Infinitesimal Interconnection Variation in Nonlinear Networked Systems

Insoon Yang  Samuel A. Burden  S. Shankar Sastry  Claire J. Tomlin

Abstract— We propose a novel infinitesimal variation for a nonlinear networked system’s behavior when its interconnection topology changes discontinuously. We introduce a variational derivative of system output with respect to the connectivity, and derive an analytic formula for the derivative using an adjoint formulation. We provide bounds relating the discontinuous change in system behavior to the proposed continuous infinitesimal variation. The variational derivative can be used as the sensitivity of the system output to the interconnection topology. The separability of the variational derivative allows us to develop a tractable algorithm for an interconnection pursuit problem applicable to optimization and inference in biochemical reaction networks.

I. INTRODUCTION

Interconnected systems naturally model the dynamic interactions of mobile sensors [1], [2], social agents [3] and biochemical species [4]. Network topology plays a critical role in determining the behavior of these systems through coupling with the nonlinear continuous dynamics. Consequently, it is important to determine the effect of variations in network connectivity.

A standard method to vary networks involves adding or deleting a subset of connections between nodes. This discrete variational approach has been applied to a number of problems including connectivity control [5] and network identification [6]. In some applications it is more appropriate to consider a small variation in the strength of an interconnection rather than its discrete variation. Such continuous variational approaches are adopted in robust consensus of interconnected systems [7] and distributed routing problems in dynamical networks [8].

Sensitivity of system output with respect to connectivity can be a useful measure of the effect of variations in the interconnection on the output. To define sensitivity with respect to a graph, we consider a variation of the interconnection changes in response to a variation in $E$. The set $V$ of vertices will remain fixed even if a vertex does not connect to any other vertices.

infinitesimal variation. The variational derivative can be used to consider a small variation in the strength of an interconnection rather than its discrete variation. Such continuous variation has been applied to a number of problems including connectivity control [5] and network identification [6]. In some applications it is more appropriate to consider a small variation in the strength of an interconnection rather than its discrete variation. Such continuous variational approaches are adopted in robust consensus of interconnected systems [7] and distributed routing problems in dynamical networks [8].

Sensitivity of system output with respect to connectivity can be a useful measure of the effect of variations in the interconnection on the output. To define sensitivity with respect to a graph, we consider a variation of the interconnection that is invariant over time with the representative graph $G = (V, E)$ where $V$ denotes the (finite) set of vertices and $E$ the set of edges. Consider the interconnected system

$$\dot{x}(t) = f_E(x(t)), \quad x(0) = x_0 \in \mathcal{X},$$

where $\mathcal{X} \subseteq \mathbb{R}^n$ is the continuous state space and $f_E : \mathcal{X} \rightarrow \mathbb{R}^n$ is the nonlinear vector field. We restrict our analysis to a subset $\Gamma \subseteq 2^{V \times V}$ of all possible graphs over the vertex set $V$ containing only the feasible interconnections.

Definition 1. The set $\Gamma \subseteq 2^{V \times V}$ of feasible interconnections for system (1) contains all $E \in 2^{V \times V}$ such that the vector field $f_E(\cdot)$ is twice differentiable, has a continuous second derivative, and is globally Lipschitz continuous in $X$.

Throughout this paper we will assume a continuously differentiable output function $h : \mathcal{X} \rightarrow \mathbb{R}$ is associated with the system (1) and we will let $y(t) = h(x(t))$ denote the output trajectory associated with any solution $x : \mathbb{R} \rightarrow \mathcal{X}$ of (1). Our aim is to investigate how $\gamma$ changes in response to a variation in $E$. The set $V$ of vertices will remain fixed even if a vertex does not connect to any other vertices.
We begin by introducing a family of dynamical systems that vary continuously between two interconnection topologies as a parameter $\epsilon$ is varied from 0 to 1. Subsequently, we develop an adjoint-based technique to compute the infinitesimal variation of trajectories to interconnection variations.

**B. Parameterizing Interconnection Variation**

For any pair $E, \tilde{E} \in \Gamma$, we construct a vector field called the variation from $E$ to $\tilde{E}$,  

$$ f_{\epsilon}(\epsilon, \tilde{E}) = f_{\tilde{E}} - f_{E}, $$  

and use this to define the $\epsilon$-variation from $E$ to $\tilde{E}$,  

$$ f_{\epsilon}(E, \tilde{E}) = f_{E} + \epsilon f_{\epsilon}(\epsilon, \tilde{E}), \quad \epsilon \in [0, 1]. $$  

An illustration of this variation is shown in Fig. 1.  

**Example 1.** Fig. 1 contains an example of interconnections $E$ and $\tilde{E}$ with an interpretation of $f_{\epsilon}(E, \tilde{E})$. The original graph in (a) consists of $V = \{A, B, C\}$ and $E = \{(A, B), (B, C), (C, A)\}$. The interconnection in (b) is given by $\tilde{E} = \{(B, C), (C, A)\}$. The $\epsilon$-variational vector field $f_{\epsilon}(\epsilon, \tilde{E})$ takes into account the small degradation of the interconnection via $\{(A, B)\}$ as depicted in (c). We will study this interconnection structure more detail in Section IV.

Let $E, \tilde{E} \in \Gamma$ and consider the dynamics of the $\epsilon$-variation from $E$ to $\tilde{E}$,  

$$ \dot{x}(t) = f_{\epsilon}(\epsilon, \tilde{E})(x(t)), \quad x(0) = x_{0} \in \mathcal{X}, $$  

which we call the $\epsilon$-variational system associated with $(E, \tilde{E})$. With $\Delta x^\epsilon(\cdot) = x^\epsilon(\cdot) - x(\cdot)$ denoting the difference between trajectories satisfying (1) and (4) with the same initial condition $x^\epsilon(0) = x(0) = x_{0} \in \mathcal{X}$,  

$$ \Delta x^\epsilon = f_{\epsilon}(\epsilon, \tilde{E})(x^\epsilon) - f_{E}(x), $$  

$$ \Delta x^\epsilon = f_{E}(\Delta x^\epsilon + x) - f_{\tilde{E}}(x) + \epsilon f_{\epsilon}(\epsilon, \tilde{E})(\Delta x^\epsilon + x). $$  

Taking the Taylor approximation,  

$$ \Delta x^\epsilon = \frac{\partial f_{E}}{\partial x} \Delta x^\epsilon + \epsilon f_{\epsilon}(\epsilon, \tilde{E})(\Delta x^\epsilon + x) + H(\Delta x^\epsilon, x) $$  

where $H := (H_{1}, \cdots, H_{n})$ and $H_{i}$ denotes the higher-order terms in the Taylor expansion of $(f_{E})(\Delta x^\epsilon + x(t))$, i.e. by applying the mean value theorem [13],  

$$ H_{i}(\Delta x^\epsilon(t), x(t)) = \int_{0}^{1} (1 - s) \frac{\partial (f_{E})(s \Delta x^\epsilon(t))}{\partial s} \Delta x^\epsilon(t) \, ds. $$

Since the vector fields in (4) and (5) are Lipschitz continuous for all $\epsilon \in [0, 1]$, they generate well-defined and bounded trajectories in a bounded time interval; this follows from Proposition 5.6.5 in [14].

**Lemma 1.** For all $\epsilon \in [0, 1]$, there exist unique trajectories $x^\epsilon(\cdot)$ and $\Delta x^\epsilon(\cdot)$ satisfying the dynamics in (4) and (6), respectively. In addition, $\|x^\epsilon(t)\|$ and $\|\Delta x^\epsilon(t)\|$ are bounded by some constant independent of $\epsilon$ for any $t \in [0, T]$.

The next question to address is how the $\epsilon$-variational system behaves as $\epsilon$ tends to zero as compared to the original system. The following Lemma shows that the difference $\Delta x^\epsilon = x^\epsilon - x$ is Lipschitz continuous in $\epsilon$; a version of this result appears as Lemma 5.6.7 in [14].

**Lemma 2.** There exists $L \in \mathbb{R}$ such that for all $t \in [0, T]$  

$$ \|\Delta x^\epsilon(t)\| = \|x^\epsilon(t) - x(t)\| \leq L\epsilon. $$

This Lemma allows us to directly deduce, for all $t \in [0, T]$,  

$$ \lim_{\epsilon \to 0^{+}} \|\Delta x^\epsilon(t)\| = 0, \quad \lim_{\epsilon \to 0^{+}} \frac{1}{\epsilon} \|\Delta x^\epsilon(t)\|^{2} = 0. $$

The previous Lemmas with the dominated convergence theorem [15] also yields the following.

**Corollary 1.** For any $t \in [0, T]$,  

$$ \lim_{\epsilon \to 0^{+}} \frac{1}{\epsilon} \int_{0}^{t} \|\Delta x^\epsilon(s)\|^{2} \, ds = 0. $$

Lemmas 1 and 2 together with Corollary 1 imply that for $\epsilon \in [0, 1]$ and $E, \tilde{E} \in \Gamma$ the $\epsilon$-variation of the interconnection as $\epsilon$ tends to zero.

**C. Infinitesimal Interconnection Variation**

We seek a first-order approximation to the change in system behavior caused by a variation of interconnection topology. Such an approximation is an intuitively appealing generalization of the sensitivity in continuous systems. Defining first-order variations with respect to the connectivity structure requires a distance metric defined over interconnection topologies. We define the distance between any two feasible connections $E$ and $\tilde{E}$ as 1 and the distance between $E$ and its $\epsilon$-variation $\epsilon(E, \tilde{E})$ as $\epsilon \in [0, 1]$. We then use the corresponding $\epsilon$-variational system (4) to obtain the $\epsilon$-variation of the system output $y' := h(x^\epsilon)$. As $\epsilon$ tends to zero, we have a well-defined derivative of the output with respect to the interconnection, which we call the variational derivative of the output from $E$ to $\tilde{E}$.
Definition 2. For $E, \tilde{E} \in \Gamma$ we define the variational derivative from $E$ to $\tilde{E}$ of the output as

$$D(E, \tilde{E})y(t) := \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} (h(x^\epsilon(t)) - h(x(t)))$$

where $x^\epsilon$ solves (4) and $x$ solves (1) with $x(0) = x(0)$.

Qualitatively, the variational derivative $D(E, \tilde{E})y(t)$ measures the sensitivity of the output with respect to the network change from $E$ to $\tilde{E}$. We aim to prove that the variational derivative (7) is well-defined and bounded for $t \in [0, T]$. We will show this by explicitly deriving an analytic formula of $D(E, \tilde{E})y(t)$ using the adjoint system of (1) at time $t$:

$$-\dot{\mu}(s) = \frac{\partial f_E(x(s))}{\partial x}^\top \left( \mu(s) - \frac{\partial h(x(t))}{\partial x} \right), s \in [0, t]$$

$$\mu(t) = 0.$$  

(8)

Here, $\mu(s)$ is called the adjoint state. Unique solutions to (8) exist for all $t \in [0, T]$ and $E \in \Gamma$ since the definition of feasible interconnections guarantees continuous differentiability of $f_E$ and boundedness of $x(s)$ for $s \in [0, t]$.

Theorem 1. The variational derivative (7) can be obtained as

$$D(E, \tilde{E})y(t) = \int_0^t \left[ \frac{\partial h(x(t))}{\partial x} - \mu(s)^\top \right] f_{(E, \tilde{E})}(x(s)) \, ds,$$

where $x$ and $\mu$ solve (1) and (8), respectively, and $f_{(E, \tilde{E})}$ is defined in (2). Hence, the variational derivative is well-defined and bounded for any $t \in [0, T]$.

Proof. Let $O := O(\int_0^t \|\Delta x^\epsilon(s)\|^2 \, ds) + O(\epsilon^2)$ denote high order terms. We then have $\lim_{\epsilon \to 0^+} O/\epsilon = 0$ due to Lemma 2 and Corollary 1. Note that $\epsilon f_{(E, \tilde{E})}(x(t) + \Delta x^\epsilon(t)) = \epsilon f_{(E, \tilde{E})}(x(t)) + O$, we can rewrite system (6) as

$$\Delta x^\epsilon(t) = \frac{\partial f_E(x(t))}{\partial x} \Delta x^\epsilon(s) + \epsilon f_{(E, \tilde{E})}(x(t)) + O,$$

which implies

$$\Delta x^\epsilon(t) = \int_0^t \frac{\partial f_E(x(s))}{\partial x} \Delta x^\epsilon(s) + \epsilon f_{(E, \tilde{E})}(x(s)) \, ds + O.$$  

We now consider the difference

$$h(x^\epsilon(t)) - h(x(t)) = \frac{\partial h(x(t))}{\partial x} \Delta x^\epsilon(t) + O,$$

in which we use the Taylor expansion of $h(\cdot)$. We now take the inner product of both sides of (10) with the adjoint state $\mu$ and subtract it from the difference (11) to obtain

$$h(x^\epsilon(t)) - h(x(t)) = \int_0^t \left[ \frac{\partial h(x(t))}{\partial x} \Delta x^\epsilon(s) + \epsilon f_{(E, \tilde{E})}(x(s)) \right] ds + O$$

Expanding

$$h(x^\epsilon(t)) - h(x(t)) = \int_0^t \left[ \frac{\partial h(x(t))}{\partial x} \Delta x^\epsilon(s) + \epsilon f_{(E, \tilde{E})}(x(s)) \right] ds + O$$

in which we let $\lambda(s, t) = \left( \frac{\partial h(x(t))}{\partial x} - \mu(s)^\top \right)$ for notational simplicity. In the second equality, we used integration by parts $\int_0^t \mu(s)^\top \Delta x^\epsilon(s) \, ds = \mu(t)^\top \Delta x^\epsilon(t) - \mu(0)^\top \Delta x^\epsilon(0) - \int_0^t \mu(s)^\top \Delta x^\epsilon(s) \, ds$ with $\mu(t) = 0$ and $\Delta x^\epsilon(0) = 0$. The second integral in the righthand side of the last equality equals zero due to the definition of $\mu$ in (8). Therefore we have the desired estimate.

From the proof, for $y^\epsilon := h(x^\epsilon)$ we have the estimate

$$y^\epsilon(t) = y(t) + \epsilon D(E, \tilde{E})y(t) + O(\epsilon^2)$$

for sufficiently small $\epsilon$, which is analogous to the Taylor expansion of $y^\epsilon(t)$ with respect to $\epsilon$. For the important case $\epsilon = 1$, we have the following.

Proposition 1. Let $\Delta_{(E, \tilde{E})}y := y^1 - y$ denote the output difference between (5) and (1) when $\epsilon = 1$. Then there exists $K \in \mathbb{R}$ such that for all $s \in (0, t)$ we have

$$|\Delta_{(E, \tilde{E})}y(t) - D_{(E, \tilde{E})}y(t)| \leq K \int_0^t \|\Delta x^1(s)\| \, ds.$$  

Proof. Choose $\epsilon = 1$ in (6). Then, instead of (10), we have

$$\Delta x^\epsilon(s) = \frac{\partial f_E(x(s))}{\partial x} \Delta x^\epsilon(s) + \epsilon f_{(E, \tilde{E})}(x(s)) + \hat{O}(s)$$

for small $\|\Delta x^1(s)\|$, where $\hat{O}(s) = O(\|\Delta x^1(s)\|)$. Therefore, (11) should be rewritten with $\hat{O}$ as follows:

$$h(x^1(t)) - h(x(t)) = \int_0^t \left[ \frac{\partial h(x(t))}{\partial x} - \mu(s)^\top \right] f_{(E, \tilde{E})}(x(s)) \, ds + \int_0^t \hat{O}(s) \, ds$$

$$D_{(E, \tilde{E})}y(t) + \int_0^t \hat{O}(s) \, ds.$$  

Recall that $\Delta_{(E, \tilde{E})}y(t) = h(x^1(t)) - h(x(t))$ by definition, we have the desired estimate.
interconnection change $(\mathcal{E} \setminus \tilde{\mathcal{E}}) \cup (\tilde{\mathcal{E}} \setminus \mathcal{E})$ in the following sense. Let $\epsilon \in [0, 1]$ parameterize the intensity of the interaction among nodes via $(\mathcal{E} \setminus \tilde{\mathcal{E}}) \cup (\tilde{\mathcal{E}} \setminus \mathcal{E})$. Then the variational derivative yields the sensitivity of the output with respect to the perturbation that slightly degrades the efficiency of $\mathcal{E} \setminus \tilde{\mathcal{E}}$ and infinitesimally generates the interaction through $\tilde{\mathcal{E}} \setminus \mathcal{E}$. For this reason, we believe the variational derivative will be useful for analyzing output sensitivity in nonlinear biochemical and power system models.

D. Computing Infinitesimal Interconnection Variation

Computation of the variational derivative $D_{(\mathcal{E}, \tilde{\mathcal{E}})} y(t)$ using (9) requires the solutions of the system (1) and its adjoint (8). Due to nonlinearity, it is not generally possible to find an analytic expression for solutions of (1) (and hence (8)). Algorithm 1 contains a straightforward algorithm to numerically approximate the variational derivative.

We discretize $[0, T]$ into $N \in \mathbb{N}$ equally-spaced time intervals, then apply the Forward Euler algorithm to obtain the numerical approximation $x_k$ to (1) for each $k \in \{0, \ldots, N\}$. To evaluate the variational derivative at time $t = iT/N$, we apply Forward Euler to the adjoint dynamics (8) backward in time to obtain $\mu_k$ for each $k \in \{0, \ldots, i\}$, then numerically integrate (i.e. sum) the formula in (9).

Algorithm 1: Variational derivative

1 Initialization:
2 Given $x_0 \in \mathbb{R}^n$ and $\mathcal{E}, \tilde{\mathcal{E}} \in \Gamma$;
3 Approximate solution of (1):
4 for $k = 0 : N - 1$ do
5 \hspace{1em} $x_{k+1} \leftarrow x_k + \frac{T}{N} f(\mathcal{E}) x_k$;
6 end
7 Variational derivative:
8 for $i = 0 : N$ do
9 \hspace{1em} Approximate solution of (8):
10 \hspace{2em} $\mu_i \leftarrow 0$;
11 \hspace{2em} for $k = i - 1 : 0$ do
12 \hspace{3em} $\mu_k \leftarrow \mu_{k+1} - \frac{1}{N} \frac{\partial f(\mathcal{E}) x_k}{\partial x}^T \left( \mu_{k+1} - \frac{\partial h(x_k)}{\partial x}^T \right)$;
13 \hspace{2em} end
14 Approximation of (9):
15 \hspace{2em} $D_{(\mathcal{E}, \tilde{\mathcal{E}})} y_i \leftarrow \frac{T}{N} \sum_{k=0}^i \left( \frac{\partial h(x_k)}{\partial x} - \mu_k^T \right) f(\mathcal{E}) x_k$;
16 end

E. Infinitesimal Interconnection Variation at Equilibria

An important special case concerns the sensitivity of the system (1) at an equilibrium $\xi \in X$ where $f(\mathcal{E}) \xi = 0$. In applications, exponentially stable equilibria may correspond to a setpoint in a biochemical reaction network or nominal load conditions in an electrical power system, and there exists a rich set of tools to ensure existence and stability of equilibria for such interconnected systems [17], [18].

Consider an exponentially stable equilibrium trajectory $x$ for interconnection $\mathcal{E}$ of (1) so that $\dot{x}(t) = \xi$ for all $t \in [0, T]$ and $\Re \lambda < 0$ for all $\lambda \in \text{spec } D(\mathcal{E}) \xi$. Then we have $\frac{\partial h(x(t))}{\partial x} := b_f \frac{\partial x(t)}{\partial x} = 0$, and $\frac{\partial f(\mathcal{E}) x(t)}{\partial x} := A \mu$ independent of $t \in [0, T]$, whence the adjoint satisfies the linear time-invariant differential equation

$$-\dot{\mu}(s) = A^T (\mu(s) - b^T), \quad s \in [0, t], \quad \mu(t) = 0$$

with solution $\mu(s) = -e^{A^T(t-s)} b^T + b^T$, $s \in [0, t]$. This leads to the simplified expression for $D_{(\mathcal{E}, \tilde{\mathcal{E}})} y(t)$.

III. INTERCONNECTION PURSUIT

Motivated by applications, we consider an optimization problem over interconnection topologies. Specifically, given an initial state $x(0) \in X$, interconnection topology $\mathcal{E} \in \Gamma$, and time $\tau \in [0, T]$, we seek a new interconnection $\tilde{\mathcal{E}} \in \Gamma$ that maximizes $\Delta_{(\mathcal{E}, \tilde{\mathcal{E}})} y(\tau) := h(x(\mathcal{E})(\tau)) - h(x(\mathcal{E})(\tau))$, the difference in the outputs at time $\tau$.

**Optimization 1 (Interconnection Pursuit).**

$$\max_{\mathcal{E} \in \Gamma} \Delta_{(\mathcal{E}, \tilde{\mathcal{E}})} y(\tau)$$

This single optimization subsumes several important problems in the context of biochemical reaction networks: design of in vitro circuits that maximize yield of a desirable species $x_i$ can be obtained by choosing $h(x) = x_i$; inference of the in vivo network most compatible with observation data $\eta(\tau)$ can be achieved using $h(x) = -||x - \eta(\tau)||$.

A. Additive Interconnections

Motivated by applications, we consider the special case where interconnections enter additively in (1). This means that for each $\mathcal{E} \in \Gamma$ the vector field $f(\mathcal{E})$ is obtained by adding one term for each edge $e \in \mathcal{E}$, and we assume all possible
interconnections are feasible for the set \( V \) of nodes, i.e. \( \Gamma = 2^{V} \times V \). In the following definition, we regard the vector field as a function \( f : X \times \Gamma \to \mathbb{R}^n \) where \( f(\cdot, \mathcal{E}) = f_{\mathcal{E}}(\cdot) \).

**Definition 3.** The vector field \( f : X \times \Gamma \to \mathbb{R}^n \) has additive interconnections if for all \( \mathcal{E} \in \Gamma \)

\[
f_{\mathcal{E}} = f_0 + \sum_{e \in \mathcal{E}} (f_{(e)} - f_0),
\]

where \( f_0 \) denotes the vector field with no interconnections.

The vector field with additive interconnections implies that the variational derivative is separable over edges, i.e.,

\[
D_{(\mathcal{E}, \xi)} y(t) = \sum_{e \in \mathcal{E}} D_{(\mathcal{E}, \mathcal{E} \cup \{e\})} y(t) 
+ \sum_{e \in \mathcal{E} \setminus \mathcal{E}} D_{(\mathcal{E}, \mathcal{E} \setminus \{e\})} y(t).
\]

As we will see in Section III-B, this decomposition will allow us to solve an approximate version of the interconnection pursuit problem with an algorithm that only requires the evaluation of \( D_{(\mathcal{E}, \mathcal{E} \cup \{e\})} y(t) \) for \( e \notin \mathcal{E} \) and \( D_{(\mathcal{E}, \mathcal{E} \setminus \{e\})} y(t) \) for \( e \in \mathcal{E} \).

**Proposition 2.** If \( f : X \times \Gamma \to \mathbb{R}^n \) has additive interconnections, then for all \( \mathcal{E}, \tilde{\mathcal{E}} \in \Gamma \) the variational derivative \( D_{(\mathcal{E}, \tilde{\mathcal{E}})} y(t) \) satisfies (18).

**Proof.** Since \( f \) has additive interconnections, the variation \( f_{(\mathcal{E}, \xi)} \) is given by

\[
f_{(\mathcal{E}, \xi)} = f_{\mathcal{E}} - f_{\mathcal{E}} = \sum_{e \in \mathcal{E}} (f_{(e)} - f_0) = \sum_{e \in \mathcal{E} \setminus \mathcal{E}} (f_{(e)} - f_0) 
+ \sum_{e \in \mathcal{E} \setminus \mathcal{E}} (f_{(e)} - f_0).
\]

Now since the variation \( f_{(\mathcal{E}, \xi)} \) enters linearly into (9) and letting \( \lambda(s, t) = \left( \frac{\partial h(x(t))}{\partial x} - \mu(s)^T \right) \) for notational simplicity,

\[
D_{(\mathcal{E}, \xi)} y(t) = \int_0^t \lambda(s, t) f_{(\mathcal{E}, \xi)}(x(s)) ds 
+ \sum_{e \in \mathcal{E} \setminus \mathcal{E}} \int_0^t \lambda(s, t) f_{(\mathcal{E}, \mathcal{E} \setminus \{e\})}(x(s)) ds,
\]

which is equivalent to (18).

**Proposition 3.** Suppose that the vector field \( f \) has additive interconnections. Let \( \xi \) and \( \tilde{\xi} \) be exponentially stable equilibria of \( \dot{x} = f_{\mathcal{E}}(x) \) and \( \dot{x} = f_{\tilde{\mathcal{E}}}(x) \), respectively. Then

\[
|\Delta_{(\mathcal{E}, \xi)} y(T) - D_{(\mathcal{E}, \xi)} y(T)| 
\leq \left\| \frac{\partial h(\xi)}{\partial x} \left( \frac{\partial f_{\mathcal{E}}(\xi)}{\partial x} \right)^{-1} \frac{\partial f_{(\mathcal{E}, \xi)}(\xi)}{\partial x} \right\| \|\xi - \tilde{\xi}\|
+ O(e^{aT}) + O(\|\xi - \tilde{\xi}\|^2)
\]

where \( a = \max \left\{ \Re \lambda : \lambda \in \text{spec} \left( \frac{\partial f_{\mathcal{E}}(\xi)}{\partial x} \right) \right\} < 0 \).

The proof is contained in the Appendix. Proposition 3 implies that for sufficiently large \( T \) and small \( \|\xi - \tilde{\xi}\| \), the variation \( D_{(\mathcal{E}, \xi)} y(t) \) approximates the actual difference in system behavior \( \Delta_{(\mathcal{E}, \xi)} y(T) \) as the interconnection changes from \( \mathcal{E} \) to \( \tilde{\mathcal{E}} \). Although the relationship between \( \Delta_{(\mathcal{E}, \xi)} y(T) \) and \( D_{(\mathcal{E}, \xi)} y(T) \) for any \( T \in (0, T] \) warrants further study beyond Proposition 1, we will use \( D_{(\mathcal{E}, \xi)} y(T) \) in the approximate interconnection pursuit problem proposed below. Additive interconnections and the separability of the variational derivative enable us to solve this approximate pursuit problem efficiently relative to the original (NP-hard) interconnection pursuit problem.

**B. Approximate Interconnection Pursuit**

The analysis in the previous section suggests a tractable approximation to the interconnection pursuit problem (16) for systems with additive interconnections. We assume that \( \tau \in (0, T] \) is fixed, all possible interconnections are feasible, and the vector field has additive interconnections.

**Optimization 2** (Approximate Interconnection Pursuit).

\[
\max_{\mathcal{E} \in \Gamma} D_{(\mathcal{E}, \xi)} y(T)
\]

To study how efficiently we can solve this problem, we first let \( \mathcal{E} := \{e_1^+, \cdots, e_L^+\} \) and \( \mathcal{E}^* \setminus \mathcal{E} := \{e_1^-, \cdots, e_L^-\} \) and , where \( \mathcal{E}^* \) is the maximal interconnection (i.e., arbitrary two nodes are connected) and \( L := |\mathcal{E}| \) and \( M := |\mathcal{E}^*| \). We also introduce the following vectors of variational derivatives

\[
c^+ := \left( D_{(\mathcal{E}, \mathcal{E} \cup \{e_i^+\})} y(T), \cdots, D_{(\mathcal{E}, \mathcal{E} \cup \{e_{M-L}^+\})} y(T) \right),
\]

\[
c^- := \left( D_{(\mathcal{E}, \mathcal{E} \cup \{e_i^-\})} y(T), \cdots, D_{(\mathcal{E}, \mathcal{E} \cup \{e_{M-L}^-\})} y(T) \right),
\]

and let \( c := [c^+, c^-] \in \mathbb{R}^M \). Due to the separability of the variational derivative, we have

\[
D_{\mathcal{E}, \xi} y(T) = c^T \alpha
\]

with \( \alpha \in \mathbb{R}^M \) such that

\[
\alpha_i = \begin{cases} 
1 & \text{if } e_i^+ \in \mathcal{E} \text{ or } e_i^- \notin \mathcal{E}, \\
0 & \text{otherwise.}
\end{cases}
\]

**Optimization 2** is then equivalent to

\[
\max_{\alpha \in \mathbb{R}^M} c^T \alpha
\]

subject to \( \alpha_i \in \{0, 1\}, \ i = 1, \cdots, M \).
This problem has an explicit solution given by
\[ \alpha_i = \begin{cases} 1 & \text{if } c_i > 0, \\ 0 & \text{otherwise}. \end{cases} \quad i = 1, \ldots, M, \]
which motivates the following tractable algorithm:

**Algorithm 2: Approximate interconnection pursuit**

1. **Initialization:**
2. \( \tilde{E} \leftarrow E; \)
3. Fix \( \tau \in (0, T]; \)
4. Compute the vectors of variational derivates \( c^+ \) and \( c^- \) given by (20);
5. **Interconnection pursuit:**
   6. for \( i = 1 : M - L \) do
      7. \[ \text{if } c_i^+ > 0 \text{ then} \]
      8. \[ \tilde{E} \leftarrow \tilde{E} \cup \{ e_i^+ \}; \]
      9. end
   10. end
   11. for \( i = 1 : L \) do
      12. \[ \text{if } c_i^- > 0 \text{ then} \]
      13. \[ \tilde{E} \leftarrow \tilde{E} \backslash \{ e_i^- \}; \]
      14. end
   15. end

Note that lines 6–15 can be completely parallelized because the order of the additions and the deletions does not matter. In practice, we choose a threshold \( \delta \geq 0, \) and add \( e_i^+ \notin \tilde{E} \) if \( D(\tilde{E}, E \cup \{ e_i^+ \})y(\tau) > \delta \) and remove \( e_i^- \notin \tilde{E} \) if \( D(\tilde{E}, E \backslash \{ e_i^- \})y(\tau) > \delta \) to account for error introduced by the approximation (3). While the original interconnection pursuit problem requires us to search all \( 2^M \) possible interconnections to find a global optimal solution, the approximate interconnection pursuit problem can be very efficiently solved with Algorithm 2 that has \( O(M) \) complexity.

**IV. Example**

The three-species biochemical circuit has played an important role in illuminating fundamental properties of complex signaling networks such as biochemical adaptation [19] and dynamic correlations in biochemical noise [20]. In this section, we consider the biochemical network in Figure 2 (a) and investigate which edges should be deleted to maximize the increase in the concentration of active species \( pB. \) We first represent the biochemical circuit in Figure 2 (a) as a graph \( \mathcal{G} = (V, E) \) with \( V = \{ A, B, C \} \) and \( E = \{ e_1, e_2, e_3 \}, \) where \( e_1 = (A, B), \) \( e_2 = (B, C) \) and \( e_3 = (C, A). \) In the biochemical network, species \( A \) activates (i.e., phosphorylates) \( B, \) \( B \) activates \( C, \) and species \( C \) inhibits (i.e., dephosphorylates) \( A. \) The following biochemical equations allow us to examine the interaction among the species:

- Activation of \( B \) by \( A \) (edge \( e_1 = (A, B) \))
  \[ pA + B \xrightarrow{k_1} pAB, \quad pAB \xrightarrow{k_2} pA + pB \]
- Inhibition of \( A \) by \( C \) (edge \( e_3 = (C, A) \))
  \[ pC + pA \xrightarrow{k_{10}} pCA, \quad pCA \xrightarrow{k_{11}} pC + A. \]

Here, \( pA \) and \( A \) denote the active (phosphorylated) version and the inactive version of \( A, \) respectively, and \( pAB \) is the complex of \( A \) and \( B. \) Others are defined in a similar way. Let \( x_1 = [pA], \) \( x_2 = [A], \) \( x_3 = [pB], \) \( x_4 = [B], \) \( x_5 = [pC], \) \( x_6 = [C], \) \( x_7 = [pAB], \) \( x_8 = [pBC] \) and \( x_9 = [pCA], \) where \( [M] \) denotes the concentration level of protein \( M. \) The dynamics of the biochemical concentrations with the signaling network \( \mathcal{G} \) in Figure 2 (a) can then be modeled by (1) with the vector field \( f_\mathcal{E} \) such that

\[
(f_\mathcal{E})_1 = -k_1 x_1 x_4 + k_2 x_7 + k_3 x_7 - k_4 x_1 x_3 - k_9 x_1 x_5 + k_{10} x_9 \\
(f_\mathcal{E})_2 = k_{11} x_9 - k_{12} x_2 x_5 \\
(f_\mathcal{E})_3 = k_{37} x_7 - k_{34} x_1 x_3 - k_{32} x_2 x_6 + k_9 x_8 + k_{78} x_7 - k_{83} x_5 \\
(f_\mathcal{E})_4 = -k_1 x_1 x_4 + k_2 x_7 \\
(f_\mathcal{E})_5 = k_{58} x_5 - k_{82} x_2 x_5 - k_{83} x_1 x_5 + k_{10} x_9 + k_{11} x_9 - k_{12} x_2 x_5 \\
(f_\mathcal{E})_6 = -k_5 x_3 x_6 + k_6 x_8 \\
(f_\mathcal{E})_7 = k_{11} x_1 x_7 - k_3 x_7 + k_4 x_1 x_3 \\
(f_\mathcal{E})_8 = k_5 x_3 x_6 - k_6 x_8 - k_7 x_5 + k_8 x_3 x_5 \\
(f_\mathcal{E})_9 = k_9 x_1 x_5 - k_1 x_9 - k_{11} x_9 + k_{12} x_2 x_5,
\]

where we used the mass-action kinetics. We can see that the vector field has additive interconnections, i.e.,

\[ f_\mathcal{E} = f_{\{e_1\}} + f_{\{e_2\}} + f_{\{e_2\}} \]

with \( f_0 = 0. \) Note that the vector field is quasi-positive, i.e.,

\[ (f_\mathcal{E})(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_9) \geq 0, \forall i \in \{1, \ldots, 9\}, \]

and the quasi-positivity guarantees the solution to be non-negative invariant. If we let \( X := x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + 2 x_7 + 2 x_8 + 2 x_9, \) then \( \dot{X} = 0. \) Hence, \( x_i \leq X(0) \) due to the non-negative invariance. Therefore, we can set the state space as the compact subset \( X = [0, X(0)]^9 \subset \mathbb{R}^9. \) Furthermore, the vector field \( f_\mathcal{E} \) is Lipschitz continuous and twice differentiable in \( X \) for any interconnection \( \tilde{E}. \)

In this example, we choose the output to be the concentration of the active protein \( pB, \) i.e.,

\[ y(t) = h(x(t)) = x_3(t). \]
We assume that $x_i(0) = 1$ for all $i = 1, \ldots, 9$, and $k_{2j-1} = 1$ and $k_{2j} = 10^{-3}$ for $j = 1, \ldots, 5$.

Our goal is to solve the interconnection pursuit problem to maximize the increase in $|pB|$, i.e., $\Delta_{(E,\tilde{E})}y(\tau)$, where the maximal interconnection is given by $E^* := \{e_1, e_2, e_3\}$. We must consider all possible $E$’s that are subsets of $\{e_1, e_2, e_3\}$. By comparing all $2^3$ cases, we have the optimal interconnection as

$$\{e_1, e_2\} = \arg \max_{E \in \Gamma} \Delta_{(E,\tilde{E})}y(\tau), \quad \tau \in (0, 10].$$

However, for the approximate interconnection pursuit problem, i.e., Optimization 2, however, it is sufficient to compute $c^- = (D_{(E,\tilde{E}\setminus \{e_1\})}y, D_{(E,\tilde{E}\setminus \{e_2\})}y, D_{(E,\tilde{E}\setminus \{e_3\})}y)$ because the vector field has a additive interconnections. Note that $E^* = E$, we have $c^+ = 0$. In Fig. 2 (b)–(d), interconnections $E \setminus \{e_1\}$, $E \setminus \{e_2\}$ and $E \setminus \{e_3\}$ are depicted. Using the adjoint-based formula for the variational derivative (9) and Algorithm 1, we compute the variational derivatives over $(0, 10]$ as shown in Fig. 3. If we use Algorithm 2 (with threshold $\delta = 0$), then the optimum of the approximate problem is as follows:

$$\{e_1, e_2\} = \arg \max_{E \in \Gamma} D_{(E,\tilde{E})}y(\tau), \quad \tau \in (0, T_1],$$

$$\{e_1\} = \arg \max_{E \in \Gamma} D_{(E,\tilde{E})}y(\tau), \quad \tau \in (T_1, 10],$$

where $T_1 \approx 5.46$. In other words, the approximate problem, Optimization 2, finds the global optimum of the original problem, Optimization 1, in $(0, T_1]$. Even in $(T_1, 10]$, the solution obtained by the approximate problem is the third best interconnection out of the total eight connections. Note that, if we choose the threshold $\delta$ as 0.075 or higher, then we can obtain the global optimum of the original problem by solving the approximate problem for all $T \in (0, 10]$.

V. CONCLUSION AND FUTURE WORK

Inspired by the directional derivative approach in functional analysis, we introduced a variational derivative for nonlinear networked systems with respect to the interconnection topology. We derived an analytical expression for the derivative by introducing an adjoint state, and provided bounds relating our derivative to the actual change in system behavior. The additivity assumption on the vector field yielded separability of the variational derivative that further enabled us to develop an efficient solution to an approximate interconnection pursuit problem with applications to optimization and inference of biochemical reaction networks. We also believe the variational derivative will be useful for analyzing output sensitivity in nonlinear biochemical and power system models with respect to the interconnection topology.

In future work we plan to further study convergence and suboptimality of the approximate interconnection pursuit problem and to generalize the variational method to the networked control systems setting where we anticipate finding applications to estimating sensitivity of power transmission systems to connectivity structure.

APPENDIX

PROOF OF PROPOSITION 3

We first note that

$$0 = f_\ell(\xi),$$

$$0 = f_\ell(\hat{\xi}) = f_\ell(\hat{\xi}) + f_\ell(\xi, \hat{\xi}).$$

Subtracting one with another, we have

$$0 = \frac{\partial f_\ell(\xi)}{\partial x} (\hat{\xi} - \xi) + f_\ell(\xi, \hat{\xi}) + O,$n

where $O := O(\|\xi - \hat{\xi}\|^2)$. This implies that

$$\hat{\xi} - \xi = -\left(\frac{\partial f_\ell(\xi)}{\partial x}\right)^{-1} f_\ell(\xi, \hat{\xi}) + O.$$

The Taylor expansion allows us to deduce

$$\Delta_{(E,\tilde{E})}y(T) = h(\hat{\xi}) - h(\xi) = \frac{\partial h(\xi)}{\partial x}(\hat{\xi} - \xi) + O$$

$$= \frac{\partial h(\xi)}{\partial x}\left(\frac{\partial f_\ell(\xi)}{\partial x}\right)^{-1} f_\ell(\xi, \hat{\xi}) + O. \quad (21)$$

On the other hand, we can rewrite (15) as

$$D_{(E,\tilde{E})}y(T) = -\frac{\partial h(\xi)}{\partial x}\left(\frac{\partial f_\ell(\xi)}{\partial x}\right)^{-1} f_\ell(\xi, \hat{\xi}) + O(e^{aT}) \quad (22)$$

because $f_\ell(\xi) = 0$ and the interconnection additivity. Comparing (21) with (22), we obtain

$$|\Delta_{(E,\tilde{E})}y(T) - D_{(E,\tilde{E})}y(T)|$$

$$= \left|\frac{\partial h(\xi)}{\partial x}\left(\frac{\partial f_\ell(\xi)}{\partial x}\right)^{-1} (f_\ell(\xi, \hat{\xi}) - f_\ell(\xi, \hat{\xi})) + O_2\right|$$

$$\leq \left|\frac{\partial h(\xi)}{\partial x}\left(\frac{\partial f_\ell(\xi)}{\partial x}\right)^{-1} \frac{\partial f_\ell(\xi)}{\partial x}\right| \left|\xi - \hat{\xi}\right| + O_2,$$

as desired, where $O_2 := O(e^{aT}) + O(\|\xi - \hat{\xi}\|^2)$. \qed

REFERENCES


