The Competition for Shortest Paths on Sparse Graphs

Chi Ho Yeung and David Saad
The Nonlinearity and Complexity Research Group,
Aston University, Birmingham B4 7ET, United Kingdom
(Dated: February 2, 2012)

Optimal paths connecting randomly selected network nodes and fixed routers are studied analytically in the presence of non-linear overlap cost that penalizes congestion. Routing becomes increasingly more difficult as the number of selected nodes increases and exhibits ergodicity breaking in the case of multiple routers. A distributed linearly-scalable routing algorithm is devised. The ground state of such systems reveals non-monotonic complex behaviors in both average path-length and algorithmic convergence, depending on the network topology, and densities of communicating nodes and routers.

Routing and path selection are at the heart of many communication and logistics applications. For instance, instant messengers, Internet telephony and payment security verification require packets to be delivered instantly or otherwise lose their functionality [1, 2]; while the efficiency of transportation networks depends crucially on effective path selection [3, 4]. Existing routing algorithms are mostly based on minimizing path lengths. Some use routing tables that register the shortest distance to various destinations but are insensitive to traffic congestion [5, 6]; others control congestion by monitoring queue length or latency heuristically [7], or merely optimize routing selfishly [8]. Devising efficient distributive principled routing algorithms which minimze route length while restricting congestion remains a challenge.

Path optimality and congestion control have been extensively studied within the physics community in other contexts, such as the research of spanning [9, 10] and Stenier trees [11] with *quenched* link weights, to mimic broadcast or multi-cast systems. However, these studies ignore interaction terms (overlap costs) that *depend on the specific choice of paths.* Other approaches such as preferential random walk and diffusion methods are used to reduce traffic congestion, but result in heuristic protocols which route packets through sub-optimal paths in a probabilistic manner [12–14].

Mapped onto a statistical physics framework, routing poses both theoretical and numerical challenges due to the multiplicity of possible routes between communicating nodes and the nonlinear costs induced by the interaction between overlapping routes, akin to a non-local repulsion force. We remark that although overlap costs have been partially addressed by assigning quenched link weights [9], they do not fully reflect the complex interaction between dynamical variables. Techniques used to analyze polymers [15], for instance of self-avoiding walks [16] and the traveling salesman problem [17], are prime candidates for the analysis of routing problems but do not consider the cost of interaction between paths.

In this Letter we study a scenario whereby numerous senders seek the shortest possible route to a few receivers while minimizing traffic congestion. The problem is relevant to node pairs on a network that communicate via designated routers; or nodes, possibly sensors, that communicate via an outlet router or base stations. It is also relevant to transportation networks where traffic gravitates towards one or several centers, such as city center or hub airports [3, 18]. Using the cavity approach [19], we examine analytically the dependence of the typical path length on the system topology, the density of communicating node-pairs and routers and their location. We identify the conditions for ergodicity breaking in solution space and observe oscillations in typical path lengths and algorithmic convergence in regular graphs. We show that allocating routers on hubs, which seems to be the natural choice, is indeed optimal in many respects. An applicable linearly-scalable local message passing algorithm is derived which optimizes, in a principled manner, individual routes subject to the mitigation of global congestion.

Model: We consider a sparse network of N nodes $i = 1, \ldots, N$ where each node i is randomly connected to k_i neighbors denoted by $\mathcal{N}_i$; where $k_i \ll N$ is randomly drawn from a distribution $\rho(k)$. We randomly select one node as a *receiver* (router) denoted by r [20], and a fraction f_s of the other nodes, with $0 \leq f_s \leq 1$, as *senders* that communicate with r through a single path each. We assign $\sigma_{ij}^{s \to r} = 1$ if the communication originated from s to r passes through the edge from i to j, and $\sigma_{ij}^{s \to r} = 0$ otherwise. We also denote the *integer* communication flow from i to j as $I_{ij} \equiv \sum_s \left[\sigma_{ij}^{s \to r} - \sigma_{ji}^{s \to r} \right]$. To facilitate the analysis, each node i is assigned with a communication load Λ_i, reflecting its role: the receiver node r is assigned with $\Lambda_r = -\infty$, senders are assigned with $\Lambda = +1$, all other nodes are assigned with $\Lambda = 0$. Flow on the network is then re-directed until excess communication load $R_i = \Lambda_i + \sum_j A_{ji} I_{ji}$ vanishes, i.e. $R_i = 0$, $\forall i \neq r$. To minimize path length subject to a nonlinear cost for overlapping paths, we introduce the Hamiltonian $\mathcal{H} = \sum_{(ij)} |I_{ij}|^\alpha$ with $\alpha > 1$; analysis and algorithm are generic for any $\alpha \geq 1$. An example of a small system attaining the ground state of $\mathcal{H}$ with $\alpha = 2$ is shown in Fig. 1(a).

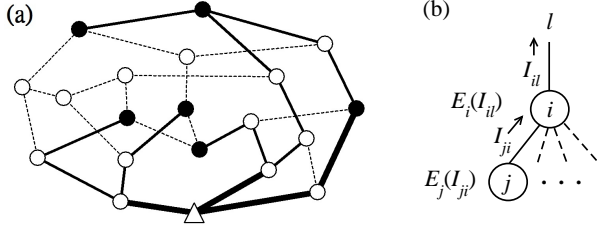This model can be adapted to accommodate various routing scenarios. For instance, senders with a positive

FIG. 1: (a) An example of the ground state of $\mathcal{H}$ with $\alpha=2$ in a simple network with 6 senders ($\bullet$) and a single receiver ($\triangle$) among 20 nodes. Dashed lines correspond to idle links, while thin and thick solid lines correspond to communication loads of 1 and 2, respectively. To minimize $\mathcal{H}$, communication from the top left sender is routed through a long path. (b) a schematic diagram representing the derivation of Eq. (1).
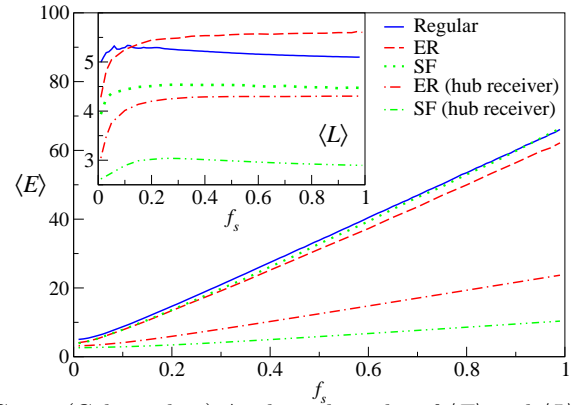

FIG. 2: (Color online) Analytical results of $\langle E \rangle$ and $\langle L \rangle$ per communication as a function of f_s on regular, Erdös-Rényi (ER) and scale-free (SF) networks, and ER and SF networks with hub receivers; all with $\langle k \rangle = 3$ and $f_r = 0.01$.

integer Λ may be introduced to model data segmentation and routing through multiple paths; multiple receivers may be considered as is the case in wireless sensor networks where *any* reachable base station will do [7]. The Hamiltonian can also be modified to accommodate different forms of cost and objectives on nodes and links.

To analyze the system's ground state behavior, we employ the cavity approach [19] and assume that only large loops exist, so that neighbors of node i become statistically independent if it is being removed. This allows one to define $E_i(I_{il})$, the energy of a tree terminated at the link from vertex i to parent l, given the flow I_{il} [21]. For any $i \neq r$, one obtains a recursion equation relating the energy $E_i(I_{il})$ to $E_j(I_{ji})$ of its neighbors (decedents) j other than (parent) l, as sketched in Fig. 1(b), given by

$$E_i(I_{il}) = \min_{\{\{I_{ji}\}|R_i=0\}}\left[|I_{il}|^\alpha + \sum_{j\in\mathcal{N}_i\setminus l} E_j(I_{ji})\right]. \quad (1)$$

For receiver nodes, being universal sinks, $E_r = |I_{jr}|^\alpha, \forall j \in \mathcal{N}_r$. For brevity we denote the right hand side of Eq. (1) by a functional $\mathcal{M}[\underline{E}_{il}; \Lambda_i, I_{il}]$ where $\underline{E}_{il} \equiv \{E_j | j \in \mathcal{N}_i \setminus l\}$. The function $E(I)$ is extensive and difficult to iterate, and is replaced by the intensive quantity $E^V(I) \equiv E(I) - E(0)$ [22] to derive the recursive relation

$$E_i^V(I_{il}) = \mathcal{M}[\underline{E}_{il}^V; \Lambda_i, I_{il}] - \mathcal{M}[\underline{E}_{il}^V; \Lambda_i, 0]. \quad (2)$$

To evaluate various physical quantities one employs population dynamics to iterate Eq. (1) and obtain a stable distribution $P[E^V(I)]$. The *ground state energy* is evaluated [23] by the average energy of an additional node and link; denoting $\underline{E}_i^V \equiv \{E_j^V | j \in \mathcal{N}_i\}$, this is given by

$$\langle E_{\text{node}} \rangle = \langle \mathcal{M}[\underline{E}_i^V; \Lambda, 0] \rangle_{\Lambda, k, \underline{E}_i^V}, \quad (3)$$

$$\langle E_{\text{link}} \rangle = \langle \min_I [E_{j_1}^V(I) + E_{j_2}^V(-I) - |I|^\alpha] \rangle_{E_{j_1}^V, E_{j_2}^V} \quad (4)$$

averaged over the sender and degree distributions $p(\Lambda)$, $\rho(k)$ and $P[E^V(I)]$. The energy per transmission is thus given by $\langle E \rangle = (\langle E_{\text{node}} \rangle - \frac{\langle k \rangle}{2}\langle E_{\text{link}} \rangle)/f_s$ [23].

Obtaining typical path length for non-integer α follows directly from Eq. (4) in the form of $\langle L \rangle = \langle|I^*|\rangle/f_s$, where $I^* = \text{argmin}_I[E_{j_1}^V(I) + E_{j_2}^V(-I) - |I|^\alpha]$, since $\langle|I^*|\rangle/f_s$ is effectively the link occupancy per communication. The case of integer α suffers from degeneracy of I^* and is best calculated by propagating, parallel to E^V, the function

$$L_i^V(I_{il}) = |I_{ji}| + \sum_{j\in\mathcal{N}_i\setminus l} L_j^V[I_{ji}^*(I_{il})], \quad (5)$$

where the optimal flow $I_{ji}^*(I_{il})$, given I_{il}, is obtained from Eq. (2). While E^V provides the optimal energy, L^V evaluates typical traffic in the optimal state. Average path length per communication is obtained using the distribution $P(E^V, L^V)$ by $\langle L \rangle = (\langle L_{\text{node}} \rangle - \frac{\langle k \rangle}{2}\langle L_{\text{link}} \rangle)/f_s$, where $\langle L_{\text{node}} \rangle$ and $\langle L_{\text{link}} \rangle$ are evaluated similarly to Eq. (5) via the corresponding I^* flow obtained in Eqs. (3-4).

Solutions: Results shown in Fig. 2, for $\alpha = 2$, $f_r = 0.01$ [20] and different connectivity distributions representing regular, Erdös-Rényi (ER) and scale-free (SF) networks, show a linear increase in $\langle E \rangle$ for most of the range of f_s regardless of network topology. This indicates that the potential quadratic increase in cost due to overlaps is mitigated through rerouting, balancing path-length and overlap costs. Typical path length $\langle L \rangle$ (inset) increases initially with f_s for all networks (SF being shorter) due to overlap costs, but saturates for ER networks while showing a *decrease* for regular and SF graphs at high f_s. This indicates longer alternative routes are also congested, making shorter congested ones more cost-effective. This suggests that simple shortest path routing schemes may achieve close-to-optimal performance for either low or high traffic densities in regular and SF networks, but are less effective for intermediate density and ER networks.

To examine the effect of receiver connectivity, we study the case where receivers are deployed to the largest hubs in ER and SF networks. Figure 2 shows a more moderate increase in energy and path length compared to random deployment. In addition, communication costs
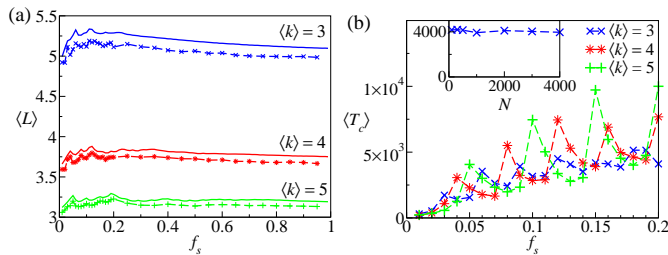
FIG. 3: (Color online) (a) Analytic $\langle L \rangle$ values (solid lines) compared to simulation results (dashed lines, symbols) for regular networks with $\langle k \rangle = 3, 4, 5$, a single receiver and $N = 100$ nodes, averaged over 1000 realizations. (b) Algorithmic convergence time $\langle T_c \rangle$ averaged over instances ($\approx 98\%$) which converge within 5×10^4 updates per node. Inset: $\langle T_c \rangle$ as a function of N for $\langle k \rangle = 3$, $f_s = 0.2$ and a single receiver.

in SF networks become lower with respect to ER networks, reversing their random case positions. This suggests that receiver local connectivity dominates network performance and potentially explains why most transportation networks are SF with hub receivers.

Careful examination of average path length $\langle L \rangle$ in regular networks as a function of f_s reveals a fine structure of small peaks, in addition to the main peak, at multiples of $\langle k \rangle / N$ as shown in Fig. 3(a). These occur when traffic is balanced predominantly around the receiver, which we term the *balanced receiver phenomenon* (BRP). For example, the long route from the top left sender in Fig. 1(a) is chosen a to ensure that links connecting the receiver are occupied exactly by 2 communications as a high cost is incurred by imbalanced traffic. This behavior is also reflected in the expected algorithmic convergence time (number of update steps per node) $\langle T_c \rangle$ shown in Fig. 3(b). BRP may appear in real-life networks; for instance, one may attempt to balance car traffic heading towards a city center by varying toll or speed limits. This phenomenon is masked in the results for ER and SF networks due to the degree variability of receivers and since BRP is more pronounced at low degree receivers.

Computation: One computational challenge in Eq. (2) is the extremization on the integer domain subject to an equality constraint. We use the convexity of $|I|^\alpha$ ($\alpha \geq 1$) to show that $E^V(I)$ is convex and denote the energy change when I_{il} increases/decreases by 1 as $\Delta_i^\pm(I_{il}) = E_i^V(I_{il} \pm 1) - E_i^V(I_{il})$. The optimality condition $\Delta_{j_1}^+[I_{j_1 i}^*(I_{il})] + \Delta_{j_2}^-[I_{j_2 i}^*(I_{il})] \geq 0$, $\forall j_1, j_2 \in \mathcal{N}_i \backslash l$, and the convexity of $E_i^V(I_{il})$ yield

$$\Delta_i^\pm(I_{il}) = |I_{il} \pm 1|^\alpha - |I_{il}|^\alpha + \min_{j \in \mathcal{N}_i \backslash l}\{\Delta_j^\pm[I_{ji}^*(I_{il})]\} \quad (6)$$

$$I_{ji}^*(I_{il} \pm 1) = \begin{cases} I_{ji}^*(I_{il}) \pm 1, & j = \underset{j \in \mathcal{N}_i \backslash l}{\mathrm{argmin}}\{\Delta_j^\pm[I_{ji}^*(I_{il})]\} \\ I_{ji}^*(I_{il}), & \text{otherwise} \end{cases}$$

This simplifies the computation (see Supplemental Material) and facilitates a study of the model's ergodicity breaking properties. Equation (6) implicitly assumes localized interdependence of flows that corresponds to

a *replica symmetry* (RS) [19] property. Small variations in *cavity field* $\Delta_i^\pm$ in Eq. (6) take the form $\delta\Delta_i^\pm(I_{il}) = \delta\Delta_{j^*}^\pm(I_{j^* i})$, where $j^* = \mathrm{argmin}_{j \in \mathcal{N}_i / l}[\Delta_j^\pm(I_{ji})]$, leading to expected perturbations of the form $\langle(\delta\Delta_i^\pm)^2\rangle = f_s\langle(\delta\Delta_j^\pm)^2\rangle$ for all system nodes. Perturbations decay for $f_s < 1$ indicating an RS [24] behavior, which leads to a unique state. This is supported by simulations as almost all cases converge in a small number of updates.

Algorithm: For optimizing individual instances one iterates Eq. (2) until convergence of the E^V's, and computes the flows using either Eqs. (3) or (4) to facilitate the identification of individual paths. The situation becomes more complicated when solutions of Eqs. (3) or (4) have degeneracy, which is typical for integer α. One can then break the degeneracy by assigning randomly to each link (ij) a small quenched bias $\epsilon_{(ij)}$, and adopts in the iteration of Eq. (2) the modified cost of $|I_{ij}|^\alpha + |I_{ij}|\epsilon_{(ij)}$. The iteration then converges to a particular solution with no degeneracy. The algorithm is linearly scalable as shown in the inset of Fig. 3(b), but the bias increases convergence time with respect to having a non-zero bias.

Breaking of ergodicity: While *replica symmetry breaking* (RSB) [19] which is associated with algorithmic hardness does not emerge in the original model, it does emerge in very similar scenarios. One such case is that of multiple receiver classes, where each sender is assigned a specific receiver type. This scenario corresponds to users/sensors with specific information for specific routers/receivers. In the case of two different receiver types, say A and B with $(\Lambda^A, \Lambda^B) = (-\infty, 0)$ and $(0, -\infty)$, respectively, senders are initialized with either $(\Lambda^A, \Lambda^B) = (+1, 0)$ or $(0, +1)$. The corresponding Hamiltonian $\mathcal{H} = \sum_{(ij)}(|I_{ij}^A| + |I_{ij}^B|)^\alpha$ and $E^V(I_{ij}^A, I_{ij}^B)$'s are two-dimensional; E^V's are no longer convex for the entire range and Eq. (6) does not hold. Nevertheless, one can derive an approximate recursion relation based on the insight gained from Eq. (6), for instance

$$\Delta_i^{A^+}(I_{il}^A, I_{il}^B) \approx (|I_{il}^A + 1| + |I_{il}^B|)^\alpha - (|I_{il}^A| + |I_{il}^B|)^\alpha$$
$$+ \min\left\{ \min_{j \in \mathcal{N}\backslash l}[\Delta_j^{A^+}], \min_{j_1, j_2 \in \mathcal{N}\backslash l}[\Delta_{j_1}^{A^{2+}} + \Delta_{j_2}^{A^-}], \right.$$
$$\left. \min_{j_1, j_2 \in \mathcal{N}\backslash l}[\Delta_{j_1}^{A^+, B^\pm} + \Delta_{j_2}^{B^\mp}] \right\} . \quad (7)$$

Superscripts A^+ and A^{2+} correspond to the increase of type A flow by 1 and 2 units, respectively (similarly for B). Various combinations of Δ_j's on the right hand side add up to give $\Delta_i^{A^+}$; the arguments of Δ_j were omitted for clarity. This approximation provides good results compared to the exact computation and can be improved by adding combinations on the right hand side.

We employ Eq. (7) to identify the RS-RSB transition by introducing perturbation to the cavity fields. The expression becomes more involved and perturbations depend on perturbations induced by a number of descendent (see Supplemental Material). Depending on system parameters, we identify decaying perturbations with RS
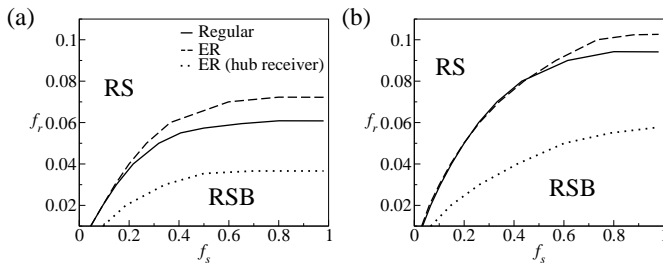
FIG. 4: Phase diagram for fractions f_s and f_r of sender and receiver (router) nodes, respectively, and two types of receivers on regular networks, and ER networks with random and hub receivers, for (a) senders with a fixed router allocation and (b) sender pairs with flexible router choice.

behavior and diverging perturbations with RSB.

Figure 4(a) shows the RS and RSB phase transitions as a function of the fractions f_s and f_r of sender and router nodes with respect to the system size, respectively, for regular networks, and ER networks with random and hub routers (f_r refers equally to each of the receiver types). Generally, RSB emerges for large f_s and small f_r as a large number of communications from distant senders are optimized simultaneously, leading to an extensive frustration of a finite fraction of all communications. While ER networks with random receiver have a larger RSB phase than the regular network, it shrinks significantly when receivers are allocated on the hubs. This shows again the advantage of hub receivers since the RS phase, where simple algorithms suffice, extends to larger systems with denser communications. Since the allocation of specified receivers to senders induces RSB we expect a more extensive RSB phase with the increase in receiver types. This is relevant to *peer-to-peer* (P2P) networks where each node represents a different receiver type.

In another scenario, senders pair up to establish a communication line via a receiver, which can be of either type A or B, by setting their respective communication loads (Λ^A, Λ^B) to either $(1,0)$ or $(0,1)$. One thus introduces a new variable representing the energy change when sender i switches from receiver A to B

$$
\begin{aligned}
\phi_i^{A^+,B^-} =& \mathcal{M}[\underline{E}_i^V; \Lambda^A = 1, \Lambda^B = 0; 0] \\
& - \mathcal{M}[\underline{E}_i^V; \Lambda^A = 0, \Lambda^B = 1; 0] \ , \quad (8)
\end{aligned}
$$

and its distribution $P(\phi^{A^+,B^-})$. This impacts on recursion equation for E^V which translates to adding a term $\delta_{\Lambda^A(I_{il}^A, I_{il}^B), 1} \min_{j \in \mathcal{N} \setminus l}[-\phi^{A^+,B^-} + \Delta_j^{A^{2+},B^-}] + \delta_{\Lambda^B(I_{il}^A, I_{il}^B), 1} \min_{j \in \mathcal{N} \setminus l}[\phi^{A^+,B^-} + \Delta_j^{B^+}]$ to the argument of the outermost minimum function in Eq. (7). The first and second term here correspond to the switch from receiver A to B and from B to A, respectively.

Identifying the RS-RSB transition is carried out in a similar manner to the previous case but one has to account also for the perturbation $\delta \phi^{A^+,B^-}$, resulting in the qualitatively similar phase diagram shown in Fig. 4(b).

Optimal routing in the presence of overlap costs is highly relevant to communication networks and exhibits a complex and interesting behavior as a function of system topology, sender and receiver densities and types. Results, obtained analytically and verified numerically, provide insight into the suitability of various network topologies as communication networks and the preferred location of network servers. While the original framework is replica symmetric, related scenarios whereby senders communicate individually, or jointly as a pair, through *specific* receivers exhibit RSB behavior, with direct impact on the algorithmic hardness of the routing problem. A generic linearly scalable algorithm which optimizes paths to central routers was also devised. The application of statistical mechanics to communication networks is highly promising both due to the insight gained and the promise of better principled routing algorithms.

[1] C. Huitema, *Routing in the Internet* (Prentice Hall, Englewood Cliffs, NJ, 1995).
[2] J. T. Moy, *OSPF: Anatomy of an Internet Routing Protocol* (Addison-Wesley, Reading, MA, 1998).
[3] Z. Wu *et al.*, Phys. Rev. Lett. 96, 148702 (2006).
[4] B. J. Kim *et al.*, Phys. Rev. E 65, 027103 (2002).
[5] R. Bellman, Q. Appl. Math. 16, 87 (1958).
[6] E. W. Dijkstra, Numerische Mathematik 1, 269 (1959).
[7] S. Rangwala *et al.*, SIGCOMM, 63 (2006).
[8] T. Roughgarden and É. Tardos J. ACM 49, 236 (2002).
[9] J. D. Noh and H. Rieger, Phys. Rev E 66, 066127 (2002).
[10] R. Dobrin and P. M. Duxbury, Phys. Rev. Lett. 86, 5076 (2001).
[11] M. Bayati *et al.*, Phys. Rev. Lett. 101, 037208 (2008).
[12] L. A. Adamic *et al.*, Phys. Rev. E 64, 046135 (2001).
[13] W. X. Wang *et al.*, Phys. Rev. E 73, 026111 (2006).
[14] B. Danila *et al.*, Phys. Rev. E 74, 046106 (2006).
[15] M. Daoud *et al.*, Macromolecules 8, 804 (1975).
[16] J. F. Stilck *et al.*, Phys. Rev. Lett. 76, 2734 (1996).
[17] M. Mézard and G. Parisi, J. Physique 47, 1285 (1986).
[18] D. R. Wuellner *et al.*, Phys. Rev. E 82, 056101 (2010).
[19] M. Mézard, G. Parisi, and M. A. Virasoro, Spin Glass Theory and Beyond (World Scientific, Singapore, 1987).
[20] Alternatively, in the large system limit, one may select a small fraction of the system nodes f_r to act as receivers, rather than selecting a single receiver.
[21] K. Y. M Wong and D. Saad, Phys. Rev E 74. 010104(R), (2006).
[22] C. H. Yeung and K. Y. M. Wong, J. Stat. Mech. P04017 (2010).
[23] M. Mézard, R. Zecchina, Phys. Rev. E 66, 056126 (2002).
[24] O. Rivoire *et al.*, Eur. Phys. J. B 37, 55-78 (2004).

Supplemental Material

The Competition for Shortest Paths on Sparse Graphs
Chi Ho Yeung and David Saad

I. PROOF OF THE CONVEXITY OF $E^V(I)$

Given $\alpha \geq 1$, we will show that $E(I)$ is convex by a self-consistent argument. To show that $E_i^V(I_{il})$ is convex, it is sufficient to show that $\Delta^+(I_{il}) \leq \Delta^+(I_{il} + 1)$ for all I_{il}, given the $E_j^V(I_{ji})$ functions of descendants, are convex. From Eq. (6), we have

$$\Delta_i^+(I_{il} + 1) = |I_{il} + 2|^\alpha - |I_{il} + 1|^\alpha + \min_j \{\Delta_j^+[I_{ji}^*(I_{il} + 1)]\}$$

$$\geq |I_{il} + 2|^\alpha - |I_{il} + 1|^\alpha + \min_j \{\Delta_j^+[I_{ji}^*(I_{il})], \Delta_j^+[I_{ji}^*(I_{il}) + 1]\}$$

$$(\because I_{ji}^*(I_{il} + 1) = I_{ji}^*(I_{il}) \text{ or } I_{ji}^*(I_{il}) + 1, \text{from Eq. (6)})$$

$$= |I_{il} + 2|^\alpha - |I_{il} + 1|^\alpha + \min_j \{\Delta_j^+[I_{ji}^*(I_{il})]\}$$

$$(\because \text{convexity of } E_j^V(I_{ji}))$$

$$\geq |I_{il} + 1|^\alpha - |I_{il}|^\alpha + \min_j \{\Delta_j^+[I_{ji}^*(I_{il})]\}$$

$$(\text{equality holds only when } \alpha = 1)$$

$$= \Delta_i^+(I_{il}).$$

Therefore, given the convexity of descendent $E_j^V(I_{ji})$'s, we can show that $E_i^V(I_{il})$ is convex. Hence, $E^V(I)$ is convex by a self-consistent argument.

II. SIMPLIFYING THE COMPUTATION OF $E^V(I)$

To compute $E^V(I)$, we first note that one has to evaluate $E^V(I)$ on a finite range of I since $E^V(I)$ cannot be expressed by simple parametrization. Fortunately, $-Nf_s \leq I \leq Nf_s$ as Nf_s is the maximum possible occupancy on any link. We can then simplify the iteration of Eq. (2) by using Eq. (6) with the following procedures: (i) evaluate $\tilde{I}_{il} = \Lambda_i + \sum_{j \in \mathcal{N} \backslash l} \hat{I}_{ji}$, where $\hat{I}_{ji} = \operatorname{argmin}_{I_{ji}} E_j^V(I_{ji})$, (ii) evaluate $E_j^V(\tilde{I}_{il}) = |\tilde{I}_{il}|^\alpha + \sum_{j \in \mathcal{N} \backslash l} E_j^V(\hat{I}_{ji})$, (iii) calculate $E_i^V(I_{il})$ in the relevant range of I_{il} by $\Delta_i^\pm$ increase or decrease of I_{il} from $\tilde{I}_{il}$. Another step which makes use of backward messages [1] can greatly simplify the iteration of Eq. (2) by limiting the relevant I to a small range; this result will be presented elsewhere [2].

For solving $P[E^V]$ numerically from Eq. (2), instead of having only one receiver as in

the simulations of finite system, one has to assume a small fraction f_r of receiver, to gauge between the infinite trees of the cavity method and the finite number of nodes/receivers studied in simulations. Such an assumption would make the cavity result slightly deviate from the finite system considered in simulations with only one receiver, since in addition to the choice of path, senders may as well choose to use different receivers. We found that this is the reason for the small offset between theory and simulations in Fig. 3(a) as the offset vanishes when we simulate larger systems with more than a single receiver.

III. IDENTIFICATION OF THE RSB PHASE IN CASES WITH MULTIPLE RECEIVER TYPE

Similar to the case with only a single type of receiver, Eq. (7) allows us to characterize the RS-RSB transition in the case with multiple receiver types. Unlike the previous case, where $\delta\Delta_i$ depends on a single perturbation $\delta\Delta_j$, the present perturbation $\delta\Delta_i$ may depend on a number of descendants $\delta\Delta_j$'s as given by Eq. (7). We denote this number of descendants as $n_i^{x,y}(I_{il}^A, I_{il}^B)$, which depends on the flow I_{il}^A, I_{il}^B and the corresponding changes x and y, for instance, $(+1, -1)$. To obtain $\langle n \rangle$, one may average n_i uniformly over a range of I_{il}^A, I_{il}^B and with any x and y, but this may not be physical. Instead, we average n_i over a distribution $P(I_{il}^{A*}, I_{il}^{B*}, x, y)$, where I_{il}^{A*} and I_{il}^{B*} are the ground state flows when E_i^V is merged according to the topology and $(I_{il}^{A*} + x, I_{il}^{B*} + y)$ represent another state with zero or minimal energy increment from the ground state. In this case, $\langle n \rangle$ characterizes the stability of the ground state with respect to a degenerate or excited state, and $\langle n \rangle < 1$ and $\langle n \rangle > 1$ characterize the RS and RSB phases, respectively.

[1] K. Y. M Wong and D. Saad, Phys. Rev E 74. 010104 (R), (2006).
[2] C. H. Yeung and D. Saad, in preparation.