Min-sum and network flows

Patrick Rebeschini

(joint work with Sekhar Tatikonda)

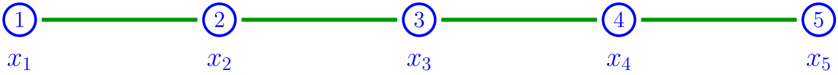
Yale Institute For Network Science

Optimization and Inference for Physical Flows on Networks
Banff International Research Station

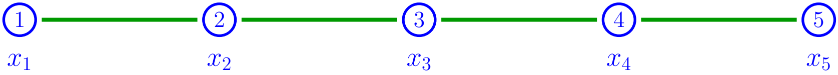
March 8, 2017

(NSF Grant: *Locality in Network Optimization*, Award no. 1609484, ECCS)

Min-sum: path graph

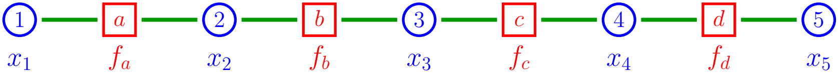


Min-sum: path graph



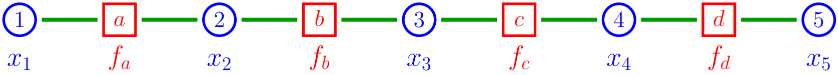
$$\min_{x_1} \min_{x_2} \min_{x_3} \min_{x_4} \min_{x_5} \underbrace{\left(f_a(x_1, x_2) + f_b(x_2, x_3) + f_c(x_3, x_4) + f_d(x_4, x_5) \right)}_{f(x)}$$

Min-sum: path graph



$$\min_{x_1} \min_{x_2} \min_{x_3} \min_{x_4} \min_{x_5} \underbrace{\left(f_a(x_1, x_2) + f_b(x_2, x_3) + f_c(x_3, x_4) + f_d(x_4, x_5) \right)}_{f(x)}$$

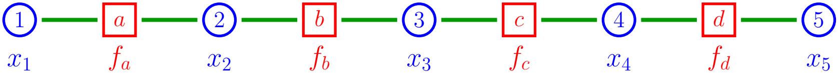
Min-sum: path graph



$$\min_{x_1} \min_{x_2} \min_{x_3} \min_{x_4} \min_{x_5} \underbrace{\left(f_a(x_1, x_2) + f_b(x_2, x_3) + f_c(x_3, x_4) + f_d(x_4, x_5) \right)}_{f(x)}$$

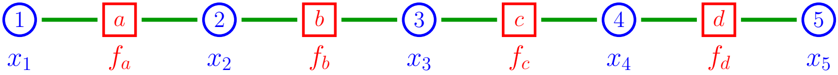
$x_i \in \{0, 1\}$ **naive algorithm** $O(2^n)$

Min-sum: path graph



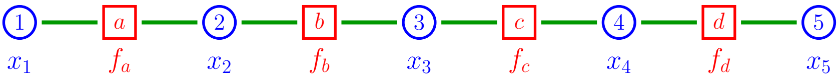
$$\min_{x_1} \min_{x_2} \min_{x_3} \min_{x_4} \boxed{\min_{x_5}} \left(f_a(x_1, x_2) + f_b(x_2, x_3) + f_c(x_3, x_4) + f_d(x_4, x_5) \right)$$

Min-sum: path graph



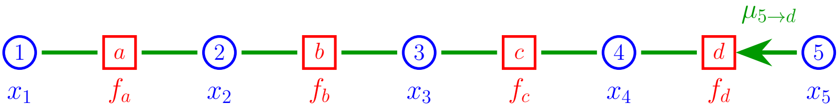
$$\min_{x_1} \min_{x_2} \min_{x_3} \min_{x_4} \left(f_a(x_1, x_2) + f_b(x_2, x_3) + f_c(x_3, x_4) + \min_{x_5} f_d(x_4, x_5) \right)$$

Min-sum: path graph



$$\min_{x_1} \min_{x_2} \min_{x_3} \min_{x_4} \left(f_a(x_1, x_2) + f_b(x_2, x_3) + f_c(x_3, x_4) + \boxed{\min_{x_5} f_d(x_4, x_5)} \right)$$

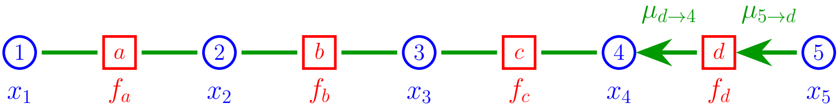
Min-sum: path graph



$$\min_{x_1} \min_{x_2} \min_{x_3} \min_{x_4} \left(f_a(x_1, x_2) + f_b(x_2, x_3) + f_c(x_3, x_4) + \boxed{\min_{x_5} f_d(x_4, x_5)} \right)$$

$\mu_{5\to d} = 0$

Min-sum: path graph

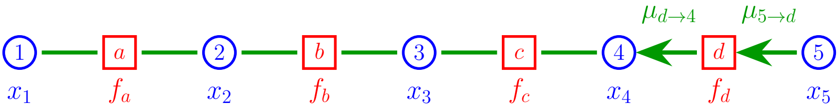


$$\min_{x_1} \min_{x_2} \min_{x_3} \min_{x_4} \left(f_a(x_1, x_2) + f_b(x_2, x_3) + f_c(x_3, x_4) + \boxed{\min_{x_5} f_d(x_4, x_5)} \right)$$

$$\mu_{5\to d} = 0$$

$$\mu_{d\to4}(\,\star\,) = \min_{x_5} \big(f_d(\,\star\,, x_5) + \mu_{5\to d}(x_5) \big)$$

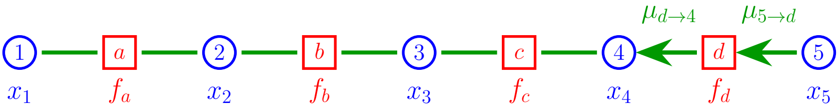
Min-sum: path graph



$$\min_{x_1} \min_{x_2} \min_{x_3} \min_{x_4} \left(f_a(x_1, x_2) + f_b(x_2, x_3) + f_c(x_3, x_4) + \mu_{d\to 4}(x_4) \right)$$

$$\mu_{5\to d} = 0 \qquad \mu_{d\to 4}(\ \star\ ) = \min_{x_5} \big(f_d(\ \star\ , x_5) + \mu_{5\to d}(x_5) \big)$$

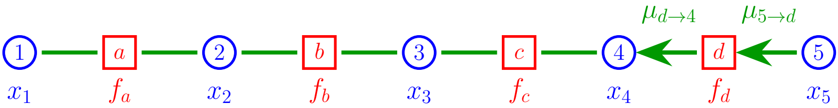
Min-sum: path graph



$$\min_{x_1} \min_{x_2} \min_{x_3} \min_{x_4} \Big(f_a(x_1, x_2) + f_b(x_2, x_3) + f_c(x_3, x_4) + \mu_{d\to 4}(x_4) \Big)$$

$$\mu_{5\to d} = 0 \qquad \mu_{d\to 4}(\,\star\,) = \min_{x_5} \big(f_d(\,\star\,, x_5) + \mu_{5\to d}(x_5) \big)$$

Min-sum: path graph

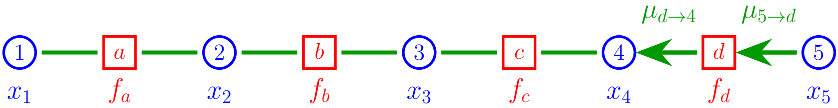


$$\min_{x_1} \min_{x_2} \min_{x_3} \left(f_a(x_1, x_2) + f_b(x_2, x_3) + \boxed{\min_{x_4}} \left(f_c(x_3, x_4) + \mu_{d\to 4}(x_4) \right) \right)$$

$$\mu_{5\to d} = 0$$

$$\mu_{d\to 4}(\,\star\,) = \min_{x_5} \left(f_d(\,\star\,, x_5) + \mu_{5\to d}(x_5) \right)$$

Min-sum: path graph

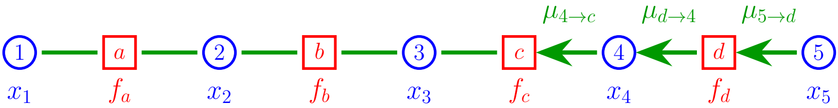


$$\min_{x_1} \min_{x_2} \min_{x_3} \left(f_a(x_1, x_2) + f_b(x_2, x_3) + \boxed{\min_{x_4} \Big(f_c(x_3, x_4) + \mu_{d\to 4}(x_4) \Big)} \right)$$

$\mu_{5\to d} = 0$

$$\mu_{d\to 4}(\star) = \min_{x_5} \big(f_d(\star, x_5) + \mu_{5\to d}(x_5) \big)$$

Min-sum: path graph



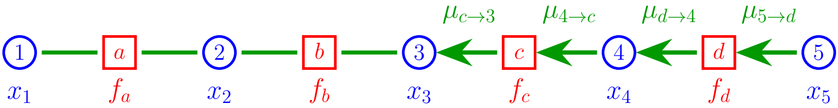
$$\min_{x_1} \min_{x_2} \min_{x_3} \left(f_a(x_1, x_2) + f_b(x_2, x_3) + \boxed{\min_{x_4} \left(f_c(x_3, x_4) + \mu_{d\to 4}(x_4) \right)} \right)$$

$$\mu_{5\to d} = 0$$

$$\mu_{d\to 4}(\,\star\,) = \min_{x_5} \left(f_d(\,\star\,, x_5) + \mu_{5\to d}(x_5) \right)$$

$$\mu_{4\to c} = \mu_{d\to 4}$$

Min-sum: path graph

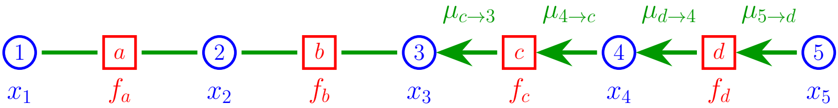


$$\min_{x_1} \min_{x_2} \min_{x_3} \left(f_a(x_1, x_2) + f_b(x_2, x_3) + \boxed{\min_{x_4} \Big(f_c(x_3, x_4) + \mu_{d\to 4}(x_4) \Big)} \right)$$

$$\mu_{5\to d} = 0 \qquad \mu_{d\to 4}(\star) = \min_{x_5} \big(f_d(\star, x_5) + \mu_{5\to d}(x_5) \big)$$

$$\mu_{4\to c} = \mu_{d\to 4} \qquad \mu_{c\to 3}(\star) = \min_{x_4} \big(f_c(\star, x_4) + \mu_{4\to c}(x_4) \big)$$

Min-sum: path graph

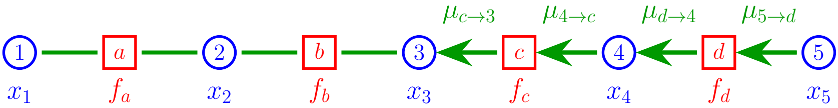


$$\min_{x_1} \min_{x_2} \min_{x_3} \left(f_a(x_1, x_2) + f_b(x_2, x_3) + \mu_{c\to 3}(x_3) \right)$$

$$\mu_{5\to d} = 0 \qquad \mu_{d\to 4}(\,\star\,) = \min_{x_5} \big(f_d(\,\star\,, x_5) + \mu_{5\to d}(x_5) \big)$$

$$\mu_{4\to c} = \mu_{d\to 4} \qquad \mu_{c\to 3}(\,\star\,) = \min_{x_4} \big(f_c(\,\star\,, x_4) + \mu_{4\to c}(x_4) \big)$$

Min-sum: path graph

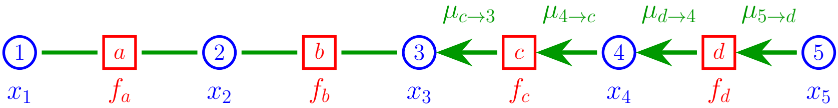


$$\min_{x_1} \min_{x_2} \min_{x_3} \Big(f_a(x_1, x_2) + f_b(x_2, x_3) + \mu_{c\to 3}(x_3) \Big)$$

$$\mu_{5\to d} = 0 \qquad \mu_{d\to 4}(\star) = \min_{x_5} \big(f_d(\star, x_5) + \mu_{5\to d}(x_5)\big)$$

$$\mu_{4\to c} = \mu_{d\to 4} \qquad \mu_{c\to 3}(\star) = \min_{x_4} \big(f_c(\star, x_4) + \mu_{4\to c}(x_4)\big)$$

Min-sum: path graph

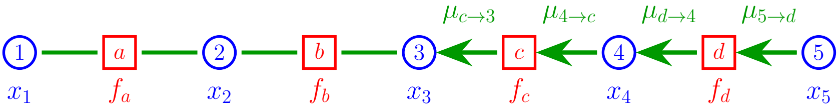


$$\min_{x_1} \min_{x_2} \left(f_a(x_1, x_2) + \boxed{\min_{x_3}} \left(f_b(x_2, x_3) + \mu_{c\to3}(x_3) \right) \right)$$

$$\mu_{5\to d} = 0 \qquad \mu_{d\to4}(\star) = \min_{x_5} \big(f_d(\star, x_5) + \mu_{5\to d}(x_5) \big)$$

$$\mu_{4\to c} = \mu_{d\to4} \qquad \mu_{c\to3}(\star) = \min_{x_4} \big(f_c(\star, x_4) + \mu_{4\to c}(x_4) \big)$$

Min-sum: path graph



$$\min_{x_1} \min_{x_2} \left(f_a(x_1, x_2) + \boxed{\min_{x_3} \left(f_b(x_2, x_3) + \mu_{c\to 3}(x_3) \right)} \right)$$

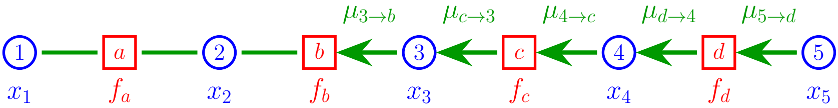
$\mu_{5\to d} = 0$

$$\mu_{d\to 4}(\,\star\,) = \min_{x_5} \left(f_d(\,\star\,, x_5) + \mu_{5\to d}(x_5) \right)$$

$\mu_{4\to c} = \mu_{d\to 4}$

$$\mu_{c\to 3}(\,\star\,) = \min_{x_4} \left(f_c(\,\star\,, x_4) + \mu_{4\to c}(x_4) \right)$$

Min-sum: path graph



$$\min_{x_1} \min_{x_2} \left(f_a(x_1, x_2) + \boxed{\min_{x_3} \left(f_b(x_2, x_3) + \mu_{c\to 3}(x_3) \right)} \right)$$

$$\mu_{5\to d} = 0$$

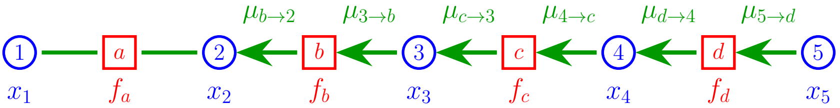
$$\mu_{4\to c} = \mu_{d\to 4}$$

$$\mu_{3\to b} = \mu_{c\to 3}$$

$$\mu_{d\to 4}(\,\star\,) = \min_{x_5} \big(f_d(\,\star\,, x_5) + \mu_{5\to d}(x_5) \big)$$

$$\mu_{c\to 3}(\,\star\,) = \min_{x_4} \big(f_c(\,\star\,, x_4) + \mu_{4\to c}(x_4) \big)$$

Min-sum: path graph



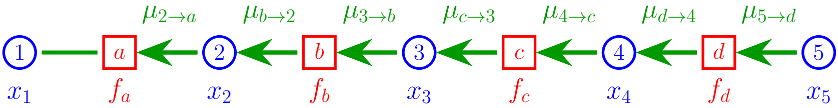
$$\min_{x_1} \min_{x_2} \left(f_a(x_1, x_2) + \boxed{\min_{x_3} \left(f_b(x_2, x_3) + \mu_{c\to 3}(x_3) \right)} \right)$$

$$\mu_{5\to d} = 0 \qquad \mu_{d\to 4}(\star) = \min_{x_5} \left(f_d(\star, x_5) + \mu_{5\to d}(x_5) \right)$$

$$\mu_{4\to c} = \mu_{d\to 4} \qquad \mu_{c\to 3}(\star) = \min_{x_4} \left(f_c(\star, x_4) + \mu_{4\to c}(x_4) \right)$$

$$\mu_{3\to b} = \mu_{c\to 3} \qquad \mu_{b\to 2}(\star) = \min_{x_3} \left(f_b(\star, x_3) + \mu_{c\to 3}(x_3) \right)$$

Min-sum: path graph



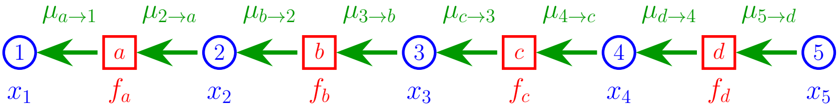
$$\min_{x_1} \min_{x_2} \left(f_a(x_1, x_2) + \mu_{b\to 2}(x_2) \right)$$

$$\mu_{5\to d} = 0 \qquad \mu_{d\to 4}(\,\star\,) = \min_{x_5} \big(f_d(\,\star\,, x_5) + \mu_{5\to d}(x_5)\big)$$

$$\mu_{4\to c} = \mu_{d\to 4} \qquad \mu_{c\to 3}(\,\star\,) = \min_{x_4} \big(f_c(\,\star\,, x_4) + \mu_{4\to c}(x_4)\big)$$

$$\mu_{3\to b} = \mu_{c\to 3} \qquad \mu_{b\to 2}(\,\star\,) = \min_{x_3} \big(f_b(\,\star\,, x_3) + \mu_{c\to 3}(x_3)\big)$$

Min-sum: path graph



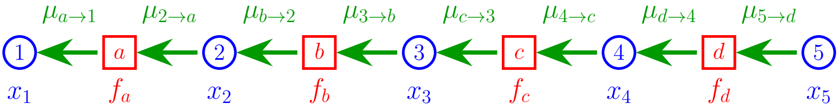
$$\min_{x_1} \boxed{\min_{x_2}} \Big(f_a(x_1, x_2) + \mu_{b\to2}(x_2) \Big)$$

$$\mu_{5\to d} = 0 \qquad \mu_{d\to4}(\,\star\,) = \min_{x_5} \big(f_d(\,\star\,, x_5) + \mu_{5\to d}(x_5) \big)$$

$$\mu_{4\to c} = \mu_{d\to4} \qquad \mu_{c\to3}(\,\star\,) = \min_{x_4} \big(f_c(\,\star\,, x_4) + \mu_{4\to c}(x_4) \big)$$

$$\mu_{3\to b} = \mu_{c\to3} \qquad \mu_{b\to2}(\,\star\,) = \min_{x_3} \big(f_b(\,\star\,, x_3) + \mu_{c\to3}(x_3) \big)$$

Min-sum: path graph



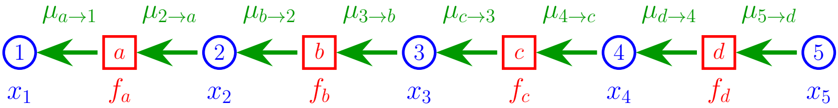
$$\min_{x_1} \boxed{\min_{x_2} \Big(f_a(x_1, x_2) + \mu_{b\to 2}(x_2) \Big)}$$

$$\mu_{5\to d} = 0 \qquad \mu_{d\to 4}(\,\star\,) = \min_{x_5} \big(f_d(\,\star\,, x_5) + \mu_{5\to d}(x_5) \big)$$

$$\mu_{4\to c} = \mu_{d\to 4} \qquad \mu_{c\to 3}(\,\star\,) = \min_{x_4} \big(f_c(\,\star\,, x_4) + \mu_{4\to c}(x_4) \big)$$

$$\mu_{3\to b} = \mu_{c\to 3} \qquad \mu_{b\to 2}(\,\star\,) = \min_{x_3} \big(f_b(\,\star\,, x_3) + \mu_{c\to 3}(x_3) \big)$$

Min-sum: path graph



$$\min_{x_1} \boxed{\min_{x_2} \left(f_a(x_1, x_2) + \mu_{b\to 2}(x_2) \right)}$$

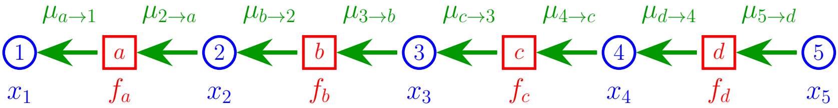
$$\mu_{5\to d} = 0 \qquad \mu_{d\to 4}(\star) = \min_{x_5} \left(f_d(\star, x_5) + \mu_{5\to d}(x_5) \right)$$

$$\mu_{4\to c} = \mu_{d\to 4} \qquad \mu_{c\to 3}(\star) = \min_{x_4} \left(f_c(\star, x_4) + \mu_{4\to c}(x_4) \right)$$

$$\mu_{3\to b} = \mu_{c\to 3} \qquad \mu_{b\to 2}(\star) = \min_{x_3} \left(f_b(\star, x_3) + \mu_{c\to 3}(x_3) \right)$$

$$\mu_{2\to a} = \mu_{b\to 2}$$

Min-sum: path graph



$$\min_{x_1} \boxed{\min_{x_2} \left(f_a(x_1, x_2) + \mu_{b\to 2}(x_2) \right)}$$

$$\mu_{5\to d} = 0 \qquad \mu_{d\to 4}(\star) = \min_{x_5} \left(f_d(\star, x_5) + \mu_{5\to d}(x_5) \right)$$

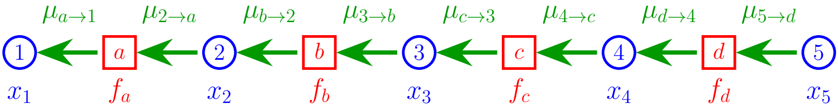
$$\mu_{4\to c} = \mu_{d\to 4} \qquad \mu_{c\to 3}(\star) = \min_{x_4} \left(f_c(\star, x_4) + \mu_{4\to c}(x_4) \right)$$

$$\mu_{3\to b} = \mu_{c\to 3} \qquad \mu_{b\to 2}(\star) = \min_{x_3} \left(f_b(\star, x_3) + \mu_{c\to 3}(x_3) \right)$$

$$\mu_{2\to a} = \mu_{b\to 2} \qquad \mu_{a\to 1}(\star) = \min_{x_2} \left(f_a(\star, x_2) + \mu_{b\to 2}(x_2) \right)$$

Min-sum: path graph

$O(n)$ **Dynamic Programming**



$$\min_{x_1} \mu_{a\to 1}(x_1)$$

$$\mu_{5\to d} = 0 \qquad \mu_{d\to 4}(\star) = \min_{x_5} \left(f_d(\star, x_5) + \mu_{5\to d}(x_5)\right)$$

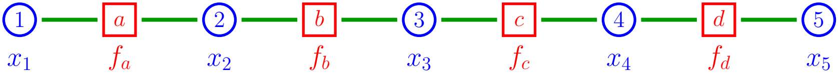
$$\mu_{4\to c} = \mu_{d\to 4} \qquad \mu_{c\to 3}(\star) = \min_{x_4} \left(f_c(\star, x_4) + \mu_{4\to c}(x_4)\right)$$

$$\mu_{3\to b} = \mu_{c\to 3} \qquad \mu_{b\to 2}(\star) = \min_{x_3} \left(f_b(\star, x_3) + \mu_{c\to 3}(x_3)\right)$$

$$\mu_{2\to a} = \mu_{b\to 2} \qquad \mu_{a\to 1}(\star) = \min_{x_2} \left(f_a(\star, x_2) + \mu_{b\to 2}(x_2)\right)$$

Min-sum: path graph

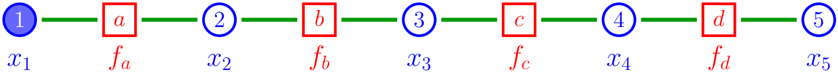
$O(n)$ **Dynamic Programming**



$$\min_{x_1} \min_{x_2} \min_{x_3} \min_{x_4} \min_{x_5} \underbrace{\left(f_a(x_1, x_2) + f_b(x_2, x_3) + f_c(x_3, x_4) + f_d(x_4, x_5) \right)}_{f(x)}$$

Min-sum: path graph

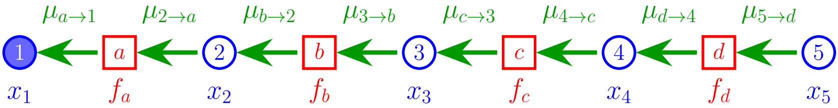
$O(n)$ **Dynamic Programming**



$$\min_{x_1}\min_{x_2}\left(f_a(x_1,x_2)+\min_{x_3}\left(f_b(x_2,x_3)+\min_{x_4}\left(f_c(x_3,x_4)+\min_{x_5}f_d(x_4,x_5)\right)\right)\right)$$

Min-sum: path graph

$O(n)$ **Dynamic Programming**

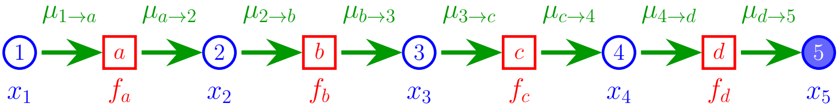


$$\min_{x_1} \underbrace{\min_{x_2}\left(f_a(x_1,x_2) + \underbrace{\min_{x_3}\left(f_b(x_2,x_3) + \underbrace{\min_{x_4}\left(f_c(x_3,x_4) + \underbrace{\min_{x_5} f_d(x_4,x_5)}_{\mu_{d\to 4}(x_4)}\right)}_{\mu_{c\to 3}(x_3)}\right)}_{\mu_{b\to 2}(x_2)}\right)}_{\mu_{a\to 1}(x_1)}$$

$$= \min_{x_1} \mu_{a\to 1}(x_1)$$

Min-sum: path graph

$O(n)$ **Dynamic Programming**

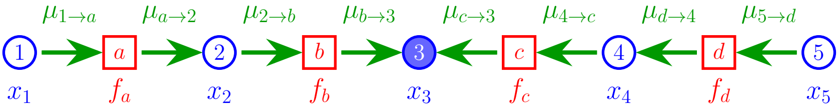


$$\underbrace{\min_{x_5} \min_{x_4} \left(\underbrace{\min_{x_3} \left(\underbrace{\min_{x_2} \left(\underbrace{\min_{x_1} f_a(x_1, x_2)}_{\mu_{a\to 2}(x_2)} + f_b(x_2, x_3) \right)}_{\mu_{b\to 3}(x_3)} + f_c(x_3, x_4) \right)}_{\mu_{c\to 4}(x_4)} + f_d(x_4, x_5) \right)}_{\mu_{d\to 5}(x_5)}$$

$$= \min_{x_5} \mu_{d\to 5}(x_5)$$

Min-sum: path graph

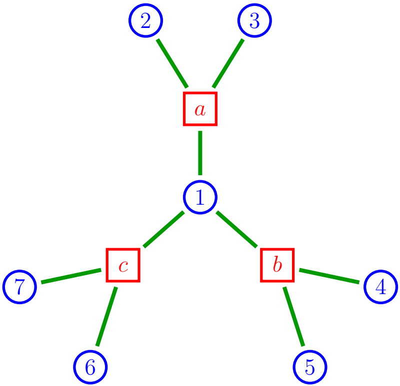
$O(n)$ **Dynamic Programming**



$$\underbrace{\min_{x_3}\left(\min_{x_2}\left(\underbrace{\min_{x_1} f_a(x_1,x_2)}_{\mu_{a\to 2}(x_2)} + f_b(x_2,x_3)\right)}_{\mu_{b\to 3}(x_3)} + \underbrace{\min_{x_4}\left(f_c(x_3,x_4) + \underbrace{\min_{x_5} f_d(x_4,x_5)}_{\mu_{d\to 4}(x_4)}\right)}_{\mu_{c\to 3}(x_3)}\right)$$

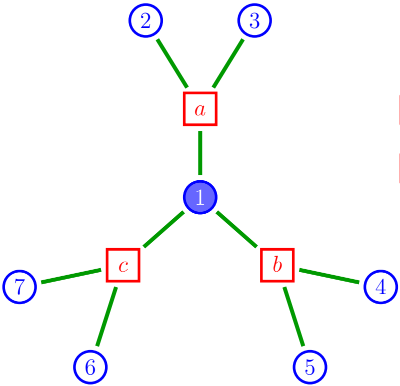
$$= \min_{x_3}\left(\mu_{b\to 3}(x_3) + \mu_{c\to 3}(x_3)\right)$$

Min-sum: trees



$$\min_{x}\left(f_a(x_1, x_2, x_3) + f_b(x_1, x_4, x_5) + f_c(x_1, x_6, x_7)\right)$$

Min-sum: trees



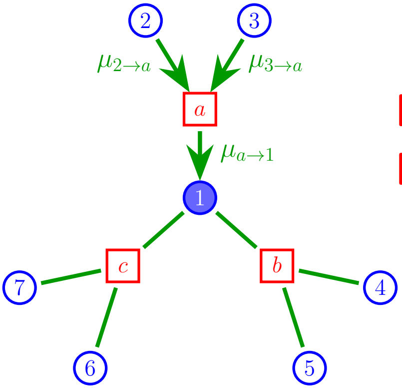
Dynamic Programming

$O(n)$

$$\min_{x}\Big(f_a(x_1,x_2,x_3)+f_b(x_1,x_4,x_5)+f_c(x_1,x_6,x_7)\Big)$$

$$=\min_{x_1}\Big(\min_{x_2,x_3} f_a(x_1,x_2,x_3)+\min_{x_4,x_5} f_b(x_1,x_4,x_5)+\min_{x_6,x_7} f_c(x_1,x_6,x_7)\Big)$$

Min-sum: trees



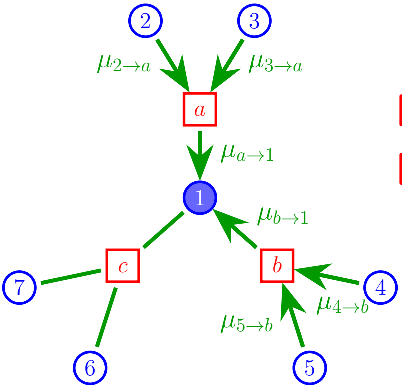
Dynamic Programming

$O(n)$

$$\min_{x}\left(f_a(x_1,x_2,x_3)+f_b(x_1,x_4,x_5)+f_c(x_1,x_6,x_7)\right)$$

$$=\min_{x_1}\left(\underbrace{\min_{x_2,x_3} f_a(x_1,x_2,x_3)}_{\mu_{a\to 1}(x_1)}+\min_{x_4,x_5} f_b(x_1,x_4,x_5)+\min_{x_6,x_7} f_c(x_1,x_6,x_7)\right)$$

Min-sum: trees



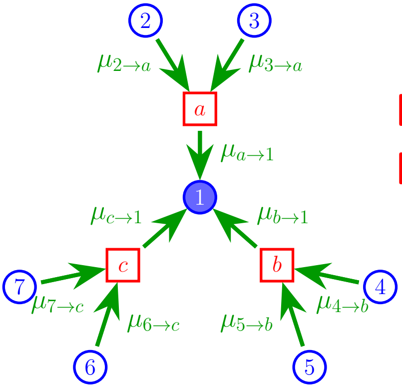
Dynamic Programming

$O(n)$

$$\min_x \left(f_a(x_1, x_2, x_3) + f_b(x_1, x_4, x_5) + f_c(x_1, x_6, x_7) \right)$$

$$= \min_{x_1} \left(\underbrace{\min_{x_2, x_3} f_a(x_1, x_2, x_3)}_{\mu_{a\to 1}(x_1)} + \underbrace{\min_{x_4, x_5} f_b(x_1, x_4, x_5)}_{\mu_{b\to 1}(x_1)} + \min_{x_6, x_7} f_c(x_1, x_6, x_7) \right)$$

Min-sum: trees



Dynamic Programming

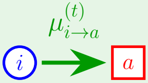
$O(n)$

$$\min_{x}\left(f_a(x_1,x_2,x_3)+f_b(x_1,x_4,x_5)+f_c(x_1,x_6,x_7)\right)$$

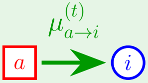
$$=\min_{x_1}\left(\underbrace{\min_{x_2,x_3} f_a(x_1,x_2,x_3)}_{\mu_{a\to 1}(x_1)}+\underbrace{\min_{x_4,x_5} f_b(x_1,x_4,x_5)}_{\mu_{b\to 1}(x_1)}+\underbrace{\min_{x_6,x_7} f_c(x_1,x_6,x_7)}_{\mu_{c\to 1}(x_1)}\right)$$

Messages

variable → factor

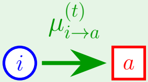


factor → variable



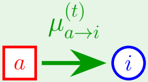
Messages

variable $\rightarrow$ factor



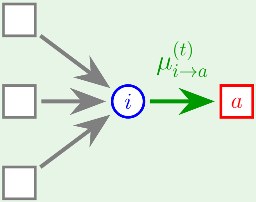
$\mu^{(t)}_{i\to a}(x_i) =$

factor $\rightarrow$ variable



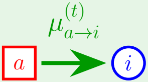
Messages

variable → factor



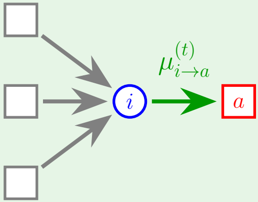
$\mu^{(t)}_{i\to a}(x_i) =$

factor → variable



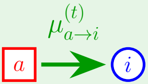
Messages

variable → factor



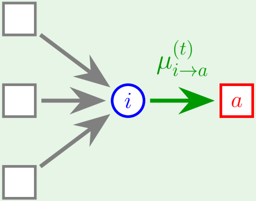
$$\mu^{(t)}_{i\to a}(x_i) = \sum_{b\in\partial i\setminus a} \mu^{(t-1)}_{b\to i}(x_i)$$

factor → variable



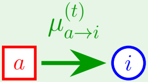
Messages

variable → factor



$$\mu^{(t)}_{i\to a}(x_i) = \sum_{b\in\partial i\setminus a} \mu^{(t-1)}_{b\to i}(x_i)$$

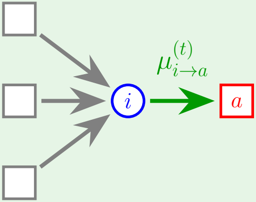
factor → variable



$$\mu^{(t)}_{a\to i}(x_i) =$$

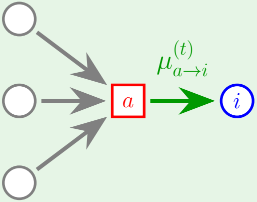
Messages

variable $\to$ factor



$$\mu^{(t)}_{i\to a}(x_i) = \sum_{b\in\partial i\setminus a} \mu^{(t-1)}_{b\to i}(x_i)$$

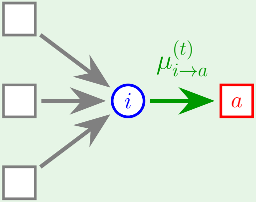
factor $\to$ variable



$$\mu^{(t)}_{a\to i}(x_i) =$$

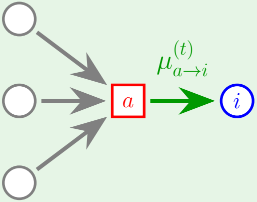
Messages

variable $\rightarrow$ factor



$$\mu_{i\to a}^{(t)}(x_i) = \sum_{b\in\partial i\setminus a} \mu_{b\to i}^{(t-1)}(x_i)$$

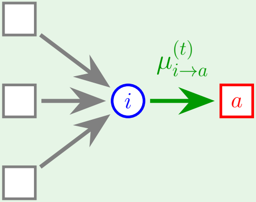
factor $\rightarrow$ variable



$$\mu_{a\to i}^{(t)}(x_i) = \sum_{j\in\partial a\setminus i} \mu_{j\to a}^{(t-1)}(x_j)$$

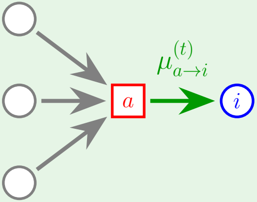
Messages

variable $\rightarrow$ factor



$$\mu_{i\rightarrow a}^{(t)}(x_i) = \sum_{b\in\partial i\setminus a} \mu_{b\rightarrow i}^{(t-1)}(x_i)$$

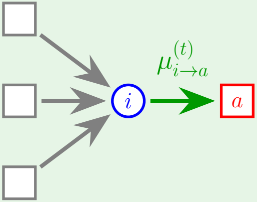
factor $\rightarrow$ variable



$$\mu_{a\rightarrow i}^{(t)}(x_i) = \sum_{j\in\partial a\setminus i} \mu_{j\rightarrow a}^{(t-1)}(x_j) + f_a(x_{\partial a\setminus i}, x_i)$$

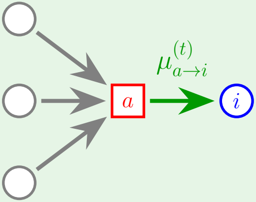
Messages

variable → factor



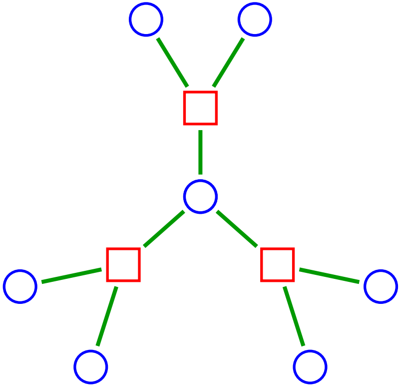
$$\mu_{i\to a}^{(t)}(x_i) = \sum_{b\in\partial i\setminus a} \mu_{b\to i}^{(t-1)}(x_i)$$

factor → variable



$$\mu_{a\to i}^{(t)}(x_i) = \min_{x_{\partial a\setminus i}} \left(\sum_{j\in\partial a\setminus i} \mu_{j\to a}^{(t-1)}(x_j) + f_a(x_{\partial a\setminus i}, x_i) \right)$$

Min-sum: beyond trees?

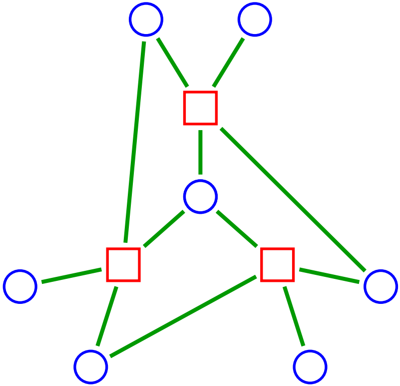


Optimal solution:

$$x^\star := \arg\min_{x} \sum_{a \in \text{factors}} f_a(x_{\partial a})$$

Min-sum: beyond trees?

?

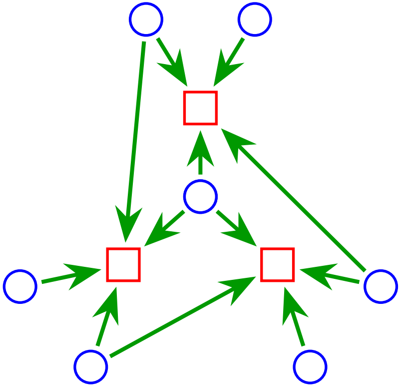


Optimal solution:

$$x^\star := \arg\min_x \sum_{a \in \text{factors}} f_a(x_{\partial a})$$

Min-sum: beyond trees?

time 1: $\{\mu^{(1)}_{i\to a} = 0\}$



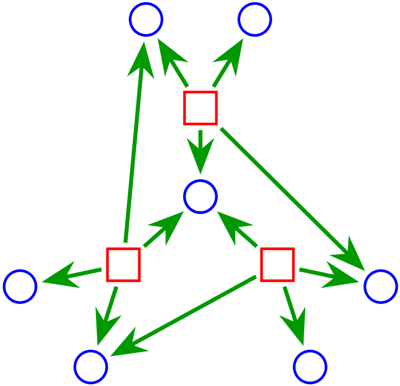
Optimal solution:

$$x^\star := \arg\min_x \sum_{a \in \text{factors}} f_a(x_{\partial a})$$

Min-sum: beyond trees?

time 1: $\{\mu^{(1)}_{i\to a} = 0\}$

time 2: $\{\mu^{(2)}_{a\to i}\}$



Optimal solution:

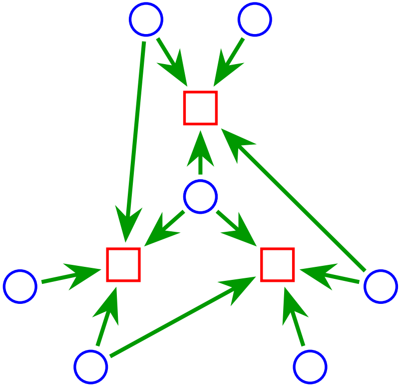
$$x^\star := \arg\min_x \sum_{a \in \text{factors}} f_a(x_{\partial a})$$

Min-sum: beyond trees?

time 1: $\{\mu^{(1)}_{i\to a} = 0\}$

time 2: $\{\mu^{(2)}_{a\to i}\}$

time 3: $\{\mu^{(3)}_{i\to a}\}$



Optimal solution:

$$x^\star := \arg\min_x \sum_{a \in \text{factors}} f_a(x_{\partial a})$$

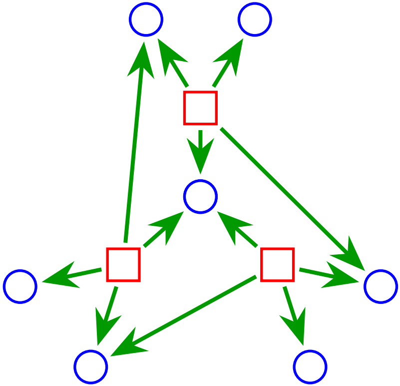
Min-sum: beyond trees?

time 1: $\{\mu^{(1)}_{i\to a} = 0\}$

time 2: $\{\mu^{(2)}_{a\to i}\}$

time 3: $\{\mu^{(3)}_{i\to a}\}$

time 4: $\{\mu^{(4)}_{a\to i}\}$



Optimal solution:

$$x^\star := \arg\min_x \sum_{a\in\text{factors}} f_a(x_{\partial a})$$

Min-sum: beyond trees?

time 1: $\{\mu^{(1)}_{i\to a} = 0\}$

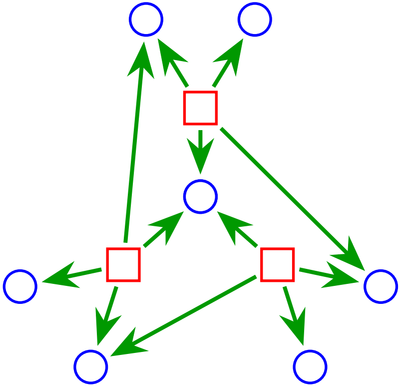
time 2: $\{\mu^{(2)}_{a\to i}\}$

time 3: $\{\mu^{(3)}_{i\to a}\}$

time 4: $\{\mu^{(4)}_{a\to i}\}$

$\vdots$

time t: $\{\mu^{(t)}_{a\to i}\}$



Optimal solution:

$$x^\star := \arg\min_x \sum_{a\in\text{factors}} f_a(x_{\partial a})$$

Min-sum: beyond trees?

time 1: $\{\mu^{(1)}_{i\to a} = 0\}$

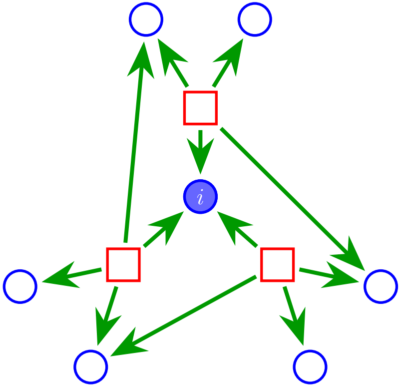
time 2: $\{\mu^{(2)}_{a\to i}\}$

time 3: $\{\mu^{(3)}_{i\to a}\}$

time 4: $\{\mu^{(4)}_{a\to i}\}$

⋮

time t: $\{\mu^{(t)}_{a\to i}\}$



Estimate time t:

$$\hat{x}^{(t)}_i := \arg\min_{x_i} \sum_{a\in\partial i} \mu^{(t)}_{a\to i}(x_i)$$

Optimal solution:

$$x^\star := \arg\min_{x} \sum_{a\in \text{factors}} f_a(x_{\partial a})$$

Min-sum: beyond trees?

time 1: $\{\mu_{i \to a}^{(1)} = 0\}$

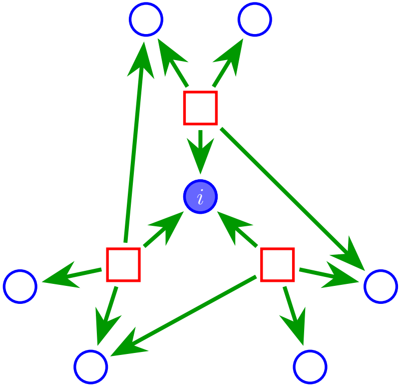
time 2: $\{\mu_{a \to i}^{(2)}\}$

time 3: $\{\mu_{i \to a}^{(3)}\}$

time 4: $\{\mu_{a \to i}^{(4)}\}$

⋮

time t: $\{\mu_{a \to i}^{(t)}\}$



?

Estimate time t:

$$\hat{x}_i^{(t)} := \arg\min_{x_i} \sum_{a \in \partial i} \mu_{a \to i}^{(t)}(x_i)$$

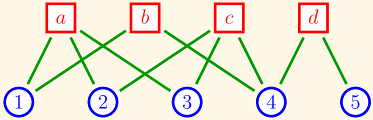
→

Optimal solution:

$$x^\star := \arg\min_{x} \sum_{a \in \text{factors}} f_a(x_{\partial a})$$

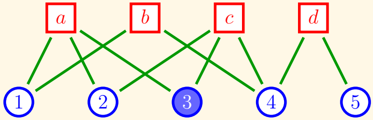
Computation tree

Graph



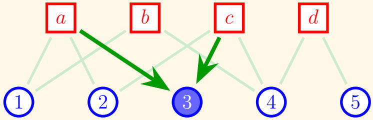
Computation tree

Graph



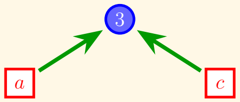
Computation tree

Graph

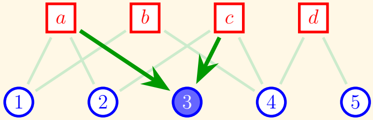


Computation tree

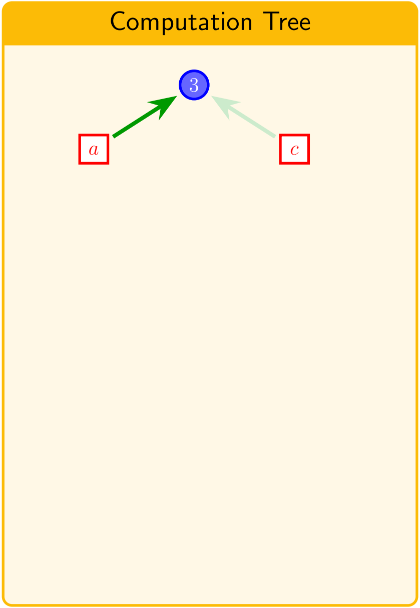
Computation Tree

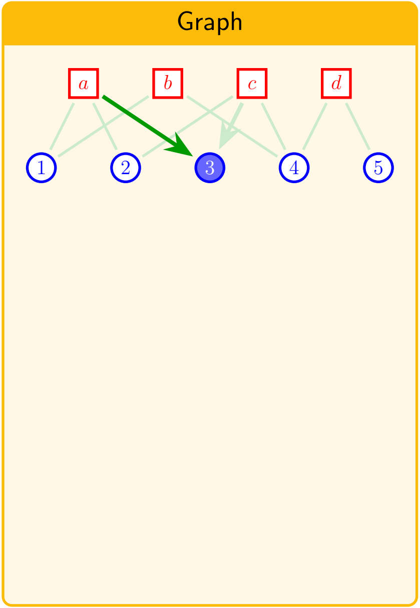


Graph



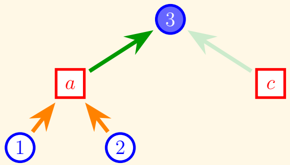
Computation tree



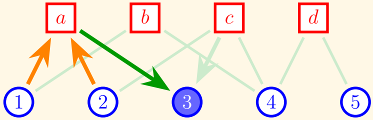


Computation tree

Computation Tree

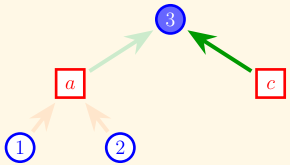


Graph

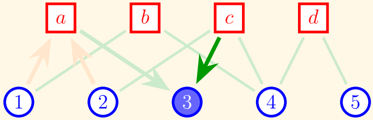


Computation tree

Computation Tree

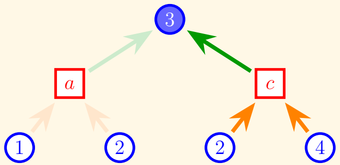


Graph

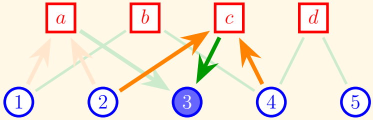


Computation tree

Computation Tree

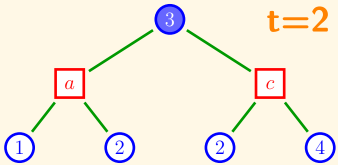


Graph

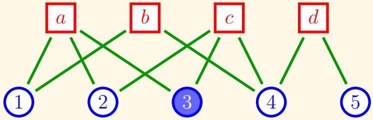


Computation tree

Computation Tree

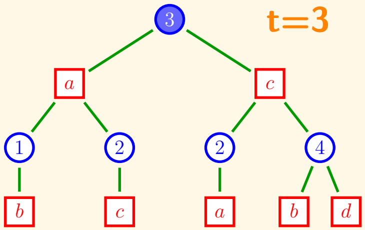


Graph

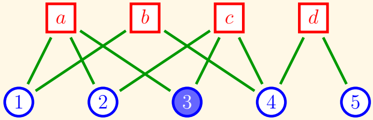


Computation tree

Computation Tree

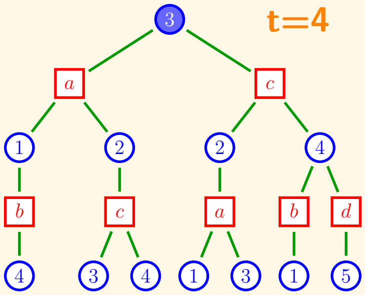


Graph

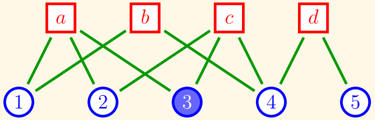


Computation tree

Computation Tree

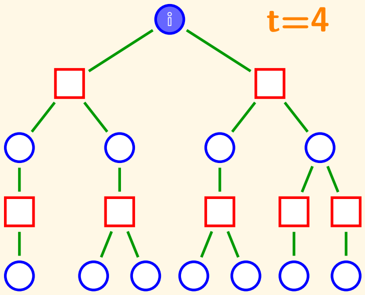


Graph

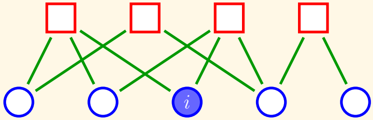


Computation tree

Computation Tree

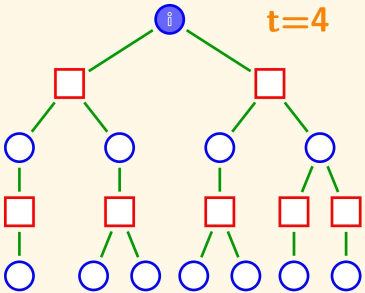


Graph

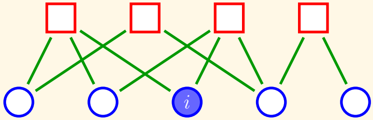


Computation tree

Computation Tree



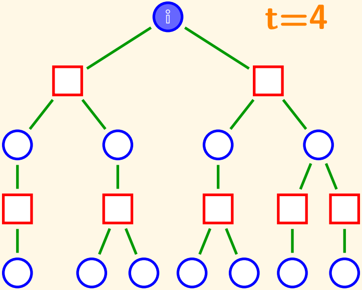
Graph



$$\overbrace{\sum_{\substack{a \in \text{factors} \\ \text{graph}}} f_a(x_{\partial a})}^{f(x)}$$

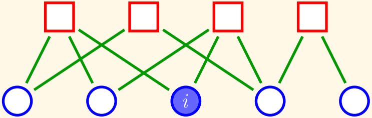
Computation tree

Computation Tree



$$\overbrace{\sum_{\mathbb{a}\in\text{factors tree}} f_{\mathbb{a}}(\mathbb{x}_{\partial\mathbb{a}})}^{\mathbb{f}(\mathbb{x})}$$

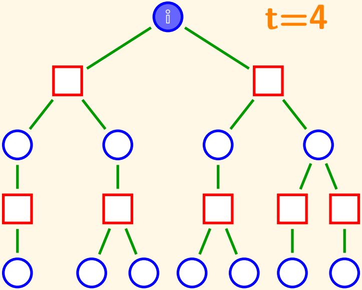
Graph



$$\overbrace{\sum_{a\in\text{factors graph}} f_a(x_{\partial a})}^{f(x)}$$

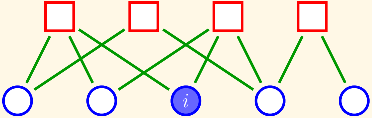
Computation tree

Computation Tree



$$\mathbb{x}^\star := \arg\min_{\mathbb{x}} \overbrace{\sum_{\substack{\mathbb{a}\in\text{factors}\\ \text{tree}}} f_{\mathbb{a}}(\mathbb{x}_{\partial\mathbb{a}})}^{\mathbb{f}(\mathbb{x})}$$

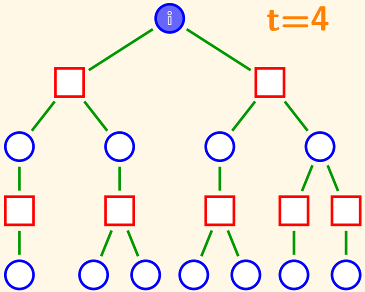
Graph



$$x^\star := \arg\min_{x} \overbrace{\sum_{\substack{a\in\text{factors}\\ \text{graph}}} f_a(x_{\partial a})}^{f(x)}$$

Computation tree

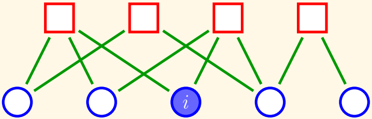
Computation Tree



$$\mathbb{x}^\star := \arg\min_{\mathbb{x}} \overbrace{\sum_{\substack{\mathbb{a}\in\text{factors}\\ \text{tree}}} f_{\mathbb{a}}(\mathbb{x}_{\partial\mathbb{a}})}^{\mathbb{f}(\mathbb{x})}$$

Lemma: $\boxed{\hat{x}_i^{(t)} = \mathbb{x}_{\mathbb{i}}^\star}$

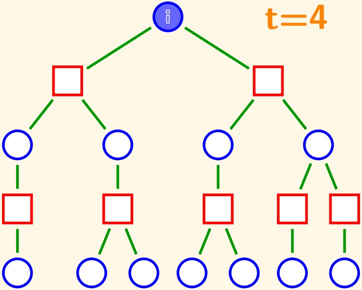
Graph



$$x^\star := \arg\min_{x} \overbrace{\sum_{\substack{a\in\text{factors}\\ \text{graph}}} f_a(x_{\partial a})}^{f(x)}$$

Computation tree

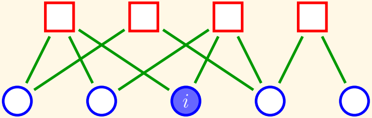
Computation Tree



$$\mathbb{x}^\star := \arg\min_{\mathbb{x}} \overbrace{\sum_{\substack{\mathbb{a}\in\text{factors}\\ \text{tree}}} f_{\mathbb{a}}(\mathbb{x}_{\partial\mathbb{a}})}^{\mathbb{f}(\mathbb{x})}$$

Lemma: $\boxed{\hat{x}_i^{(t)} = \mathbb{x}_{\mathbb{i}}^\star}$

Graph

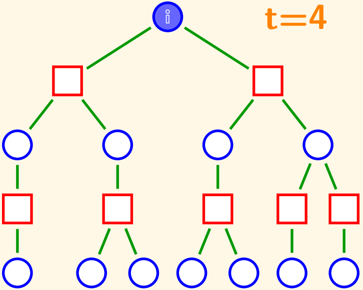


$$x^\star := \arg\min_{x} \overbrace{\sum_{\substack{a\in\text{factors}\\ \text{graph}}} f_a(x_{\partial a})}^{f(x)}$$

$x_i^\star$

Computation tree

Computation Tree

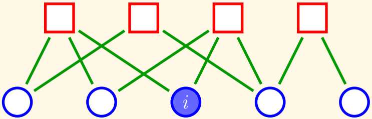


$$\mathbb{x}^\star := \arg\min_{\mathbb{x}} \overbrace{\sum_{\substack{\mathbb{a}\in\text{factors}\\ \text{tree}}} f_{\mathbb{a}}(\mathbb{x}_{\partial\mathbb{a}})}^{\mathbb{f}(\mathbb{x})}$$

?

Lemma: $\boxed{\hat{x}_i^{(t)} = \mathbb{x}_{\mathbb{i}}^\star}$ → $x_i^\star$

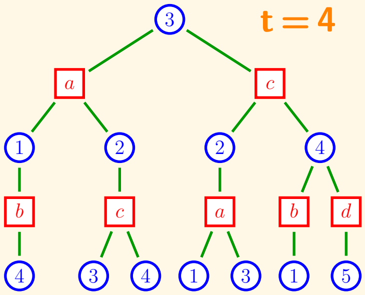
Graph



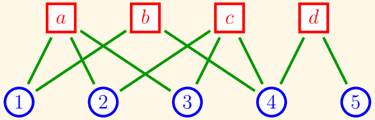
$$x^\star := \arg\min_{x} \overbrace{\sum_{\substack{a\in\text{factors}\\ \text{graph}}} f_a(x_{\partial a})}^{f(x)}$$

Correctness

Computation Tree

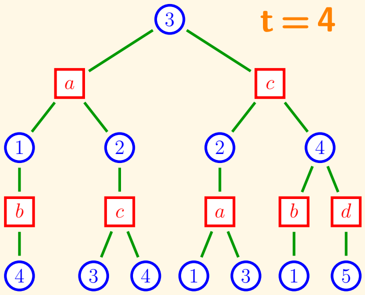


Graph

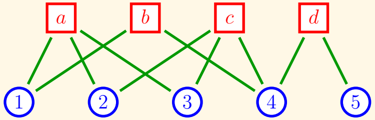


Correctness

Computation Tree



Graph

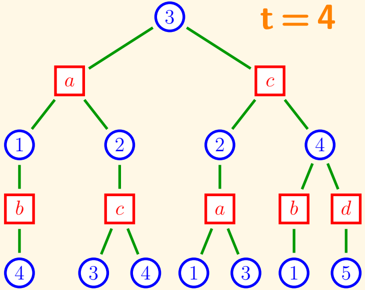


Optimality condition:

$$\overbrace{\sum_{a \in \partial j} \frac{d}{dx_j} f_a(x^\star_{\partial a})}^{\frac{d}{dx_j} f(x)} = 0 \qquad \forall j$$

Correctness

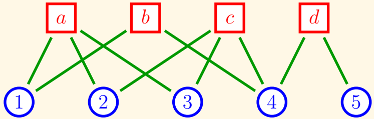
Computation Tree



Optimality condition:

$$\overbrace{\sum_{a \in \partial j} \frac{d}{dx_j} f_a(x^\star_{\partial a})}^{\frac{d}{dx_j} f(x)} = 0 \qquad \forall j$$

Graph

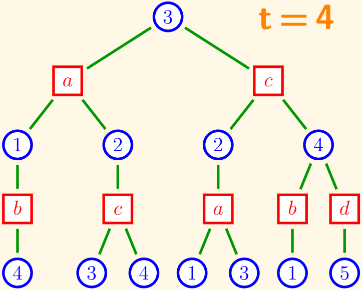


Optimality condition:

$$\overbrace{\sum_{a \in \partial j} \frac{d}{dx_j} f_a(x^\star_{\partial a})}^{\frac{d}{dx_j} f(x)} = 0 \qquad \forall j$$

Correctness

Computation Tree

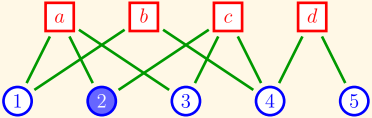


t = 4

Optimality condition:

$$\overbrace{\sum_{a\in\partial j}\frac{d}{dx_j}f_a(x^\star_{\partial a})}^{\frac{d}{dx_j}f(x)} = 0 \qquad \forall j$$

Graph



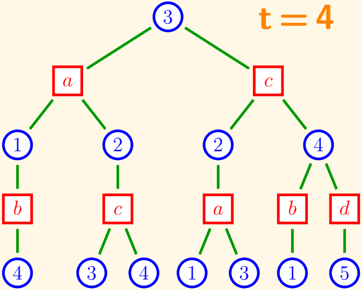
$$\frac{d f_a(x_1^\star, x_2^\star, x_3^\star)}{dx_2} + \frac{d f_c(x_2^\star, x_3^\star, x_4^\star)}{dx_2} = 0 \quad j = 2$$

Optimality condition:

$$\overbrace{\sum_{a\in\partial j}\frac{d}{dx_j}f_a(x^\star_{\partial a})}^{\frac{d}{dx_j}f(x)} = 0 \qquad \forall j$$

Correctness

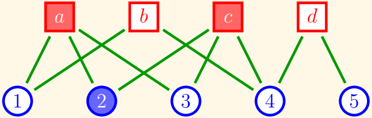
Computation Tree



Optimality condition:

$$\overbrace{\sum_{a \in \partial j} \frac{d}{dx_j} f_a(x^\star_{\partial a})}^{\frac{d}{dx_j} f(x)} = 0 \qquad \forall j$$

Graph



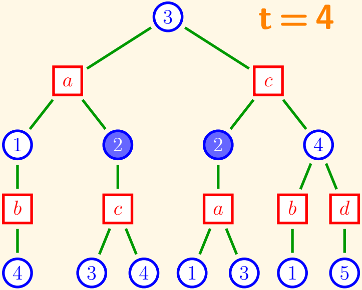
$$\frac{d f_a(x_1^\star, x_2^\star, x_3^\star)}{dx_2} + \frac{d f_c(x_2^\star, x_3^\star, x_4^\star)}{dx_2} = 0 \quad j = 2$$

Optimality condition:

$$\overbrace{\sum_{a \in \partial j} \frac{d}{dx_j} f_a(x^\star_{\partial a})}^{\frac{d}{dx_j} f(x)} = 0 \qquad \forall j$$

Correctness

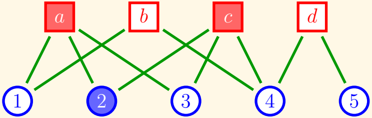
Computation Tree



Optimality condition:

$$\overbrace{\sum_{\mathbb{a} \in \partial \mathbb{j}} \frac{d}{d\mathbb{x}_{\mathbb{j}}} f_{\mathbb{a}}(\mathbb{x}^{\star}_{\partial \mathbb{a}})}^{\frac{d}{d\mathbb{x}_{\mathbb{j}}} \mathbb{f}(\mathbb{x})} = 0 \qquad \forall \mathbb{j}$$

Graph



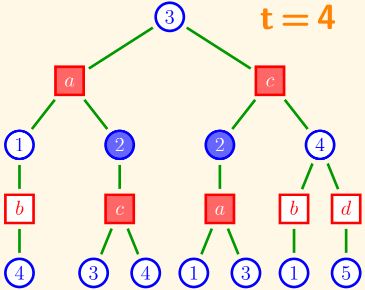
$$\frac{d f_a(x_1^\star, x_2^\star, x_3^\star)}{dx_2} + \frac{d f_c(x_2^\star, x_3^\star, x_4^\star)}{dx_2} = 0 \quad j = 2$$

Optimality condition:

$$\overbrace{\sum_{a \in \partial j} \frac{d}{dx_j} f_a(x^{\star}_{\partial a})}^{\frac{d}{dx_j} f(x)} = 0 \qquad \forall j$$

Correctness

Computation Tree

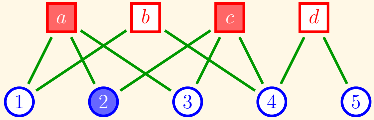


t = 4

Optimality condition:

$$\overbrace{\sum_{a \in \partial j} \frac{d}{dx_j} f_a(x^\star_{\partial a})}^{\frac{d}{dx_j} f(x)} = 0 \qquad \forall j$$

Graph



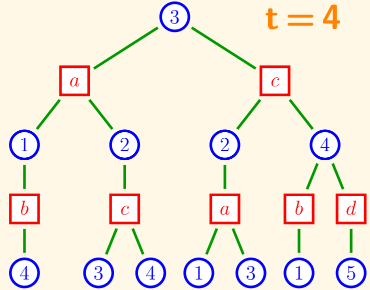
$$\frac{d f_a(x^\star_1, x^\star_2, x^\star_3)}{dx_2} + \frac{d f_c(x^\star_2, x^\star_3, x^\star_4)}{dx_2} = 0 \quad j = 2$$

Optimality condition:

$$\overbrace{\sum_{a \in \partial j} \frac{d}{dx_j} f_a(x^\star_{\partial a})}^{\frac{d}{dx_j} f(x)} = 0 \qquad \forall j$$

Correctness

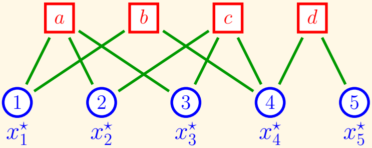
Computation Tree



Optimality condition:

$$\overbrace{\sum_{a \in \partial j} \frac{d}{dx_j} f_a(x^\star_{\partial a})}^{\frac{d}{dx_j} f(x)} = 0 \qquad \forall j$$

Graph

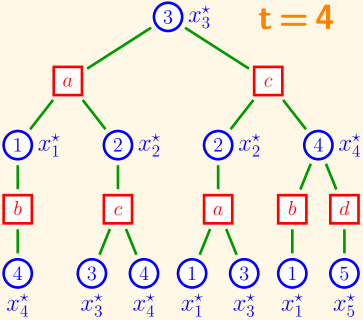


Optimality condition:

$$\overbrace{\sum_{a \in \partial j} \frac{d}{dx_j} f_a(x^\star_{\partial a})}^{\frac{d}{dx_j} f(x)} = 0 \qquad \forall j$$

Correctness

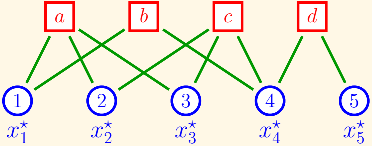
Computation Tree



Optimality condition:

$$\overbrace{\sum_{a\in\partial j} \frac{d}{dx_j} f_a(x^\star_{\partial a})}^{\frac{d}{dx_j} f(x)} = 0 \qquad \forall j$$

Graph

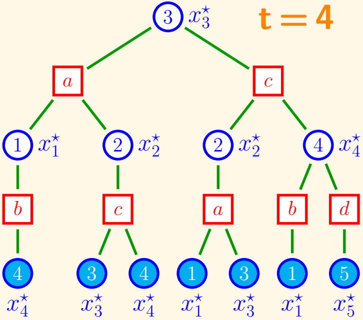


Optimality condition:

$$\overbrace{\sum_{a\in\partial j} \frac{d}{dx_j} f_a(x^\star_{\partial a})}^{\frac{d}{dx_j} f(x)} = 0 \qquad \forall j$$

Correctness

Computation Tree

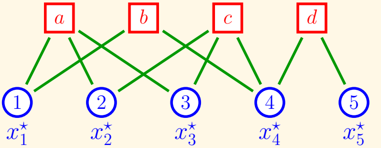


Optimality condition:

$$\overbrace{\sum_{\mathbb{a}\in\partial \mathbb{j}} \frac{d}{d\mathbb{x}_{\mathbb{j}}} f_{\mathbb{a}}(\mathbb{x}^\star_{\partial \mathbb{a}})}^{\frac{d}{d\mathbb{x}_{\mathbb{j}}} \mathbb{f}(\mathbb{x})} = 0 \qquad \forall \mathbb{j}$$

←

Graph

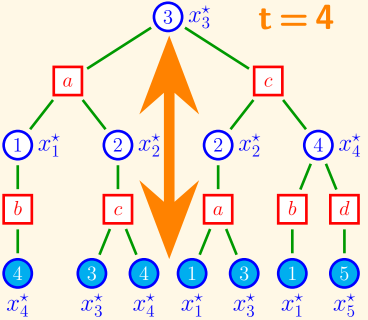


Optimality condition:

$$\overbrace{\sum_{a\in\partial j} \frac{d}{dx_j} f_a(x^\star_{\partial a})}^{\frac{d}{dx_j} f(x)} = 0 \qquad \forall j$$

Correctness

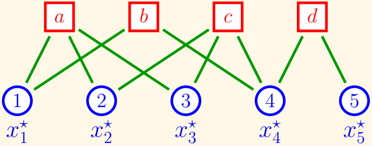
Computation Tree



Optimality condition:

$$\overbrace{\sum_{\mathbb{a} \in \partial \mathbb{j}} \frac{d}{d\mathbb{x}_{\mathbb{j}}} f_{\mathbb{a}}(\mathbb{x}^{\star}_{\partial \mathbb{a}})}^{\frac{d}{d\mathbb{x}_{\mathbb{j}}} \mathbb{f}(\mathbb{x})} = 0 \qquad \forall \mathbb{j}$$

Graph

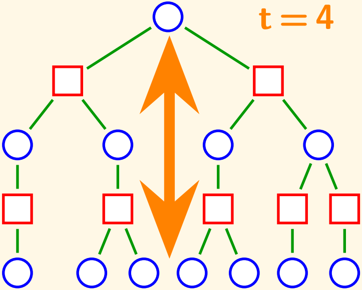


Optimality condition:

$$\overbrace{\sum_{a \in \partial j} \frac{d}{dx_j} f_a(x^{\star}_{\partial a})}^{\frac{d}{dx_j} f(x)} = 0 \qquad \forall j$$

Convergence

Computation Tree

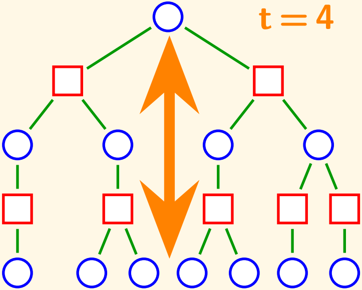


Convergence condition:

$$\sum_{j\in V\setminus i} \omega_j \left|\frac{\partial^2}{\partial x_i \partial x_j} f(x)\right| \le \lambda\, \omega_i \frac{\partial^2}{\partial x_i^2} f(x) \quad \forall i, x$$

Convergence

Computation Tree



Convergence condition:

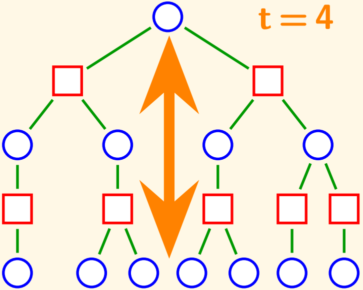
$$\sum_{\mathbb{j}\in\mathbb{V}\setminus\mathbb{i}} \omega_{\mathbb{j}} \left| \frac{\partial^2}{\partial x_{\mathbb{i}} \partial x_{\mathbb{j}}} \mathbb{f}(x) \right| \le \lambda\, \omega_{\mathbb{i}} \frac{\partial^2}{\partial x_{\mathbb{i}}^2} \mathbb{f}(x) \qquad \forall \mathbb{i}, x$$

$$\boxed{\left| \frac{\partial}{\partial p_{j\to a}} x^{\star}(p)_{\mathbb{i}} \right| \le \frac{\max_{\mathbb{i}\in\mathbb{V}} \omega_{\mathbb{i}}}{\min_{\mathbb{i}\in\mathbb{V}} \omega_{\mathbb{i}}} \frac{\lambda^t}{1-\lambda}}$$

(Algorithmic) Locality

Convergence

Computation Tree

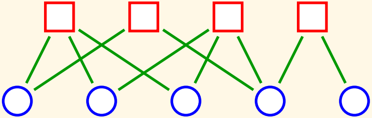


Convergence condition:

$$\sum_{\mathbb{j} \in \mathbb{V} \setminus \mathbb{i}} \omega_{\mathbb{j}} \left| \frac{\partial^2}{\partial \mathbb{x}_{\mathbb{i}} \partial \mathbb{x}_{\mathbb{j}}} \mathbb{f}(\mathbb{x}) \right| \leq \lambda \, \omega_{\mathbb{i}} \frac{\partial^2}{\partial \mathbb{x}_{\mathbb{i}}^2} \mathbb{f}(\mathbb{x}) \quad \forall \mathbb{i}, \mathbb{x}$$

$$\boxed{\left| \frac{\partial}{\partial p_{j \to a}} \mathbb{x}^{\star}(p)_{\mathbb{i}} \right| \leq \frac{\max_{\mathbb{i} \in \mathbb{V}} \omega_{\mathbb{i}}}{\min_{\mathbb{i} \in \mathbb{V}} \omega_{\mathbb{i}}} \, \frac{\lambda^t}{1 - \lambda}}$$

Graph



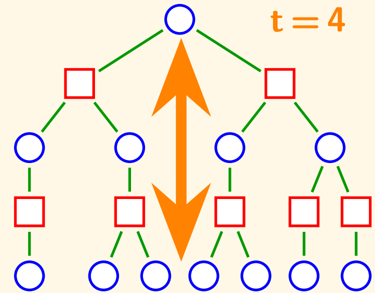
Convergence condition:

$$\sum_{j \in V \setminus i} \omega_j \left| \frac{\partial^2}{\partial x_i \partial x_j} f(x) \right| \leq \lambda \, \omega_i \frac{\partial^2}{\partial x_i^2} f(x) \quad \forall i, x$$

(λ, ω)-scaled diagonal dominance

Convergence

Computation Tree

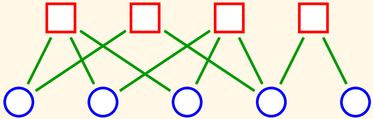


Convergence condition:

$$\sum_{\mathbb{j}\in\mathbb{V}\setminus\mathbb{i}} \omega_{\mathbb{j}} \left|\frac{\partial^2}{\partial \mathbb{x}_{\mathbb{i}}\partial \mathbb{x}_{\mathbb{j}}} \mathbb{f}(\mathbb{x})\right| \le \lambda\,\omega_{\mathbb{i}} \frac{\partial^2}{\partial \mathbb{x}_{\mathbb{i}}^2}\mathbb{f}(\mathbb{x}) \quad \forall \mathbb{i}, \mathbb{x}$$

$$\boxed{\left|\frac{\partial}{\partial p_{j\to a}} \mathbb{x}^\star(p)_{\mathbb{i}}\right| \le \frac{\max_{\mathbb{i}\in\mathbb{V}}\omega_{\mathbb{i}}}{\min_{\mathbb{i}\in\mathbb{V}}\omega_{\mathbb{i}}}\,\frac{\lambda^t}{1-\lambda}}$$

Graph



Convergence condition:

$$\sum_{j\in V\setminus i} \omega_{j} \left|\frac{\partial^2}{\partial x_{i}\partial x_{j}} f(x)\right| \le \lambda\,\omega_{i} \frac{\partial^2}{\partial x_{i}^2}f(x) \quad \forall i, x$$

(λ, ω)-scaled diagonal dominance

Scaled diagonal dominance or walk summability

Computation Tree

$$\sum_{\mathbb{j}\in\mathbb{V}\setminus\mathbb{i}} \omega_{\mathbb{j}} \left|\frac{\partial^2}{\partial \mathbb{x}_{\mathbb{i}}\partial \mathbb{x}_{\mathbb{j}}}\mathbb{f}(\mathbb{x})\right| \le \lambda\,\omega_{\mathbb{i}}\frac{\partial^2}{\partial \mathbb{x}_{\mathbb{i}}^2}\mathbb{f}(\mathbb{x}) \quad \forall \mathbb{i},\mathbb{x}$$

Graph

$$\sum_{j\in V\setminus i} \omega_{j} \left|\frac{\partial^2}{\partial x_{i}\partial x_{j}}f(x)\right| \le \lambda\,\omega_{i}\frac{\partial^2}{\partial x_{i}^2}f(x) \quad \forall i,x$$

Scaled diagonal dominance or walk summability

Computation Tree

$$\sum_{\mathbb{j} \in \mathbb{V} \setminus \mathbb{i}} \omega_{\mathbb{j}} \left| \frac{\partial^2}{\partial \mathbb{x}_{\mathbb{i}} \partial \mathbb{x}_{\mathbb{j}}} \mathbb{f}(\mathbb{x}) \right| \leq \lambda \, \omega_{\mathbb{i}} \frac{\partial^2}{\partial \mathbb{x}_{\mathbb{i}}^2} \mathbb{f}(\mathbb{x}) \quad \forall \mathbb{i}, \mathbb{x}$$

←

Graph

$$\sum_{j \in V \setminus i} \omega_j \left| \frac{\partial^2}{\partial x_i \partial x_j} f(x) \right| \leq \lambda \, \omega_i \frac{\partial^2}{\partial x_i^2} f(x) \quad \forall i, x$$

⇕

$$\rho(|R(x)|) < 1$$

with $R(x) := I - D^{-1/2} \nabla^2 f(x) D^{-1/2}$

Scaled diagonal dominance or walk summability

Computation Tree

$$\sum_{\mathbb{j} \in \mathbb{V} \setminus \mathbb{i}} \omega_{\mathbb{j}} \left| \frac{\partial^2}{\partial \mathbb{x}_{\mathbb{i}} \partial \mathbb{x}_{\mathbb{j}}} \mathbb{f}(\mathbb{x}) \right| \leq \lambda \, \omega_{\mathbb{i}} \frac{\partial^2}{\partial \mathbb{x}_{\mathbb{i}}^2} \mathbb{f}(\mathbb{x}) \quad \forall \mathbb{i}, \mathbb{x}$$

⇕

$$\rho(|\mathbb{R}(\mathbb{x})|) < 1$$

with $\mathbb{R}(\mathbb{x}) := I - \mathbb{D}^{-1/2} \nabla^2 \mathbb{f}(\mathbb{x}) \mathbb{D}^{-1/2}$

←

Graph

$$\sum_{j \in V \setminus i} \omega_j \left| \frac{\partial^2}{\partial x_i \partial x_j} f(x) \right| \leq \lambda \, \omega_i \frac{\partial^2}{\partial x_i^2} f(x) \quad \forall i, x$$

⇕

$$\rho(|R(x)|) < 1$$

with $R(x) := I - D^{-1/2} \nabla^2 f(x) D^{-1/2}$

Scaled diagonal dominance or walk summability

Computation Tree

$$\sum_{\mathbb{j} \in \mathbb{V} \setminus \mathbb{i}} \omega_{\mathbb{j}} \left| \frac{\partial^2}{\partial \mathbb{x}_{\mathbb{i}} \partial \mathbb{x}_{\mathbb{j}}} \mathbb{f}(\mathbb{x}) \right| \leq \lambda \, \omega_{\mathbb{i}} \frac{\partial^2}{\partial \mathbb{x}_{\mathbb{i}}^2} \mathbb{f}(\mathbb{x}) \quad \forall \mathbb{i}, \mathbb{x}$$

⇕

$$\rho(|\mathbb{R}(\mathbb{x})|) < 1$$

with $\mathbb{R}(\mathbb{x}) := I - \mathbb{D}^{-1/2} \nabla^2 \mathbb{f}(\mathbb{x}) \mathbb{D}^{-1/2}$

←

Graph

$$\sum_{j \in V \setminus i} \omega_j \left| \frac{\partial^2}{\partial x_i \partial x_j} f(x) \right| \leq \lambda \, \omega_i \frac{\partial^2}{\partial x_i^2} f(x) \quad \forall i, x$$

⇕

$$\rho(|R(x)|) < 1$$

with $R(x) := I - D^{-1/2} \nabla^2 f(x) D^{-1/2}$

(Weiss, Freeman, 2001) (Malioutov, Johnson, Willsky, 2006) (Moallemi, Van Roy, 2010)

Scaled diagonal dominance or walk summability

Computation Tree

$$\sum_{\mathbb{j}\in\mathbb{V}\setminus\mathbb{i}} \omega_{\mathbb{j}} \left| \frac{\partial^2}{\partial \mathbb{x}_{\mathbb{i}} \partial \mathbb{x}_{\mathbb{j}}} \mathbb{f}(\mathbb{x}) \right| \leq \lambda\, \omega_{\mathbb{i}} \frac{\partial^2}{\partial \mathbb{x}_{\mathbb{i}}^2} \mathbb{f}(\mathbb{x}) \quad \forall \mathbb{i}, \mathbb{x}$$

$\Updownarrow$

$$\rho(|\mathbb{R}(\mathbb{x})|) < 1$$

with $\mathbb{R}(\mathbb{x}) := I - \mathbb{D}^{-1/2} \nabla^2 \mathbb{f}(\mathbb{x}) \mathbb{D}^{-1/2}$

$\leftarrow$

Graph

$$\sum_{j\in V\setminus i} \omega_j \left| \frac{\partial^2}{\partial x_i \partial x_j} f(x) \right| \leq \lambda\, \omega_i \frac{\partial^2}{\partial x_i^2} f(x) \quad \forall i, x$$

$\Updownarrow$

$$\rho(|R(x)|) < 1$$

with $R(x) := I - D^{-1/2} \nabla^2 f(x) D^{-1/2}$

(Weiss, Freeman, 2001) (Malioutov, Johnson, Willsky, 2006) (Moallemi, Van Roy, 2010)

Limitations:

Scaled diagonal dominance or walk summability

Computation Tree

$$\sum_{\mathbb{j}\in\mathbb{V}\setminus\mathbb{i}} \omega_{\mathbb{j}} \left| \frac{\partial^2}{\partial \mathbb{x}_{\mathbb{i}} \partial \mathbb{x}_{\mathbb{j}}} \mathbb{f}(\mathbb{x}) \right| \leq \lambda\, \omega_{\mathbb{i}} \frac{\partial^2}{\partial \mathbb{x}_{\mathbb{i}}^2} \mathbb{f}(\mathbb{x}) \quad \forall \mathbb{i}, \mathbb{x}$$

$\Updownarrow$

$$\rho(|\mathbb{R}(\mathbb{x})|) < 1$$

with $\mathbb{R}(\mathbb{x}) := I - \mathbb{D}^{-1/2} \nabla^2 \mathbb{f}(\mathbb{x}) \mathbb{D}^{-1/2}$

$\leftarrow$

Graph

$$\sum_{j\in V\setminus i} \omega_j \left| \frac{\partial^2}{\partial x_i \partial x_j} f(x) \right| \leq \lambda\, \omega_i \frac{\partial^2}{\partial x_i^2} f(x) \quad \forall i, x$$

$\Updownarrow$

$$\rho(|R(x)|) < 1$$

with $R(x) := I - D^{-1/2} \nabla^2 f(x) D^{-1/2}$

(Weiss, Freeman, 2001) (Malioutov, Johnson, Willsky, 2006) (Moallemi, Van Roy, 2010)

Limitations:

- Inheritance does not capture convergence behavior on the **tree**.

Scaled diagonal dominance or walk summability

Computation Tree

$$\sum_{\mathbb{j}\in\mathbb{V}\setminus\mathbb{i}} \omega_{\mathbb{j}} \left| \frac{\partial^2}{\partial \mathbb{x}_{\mathbb{i}} \partial \mathbb{x}_{\mathbb{j}}} \mathbb{f}(\mathbb{x}) \right| \le \lambda\, \omega_{\mathbb{i}} \frac{\partial^2}{\partial \mathbb{x}_{\mathbb{i}}^2} \mathbb{f}(\mathbb{x}) \quad \forall \mathbb{i}, \mathbb{x}$$

$\Updownarrow$

$$\rho(|\mathbb{R}(\mathbb{x})|) < 1$$

with $\mathbb{R}(\mathbb{x}) := I - \mathbb{D}^{-1/2}\nabla^2 \mathbb{f}(\mathbb{x})\mathbb{D}^{-1/2}$

Graph

$$\sum_{j\in V\setminus i} \omega_j \left| \frac{\partial^2}{\partial x_i \partial x_j} f(x) \right| \le \lambda\, \omega_i \frac{\partial^2}{\partial x_i^2} f(x) \quad \forall i, x$$

$\Updownarrow$

$$\rho(|R(x)|) < 1$$

with $R(x) := I - D^{-1/2}\nabla^2 f(x) D^{-1/2}$

(Weiss, Freeman, 2001) (Malioutov, Johnson, Willsky, 2006) (Moallemi, Van Roy, 2010)

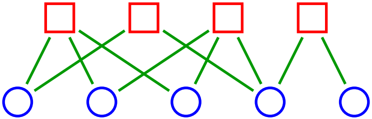
Limitations:

- Inheritance does not capture convergence behavior on the **tree**.
- Condition can **not** be applied to **constrained problems**:

$$\begin{aligned} \text{minimize} \quad & \sum_a f_a(x_{\partial a}) \\ \text{subject to} \quad & Ax = b \end{aligned}$$

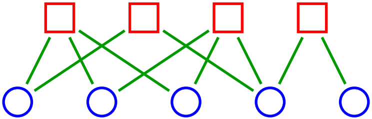
Constrained problems

minimize $\sum_{a} f_a(x_{\partial a})$



Constrained problems

minimize $\sum_{a} f_a(x_{\partial a})$

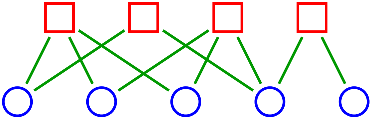


minimize $\sum_{a} f_a(x_{\partial a})$

subject to $h_b(x_{\partial b}) = 0, \quad \forall b$

Constrained problems

minimize $\sum_a f_a(x_{\partial a})$

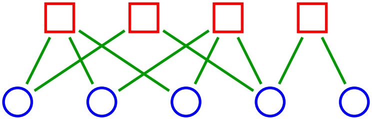


minimize $\sum_a f_a(x_{\partial a}) + \sum_b \chi(h_b(x_{\partial b}))$

$$\chi(z) := \begin{cases} 0 & \text{if } z = 0 \\ \infty & \text{if } z \neq 0 \end{cases}$$

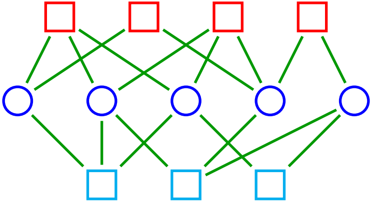
Constrained problems

minimize $\sum_{a} f_a(x_{\partial a})$



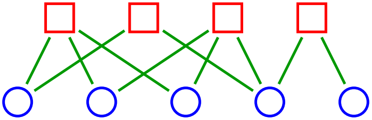
minimize $\sum_{a} f_a(x_{\partial a}) + \sum_{b} \chi(h_b(x_{\partial b}))$

$$\chi(z) := \begin{cases} 0 & \text{if } z = 0 \\ \infty & \text{if } z \neq 0 \end{cases}$$



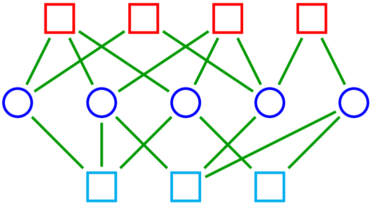
Constrained problems

minimize $\sum_{a} f_a(x_{\partial a})$



minimize $\overbrace{\sum_{a} f_a(x_{\partial a}) + \sum_{b} \chi(h_b(x_{\partial b}))}^{g(x)}$

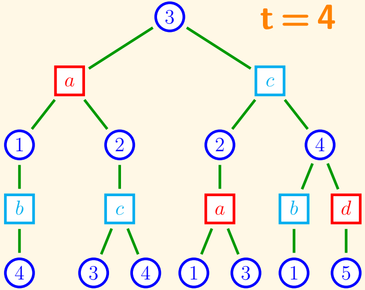
$$\chi(z) := \begin{cases} 0 & \text{if } z = 0 \\ \infty & \text{if } z \neq 0 \end{cases}$$



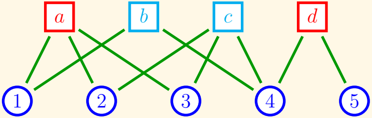
We can run Message Passing. Need different analysis!

Correctness (with constraints!)

Computation Tree

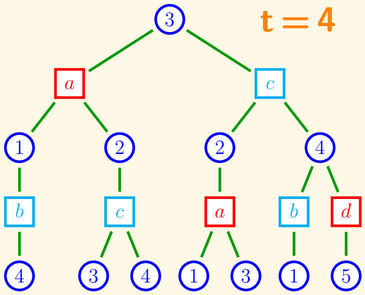


Graph

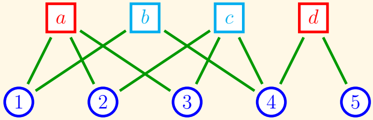


Correctness (with constraints!)

Computation Tree



Graph



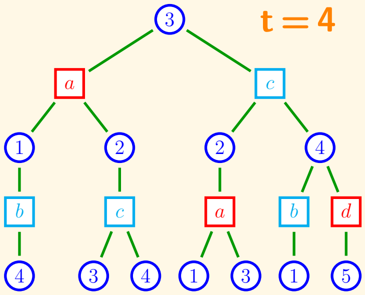
KKT optimality conditions:

$$\sum_{a \in \partial j} \frac{d}{dx_j} f_a(x^\star_{\partial a}) + \sum_{b \in \partial j} \nu^\star_b \frac{d}{dx_j} h_b(x^\star_{\partial b}) = 0 \quad \forall j$$

$$h_b(x^\star_{\partial b}) = 0 \quad \forall b$$

Correctness (with constraints!)

Computation Tree

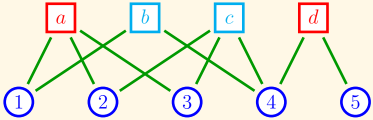


KKT optimality conditions:

$$\sum_{\mathbb{a}\in\partial \mathbb{j}} \frac{d}{dx_{\mathbb{j}}} f_{\mathbb{a}}(x^\star_{\partial \mathbb{a}}) + \sum_{\mathbb{b}\in\partial \mathbb{j}} \nu^\star_{\mathbb{b}} \frac{d}{dx_{\mathbb{j}}} h_{\mathbb{b}}(x^\star_{\partial \mathbb{b}}) = 0 \quad \forall \mathbb{j}$$

$$h_{\mathbb{b}}(x^\star_{\partial \mathbb{b}}) = 0 \quad \forall \mathbb{b}$$

Graph



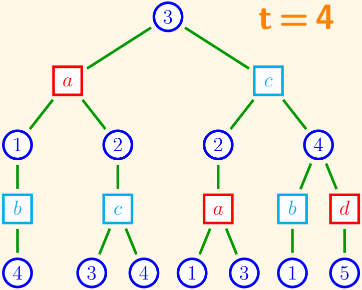
KKT optimality conditions:

$$\sum_{a\in\partial j} \frac{d}{dx_j} f_a(x^\star_{\partial a}) + \sum_{b\in\partial j} \nu^\star_b \frac{d}{dx_j} h_b(x^\star_{\partial b}) = 0 \quad \forall j$$

$$h_b(x^\star_{\partial b}) = 0 \quad \forall b$$

Correctness (with constraints!)

Computation Tree

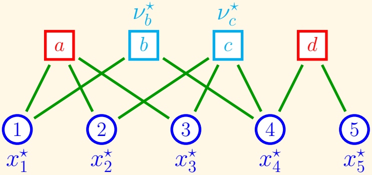


KKT optimality conditions:

$$\sum_{\mathbb{a}\in\partial\mathbb{j}} \frac{d}{dx_{\mathbb{j}}} f_{\mathbb{a}}(x^\star_{\partial\mathbb{a}}) + \sum_{\mathbb{b}\in\partial\mathbb{j}} \nu^\star_{\mathbb{b}} \frac{d}{dx_{\mathbb{j}}} h_{\mathbb{b}}(x^\star_{\partial\mathbb{b}}) = 0 \quad \forall \mathbb{j}$$

$$h_{\mathbb{b}}(x^\star_{\partial\mathbb{b}}) = 0 \quad \forall \mathbb{b}$$

Graph



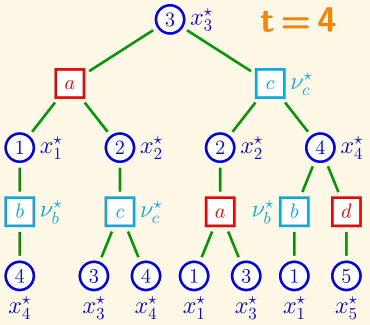
KKT optimality conditions:

$$\sum_{a\in\partial j} \frac{d}{dx_j} f_a(x^\star_{\partial a}) + \sum_{b\in\partial j} \nu^\star_b \frac{d}{dx_j} h_b(x^\star_{\partial b}) = 0 \quad \forall j$$

$$h_b(x^\star_{\partial b}) = 0 \quad \forall b$$

Correctness (with constraints!)

Computation Tree

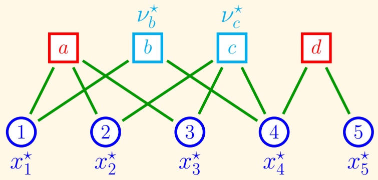


KKT optimality conditions:

$$\sum_{\mathbb{a} \in \partial \mathbb{j}} \frac{d}{dx_{\mathbb{j}}} f_{\mathbb{a}}(x^\star_{\partial \mathbb{a}}) + \sum_{\mathbb{b} \in \partial \mathbb{j}} \nu^\star_{\mathbb{b}} \frac{d}{dx_{\mathbb{j}}} h_{\mathbb{b}}(x^\star_{\partial \mathbb{b}}) = 0 \quad \forall \mathbb{j}$$

$$h_{\mathbb{b}}(x^\star_{\partial \mathbb{b}}) = 0 \quad \forall \mathbb{b}$$

Graph



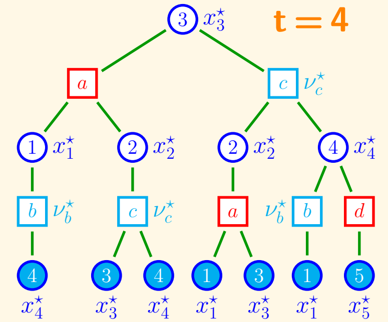
KKT optimality conditions:

$$\sum_{a \in \partial j} \frac{d}{dx_j} f_a(x^\star_{\partial a}) + \sum_{b \in \partial j} \nu^\star_b \frac{d}{dx_j} h_b(x^\star_{\partial b}) = 0 \quad \forall j$$

$$h_b(x^\star_{\partial b}) = 0 \quad \forall b$$

Correctness (with constraints!)

Computation Tree

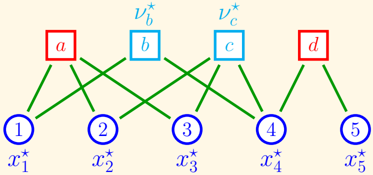


KKT optimality conditions:

$$\sum_{\mathbb{a}\in\partial\mathbb{j}} \frac{d}{dx_{\mathbb{j}}} f_{\mathbb{a}}(x^\star_{\partial\mathbb{a}}) + \sum_{\mathbb{b}\in\partial\mathbb{j}} \nu^\star_{\mathbb{b}} \frac{d}{dx_{\mathbb{j}}} h_{\mathbb{b}}(x^\star_{\partial\mathbb{b}}) = 0 \quad \forall \mathbb{j}$$

$$h_{\mathbb{b}}(x^\star_{\partial\mathbb{b}}) = 0 \quad \forall \mathbb{b}$$

Graph



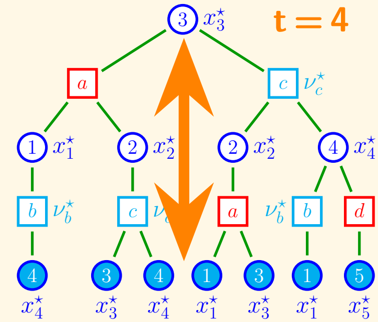
KKT optimality conditions:

$$\sum_{a\in\partial j} \frac{d}{dx_j} f_a(x^\star_{\partial a}) + \sum_{b\in\partial j} \nu^\star_b \frac{d}{dx_j} h_b(x^\star_{\partial b}) = 0 \quad \forall j$$

$$h_b(x^\star_{\partial b}) = 0 \quad \forall b$$

Correctness (with constraints!)

Computation Tree

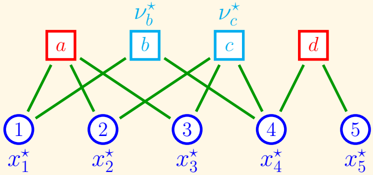


KKT optimality conditions:

$$\sum_{\mathbb{a}\in\partial\mathbb{j}} \frac{d}{dx_{\mathbb{j}}} f_{\mathbb{a}}(x^\star_{\partial\mathbb{a}}) + \sum_{\mathbb{b}\in\partial\mathbb{j}} \nu^\star_{\mathbb{b}} \frac{d}{dx_{\mathbb{j}}} h_{\mathbb{b}}(x^\star_{\partial\mathbb{b}}) = 0 \quad \forall \mathbb{j}$$

$$h_{\mathbb{b}}(x^\star_{\partial\mathbb{b}}) = 0 \quad \forall \mathbb{b}$$

Graph



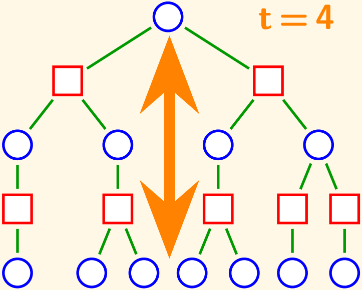
KKT optimality conditions:

$$\sum_{a\in\partial j} \frac{d}{dx_j} f_a(x^\star_{\partial a}) + \sum_{b\in\partial j} \nu^\star_b \frac{d}{dx_j} h_b(x^\star_{\partial b}) = 0 \quad \forall j$$

$$h_b(x^\star_{\partial b}) = 0 \quad \forall b$$

Convergence (with constraints!)

Computation Tree

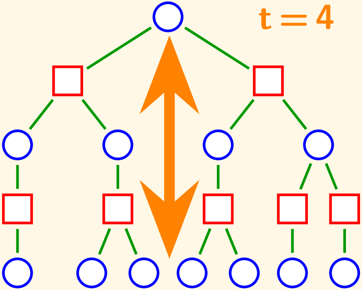


Object of interest:

$$\boxed{\frac{\partial}{\partial p_{j\to a}} \mathbb{x}^{\star}(p)_{\mathbb{i}}}$$

Convergence (with constraints!)

Computation Tree



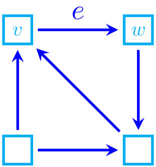
Object of interest:

$$\frac{\partial}{\partial p_{j\to a}} \mathbb{x}^\star(p)_{\mathbb{i}}$$

Beyond scale-diagonal dominance?

Application
Network Flows & Laplacian Solvers

Network Flows



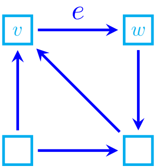
Directed graph $G = (V, E)$

$$\text{minimize} \quad \sum_{e \in E} \overbrace{f_e(x_e)}^{(x_e)^2}$$

$$\text{subject to} \quad Ax = b$$

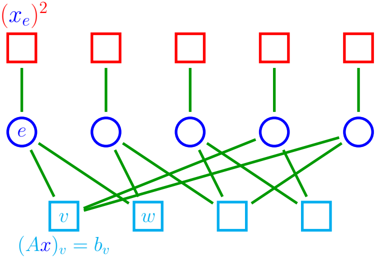
$$A_{ve} := \begin{cases} 1 & \text{if } e \text{ leaves } v, \\ -1 & \text{if } e \text{ enters } v, \\ 0 & \text{otherwise.} \end{cases}$$

Network Flows



Directed graph $G = (V, E)$

$$\begin{array}{ll} \text{minimize} & \sum_{e \in E} \overbrace{f_e(x_e)}^{(x_e)^2} \\ \text{subject to} & Ax = b \end{array}$$



$$A_{ve} := \begin{cases} 1 & \text{if } e \text{ leaves } v, \\ -1 & \text{if } e \text{ enters } v, \\ 0 & \text{otherwise.} \end{cases}$$

Connection to Laplacian solvers

- **Dual problem:** $\min_{\nu \in \mathbb{R}^V} \frac{\nu^T L \nu}{2} - \nu^T b$. **Laplacian:** $\nu^T L \nu = \sum_{\{v,w\} \in E} (\nu_v - \nu_w)^2$
- Solving $L\nu^\star = b$ is key for:
  - PDEs via Finite Element Method
  - Interior Point Methods for Optimization
  - Learning on graphs
  - Faster flow algorithms
  - Graph partitioning
  - Sampling random spanning trees
  - Graph sparsification
- State-of-the-art algorithms (Spielman-Teng, 2004 — Gödel price 2015):
  - **quasi-linear** $O(|E| \log^c(|V|) \log \frac{1}{\varepsilon})$;
  - **centralized**;
  - **involved** (randomized, many graph-theoretic constructions).

Q: What about message-passing?

Message Passing

Algorithm 5: Min-sum, flow problem, quadratic messages, no leaves

Input: Initial messages $\{R^0_{e\to v}\}, \{r^0_{e\to v}\}$, $e \in \vec{E}, v \in \partial e$;

for $s \in \{1, \dots, t\}$ **do**

Compute, for each $e \in \vec{E}$, $v \in \partial e$, with $w = \partial e \setminus v$

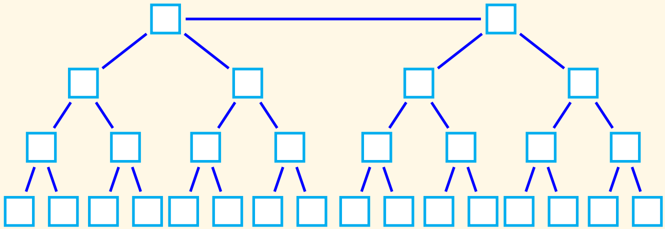
$$R^s_{e\to v} = R_{ee} + \frac{1}{\sum_{f \in \partial w \setminus e} 1/R^{s-1}_{f\to w}}, \qquad r^s_{e\to v} = -A_{we}\frac{\sum_{f \in \partial w \setminus e} A_{wf} r^{s-1}_{f\to w}/R^{s-1}_{f\to w} + b_w}{\sum_{f \in \partial w \setminus e} 1/R^{s-1}_{f\to w}};$$

Output: $\hat{x}^t_e = -\frac{r^t_{e\to v} + r^t_{e\to w}}{R^t_{e\to v} + R^t_{e\to w} - R_{ee}}$, for $e = (v, w) \in \vec{E}$.

Distributed, Simple

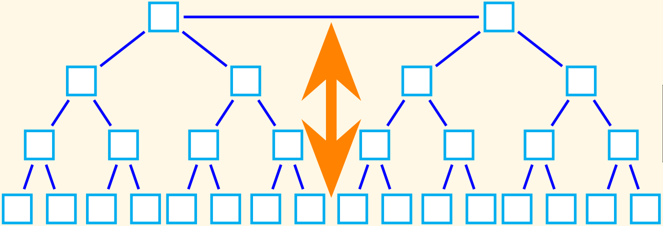
Results

Computation Tree for any d-regular graph



Results

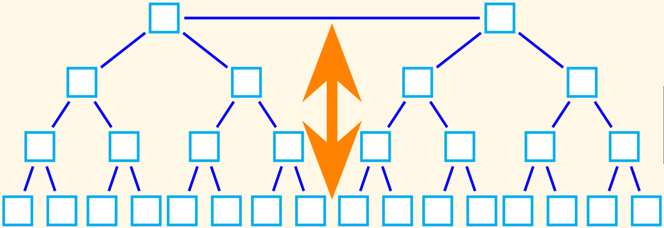
Computation Tree for any d-regular graph



$$\frac{\partial}{\partial p_{e\to v}} \mathbb{x}^{\star}(p)_{\mathbb{e}}$$

Results

Computation Tree for any d-regular graph



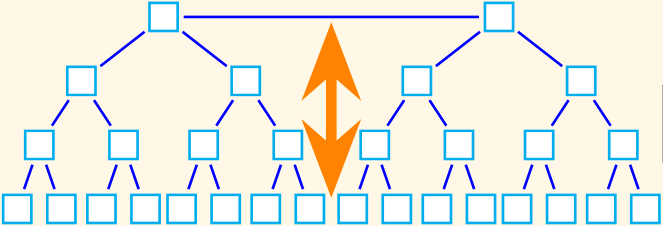
$$\frac{\partial}{\partial p_{e\to v}} \mathbb{x}^\star(p)_{\mathbb{e}}$$

Theorem (Rebeschini, Tatikonda 2016)

$$x^\star_{(v,w)} - \hat{x}^{(t)}_{(v,w)} = \sum_{z \in V} \left(P^{(t)}_{vz} - P^{(t)}_{wz} \right) \nu^\star_z$$

Results

Computation Tree for any d-regular graph



$$\frac{\partial}{\partial p_{e\to v}} \mathbb{x}^{\star}(p)_{\mathbb{e}}$$

Theorem (Rebeschini, Tatikonda 2016)

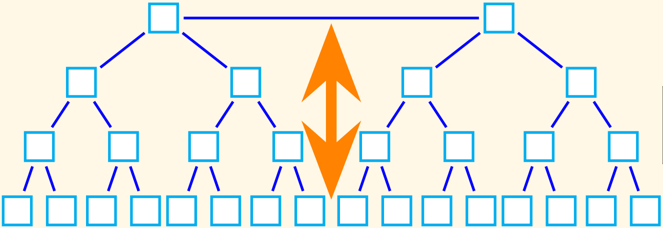
$$x^{\star}_{(v,w)} - \hat{x}^{(t)}_{(v,w)} = \sum_{z\in V} \left(P^{(t)}_{vz} - P^{(t)}_{wz}\right) \nu^{\star}_{z}$$

Corollary

$$\|x^{\star} - \hat{x}^{(t)}\|_{\infty} \le \|\nu^{\star}\|_{\infty} \sum_{z\in V} \left|P^{(t)}_{vz} - P^{(t)}_{wz}\right|$$

Results

Computation Tree for any d-regular graph



$$\frac{\partial}{\partial p_{e\to v}} \mathbb{x}^\star(p)_{\mathbb{e}}$$

Theorem (Rebeschini, Tatikonda 2016)

$$x^\star_{(v,w)} - \hat{x}^{(t)}_{(v,w)} = \sum_{z\in V} \left(P^{(t)}_{vz} - P^{(t)}_{wz}\right) \nu^\star_z$$

Corollary

$$\|x^\star - \hat{x}^{(t)}\|_\infty \le \|\nu^\star\|_\infty \sum_{z\in V} \left|P^{(t)}_{vz} - P^{(t)}_{wz}\right|$$

For some graphs (rings, grids, etc)

$$\max_{\{v,w\}\in E} \sum_{z\in V} \left|P^{(t)}_{vz} - P^{(t)}_{wz}\right| \sim \frac{1}{t^\gamma}$$

Hence, $O\left(|E| \frac{1}{\varepsilon^{1/\gamma}}\right)$

MAIN MESSAGE:

General toolbox for message passing in convex optimization.

- **General framework** for message-passing with constraints.
- **Simple, distributed, fast** algorithm for Laplacian and Network Flows algorithms.

A new approach to Laplacian solvers and flow problems, **arXiv:1611.07138**