SIAM J. CONTROL OPTIM.
Vol. 57, No. 3, pp. 2240–2268

COST-EFFECTIVE ROBUST STABILIZATION AND BIFURCATION SUPPRESSION*

KEVIN E. M. CHURCH† AND XINZHI LIU†

Abstract. A novel technique of robust stabilization and bifurcation suppression is proposed. The proposed method, the center probe method, stabilizes an equilibrium point of a delay differential equation at a bifurcation point by introducing an impulsive controller that minimizes a given cost functional. The cost functional can weight certain structural properties of the controller, such as the number of nodes controlled (in the stabilization of a complex network). The method takes advantage of the dimension reduction properties of the center manifold, which makes the method notably efficient to implement. A numerical example is provided to demonstrate its effectiveness in suppressing a Hopf bifurcation and robustly stabilizing a nonlinear network model with 100 linearly coupled nodes, while simultaneously keeping the number of controlled nodes to a minimum and minimizing a cost function that assigns higher cost to nodes with higher degree. The strengths and weaknesses of the method compared to other impulsive stabilization techniques are discussed.

Key words. impulsive stabilization, bifurcation suppression, center probe method, complex network

AMS subject classifications. 34K20, 93D09

DOI. 10.1137/18M1213142

1. Introduction.

Stabilization of complex networks and dynamical systems both large scale and small play an important role in science and industry. Since the introduction of Lyapunov's direct method and its various generalizations, the technique has seen much application in the development of sufficient conditions for the stability of steady states, which can themselves be used to derive controllers guaranteeing robust stability and synchronization. The recent survey paper [19] catalogues recent developments in the stability analysis of linear time-delay systems by Lyapunov-based methods, and one may consult the references therein for background. For a short list of specifically nonlinear results, one may consult [11, 25, 26, 27].

Suppose one has an autonomous n-dimensional retarded functional differential equation depending on a parameter $\epsilon \in \mathbb{R}^p$,

$$\dot{x} = f(x_t, \epsilon), \tag{1.1}$$

where $x_t \in C := C([-r, 0], \mathbb{R}^n)$ is the solution history with $r > 0$ finite, and it is known that $x^* = 0$ is an equilibrium point for all $\epsilon \in N \subset \mathbb{R}^p$, a neighborhood of the origin. Systems of this type include linear and nonlinear discrete time-delay systems with or without parameters:

$$\dot{x} = Ax(t) + Bx(t - r)) + f(x(t), x(t - r)),$$

*Received by the editors September 11, 2018; accepted for publication (in revised form) May 13, 2019; published electronically June 27, 2019.
http://www.siam.org/journals/sicon/57-3/M121314.html
Funding: The first author acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) through the Alexander Graham Bell Canada Graduate Scholarships Program.
†Department of Applied Mathematics, University of Waterloo, Waterloo N2L 3G1, ON, Canada (k5church@connect.uwaterloo.ca, xzliu@uwaterloo.ca).

as well as systems with multiple discrete delays and distributed delays. If f in (1.1) is Fréchet differentiable in its first variable, the *linearization* at $x^* = 0$ is the linear system

$$\dot{y} = Df(0,\epsilon)y_t, \tag{1.2}$$

where $Df(0,\epsilon) : C \to \mathbb{R}^n$ is the Fréchet derivative of f (with respect to the first variable) at the parameter ϵ. The dynamics of the linearization determine the local behavior near $x = 0$ for the nonlinear system (1.1). If one defines the *characteristic matrix*

$$\Delta(\lambda;\epsilon) = \lambda I - Df(0,\epsilon)[e^{\lambda(\cdot)}I], \tag{1.3}$$

where I the $n\times n$ identity matrix, then the *eigenvalues* (sometimes called characteristic roots) are the solutions of the generally transcedental *characteristic equation*

$$\det(\Delta(\lambda;\epsilon)) = 0. \tag{1.4}$$

x^* will be exponentially stable if and only if all eigenvalues have strictly negative real part; see [8, 18] for background.

A transition from stability to instability occurs when one or more eigenvalues cross the imaginary axis. Such a transition is an example of a *bifurcation*, and when such a bifurcation occurs, the dynamics of the nonlinear system (1.1) near $x^* = 0$ can quickly become unpredictable. One may consult the survey [6] or the textbooks [4, 7, 10] for background on bifurcation theory. In practical applications, it may be that due to some external influence or component failure, a system parameter enters a regime where stability is lost and a bifurcation occurs. The goal then shifts to *stabilization*.

One stabilization methodology that has seen a fair bit of attention in recent years is *impulsive stabilization* [9, 12, 13, 15, 16, 21, 23, 28]. These are typically proven by means of Lyapunov functionals and are stated in terms of the existence of matrices satisfying linear matrix inequalities. They often provide global stability. However, the assumptions can be somewhat strong: global Lipschitz conditions are typically needed to guarantee convergence and finding matrices satisfying the necessary inequalities can be difficult especially for large interconnected systems.

A related problem in terms of the implementation of impulsive stabilization is that, naturally, some controls may be more difficult to implement than others. In other words, there may be an explicit cost in implementing an impulsive controller. Guaranteed cost impulsive control from has been considered in [14, 24, 29] among others, where the goal is to design the impulsive controller so that a running cost is minimized. To contrast, we are interested in average costs associated to impulsive controllers, where the cost may be dictated by such factors as the control gain or its structure. The latter encompasses such factors as the amount of coupling induced by the controller, the amount of diffusivity or lack thereof, or a penalty for accessing or modifying certain system states.

Simultaneously, there may be hard constraints to the types of impulsive controllers that are permitted. In pinning control, for example, the jump functionals do not induce any additional coupling between nodes—see the aforementiond references. In an input-output setting, if one only has access to system outputs, then one might want the controller to depend only on the measurements. See later an example in section 2.1.2. We would like to incorporate such hard constraints into our stabilization methodology.

It is our goal to provide an alternative impulsive stabilization approach based on center manifold theory. Our novel method does not require global Lipschitzian constraints and provides an algorithmic way to find an impulsive controller that achieves stabilization while simultaneously guaranteeing a prescribed *local convergence rate* and *minimizing a cost functional*. The jump functionals that lead to stabilization can be chosen from a set of admissible functionals that can be set by the control designer, thereby incorporating a wide class of *hard constraints* as described in the previous paragraph. Moreover, bifurcations are *suppressed* in the sense that any nonlinear structures such as periodic orbits that could result from parameter variation will be unstable. In other words, the dynamics near the equilibrium will be robust under parameter variation.

With this in mind, our setup is as follows. We assume that at $\epsilon = 0$, the characteristic equation has some number $c > 0$ of eigenvalues on the imaginary axis, while all others have strictly negative real part. This situation corresponds to one where stability of $x^* = 0$ could potentially be either gained or lost by perturbing the parameter ϵ away from zero, and a bifurcation could therefore occur. Also, we assume $f : C \times N \to \mathbb{R}^n$ is C^2 in a neighborhood of $(0, 0)$ so that it admits a Taylor expansion of the form

$$f(\phi, \epsilon) = L_0\phi + L(\epsilon)\phi + O(||\phi||^2)$$

near $(\phi, \epsilon) = 0$. Specifically, if $Df(0, \epsilon)$ denotes the Fréchet derivative of $\phi \mapsto f(\phi, \epsilon)$ for ϵ fixed, then $L_0 = Df(0, 0)$ and $L(\epsilon) = Df(0, \epsilon) - L_0$. For simplicity, we will assume further that L_0 can be expressed in the form

$$L_0\phi = A_0\phi(0) + \sum_{k=1}^{m} C_k\phi(-r_k) + \int_{-r}^{0} C(s)\phi(s)ds$$

for $n \times n$ matrices A_0 and C_k for $j = 1, \ldots, m$, some discrete delays $r_k \in (0, r]$, and integrable $C : [-r, 0] \to \mathbb{R}^{n\times n}$. The Riesz representation theorem implies that the functional L_0 could be more general than this, but to keep things simple we will assume this form.

For each $\epsilon \in N$, we consider the problem of finding a *linear jump functional* B_ϵ such that, for the impulsive retarded functional differential equation (IRFDE)

$$\dot{x} = f(x_t, \epsilon), \qquad t \notin \frac{1}{h}\mathbb{Z}, \tag{1.5}$$

$$\Delta x = B_{\epsilon,h}(t)x_{t^-}, \qquad t \in \frac{1}{h}\mathbb{Z}, \tag{1.6}$$

the following conditions are satisfied for $|\epsilon| \le \delta$, for some positive δ.

U.1. The equilibrium $x^* = 0$ is locally asymptotically stable with local convergence rate $O(e^{-\gamma t})$.

U.2. $B_\epsilon(t)$ is optimal in the sense that it minimizes an admissible cost function.

The frequency of impulse effect, h, is chosen beforehand. The constant $\gamma > 0$ is a chosen *rate parameter*. The functional B_ϵ is typically the action of a matrix on a vector of state observations (possibly delayed) or a sum thereof. The notion of a local convergence rate and admissible cost functions will be defined later when we formalize the problem more precisely. Condition U.1 guarantees that the equilibrium is *stabilized* and any bifurcations that could lead to a loss of stability of the equilibrium are *suppressed* in the parameter regime $|\epsilon| \le \delta$, while specifying a worst-case convergence

rate. The second condition, U.2, ensures that the impulsive control (1.6) is one that minimizes an associated cost.

One might also be interested in conditions under which one can find a jump functional B_ϵ that satisfies conditions U.1 and U.2 but is *independent* of ϵ. This situation corresponds to a uniform robust cost-effective stabilization and may be desirable when the dimension of the parameter space is very high. We will consider this problem as well.

The structure of the paper is as follows. In section 2, we precisely formulate our cost-effective impulsive stabilization and bifurcation suppression problem. The existence of solutions of the problem are considered in section 3, while sections 4 and 5 are devoted to the computation of optimal solutions. The effectiveness of our stabilization method is demonstrated in section 6 by way of a numerical simulation. A discussion and conclusions follow in sections 7 and 8, respectively. All proofs are deferred to the appendix.

1.1. Notation. $\mathbb{R}^{a\times b}$ denotes the real vector space of $a \times b$ matrices with real entries. For $A \in \mathbb{R}^{a\times b}$, the notation A_{ij} denotes the entry in row i and column j.

If X is a real vector space, $Y \subset X$ is a linear subspace, and $x \in X$, we denote

$$x + Y = \{x + y : y \in Y\} \subset X$$

the affine subspace spanned by Y with translation x. If W and Z are two such affine subspaces, we define for $t \in [0, 1]$ the convex combination

$$tW + (1 - t)Z = \{tw + (1 - t)z : w \in W,\, z \in Z\}.$$

For a vector space X, we will denote $X^k = X \times X \times \cdots \times X$ the k-fold Cartesian product of X with itself, with k factors in the product. We will sometimes abuse notation and identify elements of X^k with $1 \times k$ arrays with elements in X.

For square matrix $A \in \mathbb{R}^{b\times b}$, the symbols $\rho(A)$, $\det(A)$, and $\mathrm{tr}(A)$ will respectively denote the spectral radius, determinant, and trace. We will at times suppress the parentheses and write simply ρA for the spectral radius of A.

The symbol $||\cdot||$ will be used for the norm on a relevant real vector space. When a specific choice of norm is needed (e.g., induced by an inner product), this will be stated beforehand.

$\mathcal{RCR}$ denotes the real vector space of functions $f : [-r, 0] \to \mathbb{R}^n$ that are continuous from the right and have limits on the left. In integration theory, such functions would be referred to as right-continuous and regulated. $\mathcal{RCR}$ contains the continuous functions $C = C([-r, 0], \mathbb{R}^n)$ as a proper subspace. For a function $\phi : \mathbb{R} \to \mathbb{R}^n$ and $t \in \mathbb{R}$, we define the segment $\phi_t : [-r, 0] \to \mathbb{R}^n$ by $\phi_t(\theta) = \phi(t + \theta)$. For a matrix-valued function $\Phi : \mathbb{R} \to \mathbb{R}^{n\times m}$ we use the same notation.

If X and Y are two normed vector spaces, $\mathcal{L}(X, Y)$ denotes the vector space of bounded linear maps from X to Y.

For a linear map $L : X \to Y$ between finite-dimensional vector spaces, we denote L^+ its Moore–Penrose pseudoinverse. If L is a matrix, then L^+ is its pseudoinverse. In some proofs we might use the symbol L^+ for other generalized inverses, but this will always be done with adequate warning.

If $j \in \mathbb{Z}$ and $k \in \mathbb{N}^+$, we denote $[j]_k$ the remainder of j modulo k. Specifically, if we uniquely write $j = pk + r$ for some $r \in \{0, \ldots, k - 1\}$ and $p \in \mathbb{Z}$, then we define $[j]_k = r$.

For a finite sequence of matrices $A_0, \dots, A_{k-1}$ for $k \geq 1$, we define the product $\prod_{j=0}^{k-1} A_j$ by iterative composition—that is, multiplication on the left,

$$\prod_{j=0}^{k-1} A_j = A_{k-1} \cdots A_0.$$

For a function $f : X \to Y$, we will use the symbol $f(X)$ for the image of f. For an element $y \in f(X)$, we denote $f^{-1}(y)$ its preimage:

$$f^{-1}(y) = \{x \in X : f(x) = y\}.$$

When f is one-to-one, the singleton $f^{-1}(y)$ will be identified with the unique solution x of the equation $f(x) = y$.

For a complex vector $v \in \mathbb{C}^n$, we denote $\mathrm{Re}(v)$ and $\mathrm{Im}(v)$ its real and imaginary parts, repsectively. That is, $\mathrm{Re}(v)$ and $\mathrm{Im}(v)$ are the unique elements of $\mathbb{R}^n$ such that $v = \mathrm{Re}(v) + i\mathrm{Im}(v)$.

2. Precise problem formulation. In this section we formulate the problem outlined in section 1 more precisely. We introduce the linear jump functionals and frequencies that will be considered in section 2.1. The allowable cost functionals are introduced in section 2.2.

2.1. The space $\mathcal{B}$ of discrete-delay jump functionals. We consider classes of linear jump functionals B defined by a collection of distinct discrete delays $\delta_1, \dots, \delta_\ell \in [-d, 0]$ for some $d > 0$:

$$B\phi = A\phi(0) + \sum_{i=1}^{\ell} A_i \phi(\delta_i), \tag{2.1}$$

Note that given the frequency h of impulse effect, it will need to be assumed that any nonzero delay $\delta_i \neq 0$ satisfies $\delta_i \notin \frac{1}{h}\mathbb{Z}$. This constraint is needed because of a technical assumption (the overlap conditon) of the center manifold theory and reduction principle for impulsive delay differential equations, of which our main results are based. From the right-hand side of (2.1), we can identify a jump functional with an element of the finite-dimensional vector space

$$\mathcal{B} = (\mathbb{R}^{n\times n})^{\ell+1}. \tag{2.2}$$

With this identification, we will abuse notation and write $B\phi$ for $B = (A, A_1, \dots, A_\ell) \in \mathcal{B}$ and $\phi \in \mathcal{RCR}$ as the right-hand side of (2.1). Similarly, if $\Phi = [\ \phi_1 \ \cdots \ \phi_c\ ] \in \mathcal{RCR}^c$, then we denote $B\Phi = [\ B\phi_1 \ \cdots \ B\phi_c\ ]$.

2.1.1. Pinning stabilization and diagonal jump functionals. In some situations, it may not be appropriate to work with the entire space $\mathcal{B}$ of jump functionals. For example, suppose the system (1.1) has the structure of a network of N identical linearly coupled nodes

$$\dot{x}^{(i)} = f\left(x_t^{(i)}\right) + c\sum_{j=1}^{N} a_{ij}\Gamma H\left(x^{(j)}(t)\right), \qquad i = 1, \dots, N, \qquad x^{(i)} \in \mathbb{R}^n, \tag{2.3}$$

with coupling strength $c > 0$, $\Gamma = \mathrm{diag}(\gamma_1, \dots, \gamma_n) > 0$ an inner coupling matrix, $H : \mathbb{R}^n \to \mathbb{R}^n$ a nonlinear coupling with $H(0) = 0$, and $A = (a_{ij}) \in \mathbb{R}^{N\times N}$ a diffusive Laplacian matrix representing the coupling configuration of the network, where any of these coupling terms may depend on parameters. In impulsive pinning stabilization and synchronization, one would choose jump functionals that act only on individual nodes—see the references [9, 16, 23]—and are functionally driven by an error system.

For simplicity, assume we wish to stabilize the trivial equilibrium $x^* = 0$ so that one does not need to consider a separate error system. An appropriate subspace of $\mathcal{B}$ in which one could consider pinning stabilization is the set of block diagonal operators. To define these, we note that for system (2.3), we have $\mathcal{B} = (\mathbb{R}^{nN\times nN})^{\ell+1}$. Each element of $\mathbb{R}^{nN\times nN}$ can be interpreted as an $N \times N$ block matrix with $n \times n$ blocks, so that we can write an arbitrary element of $\mathcal{B}$ in the form $B = (A, A_1, \dots, A_\ell)$ with $A_m(i,j) \in \mathbb{R}^{n\times n}$ for $m = \emptyset, 1, \dots, \ell$ and $i, j = 1, \dots, N$. Then, the diagonal subspace $\mathcal{B}_{\text{diag}} \subset \mathcal{B}$ is defined by

$$\mathcal{B}_{\text{diag}} = \{B = (A, A_1, \dots, A_\ell) \in \mathcal{B} : A_m(i,j) = 0 \ \ \forall i \neq j, \ \ m = \emptyset, 1, \dots, \ell\}. \tag{2.4}$$

This is indeed a subspace of $\mathcal{B}$ and it contains only the linear jump functionals that induce no further coupling between different nodes.

2.1.2. Proportional control. Suppose one wishes to stabilize the system $\dot{x} = f(x_t, \epsilon)$ using a proportional impulsive control, with a system output y. That is, we seek jump functionals $B_\epsilon(t)$ such that $0 \in \mathbb{R}^n$ of the system

$$\begin{aligned} \dot{x} &= f(x_t, \epsilon), && t \notin \frac{1}{h}\mathbb{Z}, \\ y(t) &= Hx(t), \\ \Delta x &= NB_\epsilon(t) y_{t^-}, && t \in \frac{1}{h}\mathbb{Z}, \end{aligned}$$

becomes stable, where $N \in \mathbb{R}^{n\times m}$ is an input matrix, $H \in \mathbb{R}^{p\times n}$ is an output matrix, and $y \in \mathbb{R}^p$ is the output. Starting with the space $\mathcal{B} = (\mathbb{R}^{(n+p)\times(n+p)})^{\ell+1}$ as introduced at the beginning of this section, our porportional control constraint on the jump functionals can be imposed by abusing notation and identifying $B_\epsilon(t)$ with an element of $\mathcal{B}$ defined in block form as

$$B_\epsilon(t) = \begin{bmatrix} 0 & NB_\epsilon(t) \\ 0 & HNB_\epsilon(t) \end{bmatrix}. \tag{2.5}$$

We need to verify that $z = \begin{bmatrix} x & y \end{bmatrix}^T$ satisfies the equation $\Delta z = B_\epsilon(t) z_{t^-}$ at times $t \in \frac{1}{h}\mathbb{Z}$:

$$\Delta z = \begin{bmatrix} \Delta x \\ H\Delta x \end{bmatrix} = \begin{bmatrix} NB_\epsilon(t) y_{t^-} \\ HNB_\epsilon(t) y_{t^-} \end{bmatrix} = \begin{bmatrix} 0 & NB_\epsilon(t) \\ 0 & HNB_\epsilon(t) \end{bmatrix} \begin{bmatrix} x_{t^-} \\ y_{t^-} \end{bmatrix} = B_\epsilon(t) z_{t^-},$$

as desired. The set of elements of $\mathcal{B}$ of the form (2.5) is indeed a proper subspace of $\mathcal{B}$, which we denote $\mathcal{B}_{\text{prop}}$. One can complete the transformation to a system of form (1.5)–(1.6) by taking the derivative of y. PID controls can be introduced in a similar way.

2.1.3. Cycles of jump functionals. In our IRFDE (1.5)–(1.6), we have left the possibility for the jump function to depend on time t. That is, $B_\epsilon(t) \in \mathcal{U}$ for each t and some subspace $\mathcal{U} \subseteq \mathcal{B}$, while being generally nonconstant. As our method is based on center manifold theory, working with infinite time horizons is difficult so we will typically take $t \mapsto B_\epsilon(t)$ to be periodic. As such, a *cycle of jump functionals* is an element of the product $\mathcal{U}^k$ for natural number $k \geq 1$ called the *period*. We can then associate a jump functional in the style of (2.1) by way of the following equivalence. If $B = (B_0, \dots, B_{k-1}) \in \mathcal{U}^k$, we define

$$B\left(\frac{j}{h}\right)\phi = B_{[j]_k}\phi. \tag{2.6}$$

This definition is sufficient to give meaning to (1.6) since we need only define $B(t)$ at the times $t = \frac{j}{h}$ for $j \in \mathbb{Z}$.

2.1.4. Cycles of jump functionals in nonidentical subspaces. In some applications, it might be that certain controls (quantified by jump functionals) can only be applied itermittently due to resource limitations. As such, it is worth considering the case where more generally we have $B_\epsilon(t) \in \mathcal{U}(t)$ for $t \in \frac{1}{h}\mathbb{Z}$ and each subspace $\mathcal{U}(t)$ is generally distinct. As in section 2.1.3 we will assume the periodicity condition

$$\mathcal{U}_{j+k} := \mathcal{U}\left(\frac{j+k}{h}\right) = \mathcal{U}\left(\frac{j}{h}\right) =: \mathcal{U}_j$$

for all $j \in \mathbb{Z}$ and associate $t \mapsto B_\epsilon(t)$ to a jump functional $B = (B_0, \dots, B_{k-1})$ in the product space

$$\mathcal{U}^{(k)} := \mathcal{U}_0 \times \cdots \times \mathcal{U}_{k-1}.$$

The construction in section 2.1.3 corresponds to the special case where $\mathcal{U}_i = \mathcal{U}_0$ for all indices i. Regardless, it remains true that $\mathcal{U}^{(k)}$ is a linear subspace of $\mathcal{B}^k$.

2.2. Allowable cost functionals. An *allowable cost functional* will be a functional $\mathcal{C} : \mathcal{U} \to \mathbb{R}^+$ satisfying the following properties:

C.1. $\mathcal{C}$ is continuous, convex, and positive-definite.

C.2. $\mathcal{C}$ is *radially unbounded*: if $B_n \in \mathcal{B}$ is a sequence with unbounded norm $||B_n|| \to \infty$, then $\mathcal{C}(B_n) \to \infty$.

These properties will eventually be used to show that a formalized version of the problem from section 1 admits a solution satisfying the conditions U.1 and U.2. One can similarly define allowable cost functions on any linear subspace $\mathcal{U} \subseteq \mathcal{B}$ using the same definition.

Given an allowable cost functional $\mathcal{C}$, we define the cost of a cycle of jump functionals of period k in a linear subspace of $\mathcal{B}^k$ as follows. For $B = (B_0, \dots, B_{k-1}) \in \mathcal{B}^k$ we set

$$\mathcal{C}(B) = \sum_{j=0}^{k-1} \mathcal{C}(B_j). \tag{2.7}$$

We will somtimes abuse notation and write $\mathcal{C} : \mathcal{B}^k \to \mathbb{R}^+$ for the associated cost functional on the cycles of jump functionals of period k. $\mathcal{B}^k$ can be replaced with any linear subspace thereof, so this definition extends naturally to encompass cycles of jump functionals in nonidentical subspaces as in section 2.1.4. Finally, the following definition will be useful later.

DEFINITION 2.1. *An allowable cost function $\mathcal{C} : \mathcal{U} \to \mathbb{R}^+$ is* projective *if there exists an inner product $\langle \cdot, \cdot \rangle$ such that $\mathcal{C}(B) = \langle B, B \rangle$.*

2.2.1. Weighted matrix norms. A typical allowable cost functional can be constructed through the introduction of a weighted matrix norm $||X||_W = ||W^{\frac{1}{2}} X W^{-\frac{1}{2}}||_2$ for a symmetric positive-definite matrix W and its principal real square root $W^{\frac{1}{2}}$, and $||\cdot||_2$ the spectral norm. Indeed, let $W_0, W_1, \dots, W_\ell$ be positive-definite matrices, let $w_0, w_1, \dots, w_\ell \in \mathbb{R}^+$ be weight constants, and define a cost function

$$\mathcal{C}((A, A_1, \dots, A_\ell)) = w_0||A||_{W_0} + \sum_{i=1}^{\ell} w_i||A_i||_{W_i}. \tag{2.8}$$

The weight matrices W_i take into account limitations and costs associated to accessing and/or modifying the states of the system by impulses. The weights $w_0, w_1, \ldots, w_\ell$ allow for a weighting of the individual factors that define the control (1.6) relative to each other.

2.3. Problem statement. Having introduced the allowable cost functionals and the space $\mathcal{B}$ of linear jump functionals, we can more precisely formulate our problem. First, we define *local convergence rates*.

DEFINITION 2.2. *Let $g : \mathbb{R}^+ \to \mathbb{R}^+$ be a function satisfying $\lim_{t\to\infty} g(t) = 0$. The asymptotic $O(g(t))$ is a* local convergence rate *of an equilibrium point x^* if there exists a neighborhood U of x^* and a constant $K > 0$ such for all $(s, \phi) \in \mathbb{R} \times U$, the solution $t \mapsto x(t; s, \phi)$ of (1.5)–(1.6) satisfying the initial condition $x_s(\cdot; s, \phi) = \phi$ satisfies the inequality $||x(t; s, \phi) - x^*|| \leq Kg(t - s)$ for all $t \geq s$.*

Next, we formalize the standing hypothesis that at parameter $\epsilon = 0$, our system is at a bifurcation point where stability could be either gained or lost by parameter variation.

SPECTRAL GAP CONDITION. *At parameter $\epsilon = 0$, the characteristic equation (1.4) has $c > 0$ eigenvalues with zero real part, and all other eigenvalues have real part less than some $\sigma < 0$. The real number σ is the* spectral gap.

Remark 2.3. If the characteristic equation has a candidate bifurcation point at a parameter $\epsilon^* \neq 0$, one can perform a change of variables to shift the bifurcation point to the origin, $\epsilon = 0$. As such, no generality is lost by assuming a bifurcation point at $\epsilon = 0$.

From this point onward, we assume the spectral gap condition. With this definition at hand, the problem whose feasibility we will study and subsequently solve is the following.

PROBLEM A. *Let $\mathcal{U}$ be a linear subspace of $\mathcal{B}^k$ for $k \geq 1$. For a given rate parameter $\gamma > 0$ and freqency h, determine whether one can, for ϵ sufficiently small, find an allowable cost functional $B_\epsilon \in \mathcal{U}$ such that the optimality condition*

$$\arg\min_{Y \in \mathcal{Y}(\epsilon, h; \gamma)} \mathcal{C}(Y) = B_\epsilon \tag{2.9}$$

is satisfied, where $\mathcal{Y}(\epsilon, h; \gamma) \subset \mathcal{U}$ is the set of all linear jump functionals for which $O(e^{-\gamma t})$ is a local convergence rate of the equilibrium point $x^* = 0$ of (1.5)–(1.6) for parameter ϵ and frequency h.

In the formulation of Problem A, we allow the optimal jump functional to depend on the parameter ϵ. However, in some settings it may be desirable to have a single jump functional provide stabilization robustly for *all* ϵ sufficiently small, or it may be compuitationally too expensive to generate optimal jump functionals for a large sample of parameters. Therefore, the following problem is of interest.

PROBLEM B. *Let $\mathcal{U}$ be a linear subspace of $\mathcal{B}^k$ for some $k \geq 1$. For a given rate parameter $\gamma > 0$ and frequency h, find $\eta > 0$ and single allowable cost functional $B \in \mathcal{U}$ such that the optimality condition*

$$\arg\min_{Y \in \mathcal{Y}^\eta(h; \gamma)} \mathcal{C}(Y) = B \tag{2.10}$$

is satisfied, where $\mathcal{Y}^\eta(h; \gamma) \subset \mathcal{U}$ is the set of all linear jump functionals for which $O(e^{-\gamma t})$ is a local convergence rate of the equilibrium point $x^* = 0$ of (1.5)–(1.6) for parameters $|\epsilon| \leq \eta$ and frequency h.

3. Existence of an optimal solution. In this section we state our solutions to Problems A and B. The main results are provided in sections 3.1 and 3.2. All proofs are deferred to the proofs appendix. Illustrative applications of our main results of this section as they apply to stabilization and bifurcation suppression are postponed until section 6.

3.1. Main existence results: Problem A. A first step toward the solution of Problem A is provided by the following proposition.

PROPOSITION 3.1. *Write $f(\phi,\epsilon)$ as a Taylor expansion near 0, so that*

$$f(\phi,\epsilon)=\underbrace{A_0\phi(0)+\sum_{k=1}^{m}C_k\phi(-r_k)+\int_{-r}^{0}C(s)\phi(s)ds}_{L_0}+L(\epsilon)\phi+O(||\phi||^2)$$

for $L:\mathbb{R}^p\to\mathcal{L}(C([-r,0],\mathbb{R}^n),\mathbb{R}^n)$ continuous and satisfying $L(0)=0$ for some discrete delays $r_k\in(0,r]$ and a matrix $A_0\in\mathbb{R}^{n\times n}$. Let $\Phi(t)$ be a real $n\times c$ matrix whose columns form a basis of the set of center generalized eigenfunctions

$$E_0=\bigcup_{n\in\mathbb{N}}\left\{z(t)=\sum_{i=1}^{n}t^{i-1}e^{\lambda t}v_i,\ :\ Re(\lambda)=0,\ \dot{z}=L_0z_t\right\},\tag{3.1}$$

and compute the $c\times c$ matrix Λ satisfying the identity $\frac{d}{dt}\Phi(t)=\Phi(t)\Lambda$. Introduce the transposed operator L_0^T,

$$L_0^T\psi=A_0\psi(0)+\sum_{k=1}^{j}\psi(r_k)C_k+\int_{-r}^{0}\psi(-s)C(s)ds,\tag{3.2}$$

acting on $C([0,r],\mathbb{R}^{n*})$. Let $\Psi(t)$ be a real $c\times n$ matrix whose rows form a basis for the set of adjoint center generalized eigenfunctions

$$E_0^T=\bigcup_{n\in\mathbb{N}}\left\{w(t)=\sum_{i=1}^{n}t^{i-1}e^{-\lambda t}v_i,\ :\ Re(\lambda)=0,\ \dot{w}=-L_0^Tw^t\right\},\tag{3.3}$$

and define the invertible matrix $\Gamma\in\mathbb{R}^{c\times c}$ by the equation

$$\begin{aligned}\Gamma^{-1}=\Psi(0)\Phi(0)-\sum_{k=1}^{j}\int_{0}^{r_k}\Psi(s)C_k\Phi(s-r_k)ds\quad\cdots\\-\int_{-r}^{0}\int_{-\theta}^{0}\Psi(s)C(s)\Phi(s+\theta)dsd\theta.\end{aligned}\tag{3.4}$$

Introduce the center monodromy map $\mathcal{M}_{\epsilon,h}:\mathcal{B}^k\to\mathbb{R}^{c\times c}$,

$$\mathcal{M}_{\epsilon,h}(B)=\prod_{j=0}^{k-1}\left(I_{c\times c}+\Gamma\Psi(0)B_j\Phi_0\right)\exp\left(\frac{1}{h}\left(\Lambda+\Gamma\Psi(0)L(\epsilon)\Phi_0\right)\right),\tag{3.5}$$

and define the set

$$\widetilde{\mathcal{Y}}(\epsilon,h;\gamma)=\{B\in\mathcal{U}\subseteq\mathcal{B}^k:e^{\gamma/h}\rho(\mathcal{M}_\epsilon(B))\le 1\}.\tag{3.6}$$

Let an allowable cost functional $\mathcal{C}$ be given. Let h be fixed. The following are true.

1. $\widetilde{\mathcal{Y}}(\epsilon, h; \gamma)$ *is closed and if the spectral gap satisfies* $\gamma < \sigma$, *there exists* $\delta = \delta(\gamma) > 0$ *such that for* $|\epsilon| \leq \delta$,

$$B_\delta(0) \cap \widetilde{\mathcal{Y}}(\epsilon, h; \gamma) = B_\delta(0) \cap \mathcal{Y}(\epsilon; \gamma). \tag{3.7}$$

2. *For each each* $\epsilon \in N$, *there exists* B_ϵ *such that* $\mathcal{C} : \widetilde{\mathcal{Y}}(\epsilon; \gamma) \to \mathbb{R}$ *attains its minimum at* B_ϵ, *assuming the domain is nonempty.*

Remark 3.2. To compute a basis for E_0, it is sufficient to compute a canonical system of Jordan chains for $\Delta(\lambda; 0)$ for all eigenvalues λ with zero real part. See Theorem 4.2, Chapter 7, in [8] for the relevant result. To compute a basis for E_0^T, one can use the connection between the transpose system $\dot{w} = -L_0^T w^t$ and the adjoint, with the result being that a basis can be computed using a canonical system of Jordan chains for the transpose $\Delta(\lambda; 0)^T$—see Theorem 5.1, Chapter 7, in [8] for the relevant theorem. Alternatively, one could exploit the characterization $\Psi(t) = e^{-\Lambda t}\Psi(0)$ for $\Psi(0) \in \mathbb{R}^{c \times n}$ and solve for the unknown coefficients of $\Psi(0)$ by imposing the equality $\Gamma = I$ in (3.4), as suggested in [1].

Remark 3.3. In the case of a Hopf bifurcation, the basis calculation is much simpler. To obtain the basis matrix for E_0, one calculates a nontrivial $v \in \mathbb{C}^n$ satisfying $\Delta(i\omega; 0)v = 0$ for the critical eigenvalue $\lambda = i\omega$. Then, the basis matrix is

$$\Phi(t) = \begin{bmatrix} \operatorname{Re}(ve^{i\omega t}) & \operatorname{Im}(ve^{i\omega t}) \end{bmatrix}.$$

For the transpose basis E_0^T, one computes a nontrivial $w \in \mathbb{C}^n$ satisfying $w^T\Delta(i\omega; 0) = 0$ and obtains the basis matrix

$$\Psi(t) = \begin{bmatrix} \operatorname{Re}(w^T e^{-i\omega t}) \\ \operatorname{Im}(w^T e^{-i\omega t}) \end{bmatrix}.$$

One must then compute Γ explicitly after.

Thus, the existence of a small solution to Problem A is equivalent to the set $\widetilde{\mathcal{Y}}(\epsilon, h; \gamma)$ being nonempty. It is therefore important that we determine conditions under which that set is nonempty. Also, it would be prudent to ensure that as $\gamma \to 0$, the cost of ensuring the convergence rate $O(e^{-\gamma t})$ for a fixed pair (ϵ) should become arbitrarily small and that the intersections (3.7) are nonempty. Our sufficient condition is the following.

THEOREM 3.4. *Let* $\mathcal{U} = \mathcal{B}^k$. $\widetilde{\mathcal{Y}}(\epsilon, h; \gamma)$ *is nonempty provided* $\Phi(0)$ *and* $\Psi(0)$ *are of rank* c. *If this is the case, for any selection* $\gamma \mapsto B_\epsilon^\gamma$ *of minimizing jump functionals for rate parameter* γ, *one has* $\lim_{\gamma \to 0^+} B_\epsilon^\gamma = 0$. *In particular, if* $\gamma > 0$ *is sufficiently small, the sets in* (3.7) *are nonempty.*

The subspace and rank condition of Theorem 3.4 allows us to quickly exclude memoryless systems and systems at fold bifurcation points. We have the following corollary.

COROLLARY 3.5. *If* $c = 1$ *or* $L_0\phi = C\phi(0)$ *for an* $n \times n$ *matrix* C, *then* $\widetilde{\mathcal{Y}}(\epsilon; \gamma)$ *is nonempty.*

We can derive a more general sufficient condition that is implied by the subspace and rank condition of Theorem 3.4. It is captured by the following corollary, whose proof is omitted since it is similar to that of the previous theorem.

COROLLARY 3.6. *The subspace condition* $\mathcal{U} = \mathcal{B}^k$ *and the rank condition on* $\Phi(0)$ *and* $\Psi(0)$ *in Theorem* 3.4 *can be replaced with the condition* $\mathcal{M}_\epsilon(\mathcal{U}) = \mathbb{R}^{c \times c}$, *and the conclusions of the theorem hold.*

3.2. Main existence results: Problem B. As it turns out, the proofs of Proposition 3.1 and Theorem 3.4 work with minimal modifications to solve Problem B. We have the following analogues and the appropriate variant of Corollary 3.5.

PROPOSITION 3.7. *With the same notation as in Proposition* 3.1*, define the set*

$$\widetilde{\mathcal{Y}}^{\eta}(h;\gamma) = \{B \in \mathcal{U} \subseteq \mathcal{B}^k : \forall |\epsilon| \leq \eta,\ e^{\gamma/h}\rho(\mathcal{M}_{\epsilon,h}(B)) \leq 1\}. \tag{3.8}$$

Let h be fixed. The following are true:

1. *$\widetilde{\mathcal{Y}}^{\eta}(h;\gamma)$ is closed and if the spectral gap satisfies $\gamma < \sigma$, there exists $\delta > 0$ such that for $\eta \leq \delta$,*

$$B_{\delta}(0) \cap \widetilde{\mathcal{Y}}^{\eta}(h;\gamma) = B_{\delta}(0) \cap \mathcal{Y}^{\eta}(h;\gamma). \tag{3.9}$$

2. *$\widetilde{\mathcal{Y}}^{\eta}(h;\gamma)$ can be written as the intersection*

$$\widetilde{\mathcal{Y}}^{\eta}(h;\gamma) = \bigcap_{|\epsilon| \leq \eta} \widetilde{\mathcal{Y}}(\epsilon;\gamma). \tag{3.10}$$

3. *There exists B such that $\mathcal{C} : \widetilde{\mathcal{Y}}^{\eta}(h;\gamma) \to \mathbb{R}$ attains its minimum at B, assuming the domain is nonempty.*

THEOREM 3.8. *Let $\mathcal{U} = \mathcal{B}^k$. $\widetilde{\mathcal{Y}}^{\eta}(h;\gamma)$ is nonempty provided $\Phi(0)$ and $\Psi(0)$ are of rank c. If this is the case, then for any selection $\gamma \mapsto B^{\gamma}$ of minimizing jump functionals for rate parameter γ, one has $\lim_{\gamma \to 0^+} B^{\gamma} = 0$. In particular, if $\gamma > 0$ is sufficiently small, the sets in* (3.9) *are nonempty.*

COROLLARY 3.9. *If $c = 1$ or $L_0\phi = C\phi(0)$ for an $n \times n$ matrix C, then $\widetilde{\mathcal{Y}}^{\eta}(h;\gamma)$ is nonempty.*

COROLLARY 3.10. *The subspace condition $\mathcal{U} = \mathcal{B}^k$ and the rank condition on $\Phi(0)$ and $\Psi(0)$ in Theorem* 3.8 *can be replaced with the condition $\mathcal{M}_{\epsilon}(\mathcal{U}) = \mathbb{R}^{c\times c}$ for $|\epsilon| \leq \eta$, and the conclusions of the theorem hold.*

Remark 3.11. Since $\epsilon \mapsto \mathcal{M}_{\epsilon}$ is continuous, the conclusion of the above corollary is guaranteed to hold for *some* $\eta > 0$, provided $\mathcal{M}_{0,h}(\mathcal{U}) = \mathbb{R}^{c\times c}$. This can be seen by vectorizing the monodromy map and recalling that rank function $X \mapsto \operatorname{rank}(X)$ is lower semicontinuous.

4. Computation of optimal solutions for Problem A: The center probe method. Solving the optimization problem

$$(Y) \qquad \begin{aligned} \text{minimize} \quad & \mathcal{C}(Y), \\ \text{subject to} \quad & Y \in \widetilde{\mathcal{Y}}(\epsilon;\gamma), \end{aligned}$$

directly appears to be very difficult. The feasible set, introduced in Proposition 3.1, is characterized as a sublevel set of $B \mapsto \rho(\mathcal{M}_{\epsilon,h}(B))$. While $B \mapsto \mathcal{M}_{\epsilon,h}(B)$ is smooth, the spectral radius is notably irregular and nonconvex. If one wishes to guarantee feasibility it is generally necessary to optimize in the space $\mathcal{B}^k$, whose dimension is $kn^2(\ell+1)$ and can be quite large even for small networks. It is therefore critical that we reduce the dimension of the problem before even considering performing an optimization task.

From this point onward, we will assume that a rate parameter γ, system parameter ϵ, and frequency h have been chosen. We will suppress all dependence on these variables unless necessary. Also, we make the following simplifying assumption on our chosen space of cycles of jump functionals of period $k \geq 1$.

Assumption. The subspace $\mathcal{U}^{(k)} \subseteq \mathcal{B}^k$ of cycles of jump functionals is given by k copies of a single subspace $\mathcal{U} \subseteq \mathcal{B}$. That is, $\mathcal{U}^{(k)} = \mathcal{U}^k$.

This assumption is not strictly needed. All constructions (e.g., probe space) can be appropriately generalized to allow the individual subspaces making up the product $\mathcal{U}^{(k)}$ to be distinct, and the major theorems (Theorems 4.11 and 4.12) have appropriate and similarly strong analogues. However, the notation can make the presentation difficult to follow. For this reason, we will specialize to this particular case.

To avoid ambiguity later, we define $\mathcal{M}_1 : \mathcal{U} \to \mathbb{R}^{c\times c}$ by

$$\mathcal{M}_1(B) = (I_{c\times c} + \Gamma\Psi(0)B\Phi_0) \exp\left(\frac{1}{h}\left(\Lambda + \Gamma\Psi(0)L(\epsilon)\Phi_0\right)\right)$$

and reserve the symbol $\mathcal{M}$ for the center monodromy operator $\mathcal{M} : \mathcal{U}^k \to \mathbb{R}^{c\times c}$. Note that this implies the factorization $\mathcal{M}(B) = \prod_{j=0}^{k-1} \mathcal{M}_1(B_j)$.

4.1. The probe space $\mathcal{P}$. The map $\mathcal{M}_1 : \mathcal{U} \to \mathbb{R}^{c\times c}$ is affine, and we can decompose it as

$$\mathcal{M}_1 = M_0 + Z, \quad Z = \exp\left(\frac{1}{h}\left(\Lambda + \Gamma\Psi(0)L(\epsilon)\Phi_0\right)\right), \quad M_0(B) = \Gamma\Psi(0)B\Phi_0 Z,$$

where $M_0 : \mathcal{U} \to \mathbb{R}^{c\times c}$ is linear.

DEFINITION 4.1. *The* probe space $\mathcal{P} \subset \mathbb{R}^{c\times c}$ *is the image of* $\mathcal{M}_1$. *That is,* $\mathcal{P} = \mathrm{im}(\mathcal{M}_1)$. *Elements* $P \in \mathcal{P}$ *are called* probe elements. *The* k-probe space *is the* k*th Cartesian power* $\mathcal{P}^k$, *and its elements are* k-probe elements.

The following lemmas describe the geometry of the probe space, k-probe space, and preimages of probe elements under the center monodromy operator. The proofs are omitted.

LEMMA 4.2. *The probe space is an affine space of dimension at most* c^2. *It can be written in the form* $\mathcal{P} = Z + \mathrm{im}(M_0)$.

Remark 4.3. If the subspace/rank condition of Theorem 3.4/Theorem 3.8 is satisfied, then $\mathcal{P}$ is a vector subspace of $\mathbb{R}^{c\times c}$ and it is precisely $\mathcal{P} = im(M_0)$.

LEMMA 4.4. *The* k*-probe space* $\mathcal{P}^k$ *is convex.*

LEMMA 4.5. *Let* $P \in \mathcal{P}$. *The preimage of* P *under the center monodromy map* $\mathcal{M}_1$ *is an affine space and can be written*

$$\mathcal{M}_1^{-1}(P) = M_0^+(P - Z) + \ker(M_0). \tag{4.1}$$

The correspondence $\mathcal{M}_1^{-1} : \mathcal{P} \rightrightarrows \mathcal{U}$ exhibits a concavity-like property that will be essential in the next section. It is summarized by the following lemma.

LEMMA 4.6. *Let* $X, Y \in \mathcal{P}$. *For all* $t \in [0, 1]$,

$$t\mathcal{M}_1^{-1}(X) + (1-t)\mathcal{M}_1^{-1}(Y) \subseteq \mathcal{M}_1^{-1}(tX + (1-t)Y). \tag{4.2}$$

At this stage we should define the precise link between the k-probe space and the image of $\mathcal{M} : \mathcal{U}^k \to \mathbb{R}^{c\times c}$. Define the product function $G : (\mathbb{R}^{c\times c})^k \to \mathbb{R}^{c\times c}$ by

$$G(X_0, \ldots, X_{k-1}) = X_{k-1} \cdots X_0.$$

In terms of the product function, the image of $\mathcal{M} : \mathcal{U}^k \to \mathbb{R}^{c\times c}$ can be written equivalently to the image of $G : \mathcal{P}^k \to \mathbb{R}^{c\times c}$. From this equivalence, we obtain the following lemma.

LEMMA 4.7. *Let $P \in \mathcal{P}^k$. For all $B \in \mathcal{M}^{-1}(G(P))$, we have $\rho(\mathcal{M}(B)) = \rho(G(P))$.*

The above lemma provides one way to pull k-probe elements back to cycles of jump functionals of period k. Namely, compute the ordered product and take the preimage under $\mathcal{M} : \mathcal{U}^k \to \mathbb{R}^{c\times c}$. However, the preimage under $\mathcal{M} : \mathcal{U}^k \to \mathbb{R}^{c\times c}$ of a given $c \times c$ matrix is not necessarily convex. To remedy this, we can alternatively define a map $\mathcal{M}_k : \mathcal{U}^k \to (\mathbb{R}^{c\times c})^k$ by

$$\mathcal{M}_k(B_0, \dots, B_{k-1}) = (\mathcal{M}_1(B_0), \dots, \mathcal{M}_1(B_{k-1})).$$

Then, an analogue of Lemma 4.7 is as follows, strengthened by the convexity of $\mathcal{M}_k^{-1}(P)$ for any $P \in \mathcal{P}^k$ due to the componentwise convexity afforded by Lemma 4.5.

LEMMA 4.8. *Let $P \in \mathcal{P}^k$. For all $B \in \mathcal{M}_k^{-1}(P)$, we have $\rho(\mathcal{M}(B)) = \rho(G(P))$. Also, $\mathcal{M}_k^{-1}(P)$ is convex and given $P = (P_0, \dots, P_{k-1})$, it can be written*

$$\mathcal{M}_k^{-1}(P) = \{(B_0, \dots, B_{k-1}) : B_i \in \mathcal{M}_1^{-1}(P_i),\ i = 0, \dots, k-1\}.$$

4.2. Compatible probe cost. Broadly speaking, our idea for solving the nonlinear program (Y) is to first solve a related optimization problem in the probe space and obtain an optimal solution P^*, pull this optimal solution into the convex space $\mathcal{M}_k^{-1}(P^*)$, and find a minimizer Y_{P^*} of the cost functional. Since every element of $\mathcal{M}^{-1}(P^*)$ will be feasible provided $\rho(P^*) \le e^{-\gamma/h}$—see Lemma 4.7—the spectral constraint does not need to be checked at this final optimization stage. The following definition allows us to define the appropriate optimization problem in $\mathcal{P}^k$.

DEFINITION 4.9. *A continuous, radially unbounded convex function $\tilde{C} : \mathcal{P}^k \to \mathbb{R}^+$ is a*

- local probe-compatible cost (LPCC) *for the nonlinear program (Y) if one of the following conditions holds:*
  1. *$\tilde{C} \circ \mathcal{M}_k(B) \le \mathcal{C}(B)$ for all $B \in \mathcal{U}$ with equality if*
  $$B \in \operatorname*{arg\,min}_{X \in \mathcal{M}^{-1}(\mathcal{M}(B))} \mathcal{C}(X), \tag{4.3}$$
  and in this case we say $\tilde{C}$ is a type 1 LPCC*,*
  2. *$\tilde{C} \circ \mathcal{M}_k(X) \le \tilde{C} \circ \mathcal{M}_k(Y)$ implies $\mathcal{C}(X) \le \mathcal{C}(Y)$, and in this case we say $\tilde{C}$ is a* type 2 LPCC*;*
- global probe-compatible cost (GPCC) *if for any global optimum B^* of the nonlinear program (Y) and any other feasible solution B, one has $\tilde{C} \circ \mathcal{M}_k(B^*) \le \tilde{C} \circ \mathcal{M}_k(B)$, with equality holding if and only if B is also a global optimum;*
- uniform probe-compatible cost (UPCC) *if it is both an LPCC and a GPCC.*

The existence of GPCC/LPCCs will be addressed in Theorem 4.12. Given a continuous, radially unbounded convex function $\tilde{C} : \mathcal{P}^k \to \mathbb{R}^+$, we define a new nonlinear program in the probe space, which we call the *probe program*:

$$(\mathcal{P}\tilde{C}) \qquad \begin{aligned} \text{minimize} \quad & \tilde{C}(X) \\ \text{subject to} \quad & X \in \mathcal{P}^k, \\ & \rho \circ G(X) \le e^{-\gamma/h}. \end{aligned}$$

The program $(\mathcal{P}\tilde{C})$ possesses a global optimum so long as $\widetilde{\mathcal{Y}}$ is nonempty. We also define a family of convex programs indexed by $P \in \mathcal{P}^k$ with feasible set given by the preimage $\mathcal{M}_k^{-1}(P)$. We call this the *inverse probe program*:

$$(Y_P) \qquad \begin{aligned} &\text{minimize} && \mathcal{C}(B) \\ &\text{subject to} && B \in \mathcal{M}_k^{-1}(P). \end{aligned}$$

Remark 4.10. As $\mathcal{M}_k^{-1}(P)$ is an external direct sum of k affine subspaces of $\mathcal{B}$—see Lemmas 4.5 and 4.8—the inverse probe program is actually unconstrained after an appropriate affine linear change of variables.

Our first theorem of this section relates the programs $(\mathcal{P}\tilde{C})$ and (Y_P) to solutions of (Y) under the assumption that $\tilde{C}$ is an LPCC or a GPCC. The second theorem establishes the existence of at least one UPCC.

THEOREM 4.11 (center probe method). *Assume $\widetilde{\mathcal{Y}}$ is nonempty. Let $\tilde{C} : \mathcal{P}^k \to \mathbb{R}^+$ be a global (resp., local) PCC for the nonlinear program (Y). Let $P \in \mathcal{P}^k$ be a global (resp., local) optimum for the program $(\mathcal{P}\tilde{C})$. Let $B^* \in \mathcal{M}_k^{-1}(P)$ be a local optimum for the convex program (Y_P). Then, B^* is a global (resp., local) optimum for the program (Y).*

We will refer to the procedure of determining a solution of the program (Y) using the probe program in conjunction with the inverse probe program collectively as the *center probe method* (CPM).

THEOREM 4.12. *Define $\tilde{C} : \mathcal{P}^k \to \mathbb{R}^+$ by*

$$\tilde{C}(P) = \min_{X \in \mathcal{M}_k^{-1}(P)} \mathcal{C}(X). \tag{4.4}$$

This function is continuous, radially unbounded, convex, and a UPCC. We will refer to it as the trivializing UPCC.

The trivializing UPCC is so named because it makes the final step of running the nonlinear program (Y_P) unnecessary. Indeed, once one has computed an optimal solution P for $(\mathcal{P}\tilde{C})$, the optimal cost for the program (Y) is precisely $\tilde{C}(P)$. In other words, running the program (Y_P) is equivalent to computing $\tilde{C}(P)$.

4.3. Explicit formula for trivializing UPCC under projective cost. Suppose the cost functional $\mathcal{C} : \mathcal{U} \to \mathbb{R}^+$ is projective. Recall this means that $\mathcal{C}(B) = \langle B, B\rangle$ for an inner product $\langle \cdot, \cdot \rangle$ on $\mathcal{U}$ (note that on $\mathcal{U}^k$ one defines the cost additively according to (2.7)). See section 4.3.1 for examples. Because of Lemma 4.5, we can calculate $\tilde{C}(P)$ for the trivializing UPCC according to

$$\tilde{C}(P) = \sum_{i=0}^{k-1} \min\left\{\langle X, X\rangle : X \in M_0^+(Z - P_i) + \ker(M_0)\right\}. \tag{4.5}$$

$\tilde{C}(P)$ and the associated minimizer in $\mathcal{M}^{-1}(P)$ in fact has an explicit solution, due to the computation of $\tilde{C}(P)$ being equivalent to squared distance minimization to an affine subspace of a Hilbert space. The proof is elementary and is omitted.

PROPOSITION 4.13. *Let $\{m_1, \ldots, m_b\}$ be an orthonormal basis for $\ker(M_0)$. For a given $P \in \mathcal{P}$, define $X(P) \in \mathcal{U}$ by*

$$X(P) = M_0^+(P - Z) - \sum_{j=1}^{b} \langle M_0^+(P - Z), m_j\rangle m_j. \tag{4.6}$$

If $\tilde{C}$ is the trivializing UPCC and the cost functional is projective, then for any $P = (P_0, \dots, P_{k-1}) \in \mathcal{P}^k$, if we define $\mathbf{X}(P) = (X(P_0), \dots, X(P_{k-1}))$, then

$$\tilde{C}(P) = \mathcal{C}(\mathbf{X}(P)).$$

Moreover, $M_0^+(P - Z)$ in (4.6) can be replaced by any solution Y of the linear equation $M_0 Y + Z = P$, and these are

$$Y = M_0^+(P - Z) + (I - M_0^+ M_0)w$$

for any $w \in \mathcal{U}$.

4.3.1. Example of projective cost functionals. A projective cost functional modeled on (2.8) and suitable for implementation (i.e., vectorized) is

$$\mathcal{C}((A, A_1, \dots, A_\ell)) = w_0 \vec{A}^T W_0 \vec{A} + \sum_{i=1}^{\ell} w_i \vec{A}_i^T W_i \vec{A}_i, \tag{4.7}$$

where $w_0, w_1, \dots, w_\ell > 0$, $W_0, W_1, \dots, W_\ell \in \mathbb{R}^{n^2 \times n^2}$ are symmetric positive-definite matrices, and for $X \in \mathbb{R}^{n \times n}$, we define $\vec{X} \in \mathbb{R}^{n^2}$ to be its standard vectorization, obtained by stacking the columns of X on top of one another from left to right. The cost (4.7) is projective with the inner product

$$\langle (A, A_1, \dots, A_\ell), (B, B_1, \dots, B_\ell) \rangle = w_0 \vec{A}^T W_0 \vec{B} + \sum_{i=1}^{\ell} w_i \vec{A}_i^T \vec{B}_i.$$

5. Computation of optimal solutions: Problem B. The CPM of section 4 can be adapted to find solutions of Problem B. Recall that our proposed solution to Problem B derived from Proposition 3.7 requires us to find a constant $\eta > 0$ such that there exists a solution of the nonlinear program

$$(Y^\eta) \qquad \begin{aligned} \text{minimize} \quad & \mathcal{C}(Y) \\ \text{subject to} \quad & Y \in \widetilde{\mathcal{Y}}^\eta(h; \gamma). \end{aligned}$$

It is important to remark that a solution of Problem B *includes* the *robustness parameter* η. That is to say, η is not chosen at the outset. This is crucial, and in general one cannot take η as arbitrary.

Suppose for the sake of argument that one could choose $\eta > 0$ at the outset. One may recall from Proposition 3.7 that a solution B^η of (Y^η) can only be guaranteed to solve the nonlinear program (2.10) if $||B^\eta|| < \delta(\eta)$; otherwise there may be secondary bifurcations involving eigenvalues crossing the imaginary axis from the left. As $\delta(\eta)$ is generally decreasing with respect to η—see the proof of the aforementioned proposition—one can only *guarantee* stabilization with an increasingly trivial linear jump functional. However, unless the equilibrium $x^* = 0$ is already exponentially stable with rate $O(e^{-\gamma t})$, a trivial jump functional should not be able to provide stabilization. Therefore, generally, it is not possible to choose the robustness parameter; a controller that stabilizes the equilibrium at the parameter $\epsilon = 0$ will generally fail to stabilize the equilibrium if the parameter is taken too large.

5.1. Uniform CPM. Because of the smoothness of the vector field (1.1), any linear jump functional that guarantees the local convergence rate $O(e^{-(\gamma+s)t})$ at parameter $\epsilon = 0$ for some $s > 0$ will automatically guarantee the local convergence rate $O(e^{-\gamma t})$ for $|\epsilon| \leq \eta$, for some $\eta > 0$ that generally depends on s. This follows from hemicontinuity arguments; see the related proof of Proposition 3.1.

THEOREM 5.1 (uniform CPM). *Let $B \in \mathcal{U}^k$ be a linear jump functional produced by the CPM (local or global) at parameter $\epsilon = 0$ and convergence rate γ' satisfying $\sigma > \gamma' = \gamma + s$ for $s > 0$ a small* safety parameter. *There exists $\eta = \eta(s) > 0$ such that B is a feasible solution of the program (Y^η).*

The above theorem does not guarantee optimality of the candidate solution. However, for typical problems where the performance (linear-order convergence rate) of the candidate B deteriorates when $|\epsilon|$ becomes large, we are guaranteed that the solution is optimal for some robustness parameter.

COROLLARY 5.2 (optimality). *With the notation from the previous theorem, suppose for some $\epsilon \neq 0$, $\rho(\mathcal{M}_\epsilon(B)) > e^{-\gamma/h}$. Then, there exists $\eta \in (0, |\epsilon|)$ such that B is an optimum (local or global) of the program (Y^η).*

6. Stabilization of a neural network near a Hopf point. In this section we consider primarily by way of example how our uniform CPM can be used to stabilize a complex network near a Hopf point. We incorporate an inhomogeneous weighting on the cost of controlling each node and address how one can minimize the number of controlled nodes while still taking advantage of the performance improvements inherent to the CPM.

The neural network we will consider in this section is a slight modification of an example considered in [16]. For $x_i \in \mathbb{R}^2$ for $i = 1, \dots, 100$, consider the nonlinear network model

$$\dot{x}_i = -x_i(t) + \epsilon\big[B\tanh(x_i(t)) + D\tanh(x_i(t-1))\big] + \sum_{j=1}^{N} a_{ij}x_j(t) \tag{6.1}$$

with connection weight matrices

$$B = \begin{bmatrix} 2 & -0.11 \\ -5 & 3.2 \end{bmatrix}, \qquad D = \begin{bmatrix} -1.6 & -0.1 \\ -0.18 & -2.4 \end{bmatrix}$$

and linear coupling determined by the matrix $A = (a_{ij})_{N\times N}$, which is the negative of a graph Laplacian associated to a small world network graph on 100 nodes. We define $\tanh(y) = [\ \tanh(y_1) \quad \tanh(y_2)\ ]^T$ for $y = (y_1, y_2) \in \mathbb{R}^2$, and $\epsilon \geq 0$ is a neural activation strength parameter. It is known [16] that when decoupled, the individual nodes determined by the dynanical system

$$\dot{y} = -y + C\tanh(y(t)) + D\tanh(y(t-1))$$

exhibit chaotic dynamics with a double-scroll-like attractor and that the origin is unstable. Thus, when $\epsilon = 1$, the diffusivity condition implies that the nonlinear network model (6.1) is chaotic and the origin is unstable. However, when $\epsilon = 0$, the origin is globally asymptotically stable. There is therefore a bifurcation at some critical activation strength $\epsilon^* \in (0, 1)$.

Our goal is to robustly stabilize the network (6.1) in a parameter neighborhood of ϵ^* with frequency $h = 1$ and period $k = 1$ using our uniform CPM. We will assume decoupled controls, so that we take our candidate jump functionals from the diagonal subspace $\mathcal{B}_{diag}$. To introduce some heterogeneity, we will assume that some nodes are more difficult to control than others, and we will incorporate this into the associated cost functional. Specifically, nodes with a higher degree of connectivity will be assigned a higher cost. We will also attempt to minimize the number of controlled nodes.

Our methodology is as follows. First, we recall the small-world network topology and introduce the cost functional we will be using in section 6.1. Then, we encode the parameter-dependent system (6.1) as the user input to DDE-BIFTOOL and use the included `GetStability` routine to determine the critical parameter ϵ^* where the bifurcation occurs, and determine its type. We also use this routine to compute the critical eigenvalus $\lambda = \pm i\omega$ on the imaginary axis. We then implement the uniform CPM at the critical parameter ϵ^* and determine a locally cost-minimizing controller B^0 and a generating probe element P^* for a target local convergence rate $O(e^{-\frac{t}{5}})$—that is, we take the target rate parameter to be $\gamma = \frac{1}{5}$. These steps are carried out in section 6.2.

The feasible jump functional B^0 is cost-minimizing in the affine space $\mathcal{M}^{-1}(P^*)$ but might not minimize the number of controlled nodes. We provide a solution to this secondary minimization problem in section 6.3.

In section 6.4 we assess the performance of the jump functional derived using the uniform CPM in conjunction with the node-minimizing process of section 6.3. Specifically, we compare the output of the neural network model with and without our pinning control at a range of parameters $\epsilon \in [\epsilon^*, \epsilon_2^*)$. The bifurcation is suppressed in this regime, but a secondary bifurcation point is identified at the parameter value $\epsilon = \epsilon_2^* < 1$. The implications of this bifurcation are discussed.

6.1. Network topology, cost functional, and the trivializing UPCC. Small-world networks [22] capture a network topology involving both a high degree of clustering and short average path-lengths. They can be constructed by starting with a ring lattice on N vertices with k edges per vertex (specifically, k nearest neighbors) and "rewiring" each edge randomly with a specified probability, p.

We wrote a script in MATLAB R2018a to generate a Watts–Strogatz small-world graph $\mathbf{G}$ on $N = 100$ vertices with parameters $k = 8$ and $p = 0.3$. The matrix A defining the linear coupling in (6.1) is then obtained by taking the negative graph laplacian: $A = -\texttt{laplacian}(\mathbf{G})$. The computed graph that was used in this example is displayed in Figure 1.

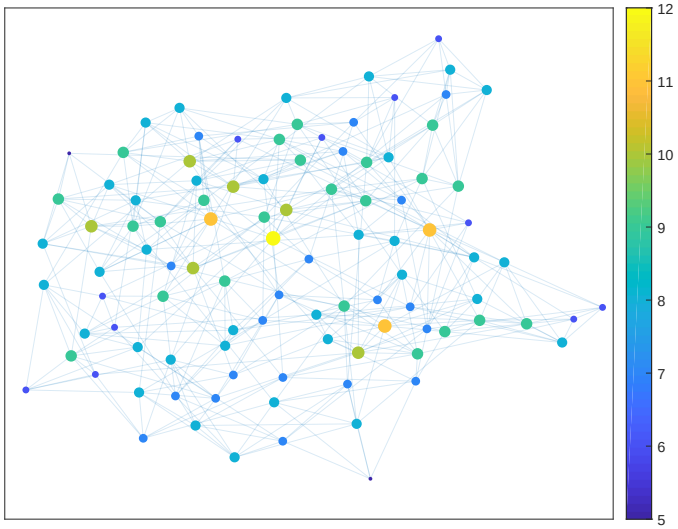


FIG. 1. *The small-world graph used in the example of section* 6. *Nodes are colored with intensity varying based on their degree, while the sizes indicate relative degree. Parameters were* $N = 100$ *nodes with initial connections to* $k = 8$ *nearest neighbors and rewiring probability* $p = 0.3$.

The degree of node i is precisely $|a_{ii}| \geq 1$. Based on this, if we uniquely write $B \in \mathcal{B}_{diag}$ as a tuple $B = (B_1, \ldots, B_{100})$ for $B_i \in \mathbb{R}^{2\times 2}$, we can define the cost functional $\mathcal{C} : \mathcal{B}_{diag} \to \mathbb{R}$,

$$\mathcal{C}(B) = \frac{1}{\Delta(\mathbf{G})} \sum_{i=1}^{100} |a_{ii}| \langle B_i, B_i \rangle_F,$$

where $\langle \cdot, \cdot \rangle_F$ is the Frobenius inner product on $\mathbb{R}^{2\times 2}$. Thus, the cost of controlling node i is linearly scaled relative to its degree, normalized with respect to the maximum degree. This cost functional is projective with inner product

$$\langle X, Y \rangle = \frac{1}{\Delta(\mathbf{G})} \sum_{i=1}^{100} |a_{ii}| \langle X_i, Y_i \rangle_F.$$

As such, the trivializing UPCC has an explicit formula from Proposition 4.13:

$$\tilde{C}(P) = \left|\left| M_0^+(Z - P) - \sum_{j=1}^{b} \langle M_0^+(Z - P), m_j \rangle m_j \right|\right|^2 \tag{6.2}$$

and $|| \cdot || = \sqrt{\langle \cdot, \cdot \rangle}$ is the norm induced by $\langle \cdot, \cdot \rangle$.

6.2. Precomputation and uniform CPM. DDE-BIFTOOL [5] was used to identify parameters where bifurcations in the neural network model (6.1) occur. The tool detected a Hopf bifurcation (center subspace of dimension $c = 2$) at parameter $\epsilon = 0.5621$. Further numerical examination revealed that the trivial equilibrium is locally stable at parameter $\epsilon \leq 0.5620$ and unstable at parameter $\epsilon \geq 0.5621$. For our purposes, we chose $\epsilon^* = 0.5621$ to be the approximate bifurcation point. The critical eigenvalues were also computed by DDE-BIFTOOL, and it was found that they are $\lambda^* = 0.0001 \pm 0.375i$. Note that the real part is positive, which is a consequence of our choosing the parameter on the unstable side of the bifurcation point. We allow ourselves to be content with this approximation.

To calculate the matrices $\Phi(t)$ and $\Psi(t)$, we numerically computed the right and left eigenvectors associated to the eigenvalue of the characteristic matrix $\Delta(\lambda^*; \epsilon^*)$ with the smallest absolute value. If ϵ^* were the true bifurcation parameter rather than a numerical approximation and λ^* was the true critical eigenvalue, we would merely compute the kernel. Then, we calculate the matrices $\Phi(t)$ and $\Psi(t)$ following Remark 3.3. The rank of both $\Phi(0)$ and $\Psi(0)$ was verified to be equal to 2. Following this we defined a three-dimensional cell array $U = \texttt{cell}(100, 2, 2)$ and populated it with basis elements for $\mathcal{B}_{diag}$ according to the assignment

$$U\{i, j, k\} = (0, \ldots, 0, E_{jk}, 0, \ldots, 0),$$

with the nonzero entry in the ith position. An ordered basis $\{V_1, \ldots, V_{400}\}$ was then defined according to the rule

$$V_{4(i-1)+2(j-1)+k} = U\{i, j, k\}, \qquad i = 1, \ldots, 100,\ j, k \in \{1, 2\}.$$

Then, relative to the aforementioned ordered basis for $\mathcal{B}_{diag}$ and the standard ordered basis $\{E_{11}, E_{21}, E_{12}, E_{22}\}$ for $\mathbb{R}^{2\times 2}$, the map $M_0 : \mathcal{B}_{diag} \to \mathbb{R}^{2\times 2}$ was vectorized as a 4×400 matrix. The result was rank 4, and it follows that our example satisfies

Corollary 3.10 with $\eta = 0$ (after a change of coordinates, treating ϵ^* as zero). Following Remark 3.11, we are guaranteed that the uniform CPM possesses at least one feasible solution.

The probe program $(\mathcal{P}\tilde{C})$ was first solved in MATLAB R2018a using the smooth constrained solver `fmincon` from the optimization toolbox. The explicit formula for the trivializing UPCC from (6.2) was used, and we supplied the solver with an explicit gradient. The solver was initialized at the infeasible matrix (vectorized)

$$P^0 = \begin{bmatrix} 100 & 100 \\ 100 & 100 \end{bmatrix}$$

and converged quickly to the feasible matrix (vectorized)

$$P^* = \begin{bmatrix} 0.7317 & 0.2672 \\ -1.4729 & 0.3783 \end{bmatrix}.$$

We then initialized the nonsmooth `patternsearch` solver at this feasible point. To standard tolerances and minimum mesh size P^* was optimal. Proposition 4.13 provides a feasible jump functional $B^0 = X(P^*)$, but as we mentioned earlier, this candidate does not minimize the number of controlled nodes.

6.3. Minimizing the number of controlled nodes. As discussed in the outline, the jump functional B^0 minimizes the cost functional, but it does not necessarily control a minimal number of nodes. To address this, we recall that because of Lemma 4.5 and Proposition 4.13, the cost function $\mathcal{C} : \mathcal{B}_{diag} \to \mathbb{R}$ is in fact constant on the hyperplane

$$\begin{aligned} H(P^*) &= \left\{ M_0^+(P^* - Z) + Qy - \sum_{j=1}^{b} \langle M_0^+(P^* - Z) + Qy, m_j \rangle m_j : y \in \mathcal{U} \right\} \\ &= \left\{ B^0 + Qy - \sum_{j=1}^{b} \langle Qy, m_j \rangle m_j : y \in \mathcal{U} \right\}, \end{aligned}$$

where $Q = I - M_0^+ M_0$. Thus, if we wish to minimize the number of controlled nodes while maintaining the minimal cost, it suffices to solve the following sequence of unconstrained nonlinear programs:

$$(N^\mu) \qquad \begin{aligned} &\text{minimize } \sum_{i=1}^{100} \tanh\left(\mu || B_i^0 + \pi_i(\mathcal{Q}y)\, ||_2\right) \\ &\text{subject to} \quad y \in \mathcal{U}, \end{aligned}$$

where we define the linear operator $\mathcal{Q} : \mathcal{U} \to \mathcal{U}$ by

$$\mathcal{Q}y = Qy - \sum_{j=1}^{b} \langle Qy, m_j \rangle m_j,$$

and $\pi_i(B_1, \ldots, B_{100}) = B_i$ is the projection onto the ith factor. The objective function of (N^μ) precisely counts the number of nodes controlled by the candidate $B^0 + \mathcal{Q}y \in \mathcal{M}^{-1}(P^*)$ in the limit $\mu \to \infty$. Consequently, if y^μ is a solution for parameter μ and $y^\mu \to y^*$, then

$$B^* = B^0 + \mathcal{Q}y^* \tag{6.3}$$

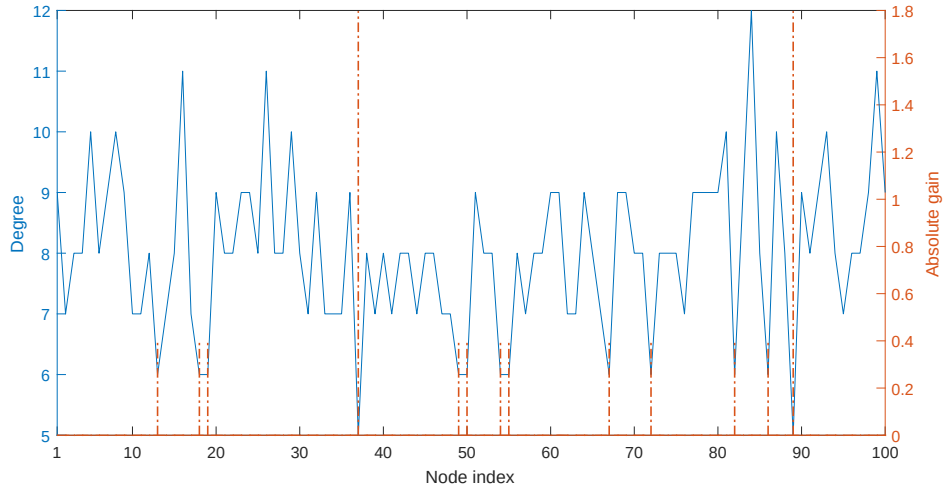


FIG. 2. *Plot of the degree (curve, solid blue) of nodes* $i = 1, \ldots, 100$ *versus the absolute gain (stems, dashed dot orange) of node* i *by the cleaned jump functional. Fourteen nodes are controlled—the small gain applied to node* 61 *is not visible under the scale of the present graph, as the associated absolute gain was* $||B^*_{61}|| = 0.01$. *Note that more expensive nodes were assigned a lower gain control, and vice versa.*

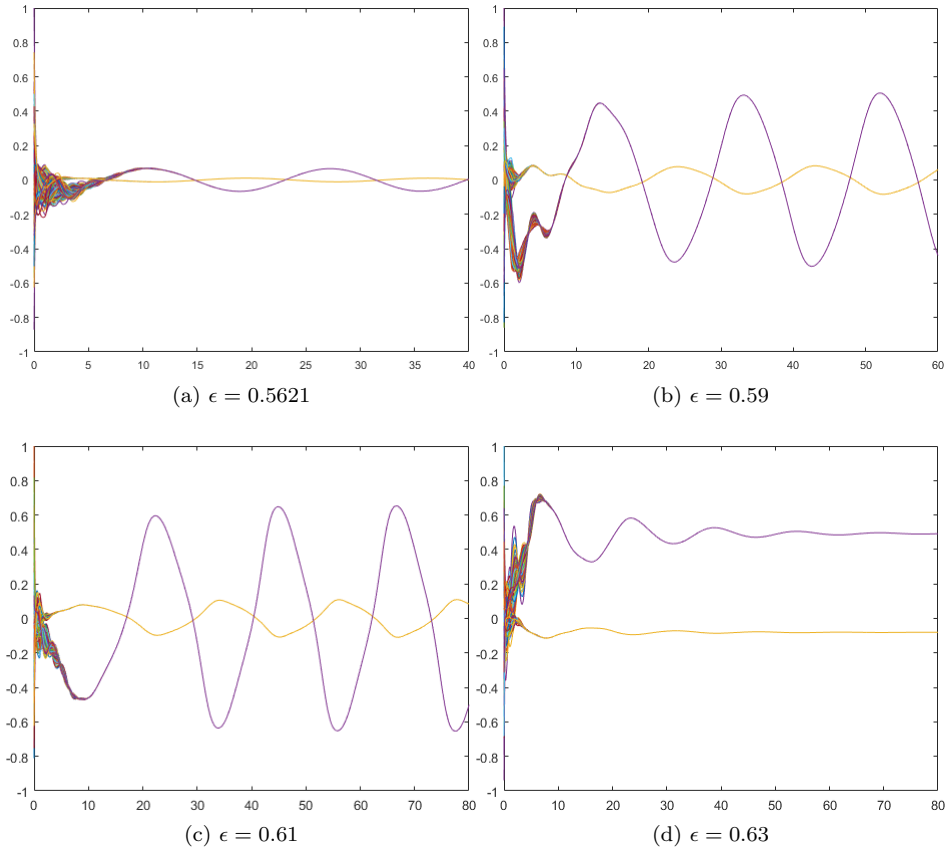
is a feasible jump functional having optimal cost that controls a minimal number of nodes.

After apropriate vectorization consistent with section 6.2, the smooth nonlinear program (N^μ) was solved using the unconstrained solver `fminunc` from MATLAB R2018a initially with parameter $\mu = 1$ and initialized at $y = 0$. Subsequently, the program was solved with increments $\mu \mapsto \mu + \frac{1}{2}$ and initialized at the previous solution. Upon reaching $\mu = 3$, the solver made step sizes smaller than 10^{-7} and we manually halted the process.

The resulting jump functional $B^* = (B^*_1, \ldots, B^*_{100})$ was cleaned by setting $B^*_i \mapsto 0$ if $||B^*_i||_F < 10^{-4}$, where $||\cdot||_F$ is the Frobenius norm. The result was a jump functional that pinned 14 nodes. Defining $||B^*_i||$ to be the absolute gain of node i, a plot of the absolute gain relative to the degree is provided in Figure 2. From this figure, it is clear that our algorithm prioritized the pinning of nodes that have a low degree. Indeed, all nodes with degrees 5 and 6 are pinned, and the only other node to be pinned was the degree 9 node at index 61.

6.4. Performance of the controller and secondary bifurcation. Plots of sample trajectories of the system without the pinning controller are given in Figure 3, while those with pinning are provided in Figure 4. Several illustrative parameters ϵ are used.

By comparing the two figures, it should be clear that our pinning control B^* derived from the uniform CPM with subsequent node minimization succeeds in exponentially stabilizing the trivial equilibrium in the activation parameter interval $[\epsilon^*, 0.61]$, while when our controller is not present, the equilibrium is unstable. Increased volatility and decreased convergence rates follow as the parameter increases, and a secondary bifurcation occurs somewhere in the parameter interval $[0.61, 0.63]$. This is consistent with our observation that the system without the pinning control appears to undergo a bifurcation in the interval $[0.61, 0.63]$; see Figure 3. We suggested



(a) $\epsilon = 0.5621$ (b) $\epsilon = 0.59$

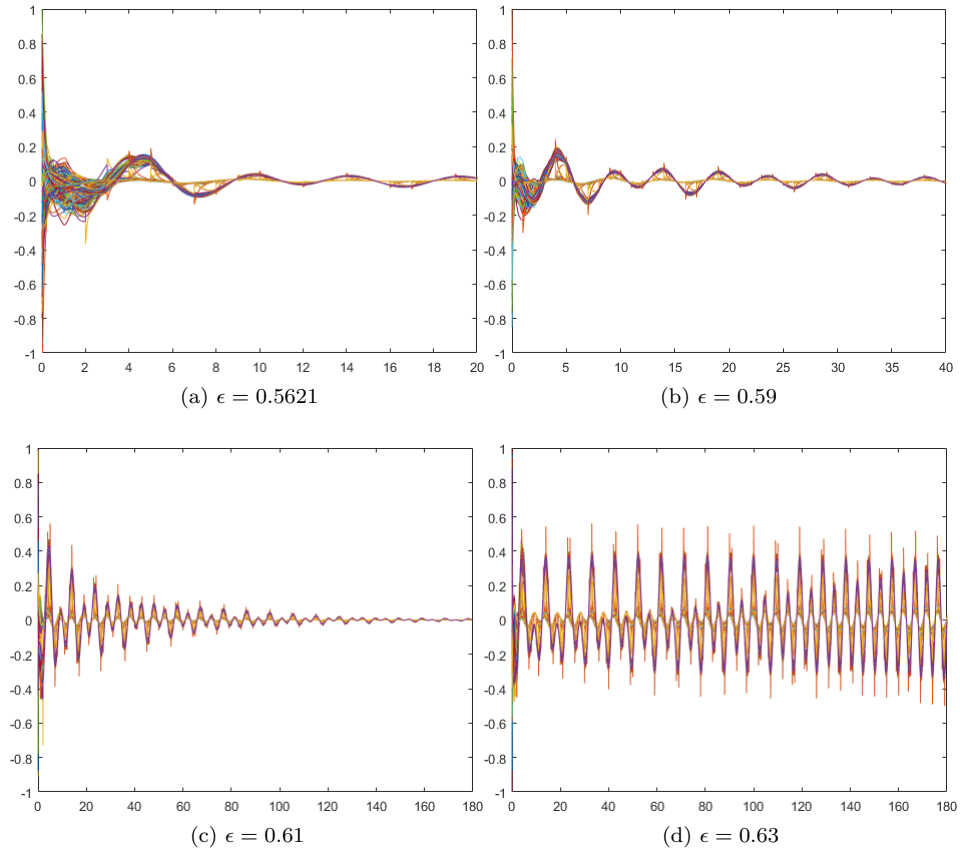
(c) $\epsilon = 0.61$ (d) $\epsilon = 0.63$

FIG. 3. *Sample trajectories (vertical axis) from random constant initial conditions drawn from the standard normal distribuition, rescaled to the interval* $[-1, 1]$ *and plotted for various representative time intervals (horizontal axis). The neural activation parameter* ϵ *for the given simulation is listed below its frame. Notice the transition from a stable periodic orbit to a stable equilibrium in the parameter interval* $[0.61, 0.63]$*, indicative of another bifurcation point.*

in the preamble to section 5 that such secondary bifurcations of the model without impulses may be responsible for poor performance or complete failure of the impulsive controller if the parameter is too far away from the bifurcation point. The presence of a secondary bifurcation in the interval $[0.61, 0.63]$ in both cases—with and without the controller—is consistent with this claim.

7. Discussion. We have proposed a method of stabilizing a candidate equilibrium point of an autonomous delay dynamical system at or near a bifurcation point using a novel impulsive stabilization approach based on invariant manifold theory. We have also introduced cost structure; our method identifies a jump functional (impulsive controller) that minimizes a given cost functional, should certain controllers have cost associated to their implementation. The method, which we call the center probe method, or CPM, takes advantage of the dimension reduction inherent to the dynamics on the center manifold. If the cost is projective, the CPM can be implemented with great efficiency.

Our method differs from the majority of impulsive stabilization methods in that one does not design the controller to satisfy a system of matrix inequalities, but rather



(a) $\epsilon = 0.5621$ (b) $\epsilon = 0.59$

(c) $\epsilon = 0.61$ (d) $\epsilon = 0.63$

FIG. 4. *Sample trajectories (vertical axis) from random constant initial conditions drawn from the standard normal distribuition, rescaled to the interval* $[-1, 1]$*, with different neural activation parameters, plotted for various representative time intervals (horizontal axis).* (a) *Activation parameter* $\epsilon = 0.5621$*; the trivial equilibrium is quickly stabilized.* (b) *Activation parameter* $\epsilon = 0.59$*; the trivial equilibrium is exponentially stabilized but with a slightly smaller rate parameter.* (c) *Activation parameter* $\epsilon = 0.61$*; the trivial equilibrium is stabilized, but the rate parameter is low.* (d) *Activation parameter* $\epsilon = 0.63$*; the control no longer stabilizes the trivial equilibrium. A secondary bifurcation has occured.*

solves a pair of optimization problems. The most computationally expensive of the two—the probe program ($\mathcal{P}\tilde{C}$)—is done in a low-dimensional space determined by the number of eigenvalues crossing the imaginary axis at the bifurcation point. The second one—the inverse probe program (Y_P)—is convex, smooth, and unconstrained (in an appropriate coordinate system) and can be solved efficiently using out-of-the-box nonlinear solvers. This is the greatest strength of our method compared to those derived from Lyapunov functionals based on linear matrix inequalities: constraint satisfaction for the CPM is done in a low-dimensional space, and if the cost is projective the low-dimensional probe program is sufficient. To contrast, the matrix inequalities of the other methods are checked in the original (potentially) high-dimensional space and must also be synthesized therein.

Being an inherently local method, the CPM does not rely on global Lipschitzian conditions on the vector field or any boundedness contraints. The higher-order terms are not important and, notably, can grow superlinearly as $||x|| \to \infty$ without

consequence. This is a strength of the CPM compared to Lyapunov functional-based linear matrix inequality sufficient conditions for impulsive stabilization. The latter universally require Lipschitzian or Lipschitz-like conditions on the vector field in order to rigorously ensure stability. An unfortunate counterpoint is that, being a local method, the CPM cannot guarantee global stability even if the vector field has a small Lipschitz constant.

As mentioned in the introduction, minimization of a cost functional defined on a space of admissible jump functionals (impulsive controllers) appears to be unique to our method. Our framework contains as a subcase pinning impulsive stabilization, so in the stabilization of a complex network one can use our method to minimize the number of controlled nodes.

The CPM was tested on an artificial neural network model with 100 identical linearly coupled nodes, in section 6. A projective cost was chosen that weighted the cost of controlling each node based on its degree, with higher degree nodes being more expensive. We identified a bifurcation point and implemented the CPM to identify a probe element and an associated affine space of jump functionals that minimize the cost functional. The number of nodes was minimized by solving a sequence of unconstrained problems in this affine space, with the result being a pinning impulsive controller that controls fourteen specific nodes.

We demonstrated by way of numerical simulations that our controller stabilizes the trivial equilibrium in a parameter interval near the bifurcation point but that a secondary bifurcation occurs when the parameter becomes too large. A similar, secondary bifurcation occurs in the system without the impulsive controller.

A weakness of our method in its current state is that such secondary bifurcations are barriers to further impulsive stabilization. Suppose, for example, that the uniform CPM is implemented at scalar parameter $\epsilon = 0$ and a controller $B(t)$ is found but that a secondary bifurcation occurs at the parameter $\epsilon^* \neq 0$. That is, the system

$$\begin{aligned} \dot{x} &= f(x_t, \epsilon), && t \notin \frac{1}{h}\mathbb{Z}, \\ \Delta x &= B(t)x_{t^-}, && t \in \frac{1}{h}\mathbb{Z}, \end{aligned}$$

undergoes a bifurcation at the parameter $\epsilon = \epsilon^*$. Technically, the CPM cannot be applied at this parameter because the system is no longer an autonomous delay differential equation. In principle, the CPM could be extended to this case, but the computation of basis functions in the center fiber bundle (the time-varying analogue of the center subspace needed to handle nonautonomous systems in infinite-dimensional spaces) becomes much more difficult, and these are absolutely essential to the CPM. Such basis functions are inherently discontinuous and, to our knowledge, there are at present no rigorous numerical methods available to compute them in general.

Another limitation of our method is that it cannot in general be used to stabilize an unstable equilibrium point in a delay differential equation. The method has a chance of success only if there is a sufficiently nearby perturbation of the system such that the equilibrium satisfies the spectral gap condition. In the absence of such a nearby perturbation, traditional impulsive stabilization approaches based on matrix inequalities derived from Lyapunov functional-based linear matrix inequality sufficient conditions for stability are more appropriate.

The nonconvexity of the probe program does pose certain difficulties insofar as implementation is concerned. These all center around the characterization of the feasible set as being the intersection of $\mathcal{P}^k$ with the sublevel set $\{X \in (\mathbb{R}^{c\times c})^k :$

$\rho \circ G(X) \leq e^{-\gamma/h}\}$. As we recalled at the beginning of section 4, the spectral radius $\rho : \mathbb{R}^{c\times c} \to \mathbb{R}$ is nonconvex and nonsmooth. However, on the set

$$V = \{X \in \mathbb{R}^{c\times c} : \text{every eigenvalue of } X \text{ is simple}\},$$

the spectral radius is continuously differentiable and locally Lipschitz continuous. V is also open and dense in $\mathbb{R}^{c\times c}$. As a consequence, methods based on gradient sampling [3, 17] can be used to solve the program $(\mathcal{P}\tilde{C})$ with guaranteed convergence results.

8. Conclusion. We have proposed the center probe method, or CPM, a novel impulsive stabilization method for delay differential equations based on invariant manifold theory. The method takes advantage of the low-dimensional properties of the center manifold at a bifurcation point of the system to be stabilized and allows an impulsive controller to be identified that minimizes a cost functional that can take into account many structural features of the controller that may affect its implementation cost. The method consists of a pair of optimization problems, with the most computationally heavy constraint satisfaction being done in a typically low-dimensional space.

The CPM has a few weaknesses. The uniform CPM is only guaranteed to provide some finte parameter interval where robust stability is guaranteed near a bifurcation point. The interval cannot be selected a priori and the controller generated by the uniform CPM will generally fail to stabilize the system if the parameter is taken too far away from the bifurcation point. Also, the CPM cannot generally be used to stabilize an unstable equilibrium point: it is necessary for there to be a sufficiently nearby perturbation of the system such that the equilibrium point undergoes a bifurcation.

Our method has several strengths compared to impulsive stabilization methods based on matrix inequalities derived from Lyapunov functionals. It does not rely on Lipschitzian (or Lipschitz-like) conditions to rigorously guarantee stability and the constraint satisfaction step is done in a low-dimensional space. Also, in complex networks of identical coupled oscillators, the dimension of this low-dimensional space remains low even while the number of nodes increases, while the number of constraints in a matrix inequality-based stability condition increases quadratically with the number of nodes. Our method allows one to consider cost structure and optimize the controller with respect to this cost, as well as impose hard constraints on the types of controllers permitted.

Appendix A. Proofs.

A.1. Proof of Proposition 3.1. 1. That the set $\widetilde{\mathcal{Y}}(\epsilon, h; \gamma)$ is closed follows by the continuity of the spectral radius function $\rho : \mathbb{R}^{c\times c} \to \mathbb{R}^+$. To prove the equality (3.7), consider the dynamics on the c-dimensional slice of the parameter-dependent center manifold at the point $x = 0$ for small parameter $(\epsilon, B) \in \mathbb{R}^p \times \mathcal{U}$ of the IRFDE

$$\dot{x} = L_0 x_t + L(\epsilon)x_t + O(||x_t||^2), \qquad t \notin \frac{1}{h}\mathbb{Z}, \tag{A.1}$$

$$\Delta x = B(t)x_{t^-}, \qquad t \in \frac{1}{h}\mathbb{Z}, \tag{A.2}$$

where we have overloaded the notation and identify B with the functional on the right-hand side of (2.1) and, if $k > 1$, using (2.6). At the parameter $(\epsilon, B) = 0$, the center fiber bundle (equivalently, center subspace) is determined by the homogeneous linear equation without impulses

$$\dot{y} = L_0 y_t,$$

so the claim concerning $\Phi(t)$ and $\Psi(t)$ as being basis matrices associated to the eigenspaces E_0 and E_0^T is true—see [8]. Moreover, we trivially have the Floquet decomposition

$$\Phi_t(\theta) = \Phi(0)e^{\Lambda\theta}e^{\Lambda t} := Q(\theta)e^{\Lambda t}$$

due to the characterization $\frac{d}{dt}\Phi(t) = \Phi(t)\Lambda$. This implies that $Q_t = \Phi_0$ is constant in t. In the notation of section 3.3 from [2], the projection operator $P_c(t)$ is constant and we therefore have

$$P_c(t)\chi_0 = \Phi_0\langle\Psi_0, \Phi_0\rangle^{-1}\langle\Psi_0, \chi_0\rangle = Q\Gamma\Psi(0) = \Phi_t e^{-\Lambda t}\Gamma\Psi(0).$$

Applying [2, Lemma 3.1] the dynamics on the c-dimensional slice of the parameter-dependent center manfiold are given to linear order in $z \in \mathbb{R}^c$ by

$$\dot{z} = \Lambda z + \Gamma\Psi(0)L(\epsilon)\Phi_0 z + O(||z||^2), \qquad t \notin \frac{1}{h}\mathbb{Z}, \tag{A.3}$$

$$\Delta z = \Gamma\Psi(0)B(t)\Phi_0 z(t^-), \qquad t \in \frac{1}{h}\mathbb{Z}, \tag{A.4}$$

for (ϵ, B) sufficiently small. The monodromy matrix associated to the equilibrium $z = 0$ is precisely $\mathcal{M}_{\epsilon,h}(B)$, and the dynamics restricted to the center manifold therefore has $O(e^{-\gamma t})$ as a convergence rate if and only if $\rho\mathcal{M}_{\epsilon,h}(B) \leq e^{-\gamma/h}$. Since all other Floquet exponents of the linearization of (A.1)–(A.2) have strictly negative real parts and the spectrum is upper hemicontinuous with respect to perturbations in (ϵ, B), the same convergence rate is achieved locally near $x = 0$ in the original nonlinear system due to linearized stability principles and the assumption $\sigma < \gamma$. The converse follows by the same argument.

2. Since the cost $\mathcal{C} : \mathcal{U} \to \mathbb{R}$ is nonnegative, the set $X = \mathcal{C}(\widetilde{\mathcal{Y}}(\epsilon, h; \gamma)) \subset \mathbb{R}$ is bounded below, so there exists a sequence $B_n \in \widetilde{\mathcal{Y}}(\epsilon, h; \gamma)$ such that $\mathcal{C}(B_n) \to \inf X$ as $n \to \infty$. The sequence B_n cannot be unbounded because $\mathcal{C}$ is radially unbounded, from which we conclude by the Bolzano–Weierstrass theorem that B_n admits a convergent subsequence having a limit $B_{\epsilon,h} \in \widetilde{\mathcal{Y}}(\epsilon, h; \gamma)$, with the latter inclusion justified by part 1. As $\mathcal{C}$ is continuous, we conclude that $\mathcal{C}(B_{\epsilon,h}) = \inf X$.

A.2. Proof of Theorem 3.4. To begin, we assume $k = 1$. If there exists a linear jump functional B^* such that $\mathcal{M}_{\epsilon,h}(B^*) = e^{-\gamma/h}I$, then the nonemptiness of $\widetilde{\mathcal{Y}}(\epsilon, h; \gamma)$ result will follow. Thus, it suffices to solve the equation

$$(I + \Gamma\Psi(0)B^*\Phi_0)\exp\left(\frac{1}{h}(\Lambda + \Psi(0)L(\epsilon)\Phi_0)\right) = e^{-\gamma/h}I. \tag{A.5}$$

If $\Phi(0)$ and $\Psi(0)$ are rank c, there exists a left-inverse $\Phi^+(0)$ and a right-inverse $\Psi^+(0)$ such that $\Phi^+(0)\Phi(0) = I_{n\times n}$ and $\Psi(0)\Psi^+(0) = I_{c\times c}$. Then, the jump functional $B^*_{\epsilon;\gamma}$ defined by

$$B^*_{\epsilon;\gamma}\xi = \Psi^+(0)\Gamma^{-1}(e^{-\gamma/h} - 1)\exp\left(-\frac{1}{h}(\Lambda + \Psi(0)L(\epsilon)\Phi_0)\right)\Phi^+(0)\xi(0)$$

satisfies (A.5). Since $B^*_{\epsilon;\gamma} \to 0$ as $\gamma \to 0^+$, it follows by minimality that also $\mathcal{C}(B^*_{\epsilon;\gamma}) \to 0$ as $\gamma \to 0^+$ for any selection $\gamma \mapsto B^*_{\epsilon;\gamma}$ of cost-minimizing jump functionals for rate parameter γ. As $\mathcal{C}$ is continuous and positive-definite, we conclude that $B^*_{\epsilon;\gamma} \to 0$.

If $k > 1$, consider the period k cycle of jump functionals $B = (B^*_{\epsilon;\gamma/k}, \ldots, B^*_{\epsilon;\gamma/k})$. Then $\mathcal{M}_{\epsilon,h}(B) = e^{-\gamma/h}I$ and the argument proceeds as before.

A.3. Proof of Corollary 3.5. If $c = 1$, then $\Phi(t)$ is an $n \times 1$ column vector and, as it constitutes a basis for E_0, it cannot be identically zero. Moreover, as $\Phi(t) = \Phi(0)e^{\Lambda t}$, we cannot have $\Phi(0) = 0$. Consequently, $\Phi(0)$ has rank one. The same argument applies to $\Psi(0)$.

If $L_0\phi = C\phi(0)$ for an $n \times n$ matrix C, then $\Phi(t) = Z(t)M$ for some $n \times c$ with matrix M and $Z(t)$ is the fundamental matrix solution of the finite-dimensional linear equation $\dot{z} = Cz$ satisfying $Z(0) = I$. As each column of $\Phi(t)$ is a linear combination of the columns of $Z(t)$, this set of columns is linearly independent if and only if M has maximal rank. As $c \leq n$, it follows that M has rank c, so that the same is true of $\Phi(0)$. The same argument applies to $\Psi(0)$.

A.4. Proof of Lemma 4.6. Let $x \in \mathcal{M}_1^{-1}(X)$ and $y \in \mathcal{M}_1^{-1}(Y)$. Then, we have

$$\begin{aligned}\mathcal{M}_1(tx + (1-t)y) &= M_0(tx + (1-t)y) + Z \\ &= t(M_0(x) + Z) + (1-t)(M_0(y) + Z) \\ &= t\mathcal{M}_1(x) + (1-t)\mathcal{M}_1(y) \\ &= tX + (1-t)Y.\end{aligned}$$

Consequently, $tx + (1-t)y \in \mathcal{M}_1^{-1}(tX + (1-t)Y)$ for all $t \in [0,1]$, $x \in \mathcal{M}_1^{-1}(X)$, and $y \in \mathcal{M}_1^{-1}(Y)$, from which the inclusion (4.2) follows.

A.4.1. Proof of Theorem 4.11. Before we begin, we remark that by virtue of convexity, any local optimum of the program (Y_P) is automatially a global optimum. As such, we will refer to these as global optimums.

Let $\tilde{C}$ be a GPCC, let P be a global optimum for (Y), and let B be a global optimum for the problem (Y_P). Suppose by way of contradiction that B is not a global optimum for (Y). By Proposition 3.1, there must exist a global optimum B^*, and as $\tilde{C}$ is a GPCC we have $\tilde{C} \circ \mathcal{M}_k(B^*) < \tilde{C} \circ \mathcal{M}_k(B)$. But $P = \mathcal{M}_k(B)$, and the latter inequality implies P is not optimal for the program $(\mathcal{P}\tilde{C})$, a contradiction. Therefore, B is a global optimum for (Y).

Now, let $\tilde{C}$ be an LPCC, let P be a local optimum for (Y), and let B be a global optimum for the problem (Y_P). Suppose by way of contradiction that B is not a local optimum for (Y). Then, there exists a sequence $B_n \to B$ such that $\mathcal{C}(B_n) < \mathcal{C}(B)$. By continuity of $\mathcal{M}_k : \mathcal{U}^k \to (\mathbb{R}^{c\times c})^k$, we have $P_n := \mathcal{M}_k(B_n) \to \mathcal{M}_k(B) = P$, but local optimality of P implies that for n sufficiently large, $\tilde{C}(P) \leq \tilde{C}(P_n)$. Using the LPCC property (type 1 or type 2) of $\tilde{C}$, it follows that $\mathcal{C}(B) \leq \mathcal{C}(B_n)$ for n sufficiently large, which is a contradiction. The result follows.

A.5. Proof of Theorem 4.12. Radial unboundedness follows from radial unboundedness of $\mathcal{C} : \mathcal{U} \to \mathbb{R}^+$. For continuity, we remark that $\tilde{C}$ is the optimal value function associated to the family of convex programs (Y_P) indexed by the parameter $P \in \mathcal{P}^k$. We have $\mathcal{M}_k^{-1} : \mathcal{P}^k \rightrightarrows \mathcal{U}^k$ is closed-valued, convex-valued, and can be shown to be continuous (see the characterization provided by Lemma 4.5). The cost $\mathcal{C}$ is convex and continuous, and the solution set is bounded for each P due to the radial unboundedness of $\mathcal{C}$. Using [20, Theorem 1] we obtain continuity of $\tilde{C}$. Using Lemma 4.6, Lemma 4.8, and the convexity of $\mathcal{C}$, if we use the subscript notation $P = (P_0, \ldots, P_{k-1})$ for arbitrary k-probe elements, then

$$\begin{aligned}\tilde{C}(tX + (1-t)Y) &= \min\{\mathcal{C}(V) : V \in \mathcal{M}_k^{-1}(tX + (1-t)Y)\} \\ &= \min\{\mathcal{C}(V_0, \ldots, V_{k-1}) : V_i \in \mathcal{M}_1^{-1}(tX_i + (1-t)Y_i)\} \\ &\leq \min\{\mathcal{C}(V_0, \ldots, V_{k-1}) : V_i \in t\mathcal{M}_1^{-1}(X_i) \oplus (1-t)\mathcal{M}_1^{-1}(Y_i)\}\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=0}^{k-1} \min\{\mathcal{C}(tx_i + (1-t)y_i) : x_i \in \mathcal{M}_1^{-1}(X_i), y_i \in \mathcal{M}_1^{-1}(Y_i)\} \\
&\leq \sum_{i=0}^{k-1} \min\{t\mathcal{C}(x_i) + (1-t)\mathcal{C}(y_i) : x_i \in \mathcal{M}_1^{-1}(X_i), y_i \in \mathcal{M}_1^{-1}(Y_i)\} \\
&\leq t\tilde{C}(X) + (1-t)\tilde{C}(Y),
\end{aligned}
$$

and we conclude that $\tilde{C}$ is convex.

Next we prove that $\tilde{C}$ is a GPCC. If B^* is a global optimum and B is feasible, then

$$
\begin{aligned}
\tilde{C} \circ \mathcal{M}_k(B^*) &= \min_{x \in \mathcal{M}_k^{-1}(\mathcal{M}_k(B^*))} \mathcal{C}(X) = \mathcal{C}(B^*) \\
&\leq \min_{x \in \mathcal{M}_k^{-1}(\mathcal{M}_k(B))} \mathcal{C}(X) = \tilde{C} \circ \mathcal{M}_k(B).
\end{aligned}
$$

B is also a global minimum if and only if $\mathcal{C}(B^*) = \mathcal{C}(B)$, and in this case the calculation above is simplified. We have

$$
\tilde{C} \circ \mathcal{M}_k(B^*) = \mathcal{C}(B^*) = \mathcal{C}(B) = \tilde{C} \circ \mathcal{M}_k(B).
$$

Our final task is to prove that $\tilde{C}$ is a LPCC. We have

$$
\tilde{C} \circ \mathcal{M}_k(B) = \min_{X \in \mathcal{M}_k^{-1}(\mathcal{M}_k(B))} \mathcal{C}(X) \leq \mathcal{C}(B)
$$

because $B \in \mathcal{M}_k^{-1}(\mathcal{M}_k(B))$. Equality holds precisely when B satisfies (4.3).

A.6. Proof of Corollary 5.2. Under the assumption of the corollary, the set $\{x \geq 0 : \forall |\epsilon| \leq x,\ \rho(\mathcal{M}_{\epsilon,h}(B)) \leq e^{-\gamma/h}\}$ is bounded above and therefore has a supremum, η. By continuity of the spectral radius, it follows that B is a feasible solution for the program (Y^η). To see that it is optimal, we recall that due to part 2 of Proposition 3.7, we have the inclusion $\widetilde{\mathcal{Y}}^\eta(h;\gamma) \subseteq \widetilde{\mathcal{Y}}(0,h;\gamma)$. Since B minimizes (locally or globally) $\mathcal{C} : \widetilde{\mathcal{Y}}(0,h;\gamma) \to \mathbb{R}$ and is feasible for (Y^η), it must also minimize (locally or globally) $\mathcal{C} : \widetilde{\mathcal{Y}}^\eta(h;\gamma) \to \mathbb{R}$.

Acknowledgment. We are grateful to the reviewers for their careful reading of the manuscript and their helpful comments.

REFERENCES

[1] S. A. Campbell and Y. Yuan, *Zero singularities of codimension two and three in delay differential equations*, Nonlinearity, 21 (2008), pp. 2671–2691, https://doi.org/10.1088/0951-7715/21/11/010.

[2] K. E. Church and X. Liu, *Computation of centre manifolds and some codimension-one bifurcations for impulsive delay differential equations*, J. Differential Equations, 267 (2019), pp. 3852–3921, https://doi.org/10.1016/j.jde.2019.04.022.

[3] F. Curtis and M. Overton, *A sequential quadratic programming algorithm for nonconvex, nonsmooth constrained optimization*, SIAM J. Optim., 22 (2012), pp. 474–500, https://doi.org/10.1137/090780201.

[4] O. Diekmann, S. M. Verduyn Lunel, S. A. van Gils, and H.-O. Walther, *Delay Equations: Functional-, Complex-, and Nonlinear Analysis*, Appl. Math. Sci. 110, Springer, New York, 1995, https://doi.org/10.1007/978-1-4612-4206-2.

[5] K. Engelborghs, T. Luzyanina, and D. Roose, *Numerical bifurcation analysis of delay differential equations using DDE-BIFTOOL*, ACM Trans. Math. Software, 28 (2002), pp. 1–21, https://doi.org/10.1145/513001.513002.

[6] S. Guo and J. Li, *Bifurcation theory of functional differential equations: A survey*, J. Appl. Anal. Comput., 5 (2015), pp. 751–766, https://doi.org/10.11948/2015057.

[7] S. Guo and J. Wu, *Bifurcation Theory of Functional Differential Equations*, Appl. Math. Sci. 184, Springer, New York, 2013, https://doi.org/10.1007/978-1-4614-6992-6.

[8] J. K. Hale and S. M. V. Lunel, *Introduction to Functional Differential Equations*, Appl. Math. Sci. 99, Springer, New York, 1993, https://doi.org/10.1007/978-1-4612-4342-7.

[9] W. He, F. Qian, and J. Cao, *Pinning-controlled synchronization of delayed neural networks with distributed-delay coupling via impulsive control*, Neural Networks, 85 (2017), pp. 1–9, https://doi.org/10.1016/j.neunet.2016.09.002.

[10] Y. A. Kuznetsov, *Elements of Applied Bifurcation Theory*, Appl. Math. Sci. 112, Springer, New York, 2004, https://doi.org/10.1007/978-1-4757-3978-7.

[11] O. M. Kwon, M. J. Park, J. H. Park, S. M. Lee, and E. J. Cha, *On stability analysis for neural networks with interval time-varying delays via some new augmented Lyapunov-Krasovskii functional*, Commun. Nonlinear Sci. Numer. Simul., 19 (2014), pp. 3184–3201, https://doi.org/10.1016/j.cnsns.2014.02.024.

[12] H. Li, C. Li, T. Huang, and W. Zhang, *Fixed-time stabilization of impulsive Cohen–Grossberg BAM neural networks*, Neural Networks, 98 (2018), pp. 203–211, https://doi.org/10.1016/j.neunet.2017.11.017.

[13] X. Li and S. Song, *Stabilization of delay systems: Delay-dependent impulsive control*, IEEE Trans. Automat. Control, 62 (2017), pp. 406–411, https://doi.org/10.1109/TAC.2016.2530041.

[14] C. Liu and W.-H. Chen, *Guaranteed cost control of linear uncertain discrete-time impulsive systems*, Trans. Inst. Measurement Control, 37 (2015), pp. 33–39, https://doi.org/10.1177/0142331214528969.

[15] X. Liu and Q. Wang, *Impulsive stabilization of high-order Hopfield-type neural networks with time-varying delays*, IEEE Trans. Neural Networks, 19 (2008), pp. 71–79, https://doi.org/10.1109/TNN.2007.902725.

[16] J. Lu, Z. Wang, J. Cao, D. W. C. Ho, and J. Kurths, *Pinning impulsive stabilization of nonlinear dynamical networks with time-varying delay*, Internat. J. Bifur. Chaos, 22 (2012), 1250176, https://doi.org/10.1142/S0218127412501763.

[17] C. ming Tang, S. Liu, J. bao Jian, and J. ling Li, *A feasible SQP-GS algorithm for non-convex, nonsmooth constrained optimization*, Numer. Algorithms, 65 (2014), pp. 1–22, https://doi.org/10.1007/s11075-012-9692-5.

[18] G. Stepan, *Retarded Dynamical Systems: Stability and Characteristic Functions*, Longman Scientific and Technical, Harlow, UK, 1989.

[19] J. Sun and J. Chen, *A survey on Lyapunov-based methods for stability of linear time-delay systems*, Frontiers Comput. Sci., 11 (2017), pp. 555–567, https://doi.org/10.1007/s11704-016-6120-3.

[20] Y. Terazono and A. Matani, *Continuity of optimal solution functions and their conditions on objective functions*, SIAM J. Optim., 25 (2015), pp. 2050–2060, https://doi.org/10.1137/110850189.

[21] X. Wang, C. Li, T. Huang, and L. Chen, *Impulsive exponential synchronization of randomly coupled neural networks with Markovian jumping and mixed model-dependent time delays*, Neural Networks, 60 (2014), pp. 25–32, https://doi.org/10.1016/j.neunet.2014.07.008.

[22] D. J. Watts and S. H. Strogatz, *Collective dynamics of 'small-world' networks*, Nature, 393 (1998), pp. 440–442, https://doi.org/Doi10.1038/30918.

[23] Z. Wu, G. Chen, and X. Fu, *Outer synchronization of drive-response dynamical networks via adaptive impulsive pinning control*, J. Franklin Inst., 352 (2015), pp. 4297–4308, https://doi.org/10.1016/j.jfranklin.2015.06.016.

[24] X. Xie, H. Xu, and R. Zhang, *Exponential stabilization of impulsive switched systems with time delays using guaranteed cost control*, Abstract Appl. Anal., 2014 (2014), https://doi.org/10.1155/2014/126836.

[25] B. Zhang, J. Lam, and S. Xu, *Stability analysis of distributed delay neural networks based on relaxed Lyapunov-Krasovskii functionals*, IEEE Trans. Neural Networks Learning Systems, 26 (2015), pp. 1480–1492, https://doi.org/10.1109/TNNLS.2014.2347290.

[26] C.-k. Zhang, Y. He, L. Jiang, and M. Wu, *Stability analysis for delayed neural networks considering both conservativeness and complexity*, IEEE Trans. Neural Networks Learning Systems, 27 (2016), pp. 1486–1501, https://doi.org/10.1109/TNNLS.2015.2449898.

[27] X. Zhao, P. Shi, Y. Yin, and S. K. Nguang, *New results on stability of slowly switched systems: A multiple discontinuous Lyapunov function approach*, IEEE Trans. Automat. Control, 62 (2017), pp. 3502–3509, https://doi.org/10.1109/TAC.2016.2614911.

[28] Y. Zhou, C. Li, T. Huang, and X. Wang, *Impulsive stabilization and synchronization of Hopfield-type neural networks with impulse time window*, Neural Comput. Appl., 28 (2017), pp. 775–782, https://doi.org/10.1007/s00521-015-2105-7.

[29] G. Zong, X. Wang, and H. Zhao, *Robust finite-time guaranteed cost control for impulsive switched systems with time-varying delay*, Internat. J. Control Automat. Systems, 15 (2017), pp. 1–9, https://doi.org/10.1007/s12555-015-0314-6.