Supporting Information

Nishikawa et al. 10.1073/pnas.0912444107

SI Text

1 Laplacian Spectrum of Optimal Networks. Consider a network whose (possibly directional) links have integer strengths, and denote its Laplacian matrix by L. Here we show that if the network is optimal, i.e., the nonidentically zero eigenvalues of L assume a common value $\bar{\lambda}$ (Eq. **3** in the main text), then $\bar{\lambda}$ must be an integer. This result will be valid even when the links are allowed to have negative integer strength.

The characteristic polynomial of L can be written as

$$\det(L - xI) = -x(\bar{\lambda} - x)^{n-1} = -\bar{\lambda}^{n-1}x + \cdots + (-1)^n x^n, \qquad \textbf{[S1]}$$

where n is the number of nodes and I is the $n \times n$ identity matrix. Since L has integer entries, all the coefficients of the characteristic polynomial are integers, and hence $\bar{\lambda}^{n-1}$ in the first term above is an integer. Denote this integer by k. Using m to denote the sum of all link strengths in the network, we have $m = \sum_i L_{ii} = (n-1)\bar{\lambda}$, and hence $k = [m/(n-1)]^{n-1}$. Writing $m/(n-1) = s/t$, where integers s and t do not have common factors, we obtain $kt^{n-1} = s^{n-1}$. Suppose p is a prime factor of k. Then p is also a factor of s^{n-1}, so in fact p is a factor of s. This implies that p^{n-1} is a factor of $kt^{n-1} = s^{n-1}$. Because s and t cannot have a common factor, p^{n-1} must be a factor of k. Thus, any prime factor of k must actually appear with multiplicity $n-1$, and hence we can write $k = q^{n-1}$ where q is an integer. Therefore, $\bar{\lambda}^{n-1} = k = q^{n-1}$, and since $\bar{\lambda}$ is real, we have $\bar{\lambda} = q$, an integer.

2 Perturbation of Laplacian Eigenvalues. Suppose that the Laplacian matrix L_0 of a given network of n nodes has an eigenvalue $\lambda_0 \neq 0$ with multiplicity $k \leq n-1$. Consider a perturbation of the network structure in the form $L = L_0 + \delta L_1$, where δ is a small parameter and L_1 is any fixed Laplacian matrix representing the perturbed links. We do not need to assume that L_0 and L_1 have nonnegative entries, making our result valid even in the presence of negative interactions. Denote the characteristic polynomial of L by $f(x,\delta) := \det(L - xI)$, where I is the $n \times n$ identity matrix. Because λ_0 is an eigenvalue of L_0 with multiplicity k, we have

$$f(x,0) = (x - \lambda_0)^k g(x), \qquad \textbf{[S2]}$$

where g is a polynomial satisfying $g(\lambda_0) \neq 0$. Denote by $\lambda = \lambda(\delta)$ an eigenvalue of L that approaches λ_0 as $\delta \to 0$. Here we show that the change $\Delta\lambda := \lambda - \lambda_0$ of the eigenvalue induced by the perturbation scales as

$$\Delta\lambda \sim \delta^{1/k} \qquad \textbf{[S3]}$$

if the derivative of the characteristic polynomial with respect to the perturbation parameter evaluated at $\delta = 0$ is nonzero:

$$\left.\frac{\partial f}{\partial \delta}\right|_{\substack{x=\lambda_0 \\ \delta=0}} \neq 0. \qquad \textbf{[S4]}$$

Through the Jacobi's formula for the derivative of determinants, this condition can be expressed as

$$\mathrm{tr}[(L_0 - \lambda_0 I)^{k-1} g(L_0) L_1] \neq 0, \qquad \textbf{[S5]}$$

where $\mathrm{tr}(A)$ denotes the trace of matrix A. We expect this condition to be satisfied for most networks and perturbations. For the optimal networks satisfying Eq. **3** in the main text, it can be shown that Eq. **S5** is violated if L_0 is diagonalizable, but L_0 is actually known to be nondiagonalizable for the majority of these optimal networks (1). The scaling **S3** shows that, for a fixed δ, the more degenerate the eigenvalue (larger k), the larger the effect of the perturbation on that eigenvalue. In particular, if the original network is optimal with $\bar{\lambda}$ having the maximum possible multiplicity $n-1$, the effect of perturbation is the largest. This, however, is so because the optimal networks have significantly smaller σ than suboptimal networks (even those with just one more or one less link), and therefore the perturbations of the optimal networks would still be more synchronizable in general than most suboptimal networks.

From Eq. **S2** it follows that

$$\left.\frac{\partial f}{\partial x}\right|_{\substack{x=\lambda_0 \\ \delta=0}} = \left.\frac{\partial^2 f}{\partial x^2}\right|_{\substack{x=\lambda_0 \\ \delta=0}} = \cdots = \left.\frac{\partial^{k-1} f}{\partial x^{k-1}}\right|_{\substack{x=\lambda_0 \\ \delta=0}} = 0, \qquad \textbf{[S6]}$$

but

$$\left.\frac{\partial^k f}{\partial x^k}\right|_{\substack{x=\lambda_0 \\ \delta=0}} = k! \cdot g(\lambda_0) \neq 0. \qquad \textbf{[S7]}$$

Using this to expand $f(x,\delta)$ around $x = \lambda_0$ and $\delta = 0$ up to the kth order terms, and setting $x = \lambda$, we obtain

$$\frac{f(\lambda,\delta)}{\delta} = \left.\frac{\partial f}{\partial \delta}\right|_{\substack{x=\lambda_0 \\ \delta=0}} + \frac{1}{k!}\left.\frac{\partial^k f}{\partial x^k}\right|_{\substack{x=\lambda_0 \\ \delta=0}} \cdot \frac{(\lambda - \lambda_0)^k}{\delta} + O(\delta), \qquad \textbf{[S8]}$$

where $O(\delta)$ includes all higher-order terms. From the characteristic equation $f(\lambda,\delta) = \det(L - \lambda I) = 0$, the left-hand side of Eq. **S8** is zero, so taking the limit $\delta \to 0$ leads to

$$\lim_{\delta \to 0} \frac{(\Delta\lambda)^k}{\delta} = -\frac{1}{g(\lambda_0)} \left.\frac{\partial f}{\partial \delta}\right|_{\substack{x=\lambda_0 \\ \delta=0}}, \qquad \textbf{[S9]}$$

which implies the scaling **S3** when condition **S4** is satisfied.

3 Complexity of Optimal Networks. Here we first describe a systematic method for increasing the number of nodes n in an optimal binary interaction network ($A_{ij} = 0, 1$), while keeping the network optimal. Given an optimal network with $\bar{\lambda} = k$ (which must be an integer by the result in Section 1 above), we construct a new network by adding a new node and connecting any k existing nodes to the new node. As a result, the Laplacian matrix has the form

$$L = \left(\begin{array}{ccc|c} & L_0 & & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline u_1 & \dots & u_n & k \end{array}\right), \qquad \textbf{[S10]}$$

where L_0 is the Laplacian matrix of the original network, each u_i is either 0 or -1, and $u_1 + \cdots + u_n = -k$. Since L is a block triangular matrix, its eigenvalue spectrum consists of the eigenvalues of L_0, which are $0, k, \ldots, k$, and an additional k, which comes from the last diagonal element. Thus, the new network is optimal with $\bar{\lambda} = k$.

We can argue that the number of optimal binary interaction networks grows combinatorially with n. To this end, we first consider $C(n)$, the number of distinct Laplacian matrices corresponding to optimal networks with n nodes. For each optimal network with n nodes and $\bar{\lambda} = k$, the above construction gives $\binom{n}{k}$ different Laplacian matrices corresponding to optimal networks with $n+1$ nodes. Using the bound $\binom{n}{k} \geq n$, which is valid for $k = 1, \ldots, n-1$, we see that $C(n+1) \geq n \cdot C(n)$, which implies $C(n) \geq (n-1)!$ and gives a combinatorially growing lower bound for $C(n)$. Since two different Laplacian matrices may represent isomorphically equivalent networks, $C(n)$ is an overestimate of the number of optimal networks with n nodes. However, given the gross underestimate coming from $\binom{n}{k} \geq n$ and the fact that we used only one out of potentially many possible schemes for adding a node while keeping the network optimal, we expect that the number of optimal binary interaction networks with n nodes also grows combinatorially with n.

4 Optimality for Networks of Heterogeneous Units. Consider a network of coupled nonidentical units whose dynamics is governed by

$$x_i(t+1) = F[x_i(t), \mu_i] + \bar{\varepsilon} \sum_{j=1}^{n} A_{ij} \Big\{ H[x_j(t), \mu_j] - H[x_i(t), \mu_i] \Big\}, \quad \textbf{[S11]}$$

where t represents the discrete time and $\bar{\varepsilon} = \varepsilon / d$ is the global coupling strength normalized by the average coupling strength per node, $d = \frac{1}{n}\sum_i \sum_{j \neq i} A_{ij}$. The dynamics of unit i follows $x_i(t+1) = F[x_i(t), \mu_i]$ in the absence of coupling with other units and is assumed to be one dimensional for simplicity. Variation in the parameter μ_i represents the dynamical heterogeneity of the network, which we measure by the standard deviation σ_μ defined by $\sigma_\mu^2 := \frac{1}{n}\sum_i \delta\mu_i^2$, where $\delta\mu_i := \mu_i - \bar{\mu}$ and $\bar{\mu} := \frac{1}{n}\sum_i \mu_i$. Here we choose the signal function to be $H(x, \mu) = F(x, \mu)$, which leads to a natural generalization of coupled map lattices (2) to arbitrary coupling topology. For example, for the one-dimensional periodic lattice in which each unit is coupled only to its two nearest neighbors with unit strength, system **S11** reduces to the well-studied system $x_i(t+1) = (1-\varepsilon)F[x_i(t), \mu_i] + \frac{\varepsilon}{2}F[x_{i-1}(t), \mu_{i-1}] + \frac{\varepsilon}{2}F[x_{i+1}(t), \mu_{i+1}]$.

We consider a nearly synchronous state in which the deviation of the states of individual units around their average is small, i.e., $\delta x_i(t) := x_i - \bar{x}(t)$ is small, where $\bar{x}(t) := \frac{1}{n}\sum_i x_i(t)$. Following the strategy used in ref. 3 for continuous-time systems, we linearize Eq. **S11** around the average state $\bar{x}(t)$ and the average parameter to obtain the variational equation in the vector form

$$\delta\mathbf{x}(t+1) = (I - \bar{\varepsilon}\tilde{L})[a_t \delta\mathbf{x}(t) + b_t \delta\boldsymbol{\mu}], \quad \textbf{[S12]}$$

where $\delta\mathbf{x}(t) = [\delta x_1(t), \ldots, \delta x_n(t)]^T$ is the state deviation vector, $\delta\boldsymbol{\mu} = (\delta\mu_1, \ldots, \delta\mu_n)^T$ is the parameter variation vector, and I is the $n \times n$ identity matrix. Matrix $\tilde{L}$ is the modified Laplacian matrix defined by $\tilde{L}_{ij} = L_{ij} - \frac{1}{n}\sum_k L_{kj}$ (3), and we denote $\{a_t, b_t\} := \left\{ \frac{\partial F}{\partial x}[\bar{x}(t), \bar{\mu}], \frac{\partial F}{\partial \mu}[\bar{x}(t), \bar{\mu}] \right\}$. As a result of the linearization, the deviation $\delta\mathbf{x}(t)$ can possibly diverge as $t \to \infty$ even when the state space for the network dynamics is bounded. Notice that $(1, \ldots, 1)^T$ is an eigenvector of the matrix $I - \bar{\varepsilon}\tilde{L}$ associated with eigenvalue one. The component of the linearized dynamics parallel to this vector is irrelevant for synchronization stability, because $\delta\mathbf{x}(t)$ by definition does not have this component. We thus remove this component, keeping all other components unchanged, by replacing I in Eq. **S12** with $\tilde{L}^*$ defined by $\tilde{L}^*_{ij} = \delta_{ij} - 1/n$, which leads to

$$\delta\mathbf{x}(t+1) = (\tilde{L}^* - \bar{\varepsilon}\tilde{L})[a_t \delta\mathbf{x}(t) + b_t \delta\boldsymbol{\mu}]. \quad \textbf{[S13]}$$

Any component along $(1, \ldots, 1)^T$ will immediately vanish under multiplication of the matrix $\tilde{L}^* - \bar{\varepsilon}\tilde{L}$, whose properties govern the evolution of synchronization error.

As a measure of synchronization error, we use the standard deviation $\sigma_x(t)$ defined by $\sigma_x^2(t) := \frac{1}{n}\sum_i \delta x_i^2(t)$. For a fixed σ_μ and $\bar{\varepsilon}$, we define the maximum asymptotic synchronization error to be

$$\Omega_{\bar{\varepsilon}}(L) := \max_{\{\mu_i\}} \limsup_{t \to \infty} \sigma_x(t), \quad \textbf{[S14]}$$

where the maximum is taken over all possible combinations of μ_i for the given σ_μ. We can explicitly compute $\Omega_{\bar{\varepsilon}}(L)$ by iterating Eq. **S13**, which leads to

$$\Omega_{\bar{\varepsilon}}(L) = \sigma_\mu \tilde{\Omega}(\bar{\varepsilon}\tilde{L}),$$
$$\text{where } \tilde{\Omega}(X) := \limsup_{T \to \infty} \left\| \sum_{t=1}^{T} \left(\prod_{k=1}^{t-1} a_{k+T-t} \right) b_{T-t} (\tilde{L}^* - X)^t \right\|, \quad \textbf{[S15]}$$

where $\|\cdot\|$ denotes the spectral norm for matrices. Notice that the synchronization error is a linear function of σ_μ, which is a consequence of using Eq. **S13**. The function $\tilde{\Omega}(X)$, whose argument is a matrix X encoding the network structure, can be interpreted as a master synchronization error function since its functional form is determined entirely by the map $F(x, \mu)$ and the averaged trajectory $\bar{x}(t)$ and is independent of the network structure. A sufficient condition for $\tilde{\Omega}(X)$ to be finite is $\rho(\tilde{L}^* - X) < e^{-\nu}$, where $\rho(\cdot)$ denotes the spectral radius of matrices, or equivalently, the maximum of the absolute values of the eigenvalues. Here ν is the Lyapunov exponent of the averaged-parameter map $F(x, \bar{\mu})$ along the average trajectory $\bar{x}(t)$, i.e., $\nu := \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \ln \left| \frac{\partial f}{\partial x}[\bar{x}(t), \bar{\mu}] \right|$. If $X = \bar{\varepsilon}\tilde{L}$, this condition reduces to the stability condition for complete synchronization of the corresponding identical units, namely, $\Lambda(\bar{\varepsilon}\lambda_i) < 0$ for $i = 2, \ldots, n$, where the stability function in this case is $\Lambda(\beta) = \nu + \ln|1 - \beta|$. For example, if $F(x, \mu) = 2x + \mu \bmod 1$, the sum in Eq. **S15** converges to $\tilde{\Omega}(X) = \|(\tilde{L}^* - X)(2X - 2\tilde{L}^* + I)^{-1}\|$ when $\rho(\tilde{L}^* - X) < 1/2$.

The set of networks with a given synchronization error tolerance E_{tol} defined by $\Omega_{\bar{\varepsilon}}(L) \leq E_{\mathrm{tol}}$ is represented by the region $\left\{ X : \tilde{\Omega}(X) \leq \frac{E_{\mathrm{tol}}}{\sigma_\mu} \right\}$ in the space of matrices. In Fig. S1 we illustrate this using the two-parameter family of networks defined by $L = c_1 L_{K_3} + c_2 (L_{C_{3,1}} - L_{C_{3,2}})$, where L_{K_3}, $L_{C_{3,1}}$, and $L_{C_{3,2}}$ are the Laplacian matrices of the fully connected network of three nodes and the two types of three cycles:

$$L_{K_3} = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}, \qquad L_{C_{3,1}} = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix},$$
$$L_{C_{3,2}} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix}. \quad \textbf{[S16]}$$

For this family of networks, we can show that for $X = \tilde{L}$, we have

$$\tilde{\Omega}(\tilde{L}) = \limsup_{T \to \infty} \left| \sum_{t=1}^{T} \left(\prod_{k=1}^{t-1} a_{k+T-t} \right) b_{T-t} (1 - \lambda)^t \right|, \quad \textbf{[S17]}$$

where $\lambda = 3c_1 + i\sqrt{3}c_2$. In particular, we have $\tilde{\Omega}(\tilde{L}) = \left| \frac{1-\lambda}{2\lambda - 1} \right|$ if $F(x, \mu) = 2x + \mu \bmod 1$, and this is used as an illustrative example in Fig. S1. As σ_μ approaches zero and the dynamical units become less heterogeneous, the region $\{X : \tilde{\Omega}(X) \leq \frac{E_{\mathrm{tol}}}{\sigma_\mu}\}$ increases in size and approaches the region of stable synchronization given by $\rho(\tilde{L}^* - X) < e^{-\nu} = 1/2$. For a given network, the change in the synchronization error $\Omega_{\bar{\varepsilon}}(L)$ with respect to $\bar{\varepsilon}$ can be understood as the change in the value of the error function $\tilde{\Omega}(X)$ as $X = \bar{\varepsilon}\tilde{L}$ moves along a straight line determined by $\tilde{L}$. From Eq. **S15**, we expect in general that $\tilde{\Omega}(X)$ is a monotonically increasing function of $\rho(\tilde{L}^* - X)$, and hence $\Omega_{\bar{\varepsilon}}(L)$ is expected to decrease to a minimum at some $\bar{\varepsilon} = \bar{\varepsilon}^*$ and increases monotonically for $\bar{\varepsilon} > \bar{\varepsilon}^*$, or monotonically decrease in the entire range of $\bar{\varepsilon}$ for which $\Omega_{\bar{\varepsilon}}(L)$ is finite. This is indeed the case for the example considered here, as illustrated by the insets in Fig. S1. Thus, in general we define $\Omega(L) := \inf_{\bar{\varepsilon}} \Omega_{\bar{\varepsilon}}(L)$ as a measure of synchronizability of a network, as it gives the lower limit on the asymptotic synchronization error for nonidentical units. For undirected networks, for which $\tilde{L}$ is symmetric and each λ_i is real, diagonalization of $\tilde{L}$ with orthogonal eigenvectors can be used to show that

$$\Omega_{\bar{\varepsilon}}(L) = \sigma_\mu \tilde{\Omega}(\bar{\varepsilon}\tilde{L}) = \sigma_\mu \max_{2 \le i \le n} \tilde{\Omega}^s(\bar{\varepsilon}\lambda_i),$$
$$\text{where } \tilde{\Omega}^s(\beta) = \limsup_{T\to\infty} \left| \sum_{t=1}^{T} \left(\prod_{k=1}^{t-1} a_{k+T-t} \right) b_{T-t} (1-\beta)^t \right| \qquad \textbf{[S18]}$$

under the stability condition $\Lambda(\bar{\varepsilon}\lambda_i) < 0$ for $i = 2, \ldots, n$. For such networks, the synchronization error can be determined visually from the error function $\tilde{\Omega}^s(\beta)$, which is a function of real numbers. This is illustrated in Fig. S2 using the example of $F(x,\mu) = 2x + \mu \bmod 1$, for which we can show that $\tilde{\Omega}^s(\beta) = |\frac{1-\beta}{2\beta-1}|$ and $\Lambda(\beta) = \ln 2 + \ln|1-\beta|$.

We now show that the class of networks with zero synchronization error for arbitrary heterogeneity of the individual units consists of those that are optimal (i.e., satisfies $\lambda_2 = \cdots = \lambda_n = \bar{\lambda} > 0$, Eq. **3** in the main text) and have diagonalizable Laplacian matrix. First, to show that any network with zero synchronization error satisfies these conditions, suppose that $\Omega(L) = 0$ for a given network. That is, for some $\bar{\varepsilon}$, we have $\delta\mathbf{x}(t) \to 0$ as $t \to \infty$ for arbitrary $\mu_1, \ldots, \mu_n$ with a given σ_μ. Then, letting $t \to \infty$ in Eq. **S13**, we conclude that we have either $\lim_{t\to\infty} b_t = \lim_{t\to\infty} \frac{\partial F}{\partial \mu}[\bar{x}(t), \bar{\mu}] = 0$ or $(\tilde{L}^* - \bar{\varepsilon}\tilde{L})\delta\boldsymbol{\mu} = \mathbf{0}$ for all possible $\delta\boldsymbol{\mu}$. The former can hold only in exceptional cases, such as when $\bar{x}(t)$ converges to a fixed point at which $\frac{\partial F}{\partial \mu}$ is zero. We thus assume a typical situation in which the latter holds. In this case, using the fact that the row sum of $\tilde{L}$ is zero and that $\sum_i \delta\mu_i = 0$, we can show that $\tilde{L}^* - \bar{\varepsilon}\tilde{L}$ must be equal to the zero matrix, and hence $\tilde{L} = \frac{1}{\bar{\varepsilon}}\tilde{L}^*$, which is diagonalizable with eigenvalues $0, \frac{1}{\bar{\varepsilon}}, \ldots, \frac{1}{\bar{\varepsilon}}$. Because in general L and $\tilde{L}$ have the same set of eigenvalues and L is diagonalizable iff $\tilde{L}$ is diagonalizable, L is diagonalizable and satisfies $\lambda_2 = \cdots = \lambda_n = \bar{\lambda} > 0$.

Conversely, suppose that the network satisfies $\lambda_2 = \cdots = \lambda_n = \bar{\lambda} > 0$ and the Laplacian matrix is diagonalizable. It can be shown that $\bar{\varepsilon}\tilde{L} = \tilde{L}^*$ if we choose $\bar{\varepsilon} = 1/\bar{\lambda}$, and therefore $\delta\mathbf{x}(t) = \mathbf{0}$ according to Eq. **S13**, but we can actually prove a stronger statement without the linear approximation involved in Eq. **S13**. From Theorem 6 in ref. 1, each node j either has equal output link strength to all other nodes ($A_{ij} = b_j \neq 0$ for all $i \neq j$) or has no output at all ($A_{ij} = b_j = 0$ for all i). This implies that the adjacency matrix satisfies $A_{ij} = b_j$ for all i and j with $i \neq j$, and we have $\sum_j b_j = \bar{\lambda}$. If we choose $\bar{\varepsilon} = 1/\bar{\lambda}$, then Eq. **S11** becomes

$$x_i(t+1) = \frac{1}{\bar{\lambda}} \sum_{j=1}^{n} b_j F[x_j(t), \mu_j] = \sum_{j=1}^{n} w_j F[x_j(t), \mu_j], \qquad \textbf{[S19]}$$

with $\sum_j w_j = 1$. Thus, the state of node i is determined by the weighted average of the signals from all the nodes that have output and, more importantly, it is independent of i for all $t \geq 1$, implying that $\delta\mathbf{x}(t) = \mathbf{0}$ for all $t \geq 1$. Therefore, the system synchronizes in one iteration with zero error, despite the presence of dynamical heterogeneity, and hence $\Omega(L) = 0$. This indicates that the largest Lyapunov exponent for the completely synchronous state is $-\infty$, which is analogous to superstable fixed points and periodic orbits observed in maps.

Since the best networks for synchronizing nonidentical maps satisfy $\lambda_2 = \cdots = \lambda_n = \bar{\lambda} > 0$, they too must have a quantized number of links: $m = k(n-1)$. For every n and every $k = 1, \ldots, n$, there is exactly one binary network ($A_{ij} = 0, 1$) that has $m = k(n-1)$ links and is capable of complete synchronization for nonidentical maps, including the directed star topology ($k = 1$) and the fully connected network ($k = n$). Note also that the above argument does not require that $b_j \geq 0$ for all j (Theorem 6 in ref. 1 remains valid without this requirement). This implies that complete synchronization is possible even for networks with negative interactions. In addition, the ability of a network to completely synchronize nonidentical units, with or without negative interactions, is invariant under the generalized complement transformation defined by Eq. **7** in the main text. To see this, suppose that for a given network we have $\lambda_2 = \cdots = \lambda_n = \bar{\lambda} > 0$ and L is diagonalizable. By Theorem 6 in ref. 1, we have $A_{ij} = b_j$ for all i and j with $i \neq j$, where $\sum_j b_j = \bar{\lambda}$. Using the definition of the complement transformation, we have $A^c_{ij} = \alpha - b_j$, and $\sum_j (\alpha - b_j) = n\alpha - \bar{\lambda} > 0$ if $\alpha > \frac{m}{n(n-1)}$. Applying Theorem 6 in ref. 1 again, we see that the complement satisfies the same property: $\lambda_2 = \cdots = \lambda_n = n\alpha - \bar{\lambda} > 0$ and its Laplacian matrix is diagonalizable. Therefore, in addition to binary networks, there are many networks with negative interactions that are guaranteed to have zero synchronization error.

5 Degree Distribution Before and After Enhancing Synchronization with Negative Directional Interactions. We describe the change in the in- and out-degree distributions of the network as negative strengths are assigned to directional links to enhance synchronization, following the algorithm presented in the main text. The in- and out-degree of node i are defined as $\sum_{k\neq i} A_{ik}$ and $\sum_{k\neq i} A_{ki}$, respectively. Fig. S3 shows the results for random scale-free networks with $\gamma = 2.6$ and $\gamma = 5$. They clearly illustrate that the large in-degree of many nodes is compensated by the negative interactions, creating a sharp cutoff in the distribution (orange arrows in *A* and *C*). In contrast, the out-degree distributions remain essentially unchanged, having a power-law tail with the same exponent (insets in *B* and *D*). Note that the algorithm can create negative out-degree nodes, as indicated by the green arrow in *B*, but this has no significant effect since the in-degree distribution is the main factor that determines the stability of synchronous states.

6 Enhancing Synchronization with Negative Bidirectional Interactions. Here we show that assigning negative strength to bidirectional links can also enhance synchronization significantly. This is implemented using two different algorithms.

The first method is fast and is based on node degrees, similarly to the algorithm used in the main text for assigning negative directional interactions. In order to create negative interactions preferentially between nodes of large degrees, we first order the bidirectional links according to the product of the degrees of the two nodes connected by each link, from high to low values. Going through all the links in this order, we change the strength of each bidirectional link from $+1$ to -1 if the degrees of the two adjacent nodes do not fall below a constant, chosen here arbitrarily to be 1.7 times the mean degree of the initial network. We applied this procedure to random scale-free networks with minimum degree 5, generated by the configuration model. Fig. S4*A* shows the degree distribution before and after assigning negative interactions for the scaling exponent $\gamma = 2.6$ and 5. In both cases, the degree distribution remains essentially scale-free with the same exponents. Denoting by λ_2 and λ_n the smallest and largest nonidentically zero eigenvalues of the Laplacian matrix L, respectively, we measure synchronization enhancement by the relative decrease in the ratio λ_n/λ_2, a standard measure of synchronization widely adopted for undirected networks (4). Fig. S4*C* shows that our method does produce significant enhancement for γ less than about 5 (bottom curve) and that the more heterogeneous the initial degree distribution (smaller γ), the more effective the algorithm.

The second algorithm is slower but more effective, and it is based on the observation that the largest eigenvalue λ_n is typically the one responsible for poor synchronizability in the undirected networks with heterogeneous degree distribution considered here (5). At each step of this algorithm, we use the first-order approximation (6, 7) for each bidirectional link to estimate the change in λ_n that would be caused by changing the strength of that link from $+1$ to -1. We then choose a link with the largest predicted reduction in λ_n and make its strength -1. Repeating this until the fraction of links with negative strength reaches a prescribed threshold (chosen here arbitrarily to be 0.2), we obtain a sequence of candidate networks for improved synchronization.* From these networks we choose one with the smallest ratio λ_n/λ_2 as the output of the algorithm. Fig. 4*B* shows the change in the degree distribution before and after applying this algorithm to random scale-free networks with minimum degree 5 for $\gamma = 2.6$ and 5, generated by the configuration model. For both values of γ, we notice a significant drop in the fraction of high-degree nodes (green arrows), accompanied by the appearance of nodes with degree less than the minimum degree of the initial network (blue arrow). This change appears to be responsible for the reduction of λ_n/λ_2 by as much as about 65%, as shown in Fig. S4*C* (top curve). Note that the synchronization enhancement achieved by this method is consistently larger than the first method based on link degrees (bottom curve). The effectiveness of this method depends on the fact that the Laplacian eigenvalues remain real when the network is kept symmetric, which would not generally hold true if negative strength were assigned to directional links.

1. Nishikawa T, Motter AE (2006) Maximum performance at minimum cost in network synchronization. *Physica D* 224:77–89.
2. Kaneko K (1992) Overview of coupled map lattices. *Chaos* 2:279–282.
3. Sun J, Bollt EM, Nishikawa T (2009) Master stability functions for coupled nearly identical dynamical systems. *Europhys Lett* 85:60011.
4. Barahona M, Pecora LM (2002) Synchronization in small-world systems. *Phys Rev Lett* 89:054101.
5. Kim DH, Motter AE (2007) Ensemble averageability in network spectra. *Phys Rev Lett* 98:248701.
6. Hagberg A, Schult DA (2008) Rewiring networks for synchronization. *Chaos* 18:037105.
7. Milanese A, Sun J, Nishikawa T (2010) Approximating spectral impact of structural perturbations in large networks. *Phys Rev E* 81:046112.

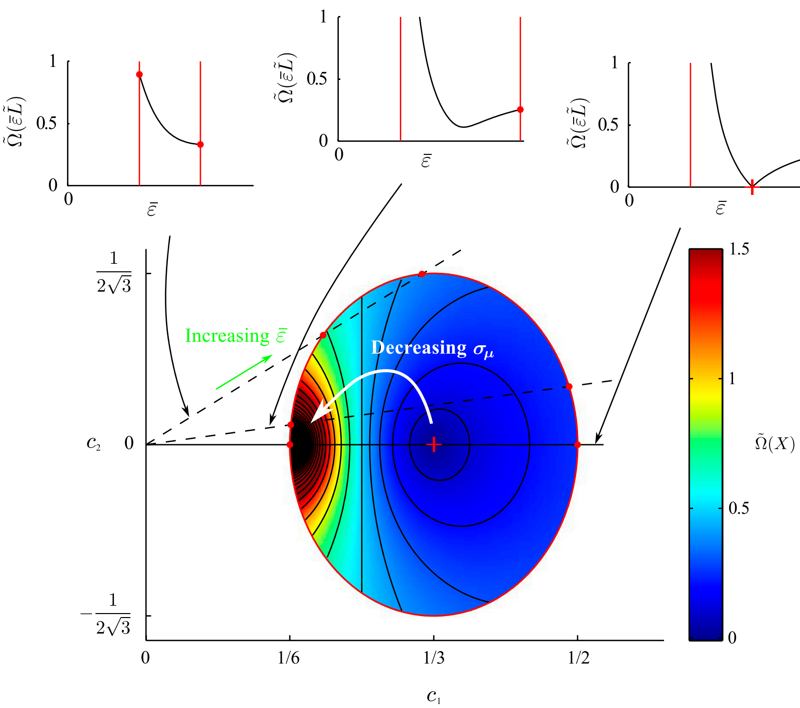


Fig. S1. Master synchronization error function $\tilde{\Omega}(X)$ defined on the space of matrices. The color indicates the values of $\tilde{\Omega}(X)$ with $X = \tilde{L}$ for the two-parameter family of networks given by $L = c_1 L_{K_3} + c_2(L_{C_{3,1}} - L_{C_{3,2}})$ and the map $F(x,\mu) = 2x + \mu \bmod 1$. Shown in black are the level curves $\tilde{\Omega}(X) = \frac{E_{\mathrm{tol}}}{\sigma_\mu}$. For fixed error tolerance E_{tol} and heterogeneity σ_μ, the region defined by $\tilde{\Omega}(X) \le \frac{E_{\mathrm{tol}}}{\sigma_\mu}$ represents the set of networks for which synchronization error is within E_{tol}. The red closed curve indicates the edge of synchronization stability defined by $\rho(\tilde{L}^* - X) < e^{-\nu}$. Each dashed line represents the path $X = \bar{\varepsilon}\tilde{L}$ for a fixed $\tilde{L}$, which indicates the locations of synchronization transition (red dots) and, more generally, how the synchronization error changes as $\bar{\varepsilon}$ is varied (insets). The same representation for the path along the c_1 axis shows that $\tilde{\Omega}(X) = 0$ when $X = \tilde{L}^*$, corresponding to the point $(c_1, c_2) = (1/3, 0)$ for this example (indicated by a red plus symbol).

*Although we have chosen to reduce λ_n here, a similar algorithm can be conceived to increase λ_2 or to simultaneously optimize λ_2 and λ_n.

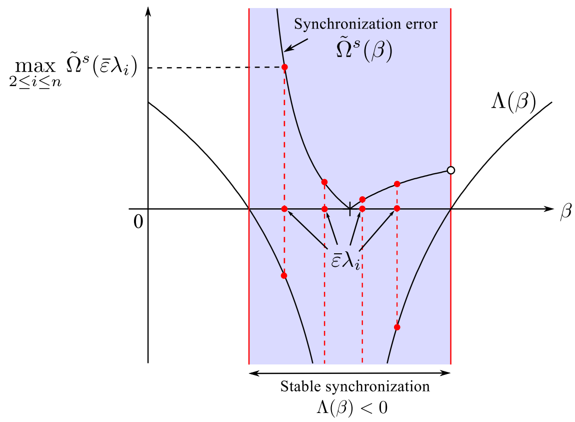


Fig. S2. Synchronization error function for undirected networks. The condition for the stability of synchronization is that the numbers $\bar{\varepsilon}\lambda_2, \ldots, \bar{\varepsilon}\lambda_n$ (the red dots on the β axis) all lie in the interval on which $\Lambda(\beta) < 0$. When this condition is satisfied, the synchronization error is determined by the highest point among the corresponding points on the curve $\tilde{\Omega}^s(\beta)$ (the red dots on the top curve). For illustration we used the curves for $F(x, \mu) = 2x + \mu \bmod 1$, which are $\tilde{\Omega}^s(\beta) = |\frac{1-\beta}{2\beta-1}|$ and $\Lambda(\beta) = \ln 2 + \ln|1 - \beta|$.

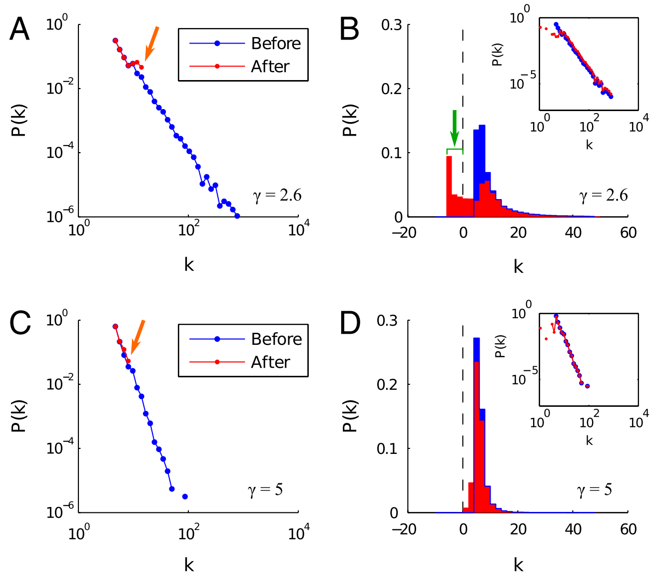


Fig. S3. Change in the degree distributions when enhancing synchronization with negative directional interactions. Distributions before (blue) and after (red) assigning negative strengths in random scale-free networks: (*A*) in-degree distribution for $\gamma = 2.6$, (*B*) out-degree distribution for $\gamma = 2.6$, (*C*) in-degree distribution for $\gamma = 5$, and (*D*) out-degree distribution for $\gamma = 5$. The in-degree distributions are plotted using logarithmic binning, and the absence of a symbol implies that we observed no node with its degree in the corresponding bin. The out-degree distributions are shown in linear scale, along with insets showing the positive part of the distributions in logarithmic scale. Note that there is a significant number of nodes with negative out-degree (green arrow). All plots are averaged over 20 network realizations with 1,000 nodes.

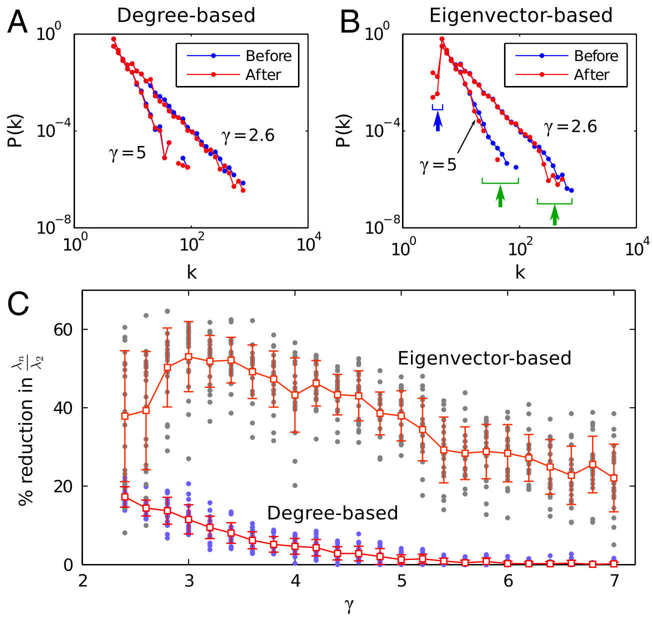
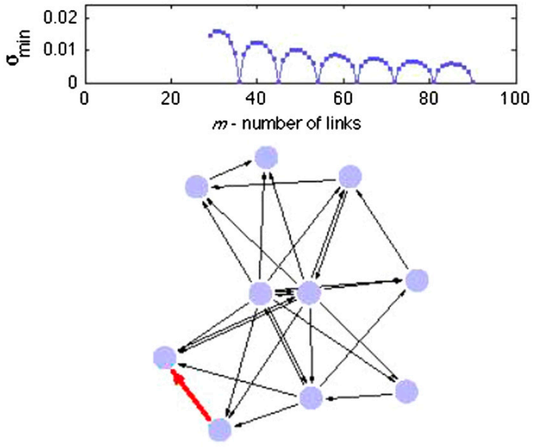


Fig. S4. Improving synchronization with negative bidirectional interactions. (*A* and *B*) Degree distributions before and after we apply two methods of assigning negative interactions to random scale-free networks with $\gamma = 2.6$ and 5: (*A*) degree-based method and (*B*) eigenvector-based method. All four plots are generated using logarithmic binning and averaged over 20 network realizations with 1,000 nodes. The absence of a symbol indicates that we observed no node with its degree in the corresponding bin. (*C*) Reduction in λ_n/λ_2 as a function of γ for the two methods. The error bars indicate the average and the standard deviation over 20 network realizations with 1,000 nodes. The individual realizations are indicated by gray and blue dots.



Movie S1. Networks with best synchronization. The first half of the movie shows the structural changes in networks with best synchronization properties (smallest σ possible for a given number of links) as directional links are removed one by one. We always choose a link that keeps the synchronizability highest (i.e., keeps σ smallest). The second half shows how σ changes in the process, revealing in particular that link removal can counterintuitively enhance synchronization. The node layout at each step was computed using the Kodama-Kawai spring layout algorithm (1). We used the implementation of the algorithm in the Boost C++ library (2) through the Matlab interface provided by MatlabBGL (3).

Movie S1 (MP4)

1. Kamada T, Kawai S (1989) An algorithm for drawing general undirected graphs. *Inform Process Lett* 31:7–15.
2. Boost c++ library, http://www.boost.org/.
3. MatlabBGL, http://www.stanford.edu/~dgleich/programs/matlab_bgl/index.html.