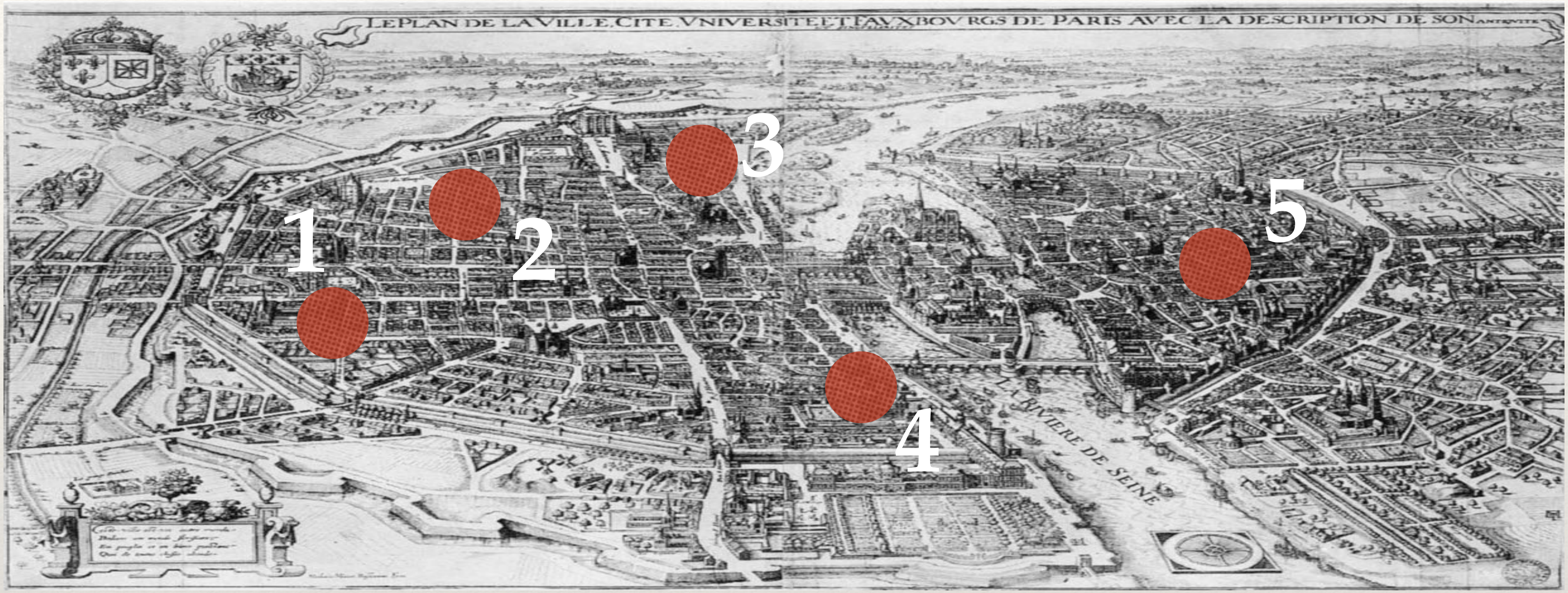
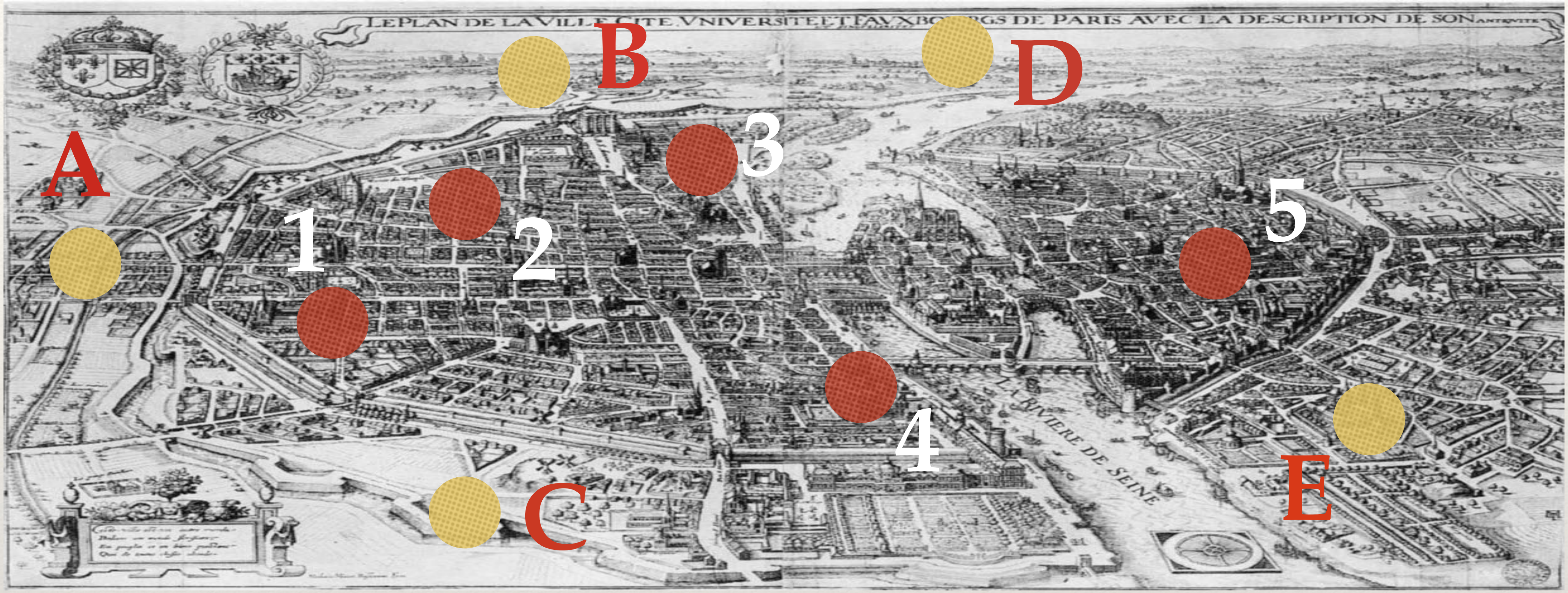
Finite Temperature Optimal Assignment

Joint work with: Henri Orland, CEA Saclay

Assignment

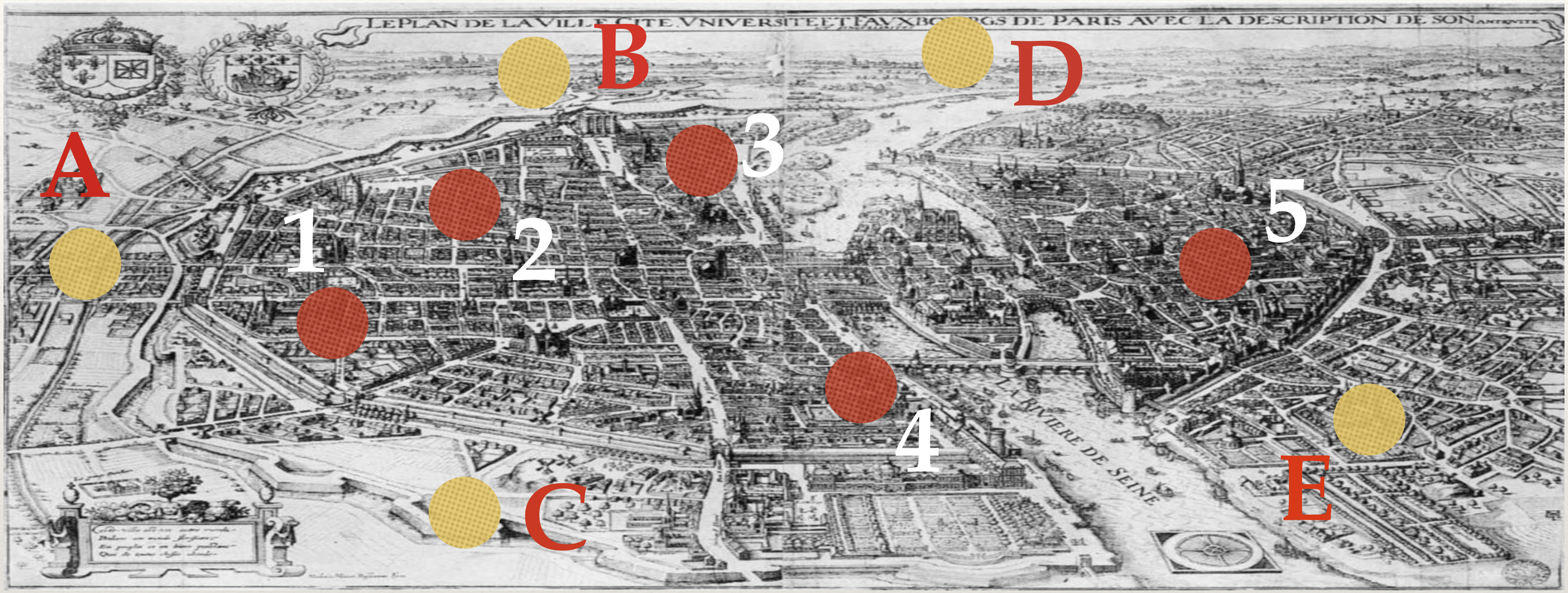

LE PLAN DE LA VILLE, CITE, VNIVERSITE ET FAVXBOVRGS DE PARIS AVEC LA DESCRIPTION DE SON
1
2
3
4
5
LA RIVIERE DE SEINE

Assignment



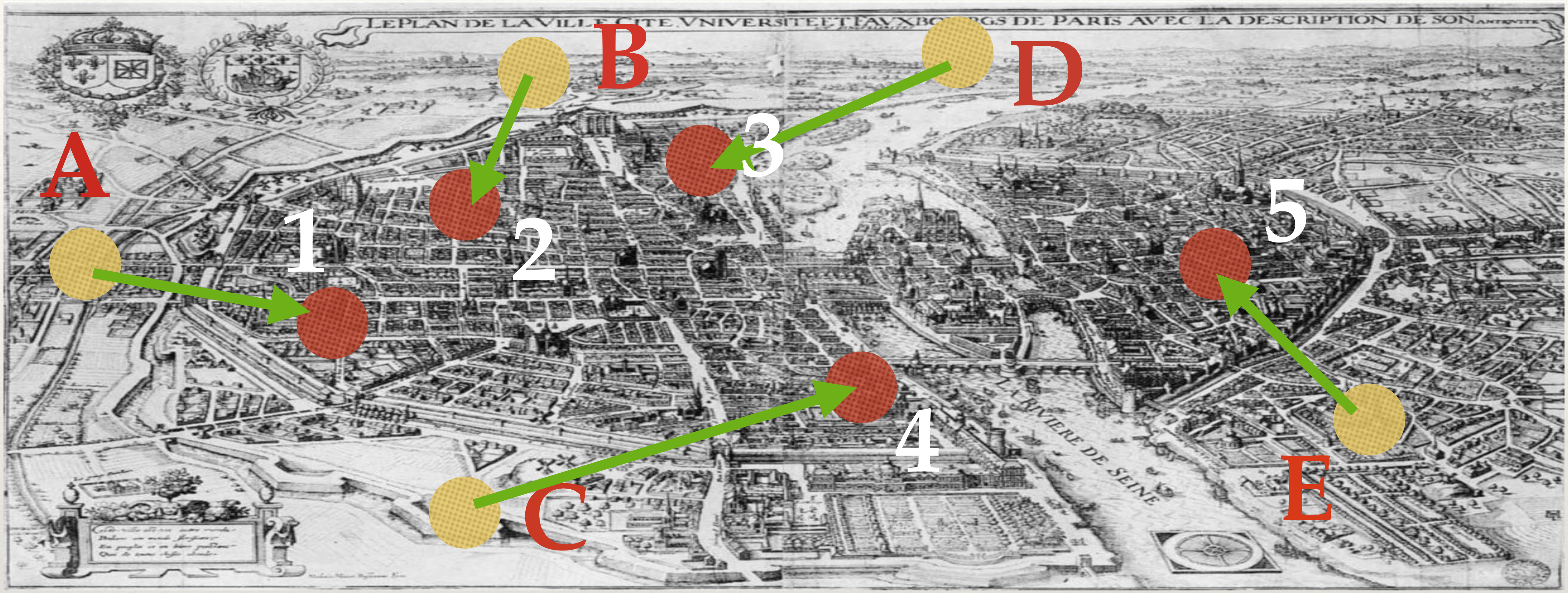
LE PLAN DE LA VILLE CITE VNIVERSITE ET FAVXBO...GS DE PARIS AVEC LA DESCRIPTION DE SON
A
B
C
D
E
1
2
3
4
5
LA RIVIERE DE SEINE

Assignment



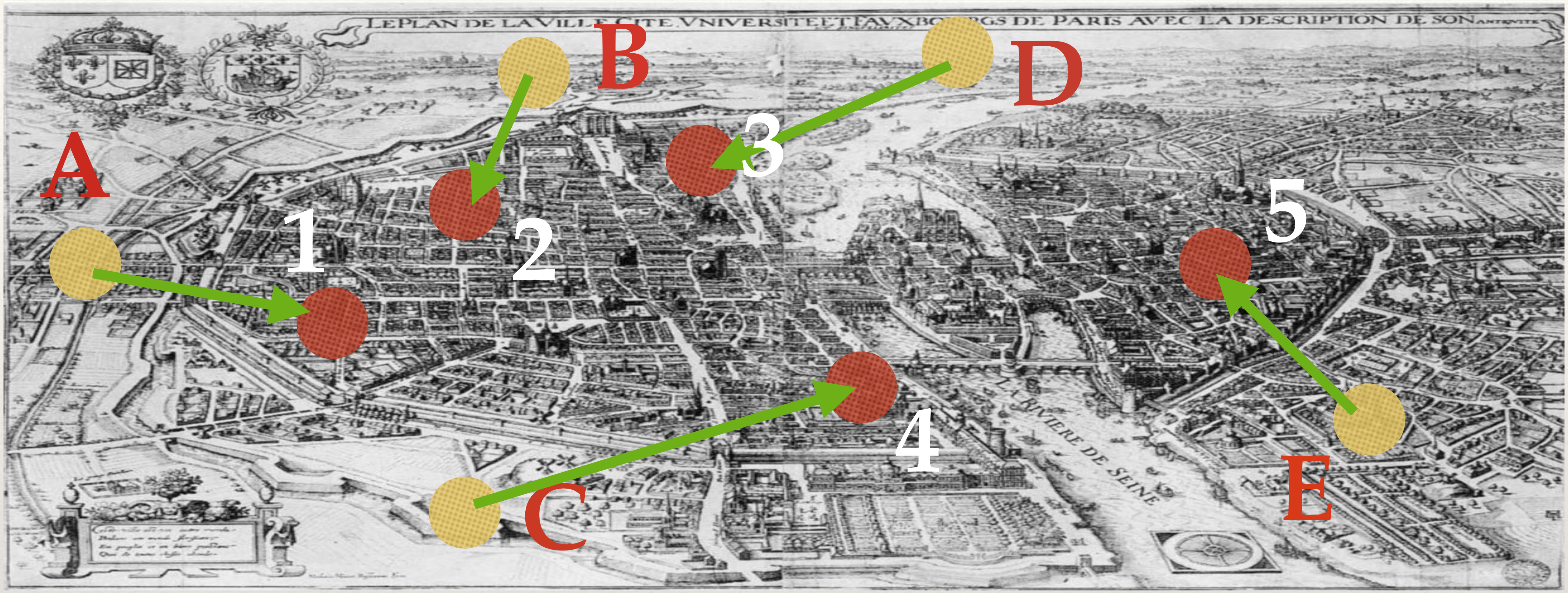
	A	B	C	D	E
1	2	3	2.5	5	8
2	3	1.5	3	4	7.5
3	4	1.5	3.5	2	4.5
4	6	4	3	3	4
5	9	6.5	7	3	2

Assignment



	A	B	C	D	E
1	**2**	3	2.5	5	8
2	3	**1.5**	3	4	7.5
3	4	1.5	3.5	**2**	4.5
4	6	4	**3**	3	4
5	9	6.5	7	3	**2**

Assignment



	A	B	C	D	E
1	**2**	3	2.5	5	8
2	3	**1.5**	3	4	7.5
3	4	1.5	3.5	**2**	4.5
4	6	4	**3**	3	4
5	9	6.5	7	3	**2**

$\odot$

	A	B	C	D	E
1	1	0	0	0	0
2	0	1	0	0	0
3	0	0	0	1	0
4	0	0	1	0	0
5	0	0	0	0	1

= 10.5

Assignment

Input:

$Set \quad S_1 : \{F_1, \ldots, F_N\}$

$Set \quad S_2 : \{B_1, \ldots, B_N\}$

$Cost \quad matrix \quad C(i,j)$

Optimal transport plan:

$$G^* = \underset{G}{\operatorname{argmin}} U(G) = \underset{G}{\operatorname{argmin}} \sum_i \sum_j C(i,j) G(i,j)$$

Optimal energy:

$$U^* = U(G^*)$$

Under the constraints:

$$Gij \in \{0,1\}$$

$$\sum_j G(i,j) = 1 \quad \forall i$$

$$\sum_i G(i,j) = 1 \quad \forall j$$

Assignment

Solving the assignment problem:

1) **Brute Force**

2) **Hungarian algorithm**

3) **Other?**

Assignment

Solving the assignment problem:

1) **Brute Force**

   O(N!)

2) **Hungarian algorithm**

3) **Other?**

Assignment

Solving the assignment problem:

1) **Brute Force**

$O(N!)$

2) **Hungarian algorithm**

$O(N^3)$ Serial

3) **Other?**

Hungarian algorithm

Lemma:

Let S_1 and S_2 be two sets of points with the same cardinality *N* and let *C* be a real-valued cost matrix between S_1 and S_2.
Let *G* be an assignment matrix between S_1 and S_2 that satisfies the constraints on row sum and column sum, namely *G* is a permutation matrix, and let *U(G,C)* be the total cost associated with *G*.

Let ***a*** and ***b*** be any two real-valued vectors of size *N*, and let $D_{\mathbf{a},\mathbf{b}}$ be the matrix defined as

$$D_{\mathbf{a},\mathbf{b}}(k,l) = C(k,l) + a(k) + b(l)$$

Then,

$$U(G, D_{\mathbf{a},\mathbf{b}}) = U(G,C) + \sum_k a(k) + \sum_l b(l)$$

Hungarian algorithm

Lemma:

Let S_1 and S_2 be two sets of points with the same cardinality *N* and let *C* be a real-valued cost matrix between S_1 and S_2.
Let *G* be an assignment matrix between S_1 and S_2 that satisfies the constraints on row sum and column sum, namely *G* is a permutation matrix, and let *U(G,C)* be the total cost associated with *G*.

Let ***a*** and ***b*** be any two real-valued vectors of size *N*, and let $D_{\mathbf{a},\mathbf{b}}$ be the matrix defined as

$$D_{\mathbf{a},\mathbf{b}}(k,l) = C(k,l) + a(k) + b(l)$$

Then,

$$U(G, D_{\mathbf{a},\mathbf{b}}) = U(G,C) + \sum_k a(k) + \sum_l b(l)$$

→ *Shift row values / column values to get a "simpler" cost matrix with one 0 per row and per column*

Hungarian algorithm

1) Rows:

	A	B	C	D	E
1	2	3	2.5	5	8
2	3	1.5	3	4	7.5
3	4	1.5	3.5	2	4.5
4	6	4	3	3	4
5	9	6.5	7	3	2

Find minimum value on each row …

Hungarian algorithm

1) Rows:

	A	B	C	D	E
1	2	3	2.5	5	8
2	3	1.5	3	4	7.5
3	4	1.5	3.5	2	4.5
4	6	4	3	3	4
5	9	6.5	7	3	2

→

	A	B	C	D	E
1	0	1	0.5	3	6
2	1.5	0	1.5	2.5	6
3	2.5	0	2	0.5	3
4	3	1	0	0	1
5	7	4.5	5	1	0

Find minimum value on each row …
and remove from the whole row

Hungarian algorithm

2) Cols:

	A	B	C	D	E
1	0	1	0.5	3	6
2	1.5	0	1.5	2.5	6
3	2.5	0	2	0.5	3
4	3	1	0	0	1
5	7	4.5	5	1	0

Find minimum value on each col …

Hungarian algorithm

2) Cols:

	A	B	C	D	E
1	0	1	0.5	3	6
2	1.5	0	1.5	2.5	6
3	2.5	0	2	0.5	3
4	3	1	0	0	1
5	7	4.5	5	1	0

	A	B	C	D	E
1	0	1	0.5	3	6
2	1.5	0	1.5	2.5	6
3	2.5	0	2	0.5	3
4	3	1	0	0	1
5	7	4.5	5	1	0

Find minimum value on each column …
and remove from the whole column

Hungarian algorithm

3) Draw lines:

	A	B	C	D	E
1	0	1	0.5	3	6
2	1.5	0	1.5	2.5	6
3	2.5	0	2	0.5	3
4	3	1	0	0	1
5	7	4.5	5	1	0

Draw minimum number of lines to cover all 0s

Hungarian algorithm

3) Draw lines:

	A	B	C	D	E
1	0	1	0.5	3	6
2	1.5	0	1.5	2.5	6
3	2.5	0	2	0.5	3
4	3	1	0	0	1
5	7	4.5	5	1	0

	A	B	C	D	E
1	0	1	0.5	3	6
2	1.5	0	1.5	2.5	6
3	2.5	0	2	0.5	3
4	*3*	*1*	0	0	1
5	*7*	*4.5*	5	1	0

Draw minimum number of lines to cover all 0s

Hungarian algorithm

4) Correct:

	A	B	C	D	E
1	0	1	0.5	3	6
2	1.5	0	1.5	2.5	6
3	2.5	0	2	0.5	3
4	3	1	0	0	1
5	7	4.5	5	1	0

Find minimum value within non covered values …

Hungarian algorithm

4) Correct:

	A	B	C	D	E
1	0	1	0.5	3	6
2	1.5	0	1.5	2.5	6
3	2.5	0	2	0.5	3
4	*3*	*1*	0	0	1
5	*7*	*4.5*	5	1	0

	A	B	C	D	E
1	0	1	0	2.5	5.5
2	1.5	0	1	2	5.5
3	2.5	0	1.5	0	2.5
4	*3.5*	*1.5*	0	0	1
5	*7.5*	*5*	5	1	0

Find minimum value within non covered values …

- *Remove from all non covered value*
- *Add to cells that are covered by two lines*

Hungarian algorithm

3) Draw lines:

	A	B	C	D	E
1	0	1	0	2.5	5.5
2	1.5	0	1	2	5.5
3	2.5	0	1.5	0	2.5
4	3.5	1.5	0	0	1
5	7.5	5	5	1	0

Draw minimum number of lines to cover all 0s

Hungarian algorithm

3) Draw lines:

	A	B	C	D	E
1	0	1	0	2.5	5.5
2	1.5	0	1	2	5.5
3	2.5	0	1.5	0	2.5
4	3.5	1.5	0	0	1
5	7.5	5	5	1	0

	A	B	C	D	E
1	0	1	0	2.5	5.5
2	1.5	0	1	2	5.5
3	2.5	0	1.5	0	2.5
4	3.5	1.5	0	0	1
5	7.5	5	5	1	0

Draw minimum number of lines to cover all 0s

If # of lines == # rows / columns then stop!!

Hungarian algorithm

5) Assign:

	A	B	C	D	E
1	0	1	0	2.5	5.5
2	1.5	0	1	2	5.5
3	2.5	0	1.5	0	2.5
4	3.5	1.5	0	0	1
5	7.5	5	5	1	0

Hungarian algorithm

1) *Rows*

2) *Cols*

3) *Draw lines*

4) *Correct*

5) *Assign:*

Hungarian algorithm

1) *Rows* **$O(N^2)$**

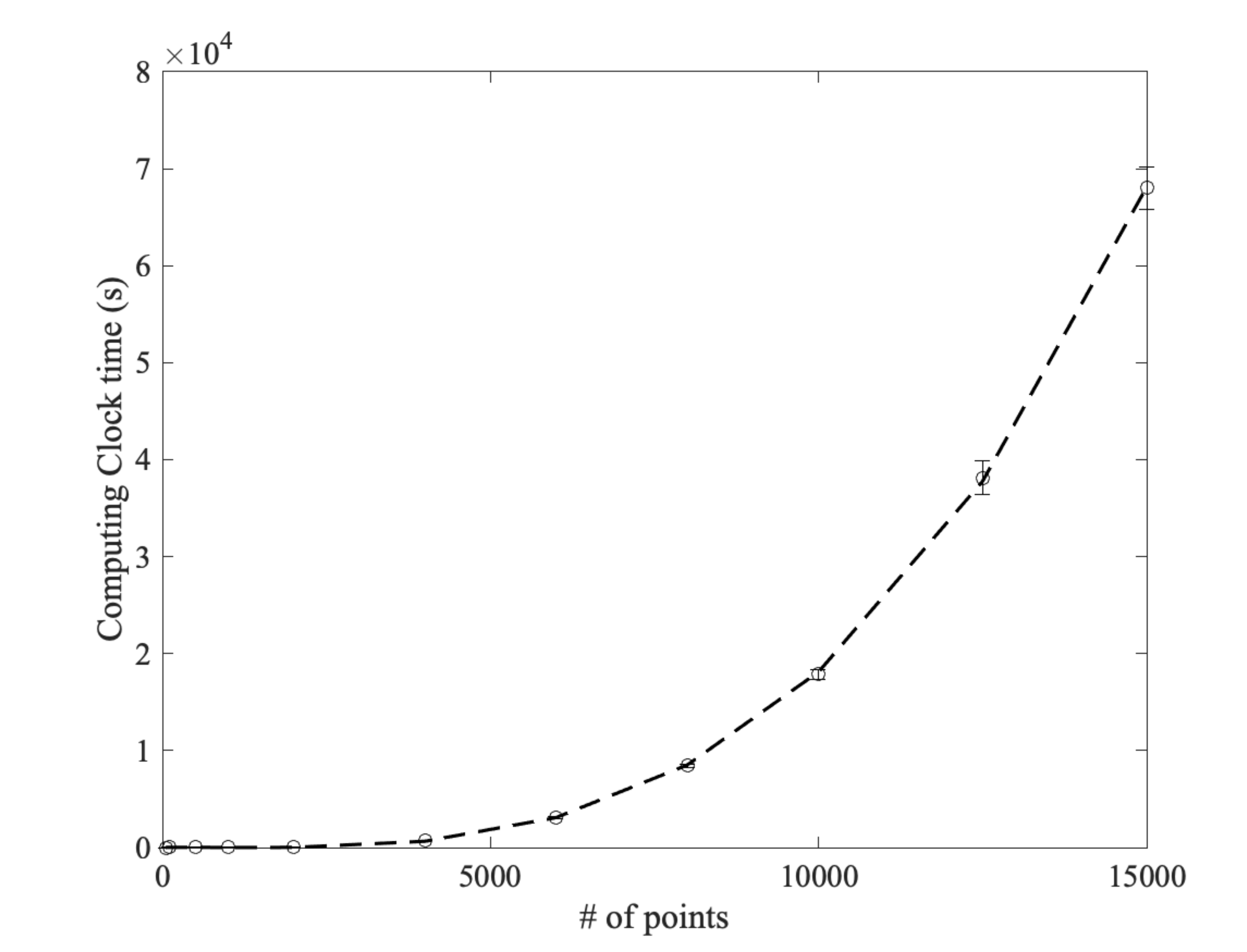
2) *Cols* **$O(N^2)$**

3) *Draw lines*

$O(N^3)$

4) *Correct*

5) *Assign:* **$O(N)$**



Assignment

Solving the assignment problem:

1) **Brute Force**

   $O(N!)$

2) **Hungarian algorithm**

   $O(N^3)$ Serial

3) **Other?**

   $O(N^3)$ Parallel

Assignment

Input:

$Set \quad S_1 : \{F_1, \ldots, F_N\}$

$Set \quad S_2 : \{B_1, \ldots, B_N\}$

$Cost \quad matrix \quad C(i,j)$

Optimal transport plan:

$$G^* = \underset{G}{\text{argmin}}\, U(G) = \underset{G}{\text{argmin}} \sum_i \sum_j C(i,j) G(i,j)$$

Optimal energy:

$$U^* = U(G^*)$$

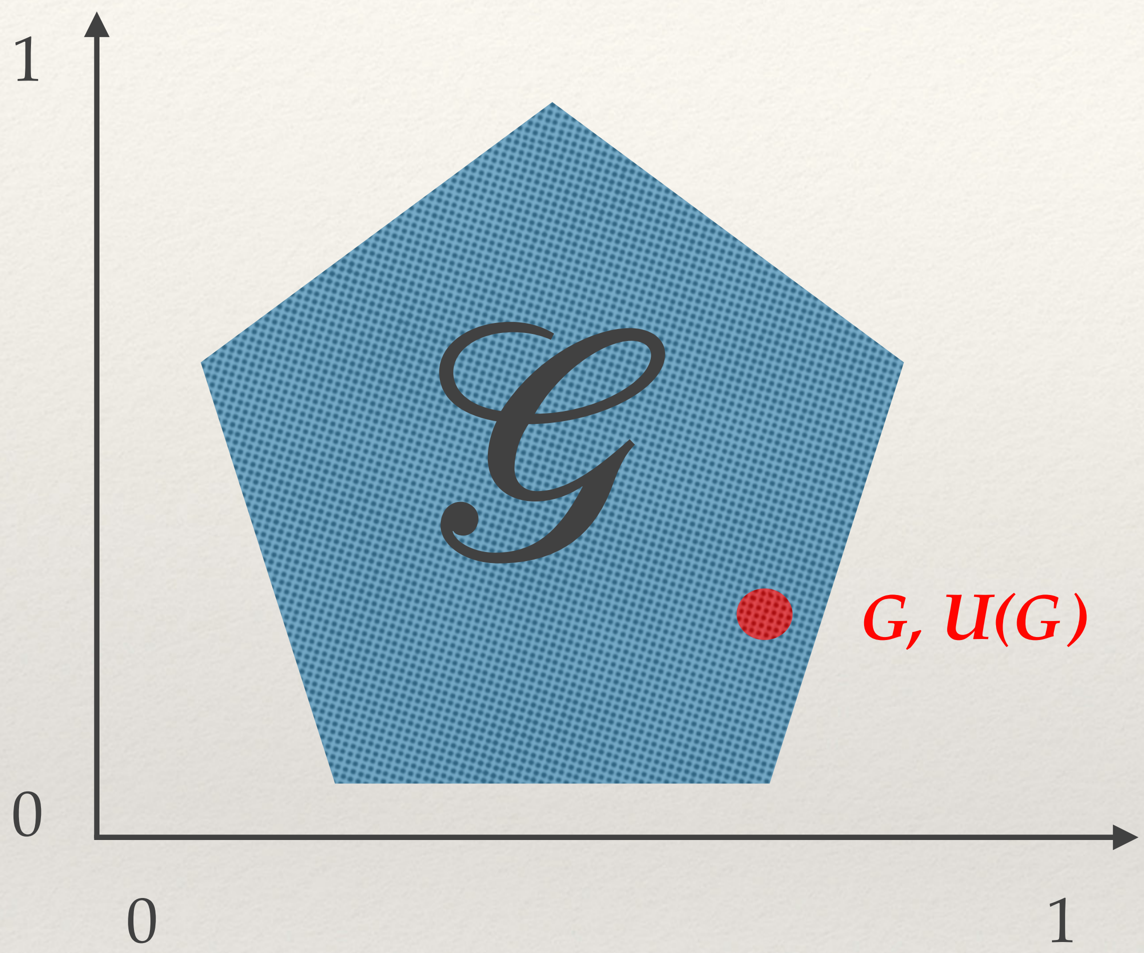
Under the constraints:

$$Gij \in \{0,1\}$$

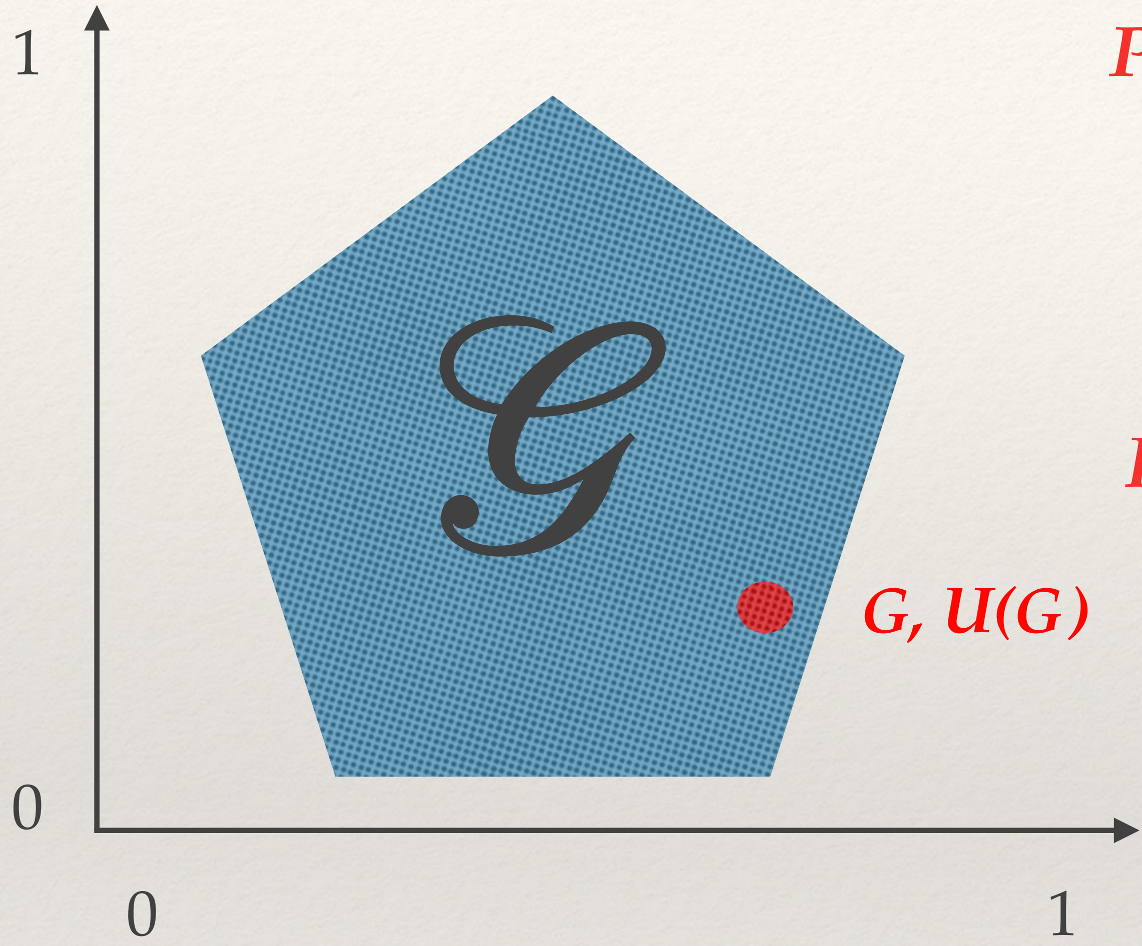
$$\sum_j G(i,j) = 1 \quad \forall i$$

$$\sum_i G(i,j) = 1 \quad \forall j$$

Finite temperature assignment



Finite temperature assignment



Statistical physics: $\left(\beta = \frac{1}{k_B T}\right)$

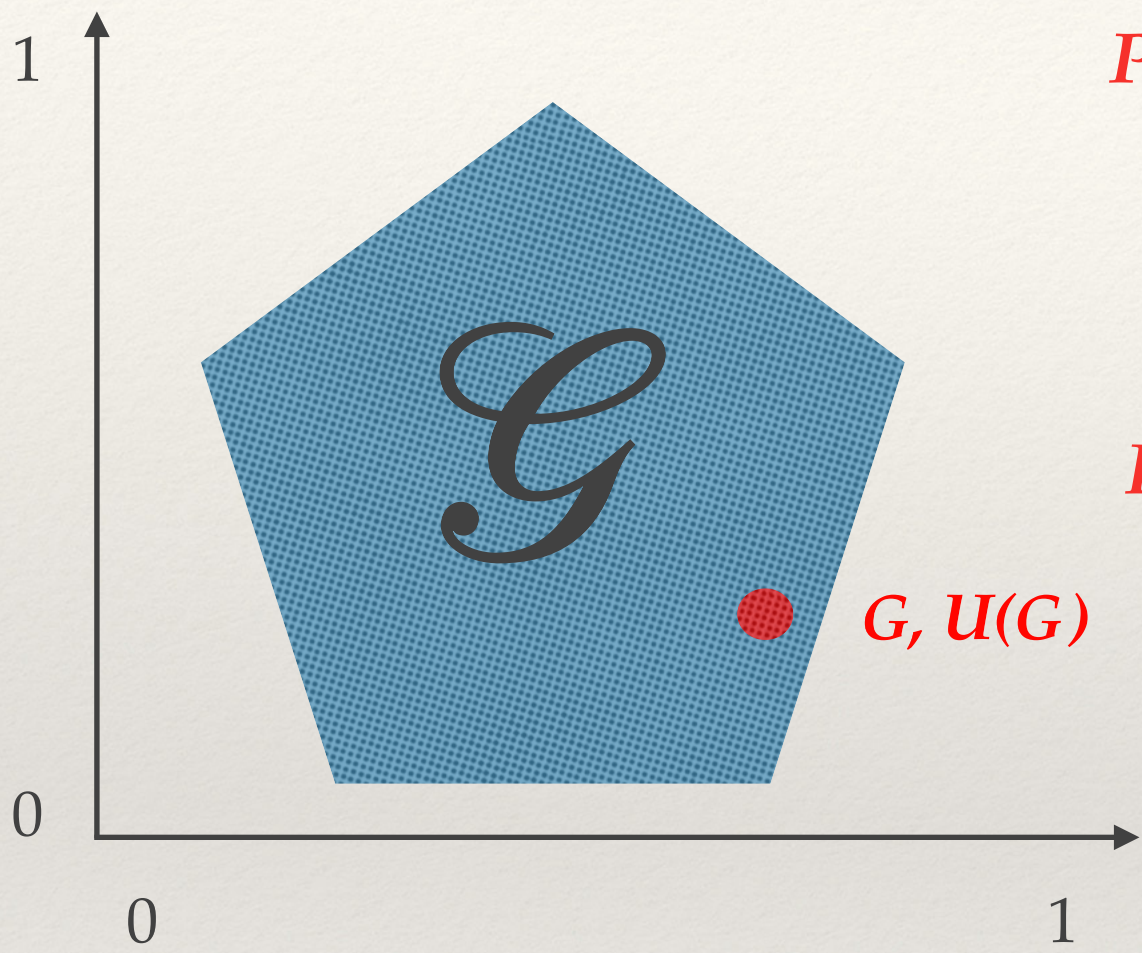
Partition function

$$Z_\beta(S_1, S_2) = \int_{G \in \mathcal{G}} e^{-\beta U(G)}$$

Free energy

$$\mathcal{F}_\beta(S_1, S_2) = -\frac{1}{\beta} \ln(Z_\beta(S_1, S_2))$$

Finite temperature assignment



Statistical physics: $\left(\beta = \frac{1}{k_B T}\right)$

Partition function

$$Z_\beta(S_1, S_2) = \int_{G \in \mathcal{G}} e^{-\beta U(G)}$$

Free energy

$$\mathcal{F}_\beta(S_1, S_2) = -\frac{1}{\beta} \ln(Z_\beta(S_1, S_2))$$

Probability of a state:

$$P(G) = \frac{e^{-\beta U(G)}}{Z_\beta(S_1, S_2)}$$

Minimizing U is equivalent to maximizing P(G)

Finite temperature assignment

$$Z_\beta(S_1, S_2) = \int_{G \in \mathcal{G}} e^{-\beta U(G)}$$

Finite temperature assignment

$$Z_\beta(S_1, S_2) = \int_{G \in \mathcal{G}} e^{-\beta U(G)}$$

$$Z_\beta(S_1, S_2) = \sum_{G(k,l) \in \{0,1\}} e^{-\beta \sum_k \sum_l C(k,l) G(k,l)}$$

$$G(k, l) \in \{0,1\}$$

Finite temperature assignment

$$Z_\beta(S_1, S_2) = \int_{G \in \mathcal{G}} e^{-\beta U(G)}$$

$$Z_\beta(S_1, S_2) = \sum_{G(k,l) \in \{0,1\}} e^{-\beta \sum_k \sum_l C(k,l) G(k,l)} \prod_k \delta\left(\sum_l G(k,l) - 1 \right)$$

$G(k, l) \in \{0,1\}$

$$\sum_l G(k, l) = 1 \quad \forall k$$

Finite temperature assignment

$$Z_\beta(S_1, S_2) = \int_{G \in \mathcal{G}} e^{-\beta U(G)}$$

$$Z_\beta(S_1, S_2) = \sum_{G(k,l) \in \{0,1\}} e^{-\beta \sum_k \sum_l C(k,l) G(k,l)} \prod_k \delta\left(\sum_l G(k,l) - 1 \right) \prod_l \delta\left(\sum_k G(k,l) - 1 \right)$$

$G(k,l) \in \{0,1\}$

$\sum_l G(k,l) = 1 \quad \forall k$

$\sum_k G(k,l) = 1 \quad \forall l$

Finite temperature assignment

Using Fourier, we can represent a delta function as an integral of an exponential,

$$\delta(x) = \frac{1}{2\pi}\int e^{-ixt}dt$$

$$Z_\beta(S_1, S_2) = \sum_{G(k,l)\in\{0,1\}} e^{-\beta\sum_{kl} C(k,l)G(k,l)} \times$$

$$\int_{-\infty}^{+\infty}\prod_k d\lambda(k) e^{-\beta\sum_{k,l}\lambda(k)G(k,l)+\beta\sum_k\lambda(k)} \times \int_{-\infty}^{+\infty}\prod_l d\mu(l) e^{-\beta\sum_{k,l}\mu(l)G(k,l)+\beta\sum_l\mu(l)}$$

$$Z_\beta(S_1, S_2) = \int_{-\infty}^{+\infty}\prod_k d\lambda(k)\int_{-\infty}^{+\infty}\prod_l d\mu(l) e^{\beta\left(\sum_k\lambda(k)+\sum_l\mu(l)\right)}$$

$$\sum_{G(k,l)\in\{0,1\}} e^{-\beta G(k,l)(C(k,l)+\lambda(k)+\mu(l))}$$

Finite temperature assignment

$$Z_\beta(S_1, S_2) = \int_{-\infty}^{+\infty} \prod_k d\lambda(k) \int_{-\infty}^{+\infty} \prod_l d\mu(l) e^{\beta\left(\sum_k \lambda(k) + \sum_l \mu(l)\right)}$$
$$\sum_{G(k,l)\in\{0,1\}} e^{-\beta G(k,l)(C(k,l)+\lambda(k)+\mu(l))}$$

$$Z_\beta(S_1, S_2) = \int_{-\infty}^{+\infty} \prod_k d\lambda(k) \int_{-\infty}^{+\infty} \prod_l d\mu(l) e^{\beta\left(\sum_k \lambda(k) + \sum_l \mu(l)\right) + \sum_{k,l} \ln\left(1 + e^{-\beta(C(k,l)+\lambda(k)+\mu(l))}\right)}$$

Finite temperature assignment

$$Z_\beta(S_1, S_2) = \int_{G \in \mathcal{G}} e^{-\beta U(G)}$$

$$Z_\beta(S_1, S_2) = \int_{-\infty}^{+\infty} \prod_k d\lambda(k) \int_{-\infty}^{+\infty} \prod_l d\mu(l) e^{-\beta F_\beta(\boldsymbol{\lambda}, \boldsymbol{\mu})}$$

where

$$F_\beta\left(\boldsymbol{\lambda}, \boldsymbol{\mu}\right) = -\left(\sum_k \lambda(k) + \sum_l \mu_l\right) - \frac{1}{\beta} \sum_{kl} \ln\left(1 + e^{-\beta\left(C(k,l) + \lambda(k) + \mu(l)\right)}\right)$$

Finite temperature assignment

$$F_\beta(\boldsymbol{\lambda}, \boldsymbol{\mu}) = -\left(\sum_k \lambda(k) + \sum_l \mu_l\right) - \frac{1}{\beta}\sum_{kl} \ln\left(1 + e^{-\beta\left(C(k,l) + \lambda(k) + \mu(l)\right)}\right)$$

Theorems:

1) The Hessian of the effective free energy $F_\beta(\lambda, \mu)$ is negative semi-definite with 2N -1 negative eigenvalues and one zero eigenvalue.

2) The eigenvector corresponding to the zero eigenvalue is (1,...,1, -1, ...-1) (with N 1s, and N -1s)

3) Setting one of the parameters $\lambda(k)$ or $\mu(l)$ as zero, the free energy function on this restricted parameter space is strictly concave.

Finite temperature assignment

Saddle Point Approximation:

$$\frac{\delta F_\beta(\lambda,\mu)}{\delta\lambda} = 0 \qquad \frac{\delta F_\beta(\lambda,\mu)}{\delta\mu} = 0$$

Finite temperature assignment

Saddle Point Approximation:

$$\frac{\delta F_\beta(\lambda,\mu)}{\delta\lambda} = 0 \qquad \frac{\delta F_\beta(\lambda,\mu)}{\delta\mu} = 0$$

$$\begin{cases} \bar{G}(k,l) = \phi(\beta(C(k,l) + \lambda(k) + \mu(l))) \\ \sum_l \bar{G}(k,l) = 1 \quad \forall k \\ \sum_k \bar{G}(k,l) = 1 \quad \forall l \end{cases}$$

where $\phi(x) = \dfrac{1}{e^x + 1}$

Finite temperature assignment

Saddle Point Approximation:

$$\frac{\delta F_\beta(\lambda,\mu)}{\delta\lambda} = 0 \qquad \frac{\delta F_\beta(\lambda,\mu)}{\delta\mu} = 0$$

$$\begin{cases} \bar{G}(k,l) = \phi(\beta(C(k,l) + \lambda(k) + \mu(l))) \\ \sum_l \bar{G}(k,l) = 1 \quad \forall k \\ \sum_k \bar{G}(k,l) = 1 \quad \forall l \end{cases}$$

where $\phi(x) = \dfrac{1}{e^x + 1}$

λ^{MF} and μ^{MF}

$$G^{MF} \qquad F_\beta^{MF} = F(\lambda^{MF}, \mu^{MF}) \qquad U_\beta^{MF} = \sum_k \sum_l C(k,l) G^{MF}(k,l)$$

Finite temperature assignment

$$F_\beta\left(\boldsymbol{\lambda},\boldsymbol{\mu}\right) = -\left(\sum_k \lambda(k) + \sum_l \mu_l\right) - \frac{1}{\beta}\sum_{kl} \ln\left(1 + e^{-\beta\left(C(k,l) + \lambda(k) + \mu(l)\right)}\right)$$

$$F_\beta\left(\boldsymbol{\lambda},\boldsymbol{\mu}\right) = U_\beta(G) - TS_\beta(G)$$
$$+\sum_k \lambda_k\left(\sum_l G(k,l) - 1\right) + \sum_l \mu_l\left(\sum_k G(k,l) - 1\right)$$

where we have defined the entropy S as

$$S_\beta(G) = -\sum_{kl} G(k,l)\ln(G(k,l)) + (1 - G(k,l))\ln(1 - G(k,l))$$

Finite temperature assignment

Theorem:

1) $$\frac{dF_\beta^{MF}}{d\beta} \geq 0 \quad \text{and} \quad \frac{dU_\beta^{MF}}{d\beta} \leq 0$$

2) $$U^* - \frac{A(N)}{\beta} \leq F^{MF}(\beta) \leq U^*$$

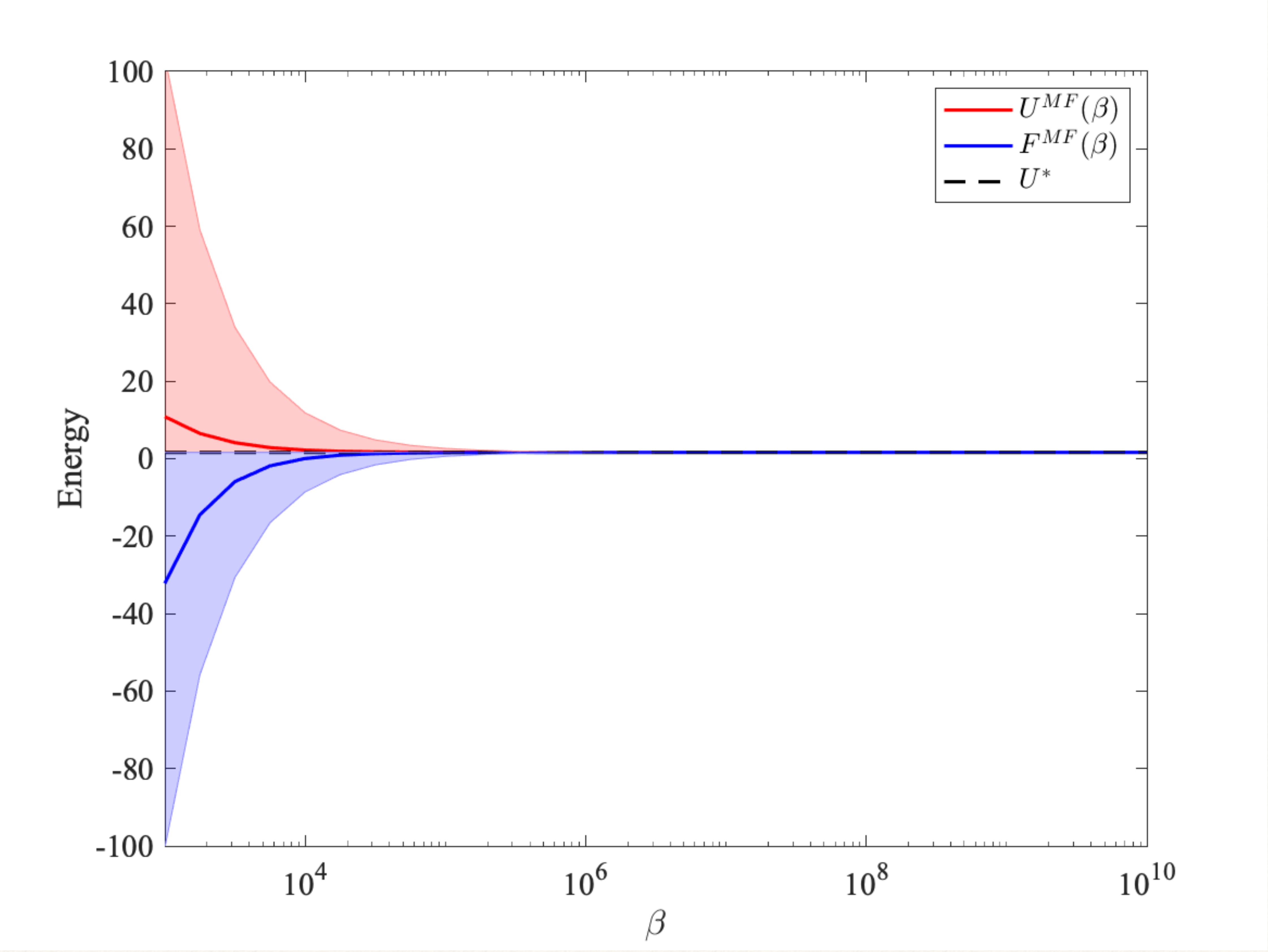
3) $$U^* \leq U^{MF}(\beta) \leq U^* + \frac{A(N)}{\beta}$$

4) $$\lim_{\beta \to \infty} F_\beta^{MF}(S_1, S_2) = U^*$$

5) $$\lim_{\beta \to \infty} U_\beta^{MF}(S_1, S_2) = U^*$$

$$A(N) = N\left(N\log(N) - (N-1)\log(N-1)\right)$$

Finite temperature assignment



Finite temperature assignment

$$U_\beta^{MF} = \sum_k \sum_l C(k,l) G^{MF}(k,l) \xrightarrow{\beta \to +\infty} U^*$$

$$G^{MF} \xrightarrow{\beta \to +\infty} ?$$

Finite temperature assignment

$$U_\beta^{MF} = \sum_k \sum_l C(k,l) G^{MF}(k,l) \xrightarrow{\beta \to +\infty} U^*$$

$$G^{MF} \xrightarrow{\beta \to +\infty} ?$$

Theorem:

If the assignment problem has a unique solution among all permutations, then there are no non integer solutions

Finite temperature assignment

$$U_\beta^{MF} = \sum_k \sum_l C(k,l) G^{MF}(k,l) \xrightarrow{\beta \to +\infty} U^*$$

$$G^{MF} \xrightarrow{\beta \to +\infty} G^*$$

Theorem:

If the assignment problem has a unique solution among all permutations, then there are no non integer solutions

Finite temperature assignment

Hypothesis: there is a unique solution

Theorem:

Let Δ be the difference in energy between the best solution and the second best solution. Then:

$$\max_{k,l} \left| G_\beta^{MF}(k,l) - G^*(k,l) \right| \leq \frac{A(N)}{\beta \Delta}$$

Furthermore, rounding off each of the entries of G_β^{MF} to the nearest integer yields the permutation matrix G^* whenever

$$\beta > \frac{2A(N)}{\Delta}$$

$A(N) = N\left(N\log(N) - (N-1)\log(N-1)\right)$

Finite temperature assignment

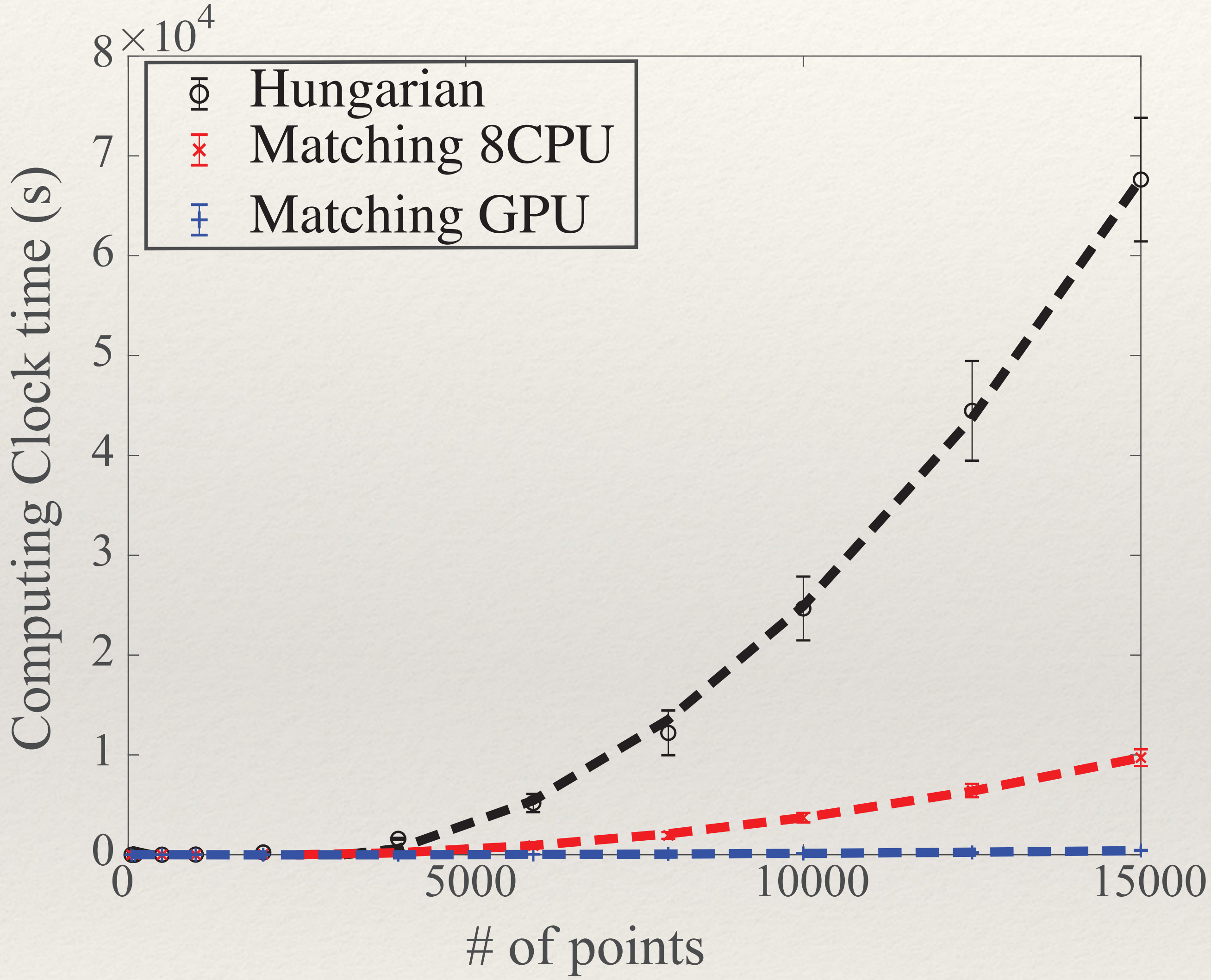
Hypothesis: there is a unique solution

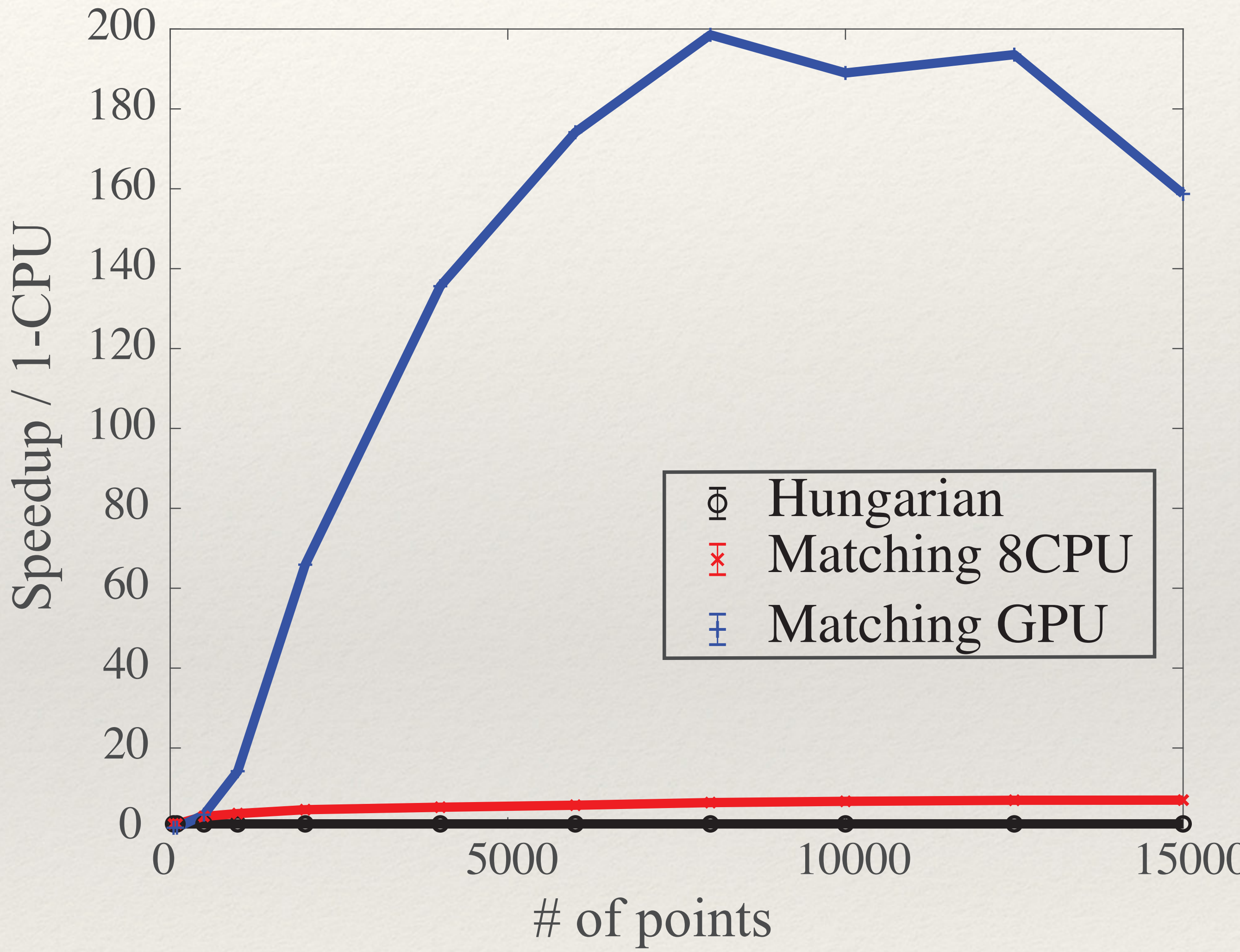
Theorem:

Let us assume that a an inverse temperature β the current solution matrix G_{β}^{MF} is row dominant. Then rounding off each of the entries of G_{β}^{MF} to the nearest integer yields the permutation matrix G^*

Finite temperature assignment

Hypothesis: there is a unique solution





Finite temperature assignment

Hypothesis:
the converged solution has some elements strictly in (0,1)

0	1	0	0	0
0.4	0	0	0.6	0
0	0	1	0	0
0.6	0	0	0.4	0
0	0	0	0	1

Finite temperature assignment

Hypothesis:
the converged solution has elements strictly in (0,1)

0	1	0	0	0
0.4	0	0	0.6	0
0	0	1	0	0
0.6	0	0	0.4	0
0	0	0	0	1

$$U = 0.4C(2,1) + 0.6C(2,4) + 0.6C(4,1) + 0.4C(4,4) + E$$

Finite temperature assignment

Hypothesis:
the converged solution has elements strictly in (0,1)

0	1	0	0	0
0.4	0	0	0.6	0
0	0	1	0	0
0.6	0	0	0.4	0
0	0	0	0	1

$$U = 0.4C(2,1) + 0.6C(2,4) + 0.6C(4,1) + 0.4C(4,4) + E$$

0	1	0	0	0
0.4-e	0	0	0.6+e	0
0	0	1	0	0
0.6+e	0	0	0.4-e	0
0	0	0	0	1

$$U = 0.4C(2,1) + 0.6C(2,4) + 0.6C(4,1) + 0.4C(4,4) + E - e(C(2,1) - C(2,4) + C(4,4) - C(4,1))$$

Finite temperature assignment

Hypothesis:
the converged solution has elements strictly in (0,1)

0	1	0	0	0
0.4-e	0	0	0.6+e	0
0	0	1	0	0
0.6+e	0	0	0.4-e	0
0	0	0	0	1

Necessarily:

A) *C(2,1)-C(2,4)+C(4,4)-C(4,1) = 0*

B) Setting **e = 0.4** leads to an integer solution

$$U = U^* - e(C(2,1) - C(2,4) + C(4,4) - C(4,1))$$

Finite temperature assignment

Hypothesis:
the converged solution has elements strictly in (0,1)

Theorem : *(Gartner and Matousek, 2006)*

If the relaxed assignment problem has at least one feasible solution, then it has at least one integral optimal solution. This solution is an optimal solution for the corresponding integer assignment program.

Property:

If the relaxed assignment problem has an optimum solution that contains fractional elements, then there exists (at least) one cycle $A = \{(a_1, b_1), (a_2, b_2), \ldots, (a_{2M}, b_{2M})\}$ in the cost matrix *C* for which $\Gamma = \sum_{i=1}^{2M} (-1)^i C(a_i, b_i) = 0$.

Finite temperature assignment

Hypothesis:
the converged solution has elements strictly in (0,1)

Theorem

Let Δ be the difference in total cost between the optimal solution and the second best solution. Then, adding random uniform noise with support $(0,\alpha]$ to each value of *C* and solving the assignment problem on this perturbed matrix will generate one integer solution that is also solution to the unperturbed assignment problem with probability **1**, whenever

$$\alpha < \frac{\Delta}{2N}$$

If all the entries of the cost matrix *C* are scaled to be integers, then $\Delta \geq 1$ and it suffices to have

$$\alpha < \frac{1}{2N}$$

Thank you!

Marc Delarue, Institut Pasteur

Henri Orland

NSF, CEA Saclay