Supplementary Information for

Prevalence and scalable control of localized networks

Chao Duan, Takashi Nishikawa, and Adilson E. Motter

This PDF file includes:

  - Supplementary text
  - Fig. S1
  - Table S1
  - SI References
1. Basic Implications of Locality

Locality of the Solutions for Linear Equations. Here we show that, if \( M \) is localized, then the locality of \( b \) implies the locality of the solution of \( Mx = b \). To see this, suppose that an invertible \( M \) belongs to \( \mathcal{L}_{\nu,\rho} \) (i.e., \( M \) is localized) and that \( |b_j| \leq \kappa \cdot v(\rho(i,j))^{-1} \) for all \( j \) (i.e., \( b \) is localized around a given node \( i \)). Since \( \mathcal{L}_{\nu,\rho} \) is closed under the inverse operation, we have \( M^{-1} \in \mathcal{L}_{\nu,\rho} \). It then follows that the solution \( x^* = M^{-1}b \) is also localized around node \( i \), i.e., there is a constant \( \kappa' \) such that \( \|x^*\| \leq \kappa' \cdot v(\rho(i))^{-1} \). In the special case of \( b = e_ib' \) (recalling that \( e_i \in \mathbb{R}^{n_i \times 1} \) is a matrix mapping the state space of node \( i \) to that of the entire network), this implies that the solution of \( Mx = e_ib' \) is localized around node \( i \) for any \( n_i \)-dimensional vector \( b' \).

The locality of the solution \( x^* \) has further implications. Consider the projection of the equation \( Mx = b \) from the state space of the entire network to the subspace corresponding to the information neighborhood \( N_i(\tau) \):

\[
N_iM_iN_i^T z = N_i b, \tag{S1}
\]

where \( N_i \) is the corresponding projection matrix. The solution \( z^*(\tau) \) of this equation satisfies

\[
N_iM_iN_i^T(z^*(\tau) - N_i x^*) = -N_iM(N_i^TN_i - I)x^*. \tag{S2}
\]

The locality of \( x^* \) shown above ensures that the right hand side of this equation quickly decreases to zero as \( \tau \) is increased. More precisely, we have

\[
\|N_iM(N_i^TN_i - I)x^*\|_\infty = \mathcal{O}(v(\tau)^{-1}), \tag{S3}
\]

which implies

\[
\|z^*(\tau) - N_i x^*\|_\infty = \mathcal{O}(v(\tau)^{-1}) \tag{S4}
\]

and

\[
\|N_i^T z^*(\tau) - x^*\|_\infty = \mathcal{O}(v(\tau)^{-1}). \tag{S5}
\]

Thus, the locality of \( M \) and \( b \) guarantees that the solution of \( Mx = b \) can be well approximated by solving its projection, Eq. \( S1 \). Indeed, by decomposing any given vector \( b \) into individual nodes as \( b = \sum_i e_ib_i \) with \( b_i \in \mathbb{R}^{n_i \times 1} \) and solving \( N_iM_iN_i^T z = N_i e_i \) to obtain the \( m \times n_1 \) solution matrix \( z_i \), for each \( i \), we have an approximate solution of the original equation as \( x = \sum_i N_i^T z_i b_i \). If the sizes of the information neighbourhoods are chosen to be independent of the system size, this leads to a linear-time algorithm for computing approximate solutions of large-scale localized linear equations. In addition, the quadratic form \( b'^T M^{-1} b \) for an arbitrary vector \( b' \) can be well approximated by \( (N_i b')^T (N_iM_iN_i^T)^{-1} (N_i b) \), which can be seen by noting that

\[
\|(N_i b')^T (N_iM_iN_i^T)^{-1} (N_i b) - b'^T M^{-1} b\|_1 \leq \|b'\|_1 \|N_i^T z^*(\tau) - x^*\|_\infty = \mathcal{O}(v(\tau)^{-1}). \tag{S6}
\]

Here, we have assumed that the projected matrix \( N_iM_iN_i^T \) is invertible. However, even when this matrix is not invertible, this analysis still applies to the minimum norm solution of Eq. \( S1 \) obtained by using Moore–Penrose inverse of matrix \( N_iM_iN_i^T \).
Locality of the Matrix Exponential, Riccati Equation, and Lyapunov Equation. An implication of $\mathcal{L}_{v,\rho}$ being a Banach algebra is that it is closed under matrix exponential, i.e., $e^C \in \mathcal{L}_{v,\rho}$ if $C \in \mathcal{L}_{v,\rho}$. This is easy to see from the polynomial expansion of matrix exponential $e^C = \sum_{k=0}^{\infty} \frac{1}{k!} C^k$ and the fact that a Banach algebra is complete and closed under addition and multiplication.

In control theory, the Riccati equation,

$$ C^T P + PC - PBR^{-1}B^TP + Q = 0, \quad [S7] $$

and the Lyapunov equation,

$$ C^T P' + P'C + Q' = 0, \quad [S8] $$

are of central importance for system analysis and synthesis. The Lyapunov equation can be seen as a special case of the Riccati equation in the limit of $B^{-1}$ approaching zero. The fact that $\mathcal{L}_{v,\rho}$ is an inverse-closed Banach algebra has significant implications for the solutions of both equations. It has been shown in ref. (1) that, if the matrices $C$, $BR^{-1}B^T$, and $Q$ belong to an inverse-closed Banach algebra, then the unique stabilizing solution of the Riccati equation, Eq. $S7$, also belongs to the same algebra, as shown in ref. (2). Applying this result to $\mathcal{L}_{v,\rho}$ with the characteristic function $v(z) = e^{\alpha z^\beta} (1+z)^q$, we have the following result.

**Lemma 1** Assume that the characteristic function is $v(z) = e^{\alpha z^\beta} (1+z)^q$ with $\alpha > 0$, $0 < \beta < 1$, and $q > 1$. If $i)$ the matrices $C$, $BR^{-1}B^T$, and $Q$ belong to $\mathcal{L}_{v,\rho}$, $ii)$ the matrix pair $(C, B)$ is stabilizable, and $iii)$ the matrix pair $(C, Q^{1/2})$ is detectable, then the stabilizing solution $P$ of the Riccati equation (Eq. $S7$) belongs to $\mathcal{L}_{v,\rho}$. As a special case, if $Q' \in \mathcal{L}_{v,\rho}$ and $(C, Q^{1/2})$ is detectable, then the solution $P'$ of the Lyapunov equation (Eq. $S8$) also belongs to $\mathcal{L}_{v,\rho}$.

In this Lemma, $(C, B)$ is said to be stabilizable if the matrix $[C - \lambda I \ B]$ has full row rank for all $\Re \lambda \geq 0$, and $(C, Q^{1/2})$ is said to be detectable if the matrix $[C - \lambda I \ Q^{1/2}]$ has full column rank for all $\Re \lambda \geq 0$.

There are other notions of network locality developed in the literature, such as those in refs. (3, 4). However, those notions of locality are not developed for applications in network control and do not lead to the basic implications described in this section.

### 2. Network Data

The network data used in this paper is summarized in the following table.

| Networks                                      | Figures | References |
|-----------------------------------------------|---------|------------|
| KONECT dataset\(^a\)                          | Fig. 2  | (5)        |
| Advoqato, Air traffic control, arXiv astro-ph, arXiv cond-mat, arXiv hep-ph (citation), arXiv hep-ph (coauthor), arXiv hep-th (citation), arXiv hep-th (coauthor), Blogs, Brightkite, CAIDA, Chicago, Cora citation, DBLP, Digg, DNC co-recipient, DNC emails co-recipient, Enron, Epinions, Euroroad, Facebook (NIPS), Facebook friendships, Facebook wall posts, FOLDOC, Gnutella, Google.com internal, Google+, Hamsterster friendships, Hamsterster full, Human protein (Figeys), Human protein (Stelzl), Human protein (Vidal), Internet topology, Linux kernel mailing list replies, OpenFlights, Pretty Good Privacy, Protein, Reactome, Route views, Slashdot threads, Slashdot Zoo, Twitter lists, U. Rovira i Virgili, UC Irvine messages, US airports, US power grid, Wikibooks (fr), Wikinews (fr), Wikipedia elections, Wikipedia threads (de) | Fig. 4, Fig. 5D, Fig. 6A-C, Fig. 8 | (6) |
| Eastern U.S. power grid                       | Fig. 5E, Fig. 6A, Fig. 6C, Fig. 9 | (7) |
| Global air transportation network             | Fig. 5F, Fig. 6A, Fig. 6C, Fig. 10 | (8) |
| Whole brain network                           |         |            |

\(^a\)Network names reproduced as in the KONECT dataset.

For the KONECT dataset (5), we downloaded the edge list for each network that has $10^3$—$10^5$ nodes and is in one of the following categories: communication, social, online contact, infrastructure, computer, hyperlink, authorship, citation & coauthorship, and metabolic. From each list, we constructed the adjacency and Laplacian matrices of the (possibly directed and/or weighted) network. For these networks, unavailable edge weights are set to one, self-links are ignored, and the weights of parallel edges are combined. For the Eastern U.S. power grid, the network considered represents a snapshot of the summer of 2017 obtained from Federal Energy Regulatory Commission (FERC) (6). This network was analyzed using MATPOWER (10), a MATLAB-based power system analysis toolbox. The effective admittance matrix representing the coupling among generators was obtained from the full admittance matrix $Y_{\text{full}}$ through the Kron reduction (11) given by $Y = Y_{66} - Y_{6l} Y_{ll}^{-1} Y_{6l}$, where $Y_{66}$ and $Y_{ll}$ are the principal submatrices of $Y_{\text{full}}$ induced by the set of generator nodes and load nodes, respectively; $Y_{6l}$ ($Y_{ll}$) is the submatrix of $Y_{\text{full}}$ whose rows correspond to generator (load) nodes and whose columns correspond to load (generator) nodes. The steady state $(E^*, \delta^*)$ of the system was obtained by solving the power flow equation with the MATPOWER function `runpf`. Since the FERC data does not contain generator dynamic parameters, we sampled the values for the inertia constants.
\( \Gamma \), uniformly from the interval \([4, 8]\) p.u. and set \( D_i = 0.1 \) p.u. for each generator, which are typical values for these parameters \((12)\) in power systems. To create testing scenarios, we also assume that generators with output power between 0 and 200MW are renewable generation units. For the global air transportation network, the OpenFlights data \((7)\) provides information on 67663 airline routes, including the source airport, the destination airport, and the aircraft type designator. We identified the city served by each airport using either the International Air Transport Association (IATA) code or the International Civil Aviation Organization (ICAO) code provided in the data set, along with the mapping between the airport codes and city names available at [https://www.world-airport-codes.com](https://www.world-airport-codes.com). We also obtained the population and geographical location of the cities from the World Cities Database available at [https://simplemaps.com/data/world-cities](https://simplemaps.com/data/world-cities). We consider only the 2219 cities with a population larger than 10000, corresponding to a total population of 2.42 billion in this model. We obtained the seating capacity of each airplane type from the maker's website (as in the case of Airbus and Boeing) or other Internet sources; the seating capacity was then used to estimate the number of passengers in each flight (considering full occupancy for simplicity, which is an assumption that does not impact the qualitative results). The passenger flow network constructed connects all 2219 cities considered, where the entry \( C_{ijk} \) of the adjacency matrix \( C \) in the dynamical model represents the fraction of the population in city \( j \) that travel to city \( k \) on average per day. For the whole brain network, the coupling matrix \( W \) in ref. \((8)\) is readily available from the Brain Connectivity Toolbox website [https://sites.google.com/site/bctnet/datasets](https://sites.google.com/site/bctnet/datasets) (as Coactivation_matrix.mat). It represents the functional coactivation strengths among 638 similarly sized regions of a human brain. The parameters \( \alpha_h, \alpha_p, \beta, \) and \( \gamma \) of the associated dynamical model were obtained from ref. \((9)\).

3. Target Controllability of Localized Networks

In many practical problems of controlling dynamical networks, it is not necessary to control the state of all nodes; instead, the goal is to steer a target subset of nodes \( S \subseteq \mathcal{N} \) to desired states, regardless of the states of the other nodes. This was the motivation for the concept of target controllability \((13, 14)\). Let \( S \) be the projection matrix from the entire state space to that associated with the subset \( S \). The target subset \( S \) is said to be controllable if, for any initial state \( x_0 \) at \( t = t_0 \), final target-set state \( y_1 \), and finite time \( t_1 > t_0 \), there exists an input \( u \) that drives the system from \( x(t_0) = x_0 \) to \( x(t_1) = x_1 \) for some \( x_1 \) such that \( Sx_1 = y_1 \); that is, the nodes in \( S \) can be steered to the desired states. When \( S = \mathcal{N} \), target controllability reduces to the usual notion of controllability. It is shown in ref. \((14)\) that \( S \) is controllable if and only if the principal minor of the controllability Gramian indexed by \( S \), i.e., the matrix \( SW_c S^T \), is positive definite for any \( t > 0 \) (from this point on we assume that \( t_0 = 0 \)). It is straightforward to verify that, for any \( y \in \text{Null}(W_c) \), the control input

\[
\begin{align*}
    u(t) &= B^T e^{G(t-t_1)} \left( S^T W_c^1 S^T \right)^{-1} S x_1 + y, \quad 0 < t < t_1,
\end{align*}
\]

steers the system from \( x_0 = 0 \) to a state \( x(t_1) \) such that \( Sx(t_1) = y_1 \). In addition, the control input given by Eq. \((9)\) has the minimum possible control energy:

\[
\begin{align*}
    \int_0^{t_1} \| u(\tau) \|^2 d\tau &= (S x_1)^T \left( S W_c^1 S^T \right)^{-1} (S x_1).
\end{align*}
\]

In the special case of full controllability, \( S = \mathcal{N} \) and \( S = I \), and hence \( \int_0^{t_1} \| u(\tau) \|^2 d\tau = x_1^T (W_c^1)^{-1} x_1 \leq \lambda_{\text{min}}^{-1} (W_c^1) \cdot \| x_1 \|^2 \), where the equality is attained when \( x_1 \) becomes parallel with the eigenvector of \( W_c^1 \) corresponding to its smallest eigenvalue. This implies that the worst-case control energy is inversely proportional to \( \lambda_{\text{min}}(W_c^1) \). Therefore, the smallest eigenvalue of controllability Gramian can be considered a controllability measure: the larger the value of \( \lambda_{\text{min}}(W_c^1) \), the more controllable the system is.

Therefore, analogously to the case of full controllability, the worst-case minimum control energy needed to steer the target subset of nodes to the desired states is inversely proportional to the smallest eigenvalue of the projected Gramian \( SW_c^1 S^T \). Since the smallest eigenvalue of \( SW_c^1 S^T \) is upper-bounded by the smallest diagonal element of the matrix, we have

\[
\begin{align*}
    \lambda_{\text{min}}(SW_c^1 S^T) &\leq \kappa_{t_1} \eta^2 \cdot \min_{j \in S} \sum_{i \in \mathcal{D}} v\left( \rho(i, j) \right)^{-2} \\
    &\leq \kappa_{t_1} \eta^2 \cdot |D| \cdot \min_{j \in S} \{ v\left( \rho_H(D, \{ j \}) \right) \}^{-2} \\
    &= \kappa_{t_1} \eta^2 \cdot |D| \cdot \{ v\left( \rho_H(D, S) \right) \}^{-2},
\end{align*}
\]

where \( \rho_H(D, S) = \max_{j \in S} \min_{i \in \mathcal{D}} \rho(i, j) \). This inequality also establishes a crucial implication of locality for target controllability: the target subset of nodes \( S \) can be controlled using a smaller amount of energy if \( S \) lies closer to the driver set \( D \) in terms of the information distance \( \rho_H(\cdot, \cdot) \). This is a more general form of the result in Eq. \((5)\) of the main text.

4. Local Approximability of Controllability Measure

For notational convenience, we use \( W_c \) to denote \( W_c^1 \) for any \( t \in (0, +\infty) \) since the analysis below applies to both finite- and infinite-time Gramian matrices. Let \( W_c(\mathcal{N}, \mathcal{M}) \) be the submatrix of the controllability Gramian \( W_c \) whose rows and columns are induced by node sets \( \mathcal{N} \) and \( \mathcal{M} \), respectively. If \( \mathcal{M} \) is a singleton, i.e., \( \mathcal{M} = \{ i \} \), we simply write \( W_c(\mathcal{N}, i) \) (likewise, when \( \mathcal{N} \) is a singleton). Here, we show that, when a network is localized, the smallest eigenvalue \( \lambda_{\text{min}}(W_c) \) of
the entire Gramian $W_c$ can be well approximated by $\lambda_{\min}(W_c(N_i,N_i))$, where $W_c(N_i,N_i)$ is the principal submatrix of $W_c$ induced by an information neighborhood $N_i$ of some node $i$. Indeed, we show that there exist a node $i$, such that $\lambda_{\min}(W_c(N_i,N_i)) - \lambda_{\min}(W_c) = O(\varepsilon(\tau)^{-1})$.

Suppose that the controllability Gramian $W_c$ is localized (in addition to being symmetric and positive semi-definite by definition). Without loss of generality, we can assume that $\lambda_{\min}(W_c) = 0$; otherwise, we can instead consider $W_c - \lambda_{\min}(W_c)I$ since locality, symmetry, and semi-definiteness are all preserved under the subtraction of $\lambda_{\min}(W_c)I$. If there exists a diagonal block $W_c(i,i)$ of $W_c$ for which $\lambda_{\min}(W_c(i,i)) = 0$, the desired information neighborhood is trivially $N_i = \{i\}$. If not, we can show that the desired information neighborhood is that of a node $i$ satisfying the following condition.

**Condition 1** There exist a subset of nodes $M$ and a information radius $\hat{\tau} < \infty$ such that $i \in M \subseteq N_i(\hat{\tau})$, $W_c(M,M)$ is singular, and $W_c(M \setminus \{i\}, M \setminus \{i\})$ is non-singular.

We now show how this condition leads to the desirable information neighbourhood when $\lambda_{\min}(W_c(i,i)) > 0$ for all $i$. For any given $\tau \leq \hat{\tau}$, we define the set $\tilde{N}_i(\tau) = N_i(\tau) \cap M \setminus \{i\}$ and consider the following function of the radius $\tau$:

$$f(\tau) = \lambda_{\min}(W_c(i,i) - W_c(i,\tilde{N}_i(\tau))W_c(\tilde{N}_i(\tau),\tilde{N}_i(\tau))^{-1}W_c(\tilde{N}_i(\tau),i)),$$

i.e., the smallest eigenvalue of the entire matrix. Combined with Eq. S15, this implies $\lambda_{\min}(W_c(\tilde{N}_i(\tau),\tilde{N}_i(\tau))) = O(\varepsilon(\tau)^{-1})$, i.e., the smallest eigenvalue of the entire Gramian can be well approximated by the smallest eigenvalue of the projected Gramian in an information neighborhood of node $i$.

The analysis above relies on the existence of a node satisfying the three properties in Condition 1. We now show that such a node indeed exists by providing an iterative procedure to find it. We note that this procedure only serves to prove the validity of the assumption and is not intended as an efficient algorithm for practical use. The procedure is as follows. First, initialize $\mathcal{M}$ as $\mathcal{M} = N$.

1. Pick any $i \in \mathcal{M}$.

2. If $W_c(M \setminus \{i\}, M \setminus \{i\})$ is non-singular, stop and output $i$ and $\mathcal{M}$; otherwise, set $\mathcal{M} := \mathcal{M} \setminus \{i\}$ and go back to step 1.
The procedure always terminates in a finite number of steps, since \(|M|\) decreases by 1 at each iteration and \(W_c(M \setminus \{i\}, M \setminus \{i\})\) is guaranteed to be non-singular when \(M \setminus \{i\}\) contains only one node. From the assumption that \(W_c(N, N)\) is singular, it follows that the matrix \(W_c(M, M)\) is always singular. As a result, when the procedure terminates, it is guaranteed to output the desirable \(i\) and \(M\) that satisfy Condition 1. The parameter \(\tau\) can then be determined as \(\tau = \max_{j \in M} \rho(i, j)\).

Summarizing all of the above, given a localized network, we have shown that there exists a node \(i\) for which the smallest eigenvalue of the controllability Gramian can be well approximated by \(\lambda_{\text{min}}(W_c(N_i, N_i))\). In practice, it is generally difficult to determine the desirable node \(i\) \emph{a priori}. We thus consider instead the minimum among the smallest eigenvalues of all locally projected Gramians,

\[
\lambda_{\text{min}}(\tau) = \min_i \lambda_{\text{min}}(W_c(N_i, N_i)) \tag{[S18]}
\]

which enjoys the same convergence \(O(v(\tau)^{-1})\) due to the (guaranteed) existence of a special node \(i\) satisfying Condition 1.

5. Controllability Gramian of Diffusively Coupled Networks

A diffusively coupled network always has a zero eigenvalue, which causes the divergence of the integral defining the infinite-horizon controllability Gramian (i.e., \(W_c(t)\) for \(t \to \infty\)). This prevents the use of the Lyapunov equation to compute the infinite-horizon Gramian and its smallest eigenvalue, even though such a divergence does not affect the smallest eigenvalue itself. To overcome this issue, we perturb the system matrix as \(C - \epsilon I\) with a small \(\epsilon > 0\) to remove the zero eigenvalue from the system and consider the Lyapunov equation for the perturbed system,

\[
(C - \epsilon I)W_c(\tau) + W_c(\tau)(C - \epsilon I)^T + BB^T = 0, \tag{[S19]}
\]

which has a unique stabilizing solution and thus is solvable by the Bartels–Stewart algorithm (16). Once the solution \(W_c(\cdot)\) is obtained, we can project it to the orthogonal complement of the eigenvector \(v\) associated with the zero eigenvalue of \(C\). Denoting by \(\Pi\) an orthogonal basis for the orthogonal complement, we compute

\[
W_c^{\infty} = \Pi \Pi^T W_c(\tau) \Pi \Pi^T \tag{[S20]}
\]

as a projected version of the controllability Gramian. This \(W_c^{\infty}\) is insensitive to the choice of parameter \(\epsilon\) when \(\epsilon \ll |\text{Re}\lambda_1|\), where \(\lambda_1\) is the rightmost nonzero eigenvalue of \(C\). The matrix \(W_c^{\infty}\) calculated using Eq. [S20] approaches the projected infinite-horizon Gramian of the original system (Eq. 3 in the main text) as \(\epsilon\) approaches zero.

6. System-Level Synthesis

Before presenting the main result of the System-Level Synthesis (SLS) theory, we briefly introduce the control-theoretic set-up for signals and systems. The time-domain signals are assumed to be in \(L_2[0, +\infty)\), i.e., the space of square integrable functions supported on \(t \geq 0\) with inner product

\[
\langle F(t), E(t) \rangle = \int_{-\infty}^{+\infty} \text{Trace}[F^\dagger(\tau)E(\tau)] d\tau \tag{[S21]}
\]

and norm

\[
\|F(t)\|_{L_2} = \sqrt{\langle F(t), F(t) \rangle}, \tag{[S22]}
\]

where \(F^\dagger\) denotes the conjugate transpose of matrix \(F\). Through the Laplace transform

\[
F(s) = \int_{0}^{+\infty} F(t)e^{-st} dt, \tag{[S23]}
\]

one can see that the time-domain signal space \(L_2[0, +\infty)\) is isomorphic to \(H_2\). Here, \(H_2\) denotes the space of complex functions that are analytic in the open right-half plane \(\text{Re}(s) > 0\), which is endowed with inner product

\[
\langle F(s), E(s) \rangle = \frac{1}{2\pi i} \int_{-\infty}^{+\infty} \text{Trace}[F^\dagger(jw)E(jw)] dw \tag{[S24]}
\]

and norm

\[
\|F(s)\|_{H_2} = \sqrt{\langle F(s), F(s) \rangle}, \tag{[S25]}
\]

where \(j\) denotes the imaginary unit. The Plancherel theorem (2) states that \(\|F(t)\|_{L_2} = \|F(s)\|_{H_2}\), where \(F(s) \in H_2\) is the Laplace transform of \(F(t) \in L_2[0, +\infty)\). Linear dynamical systems can then be regarded as linear maps \(G\) from \(L_2[0, +\infty)\) to \(L_2[0, +\infty)\) or, equivalently, from \(H_2\) to \(H_2\). The induced norm is denoted by

\[
\|G\|_{H_\infty} = \text{ess sup}_{w \in \mathbb{R}} \sigma(G(jw)), \tag{[S26]}
\]
where $\sigma(\cdot)$ is the largest singular value and $\mathcal{H}_\infty$ denotes the associated normed space of all linear maps (2). Furthermore, the subspaces of $\mathcal{H}_2$ and $\mathcal{H}_\infty$ consisting of proper rational matrix functions (i.e., matrices whose elements are rational functions in the complex variable $s$) are denoted by $\mathcal{RH}_2$ and $\mathcal{RH}_\infty$, respectively. In addition, the strictly proper rational function subspaces of $\mathcal{RH}_2$ and $\mathcal{RH}_\infty$ are denoted by $1/\mathcal{RH}_2$ and $1/\mathcal{RH}_\infty$, respectively.

By the Laplace transform, the optimal control problem in Eq. 9 of the main text,

$$
\min_{u \in \mathcal{L}_2(0, \infty)} J = \int_0^\infty x(\tau)^T Q x(\tau) + u(\tau)^T R u(\tau) d\tau
$$

s.t. $x = C x + B u$, $x(0) = x_0,$

[S27]

can be equivalently written in the $s$-domain as

$$
\min_{u \in \mathcal{H}_2} \left\| \begin{bmatrix} Q^{1/2} & R^{1/2} \\ R^{1/2} & I \end{bmatrix} \begin{bmatrix} x(s) \\ u(s) \end{bmatrix} \right\|_{\mathcal{H}_2}^2
$$

s.t. $x(s) = (sI - C)^{-1}(B u(s) + x_0).$

[S28]

Given a feedback controller $K(s)$, the transfer function from the initial state $x_0$ to the state response $x(s)$ is given by $\Phi(s) = (sI - C - BK(s))^{-1}$, and the response of the controller output $u(s)$ is given by the transfer function $H(s) = K(s) (sI - C - BK(s))^{-1}$. The parameterization of all $\Phi(s)$ and $H(s)$ that are achievable by some stabilizing controller $K(s)$ is given by the following theorem.

**Theorem 1 (System-Level Parameterization for State Feedback Systems (17))** The following are true for any linear system $\dot{x} = Cx + Bu$: 1) The affine space defined by

$$
\begin{bmatrix} sI - C & -B \\ H(s) \end{bmatrix} = I, \; \Phi, H \in \frac{1}{s} \mathcal{RH}_\infty
$$

[S29]

parameterizes all system responses from $x_0$ to $x(s)$ and control responses from $x_0$ to $u(s)$ that are achievable by an internally stabilizing state feedback controller; 2) For all transfer matrices $\{\Phi(s), H(s)\}$ satisfying Eq. S29, the controller $K(s) = H(s) \Phi(s)^{-1}$ is internally stabilizing and achieves the desired responses $x(s) = \Phi(s) x_0$ and $u(s) = H(s) x_0$.

This theorem shows that there is a bijection between the stabilizing controllers and the responses in the affine space defined by Eq. S29. Therefore, instead of directly designing the feedback control law $K(s)$, we can equivalently design the transfer matrices $\Phi(s)$ and $H(s)$ under the affine constraint given by Eq. S29. With this parameterization, the optimal control problem can be rewritten as

$$
\min_{\Phi, H \in \frac{1}{s} \mathcal{RH}_\infty} J = \left\| \begin{bmatrix} Q^{1/2} & R^{1/2} \\ R^{1/2} & I \end{bmatrix} \begin{bmatrix} \Phi(s) \\ H(s) \end{bmatrix} \begin{bmatrix} x_0 \end{bmatrix} \right\|_{\mathcal{H}_2}^2
$$

s.t. $\begin{bmatrix} sI - C & -B \end{bmatrix} \begin{bmatrix} \Phi(s) \\ H(s) \end{bmatrix} = I.$

[S30]

The solution of this problem for any given $x_0$ is given by

$$
\Phi(s) = (sI - C + BR^{-1}B^T P)^{-1},
$$

[S31]

$$
H(s) = -R^{-1}B^T P (sI - C + BR^{-1}B^T P)^{-1},
$$

[S32]

where $P$ is the solution of the Riccati equation, Eq. S7. The optimal controller thus becomes a static feedback law and is given by

$$
K(s) = H(s) \Phi(s)^{-1} = -R^{-1}B^T P.
$$

[S33]

We note that $\Phi(s)$ and $H(s)$ in Eqs. S31 and S32, respectively, are independent of $x_0$ and hence serve as the “fundamental solution” of the optimal control problem in Eq. S30 for arbitrary initial conditions. The columns of $\Phi(s)$ and $H(s)$ can be computed independently. To see this, we set $x_0 = e_j$ in Eq. S30, rewrite the problem as the minimization over $\phi_j(s) = \Phi(s) e_j$ and $h_j(s) = H(s) e_j$, and eliminate the unnecessary constraints corresponding to all but the $j$th columns of $\Phi(s)$ and $H(s)$. This leads to Eq. 10 in the main text.

7. Disturbance-Oriented Localization

By projecting the original optimal control problem in Eq. 10 of the main text onto the information neighborhood $N_j$ of node $j$, we have the following projected optimization problem:

$$
\min_{\phi_j, h_j \in \frac{1}{s} \mathcal{RH}_\infty} \left\| \begin{bmatrix} \tilde{Q}^{1/2} \\ \tilde{R}^{1/2} \end{bmatrix} \begin{bmatrix} \phi_j(s) \\ h_j(s) \end{bmatrix} \begin{bmatrix} \bar{e}_j \end{bmatrix} \right\|_{\mathcal{H}_2}^2
$$

s.t. $\begin{bmatrix} sI - \tilde{C}_j & -\tilde{B}_j \end{bmatrix} \begin{bmatrix} \phi_j(s) \\ h_j(s) \end{bmatrix} = \bar{e}_j.$

[S34]
To solve this projected problem, we invoke the inverse Laplace transform to go back to the time domain, which transforms the problem in Eq. S34 into

$$\min_{u_j \in L_2(0, \infty)} \tilde{J} = \int_0^{\infty} \left[ \tilde{x}_j(\tau)^T \tilde{Q}_j \tilde{x}_j(\tau) + \tilde{u}_j(\tau)^T \tilde{R}_j \tilde{u}_j(\tau) \right] d\tau$$

subject to

$$\tilde{x}_j = \tilde{C} \tilde{x}_j + \tilde{B} \tilde{u}_j, \quad \tilde{x}_j(0) = \tilde{c}_j^T x_0.$$  

This is simply a projected version of the optimal control problem defined by Eq. S27. The optimal control law is given by the static feedback matrix

$$\tilde{K}_j = -\tilde{R}_j^{-1} \tilde{B}_j^T \tilde{P}_j,$$

where \(\tilde{P}_j\) is the solution to the Riccati equation

$$\tilde{C}_j^T \tilde{P}_j + \tilde{P}_j \tilde{C}_j - \tilde{P}_j \tilde{B}_j \tilde{R}_j^{-1} \tilde{B}_j^T \tilde{P}_j + \tilde{Q}_j = 0.$$  

The \(s\)-domain solution corresponding to the problem in Eq. S34 is then given by

$$\tilde{\phi}_j(s) = (sI - \tilde{C}_j - \tilde{B}_j \tilde{K}_j)^{-1} \tilde{e}_j,$$

$$\tilde{h}_j(s) = \tilde{K}_j \tilde{\phi}_j(s).$$

After solving the projected problem for all \(1 \leq j \leq N\), we can construct

$$\tilde{\Phi}(s) = \begin{bmatrix} N_1^T \tilde{\phi}_1(s) & N_2^T \tilde{\phi}_2(s) & \cdots & N_n^T \tilde{\phi}_n(s) \end{bmatrix},$$

$$\tilde{H}(s) = \begin{bmatrix} T_1^T \tilde{h}_1(s) & T_2^T \tilde{h}_2(s) & \cdots & T_n^T \tilde{h}_n(s) \end{bmatrix}.$$  

From Theorem 1, it then follows that the overall optimal control law takes the form

$$u(s) = \tilde{K}(s)x(s) = \tilde{H}(s)\tilde{\Phi}(s)^{-1}x(s).$$

Crucially, we now show that the controller \(\tilde{K}(s)\) is guaranteed to be stabilizing and its associated control objective value approaches that of the global controller when the system is sufficiently localized. That is, the controller designed based on the projected model, \(\tilde{K}(s)\), will enjoy certain stability and near-optimal performance guarantees when implemented on the original system. Due to the locality of the system, in view of Eq. S29, the following equality shows that the projected problem can be regarded as a perturbed version of the original one:

$$(sI - C)N_j^T \tilde{\phi}_j(s) - B \tilde{H}_j(s) - \epsilon_j = (I - N_j^T N_j)(-C)N_j^T \tilde{\phi}_j(s) := \epsilon_j(s).$$

The perturbation term \(\epsilon_j(s)\) is expected to be very small in magnitude when \(N_j\) is sufficiently large. This is the case because: 1) the non-zeros elements of vector \(N_j^T \tilde{\phi}_j(s)\) are limited to those corresponding to the neighbourhood \(N_j\), and their magnitude decays as \(O(\nu(\tau)^{-1})\) with the information distance \(\tau\); 2) the multiplication by \(-C\) preserves the decay pattern due to the locality of the system; and 3) the further multiplication by \((I - N_j^T N_j)\) turns the elements inside the information neighborhood \(N_j\) into zeros and leaves only negligible elements outside \(N_j\). That is, the solution of the projected problem (Eq. S34 with mismatch \(\epsilon_j(s)\)) when lifted to the original space is an approximate solution of the original problem in Eq. S30. Concatenating Eq. S43 for all \(1 \leq j \leq n\) and accounting for Eq. S40 and Eq. S41, we have

$$(sI - C)\tilde{\Phi}(s) - B\tilde{H}(s) = I + \Sigma(s),$$

where \(\Sigma(s) = [\epsilon_1(s), \epsilon_2(s), \cdots, \epsilon_n(s)]\). If the perturbation \(\Sigma(s)\) is sufficiently small such that

$$(I + \Sigma(s))^{-1} \in \frac{1}{8} \mathcal{R}_{H,\infty},$$

then the matrices

$$\tilde{\Phi}'(s) = \tilde{\Phi}(s)(I + \Sigma(s))^{-1},$$

$$\tilde{H}'(s) = \tilde{H}(s)(I + \Sigma(s))^{-1}$$

would satisfy exactly the condition given by Eq. S29 in Theorem 1. Hence, Eqs. S46 and S47 constitute achievable system and control responses with the controller \(\tilde{K}(s) = H'(s)\tilde{\Phi}'(s)^{-1} = H(s)\tilde{\Phi}(s)^{-1}\), which stabilizes the original system even though this controller is designed based on the projected model.

A sufficient condition for Eq. S45, according to the small-gain theorem (18, 19), is given by any of the following:

$$\|\Sigma\|_{\mathcal{H},\infty} < 1,$$
This shows that, when the system is more localized, where we define $\epsilon$ and $\|\cdot\|_{L_1}$ denotes the $L_1$ system norm. The $L_1$ system norm is defined through the impulse response in the time domain as

$$
\|\Sigma^T(t)\|_{L_1} = \max_{1 \leq i \leq n} \sum_{i=1}^{n} \int_{0}^{\infty} |\epsilon_{ij}(t)|dt,
$$

where $\epsilon_{ij}(t)$ is the $i$th component of the perturbation (column) vector $\epsilon_j$ in Eq. S43 under inverse Laplace transform. The computation required to verify the small-gain condition in Eq. S50 can be distributed over the columns of the impulse response matrix $\Sigma(t)$, since Eq. S50 is equivalent to

$$
\sum_{i=1}^{n} \int_{0}^{\infty} |\epsilon_{ij}(t)|dt < 1, \quad \forall 1 \leq j \leq n.
$$

Once the controller defined by Eq. S42 stabilizes the system, the next question is to determine the extent to which the dynamical performance of this controller compares with that of the theoretical global optimal control. Let $J^*$ be the optimal objective value for the original optimal control problem in Eq. S30, and define

$$
\tilde{J}^* = \left\| Q^{1/2} \bar{\Phi}(s) \right\|_{\mathcal{H}_2} + \left\| R^{1/2} \bar{H}(s) \right\|_{\mathcal{H}_2},
$$

i.e., the optimal objective value obtained from the solutions of the projected problems in Eq. S34. Since $\bar{\Phi}(s)$ and $\bar{H}'(s)$ form a feasible solution of the optimal control problem in Eq. S30, we have an upper bound for $J^*$:

$$
J^* \leq \left\| Q^{1/2} \bar{\Phi}'(s) \right\|_{\mathcal{H}_2} + \left\| R^{1/2} \bar{H}'(s) \right\|_{\mathcal{H}_2} \leq \left\| Q^{1/2} \bar{\Phi}(s) \right\|_{\mathcal{H}_2} + \| I + \Sigma(s) \|_{\mathcal{H}_\infty} + \left\| R^{1/2} \bar{H}(s) \right\|_{\mathcal{H}_2} \| I + \Sigma(s) \|_{\mathcal{H}_\infty} \leq \tilde{J}^* \| I + \Sigma(s) \|_{\mathcal{H}_\infty} \leq \tilde{J}^* \| I + \Sigma(s) \|_{\mathcal{H}_\infty}. \tag{S54}
$$

To derive a lower bound for $J^*$, we interchange the roles of the original and projected systems in Eq. S43 and obtain

$$(sI - C)N_j \phi_j(s) - BT_j h_j(s) - \bar{e}_j = -N_j C(N_j^T N_j - I) \phi_j(s) := \xi_j(s), \tag{S55}$$

which implies

$$(sI - C) \Phi(s) - BH(s) = I + \Xi(s), \tag{S56}$$

where we define $\Xi(s) = [\xi_1(s), \xi_2(s), \ldots, \xi_n(s)]$. Due to the locality, the perturbation term can be small enough such that $(I + \Xi(s))^{-1} \in \frac{1}{\rho} \mathcal{R}_{\mathcal{H}_\infty}$, which is guaranteed by any of the inequalities in Eqs. S48–S50. Under this condition, $\Phi(s)(I + \Xi(s))^{-1}$ and $H(s)(I + \Xi(s))^{-1}$ form a feasible solution of the projected problem in Eq. S34. This implies

$$
\tilde{J}^* \leq \left\| Q^{1/2} \Phi(s)(I + \Xi(s))^{-1} \right\|_{\mathcal{H}_2} + \left\| R^{1/2} H(s)(I + \Xi(s))^{-1} \right\|_{\mathcal{H}_2} \leq \left( \left\| Q^{1/2} \Phi(s) \right\|_{\mathcal{H}_2} + \left\| R^{1/2} H(s) \right\|_{\mathcal{H}_2} \right) \| (I + \Xi(s))^{-1} \|_{\mathcal{H}_\infty} = J^* \| (I + \Xi(s))^{-1} \|_{\mathcal{H}_\infty} \leq \frac{J^*}{1 - \| \Xi(s) \|_{\mathcal{H}_\infty}}. \tag{S57}
$$

Combining the inequalities in Eqs. S54 and S57, we have

$$
\tilde{J}^* (1 - \| \Xi(s) \|_{\mathcal{H}_\infty}) \leq J^* \leq \frac{\tilde{J}^*}{1 - \| \Xi(s) \|_{\mathcal{H}_\infty}}. \tag{S58}
$$

This shows that, when the system is more localized, $\tilde{J}^* \to J^*$ as $\Sigma(s) \to 0$ and $\Xi(s) \to 0$. Thus, when the residual terms $\Sigma(s)$ and $\Xi(s)$ are small in magnitude, the optimal objective value for the projected problem is close to that of the original problem, meaning that solving the projected problem provides a near-optimal solution to the original problem.
8. Controller-Oriented Localization

Although the controller in Eq. S42 obtained by the disturbance-oriented localization is guaranteed to be stabilizing and near-optimal, this controller is a dynamic feedback controller, whereas the global optimal controller given by Eq. S33 is static. As a consequence, the controller given by Eq. S42 is a dynamical system for each driver node with the same state-space dimension as the entire network system, which makes the controller impractical for large networks. Here, we show how to convert the dynamical controller into static ones that are scalable to large networks.

To understand why the localized solutions in Eqs. S40 and S41 lead to dynamical controllers and how they can be converted into static ones, we take the viewpoint of the controller installed at each node individually. Recall that the projected problem in Eq. S34 at node \( j \) is obtained by projecting the original problem onto the information neighborhood \( \mathcal{N}_j \). Also recall that, since the controller at node \( i \) responds to disturbances at all node \( j \) for which \( i \in \mathcal{N}_j \), we define the set of all such nodes \( j \) as the control neighborhood of \( i \):

\[
    \mathcal{C}_i = \{ 1 \leq j \leq N \mid i \in \mathcal{N}_j \}. \tag{S59}
\]

By using the projection matrix \( f_i^T \) defined in the main text, the controller installed at the \( i \)th node is \( f_i^T \tilde{K}(s) \), which is a linear map from the entire state space to the input of that node. The nonzero columns of \( f_i^T \tilde{K}(s) \) correspond to nodes belonging to \( \mathcal{C}_i \). From Eq. S42, we have

\[
f_i^T \tilde{K}(s) \Phi(s) = f_i^T \tilde{H}(s), \tag{S60}
\]

which is equivalent to

\[
f_i^T \tilde{K}(s) \mathcal{N}_j^T \tilde{\phi}_j(s) = f_i^T \tilde{T}_j^T \tilde{K}_j \tilde{\phi}_j(s), \quad \forall \ j \in \mathcal{C}_i, \tag{S61}
\]

as verified using Eqs. S39–S41. Suppose that the set of feedback matrices \( \{ \tilde{K}_j \}_{j=1}^N \) for the projected problem satisfies

\[
f_i^T \tilde{T}_j^T \tilde{K}_j \mathcal{N}_j e_p = f_i^T \tilde{T}_k^T \tilde{K}_j \mathcal{N}_k e_p, \quad \forall \ j, k \in \mathcal{C}_i, \ p \in \mathcal{N}_j \cap \mathcal{N}_k. \tag{S62}
\]

That is, for any given node \( p \), the feedback matrix from the state of node \( p \) to the control input of node \( i \) take the same value for all node pairs \( (j, k) \) in \( \mathcal{C}_i \) whose information neighborhoods \( \mathcal{N}_j \) and \( \mathcal{N}_k \) both contain node \( p \). We denote this common matrix \( f_i^T \tilde{T}_j^T \tilde{K}_j \mathcal{N}_j e_p \) in Eq. S62 as \( \tilde{K}_{ip} \). It follows that

\[
f_i^T \tilde{K}(s) = \sum_{j \in \mathcal{C}_i} \tilde{K}_{ij} e^*_j \tag{S63}
\]

satisfies Eq. S60. Thus, when the condition in Eq. S62 holds, the controller at node \( i \) becomes a static feedback matrix.

However, since \( \tilde{K}_j \) is determined independently for each \( j \in \mathcal{C}_i \) by solving the projected Riccati equation (Eq. S37), the condition in Eq. S62 is not guaranteed to hold. To ensure that condition is satisfied by the set of feedback matrices \( \{ \tilde{K}_j \}_{j=1}^N \), we modify the way in which the projected problems defined by Eq. S34 are formulated. To do this, we merge \( \mathcal{N}_j \) for all \( j \in \mathcal{C}_i \) to form the set

\[
    \mathcal{\hat{N}}_i = \bigcup_{j \in \mathcal{C}_i} \mathcal{N}_j, \tag{S64}
\]

which is a superset of the information neighborhood of any \( j \in \mathcal{C}_i \). Now, if we project the original problem onto a superset of the information neighborhood, the quality of the approximation of the projected problem would not be compromised. This gives an extra degree of freedom to construct the projected problems, so that the condition in Eq. S62 can be satisfied. Let \( \mathcal{\hat{N}}_i \) be the projection matrix from the entire state space to the state subspace corresponding to nodes in \( \mathcal{\hat{N}}_i \). In analogy with \( \mathcal{T}_i \) and \( \mathcal{T}_i \) in the main text, we introduce

\[
    \mathcal{\hat{T}}_i = \{ 1 \leq k \leq r \mid |B|_{jk} \neq 0 \text{ for some } j \in \mathcal{\hat{N}}_i \} \tag{S65}
\]

and define \( \mathcal{\hat{T}}_i \) as the projection matrix from the entire input space \( \mathbb{R}^r \) to the input subspace \( \mathbb{R}^{\mathcal{\hat{T}}_i} \) associated with \( \mathcal{T}_i \), where \( \mathcal{L} = |\mathcal{T}_i| \). We can then project the original problem onto \( \mathcal{\hat{N}}_i \) to obtain a new set of projected problems, one for each \( j \in \mathcal{C}_i \):

\[
    \min_{\tilde{\phi}_j, \tilde{h}_j} \left\| \begin{bmatrix} \tilde{Q}_j^{1/2} & \tilde{R}_j^{1/2} \end{bmatrix} \begin{bmatrix} \tilde{\phi}_j(s) \\ \tilde{h}_j(s) \end{bmatrix} \right\|_{\mathcal{H}_2}^2 \tag{S66}
\]

s.t. \[
    \begin{bmatrix} sI - \tilde{C}_j & -\tilde{B}_j \\ \tilde{C}_j - \tilde{C}_i \end{bmatrix} \begin{bmatrix} \tilde{\phi}_j(s) \\ \tilde{h}_j(s) \end{bmatrix} = \tilde{e}_j,
\]

where \( \tilde{\mathcal{C}}_i = \mathcal{\hat{N}}_i C \mathcal{\hat{N}}_i^T, \tilde{B}_i = \mathcal{\hat{N}}_i B \mathcal{\hat{T}}_i^T, \tilde{Q}_i = \mathcal{\hat{N}}_i Q \mathcal{\hat{N}}_i^T, \tilde{R}_i = \mathcal{T}_i R \mathcal{T}_i^T, \) and \( \tilde{e}_j = \mathcal{\hat{N}}_i e_j \). We call this set of projected problems the controller-oriented localization of the original problem, since each is a local version of the original problem around a driver node. The solution to Eq. S66 is obtained by solving the Riccati equation

\[
    \tilde{C}_i^T \tilde{P}_i + \tilde{P}_i \tilde{C}_i - \tilde{P}_i \tilde{B}_i \mathcal{\hat{T}}_i^{-1} \mathcal{\hat{T}}_i^T \tilde{B}_i^T \tilde{P}_i + \tilde{Q}_i = 0. \tag{S67}
\]
which yields the optimal feedback $\hat{K}_j = -\hat{R}_j^{-1}\hat{B}_j^T\hat{P}_j$. Hence, for each $j \in \mathcal{C}_i$,

$$\hat{\phi}_j(s) = (sI - \hat{C}_j - \hat{B}_j\hat{K}_j)^{-1}\hat{\epsilon}_j,$$

$$\hat{h}_j(s) = \hat{K}_j\hat{\phi}_j(s).$$  \[S68\]

Compared to Eqs. S38 and S39, the optimal response in Eqs. S68 and S69 is obtained by projecting the original problem onto a superset of the information neighborhood $\mathcal{N}_j$, and hence this solution provides at least the same accuracy as the solution in Eqs. S38 and S39. The resulting controller is static because

$$\hat{K}_j = T_j\bar{T}_j^T\hat{K}_j\hat{N}_j\mathcal{N}_j^T$$  \[S70\]

satisfies the condition in Eq. S62. The corresponding responses in $\mathcal{N}_j$ achieved by $\hat{K}_j$ are given by

$$\hat{\phi}_j(s) = N_j\hat{N}_j^T\hat{\phi}_j(s),$$  \[S71\]

$$\hat{h}_j(s) = T_j\bar{T}_j^T\hat{h}_j(s).$$  \[S72\]

It follows that

$$(f_j^T\hat{T}_j^T\hat{K}_j\hat{N}_j)\mathcal{N}_j^T\hat{\phi}_j(s) = (f_j^T\bar{T}_j^T\hat{K}_j\hat{N}_j)\mathcal{N}_j^T\hat{\phi}_j(s) = f_j^T\bar{T}_j^T\hat{K}_j\hat{\phi}_j(s),$$  \[S73\]

i.e., $f_j^T\hat{K}(s) = f_j^T\bar{T}_j^T\hat{K}_j\hat{N}_j$ satisfies Eq. S61. This implies that $f_j^T\bar{T}_j^T\hat{K}_j\hat{N}_j$ is the appropriate feedback matrix for the controller at node $i$, and thus

$$f_j^T u(t) = f_j^T\bar{T}_j^T\hat{K}_j\hat{N}_j x(t).$$  \[S74\]

Therefore, the design of the feedback law of each controller only requires solving the projected Riccati equation locally around that driver node and the computation can be performed in parallel for all drivers across the network. This provides a decentralized method for designing a near-optimal control strategy for the entire network.

9. Basic Control Tasks and Linearization Methods

In scientific and engineering applications, one is often required to actively control the dynamics of complex networks so that they exhibit certain desirable behaviors and functionalities. Different applications require addressing different control tasks, and the most often encountered in practice are equilibrium stabilization, trajectory tracking, and command following. All these control tasks can be accomplished by proper design of feedback control laws, as described next. To proceed, we consider a general system

$$\dot{x} = f(x) + Bu$$  \[S75\]

in which the function $f(\cdot)$ can be nonlinear.

Equilibrium stabilization refers to the control task in which the system is driven from a given initial condition to a desired equilibrium and held stably there. In power grids, for example, each power flow solution corresponds to an equilibrium of the system. A major task of power system controllers is to bring the system towards the most efficient and reliable power flow equilibrium and to maintain the system at the equilibrium in the presence of disturbances. Consider the system in Eq. S75 with a desired equilibrium $(x^*, u^*)$, i.e., $f(x^*) + Bu^* = 0$. The problem is to design a feedback law of the form $u(t) = K(x(t) - x^*) + u^*$ that drives the system from the initial state $x_0$ to the target equilibrium $x^*$. In general, the map $K(\cdot)$ can be nonlinear but is assumed to be homogeneous, i.e., $K(\theta x) = \theta K(x)$. By defining $\Delta x = x - x^*$ and $\Delta u = u - u^*$, we have $\Delta \dot{x} = f(x^* + \Delta x) - f(x^*) + B\Delta u$ (throughout the Article we use $\Delta x$ to denote the time derivative of variable $\Delta x$ even in cases where $x^*$ is time dependent). This reduces the original problem to the problem of designing a control $\Delta u$ to stabilize the system at the origin. To apply the linear-quadratic optimal control theory, we replace the nonlinear term $f(x^* + \Delta x) - f(x^*)$ with $C(x, x^*)\Delta x$, which is linear in $\Delta x$. Linearization methods to obtain $C(x, x^*)$ are presented at the end of this section. Accordingly, the problem further reduces to $\Delta \dot{x} = C(x, x^*)\Delta x + B\Delta u$, which shares the same structure as the linear system in Eq. 3 of the main text. Thus, Algorithm 3 (in Materials and Methods) can be used to design $K$ for this problem, which can be time independent or time varying, depending on the linearization method employed.

Trajectory tracking aims to drive the system to a given feasible trajectory and then force it to follow the trajectory stably. A feasible trajectory of the system is a pair of curves $(x^*(t), u^*(t))$ that satisfy Eq. S75. For example, it can be the set of planned optimal trajectories for a group of autonomous spacecrafts in formation, or a selected common orbit for a group of periodic oscillators in a synchronous state. The task is to design a feedback control $u(t) = K(x(t) - x^*) + u^*$ such that the system starting from an given initial condition converges to and then stays on the target trajectory. The design approach using Algorithm 3 is essentially the same as for equilibrium stabilization, except that $(x^*(t), u^*(t))$ is a time-varying target.

Command following differs from the two tasks above in that a desired equilibrium or feasible trajectory is not known a priori. The goal is to design a control law in which a subset $x_1$ of the state vector elements will quickly follow any (possibly time-varying) control command $r$ given in real time. Thus, $r$ is unspecified at the design stage and the state vector can be written as $x = [x_1^T, x_2^T]^T$, where $x_2$ represents the other elements of the state vector. In order to completely track the command $r$, there must exist an equilibrium $(x^*, u^*)$ such that $f(x^*) + Bu^* = 0$, $x_1^* - r = 0$. Because $r$ is unknown a priori, $(x^*, u^*)$ is
also not available in advance and therefore the controller cannot be designed against a known \((x^*, u^*)\) as done above. This problem is solved by augmenting the system with an internal state of the controller \(z\), representing the integral of the error in following \(r\). The augmented system reads \(\dot{x} = f(x) + Bu, \dot{z} = x_1 - r\). We seek to design a feedback control law in the form 
\[
 u = K_1(x_1 - r) + K_2x_2 + K_3z,
\]
which can be seen as a high-dimension generalization of the proportional-integral control widely used in industrial applications \((21)\). The closed-loop system is then given by \(\dot{x} = f(x) + BK_1(x_1 - r) + BK_2x_2 + BK_3z, \dot{z} = x_1 - r\), and the equilibrium equation is rewritten as \(f(x^*) + BK_1x_1^* + BK_2x_2^* + BK_3z^* = 0, x_1^* - r = 0\). Defining \(\Delta x_1 = x_1 - r, \Delta x_2 = x_2 - x_2^*, \Delta z = z - z^*\), we have \(\Delta \dot{x} = f(x) - f(x - \Delta x) + BK_1\Delta x_1 + BK_2\Delta x_2 + BK_3\Delta z, \Delta z = \Delta x_1\). By linearizing \(f(x) - f(x - \Delta x)\) as \(C(x)\Delta x\), the system becomes
\[
\begin{bmatrix}
\Delta x_1 \\
\Delta x_2 \\
\Delta z
\end{bmatrix} =
\begin{bmatrix}
C(x) & 0 & 0 \\
I & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
\Delta x_1 \\
\Delta x_2 \\
\Delta z
\end{bmatrix} +
\begin{bmatrix}
B \\
0 \\
0
\end{bmatrix}
\begin{bmatrix}
K_1 \\
K_2 \\
K_3
\end{bmatrix}
\begin{bmatrix}
\Delta x_1 \\
\Delta x_2 \\
\Delta z
\end{bmatrix},
\]
\[\text{[S76]}\]

independently of the unknown equilibrium. The feedback law \([K_1 K_2 K_3]\) can then be designed according to Algorithm 3.

All three control tasks just described rely on linearization of \(f(x^* + \Delta x) - f(x^*)\) or, equivalently, of \(f(x) - f(x - \Delta x)\). The most basic linearization approach is the Jacobian linearization, in which the first-order Taylor expansion around a known equilibrium or feasible trajectory is used as an approximation: \(f(x^* + \Delta x) - f(x^*) \approx \frac{\partial f}{\partial x}\big|_{x^*}\Delta x\). When the linearized system is stable, the approach guarantees that the original nonlinear system has a non-empty basin of attraction around \(x^*\), which can be enlarged by using high-gain feedback. Moreover, the feedback matrix does not change while the system is controlled (unless \(x^*\) varies in time), which is computationally appealing. Another common approach is the so-called extended linearization, in which the nonlinear function is algebraically factorized as \(f(x^* + \Delta x) - f(x^*) = C(x^*, \Delta x)x\). The specific form of \(C(x^*, \Delta x)\) depends on the problem formulation and is in general not unique. An existence condition and a complete parameterization of all such factorizations can be found in \((22, 23)\). Due to the dependence of the linear coefficient matrix on the state, different Riccati equations are solved at different time steps to obtain a time-varying feedback law (owing to this, the method is also called the state-dependent Riccati equation approach). The appeal of this approach is that it not only guarantees local stability but also generates a system trajectory that satisfies the Hamilton–Jacobi–Bellman equation, which is a necessary condition for the optimality of the trajectory \((24)\). For the command following problem, which requires a linearization coefficient matrix that does not depend on the unknown equilibrium \(x^*\), a common practice is to first use the extended linearization to obtain \(f(x) = C(x)x\) and then approximate the nonlinear term as \(f(x) - f(x - \Delta x) \approx C(x)\Delta x\). Despite a certain lack of theoretical justification, this approximation has been successfully used in many practical problems \((25, 26)\).

References
1. R. Curtain, Riccati equations on noncommutative Banach algebras. SIAM J. Control. Optim. 49, 2542–2557 (2011).
2. K. Zhou, J.C. Doyle, K. Glover, Robust and Optimal Control (Prentice Hall, New Jersey, 1996).
3. S. Bartolucci, F. Caravelli, F. Caccioli, P. Vivo. Emerging locality of network influence. arXiv:2009.06307 (2020).
4. A. Favero, F. Cagnetta, M. Wyart. Locality defeats the curse of dimensionality in convolutional teacher-student scenarios. Adv. Neural Inf. Process. Syst. Conf. (2021).
5. J. Kunegis, KONECT: The Koblenz network collection in Proceedings of the 22nd International Conference on World Wide Web. (2013). Available at http://konect.cc.
6. Federal Energy Regulatory Commission Form 715 (2017). Data obtained under a non-disclosure agreement by following the procedure described at https://www.ferc.gov/legal/ceiifoia/cei.asp.
7. Openflights (2014). Available at https://openflights.org/data.html.
8. N.A. Crossley, et al., Cognitive relevance of the community structure of the human brain functional coactivation network. Proc. Natl. Acad. Sci. U.S.A. 110, 11583–11588 (2013). Available from the Brain Connectivity Toolbox website at https://sites.google.com/site/bctnet/datasets-and-demos.
9. L.M. Sanchez-Rodriquez, et al., Design of optimal nonlinear network controllers for Alzheimer’s disease. PLoS Comput. Biol. 14, e1006136 (2018).
10. R.D. Zimmerman, C.E. Murillo-Sánchez, R.J. Thomas, MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. IEEE T. Power Syst. 26, 12–19 (2010).
11. J. Machowski, J. Bialek, J. Bumby, Power System Dynamics: Stability and Control (John Wiley & Sons, 2011).
12. P. Kundur, Power System Stability and Control (McGraw-Hill, 1994).
13. J. Gao, Y.Y. Liu, R.M. D’Souza, A.L. Barabási. Target control of complex networks. Nat. Commun. 5, 5415 (2014).
14. B. Zhao, Y. Guan, L. Wang. Non-fragility and partial controllability of multi-agent systems. arXiv:1607.07753 (2016).
15. R.A. Horn, C.R. Johnson, Matrix Analysis (Cambridge University Press, 2012).
16. R.H. Bartels, G.W. Stewart. Solution of the matrix equation \(AX + XB = C\). Commun. ACM 15, 820–826 (1972).
17. Y.S. Wang, N. Matni, J.C. Doyle, A system-level approach to controller synthesis. IEEE T. Autom. Contr. 64, 4079–4093 (2019).
18. G.E. Dullerud, F. Paganini, A Course in Robust Control Theory: A Convex Approach (Springer Science & Business Media, 2013).
19. J. Anderson, J.C. Doyle, S.H. Low, N Matni, System level synthesis. Annual Reviews in Control 47, 364–393 (2019).
20. J. Zhu, J. Chen, Stability of systems with time-varying delays: An \(L_1\) small-gain perspective. Automatica 52, 260–265 (2015).
21. K.J. Åström, T. Hägglund, *PID Controllers: Theory, Design, and Tuning* (Instrument society of America Research Triangle Park, 1995).

22. Y.W. Liang, L.G. Lin, Analysis of SDC matrices for successfully implementing the SDRE scheme. *Automatica* 49, 3120–3124 (2013).

23. L.G. Lin, J. Vandewalle, Y.W. Liang, Analytical representation of the state-dependent coefficients in the SDRE/SDDRE scheme for multivariable systems. *Automatica* 59, 106–111 (2015).

24. T. Çimen, State-dependent Riccati equation (SDRE) control: A survey. *IFAC Proc.* 41, 3761–3775 (2008).

25. J.R. Cloutier, D.T. Stansbery, “The capabilities and art of state-dependent Riccati equation-based design” in *Proc. American Control Conference* (IEEE, 2002), pp. 86–91.

26. T. Çimen, Systematic and effective design of nonlinear feedback controllers via the state-dependent Riccati equation (SDRE) method. *Annu. Rev. Control* 34, 32–51 (2010).
Fig. S1. The scaling of the off-diagonal decay constant for the inverse of matrix $C$. Here, we consider $C = -L + I$, where $L$ is the Laplacian matrix of an ER random network with $d = 10$ and $I$ is a identity matrix to ensure that matrix $C$ is invertible. According to our construction of information distance, given a characteristic function $v(\cdot)$, we can identify an information distance $\rho(i, j)$ such that $\|C_{ij}\| \leq \kappa \cdot v(\rho(i, j))^{-1}$ for all $i, j = 1, 2, \ldots, N$. The constant $\kappa$ can be chosen as the spectral radius $\lambda_r$ of matrix $C$. We now consider the off-diagonal decay of $C^{-1}$ and seek to identify a number $c'$ such that $\| (C^{-1})_{ij} \| \leq c' \lambda_r \cdot v(\rho(i, j))^{-1}$ for all $i, j = 1, 2, \ldots, N$. When we use a characteristic function $v(\cdot)$ that satisfies the GRS condition (e.g., the one we use in this paper), the constant $c'$ can be chosen to be independent of the network size $N$ (blue curve). In contrast, if we use a characteristic function $v(\cdot)$ that violates the GRS condition (e.g., $v(z) = e^z$), the constant $c'$ may necessarily grow with $N$ (red curve).
Table S1. Relative approximation error* for the smallest eigenvalue of the controllability Gramian.

| model networks† | ER       | BA       | WS       |
|-----------------|----------|----------|----------|
| (9.3 ± 7.7) × 10^{-3} | (2.2 ± 0.7) × 10^{-2} | (6.1 ± 5.0) × 10^{-3} |
| empirical networks‡ | power grid | air transportation | human brain |
| 2.13 × 10^{-6} | 7.27 × 10^{-4} | 1.21 × 10^{-2} |

*Computed as |λ_{min}(\tilde{W}_c^\infty) - λ_{min}(W_c^\infty)|/λ_{min}(W_c^\infty), with each node chosen as a driver and assigned an information neighborhood of size L = \lceil N/100 \rceil.

†Mean and standard deviation of the relative approximation error over 100 network realizations for the same parameters as in Fig. 3.

‡Relative approximation error for the empirical networks and dynamics used in Fig. 5D–F.