‘math+econ+code’ masterclass on equilibrium transport and matching models in economics

Alfred Galichon (NYU)

Day 4: matching with nontransferable utility

- NTU stable matchings
- The deferred acceptance algorithm
- Lattice structure of NTU stable matchings
- Aggregate NTU stable matchings
- Approximate NTU stable matchings
- Adachi's algorithm

- Gale and Shapley (1962). "College Admissions and the Stability of Marriage." *The American Mathematical Monthly*.
- Knuth (1976). Mariages stables. Presses de l'Université de Montréal.
- Roth, Sotomayor (1990). *Two-Sided Matching A study in Game-Theoretic Modeling and Analysis*. Cambridge.
- Adachi (2000). On a characterization of stable matchings. *Economic Letters*.
- Dagsvik (2000). "Aggregation in Matching Markets." *International Economic Review*.
- Menzel (2015). "Large Matching Markets as Two-Sided Demand Systems." *Econometrica*.
- Galichon, Hsieh (2019). ""Aggregate stable matching with money burning." SSRN id=2887732.."

Section 1

Gale and Shapley's stable marriages

- Consider "men" $i \in \mathcal{I}$ and "women" $j \in \mathcal{J}$. One of each type. If i and j match, then i gets α_{ij} and j gets γ_{ij}. Unmatched agent's utility normalized to zero. Let μ_{ij} be such that

$$\mu_{ij} \in \{0, 1\}, \ \sum_j \mu_{ij} \leq 1 \text{ and } \sum_i \mu_{ij} \leq 1$$

- **Gale-Shapley stable matching**. μ is a Gale-Shapley stable matching (GS-SM) if, when definining $u_i := \sum_{j'} \mu_{ij'} \alpha_{ij'}$ and $v_j := \sum_{i'} \mu_{i'j} \gamma_{i'j}$, the following stability inequalities holds

$$\forall i, j : \max \{u_i - \alpha_{ij}, v_j - \gamma_{ij}\} \geq 0, \ u_i \geq 0, \ v_j \geq 0.$$

- Intuition: if on the contrary, $\max \{u_i - \alpha_{ij}, v_j - \gamma_{ij}\} < 0$, then both i and j would achieve better by matching than what they get with their current partner. They would thus form a *blocking pair*.

For a matching μ, denote u^{μ} and v^{μ} the corresponding payoffs, i.e. $u_i^{\mu} := \sum_{j'} \mu_{ij'} \alpha_{ij'}$ and $v_j^{\mu} := \sum_{i'} \mu_{i'j} \gamma_{i'j}$.
Define a partial order $\leq_{\mathcal{I}}$ on matchings such that $\mu \leq_{\mathcal{I}} \mu'$ if and only if $u_i^{\mu} \leq u_i^{\mu'}$ for all $i \in \mathcal{I}$, and $\mu \leq_{\mathcal{J}} \mu'$ if and only if $u_i^{\mu} \leq u_i^{\mu'}$. The following theorem is due to Conway, but first appeared in Knuth (1976):
Theorem (Lattice structure of stable matchings). One has:
(i) The set of GS-stable matchings is a lattice.
(ii) If μ and μ' are stable matchings, then $\mu \leq_{\mathcal{I}} \mu'$ if and only if $\mu' \leq_{\mathcal{J}} \mu$.

Conway's lemma. One has:
Fact 1: $1\left\{u_i^{\mu} \leq u_i^{\mu'}\right\}\mu_{ij} = 1\left\{v_j^{\mu} \geq v_j^{\mu'}\right\}\mu_{ij}$
Fact 2: $\left\{u_i^{\mu} > u_i^{\mu'}\right\}\mu'_{ij} = 1\left\{v_j^{\mu} < v_j^{\mu'}\right\}\mu'_{ij}$
Proof. To show fact 1, note that $\mu_{ij} > 0$ implies $\max\left\{u_i^{\mu} - \alpha_{ij}, v_j^{\mu} - \gamma_{ij}\right\} \geq 0$ thus $\max\left\{u_i^{\mu} - u_i^{\mu'}, v_j^{\mu} - v_j^{\mu'}\right\} \geq 0$. In particular if $u_i^{\mu} \leq u_i^{\mu'}$, then either $=$ and thus $v_j^{\mu} = v_j^{\mu'}$; or $u_i^{\mu} < u_i^{\mu'}$ which implies $v_j^{\mu} \geq v_j^{\mu'}$. QED. The proof of fact 2 is similar. Indeed, $\mu'_{ij} > 0$ implies $\max\left\{u_i^{\mu'} - \alpha_{ij}, v_j^{\mu'} - \gamma_{ij}\right\} \geq 0$ thus $\max\left\{u_i^{\mu'} - u_i^{\mu}, v_j^{\mu'} - v_j^{\mu}\right\} \geq 0$, hence in conjuction with $u_i^{\mu} > u_i^{\mu'}$, it follows that $v_j^{\mu'} \geq v_j^{\mu}$. But $v_j^{\mu'} = v_j^{\mu}$ would imply $u_i^{\mu} = u_i^{\mu'}$ which would contradict our assumption, hence $v_j^{\mu'} > v_j^{\mu}$.

Proof of (i). Let μ and μ' be two stable matchings. Consider $\mu^{\wedge}$ defined by $\mu^{\wedge}_{ij} = 1\left\{u_i^{\mu} \leq u_i^{\mu'}\right\}\mu_{ij} + 1\left\{u_i^{\mu} > u_i^{\mu'}\right\}\mu'_{ij}$. Then by facts 1 and 2, $u_i^{\mu^{\wedge}} = u_i^{\mu} \wedge u_i^{\mu'}$ and $v_j^{\mu^{\wedge}} = v_j^{\mu} \wedge v_j^{\mu'}$ and $\mu^{\wedge}$ is a stable matching.

Proof of (ii). If $\mu \leq_{\mathcal{I}} \mu'$, then by fact 1, $1\left\{v_j^{\mu} \geq v_j^{\mu'}\right\}\mu_{ij} = 1\left\{u_i^{\mu} \leq u_i^{\mu'}\right\}\mu_{ij} = \mu_{ij}$. Summation over i yields $1\left\{v_j^{\mu} \geq v_j^{\mu'}\right\} = 1$.

Section 2

Deferred acceptance algorithm

- At each step, each man proposes to his favorite woman among those who have not rejected him yet. Each woman who has several offers gets engaged to her favorite man among those who have proposed, and rejects the other offers.
- The algorithm stops when no more offer is rejected. Engaged pairs become married.

Let $\mathcal{A}_t(i) \subseteq \mathcal{J}$ be the set of available women to man $i \in \mathcal{I}$ at step t. Define $\mathcal{P}_t(i)$ as the set of woman to whom man i has proposed to at step t (a singleton or empty); define $\mathcal{P}_t^{-1}(j) = \{i \in \mathcal{I} : j \in \mathcal{P}_t(i)\}$; define $\mathcal{E}_t(j)$ as the set of men i to whom woman j is engaged at the end of step t and $\mathcal{E}_t^{-1}(i)$ accordingly.

- Algorithm (Gale-Shapley).
  - Set $\mathcal{A}_0(i) = \mathcal{J}$. (Initially, all the women are available to all men)
  - At step $t \geq 0$, assume $\mathcal{A}_t(i)$ has been defined and set

$$\begin{cases} \mathcal{P}_t(i) = \arg\max_j \left\{\alpha_{ij} : j \in \mathcal{A}_t(i)\right\} \text{ (men propose)} \\ \mathcal{E}_t(j) = \arg\max_i \left\{\gamma_{ij} : i \in \mathcal{P}_t^{-1}(j)\right\} \text{ (women dispose)} \end{cases}$$

    and update the available offers

$$\mathcal{A}_{t+1}(i) = \mathcal{A}_t(i) \setminus \left(\mathcal{P}_t(i) \setminus \mathcal{E}_t^{-1}(i)\right).$$

  - When $\mathcal{A}_t(i) = \mathcal{A}_{t+1}(i)$, stop.
- In the sequel it will be useful to define u_i^t and v_j^t as the value of the maximization problems above.

Galichon and Hsieh (2019) reformulated the previous algorithm in order to be able to introduce unobserved heterogeneity.
Let $\mu_{ij}^{A,t} := 1\left\{j \in \mathcal{A}_t\left(i\right)\right\}$, $\mu_{ij}^{P,t} := 1\left\{j \in \mathcal{P}_t\left(i\right)\right\}$, and $\mu_{ij}^{E,t} := 1\left\{i \in \mathcal{E}_t\left(j\right)\right\}$.

- Algorithm (Gale-Shapley).
  - Set $\mu_{ij}^{A,0} = 1$. (Initially, all the women are available to all men)
  - At step t, pick

$$\begin{cases} \mu_{ij}^{P,t} \in \arg\max_{\mu_{ij} \geq 0} \left\{ \sum_j \mu_{ij}\alpha_{ij} : \mu_{ij} \leq \mu_{ij}^{A,t}, \sum_{j\in\mathcal{J}} \mu_{ij} \leq 1 \right\} \\ \mu_{ij}^{E,t} \in \arg\max_{\mu_{ij} \geq 0} \left\{ \sum_i \mu_{ij}\gamma_{ij} : \mu_{ij} \leq \mu_{ij}^{P,t}, \sum_{i\in\mathcal{I}} \mu_{ij} \leq 1 \right\} \end{cases}$$

    and update the available offers

$$\mu_{ij}^{A,t+1} = \mu_{ij}^{A,t} - \left(\mu_{ij}^{P,t} - \mu_{ij}^{E,t}\right)$$

  - When $\mu_{ij}^{A,t+1} = \mu_{ij}^{A,t}$, stop.

Theorem. The algorithm converges toward a GS-SM.

Fact 1. $u_i^{t+1} \leq u_i^{t+1}$. This follows from $\mathcal{A}_{t+1}(i) \subseteq \mathcal{A}_t(i)$.
Fact 2. It i is engaged to j at the end of phase t, then i will propose to j at phase $t+1$. Indeed, if i is engaged to j at the end of phase t, then in particular i proposed to j in phase t. Thus $\forall j' \in \mathcal{A}_t(i)$, $\alpha_{ij} \geq \alpha_{ij'}$. In particular as $\mathcal{A}_{t+1}(i) \subseteq \mathcal{A}_t(i)$, we get that $\forall j' \in \mathcal{A}_{t+1}(i)$, $\alpha_{ij} \geq \alpha_{ij'}$, and as $j \in \mathcal{A}_{t+1}(i)$ because $i \in \mathcal{E}_t(j)$, it follows that $j \in \mathcal{P}_{t+1}(i)$.
Fact 3. $v_j^{t+1} \geq v_j^t$. Indeed, $v_j^t = \gamma_{ij}$ where $i \in \mathcal{E}_t(j)$. At step t+1, that i proposed again by the virtues of fact 2; hence $v_j^{t+1} \geq \gamma_{ij} = v_j^t$.
Fact 4. If $\mathcal{A}_{t+1} = \mathcal{A}_t$, then then $\mathcal{E}_t$ defines a stable matching.
Suppose $\mathcal{A}_{t+1} = \mathcal{A}_t$, and E^t is not stable. Then there is a pair (i,j) such that $\alpha_{ij}^t > u_i^t$ and $\gamma_{ij} > v_j^t$. At step t, i has never proposed to j; indeed, if i had ever proposed to j, he would have been rejected at an earlier phase $s < t$, which would contradict $\gamma_{ij} > v_j^t \geq v_j^s$. Hence $j \in \mathcal{A}_t(i)$. However, $\mathcal{A}_{t+1} = \mathcal{A}_t$ means that at step t, i has made an offer that has been accepted. This in conjunction with $j \in \mathcal{A}_t(i)$ implies that $u_i^t > \alpha_{ij}$, a contradiction.

Section 3

Aggregate NTU stable matching

The reference for this section is Galichon and Hsieh (2019):
NTU matching with free disposal. (μ, u, v) is a free-disposal stable matching (FD-SM) if

$$\begin{cases} \forall i, j : \max\left\{u_i - \alpha_{ij}, v_j - \gamma_{ij}\right\} \geq 0,\ \ u_i \geq 0,\ \ v_j \geq 0 \\ \mu_{ij} > 0 \Longrightarrow \max\left\{u_i - \alpha_{ij}, v_j - \gamma_{ij}\right\} = 0 \\ \sum_j \mu_{ij} = 0 \Longrightarrow u_i = 0,\ \ \sum_i \mu_{ij} = 0 \Longrightarrow v_j = 0 \end{cases}$$

- The frontier of the feasible set of utilities achievable by i and j is $\left\{(u, v) : \max\left\{u - \alpha_{ij}, v - \gamma_{ij}\right\} = 0\right\}$. Of this set, only the point $(\alpha_{ij}, \gamma_{ij})$ is efficient.
- It may be awkard to allow for the possibility of not attaining the efficient point (=burning money). But burning money may be induced by competition (rat race, competitive overinvestment, waiting lines, etc).

Proposition. In the setting above:
If μ is a GS-SM, then (μ, U^{μ}, V^{μ}) is an FD-SM.
Convertly, if (μ, u, v) is a FD-SM, then μ is a GS-SM.

The direct implication of (1) is obvious. Let us show the converse of (1). Consider (μ, u, v) a CSM, and assume μ is not an OSM. Then there is a blocking pair, or a blocking individual. In the first case one has

$$\max\left\{U_i^\mu - \alpha_{ij}, V_j^\mu - \gamma_{ij}\right\} < 0$$

Assume $\sum_j \mu_{ij} = 1$. Then let j' be such that $\mu_{ij'} = 1$; we have $U_i^\mu = \alpha_{ij'}$ and $\max\left\{u_i - \alpha_{ij'}, v_{j'} - \gamma_{ij'}\right\} = 0$, hence $u_i \leq \alpha_{ij'} = U_i^\mu$. If on the contrary $\sum_j \mu_{ij} = 0$, then we have $U_i^\mu = 0 = u_i$. Similarly one can show that $v_j \leq V_j^\mu$. Therefore, we have

$$\max\left\{u_i - \alpha_{ij}, v_j - \gamma_{ij}\right\} \leq \max\left\{U_i^\mu - \alpha_{ij}, V_j^\mu - \gamma_{ij}\right\} < 0$$

so the existence of a blocking pair leads to a contradiction. If there is a blocking individual $U_i^\mu < 0$, but in that case a similar logic implies that $u_i \leq U_i^\mu < 0$, a contradiction as well.

- Why bother introducing FD-SMs if they are essentially equivalent to the classical GS-SMs?
- The reason is that FD-SMs allow for a natural notion of *aggregate* decentralized matching, which GS-SMs don't.
- If there are multiple indistinguishable agents, a natural requirement of decentralized equilibrium is to satisfy equal treatment – i.e. that identical individuals should get the same payoffs at equilibrium.

- Assume that there are n_x men's types, $x \in \mathcal{X}$ and m_y women's types, $y \in \mathcal{Y}$. If x and y match, then x gets α_{xy} and y gets γ_{xy}. Unmatched agent's utility normalized to zero. Let μ_{xy} be such that

$$\mu_{xy} \in \mathbb{N}, \ \sum_{y \in \mathcal{Y}} \mu_{xy} \leq n_x \text{ and } \sum_{x \in \mathcal{X}} \mu_{xy} \leq m_y$$

- (μ, u, v) is an aggregate FD-SM if

$$\begin{cases} \forall x, y : \max \{u_x - \alpha_{xy}, v_y - \gamma_{xy}\} \geq 0, \ u_x \geq 0, \ v_y \geq 0 \\ \mu_{xy} > 0 \Longrightarrow \max \{u_x - \alpha_{xy}, v_y - \gamma_{xy}\} = 0 \\ \sum_{y \in \mathcal{Y}} \mu_{xy} = 0 \Longrightarrow u_x = 0, \ \sum_{x \in \mathcal{X}} \mu_{xy} = 0 \Longrightarrow v_y = 0 \end{cases}$$

- Assume that there are 2 identical passengers and 1 driver. The value of being unmatched (for the passengers and the driver alike) is 0. The value of being matched is 1, both for the passengers and driver.
- In a model with prices (Uber model–transferable utility), the price of the ride will be 1, so that the driver's payoff is 2, and both passengers' payoffs is zero. Thus, passengers are indifferent between being matched and unmatched.
- In a classical model without transfers (taxi model–nontransferable utility), there are two stable matchings in each of which the matched passenger gets one, while the unmatched gets zero. Thus in this Gale-Shapley solution, one passenger is happier than the other one.

- However, people don't like to be unhappier than their peers!
  - For example, passengers will fight for the only available taxi...
  - ... or they will wait in line, and the length of the line will make each passenger indifferent between waiting in line and opting out.
  - In both cases, the driver is not better off, but both passengers have destroyed utility so that they are indifferent between being matched or unmatched, and both passengers have the same payoff (i.e., zero) at equilibrium.
- If, on the contrary, there are two drivers and one passengers, the story is reversed: drivers will fight / wait in line, and destroy utility so that both drivers get zero payoff; in this case, the passenger gets surplus one.
- To study these problems, we shall need to develop a theory of *multinomial choice under rationing*.

Section 4

Multinomial choice under rationing

- Consider the problem of allocation under capacity constraints

$$
\begin{aligned}
&\max_{\mu \geq 0} \sum_{y \in \mathcal{Y}} \mu_y \alpha_y \\
&s.t.\ \mu_y \leq \bar{\mu}_y \ \left[\tau^\alpha \geq 0\right] \\
&\qquad \sum_{y \in \mathcal{Y}} \mu_y \leq n \ \left[u \geq 0\right]
\end{aligned}
$$

- This is a linear programming problem; the dual is

$$
\begin{aligned}
&\min_{u \geq 0, \tau_y \geq 0} \left\{ nu + \sum_{y \in \mathcal{Y}} \bar{\mu}_y \tau_y \right\} \\
&s.t.\ u \geq \alpha_y - \tau_y
\end{aligned}
$$

which rewrites

$$
\min_{\tau_y \geq 0} \left\{ n \max_{y \in \mathcal{Y}} \left\{ \alpha_y - \tau_y, 0 \right\} + \sum_{y \in \mathcal{Y}} \bar{\mu}_y \tau_y \right\}.
$$

- **Proposition**. $\mu \geq 0$ is a primal solution if and only $\sum_y \mu_y \leq 1$ and if there is some real number U such that, letting $\mathcal{Y}_0 = Y \cup \{0\}$, $\alpha_0 = 0$, and $\mu_0 = n - \sum_{y \in \mathcal{Y}} \mu_y$, one has for all $y \in \mathcal{Y}_0$

$$\left\{ \begin{array}{l} \alpha_y < U \implies \mu_y = 0 \\ \alpha_y = U \implies \mu_y \in [0, \bar{\mu}_y] \\ \alpha_y = U \implies \mu_y = \bar{\mu}_y \end{array} \right. .$$

- **Notation**. we shall denote $y_K \in \mathcal{Y}_0$ the alternative such that $\alpha_{y_K} = U$.

- There is a simple algorithm to solve the problem. Assume from now on that $\alpha_y \neq \alpha_{y'}$ for $y \neq y'$. Then one can order the set $\mathcal{Y}_0$ so that

$$\alpha_{y_1} > \alpha_{y_2} > ... > \alpha_{y_M}.$$

- Consider the greedy algorithm that consists of:

Algorithm. At step k:
If $n - \sum_{i=1}^{k-1} \mu_{y_i} \leq \bar{\mu}_{y_k}$, then set $\mu_{y_k} = n - \sum_{i=1}^{k-1} \bar{\mu}_{y_i}$, set $K = k$, set $\mu_{y_j} = 0$ for $j > k$ and stop.
Else, set $\mu_{y_k} = \bar{\mu}_{y_k}$. If $k = M$ then set stop; else go to step $k+1$.

- **Corollary**. If all the α_y's are all distinct, then μ is unique.

- Once the primal variable μ has been determined, the dual variables τ and u can simply be obtained. Recall that K is the largest k such that $\mu_{y_k} > 0$. The set Lagrange multipliers is given by:

$$\begin{cases} \tau_{y_k} = \alpha_{y_k} - \alpha_{y_K} \text{ for } k \leq K \\ \tau_{y_k} \in \left[0, \alpha_{y_k} - \alpha_{y_K}\right] \text{ for } k > K \end{cases}$$

and u is given by

$$u = \alpha_{y_K}.$$

- Interpretation:
  - the Lagrange multipliers associated with the chosen options $y_1, ..., y_K$ are set to equate the utility of chosing any of these with the utility to choose the least attractive chosen option y_K. They are uniquely defined.
  - the Lagrange multipliers associated with the nonchosen options $y_{K+1}, ..., y_M$ are set so that these choices are (weakly) dominated by the chosen options. They are not uniquely defined.

- See Galichon and Hsieh (2019). τ_y can be interpreted as a shadow price of the capacity constraint $\mu_y \leq \bar{\mu}_y$. Consider the constrained maximum welfare problem

$$\bar{G}(\alpha, \bar{\mu}) = \max_{\mu \geq 0} \sum_{y \in \mathcal{Y}} \alpha_y \mu_y - G^* (\mu)$$
$$s.t.\ \mu_y \leq \bar{\mu}_y \ \ [\tau_y \geq 0]$$

- Then, classically

$$\bar{G}(\alpha, \bar{\mu}) = G(\alpha - \tau) + \sum_y \bar{\mu}_y \tau_y, \text{ and}$$
$$\partial \bar{G}(\alpha, \bar{\mu}) / \partial U_y = \partial G(\alpha - \tau) / \partial U_y.$$

- A natural measure of the market inefficiency is the total time waited in line: $\sum_y \mu_y \tau_y$. It is lost to the passengers, and not appropriated by the taxi drivers.

- **Theorem 1**. The shadow price vector τ is an antitone function of the vector of number of available offers $\bar{\mu}$.
- This result says that when the constraint becomes tighter ($\bar{\mu}$ decreases), the vector of Lagrange multipliers τ increases.
- **Proof**. $\bar{G}(\alpha, \bar{\mu}) = \min_{\tau \geq 0} \left\{ G(\alpha - \tau) + \sum_y \tau_y \bar{\mu}_y \right\}$, hence $\tau = \arg\max_{\tau \geq 0} \left\{ -G(\alpha - \tau) + \sum_y \tau_y (-\bar{\mu}_y) \right\}$. By Topkis' theorem, τ is an isotone function of $-\bar{\mu}$, hence an antitone function of $\bar{\mu}$.

- In the logit case, we look for $\tau_y \geq 0$ such that

$$\alpha_y - \tau_y = \log \frac{\mu_y}{\mu_0}$$
$$\tau_y > 0 \Longrightarrow \mu_y = \bar{\mu}_y$$

- Thus, the demand is given by $\mu_y = \min\left(\bar{\mu}_y, \mu_0 e^{\alpha_y}\right)$, where μ_0 solves the scalar equation

$$\mu_0 + \sum_{y \in \mathcal{Y}} \min\left(\bar{\mu}_y, \mu_0 e^{\alpha_y}\right) = 1.$$

(very easy to solve for numerically).

Section 5

The aggregate deferred acceptance algorithm

- Algorithm.
  - Let $\mu_{xy}^{A,0} = \min\left(n_x, m_y\right)$.
  - At step t, pick

$$\left\{ \begin{array}{l} \mu_{xy}^{P,t} \in \arg\max_{\mu \in \mathbb{N}^{\mathcal{X}\times\mathcal{Y}}} \left\{ \sum_{xy} \mu_{xy}\alpha_{xy} : \mu_{xy} \leq \mu_{xy}^{A,t-1}, \sum_{y\in\mathcal{Y}} \mu_{xy} \leq n_x \;\; \left[u_x^t\right] \right\} \\ \mu_{xy}^{E,t} \in \arg\max_{\mu \in \mathbb{N}^{\mathcal{X}\times\mathcal{Y}}} \left\{ \sum_{xy} \mu_{xy}\gamma_{xy} : \mu_{xy} \leq \mu_{xy}^{P,t}, \sum_{x\in\mathcal{X}} \mu_{xy} \leq m_y \;\; \left[v_y^t\right] \right\} \end{array} \right.$$

    and update the available offers

$$\mu_{xy}^{A,t} = \mu_{xy}^{A,t-1} - \left(\mu_{xy}^{P,t} - \mu_{xy}^{E,t}\right)$$

  - When $\mu_{xy}^{E,t} = \mu_{xy}^{P,t}$, stop.
- Note that when $n_x = 1$ for all x and $m_y = 1$ for all y, this is *exactly* Gale and Shapley.
- **Theorem**. The algorithm converges in a finite number T of steps and $\left(\mu_{xy}^{E,T}, u_x^T, v_y^T\right)$ is an aggregate FD-SM.

- We show convergence by showing a series of facts.
  - Fact 1: Tentatively accepted offers remain in place at the next period: $\mu^{E,t} \leq \mu^{P,t+1}$.
  - Fact 2: As t grows, $\tau^{\alpha,t}$ weakly increases and $\tau^{\gamma,t}$ weakly decreases.
  - Fact 3: At every step t, $\min\left(\tau_{xy}^{\alpha,t}, \tau_{xy}^{\gamma,t}\right) = 0$.
  - Fact 4: As $t \rightarrow \infty$, $\lim \mu^{P} = \lim \mu^{E} =: \mu$.
- As a result, $\left(\mu_{xy}, \tau_{xy}^{\alpha,t}, \tau_{xy}^{\gamma,t}\right)$ is an equilibrium with non-price rationing.

- Tentatively accepted offers remain proposed at the next period: $\mu^{E,t} \leq \mu^{P,t+1}$.
- **Proof**: By theorem 2, $\mu^{A,t} \leq \mu^{A,t-1}$ implies $\mu^{A,t} - \mu^{P,t+1} \leq \mu^{A,t-1} - \mu^{P,t}$, thus $\mu^{A,t} - \mu^{A,t-1} + \mu^{P,t} \leq \mu^{P,t+1}$. Thus, $\mu^{E,t} \leq \mu^{P,t+1}$.

- As t grows, $\tau^{\alpha,t}$ weakly increases and $\tau^{\gamma,t}$ weakly decreases.
- **Proof**:
  - One has $\mu_{xy}^{A,t-1} \leq \mu_{xy}^{A,t}$, hence $\tau_{xy}^{\alpha,t} \geq \tau_{xy}^{\alpha,t+1}$.
  - To see that $\tau^{\gamma,t+1} \geq \tau^{\gamma,t}$, note that the $\tau^{\gamma,t}$ is still the solution to the constraint choice problem if one replaces $\mu^{P,t}$ by $\mu^{E,t}$, and $\tau^{\gamma,t+1}$ is the solution to the constraint choice problem associated with capacity $\mu^{P,t+1}$. As $\mu^{E,t} \leq \mu^{P,t+1}$, it follows from theorem 1 that $\tau^{\gamma,t+1} \geq \tau^{\gamma,t}$.

- At every step t, $\min\left(\tau_{xy}^{\alpha,t}, \tau_{xy}^{\gamma,t}\right) = 0$.
- **Proof**: $\tau_{xy}^{\gamma,t} > 0$ implies $\tau_{xy}^{\gamma,s} > 0$ for $s \in \{1, ..., t\}$; hence $\mu_{xy}^{P,s} = \mu_{xy}^{E,s}$, hence $\mu_{xy}^{A,t-1} = \mu_{xy}^{A,0} = \min(n_x, m_y)$. Assume $\tau_{xy}^{\alpha,t} > 0$. Then it means that the corresponding constraint is saturated, which means $\mu_{xy}^{P,t} = \mu_{xy}^{E,t-1} = \min(n_x, m_y)$. But $\tau_{xy}^{\alpha,t} > 0$ implies that x proposes to other y', and $\tau_{xy}^{\gamma,t} > 0$ implies that y proposes to other x', which together contradict that $\mu_{xy}^{P,t} = \mu_{xy}^{E,t-1} = \min(n_x, m_y)$.

Fact 4

- As $t \to \infty$, $\lim \mu^P = \lim \mu^E =: \mu$.
- **Proof**: One has $\mu^{A,t-1} - \mu^{A,t} = \mu^{P,t} - \mu^{E,t}$, but as $\mu^{A,t}$ is nonincreasing and bounded, this quantity tends to zero.

See Galichon and Hsieh (2019).

- Step 0. Initialize by
$$\mu_{xy}^{A,0} = n_x.$$
- Step $t \geq 1$.
  - Proposal phase: Passengers make proposals subject to availability constraint:
$$\mu^{P,t} \in \arg\max_{\mu} \left\{ \sum \mu_{xy} \alpha_{xy} - G^* \left(\mu \right) : \mu \leq \mu^{A,t-1} \ \left[\tau^{G,t} \geq 0 \right] \right\}.$$
  - Disposal phase: Taxis pick up their best offers among the proposals:
$$\mu^{E,t} \in \arg\max_{\mu} \left\{ \sum \mu_{xy} \gamma_{xy} - H^* \left(\mu \right) : \mu \leq \mu^{P,t} \ \left[\tau^{H,t} \geq 0 \right] \right\}.$$
  - Update phase: The number of available offers is decreased according to the number of rejected ones
$$\mu^{A,t} = \mu^{A,t-1} - \left(\mu^{P,t} - \mu^{E,t} \right).$$

- We show convergence by showing a series of facts.
  - Fact 1: Tentatively accepted offers remain in place at the next period: $\mu^{E,t} \leq \mu^{P,t+1}$.
  - Fact 2: As t grows, $\tau^{G,t}$ weakly increases and $\tau^{H,t}$ weakly decreases.
  - Fact 3: At every step t, $\min\left(\tau_{xy}^{G,t}, \tau_{xy}^{H,t}\right) = 0$.
  - Fact 4: As $t \to \infty$, $\lim \nabla G\left(\alpha - \tau^{G,t}\right) = \lim \nabla H\left(\gamma - \tau^{H,t}\right) =: \mu$.
- As a result, $\left(\mu_{xy}, \tau_{xy}^{G,t}, \tau_{xy}^{H,t}\right)$ is an equilibrium with non-price rationing.

- Tentatively accepted offers remain proposed at the next period: $\mu^{E,t} \leq \mu^{P,t+1}$.
- **Proof**: By theorem 2, $\mu^{A,t} \leq \mu^{A,t-1}$ implies $\mu^{A,t} - \mu^{P,t+1} \leq \mu^{A,t-1} - \mu^{P,t}$, thus $\mu^{A,t} - \mu^{A,t-1} + \mu^{P,t} \leq \mu^{P,t+1}$. Thus, $\mu^{E,t} \leq \mu^{P,t+1}$.

- As t grows, $\tau^{G,t}$ weakly increases and $\tau^{H,t}$ weakly decreases.
- **Proof**:
  - One has $\mu_{xy}^{A,t-1} \leq \mu_{xy}^{A,t}$, thus as ∇G^{*} is isotone, $\nabla G^{*}\left(\mu^{A,t-1}\right) \leq \nabla G^{*}\left(\mu^{A,t}\right)$, hence $\alpha_{xy} - \tau_{xy}^{G,t-1} \leq \alpha_{xy} - \tau_{xy}^{G,t}$.
  - To see that $\tau^{H,t} \geq \tau^{H,t-1}$, note that

$$\tau_{xy}^{H,t} = \partial H\left(\gamma, \mu^{E,t}\right) / \partial \bar{\mu}_{xy}$$

$$\tau_{xy}^{H,t+1} = \partial H\left(\gamma, \mu^{P,t+1}\right) / \partial \bar{\mu}_{xy}$$

and $\mu^{E,t} \leq \mu^{P,t+1}$ along with the fact that $\partial H(\gamma, \bar{\mu}) / \partial \bar{\mu}_{xy}$ is antitone in $\bar{\mu}$ (Theorem 1) allows to conclude.

- At every step t, $\min\left(\tau_{xy}^{G,t}, \tau_{xy}^{H,t}\right) = 0$.
- **Proof**: $\tau_{xy}^{H,t} > 0$ implies $\tau_{xy}^{H,s} > 0$ for $s \in \{1, ..., t\}$; hence $\mu_{xy}^{P,s} = \mu_{xy}^{E,s}$, hence $\mu_{xy}^{A,t-1} = \mu_{xy}^{A,0} = n_x$. Assume $\tau_{xy}^{G,t} > 0$. Then it means that the corresponding constraint is saturated, which means $\mu_{xy}^{P,t} = \mu_{xy}^{E,t-1} = n_x$, a contradiction.

- As $t \to \infty$, $\lim \nabla G\left(\alpha - \tau^{G,t}\right) = \lim \nabla H\left(\gamma - \tau^{H,t}\right) =: \mu$.
- **Proof**: One has $\mu^{A,t-1} - \mu^{A,t} = \mu^{P,t} - \mu^{D,t} = \nabla G\left(\alpha - \tau^{G,t}\right) - \nabla H\left(\gamma - \tau^{H,t}\right)$, but as $\mu^{A,t}$ is nonincreasing and bounded, this quantity tends to zero. Further, $\tau^{G,t}$ and $\tau^{H,t}$ converge monotonically, which shows that $\lim_t \nabla G\left(\alpha - \tau^{G,t}\right) = \lim_t \nabla H\left(\gamma - \tau^{H,t}\right)$.

- Assume $\alpha_{ij} = \alpha_{x_i y_j} + \varepsilon_{iy}$ and $\gamma_{ij} = \gamma_{x_i y_j} + \eta_{xj}$ where ε and η iid Gumbel. The aggregate stability equations write

$$\begin{cases} \sum \mu_{ij} + \mu_{i0} = 1 \text{ and } \sum \mu_{ij} + \mu_{0j} = 1 \\ \max\left(u_i - \alpha_{x_i y_j} - \varepsilon_{i y_j}, v_j - \gamma_{x_i y_j} - \eta_{x_i j}\right) \geq 0 \text{ with equality if } \mu_{ij} > 0 \\ u_i \geq \varepsilon_{i0} \text{ with equality if } \mu_{i0} > 0 \\ v_j \geq \eta_{0j} \text{ with equality if } \mu_{0j} > 0 \end{cases}$$

- Take the third equation, and take the minimum over i such that $x_i = x$ and j such that $y_j = y$. Let $U_{xy} = \min_{i: x_i = x} \left\{ u_i - \varepsilon_{iy} \right\}$ and $V_{xy} = \min_{j: y_j = y} \left\{ v_j - \gamma_{xy} \right\}$. One has then

$$\begin{cases} \sum_{y \in \mathcal{Y}} \mu_{xy} + \mu_{x0} = n_x \text{ and } \sum_{x \in \mathcal{X}} \mu_{xy} + \mu_{0y} = m_y \\ \max\left(U_{xy} - \alpha_{xy}, V_{xy} - \gamma_{xy}\right) \geq 0 \text{ with equality if } \mu_{xy} > 0 \\ U_{x0} \geq 0 \text{ with equality if } \mu_{x0} > 0 \\ V_{0y} \geq 0 \text{ with equality if } \mu_{0y} > 0 \end{cases}$$

and $u_i = \max_{y \in \mathcal{Y}} \left\{ U_{xy} + \varepsilon_{iy}, \varepsilon_{i0} \right\}$ and $v_j = \max_{x \in \mathcal{X}} \left\{ V_{xy} + \eta_{xj}, \eta_{0j} \right\}$.

- Therefore in the logit model, $U_{xy} = \ln \mu_{xy} / \mu_{x0}$ and $V_{xy} = \ln \mu_{xy} / \mu_{0y}$. Thus one has existence and uniqueness of an equilibrium, and

$$\mu_{xy} = \min\left(\mu_{x0} e^{\alpha_{xy}}, \mu_{0y} e^{\gamma_{xy}}\right). \tag{1}$$

where

$$\begin{cases} \mu_{x0} + \sum_{y \in \mathcal{Y}} \min\left(\mu_{x0} e^{\alpha_{xy}}, \mu_{0y} e^{\gamma_{xy}}\right) = n_x \\ \mu_{0y} + \sum_{x \in \mathcal{X}} \min\left(\mu_{x0} e^{\alpha_{xy}}, \mu_{0y} e^{\gamma_{xy}}\right) = m_y \end{cases}$$

and this system can be efficiently solved by coordinate updates.

- Computationally, scales extremely well – easily parallelizable.

Section 6

Adachi's algorithm

The reference for this section is Adachi (2000).

- We keep assuming all preferences are strict, and note that

$$\begin{cases} u_i = \max\left\{\max_j \left\{\alpha_{ij} : \gamma_{ij} \geq v_j\right\}, 0\right\} \\ v_j = \max\left\{\max_i \left\{\gamma_{ij} : \alpha_{ij} \geq u_i\right\}, 0\right\} \end{cases} \tag{2}$$

  where $\{j \in \mathcal{J} : \gamma_{ij} \geq v_j\}$ is i's consideration set: it is the set of partners that would offer i some utility amount greater than u_i.

- Note that the map defined by (2) is not isotone. Indeed, $\max_j \left\{\alpha_{ij} : \gamma_{ij} \geq v_j\right\}$ is actually nonincreasing in v_j (higher v_j's means smaller consideration set for j). Hence, change the sign of u_i take $p = (u, -v)$ and $\kappa_{ij} = -\gamma_{ij}$ so to define

$$\begin{cases} p_i' = \max_{j \in \mathcal{J}} \left\{\alpha_{ij} : p_j \geq \kappa_{ij}, 0\right\}, i \in \mathcal{I} \\ p_j' = \min_{i \in \mathcal{I}} \left\{\kappa_{ij} : \alpha_{ij} \geq p_i, 0\right\}, j \in \mathcal{I} \end{cases}$$

  and we see that the map T defined by $p' = F(p)$ above is isotone. Further, the image of T is contained in the set of p such that $\max_j \alpha_{ij} \geq p_i \geq 0$, and $-\max_i \gamma_{ij} \leq p_j \leq 0$.

One has:

- **Theorem (Adachi)**. Assume $\alpha_{ij} \neq \alpha_{ij'} \neq 0$ for any $j \neq j'$ and $\gamma_{ij} \neq \gamma_{i'j} \neq 0$ for any $i \neq i'$. Then:
(i) If μ is a GS-SM, then defining $u_i = \sum_j \mu_{ij}\alpha_{ij}$ and $v_j = \sum_i \mu_{ij}\gamma_{ij}$, implies that u and v satisfy

$$\begin{cases} u_i = \max_{j:\gamma_{ij}\geq v_j} \{\alpha_{ij}, 0\} \\ v_j = \max_{i:\alpha_{ij}\geq u_i} \{\gamma_{ij}, 0\} \end{cases} \tag{3}$$

(ii) Conversely, if (u, v) satisfy (3), then letting $\mu_{ij} = 1\{u_i = \alpha_{ij}\} 1\{v_j = \gamma_{ij}\}$, it follows that μ is a GS-SM.

- In other words, taking $p_i = u_i$ and $p_j = -v_j$, we have that p is such that $e(p) = 0$, where

$$\begin{cases} e_i(p) := p_i - \max_{j\in J_0} \{\alpha_{ij} : p_j \geq \kappa_{ij}\} \\ e_j(p) := p_j - \min_{i\in I_0} \{\kappa_{ij} : p_i \leq \alpha_{ij}\} \end{cases} \tag{4}$$

μ is a Gale-Shapley stable matching (GS-SM) if, when definining $u_i := \sum_{j'} \mu_{ij'} \alpha_{ij'}$ and $v_j := \sum_{i'} \mu_{i'j} \gamma_{i'j}$, the following stability inequalities holds

$$\forall i, j : \max\{u_i - \alpha_{ij}, v_j - \gamma_{ij}\} \geq 0, \;\; u_i \geq 0, \;\; v_j \geq 0.$$

Proof. Direct implication. If μ is a GS-SM, then $u_i \geq 0$; assume i is assigned to j^* under μ. Then $\gamma_{ij^*} = v_{j^*}$, and $u_i = \alpha_{ij^*}$, thus $u_i \geq \max_{j:\gamma_{ij} \geq v_j} \{\alpha_{ij}, 0\}$. But $\gamma_{ij} \geq v_j$ and $j \neq j^*$ implies $\gamma_{ij} > v_j$, which in turns implies $u_i < \alpha_{ij}$; as a result $u_i = \max_{j:\gamma_{ij} \geq v_j} \{\alpha_{ij}, 0\}$. The case when i is unassigned is treated in a similar fashion. A similar argument shows $v_j = \max_{i:\alpha_{ij} \geq u_i} \{\gamma_{ij}, 0\}$.
Conversely, assume that (u, v) satisfy (3), and let $\mu_{ij} = 1\{u_i = \alpha_{ij}\} 1\{v_j = \gamma_{ij}\}$. One has $u_i = \alpha_{ij}$ if and only if $v_j = \gamma_{ij}$; indeed, $u_i > \alpha_{ij}$ implies $\gamma_{ij} < v_j$, and $v_j > \alpha_{ij}$ implies $u_i < \alpha_{ij}$. Therefore, $u_i = \alpha_{ij}$ implies $v_j = \gamma_{ij}$, but by symmetry, equivalence holds. This implies also that $u_i > \alpha_{ij}$ if and only if $v_j < \gamma_{ij}$. As a result, $u_i := \sum_{j'} \mu_{ij'} \alpha_{ij'}$ and $v_j := \sum_{i'} \mu_{i'j} \gamma_{i'j}$, and clearly $u_i \geq 0$ and $v_j \geq 0$, while $\max\{u_i - \alpha_{ij}, v_j - \gamma_{ij}\} \geq 0$.

- Start by $p_i^0 = \max_{i \in I_0} \{\alpha_{ij}\}$ and $p_j^0 = 0$, which are such that $e(p) \geq 0$.
- Adachi reintepresets as a blockwise Gauss-Seidel algorithm

$$\begin{cases} p_i^{t+1} : e_i\left(p_i^{t+1}, \left(p_j^t\right)_j\right) = 0, \\ p_j^{t+1} : e_j\left(p_j^{t+1}, \left(p_i^{t+1}\right)_i\right) = 0. \end{cases}$$

- Gale and Shapley interpets as

$$\begin{cases} p_i^{t+1} = decr_i\left(p_i^t\right) \text{ if } e_i\left(p_i^{t+1}, \left(p_j^t\right)_j\right) > 0, \; = p_i^t \text{ if } e_i\left(p_i^{t+1}, \left(p_j^t\right)_j\right) = 0, \\ p_j^{t+1} : e_j\left(p_j^{t+1}, \left(p_i^{t+1}\right)_i\right) = 0 \end{cases}$$

where $decr_i(p) = \max_j \{\alpha_{ij} : \alpha_{ij} < p\}$ is the "next value below p_i" in terms of the α_{ij}'s.

- Similarities:
  - In both algorithms, start with the highest possible $u_i = \max_j \alpha_{ij}$.
  - u_i keeps decreasing and v_j keeps increasing
- Differences:
  - In Gale and Shapley, μ^E_{ij} is a feasible matching, even though it is not stable until convergence. Indeed, any newly engaged pair at step $t+1$ is a blocking pair for the matching at step t. On the contrary, in Adachi, we retain stability throughout, but none of the "pre-matchings" involved is a feasible matching before the last step.
  - In GS, the primitive object at the start of each iteration is μ^A_{ij}, i.e. who is no longer available to whom – an object of size $\mathcal{I} \times \mathcal{J}$; in A, it is u_i, an object of size $\mathcal{I}$.

Basis of Dagsvik-Menzel's model approximations:

1. replace $\max\{a, b\}$ by the smooth-max $\sigma \log\left(e^{a/\sigma} + e^{b/\sigma}\right)$, and get

$$\begin{cases} \exp(u_i/\sigma) = 1 + \sum_j 1\{\gamma_{ij} \geq v_j\} \exp(\alpha_{ij}/\sigma) \\ \exp(v_j/\sigma) = 1 + \sum_i 1\{\alpha_{ij} \geq u_i\} \exp(\gamma_{ij}/\sigma) \end{cases}$$

2. replace further $1\{a \geq b\}$ by its smoothed version $e^{a/\sigma} / \left(e^{a/\sigma} + e^{b/\sigma}\right)$, and get

$$\begin{cases} \exp(u_i/\sigma) = 1 + \sum_j \left(\exp\left(\frac{v_j - \gamma_{ij}}{\sigma}\right) + 1\right)^{-1} \exp(\alpha_{ij}/\sigma) \\ \exp(v_j/\sigma) = 1 + \sum_i \left(\exp\left(\frac{u_i - \alpha_{ij}}{\sigma}\right) + 1\right)^{-1} \exp(\gamma_{ij}/\sigma) \end{cases}$$

3. Assume $u_i \gg \alpha_{ij}$ and $v_j \gg \gamma_{ij}$, to get

$$\begin{cases} \exp(u_i/\sigma) = 1 + \sum_j \exp\left(\frac{\gamma_{ij} - v_j}{\sigma}\right) \exp(\alpha_{ij}/\sigma) \\ \exp(v_j/\sigma) = 1 + \sum_i \exp\left(\frac{\alpha_{ij} - u_i}{\sigma}\right) \exp(\gamma_{ij}/\sigma) \end{cases}$$

- Therefore, we get

$$\begin{cases} \exp(-u_i/\sigma) + \sum_j \exp\left(\frac{\alpha_{ij}+\gamma_{ij}-u_i-v_j}{\sigma}\right) = 1 \\ \exp(-v_j/\sigma) + \sum_i \exp\left(\frac{\alpha_{ij}+\gamma_{ij}-u_i-v_j}{\sigma}\right) = 1 \end{cases}$$

- In block 7, we will see how to solve for this type of systems.
- The fraction of pairs ij resp single i and single j is then

$$\begin{cases} \mu_{ij} = \exp\left(\frac{\alpha_{ij}+\gamma_{ij}-u_i-v_j}{\sigma}\right) \\ \mu_{i0} = \exp(-u_i/\sigma) \\ \mu_{0j} = \exp(-v_j/\sigma) \end{cases}$$

- In the case of matching with transferable utility, a vector (μ, u, v) is a stable outcome if $\sum_j \mu_{ij} + \mu_{i0} = 1$, $\sum_i \mu_{ij} + \mu_{0j} = 1$
  $u_i + v_j \geq \alpha_{ij} + \gamma_{ij} =: \Phi_{ij}$, $u_i \geq 0$, $v_j \geq 0$, and
  $\mu_{ij} > 0 \implies u_i + v_j = \Phi_{ij}$, $\mu_{i0} > 0 \implies u_i = 0$, $\mu_{0j} > 0 \implies v_j = 0$.
- In particular, if (μ, u, v) is stable, then

$$u_i = \max_j \{\Phi_{ij} - v_j, 0\}$$

$$v_j = \max_i \{\Phi_{ij} - u_i, 0\}$$

  which with the change of variables $p = (u, -v)$, rewrites $p = F_0(p)$, where

$$\begin{cases} F_0(p)_i = \max_{j \in \mathcal{J}} \{\Phi_{ij} + p_j, 0\}, & i \in \mathcal{I} \\ F(p)_j = \min_{i \in \mathcal{I}} \{-\Phi_{ij} + p_i, 0\}, & j \in \mathcal{J} \end{cases}$$

- Thus, one could hope to characterize TU stable matchings as a fixed point of the previous operator, à-la Adachi. Why?

- In the TU case, stability as a solution concept extends naturally to *aggregate matchings*, when there are several identical copies of an individual in the population. Assume there are n_i copies of $i \in \mathcal{I}$ and m_j of $j \in \mathcal{J}$. A vector (μ, u, v) is a stable outcome if $\sum_j \mu_{ij} + \mu_{i0} = n_i$, $\sum_i \mu_{ij} + \mu_{0j} = m_j$, $u_i + v_j \geq \alpha_{ij} + \gamma_{ij} =: \Phi_{ij}$, $u_i \geq 0$, $v_j \geq 0$, and $\mu_{ij} > 0 \implies u_i + v_j = \Phi_{ij}$, $\mu_{i0} > 0 \implies u_i = 0$, $\mu_{0j} > 0 \implies v_j = 0$.
- Assume (μ, u, v) is a stable outcome associated with (n, m), and (μ', u', v') associated with (n', m'), then both (u, v) and (u', v') are fixed point of F.
- We need to remember that for later: simple fixed point characterization of stable matchings (Adachi) are difficult to extend to aggregate matchings. We'll see another illustration of this fact when we'll talk about one-to-many matching algorithms.

- However, one can write a model that provides an approximation of a solution to the TU stable matching problem up to arbitrary precision. This model (actually Dagsvik's model) has:

$$\left\{ \begin{array}{c} n_i \exp\left(\frac{-p_i}{T}\right) + \sum_j n_i m_j \exp\left(\frac{\Phi_{ij} - p_i + p_j}{T}\right) = n_i \\ m_j \exp\left(\frac{p_j}{T}\right) + \sum_i n_i m_j \exp\left(\frac{\Phi_{ij} - p_i + p_j}{T}\right) = m_j \end{array} \right.$$

- Therefore, the analog of Adachi is now $p' = F_T(p)$, where

$$\left\{ \begin{array}{c} F_T(p)_i = T \log\left(1 + \sum_j \exp\left(\frac{\Phi_{ij} + T \log m_j + p_j}{T}\right)\right) \\ F_T(p)_j = -T \log\left(1 + \sum_i \exp\left(\frac{\Phi_{ij} + T \log n_i + p_i}{T}\right)\right) \end{array} \right.$$

- Denote p_T that fixed point, which can easily be shown to be unique by BGH. Then one has

$$F_T\left(p_T\right) = p_T$$

- Hence if we let $T \to 0$, we see that $F_0\left(p_0\right) = p_0$, but if F_0 has multiple fixed points, the regularization will select one. This is understood by the following remark:
  Lemma. p is a fixed point of F_T if and only it is a minimizer of

$$\min_{u,v}\left\{\begin{array}{c}\sum_i n_i p_i - \sum_j m_j p_j \\ +T\sum_{ij} n_i m_j \exp\left(\frac{\Phi_{ij}-p_i+p_j}{T}\right) \\ +T\sum_i n_i \exp\left(\frac{-p_i}{T}\right) + T\sum_j m_j \exp\left(\frac{p_j}{T}\right)\end{array}\right\}.$$