

元素为 $M_{a \rightarrow i, a \rightarrow i} = \sqrt{2}\delta, M_{a \rightarrow i, b \rightarrow j} = 1_{V(j) \setminus a}(b)1_{V(a) \setminus i}(j)\theta(-J_j^a J_j^b), 1_X(x)$ 是标识函数. 若 $x \in X$, 则 $1_X(x)=1$; 若 $x \notin X$, 则 $1_X(x)=0, \delta$ 为非负常数.

在 WP 算法中, 信息 $u_{a \rightarrow i}$ 取值为 0 或 1. 为了证明定理 2, 我们将 WP 算法的消息取值从 $\{0, 1\}$ 松弛为 $[0, 1]$. 具体地讲: 对于因子图的每条边 (a, i) , 定义消息传递 $\eta_{a \rightarrow i} \in [0, 1]$, 将警示传播算法的消息更新迭代方程(1)更改为

$$\eta_{a \rightarrow i}(t) = (\delta(\eta_{a \rightarrow i}(t-1))^2 - \delta(\eta_{a \rightarrow i}(t-1)) + 1) \prod_{j \in V(a) \setminus i} \theta \left(-J_j^a \left(\sum_{b \in V(j) \setminus a} J_j^b \eta_{b \rightarrow j}(t-1) \right) \right) \quad (5)$$

其中, δ 为非负常数.

相对于公式(1), 在公式(5)中, 我们构造了项 $H(\eta_{a \rightarrow i}(t-1)) = \delta(\eta_{a \rightarrow i}(t-1))^2 - \delta(\eta_{a \rightarrow i}(t-1)) + 1$. 注意到: 当 $\eta_{a \rightarrow i} \in \{0, 1\}$ 时, $H(\eta_{a \rightarrow i}(t-1)) = 1$. 进一步, 如果公式(5)中消息 $\eta_{a \rightarrow i}(t=0)$ 的初始取值仍然为 0 或 1, 则公式(5)退化为公式(1). 我们的想法是: 通过分析公式(5)的收敛性来解释公式(1)的收敛性, 进而给出 WP 算法收敛的有效条件.

设实例的因子图中边的序列集合为 $D = \{a \rightarrow i: (a, i) \in E\}$, 定义向量空间 $V = [0, 1]^{|D|}, (V, \|\cdot\|)$ 是一个赋范空间. 设向量 $x \in V, l_1$ -范数定义为

$$\|x\|_1 = \sum_{i=1}^{|D|} |x_i| \quad (6)$$

l_∞ -范数定义为

$$\|x\|_\infty = \max_{i \in \{1, \dots, |D|\}} |x_i| \quad (7)$$

向量空间 V 上的范数可诱导一个 V 上的度量 $d(v, w) = \|v - w\|, v, w \in V$. 显然, 度量空间 (V, d) 是完备的^[40]. 设函数 $f: V \rightarrow V$, 如果存在 $0 \leq K < 1$, 对于任意的 $v, w \in V$, 使得 $d(f(v) - f(w)) \leq Kd(v, w)$, 则称函数 f 是压缩函数. 如果 $f: V \rightarrow V$ 是完备度量空间 (V, d) 中的压缩函数, 那么对于任意的 $x \in V, f$ 收敛到唯一固定点 x_∞ . 也即, $x, f(x), f^2(x), \dots, x_\infty$.

引理 1. 设函数 $f: V \rightarrow V, d$ 是 V 上的度量. 对于某个 $N \in \mathbb{N}$, 假设 f^N 是关于 d 的压缩函数, 那么 f 收敛到唯一固定点 x_∞ , 且与初始信息无关.

证明: 对于任意的 $x \in V$, 考虑如下 $N-1$ 个序列:

$$\begin{aligned} &x, f^N(x), f^{2N}(x), \dots, \\ &f(x), f^{N+1}(x), f^{2N+1}(x), \dots, \\ &\vdots \\ &f^{N-1}(x), f^{2N-1}(x), f^{3N-1}(x), \dots \end{aligned}$$

因为 f^N 是关于 d 的压缩函数, 所以上述每一个序列都收敛到 x_∞ . 因此, 序列 $x, f(x), f^2(x), \dots$ 一定收敛到 x_∞ . 所以函数 f 收敛到唯一固定点 x_∞ , 且与初始信息无关. □

设映射 $A: V \rightarrow V, (V, \|\cdot\|)$ 是一个赋范空间, 诱导 A 的矩阵范数:

$$\|A\| = \sup_{v \in V} \|Av\| \quad (8)$$

$\|A\|_1$ -范数定义为

$$\|A\|_1 = \max_{j \in \{1, \dots, |D|\}} \sum_{i=1}^{|D|} |A_{ij}| \quad (9)$$

$\|A\|_\infty$ -范数定义为

$$\|A\|_\infty = \max_{i \in \{1, \dots, |D|\}} \sum_{j=1}^{|D|} |A_{ij}| \quad (10)$$

证毕. □

若函数 $f: V \rightarrow V$ 可微, 构造雅可比矩阵 A . 依据中值定理有如下引理.

引理 2^[40]. 设 $(V, \|\cdot\|)$ 是赋范空间, 映射 $f: V \rightarrow V$ 可微, 那么对于任意的 $x, y \in V$, 有公式(11)成立:

$$\|f(x) - f(y)\| \leq \|x - y\| \cdot \sup_{z \in V} \|f'(z)\| \quad (11)$$

根据引理 2,如果 $\sup_{z \in V} \|f'(z)\| < 1$,那么函数 f 收敛到固定点 x_∞ .假设算法中的消息更新并行执行,对于任意的向量 $\eta \in V$,函数 f 的分量 $f(\eta)_{a \rightarrow i}$ 定义为 $f(\eta)_{a \rightarrow i} = \eta_{a \rightarrow i}$.自然,消息更新方程(5)的并行执行过程可以用函数 f 的迭代过程来刻画.下面从引理 2 出发导出 WP 算法收敛的一个充分条件.

为了求解需要,我们对公式(5)进行参数化.通过观察,取 $\sinh \lambda_{a \rightarrow i} = \eta_{a \rightarrow i}$,则公式(5)变为如下形式:

$$\sinh \lambda_{a \rightarrow i}(t) = (\delta \sinh^2 \lambda_{a \rightarrow i}(t-1) - \delta(\sinh \lambda_{a \rightarrow i}(t-1) + 1) \prod_{j \in V(a) \setminus i} \theta \left(-J_j^a \sum_{b \in V(j) \setminus a} J_j^b \sinh \lambda_{b \rightarrow j}(t-1) \right)) \quad (12)$$

由于 $\eta_{a \rightarrow i} \in [0, 1]$,所以 $\lambda_{a \rightarrow i} \in [0, \log(1 + \sqrt{2})]$.定义向量空间 $V' = [0, \log(1 + \sqrt{2})]^{|D|}$,设函数 $h: V' \rightarrow V'$,对于任意的向量 $\lambda \in V'$,函数 h 的分量 $h(\lambda)_{a \rightarrow i}$ 定义为 $h(\lambda)_{a \rightarrow i} = \lambda_{a \rightarrow i}$.设 $i, j \in V(a), i \neq j, b \in V(j) \setminus a$.令:

$$G(\lambda_{a \rightarrow i}(t-1)) = \delta \sinh^2 \lambda_{a \rightarrow i}(t-1) - \delta(\sinh \lambda_{a \rightarrow i} F(\lambda_{c \rightarrow k}(t-1))) = \prod_{k \in V(a) \setminus i, j} \theta \left(-J_k^a \sum_{c \in V(k) \setminus a} J_k^c \sinh \lambda_{c \rightarrow k}(t-1) \right) (t-1) + 1.$$

对函数 h 求导,根据公式(12),我们有:

$$(h'(\lambda))_{b \rightarrow j}^{a \rightarrow i} = \frac{\partial \lambda_{a \rightarrow i}(t)}{\partial \lambda_{b \rightarrow j}(t-1)} = \frac{G(\lambda_{a \rightarrow i}(t-1)) \cdot F(\lambda_{c \rightarrow k}(t-1))}{\sqrt{1 + \sinh^2 \lambda_{a \rightarrow i}(t)}} \theta(-J_j^a J_j^b \cosh \lambda_{b \rightarrow j}(t-1)) \quad (13)$$

注意,上述是关于变量 $\lambda_{b \rightarrow j}$ 的求导,对于所有的子句 $c \in V(j) \setminus \{a, b\}, \lambda_{c \rightarrow j}$ 在公式(12)中均被看作常量.同样,对于所有的 $k \in V(a) \setminus \{i, j\}, d \in V(k) \setminus a, \lambda_{d \rightarrow k}$ 也被看做常量处理.令 $y(\lambda_{b \rightarrow j}) = -J_j^a \sum_{b \in V(j) \setminus a} J_j^b \sinh \lambda_{b \rightarrow j}$,考虑到截尾函数 $\theta(\cdot)$ 在公式(12)中的实际作用,先通过对函数 $y(\lambda_{b \rightarrow j})$ 进行求导,再进行截尾函数 $\theta(\cdot)$ 处理.因此,公式(12)通过求导后,截尾函数为 $\theta(-J_j^a J_j^b \cosh \lambda_{b \rightarrow j}(t-1))$.

又因为 $\cosh \lambda_{b \rightarrow j}(t-1) > 0$,所以 $\theta(-J_j^a J_j^b \cosh \lambda_{b \rightarrow j}(t-1)) = \theta(-J_j^a J_j^b)$.注意到:

$$\sup_{\lambda \in [0, \log(1 + \sqrt{2})]^{|D|}} \frac{G(\lambda_{a \rightarrow i}(t-1)) \cdot F(\lambda_{c \rightarrow k}(t-1))}{\sqrt{1 + \sinh^2 \lambda_{a \rightarrow i}(t)}} = 1 \quad (14)$$

因此,

$$|(h'(\lambda))_{b \rightarrow j}^{a \rightarrow i}| = \left| \frac{\partial \lambda_{a \rightarrow i}(t)}{\partial \lambda_{b \rightarrow j}(t-1)} \right| \leq \theta(-J_j^a J_j^b) \quad (15)$$

而

$$\begin{aligned} h'(\lambda)_{a \rightarrow i}^{a \rightarrow i} &= \frac{\partial \lambda_{a \rightarrow i}(t)}{\partial \lambda_{a \rightarrow i}(t-1)} \\ &= \frac{\delta \cosh \lambda_{a \rightarrow i}(t-1) (2 \sinh \lambda_{a \rightarrow i}(t-1) - 1) \prod_{j \in V(a) \setminus i} \theta \left(-J_j^a \sum_{b \in V(j) \setminus a} J_j^b \sinh \lambda_{b \rightarrow j}(t-1) \right)}{\sqrt{1 + \sinh^2 \lambda_{a \rightarrow i}(t)}} \\ &\leq \delta \cosh \lambda_{a \rightarrow i}(t-1) (2 \sinh \lambda_{a \rightarrow i}(t-1) - 1) \end{aligned} \quad (16)$$

由于 $\sinh \lambda_{a \rightarrow i} = \eta_{a \rightarrow i} \in [0, 1]$,因此 $h(\lambda)_{a \rightarrow i}^{a \rightarrow i} = \frac{\partial \lambda_{a \rightarrow i}(t)}{\partial \lambda_{a \rightarrow i}(t-1)} \leq \delta \cosh \lambda_{a \rightarrow i}(t-1)$.

又因为 $\cosh \lambda_{a \rightarrow i}(t-1) = \sqrt{1 + \sinh^2 \lambda_{a \rightarrow i}(t-1)} \leq \sqrt{2}$,

所以,

$$|(h'(\lambda))_{a \rightarrow i}^{a \rightarrow i}| = \left| \frac{\partial \lambda_{a \rightarrow i}(t)}{\partial \lambda_{a \rightarrow i}(t-1)} \right| \leq \sqrt{2} \delta \quad (17)$$

根据代数知识,利用函数 h 的导数构造雅可比矩阵 A .由压缩函数的性质及引理 2 可知:如果 A 的范数严格小于 1,则 h 为压缩函数,进而收敛到唯一固定点,且与初始信息无关.有如下结论:

定理 3. 对于 $\|A\|_1$ -范数,如果

$$\max_b \max_{j \in V(b)} \sum_{a \in V(j), b} \theta(-J_j^a J_j^b) < 1 - \sqrt{2}\delta \tag{18}$$

则公式(5)在 $\|A\|_1$ -范数下收敛,且与初始信息无关.

证明:根据公式(9)、公式(15)、公式(17),有:

$$\|A\|_1 = \|h'(\lambda)\|_1 = \max_{b \rightarrow j} \sum_{a \rightarrow i} \left| \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} \right| \leq \sqrt{2}\delta + \max_b \max_{j \in V(b)} \sum_{a \in V(j), b} \theta(-J_j^a J_j^b) \tag{19}$$

由压缩函数的性质和引理 2 可知:如果公式(20)成立,则函数 h 收敛到唯一固定点,且与初始信息无关:

$$\max_b \max_{j \in V(b)} \sum_{a \in V(j), b} \theta(-J_j^a J_j^b) < 1 - \sqrt{2}\delta \tag{20}$$

根据代数知识,函数 f 收敛到某一固定点,且与初始信息无关.所以,公式(5)在 $\|A\|_1$ -范数下收敛,且与初始信息无关. \square

同理可证如下结论.

定理 4. 对于 $\|A\|_\infty$ -范数,如果:

$$\max_a \max_{j \in V(a)} \sum_{b \in V(j), a} \theta(-J_j^a J_j^b) < 1 - \sqrt{2}\delta \tag{21}$$

则公式(5)在 $\|A\|_\infty$ -范数下收敛,且与初始信息无关.

证明:根据公式(10)、公式(15)、公式(17),有:

$$\|A\|_\infty = \|h'(\lambda)\|_\infty = \max_{a \rightarrow i} \sum_{b \rightarrow j} \left| \frac{\partial \lambda_{a \rightarrow i}}{\partial \lambda_{b \rightarrow j}} \right| \leq \sqrt{2}\delta + \max_a \max_{j \in V(a)} \sum_{b \in V(j), a} \theta(-J_j^a J_j^b) \tag{22}$$

由压缩函数的性质和引理 2 可知:如果公式(23)成立,则函数 h 收敛到唯一固定点,且与初始信息无关:

$$\max_a \max_{j \in V(a)} \sum_{b \in V(j), a} \theta(-J_j^a J_j^b) < 1 - \sqrt{2}\delta \tag{23}$$

根据代数知识,函数 f 收敛到某一固定点,且与初始信息无关.所以,则公式(5)在 $\|A\|_\infty$ -范数下收敛,且与初始信息无关. \square

设 M 是一个方阵, $\gamma_1, \dots, \gamma_n$ 是 M 的特征值,称 $\gamma(M) = \{\gamma_1, \dots, \gamma_n\}$ 为方阵 M 的谱, $\rho(M) = \sup\{|\gamma_i| : \gamma_i \in \gamma(M)\}$ 为方阵 M 的谱半径.

定理 5. 设函数 $h: V' \rightarrow V'$ 可微,进一步假设 $|h'(x)| \leq M$, M 是不依赖于 x 的非负矩阵.如果 $\rho(M) < 1$,那么对于任意 $x \in V'$, h 收敛到唯一固定点 x_∞ .也即, $x, h(x), h^2(x), \dots, x_\infty$ (注意: $|h'(x)| \leq M$ 表示由 $h'(x)$ 构成的雅可比矩阵中的每个元素的绝对值不大于 M 中对应元素的值).

证明:对于任意的 $x \in V'$, 假设 $n=1, 2, \dots$, 我们有:

$$\|(h^n)'(x)\| = \left\| \prod_{i=1}^n h'(h^{i-1}(x)) \right\| = \left\| \prod_{i=1}^n h'(h^{i-1}(x)) \right\| \leq \left\| \prod_{i=1}^n |h'(h^{i-1}(x))| \right\| \leq \left\| \prod_{i=1}^n M \right\| = \|M^n\| \tag{24}$$

由 Gelfand 定理^[40]: $\lim_{n \rightarrow \infty} \|M^n\|^{1/n} = \rho(M)$. 我们选取 $\varepsilon > 0$, 使得 $\rho(M) + \varepsilon < 1$. 对于某个 $N \in \mathbb{N}$, $\|M^N\| = (\rho(M) + \varepsilon)^N < 1$.

因此, $\|(h^N)'(x)\| < 1$. 所以, $\|(h^N)(x)\|$ 是压缩函数.

由引理 1 可知:函数 $h(x)$ 收敛到唯一固定点 x_∞ , 且与初始信息无关. \square

对于任意的 $i, j \in V(a), i \neq j, b \in V(j), a$, 定义 $|D| \times |D|$ 的非负方阵 M . M 中的元素:

$$M_{a \rightarrow i, a \rightarrow i} = \sqrt{2}\delta, M_{a \rightarrow i, b \rightarrow j} = 1_{V(j), a}(b) 1_{V(a), i}(j) \theta(-J_j^a J_j^b).$$

其中, $1_X(x)$ 是标识函数, 若 $x \in X$, 则 $1_X(x) = 1$; 若 $x \notin X$, 则 $1_X(x) = 0$, δ 为非负常数. 根据公式(15)、公式(17)和定理 5, 我们有如下结论:

定理 6. 如果方阵 M 的谱半径 $\rho(M)$ 严格小于 1, 则迭代方程(5)式收敛, 且与初始信息无关.

证明:根据公式(15)、公式(17)和定理 5, 函数 h 收敛到唯一固定点, 且与初始信息无关. 进一步地, 函数 f 收敛到唯一固定点, 且与初始信息无关. 所以, 方程(5)收敛, 且与初始信息无关. \square

假设 M 的特征值 γ , 根据代数知识, 我们有: $|\gamma| \cdot \|x\| = \|\gamma x\| = \|Mx\| \leq \|M\| \cdot \|x\|$. 因此, 定理 6 的条件弱于定理 3 和定理 4.

设向量空间 $V^m = \{0, 1\}^D \subset V$, 定义函数 $g: V^m \rightarrow V^m$. 假设 WP 算法中的信息更新是并行执行, 对于任意的向量 $u \in V^m$, 函数 g 的分量 $g(u)_{a \rightarrow i}$ 定义为 $g(u)_{a \rightarrow i} = u_{a \rightarrow i}$. 自然, WP 算法的执行过程可用函数 g 的迭代过程来刻画. 根据函数 f 和 g 的构造, 显然, 对于任意的 $x \in V^m$, 有 $f(x) = g(x)$.

引理 3. 如果函数 f 是一个压缩映射, 那么函数 g 也为压缩映射.

证明: 由于 f 是一个压缩映射, 对于任意的 $x \in V$, 存在唯一固定点 x_∞ , 有 $x, f(x), f^2(x), \dots, x_\infty$, 也即: 函数 f 收敛到 x_∞ , 且与初始信息无关. 因此, 对于任意的 $y \in V^m$, 仍然有 $y, f(y), f^2(y), \dots, x_\infty$. 而当 $y \in V^m$ 时, 有 $f(y) = g(y)$. 进而有 $y, g(y), g^2(y), \dots, x_\infty$. 因此, 函数 g 为压缩映射, 收敛到固定点 x_∞ , 且与初始信息无关. \square

根据定理 6 和引理 3, 至此, 定理 2 得证.

注意到: 定理 2 中非负矩阵的元素由公式的结构决定. 因此, 参数 δ 的值取决于公式的结构. 显然, 对于定理 1 中的结论, 容易使用定理 2 来证明. 为了方便理解, 下面举例说明定理 2 的使用.

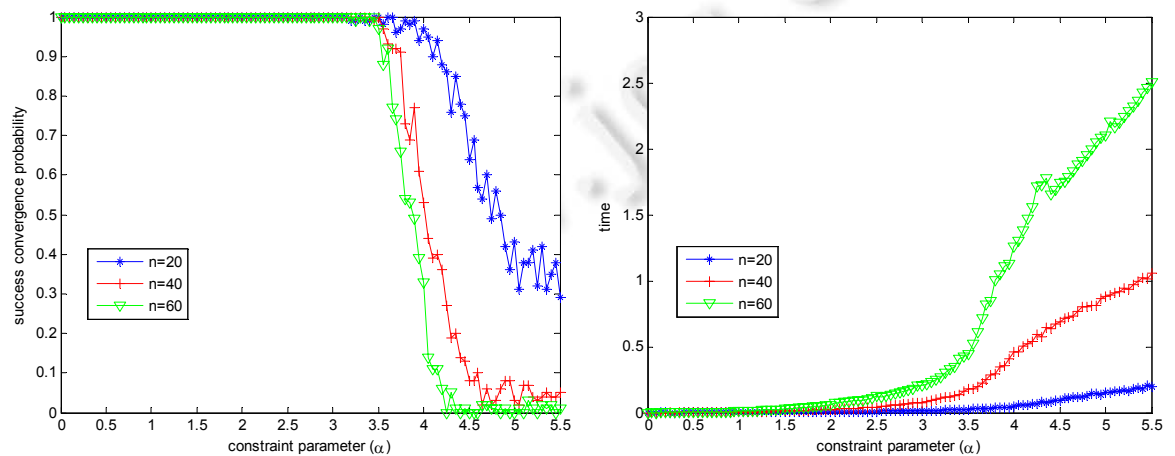
例 1: 设 CNF 公式 $F_1 = \{(x_1 \vee x_2), (x_2 \vee x_3), \dots, (x_{n-1} \vee x_n), (x_n \vee x_1)\}$, 其因子图含有一个环. 因此, 矩阵 M 的谱半径 $\rho(M) = \sqrt{2}\delta$. 若取 $0 < \delta < \frac{1}{\sqrt{2}}$, 则 $\rho(M) < 1$. 根据定理 2, 则 WP 算法在公式 F_1 上收敛.

同样, WP 算法在公式 $F_2 = \{(-x_1 \vee -x_2), (-x_2 \vee -x_3), \dots, (-x_{n-1} \vee -x_n), (-x_n \vee -x_1)\}$ 上也收敛.

3 数值实验及分析

在数值实验中, 使用模型 $G(n, 3, m)$ 产生随机 3-SAT 实例, 其中, n 表示变元个数, m 表示子句个数, 子句大小为 3. 模型 $G(n, 3, m)$ 按如下方式产生随机 3-SAT 实例: 随机均匀地在可能的子句 (这样的子句共有 $2^3 \cdot \binom{n}{3}$ 个) 中选择 $m = \alpha n$ 个子句, 被选中的子句之集构成实例. 在算法中, 取最大迭代次数 $t_{\max} = 10^3$.

我们选取 $G(n, 3, m)$ 随机实例进行模拟实验. 图 1 是 WP 算法分别在变元规模 $n=20, n=40, n=60$ 上的运行情况. 图 1(a) 是算法的收敛性随着约束参数 α 的变化情况, 图 1(b) 是算法的计算时间随着约束参数 α 的变化情况. 图中每个数据点由 $G(n, 3, m)$ 模型产生的 100 个随机实例构成. 可以看出: 当参数 $\alpha < 3.5$ 时, WP 算法高概率收敛; 当 $\alpha > 4.5$ 时, 算法高概率不收敛. 在 $\alpha \approx 3.5$ 点, WP 算法的收敛性发生了突变. 同样, 当参数 $\alpha < 3.5$ 时, WP 算法的运行时间较小; 当 $\alpha > 3.5$ 时, WP 算法的运行时间突然增大.



(a) 算法的收敛性

(b) 算法的计算时间

Fig.1 Change of convergence probability and time of WP with parameter α

图 1 WP 算法的收敛性和计算时间随约束参数 α 的变化情形

对于随机 3-SAT 实例, α 是重要的参数, 它影响实例的可满足性和判定难度. 存在可满足性的相变点 a_d , 当 $\alpha < a_d$ 时, 实例高概率可满足; 当 $\alpha > a_d$ 时, 实例高概率不可满足. 同时, 存在第 2 相变点 a_c , $a < a_c$ 为易解 SAT 区域, $a > a_c$ 为难解 SAT 区域. 大量的材料表明: a_d 至少为 3.52, 至多为 4.506, $a_c \approx 3.921$. 注意到, WP 算法在 $3.52 < \alpha < 4.506$ 区域的收敛情况不能从概率上给出确切的回答. 因此, 警示传播算法能够有效判定的随机 3-SAT 实例最多为 $\alpha < 3.52$ 区域. 从图 1(a) 中可以看出: 随着问题规模 n 的增大, WP 算法在 $G(n, 3, m)$ 实例上收敛的概率趋向于一个稳定值, 即, 收敛的概率曲线随着 n 的增大几乎趋于不变. 我们知道, 因子图的结构直接影响 WP 算法的收敛性. 事实上, 当参数 $\alpha > a_c$ 时, $G(n, 3, m)$ 模型产生的随机实例的因子图结构比较复杂, 求解难度较大. 从图 1(b) 中我们也可以看出: 大约在 $\alpha > 3.5$ 处, 算法的运行时间急剧上升.

为了验证定理 2 中的判定条件, 我们分别取了 $n=20$ 和 $n=60$ 的 $G(n, 3, m)$ 实例数据集进行模拟实验. 从图 2 中可以看出: 当 $\alpha < 1.8$ 时, 定理 2 能够判定 WP 算法的收敛性. 事实上, 文献[39]已经表明: 当约束参数 α 较小时, $G(n, 3, m)$ 实例的因子图的结构内部呈现树形状. 因此, 根据定理 2, 容易判定 WP 算法收敛. 随着 α 的增大, 因子图结构中开始出现圈, 并且圈的数量随着 α 的进一步增大而急剧增加. 由于定理 2 中的非负矩阵的元素构成与因子图的结构有关, 当因子图的结构变得复杂时, 非负矩阵的谱半径不再小于 1. 这时, 定理 2 中的判定条件失效. 如图 2 所示: 当 $\alpha > 1.8$ 时, 定理 2 中的判定条件逐渐失效.

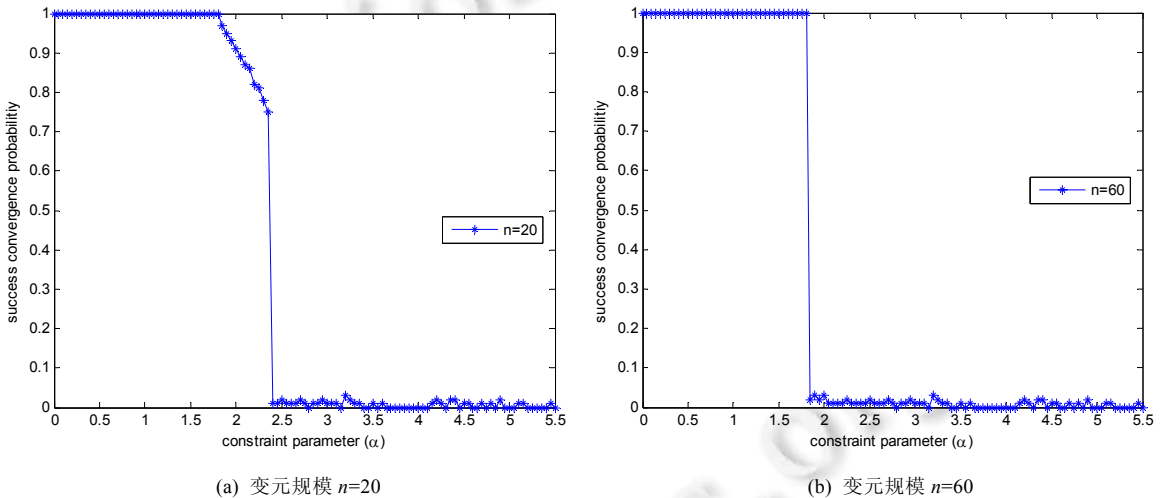


Fig.2 Results of Th.2

图 2 定理 2 的实验结果

如果一个因子图包含环, 且只含正文字或只含负文字的实例 F , 根据定理 2 容易判断, 则 WP 算法在 F 上收敛. 我们分别选取 3 组不同变元规模的由 $G(n, 3, m)$ 模型产生的随机 3SAT 可满足实例进行实验模拟, 这里取 $\alpha=2.0$. 这些实例只含正文字或只含负文字, 每组数据有 200 个实例(其中, 有 100 个只包含正文字的实例和 100 个只包含负文字的实例). 实验结果表明: 在这种随机实例上, WP 算法收敛. 表 1 是 WP 算法的收敛情况统计表. 用符号“+”表示只含正文字的实例; 用符号“-”表示只含负文字的实例. Var_num 表示实例的变元规模, Instance 表示实例集, Convergence 表示 WP 算法收敛的概率.

Table 1 The convergence of WP (Theory 2)

表 1 WP 算法的收敛情况(定理 2)

Var_num	Instance	Convergence
20	+ -	1
40	+ -	1
60	+ -	1

4 结束语

本文分析了 WP 算法的收敛性,通过引入参数,将 WP 算法的信息取值从 $\{0,1\}$ 松弛为 $[0,1]$,利用向量空间中压缩函数的性质,给出了 WP 算法收敛的一个充分条件.大量的实验结果表明,该方法有效.同时,该方法提供了理解 WP 算法性能的理论依据,并为分析其他信息传播算法的收敛性奠定了理论基础.进一步工作是:给出 WP 算法更弱的收敛条件.

References:

- [1] Martin D, Alan F, Michael M. A probabilistic analysis of randomly generated binary constraint satisfaction. *Theory Computer Science*, 2003,290:1815–1828. [doi: 10.1016/S0304-3975(02)00317-1]
- [2] Creignou N, Eaude H. The SAT-UNSAT transition for random constraint satisfaction problems. *Discrete Mathematics*, 2009,309:2085–2099. [doi: 10.1016/j.disc.2008.04.025]
- [3] Martin OC, Monasson R, Zecchina R. Statistical mechanics methods and phase transitions in optimization. *Theory Computer Science*, 2001,265:3–67. [doi: 10.1016/S0304-3975(01)00149-9]
- [4] Kirkpatrick S, Selman B. Critical behavior in the satisfiability of random Boolean formulae. *Science*, 1994,264(5163):1297–1301. [doi: 10.1126/science.264.5163.1297]
- [5] Mertens S, Mezard M, Zecchina R. Threshold values of random k -SAT from the cavity method. *Random Structure Algorithms*, 2006,28:340–373. [doi: 10.1002/rsa.v28:3]
- [6] Kaporis A, Kirousis L, Lalas E. The probabilistic analysis of a greedy satisfiability algorithm. *Random Structures & Algorithms*, 2006,28(4):444–480. [doi: 10.1002/rsa.v28:4]
- [7] Dubois O, Boufkhad Y, Mandler J. Typical random 3-sat formulae and the satisfiability threshold. *Electronic Colloquium on Computational Complexity*, 2003,10(007):1–2.
- [8] Moskewicz MW, Madigan CF, Zhao Y. Chaff: Engineering an efficient SAT solver. In: *Proc. of the 38th Annual Design Automation Conf.* New York: ACM Press, 2001. 530–535. [doi: 10.1145/378239.379017]
- [9] Aurell E, Gordon U, Kirkpatrick S. Comparing beliefs, surveys, and random walks. *Advances in Neural Information Processing Systems*, 2004,17(1):1–8.
- [10] Mezard M, Parisi G. The cavity method at zero temperature. *Journal of Statistical Physics*, 2003,111(1-2):1–34.
- [11] Mezard M, Zecchina R. Random k -satisfiability problem: From an analytic solution to an efficient algorithm. *Physical Review E*, 2002,66(5):056126.
- [12] Maneva E, Mossel E, Wainwright M. A new look at survey propagation and its generalizations. *Journal of the ACM*, 2007,54(4):1089–1098.
- [13] Braunstein A, Mezard M, Zecchina R. Survey propagation: An algorithm for satisfiability. *Random Structures and Algorithms*, 2005,27(2):201–226.
- [14] Yedidia JS, Freeman WT, Weiss Y. Understanding belief propagation and its generalizations. *Artificial Intelligence*, 2003,8(1):239–269.
- [15] Braunstein A, Zecchina R. Survey and belief propagation on random k -SAT. *Lecture Notes in Computer Science*, 2004,2919(1):519–528.
- [16] Xu K, Li W. Exact phase transitions in random constraint satisfaction problems. *Journal of Artificial Intelligence Research*, 2000,12(1):93–103.
- [17] Xu K, Li W. Many hard examples in exact phase transitions. *Theory Computer Science*, 2006,355(1):291–302. [doi: 10.1016/j.tcs.2006.01.001]
- [18] Xu K, Boussemart F, Hemery F, Lecoutre C. Random constraint satisfaction: Easy generation of hard instances. *Artificial Intelligence*, 2007,171(1):514–534. [doi: 10.1016/j.artint.2007.04.001]
- [19] 赵春艳,郑志明.一种基于变量熵求解约束满足问题的置信传播算法. *中国科学(F:信息科学)*,2012,42(9):1170–1180. [doi: 10.1360/112011-693]
- [20] Yin MH, Zhou JP, Sun JG, Gu W. Heuristic survey propagation algorithm for solving QBF problem. *Ruan Jian Xue Bao/Journal of Software*, 2011,22(7):1538–1550 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3859.htm> [doi: 10.3724/SP.J.1001.2011.03859]
- [21] Ravanbakhsh S, Greiner R. Perturbed message passing for constraint satisfaction problems. *Artificial Intelligence*, 2014. arXiv:1401.6686
- [22] Chieu HL, Lee WS. Relaxed survey propagation for the weighted maximum satisfiability problem. *Journal of Artificial Intelligence Research*, 2009,36:229–266.

- [23] Bayati M, Shah D, Sharma M. Max product for maximum weight matching: Convergence, correctness, and LP duality. *IEEE Trans. on Information Theory*, 2007,54(3):1241–1251.
- [24] Coja-Oghlan A, Mossel E, Vilenchik D. A spectral approach to analyzing belief propagation for 3-colouring. *Combinatorics, Probability and Computing*, 2009,18(6):881–912.
- [25] Gamarnik D, Shah D, Wei Y. Belief propagation for min-cost network flow: Convergence and correctness. *Operations Research*, 2012,60(2):410–428.
- [26] Sanghavi S, Malioutov DM, Willsky AS. Belief propagation and LP relaxation for weighted matching in general graphs. *IEEE Trans. on Information Theory*, 2011,57(4):2203–2212.
- [27] Sanghavi S, Shah D, Willsky AS. Message passing for maximum weight independent set. *IEEE Trans. on Information Theory*, 2009, 55(11):4822–4834.
- [28] Kschischang F, Frey B, Loeliger H. Factor graph and the sum-product algorithm. *Information Theory*, 2001,47(1):498–519.
- [29] Weiss Y. Correctness of local probability propagation in graphical models with loops. *Neural Computation*, 2000,12(1):1–41.
- [30] Weiss Y, Freeman WT. Correctness of belief propagation in Gaussian graphical models of arbitrary topology. *Neural Computation*, 2001,13(10):2173–2200. [doi: 10.1162/089976601750541769]
- [31] Tatikonda S, Jordan MI. Loopy belief propagation and Gibbs measure. In: *Proc. of the 18th Annual Conf. on Uncertainty in Artificial Intelligence*. 2002. 493–500.
- [32] Heskes T. On the uniqueness of loopy belief propagation fixed points. *Neural Computation*, 2004,16:2379–2413. [doi: 10.1162/0899766041941943]
- [33] Ihler AT. Loopy belief propagation: Convergence and effects of message errors. *Machine Learning Research*, 2005,6:905–936.
- [34] Mooij JM, Kappen HJ. Sufficient conditions for convergence of the sum-product algorithm. *IEEE Trans. on Information Theory*, 2007,53:4422–4437.
- [35] Shi XQ, Schonfeld D, Tuninetti D. Message error analysis of loopy belief propagation for the sum-product algorithm. *Computer and Information Science*, 2010,1009:1–30.
- [36] Brunsch T, Cornelissen K, Manthey B, Roglin H. Smoothed analysis of BP for minimum cost flow and matching. *Journal of Graph Algorithms and Applications*, 2013,17(6):647–670.
- [37] Norshams N, Wainwright MJ. Stochastic belief propagation: A low complexity alternative to the sum product algorithm. *IEEE Trans. on Information Theory*, 2013,59(4):1981–2000.
- [38] Feige U, Mossel E, Vilenchik D. Complete convergence of message passing algorithms for some satisfiability problems. *Theory of Computing*, 2013,9(19):617–651. [doi: 10.4086/toc.2013.v009a019]
- [39] Wang XF, Xu DY, Wei L. Convergence of warning propagation algorithms for random satisfiable instances. *Ruan Jian Xue Bao/ Journal of Software*, 2013,24(1):1–11 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4213.htm> [doi: 10.3724/SP.J.1001.2013.04213]
- [40] Dieudonné J. *Foundations of Modern Analysis*. New York: Academic Press, 1969.

附中中文参考文献:

- [20] 殷明浩,周俊萍,孙吉贵,谷文祥.求解 QBF 问题的启发式调查传播算法. *软件学报*,2011,22(7):1538–1550. <http://www.jos.org.cn/1000-9825/3859.htm> [doi: 10.3724/SP.J.1001.2011.03859]
- [39] 王晓峰,许道云,韦立.随机可满足实例集上警示传播算法的收敛性. *软件学报*,2013,24(1):1–11. <http://www.jos.org.cn/1000-9825/4213.htm> [doi: 10.3724/SP.J.1001.2013.04213]



王晓峰(1980—),男,甘肃会宁人,博士,讲师,CCF 会员,主要研究领域为算法设计与分析,计算复杂性,人工智能.



许道云(1959—),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为计算复杂性,可计算性分析.