Multilevel Algorithms for Linear Ordering Problems

Il'ya Safro

Joint work with Dorit Ron and Achi Brandt

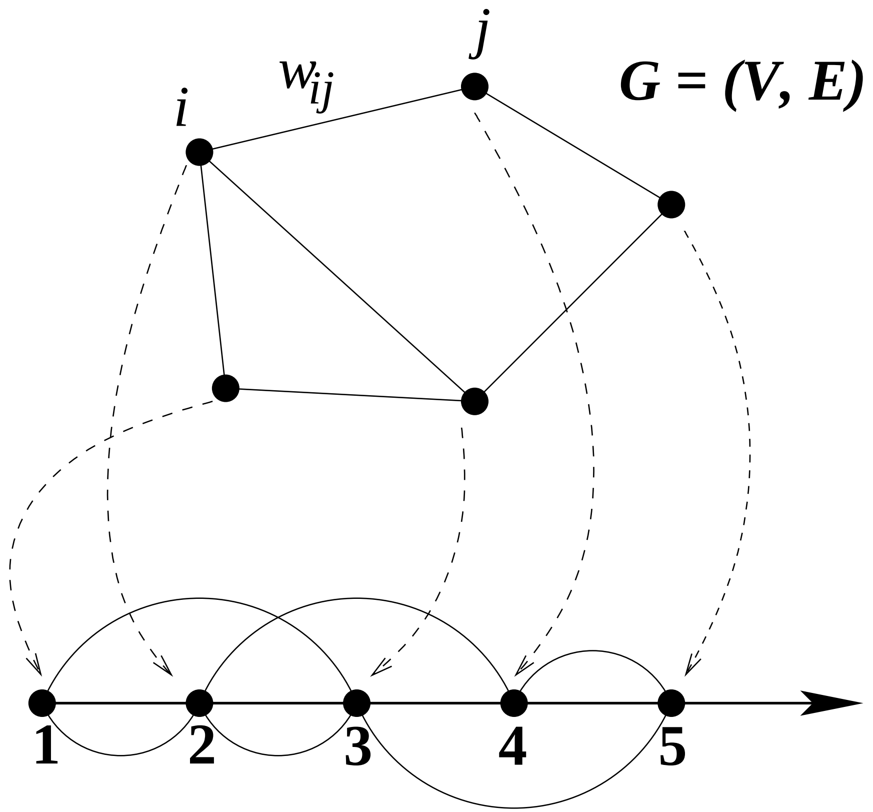
Faculty of Mathematics and Computer Science

Weizmann Institute of Science

June 21, 2005

Problem definition : Minimum p-sum

- Given graph $G = (V, E)$
- E - set of *weighted* edges



Goal : minimize over all π $\quad \sigma_p(G) = \Big(\sum_{ij \in E} w_{ij} |\pi(i) - \pi(j)|^p\Big)^{1/p}$

If $p = \infty$: minimize over all π $\quad \sigma_\infty(G) = \max_{ij \in E} w_{ij} |\pi(i) - \pi(j)|$

General facts

Motivation : VLSI design, sparse matrix computations, graph drawing, etc.

Status : NP-hard on general graphs

Heuristics :

For $p = 1$:

Diaz et al. : Simulated annealing, Hillclimbing, Spectral approach;

Koren et al. : Spectral approach + Multiscale approach;

Poranen : Genetic Hillclimbing;

Bar-Yehuda et al. : Decomposition Tree + Simulated annealing.

For $p = 2$:

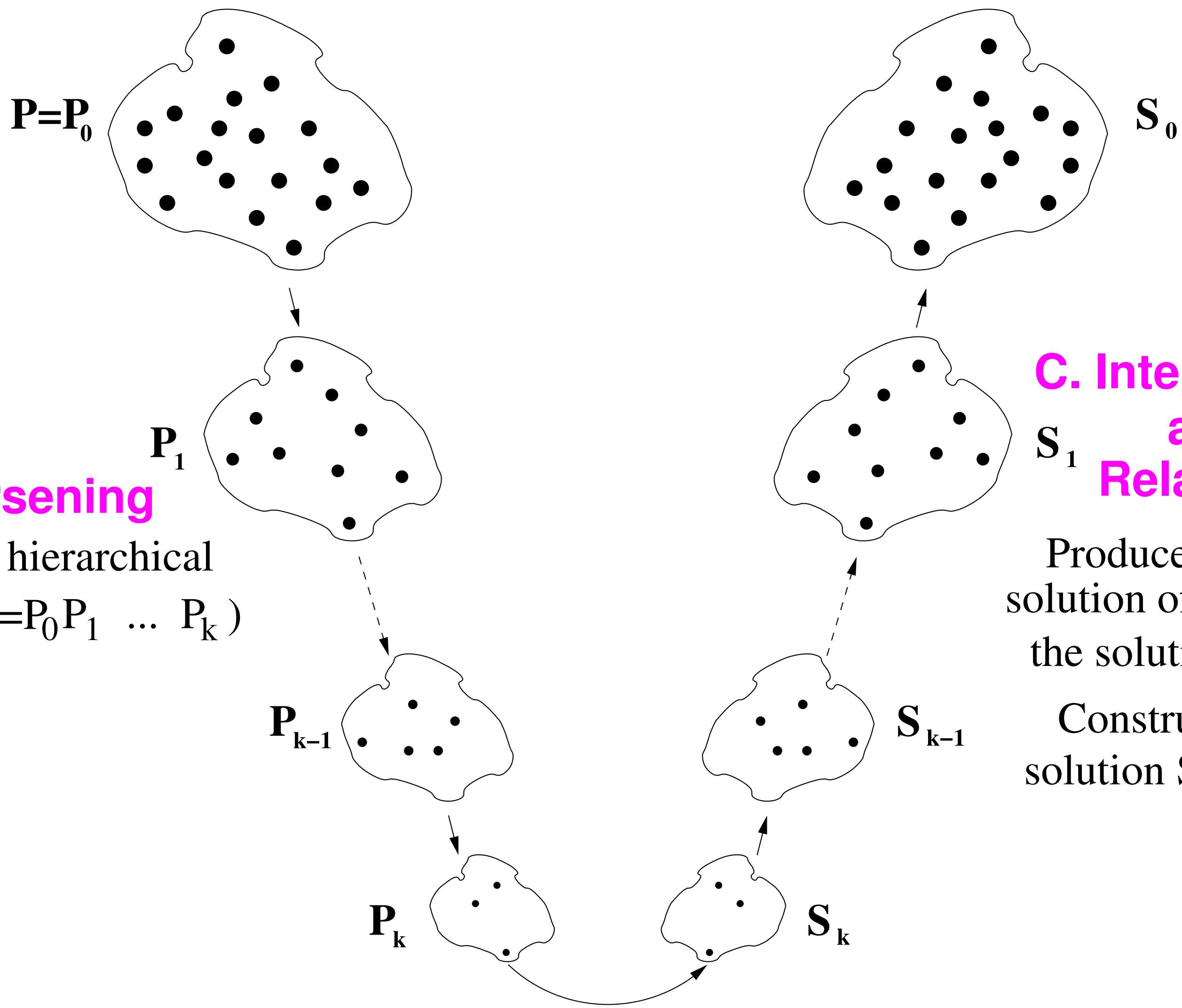
George et al. : Spectral approach

For $p = \infty$:

Del Corso et al. : Spectral approach;

Barnard et al. : Spectral approach.

General idea of the multiscale techniques



A. Coarsening

Defines the hierarchical structure ($P=P_0 P_1 \ \dots \ P_k$)

B. Exact solution

C. Interpolation and Relaxation

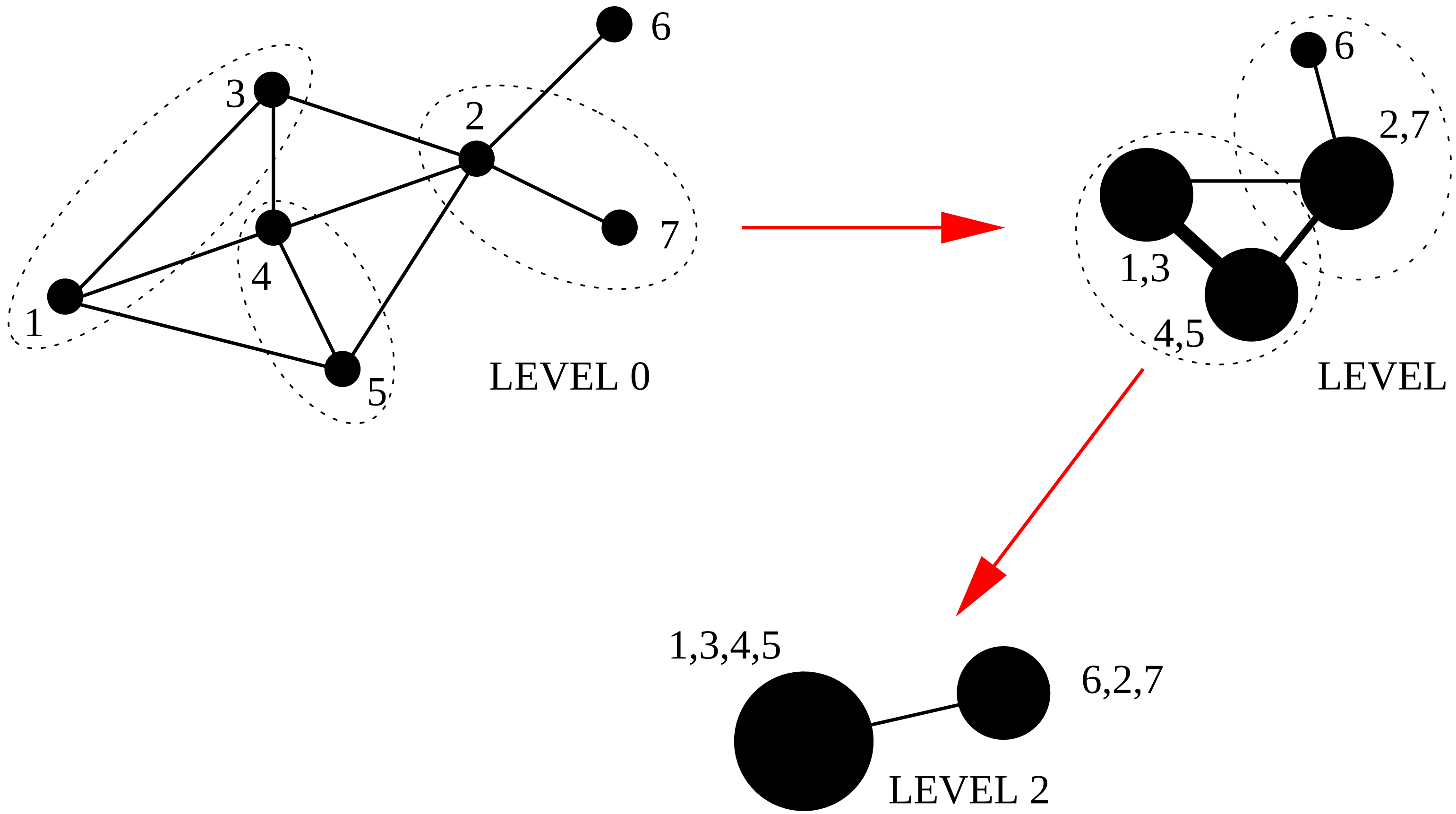
Produces an initial solution of P_{i-1} from the solution S_i of P_i

Constructs final solution S_{i-1} from P_{i-1}

Short survey on existing multiscale algorithms for Combinatorial Optimization Problems

- Examples : Placement, Partitioning, Minimum Linear Arrangement, Wavefront reduction.
- Quality : Usually exhibit superior results to other methods on practical test suites.
- Time : Usually exhibit **linear** time complexity.
- Most common algorithmic approach : “strict” coarsening (like matching, first choice, etc.).

Example of strict coarsening



Disadvantage : makes local decisions before accumulating the relevant global information.

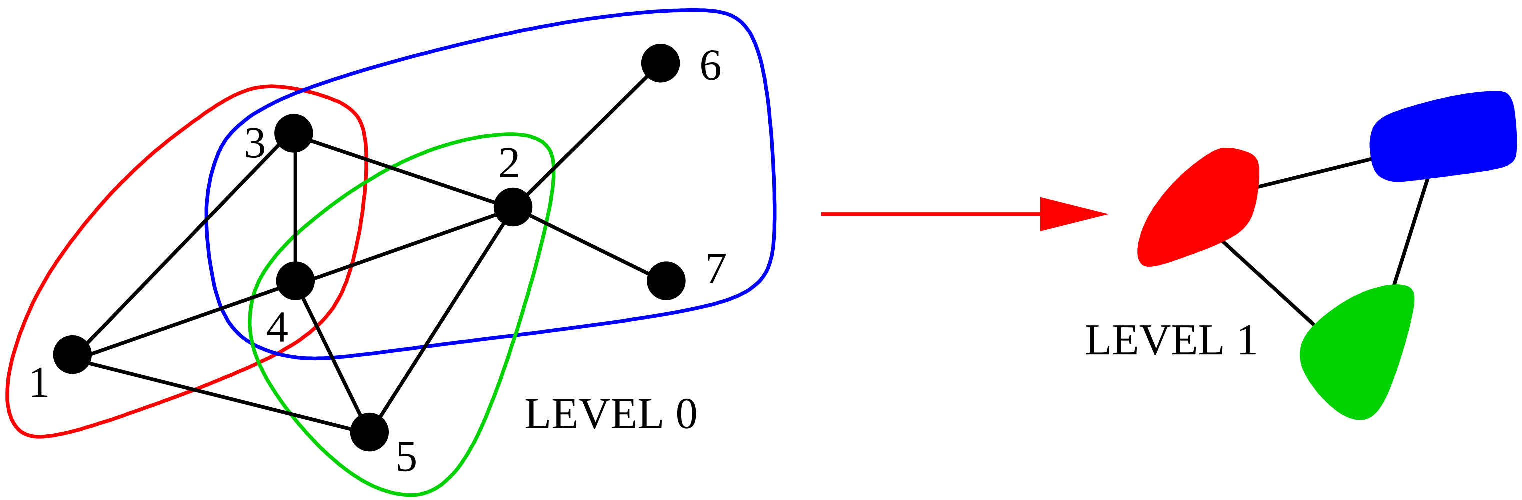
General idea : Weighted aggregation (inspired by Algebraic Multigrid)

Algebraic Multigrid

- Well established *linear* time method to solve $Ax = b$ for symmetric, positive definite A.

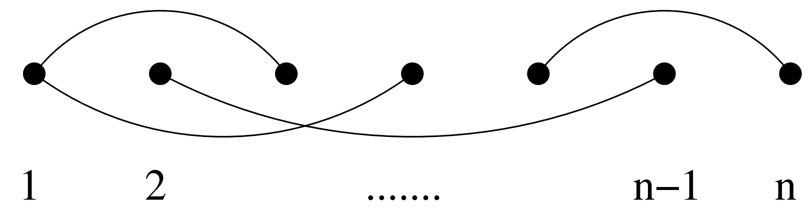
Weighted Aggregation

- The main new algorithmic property (inspired by AMG) is to make a ”soft” coarsening in which each node may be divided into *fractions* and different fractions form a coarse node.



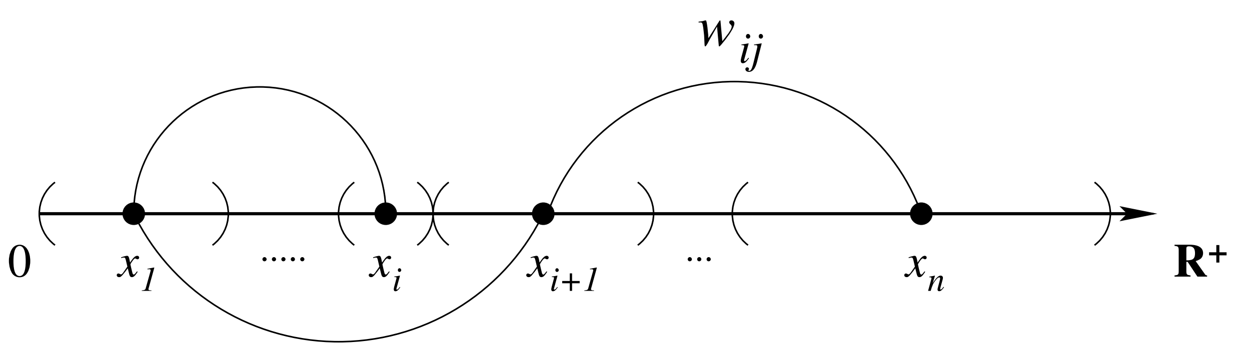
Generalized problem definition

Original problem



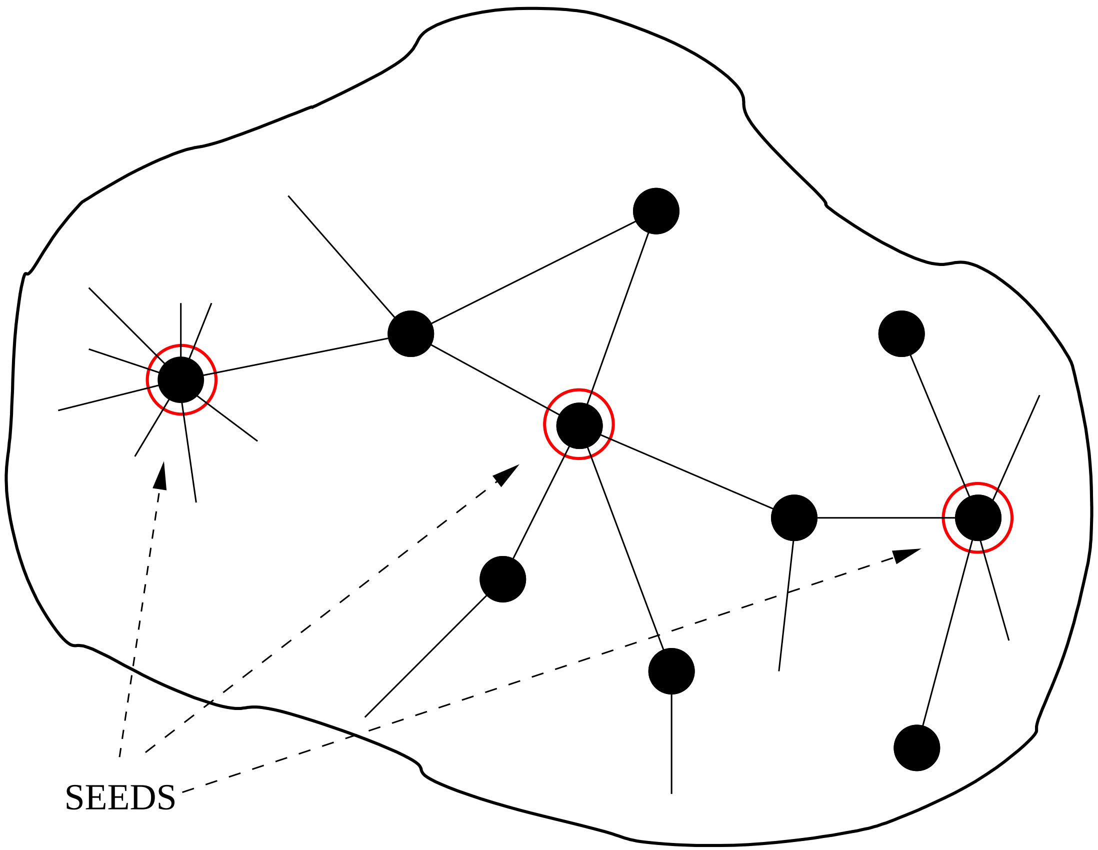
But now each node has a volume !

Generalized problem



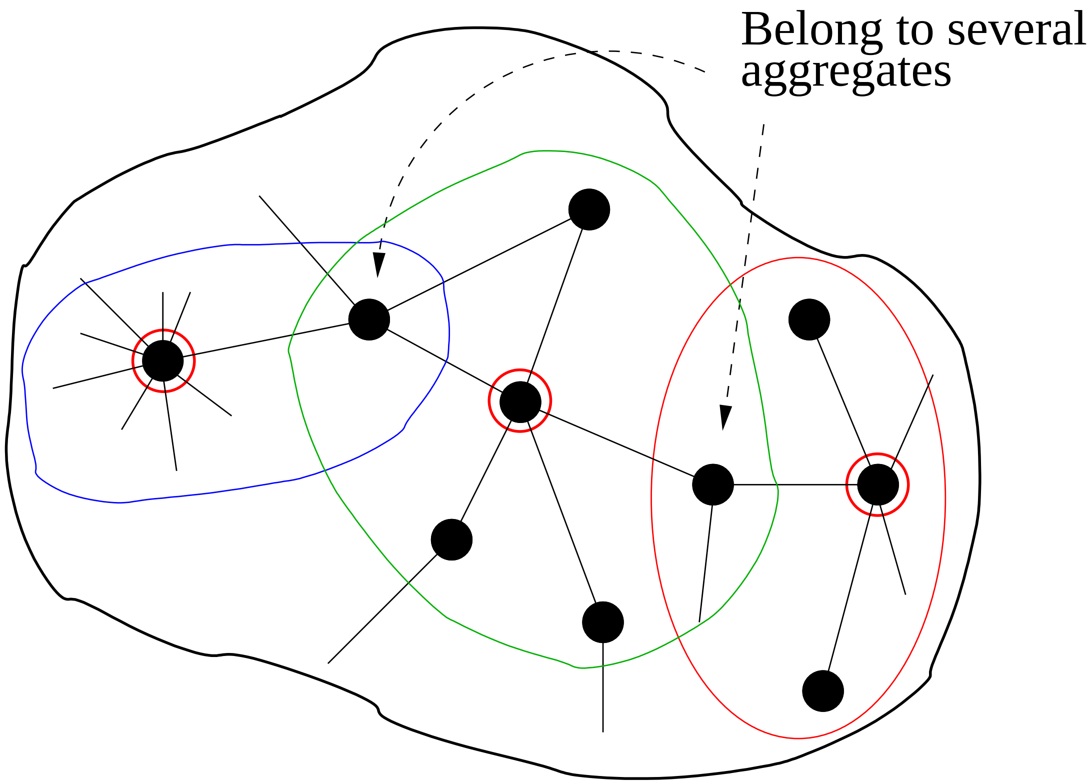
$$\mathcal{E}_\pi(x) \;=\; \Big(\sum_{ij\in E} w_{ij}\cdot|x_i - x_j|^p\Big)^{1/p}$$

Coarse nodes construction : the SEEDS selection



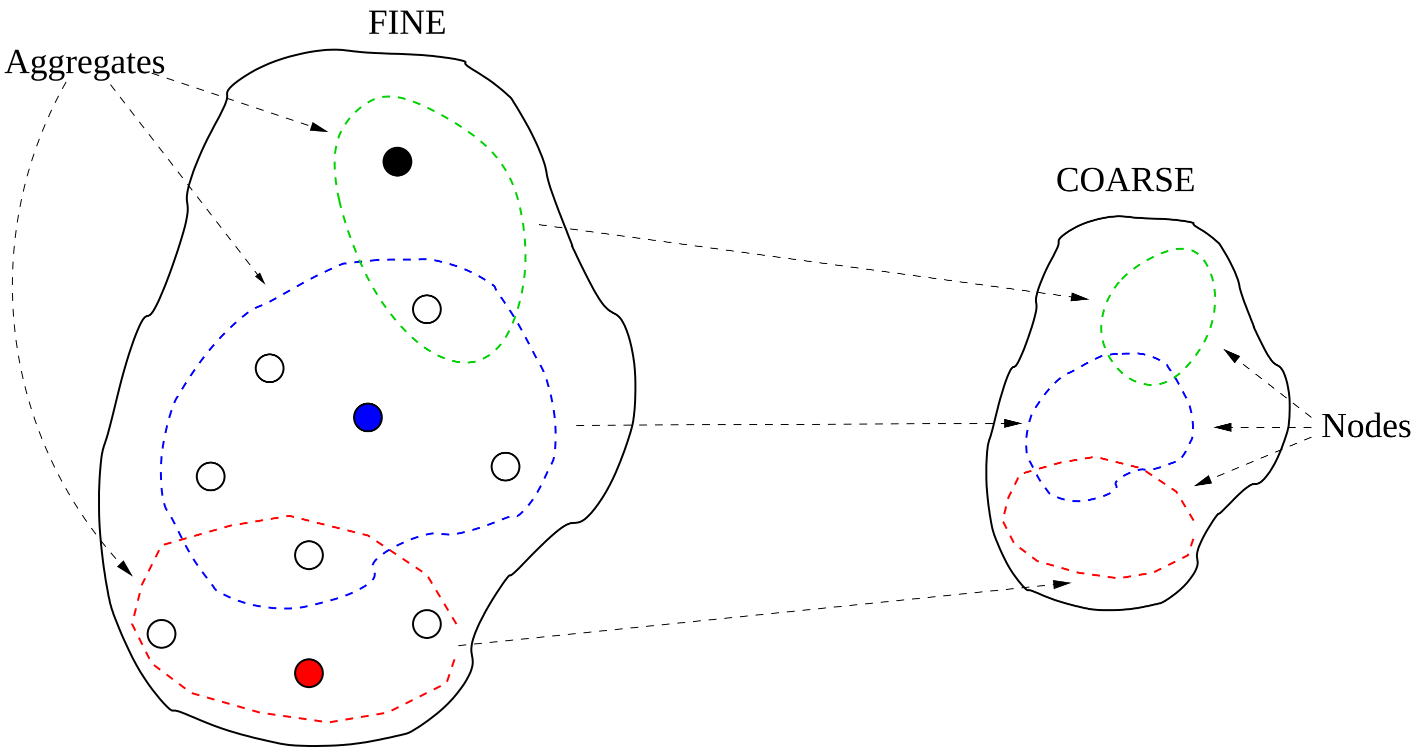
- Choose a dominating set $C \subset V$ s.t. all others from $F = V \setminus C$ are "strongly coupled" to C.
- Each such chosen node will serve as the *SEED* of a coarse aggregate.

Coarse nodes construction : the aggregation weights



- Define the aggregation weights of all vertices
- In some sense, the aggregation weights are the probabilities of a vertex to share a common property with the aggregates it belongs to.

Weighted aggregation : the construction of the Coarse Graph



The volumes of the coarse nodes satisfy

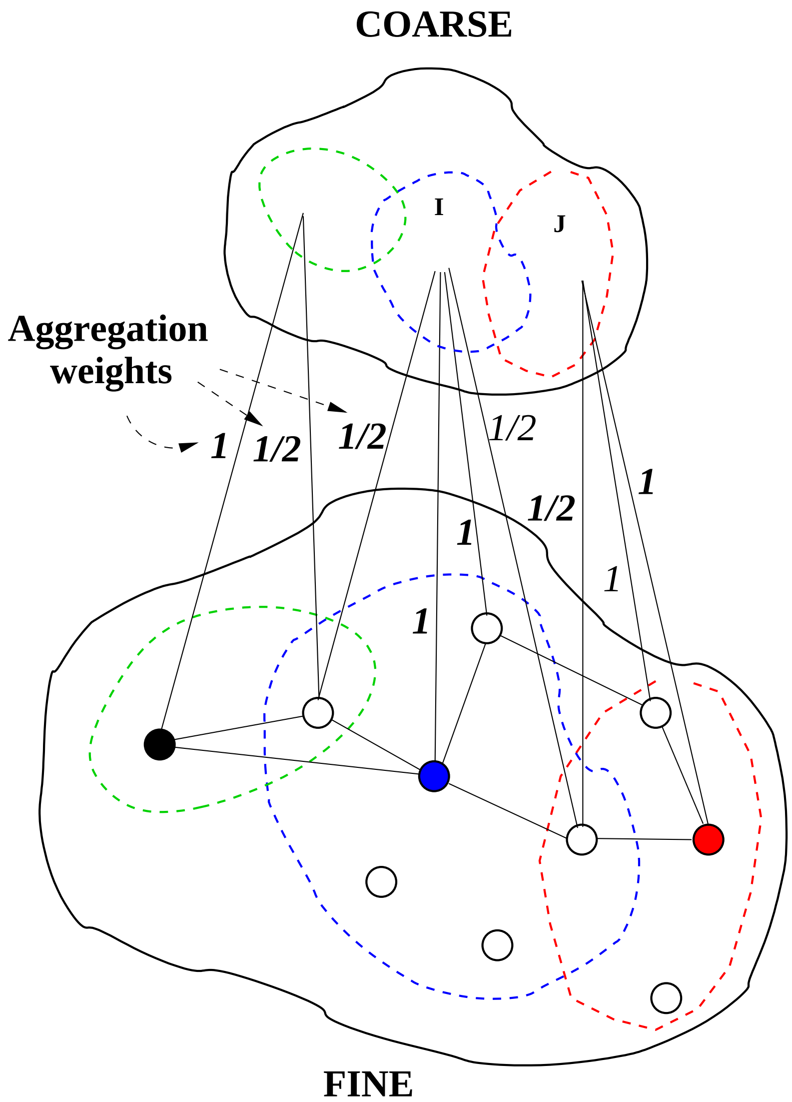
$$\sum_{v \in G_c} vol(v) \; = \; \sum_{u \in G_f} vol(u)$$

The edge construction of the Coarse Graph

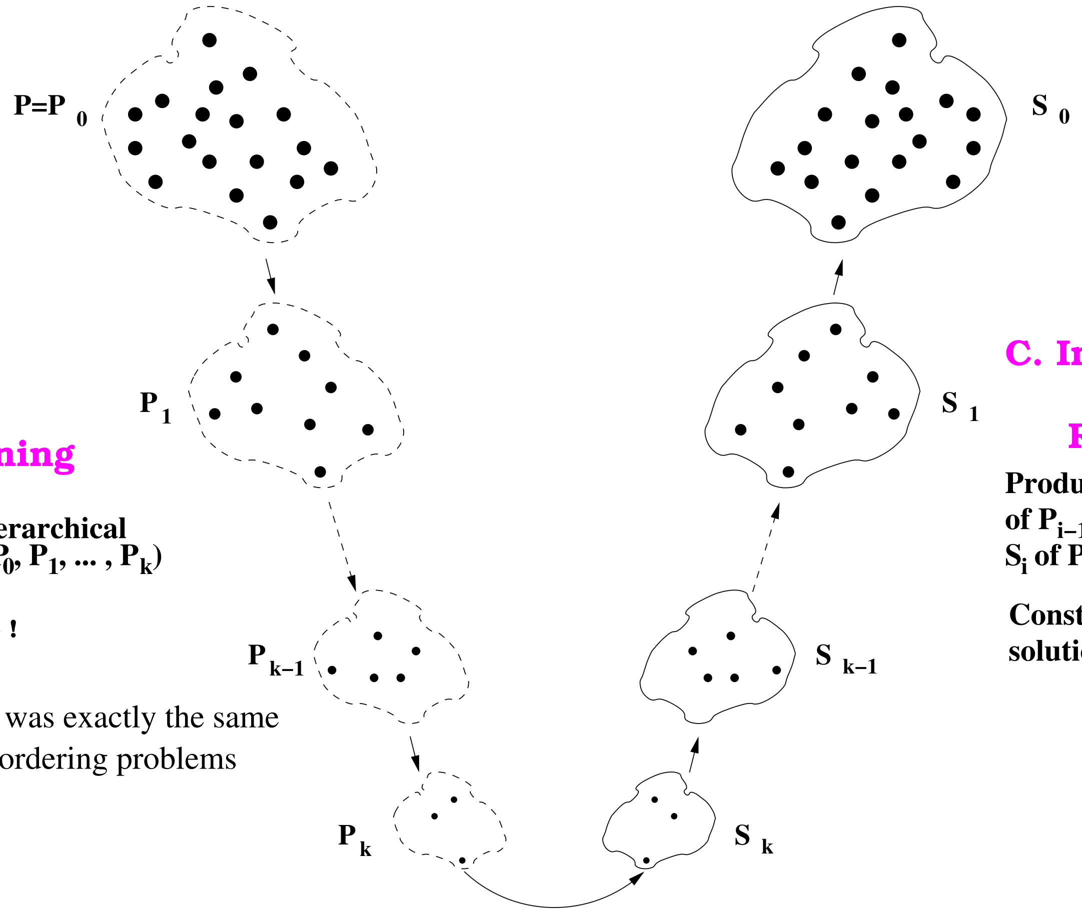
- Let $a(Ik)$ be the aggregation weight of the fine vertex k to aggregate I, then

$$w_{IJ} = \sum_{l,k} a(Il) \cdot w_{lk} \cdot a(kJ)$$

- Time : $O\big(|E_f|Q^2\big)$, where Q is the interpolation order.



Uncoarsening



A. Coarsening

Defines the hierarchical structure ($P=P_0, P_1, \dots, P_k$)

✓ **DONE !**

The coarsening was exactly the same for all linear ordering problems

B. Exact solution

1) Solve the problem exactly (energy dependent)
2) Can propogate many "different" solutions

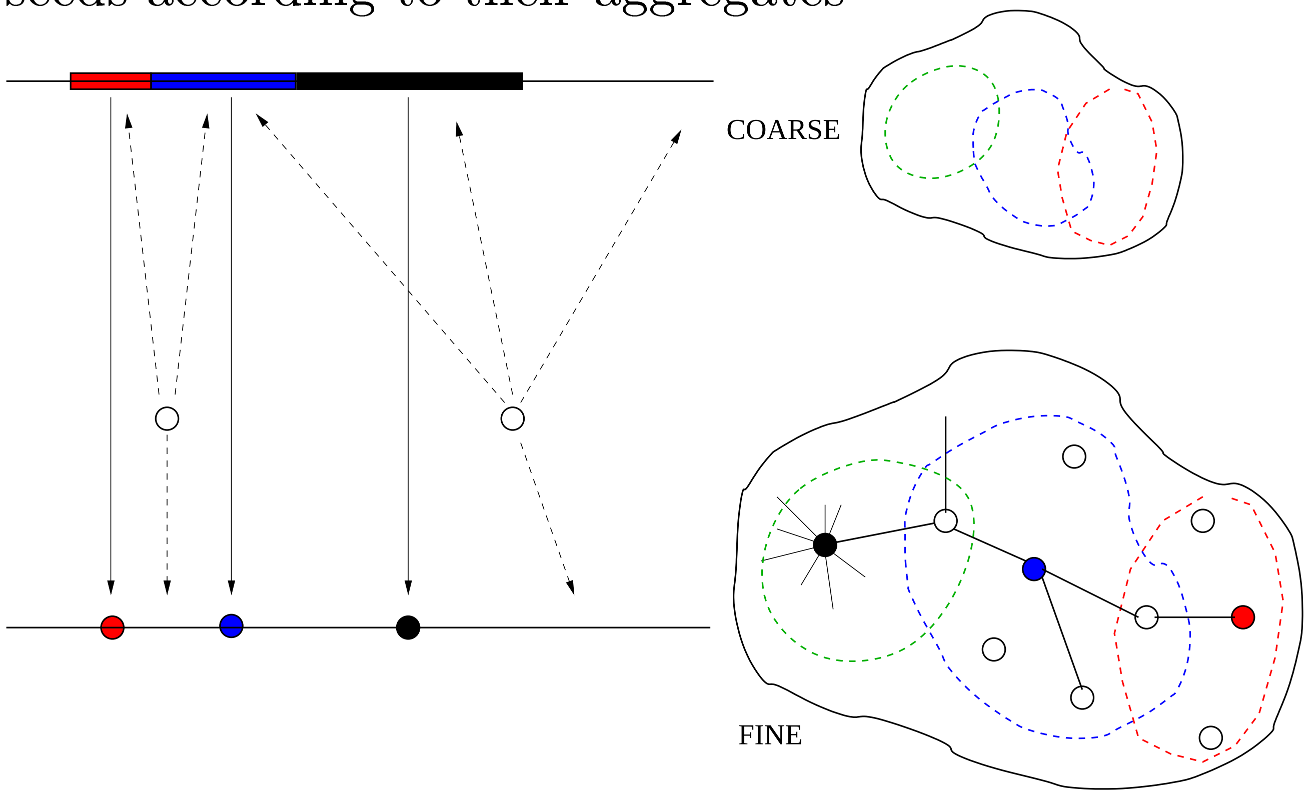
C. Interpolation and Relaxation

Produces an initial solution of P_{i-1} from the solution S_i of P_i

Constructs final solution S_{i-1} of P_{i-1}

Uncoarsening : interpolation by stages

1) Place the seeds according to their aggregates



2) Place other vertices by **minimizing** their local contribution to the total energy :

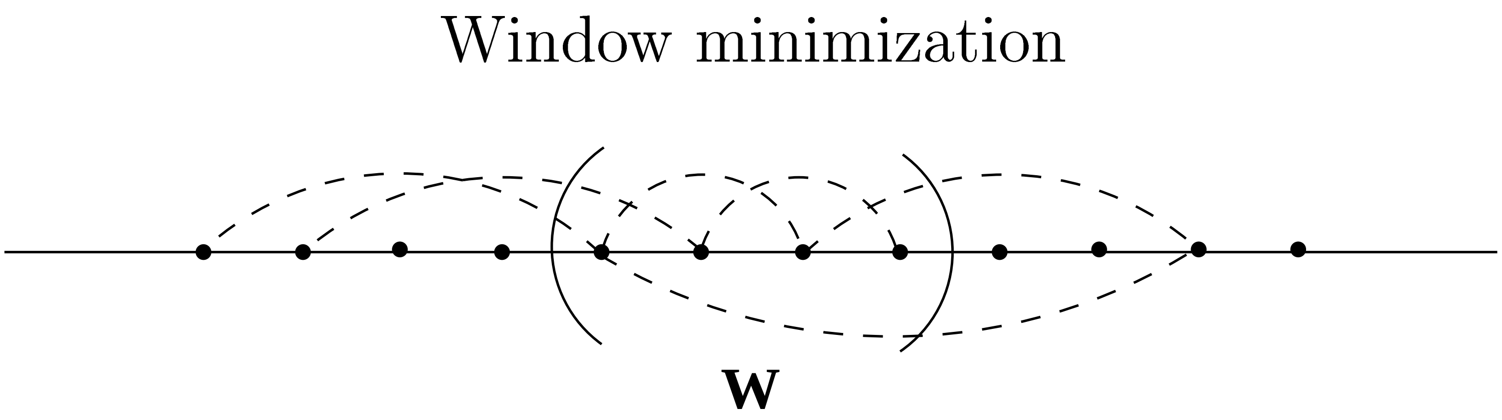
- $p = 1$: at their medians
- $p = 2$: at their weighted averages
- $p > 2$: solve minimization numericaly

Uncoarsening : relaxation

- The arrangement obtained after interpolation is a *first feasible solution.*
- Let us improve it by two types of relaxation which are very similar to the interpolation :
  - Compatible Relaxation Scheme : **keep coarse-level variables invariant** minimizing the energy of other vertices one-by-one wr to the problem,
  - Gauss-Seidel-like : Improve *all* vertices.

At the end of the relaxation legalize the order according to the volumes of the vertices.

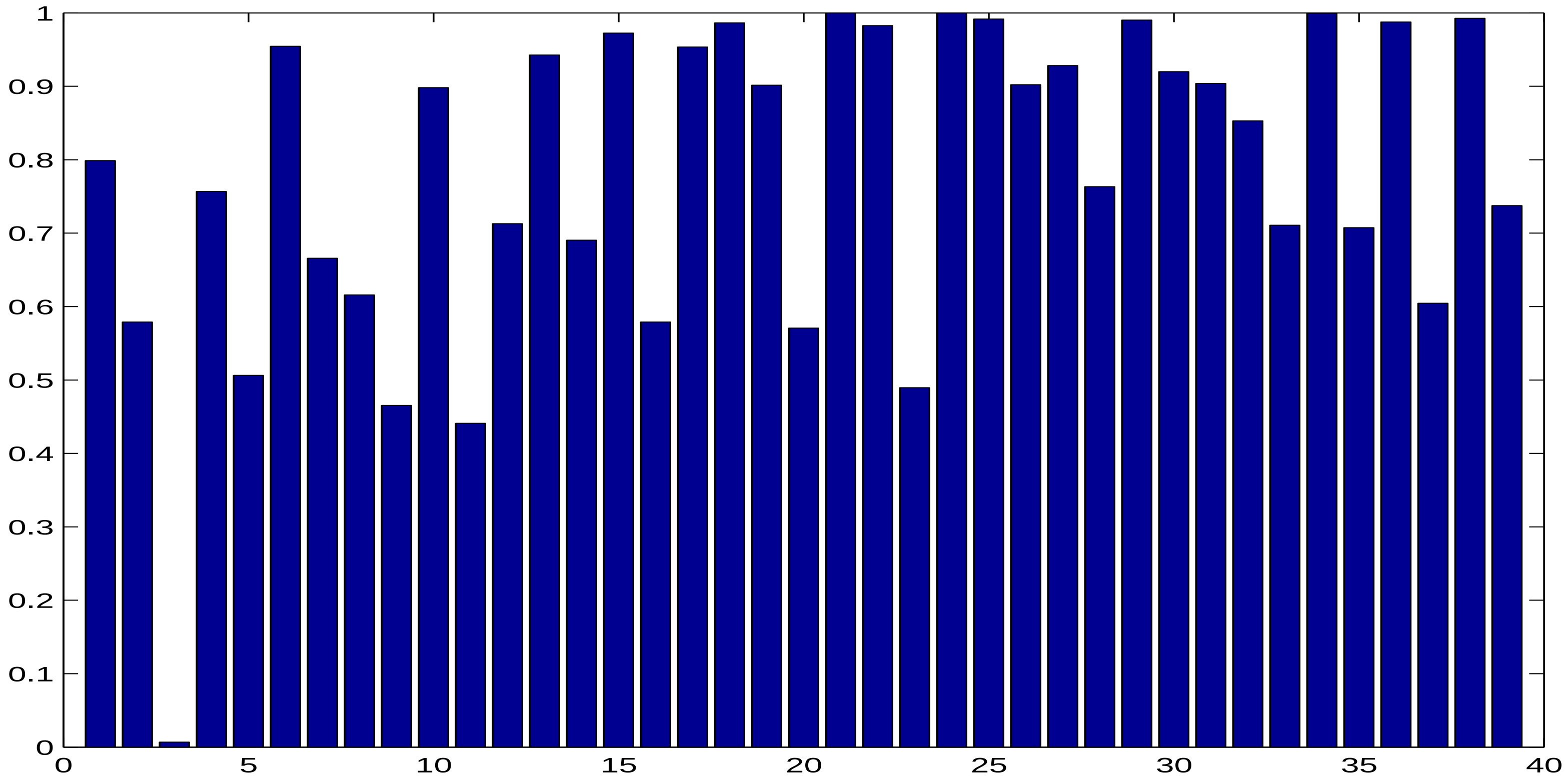
Uncoarsening : local improvements, $p = 2$



minimize $\sigma_2(W, \tilde{x}, \delta) = \sum_{i,j \in W} w_{ij}(\tilde{x}_i + \delta_i - \tilde{x}_j - \delta_j)^2 + \sum_{\substack{i \in W \\ j \notin W}} w_{ij}(\tilde{x}_i + \delta_i - \tilde{x}_j)^2.$

Current approximation Correction inside W outside W

Experimental results : $p = 2$



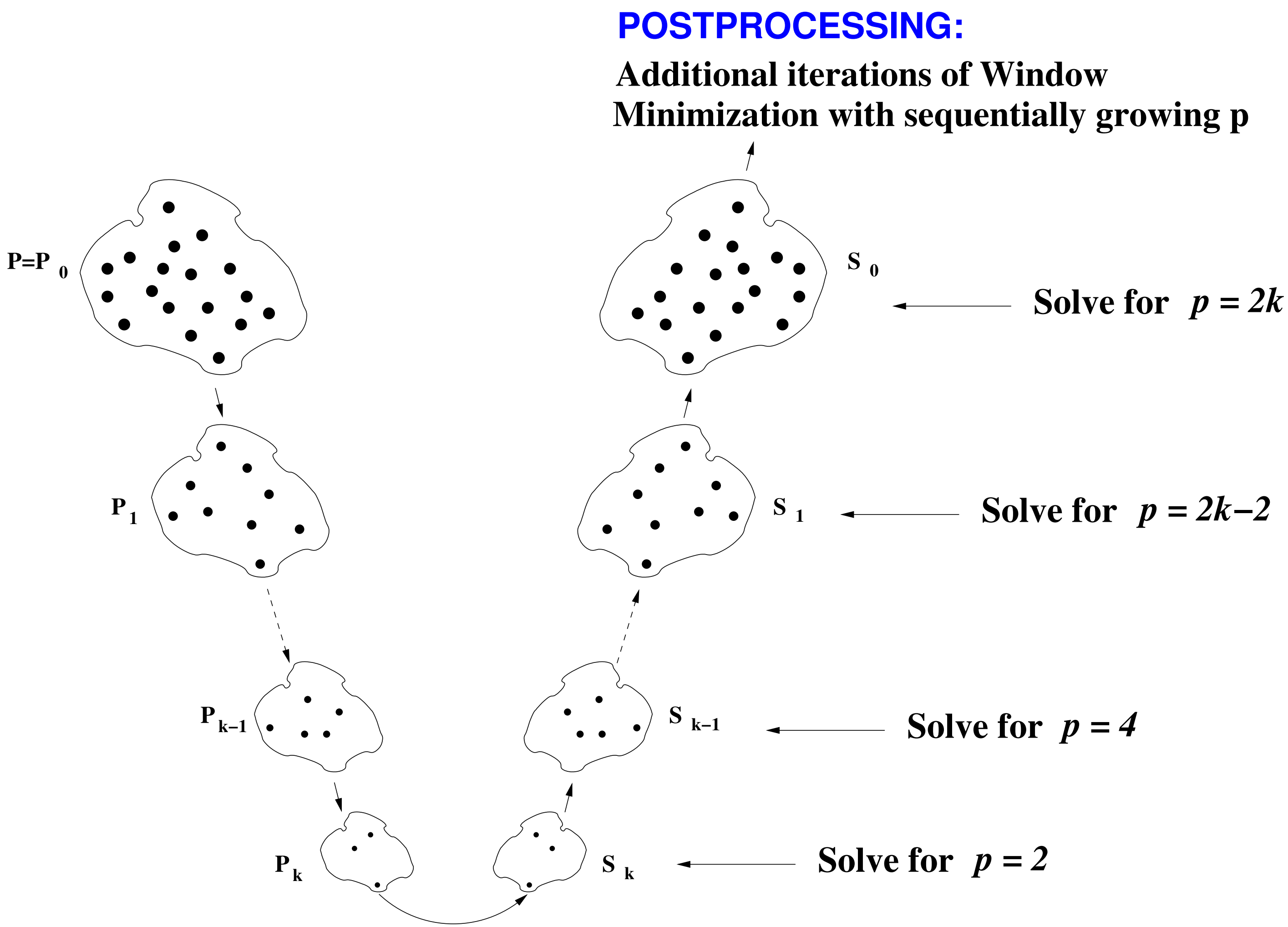
Ratio between multilevel algorithm (without postprocessing) and spectral approach

Experimental results : $p = 1$

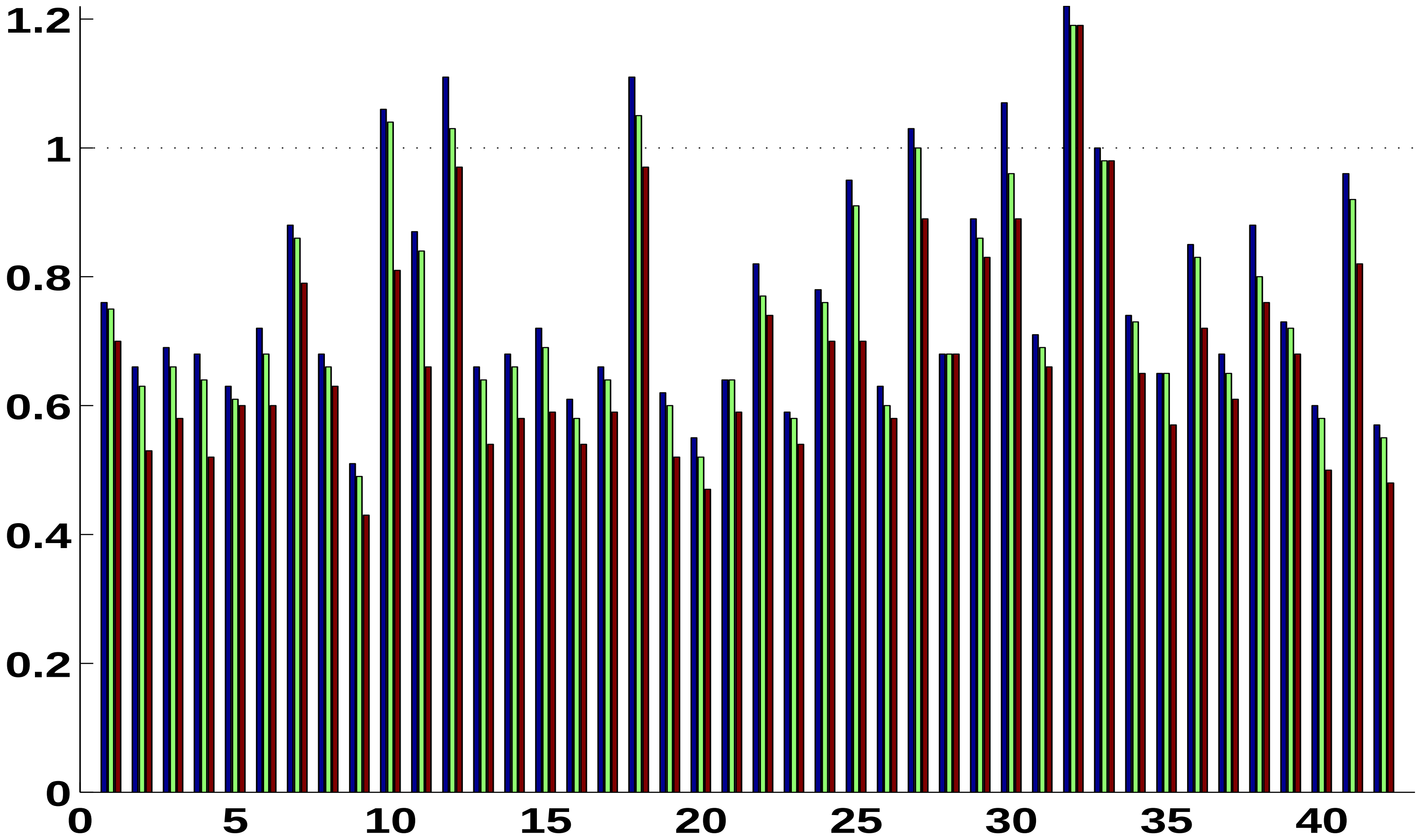
Graph	$\|V\|$	$\|E\|$	KH Time (in min)	Ours Time (in min)	Improvement for σ_1 Ours / KH
tooth	7.81E+04	4.53E+05	10.5	1.2	0.92
ocean	1.43E+05	4.10E+05	13.5	3.2	0.93
mrngA	2.57E+05	5.05E+05	23.5	6	0.91
rotor	9.96E+04	6.62E+05	16.5	1.9	0.93
598	1.11E+05	7.42E+05	19	3	0.84
144	1.45E+05	1.07E+06	28.5	4.4	0.98
m14b	2.15E+05	1.68E+06	40	6.8	0.87
mrngB	1.02E+06	2.02E+06	98	38	0.94
auto	4.49E+05	3.31E+06	100	18	0.88

[KH] - Y. Koren and D. Harel, *A Multi-Scale Algorithm for the Linear Arrangement Problem*, 2002.

Uncoarsening : Continuation scheme for $p > 2$



Experimental results : bandwidth



Suggestion : general scheme for other linear ordering problems for general p

- First approximation : Solve $\sigma_2(G)$
- Postprocessing : Do several iterations of local relaxation (like Window Minimization) with sequentialy growing p

---

Example 1: Minimum workbound problem

Example 2: Minimum wavefront problem

The minimum workbound problem

Goal : minimize over all π

$$wb(G,\pi) = \sum_i \max_{\substack{j \\ \pi(j)<\pi(i)}} w_{ij}(\pi(i)-\pi(j))^2$$

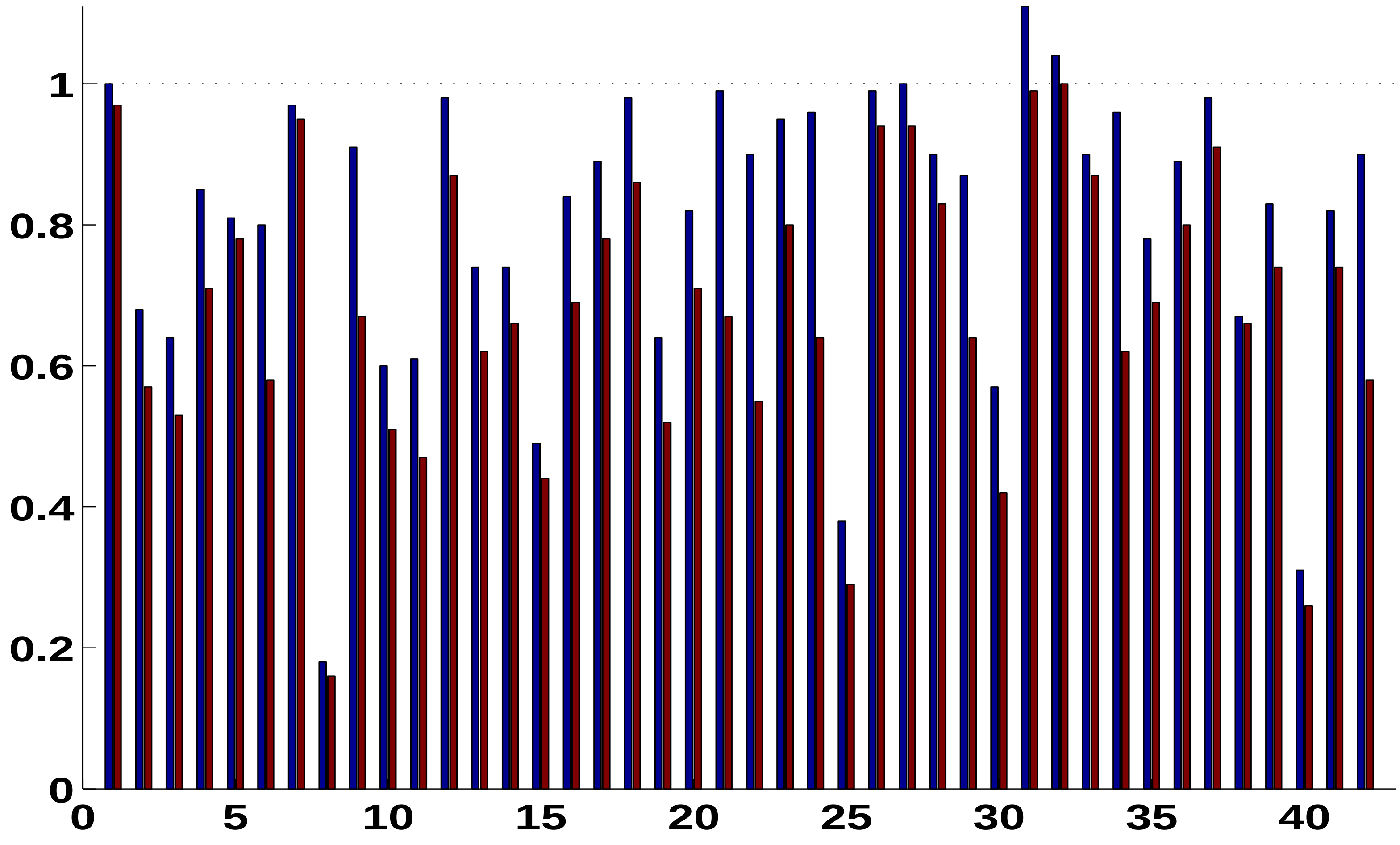
Generalization :

$$wb(G,x) = \sum_i \max_{j:x_j<x_i} w_{ij}(x_i - x_j)^2 \approx \sum_i \Big(\sum_{j:x_j<x_i} w_{ij}(x_i-x_j)^p \Big)^{2/p}$$

Window Minimization for the minimum workbound problem (Taylor exp.) :

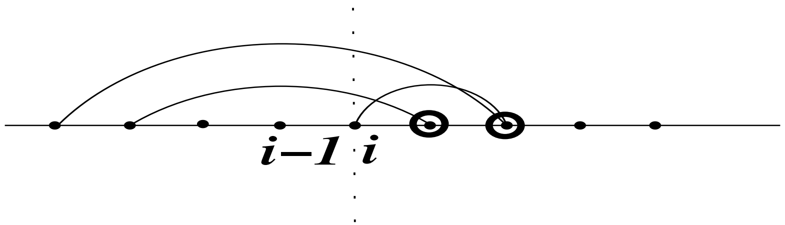
$$wb_p(W,\tilde{x},\delta) \approx wb_p(W,\tilde{x},\underline{0}) + \sum_{i\in W} \frac{\partial wb_p}{\partial \delta_i}(W,\tilde{x},\underline{0})\delta_i + \sum_{i,j\in W} \frac{\partial^2 wb_p}{\partial \delta_i \partial \delta_j}(W,\tilde{x},\underline{0})\delta_i\delta_j$$

Experimental results : workbound



The minimum wavefront problem

Goal : Minimize over all π $\left(\sum_i (f_{i,\pi})^2/n\right)^{1/2}$, where $f_{i,\pi}$



By just *evaluating* the wavefront functional on the arrangement provided by σ_2, we have obtained comparable results to the multilevel algorithm of Hu and Scott.

Conclusions

- A strategy to develop Multilevel algorithms for different combinatorial optimization problems
  - Linear time
  - Small standard deviation
  - Good scalability
- Using σ_2 instead of spectral ordering as a first approximation
- Other problems in our group : Placement for VLSI design, Partitioning of general graphs