A $(1 + \epsilon)$ approx for max cardinality matching in bipartite graphs.

[Ahn, Guha ICALP 11]
Today’s Result

Theorem
There’s a $O(\epsilon^{-4} n \text{ polylog } n)$-space algorithm, using $O(\epsilon^{-3} \log n)$ passes that $1 + \epsilon$ approximates max cardinality matching in bipartite graphs.

The results follows from the following lemma by trying values $\alpha = 1, (1 + \epsilon), (1 + \epsilon)^2 \ldots n$ in parallel.

Lemma
There’s a $O(\epsilon^{-3} \log n)$-pass, $O(\epsilon^{-4} n \text{ polylog } n)$-space algorithm that given value $\alpha$ either:

1. Finds a matching of size $\frac{\alpha}{1+4\epsilon}$
2. Or determines the maximum matching has size at most $\alpha(1 + 2\epsilon)$. 
Ideas behind algorithm: Fractional Matchings

Fact (Fractional Matchings and Vertex Cover)

For any graph, the size of max fractional matching equals for the size of the min fractional vertex cover, i.e.,

$$\max \left\{ \sum_{(i,j) \in E} y_{ij} : \sum_{j \in \Gamma(i)} y_{ij} \leq 1 \ \forall i \in V \right\} = \min \left\{ \sum_{i \in V} x_i : x_i + x_j \geq 1 \ \forall (i,j) \in E \right\}$$

For a bipartite graph, this also equals the size of max (integral) matching.
Ideas behind algorithm: Multiplicative Weights

- $y$ is a fractional matching if for all $i \in V$, $M_i(y) \leq 0$ where

$$M_i(y) = \left( \sum_{j \in \Gamma(i)} y_{ij} \right) - 1.$$

- Rather than trying to satisfy all $n$ constraints, it’s easier to satisfy them “on average”, i.e., given a distribution $\mu$ over $[n]$ find $y$ with

$$\sum_i \mu_i M_i(y) \leq 0.$$

- The algorithm defines sequence of distributions $\mu_1, \mu_2, \mu_3, \ldots, \mu_T$ and finds a set of weights $y^1, y^2, \ldots, y^T$ for each.

- If $M_i(y^t) > 0$ then $\mu^{t+1}$ increases weight on that constraint such that average of $y^1, \ldots, y^T$ nearly satisfies all constraints.
Algorithm and Analysis

1. Define $\mu_i^1 = 1/n$ and $u_i^1 = 1$ for all $i \in V$
2. For $t = 1$ to $T$:
   2.1 **Construct** $y^t$: Let $x_i = \alpha \mu_i^t$ and find maximal matching $S$ in
      $$\{(i, j) \in E : x_i + x_j < 1\}$$
      If $|S| < \delta \alpha$ output “Fail” and otherwise define
      $$y_{ij}^t = \alpha / |S|$$ for each $(i, j) \in S$ and 0 otherwise.
   2.2 **Define** $\mu^{t+1}$: $u_i^{t+1} = u_i^t(1 + \epsilon \delta M_i(y^t))$ and $\mu_i^t = u_i^t / \|u^t\|$ $\forall i \in V$.
3. Output $\frac{1}{1 + 4\epsilon} \sum_t y^t / T$.

**Lemma**

If fail, max fractional matching is $< (1 + 2\delta)\alpha$. Otherwise, $y^t$ satisfies:

$$\sum_{i \in V} \mu_i M_i(y^t) \leq 0 \quad \sum_{e \in E} y_e = \alpha \quad -1 \leq M_i(y^t) \leq 1/\delta \quad \forall i \in V$$

**Theorem**

*Output is fractional matching size $\alpha / (1 + 4\epsilon)$ if $T = O(\epsilon^{-3} \log n)$, $\delta = \epsilon$.**
Proof of Lemma

- If we fail, for each $i$ that is an endpoint of an edge in $S$, consider adding 1 to each $x_i$. Then gives a fractional vertex cover of size
  \[ \sum_i \alpha \mu_i + 2|S| < \alpha + 2\delta \alpha \]
  and hence that max fractional matching has size $< \alpha(1 + 2\delta)$.
- To prove $y^t$ satisfies constraints on average
  \[ \sum_{i \in V} x_i \left( \sum_{j \in \Gamma(i)} y_{ij} - 1 \right) = \sum_{i \in V} x_i \sum_{j \in \Gamma(i)} y_{ij} - \sum_{i \in V} x_i = \sum_{ij \in E} y_{ij}(x_i + x_j) - \alpha < \alpha - \alpha \]
  where last inequality follows since, $x_i + x_j < 1$ for $(i, j) \in S$.
- To prove size of $y^t$ is $\alpha$
  \[ \sum e \in E \ y_e = \sum e \in S \alpha / |S| = \alpha \]
- To prove upper and lower bounds on $M_i(y^t)$,
  \[ -1 \leq \sum_{j \in \Gamma(i)} y_{ij}^t - 1 < \alpha / |S| \leq 1 / \delta \]
Proof of Theorem

- $\sum_i u_i^t$ decreases as $t$ increases:

$$\sum_i u_i^{t+1} = \sum_i u_i^t (1 + \delta \epsilon M_i(y^t)) = \sum_i u_i^t + \delta \epsilon \sum_i u_i M_i(y^t) \leq \sum_i u_i^t$$

- Use this to bound $u_i^{T+1}$:

$$n \geq \|u^{T+1}\| \geq u_i^{T+1} = \prod_{t \in [T]} (1 + \delta \epsilon M_i(y^t))$$

- After some algebra and using fact $-1 \leq M_i(y^t) \leq 1/\delta$:

$$\ln n \geq (\epsilon - \epsilon^2)\delta \sum_{t \in [T]} M_i(y^t) - 2\epsilon^2 T \delta$$

- Finally, setting $y = \sum_t y^t / T$ and $T = \frac{\ln n}{\delta \epsilon^2}$ gives,

$$\sum_{j \in \Gamma(i)} y_{ij} - 1 = M_i(y) = \sum_{t \in [T]} \frac{M_i(y^t)}{T} \leq \frac{\ln n + 2\epsilon^2 T \delta}{T \delta (\epsilon - \epsilon^2)} \leq 4\epsilon$$

and therefore $y' = y / (1 + 4\epsilon)$ satisfies $\sum_{j \in \Gamma(i)} y'_{ij} \leq 1$ for all $i \in [n]$. 
Missing Algebra for Proof of Theorem

Using $1 + \epsilon x \geq\begin{cases} (1 + \epsilon)^x & \text{if } 0 \leq x \leq 1 \\ (1 - \epsilon)^{-x} & \text{if } -1 \leq x \leq 0 \end{cases}$ we get:

$$n \geq \prod_{t \in [T]} (1 + \delta \epsilon M_i(y^t)) \geq (1 + \epsilon)^\delta \sum_{t : M_i(y^t) > 0} M_i(y^t) (1 - \epsilon)^{-\delta} \sum_{t : M_i(y^t) < 0} M_i(y^t)$$

Taking logs of both sides gives:

$$\ln n \geq \ln(1 + \epsilon) \sum_{t : M_i(y^t) > 0} \delta M_i(y^t) - \ln(1 - \epsilon) \sum_{t : M_i(y^t) < 0} \delta M_i(y^t)$$

\[
\geq (\epsilon - \epsilon^2) \sum_{t : M_i(y^t) > 0} \delta M_i(y^t) + (\epsilon + \epsilon^2) \sum_{t : M_i(y^t) < 0} \delta M_i(y^t)
\]

\[
= (\epsilon - \epsilon^2) \sum_{t \in [T]} \delta M_i(y^t) + 2\epsilon^2 \sum_{t : M_i(y^t) < 0} \delta M_i(y^t)
\]

\[
\geq (\epsilon - \epsilon^2) \sum_{t \in [T]} \delta M_i(y^t) - 2\epsilon^2 T \delta
\]