Controlling Linear Networks with Minimally Novel Inputs

Gautam Kumar, Delsin Menolascino, MohammadMehdi Kafashan and ShiNung Ching

Abstract—In this paper, we propose a novelty-based metric for quantitative characterization of the controllability of complex networks. This inherently bounded metric describes the average angular separation of an input with respect to the past input history. We use this metric to find the minimally novel input that drives a linear network to a desired state using unit average energy. Specifically, the minimally novel input is defined as the solution of a continuous time, non-convex optimal control problem based on the introduced metric. We provide conditions for existence and uniqueness, and an explicit, closed-form expression for the solution. We support our theoretical results by characterizing the minimally novel inputs for an example of a recurrent neuronal network.

I. INTRODUCTION

In its most basic form, the systems-theoretic notion of controllability carries a binary definition: a dynamical system either is, or is not, controllable, with respect to its exogenous inputs. Naturally, such a notion has the deficiency of not grading the ease or difficulty associated with effecting such control. To obviate this issue, consistent research effort has been directed at the characterization of controllability using systems-theoretic metrics. Roughly, these metrics can be grouped into two categories

1) Those that characterize the minimum energy parametric perturbations that result in a loss of controllability [1], [2]. These are related to basic characterizations of the robustness of linear systems [3].
2) Those that characterize the controllability of a system in terms of the minimum energy excitation required to achieve a unit length state trajectory [4]–[7].

The latter, in particular, is a natural paradigm that is directly relatable to the celebrated Kalman rank condition (or the controllability gramian) used to ascertain the controllability of linear systems [8]. Recently, energy-based controllability metrics have been successfully used in the emerging domain of network science to assess the putative controllability of large-scale linear systems, formulated as complex networks of various topologies [5], [6]. However, for complex networks in general and, in particular, for biological neuronal networks, an energy-based metric offers insight into only one aspect of the overall system’s controllability.

We appeal, specifically, to the domain of neural coding and the dynamics of sensory neural circuits. Consider the simple, prototypical layered model of a sensory network shown in Figure 1 wherein sensory neurons are tuned to a high dimensional feature space (i.e., environmental variables from the sensory periphery; say, different molecules corresponding to tastes). Those sensory neurons impinge on a complex, interconnected sensory network that performs intermediate transformations en route to higher brain areas.

One may put forth a supposition that the ‘controllability’ of such a sensory network, with respect to the afferent input from the sensory neurons, is critical in mediating the ability to perceive minute changes in the environment. But as much as energy is important in mediating such a response, orientation, i.e., the alignment of an input with certain features, may be even more so. Indeed, a weak, but highly novel input may be more easily perceived than an intense, but more familiar, stimulus. The ability to assess the responsiveness of neuronal networks to novelty – at a particular moment in time, relative to past inputs – has immediate implications in the analysis and control of biophysiological neuronal network dynamics in different behavioral and clinical regimes [9]–[11].

Here, as a first step, we seek to characterize the controllability of linear systems (linear networks) possessing high dimensional input-spaces, with respect to input novelty. In particular, we ask how responsive are the state (node) trajectories to inputs that differ in orientation from those that have previously been applied. Figure 2 illustrates the
basic notion of input novelty for a simple two-dimensional linear system with three inputs. A particular input drives the system to an intermediate point in the phase space; from this point emerge two trajectories, both of which reach a common endpoint; one minimizes input novelty (note the similarity between the input from $t \in [0, 2]$ and that from $t \in [2, 4]$), the other minimizes energy.

Specifically, we: (i) analytically derive the \textit{minimum novelty control} for linear networks by formulating a non-convex optimization problem. The problem seeks the minimum angular separation, defined in terms of an inner product in the input feature space, required in order to create a desired angular separation, defined in terms of an inner product in the input feature space, required in order to create a desired orientation in (2) and (4), respectively.

From (3), we note that the novelty of the input $u(t)$ decreases as $J(t)$ increases and is minimum when $J(T) = 1$ i.e. when $u(t) = v(t - T)$ for all $t \in [0, T]$.

Remark 3: We observe that, due to the energy normalization in (2) and (4),

$$\frac{1}{T} \int_0^T \|v(t - T) - u(t)\|^2 dt = 2(1 - J(T))$$

Thus, the average Euclidean distance, i.e. the left hand side of (5), between two inputs can equivalently be used as an alternate measure of input novelty in our context.

C. \textit{Minimum novelty problem}

From the conceptual formulation introduced above, we can develop a control problem to design the minimally novel input $u(t), t \in [0, T]$ such that a desired directional change in the state of the system can be achieved under the constraint

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t)$$

where $x(t) \in \mathbb{R}^{d \times 1}$ represents the state of the system at time $t$, $A \in \mathbb{R}^{n \times n}$ is the state transition matrix, $B \in \mathbb{R}^{n \times m}$ is the input matrix, and $u(t) \in \mathbb{R}^{m \times 1}$ is the input to the system. Without loss of generality, we say that (1) describes the time evolution of linear networks in the presence of external inputs.

Let us assume an input $v(t - T) \in \mathbb{R}^{m \times 1}, t \in [0, T]$, with total energy $T$, i.e.

$$\frac{1}{T} \int_0^T \|v(t - T)\|^2 dt = 1$$

We assume that $v(t - T)$ can drive $x(t)$ from $x(0)$ to $x(T)$, where $\|x(0)\|_2 = 1$, subject to the dynamics (1). Here $T > 0$ is a constant. We introduce the inner-product based metric

$$J(T) = \frac{1}{T} \int_0^T v'(t - T)u(t) dt$$

where

$$\frac{1}{T} \int_0^T \|u(t)\|^2 dt = 1,$$
of fixed energy subject to the system dynamics (1). For this, we formulate the following optimal control problem:

\[
\begin{align*}
\min_{\mathbf{u}(t), t \in [0, T]} & \quad -J(T) \\
\text{s.t.} & \quad \frac{1}{T} \int_0^T \|\mathbf{u}(t)\|^2 dt = 1 \quad (6a) \\
& \quad \mathbf{x}(T) = e^{A T} \mathbf{x}(0) + \int_0^T e^{A(T-t)} \mathbf{B} \mathbf{u}(t) dt \quad (6b)
\end{align*}
\]

It should be noted here that the constraint (6c) is obtained by integrating (1) with respect to \(t\) over the period of \([0, T]\). Immediately, we note that the quadratic equality constraint (6b) makes the optimization problem (6) non-convex. Furthermore, we note that our optimal control problem formulation (6) is different from the classical minimum effort problems where the \(L^1\)-norm of control inputs is minimized under the constraints of explicit lower and upper bounds on the inputs.

III. RESULTS

We derive conditions for the existence of a unique global optimal solution of the non-convex optimization problem (6). Based on this, we provide a closed-form expression for the optimal \(\mathbf{u}(t), t \in [0, T]\).

A. Existence of a Minimally Novel Input

**Lemma 1:** A solution of the non-convex optimization problem (6) exists if

\[
T > \max \{ s'(T) W_c^{-1}(T) s(T), r'(T) W_c^{-1}(T) r(T) \} \quad (7)
\]

where

\[
\begin{align*}
s(T) &= \int_0^T e^{A(T-t)} \mathbf{B} \mathbf{v}(t-T) dt \quad (8a) \\
r(T) &= \mathbf{x}(T) - e^{A T} \mathbf{x}(0) \quad (8b)
\end{align*}
\]

Here, \(W_c(T)\) is the controllability gramian at time \(T\) and is defined as

\[
W_c(T) = \int_0^T e^{A(T-t)} \mathbf{B} \mathbf{B}^T e^{A^T(T-t)} dt
\]

Recall that by our formulation, \(T\) is the total energy available to the system (1).

**Remark 4:** The arguments \(s'(T) W_c^{-1}(T) s(T)\) and \(r'(T) W_c^{-1}(T) r(T)\) in (7) are the minimum energy required to drive the system (1) from \(\mathbf{x}(-T)\) to \(\mathbf{x}(0)\) and \(\mathbf{x}(0)\) to \(\mathbf{x}(T)\) respectively [7], [8].

**Proof:** Define \(y(t)\) as

\[
y(t) = \frac{1}{T} \int_0^t \|\mathbf{u}(\tau)\|^2 d\tau \quad (10)
\]

Clearly, \(y(0) = 0\) and \(y(T) = 1\) from (6b). Thus, we can replace the constraint (6b) by

\[
y(T) = 1 \quad (11)
\]

In differential form, we can write (10) as

\[
\frac{dy(t)}{dt} = \frac{1}{T} \|\mathbf{u}(t)\|^2 \quad (12)
\]

To solve the dynamic optimization problem (6a), (6b) and (6c) in continuous time, we write the Hamiltonian \(\mathcal{H}(\mathbf{x}(t), y(t), \mathbf{u}(t), \lambda(t), \mu(t), t)\) as

\[
\mathcal{H}(\mathbf{x}(t), y(t), \mathbf{u}(t), \lambda(t), \mu(t), t) = -\frac{1}{T} y'(t-T) \mathbf{u}(t) + \lambda'(t)(A \mathbf{x}(t) + B \mathbf{u}(t)) + \frac{\mu(t)}{T} \|\mathbf{u}(t)\|^2 \quad (13)
\]

Here, \(\lambda(t)\) and \(\mu(t)\) are the costate variables associated with the dynamics (1) and (12) respectively. We derive the following optimality conditions (i.e. the Euler-Lagrange equations [12]):

\[
\begin{align*}
\frac{d\lambda(t)}{dt} &= -(\partial \mathcal{H}(\mathbf{x}(t), y(t), \mathbf{u}(t), \lambda(t), \mu(t), t))' = -A' \lambda(t) \quad (14a) \\
\frac{d\mu(t)}{dt} &= -\partial \mathcal{H}(\mathbf{x}(t), y(t), \mathbf{u}(t), \lambda(t), \mu(t), t) = 0 \quad (14b) \\
\frac{\partial \mathcal{H}(\mathbf{x}(t), y(t), \mathbf{u}(t), \lambda(t), \mu(t), t)}{\partial \mathbf{u}(t)} &= 0
\end{align*}
\]

By integrating the costate equations (14a) and (14b) over \(t\), we obtain

\[
\lambda(t) = e^{-A'T} \lambda(0) \quad (15a) \\
\mu(t) \equiv \mu \quad \forall t \in [0, T] \quad (15b)
\]

Here, \(\lambda(0)\) is the initial condition (at \(t = 0\)) of (14a). From (14c), (15a) and (15b), we derive the optimal control law as

\[
\mathbf{u}(t) = \frac{1}{2\mu} \mathbf{v}(t-T) - \frac{T}{2\mu} B' e^{-A'T} \lambda(0) \quad (16)
\]

By substituting (16) into (6a), we obtain \(\lambda(0)\) as

\[
\lambda(0) = e^{A'T} W_e^{-1}(T) \left(\frac{1}{T} s(T) - \frac{2\mu}{T} r(T) \right) \quad (17)
\]

By substituting (16) and (17) in (11) and using (2), we obtain

\[
\mu = \pm \frac{1}{2} \sqrt{\frac{T - s(T) W_e^{-1}(T) s(T)}{T - r(T) W_e^{-1}(T) r(T)}} \quad (18)
\]

For the existence of a solution, \(\mu\) must be a real number. Thus, either \(T < \min \{ s'(T) W_c^{-1}(T) s(T), r'(T) W_c^{-1}(T) r(T) \}\) or \(T > \max \{ s'(T) W_c^{-1}(T) s(T), r'(T) W_c^{-1}(T) r(T) \}\). Now it follows directly from Remark 3 that the total energy \(T\) must satisfy (7) for the existence of a solution i.e. \(T > \max \{ s'(T) W_c^{-1}(T) s(T), r'(T) W_c^{-1}(T) r(T) \}\).
B. Uniqueness of the Minimally Novel Input

Lemma 2: Under the hypothesis of Lemma 1, the solution of the non-convex optimization problem (6) is unique.

Proof: By substituting (16) and (17) in (3), we obtain the optimal value of \( J(T) \) as a function of \( \mu \) as

\[
J(T) = \frac{1}{T} s'(T) W T^{-1} (T) r(T) + \frac{1}{2} \mu \left( 1 - \frac{1}{T} s'(T) W T^{-1} (T) s(T) \right)
\]

It follows from Lemma 1 that \( \frac{1}{T} s'(T) W T^{-1} (T) s(T) \in (0,1) \) (see (7)). Thus, the maximum of \( J(T) \) occurs when \( \mu > 0 \) in (18), i.e.

\[
\mu = \frac{1}{2} \sqrt{\frac{T - s(T) W T^{-1} (T) s(T)}{T - r(T) W T^{-1} (T) r(T)}}
\]

Thus, a unique optimal control input \( u(t) \) exists and is given by

\[
u(t) = \frac{1}{2 \mu} \left( v(t-T) - B' e^{A'(T-t)} W T^{-1} (T) s(T) \right) + B' e^{A'(T-t)} W T^{-1} (T) r(T)
\]

C. Euclidean - Inner Product Equivalence

As noted in Remark 3, it is an interesting and notable consequence of our cost formulation that the problem can exactly recast in terms of a Euclidean norm. Specifically, if we consider

\[
J_1(T) = \frac{1}{T} \int_0^T \| v(t-T) - u(t) \|^2 dt
\]

as the cost function in (6a), we obtain the optimal solution as

\[
\mu = -1 + \sqrt{\frac{T - s(T) W T^{-1} (T) s(T)}{T - r(T) W T^{-1} (T) r(T)}}
\]

\[
u(t) = \frac{1}{1 + \mu} \left( v(t-T) - B' e^{A'(T-t)} W T^{-1} (T) s(T) \right) + B' e^{A'(T-t)} W T^{-1} (T) r(T)
\]

IV. Example

We consider a recurrent network of \( n \) neurons with linearized firing rate dynamics of the form [13]

\[
\frac{dx(t)}{dt} = -x(t) + Tx(t) + Bu(t)
\]

Here, \( x(t) \in \mathbb{R}^{n \times 1} \) represents the firing rate of the neurons at time \( t \), \( S \in \mathbb{R}^{n \times n} \) is a diagonal matrix whose diagonal elements are the (positive) time constants of the neurons, \( W \in \mathbb{R}^{n \times n} \) defines the interaction among neurons in the network (weight matrix), \( B \in \mathbb{R}^{n \times n} \) is the input matrix, and \( u(t) \) is \( \mathbb{R}^{n \times 1} \) is the afferent input. Since \( S \) is invertible, (24) can be represented in the form of (1) by considering \( A = S^{-1} (I + W) \) where \( I \) is the \( n \times n \) identity matrix.

For illustrative purposes, we consider a recurrent network of \( n = 100 \) neurons where 80 neurons are excitatory and every 5th neuron is inhibitory. We choose the time constants (in ms) of the neurons, i.e. the elements of the diagonal matrix \( S \), from a uniform distribution \( \mathcal{U}(5,10) \). For every excitatory neuron \( i \), we choose the connectivity weight \( w_{i,j} \) (in essence, a time constant for excitation from the neuron \( i \) to \( j \)) from a uniform distribution \( \mathcal{U}(0.1) \). Similarly, for every inhibitory neuron \( i \), we choose the connectivity weight \( w_{i,j} \) (from the neuron \( i \) to \( j \)) from a uniform distribution \( \mathcal{U}(-1,0) \). We assume that \( w_{i,j} = 0 \) for \( i = j \), i.e. neurons do not possess direct feedback. Assuming \( B \) as an identity matrix, we proceed to compute the minimum directional change in inputs (i.e. minimally novel inputs) required to make a desired directional change in firing rates of neurons using (8)-(9), (19)-(21).

To complete the example, we specify \( T = 3 \) ms. The initial and terminal states \( x(0) \) and \( x(T) \), respectively, are specified to satisfy \( \| x(0) \|_2 = \| x(T) \|_2 = 1 \) with \( x(0)' x(T) = r \), where in this particular case we specify \( r = 0.7645 \). The prior input \( v(t - T) \) is specified to be constant over the interval \( t \in [0,T] \). Figure 3 illustrates the outcome of the example for \( n = 1000 \) random realizations of the system. Each red dot on the figure depicts the novelty associated with the solution to (18)-(19), i.e., the minimum novelty. Note, again, that by formulation, these inputs all have unit average energy. Each blue dot corresponds to the minimum energy solution. As a verification of our theoretical development, we note that the minimum energy solution consistently requires an injection of novelty (angular orientation) relative to the prior input and relative to the optimum.

V. Conclusions and Future Work

In this paper, we have introduced a systems-theoretic analysis to characterize the minimum input novelty required to effect a change to the trajectory of a linear system. We have focused this paper on introducing the key conceptual notion and on exact analytical characterization of the minimum novelty solution for the case of linear systems. Naturally, several extensions are possible and some are immediate. For instance, the analysis readily extends to the case of linear-time varying systems, with appropriate replacement of the static \( A \) and \( B \) matrices with their time-varying equivalents.
Fig. 3. Comparison of minimum novelty control with minimum energy control for $n = 1000$ random realizations of the recurrent neuronal network: Each red dot on the figure depicts the novelty associated with the solution to the minimum novelty control. Each blue dot corresponds to the minimum energy solution.

in the controllability gramian. Depending on the domain example at hand, one may also modify the novelty metric itself, for instance by weighting novelty in certain segments of the state-space.

The compelling aspect of this analysis is its direct interpretability in the context of sensory neuronal networks where, as stated in the Introduction, energy alone does not provide a full controllability characterization. With suitable adaptation, it is expected that the analysis herein can be used for both the analysis of biophysical neuronal networks in clinically relevant regimes [14] and, eventually, for design and synthesis of sensory inputs [15].

VI. ACKNOWLEDGEMENT

S. Ching Holds a Career Award at the Scientific Interface from the Burroughs-Wellcome Fund.

REFERENCES

[1] G. Hu and E. Davison, “A real radius measure for controllability,” in American Control Conference, 2001. Proceedings of the 2001, vol. 4, 2001, pp. 3144–3148. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=946404

[2] G. Hu and E. J. Davison, “Real controllability/stabilizability radius of lti systems,” IEEE Transactions on Automatic Control, vol. 49, no. 2, pp. 254–257, 2004.

[3] D. Hinrichsen and A. Pritchard, Real and Complex Stability Radii: A Survey, ser. Institut für Dynamische Systeme, Universität Bremen. Inst. für Dynam. Systeme, 1989. [Online]. Available: http://books.google.com/books?id=EpC8AAACAAJ

[4] G. Yan, J. Ren, Y.-C. Lai, C.-H. Lai, and B. Li, “Controlling complex networks: How much energy is needed?” Phys. Rev. Lett., vol. 108, p. 218703, May 2012. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevLett.108.218703

[5] F. Pasqualetti, S. Zampieri, and F. Bullo, “Controllability metrics, limitations and algorithms for complex networks,” IEEE Transactions on Control of Network Systems, vol. 1, no. 1, pp. 40–52, 2014.

[6] F. Pasqualetti and S. Zampieri, “On the controllability of isotropic and anisotropic networks,” 2014. [Online]. Available: http://www.fabiopas.it/papers/FP-SZ-14.pdf

[7] F. L. Cortesi, T. H. Summers, and J. Lygeros, “Submodularity of energy related controllability metrics,” ArXiv preprint arXiv:1403.6351, 2014.

[8] H. Kwakernaak and R. Sivan, Linear optimal control systems. Wiley-Interscience, 1972.

[9] S. Ching and J. T. Ritt, “Control strategies for underactuated neural ensembles driven by optogenetic stimulation.” Front Neural Circuits, vol. 7, p. 54, 2013. [Online]. Available: http://dx.doi.org/10.3389/fncir.2013.00054

[10] S. Ching, E. N. Brown, and M. A. Kramer, “Distributed control in a mean-field cortical network model: implications for seizure suppression.” Phys Rev E Stat Nonlin Soft Matter Phys, vol. 86, no. 2 Pt 1, p. 021920, Aug 2012.

[11] K. Q. Lepage, S. Ching, and M. A. Kramer, “Inferring evoked brain connectivity through adaptive perturbation.” J Comput Neurosci, vol. 34, no. 2, pp. 303–318, Apr 2013. [Online]. Available: http://dx.doi.org/10.1007/s10877-012-0422-8

[12] D. E. Kirk, Optimal Control Theory: An Introduction. Dover, 2004.

[13] P. Dayan and L. F. Abbott, Theoretical Neuroscience. The MIT Press, 2001.

[14] S. Ching and E. N. Brown, “Modeling the dynamical effects of anesthesia on brain circuits.” Curr Opin Neurobiol, vol. 25, pp. 116–122, Apr 2014. [Online]. Available: http://dx.doi.org/10.1016/j.conb.2015.12.011

[15] G. Kumar, M. H. Schieber, N. V. Thakor, and M. V. Kothare, “Designing closed-loop brain-machine interfaces using optimal receding horizon control;” In the Proceedings of the 2013 American Control Conference, Washington D.C., DC, pp. 5029–5034, 2013.