**CSE 525: Randomized Algorithms** **Spring 2023**

# Lecture 10: Pipage Rounding Method

*Lecturer: Shayan Oveis Gharan* *April 27th*

**Disclaimer**: *These notes have not been subjected to the usual scrutiny reserved for formal publications.*

In this lecture we study a randomized rounding method called *pipage rounding method*. The materials of this lecture are based on the work of Chekuri, Vondrák and Zenklusen [**CVZ10**].

From a high-level point of view, one can use this method to round a fractional point in (some families of) integral polytope, i.e., polytopes whose vertices are integer points, to a vertex while making sure that the underlying elements are negatively correlated.

We will instantiate a method for the case of spanning trees, but the method works for a more general family of polytopes called matroid base polytopes and beyond.

First, let us define the spanning tree polytope. Given a graph $G = (V, E)$, the spanning tree polytope is the **convex hull** of the indicator vectors, $1_T \in \mathbb{R}^E$ for spanning tree $T$, of all spanning trees of $G$. Edmonds proved that one can give a simple description of the faces of this polytope: Say we have a variable $x_e$, for every edge $e \in E$. Then,

$$\begin{aligned} &\sum_{e \in E(S)} x_e \leq |S| - 1, \quad \forall S \subseteq V, \\ &\sum_e x_e = |V| - 1, \\ &x_e \geq 0, \qquad\qquad\qquad \forall e \in E. \end{aligned} \tag{10.1}$$

It is not hard to see that $\mathbf{1}_T$ for any (integral) spanning tree $T$ of $G$ indeed satisfies all of the above constraints. In particular, $T$ can have at most $|S| - 1$ edges in any set $S$ of vertices (otherwise $T$ has a cycle). Edmonds proved that indeed that is all one needs to check. We don't prove here, but it is not hard to see that given a point $x \in \mathbb{R}^E$, we can check in polynomial whether $x$ is in the spanning tree polytope or find a violating constraint. Because of that we can optimize any linear function (or minimize any convex function) over this polytope. See course notes of CSE 521 for more info. In the rest of this note we use the shorthand $x(E(S)) = \sum_{e \in E(S)} x_e$.

**Tight Sets** Given a point $x \in \mathbb{R}^E_{\geq 0}$, we say a set $S \subseteq V$ is a tight set, if $\sum_{e \in E(S)} x_e = |S| - 1$. The following fact is immediate:

**Fact 10.1.** *Suppose $S, T \subseteq V$ are two tight sets for a given $x \in \mathbb{R}^E_{\geq 0}$. Then, $S \cap T$, if non-empty, is also a tight set.*

*Proof.* Suppose $E(S \cap T) \neq \emptyset$. The claim is obvious if $S \subseteq T$ or $T \subseteq S$. Since $S, T$ are tight,

$$\begin{aligned} |S| - 1 + |T| - 1 &= x(E(S)) + x(E(T)) \\ &\leq x(E(S \cap T)) + x(E(S \cup T)) \\ &\leq |S \cap T| - 1 + |S \cup T| - 1 = |S| - 1 + |T| - 1 \end{aligned}$$

The first inequality follows from the fact there may some edges in $S \cup T$ that are not in $S$ or $T$, i.e., those with one endpoint in $S \smallsetminus T$ and the other in $T \smallsetminus T$. The second inequality follows from feasiblity of $x$. So all of the inequalities must be equalities and in particular, $x(E(S \cap T)) = |S \cap T|$. ∎

## 10.1 The Pipage Rounding Method

In this section we describe the pipage rounding algorithm. Given a point $x$ in the spanning tree polytope, there are many ways to write $x$ as a convex combination of the vertices of the polytope,

$$x = \sum_T p_T \mathbf{1}_T.$$

No matter which distribution we choose we always satisfy linear functions in expectations. A quantity of interest is whether this "rounding" introduces correlations between coordinates. In other words, if we can use strong concentration bounds (such as multiplicative Chernoff bound) to the rounded solution. More concretely, is it possible to write $x$ as a distribution that is negatively correlated? We will see that yes, for spanning trees, and more generally for any matroid this is possible.

This algorithm is first proposed by Ageev and Sviridenco [**AS04**] to round a fractional matching into an integral one. Chekuri, Vondrák and Zenklusen [**CVZ10**] observed that a randomized version of the pipage rounding algorithm gives a negatively correlated distribution; among other things, the method can be used to get an $O(\log(n)/\log\log(n))$ approximation for ATSP.

In the basic version of this method we start from a fractional point $x$ in the spanning tree polytope and we round to an integral spanning tree in a number of steps $T$. Let $X_i \in \mathbb{R}^E$ be our vector in the $i$-th step.

i) At time zero we start with the given $x$, $X_0 = x$ deterministically and at any $i$, $X_i$ is in the spanning tree polytope.

ii) $X_\infty$ must be an integral spanning tree. In fact, we will see that $X_t$ for some $t = \text{poly}(n)$ will be integral.

iii) $X$ is a martingale: For any $i$,
$$\mathbb{E}\left[X_i | X_0, \ldots, X_{i-1}\right] = X_{i-1}.$$
This in particular implies that for any $i$, $\mathbb{E}\left[X_i\right] = x$, i.e., the original marginals are preserved.

iv) The most important property is negative correlation: We will see that products of coordinates of $X_i$ form a "sub-martinagle". For any set $F \subseteq E$ and any $i \geq 1$,
$$\mathbb{E}\left[\prod_{e\in F} X_i(e)\right] \leq \mathbb{E}\left[\prod_{e\in F} X_{i-1}(e)\right].$$
Using induction, this property in particular implies that we have multiplicative Chernoff bound for sums of coordinates of $X_i$ (for any $i$).

Now, we are ready to describe the pipage rounding algorithm. Given $X_{i-1}$, if it is integral we are done. Otherwise, there must be at least two non-integral coordinates. We choose the smallest tight set $S$ with at least two fractional elements say $0 < X_{i-1}(e), X_{i-1}(f) < 1$. We consider the line $\mathbf{1}_e - \mathbf{1}_f$ and extend it until we hit the faces of the polytope (see Figure 10.1). We move to one of these two points randomly such that in expectation is $X_{i-1}$. In particular say these two points are $x_1, x_2$ (see Figure 10.1). We let $X_i = x_1$ with probability $p$ and $X_i = x_2$ with probability $1 - p$ for $p$ chosen such that

$$px_1 + (1-p)x_2 = X_{i-1}.$$

So, this immediately proves the martingale property.

**Lemma 10.2.** *$X_t$ is integral (with probability 1) for $t \geq |E|(n-1)$.*

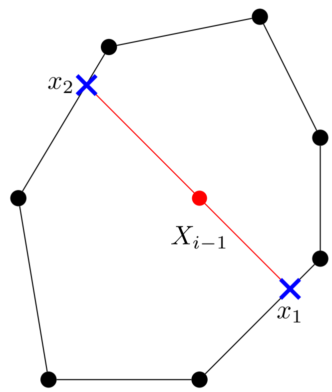


Figure 10.1: An iteration of the pipage rounding method. given the point $x$, we choose two variables say $x(e), x(f)$ and move randomly along the direction of the line that keeps the sum of $x(e), x(f)$ invariant until we hit the polytope, i.e., the blue crosses.

*Proof.* The first observations is that if a set $T \subseteq V$ is a tight set in $X_{i-1}$ it remains a tight set in $X_i$. To see that, notice in step $i$ we only change $X_{i-1}(e), X_{i-1}(f)$, and that $X_i(e) + X_i(f) = X_{i-1}(e) + X_{i-1}(f)$. So, there are the following cases:

**Case 1:** $e, f \notin E(T)$ Then, $X_i(E(T)) = X_{i-1}(E(T))$ so $T$ remains tight. In particular, if $S \subseteq T$ or $S \cap T = \emptyset$, $T$ remains tight.

**Case 2:** $T \subsetneq S$ This case cannot happen because $S$ is the smallest tight set.

**Case 3: Only one of $e, f$ in $T$.** In that case we must have $E(S \cap T) \neq \emptyset$. But then by Fact 10.1, $S \cap T$ was a tight set in $X_{i-1}$ so $S$ was not the smallest tight set.

Now, consider changes in step $i$. By definition of the algorithm, we move $X_{i-1}$ to one of two faces $x_1$ or $x_2$. In either of these point we must hit a new face of the polytope. If we hit an integrality face, for either of $e, f$ i.e., if $X_i(e) = 1$ or $X_i(f) = 1$ then we make progress and we have one more integral variable (observe that once a coordinate becomes integral it remains integral. Otherwise, a new set, say $T$, becomes tight. But such a $T$ must have exactly one of $e, f$ otherwise $X_{i-1}(E(T)) = X_i(E(T))$. So, by Fact 10.1, $S \cap T$ is a (new) smaller tight set. Repeating this procedure after at most $n-1$ iterations we get a tight set with a single variable, i.e., an integer coordinate. So, all in all, the procedure gets to an integer vector in at most $(n-1)|E|$ many steps. $\square$

## 10.2 Negative Correlation

In this section we prove property (iv) of the pipage rounding method. Fix a set $F$, We prove a stronger claim,

$$\mathbb{E}\left[\prod_{e \in F} X_i(e) \middle| X_{i-1}\right] \leq \prod_{e \in F} X_{i-1}(e). \tag{10.2}$$

Taking expectations from both sides of the above inequality proves (iv).

Say in step $i$ we move along the line $\mathbf{1}_{f_1} - \mathbf{1}_{f_2}$. Then, we for all $e \neq f_1, f_2$, $X_i(e) = X_{i-1}(e)$. If $f_1, f_2 \notin F$, then (10.2) holds trivially. If $f_1 \in F$ and $f_2 \notin F$, then

$$\begin{aligned}\mathbb{E}\left[\prod_{e\in F} X_i(e)\middle| X_{i-1}\right] &= \mathbb{E}\left[X_i(f_1)|X_i\right] \cdot \prod_{e\in F\setminus\{f_1\}} X_{i-1}(e) \\ &= \prod_{e\in F} X_{i-1}(e),\end{aligned}$$

where the second equality follows by the martingale property (iii).

Now, assume $f_1, f_2 \in F$. First, note that similar to above all we need to show is that

$$\mathbb{E}\left[X_i(f_1)\cdot X_i(f_2)|X_{i-1}\right] \leq X_{i-1}(f_1)\cdot X_{i-1}(f_2). \tag{10.3}$$

The main observation is that $X_i(f_1) + X_i(f_2)$ is invariant (w.p. 1). Therefore,

$$\mathbb{E}\left[(X_i(f_1)+X_i(f_2))^2|X_{i-1}\right] = (X_{i-1}(f_1)+X_{i-1}(f_2))^2. \tag{10.4}$$

Now, since

$$\mathbb{E}\left[X_i^2(f_1)|X_{i-1}\right] \geq \mathbb{E}\left[X_i(f_1)|X_{i-1}\right]^2 = X_{i-1}(f_1)^2$$

and similarly for $f_2$, subtracting this from (10.4) implies (10.3). This concludes the proof of (10.2).

## 10.3 Conclusion

Chekuri, Vondrak and Zenklusen also introduce another algorithm that is similar to pipage rounding with some additional properties in some aspects, called the *randomized swap rounding*. They extend the above ideas to round a fractional point in the intersection of two matroids, e.g., a fractional matching in a bipartite graph.

An open problem here is to understand further properties of the distribution of spanning trees defined by the pipage rounding method. For example, are they negative associated? Are Lipschitz functions concentrated?