 0 & 0 & 0 \\ \Sigma_{12} & \Sigma_{22} & \Sigma_{23} & \Sigma_{12} & \Sigma_{13} & 0 \\ \Sigma_{13} & \Sigma_{23} & \Sigma_{33} & 0 & 0 & \Sigma_{13} \\ 0 & 0 & 0 & 2\Sigma_{22} & 2\Sigma_{23} & 0 \\ 0 & 0 & 0 & \Sigma_{32} & \Sigma_{33} & \Sigma_{23} \\ 0 & 0 & 0 & 0 & 0 & 2\Sigma_{33} \end{pmatrix} \end{array}.$$

Up to some rows being scaled by 2, the three diagonal blocks are principal minors of the positive definite matrix Σ. Therefore, it holds for all $\Sigma \in \mathrm{PD}_3$ that

$$\begin{aligned} \det A(\Sigma)_{\cdot, E^*} &= \begin{vmatrix} 2\Sigma_{11} & 2\Sigma_{12} & 2\Sigma_{13} \\ \Sigma_{12} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{13} & \Sigma_{23} & \Sigma_{33} \end{vmatrix} \cdot \begin{vmatrix} 2\Sigma_{22} & 2\Sigma_{23} \\ \Sigma_{23} & \Sigma_{33} \end{vmatrix} \cdot |2\Sigma_{33}| \\ &= 2^3 \cdot \det(\Sigma) \cdot \det(\Sigma_{\{2,3\},\{2,3\}}) \cdot \Sigma_{33} > 0. \end{aligned}$$

The block structure found in Example 5.2 generalizes and gives the main result of this section.

THEOREM 5.3. *Let $G = (V, E)$ be a DAG with $V = [p]$. Then the model $\mathcal{M}_{G,C}$ is globally identifiable for every matrix $C \in \mathrm{PD}_p$.*

Proof. As noted above, it suffices to consider the complete DAG $G^* = (V, E^*)$ whose edges are $i \to j$ for $i \geq j$.

Our proof then applies Lemma 4.3, which states that model $\mathcal{M}_{G^*,C}$ is globally identifiable if and only if $\det(A(\Sigma)_{\cdot,E^*}) \neq 0$ for all $\Sigma \in \mathcal{M}_{G^*,C}$.

In what follows, let $\Sigma \in \mathrm{PD}_p$. Partition the edge set as $E^* = E_1^* \cup E_2^* \cup \cdots \cup E_p^*$, where $E_i^* = \{j \to i : j \geq i\}$. Similarly, partition the row index set of $A(\Sigma)$ into the disjoint union of the sets $R_k = \{(k,l) : l \geq k\}$, $k = 1, \ldots, p$. Inspecting (4.3), we see that the submatrix

$$A(\Sigma)_{R_k, E_i^*} = 0 \quad \text{if } k > i.$$

Hence, the matrix $A(\Sigma)$ can be arranged in block upper-triangular form, and

$$\det\left(A(\Sigma)_{\cdot,E^*}\right) = \prod_{i=1}^{p} \det\left(A(\Sigma)_{R_i,E_i^*}\right).$$

Note that arrangement of columns and rows is consistent with the ordering in (2.3). Inspecting again (4.3), we find that $A(\Sigma)_{R_i,E_i^*}$ is equal to the principal submatrix $P(\Sigma)_{\geq i} := \Sigma_{\{i,\ldots,p\},\{i,\ldots,p\}}$ but with the first row of $P(\Sigma)_{\geq i}$ (the one indexed by i) being multiplied by 2 in $A(\Sigma)_{R_i,E_i^*}$. Since all principal minors of a positive definite matrix Σ are positive, we obtain that

$$\left|\det\left(A(\Sigma)_{\cdot,E^*}\right)\right| = 2^p \prod_{i=1}^{p} \det\left(P(\Sigma)_{\geq i}\right) > 0 \quad \text{for all } \ \Sigma \in \mathrm{PD}_p.$$

In particular, $A(\Sigma)_{\cdot,E^*}$ has nonvanishing determinant for all $\Sigma \in \mathcal{M}_{G^*,C}$. ∎

The proof of Theorem 5.3 shows that for any complete DAG $G = (V, E)$ the matrix $A(\Sigma)_{\cdot,E}$ is invertible for all $\Sigma \in \mathrm{PD}_p$. Using this fact, the proof of the theorem reveals more information about Lyapunov models arising from DAGs.

COROLLARY 5.4. *Let $G = (V, E)$ be a DAG with $V = [p]$. Then $\mathcal{M}_{G,C}$ is an algebraic and thus closed subset of PD_p. If G is complete, then $\mathcal{M}_{G,C} = \mathrm{PD}_p$.*

Proof. Let G be a complete DAG. By Theorem 5.3, the square matrix $A(\Sigma)_{\cdot,E}$ has full rank for all $\Sigma \in \mathrm{PD}_p$. Therefore, the solution $\mathrm{vec}(M)$ to the vectorized Lyapunov equation (4.4) exists uniquely for all $\Sigma \in \mathrm{PD}_p$. The resulting drift matrix M has the right support by construction, hence $\mathcal{M}_{G,C} = \mathrm{PD}_p$.

If G is a noncomplete DAG, then we may add edges to obtain a complete DAG $\bar{G} = (V, \bar{E})$. As $A(\Sigma)_{\cdot,\bar{E}}$ has full column rank for all $\Sigma \in \mathrm{PD}_p$ the same is true for $A(\Sigma)_{\cdot,E}$; recall Proposition 3.6. Hence, a matrix $\Sigma \in \mathrm{PD}_p$ is in $\mathcal{M}_{G,C}$ if and only if $\mathrm{vech}(C)$ is in the column span of $A(\Sigma)_{\cdot,E}$ if and only if the $(|E|+1)$-minors of the augmented matrix $(A(\Sigma)_{\cdot,E} \mid \mathrm{vech}(C))$ vanish. The model $\mathcal{M}_{G,C}$ is thus an algebraic subset: it is the set of positive definite matrices at which these minors vanish. ∎

6. Sum of squares decompositions and finer rank conditions. Directed cycles break the block-diagonal structure found for DAGs (Theorem 5.3), making it difficult to check rank conditions on $A(\Sigma)$. In this section we show that small cyclic graphs can nevertheless be handled by applying SOS decompositions to certify positivity of subdeterminants. Moreover, we show that our rank conditions may be placed on a smaller matrix containing a basis for the kernel of $A(\Sigma)$.

In Example 2.6, we proved global identifiability for the 3-cycle by showing that the key factor $\Sigma_{11}\Sigma_{22}\Sigma_{33} - \Sigma_{12}\Sigma_{13}\Sigma_{23}$ in the determinant of $A(\Sigma)_{\cdot,E}$ is positive on

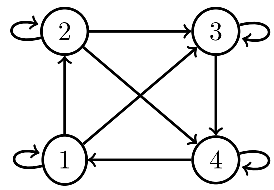

FIG. 3. *A completion of the* 4*-cycle.*

PD_3. We were able to argue this via the positivity of 2×2 principal minors of Σ. However, a direct extension of this approach to cyclic graphs with a larger number of nodes is difficult. Nevertheless, some headway can be made by exploiting the positive-definiteness of Σ via its Cholesky decomposition.

Example 6.1. Let $G = (V, E)$ be the completion of the 4-cycle with $V = [4]$ and $E = \{1 \to 1, 2 \to 2, 3 \to 3, 4 \to 4, 1 \to 2,\ 1 \to 3,\ 2 \to 3, 2 \to 4,\ 3 \to 4,\ 4 \to 1\}$. It is displayed in Figure 3. Let $\Sigma = LL^\top$ be the Cholesky decomposition of $\Sigma \in \mathrm{PD}_4$ in terms of the lower-triangular matrix

$$L = \begin{pmatrix} l_{11} & 0 & 0 & 0 \\ l_{12} & l_{22} & 0 & 0 \\ l_{13} & l_{23} & l_{33} & 0 \\ l_{14} & l_{24} & l_{34} & l_{44} \end{pmatrix}$$

with $l_{11}, l_{22}, l_{33}, l_{44} > 0$. Then

$$|\det(A(LL^\top)_{\cdot,E})| = 16\, l_{44}^2 l_{33}^2 l_{22}^4 l_{11}^6 \cdot |f(L)|,$$

where the key factor is

$$\begin{aligned} f(L) = {} & l_{14}^2 l_{22}^2 l_{33}^2 - l_{12} l_{14} l_{22} l_{24} l_{33}^2 + l_{12}^2 l_{24}^2 l_{33}^2 + l_{22}^2 l_{24}^2 l_{33}^2 - l_{13} l_{14} l_{22}^2 l_{33} l_{34} \\ & + l_{12} l_{14} l_{22} l_{23} l_{33} l_{34} + l_{12} l_{13} l_{22} l_{24} l_{33} l_{34} - l_{12}^2 l_{23} l_{24} l_{33} l_{34} + l_{13}^2 l_{22}^2 l_{34}^2 \\ & - 2 l_{12} l_{13} l_{22} l_{23} l_{34}^2 + l_{12}^2 l_{23}^2 l_{34}^2 + l_{12}^2 l_{33}^2 l_{34}^2 + l_{22}^2 l_{33}^2 l_{34}^2 + l_{13}^2 l_{22}^2 l_{44}^2 \\ & - 2 l_{12} l_{13} l_{22} l_{23} l_{44}^2 + l_{12}^2 l_{23}^2 l_{44}^2 + l_{12}^2 l_{33}^2 l_{44}^2 + l_{22}^2 l_{33}^2 l_{44}^2. \end{aligned}$$

A computer algebra system such as `Macaulay2` with the package from Cifuentes et al. (2020) quickly finds an SOS decomposition for f as

$$\begin{aligned} f(L) = {} & \left(\frac{1}{2} l_{14} l_{22} l_{33} - \frac{1}{2} l_{12} l_{24} l_{33} - l_{13} l_{22} l_{34} + l_{12} l_{23} l_{34} \right)^2 \\ & + (-l_{13} l_{22} l_{44} + l_{12} l_{23} l_{44})^2 + (l_{12} l_{33} l_{34})^2 + (l_{12} l_{33} l_{44})^2 + (l_{22} l_{24} l_{33})^2 \\ & + (l_{22} l_{33} l_{34})^2 + (l_{22} l_{33} l_{44})^2 + \frac{3}{4} \left(l_{14} l_{22} l_{33} - \frac{1}{3} l_{12} l_{24} l_{33} \right)^2 + \frac{2}{3} (l_{12} l_{24} l_{33})^2. \end{aligned}$$

Since $l_{22} l_{33} l_{44} > 0$, it follows that f is strictly positive for any Cholesky factor L. Therefore, $|\det(A(\Sigma)_{\cdot,E})| > 0$ and we conclude that $\mathcal{M}_{G,C}$ is globally identifiable, no matter the choice of $C \in \mathrm{PD}_4$.

Remark 6.2. A polynomial being an SOS is a stronger requirement than the polynomial being nonzero. Therefore, we could have a nonvanishing determinant even if the considered polynomial factor failed the SOS test. However, we do not know of an example where this might be the case.

Observe that $\det(\Sigma) = (\det L)^2 = l_{11}^2 l_{22}^2 l_{33}^2 l_{44}^2$ appears as a factor of $\det(A(\Sigma)_{\cdot,E})$ in all our examples so far (recall Examples 2.6, 5.2, and 6.1). This phenomenon actually occurs for any complete simple graph (see Corollary B.2 in the appendix) and suggests that identifiability should be encoded in a smaller matrix. Indeed, this information is carried by a specific row restriction of a matrix whose columns form a basis of the kernel of $A(\Sigma)$.

The kernel of $A(\Sigma)$ is described by the following fact, straightforward to verify; see also Barnett and Storey (1967). It parametrizes the stable matrices M that are solutions to the Lyapunov equation in terms of skew-symmetric matrices (matrices K with $K^\top = -K$).

LEMMA 6.3. *Consider the continuous Lyapunov equation from* (1.2) *for given* $\Sigma, C \in \mathrm{PD}_p$. *Then a matrix* $M \in \mathbb{R}^{p\times p}$ *solves the Lyapunov equation if and only if there exists a skew-symmetric matrix* K *such that*

$$M = \left(K - \frac{1}{2}C\right)\Sigma^{-1}.$$

The proof can be found in section 2 of Barnett and Storey (1967) and is included here for completeness as Lemma 6.3 plays an important role for the following results.

Proof. Substituting M into (1.2) and using the symmetry of Σ, C and that $K^\top = -K$, we obtain

$$\begin{aligned} M\Sigma + \Sigma M^\top &= \left(K - \frac{1}{2}C\right)\Sigma^{-1}\Sigma + \Sigma(\Sigma^{-1})^\top\left(K - \frac{1}{2}C\right)^\top \\ &= \left(K - \frac{1}{2}C\right) + \left(-K - \frac{1}{2}C\right) = -C. \end{aligned}$$

Conversely, since M and C are both symmetric matrices, we can write (1.2) as

$$(M\Sigma)^\top + \frac{1}{2}C^\top = -M\Sigma - \frac{1}{2}C.$$

Therefore, the matrix $K = M\Sigma + \frac{1}{2}C$ is skew-symmetric. ◻

The space of skew-symmetric matrices has dimension $p(p-1)/2$. Hence, for $\Sigma \in \mathrm{PD}_p$, the kernel of $A(\Sigma)$ also has dimension $p(p-1)/2$. We give further details about the spectral properties of $A(\Sigma)$ in Theorem B.1. The following result now gives simplified rank conditions for identifiability.

LEMMA 6.4. *Let* $G = (V, E)$ *be a directed graph with* $V = [p]$, *and let* $C \in \mathrm{PD}_p$. *For every* $\Sigma \in \mathrm{PD}_p$, *let* $H(\Sigma)$ *be a* $p^2 \times p(p-1)/2$ *matrix whose columns form a basis of the kernel of* $A(\Sigma)$, *and let* $H(\Sigma)_{E^c,\cdot}$ *be the submatrix obtained by restriction to rows corresponding to nonedges* E^c *of* G. *Then the associated model* $\mathcal{M}_{G,C}$ *is*

(i) *globally identifiable if and only if* $H(\Sigma)_{E^c,\cdot}$ *has full column rank* $p(p-1)/2$ *for all* $\Sigma \in \mathcal{M}_{G,C}$;

(ii) *generically identifiable if and only if there exists a matrix* $\Sigma \in \mathcal{M}_{G,C}$ *such that* $H(\Sigma)_{E^c,\cdot}$ *has full column rank* $p(p-1)/2$.

Proof. Recall from Lemma 4.3 that the elements of the fiber are solutions of the equation system (4.4), which has a unique solution for a given (positive definite) matrix $\Sigma \in \mathcal{M}_{G,C}$ if and only if $A(\Sigma)_{\cdot,E}$ has linearly independent columns. The latter condition can be rephrased as follows: the kernel of $A(\Sigma)$ does not contain any element $\mathrm{vec}(M) \neq 0$ such that $M \in \mathbb{R}^E$. Put differently, (4.4) admits a unique solution if and only if the column span of $H(\Sigma)$ does not contain any element $\mathrm{vec}(M) \neq 0$ for $M \in \mathbb{R}^E$. As $H(\Sigma)$ has linearly independent columns, this latter condition is equivalent to the linear independence of the columns of the extended matrix $(H(\Sigma) \mid \mathrm{vec}(M))$ for any nontrivial $M \in \mathbb{R}^E$. It remains to be proven that this, in turn, is equivalent to the $|E^c| \times p(p-1)/2$ submatrix $H(\Sigma)_{E^c,\cdot}$ having rank $p(p-1)/2$.

Assume that $H(\Sigma)_{E^c,\cdot}$ has rank $p(p-1)/2$, and consider one of its nonvanishing maximal minors. This minor can always be extended to a nonvanishing maximal minor of $(H(\Sigma) \mid \mathrm{vec}(M))$ by adding one of the rows corresponding to $m_{ji} \neq 0$. Therefore, the extended matrix has full rank.

For the converse implication, note that if $H(\Sigma)_{E^c,\cdot}$ has rank strictly less than $p(p-1)/2$, then there exists a (not unique) nontrivial $M \in \mathbb{R}^E$ such that $\mathrm{vec}(M)$ belongs to the kernel of $A(\Sigma)$. ◻

For a convenient choice of a basis of the kernel of $A(\Sigma)$ we may appeal to the following fact.

LEMMA 6.5. *For a matrix $\Sigma \in \mathrm{PD}_p$, the kernel of $A(\Sigma)$ equals*

$$\begin{aligned}\ker A(\Sigma) &= \{\mathrm{vec}(K\Sigma^{-1}) : K \textit{ skew-symmetric}\} \\ &= \{\mathrm{vec}(\Sigma K) : K \textit{ skew-symmetric}\}.\end{aligned}$$

Proof. The first equality holds by Lemma 6.3. The second equality follows from the fact that K is skew-symmetric if and only if $\Sigma K \Sigma$ is skew-symmetric. ◻

For $1 \leq k, l \leq p$, let $K^{(k,l)} = e_k \otimes e_l - e_l \otimes e_k$ be the skew-symmetric matrix whose only nonzero entries are 1 in place (k,l) and -1 in place (l,k). Then the set $\{K^{(k,l)} : k < l\}$ is a basis of the space of $p \times p$ skew-symmetric matrices and, thus, the set $\{\mathrm{vec}(\Sigma K^{(k,l)}) : k < l\}$ is a basis of $\ker A(\Sigma)$. We may thus choose the matrix $H(\Sigma)$ in Lemma 6.5 as the matrix with entries

$$H(\Sigma)_{i\to j,(k,l)} = \mathrm{vec}(\Sigma K^{(k,l)})_{ji} = \begin{cases} -\Sigma_{lj} & \text{if } i = k, \\ \Sigma_{kj} & \text{if } i = l, \\ 0 & \text{otherwise.} \end{cases} \tag{6.1}$$

Note that we index the rows of $H(\Sigma)$ by all possible edges of a directed graph (including self-loops), in accordance with the indexing of the columns of $A(\Sigma)$.

Example 6.6. Consider the 6×9 matrix $A(\Sigma)$ in Example 2.6 corresponding to $p = 3$. Then the matrix from (6.1) is

$$H(\Sigma) = \begin{array}{cc} \begin{pmatrix} -\Sigma_{12} & -\Sigma_{13} & 0 \\ -\Sigma_{22} & -\Sigma_{23} & 0 \\ -\Sigma_{23} & -\Sigma_{33} & 0 \\ \Sigma_{11} & 0 & -\Sigma_{13} \\ \Sigma_{12} & 0 & -\Sigma_{23} \\ \Sigma_{13} & 0 & -\Sigma_{33} \\ 0 & \Sigma_{11} & \Sigma_{12} \\ 0 & \Sigma_{12} & \Sigma_{22} \\ 0 & \Sigma_{13} & \Sigma_{23} \end{pmatrix} & \begin{array}{c} 1\to 1 \\ 1\to 2 \\ 1\to 3 \\ 2\to 1 \\ 2\to 2 \\ 2\to 3 \\ 3\to 1 \\ 3\to 2 \\ 3\to 3 \end{array} \end{array}.$$

Consider the DAG on 3 nodes given in Figure 2, for which the set of nonedges is $E^c = \{1 \to 2, 1 \to 3, 2 \to 3\}$. Then

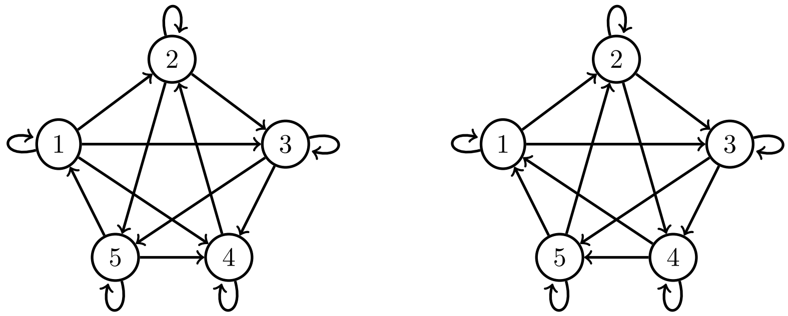

Fig. 4. *The two simple cyclic graphs on* 5 *nodes for which Proposition* 6.7 *could not be computationally proved with the techniques employed.*

$$|\det H(\Sigma)_{E^c,\cdot}| = \left|\det\begin{pmatrix} -\Sigma_{22} & -\Sigma_{23} & 0 \\ -\Sigma_{23} & -\Sigma_{33} & 0 \\ \Sigma_{13} & 0 & -\Sigma_{33} \end{pmatrix}\right| = \Sigma_{33}(\Sigma_{22}\Sigma_{33} - \Sigma_{23}^2)$$

is a product of two principal minors of Σ, as expected from Theorem 5.3.

Next, let $E^c = \{2 \to 1, 1 \to 3, 3 \to 2\}$ be the set of nonedges of the 3-cycle from Figure 1. Then

$$|\det H(\Sigma)_{E^c,\cdot}| = \left|\det\begin{pmatrix} \Sigma_{11} & 0 & -\Sigma_{13} \\ -\Sigma_{23} & -\Sigma_{33} & 0 \\ 0 & \Sigma_{12} & \Sigma_{22} \end{pmatrix}\right| = \Sigma_{11}\Sigma_{22}\Sigma_{33} - \Sigma_{12}\Sigma_{13}\Sigma_{23},$$

which is what we obtained in (2.4).

Following Example 6.1, we can establish global identifiability by computing an SOS decomposition of the determinant of the restricted kernel $H(\Sigma)_{E^c,\cdot}$ using the Cholesky decomposition of Σ. Such computations allowed us to establish the following proposition.

Proposition 6.7. *Let $G = (V, E)$ be a simple graph with $V = [p]$, and let $C \in \mathrm{PD}_p$. Let $L \in \mathbb{R}^{p\times p}$ be lower-triangular. If $p \le 4$, then there exists a permutation matrix P such that $\det H(PLL^\top P^\top)_{E^c,\cdot}$ is an everywhere positive SOS in the entries of L, implying that $\mathcal{M}_{G,C}$ is globally identifiable. The same is true for $p = 5$ with the possible exception of two types of graphs, depicted in Figure* 4.

For our computer proof of the claims in the proposition, we applied the computer algebra system `Macaulay2`. For the graphs in Figure 4, we additionally employed MATLAB toolboxes, but our computations did not terminate. It is natural to conjecture that Proposition 6.7 holds for all graphs with $p = 5$, and even all simple graphs.

7. Simple cyclic graphs. In this section we establish our main result: global identifiability of all Lyapunov models given by simple cyclic graphs. Moreover, we can show that simple cyclic graphs give models that are algebraic subsets of the positive definite cone. Our proofs exploit the parametrization of stable matrices M that are solutions to the Lyapunov equation in terms of skew-symmetric matrices (matrices K with $K^\top = -K$); recall Lemma 6.3.

THEOREM 7.1. *Let $G=(V,E)$ be a directed graph with $V=[p]$.*
(i) *If G is simple, then the model $\mathcal{M}_{G,C}$ is globally identifiable for all $C\in \mathrm{PD}_p$.*
(ii) *If $C\in \mathrm{PD}_p$ is diagonal, then the model $\mathcal{M}_{G,C}$ is globally identifiable if and only if G is simple.*

Proof. It suffices to prove (i), as (ii) then follows from Proposition 3.10.

To prove (i), suppose G is indeed simple. Let $M_1, M_2 \in \mathrm{Stab}(E)$ be any two matrices that solve the Lyapunov equation (1.2) for the same $\Sigma\in\mathcal{M}_{G,C}$. According to Lemma 6.3 there exist two skew-symmetric matrices K_1 and K_2 such that $M_1 = (K_1 - \frac{1}{2}C)\Sigma^{-1}$ and $M_2 = (K_2-\frac{1}{2}C)\Sigma^{-1}$. For the difference we obtain

$$M := M_1 - M_2 = \left(K_1 - \frac{1}{2}C\right)\Sigma^{-1} - \left(K_2 - \frac{1}{2}C\right)\Sigma^{-1} = (K_1-K_2)\Sigma^{-1}.$$

The difference $K = K_1 - K_2$ is again skew-symmetric, so that M is the product of a skew-symmetric matrix K and the positive definite matrix Σ^{-1}.

As Σ and Σ^{-1} are positive definite, we may form the square roots $\Sigma^{\frac{1}{2}}$ and $\Sigma^{-\frac{1}{2}}$. Observe that $M = K\Sigma^{-1}$ is similar to $\tilde{M} = \Sigma^{-\frac{1}{2}}K\Sigma^{-1}\Sigma^{\frac{1}{2}}$, which is skew-symmetric since

$$\tilde{M}^\top = (\Sigma^{-\frac{1}{2}}K\Sigma^{-\frac{1}{2}})^\top = \Sigma^{-\frac{1}{2}}K^\top\Sigma^{-\frac{1}{2}} = -\Sigma^{-\frac{1}{2}}K\Sigma^{-\frac{1}{2}} = -\tilde{M}.$$

The nonzero eigenvalues of real skew-symmetric matrices are purely imaginary. Let $i\lambda_1,\dots,i\lambda_p$ with $\lambda_i\in\mathbb{R}$ be the eigenvalues of $\tilde{M}$. Then the similarity implies that M has the same eigenvalues.

Observe now that the eigenvalues of M^2 are $-\lambda_1^2,\dots,-\lambda_p^2$ and thus $\mathrm{tr}(M^2)\le 0$. As M is supported over a simple graph, it holds for all pairs of indices $i\neq j$ that $m_{ij}m_{ji}=0$. Hence, the diagonal of M^2 is given by the squared diagonal elements of M, i.e., $(M^2)_{ii} = m_{ii}^2$. It follows that

$$0 \le \sum_{i=1}^p m_{ii}^2 = \mathrm{tr}(M^2) = \sum_{i=1}^p -\lambda_i^2 \le 0,$$

which implies that $\lambda_i^2 = 0$ for all $i=1,\dots,p$. But this is only true if $\lambda_i = 0$ for all $i=1,\dots,p$. Therefore, all eigenvalues of M are zero. Using the similarity of M with $\tilde{M}$ and that skew-symmetric matrices are diagonalizable, we deduce that M is similar to the zero matrix. But then $M=0$ and consequently $M_1 = M_2$, which shows that the Lyapunov equation admits a unique solution in $\mathrm{Stab}(E)$. ∎

In addition to global identifiability, we have a generalization of Corollary 5.4 to general simple graphs.

COROLLARY 7.2. *Let $G=(V,E)$ be a simple graph with $V=[p]$. Then $\mathcal{M}_{G,C}$ is an algebraic and thus closed subset of PD_p. If G is complete, then $\mathcal{M}_{G,C} = \mathrm{PD}_p$.*

Proof. Consider first the case where G is complete (with an edge between every pair of nodes). Let $\Sigma_0\in\mathrm{PD}_p$ be an arbitrary positive definite matrix. Choosing $M=-I_p$, the negated identity matrix, shows that Σ_0 belongs to the model $\mathcal{M}_{G,C_0}$ for $C_0 = 2\Sigma_0$. By Theorem 7.1 and Lemma 4.3, we obtain that the determinant of $A(\Sigma)_{\cdot,E}$ is nonzero at every matrix in $\mathcal{M}_{G,C_0}$ and, in particular, at Σ_0. We conclude that $\det(A(\Sigma)_{\cdot,E})\neq 0$ on all of PD_p. As in the proof of Corollary 5.4, we deduce that $\mathcal{M}_{G,C} = \mathrm{PD}_p$ for all $C\in\mathrm{PD}_p$.

If G is not complete, then it can be augmented to a complete graph $\bar{G} = (V,\bar{E})$, and we may complete the proof in analogy to the proof of Corollary 5.4. ∎

8. Nonsimple graphs. In this section, we consider directed graphs $G=(V,E)$ that are allowed to be nonsimple, i.e., may contain a 2-cycle. In our study, we restrict attention to the case where $C\in \mathrm{PD}_p$ is diagonal. Proposition 3.10 tells us that, for C diagonal, a model $\mathcal{M}_{G,C}$ given by a nonsimple graph G can never be globally identifiable. However, nonsimple graphs with at most $p(p+1)/2$ edges may still give generically identifiable models (Definition 3.1, Lemma 3.5). We are able to provide a combinatorial condition that is necessary for generic identifiability, and we computationally classify all graphs with $p\le 5$ nodes. Our study reveals examples for which generic identifiability depends in subtle ways on the pattern of edges.

We begin with a small example.

Example 8.1. Let $G=(V,E)$ be the graph from Figure 5, a 2-cycle with an additional edge pointing to a third node, and let $C\in \mathrm{PD}_3$ be a diagonal matrix. To inspect identifiability of $\mathcal{M}_{G,C}$, we may use the kernel basis of Example 6.6 with the set of nonedges $E^c=\{1\to 3, 3\to 1, 3\to 2\}$. We find

$$\det H(\Sigma)_{E^c,\cdot} = \det\begin{pmatrix} -\Sigma_{23} & -\Sigma_{33} & 0\\ 0 & \Sigma_{11} & \Sigma_{12}\\ 0 & \Sigma_{12} & \Sigma_{22}\end{pmatrix} = \Sigma_{23}\left(\Sigma_{12}^2-\Sigma_{11}\Sigma_{22}\right).$$

Since $\mathcal{M}_{G,C}$ contains positive definite matrices with both vanishing and nonvanishing Σ_{23}, we conclude that $\mathcal{M}_{G,C}$ is generically (but not globally) identifiable.

Note that the matrices $\Sigma\in\mathcal{M}_{G,C}$ with $\Sigma_{23}=0$ are obtained precisely from the drift matrices in the lower-dimensional set $\{M\in \mathrm{Stab}(E) : m_{32}=0\}$. Indeed, if $m_{32}=0$, then the situation is as if the $2\to 3$ edge were removed, and we will see in Proposition 8.3 that this implies $\Sigma_{23}=0$ when C is diagonal. Conversely, when solving for Σ given a drift matrix $M\in\mathbb{R}^E$ we find that Σ_{23} is a rational function of (M,C) whose numerator is

$$m_{32}\left(c_{11}m_{21}^2\mathrm{tr}(M)+c_{22}m_{11}^2\mathrm{tr}(M)+c_{22}\det(M)\right).$$

As C is positive definite and M stable, the second factor is negative. Thus, if $\Sigma=\Sigma(M,C)$ is a positive definite matrix in $\mathcal{M}_{G,C}$, then $\Sigma_{23}=0$ implies $m_{32}=0$.

By Lemma 3.5, $|E|\le p(p+1)/2$ is a necessary condition for generic identifiability of the model of a graph $G=(V,E)$. We now show how this bound may be improved by accounting for knowledge about vanishing covariances.

Definition 8.2. *A* trek *is a sequence of edges of the form*

$$l_m\leftarrow l_{m-1}\leftarrow\cdots\leftarrow l_1\leftarrow t\to r_1\to\cdots\to r_{n-1}\to r_n.$$

The node t is the top node of the trek. The directed paths $l_m\leftarrow l_{m-1}\leftarrow\cdots\leftarrow l_1$ and $r_1\to\cdots\to r_{n-1}\to r_n$ are the left and the right side of the trek, respectively. The definition allows for one or both sides to be trivial, so directed paths and also single nodes are treks.

From Varando and Hansen (2020, Corollary 2.3), we deduce the following fact.

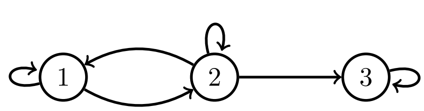

Fig. 5. *Nonsimple graph on 3 nodes.*

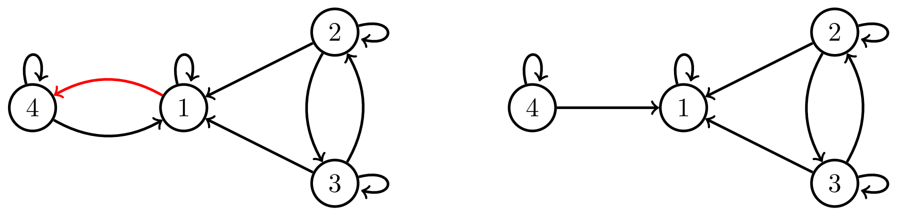

FIG. 6. *Left: Graph G_1 on 4 nodes with $\mathcal{M}_{G_1,C}$ generically identifiable. Right: Subgraph G_2 of G_1 such that $\mathcal{M}_{G_2,C}$ is nonidentifiable. $C \in \mathrm{PD}_4$ is diagonal.*

PROPOSITION 8.3. *Let $G = (V, E)$ be a directed graph with $V = [p]$, and let $C \in \mathrm{PD}_p$ be diagonal. If there is no trek from i to j in G, then $\Sigma_{ij} = 0$ in all matrices $\Sigma \in \mathcal{M}_{G,C}$.*

Example 8.4. Let $C \in \mathrm{PD}_4$ be diagonal. Then the left graph $G_1 = (V, E_1)$ in Figure 6 defines a generically identifiable model but its subgraph $G_2 = (V, E_2)$ does not. This example stresses that global identifiability is needed in Proposition 3.6. But why is $\mathcal{M}_{G_2,C}$ nonidentifiable despite G_2 having fewer edges? We observe that G_2 contains no trek between 2 and 4 and no trek between 3 and 4. Proposition 8.3 yields $\Sigma_{24} = \Sigma_{34} = 0$. Although the PD_4-cone has dimension $\binom{4+1}{2} = 10$, the existence of the constraints $\Sigma_{24} = \Sigma_{34} = 0$ implies that $\dim(\mathcal{M}_{G_2,C}) \le 10 - 2 = 8$. Since $|E_2| = 9 > 8$, nonidentifiability follows from Lemma 3.5.

As a last subtlety, we emphasize that if we remove one of the edges $2 \to 1$, $3 \to 1$, or $4 \to 1$ of G_2, we are left again with a generically identifiable model.

The ideas in Example 8.4 can be generalized into a sharper necessary condition for identifiability that is a consequence of Lemma 3.5 and Proposition 8.3.

COROLLARY 8.5. *Let $G = (V, E)$ be a directed graph with $V = [p]$. If $\mathcal{M}_{G,C}$ is generically identifiable for a diagonal matrix $C \in \mathrm{PD}_p$, then it has to hold that*

$$|E| \le \frac{p(p+1)}{2} - \#\big\{ \{i,j\} : i, j \in V \text{ with no trek between them} \big\}. \tag{8.1}$$

With this criterion in hand, we can construct graphs of arbitrary size p and fewer than $p(p+1)/2$ edges that yield nonidentifiable models.

COROLLARY 8.6. *Consider the graph $G = (V, E)$ with $p \ge 4$ nodes displayed in Figure 7. The model $\mathcal{M}_{G,C}$ is nonidentifiable for any diagonal $C \in \mathrm{PD}_p$.*

Proof. The number of parameters $|E|$ is

$$\begin{aligned} &2 \ (\text{edges from 2-cycle}) \ + \ p-1 \ (\text{edges pointing to node 1}) \\ &\qquad\qquad + p \ (\text{parameters due to the self-loops}) = \ 2p+1. \end{aligned}$$

There are no treks between any pair of nodes $\{2, \ldots, p\}$ except for the pair $(2, 3)$. This results in $\binom{p-1}{2} - 1$ (unordered) pairs of nodes with no trek. Corollary 8.5 implies that

$$\dim(\mathcal{M}_{G,C}) \le \frac{p(p+1)}{2} - \binom{p-1}{2} + 1 = 2p. \qquad \square$$

Unfortunately, the criterion in Corollary 8.5 is not sufficient.

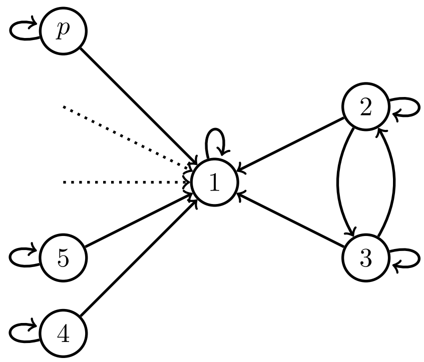

FIG. 7. *Graph G with $V=[p]$ such that $\mathcal{M}_{G,C}$ is nonidentifiable for diagonal $C \in \mathrm{PD}_p$.*

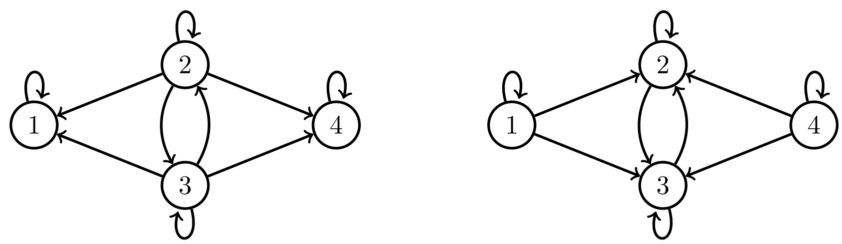

FIG. 8. *Left: Graph fulfilling the criterion in Corollary* 8.5*, yet yields a nonidentifiable model. Right: Reversing edges retains nonidentifiability, due to Corollary* 8.5*, as $\Sigma_{14}=0$.*

Example 8.7. Let $G_1=(V,E)$ be the left graph in Figure 8. Graph G_1 fulfills the necessary condition of Corollary 8.5 as the number of parameters is $6+4=10$ and all pairs of nodes are connected with a trek, which is why the right side of equation (8.1) is also $\binom{4+1}{2}=10$. However, $A(\Sigma)_{\cdot,E} \in \mathbb{R}^{10\times 10}$ does not have full rank because the columns of $A(\Sigma)$ may be linearly combined to

$$\begin{aligned}&\Sigma_{13}A(\Sigma)_{\cdot,2\to 1}+\Sigma_{23}A(\Sigma)_{\cdot,2\to 2}+\Sigma_{33}A(\Sigma)_{\cdot,2\to 3}+\Sigma_{34}A(\Sigma)_{\cdot,2\to 4}\\ -&\Sigma_{12}A(\Sigma)_{\cdot,3\to 1}-\Sigma_{22}A(\Sigma)_{\cdot,3\to 2}-\Sigma_{23}A(\Sigma)_{\cdot,3\to 3}-\Sigma_{24}A(\Sigma)_{\cdot,3\to 4}=0.\end{aligned}$$

Therefore, the model $\mathcal{M}_{G_1,C}$ is nonidentifiable for C diagonal despite fulfilling the necessary criterion. The right graph in Figure 8 yields a nonidentifiable model for the simple reason that the necessary condition of Corollary 8.5 is violated due to the absence of a trek between nodes 1 and 4.

For smaller examples, we may check generic identifiability by choosing random drift matrices and determining whether the resulting matrix Σ satisfies the rank condition from Lemma 4.3. When this does not succeed we can check symbolically whether the corresponding restriction of the coefficient matrix $A(\Sigma)$ or the restricted kernel basis $H(\Sigma)$ from Lemma 6.4 is rank-deficient, thus implying nonidentifiability. We implemented this strategy for all nonsimple graphs with $p\leq 5$ nodes and fewer than $p(p+1)/2$ parameters. As justified by Proposition 3.7, we took $C=I_p$ in our computations. This led to the results displayed in Table 1, which shows that the majority of graphs are generically identifiable. The details of the computations can be found at https://mathrepo.mis.mpg.de/LyapunovIdentifiability.

TABLE 1

Classification of models with $p = 3, 4, 5$ nodes and $C = I_p$. The last column displays the number of nonidentifiable models whose underlying graphs satisfy the necessary criterion for generic identifiability in Corollary 8.5.

Nodes	Total nonsimple	Nonidentifiable	Nonidentifiable satisfying (8.1)
3	2	0	0
4	80	3	2
5	4862	68	37

9. Conclusion. Graphical continuous Lyapunov models offer a new perspective on modeling the covariance structure of multivariate data by relating each observation to an underlying continuous-time dynamic process. The resulting covariance structure is determined by the continuous Lyapunov equation. Our work addresses the fundamental problem of whether, up to joint scaling, the parameters of the dynamic process can be identified from the covariance matrix of the cross-sectional equilibrium observations. Our main contribution shows that simple graphs yield globally identifiable models, and that the graph being simple is necessary and sufficient for global identifiability in the case where the volatility matrix C is diagonal. Moreover, we are able to show that the models of simple graphs are closed algebraic subsets of the positive definite cone. In particular, the models of complete simple graphs equal the entire positive definite cone.

Our analysis of directed acyclic graphs (DAGs) highlights block structure in the coefficient matrix for the Lyapunov equation. This leads to a straightforward proof of global identifiability and also reveals that the determinant studied in our rank conditions is a positive SOS in the entries of a Cholesky factor. This SOS property was also observed in small cyclic graphs.

While we were able to characterize global identifiability, we know less about generic identifiability of graphical Lyapunov models. Our results include an effective necessary but not sufficient graphical criterion for nonsimple graphs to be generically identifiable. We also obtain a computational classification of graphs with up to 5 nodes, and we hope that future research will lead to an improved understanding of generic identifiability of the models we considered.

Appendix A. Volatility matrix: Diagonal versus nondiagonal PD matrix. This section aims at providing insight into the need of the diagonality constraint on the volatility matrix $C \in \mathrm{PD}_p$ of the Lyapunov equation to ensure that some of the stronger results of the paper hold.

Example A.1. Let G be the 2-cycle with an additional third node, so $V = [3]$ and $E = \{1 \to 1, 1 \to 2, 2 \to 1, 2 \to 2, 3 \to 3\}$, which encodes drift matrices

$$M = \begin{pmatrix} m_{11} & m_{12} & 0 \\ m_{21} & m_{22} & 0 \\ 0 & 0 & m_{33} \end{pmatrix}.$$

Let $C = (c_{ij}) \in \mathrm{PD}_3$. Clearly, the graph does not contain any treks between nodes 1 and 3, or between nodes 2 and 3. However, a matrix $\Sigma = \phi_{G,C}(M)$ has

$$\Sigma_{13} = \frac{c_{23}m_{12} - c_{13}(m_{22} + m_{33})}{(m_{11}m_{22} - m_{12}m_{21}) + m_{33}(m_{11} + m_{22} + m_{33})},$$

with a denominator that is positive on $\operatorname{Stab}(E)$ and a numerator that is constant zero only if $c_{13} = c_{23} = 0$. The same holds for Σ_{23} by symmetry. This example serves to highlight that Proposition 8.3 may be false when C is not diagonal. Indeed, the treks would need to be allowed to move along new edges that reflect presence of nonzero diagonal entries in C; compare Varando and Hansen (2020).

Example A.2. Consider again the 2-cycle with an additional third node from the previous example. Again, consider an arbitrary matrix $C = (c_{ij}) \in \mathrm{PD}_3$. The kernel basis of Example 6.6 restricted to the set of nonedges $E^c = \{1 \to 3, 3 \to 1, 2 \to 3, 3 \to 2\}$, namely

$$H(\Sigma)_{E^c,\cdot} = \begin{pmatrix} -\Sigma_{23} & -\Sigma_{33} & 0 \\ 0 & \Sigma_{11} & \Sigma_{12} \\ \Sigma_{13} & 0 & -\Sigma_{33} \\ 0 & \Sigma_{12} & \Sigma_{22} \end{pmatrix},$$

is rank-deficient for any $\Sigma \in \mathrm{PD}_3$ if and only if $\Sigma_{13} = \Sigma_{23} = 0$. Adding this constraint to the Lyapunov equation yields $c_{13} = c_{23} = 0$. Therefore, it follows from Lemma 6.4 that $\mathcal{M}_{G,C}$ is globally identifiable for any $C = (c_{ij}) \in \mathrm{PD}_3$ in which c_{13} and c_{23} do not vanish simultaneously.

Observe that this provides a counterexample to Propositions 3.10 and 3.8 when dropping the diagonality assumption. To begin with, such $\mathcal{M}_{G,C}$ is an instance of a globally identifiable model associated to a nonsimple graph. Moreover, the subgraph H obtained by removing node 3 from G defines a nonidentifiable model for all positive definite volatility matrices by Lemma 3.5.

For the sake of completeness, note that, by Example A.1, $c_{13} = c_{23} = 0$ completely describes when the rank of $H(\Sigma)_{E^c,\cdot}$ drops for all $\Sigma \in \mathcal{M}_{G,C}$. In other words, the model is nonidentifiable if and only if $c_{13} = c_{23} = 0$ and globally identifiable otherwise.

Appendix B. Spectral description, kernel, and factorization. Here, we collect spectral properties of $A(\Sigma)$, derived more conveniently for its square $p \times p$ version

$$\tilde{A}(\Sigma) = \Sigma \otimes I_p + (I_p \otimes \Sigma) K_p,$$

which features in Lemma 4.1. We will then use this information to clarify that $\det(\Sigma)$ is a factor of $\det(A(\Sigma)_{\cdot,E})$ for complete graphs which have edge sets of size $|E| = p(p+1)/2$; see Corollary B.2.

THEOREM B.1. *Let $\Sigma \in \mathrm{PD}_p$, and let $(\lambda_i)_{i\in[p]}$ be its eigenvalues with corresponding orthogonal eigenvectors $(z_i)_{i\in[p]}$.*

(i) *The matrix $\tilde{A}(\Sigma)$ has rank $p(p+1)/2$, and* (6.1) *gives a basis for its kernel.*

(ii) *The transposed matrix $\tilde{A}(\Sigma)^\top$ has rank $p(p+1)/2$, and a basis for its kernel is given by* $\operatorname{vec}(e_i \otimes e_j - e_j \otimes e_i)$ *for $1 \le i < j \le p$.*

(iii) *Counting with multiplicities, the $p(p+1)/2$ nonzero eigenvalues of $\tilde{A}(\Sigma)$ and of $\tilde{A}(\Sigma)^\top$ are given by the sums $\lambda_i + \lambda_j$ for $1 \le i \le j \le p$, and for either matrix the associated set of orthogonal eigenvectors is* $\operatorname{vec}(z_i \otimes z_j + z_j \otimes z_i)$ *for $1 \le i \le j \le p$.*

Proof. (i) follows from (6.1), and (ii) follows from the symmetry of the Lyapunov (matrix) equation.

For (iii), the claim about $\tilde{A}(\Sigma)$ follows from the calculation

$$\begin{aligned}&(z_i \otimes z_j + z_j \otimes z_i)\Sigma + \Sigma(z_i \otimes z_j + z_j \otimes z_i)^\top \\ &\quad = [\lambda_j(z_i \otimes z_j) + \lambda_i(z_j \otimes z_i)] + [\lambda_i(z_i \otimes z_j) + \lambda_j(z_j \otimes z_i)] \\ &\quad = (\lambda_i + \lambda_j)(z_i \otimes z_j + z_j \otimes z_i).\end{aligned}$$

The transpose $\tilde{A}(\Sigma)^\top = \Sigma \otimes I_p + K_p(I_p \otimes \Sigma)$ encodes the Lyapunov equation with M replaced by $M^\top$, and the claim about $\tilde{A}(\Sigma)^\top$ follows from the symmetry of the matrices $z_i \otimes z_j + z_j \otimes z_i$. The orthogonality of the eigenvectors holds because

$$\operatorname{tr}((z_i \otimes z_j + z_j \otimes z_i)(z_k \otimes z_l + z_l \otimes z_k)) = 0$$

unless $\{i,j\} = \{k,l\}$. ∎

As a consequence of Theorem B.1, we can conclude information regarding the factorization of the determinant of $A(\Sigma)_{\cdot,E}$ when $|E| = p(p+1)/2$ such that $A(\Sigma)_{\cdot,E}$ is a square matrix.

COROLLARY B.2. *Let $G = (V,E)$ be a directed graph with $V = [p]$ and $|E| = p(p+1)/2$. The polynomials $\det(\Sigma)$ and $\det(H(\Sigma)_{E^c,\cdot})$ are factors of $\det(A(\Sigma)_{\cdot,E})$.*

Proof. The zero set of the determinant $\det(\Sigma)$ is the set of singular symmetric matrices. Since $\det(\Sigma)$ is an irreducible polynomial, every polynomial that vanishes at all singular matrices must be a polynomial multiple of $\det(\Sigma)$. Hence, it suffices to show that $\det(A(\Sigma)_{\cdot,E}) = 0$ for all singular matrices Σ. So let Σ be a singular matrix. Then there exists an eigenvalue $\lambda_i = 0$ with $i \in [p]$. Using Theorem B.1 this implies that the eigenvalue $\lambda_i + \lambda_i$ of $\tilde{A}(\Sigma)$ is zero (the theorem is written for Σ positive definite, but the fact we used also holds for Σ singular). Hence, $\operatorname{rank}(\tilde{A}(\Sigma)) \le p(p+1)/2 - 1$, which implies that $\operatorname{rank}(A(\Sigma)_{\cdot,E}) \le p(p+1)/2 - 1$ and thus $\det(A(\Sigma)_{\cdot,E}) = 0$.

The fact that $\det(H(\Sigma)_{E^c,\cdot})$ is a factor of $\det(A(\Sigma)_{\cdot,E})$ follows from the proof of Lemma 6.4. ∎

REFERENCES

R. F. Barber, M. Drton, N. Sturma, and L. Weihs (2022), *Half-Trek Criterion for Identifiability of Latent Variable Models*, preprint, https://arxiv.org/abs/2201.04457.

S. Barnett and C. Storey (1967), *Analysis and synthesis of stability matrices*, J. Differential Equations, 3, pp. 414–422.

D. S. Bernstein (2011), *Matrix Mathematics: Theory, Facts, and Formulas*, 2nd ed., Princeton University Press.

A. Bhaya, E. Kaszkurewicz, and R. Santos (2003), *Characterizations of classes of stable matrices*, Linear Algebra Appl., 374, pp. 159–174.

C. Brito and J. Pearl (2006), *Graphical condition for identification in recursive SEM*, in Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence, pp. 47–54.

D. Cifuentes, T. Kahle, and P. Parrilo (2020), *Sums of squares in Macaulay*2, J. Softw. Algebra Geom., 10, pp. 17–24.

P. Dettling, M. Drton, and M. Kolar (2022), *On the Lasso for Graphical Continuous Lyapunov Models*, preprint, https://arxiv.org/abs/2208.13572.

M. Drton (2018), *Algebraic problems in structural equation modeling*, in The 50th Anniversary of Gröbner Bases, Adv. Stud. Pure Math. 77, Mathematical Society of Japan, Tokyo, pp. 35–86.

M. Drton, R. Foygel, and S. Sullivant (2011), *Global identifiability of linear structural equation models*, Ann. Statist., 39, pp. 865–886.

M. Drton and L. Weihs (2016), *Generic identifiability of linear structural equation models by ancestor decomposition*, Scand. J. Stat., 43, pp. 1035–1045.

K. E. Fitch (2019), *Learning Directed Graphical Models from Gaussian Data*, preprint, https://arxiv.org/abs/1906.08050.

R. Foygel, J. Draisma, and M. Drton (2012), *Half-trek criterion for generic identifiability of linear structural equation models*, Ann. Statist., 40, pp. 1682–1713.

R. A. Horn and C. R. Johnson (1991), *Topics in Matrix Analysis*, Cambridge University Press.

D. Kumor, B. Chen, and E. Bareinboim (2019), *Efficient identification in linear structural causal models with instrumental cutsets*, in Advances in Neural Information Processing Systems 32, Curran Associates, https://proceedings.neurips.cc/paper/2019/hash/2a845d4d23b883acb632fefd814e175f-Abstract.html.

M. Maathuis, M. Drton, S. Lauritzen, and M. Wainwright, eds., (2019), *Handbook of Graphical Models*, Chapman & Hall/CRC Handbooks of Modern Statistical Methods, CRC Press, Boca Raton, FL.

J. R. Magnus and H. Neudecker (1999), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Wiley Ser. Probab. Stat., John Wiley & Sons, Chichester.

J. Pearl (2009), *Causality. Models, Reasoning, and Inference*, 2nd ed., Cambridge University Press.

J. Peters, D. Janzing, and B. Schölkopf (2017), *Elements of Causal Inference. Foundations and Learning Algorithms*, Adapt. Comput. Mach. Learn., MIT Press, Cambridge, MA.

H. Risken (1996), *The Fokker-Planck Equation. Methods of Solution and Applications*, Springer.

K. Sachs, O. Perez, D. Pe'er, D. A. Lauffenburger, and G. P. Nolan (2005), *Causal protein-signaling networks derived from multiparameter single-cell data*, Science, 308, pp. 523–529.

P. Spirtes, C. Glymour, and R. Scheines (2000), *Causation, Prediction, and Search*, 2nd ed., MIT Press, Cambridge, MA.

S. Sullivant (2018), *Algebraic Statistics*, Grad. Stud. Math. 194, American Mathematical Society, Providence, RI.

G. Varando and N. R. Hansen (2020), *Graphical continuous Lyapunov models*, in Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence, pp. 989–998.

W. C. Young, K. Y. Yeung, and A. E. Raftery (2019), *Identifying dynamical time series model parameters from equilibrium samples, with application to gene regulatory networks*, Stat. Model., 19, pp. 444–465.