

Chien-Chung Huang¹ François Sellier^{2,3}

 $^1\mathrm{CNRS},$ DI ENS, École normale supérieure, Université PSL, France

²Université Paris Cité, CNRS, IRIF, Paris, France

³MINES ParisTech, Université PSL, Paris, France





Definition of a *b*-Matching



Figure: An input graph with capacities $b_{v_1} = 2, b_{v_2} = 3$, and $b_{v_3} = b_{v_4} = 1$

Definition of a *b*-Matching



Figure: An input graph with capacities $b_{v_1} = 2$, $b_{v_2} = 3$, and $b_{v_3} = b_{v_4} = 1$ An output *b*-matching *M*, in red

The Maximum Weight *b*-Matching Problem



Figure: Finding a maximum weight *b*-matching

Chien-Chung Huang, François Sellier Weighted b-Matchings in Random-Order Streams

The Maximum Weight *b*-Matching Problem



Figure: Finding a maximum weight b-matching The weight of the highlighted matching M is 2 + 3 + 4 = 9.

The Maximum Weight *b*-Matching Problem



Figure: Finding a maximum weight b-matching The weight of the highlighted matching M is 2 + 3 + 4 = 9.

In this work, we assume the weights are small integers in a range $\{1, \ldots, W\}$.

Degree and Weighted Degree



Figure: Defining the degree and weighted degree of vertex v_2 in some subgraph ${\cal H}$

Degree and Weighted Degree



Figure: Defining the degree and weighted degree of vertex v_2 in some subgraph ${\cal H}$

•
$$\deg_H(v_2) = 2$$
 (and $\deg_G(v_2) = 3$)

Degree and Weighted Degree



Figure: Defining the degree and weighted degree of vertex v_2 in some subgraph H

- $\deg_H(v_2) = 2$ (and $\deg_G(v_2) = 3$)
- $\mathbf{w} \deg_H(v_2) = 3 + 4 = 7$ (and $\mathbf{w} \deg_G(v_2) = 3 + 4 + 2 = 9$)

Some Definitions

In the *semi-streaming* model:

• edges of E arrive over time in a stream $S = \langle e_1, \ldots, e_m \rangle$

Some Definitions

In the *semi-streaming* model:

- edges of E arrive over time in a stream $S = \langle e_1, \ldots, e_m \rangle$
- we have access only to a **limited memory**

Some Definitions

In the *semi-streaming* model:

- edges of E arrive over time in a stream $S = \langle e_1, \ldots, e_m \rangle$
- we have access only to a **limited memory**

In the *random-order semi-streaming* model:

• the permutation of the edges given in the stream is assumed to be chosen **uniformly at random** among all the possible permutations

Previous Results

In the *adversarial* (worst-case order) semi-streaming model:

- For unweighted maximum (*b*-)matching approximation:
 - 2 approximation using the trivial greedy algorithm
 - $1 + \ln 2 \approx 1.69$ inapproximability [Kapralov, 2021]

Previous Results

In the *adversarial* (worst-case order) semi-streaming model:

- For unweighted maximum (*b*-)matching approximation:
 - 2 approximation using the trivial greedy algorithm
 - $1 + \ln 2 \approx 1.69$ in approximability [Kapralov, 2021]
- For maximum weight matching approximation:
 - 6 [Feigenbaum, Kannan, McGregor, Suri, and Zhang, 2005]
 - 5.83 [McGregor, 2005]
 - 5.58 [Zelke, 2010]
 - 4.91 + ε [Epstein, Levin, Mestre, and Segev, 2010]
 - $4 + \varepsilon$ [Crouch and Stubbs, 2014]
 - $2 + \varepsilon$ [Paz and Schwartzman, 2018]

Previous Results

In the *adversarial* (worst-case order) semi-streaming model:

- $\bullet\,$ For unweighted maximum (b-)matching approximation:
 - 2 approximation using the trivial greedy algorithm
 - $1 + \ln 2 \approx 1.69$ in approximability [Kapralov, 2021]
- For maximum weight matching approximation:
 - 6 [Feigenbaum, Kannan, McGregor, Suri, and Zhang, 2005]
 - 5.83 [McGregor, 2005]
 - 5.58 [Zelke, 2010]
 - 4.91 + ε [Epstein, Levin, Mestre, and Segev, 2010]
 - $4 + \varepsilon$ [Crouch and Stubbs, 2014]
 - $2 + \varepsilon$ [Paz and Schwartzman, 2018]
- For maximum weight *b*-matching approximation:
 - $4 + \varepsilon$ [Crouch and Stubbs, 2014]
 - $3 + \varepsilon$ [Levin and Wajc, 2021]
 - $2 + \varepsilon$ [Huang and Sellier, 2021]

Previous Results

In the *random-order* semi-steaming model:

- For maximum unweighted matching approximation:
 - 1.98 (bipartite), 1.99 (general) [Konrad, Magniez, and Mathieu, 2012]
 - 1.86 (bipartite) [Konrad, 2018]
 - $5/3 \approx 1.67$ (bipartite), $11/6 \approx 1.84$ (general) [Farhadi, Hajiaghayi, Mai, Rao, and Rossi, 2020]
 - $3/2 + \varepsilon$ (general) [Bernstein, 2020]
 - $3/2 \delta$ (general, $\delta \sim 10^{-14}$) [Assadi and Behnezhad, 2021]

Previous Results

In the *random-order* semi-steaming model:

- For maximum unweighted matching approximation:
 - 1.98 (bipartite), 1.99 (general) [Konrad, Magniez, and Mathieu, 2012]
 - 1.86 (bipartite) [Konrad, 2018]
 - $5/3 \approx 1.67$ (bipartite), $11/6 \approx 1.84$ (general) [Farhadi, Hajiaghayi, Mai, Rao, and Rossi, 2020]
 - $3/2 + \varepsilon$ (general) [Bernstein, 2020]
 - $3/2 \delta$ (general, $\delta \sim 10^{-14}$) [Assadi and Behnezhad, 2021]
- For maximum weight matching approximation:
 - 2 δ (general, $\delta \sim 10^{-17})$ [Gamlath, Kale, Mitrovic, and Svensson, 2019]

Previous Results

In the *random-order* semi-steaming model:

- For maximum unweighted matching approximation:
 - 1.98 (bipartite), 1.99 (general) [Konrad, Magniez, and Mathieu, 2012]
 - 1.86 (bipartite) [Konrad, 2018]
 - $5/3 \approx 1.67$ (bipartite), $11/6 \approx 1.84$ (general) [Farhadi, Hajiaghayi, Mai, Rao, and Rossi, 2020]
 - $3/2 + \varepsilon$ (general) [Bernstein, 2020]
 - $3/2 \delta$ (general, $\delta \sim 10^{-14}$) [Assadi and Behnezhad, 2021]
- For maximum weight matching approximation:
 - 2 δ (general, $\delta \sim 10^{-17})$ [Gamlath, Kale, Mitrovic, and Svensson, 2019]
- For maximum (weight) *b*-matching approximation:
 - no previous result in the random-order setting

Previous Results and Our Contribution

Maximization problem	Adversarial	Random-Order
Unweighted matching	2	$3/2 - \delta$
Weighted matching	2	$2-\delta$
Unweighted <i>b</i> -matching	2	2
Weighted <i>b</i> -matching	2	2

Table: Approximation ratios for adversarial and random-order streams (we assume weights are small integers in $\{1, \ldots, W\}$).

Previous Results and Our Contribution

Maximization problem	Adversarial	Random-Order
Unweighted matching	2	$3/2 - \delta$
Weighted matching	2	$2-\delta (2-1/(2W)+\varepsilon)$
Unweighted <i>b</i> -matching	2	$2 (3/2 + \varepsilon)$
Weighted b -matching	2	$2 (2 - 1/(2W) + \varepsilon)$

Table: Approximation ratios for adversarial and random-order streams (we assume weights are small integers in $\{1, \ldots, W\}$). Comparison with our results.

Our Contribution

We generalize the result of [Bernstein, 2020] $(3/2 + \varepsilon)$ approximation for simple matchings in random-order streams) to integer-weighted *b*-matchings.

Theorem

We can extract with high probability from a randomly-ordered stream of edges having integer weights in $\{1, \ldots, W\}$ a weighted b-matching with an approximation ratio of

$$2 - \frac{1}{2W} + \varepsilon$$

using $O(\max(|M_G|, n) \cdot poly(\log(m), W, 1/\varepsilon))$ memory.

Technique

To obtain algorithms for random-order streams:

• for unweighted simple matching [Bernstein, 2020]: use of Edge-Degree Constrained Subgraph

Technique

To obtain algorithms for random-order streams:

- for unweighted simple matching [Bernstein, 2020]: use of Edge-Degree Constrained Subgraph
- for integer-weighted *b*-matching [our paper]: introduce a generalization of the *Edge-Degree Constrained Subgraph*

Edge-Degree Constrained Subgraphs

Definition (EDCS [Bernstein and Stein, 2015])

Let G = (V, E) be a graph, and H a subgraph of G. Given any integer parameters $\beta > \beta^-$, we say that H is a (β, β^-) -EDCS of G if H satisfies the following properties:

- For any $(u, v) \in H$, $\deg_H(u) + \deg_H(v) \le \beta$
- **(b)** For any $(u, v) \in G \setminus H$, $\deg_H(u) + \deg_H(v) \ge \beta^-$.



Figure: Example with $\beta = 5$ and $\beta^- = 4$

Edge-Degree Constrained Subgraphs

Simple facts on about EDCSes:

- they always exist (can be built in polynomial time)
- they are of size $O(n \cdot \beta)$

Edge-Degree Constrained Subgraphs

Simple facts on about EDCSes:

- they always exist (can be built in polynomial time)
- they are of size $O(n \cdot \beta)$

Main result [Bernstein and Stein, 2015]:

Theorem (Main theorem about EDCSes)

For β/β^- close enough to 1, a (β, β^-) -EDCS always contain a $3/2 + \varepsilon$ approximation of the maximum cardinality matching.

This bound is tight.

Definition (EDCS)

Let $\beta > \beta^-$, then H is a (β, β^-) -EDCS of G if H satisfies the following properties:

- For any $(u, v) \in H$,

 $\deg_H(u) + \deg_H(v) \le \beta$ $\deg_H(u) + \deg_H(v) \ge \beta^-$

14/28

Some properties we want to have for weighted *b*-matchings:

Definition (EDCS)

Let $\beta > \beta^-$, then H is a (β, β^-) -EDCS of G if H satisfies the following properties:

- For any $(u, v) \in H$,

 $\deg_H(u) + \deg_H(v) \le \beta$ $\deg_H(u) + \deg_H(v) \ge \beta^-$

Some properties we want to have for weighted b-matchings:

• favoring edges having large weights

Definition (EDCS)

Let $\beta > \beta^-$, then H is a (β, β^-) -EDCS of G if H satisfies the following properties:

- For any $(u, v) \in H$,

 $\deg_H(u) + \deg_H(v) \le \beta$ $\deg_H(u) + \deg_H(v) \ge \beta^-$

Some properties we want to have for weighted b-matchings:

- favoring edges having large weights
- homogeneity

Definition (EDCS)

Let $\beta > \beta^-$, then H is a (β, β^-) -EDCS of G if H satisfies the following properties:

- For any $(u, v) \in H$,

 $\deg_H(u) + \deg_H(v) \le \beta$ $\deg_H(u) + \deg_H(v) \ge \beta^-$

14/28

Some properties we want to have for weighted b-matchings:

- favoring edges having large weights
- homogeneity
- taking into account the capacities b_v

Definition (w-b-EDCS)

Let $\beta > \beta^-$, then H is a (β, β^-) -w-b-EDCS of G if H satisfies the following properties:

• For any $(u, v) \in H$, $\deg_H(u) + \deg_H(v) \le \beta \cdot w(u, v)$

(b) For any $(u, v) \in G \setminus H$, $\deg_H(u) + \deg_H(v) \ge \beta^- \cdot w(u, v)$



Figure: Favoring larger edges

Definition (w-b-EDCS)

Let $\beta > \beta^-$, then H is a (β, β^-) -w-b-EDCS of G if H satisfies the following properties:

- For any $(u, v) \in H$, $\mathbf{wdeg}_H(u) + \mathbf{wdeg}_H(v) \le \beta \cdot w(u, v)$
- **(b)** For any $(u, v) \in G \setminus H$, $wdeg_H(u) + wdeg_H(v) \ge \beta^- \cdot w(u, v)$

We want to have an homogeneous definition, so that multiplying all the weights of the graph G by a factor λ does not change the EDCS structure in the graph:

$$w'(u,v) = \lambda \cdot w(u,v)$$

Definition (weighted b-EDCS)

Let $\beta > \beta^-$, then H is a (β, β^-) -w-b-EDCS of G if H satisfies the following properties:

• For any
$$(u, v) \in H$$
, $\frac{\operatorname{wdeg}_H(u)}{b_u} + \frac{\operatorname{wdeg}_H(v)}{b_v} \le \beta \cdot w(u, v)$
• For any $(u, v) \in G \setminus H$, $\frac{\operatorname{wdeg}_H(u)}{b_u} + \frac{\operatorname{wdeg}_H(v)}{b_v} \ge \beta^- \cdot w(u, v)$

We take into account the capacity of a vertex:

• a vertex with a large capacity should be able to have more edges in the w-b-EDCS

Definition (EDCS)

Let $\beta > \beta^-$, then H is a (β, β^-) -EDCS of G if H satisfies the following properties:

- For any $(u, v) \in H$, $\deg_H(u) + \deg_H(v) \le \beta$
- **(b)** For any $(u, v) \in G \setminus H$, $\deg_H(u) + \deg_H(v) \ge \beta^-$

Definition (weighted b-EDCS)

Let $\beta > \beta^-$, then H is a (β, β^-) -w-b-EDCS of G if H satisfies the following properties:

• For any
$$(u, v) \in H$$
, $\frac{\operatorname{wdeg}_H(u)}{b_u} + \frac{\operatorname{wdeg}_H(v)}{b_v} \le \beta \cdot w(u, v)$

(b) For any
$$(u, v) \in G \setminus H$$
, $\frac{\operatorname{wdeg}_H(u)}{b_u} + \frac{\operatorname{wdeg}_H(v)}{b_v} \ge \beta^- \cdot w(u, v)$

Similarly, for weighted b-EDCSes:

- they always exist (can be built in pseudo-polynomial time)
- they are of size $O(|M_G| \cdot \beta)$

Similarly, for weighted b-EDCSes:

- they always exist (can be built in pseudo-polynomial time)
- they are of size $O(|M_G| \cdot \beta)$

Main result:

Theorem (Main theorem about weighted *b*-EDCSes)

For β/β^- close enough to 1, a (β, β^-) -w-b-EDCS always contain a $2 - \frac{1}{2W} + \varepsilon$ approximation of the maximum weight matching.

This bound is tight.

In the following:

- we prove the result for bipartite weighted matchings, as the generalization to weighted *b*-matchings will be a reduction to this case
- we will explain what adaptations of the proof of [Assadi and Bernstein, 2018] is required in that case

Lemma

For β/β^- close enough to 1, a (β, β^-) -w-b-EDCS always contain a $2 - \frac{1}{2W} + \varepsilon$ approximation of the maximum weight matching in a bipartite graph.

Existing Proof for Unweighted Simple Matchings



Figure: Witness set $\overline{A} \cup B$ from Hall's marriage theorem — can also be seen as a vertex cover (picture from [Assadi and Bernstein, 2018])

Existing Proof for Unweighted Simple Matchings



Figure: Witness set $\overline{A} \cup B$ from Hall's marriage theorem — can also be seen as a vertex cover (picture from [Assadi and Bernstein, 2018])

We want to bound $|S| \ge 2(|M_G| - |M_H|)$.

Definition (EDCS)

Let $\beta \ge \beta^- + 1$, then H is a (β, β^-) -EDCS of G if H satisfies the following properties:

- For any $(u, v) \in H$,
- (i) For any $(u, v) \in G \setminus H$,

 $\deg_H(u) + \deg_H(v) \le \beta$ $\deg_H(u) + \deg_H(v) > \beta^-.$

22/28

Consider the edges \tilde{E} incident to S in H.

Definition (EDCS)

Let $\beta \ge \beta^- + 1$, then H is a (β, β^-) -EDCS of G if H satisfies the following properties:

- For any $(u, v) \in H$,
- (i) For any $(u, v) \in G \setminus H$,

 $\deg_H(u) + \deg_H(v) \le \beta$ $\deg_H(u) + \deg_H(v) > \beta^-.$

Consider the edges \tilde{E} incident to S in H. By Property ii:

 $|S| \cdot \beta^-/2 \le |\tilde{E}|$

Definition (EDCS)

Let $\beta \ge \beta^- + 1$, then *H* is a (β, β^-) -EDCS of *G* if *H* satisfies the following properties:

- For any $(u, v) \in H$, $\deg_H(u) + \deg_H(v) \le \beta$

Consider the edges \tilde{E} incident to S in H. By Property ii:

$$|S|\cdot\beta^-/2\leq |\tilde{E}|$$

and using convexity and averaging arguments,

$$|\tilde{E}| \le (\beta - \beta^-/2) \cdot |M_H|$$

Definition (EDCS)

Let $\beta \ge \beta^- + 1$, then *H* is a (β, β^-) -EDCS of *G* if *H* satisfies the following properties:

- For any $(u, v) \in H$, $\deg_H(u) + \deg_H(v) \le \beta$
- **(b)** For any $(u, v) \in G \setminus H$, $\deg_H(u) + \deg_H(v) \ge \beta^-$.

Consider the edges \tilde{E} incident to S in H. By Property ii:

$$|S|\cdot\beta^-/2\leq |\tilde{E}|$$

and using convexity and averaging arguments,

$$|\tilde{E}| \le (\beta - \beta^-/2) \cdot |M_H|$$

hence,

$$|S| \le (2\beta/\beta^- - 1) \cdot |M_H| \le (1 + \varepsilon) \cdot |M_H|$$

Theorem (Kőnig-Egerváry duality theorem)

In any edge-weighted bipartite subgraph H of G, the maximum weight of a matching equals the smallest weight of a w-vertex-cover.

In other words, there exist integers $(\alpha_v)_{v \in V}$ such that:

- $\sum_{v \in V} \alpha_v = w(M_H)$
- for all $(u, v) \in H$, $w(u, v) \le \alpha_u + \alpha_v$

Theorem (Kőnig-Egerváry duality theorem)

In any edge-weighted bipartite subgraph H of G, the maximum weight of a matching equals the smallest weight of a w-vertex-cover.

In other words, there exist integers $(\alpha_v)_{v \in V}$ such that:

- $\sum_{v \in V} \alpha_v = w(M_H)$
- for all $(u, v) \in H$, $w(u, v) \le \alpha_u + \alpha_v$

We now consider the optimal matching M_G in G. The first idea is to use this duality theorem to relate $w(M_G)$ to $w(M_H)$, with a leftover term that will be analyzed in the second part of the proof (just like the set S previously).

We introduce the notion of *good* and *bad* edges:

- $(u, v) \in M_G$ such that $\beta^- \cdot w(u, v) \leq \beta \cdot (\alpha_u + \alpha_v)$ is called a *good edge*; the set of good edges is denoted as M_{qood} ;
- $(u, v) \in M_G$ such that $\beta^- \cdot w(u, v) > \beta \cdot (\alpha_u + \alpha_v)$ is called a *bad edge*; the set of bad edges is denoted as M_{bad} .

Proof for Weighted EDCS

Now, denoting by V_{bad} the set of vertices which are the endpoints of a bad edge and such that $\mathbf{w} \deg_H(u) - \beta \cdot \alpha_u > 0$, we can get

$$\beta^{-} \cdot w(M_G) \le \beta \cdot w(M_H) + \sum_{v \in V_{bad}} (\mathbf{w} \deg_H(v) - \beta \cdot \alpha_v)$$

Proof for Weighted EDCS

Now, denoting by V_{bad} the set of vertices which are the endpoints of a bad edge and such that $\mathbf{w} \deg_H(u) - \beta \cdot \alpha_u > 0$, we can get

$$\beta^{-} \cdot w(M_G) \le \beta \cdot w(M_H) + \sum_{v \in V_{bad}} (\mathbf{w} \deg_H(v) - \beta \cdot \alpha_v)$$

Claim

We have the inequality

$$\sum_{v \in V_{bad}} (\mathbf{w} \deg_H(v) - \beta \cdot \alpha_v) \le \left(\beta + W - \frac{\beta^-}{2W \cdot (1 + \varepsilon/4)}\right) \cdot w(M_H)$$

Proof for Weighted EDCS

Claim

We have the inequality

$$\sum_{v \in V_{bad}} (\mathbf{w} \deg_H(v) - \beta \cdot \alpha_v) \le \left(\beta + W - \frac{\beta^-}{2W \cdot (1 + \varepsilon/4)}\right) \cdot w(M_H)$$

• V_{bad} plays somehow here the role of S

Claim

We have the inequality

$$\sum_{v \in V_{bad}} (\mathbf{w} \deg_H(v) - \beta \cdot \alpha_v) \le \left(\beta + W - \frac{\beta^-}{2W \cdot (1 + \varepsilon/4)}\right) \cdot w(M_H)$$

- V_{bad} plays somehow here the role of S
- however we cannot just use the set of edges in *H* incident to V_{bad} some vertices are partially covered

Claim

We have the inequality

$$\sum_{v \in V_{bad}} (\mathbf{w} \deg_H(v) - \beta \cdot \alpha_v) \le \left(\beta + W - \frac{\beta^-}{2W \cdot (1 + \varepsilon/4)}\right) \cdot w(M_H)$$

- V_{bad} plays somehow here the role of S
- however we cannot just use the set of edges in H incident to V_{bad} — some vertices are partially covered
- we need to build an auxiliary graph H_{bad} made of a well-chosen set of edges \tilde{E}

Claim

We have the inequality

$$\sum_{v \in V_{bad}} (\mathbf{w} \deg_H(v) - \beta \cdot \alpha_v) \le \left(\beta + W - \frac{\beta^-}{2W \cdot (1 + \varepsilon/4)}\right) \cdot w(M_H)$$

- V_{bad} plays somehow here the role of S
- however we cannot just use the set of edges in H incident to V_{bad} — some vertices are partially covered
- we need to build an auxiliary graph H_{bad} made of a well-chosen set of edges \tilde{E}
- then a similar averaging/convexity argument can be used

Lemma

For β/β^- close enough to 1, a (β, β^-) -w-b-EDCS always contain a $2 - \frac{1}{2W} + \varepsilon$ approximation of the maximum weight matching in a bipartite graph.

Theorem

For β/β^- close enough to 1, a (β, β^-) -w-b-EDCS always contain a $2 - \frac{1}{2W} + \varepsilon$ approximation of the maximum weight matching.

Using the probabilistic method and Lovasz Local Lemma, as in [Assadi and Bernstein, 2018].

Conclusion

Definition (weighted b-EDCS)

Let $\beta > \beta^-$, then H is a (β, β^-) -w-b-EDCS of G if H satisfies the following properties:

• For any
$$(u, v) \in H$$
,

For any
$$(u, v) \in G \setminus H$$
, $\frac{\operatorname{wdeg}_H(u)}{h_v} + \frac{\operatorname{wdeg}_H(v)}{h_v} \geq \beta^- \cdot w(u, v)$

 $\frac{\mathbf{w} \mathrm{deg}_H(u)}{\mathbf{v}} + \frac{\mathbf{w} \mathrm{deg}_H(v)}{\mathbf{v}} < \beta \cdot w(u, v)$

Theorem (Main theorem about weighted *b*-EDCSes)

For β/β^- close enough to 1, a (β, β^-) -w-b-EDCS always contain a $2 - \frac{1}{2W} + \varepsilon$ approximation of the maximum weight matching.

Can be used to design an algorithm for random-order streams.