independently and uniformly over N_e, where N_e is a discrete set of $4m$ positive numbers ($m = |E|$). Then, the probability that $\overline{\mathcal{MCF}}$ has a unique optimal solution is at least $\frac{1}{2}$.

Proof. Fix an arc $e_1 \in E$, and fix $\bar{c}_e$ for all $e \in E \setminus e_1$. Suppose that there exists $\alpha \geq 0$, such that when $\bar{c}_{e_1} = \alpha$, $\overline{\mathcal{MCF}}$ have optimal solutions x^*, x^{**}, where $x^*_{e_1} = 0$ and $x^{**}_{e_1} > 0$. Then, if $\bar{c}_{e_1} > \alpha$, for any feasible solution x of $\overline{\mathcal{MCF}}$, where $x_{e_1} > 0$, we have

$$\begin{aligned}\sum_{e\in E} \bar{c}_e x^*_e &= \sum_{e\in E, e\neq e_1} \bar{c}_e x^*_e \\ &\leq \sum_{e\in E, e\neq e_1} \bar{c}_e x_e + x_{e_1}\alpha \\ &< \sum_{e\in E} \bar{c}_e x_e\end{aligned}$$

And if $\bar{c}_{e_1} < \alpha$, for any feasible solution x of $\overline{\mathcal{MCF}}$ where $x_{e_1} = 0$, we have

$$\begin{aligned}\sum_{e\in E} \bar{c}_e x^{**}_e &< \sum_{e\in E, e\neq e_1} \bar{c}_e x^{**}_e + \alpha x^{**}_{e_1} \\ &\leq \sum_{e\in E, e\neq e_1} \bar{c}_e x_e + \alpha x_{e_1} \\ &= \sum_{e\in E} \bar{c}_e x_e\end{aligned}$$

This implies there exists at most one value for α, such that if $\bar{c}_{e_1} = \alpha$ then $\overline{\mathcal{MCF}}$ have optimal solutions x^*, x^{**}, where $x^*_{e_1} = 0$ and $x^{**}_{e_1} > 0$. Similarly, we can also deduce that there exists at most one value for β, such that if $\bar{c}_{e_1} = \beta$, $\overline{\mathcal{MCF}}$ have optimal solutions x^*, x^{**}, where $x^*_{e_1} < u_{e_1}$ and $x^{**}_{e_1} = u_{e_1}$.

Let OS denote the set of all optimal solutions of $\overline{\mathcal{MCF}}$, and let $D(e)$ be the condition of either: $0 = x_e$, $\forall x \in OS$, or $0 < x_e < u_e$, $\forall x \in OS$, or $x_e = u_e$, $\forall x \in OS$. Since $\bar{c}_{e_1}$ has $4m$ possible values, where each value is chosen with equal probability, we conclude that the probability of $D(e_1)$ is satisfied is at least $\frac{4m-2}{4m} = \frac{2m-1}{2m}$. By the union bound of probability, we have that the probability of $D(e)$ is satisfied for all $e \in E$ is at least $1 - \sum_{e\in E} \frac{1}{2m} = \frac{1}{2}$. Now, we state the following lemma:

Lemma 7.1. *If $\forall e \in E$, condition $D(e)$ is satisfied, then $\overline{\mathcal{MCF}}$ have a unique optimal solution.*

Then by Lemma 7.1, we conclude that the probability that $\overline{\mathcal{MCF}}$ has a unique optimal solution is at least $\frac{1}{2}$.

Proof. [Proof of Lemma 7.1] Suppose x^* and x^{**} are two distinct optimal solutions of $\overline{\mathcal{MCF}}$. Let $d = x^{**} - x^*$, then $x^* + \lambda d$ is an optimal solution of $\overline{\mathcal{MCF}}$ iff $0 \leq (x^* + \lambda d)_e \leq u_e$, $\forall e \in E$. As $\bar{c}_e > 0$ for any $e \in E$, and $\bar{c}^T d = \bar{c}^T x^{**} - \bar{c}^T x^* = 0$, there exists some e' such that $d_{e'} < 0$. Let $\lambda^* = \sup\{\lambda \mid x^* + \lambda d$ is an optimal solution of $\overline{\mathcal{MCF}}\}$, since $d_{e'} < 0$, λ^* is bounded; and since $x^* + d = x^{**}$, $\lambda^* > 0$. By optimality of λ^*, there must exists some e'' such that either $(x^* + \lambda^* d)_{e''} = 0$ or $u_{e''}$. Since $\lambda^* > 0$, $x^*_{e''} \neq (x^* + \lambda^* d)_{e''}$, this contradicts the assumption that $D(e'')$ is satisfied. Thus, $\overline{\mathcal{MCF}}$ must have a unique optimal solution.

We note that Theorem 7.1 can be easily modified for LP in standard form.

Corollary 7.1. *Let $\overline{\mathcal{LP}}$ be an LP problem with constraint $Ax = b$, where A is a $m \times n$ matrix, $b \in \mathbb{R}^m$. The cost vector $\bar{c}$ of $\overline{\mathcal{LP}}$ is generated as follows: for each $e \in E$, $\bar{c}_e$ is chosen independently and uniformly over N_e, where N_e is a discrete set of $2n$ elements. Then, the probability that $\overline{\mathcal{LP}}$ has a unique optimal solution is at least $\frac{1}{2}$.*

7.2 Finding the Correct Modified Cost Vector $\bar{c}$

Next, we provide a randomly generate $\bar{c}$ with the desired properties stated in the beginning of this section. Let $X : E \to \{1, 2, ..., 4m\}$ be a random function where for each $e \in E$, $X(e)$ is chosen independently and uniformly over the range. Let $t = \frac{c_{\max}\epsilon}{4mn}$, and generate $\bar{c}$ as: for each $e \in E$, let $\bar{c}_e = 4m \cdot \lfloor \frac{c_e}{t} \rfloor + X(e)$. Then, for any $\bar{c}$ generated at random, $\bar{c}_{\max}$ is polynomial in m, n and $\frac{1}{\epsilon}$, and by Theorem 7.1, the probability of $\overline{\mathcal{MCF}}$ having a unique optimal solution is greater than $\frac{1}{2}$.

Now, we introduce algorithm APRXMT($\mathcal{MCF}, \epsilon$), which works as follows: select a random $\bar{c}$; try to solve $\overline{\mathcal{MCF}}$ using BP. If BP discovers that $\overline{\mathcal{MCF}}$ has no unique optimal solution, then we start the procedure by selecting another $\bar{c}$ at random, otherwise, return the unique optimal solution found by BP. Formally, we present APRXMT($\mathcal{MCF}, \epsilon$) as Algorithm 3.

Corollary 7.2. *The expected number of operations for APRXMT($\mathcal{MCF}, \epsilon$) to terminate is $O(\frac{n^8 m^7 \log n}{\epsilon^3})$.*

Proof. With Theorem 7.1, when we call APRXMT($\mathcal{MCF}, \epsilon$), the expected number of $\overline{\mathcal{MCF}}$ BP tried to solve is bounded by 2. For each selection of $\bar{c}$, we run Algorithm 2 for $2\bar{c}_{\max} n^2$ iterations. As $\bar{c}_{\max} = O(\frac{m^2 n}{\epsilon})$, by Lemma 6.1, the expected number of operations for APRXMT($\mathcal{MCF}, \epsilon$) to terminate is $O(\bar{c}^3_{\max} n^5 m \log n) = O(\frac{n^8 m^7 \log n}{\epsilon^3})$.

Now let $\bar{c}$ be a randomly chosen vector such that $\overline{\mathcal{MCF}}$ has a unique optimal solution x^2. The next thing we want is to show that x^2 is a "near optimal" solution of $\mathcal{MCF}$. To accomplish this, let $e' = \arg\max\{c_e\}$ and define

Algorithm 3 APRXMT($\mathcal{MCF}, \epsilon$)

Let $t = \frac{c_{\max}\epsilon}{4mn}$, for any $e \in E$, assign $\bar{c}_e = 4m \cdot \lfloor \frac{c_e}{t} \rfloor + p_e$, where p_e is an integer chosen independently, uniformly random from $\{1, 2, \ldots, 4m\}$
Let $\overline{\mathcal{MCF}}$ be the problem with modified cost $\bar{c}$.
Run Algorithm 2 on $\overline{\mathcal{MCF}}$ for $N = 2\bar{c}_{\max}n^2$ iterations.
Use Corollary 4.1 to determine if $\overline{\mathcal{MCF}}$ has a unique solution.
if $\overline{\mathcal{MCF}}$ does not have a unique solution **then**
    Restart the procedure APRXMT($\mathcal{MCF}, \epsilon$).
else
    Terminate and return $x^2 = \hat{x}^N$, where $\hat{x}^N$ if the "belief" vector found in Algorithm 2.
end if

$$\min \sum_{e \in E} c_e x_e \qquad (\underline{\mathcal{MCF}})$$

$$\sum_{e \in E_v} \Delta(v, e) x_e = b_v, \quad \forall v \in V \quad \text{(demand constraints)}$$

$$x_{e'} = x^2_{e'}$$

$$0 \le x_e \le u_e, \qquad \forall e \in E \quad \text{(flow constraints)}$$

Suppose x^3 is an optimal solution for ($\underline{\mathcal{MCF}}$), and x^1 is an optimal solution of $\mathcal{MCF}$, then

LEMMA 7.2. *$c^T x^3 - c^T x^1 \le |x^2_{e'} - x^1_{e'}|nt$.*

Proof. Consider $d = x^2 - x^1$, we call a vector $k \in \{-1, 0, 1\}^{|E|}$ an *synchronous cycle vector* of d if for any $e \in E$, $k_e = 1$ only if $d_e > 0$, $k_e = -1$ only if $d_e < 0$, and the set $\{(i, j) | k_{ij} = 1 \text{ or } k_{ji} = -1\}$ forms exactly one directly cycle in G. Since d is a integral vector of circulation (i.e., d send 0 unti amount of flow to every vertex $v \in V$). Hence that we can decompose d so that $\sum_{k \in K} k = d$, for some set K of consisting of synchronous cycle vectors.

Now, let $K' \subset K$ be the subset of K containing all vectors $k \in K$ such that $k_{e'} \ne 0$. Then, for any $k \in K'$, observe that $x^2 - k$ is a feasible solution for $\overline{\mathcal{MCF}}$. Because x^2 is the optimal solution for $\overline{\mathcal{MCF}}$, we have that $\bar{c}'k \le 0$. And

$$\bar{c}_e = 4m\lfloor \frac{c_e}{t} \rfloor + p_e, \; 1 \le p_e \le 4m,$$
$$\implies \bar{c}_e, \frac{4mc_e}{t} \in [4m\lfloor \frac{c_e}{t} \rfloor, 4m(\lfloor \frac{c_e}{t} \rfloor + 1)],$$
$$\implies |\bar{c}_e - \frac{4mc_e}{t}| \le 4m,$$
$$\implies \sum_e |\frac{4mc_e}{t} - \bar{c}_e||k_e| \le \sum_e |4m||k_e| \le 4mn,$$
$$\text{but } \frac{4mc'k}{t} \le \frac{4mc'k}{t} - \bar{c}'k \le \sum_e |\frac{4mc_e}{t} - \bar{c}_e||k_e|,$$
$$\text{thus, we have } \frac{4mc'k}{t} \le 4mn,$$
$$\implies c'k \le nt.$$

Clearly, $x^1 + \sum_{k \in K'} k$ satisfy the demand/supply constraints of $\underline{\mathcal{MCF}}$. Since $\min\{x^1_e, x^2_e\} \le x^1_e + \sum_{k \in K'} k_e \le \max\{x^1_e, x^2_e\}$ for all $e \in E$, and each $k \in K'$, we have that $x^1 + \sum_{k \in K'} k$ is a feasible solution for $\underline{\mathcal{MCF}}$. Since x^3 is the optimal solution of $\underline{\mathcal{MCF}}$, $c^T x^3 \le c^T x^1 + \sum_{k \in K'} c^T k \le c^T x^1 + |K'|nt$. Since $|K'| = |x^2_{e'} - x^1_{e'}|$, we have $c^T x^3 - c^T x^1 \le |x^2_{e'} - x^1_{e'}|nt$.

COROLLARY 7.3. *$c^T x^3 \le (1 + \frac{\epsilon}{2m}) c^T x^1$, for any $\epsilon \le 2$.*

Proof.

$$\text{By Lemma 7.2,} \quad \frac{c^T x^3 - c^T x^1}{c^T x^3} \le \frac{|x^2_{e'} - x^1_{e'}|nt}{c^T x^3} \le \frac{|x^2_{e'} - x^1_{e'}|nt}{|x^2_{e'} - x^1_{e'}|c_{e'}} = \frac{\epsilon}{4m}, \text{ as } t = \frac{c_{e'}\epsilon}{4mn}.$$

Thus,

$$\frac{c^T x^3 - c^T x^1}{c^T x^3} \le \frac{\epsilon}{4m}$$
$$\implies c^T x^3 \le \frac{1}{1 - \frac{\epsilon}{4m}} c^T x^1$$
$$\implies c^T x^3 \le (1 + \frac{\epsilon}{2m}) c^T x^1$$

where the last inequality holds because $(1 - \frac{\epsilon}{4m})(1 + \frac{\epsilon}{2m}) = 1 + \frac{\epsilon}{4m} - \frac{\epsilon^2}{8m} \ge 1$, as $\frac{\epsilon}{2m} \le 1$.

7.3 A $(1 + \epsilon)$ Approximation Scheme Loosely speaking, Corollary 7.3 shows that x^2 at arc e' is "near optimal", since fixing the flow at arc e' to $x^2_{e'}$ helps us in finding a feasible solution of $\mathcal{MCF}$ which is close to optimal. This leads us to approximate algorithm AS($\mathcal{MCF}$, ϵ) (see Algorithm 4); at each iteration, it uses APRXMT (Algorithm 3), and iteratively fixes the flow at the arc with the largest cost.

Algorithm 4 AS($\mathcal{MCF}$, ϵ)

1: Let $G = (V, E)$ be the underlying directed graph of ($\mathcal{MCF}$), $m = |E|$, and $n = |V|$.
2: **while** ($\mathcal{MCF}$) contains at least 1 arcs **do**
3: Run APRXMT ($\mathcal{MCF}, \epsilon$), let x^2 be the solution returned.
4: Find $(i', j') = e' = \arg\max_{e \in E}\{c_e\}$, modify $\mathcal{MCF}$ by fix the flow on arc e' by $x^2_{e'}$, and change the demands/supply on nodes i', j' accordingly.
5: **end while**

THEOREM 7.2. *The expected number of operations for AS($\mathcal{MCF}$, ϵ) to terminate is $O(\frac{n^8 m^8 \log n}{\epsilon^3})$. Suppose x^* be an answer returned by AS($\mathcal{MCF}$, ϵ), then $c^T x^* \leq (1+\epsilon)c^T x^1$, for any $\epsilon \leq 1$.*

Proof. By Corollary 7.2, the expected number of operations for APRXMT($\mathcal{MCF}, \epsilon$) to terminates is $O(\frac{n^8 m^7 \log n}{\epsilon^3})$. Since AS($\mathcal{MCF}$, ϵ) calls the method APRXMT($\mathcal{MCF}, \epsilon$) m times, the expected number of operations for AS($\mathcal{MCF}$, ϵ) to terminate is $O(\frac{n^8 m^8 \log n}{\epsilon^3})$. And by Corollary 7.3,

$$c^T x^* \leq (1 + \frac{\epsilon}{2m})^m c^T x^1 \leq e^{\frac{\epsilon}{2}} c^T x^1 \leq (1+\epsilon) c^T x^1$$

where the last two inequalities follows using standard analysis arguments.

And thus, AS($\mathcal{MCF}$, ϵ) is an $(1+\epsilon)$ approximation scheme for $\mathcal{MCF}$ with polynomial running time.

8 Discussions

In this paper, we formulated belief propagation (BP) for $\mathcal{MCF}$, and proved that BP solves $\mathcal{MCF}$ exactly when the optimal solution is unique. This result generalizes the result from [6], and provides new insights for understanding BP as an optimization solver. Although the running time of BP for $\mathcal{MCF}$ is slower than other existing algorithms for $\mathcal{MCF}$, BP is rather a generic and flexible heuristic, and one should be able to modify it to solve other variants of network flow problems. One such example, the convex-cost flow problem, was analyzed in Section 5. Moreover, the distributed nature of BP can be taken advantage of using parallel computing.

This paper also presents the first fully-polynomial running time scheme for optimization problems using BP. Since the correctness of BP often rely on the optimization problem having a unique optimal solution, Corollary 7.1 can be used for modifying BP in general to work in the absence of a unique optimal solution. In the approximation scheme, the heuristic of fixing values on variables while running BP is commonly known as 'decimation' (see [14]). To the best of out knowledge, this is the first disciplined, provable instance of decimation procedure. The 'decimation' in our scheme is extremely conservative, and a natural question is if there exists a more aggressive 'decimation' heuristic which can improve the running time.

Acknowledgements

While working on this paper, D. Gamarnik was partially supported by NSF Project CMMI-0726733; D. Shah was supported in parts by NSF EMT Project CCF 0829893 and NSF CAREER Project CNS 0546590; and Y. Wei was partially supported NSERC PGS M. The authors would also like to thank the anonymous referees for the helpful comments.

References

[1] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin. *Prentice-Hall Inc.* Addison-Wesley, 1993.

[2] S. M. Aji and R. J. McEliece. The generalized distributive law. *IEEE Transaction on Information Theory*, 46(2):325–343, March 2000.

[3] S. Arora, C. Daskalakis, and D. Steurer. Message passing algorithms and improved lp decoding. 2009.

[4] M. Bayati, C. Borgs, J. Chayes, and R. Zecchina. Belief-propagation for weighted b-matchings on arbitrary graphs and its relation to linear programs with integer solutions. Technical report, 2008. arXiv:0709.1190v2.

[5] M. Bayati, A. Braunstein, and R. Zecchina. A rigorous analysis of the cavity equations for the minimum spanning tree. *Journal of Mathematical Physics*, 49(12):125206, 2008.

[6] M. Bayati, D. Shah, and M. Sharma. Max-product for maximum weight matching: Convergence, correctness, and lp duality. *IEEE Transaction on Information Theory*, 54(3):1241–1251, March 2008.

[7] D. Bertsimas and J. Tsitsiklis. *Introduction to Linear Optimization*, pages 289–290. Athena Scientific, third edition, 1997.

[8] R. Gallager. *Low Density Parity Check Codes.* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 1963.

[9] G. B. Horn. *Iterative Decoding and Pseudocodewords.* PhD thesis, California Institute of Technology, Pasadena, CA, 1999.

[10] D. M. Malioutov, J. K. Johnson, and A. S. Willsky. Walk-sums and belief propagation in gaussian graphical models. *J. Mach. Learn. Res.*, 7:2031–2064, 2006.

[11] M. Mezard, G. Parisi, and R. Zecchina. Analytic and algorithmic solution of random satisfiability problems. *Science*, 297:812, 2002.

[12] C. Moallemi and B. V. Roy. Convergence of min-sum message passing for convex optimization. In *45th*

Allerton Conference on Communication, Control and Computing, 2008.

[13] C. C. Moallemi and B. V. Roy. Convergence of the min-sum message passing algorithm for quadratic optimization. *CoRR*, abs/cs/0603058, 2006.

[14] A. Montanari, F. Ricci-Tersenghi, and G. Semerjian. Solving constraint satisfaction problems through belief propagation-guided decimation.

[15] K. Mulmuley, U. Vazirani, and V. Vazirani. Matching is as easy as matrix inversion. *Combinatorica*, 7(1):105–113, March 1987.

[16] J. Pearl. *Causality: Models, reasoning, and inference*. Cambridge University Press, 2000.

[17] T. Richardson and R. Urbanke. The capacity of low-density parity check codes under message-passing decoding. *IEEE Transaction on Information Theory*, 47(2):599–618, February 2001.

[18] S. Sanghavi, D. Malioutov, and A. Willsky. Linear programming analysis of loopy belief propagation for weighted matching. In *Proc. NIPS Conf*, Vancouver, Canada, 2007.

[19] S. Sanghavi, D. Shah, and A. Willsky. Message-passing for maximum weight independent set. July 2008.

[20] A. Schrijver. *Combinatorial Optimization*. Springer, 2003.

[21] Y. Weiss and W. Freeman. On the optimality of solutions of the max-product belief-propagation algorithm in arbitrary graphs. *IEEE Transactions on Information Theory*, 47(2), February 2001.

[22] J. Yedidia, W. Freeman, and Y. Weiss. Understanding belief propagation and its generalizations. Technical Report TR-2001-22, Mitsubishi Electric Research Lab, 2002.