Probabilistic Graphical Models & Applications
Pseudo Boolean Optimization II/III

Bjoern Andres and Bernt Schiele

Max Planck Institute for Informatics
For any $n \in \mathbb{N}$, consider $n$-variate **quadratic** forms:

- any **multi-linear polynomial form** $c \in C_{n^2}$ and $f_c : \{0, 1\}^2 \to \mathbb{R}$, i.e., for all $x \in \{0, 1\}^n$,

  $$f_c(x) = c_\emptyset + \sum_{j \in [n]} c_{\{j\}} x_j + \sum_{\{j,k\} \in \binom{[n]}{2}} c_{\{j,k\}} x_j x_k$$

- any **posiform** $c' \in C_{n^2}^+$ and $f'_{c} : \{0, 1\}^2 \to \mathbb{R}$, i.e., for all $x \in \{0, 1\}^n$,

  $$f'_{c}(x) = c'_{\emptyset \emptyset} + \sum_{j \in [n]} \left( c'_{\{j\} \emptyset} x_j + c'_{\emptyset \{j\}} (1 - x_j) \right)$$

  $$\quad + \sum_{\{j,k\} \in \binom{[n]}{2}} \left( c'_{\{j,k\} \emptyset} x_j x_k + c'_{\{j\}\{k\}} x_j (1 - x_k) \right.$$ 

  $$\quad \left. + c'_{\{k\}\{j\}} x_k (1 - x_j) + c'_{\emptyset \{j,k\}} (1 - x_j)(1 - x_k) \right)$$
Lemma 1

For any $n \in \mathbb{N}$, any QPBF $f : \{0, 1\}^n \to \mathbb{R}$, the $c \in C_{n2}$ such that $f_c = f$ and any $c' \in C_{n2}^+(f)$ holds

$$c_\emptyset = c'_\emptyset + \sum_{j=1}^{n} c'_\{j\} + \sum_{\{j,k\}\in\binom{[n]}{2}} c'_\{j,k\}$$

$$\forall j \in [n] \quad c_{\{j\}} = c'_{\{j\} \emptyset} - c'_\emptyset\{j\} + \sum_{k \in [n]-\{j\}} \left( c'_{\{j\}\{k\}} - c'_\emptyset\{j,k\} \right)$$

$$\forall \{j,k\} \in \binom{[n]}{2} \quad c_{\{j,k\}} = c'_{\{j,k\} \emptyset} + c'_\emptyset \{j,k\} - c'_{\{j\}\{k\}} - c'_{\{k\}\{j\}}$$
Proof.

- Expansion of the posiform $c'$ yields a quadratic multi-linear polynomial form.
- Comparison with $c$ yields the conditions stated in the Lemma.
Definition 1 (Complementation)

For any \( n \in \mathbb{N} \) and any QPBF \( f : \{0, 1\}^n \to \mathbb{R} \),

\[
rf := \max_{c' \in C_{n2}^+(f)} c'_{000}
\]  

(1)

is called the **floor dual** of \( f \).
Lemma 2

For any $n \in \mathbb{N}$ and any QPBF $f : \{0, 1\}^n \rightarrow \mathbb{R}$, the floor dual can be computed in polynomial time.
**Proof.** For the multi-linear polynomial form $c \in C_{n2}$ such that $f_c = f$, $r_f$ is the solution of the linear programming problem below (by Lemma 1).

$$\max_{c': J^+_{n2} \to \mathbb{R}} c_\emptyset - \sum_{j=1}^{n} c'_{\emptyset\{j\}} - \sum_{\{j,k\} \in \left(\left[\begin{array}{c} n \end{array}\right]^{2}\right)} c'_{\emptyset\{j,k\}}$$

subject to

\[ \forall j \in [n] \quad c_{\{j\}} = c'_{\{j\}\emptyset} - c'_{\emptyset\{j\}} + \sum_{k \in [n] - \{j\}} (c'_{\{j\}\{k\}} - c'_{\emptyset\{j,k\}}) \]

\[ \forall \{j, k\} \in \left(\left[\begin{array}{c} n \end{array}\right]^{2}\right) \quad c_{\{j,k\}} = c'_{\{j,k\}\emptyset} + c'_{\emptyset\{j,k\}} - c'_{\{j\}\{k\}} - c'_{\{k\}\{j\}} \]

\[ \forall J \in J^+_{n2} - \{ (\emptyset, \emptyset) \} \quad 0 \leq c'_{J} \]
Can the floor dual be computed more efficiently than by an algorithm for general LPs?
Probabilistic Graphical Models & Applications

Excursus: Maximum st-Flow and Minimum st-Cut

Bjoern Andres and Bernt Schiele

Max Planck Institute for Informatics
Outline

- Definitions: Flux, preflow, flow, network, $st$-preflow and $st$-flow
- Maximum $st$-Flow Problem
- Residual network and augmenting paths
- $st$-cut
- Minimum $st$-Cut Problem
- Maximum $st$-Flow/Minimum $st$-Cut Theorem
- Algorithms
Definition 2
A pair \((V, E)\) is called
- **an undirected graph** iff \(E \subseteq \binom{V}{2}\)
- **a directed graph** iff \(E \subseteq V^2\).

The elements of \(V\) are called **nodes**. The elements of \(E\) are called **edges**.
Definition 3

In any directed graph \((V, E)\),

- an edge \(e \in E\) is called a **self-edge** iff
  \[
  \exists v \in V \quad e = vv .
  \]

- a pair \(ee' \in E^2\) of edges is called a **digon** iff
  \[
  \exists v, v' \in V \quad v \neq v' \land e = vv' \land e' = v'v .
  \]
Below, the term **directed graph** shall always mean directed graph **without self-edges** and **without digons**, i.e., a pair \((V, E)\) such that \(E \subseteq V^2\) and

\[
\forall vw \in V^2 \quad vw \notin E \lor wv \notin E.
\]
For any directed graph \((V, E)\), any \(U \subseteq V\) and any \(W \subseteq V\) let

\[ UW := \{ uv \in E \mid u \in U \land w \in W \} \, . \]
Definition 4

For any directed graph \((V, E)\) and any \(n \in \mathbb{N}\), a map \(p \in E^n\) is called a path in \((V, E)\) of length \(n\) iff

\[
\forall j, k \in [n] \quad j = k \lor p_j \neq p_k
\]

\[
\forall j \in [n - 1] \exists v, w, x \in V \quad p_j = vw \land p_{j+1} = wx
\].

A path is called simple iff the nodes along the path are pairwise distinct, except for, possibly, the first and the last node.

Below, path shall always mean simple path.
Definition 5

For any directed graph \((V, E)\) and any \(f \in \mathbb{N}_0^E\), the maps \(\varphi^+, \varphi^-, \varphi : 2^V \to \mathbb{Z}\) such that

\[
\forall U \in 2^V \quad \varphi^+_U = \sum_{uv \in UU^c} f_{uv}
\]

\[
\varphi^-_U = \sum_{vu \in U^c U} f_{vu}
\]

\[
\varphi_U = \varphi^+_U - \varphi^-_U
\]

are called the **outflux**, **influx** and **flux** in \((V, E)\) w.r.t. \(f\).
For any \( u \in V \),

\[
\varphi_u^+ := \varphi_{\{u\}} \\
\varphi_u^- := \varphi_{\{u\}} \\
\varphi_u := \varphi_{\{u\}}
\]

are called the **outflux**, **influx** and **flux** of \( u \) in \((V, E)\) w.r.t. \( f \).
Lemma 3

For any directed graph \((V, E)\), any \(f \in \mathbb{N}_0^E\) and any \(U \subseteq V\)

\[
\varphi_U = \sum_{u \in U} \varphi_u.
\] (5)
Proof.

\[
\varphi_U = \sum_{uv \in UU^c} f_{uv} - \sum_{vu \in U^c U} f_{vu} \\
= \left( \sum_{uv \in UV} f_{uv} - \sum_{u'u \in UU} f_{uu'} \right) - \left( \sum_{vu \in V U} f_{vu} - \sum_{u'u \in UU} f_{uu'} \right) \\
= \sum_{uv \in UV} f_{uv} - \sum_{vu \in V U} f_{vu} \\
= \sum_{u \in U} \left( \sum_{vw \in \{u\}\{u\}^c} f_{vw} - \sum_{vw \in \{u\}^c\{u\}} f_{vw} \right) \\
= \sum_{u \in U} \varphi_u .
\]

□
Definition 6

A 5-tuple \( N = (V, E, s, t, c) \) is called a network iff \((V, E)\) is a directed graph and \( s \in V \) and \( t \in V \) and \( s \neq t \) and \( c \in \mathbb{N}^E \).

The nodes \( s \) and \( t \) are called the source and the sink of \( N \), respectively.

For any edge \( e \in E \), \( c_e \) is called the capacity of \( e \) in \( N \).
Definition 6

A 5-tuple $N = (V, E, s, t, c)$ is called a network iff $(V, E)$ is a directed graph and $s \in V$ and $t \in V$ and $s \neq t$ and $c \in \mathbb{N}^E$.

The nodes $s$ and $t$ are called the **source** and the **sink** of $N$, respectively.

For any edge $e \in E$, $c_e$ is called the **capacity** of $e$ in $N$. 
Definition 6

A 5-tuple \( N = (V, E, s, t, c) \) is called a network iff \( (V, E) \) is a directed graph and \( s \in V \) and \( t \in V \) and \( s \neq t \) and \( c \in N^E \).

The nodes \( s \) and \( t \) are called the source and the sink of \( N \), respectively.

For any edge \( e \in E \), \( c_e \) is called the capacity of \( e \) in \( N \).
Definition 7

A map $f \in \mathbb{N}_0^E$ is called an **st-preflow** in a network $N = (V, E, s, t, c)$ iff

$$\forall e \in E \quad 0 \leq f_e \leq c_e$$

$$\forall v \in V - \{s\} \quad \varphi_v \leq 0$$ \hspace{1cm} (6, 7)

An st-preflow $f$ in $N$ is called an **st-flow** in $N$ iff, in addition,

$$\forall v \in V - \{s, t\} \quad \varphi_v = 0$$ \hspace{1cm} (8)
Definition 7

A map $f \in \mathbb{N}_0^E$ is called an $st$-preflow in a network $N = (V, E, s, t, c)$ iff

$$\forall e \in E \quad 0 \leq f_e \leq c_e$$

$$\forall v \in V - \{s\} \quad \varphi_v \leq 0 .$$

(6) 

(7)

An $st$-preflow $f$ in $N$ is called an $st$-flow in $N$ iff, in addition,

$$\forall v \in V - \{s, t\} \quad \varphi_v = 0 .$$

(8)
Definition 8

The instance of the **Maximum $st$-Flow Problem** w.r.t. a network $N = (V, E, s, t, c)$ is to

$$\max_{f \in \mathbb{N}^E_0} \sum_{sv \in E} f_{sv} - \sum_{vs \in E} f_{vs}$$ (9)

subject to

$$\forall e \in E \quad 0 \leq f_e \leq c_e$$ (10)

$$\forall v \in V - \{s, t\} \quad \sum_{vw \in E} f_{vw} = \sum_{uv \in E} f_{uv}.$$ (11)

Note:

$$\sum_{sv \in E} f_{sv} - \sum_{vs \in E} f_{vs} = \varphi_s$$
Definition 9

For any network \( N = (V, E, s, t, c) \) and any \( st \)-preflow \( f \) in \( N \), the **residual network** of \( N \) w.r.t. \( f \) is the network \( N' = (V, E', s, t, c') \) such that

\[
E' = E^+ \cup E^-
\]

\[
E^+ = \{ vw \in E \mid c_{vw} - f_{vw} > 0 \}
\]

\[
E^- = \{ vw \in V^2 \mid wv \in E \land f_{wv} > 0 \}
\]

and

\[
\forall vw \in E' \quad c'_{vw} = \begin{cases} c_{vw} - f_{vw} & \text{if } vw \in E^+ \\ f_{wv} & \text{if } vw \in E^- \end{cases}.
\]

For any \( e \in E' \), \( c'_e \) is called the **residual capacity** of \( e \) w.r.t. \( f \).

Any path in \((V, E')\) from \( s \) to \( t \) (if such a path exists) is called an **augmenting path** of \( f \).
Definition 9

For any network \( N = (V, E, s, t, c) \) and any \( st \)-preflow \( f \) in \( N \), the **residual network** of \( N \) w.r.t. \( f \) is the network \( N' = (V, E', s, t, c') \) such that

\[
E' = E^+ \cup E^-
\]

\[
E^+ = \{vw \in E \mid c_{vw} - f_{vw} > 0\}
\]

\[
E^- = \{vw \in V^2 \mid wv \in E \land f_{wv} > 0\}
\]

and

\[
\forall vw \in E' \quad c'_{vw} = \begin{cases} 
  c_{vw} - f_{vw} & \text{if } vw \in E^+ \\
  f_{wv} & \text{if } vw \in E^-
\end{cases}
\]

(12)

For any \( e \in E' \), \( c'_e \) is called the **residual capacity** of \( e \) w.r.t. \( f \).

Any path in \((V, E')\) from \( s \) to \( t \) (if such a path exists) is called an **augmenting path** of \( f \).
Definition 9

For any network \( N = (V, E, s, t, c) \) and any \( st \)-preflow \( f \) in \( N \), the \textbf{residual network} of \( N \) w.r.t. \( f \) is the network \( N' = (V, E', s, t, c') \) such that

\[
E' = E^+ \cup E^-
\]

\[
E^+ = \{vw \in E \mid c_{vw} - f_{vw} > 0\}
\]

\[
E^- = \{vw \in V^2 \mid wv \in E \land f_{wv} > 0\}
\]

and

\[
\forall vw \in E' \quad c'_{vw} = \begin{cases} 
  c_{vw} - f_{vw} & \text{if } vw \in E^+ \\
  f_{wv} & \text{if } vw \in E^-
\end{cases}
\]  \quad (12)

For any \( e \in E' \), \( c'_e \) is called the \textbf{residual capacity} of \( e \) w.r.t. \( f \).

Any path in \((V, E')\) from \( s \) to \( t \) (if such a path exists) is called an \textbf{augmenting path} of \( f \).
Lemma 5

Let $N = (V, E, s, t, c)$ be a network and $f$ an $st$-preflow in $N$. Assume that an $n \in \mathbb{N}$ and an augmenting path $p = (v_1 w_1, \ldots, v_n w_n)$ of $f$ exist. Let

$$\delta := \min_{vw \in p([n])} c'_{vw}.$$  \hfill (13)

Then, $f' \in N_0^E$ such that

$$\forall vw \in E' : \quad f'_{vw} = \begin{cases} f_{vw} + \delta & \text{if } vw \in p([n]) \land vw \in E \\ f_{vw} - \delta & \text{if } vw \in p([n]) \land wv \in E \\ f_{vw} & \text{otherwise} \end{cases}$$  \hfill (14)

is an $st$-preflow in $N$ w.r.t. which

$$\varphi'_s = \varphi_s + \delta.$$  \hfill (15)

Moreover, if $f$ is an $st$-flow in $N$, so is $f'$. 
Lemma 5

Let $N = (V, E, s, t, c)$ be a network and $f$ an $st$-preflow in $N$. Assume that an $n \in \mathbb{N}$ and an augmenting path $p = (v_1w_1, \ldots, v_nw_n)$ of $f$ exist. Let

$$\delta := \min_{vw \in p([n])} c'_{vw}.$$  \hspace{1cm} (13)

Then, $f' \in \mathbb{N}_0^E$ such that

$$\forall vw \in E' : \quad f'_{vw} = \begin{cases} 
  f_{vw} + \delta & \text{if } vw \in p([n]) \land vw \in E \\
  f_{vw} - \delta & \text{if } vw \in p([n]) \land wv \in E \\
  f_{vw} & \text{otherwise}
\end{cases}$$  \hspace{1cm} (14)

is an $st$-preflow in $N$ w.r.t. which

$$\varphi'_s = \varphi_s + \delta.$$  \hspace{1cm} (15)

Moreover, if $f$ is an $st$-flow in $N$, so is $f'$. 
Lemma 5

Let $N = (V, E, s, t, c)$ be a network and $f$ an $st$-preflow in $N$. Assume that an $n \in \mathbb{N}$ and an augmenting path $p = (v_1w_1, \ldots, v_nw_n)$ of $f$ exist. Let

$$\delta := \min_{vw \in p([n])} c'_{vw}.$$ (13)

Then, $f' \in \mathbb{N}_0^E$ such that

$$\forall vw \in E' : \quad f'_{vw} = \begin{cases} f_{vw} + \delta & \text{if } vw \in p([n]) \land vw \in E \\ f_{vw} - \delta & \text{if } vw \in p([n]) \land wv \in E \\ f_{vw} & \text{otherwise} \end{cases}$$ (14)

is an $st$-preflow in $N$ w.r.t. which

$$\varphi'_s = \varphi_s + \delta.$$ (15)

Moreover, if $f$ is an $st$-flow in $N$, so is $f'$. 
Lemma 5

Let $N = (V, E, s, t, c)$ be a network and $f$ an $st$-preflow in $N$. Assume that an $n \in \mathbb{N}$ and an augmenting path $p = (v_1w_1, \ldots, v_nw_n)$ of $f$ exist. Let

$$\delta := \min_{vw \in p([n])} c'_{vw}.$$  \hspace{1cm} (13)

Then, $f' \in N^E_0$ such that

$$\forall vw \in E' : \hspace{1cm} f'_{vw} = \begin{cases} f_{vw} + \delta & \text{if } vw \in p([n]) \land vw \in E \\ f_{vw} - \delta & \text{if } vw \in p([n]) \land wv \in E \\ f_{vw} & \text{otherwise} \end{cases}$$ \hspace{1cm} (14)

is an $st$-preflow in $N$ w.r.t. which

$$\varphi'_s = \varphi_s + \delta.$$  \hspace{1cm} (15)

Moreover, if $f$ is an $st$-flow in $N$, so is $f'$. 
Definition 10

Let $(V, E)$ be a directed graph. Let $s \in V$ and $t \in V$ and $s \neq t$.

- $X \subseteq V$ is called an \textbf{st-cutset} of $(V, E)$ iff $s \in X$ and $t \notin X$.
- $Y \subseteq E$ is called an \textbf{st-cut} of $(V, E)$ iff there exists an \textbf{st-cutset} $X$ such that $Y = \{vw \in E | v \in X \land w \notin X\}$. 

\[ 
\begin{array}{cccc}
| & s & | & s \\
\hline
v_1 & \rightarrow v_2 & v_1 & \rightarrow v_2 \ \\
\downarrow & & \downarrow & \ \\
t & & t & \\
\end{array}
\]
Definition 11

The instance of the **Minimum st-Cut Problem w.r.t. a network** \(N = (V, E, s, t, c)\) is to

\[
\min_{x \in \{0,1\}^V} \sum_{vw \in E} x_v (1 - x_w) c_{vw}
\]

subject to

\[
x_s = 1
\]
\[
x_t = 0
\]

Note: With \(X := \{v \in V | x_v = 1\}\), we have

\[
\sum_{vw \in E} x_v (1 - x_w) c_{vw} = \sum_{vw \in X \times X^c} c_{vw}
\]
Lemma 7

For every network $N = (V, E, s, t, c)$, every $st$-flow $f$ in $N$, and every $st$-cutset $X \subseteq V$, 

$$\phi_s \leq \sum_{vw \in XX^c} c_{vw} .$$  (19)
Theorem 1

For any network $N = (V, E, s, t, c)$, any $s, t \in V$ such that $s \neq t$, and any $st$-flow $f$ in $N$, the following three conditions are equivalent

1. There exists an $st$-cut whose capacity is equal to $\varphi_s$.
2. The $st$-flow $f$ is optimal, i.e., a solution of (9)–(11).
3. No augmenting path of $f$ exists.
Proof.
(1) implies (2) by virtue of Lemma 7.
(2) implies (3) by virtue of Lemma 5.
We prove that (3) implies (1):

▶ Let \( f \) be an \( st \)-flow such that no augmenting path exists.
▶ Let \( S \) be the set of all nodes \( v \in V \) such that there exists a path in the residual network w.r.t. \( f \) from \( s \) to \( v \). Let \( S \) also include \( s \) itself.
▶ Then, \( t \notin S \) (otherwise, the path from \( s \) to \( t \) in the residual network would be an augmenting path).
▶ Moreover, . . .
Moreover,

\[ \varphi_s = \sum_{v \in S} \varphi_v \]

by (8) and \( t \notin S \)

\[ = \varphi_S \]

by Lemma 3

\[ = \sum_{vw \in SS^c} f_{vw} - \sum_{vw \in S^cS} f_{vw} \]

by definition of \( \varphi_S \)

\[ = \sum_{vw \in SS^c} c_{vw} \]

by the arguments below.

- For any \( vw \in SS^c \), we have \( f_{vw} = c_{vw} \) (otherwise, the contradiction \( w \in S \) follows by construction of \( S \) and by definition of the residual network).

- For any \( vw \in S^cS \), we have \( f_{vw} = 0 \) (otherwise, the contradiction \( v \in S \) follows by construction of \( S \) and by definition of the residual network).
Algorithm 1 (Ford and Fulkerson, 1956)

**Input:** Network $N = (V, E, s, t, c)$

**Output:** $f : E \rightarrow \mathbb{N}_0$

for all $vw \in E$

$f_{vw} := 0$

**while** $\exists n \in \mathbb{N}$ $\exists$ augmenting path $p = (v_1w_1, \ldots, v_nw_n)$ of $f$

$\delta := \min_{vw \in p([n])} c'_{vw}$

for all $vw \in E$

$f_{vw} := \begin{cases} 
  f_{vw} + \delta & \text{if } vw \in P \land vw \in E \\
  f_{vw} - \delta & \text{if } vw \in P \land wv \in E \\
  f_{vw} & \text{otherwise}
\end{cases}$
Theorem 2

Algorithm 1 terminates. The output $f$ is a maximum $st$-flow in $N$. 
Proof. Termination.

- For every augmenting path processed, $\phi_s$ increases by at least 1.
- Moreover,

$$\phi_s \leq \sum_{vw \in \{s\}\{s\}^c} c_{vw} \quad \text{(by Lemma 7)}$$

- Therefore, only finitely many augmenting paths are processed.
- Thus, the algorithm terminates.

Optimality:

- Throughout the execution, $f$ is an $st$-flow in $N$.
- When the algorithm terminates, no augmenting path exists.
- Thus, $f$ is a maximum $st$-flow in $N$ (by Theorem 1).

Note: An implementation with runtime complexity $O(|E|\phi_s)$ exists.
Ford and Fulkerson (1956) – $O(|E| \hat{\phi}_s)$
- Dinic (1970) – $O(|V|^2 |E|)$
- Edmonds and Karp (1972) – $O(|V||E|^2)$
- Sleator and Tarjan (1983, 1985) – $O(|V||E| \log |V|)$
- ...
Definition 12

For any network \( N = (V, E, s, t, c) \) and any \( st\)-preflow \( f \) in \( N \), \( h : V \rightarrow \mathbb{N}_0 \) is said to be **compatible** with \( f \) iff

1. \( h_s = |V| \) and \( h_t = 0 \)

2. For every edge \( vw \in E' \) in the residual network \((V, E', s, t, c')\), \( h(v) \leq h(w) + 1 \).
Algorithm 2 (Goldberg and Tarjan, 1986, 1988)

**Input**: Network $N = (V, E, s, t, c)$  
**Output**: $f : E \rightarrow \mathbb{N}_0$

for all $vw \in E$  
\[ f_{vw} := \begin{cases}  
  c_{vw} & \text{if } v = s \\
  0 & \text{otherwise} 
\end{cases} \]

for all $v \in V$  
\[ h_v = \begin{cases}  
  |V| & \text{if } v = s \\
  0 & \text{otherwise} 
\end{cases} \]

while $\exists v \in V : v \neq t \land \varphi_v < 0$

  if $\exists w \in V : vw \in E' \land h_w < h_v$
  
    if $vw \in E$
    
      \[ f_{vw} := f_{vw} + \min\{-\varphi_v, c_{vw} - f_{vw}\} \]
    
    else
    
      \[ f_{vw} := f_{vw} - \min\{-\varphi_v, f_{vw}\} \]
  
    else
    
      $h_v := h_v + 1$. 


Goldberg and Tarjan (1986, 1988) – $O(|V|^2 E)$

- FIFO node selection – $O(|V|^3)$
- highest node selection – $O(|V|^2 \sqrt{|E|})$
- Sleator and Tarjan (1983, 1985) – $O(|V||E| \log(|V|^2 / |E|))$

- ...