Distributed algorithms for global optimization on sparse networks of arbitrary bandwidths

K. Y. Michael Wong[1], C. H. Yeung[1], and David Saad[2]

[1]Department of Physics, Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong, China
[2]NCRG, Aston University, Birmingham B4 7ET, UK
phkywong@ust.hk, phbill@ust.hk, d.saad@aston.ac.uk

Abstract. The optimization of resource allocation in sparse networks with *real* variables is studied using methods of statistical physics. Efficient distributed algorithms are devised on the basis of insight gained from the analysis and are examined using numerical simulations, showing excellent performance and full agreement with the theoretical results.

1 Introduction

The optimization of resource allocation is a well known problem in the area of distributed computing [1, 2] to which significant effort has been dedicated within the computer science community. It is representative of a large class of problems in many other areas where a large number of nodes are required to balance their resources and redistribute tasks, such as reducing internet traffic congestion and streamlining network flows of commodities [3, 4]. Many attempts were made in the computer science community, to find practical heuristic solutions to the distribution of computational load between computers connected by networks.

The traditional approach to network optimization is to adopt computationally demanding global optimization techniques, such as linear or quadratic programming [5]. On the other hand, message-passing approaches have gained recent success in problems with *discrete* variables connected by network structures, such as error-correcting codes [6] and probabilistic inference [7]. These approaches have the potential to solve global optimization problems via local updates, thereby reducing the computational complexity. For example, the computational complexity of quadratic programming for the load balancing task typically scales as the cube of the system size, whereas capitalizing on the network topology underlying the connectivity of the variables, message-passing scales linearly with the system size. An even more important advantage of message passing techniques, relevant to practical implementation, is their distributive nature. Since they do not require a global optimizer, they are particularly suitable for distributive control in large or evolving networks. While most analyses so far have focused on cases of discrete variables, here we explore networks of continuous variables.

To study the principles and ingredients in realizing global optimization through distributed algorithms, we will examine here, as a vehicle, the task of resource

allocation. We address a generic version of the problem, which is represented by nodes of some computational power that should carry out tasks. Both computational powers and tasks will be chosen at random from some arbitrary distribution. The nodes are located on a randomly chosen sparse network of some given connectivity. The goal is to migrate tasks on the network such that demands will be satisfied while minimizing the migration of (sub-)tasks. This formulation of the problem is reminiscent of many disordered systems in physics [8], and methods of statistical physics can be used to generate insights. Early and partial work in this direction was presented in [9].

In Section 2, we analyze the problem using the Bethe approximation of statistical mechanics. We then present numerical results in Section 3, and derive new distributed algorithms on the basis of the analysis in Sections 4 and 5. The study is extended to the unsatisfiable case in Section 18. We conclude the paper in Section 7.

2 The theoretical framework

We consider a typical resource allocation task on a sparse network of N nodes, labelled $i = 1, .., N$. Each node i is randomly connected to c other nodes[1], and has a capacity λ_i randomly drawn from a distribution $\rho(\lambda_i)$. The objective is to migrate tasks between nodes such that each node will be capable of carrying out its tasks. The *current* $y_{ij} \equiv -y_{ji}$ drawn from node j to i is a continuous variable aimed at satisfying the node capacity constraints for all i,

$$\sum_j \mathcal{A}_{ij} y_{ij} + \lambda_i \geq 0 \ , \tag{1}$$

where $\mathcal{A}_{ij} = 1$ or 0 for connected or unconnected node pairs i and j, respectively. The currents y_{ij} satisfy the link bandwidth constraints $-W \leq y_{ij} \leq W$ for all connected pairs with $\mathcal{A}_{ij} = 1$.

We consider the load balancing task of minimizing the energy function (cost) $E = \sum_{(ij)} \mathcal{A}_{ij} \phi(y_{ij})$, where the summation (ij) runs over all pairs of nodes, subject to the constraints (1); $\phi(y)$ is a general function of the current y. For load balancing tasks, $\phi(y)$ is typically a convex function, which will be assumed in our study.

For sufficiently large W and capacity distributions with non-negative average λ, the optimal solution of the problem exists for sufficiently large networks. We call this the *satisfiable* case, which will be considered in Sections 3 to 5 for unconstrained links ($W = \infty$) and $\langle \lambda \rangle > 0$. The unsatisfiable case will be considered in Section 18.

[1] Although we focus here on graphs of fixed connectivity, one can easily accommodate any connectivity profile within the same framework; the algorithms presented later are completely general.

The analysis of the network is done by introducing the free energy $F = -T \ln \mathcal{Z}_y$ for a temperature $T \equiv \beta^{-1}$, where $\mathcal{Z}_y$ is the partition function

$$\mathcal{Z}_y = \prod_{(ij)} \int_{-W}^{W} dy_{ij} \prod_i \Theta \left(\sum_j \mathcal{A}_{ij} y_{ij} + \lambda_i \right) \exp \left[-\beta \sum_{(ij)} \mathcal{A}_{ij} \phi(y_{ij}) \right] . \qquad (2)$$

The Θ function returns 1 for a non-negative argument and 0 otherwise.

When the connectivity c is low, the probability of finding a loop of finite length on the graph is low, and the Bethe approximation well describes the local environment of a node. In the approximation, a node is connected to c branches in a tree structure, and the correlations among the branches of the tree are neglected. In each branch, nodes are arranged in generations. A node is connected to an ancestor node of the previous generation, and another $c-1$ descendent nodes of the next generation. Thus, the node is the *vertex* of the tree structure formed by its descendents.

Consider a vertex $V(\mathbf{T})$ of a tree $\mathbf{T}$ having a capacity $\lambda_{V(\mathbf{T})}$, and a current y is drawn from the vertex by its ancestor. One can write an expression for the free energy $F(y|\mathbf{T})$ as a function of the free energies $F(y_k|\mathbf{T}_k)$ of its descendants, that branch out from this vertex

$$\begin{aligned} F(y|\mathbf{T}) = \; & -T \ln \Bigg\{ \prod_{k=1}^{c-1} \left(\int_{-W}^{W} dy_k \right) \Theta \left(\sum_{k=1}^{c-1} y_k - y + \lambda_{V(\mathbf{T})} \right) \\ & \times \exp \left[-\beta \sum_{k=1}^{c-1} \left(F(y_k|\mathbf{T}_k) + \phi(y_k) \right) \right] \Bigg\}, \end{aligned} \qquad (3)$$

where $\mathbf{T}_k$ represents the tree terminated at the k^{th} descendent of the vertex. The free energy can be considered as the sum of two parts, $F(y|\mathbf{T}) = N_{\mathbf{T}} F_{\text{av}} + F_V(y|\mathbf{T})$, where $N_{\mathbf{T}}$ is the number of nodes in the tree $\mathbf{T}$, F_{av} is the average free energy per node, and $F_V(y|\mathbf{T})$ is referred to as the *vertex free energy*. Note that when a vertex is added to a tree, there is a change in the free energy due to the added vertex. Since the number of nodes increases by 1, the vertex free energy is obtained by subtracting the free energy change by the average free energy. This allows us to obtain the recursion relation

$$\begin{aligned} F_V(y|\mathbf{T}) = \; & -T \ln \Bigg\{ \prod_{k=1}^{c-1} \left(\int_{-W}^{W} dy_k \right) \Theta \left(\sum_{k=1}^{c-1} y_k - y + \lambda_{V(\mathbf{T})} \right) \\ & \times \exp \left[-\beta \sum_{k=1}^{c-1} \left(F_V(y_k|\mathbf{T}_k) + \phi(y_k) \right) \right] \Bigg\} - F_{\text{av}} . \end{aligned} \qquad (4)$$

The average free energy per node is obtained by considering the average increase in the total free energy when a node is added to the network,

$$F_{\rm av} = -T\left\langle \ln\left\{ \prod_{k=1}^{c}\left(\int_{-W}^{W} dy_k\right)\Theta\left(\sum_{k=1}^{c} y_k + \lambda_V\right) \times \exp\left[-\beta\sum_{k=1}^{c}\left(F_V(y_k|\mathbf{T}_k)+\phi(y_k)\right)\right]\right\}\right\rangle_{\lambda}, \quad (5)$$

where λ_V is the capacity of the vertex V fed by c trees $\mathbf{T}_1,\ldots,\mathbf{T}_c$, and $\langle\bullet\rangle_\lambda$ represents the average over the distribution $\rho(\lambda)$. For optimization, we take the zero temperature limit of Eq. (4), in which the free energy reduces to the minimum energy, yielding

$$F_V(y|\mathbf{T}) = \min_{\{y_k\,|\,\sum_{k=1}^{c-1} y_k - y + \lambda_{V(\mathbf{T})}\geq 0\}}\left[\sum_{k=1}^{c-1}\left(F_V(y_k|\mathbf{T}_k)+\phi(y_k)\right)\right] - F_{\rm av}. \quad (6)$$

These iterative equations can be directly linked to those obtained from a principled Bayesian approximation, where the logarithms of the messages passed between nodes are proportional to the vertex free energies.

The current distribution and the average free energy per link can be derived by integrating the current y' in a link from one vertex to another, fed by the trees $\mathbf{T}_1$ and $\mathbf{T}_2$, respectively; the obtained expressions are $P(y)=\langle\delta(y-y')\rangle_\star$ and $\langle E\rangle=\langle\phi(y')\rangle_\star$ where

$$\langle\bullet\rangle_\star = \left\langle \frac{\int dy' \exp\left[-\beta\left(F_V(y'|\mathbf{T}_1)+F_V(-y'|\mathbf{T}_2)+\phi(y')\right)\right](\bullet)}{\int dy' \exp\left[-\beta\left(F_V(y'|\mathbf{T}_1)+F_V(-y'|\mathbf{T}_2)+\phi(y')\right)\right]}\right\rangle_{\lambda}. \quad (7)$$

Before closing this section, we mention the alternative analysis of the problem using the replica method [10,11], which was successfully applied in the physics of disordered systems. The derivation is rather involved (details will be provided elsewhere), but gives rise to the same recursive equation Eq. (4) as in the Bethe approximation.

3 Numerical solution

The Bethe approximation provides a theoretical tool to analyze the properties of optimized networks. The solution of Eq. (6) is free from finite size effects inherent in Monte Carlo simulations, and can be obtained numerically. Since the vertex free energy of a node depends on its own capacity and the disordered configuration of its descendants, we generate 1000 nodes at each iteration of Eq. (6), with capacities randomly drawn from the distribution $\rho(\lambda)$, each being fed by $c-1$ nodes randomly drawn from the previous iteration.

We have discretized the vertex free energies $F_V(y|\mathbf{T})$ function into a vector, whose $i^{\rm th}$ component takes the value $F_V(y_i|\mathbf{T})$. To speed up the optimization

search at each node, we first find the *vertex saturation current* drawn from a node such that: (a) the capacity of the node is just used up; (b) the current drawn by each of its descendant nodes is just enough to saturate its own capacity constraint. At this saturation point, we can separately optimize the current drawn by each descendant node, providing a convenient starting point for searching the optimal solutions.

To compute the average energy, we randomly draw 2 nodes, compute the optimal current flowing between them, and repeat the process 1000 times to obtain the average. Figure 1(a) shows the results as a function of iteration step t, for a Gaussian capacity distribution $\rho(\lambda)$ with variance 1 and average $\langle\lambda\rangle$. Each iteration corresponds to adding one extra generation to the tree structure, such that the iterative process corresponds to approximating the network by an increasingly extensive tree. We observe that after an initial rise with iteration steps, the average energies converge to steady-state values, at a rate which increases with the average capacity.

To study the convergence rate of the iterations, we fit the average energy at iteration step t using $\langle E(t) - E(\infty)\rangle \sim \exp(-\gamma t)$ in the asymptotic regime. As shown in the inset of Fig. 1(a), the relaxation rate γ increases with the average capacity. It is interesting to note that a cusp exists at the average capacity of about 0.45. Below that value, convergence of the iteration is slow, since the average energy curve starts to develop a plateau before the final convergence. On the other hand, the plateau disappears and the convergence is fast above the cusp. The slowdown of convergence below the cusp is probably due to the appearance of increasingly large clusters of nonzero currents on the network, since clusters of nodes with negative capacities become increasingly extensive, and need to draw currents from increasingly extensive regions of nodes with excess capacities to satisfy the demand.

4 Distributed algorithms: message-passing

The local nature of the recursion relation Eq. (6) points to the possibility that the network optimization can be solved by local iterative approaches. However, in contrast to other message-passing algorithms which pass conditional probability estimates of *discrete variables* to neighboring nodes, the messages in the present context are more complex, since they are *functions* $F_V(y|\mathbf{T})$ of the current y. We simplify the message to 2 parameters, namely, the first and second derivatives of the vertex free energies. For the quadratic load balancing task, it can be shown that a self-consistent solution of the recursion relation, Eq. (6), consists of vertex free energies which are piecewise quadratic with continuous slopes. This makes the 2-parameter message a very precise approximation.

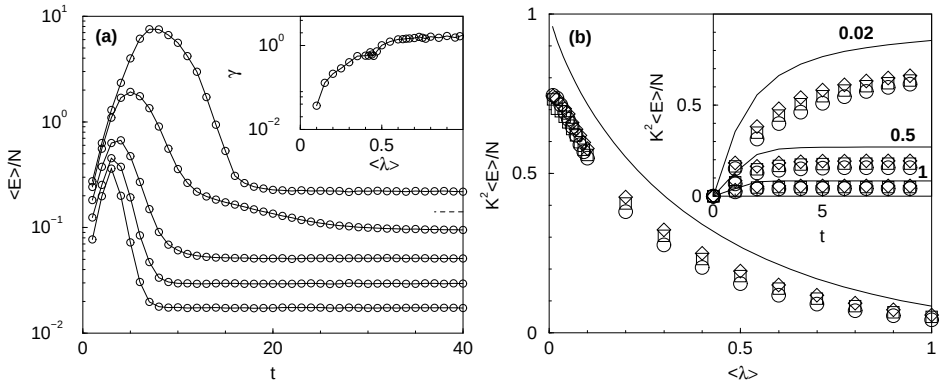

Fig. 1. Results for $N=1000$, $\phi(y)=y^2/2$ and $W=\infty$. (a) $\langle E\rangle/N$ obtained by iterating Eq. (4) as a function of t for $\langle\lambda\rangle=0.1, 0.2, 0.4, 0.6, 0.8$ (top to bottom), $c=3$ and 200-800 samples. Dashed line: the asymptotic $\langle E\rangle/N$ for $\langle\lambda\rangle=0.1$. Inset: γ as a function of $\langle\lambda\rangle$. (b) $K^2\langle E\rangle/N$ as a function of $\langle\lambda\rangle$ for $c=3$ ($\bigcirc$), 4 ($\Box$), 5 ($\Diamond$) and 1000 samples. Line: large K. Inset: $K^2\langle E\rangle/N$ as a function of time for random sequential update of Eqs. (8-9). Symbols: as in (b) for $\langle\lambda\rangle=0.02, 0.1, 0.5$ (top to bottom).

Let $(A_{ij}, B_{ij}) \equiv (\partial F_V(y_{ij}|\mathbf{T}_j)/\partial y_{ij}, \partial^2 F_V(y_{ij}|\mathbf{T}_j)/\partial y_{ij}^2)$ be the message passed from node j to i; using Eq.(6), the recursion relation of the messages become

$$A_{ij} \leftarrow -\mu_{ij}, \quad B_{ij} \leftarrow \Theta(-\mu_{ij}) \left[\sum_{k\neq i} \mathcal{A}_{jk}(\phi''_{jk}+B_{jk})^{-1}\right]^{-1},$$
$$\mu_{ij} = \min\left[\frac{\sum_{k\neq i}\mathcal{A}_{jk}[y_{jk}-(\phi'_{jk}+A_{jk})(\phi''_{jk}+B_{jk})^{-1}]-y_{ij}+\lambda_j}{\sum_{k\neq i}\mathcal{A}_{jk}(\phi''_{jk}+B_{jk})^{-1}}, 0\right], \quad (8)$$

with ϕ'_{jk} and ϕ''_{jk} representing the first and second derivatives of $\phi(y)$ at $y=y_{jk}$ respectively. The forward passing of the message from node j to i is followed by a backward message from node j to k for updating the currents y_{jk} according to

$$y_{jk} \leftarrow y_{jk} - \frac{\phi'_{jk}+A_{jk}+\mu_{ij}}{\phi''_{jk}+B_{jk}}. \quad (9)$$

We note that Eqs. (8-9) differ from conventional message-passing algorithms in that backward messages of the currents are present. As a consequence of representing the messages by the first and second derivatives, the backward messages serve to inform the descendent nodes of the particular arguments they should use in calculating the derivatives for sending the next messages. Furthermore, the criterion that $y_{ij}=-y_{ji}$ provides a check for the convergence of the algorithm.

The message-passing equations further enable us to study the properties of the optimized networks in the limit of large $K\equiv c-1$, and hence consider the convergence to this limit when the connectivity increases. Given an arbitrary

cost function ϕ with nonvanishing second derivatives for all arguments, Eq. (6) converges in the large K limit to the steady-state results

$$A_{ij} = \max\left(\frac{1}{K}\left[\sum_{k\neq i}\mathcal{A}_{jk}A_{jk} - \lambda_j\right], 0\right), \quad B_{ij} \sim \frac{1}{K}. \tag{10}$$

Then, $\sum_{k\neq i}\mathcal{A}_{jk}A_{jk}$ becomes self-averaging and equal to Km_A, where $m_A \sim K^{-1}$ is the mean of the messages A_{ij} given by

$$Km_A = I_1(Km_A); \quad I_n(x) = \langle\Theta(x-\lambda)(x-\lambda)^n\rangle_\lambda. \tag{11}$$

Thus, $y_{ij} \sim \mu_i \sim K^{-1}$. The physical picture of this scaling behavior is that the current drawn by a node is shared among the K descendent nodes. After rescaling, quantities such as $K^2\langle E\rangle/N$, $P(Ky)/K$ and $P(K\mu)/K$ become purely dependent on the capacity distribution $\rho(\lambda)$. For instance, we find

$$\frac{K^2\langle E\rangle}{N} = I_2(Km_A) - I_1(Km_A)^2, \tag{12}$$

$$\frac{P(Ky)}{K} = \int d\lambda_1\rho(\lambda_1)\int d\lambda_2\rho(\lambda_2)\delta\big(\big|(Km_A-\lambda_1)\Theta(Km_A-\lambda_1) - (Km_A-\lambda_2)\Theta(Km_A-\lambda_2)\big| - Ky\big). \tag{13}$$

For increasing finite values of K, Fig. 1(b) shows the common trend of $K^2\langle E\rangle/N$ decreasing with $\langle\lambda\rangle$ exponentially, and gradually approaching the large K limit. The scaling property extends to the optimization dynamics (Fig. 1(b) inset). As shown in Fig. 2(a), the current distribution $P(Ky)/K$ consists of a delta function component at $y=0$ and a continuous component, whose breadth decreases with $\langle\lambda\rangle$. Remarkably, the distributions for different connectivities collapse almost perfectly after the currents are rescaled by K^{-1}, with a very mild dependence on K and gradually approaching the large K limit. As shown in the inset of Fig. 2(a), the fraction of idle links increases with $\langle\lambda\rangle$. Hence the current-carrying links form a percolating cluster at a low $\langle\lambda\rangle$, and breaks into isolated clusters at a high $\langle\lambda\rangle$. The fraction has a weak dependence on the connectivity, confirming the almost universal distributions rescaled for different K.

Since the current on a link scales as K^{-1}, the allocated resource of a node should have a weak dependence on the connectivity. Defining the resource at node i by $r_i \equiv \sum_j \mathcal{A}_{ij}y_{ij} + \lambda_i$, the resource distribution $P(r)$ shown in Fig. 2(b) confirms this behavior even at low connectivities. The fraction of nodes with unsaturated capacity constraints increases with the average capacity, and is weakly dependent on the connectivity (Fig. 2(b) inset). Hence the saturated nodes form a percolating cluster at a low average capacity, and breaks into isolated clusters at a high average capacity. It is interesting to note that at the average capacity of 0.45, below which a plateau starts to develop in the relaxation rate of the recursion relation, Eq. (6), the fraction of unsaturated nodes is about 0.53, close to the theoretical percolation threshold of 0.5 for $c=3$.

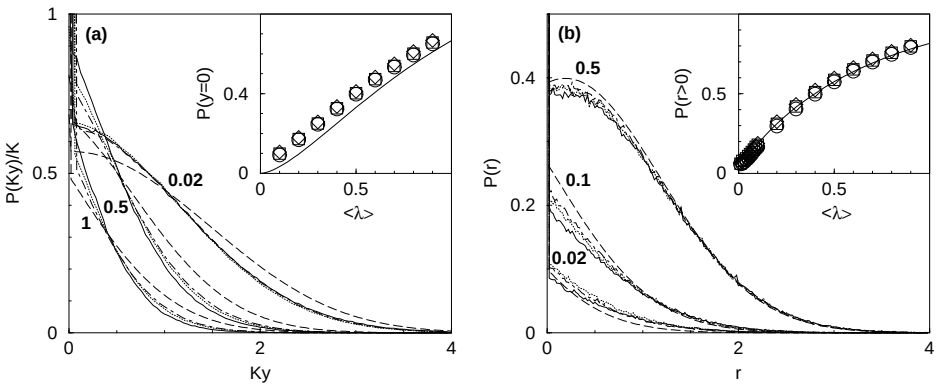

Fig. 2. Results for $N = 1000$, $\phi(y) = y^2/2$, $W = \infty$ and 1000 samples. (a) The current distribution $P(Ky)/K$ for $\langle A \rangle = 0.02, 0.5, 1$, and $c = 3$ (solid lines), 4 (dotted lines), 5 (dot-dashed lines), large K (long dashed lines). Inset: $P(y=0)$ as a function of $\langle \lambda \rangle$ for $c = 3$ ($\bigcirc$), 4 ($\square$), 5 ($\Diamond$), large K (line). (b) The resource distribution $P(r)$ for $\langle \lambda \rangle = 0.02, 0.1, 0.5$, large K. Symbols: as in (a). Inset: $P(r > 0)$ as a function of $\langle \lambda \rangle$. Symbols: as in the inset of (a).

5 Distributed algorithms: price iteration

An alternative distributed algorithm can be obtained by iterating the chemical potentials of the node. Introducing Lagrange multipliers μ_i for the capacity constraints in Eq. (1) we get, for the case of unconstrained links,

$$L = \sum_{(ij} \mathcal{A}_{ij} \phi(y_{ij}) + \sum_i \left(\sum_j \mathcal{A}_{ij} y_{ij} + \lambda_i \right). \tag{14}$$

The extremum condition yields

$$y_{ij} = \phi'^{-1}(\mu_j - \mu_i), \tag{15}$$

and using the Kühn-Tucker condition, μ_i can be solved in terms of μ_j of its neighbours, namely, $\mu_i = \min(g_i^{-1}(0), 0)$, where

$$g_i(x) = \sum_j \mathcal{A}_{ij} \phi'^{-1}(\mu_j - x) + \lambda_i. \tag{16}$$

This provides a local iterative method for the optimization problem. We may interpret this algorithm as a price iteration scheme, by noting that the Lagrangian can be written as $L = \sum_{(ij)} \mathcal{A}_{ij} L_{ij} + \text{constant}$, where

$$L_{ij} = \phi(y_{ij}) + (\mu_i - \mu_j) y_{ij}. \tag{17}$$

Therefore, the problem can be decomposed into independent optimization subproblems, each for a current on a link. μ_i is the storage price at node i, and

Eq. (17) involves balancing the transportation cost on the link, and the storage cost at node i less that at node j, yielding the optimal solution given by Eq. (15). This provides a pricing scheme for the individual links to optimize, which simultaneously optimize the global performance [12]. Simulations show that it yields excellent agreement with the theory Eq. (6) and message-passing Eqs. (8-9).

6 The unsatisfiable case

For links with small bandwidth W, or nodes with negative average capacity, there exist nodes which violate the capacity constraint Eq. (1). In these unsatisfiable cases, it is expedient to relax the constraints and search for optimal solutions which limit the violations. Hence we consider the energy function

$$E = \sum_{ij} \mathcal{A}_{ij}\phi(y_{ij}) + \sum_i \frac{\xi_i^2}{2}, \tag{18}$$

where $\xi_i \equiv \max(-\sum_j \mathcal{A}_{ij}y_{ij} - \lambda_i, 0)$ is the *violation* at node i. The analysis and the distributed algorithms for the satisfiable case can be generalized in the present context as follows.

The recursion relation Eq. (6) in the Bethe approximation is modified to

$$\begin{aligned} F_V(y|\mathbf{T}) = \min_{\{y_k\}} \Bigg[& \frac{1}{2}\left(-\sum_{k=1}^{c-1} y_k + y - \lambda_{V(\mathbf{T})}\right)^2 \Theta\left(-\sum_{k=1}^{c-1} y_k + y - \lambda_{V(\mathbf{T})}\right) \\ & + \sum_{k=1}^{c-1} \left(F_V(y_k|\mathbf{T}_k) + \phi(y_k)\right)\Bigg] - F_{\rm av}. \end{aligned} \tag{19}$$

The message-passing algorithm now becomes

$$\begin{aligned} A_{ij} &\leftarrow -\mu_{ij}, \\ B_{ij} &\leftarrow \left\{1 + \sum_{k\neq i} \mathcal{A}_{jk}(\phi''_{jk} + B_{jk})^{-1}\Theta\left[W - \left|y_{jk} - \frac{\phi'_{jk} + A_{jk} + \mu_{ij}}{\phi''_{jk} + B_{jk}}\right|\right]\right\}^{-1} \end{aligned} \tag{20}$$

where $\mu_{ij} = \min(g_{ij}^{-1}(0), 0)$, with

$$g_{ij}(x) = \sum_{k\neq i} \mathcal{A}_{jk} \max\left\{-W, \min\left[W, \phi'^{-1}(\mu_{jk} - x)\right]\right\} - y_{ij} + \lambda_j - x, \tag{21}$$

The backward message is given by

$$y_{jk} \leftarrow \max\left[-W, \min\left(W, y_{jk} - \frac{\phi'_{jk} + A_{jk} + \mu_{ij}}{\phi''_{jk} + B_{jk}}\right)\right]. \tag{22}$$

The price iteration algorithm now uses $\mu_i = \min(g_i^{-1}(0), 0)$, where

$$g_i(x) = \sum_j \mathcal{A}_{ij} \max\left\{-W, \min\left[W, \phi'^{-1}(\mu_j - x)\right]\right\} + \lambda_i - x, \tag{23}$$

Figure 3(a) shows the simulation results when $\langle\lambda\rangle$ varies. The average energy increases rapidly when $\langle\lambda\rangle$ enters the unsatisfiable regime, and the results obtained by the theory, the message-passing and price iteration algorithms show excellent agreement. There are 3 types of links in the network: idle ($|y_{ij}| = 0$), unsaturated ($|y_{ij}| < W$) and saturated ($|y_{ij}| = W$). When $\langle\lambda\rangle$ enters the unsatisfiable regime, the fraction of idle links vanishes rapidly, while that of saturated links increases to a steady level, implying that more resources are transported in the links in response to the networkwide demand on resources (Fig. 3(a) inset).

In the limit of very negative $\langle\lambda\rangle$, almost all nodes have violations, and the lower limit of 0 for ξ becomes irrelevant, that is, $\xi_i = -\sum_j \mathcal{A}_{ij} y_{ij} - \lambda_i$. This reduces Eq. (18) to a sum of independent optimization problems, one for each current variable. Thus, the optimal solution of y_{ij} depends only on $\lambda_j - \lambda_i$. The average energy and the fraction of saturated links can then be determined from the distribution $\rho(\lambda)$. The theoretical predictions are consistent with the simulation results in Fig. 3(a) and inset.

Figure 3(b) shows the simulation results when W varies. For large values of W, the average energy is effectively constant, since the link bandwidth constraints become irrelevant. On the other hand, when W decreases, the average energy increases rapidly, since the links become increasingly ineffective in allocating resources in the network.

As shown in Fig. 3(b) inset, the fraction of saturated links increases when W decreases. It is interesting to note that the fraction of idle links *increases* when W decreases, contrary to the expectation that more links are involved in resource provision. This can be attributed to what we call a *one-step effect*. If the links in the network were unconstrained, nodes with sufficiently large violations would have drawn currents from distant neighbours, causing currents to flow through many intermediate nodes. However, when W is small, the currents drawn by nodes with violations from their nearest neighbours may have already saturated the links, and there is no use to draw currents from further neighbours. In the limit of vanishing W, the links are exclusively either idle or saturated. In this limit, a link is idle only when both nodes at its ends have positive λ. Hence the fraction of idle links is $f_{\text{idle}} = 1 - f_{\text{sat}} = [P(\lambda > 0)]^2$. Since the transportation cost is negligible in this limit, the contribution to the average energy only comes from the violated nodes, given by $\langle E\rangle/N = \langle\Theta(-\lambda)\lambda^2/2\rangle_\lambda$. These predictions are consistent with the simulation results in Fig. 3(b).

7 Conclusion

We have studied a prototype problem of resource allocation on sparsely connected networks. The resultant recursion relation leads to a message-passing algorithm and a price iteration algorithm for optimizing the average energy, which significantly reduces the computational complexity of the global optimization task and is suitable for online distributive control. The suggested 2-parameter approximation produces results with excellent agreement with the original recursion relation. The Bethe approximation also reveals the scaling properties of

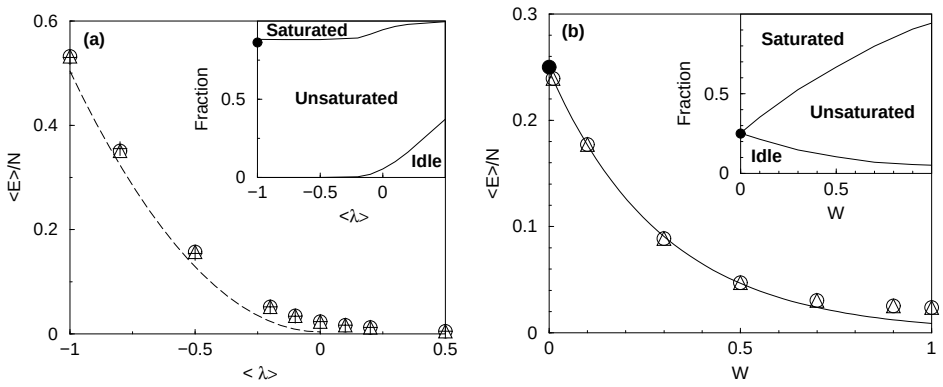

Fig. 3. Results for $N = 1000$, $c = 3$, $\phi = 0.05y^2$ and 100 samples. (a) $\langle E\rangle/N$ as a function of $\langle\lambda\rangle$ for $W = 1$. Symbols: Bethe approximation (+), message-passing ($\triangle$), price iteration ($\bigcirc$). Dashed line: theoretical limit for largely negative $\langle\lambda\rangle$. Inset: the fraction of idle, unsaturated and saturated links as a function of $\langle\lambda\rangle$ for $W = 1$; the vertical height of each region for a given $\langle\lambda\rangle$ corresponds to the respective fraction. Symbol: Theoretical limit of 1 minus the fraction of saturated links for largely negative $\langle\lambda\rangle$ ($\bullet$). (b) $\langle E\rangle/N$ as a function of W for $\langle\lambda\rangle = 0$. Symbols: message-passing ($\triangle$), price iteration ($\bigcirc$), $W \to 0$ theoretical limit ($\bullet$). Line: exponential fit for small values of W. Inset: the fraction of idle, unsaturated and saturated links as a function of W for $\langle\lambda\rangle$. Symbol: $W \to 0$ theoretical limit of the fraction of idle links ($\bullet$).

this model, showing that the resource distribution on the nodes depends principally on the networkwide availability of resources, and depends only weakly on the connectivity. Links share the task of resource provision, leading to current distributions that are almost universally dependent on the resource availability after rescaling.

While the analysis focused on fixed connectivity and zero temperature for optimization, it can accommodate any connectivity profile and temperature parameter and may be used for analyzing a range of inference problems besides optimization. Consider, for instance, an energy function (cost) $E = \sum_{(ij)} \mathcal{A}_{ij}\phi(y_{ij}) + \sum_i \psi(\lambda_i, \{y_{ij}|\mathcal{A}_{ij} = 1\})$, where λ_i is a quenched variable defined on node i. In the context of probabilistic inference, y_{ij} may represent the coupling between observables in nodes j and i, $\phi(y_{ij})$ may correspond to the logarithm of the prior distribution of y_{ij}, and $\psi(\lambda_i, \{y_{ij}|\mathcal{A}_{ij} = 1\})$ the logarithm of the likelihood of the observables λ_i. Both analysis and algorithm extend the use of current message-passing techniques to inference in problems with continuous variables. These advances open up a rich area for further investigations with many potential applications in optimization and inference.

Acknowledgments This work is partially supported by research grants HKUST 6062/02P, DAG04/05.SC25 and DAG05/06.SC36 of the Research Grant Coun-

cil of Hong Kong and by EVERGROW, IP No. 1935 in the FET, EU FP6 and STIPCO EU FP5 contract HPRN-CT-2002-00319.

References

1. Peterson L. and Davie B.S., *Computer Networks: A Systems Approach*, Academic Press, San Diego, CA (2000).
2. Ho Y.C., Servi L. and Suri R., *Large Scale Systems* **1** (1980) 51.
3. Shenker S., Clark D., Estrin D. and Herzog S., *ACM Computer Comm. Review* **26** (1996) 19.
4. Rardin R. L., *Optimization in Operations Research*, Prentice Hall, Upper Saddle River, NJ (1998).
5. Bertsekas D., *Linear Network Optimzation*, MIT Press, Cambridge, MA (1991).
6. Opper M. and Saad D., *Advanced Mean Field Methods*, MIT press (2001).
7. MacKay D.J.C., *Information Theory, Inference and Learning Algorithms*, CUP UK (2003).
8. Nishimori H., *Statistical Physics of Spin Glasses and Information Processing*, OUP UK (2001).
9. Wong K.Y.M. and Saad D., arXiv:cond-mat/0509794 (2005).
10. Mézard M., Parisi P. and Virasoro M., *Spin Glass Theory and Beyond*, World Scientific, Singapore (1987).
11. Wong K.Y.M. and Sherrington D., *J. Phys. A* **20** (1987) L793.
12. Kelly F.P., *Euro. Trans. on Telecommunications & Related Technologies* **8** (1997) 33.