Energetic Lattice Based Optimization

Bangalore Ravi KIRAN

ESIEE, Université Paris-Est

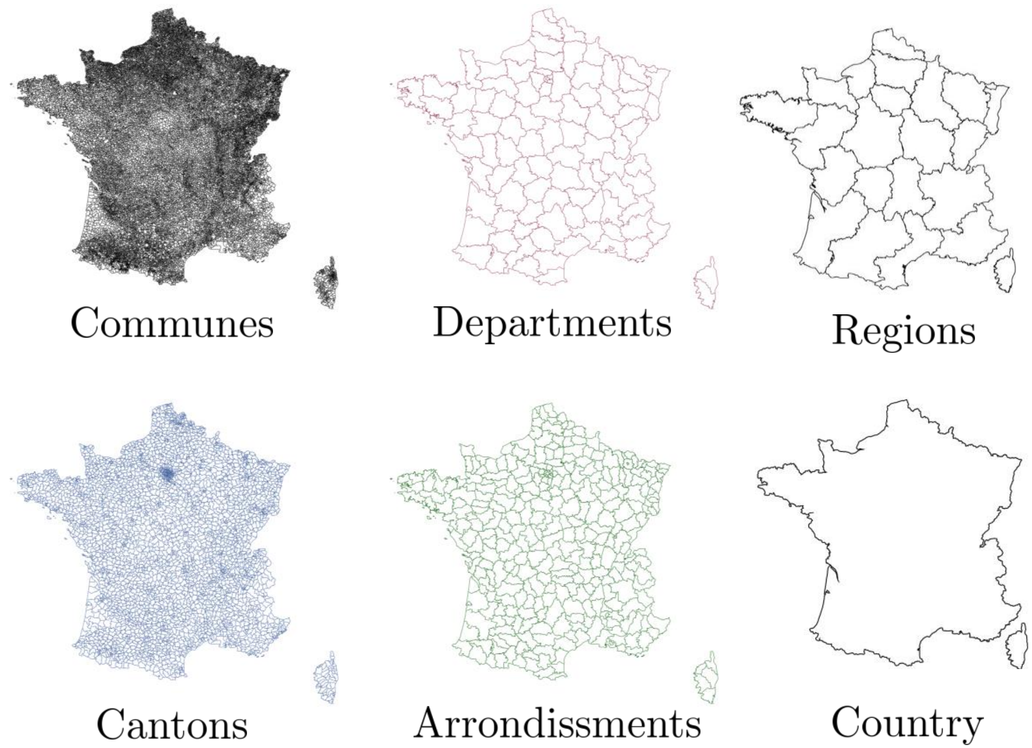
A3SI-LIGM

31 October 2014

Directed by: Jean SERRA

Introduction: Hierarchical Segmentation

Ultrametric Contour Map Hierarchy



Problem Formulation

Input

Hierarchy of Partitions

Mumford-shah
Haussdorf-distance
Total variation

Energy ω

Luminance/Chrominance
Proximity measure
Stereo disparity

Functions

Goal

Extract partition from hierarchy with least energy

- Tractable global solution
- Conditions for uniqueness
- Conditions for Increasing unique solutions
- Optimality of solutions

Overview

1. **Notations and structures**
    - Partitions & Partial Partitions
    - Hierarchy of Partitions
    - Representations
    - Energy
2. Review on Optimization on Hierarchies
3. Dynamic Programming
4. Braids of Partitions
5. Energetic-Lattices
6. Constrained Optimization
7. Conclusion

Partitions & Partial Partitions

Non-void disjoint union: A family π of subsets of E, that are non-empty, mutually disjoint, and whose union covers E.

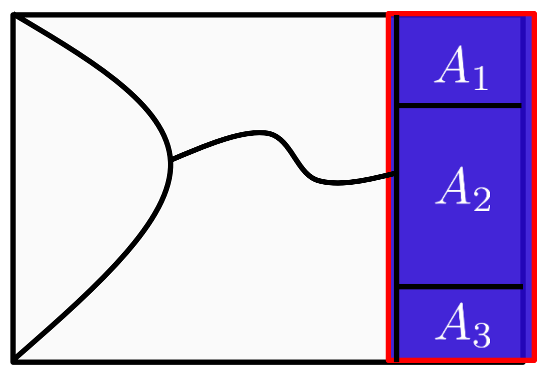
$$\pi = \{S_i \subseteq E\}, \qquad \cup S_i = E, \qquad S_i \cap S_j = \emptyset$$

A partition of a subset $S \subseteq E$ is defined as:

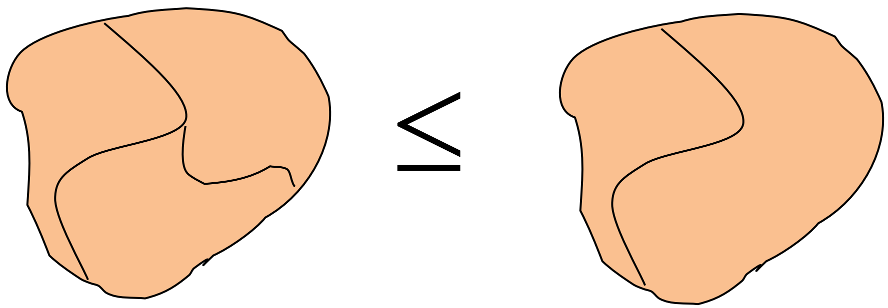
$$\pi(S) = \{A_i \mid A_i \subseteq S, A_i \cap A_j = \emptyset\}$$

$S = \cup A_i$ is called the support of $\pi(S)$

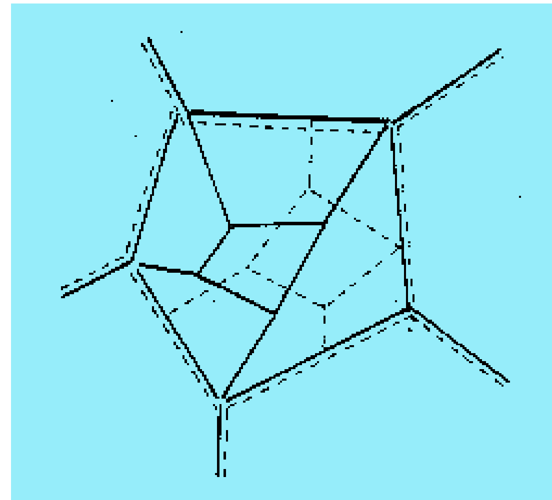
$$S = A_1 \cup A_2 \cup A_3$$
$$E := \mathbb{R}^2 \text{ or } \mathbb{Z}^2$$



Partial Partition Lattice



Refinement Ordering



Each two partitions admit:

- a lowest upper bound,
- a greatest lower bound.
- Forms a Complete lattice

$\mathcal{D}(E)$ Set of all partial partitions of E

Hierarchy of Partitions (HOP)

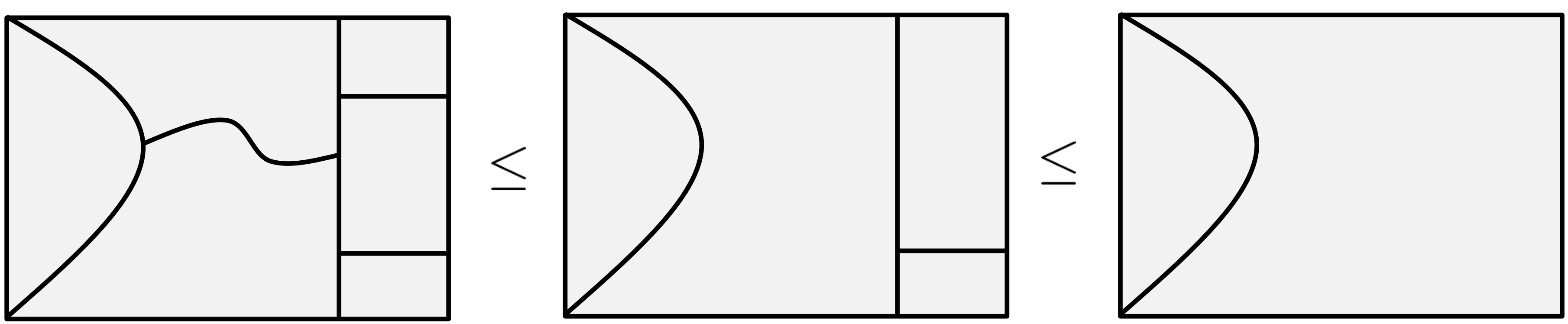
An indexed family $\{\pi_i, i \in I \subseteq \overline{\mathbb{Z}}\}$ of partitions of E defines a hierarchy when,

(i) π_i are nested, forming a chain:

$$H = \{\pi_i, i \in I\} \quad \text{with} \quad i \leq k \ \Rightarrow \ \pi_i \leq \pi_k, \qquad I \subseteq \overline{\mathbb{Z}},$$

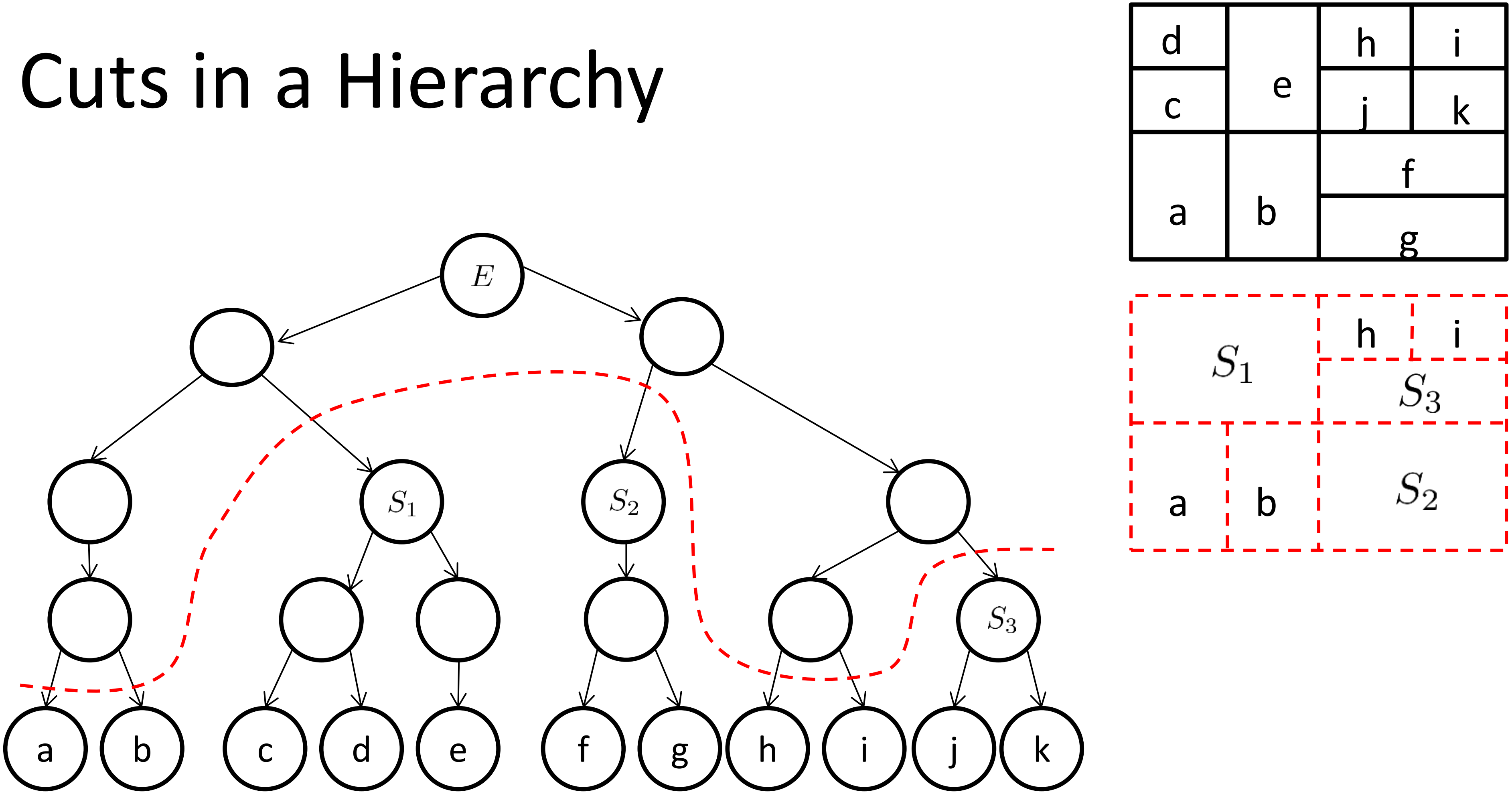
where π_0 is finest partition called leaves, and the coarsest one, is the root

(ii) finite leaves in any class of H



Elements $S \in \pi_i, \pi_i \in H$ are called classes of the HOP

Cuts in a Hierarchy



A cut of H is a partition of E with classes of H

$\Pi(E, H)$:= Set of all cuts composed by classes from H

Energies on Partitions and P.P.

Energy/Function on partial partitions:

$$\omega : \mathcal{D} \to \mathbb{R}$$

Energy of a partial partition can be written by composing energies of its classes

$$\omega(\pi(S)) = \sum_{A_i \in \pi(S)} \omega(A_i)$$

[Guigues 2003] Separable energies can be rewritten in the same form by additive composition of energies.

$\{S\}$:= Partial partition of support S into a single class

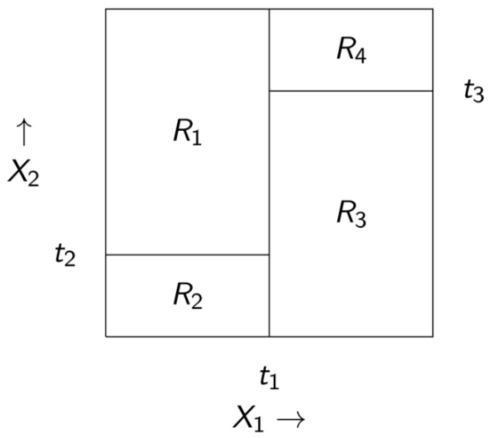
Energy over $\{S\}$ is written shortly as $\omega(S)$

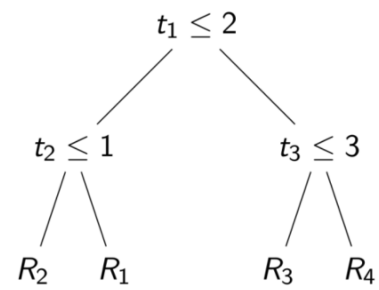
Overview

- Notations and structures
- **Review on Optimization on Hierarchies**
  - [Breiman et al 1984]: Classification & Regression Tree Pruning
  - [Salembier-Garrido 2000]: Binary Partition Tree pruning
  - [Guigues 2003]: Scale Sets and Scale climbing
- Dynamic Programming
- Braids of Partitions
- Energetic Order and Energetic-Lattices

Please refer the thesis for a more complete review.

Constrained Optimization on Trees



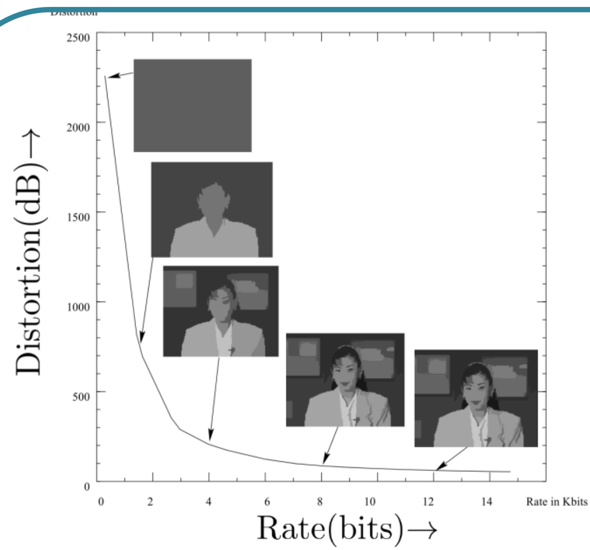


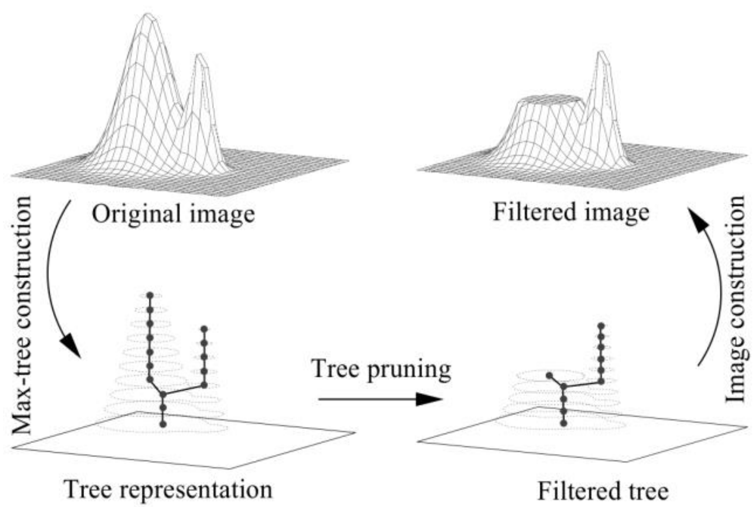
Classifier Complexity-Training Error

$$\omega(S, \lambda) = \int_S \text{Error(x)} + \lambda \cdot \ \# \ \text{Classes}$$

[Breiman et al. 1984]

$\hat{\lambda}$ minimizes cross-validated error.





[Salembier-Garrido 2000]

Minimal distortion, for λ-dependent rate constraint.

Scale-Set Representation

[Guigues 2003]

Extraction sequence of λ-cuts given

- Hierarchy of Partitions
- Energy, like Mumford-shah functional
- Scale parameters λ_i

The set of λ-cuts form a hierarchy.

Optimally pruning (Breiman, Salembier), and Guigues λ-cut calculated by dynamic programming.

Questions

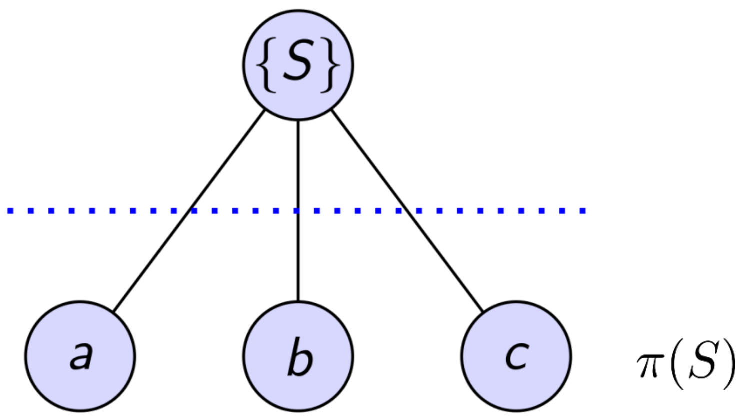
Dynamic program aggregates local comparisons.

- What are the necessary conditions for global minima to exist?
- Which class of energies enable local comparisions aggregate to reach global minimum ?
- Do optima in these studies use only the numerical ordering in energy?
- Is additivity necessary condition to answer these questions ?

Overview

- Notations and structures
- Review on Optimization on Hierarchies
- **Dynamic Programming**
  - Dynamic Program Sub-structure
  - Examples
  - h-increasingness
  - Minkowski norm based generalization
  - Other h-increasing energies
- Braids of Partitions
- Energetic Order and Energetic-Lattices
- Constrained Optimization
- Conclusion

Dynamic Program



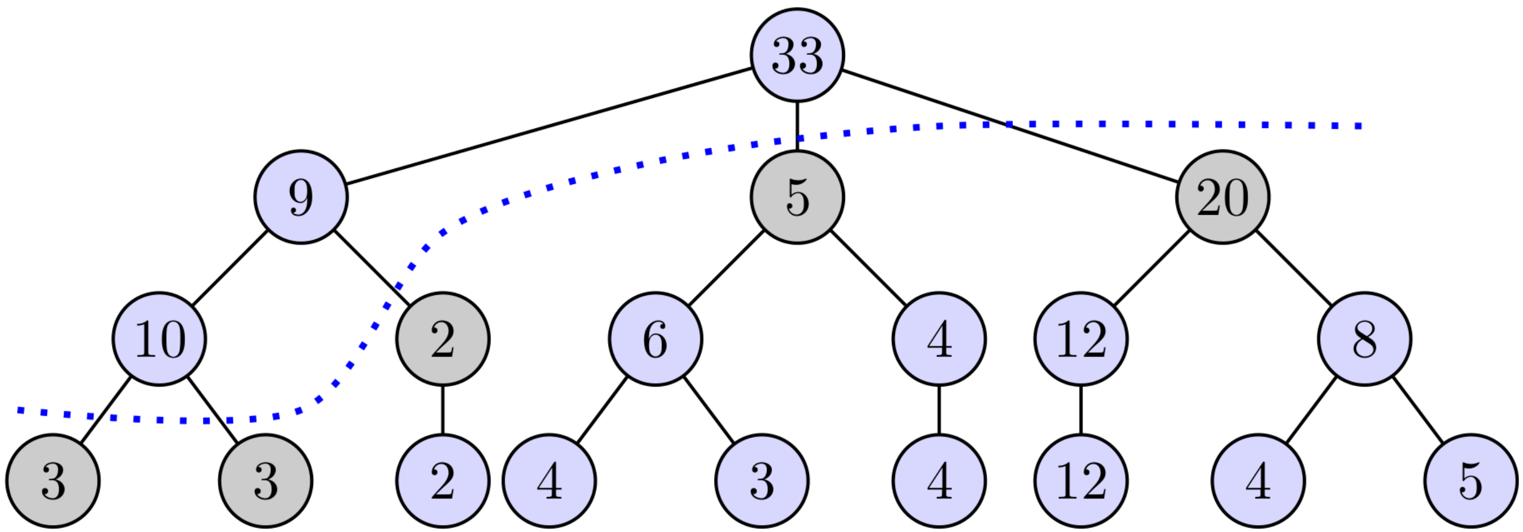
$$\omega^*(\pi(S)) = \min\{\omega(\{S\}, \sum_{a\in\pi(S)} \omega(a)\}$$

$$\pi^*(S) = \begin{cases} \{S\}, & \text{if } \omega(S) \leq \sum_{a\in\pi(S)} \omega(a) \\ \pi(S), & \text{otherwise} \end{cases}$$

π^* is the optimal cut given ω after DP

$\pi^*(\lambda)$ is optimal λ-cut given $\omega(\lambda)$ after DP

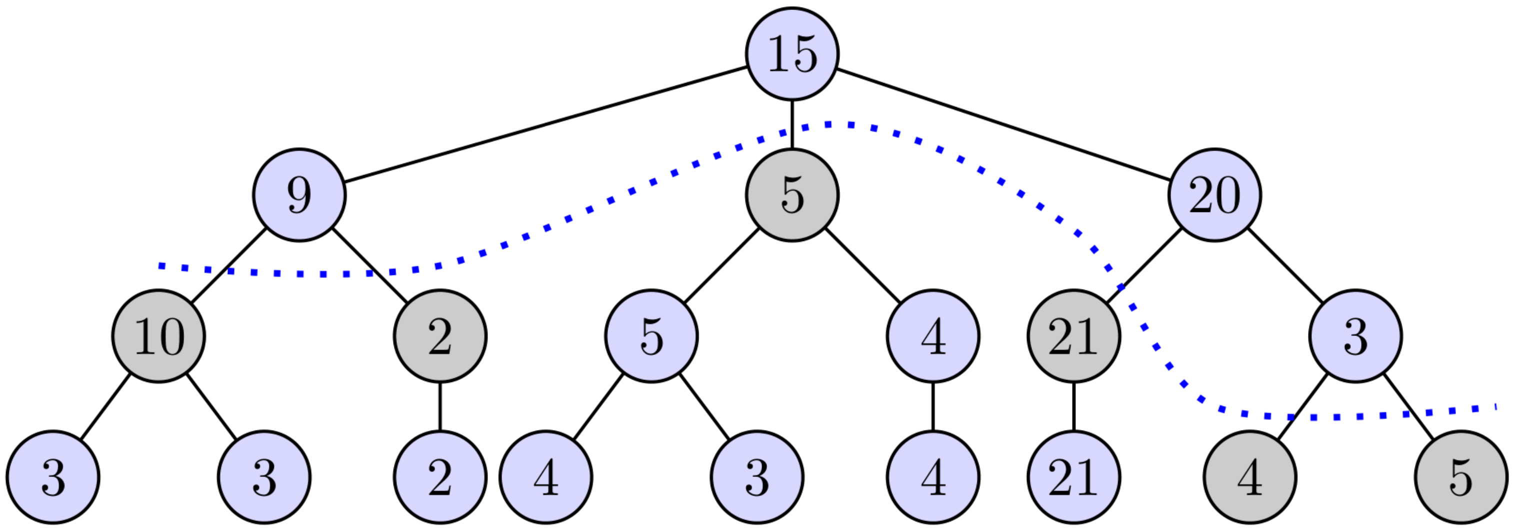
Salembier-Garrido & Guigues (Additive)



$$\pi^*(S) = \begin{cases} \{S\}, & \text{if } \omega(S) \leq \sum_{a \in \pi(S)} \omega(a) \\ \pi(S), & \text{otherwise} \end{cases}$$

[Salembier Garrido 2000, Guigues 2003]

Dominant Ancestor



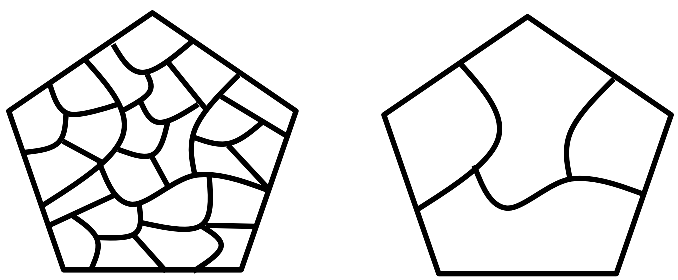
- Optimal class S^*: smallest class more energetic than all its descendants.
- $\omega(S^*) \leq \bigvee_{\pi(S)} \omega(T_i)$

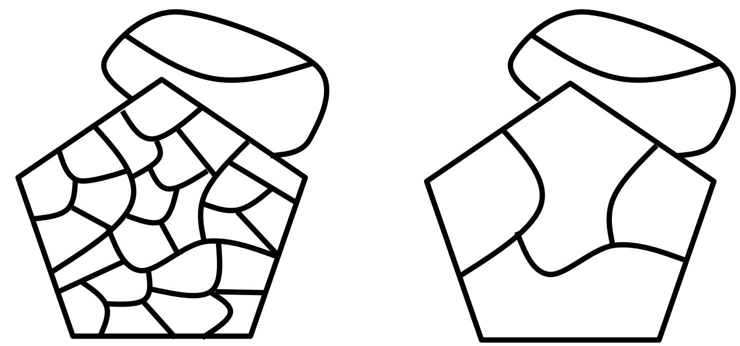
[Akcay-Akcoy 2008]

Generalizing the DP
h-increasingness on HOP

[Serra DGCI 2011, Kiran-Serra PR 2013]

$\pi_1(S) \quad \leq \quad \pi_2(S)$ $\qquad\qquad$ $\pi_1(S) \sqcup \pi_0 \quad \leq \quad \pi_2(S) \sqcup \pi_0$





$$\omega(\pi_1(S)) \leq \omega(\pi_2(S)) \quad \Longrightarrow \quad \omega(\pi_1(S) \sqcup \pi_0) \leq \omega(\pi_2(S) \sqcup \pi_0)$$

$\sqcup$: disjoint union to concatenate partial partitions during DP

Local optimum $\Longrightarrow$ Global optimum

h-increasing energy compositions

- **Additive**

  [Breiman et al. 1984, Salembier-Garrido 2000, Guigues 2003]

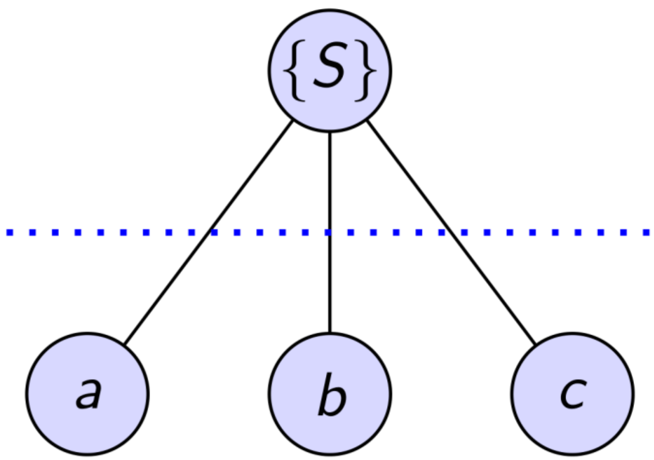
- **Supremum & Dominant Ancestor**

  [Akcay-Akcoy 2008, Soille 2008, Valero 2011, Veganzones-Chanussot 2014]

- **Minkowski norm generalization**

- **Max-pooling type, alternating compositions**

Generalized Minkowski composition



$$\omega(\pi(S)) = \left[\sum_{u \in [1,q]} \omega(T_u)^{\alpha} \right]^{\frac{1}{\alpha}}$$

$$\omega^*(\pi(S)) = \min\{\omega(\{S\}, \omega(\pi(S))\}$$

α	$\omega(T_i)$ Composition Law
$-\infty$	infimum
-1	harmonic sum
0	number of classes
$+1$	sum
$+2$	quadratic sum
$+\infty$	supremum

Mumford-Shah Energy

Initial Image

Initial watershed hierarchy H on luminance l

Mumford-Shah Energy

$$\omega(\pi(S), \lambda) = \sum_{1 \leq k \leq p} \omega_{\varphi}(T_k) + \lambda \sum_{1 \leq k \leq p} \omega_{\partial}(T_k)$$

Optimal Cuts

luminance fidelity term

$\omega_\varphi(T) = \int_T ||l(x) - \mu(T)||^2 \, \mathrm{d}x$

chrominance fidelity term

$\omega_\varphi(T) = \sum_i \int_T ||c_i(x) - \mu_i(T)||^2 \, \mathrm{d}x$

Contour length

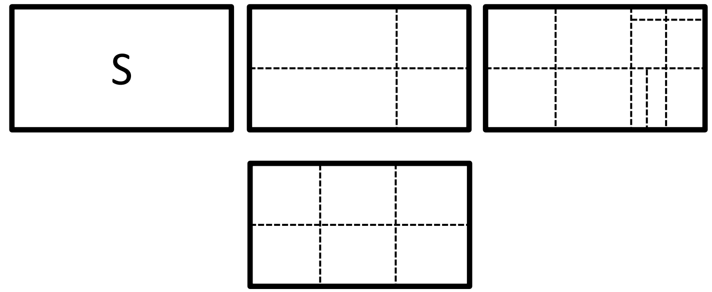
$\omega_\partial(T_k) = \partial T_k$

$$\omega(\pi(S), \lambda) = \sum_{1 \leq k \leq p} \omega_\varphi(T_k) + \lambda \sum_{1 \leq k \leq p} \omega_\partial(T_k)$$

λ fixed to have partition with same coding cost

Another example: color and texture

Initial Image



Partition with least variation in component sizes

$$\omega(\pi(S),\lambda) = \sum_{1\leq k\leq p} \omega_{\varphi}(T_k) + \lambda \sum_{1\leq k\leq p} \omega_{\partial}(T_k) + \mu\omega_{\rho}(T_k)$$

$$\omega_{\rho}(T) = |T| - \left(\frac{\sum(|T_i|)}{|\pi(S)|}\right)^2$$ Texture: deviation from average sibling size

Another example: color and texture

Initial Image

$$\omega(\pi(S),\lambda) = \omega_{\varphi}(\pi(S)) + \lambda\omega_{\partial}(\pi(S)) + \mu\omega_{\rho}(\pi(S))$$

High μ

Low μ

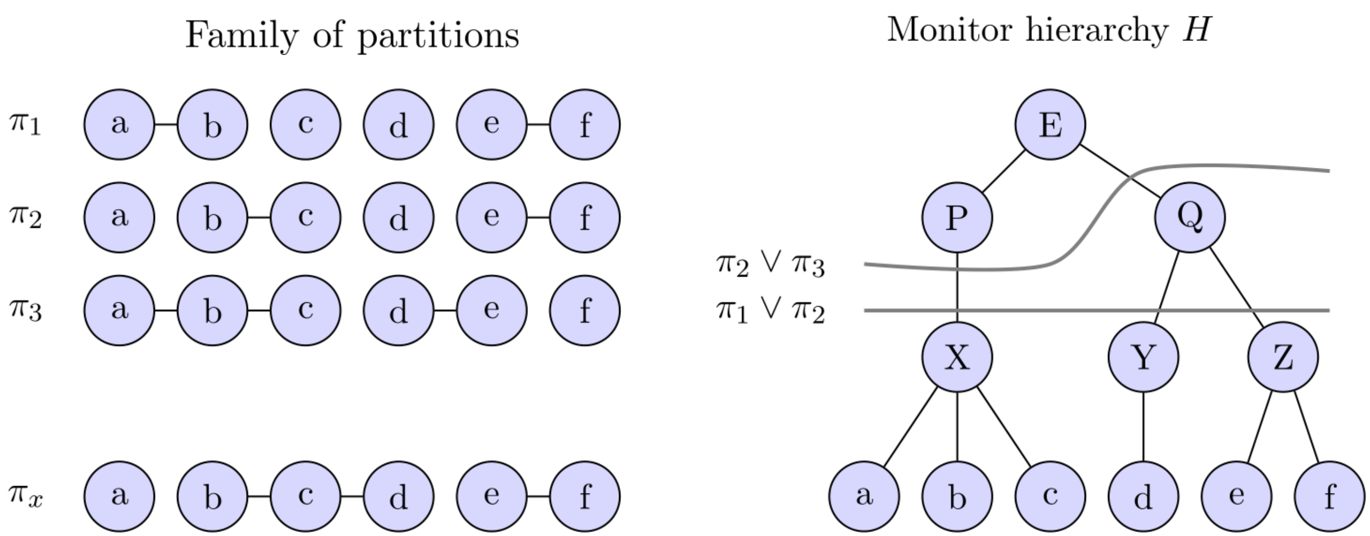
Overview

- Notations and structures
- Review on Optimization on Hierarchies
- Dynamic Programming
- **Braids of Partitions**
  - Composing Hierarchies
  - Binary net opening
  - h-increasingness for Braids
  - Braid Dynamic Program
- Energetic Order and Energetic-Lattices
- Constrained Optimization

Braid of Partitions

A Braid B with a monitoring hierarchy H, is a family of partitions, where the refinement supremum between two partitions in B is an element of H.

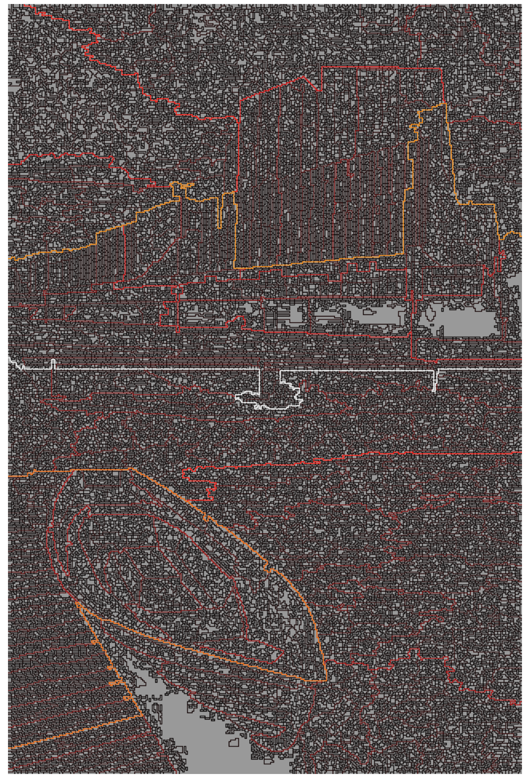
$$\forall\ \pi_1, \pi_2 \in B \quad \Rightarrow \quad \pi_1 \vee \pi_2 \in H \setminus \{E\}$$



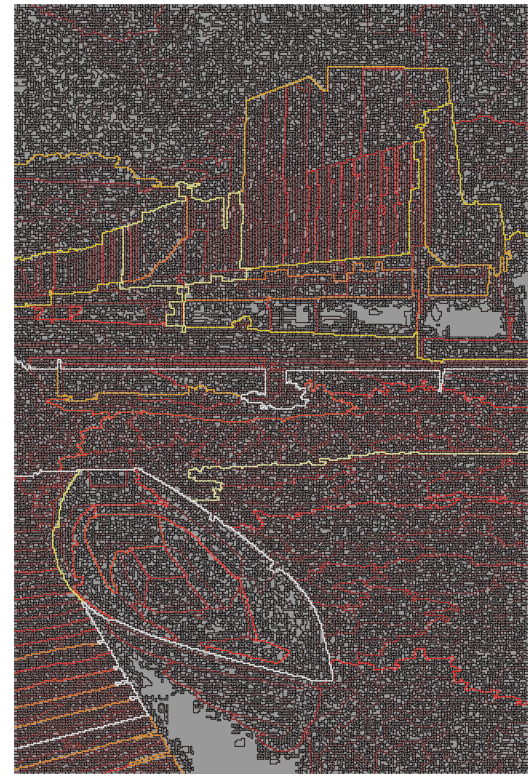
$B_1 = \{\pi_1, \pi_2, \pi_3\}$ ✓ $\quad B_2 = B_1 \cup \pi_x$ ✗

Braid: Composing hierarchies

Input Image

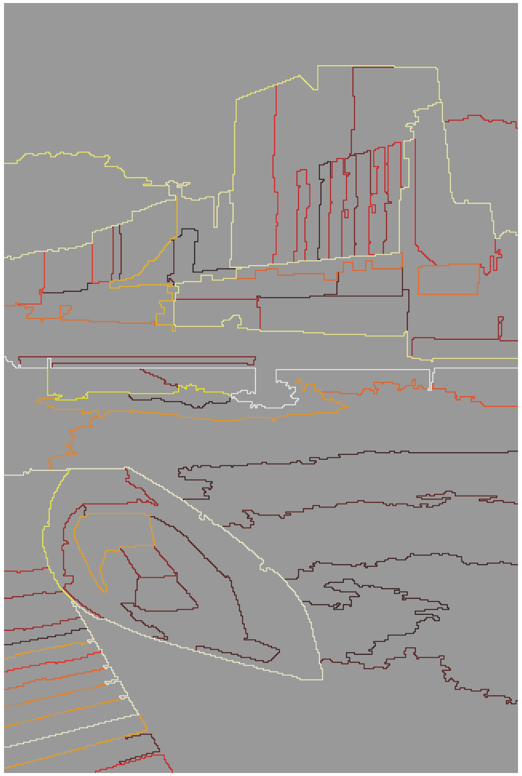


Watershed hierarchy
(Area Attribute)

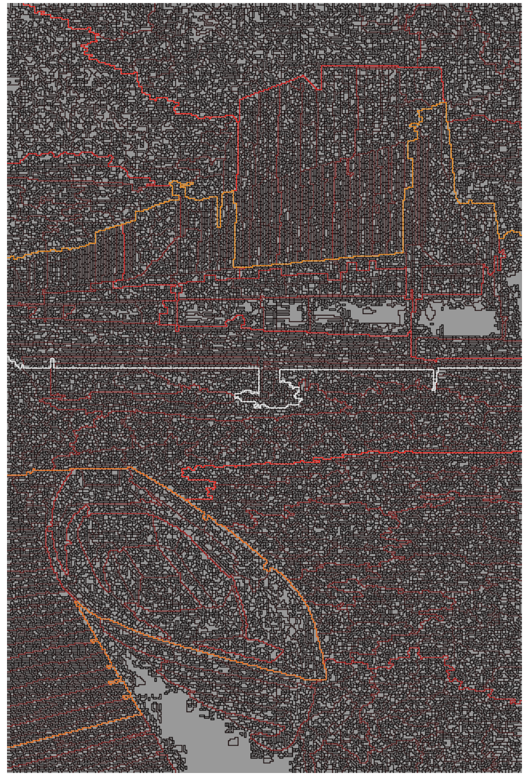


Watershed hierarchy
(Volume Attribute)

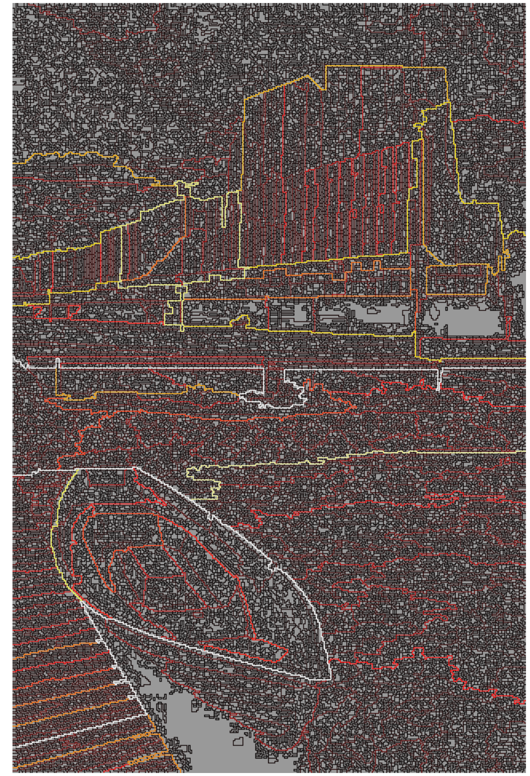
Braid: Composing hierarchies



Monitor Hierarchy



Watershed hierarchy
(Area Attribute)

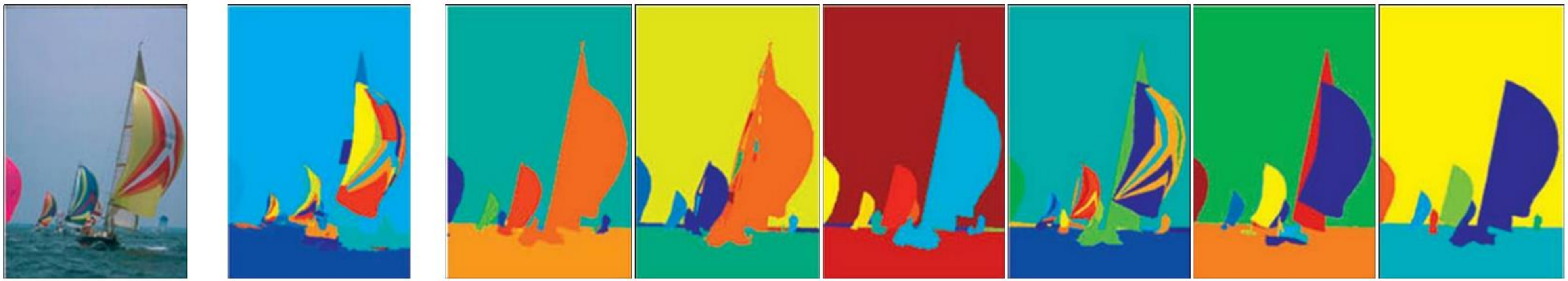


Watershed hierarchy
(Volume Attribute)

Why Braids

- Uncertain partition boundaries $\implies$ many possible partial partitions
- Multivariate segmentations
- Composition of hierarchical segmentations
- The dynamic program works for the family of braids, and ensures better infimum for over composition of hierarchies with non-trivial monitors.

No single GT is a refinement of the mean-shift segmentation, but their suprema are!

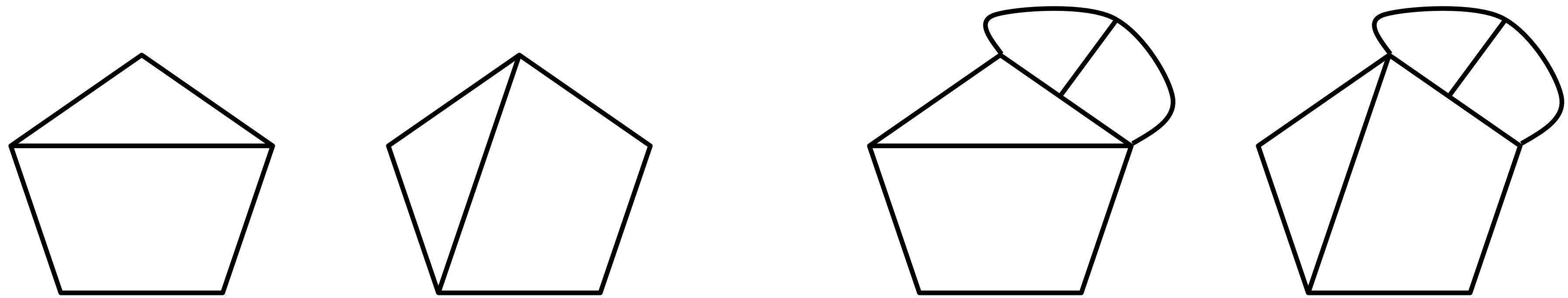


Input MeanShift Segmentation ⟵ Berkeley Expert Ground truths ⟶

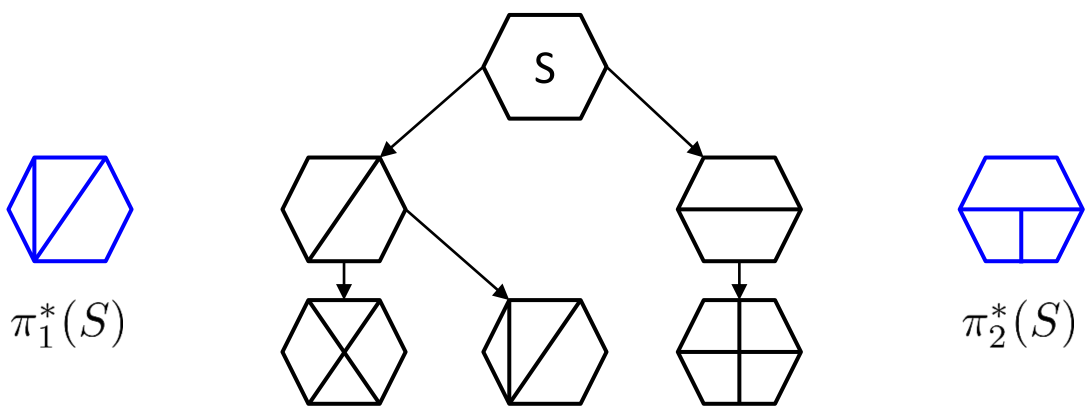
[Unnikrishnan et al. 2007]

h-increasingness on BOP

$\pi_1(S) \quad \nleq \quad \pi_2(S) \qquad\qquad \pi_1(S) \sqcup \pi_0 \quad \nleq \quad \pi_2(S) \sqcup \pi_0$



$$\omega(\pi_1(S)) \leq \omega(\pi_2(S)) \quad \Rightarrow \quad \omega(\pi_1(S) \sqcup \pi_0) \leq \omega(\pi_2(S) \sqcup \pi_0)$$



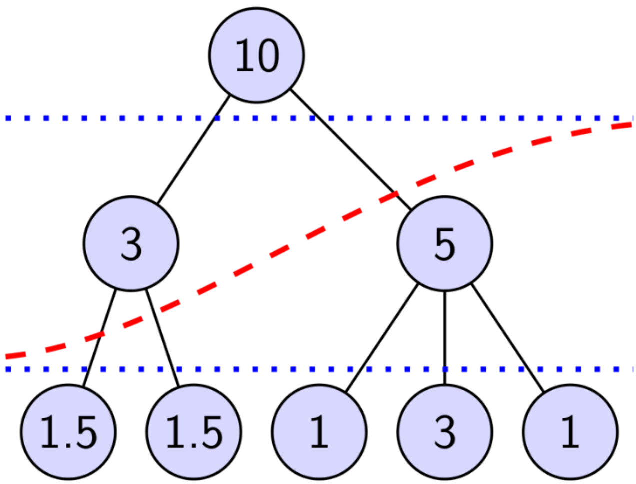
[PhD Thesis]

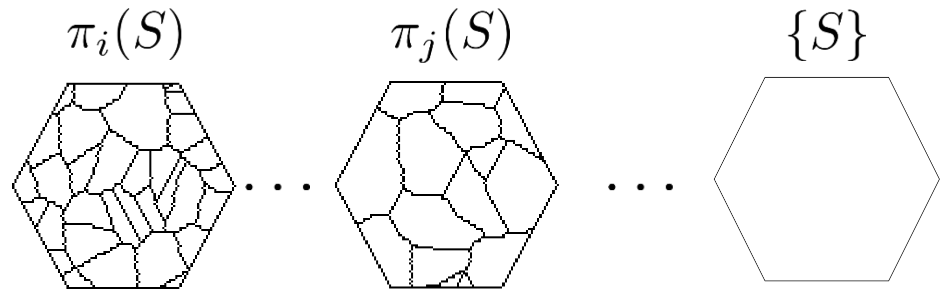
Overview

- Notations and structures
- Review on Optimization on Hierarchies
- Dynamic Programming
- Braids of Partitions
- **Energetic-Lattices**
  - Singular Energies
  - Energetic Order
  - Energetic Lattices
  - Scale-Increasingness
- Constrained Optimization
- Conclusion

Uniqueness and Singular Energy

[Kiran-Serra PR 2013]



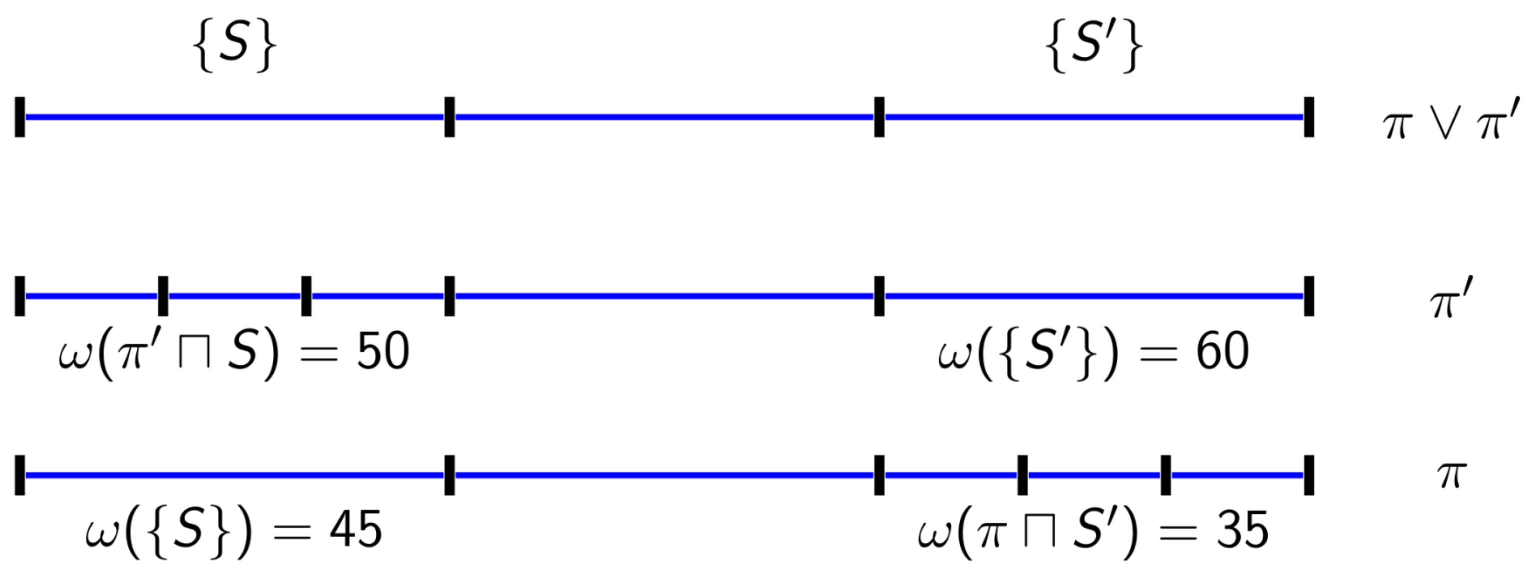


$$\omega(\{S\}) \neq \omega(\pi_i(S)), \forall i \in [1, n]$$

Various authors indirectly use the singularity condition for a unique solution.

Energetic Order

[Kiran-Serra PR 2013]



$$\pi \preceq_\omega \pi' \;\Leftrightarrow\; \forall S \in \pi \vee \pi' \text{ we have } \omega(\pi \sqcap \{S\}) \leq \omega(\pi' \sqcap \{S\})$$

Energetic Lattice

[Kiran-Serra PR 2013]

- The energetic lattice $(\preceq_\omega, \vee_\omega)$ derives from the energetic order.
- Existence of unique solution when ω singular.
- local minimum $\implies$ global minimum.
- Given ω and the family of partitions $\Pi(E, B)$ generate an energetic lattice iff ω is singular.

Scale Increasing Energies

[Kiran-Serra PR 2013]

A family $\{\omega(\lambda), \lambda \in \overline{\mathbb{R}}\}$ of energies on $\mathcal{D}(E)$ is scale increasing when:

$$\lambda \leq \mu \ \text{ and } \ \omega(\{S\}, \lambda) \leq \omega(a, \lambda) \Rightarrow \ \omega(\{S\}, \mu) \leq \omega(a, \mu), \quad S \in \mathcal{P}(E), \ a \sqsubseteq \{S\}$$

These energies produce a chain of λ-cuts which increase with λ.

- λ-Set is a descriptor dependent on energy
- λ-Set also provides λ's to perform constrained optimization.

Hierarchy of optimal cuts

[Kiran-Serra PR 2013]

Given a parametrized energy: $\{\omega(\lambda, \pi), \lambda > 0\}$.

A family $\{\omega_\lambda, \lambda > 0\}$ is said to be climbing when:

- $\{\omega(\pi, \lambda)$ is scale increasing,
- $\forall \lambda, \{\omega_\lambda\}$ is singular and h-increasing.

Then, λ-cuts $\{\pi^*(\lambda_1) \leq \pi^*(\lambda_2)\}$, for $\lambda_1 \leq \lambda_2$ produce a hierarchy.

Scale-Increasingness: Example Hierarchy

Input Image

Input Hierarchy (UCM)

$\lambda = 0$

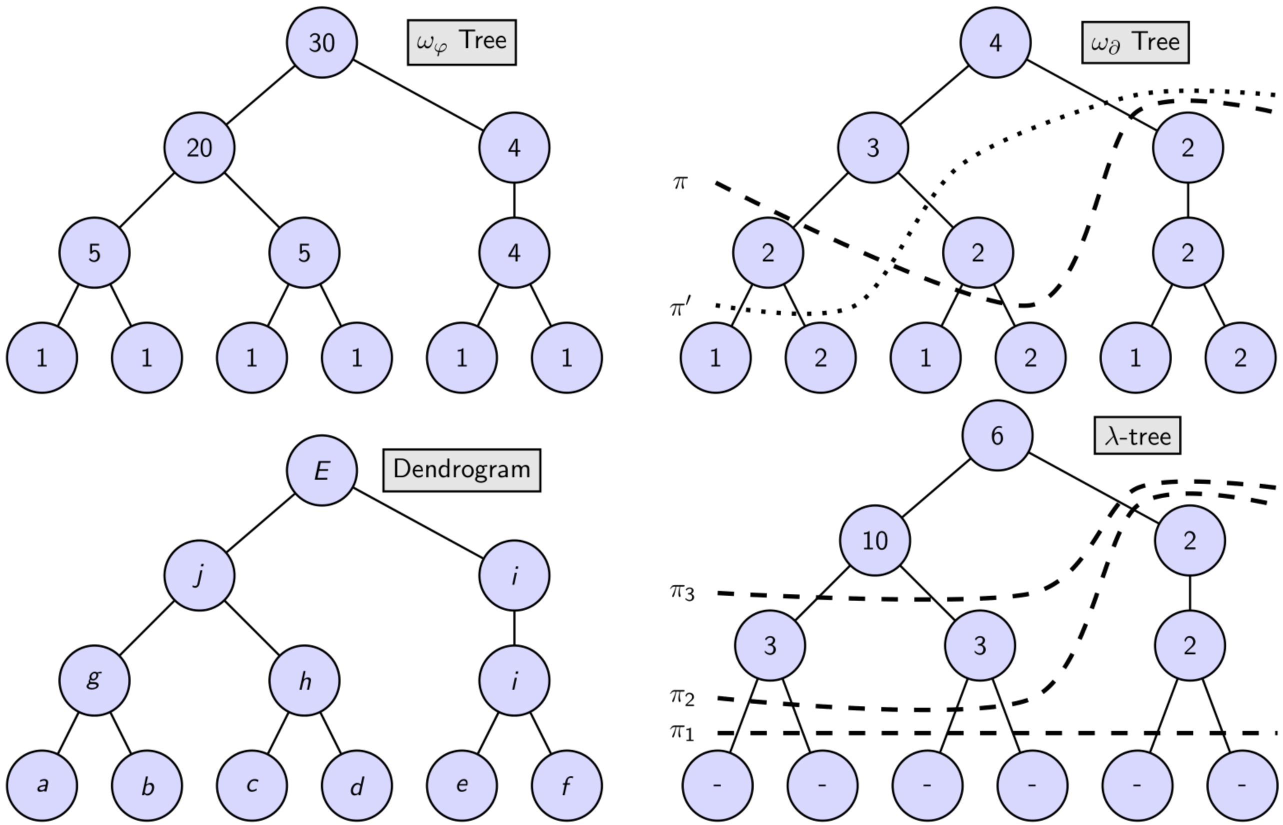
$\lambda = 400$

$\lambda = 10000$

Overview

- Notations and structures
- Review on Optimization on Hierarchies
- Dynamic Programming
- Braids of Partitions
- Energetic-Lattices
- **Constrained Optimization**
  - Counter Example: λ-cuts are lower bounds
  - Recall on Lagrangian Primal and Dual Problems
  - Guigues-Salembier search dual domain
  - Energetic Lattice based Constrained Optimization
- Conclusion

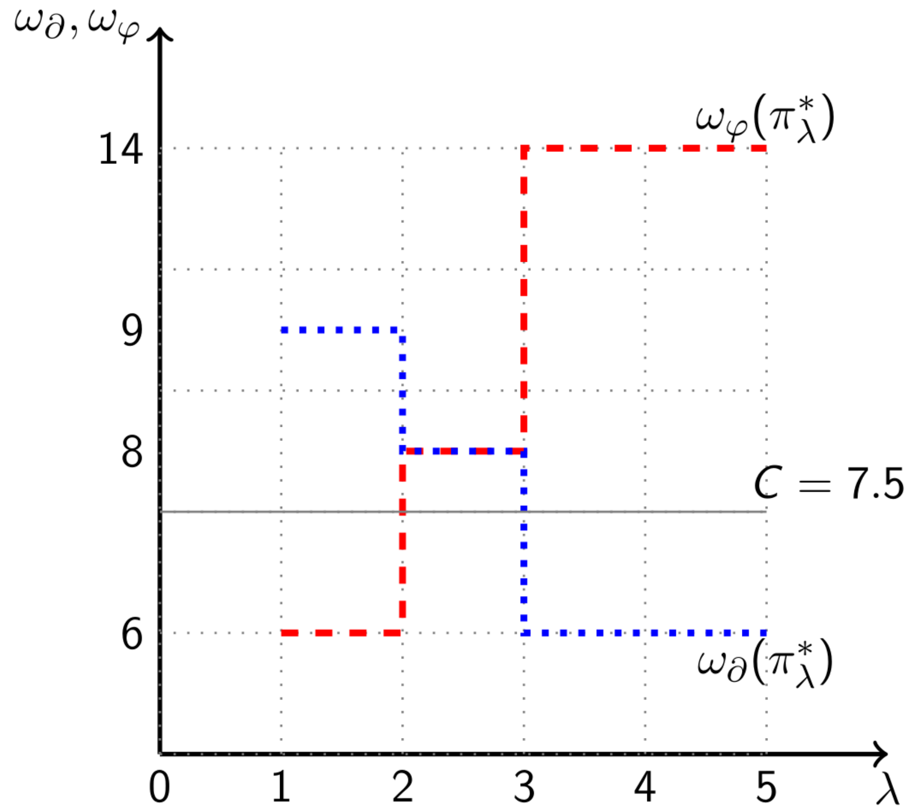
Scale-sets are Upper bounds



$$\forall \text{ Parents } \lambda = -\frac{\Delta\omega_\varphi}{\Delta\omega_\partial}$$

λ-cuts π_1, π_2, π_3 for $\lambda = 1, 2, 3$ **Vs** Minimal Cuts π, π'

Scale-sets are Upper bounds



- Other Cut Energies: $\omega_\varphi(\pi') = 11\ \omega_\partial(\pi') = 7$
- ω_∂ is never equal to the cost $C = 7.5$ at any time.

Observations

- Lack of $C \rightarrow \lambda$ mapping: For a given cost $\omega_\partial \leq C$ one is not assured a corresponding multipler λ.
- Uniqueness is lost, even when ω_φ is strictly h-increasing.
- $\pi(\lambda^*)$ is only the upper-bound for a given C.
- $\mid \omega_\partial(\pi^*(\lambda^*)) - C \mid$ gives no information about $\mid \omega_\varphi(\pi^*(\lambda^*)) - \omega_\varphi(\pi) \mid$ where π is a constrained minimal cut.

Perturbed Primal Problem

$$
\begin{array}{ll}
\underset{\pi \in \Pi(E,B)}{\text{minimize}} & \omega_\varphi(\pi) \\
\text{subject to} & \omega_\partial(\pi) \leq 0,
\end{array}
$$

Given the Lagrangian $\omega(\pi, \lambda) = \omega_\varphi(\pi) + \lambda \cdot \omega_\partial(\pi)$, and multipler λ:

let $\pi^*(\lambda)$ minimize the Lagrangian.

$\pi^*(\lambda)$ solves the constrained problem, where the constraint is **λ-dependent**

$$
\begin{array}{ll}
\underset{\pi \in \Pi(E,B)}{\text{minimize}} & \omega_\varphi(\pi) \\
\text{subject to} & \omega_\partial(\pi) \leq \omega_\partial(\pi^*(\lambda)),
\end{array}
$$

[Everett 1963]

Optimal Dual Parameter

- Changing to the dual domain does not aid solving a combinatorial problem, it provides some upper-bound.

- [Salembier-Garrido 2001] searches λ-cuts to approximate constraint value $\omega_\partial(\pi) \approx C$

- [Guigues 2003] produces causal scale-set description of image using defined energy.

- By choosing multiplier λ from possible λ's one achieves an optimum for a constraint value is λ-dependent (Everett's Theorem [Everett 1963]).

Overview

- Notations and structures
- Review on Optimization on Hierarchies
- Dynamic Programming
- Braids of Partitions
- Energetic-Lattices
- Constrained Optimization
- Conclusion

Conclusion

- Expanding solution space to Braids, which is the largest family for which the energetic lattice structure holds.
- Singularity: Necessary & Sufficient conditions to find unique solutions
- h-increasingness of energies preserves DP substructure
- Energetic Lattice Infimum characterizes minimal cut