Lecture 5 (27.03.2013) Scribe: Tomasz Kociumaka

Lecturer: Marcin Mucha

# Iterative & Dependent Rounding

## 1 Toy problem: minimum spanning tree

The following LP is a relaxation of the ILP for the MST problem.

$$\begin{aligned} \text{minimize} \sum_{e\in E} w_e x_e& \\ \forall_{\emptyset\neq S\subseteq V},\ x(E(S)) \le |S|-1& \\ x(E) = |V|-1& \\ \forall_{e\in E}\ x_e \ge 0& \end{aligned} \tag{1}$$

We shall prove that the vertices of the corresponding polytope are integral, i.e. they are spanning trees of the graph. This way we get some insight on the combinatorial structure behind this LP, which is going to be useful later. A general approach is going to be similar to the one for the maximum-weight matching problem investigated in lecture 4. Here this is going to be more complicated though, since the number of constraints fails to be polynomial. There is a separation oracle for the LP, so we know that we can solve it in poly-time, but the structure of tight constraints is going to be much more involved. The techniques we are to develop are very general (an exposition can be found under the keyword 'submodular (supermodular) functions').

**Fact 1** (see lecture 4)**.** *Let $x^\star$ be an extremal solution of* (1)*. Assume that $x^\star(e) > 0$ for all $e \in E$. Then there exists a family $\mathcal{S} \subseteq 2^V$ of $|E|$ subsets of $V$ such that $\forall_{S\in\mathcal{S}}\ x^\star(E(S)) = |S| - 1$ and the these equations are linearly independent.*

**Definition 2.** Sets $A, B$ are *crossing* if $A \cap B$, $A \setminus B$ and $B \setminus A$ are all non-empty.

**Definition 3.** A family $\mathcal{F} \subseteq 2^V$ is *laminar* if no two $F, F' \in \mathcal{F}$ are crossing, see Fig. 1

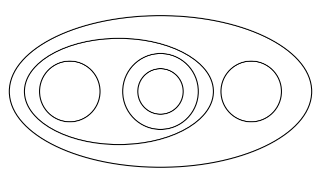

Figure 1: A sample laminar family; note that the sets are either disjoint or nested.

**Theorem 4.** *There exists a laminar family $\mathcal{S}$ satisfying the conditions of Fact 1*

*Proof.* Before we proceed, let us introduce some notation. For $F \subseteq E$ let us denote the *characteristic vector* of $F$ by $\chi(F)$, i.e.

$$\chi(F)(e) = \begin{cases} 1 & \text{if } e \in F, \\ 0 & \text{otherwise.} \end{cases}$$

Note that linear dependence of tight constraints is actually linear dependence of the corresponding vectors $\chi(E(S))$ in $\mathbb{R}^E$.

For a function $f : E \to \mathbb{R}_{\geq 0}$ let us define a function $i_f : 2^V \to \mathbb{R}_{\geq 0}$ by

$$i_f(X) = \sum_{e \in E(X)} f(e).$$

Moreover for $X, Y \subseteq V$ let us define $d_f(X, Y)$ as

$$d_f(X, Y) = \sum_{e \in E(X \setminus Y, Y \setminus X)} f(e).$$

**Lemma 5.** *The following equality on vectors holds for arbitrary $X, Y \subseteq V$:*

$$\chi(E(X)) + \chi(E(Y)) + \chi(E(X \setminus Y, Y \setminus X)) = \chi(E(X \cap Y)) + \chi(E(X \cup Y)). \tag{2}$$

*Consequently for any $f : E \to \mathbb{R}_{\geq 0}$*

$$i_f(X) + i_f(Y) + d_f(X, Y) = i_f(X \cap Y) + i_f(X \cup Y).$$

*In particular*

$$i_f(X) + i_f(Y) \leq i_f(X \cap Y) + i_f(X \cup Y).$$

**Remark 6.** The latter property of $i_f$ is called *supermodularity.*

*Proof.* For each $e \in E$ we check that (2) holds for the corresponding coordinate. One needs to consider several cases depending on how endpoints of $e$ are aligned with respect to $X$ and $Y$. The other equations are simple consequences of (2). $\square$

Let $\mathcal{F}$ denote the family of all tight sets, i.e.

$$\mathcal{F} = \{S \subseteq V : i_{x^\star}(S) = |S| - 1\}.$$

**Lemma 7.** *If $A, B \in \mathcal{F}$ are intersecting (i.e. $A \cap B \neq \emptyset$), then both $A \cup B \in \mathcal{F}$ and $A \cap B \in \mathcal{F}$. Moreover $d_{x^\star}(A, B) = 0$ and consequently $E(A \setminus B, B \setminus A) = \emptyset$.*

*Proof.* We have

$$\begin{aligned} |A| - 1 + |B| - 1 = i_{x^\star}(A) + i_{x^\star}(B) \leq i_{x^\star}(A) + i_{x^\star}(B) + d_{x^\star}(A, B) = \\ i_{x^\star}(A \cap B) + i_{x^\star}(A \cup B) \leq |A \cap B| - 1 + |A \cup B| - 1 = |A| - 1 + |B| - 1. \end{aligned}$$

We conclude that all the inequalities are actually equalities. In particular $i_{x^\star}(A \cap B) = |A \cap B| - 1$ and $i_{x^\star}(A \cup B) = |A \cup B| - 1$, i.e. $A \cap B$ and $A \cup B$ are both tight. Moreover $d_{x^\star}(A, B) = 0$ and, as we assume $x^\star(e) > 0$, this implies $E(A \setminus B, B \setminus A) = \emptyset$.

Note that we have used the assumption that $A \cap B \neq \emptyset$ as $i_{x^\star}(\emptyset) = 0 > -1 = |\emptyset| - 1$ while for non-empty set the converse inequality is provided by (1). $\square$

We shall prove Theorem 4 in the following form, which is clearly equivalent to the original.

**Lemma 8.** *Let $\mathcal{L}$ be an inclusion-wise maximal laminar subfamily of $\mathcal{F}$. Then $\mathcal{L}$ spans $\mathcal{F}$, where the linear structure is given by vectors $\chi(E(S))$ for $S \in \mathcal{F}$.*

*Proof.* For a proof by contradiction assume that $\mathcal{L}$ is inclusion-wise maximal but fails to span $\mathcal{F}$. Then there exists $S \in \mathcal{F} \setminus \operatorname{span} \mathcal{L}$. Let us choose $S$ to minimize number of sets $T \in \mathcal{L}$ that cross $S$. As $\mathcal{L}$ is maximal, this value is at least one; therefore let $T \in \mathcal{L}$ cross $S$.

Let us consider sets $T \cap S$ and $T \cup S$, which are tight by Lemma 7. Observe that $T \cap S$ or $T \cup S$ cross only those $T' \in \mathcal{L}$ that $S$ crosses. For a proof, we assume that $S$ does not cross $T'$, i.e. these sets are disjoint or nested. As $T, T' \in \mathcal{L}$ these sets are also disjoint or nested. A straightforward inspection of all cases shows that $T \cap S$ and $T \cup S$ do not cross $T'$.

Obviously $T \cap S$ and $T \cup S$ do not cross $T$, so these sets cross less sets in $\mathcal{L}$ than $S$. Hence, $T, T \cap S, T \cup S \in \operatorname{span} \mathcal{L}$. However, by Lemmas 5 and 7 we get $\chi(E(S)) = \chi(E(S \cap T)) + \chi(E(S \cup T)) - \chi(E(T))$, so $S \in \operatorname{span}(\mathcal{L})$, a contradiction with the definition of $S$. □

The technique we have used is called *uncrossing*. Below, we present its constructive variant.

**Remark 9.** A laminar family $\mathcal{L}$ satisfying the conditions of Lemma 8 can be found in polynomial time.

*Proof.* The ellipsoid method, as well as other LP solvers, along with an extremal optimal solution return a family of tight constraints defining this solution. These constraints give a (linear) base $\mathcal{S}$ of $\operatorname{span} \mathcal{F}$. There is no reason for $\mathcal{S}$ to be laminar, but we can transform $\mathcal{S}$ preserving $\operatorname{span} \mathcal{S}$ so that finally we obtain a laminar family that still spans $\mathcal{F}$.

The construction we are to give is going to mimic the non-constructive proof presented above. We maintain a linearly-independent laminar family $\mathcal{L} \subseteq \mathcal{F}$. If $\mathcal{L}$ spans $\mathcal{S}$, then $\mathcal{L}$ also spans $\mathcal{F}$, so we are done. Otherwise we can find $S \in \mathcal{S}$ such that $S \notin \operatorname{span} \mathcal{L}$. Then, we modify $S$ several times maintaining an invariant $S \in \mathcal{F} \setminus \operatorname{span} \mathcal{L}$. In each step we decrease the number of sets $T \in \mathcal{L}$ that cross $S$ so that finally $\mathcal{L} \cup \{S\}$ is laminar, i.e. we can extend $\mathcal{L}$ by adding $S$. Note that we modify a 'local' copy of $S$, we never change the base $\mathcal{S}$.

In a single step we take $T \in \mathcal{L}$ crossing $S$ and replace $S$ with $S \cap T$ or $S \cup T$, depending on which set is not generated by $\mathcal{L}$. By Lemma 7, $S$ is generated by $T, S \cap T$ and $S \cup T$, so it is not possible that both $S \cap T$ and $S \cup T$ are generated by $\mathcal{L}$. The argument that the number of sets in $\mathcal{L}$ crossing $S$ decreases is exactly the same as in the non-constructive proof. □

□

**Theorem 10.** *The linear program* (1) *is integral.*

As usually, edges $e$ such that $x^\star(e) = 0$ can be removed and the solution remains optimal and extremal, so we assume that $x^\star(e) > 0$ for all $e \in E$.

We present two proofs of the theorem above. While the first one is simpler, it is the second that gives more insight.

*First proof.* Let us start with the following lemma.

**Lemma 11.** *A laminar family of subsets of an $n$-element universe has up to $2n$ elements. At most $n-1$ of them may contain 2 or more elements.*

*Proof.* Simple induction. □

Together with Theorem 4 this shows that $|E| = |\mathcal{L}| \leq |V| - 1$, since sets $S$ of cardinality less than two correspond to null vectors, i.e. cannot be contained in a linearly-independent family $\mathcal{L}$. However $i_{x^\star}(V) = |V| - 1$ and applying the constraint $i_{x^\star}(S) \leq |S| - 1$ to the endpoints of an edge $e$, we obtain $x^\star(e) \leq 1$ for each edge. Consequently the LP (1) is integral. □

*Second proof.* We argue using a discharging method. Initially, we charge each edge with one unit (*charging*). Then each edge $e$ gives $x^\star(e)$ of its charge to the smallest set $S \in \mathcal{L}$ such that $e \in E(S)$ (*discharging*). Since $\mathcal{L}$ is laminar, such a set is uniquely determined unless $e \notin E(S)$ for all $S \in \mathcal{L}$. In this case, $e$ does not give its charge to anybody.

By Theorem 4, $|E| = |\mathcal{L}|$, so the initial charge is $|\mathcal{L}|$. We shall prove that each $S \in \mathcal{L}$ receives at least one unit of its potential. Consequently the sets of $\mathcal{L}$ in total receive $|\mathcal{L}|$ units of charge. This means that no potential is left in the edges, in particular $x^\star(e) = 1$ for all $e \in E$ since each edge is left with $1 - x^\star(e)$ units of charge.

**Claim 12.** Each $S \in \mathcal{L}$ receives at least one unit of charge.

*Proof.* Let $S_1, \ldots, S_k$ be inclusion-wise maximal (proper) subsets of $S$ present in $\mathcal{L}$. Observe that the charge that $S$ receives is

$$i_{x^\star}(S) - \sum_{i=1}^{k} i_{x^\star}(S_i).$$

This is because $S$ receives charge from edges that are induced by $S$ but not by any subset of $S$ present in $\mathcal{L}$. Sets $S, S_1, \ldots, S_k$ are all tight and sets $S_i$ are pairwise disjoint, which gives:

$$i_{x^\star}(S) - \sum_{i=1}^{k} i_{x^\star}(S_i) = |S| - 1 - \sum_{i=1}^{k}(|S_i| - 1) = \left| S \setminus \bigcup_{i=1}^{k} S_i \right| + k - 1.$$

If $k \geq 2$ this already proves the claim. If $k = 1$, then $|S \setminus S_1| \geq 1$, since $S_1$ is a proper subset of $S$, so the claim remains valid. Finally, for $k = 0$ we repeat what we have already seen in the first proof: sets of cardinality $\leq 1$ correspond to null vectors, so they cannot be present in $S$. Consequently $|S| \geq 2$, which implies the claim. □

□

## 2 Minimum degree-bounded spanning tree

| |
|---|
| MINIMUM DEGREE-BOUNDED SPANNING TREE |
| **Input:** A connected graph $G$, a weight function $w : E \to \mathbb{R}_{\geq 0}$ and a bound function $b : B \to \mathbb{Z}_+$ |
| **Output:** A tree $T$ spanning $G$ such that $\forall_{v \in V} \deg_T(v) \leq b(v)$ and $\sum_{e \in T} w_e$ is minimum. |

Note that we cannot hope for traditional approximation algorithms for this problem. In order to overcome some issues with definitions, we turn $G$ into a complete graph by adding edges of huge cost, so that any solution using these edges is worse than any solutions avoiding them. This way some solution always exists, so we can start speaking about approximation algorithms. Nevertheless, if we set $w_e \in \{0, 1\}$ and $b_v = 2$ respectively for all $e \in E$ and $v \in V$, then any approximation algorithm could be used to decide whether the graph formed by the zero-weight edges contains a Hamiltonian path, i.e. to solve an NP-complete problem.

This is why we develop a different kind of approximation algorithm. We shall give an algorithm that gives a spanning tree of weight not exceeding the optimum, but violating some of the degree constraints by one. The following results are due to Singh and Lau [6].

Before we proceed, let us take a look at the following LP, which is a relaxation of the ILP for the MDBST problem.

$$
\begin{aligned}
\text{minimize } \sum_{e \in E} w_e x_e & \\
\forall_{\emptyset \neq S \subseteq V},\ x(E(S)) &\leq |S| - 1 \\
x(E) &= |V| - 1 \\
\forall_{v \in V}\ x(\delta(v)) &\leq b_v \\
\forall_{e \in E}\ x_e &\geq 0
\end{aligned}
\tag{3}
$$

Here $\delta(v)$ denotes the set of edges incident to $v$. In contrast with a similar LP for the MST problem, this LP is not integral for some graphs. (This is nothing unexpected; otherwise we could solve MDBST in polynomial time and prove $P = NP$).

The algorithm is going to iteratively remove edges and degree-constraints. We will prove that it finally obtains a tree $T$ such that $\sum_{e \in T} w_e \leq OPT$ and $\forall_{v \in V} \deg_T(v) \leq b_v + 1$, where $OPT$ is the optimum of the LP (3). Since some the degree constraints are removed, formally we work with the following $LP(E, W)$ with $W \subseteq V$ being a set of vertices on which we still impose the constraints:

$$
\begin{aligned}
\text{minimize } \sum_{e \in E} w_e x_e & \\
\forall_{\emptyset \neq S \subseteq V},\ x(E(S)) &\leq |S| - 1 \\
x(E) &= |V| - 1 \\
\forall_{v \in W}\ x(\delta(v)) &\leq b_v \\
\forall_{e \in E}\ x_e &\geq 0
\end{aligned}
\tag{4}
$$

The algorithm works according to the following pseudocode:

**Algorithm 1**: Minimum Degree-Bounded Spanning Tree

```
while W ≠ ∅ do
    x* := an extremal optimal solution of LP(E, W);
    E := {e ∈ E : x*(e) > 0};
    foreach v ∈ W : deg_G(v) ≤ b_v + 1 do
        W := W − v;
x* := an extremal optimal solution of LP(E, ∅);
return x*
```

Finally we have $W = \emptyset$, i.e. we get a linear program, that we have already proved in Theorem 10 to give the minimum spanning tree. Moreover, we never increase the optimum of the LP by removing constraints and edges $e$ with $x^\star(e) = 0$. That is why we always return a spanning tree of weight not exceeding the optimum of the original LP, which in turn is not greater than the optimum integral solution, i.e. the solution of MDBST. While removing a constraint for $v$, we have just $b_v + 1$ edges incident to $v$, so it is also evident that we never violate any degree-constraint by more than one.

What needs to be shown, however, is that Algorithm 1 always terminates. This is a simple consequence of the following lemma:

**Lemma 13.** *If $W \neq \emptyset$ and $G$ is not a tree, then there exists $v \in W$ such that* $\deg(v) < b_v + 2$.

*Proof.* For a proof by contradiction we assume that $\deg(v) \geq b_v + 2$ for all $v \in W$. Again, we assume that we have an extremal optimal solution $x^\star$ which is positive on all edges.

**Claim 14.** There exists a set $T \subseteq W$ and a laminar family $\mathcal{L} \subseteq 2^V$ such that

- $\forall_{v \in T}\ x^\star(\delta(v)) = b_v$,
- $\forall_{S \in \mathcal{L}}\ i_{x^\star}(S) = |S| - 1$,
- $\operatorname{span}(\mathcal{L}) = \{S \subseteq V : i_{x^\star}(S) = |S| - 1\}$,
- vectors $\chi(E(S))$ for $S \in \mathcal{L}$ and $\chi(\delta(v))$ for $v \in T$ are linearly independent,
- $|E| = |\mathcal{L}| + |T|$.

*Proof.* Due to the fact we've seen in lecture 4, the linear space spanned by non-trivial tight constraints has dimension $|E|$. The tight constraints are of two forms: $i_{x^\star}(S) = |S| - 1$ and $x^\star(\delta(v)) = b_v$. Let $\mathcal{F}$ be the family of tight constraint of the first type. By Lemma 8 $\mathcal{F}$ is generated by a linearly-independent laminar family $\mathcal{L}$.

A simple fact from linear algebra states that if $B$ is a base of $V$ and $B'$ spans $V'$, then $B$ can be extended by a subset of $T \subseteq B'$ so that $B \cup T$ is a base of $V + V'$. We apply it for $B = \mathcal{L}$ and $B'$ corresponding to the tight conditions of the second type. This way we get $T \subseteq V$, which together with $\mathcal{L}$ satisfies all conditions of the claim. □

Now, we apply a discharging argument similar to the one in the proof of Theorem 10. Again each edge $e$ is charged with one unit. Then, if $e$ is induced by some $S \in \mathcal{L}$, then $e$ gives $x^\star(e)$ to the smallest such set $S$. Moreover, it splits the remaining charge evenly between its endpoints, i.e. each endpoint receives $\frac{1-x^\star(e)}{2}$ charge.

We shall see that each $S \in \mathcal{L}$ and each $v \in T$ receives at least one unit of charge. Then, by a more involved argument, we shall prove that some charge is not counted this way.

A proof that each $S \in \mathcal{L}$ gets at least 1 unit of charge is identical to the one in the proof of Theorem 10. Moreover, each vertex $v \in V$ receives exactly $\frac{1-x^\star(e)}{2}$ from each incident edge $e$, which for $v \in W$ in total gives

$$\sum_{e \in \delta(v)} \frac{1 - x^\star(e)}{2} = \frac{\deg(v) - x^\star(\delta(v))}{2} \geq \frac{\deg(v) - b_v}{2} \geq \frac{b_v + 2 - b_v}{2} \geq 1.$$

It remains to find some charge not counted this way. Let $v \in V \setminus T$. If $v$ receives some charge we are done, so we may assume that each edge incident to $v$ satisfies $x^\star(e) = 1$. Thus, if $vw \in E$, then $i_{x^\star}\{v, w\} = 1$, so $\{v, w\} \in \operatorname{span} \mathcal{L}$ as a tight set. Moreover $V$ is forced by (4) to be tight, so $V \in \operatorname{span} \mathcal{L}$.

The characteristic vectors clearly satisfy the following equality

$$2\chi(E) = \sum_{v \in V} \chi(\delta(v)) = \sum_{v \in T} \chi(\delta(v)) + \sum_{v \in V \setminus T} \sum_{w : vw \in E} \chi(\{vw\}). \tag{5}$$

Assume that $T$ is not empty and fix $t \in T$. Then (5) allows to express $\chi(\delta(t))$ as a linear combination of vectors $\chi(\delta(v))$ for other $v \in T$ and vectors in $\operatorname{span} \mathcal{L}$ (i.e. vectors in $\mathcal{L}$ as well). This is a contradiction with our choice of $\mathcal{L}$ and $T$ and thus $T = \emptyset$. However for $v \in V \setminus T$ we have shown that $x^\star(\delta(v)) = \deg(v)$. In particular for $v \in W$ this implies $\deg(v) = b_v$, a contradiction with the assumption that $\deg(v) \geq b_v + 2$ for all $v \in W$, since $W \neq \emptyset$. $\square$

**Remark 15.** Algorithm 1 can be changed so that an LP solver is called only once.

*Proof.* Note that just in the first phase we need $x^\star$ to be optimal. In the further ones it suffices for $x^\star$ to be extremal; we also need to make sure that $\sum_e x_e^\star w_e$ does not increase.

In order to make the analysis simpler, we assume that just one vertex is removed from $W$ in a single phase. This might increase the number of phases but does not influence the correctness of the algorithm. Let us investigate what may happen with $x^\star$, when a vertex $v$ is removed from $W$. Note that $x^\star$ as an extremal vertex is defined by a collection $C$ of $|E|$ (linearly-independent) tight constraints. We are given these constraints by the LP solver and we shall maintain such a set. The only effect of removing $v$ from $W$ is the deletion the corresponding constraint. If this constraint does not belong to $C$, then $x^\star$ is still extremal. Otherwise, the remaining constraints define a line, which can be determined with a single Gaussian elimination. When we move $x^\star$ along this line, then in one of the directions the objective function does not increase.

The idea is to move $x^\star$ in this direction as long as it does not violate any constraints. Then, we add to $C$ any of the constraints that prevent $x^\star$ from moving further. The polytope we consider is bounded, so this way we clearly obtain an extremal feasible solution. Technically, we perform a binary search algorithm and use the separation oracle to check feasibility and obtain a constraint that is violated. Note that we need to use high-precision numbers, as the solution might be a fraction with denominator as large as $\Theta(n!)$. $\square$

# 3 Dependent rounding

Ultimately, we are going to give an $O\left(\frac{\log n}{\log\log n}\right)$ approximation algorithm for Asymmetric Travelling Salesmen Problem (with triangle inequality). Note there is a simple $O(\log n)$ approximation algorithm (see the Algorithmics course). A few years ago Asadpour et al. [2] have broken a long-standing barrier providing the $O\left(\frac{\log n}{\log\log n}\right)$-approximation. Nevertheless many people believe that a 2-approximation exists.

However, before presenting the new result for ATSP, we first introduce the technique of dependent rounding using a simpler problem as an example.

## 3.1 Minimum-capacity integer multicommodity flow

> Minimum-Capacity Integer Multicommodity Flow
> **Input:** A graph $G$, a family of triples $(s_i, t_i, \mathcal{P}_i)$, where $s_i, t_i \in V(G)$ and $\mathcal{P}_i$ is a collection of $s_i$-$t_i$-paths in $G$
> **Output:** Choose a single path from each $\mathcal{P}_i$ to minimize the maximum number of paths passing through a single edge.

Note that one might consider a capacitated version of this problem and the algorithm we develop can still be used to solve it. In this version we minimize $\max_{e\in E} \frac{p_e}{c_e}$, where $p_e$ is the number of paths passing through $e$ and $c_e$ is the capacity of $e$.

The following LP is a relaxation of the ILP solving the problem. Let $\mathcal{P} = \bigcup_i \mathcal{P}_i$.

$$\begin{aligned} &\text{minimize } y \\ \forall_{e\in E} &\sum_{P\in\mathcal{P}\,:\,e\in P} x_P \le y \\ \forall_i &\sum_{P\in\mathcal{P}_i} x_P = 1 \\ &\forall_{P\in\mathcal{P}}\, x_P \ge 0 \end{aligned} \tag{6}$$

Let $x^\star$ be an optimal solution. Observe that $x^\star$ gives a probability distribution on each family $\mathcal{P}_i$. Let us analyze a simple randomized algorithm, which computes $x^\star$ and then, for each $i$ draws a random path $P \in \mathcal{P}_i$ from this distribution.

Let $ALG$ be the solution obtained by the algorithm and let $OPT$ be the optimum of the ILP (which is clearly larger than the optimum of the LP). Note that $ALG$ is formally a random variable. Let $X_P$ be a binary random variable, $X_P = 1$ means that the algorithm includes $P$ in the solution.

Observe that $\mathbb{E}[X_P] = x^\star_P$. Consequently for each $e \in E$ we have

$$\mathbb{E}\left[\sum_{P\in\mathcal{P}\,:\,e\in P} X_P\right] = \sum_{P\in\mathcal{P}\,:\,e\in P} \mathbb{E}[X_P] = \sum_{P\in\mathcal{P}\,:\,e\in P} x^\star_P \le OPT.$$

Nevertheless we care about the worst edge, so bounding the expectation for each edge does suffice for the analysis.

**Definition 16.** Binary variables $X_1, \ldots, X_n$ are *negatively correlated* if for any $I \subseteq \{1, \ldots, n\}$

$$\mathbb{P}\left[\forall_{i\in I} X_i = 0\right] \le \prod_{i\in I} \mathbb{P}[X_i = 0]$$
$$\mathbb{P}\left[\forall_{i\in I} X_i = 1\right] \le \prod_{i\in I} \mathbb{P}[X_i = 1]$$

**Theorem 17** (Chernoff-Hoeffding bounds, generalized version)**.** *Let $X_1, \ldots, X_n$ be negatively correlated binary variables, $X = \sum_{i=1}^n X_i$ and $\mu = \mathbb{E}[X]$. Then for any $\delta > 0$ and $U \ge \mu$*

$$\mathbb{P}[X \ge (1+\delta)U] \le \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^U$$

*In particular, for $\delta \le 1$*

$$\mathbb{P}[X \ge (1+\delta)U] \le e^{-\frac{U\delta^2}{3}}$$

For a proof and more details, see [4, 5].

**Lemma 18.** *Variables $X_P$ for $P \in \mathcal{P}$ are negatively correlated.*

*Proof.* Observe that variables $P$ from different $\mathcal{P}_i$ are independent. That is for any $\mathcal{Q} \subseteq \mathcal{P}$

$$\mathbb{P}\left[\forall_{P\in\mathcal{Q}} X_P = 1\right] = \prod_i \mathbb{P}\left[\forall_{P\in\mathcal{Q}\cap\mathcal{P}_i} X_P = 1\right].$$

If $P, P' \in \mathcal{P}_i$, then $\mathbb{P}[X_P = 1 \wedge X_{P'} = 1] = 0$, so

$$\mathbb{P}\left[\forall_{P \in \mathcal{Q} \cap \mathcal{P}_i} X_P = 1\right] \leq \prod_{P \in \mathcal{Q} \cap \mathcal{P}_i} \mathbb{P}[X_P = 1]$$

and consequently

$$\mathbb{P}\left[\forall_{P \in \mathcal{Q}} X_P = 1\right] = \prod_i \mathbb{P}\left[\forall_{P \in \mathcal{Q} \cap \mathcal{P}_i} X_P = 1\right] \leq \prod_i \prod_{P \in \mathcal{Q} \cap \mathcal{P}_i} \mathbb{P}[X_P = 1] = \prod_{P \in \mathcal{Q}} \mathbb{P}[X_P = 1].$$

Moreover for any $\mathcal{Q}_i \subseteq \mathcal{P}_i$ we have

$$\mathbb{P}\left[\forall_{P \in \mathcal{Q}_i} X_P = 0\right] = 1 - \sum_{P \in \mathcal{Q}_i} x^\star_P \leq \prod_{P \in \mathcal{Q}_i} (1 - x^\star_P) = \prod_{P \in \mathcal{Q}_i} \mathbb{P}[X_P = 0]$$

and thus

$$\mathbb{P}\left[\forall_{P \in \mathcal{Q}} X_P = 0\right] = \prod_i \mathbb{P}\left[\forall_{P \in \mathcal{Q} \cap \mathcal{P}_i} X_P = 0\right] \leq \prod_i \prod_{P \in \mathcal{Q} \cap \mathcal{P}_i} \mathbb{P}[X_P = 0] = \prod_{P \in \mathcal{Q}} \mathbb{P}[X_P = 0].$$

□

**Theorem 19.** *Let $c \geq 12$ be a real constant. If $OPT > c \ln n$, then $ALG \leq OPT + \sqrt{cOPT \ln n}$ with high probability (i.e. the converse holds with inverse polynomial probability, here $\mathcal{O}(n^{-2})$).*

*Proof.* Let us fix a single edge $e \in E$. Let $X_e = \sum_{P \in \mathcal{P} : e \in P} X_P$. As we have already shown $\mathbb{E}[X_e] \leq OPT$, moreover $\frac{c \ln n}{OPT} \leq 1$. Then, using the second part of Theorem 17, we get

$$\mathbb{P}\left[X_e \geq OPT + \sqrt{OPT c \ln n}\right] = \mathbb{P}\left[X_e \geq OPT\left(1 + \sqrt{\frac{c \ln n}{OPT}}\right)\right] \leq$$
$$e^{-OPT \frac{c \ln n}{3OPT}} = e^{-\frac{c}{3} \ln n} = n^{-\frac{c}{3}} \leq n^{-4}.$$

Using the union bound we obtain

$$\mathbb{P}\left[ALG \geq OPT + \sqrt{OPT c \ln n}\right] = \mathbb{P}\left[\exists_{e \in E} X_e \geq OPT + \sqrt{OPT c \ln n}\right] \leq mn^{-4} \leq n^{-2}.$$

□

**Corollary 20.** *The algorithm is an $O(\log n)$ approximation with high probability.*

### 3.2 Dependent rounding

Let us consider the following extension of the Minimum-Capacity Integer Multicommodity Flow problem.

| |
|---|
| Extended Minimum-Capacity Integer Multicommodity Flow<br>**Input:** A graph $G$, a family of tuples $(s_i, t_i, \mathcal{P}_i, k_i)$, where $s_i, t_i \in V(G)$, $\mathcal{P}_i$ is a collection of $s_i$-$t_i$-paths in $G$, and $k_i \in \mathbb{Z}_+$<br>**Output:** Choose $k_i$ paths from each $\mathcal{P}_i$ to minimize the maximum number of paths passing through a single edge. |

The following LP is a relaxation of the ILP formulation of this problem

$$\begin{aligned} &\text{minimize } y \\ \forall_{e\in E} &\sum_{P\in\mathcal{P}\,:\,e\in P} x_P \le y \\ \forall_i &\sum_{P\in\mathcal{P}_i} x_P = k_i \\ &\forall_{P\in\mathcal{P}}\ x_P \ge 0 \\ &\forall_{P\in\mathcal{P}}\ x_P \le 1 \end{aligned} \tag{7}$$

Provided that we can draw paths so that $\mathbb{E}[X_P] = x^\star_p$ for an optimal solution $x^\star$ of (7), the analysis for the original problem still holds. Nevertheless $x^\star$ does not give a probability distribution over $\mathcal{P}_i$, so it is not clear how to draw paths, in particular several naive approaches fail.

The following technique provides a solution in much more general setting, see [3, 1].

**Theorem 21.** *Let $G = (V, E)$ be a bipartite graph and $x : E \to (0, 1)$ be an arbitrary function. Then we can (algorithmically) construct binary random variables $X_e$ such that*

- $\mathbb{E}[X_e] = x_e$ *for each* $e \in E$,
- $\sum_{e\in\delta(v)} X_e \in \{\lfloor d_v \rfloor, \lceil d_v \rceil\}$ *for each* $v \in V$ *where* $d_v = \sum_{e\in\delta(v)} x_e$,
- *variables* $X_e$ *for* $e \in \delta(v)$ *are negatively correlated for each* $v \in V$.

Note for our purposes is suffices to consider a bipartite graph with a single vertex in one of the colour classes. Nevertheless, we will prove the result in full generality.

*Proof of Theorem 21.* Let us consider the following algorithm

**Algorithm 2**: Dependent Rounding

```
while E ≠ ∅ do
    Let C ⊆ E by a cycle or an inclusion maximal simple path in G; Let
    C = M_1 ∪ M_2, where M_i are matchings;
    α := min{x_e : e ∈ M_1} ∪ {1 − x_e : e ∈ M_2};
    β := min{1 − x_e : e ∈ M_1} ∪ {x_e : e ∈ M_2};
    with probability β/(α+β) do
        foreach e ∈ M_1 do x_e := x_e − α;
        foreach e ∈ M_2 do x_e := x_e + α;
    otherwise
        foreach e ∈ M_2 do x_e := x_e − β;
        foreach e ∈ M_1 do x_e := x_e + β;
    foreach e ∈ E do
        if x_e = 0 then set X_e = 0; E := E − e;
        if x_e = 1 then set X_e = 1; E := E − e;
```

See lecture 6 for a remainder of the proof. ∎

# References


[1] A. A. Ageev and M. Sviridenko. Pipage rounding: A new method of constructing algorithms with proven performance guarantee. *J. Comb. Optim.*, 8(3):307–328, 2004.

[2] A. Asadpour, M. X. Goemans, A. Madry, S. O. Gharan, and A. Saberi. An O(log n/ log log n)-approximation algorithm for the asymmetric traveling salesman problem. In M. Charikar, editor, *SODA*, pages 379–389. SIAM, 2010.

[3] R. Gandhi, S. Khuller, S. Parthasarathy, and A. Srinivasan. Dependent rounding and its applications to approximation algorithms. *J. ACM*, 53(3):324–360, 2006.

[4] M. Mitzenmacher and E. Upfal. *Probability and computing - randomized algorithms and probabilistic analysis*. Cambridge University Press, 2005.

[5] A. Panconesi and A. Srinivasan. Randomized distributed edge coloring via an extension of the chernoff-hoeffding bounds. *SIAM J. Comput.*, 26(2):350–368, 1997.

[6] M. Singh and L. C. Lau. Approximating minimum bounded degree spanning trees to within one of optimal. In D. S. Johnson and U. Feige, editors, *STOC*, pages 661–670. ACM, 2007.