

Network synchronization landscape reveals compensatory structures, quantization, and the positive effect of negative interactions

Takashi Nishikawa * † and Adilson E. Motter * ‡

* Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208, USA, † Department of Mathematics, Clarkson University, Potsdam, NY 13699, USA, and ‡ Northwestern Institute on Complex Systems, Northwestern University, Evanston, IL 60208, USA

Published in PNAS 107, 10342-10347 (2010)

Synchronization, in which individual dynamical units keep in pace with each other in a decentralized fashion, depends both on the dynamical units and on the properties of the interaction network. Yet, the role played by the network has resisted comprehensive characterization within the prevailing paradigm that interactions facilitating pair-wise synchronization also facilitate collective synchronization. Here we challenge this paradigm and show that networks with best complete synchronization, least coupling cost, and maximum dynamical robustness, have arbitrary complexity but *quantized* total interaction strength, which constrains the allowed number of connections. It stems from this characterization that negative interactions as well as link removals can be used to systematically improve and optimize synchronization properties in both directed and undirected networks. These results extend the recently discovered compensatory perturbations in metabolic networks to the realm of oscillator networks and demonstrate why “less can be more” in network synchronization.

complex networks | nonlinear dynamics | stability analysis | network interactions | collective behavior

Flocking animals [1, 2], self-coordinated moving sensors [3], bursting neurons [4], pace-matching chemical oscillators [5], and frequency-locked power generators [6] are some of many physical manifestations of spontaneous synchronization. Like other forms of collective phenomena [7], synchronization depends critically on the properties of the interaction network [8]. Common wisdom suggests that synchronization is generally easier to achieve with more interactions, that synchronization properties change monotonically as the number of available interactions is varied, and that certain network structures facilitate while others inhibit synchronization. These three expectations, however, are all false because they ignore the possibility of compensatory structural effects. For example, removing a link from a globally connected network makes it less likely to synchronize, but targeted removal of additional links can enhance synchronization [9, 10, 11, 12, 13, 14, 15]. Heterogeneous distribution of coupling strengths or connectivity generally inhibits synchronization [16, 17], but when combined they can compensate for each other [18, 19]. Bringing this argument one step further, while previous studies have focused mainly on positive interactions (but see Refs. [20, 21, 22]) — presumably because negative interactions alone generally do not support synchrony — it is actually easy to provide examples in which negative interactions help stabilize synchronous states. This is illustrated in Fig. 1, where the network composed of black and blue links is not optimal for synchronization but the removal of the blue interactions or, alternatively, the addition of interactions with negative strengths (red links) makes it optimal; the same is achieved by weighting the strengths of all three input interactions of each purple node by a factor of 2/3. However, the counter-intuitive properties that start to emerge from such case studies currently lack a common in-depth explanation that is both comprehensive and predictive in nature.

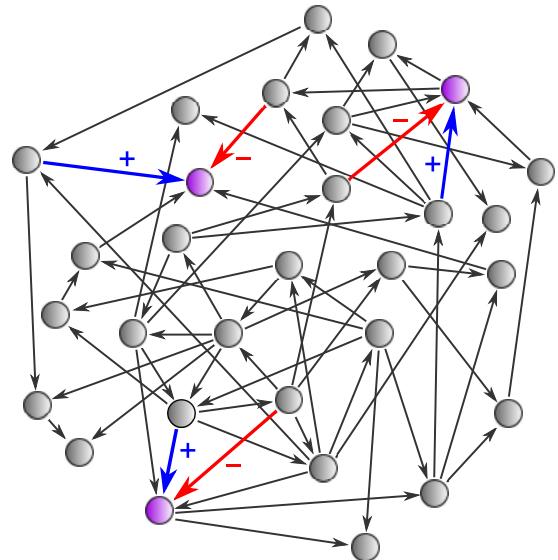


Fig. 1. Examples of compensatory network structures. Nodes represent identical dynamical units and links diffusive couplings. The network consisting of the black and blue links, all having strength +1, is not an optimal network for synchronization, namely one that synchronizes over the widest range of global coupling strength (to be formalized in Results section). However, it can be made optimal either by 1) removing the blue links, 2) adding the red links with negative strength -1 , or 3) scaling the strengths of all in-links of the purple nodes by a factor of $2/3$.

Here we show that these and other apparently disparate properties follow from the discovery we present below that networks optimal for synchronization have a quantized number of links, in multiples of a constant that depends only on the number of nodes and on the connection strengths. We derive our results for the local stability of synchronous states in networks of identical units and provide evidence that our main results remain valid for networks of non-identical units. We choose to focus on optimal networks because, as we show, this class is very rich, can be dealt with analytically, and forms a multi-cusped synchronization landscape, which underlies all

E-mail addresses: tnishika@clarkson.edu (TN), motter@northwestern.edu (AEM)

synchronizable networks and from which suboptimal networks can be studied using perturbation and numerical methods. An immediate consequence of our quantization result is that the addition of links to an optimal network generally results in a suboptimal network, providing systematic examples of link removals that enhance synchronizability. Similar quantization results hold true for optimal networks with negative interactions, which we derive using a generalized complement transformation that maps them into networks with only positive interactions. We show that negative interactions can reveal antagonistic structural relations and counterbalance structural heterogeneities, with potential implications for inhibitory interactions in neural [23, 24, 25], power-grid [6, 26], and cell-regulatory networks [27]. The interactions between power-grid generators, for example, can be positive or negative depending on the inductive versus capacitive nature of the corresponding network elements (e.g., for the Western U.S. power grid they are split 97% versus 3% [26]).

Results

Optimal networks for synchronization. We represent the structure of a network with n nodes using its adjacency matrix $A = (A_{ij})_{1 \leq i,j \leq n}$, where A_{ij} is the strength of the link from node j to node i . We consider the network dynamics governed by

$$\dot{\mathbf{x}}_i = \mathbf{F}(\mathbf{x}_i) + \bar{\varepsilon} \sum_{j=1}^n A_{ij} [\mathbf{H}(\mathbf{x}_j) - \mathbf{H}(\mathbf{x}_i)], \quad [1]$$

where \mathbf{x}_i is the state vector of the i th dynamical unit, \mathbf{F} represents the dynamics of an isolated unit, $\mathbf{H}(\mathbf{x}_j)$ is the signal that the j th unit sends to other units [28], and $\bar{\varepsilon} = \varepsilon/d$ is the global coupling strength $\varepsilon \geq 0$ normalized by the average coupling strength per node $d := \frac{1}{n} \sum_i \sum_{j \neq i} A_{ij}$. As a result of this normalization, the dynamics for a given ε is invariant under scaling of A by a constant, which does not change the network structure. This system has served as a workforce model to study synchronization because it allows analysis of the network influence without detailed specification of the properties of the dynamical units. For example, using a stability function $\Lambda(\cdot)$ that is independent of the network structure [7, 28, 29, 30], the condition for a synchronous state $\mathbf{x}_1(t) = \dots = \mathbf{x}_n(t) = \mathbf{s}(t)$ to be linearly stable is $\Lambda(\bar{\varepsilon}\lambda_i) < 0$ for $i = 2, \dots, n$, where $\lambda_2, \dots, \lambda_n$ are the non-identically zero eigenvalues of the Laplacian matrix L defined by $L_{ij} = \delta_{ij} \sum_k A_{ik} - A_{ij}$ (see *Materials and Methods*). Thus, the smaller the normalized spread of the eigenvalues in the complex plane, which we quantify using

$$\sigma^2 := \frac{1}{d^2(n-1)} \sum_{i=2}^n |\lambda_i - \bar{\lambda}|^2, \text{ where } \bar{\lambda} := \frac{1}{n-1} \sum_{i=2}^n \lambda_i, \quad [2]$$

the more synchronizable the network will generally be. Another measure of synchronizability is the coupling cost $K := \bar{\varepsilon}_s \sum_i \sum_{j \neq i} A_{ij} = \bar{\varepsilon}_s \cdot m$ at the synchronization threshold $\bar{\varepsilon} = \bar{\varepsilon}_s$, whose minimization is equivalent to the condition

$$\lambda_2 = \lambda_3 = \dots = \lambda_n = \bar{\lambda} > 0, \quad [3]$$

which is also equivalent to the condition $\sigma = 0$. This condition constrains all the eigenvalues to be real, even though the networks can be directed and directed networks have complex eigenvalues in general. Condition 3 is also equivalent to the maximization of the range of $\bar{\varepsilon}$ that allows for stable synchronization as well as to the maximization of dynamical robustness, in that it can achieve the fastest exponential convergence

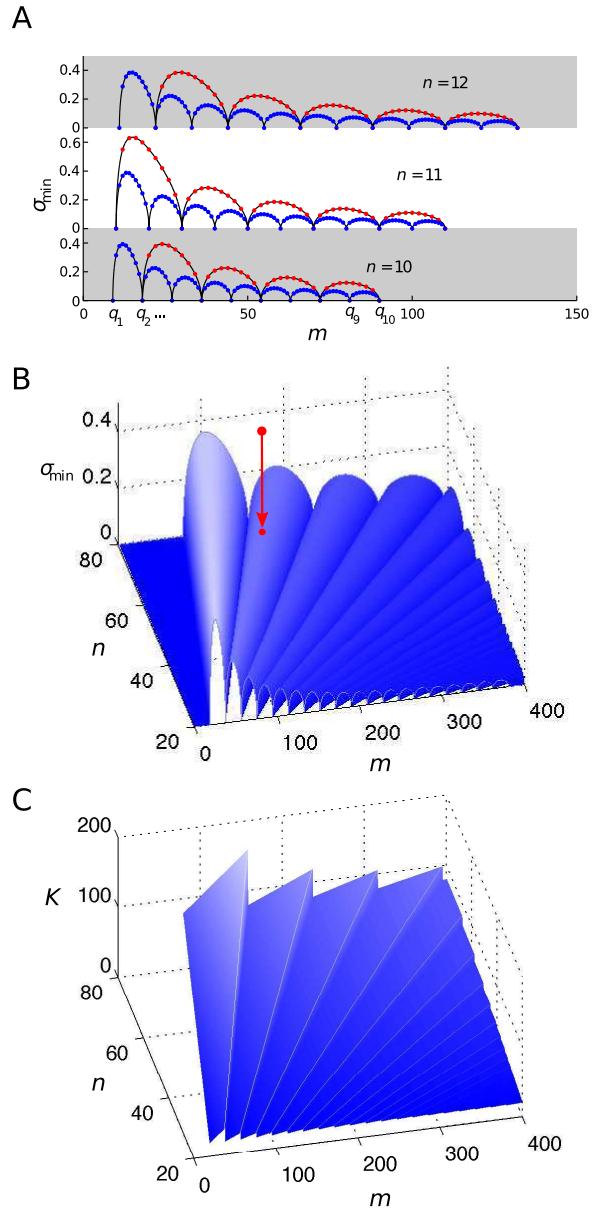


Fig. 2. Quantization in the synchronization and cost landscapes. (A) Estimated by simulated annealing for networks with $n = 10\text{--}12$ nodes, the minimum value σ_{\min} (blue dots) of σ (Eq. 2) shows cusp points at periodic values of the total interaction strength, $m = q_k$ (Eq. 4). In this representation as a function of m , the results for networks with positive ($A_{ij} = 0, 1$) and mixed positive/negative ($A_{ij} = 0, \pm 1$) interactions are identical (blue dots) and coincide with the theoretical prediction (black curves) (Eq. 6). The results for the case $A_{ij} = \pm 1$ (red dots) shows a similar but different quantization behavior, following Eq. 8. (B) Synchronization landscape σ_{\min} as a function of both m and n , computed from Eq. 6. Through simulated annealing, a network with initially high σ (top red dot) evolves to a network satisfying Eq. 5 and $\sigma = \sigma_{\min}$ (bottom red dot). (C) The coupling cost K as a function of m and n in units of K_2 (the cost for a network of two coupled units). The cost takes the lower values along the lines of discontinuity ($m = q_k$).

to synchronization (see *Materials and Methods*). (This may relate, for example, to the finding that heterogeneous distribution of links, which tend to make the distribution of λ_i 's heterogeneous, leads to longer transients to synchronization in ecological network models with strong couplings [31].) We emphasize that the equivalence of these conditions holds for a

wide variety of possible stability functions, including those for which the region defined by $\Lambda(\cdot) < 0$ may be finite or semi-infinite, or have multiple disconnected components. Having the maximum range, minimum coupling cost, and maximum dynamical robustness, the networks satisfying Eq. 3 are called the optimal networks for synchronization. Similar results can be derived for networks of non-identical units, in which the functions \mathbf{F} and \mathbf{H} are possibly different for different nodes, and this more general case will be discussed below.

Quantized number of links. A surprising consequence of having these optimal synchronization properties is the quantization of the number of links in the networks. We find, for example, that for binary interaction networks (i.e., $A_{ij} = 0, 1$) satisfying condition 3, the number of links $m = \sum_i \sum_{j \neq i} A_{ij}$ is quantized to multiples of $n - 1$. That is,

$$m = q_k := k(n - 1), \quad \text{where } k = 1, 2, \dots, n. \quad [4]$$

This follows from the identity $m = \sum_i L_{ii} = \sum_{i=2}^n \lambda_i = (n - 1)\bar{\lambda}$, combined with the fact that condition 3 constrains the real eigenvalue $\bar{\lambda}$ further to be an integer for networks with integer-valued interactions (see Supporting Information, Section 1). Consequently, any network with m strictly between these quantized values must have $\sigma > 0$, and hence cannot be optimal. What is then the minimum σ for all such networks with a given m ? We denote this minimum by $\sigma = \sigma_{\min}(m)$. Based on our analysis of all networks with $n \leq 6$, we conjecture that the condition to achieve this minimum is that the Laplacian eigenvalues (counting multiplicity) have the form

$$0, \underbrace{k, \dots, k}_{q_{k+1}-m}, \underbrace{k+1, \dots, k+1}_{m-q_k}, \quad [5]$$

where k is the integer satisfying $q_k \leq m \leq q_{k+1}$. Note that, analogously to Eq. 4 for $m = q_k$, this condition asserts that the eigenvalues are not only real but also integer for any m . This leads to our prediction that

$$\sigma_{\min}(m) = \frac{1}{(n-1)d} \sqrt{(m - q_k)(q_{k+1} - m)}, \quad [6]$$

which is expected to be valid for binary interaction networks with arbitrary number of nodes and links. Indeed, Fig. 2A shows that for $10 \leq n \leq 12$, a simulated annealing procedure identified networks (blue dots) with σ within 10^{-3} of (but not smaller than) $\sigma_{\min}(m)$ predicted by Eq. 6 (black curve). Starting from the (optimal) fully connected network [with $m = n(n-1)$] at the right end of the curve $\sigma_{\min}(m)$, any initial link deletion necessarily makes synchronization more difficult. Further deletions, however, can make it easier, bringing the networks back to optimal periodically as a function of m at the cusp points and eventually leading to (optimal) directed spanning trees with $m = n - 1$ (see Supporting Video). The optimization of synchronization at these cusp points is analogous to the optimization of the dynamical range of excitable networks at the critical points of phase transitions [32]. Similar cusps are also observed for the Laplacian eigenvalues of a structurally perturbed optimal network as a function of the perturbation (see Supporting Information, Section 2). Note that, although the cost $K = \bar{\varepsilon}_s \cdot m$ generally depends on m , it is actually independent of m for optimal networks of given size n because the synchronization threshold $\bar{\varepsilon}_s = \bar{\varepsilon}_s(m)$ compensates for any change in m .

To reveal the intricate dependence of synchronization on the number of nodes and links, we now consider σ_{\min} as a function of both m and n (Fig. 2B) based on our prediction 6. Each point on this synchronization landscape represents *all* networks with $\sigma = \sigma_{\min}$ for a given m and n , and the number of such networks is expected to grow combinatorially with n (see Supporting Information, Section 3). All other

networks with the same m and n are represented by points directly above that point. The evolution of one such network by rewiring links under pressure to optimize synchronizability can be visualized as a vertical descent from a point above the landscape toward a point on the landscape that minimizes σ (red arrow in Fig. 2B). The sharp valleys are observed along the lines $m = q_k$, and therefore correspond to the cusp points in Fig. 2A. Because two adjacent points on the landscape may include networks that do not have similar structures, it is surprising that one can actually move from one point to the next often by a simple structural modification (see Supporting Video). Indeed, moving in the direction of the m -coordinate can be achieved by simply adding or removing a link that induces the smallest increase or largest decrease in σ . Along the line $m = q_k$, an optimal network can be “grown” and kept optimal, by adding a node and connecting any k existing nodes to the new node (see Supporting Information, Section 3). The flexibility of choosing new links in this construction strongly suggests that optimal networks can have arbitrarily complex connectivity structures as the network size grows. Another interesting landscape appears when we compute the cost K as a function of m and n (Fig. 2C) based on condition 5. In this landscape, optimal networks lie along the lines of discontinuity, resulting from the discontinuous change in $\bar{\varepsilon}_s$ that occurs when m changes from $q_k - 1$ to q_k and defines a saw-tooth function along the m -coordinate. Note that for optimal networks, the cost K is independent of m , as mentioned above, but increases linearly with n and can be expressed as $K = K_2(n - 1)$, where K_2 is the coupling cost for a network of two coupled units. A different but potentially related structure is the roughness of the time horizon considered in the synchronization of parallel processing units in distributed computing [33, 34].

While the presentation above focused on networks of identical units, the main results also hold true for networks of non-identical units. Adopting the framework of Ref. [35] and developing further analysis for networks of one-dimensional maps, we show in Supporting Information (Section 4) that complete synchronization is possible even for non-identical units. Moreover, we show that this is possible only for networks satisfying Eq. 3 in addition to the condition that the Laplacian matrix L is diagonalizable. Since any such networks must also exhibit the same quantization expressed in Eq. 4, we also expect cusps similar to those shown in Fig. 2. For each quantized number of links q_k , we can show that there is a network that can synchronize completely despite the heterogeneity of the dynamical units. Therefore, for both identical *and* non-identical units, condition 5 can be used to systematically construct examples of suboptimal networks that can be made optimal by either adding or removing links.

Stabilizing effect of negative interactions. Interestingly, the exact same quantization effect described above for binary interaction networks is also observed when we allow for negative interactions and interpret $m = \sum_i \sum_{j \neq i} A_{ij}$ as the *net* number of links. To see this, we use a generalization of the complement network, which we define for a given constant α to be the network with adjacency matrix A^c given by

$$A_{ij}^c := (\alpha - A_{ij})(1 - \delta_{ij}). \quad [7]$$

(This includes the special case $\alpha = 1$, which for undirected unweighted networks corresponds to the previously studied case of complement graphs [36].) The transformation from a network to its complement maps m to $\alpha n(n - 1) - m$, λ_i to $\alpha n - \lambda_i$, σ to $\frac{m\sigma}{\alpha n(n - 1) - m}$, and thus an optimal network

to another optimal network when $\alpha > \frac{m}{n(n-1)}$ (see *Materials and Methods*). This also establishes a mapping between networks capable of complete synchronization of non-identical units, since the Laplacian matrix for such a network remains diagonalizable under the transformation (see Supporting Information, Section 4). The condition on α avoids (nonsynchronizable) networks having eigenvalues with negative real part as an artifact of the transformation. We argue that this generalized complement transformation is a powerful tool in analyzing networks with negative interactions because it reduces problems involving negative interactions to those involving only positive interactions when we choose $\alpha \geq \max A_{ij}$.

As an example of quantization with negative interactions, we consider the class of networks for which $A_{ij} = 0, \pm 1$ and assume that $\lambda_2, \dots, \lambda_n$ have positive real parts to ensure that the network can synchronize. Mapping these networks under the complement transformation with $\alpha = 1$, we obtain positive interaction networks with $A_{ij} = 0, 1, 2$. Conjecture 5 applied to the resulting networks then leads to a prediction *identical* to Eq. 6, which we validated by simulated annealing for networks of size up to 12 (blue dots, Fig. 2A). Thus, all the results discussed above based on Eq. 6, including Eq. 4 and the shape of the synchronization landscape, are predicted to be valid even when negative interactions are allowed. Whether the networks with $\sigma = \sigma_{\min}(m)$ actually do have negative interactions is not a priori clear because the synchronization landscape can be built entirely by using the subset of networks with only positive interactions. Our simulated annealing shows, however, that many optimal networks (i.e., with $\sigma = 0$) have negative interactions, as illustrated by an example in Fig. 1. In addition, complete synchronization of non-identical units is possible in the presence of negative interactions (see Supporting Information, Section 4). These networks provide clear and systematic examples of negative interactions that improve synchronization, as removing negative interactions from any such network would in general push m off the quantized values ($m = q_k$) and make the network suboptimal.

The quantization of m and the shape of the synchronization landscape, though they were identical for the two examples above, do depend critically on the constraints imposed on the interactions. Consider, for example, the fully connected networks with interaction strengths ± 1 . It can be shown that the implied constraint $A_{ij} \neq 0$ leads to a different quantization compared to Eq. 6, which involves multiples of $2(n-1)$ rather than $n-1$,

$$\sigma_{\min}(m) = \frac{1}{(n-1)d} \sqrt{(m - q_{n-2(k+1)})(q_{n-2k} - m)} \quad [8]$$

for $q_{n-2(k+1)} \leq m \leq q_{n-2k}$ (red dots, Fig. 2A). The synchronization landscape can be drastically different in some extreme cases. On the one hand, within the widely studied class of undirected unweighted networks (corresponding to a stronger set of constraints, $A_{ij} = A_{ji}, A_{ij} = 0, 1$), no network with $m < n(n-1)$ satisfies the optimality condition 3 [37, 10], which indicates that $\sigma_{\min}(m) > 0$ for all $m < n(n-1)$. On the other hand, within the class of weighted networks corresponding to having no constraint on the interaction strengths, an optimal network can be constructed for any number of links $\geq n-1$ (e.g., the hierarchical networks in Ref. [10]), which implies that σ_{\min} is identically zero in this particular case.

To demonstrate that the synchronization enhancement by negative interactions goes much beyond the realm of optimal networks, we propose a simple algorithm for assigning strength -1 to directional links in an arbitrary network with all link strengths initially equal to $+1$. Our strategy is based

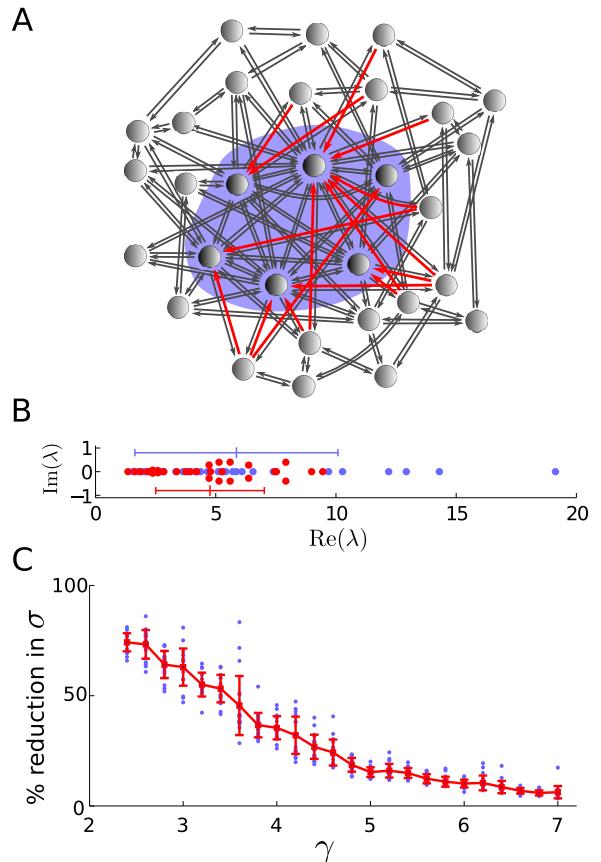


Fig. 3. Improving synchronization with negative interactions. (A) Strength -1 is preferentially assigned to directional links (red arrows) connecting to high in-degree nodes (those in the shaded area) in an initially undirected network. The black links have strength $+1$. (B) The initial spread of the Laplacian eigenvalues is significantly reduced as a result (from blue to red dots). The error bars indicate the mean and one standard deviation of the corresponding distributions. (C) The mean and the standard deviation (red error bars) of the percentage reduction in σ over 20 realizations (blue dots) of random undirected scale-free networks with 1,000 nodes, minimum degree 5, and the scaling exponent γ . The networks were generated by connecting nodes randomly according to a power-law degree distribution $P(\kappa) \sim \kappa^{-\gamma}$.

on the observation that the in-degree distribution is a main factor determining synchronizability [17, 18, 19, 38, 39], where the in-degree (or the total input strength) of node i is defined to be $\sum_{j \neq i} A_{ij}$. Since heterogeneity in the in-degree distribution tends to inhibit synchronization [17, 18, 19, 38, 40], here we use negative interactions to compensate for positive interactions and to homogenize the in-degree distribution. For a given network, we first choose randomly a node with the smallest in-degree and change the strength of each out-link of that node to -1 , unless it makes the in-degree of the target node smaller than the mean in-degree of the original network. We keep strength $+1$ for the other out-links, as well as all the in-links. Having treated this node, we repeat this process for the subnetwork of untreated nodes, considering links and hence degrees only within that subnetwork. We continue this process until all nodes are treated. Applying this algorithm to the network in Fig. 3A, we see that the high in-degree nodes of the initial network (those in the shaded area) receive all of the compensatory negative interactions (red arrows), reducing σ by nearly 35% (Fig. 3B). The effectiveness of the method was further validated (Fig. 3C) using random scale-free networks [41] generated by the standard configuration model [42], where we see more dramatic effect for more heterogeneous networks,

reducing σ by as much as 85%. The synchronization enhancement was indeed accompanied by the homogenization of the in-degree distribution (see Supporting Information, Section 5). The use of negative directional interactions in our algorithm suggests that the enhancement is partly due to link directionality (see Ref. [43] for an enhancement method purely based on link directionality), which generally plays an important role in synchronization (see for example Refs. [21] and [44]). However, negative strength of interactions alone can also produce similar enhancement in random scale-free networks when they are assigned to bidirectional links between hubs (see Supporting Information, Section 6).

Conclusions

Even though negative interactions and link removals by themselves tend to destabilize synchronous states, we have shown that they can compensate for other instabilities, such as those resulting from a “forbidden” number of interactions or from a heterogeneous in-degree distribution. This establishes an unexpected relation between network synchronization and recent work on metabolic networks, where locally deleterious perturbations have been found to generate similar compensatory effects that are globally beneficial in the presence of other deleterious perturbations [45, 46]. For example, the removal of a metabolic reaction — or, equivalently, of its enzyme-coding gene(s) — can often be partially or totally compensated by the targeted removal of a second metabolic reaction. That is, the inactivation of specific metabolic reactions can improve the performance of defective or sub-optimally operating metabolic networks, and in some cases it can even rescue cells that would be nonviable otherwise [45]. Other research has shown that similar reaction inactivation occurs spontaneously for cells evolved to optimize specific metabolic objectives [47]. These apparently disparate examples share the common property that the collective dynamics is controlled, and in fact enhanced, by *constraining* rather than augmenting the underlying network. Such network-based control, including the conditionally positive impact of otherwise detrimental interactions and constraints, is most likely not unique to these contexts.

In neuronal networks, for example, inhibitory interactions have been predicted to facilitate synchronous bursting [23, 24, 25, 48, 49, 50]. Even more compelling, it has been established for both animal [51] and human [52, 53] cerebral cortices that the density of synapses first increases and then decreases during early brain development, suggesting a positive role played by link removal also in neuronal networks. More generally, we expect that the eigenvalue spectrum analysis that lead to our conclusions will help investigate analogous behavior in yet other classes of network processes governed by spectral properties, such as epidemic spreading [54] and cascading failures [55, 56].

Taken together, the highly structured characteristics revealed by the synchronization landscape explain why the characterization of the network properties that govern synchronization has been so elusive. Numerous previous studies performed under comparable conditions have sometimes led to apparently conflicting conclusions about the role played by specific network structural properties such as randomness, betweenness centrality, and degree distribution [30]. Our results indicate that at least part of these disagreements can be attributed to the sensitive dependence of the synchronization properties on the specific combination of nodes and links, as clearly illustrated by the non-monotonic, periodic structure of cusps exhibited by the synchronization landscape. We suggest that insights provided by these findings will illuminate the de-

sign principles and evolution mechanisms of both natural and engineered networks in which synchronization is functionally important [57].

Materials and Methods

Synchronization stability analysis. The stability analysis can be carried out using a master stability approach based on linearizing Eq. 1 around synchronous states [28]. We apply a coordinate transformation to reduce the Laplacian matrix L to its Jordan canonical form and decompose the variational equations into components along the corresponding eigenspaces [9, 10]. This leads to a master stability equation,

$$\dot{\mathbf{y}} = [D\mathbf{F}(\mathbf{s}) - aD\mathbf{H}(\mathbf{s})]\mathbf{y}, \quad [9]$$

where $a = \bar{\varepsilon}\lambda_i$ for the eigenspace corresponding to λ_i . The stability function $\Lambda(a)$ is the maximum Lyapunov exponent for the solution $\mathbf{y}(t) = \mathbf{0}$ of Eq. 9. The A_{ij} are not required to be nonnegative and, since the derivation of $\Lambda(a)$ is based on a Jordan form, L is not required to be diagonalizable either [9, 10]. Geometrically, the stability condition is that all the (possibly complex) numbers $\bar{\varepsilon}\lambda_2, \dots, \bar{\varepsilon}\lambda_n$ lie in the region $\{a \in \mathbb{C} : \Lambda(a) < 0\}$ of the complex plane. Since $\Lambda(a)$ is generally positive on the left side of the imaginary axis, we consider only the networks whose Laplacian eigenvalues have nonnegative real part, which ensures that complete synchronization is possible.

Dynamical robustness of optimal networks. For a given stability function $\Lambda(a)$, a network satisfying condition 3 has the maximum rate of exponential convergence to synchronization among all networks. This is so under the assumption that there is a physical limit M to the coupling strength of individual links, $\bar{\varepsilon}A_{ij} \leq M$, and that the stability function $\Lambda(a)$ is continuous and monotonically increasing with respect to the distance from the real axis, which appears to hold true for all known examples of $\Lambda(a)$ in the literature [58]. From the limit on coupling strengths, it follows that the real part of $\bar{\varepsilon}\lambda_i$ is bounded by a constant $Mn(n-1)$, which, combined with the assumption on $\Lambda(a)$, implies that the exponential rate of convergence is completely determined by the value of $\Lambda(a)$ in the interval $[0, Mn(n-1)]$ on the real axis. In this interval $\Lambda(a)$ has a global minimum at some a^* , and thus $r^* := -\Lambda(a^*) > 0$ is the maximum possible rate of exponential convergence. If the network satisfies condition 3, all perturbation eigenmodes converge at the maximum rate r^* when we choose $\bar{\varepsilon} = \frac{a^*}{\lambda}$. In contrast, if the network violates condition 3, there must be at least one eigenvalue λ_i such that $\bar{\varepsilon}\lambda_i \neq a^*$ [excluding the exceptional situation where multiple values of $\bar{\varepsilon}\lambda_i$ fall precisely on multiple global minima of $\Lambda(a)$], resulting in an eigenmode that converges at a slower rate $r := -\Lambda(\bar{\varepsilon}\lambda_i) < r^*$. Therefore, although optimal networks may suffer from initially slower convergence to synchronization that is polynomial in time [9, 10], the long-term convergence rate is dominated by r^* and is faster than for any other network. Indeed, the deviation from the synchronous state can be written as $P(t)e^{-r^*t}$ for an optimal network and as $Q(t)e^{-rt}$ for a suboptimal network, where $P(t)$ and $Q(t)$ are polynomials. The ratio between the deviations in the two cases is then

$$\frac{P(t)}{Q(t)}e^{-(r^*-r)t} \rightarrow 0 \quad \text{as } t \rightarrow \infty \quad [10]$$

and, in particular, is less than 1 for sufficiently large t , implying that the deviation will eventually become smaller for the optimal network.

Laplacian spectrum of generalized complements. We show that if the eigenvalues of L are $0, \lambda_2, \dots, \lambda_n$ (counting multiplicity), then the eigenvalues of the Laplacian matrix L^c of the generalized complement (defined by Eq. 7) are $0, n\alpha - \lambda_2, \dots, n\alpha - \lambda_n$. This result follows from the relation

$$\mu(L^c, x) = (-1)^{n+1} \frac{x}{n\alpha - x} \mu(L, n\alpha - x), \quad [11]$$

where $\mu(L, x) = \det(L - xI)$ is the characteristic polynomial of the matrix L and I is the $n \times n$ identity matrix. We derive this relation by following the strategy of the proof of Lemma 2.3 in Ref. [59] for undirected networks with non-negative link strengths, to now consider directional and weighted links, possibly with negative strengths. From the definition of the complement transformation, we can write $L + L^c = n\alpha I - \alpha J$, where J is the $n \times n$ matrix with every entry equal to one. Using this and well-known properties of the determinant, we have

$$\begin{aligned} \mu(L^c, x) &= \det(L^c - xI) \\ &= \det(n\alpha I - \alpha J - L - xI) \\ &= (-1)^n \det[L^T + \alpha J - (n\alpha - x)I] \\ &= (-1)^n \mu(L^T + \alpha J, n\alpha - x), \end{aligned} \quad [12]$$

where L^T denotes the transpose of L . Eq. 11 then follows from the fact that $\mu(L^T + \alpha J, z)/(n\alpha - z) = -\mu(L^T, z)/z$ whenever L has row sums equal to zero (which is the case, since L is the Laplacian matrix of a network). To prove this fact, we will use elementary row operations, which do not change the determinants. First, we replace the first row of the matrix $L^T + \alpha J - zI$ by the sum of all rows, making each entry in this row $n\alpha - z$. Next, we subtract this row multiplied by $\alpha/(n\alpha - z)$ (which makes it a row with all entries equal to α) from the remaining rows, canceling out the contribution from the term αJ in these rows. We denote the resulting matrix by M_1 . Finally, we take the matrix $L^T - zI$ and replace the first row by the sum of all rows. This results in a matrix M_2 with exactly the same entries as M_1 except for the first row. For this matrix this is a row with all entries equal to $-z$ rather than $n\alpha - z$. Dividing the first rows of M_1 and M_2 by $n\alpha - z$ and $-z$, respectively, which will scale the determinants accordingly, we

see that

$$\begin{aligned} \frac{\mu(L^T + \alpha J, z)}{n\alpha - z} &= \frac{\det(L^T + \alpha J - zI)}{n\alpha - z} = \frac{\det(M_1)}{n\alpha - z} \\ &= \frac{\det(M_2)}{-z} = \frac{\det(L^T - zI)}{-z} = -\frac{\mu(L^T, z)}{z}. \end{aligned} \quad [13]$$

As an immediate consequence of this result, the normalized standard deviation σ changes to $\frac{m\sigma}{\alpha n(n-1)-m}$ under the complement transformation, while the total link strength m is mapped to $\alpha n(n-1) - m$. In particular, this implies that the complement of any optimal network is also optimal if $\alpha > \frac{m}{n(n-1)}$.

ACKNOWLEDGMENTS. The authors thank Marian Anghel for discussions on the power-grid network and Sara Solla for providing important references on the role of inhibitory neurons. This work was supported by NSF under Grant No. DMS-0709212.

1. Vicsek T, Czirók A, Ben-Jacob E, Shochet O (1995) Novel type of phase transition in a system of self-driven particles. *Phys Rev Lett* 75:1226–1229.
2. Cucker F, Smale S (2007) Emergent behavior in flocks. *IEEE Trans Automat Contr* 52:852–862.
3. Yang P, Freeman R, Lynch K (2008) Multi-agent coordination by decentralized estimation and control. *IEEE Trans Automat Contr* 53:2480–2496.
4. Izhikevich E (2001) Synchronization of elliptic bursters. *SIAM Rev* 43:315–344.
5. Taylor AF, Tinsley MR, Wang F, Huang Z, Showalter K (2009) Dynamical quorum sensing and synchronization in large populations of chemical oscillators. *Science* 323:614–617.
6. Grainger J, Stevenson Jr W (1994) *Power System Analysis* (McGraw-Hill Co., Singapore).
7. Barrat A, Barthelemy M, Vespignani A (2008) *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, UK).
8. Strogatz SH (2001) Exploring complex networks. *Nature* 410:268–276.
9. Nishikawa T, Motter AE (2006) Synchronization is optimal in non-diagonalizable networks. *Phys Rev E* 73:065106(R).
10. Nishikawa T, Motter AE (2006) Maximum performance at minimum cost in network synchronization. *Physica D* 224:77–89.
11. Yin CY, Wang WX, Chen G, Wang BH (2006) Decoupling process for better synchronizability on scale-free networks. *Phys Rev E* 74:047102.
12. Duan Z, Chen G, Huang L (2007) Complex network synchronizability: Analysis and control. *Phys Rev E* 76:056103.
13. Gu Y, Sun J (2009) Altering synchronizability by adding and deleting edges for scale-free networks. *Physica A* 388:3261–3267.
14. Hagberg A, Schult DA (2008) Rewiring networks for synchronization. *Chaos* 18:037105.
15. Zhi-Sheng D, Wen-Xu W, Chao L, Guan-Rong C (2009) Are networks with more edges easier to synchronize, or not? *Chinese Phys B* 18:3122–3130.
16. Nishikawa T, Motter AE, Lai YC, Hoppnerstadt FC (2003) Heterogeneity in oscillator networks: Are smaller worlds easier to synchronize? *Phys Rev Lett* 91:014101.
17. Zhou C, Motter AE, Kurths J (2006) Universality in the synchronization of weighted random networks. *Phys Rev Lett* 96:034101.
18. Motter AE, Zhou C, Kurths J (2005) Enhancing complex-network synchronization. *Europhys Lett* 69:334–340.
19. Motter AE, Zhou C, Kurths J (2005) Network synchronization, diffusion, and the paradox of heterogeneity. *Phys Rev E* 71:016116.
20. Restrepo JG, Ott E, Hunt BR (2006) Characterizing the dynamical importance of network nodes and links. *Phys Rev Lett* 97:094102.
21. Restrepo JG, Ott E, Hunt BR (2006) Synchronization in large directed networks of coupled phase oscillators. *Chaos* 16:015107.
22. Bauer F, Atay FM, Jost J (2010) Synchronized chaos in networks of simple units. *Europhys Lett* 89:20002.
23. Bragin A, et al. (1995) Gamma (40–100 Hz) oscillation in the hippocampus of the behaving rat. *J Neurosci* 15:47–60.
24. Whittington M, Traub R, Jefferys J (1995) Synchronized oscillations in interneuron networks driven by metabotropic glutamate receptor activation. *Nature* 373:612–615.
25. Ylinen A, et al. (1995) Sharp wave-associated high-frequency oscillation (200 Hz) in the intact hippocampus: network and intracellular mechanisms. *J Neurosci* 15:30–46.
26. Myers SA, Anghel M, Motter AE (2009) Spontaneous synchrony in power grids. *Tech. Rep. LA-UR-09-02778*, Los Alamos National Laboratory.
27. McAdams HH, Shapiro L (2003) A bacterial cell-cycle regulatory network operating in time and space. *Science* 301:1874–1877.
28. Pecora LM, Carroll TL (1998) Master stability functions for synchronized coupled systems. *Phys Rev Lett* 80:2109–2112.
29. Dorogovtsev SN, Goltsev AV, Mendes JFF (2008) Critical phenomena in complex networks. *Rev Mod Phys* 80:1275–1335.
30. Arenas A, Díaz-Guilera A, Kurths J, Moreno Y, Zhou C (2008) Synchronization in complex networks. *Phys Rep* 469:93–153.
31. Holland MD, Hastings A (2008) Strong effect of dispersal network structure on ecological dynamics. *Nature* 456:792–794.
32. Kinouchi O, Copelli M (2006) Optimal dynamical range of excitable networks at criticality. *Nat Phys* 2:348–351.
33. Guclu H, Korniss G, Novotny MA, Toroczkai Z, Rácz Z (2006) Synchronization landscapes in small-world-connected computer networks. *Phys Rev E* 73:066115.
34. Korniss G, Novotny MA, Guclu H, Toroczkai Z, Rikvold PA (2003) Suppressing roughness of virtual times in parallel discrete-event simulations. *Science* 299:677–679.
35. Sun J, Bolt EM, Nishikawa T (2009) Master stability functions for coupled nearly identical dynamical systems. *Europhys Lett* 85:60011.
36. Diestel R (2005) *Graph theory*, vol. 173 of *Graduate Texts in Mathematics* (Springer-Verlag, Berlin), third edn.
37. Donetti L, Hurtado PI, Muñoz MA (2005) Entangled networks, synchronization, and optimal network topology. *Phys Rev Lett* 95:188701.
38. Motter AE (2007) Bounding network spectra for network design. *New J Phys* 9:182.
39. Belykh I, de Lange E, Hasler M (2005) Synchronization of bursting neurons: What matters in the network topology. *Phys Rev Lett* 94:188101.
40. Denker M, Timme M, Diesmann M, Wolf F, Geisel T (2004) Breaking synchrony by heterogeneity in complex networks. *Phys Rev Lett* 92:074103.
41. Barabási AL, Albert R (1999) Emergence of scaling in random networks. *Science* 286:509–512.
42. Newman MEJ, Strogatz SH, Watts DJ (2001) Random graphs with arbitrary degree distributions and their applications. *Phys Rev E* 64:026118.
43. Son SW, Kim BJ, Hong H, Jeong H (2009) Dynamics and directionality in complex networks. *Phys Rev Lett* 103:228702.
44. Belykh I, Belykh V, Hasler M (2006) Synchronization in asymmetrically coupled networks with node balance. *Chaos* 16:015102.
45. Motter AE, Gulbahce N, Almaas E, Barabási AL (2008) Predicting synthetic rescues in metabolic networks. *Mol Syst Biol* 4:168.
46. Motter AE (2010) Improved network performance via antagonism: From synthetic rescues to multi-drug combinations. *BioEssays* 32:236–245.
47. Nishikawa T, Gulbahce N, Motter AE (2008) Spontaneous reaction silencing in metabolic optimization. *PLoS Comp Bio* 4:e1000236.
48. Belykh I, Shilnikov A (2008) When weak inhibition synchronizes strongly desynchronizing networks of bursting neurons. *Phys Rev Lett* 101:078102.
49. Vreeswijk C, Abbott LF, Bard Ermentrout G (1994) When inhibition not excitation synchronizes neural firing. *J Comput Neurosci* 1:313–321.
50. Wang XJ, Buzsaki G (1996) Gamma oscillation by synaptic inhibition in a hippocampal interneuronal network model. *J Neurosci* 16:6402–6413.
51. Bourgeois J, Rakic P (1993) Changes of synaptic density in the primary visual cortex of the macaque monkey from fetal to adult stage. *J Neurosci* 13:2801–2820.
52. Huttenlocher P (1979) Synaptic density in human frontal cortex - developmental changes and effects of aging. *Brain Res* 163:195–205.
53. Huttenlocher PR, Dabholkar AS (1997) Regional differences in synaptogenesis in human cerebral cortex. *J Comp Neurol* 387:167–178.
54. Boguñá M, Pastor-Satorras R (2002) Epidemic spreading in correlated complex networks. *Phys Rev E* 66:047104.
55. Motter AE (2004) Cascade control and defense in complex networks. *Phys Rev Lett* 93:098701.
56. Simonsen I, Buzna L, Peters K, Bornholdt S, Helbing D (2008) Transient dynamics increasing network vulnerability to cascading failures. *Phys Rev Lett* 100:218701.
57. Kiss IZ, Rusin CG, Kori H, Hudson JL (2007) Engineering complex dynamical structures: Sequential patterns and desynchronization. *Science* 316:1886–1889.
58. Huang L, Chen Q, Lai YC, Pecora LM (2009) Generic behavior of master-stability functions in coupled nonlinear dynamical systems. *Phys Rev E* 80:036204.
59. Mohar B, Poljak S (1990) Eigenvalues and the max-cut problem. *Czech Math J* 40:343–352.

Supporting Information

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 non-identically 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

$$\begin{aligned}\det(L - xI) &= -x(\bar{\lambda} - x)^{n-1} \\ &= -\bar{\lambda}^{n-1}x + \cdots + (-1)^nx^n,\end{aligned}\quad [\text{S1}]$$

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 $kt^{n-1} = s^{n-1}$. Since 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. Since λ_0 is an eigenvalue of L_0 with multiplicity k , we have

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

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 $\delta \rightarrow 0$. Here we show that the change $\Delta\lambda := \lambda - \lambda_0$ of the eigenvalue induced by the perturbation scales as

$$\Delta\lambda \sim \delta^{1/k} \quad [\text{S3}]$$

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

$$\frac{\partial f}{\partial \delta} \Big|_{\substack{x=\lambda_0 \\ \delta=0}} \neq 0. \quad [\text{S4}]$$

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

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

where $\text{tr}(A)$ 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

$$\frac{\partial f}{\partial x} \Big|_{\substack{x=\lambda_0 \\ \delta=0}} = \frac{\partial^2 f}{\partial x^2} \Big|_{\substack{x=\lambda_0 \\ \delta=0}} = \cdots = \frac{\partial^{k-1} f}{\partial x^{k-1}} \Big|_{\substack{x=\lambda_0 \\ \delta=0}} = 0, \quad [\text{S6}]$$

but

$$\frac{\partial^k f}{\partial x^k} \Big|_{\substack{x=\lambda_0 \\ \delta=0}} = k! \cdot g(\lambda_0) \neq 0. \quad [\text{S7}]$$

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

$$\frac{f(\lambda, \delta)}{\delta} = \frac{\partial f}{\partial \delta} \Big|_{\substack{x=\lambda_0 \\ \delta=0}} + \frac{1}{k!} \frac{\partial^k f}{\partial x^k} \Big|_{\substack{x=\lambda_0 \\ \delta=0}} \cdot \frac{(\lambda - \lambda_0)^k}{\delta} + O(\delta), \quad [\text{S8}]$$

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 $\delta \rightarrow 0$ leads to

$$\lim_{\delta \rightarrow 0} \frac{(\Delta\lambda)^k}{\delta} = -\frac{1}{g(\lambda_0)} \frac{\partial f}{\partial \delta} \Big|_{\substack{x=\lambda_0 \\ \delta=0}}, \quad [\text{S9}]$$

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 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

$$L = \left(\begin{array}{cc|c} & & 0 \\ & L_0 & \vdots \\ \hline u_1 & \cdots & u_n & k \end{array} \right), \quad [\text{S10}]$$

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 $0, k, \dots, k$, 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 $k = 1, \dots, n-1$, 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 non-identical units whose dynamics is governed by

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

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 [3], 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 $\bar{\mu}$ to obtain the variational equation in the vector form

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

where $\delta\mathbf{x}(t) = (\delta x_1(t), \dots, \delta x_n(t))^T$ is the state deviation vector, $\delta\boldsymbol{\mu} = (\delta\mu_1, \dots, \delta\mu_n)^T$ 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 $t \rightarrow \infty$ even when the state space for the network dynamics is bounded. Notice that $(1, \dots, 1)^T$ 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, since $\delta\mathbf{x}(t)$ 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

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

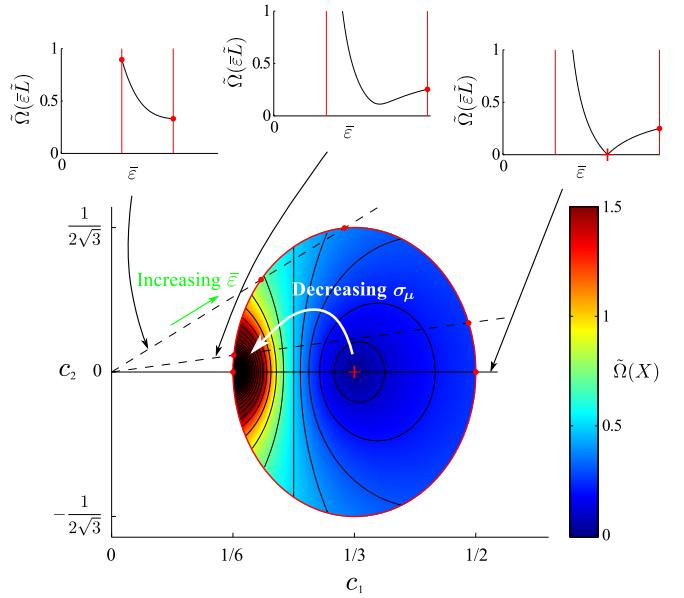


Fig. S1. Master synchronization error function $\tilde{\Omega}(X)$ defined on the space of matrices. The color indicates the values of $\tilde{\Omega}(X)$ with $X = \tilde{L}$ for the two-parameter family of networks given by $L = c_1 L_{K_3} + c_2 (L_{C_{3,1}} - L_{C_{3,2}})$ and the map $F(x, \mu) = 2x + \mu \bmod 1$. Shown in black are the level curves $\tilde{\Omega}(X) = \frac{E_{\text{tol}}}{\sigma_\mu}$. For fixed error tolerance E_{tol} and heterogeneity σ_μ , the region defined by $\tilde{\Omega}(X) \leq \frac{E_{\text{tol}}}{\sigma_\mu}$ 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 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).

Any component along $(1, \dots, 1)^T$ 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

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

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

$$\Omega_{\bar{\varepsilon}}(L) = \sigma_\mu \tilde{\Omega}(\bar{\varepsilon}\tilde{L}), \quad \text{where}$$

$$\tilde{\Omega}(X) := \limsup_{T \rightarrow \infty} \left\| \sum_{t=1}^T \left(\prod_{k=1}^{t-1} a_{k+T-t} \right) b_{T-t} (\tilde{L}^* - X)^t \right\|, \quad [\text{S15}]$$

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 ra-

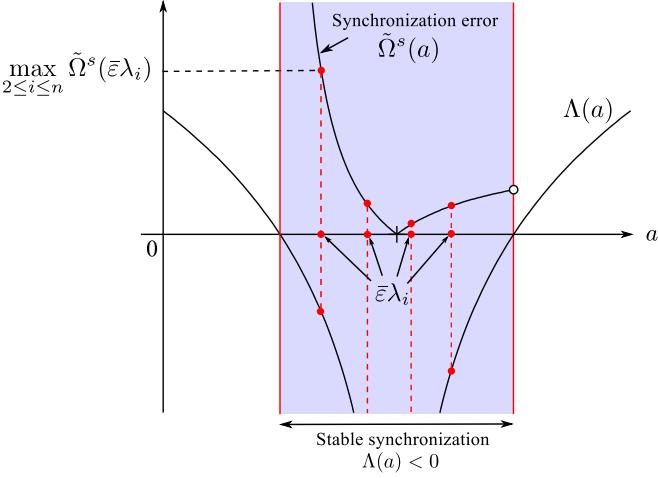


Fig. S2. Synchronization error function for undirected networks. The condition for the stability of synchronization is that the numbers $\tilde{\varepsilon}\lambda_2, \dots, \tilde{\varepsilon}\lambda_n$ (the red dots on the a -axis) all lie in the interval on which $\Lambda(a) < 0$. When this condition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

dius 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., $\nu := \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \ln |\frac{\partial F}{\partial x}(\bar{x}(t), \bar{\mu})|$.

If $X = \tilde{\varepsilon}\tilde{L}$, this condition reduces to the stability condition for complete synchronization of the corresponding identical units, namely, $\Lambda(\tilde{\varepsilon}\lambda_i) < 0$ for $i = 2, \dots, n$, where the stability function in this case is $\Lambda(a) = \nu + \ln|1-a|$. 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 $\Omega_{\tilde{\varepsilon}}(L) \leq E_{\text{tol}}$ is represented by the region $\{X : \tilde{\Omega}(X) \leq \frac{E_{\text{tol}}}{\sigma_\mu}\}$ 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 3 nodes and the two types of 3-cycles:

$$L_{K_3} = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}, \quad L_{C_{3,1}} = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix},$$

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

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

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

where $\lambda = 3c_1 + i\sqrt{3}c_2$. In particular, we have $\tilde{\Omega}(\tilde{L}) = |\frac{1-\lambda}{2\lambda-1}|$ 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 $\{X : \tilde{\Omega}(X) \leq \frac{E_{\text{tol}}}{\sigma_\mu}\}$ 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 $\Omega_{\tilde{\varepsilon}}(L)$ with respect to $\tilde{\varepsilon}$ can be understood as the change in the value of the error function $\tilde{\Omega}(X)$ as $X = \tilde{\varepsilon}\tilde{L}$ 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 $\Omega_{\tilde{\varepsilon}}(L)$ is expected to decrease to a minimum at some $\tilde{\varepsilon} = \tilde{\varepsilon}^*$ and increases monotonically for $\tilde{\varepsilon} > \tilde{\varepsilon}^*$, or monotonically decrease in the entire range of $\tilde{\varepsilon}$ for which $\Omega_{\tilde{\varepsilon}}(L)$ 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 $\Omega(L) := \inf_{\tilde{\varepsilon}} \Omega_{\tilde{\varepsilon}}(L)$ as a measure of synchronizability of a network, as it gives the lower limit on the asymptotic synchronization error for non-identical 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

$$\Omega_{\tilde{\varepsilon}}(L) = \sigma_\mu \tilde{\Omega}(\tilde{\varepsilon}\tilde{L}) = \sigma_\mu \max_{2 \leq i \leq n} \tilde{\Omega}^s(\tilde{\varepsilon}\lambda_i), \quad \text{where}$$

$$\tilde{\Omega}^s(a) = \limsup_{T \rightarrow \infty} \left| \sum_{t=1}^T \left(\prod_{k=1}^{t-1} a_{k+T-t} \right) b_{T-t} (1-a)^t \right| \quad [\text{S18}]$$

under the stability condition $\Lambda(\tilde{\varepsilon}\lambda_i) < 0$ for $i = 2, \dots, n$. For such networks, the synchronization error can be determined visually from the error function $\tilde{\Omega}^s(a)$, 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 $\tilde{\Omega}^s(a) = |\frac{1-a}{2a-1}|$ and $\Lambda(a) = \ln 2 + \ln|1-a|$.

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 $\lambda_2 = \dots = \lambda_n = \bar{\lambda} > 0$, 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 $\tilde{\varepsilon}$, we have $\delta\mathbf{x}(t) \rightarrow 0$ as $t \rightarrow \infty$ for arbitrary μ_1, \dots, μ_n with a given σ_μ . Then, letting $t \rightarrow \infty$ in Eq. S13, we conclude that we either have $\lim_{t \rightarrow \infty} b_t = \lim_{t \rightarrow \infty} \frac{\partial F}{\partial \mu}(\bar{x}(t), \bar{\mu}) = 0$ or $(\tilde{L}^* - \tilde{\varepsilon}\tilde{L})\delta\mu = \mathbf{0}$ for all possible $\delta\mu$. 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 $\tilde{L}^* - \tilde{\varepsilon}\tilde{L}$ must be equal to the zero matrix, and hence $\tilde{L} = \frac{1}{\tilde{\varepsilon}}\tilde{L}^*$, which is diagonalizable with eigenvalues $0, \frac{1}{\tilde{\varepsilon}}, \dots, \frac{1}{\tilde{\varepsilon}}$. Since 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 $\lambda_2 = \dots = \lambda_n = \bar{\lambda} > 0$.

Conversely, suppose that the network satisfies $\lambda_2 = \dots = \lambda_n = \bar{\lambda} > 0$ and the Laplacian matrix is diagonalizable. It can be shown that $\tilde{\varepsilon}\tilde{L} = \tilde{L}^*$ if we choose $\tilde{\varepsilon} = 1/\bar{\lambda}$, and therefore $\delta\mathbf{x}(t) = \mathbf{0}$ according to Eq. S13, but we can actually prove 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 $\tilde{\varepsilon} = 1/\bar{\lambda}$, then Eq. S11 becomes

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

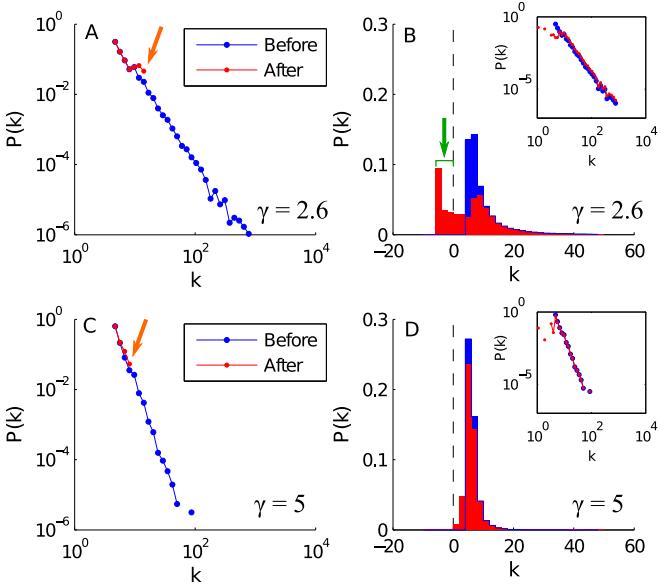


Fig. S3. Change in the degree distributions when enhancing synchronization with negative directional interactions. Distributions before (blue) and after (red) assigning negative strengths in random scale-free networks: (A) in-degree distribution for $\gamma = 2.6$, (B) out-degree distribution for $\gamma = 2.6$, (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.

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 super-stable fixed points and periodic orbits observed in maps.

Since the best networks for synchronizing non-identical maps satisfy $\lambda_2 = \dots = \lambda_n = \bar{\lambda} > 0$, they too must have a quantized number of links: $m = k(n - 1)$. For every n and every $k = 1, \dots, n$, there is exactly one binary network ($A_{ij} = 0, 1$) that has $m = k(n - 1)$ links and is capable of complete synchronization for non-identical 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 non-identical 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 $\lambda_2 = \dots = \lambda_n = \bar{\lambda} > 0$ 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 $A_{ij}^c = \alpha - b_j$, 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: $\lambda_2 = \dots = \lambda_n = n\alpha - \bar{\lambda} > 0$ and its Lapla-

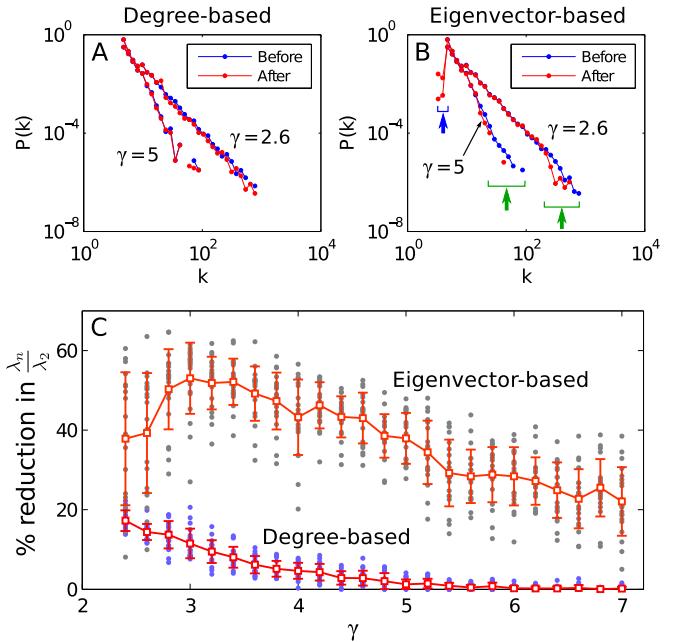


Fig. S4. Improving synchronization with negative bidirectional interactions. (A and B) Degree distributions before and after we apply two methods of assigning negative interactions to random scale-free networks with $\gamma = 2.6$ and 5 : (A) degree-based method and (B) eigenvector-based method. All 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.

cian 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. Figure 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 are compensated by the negative interactions, creating a sharp cut-off in the distribution (orange arrows in panels A and C). In contrast, the out-degree distributions remain essentially unchanged, having a power-law tail with the same exponent (insets in panels B and D). Note that the algorithm can create negative out-degree nodes, as indicated by the green arrow in panel 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. Figure S4A 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 non-identically 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]. Figure S4C 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 synchronization¹. From these networks we choose one with the smallest ratio λ_n/λ_2 as the output of the algorithm. Figure S4B 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,

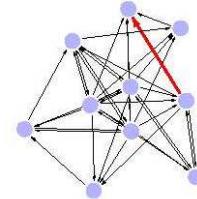
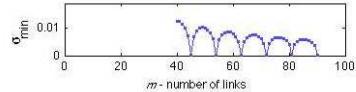
generated by the configuration model. For both values of γ , we notice a significant drop in the fraction of high-degree nodes (green arrows), accompanied by the appearance of nodes with degree less than the minimum degree of the initial network (blue arrow). This change appears to be responsible for the reduction of λ_n/λ_2 by as much as about 65%, as shown in Fig. S4C (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.

Supporting Video

<http://www.youtube.com/watch?v=3dMI1Yyxmbw>

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 [8]. We used the implementation of the algorithm in the Boost C++ library [9] through the Matlab interface provided by MatlabBGL [10].

After that, the curve goes through a series of cusps corresponding to optimal networks ($\sigma_{\min} = 0$).



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, Boltt 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.
8. Kamada T, Kawai S (1989) An algorithm for drawing general undirected graphs. *Inform Process Lett* 31:7–15.
9. Boost C++ library. <http://www.boost.org/>.
10. MatlabBGL. http://www.stanford.edu/~dgleich/programs/matlab_bgl/index.html.

¹ 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 .