Asymmetric Coupling Optimizes Interconnected Consensus Systems

Zhao Song1,∗ and Dane Taylor1,†

1Department of Mathematics, University at Buffalo, State University of New York, Buffalo, NY 14260, USA

(Dated: June 25, 2021)

Networks are often interconnected with one system wielding greater influence over another, yet we lack basic insights for how such an asymmetry affects collective dynamics. This is particularly important for group decision making within social networks supported by AI agents, which we model as coupled consensus systems. We show that coupling asymmetry can monotonically increase/decrease the convergence rate $\text{Re}(\lambda_2)$ of consensus or give rise to different types of optima that maximally accelerate consensus. For the coupling of faster and slower systems, we identify a bifurcation: if their timescales are very dissimilar, then the optimal asymmetry involves the faster system “dominating” the slower one; otherwise, they are optimized with an intermediate amount of asymmetry, and it is less clear which system should be more influential. Our findings support the design of human-AI decision systems and asymmetrically coupled networks, more broadly.

PACS numbers: 89.75.Fb, 87.23.Ge, 45.30.+s

Asymmetric plays a crucial role in shaping self-organization for network-coupled dynamical systems [1–5] and provides opportunities for optimization [6]. However, the affects of coupling asymmetry on interconnected systems is not well understood. This issue is particularly important for coupled decision system including, e.g., the scenario in which a group of persons interact via a social network and collectively make decisions with the support of AI agents (which themselves interact and cooperatively learn). A distopian-minded engineer would naturally design the social network to wield greater influence over the network of AI agents, but how can this be achieved? And how might such a system be optimized?

Thus motivated, we study asymmetrically coupled consensus systems using multiplex networks [7], which are composed of layers that encode different types of interactions. Consensus is a popular model for collective decision making in the cognitive, social and biological sciences [8–12], and it also provides a foundation for decentralized algorithms for neural networks and machine learning [13–18]. Thus, network layers in our model can encode different types of social ties (e.g., friendships and mentorships), technological communications (e.g., shared data or model parameters), or a combination thereof. We consider possibly directed multiplex networks using a generalization of the Laplacian matrix called a supralaplacian matrix [19–23] $\mathbb{L}(\omega, \delta)$, which we define in a novel way to isolate and study the effects of asymmetry. As illustrated in Fig. 1(A), asymmetry parameter $\delta \in [-1, 1]$ tunes the extent to which interlayer couplings are biased in a particular direction, and $\omega \geq 0$ tunes their coupling strength.

When $\mathbb{L}(\omega, \delta)$ is an irreducible matrix, coupled consensus systems converge to a fixed point [see Fig. 1(B)] at a rate that is bounded by the real part, $\text{Re}(\lambda_2)$, of the second-smallest eigenvalue $\lambda_2(\omega, \delta)$ of $\mathbb{L}(\omega, \delta)$ [see Fig. 1(C)]. Herein, we show that $\lambda_2$ can depend on $\delta$ in unintuitive ways. In the case of two coupled consensus systems, we show that if one system is much faster than another, then the optimal asymmetry is either $\delta = 1$ or $\delta = -1$, which allows the faster system to “dominate” in that it drives the slower one without feedback. Otherwise, there exists an optimum $\delta \in (-1, 1)$ at which consensus is maximally accelerated through an asymmetric balancing of the systems, and it’s less intuitive which system should be more influential. We provide theory for such a bifurcation to support the design of asymmetrically coupled systems, including emerging technologies that interface social networks with AI agents.

We formulate interconnected consensus systems using multiplex networks with $T$ network “layers”, each consisting of $N$ nodes. For each layer $t \in \{1, \ldots, T\}$, we let $A^{(t)}$ be its “intralayer” adjacency matrix and $L^{(t)} = D^{(t)} - A^{(t)}$ be its intralayer unnormalized Laplacian, where $D^{(t)}$ is a diagonal matrix that encodes the nodes’ weighted in-degrees $D_{ij}^{(t)} = \sum_j A_{ij}^{(t)}$ (also called ‘strengths’). Note that our con-

FIG. 1. Asymmetric coupling biases and accelerates consensus. (A) Two consensus systems are asymmetrically coupled with “interlayer” edges of weight $(1 \pm \delta)\omega$. (B) Nodes’ states $x_p(\tau)$ reach consensus following Eq. (2), which uses a supraLaplacian matrix $\mathbb{L}(\omega, \delta)$ with $(\omega, \delta) = (30, 0.5)$. (C) Convergence $x_p(\tau) \rightarrow \bar{x}$ occurs faster for $\delta = -0.5$ than for $\delta \in \{0, 0.5\}$. For each $\delta$, the convergence rate $\text{Re}(\lambda_2)$ gives the line’s slope when $\tau$ is large.
structure is to let $A_{ij} > 0$ encode the weight for an edge from node $j$ to $i$. Matrices $A(t)$ and $L(t)$ are size $N \times N$ and are asymmetric if layer $t$ contains directed edges.

We couple the layers using an “interlayer” adjacency matrix $A^{(t)}$, which has an interlayer unnormalized Laplacian $L^{(t)}(\delta) = D^{(t)}(\delta) - A^{(t)}(\delta)$, where $D^{(t)}(\delta)$ is a diagonal matrix with entries $D^{(t)}_{ii}(\delta) = \sum_{j} A^{(t)}_{ij}(\delta)$. Matrices $A^{(t)}(\delta)$ and $L^{(t)}(\delta)$ are size $T \times T$ and are asymmetric when the layers are asymmetrically coupled. We separate $L^{(t)}(\delta)$ into its symmetric and asymmetric parts: $L^{(t)}(\delta) = T^{(t)} + \delta \tilde{L}^{(t)}$, where $T^{(t)} = (L^{(t)} + [L^{(t)}]^T)/2$ and $\tilde{L}^{(t)} = (L^{(t)} - [L^{(t)}]^T)/2$. The interlayer coupling is “symmetric” when $\delta = 0$. We scale each $L^{(t)}(\delta)$ by $\omega \geq 0$ to construct a supralaplacian matrix [19–23]

$$L(\omega, \delta) = \mathbb{I}^{(t)} + \omega L^{(t)}(\delta),$$

where $\mathbb{I}^{(t)} = \text{diag}(L^{(1)}, \ldots, L^{(T)})$ contains intralayer Laplacians as diagonal blocks, and $L^{(t)}(\delta) = L^{(t)}(\delta) \otimes \mathbb{I}$ couples the layers in a way that is uniform (i.e., any coupling between two given layers is the same) and diagonal (i.e., any coupling between layers connects a node in one layer to itself in another layer). Symbol $\otimes$ indicates the Kronecker product.

We define the interconnected consensus dynamics by

$$\frac{d}{dt} \xi(t) = -L(\omega, \delta) \xi(t),$$

where $\xi(t) = [x_1(t), \ldots, x_N(t)]^T$ is a length-$NT$ vector. Each $x_p(t)$ encodes the state of node $i_p = (p \mod N)$ in layer $\ell_p = \lfloor p/N \rfloor$ at time $t$ for $p \in \{1, \ldots, NT\}$. We assume that the smallest eigenvalue $\lambda_1 = 0$ of $L(\omega, \delta)$ is simple, which is guaranteed, e.g., if the interlayer and intralayer adjacency matrices are associated with strongly connected networks [23]. Then the system converges to an equilibrium $\xi(t) \rightarrow \hat{\xi} = [\bar{\xi}, \ldots, \bar{\xi}]^T$, which is the right eigenvector of $\lambda_1$. Consensus is reached at a scalar value $\bar{\xi} = \sum_{p} u^{(1)}_{p0} x_p(0)/\sum_{p} u^{(1)}_{p0}$, which is a weighted average of the initial states, and the weights $u^{(1)}_{p0}$ are entries of the associated left eigenvector $\mathbf{u}^{(1)}$. The convergence rate, $-\limsup_{\epsilon(t) \rightarrow 0} \frac{1}{\epsilon(t)} \int_{0}^{T} \log \left(\frac{\xi(t) - \bar{\xi}}{|\xi(t) - \bar{\xi}|} \right) dt$, is bounded by the real part $\text{Re}(\lambda_2)$ of the eigenvalue of $L(\omega, \delta)$ that has second-smallest real part.

In Fig. 1, we show that asymmetry can bias the state to which convergence occurs as well as accelerate convergence. We study Eq. (2) for $T = 2$ layers with

$$L^{(t)}(\delta) = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \delta \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{3}$$

System 1 models group decision making with an empirical social network that encodes mentoring relationships among corporate executives [24]. It contains $N = 21$ nodes and 190 directed edges and was downloaded from [25]. System 2 represents AI agents that support the executives’ decisions, and we model their communication by a directed 4-regular graph.

In Fig. 1(B), we show convergent trajectories $x_p(t)$ with $(\omega, \delta) = (30, 0.5)$ for an initial condition where $x_p(t)$ are positive (negative) for nodes in layer 1 (2). Observe that the limit $\bar{\xi} > 0$ is biased positive, implying that the social network has a “stronger say” than the AI agents, which occurs in this case because $\delta > 0$. However, observe in Fig. 1(C) that convergence is slowest for this value of $\delta$, and consensus is faster if $\delta \in [0, -0.5]$. In a real-world setting, it would be advantageous to design the system so that both the convergence rate is optimal and the social network is more influential.

Next, we study the nonlinear effects on $\text{Re}(\lambda_2)$ of $\delta$ for the five interconnected consensus systems shown in Fig. 2(A). For each system, we plot $\text{Re}(\lambda_2)$ versus $\delta$ in Fig. 2(B), and we are particularly interested in the optimal asymmetry $\delta = \text{argmax}_\delta \text{Re}(\lambda_2)$ that maximally accelerates consensus. The columns of Fig. 2 highlight five behaviors:

(i) *layer dominance*: $\text{Re}(\lambda_2)$ monotonically increases or decreases with $\delta$ and obtains a maximum at either $\delta = \pm 1$;

(ii) *robust optimum*: $\text{Re}(\lambda_2)$ obtains a maximum at some value $\delta \in (-1, 1)$, and $\text{Re}(\lambda_2)$ is differentiable with respect to $\delta$ at the optimum;

(iii) *nonrobust optimum*: $\text{Re}(\lambda_2)$ obtains a maximum at some value $\delta \in (-1, 1)$, but $\text{Re}(\lambda_2)$ is not differentiable with respect to $\delta$ at the optimum;

(iv) *positive effect*: $\text{Re}(\lambda_2)$ is a non-decreasing function of $|\delta|$ (i.e., slowest convergence at $\delta = 0$);

(v) *negative effect*: $\text{Re}(\lambda_2)$ is a non-increasing function of $|\delta|$ (i.e., fastest convergence at $\delta = 0$).

Layer dominance occurs for the system shown in the first column of Fig. 2, since consensus over a star graph (layer 2) is much faster than consensus over a chain (layer 1). Specifically, $\lambda_2(2) > \lambda_2(1)$, where $\lambda_2(1) \approx 0.26$ and $\lambda_2(2) = 1$ are the second-smallest eigenvalues of their intralayer Laplacians $L^{(1)}$ and $L^{(2)}$, respectively. We point out that for all five systems, $\lambda_2(\omega, \delta)$ converges to $\lambda_2(1)$ and $\lambda_2(2)$ in the limits $\delta \rightarrow -1$ and $\delta \rightarrow 1$, respectively.

Importantly, when a system exhibits layer dominance, the optimal convergence involves one layer having no affect on the other (i.e., $\bar{\xi} \rightarrow \pm 1$), and so the coupled systems do not mutually interact. In contrast, consensus systems designed near an optimum $\delta \in (-1, 1)$ quickly converge and do mutually cooperate. The optima shown in the second and third columns of Fig. 2 imply that consensus occurs fastest when there is a precise amount of asymmetry, and one system is slightly more influential. We call an optimum “robust” if $\frac{d}{d\delta} \text{Re}(\lambda_2)$ is zero at the optimum, since near the optimum, the convergence rate would be minimally impacted by a small change to $\delta$.

SupraLaplacian matrices associated with the systems shown in the first, second and fifth columns were previously studied in [23], [22] and [19], respectively, but those works studied diffusion and were restricted to symmetric coupling, $\delta = 0$. In Supplementary Material, we plot $\text{Re}(\lambda_2)$ versus $\omega$ to reveal that $\text{Re}(\lambda_2)$ has a “peak” at an intermediate value of $\omega$ for the third, fourth and fifth system. This phenomenon has been previously called “superdiffusion” [19, 21, 26], and we show that asymmetry can amplify or inhibit such a peak.
(which has practical implications for both coupled consensus systems and the study of diffusion). We also show that the nonrobust optima arise here due to a spectral bifurcation in which two eigenvalues, $\lambda_2$ and $\lambda_3$, of $L(\omega, \delta)$ collide to create a complex pair of eigenvalues.

To provide analytical guidance, we characterize the dependence of $\text{Re}(\lambda_2)$ on $\delta$ using perturbation theory for directed multilayer networks [23, 27]. We present the derivations in Supporting Material and summarize our findings here. The black curves in Fig. 2(B) depict our predictions for large $\omega$:

$$\lim_{\omega \to \infty} \lambda_2(\omega, \delta) = \overline{\lambda}_2(\delta),$$

where $\overline{\lambda}_2(\delta)$ is the eigenvalue of $\overline{L}(\delta)$ that has the smallest real part, and

$$\overline{L}(\delta) = \sum_{i=1}^{T} \omega_i(\delta) L_i^{(i)}$$

is a weighted average of the layers’ Laplacian matrices. The weights $\omega_i(\delta) = u_i(\delta)/\sum_i u_i(\delta)$ come from the entries of the left eigenvector $u(\delta) = [u_1(\delta), \ldots, u_T(\delta)]^T$ that is associated with the zero-valued (i.e., trivial) eigenvalue of $L^I(\delta)$.

Equations (4)-(5) imply that when consensus systems are strongly coupled, the convergence rate is identical to that for consensus on an “effective” network that is associated with Laplacian $\overline{L}(\delta)$, and the effects of $\delta$ can be examined by considering the dependence of $\omega_i(\delta)$ on $\delta$. For example, for $T = 2$ layers, the interlayer Laplacian $L^I(\delta)$ is given by Eq. (3), $u(\delta) = [1 + \delta, 1 - \delta]^T$ and

$$\overline{L}(\delta) = \left(\frac{1 + \delta}{2}\right) L^{(1)} + \left(\frac{1 - \delta}{2}\right) L^{(2)}.$$  

Despite this simple form, the associated convergence rate $\text{Re}(\overline{\lambda}_2)$ can exhibit a complicated dependence on $\delta$. Moreover, observe in Fig. 2(B) that in addition to being accurate for large $\omega$, this theory predicts the qualitative behavior of the relationship between $\text{Re}(\lambda_2)$ and $\delta$ for a broad range of $\omega$.

Understanding whether a system exhibits an optimum versus layer dominance is important for human-AI systems, and likely other applications including diffusion. We now show that the presence of an optimum depends crucially on the coupled systems’ relative timescales. Again focusing on $T = 2$ layers, we introduce a rate-scaling parameter $\chi \in (0, 1)$ to vary the relative convergence rate for each system. We replace the intralayer Laplacians by $L^{(1)} \mapsto \chi L^{(1)}$ and $L^{(2)} \mapsto (1 - \chi) L^{(2)}$. Varying $\chi$ controls whether the systems’ consensus rates are balanced ($\chi \approx 0.5$), whether system 1 is much faster than system 2 ($\chi \approx 1$), or vice versa ($\chi \approx 0$). We insert these weighted Laplacians into Eq. (6) to obtain $\overline{L}(\delta) = \frac{1+\delta}{2} \chi L^{(1)} + \frac{1-\delta}{2} (1 - \chi) L^{(2)}$, and then study how a system’s behavior (i)-(v) depends on both $\chi$ and $\delta$.

We predict the existence/nonexistence of an optimum by considering the derivative $\frac{d}{d\delta} \overline{\lambda}_2(\delta)$ and by invoking Rolle’s theorem [28] for a continuous function: if $\overline{\lambda}_2(-1) > 0$ and $\overline{\lambda}_2(1) < 0$, then there exists at least one optimum $\delta \in (-1, 1)$ that maximizes $\text{Re}(\overline{\lambda}_2)$. As $\delta \to \pm 1$, we find that $\overline{\lambda}_2(\delta)$ converges to a simplified form:

$$\overline{\lambda}_2(1) = \frac{-\sqrt{u^{(1)}} L^{(2)} v^{(1)}}{2u^{(1)} v^{(1)}} + \chi u^{(1)*} \left( L^{(1)} + L^{(2)} \right) v^{(1)}$$

$$\overline{\lambda}_2(-1) = \frac{-\lambda_2^{(2)}}{2} + \chi \frac{u^{(2)*} \left( L^{(1)} + L^{(2)} \right) v^{(2)}}{2u^{(2)} v^{(2)}},$$

where $u^{(i)}$ and $v^{(i)}$ are the left and right eigenvector for the eigenvalue $\lambda_2^{(i)}$ of $L^{(i)}$ that has the second-smallest real part (assumed to be nonzero). Symbol $\ast$ denotes a vector’s complex conjugate. To apply Rolle’s theorem, for both $\delta \to 1$ and $\delta \to -1$, we identify the value of $\chi$ where $\text{Re}(\overline{\lambda}_2(\pm 1))$
changes sign by setting $\text{Re}(\lambda_2' (\pm 1)) = 0$. This reveals a range for which an optimum exists: $\chi \in (\hat{\chi}(-1), \hat{\chi}(1))$, where

$$\hat{\chi}(1) = \text{Re} \left( \frac{u^{(1)*} L^{(2)} v^{(1)}}{u^{(1)*} (L^{(1)} + L^{(2)}) v^{(1)}} \right),$$

$$\hat{\chi}(-1) = \text{Re} \left( \frac{u^{(2)*} L^{(2)} v^{(2)}}{u^{(2)*} (L^{(1)} + L^{(2)}) v^{(2)}} \right).$$

That is, coupled consensus systems must have sufficiently similar timescales (i.e., $\chi$ is neither too large or small) if their optimal coupling is not the situation where one system dominates the other. Note that this criterion also guarantees that the optimal convergence rate is faster that that of either system.

In Fig. 3, we study this phenomenon for the human-AI system from Fig. 1. $\chi$ tunes the relative timescale for communications among humans and among AI agents. In Fig. 3(A), we plot $\bar{\lambda}_2(\delta)$ versus $\delta$ for several choices of $\chi$. Observe the property of layer dominance when $\chi \in [0.05, 0.85]$ and optima when $\chi \in [0.25, 0.45, 0.65]$, which is in agreement with our theory. See the shaded region in Fig. 3(B) for the range $\chi \in (\hat{\chi}(-1), \hat{\chi}(1))$. Solid and dashed black lines in Fig. 3(B) illustrate Eqs. (7), and their intersections with the horizontal axis give rise to Eqs. (8). An optimum exists when $\chi \in (\hat{\chi}(-1), \hat{\chi}(1))$, as shown by the shaded region. (C) Color depicts $\text{Re}(\hat{\lambda}_2)$ for the $(\delta, \chi)$ parameter space. The dashed white curve shows the optimal asymmetry $\delta$ for each value of $\chi$, and the purple arrow highlights the overall optimum $(\delta, \hat{\chi})$. (D) Convergence rate for each separate system versus $\chi$.

In summary, emergent technologies that interface social networks and AI demand a deeper understanding for how coupling asymmetry impacts interconnected consensus systems. We provided a categorization of some behaviors (i)–(v) and showed how these crucially depend on asymmetric coupling of systems and their relative timescales. Our findings can aid the design of human-AI decision systems, and they open the door for further theory development studying the nonlinear effects of coupling asymmetry on interconnected network dynamics. See [29] for a codebase that models interconnected consensus systems and reproduces our findings.

We thank Sarah Muldoon, Naoki Masuda, Per Sebastian Skardal, and Malbor Asllani for helpful feedback. ZS and DT were supported in part by the NSF (DMS-2052720), Simons Foundation (award #578333) and Nakatani Foundation.

* zhaosong@buffalo.edu
† danet@buffalo.edu
\[1\] M. Timme, EPL (Europhysics Letters) 76, 367 (2006).
\[2\] J. G. Restrepo and E. Ott, EPL (Europhysics Letters) 107, 60006 (2014).
\[3\] P. S. Skardal, D. Taylor, J. Sun, and A. Arenas, Physical Review E 91, 010802(R) (2015).
\[4\] T. Nishikawa and A. E. Motter, Physical Review Letters 117, 114101 (2016).
\[5\] M. Asllani, R. Lambiotte, and T. Carletti, Science Advances 4, eaau9403 (2018).
\[6\] D. Taylor, P. S. Skardal, and J. Sun, Chaos 060401 (2020).
\[7\] E. Cozzo, G. F. De Arruda, F. A. Rodrigues, and Y. Moreno, Multiplex Networks: Basic Formalism and Structural Properties (Springer, 2018).
\[8\] V. B. Hinsz, Journal of Personality and Social Psychology 59, 705 (1990).
\[9\] C. M. Fiol, Organization Science 5, 403 (1994).
\[10\] P. C. Flood, E. Hannan, K. G. Smith, T. Turner, M. A. West, and J. Dawson, European Journal of Work and Organizational Psychology 9, 401 (2000).
\[11\] S. Mohammed, The Journal of Applied Behavioral Science 37,
[12] L. Conradt and T. J. Roper, Trends in Ecology & Evolution 20, 449 (2005).
[13] A. S. Bijral, A. D. Sarwate, and N. Srebro, IEEE Transactions on Automatic Control 62, 4483 (2017).
[14] M. Assran, N. Loizou, N. Ballas, and M. Rabbat, in International Conference on Machine Learning (PMLR, 2019) pp. 344–353.
[15] K. Niwa, N. Harada, G. Zhang, and W. B. Kleijn, in Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (2020) pp. 668–678.
[16] T. Vogels, S. P. Karimireddy, and M. Jaggi, arXiv preprint arXiv:2008.01425 (2020).
[17] B. Huynh, H. Dutta, and D. Taylor, arXiv preprint arXiv:2011.01334 (2021).
[18] L. Kong, T. Lin, A. Koloskova, M. Jaggi, and S. U. Stich, arXiv preprint arXiv:2102.04828 (2021).
[19] A. Tejedor, A. Longjas, E. Foufoula-Georgiou, T. T. Georgiou, and Y. Moreno, Physical Review X 8, 031071 (2018).
[20] A. Sole-Ribaltà, M. De Domenico, N. E. Kouvaris, A. Díaz-Guilera, S. Gomez, and A. Arenas, Physical Review E 88, 032807 (2013).
[21] X. Wang, A. Tejedor, Y. Moreno, et al., New Journal of Physics (2021).
[22] S. Gómez, A. Díaz-Guilera, J. Gómez-Gardeñes, C. J. Pérez-Vicente, Y. Moreno, and A. Arenas, Physical Review Letters 110, 028701 (2013).
[23] D. Taylor, Physical Review Research 2, 033164 (2020).
[24] D. Krackhardt, Social Networks 9, 109 (1987).
[25] M. Domenico, “Datasets released for reproducibility; available at https://manliodedomenico.com/data.php.”.
[26] G. Cencetti and F. Battiston, New Journal of Physics 21, 035006 (2019).
[27] D. Taylor, M. A. Porter, and P. J. Mucha, Multiscale Modeling & Simulation 15, 537 (2021).
[28] P. Sahoo and T. Riedel, Mean Value Theorems and Functional Equations (World Scientific, 1998).
[29] Z. Song, “Codebase for Asymmetrically coupled consensus systems https://github.com/zhaos426/Interconnected_Consensus_Systems_with_Asymmetric_Coupling.”.
This supplementary document contains the following extended results: (Appendix A) further investigation of the nonrobust optima; (Appendix B) impact of asymmetric coupling on “superdiffusion” phenomenon; (Appendix C) derivation of Eqs. (4) and (5) in the main text; and (Appendix D) derivation of Eqs. (6)–(8) in the main text.

APPENDIX A: FURTHER EXPERIMENTS FOR NONROBUST OPTIMA

Here, we further investigate Re(\(\lambda_2\)) for the multiplex network shown in the center column of Fig. 2(A). In the center column of Fig. 2(B), we plotted Re(\(\lambda_2\)) versus \(\delta\) for these interconnected consensus systems and observed two optima near \(\delta = \pm 0.35\). We classify these as “nonrobust” optima, because the derivative \(\frac{d}{d\delta}\) Re(\(\lambda_2\)) is undefined (i.e., discontinuous) at these optima. Here, we show that these two nonrobust optima (and the lack of differentiability) arise due to a spectral bifurcation in which the two eigenvalues \(\lambda_2\) and \(\lambda_3\) of \(L(\omega, \delta)\) collide and give rise to a complex pair of eigenvalues. Here, we have defined \(\lambda_2\) and \(\lambda_3\) as the eigenvalues of \(L(\omega, \delta)\) with second-smallest and third-smallest real part, respectively.

In Fig. SM1, we plot \(\lambda_2\) and \(\lambda_3\) for a supraLaplacian \(L(\omega, \delta)\) associated with the multiplex network shown in the center column of Fig. 2(A) with \(\omega = 30\) for various \(\delta\). In panels (A) and (B) of Fig. SM1, we depict the real and imaginary parts, respectively, of these eigenvalues. The symbols depict the observed values that are directly computed using \(L(\omega, \delta)\), whereas the solid and dashed black curves indicate our theoretical predictions for \(\lambda_2\) and \(\lambda_3\), respectively, when \(\omega\) is large. The prediction for \(\lambda_2\) is given by Eqs. (4)–(6) in the main text and is further discussed below in Appendix B. Our prediction for \(\lambda_3\) is obtained using a similar technique; we find that it is the eigenvalue of \(L(\delta)\) that has the third-smallest real part. Note that the observed and predicted eigenvalues are in excellent agreement.

Observe in Fig. SM1(A) that the theoretical curve for \(\lambda_2\) is identical to the one that is in the center column of Fig. 2(B). The curve has two optima near \(\delta = \pm 0.35\), and these values coincide with spectral bifurcations in which \(\lambda_2\) and \(\lambda_3\) change from being distinct real-valued eigenvalues to being a complex pair of eigenvalues (or vice versa). That is, for \(|\delta| \leq 0.35\), the eigenvalues \(\lambda_2\) and \(\lambda_3\) are purely real (i.e., \(\text{Im}(\lambda_2) = \text{Im}(\lambda_3) = 0\) and \(\text{Re}(\lambda_3) > \text{Re}(\lambda_2)\)). In contrast, for \(|\delta| > 0.35\), the eigenvalues \(\lambda_2\) and \(\lambda_3\) are complex numbers and \(\text{Re}(\lambda_3) = \text{Re}(\lambda_2)\). Thus, the nonrobust optima at \(\delta = \pm 0.35\) arise for these interconnected consensus systems because of a spectral bifurcation.

![Fig. SM1](image-url)

**FIG. SM1.** A collision between eigenvalues \(\lambda_2\) and \(\lambda_3\) yields a spectral bifurcation and nonrobust optima. We plot (A) the real parts and (B) imaginary parts of \(\lambda_2\) and \(\lambda_3\) versus \(\delta\) for the interconnected consensus systems shown in the center column of Fig. 2(A). The symbols depict the observed eigenvalues that are directly computed using \(L(\omega, \delta)\) with \(\omega = 30\), whereas the black curves depict our theoretical predictions. The nonrobust optima at \(\delta = \pm 0.35\) coincide with a spectral bifurcation in which \(\lambda_2\) and \(\lambda_3\) collide to yield a complex pair of eigenvalues.
APPENDIX B: FURTHER EXPERIMENTS FOR OPTIMAL COUPLING

Here, we turn our attention to studying the effects of interlayer coupling asymmetry on a recently observed phenomenon for supraLaplacians called “superdiffusion” [1–3], whereby the convergence rate Re(λ₂) exhibits a maximum (or “peak”) for some intermediate value of coupling strength ω. Although we do not study diffusion in this paper, our formulation for interconnected consensus also uses a supraLaplacian matrix, and so the effects of ω on Re(λ₂) are also very relevant for coupled consensus systems. Because this graph-spectral phenomenon has broader applications than diffusion, herein we will refer to this phenomenon as the existence of an intermediate coupling optimum (ICO). We also highlight that until now, this phenomenon was only studied for supraLaplacians associated with symmetrically coupled network layers. We now extend the study of superdiffusion to the setting of asymmetrically coupled network layers.

In Fig. SM2 below, we show that the ICO phenomenon can be either exaggerated or inhibited by coupling asymmetry, which we tune using an asymmetry parameter δ. We study the same five systems that are shown in Fig. 2(A) of the main text, and we show results that are similar to those in Fig. 2(B). However, we now plot Re(λ₂) versus ω so that the ICO phenomenon can be more easily observed.

![Fig. SM2. Effects of asymmetry on the phenomenon of an intermediate coupling optimum (ICO). For each of 5 different systems shown in Fig. 2(A) of the main text, we now plot the convergence rate Re(λ₂) versus coupling strength ω. Each panel corresponds to a different system, and the different colored curves correspond to different choices for the asymmetry parameter δ, as shown in the legend.](image)

As discussed for Fig. 2 in the main text, the first and second columns in Fig. SM2 corresponds to systems that exhibit layer dominance and a robust optimum, respectively. These systems do not exhibit an ICO phenomenon since Re(λ₂) monotonically increases with ω for all curves in these two panels. However, observe in the third, fourth and fifth panels that many curves exhibit a “peak” at some intermediate value of ω. In the context of interconnected consensus systems, such a peak corresponds to an optimal choice of coupling strength at which the convergence rate is maximal and consensus is maximally accelerated. Thus, in addition to optimizing δ and χ (see, e.g., Fig. 3 in the main text), it can also be beneficial to simultaneously optimize ω for an interconnected consensus system. Finally, observe that the peaks are largest when δ ≈ −1 and smallest when δ ≈ 1, implying that asymmetry can either enhance or inhibit the ICO phenomenon.

We conclude by highlighting one final phenomenon for this experiment. Namely, observe that when ω is small, all five systems exhibit the same behavior: λ₂ converges to 0 as ω → 0⁺. Interestingly, for each system there exists some critical value of ω below which δ has no observable effect on Re(λ₂). In contrast, δ has a significant effect on all five systems when ω is sufficiently large. This property will be explored in future work.

APPENDIX C: DERIVATION OF EQUATIONS (4) AND (5) IN THE MAIN TEXT

Here, we provide our derivation of Eqs. (4) and (5) in the main text, which predict the eigenvalue λ₂ of L(ω, δ) in the limit of large ω. We will use perturbation theory for directed multiplex networks that is similar to that which was developed in [4, 5]. First, we introduce a change of variables ε = 1/ω and multiply both sides of Eq. (1) in the main text by ε to obtain

\[
\hat{L}(\epsilon, \delta) = \epsilon L(\epsilon^{-1}, \delta) = \epsilon L^1 + \epsilon L^2(\delta).
\]  

(SM1)

Because we’ve only scaled the matrix L(ω, δ) by ε, it follows that \( \lambda_2(\epsilon, \delta) = \epsilon \lambda_2 \) is an eigenvalue of \( \hat{L}(\epsilon, \delta) \). Let \( \tilde{u}(\epsilon, \delta) \) and \( \tilde{v}(\epsilon, \delta) \) be the associated left and right eigenvectors of \( \hat{L}(\epsilon, \delta) \). Note that scalar multiplication does not change the eigenvectors of a matrix, and so \( \tilde{u}(\epsilon, \delta) \) and \( \tilde{v}(\epsilon, \delta) \) are also the eigenvectors of \( L(\omega, \delta) \) that are associated with \( \lambda_2 \). The main motivation for this transformation is that we can more easily study \( \lambda_2 \) in \( \omega \to \infty \) limit by instead studying \( \lambda_2(\epsilon, \delta) \) as \( \epsilon \to 0^+ \). In this limit,

\[
\hat{L}(\epsilon, \delta) \to L^1(\delta) = L^1(\delta) \otimes I.
\]  

(SM2)
To proceed, we first establish some properties about the eigenvalues and eigenvectors of \( L^1(\delta) \otimes I \).

**Lemma 1.** Let \( \mu \) be an eigenvalue of \( L^1(\delta) \in \mathbb{R}^{T \times T} \) and \( \mathbf{u} \) and \( \mathbf{v} \), respectively, be its associated left and right eigenvectors. Furthermore, let \( \mathbf{e}^{(i)} \in \mathbb{R}^N \) denote the \( i \)-th unit vector such that all entries are zeros, expect for entry \( i \), which is a one. It then follows that \( \mu \) is an eigenvalue of \( L^1(\delta) \otimes I \) and it has associated left and right eigenvectors given by

\[
\mathbf{u}^{(i)} = \mathbb{P} \left( \mathbf{e}^{(i)} \otimes \mathbf{u} \right), \quad \mathbf{v}^{(i)} = \mathbb{P} \left( \mathbf{e}^{(i)} \otimes \mathbf{v} \right),
\]

where \( \mathbb{P} \) is a “stride permutation matrix” that contains entries

\[
P_{ij} = \begin{cases} 1, & j = \lfloor i/N \rfloor + T[(i-1) \mod N] \\ 0, & \text{otherwise}. \end{cases}
\]

**Remark 1.** As discussed in [4, 6], this stride permutation is a unitary matrix that changes the ordering of indices for a supramatrix associated with a multiplex network. That is, the indices originally count by nodes and then layers, but after applying \( \mathbb{P} \), the counting is first by layers and then by nodes.

**Remark 2.** Lemma 1 is true for any choice \( i \in \{1, \ldots, N\} \), and so each eigenvalue \( \mu \) of \( L^1(\delta) \otimes I \) has an eigenspace that is at least \( N \)-dimensional. [It may be larger if \( \mu \) is a repeated eigenvalue of \( L^1(\delta) \).]

**Proof.** The stride permutation yields an identity

\[
L^1(\delta) \otimes I = \mathbb{P} \left( I \otimes L^1(\delta) \right) \mathbb{P}^T,
\]

where

\[
I \otimes L^1(\delta) = \begin{pmatrix}
L^1(\delta) & 0 & \cdots & 0 \\
0 & L^1(\delta) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & L^1(\delta)
\end{pmatrix}.
\]

Since \( \mathbb{P} \) is a unitary matrix, the eigenvalues of \( L^1(\delta) \otimes I \) are identical to those of \( I \otimes L^1(\delta) \). Moreover, if \( \mathbf{v} \) is a right eigenvector of \( I \otimes L^1(\delta) \) with eigenvalue \( \mu_2 \), then \( \mathbb{P} \mathbf{v} \) is a right eigenvector of \( L^1(\delta) \otimes I \). This can easily be checked:

\[
\left( L^1(\delta) \otimes I \right) \mathbb{P} \mathbf{v} = \mathbb{P} \left( I \otimes L^1(\delta) \right) \mathbb{P}^T \mathbb{P} \mathbf{v} = \mathbb{P} \left( I \otimes L^1(\delta) \right) \mathbf{v} = \mu \mathbb{P} \mathbf{v}.
\]

One can similarly show that if \( \mathbf{u} \) is a left eigenvector for \( I \otimes L^1(\delta) \), then \( \mathbb{P} \mathbf{u} \) is one for \( L^1(\delta) \otimes I \). What remains for us to show is that \( \mathbf{u} = \mathbf{e}^{(i)} \otimes \mathbf{u} \) and \( \mathbf{v} = \mathbf{e}^{(i)} \otimes \mathbf{v} \) are left and right eigenvectors of \( I \otimes L^1(\delta) \). We prove this using a standard property for the product of two Kronecker products:

\[
\left( I \otimes L^1(\delta) \right) \left( \mathbf{e}^{(i)} \otimes \mathbf{v} \right) = \left( I \mathbf{e}^{(i)} \right) \otimes \left( L^1(\delta) \mathbf{v} \right) = \mathbf{e}^{(i)} \otimes \mu \mathbf{v} = \mu \left( \mathbf{e}^{(i)} \otimes \mathbf{v} \right).
\]

A similar result can be obtained for the left eigenvector \( \mathbf{e}^{(i)} \otimes \mathbf{u} \).

Having established basic results for the spectral properties of \( L^1(\delta) \otimes I \) in Lemma 1, we are now ready to study the eigenvalue \( \tilde{\lambda}_2(\epsilon, \delta) \) of \( L(\epsilon, \delta) \) that has the second-smallest real part in the limit \( \epsilon \to 0^+ \). We formalize this result with in following theorem.

**Theorem 3.** Assume that a supraLaplacian \( L(\omega, \delta) \) and its associated intralayer Laplacians \( \{L^{(i)}\} \) all correspond to strongly connected graphs. Further, let \( \mathbf{u}(\delta) = [u_1(\delta), \ldots, u_T(\delta)]^T \) be the left eigenvector of an interlayer Laplacian \( L^1(\delta) \) that is associated with the zero-valued (i.e., trivial) eigenvalue. Then the eigenvalue \( \lambda_2 \) of \( L(\omega, \delta) \) that has second-smallest real part has the following limit:

\[
\lim_{\omega \to \infty} \lambda_2(\omega, \delta) = \tilde{\lambda}_2(\delta),
\]

where \( \tilde{\lambda}_2(\delta) \) is the second-smallest eigenvalue of \( L^1(\delta) \).
where $\tilde{\lambda}_2(\delta)$ is the eigenvalue of matrix

$$\tilde{L}(\delta) = \sum_{t=1}^{T} w_t(\delta) L(t) \quad \text{(SM9)}$$

that has the smallest real part, and $w_t(\delta) = u_t(\delta)/\sum u_t(\delta)$.

**Proof.** Consider second-order Taylor expansions for the eigenvalue $\tilde{\lambda}_2(\epsilon, \delta)$ of $\tilde{L}(\epsilon, \delta)$ and its associated left and right eigenvectors:

$$\tilde{\lambda}_2(\epsilon, \delta) = \tilde{\lambda}_2(0, \delta) + \epsilon \tilde{\lambda}'_2(0, \delta) + \mathcal{O}(\epsilon^2),$$

$$\tilde{u}(\epsilon, \delta) = \tilde{u}(0, \delta) + \epsilon \tilde{u}'(0, \delta) + \mathcal{O}(\epsilon^2),$$

$$\tilde{v}(\epsilon, \delta) = \tilde{v}(0, \delta) + \epsilon \tilde{v}'(0, \delta) + \mathcal{O}(\epsilon^2). \quad \text{(SM10)}$$

Note that we have defined the derivatives

$$\tilde{\lambda}'_2(\epsilon, \delta) \equiv \frac{d}{d\epsilon} \tilde{\lambda}_2(\epsilon, \delta), \quad \tilde{u}'(\epsilon, \delta) \equiv \frac{d}{d\epsilon} \tilde{u}(\epsilon, \delta), \quad \tilde{v}'(\epsilon, \delta) \equiv \frac{d}{d\epsilon} \tilde{v}(\epsilon, \delta). \quad \text{(SM11)}$$

We first consider the term $\tilde{\lambda}_2(0, \delta)$. Since we assumed $L^1$ to be the Laplacian of a strongly connected graph, the smallest eigenvalue $\mu_1 = 0$ of $L^1$ is guaranteed to be a simple eigenvalue with multiplicity 1. Lemma 1 implies $\mu_1 = 0$ is an eigenvalue of $L^1(\delta) \otimes I$ with an $N$-dimensional eigenspace. Hence, matrix $L(\epsilon, \delta)$ has $N$ eigenvalues that converge to 0 as $\epsilon \to 0^+$ (and in fact, one of these eigenvalues is always equal to zero). By definition (i.e., since it has the smallest, positive real part), the eigenvalue $\tilde{\lambda}_2(\epsilon, \delta)$ of $\tilde{L}(\epsilon, \delta)$ must one of these eigenvalues, which implies that $\tilde{\lambda}_2(0, \delta) = 0$.

Next, we consider the derivative term $\tilde{\lambda}'_2(0, \delta)$. We will show that it equals the second smallest eigenvalue of the matrix defined in Eq. (SM9). To this end, we consider the eigenvalue equation

$$\tilde{L}(\epsilon, \delta)\tilde{v}(\epsilon, \delta) = \tilde{\lambda}_2(\epsilon, \delta)\tilde{v}(\epsilon, \delta), \quad \text{(SM12)}$$

and we expand all terms to first order to obtain

$$\left[\epsilon I + L^1(\delta)\right] [\tilde{v}(0, \delta) + \epsilon \tilde{v}'(0, \delta)] = \left[\tilde{\lambda}_2(0, \delta) + \epsilon \tilde{\lambda}'_2(0, \delta)\right] [\tilde{v}(0, \delta) + \epsilon \tilde{v}'(0, \delta)]. \quad \text{(SM13)}$$

The zeroth-order and first-order terms must be consistent, which gives rise to two separate equations. The equation associated with the zeroth-order terms yields an eigenvalue equation

$$L^1(\delta)\tilde{v}(0, \delta) = \tilde{\lambda}_2(0, \delta)\tilde{v}(0, \delta). \quad \text{(SM14)}$$

Lemma 1 implies that the eigenvalue $\tilde{\lambda}_2(0, \delta) = 0$ has an $N$-dimensional right eigenspace spanned by right eigenvectors having the form $\tilde{v}^{(i)} = \mathbb{P} \left(e^{(i)} \otimes v\right)$ for $i \in \{1, 2, \ldots, N\}$. We similarly define $\tilde{u}^{(i)} = \mathbb{P} \left(e^{(i)} \otimes u\right)$ for the left eigenspace. Note that the vectors $u$ and $v$ contain entries that are nonnegative, which can be proved using the Perron–Frobenius theorem.

Next, we expand $\tilde{v}(0, \delta)$ in this eigenbasis as

$$\tilde{v}(0, \delta) = \sum_{i} \hat{\alpha}_i \tilde{v}^{(i)}. \quad \text{(SM15)}$$

We define a vector of coordinates $\hat{\alpha} = [\hat{\alpha}_1, \ldots, \hat{\alpha}_N]^T$ that must be determined, and we note that the vector must be normalized with $||\hat{\alpha}||_2 = 1$. The first-order terms in Eq. (SM13) give rise to a linear equation

$$L^1(\delta)\tilde{v}(0, \delta) + L^1(\delta)\tilde{v}'(0, \delta) = \tilde{\lambda}'_2(0, \delta)\tilde{v}(0, \delta), \quad \text{(SM16)}$$

which has used that $\tilde{\lambda}_2(0, \delta) = 0$.

To solve for $\tilde{\lambda}'_2(0, \delta)$, we left multiple by a left eigenvector $\tilde{u}^{(i')} = \mathbb{P} \left(e^{(i')} \otimes u\right)$ of $L^1(\delta)$ and again use $\tilde{\lambda}_2(0, \delta) = 0$ to obtain

$$\left[\tilde{u}^{(i')}\right]^T L^1(\delta)\tilde{v}(0, \delta) = \tilde{\lambda}'_2(0, \delta) \left[\tilde{u}^{(i')}\right]^T \tilde{v}(0, \delta). \quad \text{(SM17)}$$
Using the general form of $\hat{\psi}(0, \delta)$ from Eq. (SM15), we obtain

$$
\sum_{i=1}^{N} \tilde{\alpha}_i \left[ \hat{u}(i') \right]^T \mathbb{L} \hat{\psi}(i) = \tilde{\lambda}_2(0, \delta) \sum_{i=1}^{N} \tilde{\alpha}_i \left[ \hat{u}(i') \right]^T \hat{\psi}(i)
$$

$$
= \tilde{\lambda}_2(0, \delta) \sum_{i=1}^{N} \tilde{\alpha}_i \left[ \hat{u}(i') \right]^T \hat{\psi}(i). \tag{SM18}
$$

This system is identical to the following eigenvalue equation \( \mathbf{L}(\delta) \tilde{\alpha} = \tilde{\lambda}_2(0, \delta) \tilde{\alpha} \), where

$$
\mathbf{L}(\delta) \equiv \left[ \left[ \hat{u}(i') \right]^T \mathbb{L} \hat{\psi}(i) \right] \tag{SM19}
$$

and \( \tilde{\lambda}_2(0, \delta) = \lambda_2(\delta) \) is the eigenvalue with smallest real part. We can further simplify this result using the definitions of \( \hat{u}(i') = \mathbb{P} \left( \mathbf{e}(i') \otimes u \right) \) and \( \hat{\psi}(i) = \mathbb{P} \left( \mathbf{e}(i) \otimes v \right) \) to obtain

$$
\mathbf{L}(\delta) \equiv \left[ \left[ \hat{u}(i') \right]^T \mathbf{L}(i', i) \hat{\psi}(i) \right] \tag{SM20}
$$

where

$$
\mathbf{L}(i', i) = \text{diag} \left( \mathbf{L}_1^{(1)}, \mathbf{L}_2^{(2)}, \ldots, \mathbf{L}_1^{(T)} \right) \tag{SM21}
$$

and \( \mathbf{L}_1^{(t)} \) denotes the \((i', i)\)-component of intralayer Laplacian \( \mathbf{L}_1^{(t)} \). Eq. (SM20) follows after using the definition of \( \mathbb{P} \), which is a unitary matrix that permutes the enumeration of nodes and layers as described for Lemma 1. Eq. (SM21) follows after using that the right eigenvector \( v \) that is associated with the zero eigenvalue is spanned by the all-ones vector, \( v \propto [1, \ldots, 1]^T \).

Finally, we recall that \( \epsilon = 1/\omega \) and \( \epsilon \lambda_2(\omega, \delta) = \left[ \tilde{\lambda}_2(0, \delta) + \epsilon \tilde{\lambda}_2'(0, \delta) + \mathcal{O}(\epsilon^2) \right] \), which implies

$$
\lim_{\omega \to \infty} \lambda_2(\omega, \delta) = \lim_{\epsilon \to 0^+} \left[ 0 + \tilde{\lambda}_2'(0, \delta) + \mathcal{O}(\epsilon) \right] = \lambda_2(\delta). \tag{SM22}
$$

\[ \Box \]

**APPENDIX D: DERIVATION OF EQUATIONS (6)–(8) IN THE MAIN TEXT**

**Lemma 2.** Let \( \mathbf{L}_1^{(1)} \) and \( \mathbf{L}_1^{(2)} \) be the intralayer Laplacians of a two-layer multiplex network, and define \( \chi \in [0, 1] \) to be a time-scaling parameter that varies the relative timescale of dynamics for the two layers through the mapping: \( \mathbf{L}_1^{(1)} \mapsto \chi \mathbf{L}_1^{(1)} \) and \( \mathbf{L}_2^{(2)} \mapsto \left( 1 - \chi \right) \mathbf{L}_2^{(2)} \). (Note that the dynamics of layer 1 is much faster as \( \chi \to 1 \), whereas layer 2 is much faster as \( \chi \to 0 \).) Further, let \( \mathbf{L}(\delta, \chi) \) be the weighted-average Laplacian given in Eq. (SM9) under this mapping. It then follows that

$$
\mathbf{L}(\delta, \chi) = \frac{1 + \delta}{2} \mathbf{L}_1^{(1)} + \frac{1 - \delta}{2} (1 - \chi) \mathbf{L}_2^{(2)}. \tag{SM23}
$$

**Proof.** The interlayer Laplacian for \( T = 2 \) asymmetrically coupled layers is given by

$$
\mathbf{L}_1^{(\delta)} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \delta \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{SM24}
$$

and its zero-valued eigenvalue has left eigenvector \( \mathbf{u}(\delta) = [1 + \delta, 1 - \delta]^T \) and right eigenvector \( \mathbf{v}(\delta) = [1, 1]^T \). The result follows after using that \( u_1(\delta) = 1 + \delta, u_2(\delta) = 1 - \delta, \) and \( u_1(\delta) + u_2(\delta) = 2 \). \[ \Box \]
Theorem 4. Let \( \lambda_2' (\delta) \equiv \frac{d}{d\delta} \text{Re}(\lambda_2) \) be the derivative of the real part of the eigenvalue \( \lambda_2 (\delta) \) of \( L(\delta, \chi) \) that has the second-smallest real part. Further, let \( \mathbf{u}^{(t)} \) and \( \mathbf{v}^{(t)} \) be the left and right eigenvectors for the second smallest eigenvalue \( \lambda_2^{(t)} \) of \( L^{(t)} \) for \( t \in \{1, 2\} \). We then find the following limits as \( \delta \to \pm 1 \),

\[
\begin{align*}
\lambda_2'(1) &= \frac{-\mathbf{u}^{(1)} \mathbf{L}^{(2)} \mathbf{v}^{(1)} + \chi \mathbf{u}^{(1)*} (\mathbf{L}^{(1)} + \mathbf{L}^{(2)}) \mathbf{v}^{(1)}}{2 \mathbf{u}^{(1)*} \mathbf{v}^{(1)}} , \\
\lambda_2'(-1) &= -\frac{\lambda_2^{(2)}}{2} + \frac{-\mathbf{u}^{(2)*} (\mathbf{L}^{(1)} + \mathbf{L}^{(2)}) \mathbf{v}^{(2)}}{2 \mathbf{u}^{(2)*} \mathbf{v}^{(2)}},
\end{align*}
\]

where \( \mathbf{u}^* \) denotes the conjugate transpose of vector \( \mathbf{u} \).

Proof. We first consider \( \delta \to -1 \) and note the identity

\[
\left[ \frac{1 + \delta}{2} \chi \mathbf{L}^{(1)} + \frac{1 - \delta}{2} (1 - \chi) \mathbf{L}^{(2)} \right] = (1 - \chi) \mathbf{L}^{(2)} + (1 + \delta) \chi \frac{\mathbf{L}^{(1)} - (1 - \chi) \mathbf{L}^{(2)}}{2} ,
\]

We Taylor expand the eigenvalue \( \lambda_2 (\delta) \) and its associated left and right eigenvectors \( \mathbf{u}(\delta) \) and \( \mathbf{v}(\delta) \), respectively, to find

\[
\lambda_2 (\delta, \chi) = \lambda_2 (-1) + (1 + \delta) \lambda_2' (-1) + \mathcal{O}(1 + \delta)^2 ,
\]

\[
\mathbf{u}(\delta) = \mathbf{u}(-1) + (1 + \delta) \mathbf{u}' (-1) + \mathcal{O}(1 + \delta)^2 ,
\]

\[
\mathbf{v}(\delta) = \mathbf{v}(-1) + (1 + \delta) \mathbf{v}' (-1) + \mathcal{O}(1 + \delta)^2 .
\]

We substitute the first-order approximations into the eigenvalue equation \( L(\delta, \chi) \mathbf{v}(\delta) = \lambda_2 (\delta) \mathbf{v}(\delta) \) to obtain

\[
\left[ \frac{1 + \delta}{2} \chi \mathbf{L}^{(1)} + \frac{1 - \delta}{2} (1 - \chi) \mathbf{L}^{(2)} \right] \left[ \mathbf{v}(-1) + (1 + \delta) \mathbf{v}' (-1) \right] = \left[ \lambda_2(-1) + (1 + \delta) \lambda_2' (-1) \right] \left[ \mathbf{v}(-1) + (1 + \delta) \mathbf{v}' (-1) \right] .
\]

The zeroth-order and first-order terms must both be consistent, which gives rise to two equations. The zeroth-order terms yield an eigenvalue equation

\[
(1 - \chi) \mathbf{L}^{(2)} \mathbf{v}(-1) = \lambda_2(-1) \mathbf{v}(-1) ,
\]

which implies that \( \mathbf{v}(-1) = \mathbf{v}^{(2)} \) and \( \lambda_2 (-1) = (1 - \chi) \lambda_2^{(2)} \). The first-order terms yield a linear equation

\[
(1 - \chi) \mathbf{L}^{(2)} \mathbf{v}'(-1) + \chi \mathbf{L}^{(1)} - (1 - \chi) \mathbf{L}^{(2)} \mathbf{v}^{(2)} = (1 - \chi) \lambda_2^{(2)} \mathbf{v}'(-1) + \lambda_2'(-1) \mathbf{v}^{(2)} .
\]

We left multiply both sides of this equation by \( \mathbf{u}^{(2)*} \) to obtain

\[
\chi \frac{\mathbf{u}^{(2)*} [\mathbf{L}^{(1)} + \mathbf{L}^{(2)}] \mathbf{v}^{(2)}}{2 \mathbf{u}^{(2)*} \mathbf{v}^{(2)}} - \frac{\lambda_2^{(2)}}{2} = \lambda_2'(-1) .
\]

This completes the analysis for \( \delta \to -1 \). We repeat this procedure for \( \delta \to 1 \) to complete the proof.

Theorem 5. For a 2-layer multiplex network, \( \text{Re}(\lambda_2(\delta)) \) is guaranteed to have a maximum at some optimum value of \( \delta \) if \( \chi \in (\hat{\chi}(-1), \hat{\chi}(1)) \), where

\[
\begin{align*}
\hat{\chi}(1) &= \text{Re} \left( \frac{\mathbf{u}^{(1)*} \mathbf{L}^{(2)} \mathbf{v}^{(1)}}{\mathbf{u}^{(1)*} (\mathbf{L}^{(1)} + \mathbf{L}^{(2)}) \mathbf{v}^{(1)}} \right) , \\
\hat{\chi}(-1) &= \text{Re} \left( \frac{\mathbf{u}^{(2)*} \mathbf{L}^{(2)} \mathbf{v}^{(2)}}{\mathbf{u}^{(2)*} (\mathbf{L}^{(1)} + \mathbf{L}^{(2)}) \mathbf{v}^{(2)}} \right) .
\end{align*}
\]

Proof. Our proof relies on Rolle’s theorem [7] for a continuous function \( f(\delta) \) on some domain \( \delta \in [a, b] \): if \( f'(a) > 0 \) and \( f'(b) < 0 \), then there exists at least one value of \( \delta \) at which the function \( f(\delta) \) obtains its maximum. In our case, \( f(\delta) = \text{Re}(\lambda_2(\delta)) \)
and \([a, b] = [-1, 1]\). Thus, the maximum is guaranteed to exists provided that \(\lambda_2'(−1) > 0\) and \(\lambda_2'(1) < 0\). We apply these bounds the right-hand-sides of Eqs. (SM25) and solve for \(\chi\) to obtain the desired results.

\[\]  

\footnote{ziaosong@buffalo.edu} \footnote{danet@buffalo.edu}  

[1] A. Tejedor, A. Longjas, E. Foufoula-Georgiou, T. T. Georgiou, and Y. Moreno, Physical Review X \textbf{8}, 031071 (2018).  
[2] G. Cencetti and F. Battiston, New Journal of Physics \textbf{21}, 035006 (2019).  
[3] X. Wang, A. Tejedor, Y. Moreno, \textit{et al.}, New Journal of Physics (2021).  
[4] D. Taylor, M. A. Porter, and P. J. Mucha, Multiscale Modeling \& Simulation \textbf{15}, 537 (2021).  
[5] D. Taylor, Physical Review Research \textbf{2}, 033164 (2020).  
[6] D. Taylor, S. A. Myers, A. Clauset, M. A. Porter, and P. J. Mucha, Multiscale Modeling \& Simulation \textbf{15}, 537 (2017).  
[7] P. Sahoo and T. Riedel, \textit{Mean Value Theorems and Functional Equations} (World Scientific, 1998).