Discrete Inference and Learning

Lecture 2
Primal-dual schema, dual decomposition

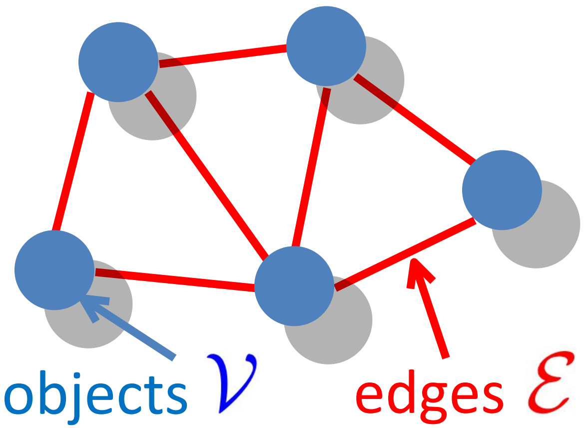
Demian Wassermann
demian.wassermann@inria.fr

Slides progressivevly constructed by
N. Komodakis, Y. Tarabalka, G. Charpiat, and me

Part I
Recap: MRFs and Convex Relaxations

Discrete MRF setting

- Given:
  - Objects $\mathcal{V}$ from a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$
  - The edges are undirected
  - A probability function



$$P(\mathcal{G}) = \Pi_{v \in \mathcal{V}} P(v | \mathcal{N}_v)$$

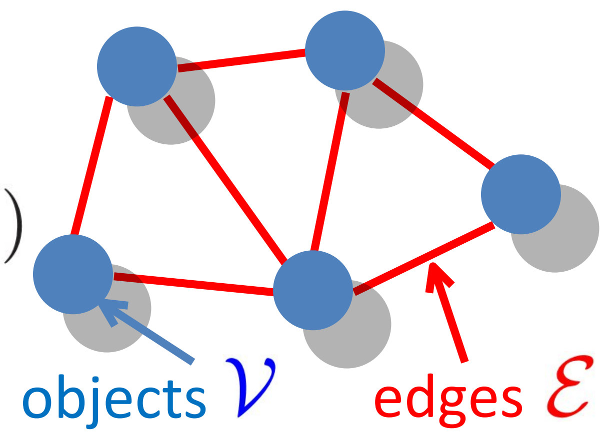
- We can then state a wide range of problems on finding a set of assignments to maximize the probability P

$$\arg\max_{\mathcal{X}} P(\mathcal{G}|X) = \Pi_{v_p \in \mathcal{V}} P(v = x_p | \mathcal{N}_v)$$

$$\arg\min_{\mathcal{X}} - \log P(\mathcal{G}|X) = \sum_p g_p(x_p) + \sum_{p,q : v_q \in \mathcal{N}_{v_p}} f_{p,q}(x_p, x_q)$$

Discrete MRF optimization

- Given:
  - Objects $\mathcal{V}$ from a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$
  - If X can be an assignment of discrete or continuous values



- Assign labels (to objects) that minimize MRF energy:

$$\arg\min_{\mathcal{X}} -\log P(\mathcal{G}|X) = \sum_{p} \underbrace{g_p(x_p)}_{\text{unary potential}} + \sum_{p,q:v_q \in \mathcal{N}_{v_p}} \underbrace{f_{p,q}(x_p, x_q)}_{\text{pairwise potential}}$$

Continuous MRF optimization

- Can be seen as a particular case of several machine learning scenarios with a specific prior

- Examples in computer vision and neuroimaging

  - Restoration
    Functional brain activation
    optical flow
    ...

- and beyond, connections with graph deep learning
- Comfortable way to express spatial priors

- Really powerful sound formulation

Continuous MRF optimization as common Regression (ML?) Problems

$$\arg\min_{\mathcal{X}} -\log P(\mathcal{G}|X) = \sum_p g_p(x_p) + \sum_{p,q:v_q \in \mathcal{N}_{v_p}} f_{p,q}(x_p, x_q)$$

Regularized $\quad \arg\min_{\mathcal{X}} \quad d(Y, g(\mathcal{X})) + f(\mathcal{X})$

Ridge /Tik $\quad \arg\min_{\mathcal{X}} \quad \|Y - A\mathcal{X}\|_2^2 + \lambda \|\Gamma \mathcal{X}\|_2^2$

Lasso $\quad \arg\min_{\mathcal{X}} \quad \|Y - A\mathcal{X}\|_2^2 + \lambda \|\Gamma \mathcal{X}\|_1^1$

Elastic Net $\quad \arg\min_{\mathcal{X}} \quad \|Y - A\mathcal{X}\|_2^2 + \lambda_l \|\Gamma \mathcal{X}\|_1^1 + \lambda_r \|\Gamma \mathcal{X}\|_2^2$

…

These can be solved through quadratic programming
[Hastie et al, Elements of Statistical Learning 2017]

Continuous MRF optimization as common Regression (ML?) Problems

$$\arg\min_{\mathcal{X}} -\log P(\mathcal{G}|X) = \sum_p g_p(x_p) + \sum_{p,q: v_q \in \mathcal{N}_{v_p}} f_{p,q}(x_p, x_q)$$

$$\begin{aligned}\arg\min_{\mathcal{X}} \quad & \|Y - A\mathcal{X}\|_2^2 + \lambda\|\Gamma\mathcal{X}\|_2^2 \\ =& Y^T Y + \mathcal{X}^T A^T A \mathcal{X} - 2Y^T A\mathcal{X} + \lambda \mathcal{X}^T \Gamma^T \Gamma \mathcal{X} \\ =& \frac{1}{2}\mathcal{X}^T (A^T A + \lambda \Gamma^T \Gamma)\mathcal{X} - Y^T A \mathcal{X}\end{aligned}$$

Discrete MRF optimization

- Extensive research for more than 30 years

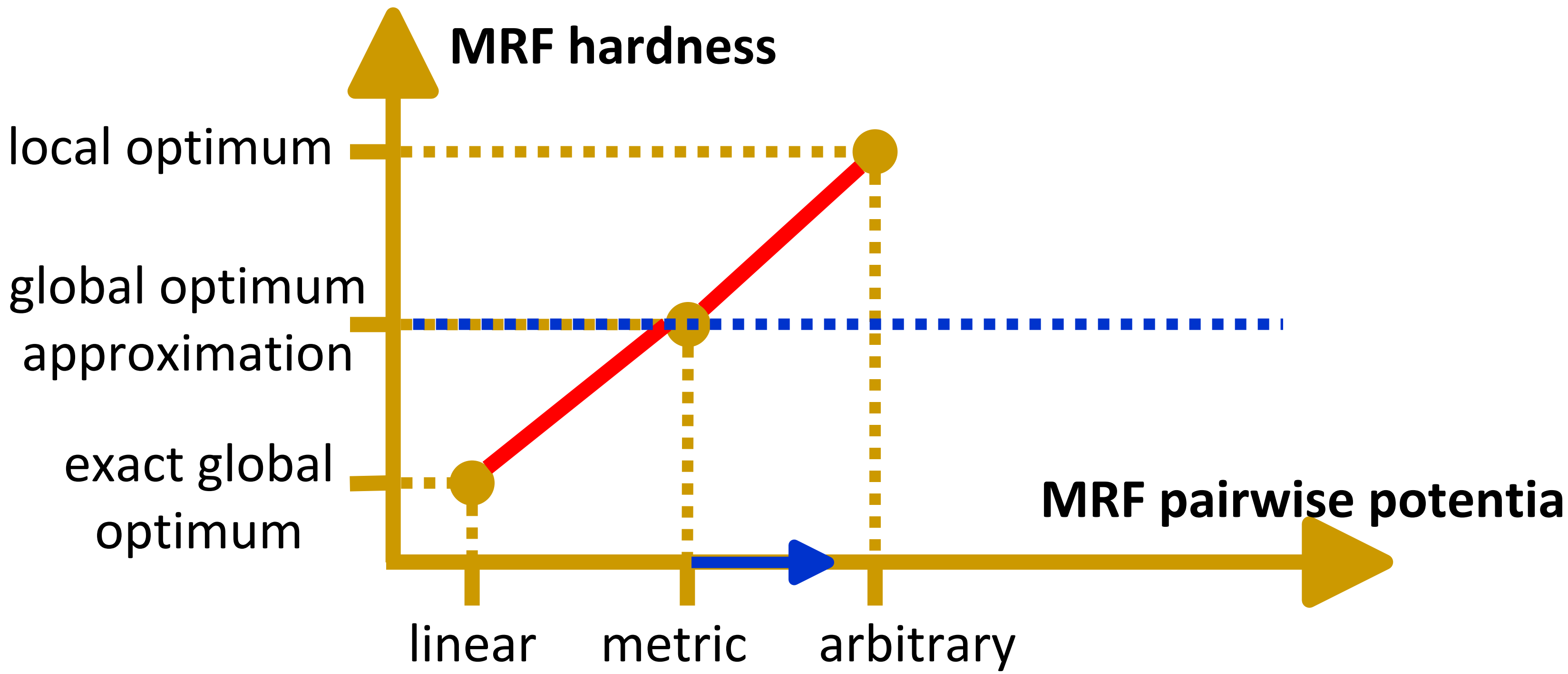
- MRF optimization ubiquitous in computer vision
  - segmentation, stereo matching, optical flow, image restoration, image completion, object detection/localization
    ...

- and beyond
  - medical imaging, computer graphics, digital communications, physics…

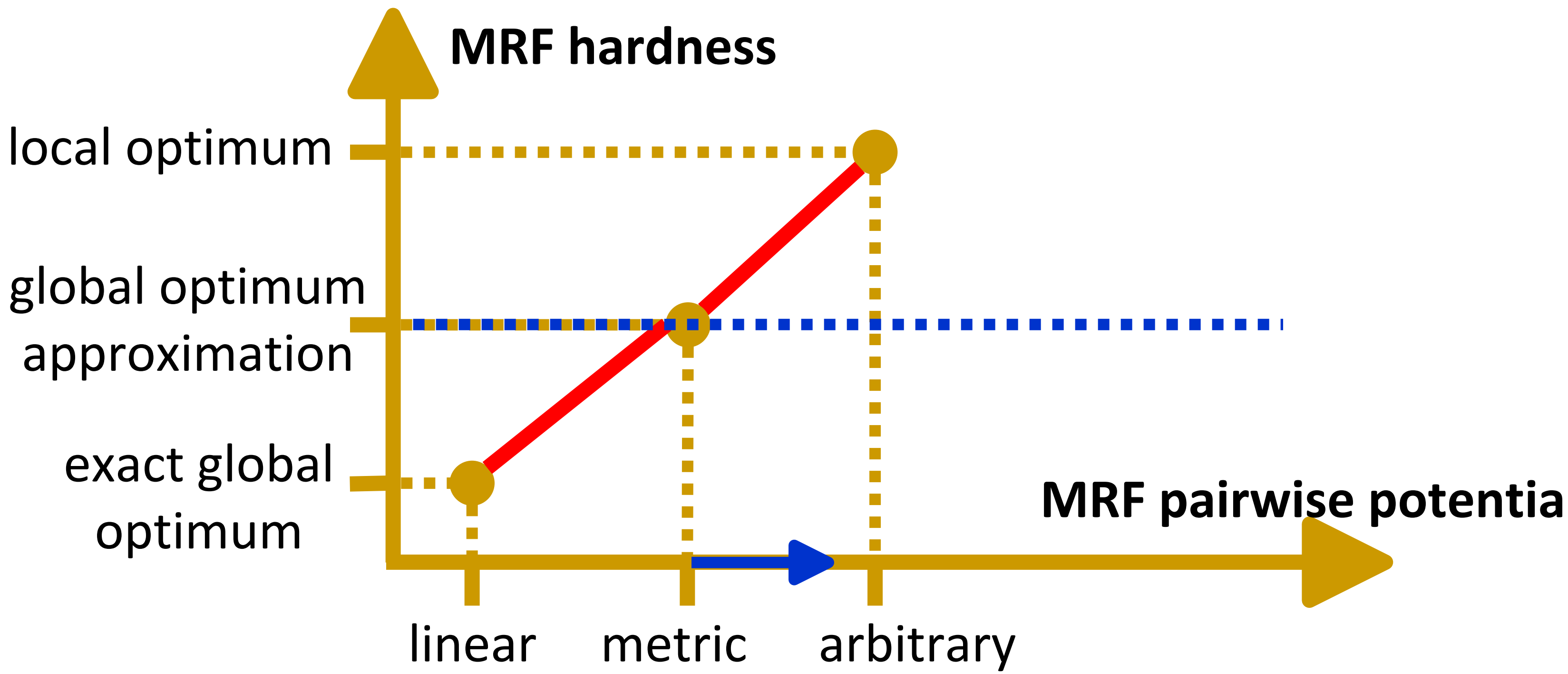
- Really powerful formulation

How to handle MRF optimization?

- Unfortunately, discrete MRF optimization is extremely hard (a.k.a. NP-hard)
  - E.g., highly non-convex energies



How to handle MRF optimization?



We want:
Move right in the horizontal axis,
And remain low in the vertical axis
(i.e. still be able to provide approximately optimal solution)

We want to do it efficiently (fast)!

MRFs and Optimization

- Deterministic methods
  - Iterated conditional modes
- Non-deterministic methods
  - Mean-field and simulated annealing
- Graph-cut based techniques such as alpha-expansion
  - Min cut/max flow, etc.
- Message-passing techniques
  - Belief propagation networks, etc.

- We would like to have a method which provides theoretical guarantees to obtain a good solution
- Within a reasonably fast computational time

Discrete optimization problems

$$\min_x f(x)$$ (optimize an objective function)

$$\text{s.t. } x \in C$$ (subject to some constraints)

this is the so called **feasible set**, containing all x satisfying the constraints

- Typically x lives on a very high dimensional space

How to handle MRF optimization?

- Unfortunately, discrete MRF optimization is extremely hard (a.k.a. NP-hard)
  - E.g., highly non-convex energies

- So what do we do?
  - Is there a principled way of dealing with this problem?

- Well, first of all, we don't need to panic.
  Instead, we have to stay calm and **RELAX!**

- Actually, this idea of relaxing may not be such a bad idea after all…

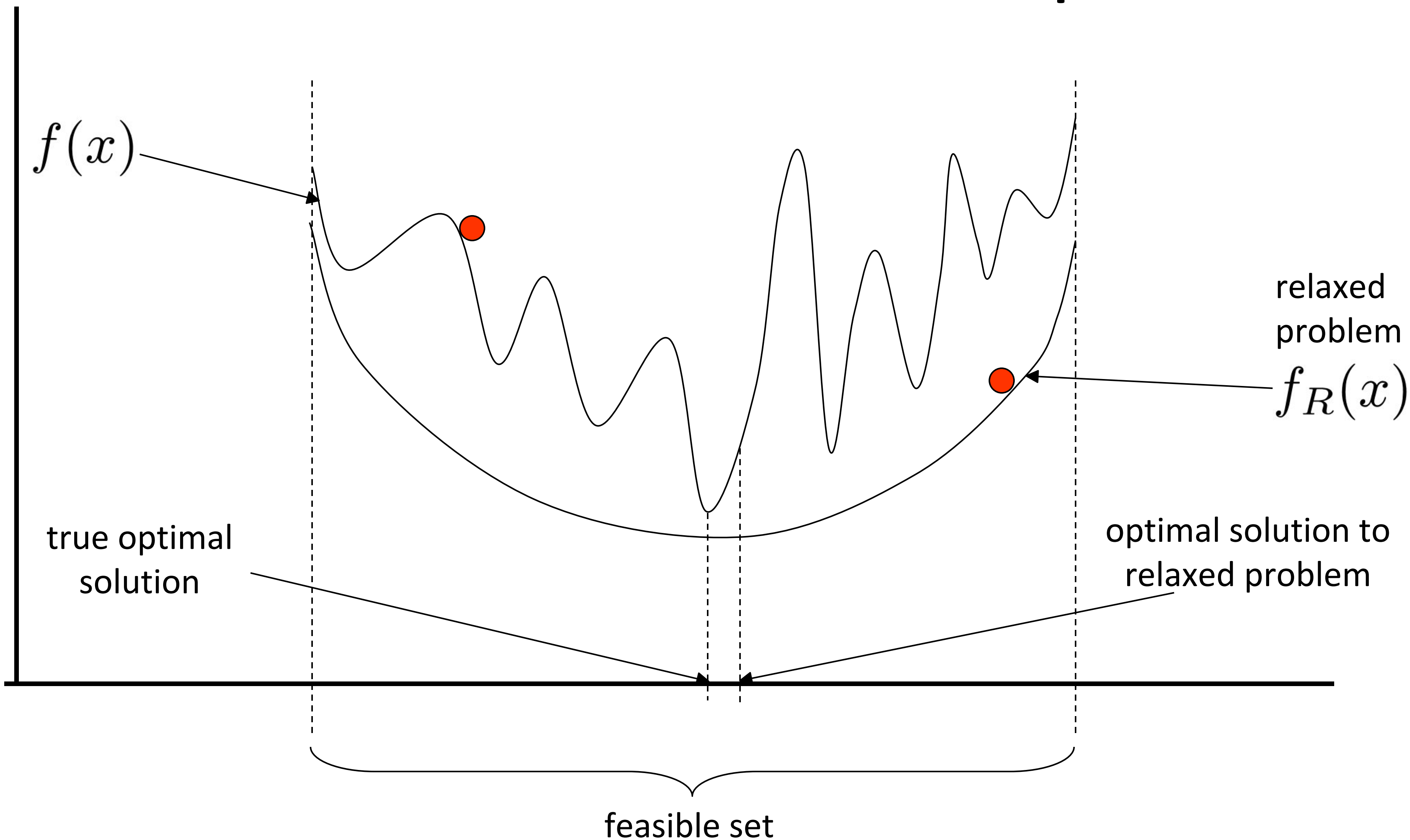
The relaxation technique

- Very successful technique for dealing with difficult optimization problems

- It is based on the following simple idea:
  - try to approximate your original difficult problem with another one (the so called **relaxed problem**) which is easier to solve

- Practical assumptions:
  - Relaxed problem must always be easier to solve
  - Relaxed problem must be related to the original one

The relaxation technique



How do we find easy problems?

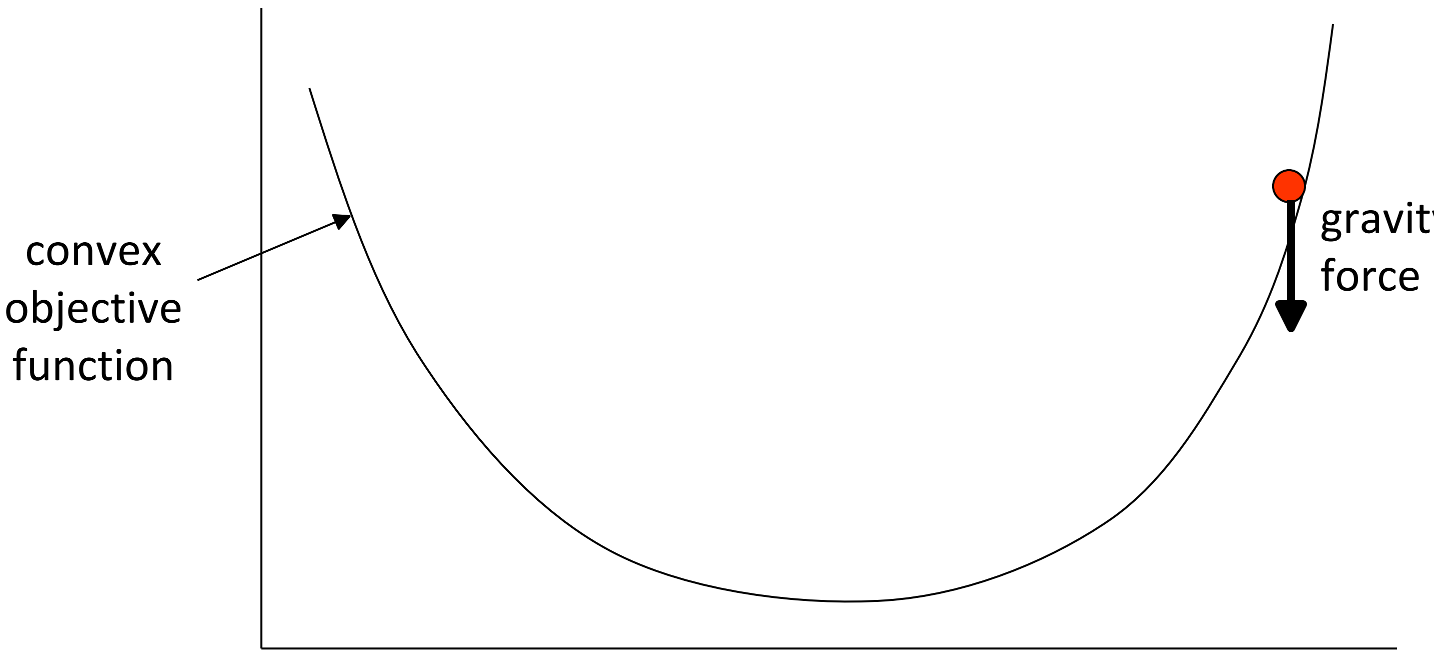
- Convex optimization to the rescue

  "...in fact, the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity" - R. Tyrrell Rockafellar, in SIAM Review, 1993

- Two conditions for an optimization problem to be convex:
  - convex objective function
  - convex feasible set

Why is convex optimization easy?

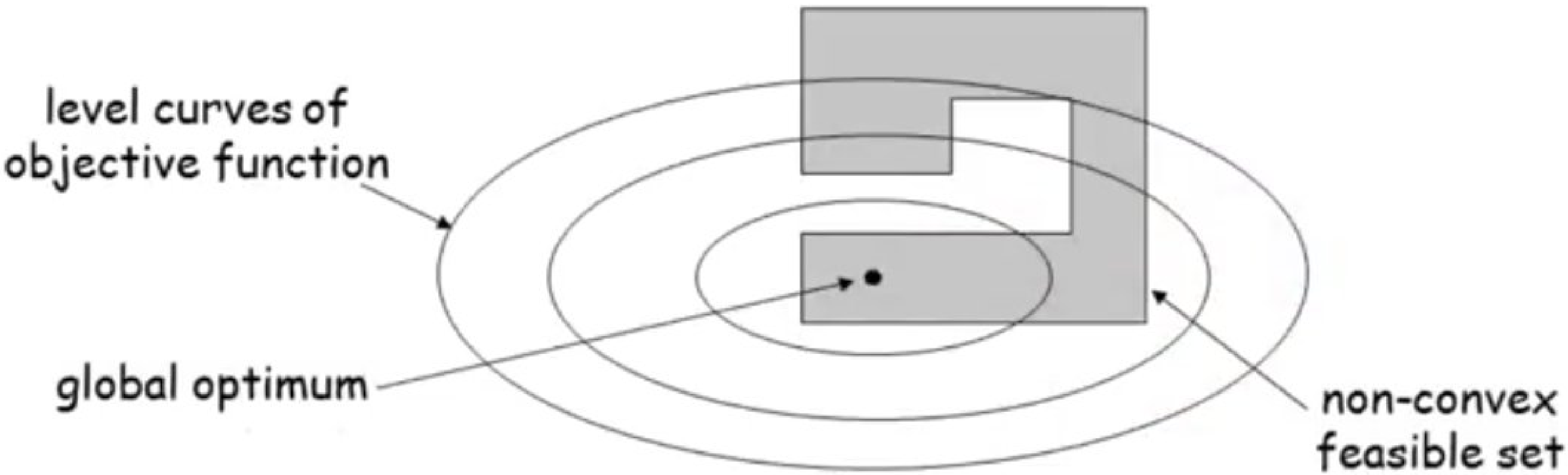
- Because we can simply let gravity do all the hard work for us



- More formally, we can let gradient descent do all the hard work for us

Why do we need the feasible set to be convex as well?

- Because, otherwise we may get stuck in a local optimum of we simply "follow gravity"



How do we get a convex relaxation?

- By dropping some constraints
  (so that the enlarged feasible set is convex)

- By modifying the objective function
  (so that the new function is convex)

- By combining both of the above

Linear programming (LP) relaxations

- Optimize linear function subject to linear constraints, i.e.:

$$\begin{aligned} &\min \ \mathbf{c}^T \mathbf{x} \\ &\text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned}$$

- Very common form of a convex relaxation, because:
  - Typically leads to very efficient algorithms (important due to large scale nature of problems in computer vision)
  - Also often leads to combinatorial algorithms
  - Surprisingly good approximation for many problems

Geometric interpretation of LP

Max Z = 5X + 10Y

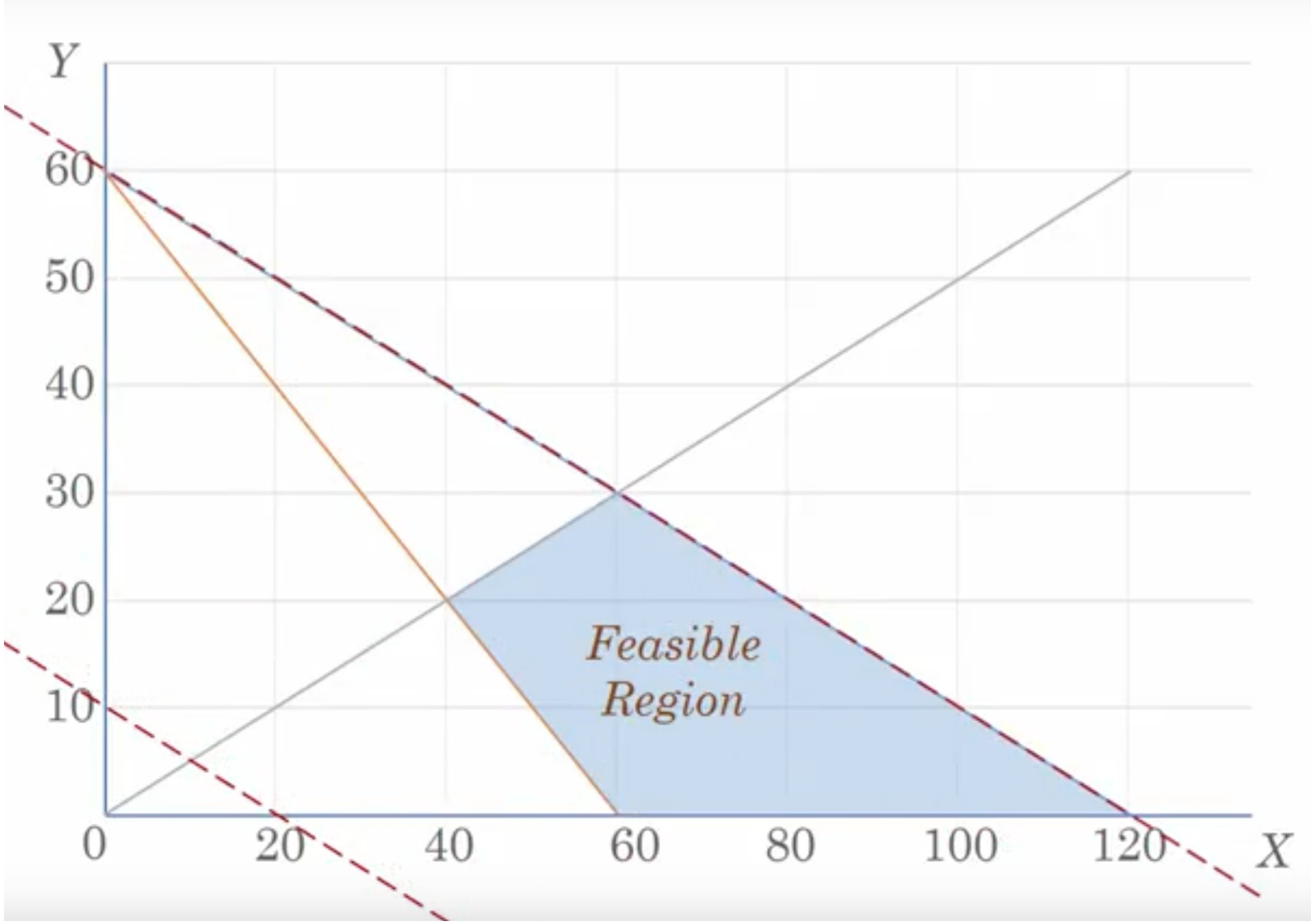
s.t.

X + 2Y <= 120

X + Y >= 60

X – 2Y >= 0

X, Y >= 0



MRFs and Linear Programming

- Tight connection between MRF optimization and Linear Programming (LP) recently emerged

- Active research topic with a lot of interesting work:

  – MRFs and LP-relaxations [Schlesinger] [Boros] [Wainwright et al. 05] [Kolmogorov 05] [Weiss et al. 07] [Werner 07] [Globerson et al. 07] [Kohli et al. 08]…

  – Tighter/alternative relaxations [Sontag et al. 07, 08] [Werner 08] [Kumar et al. 07, 08]

MRFs and Linear Programming

- E.g., state of the art MRF algorithms are now known to be directly related to LP:
  - <u>Graph-cut based techniques such as *a*-expansion:</u>
    generalized by primal-dual schema algorithms (Komodakis et al. 05, 07)
  - <u>Message-passing techniques:</u>
    further generalized by Dual-Decomposition (Komodakis 07)

- The above statement is more or less true for almost all state-of-the-art MRF techniques

Part II
Primal-dual schema

The primal-dual schema

- Highly successful technique for exact algorithms. Yielded exact algorithms for cornerstone combinatorial problems:

matching	network flow
minimum spanning tree	minimum branching
shortest path	...

- Soon realized that it's also an extremely powerful tool for deriving approximation algorithms [Vazirani]:

set cover	steiner tree
steiner network	feedback vertex set
scheduling	...

The primal-dual schema

- **Conjecture:**

Any approximation algorithm can be derived using the primal-dual schema

(has not been disproved yet)

The primal-dual schema

- Say we seek an optimal solution x* to the following integer program (this is our primal problem):

$$\begin{aligned} &\min \ \mathbf{c}^T\mathbf{x} \\ &\text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \in \mathbb{N} \end{aligned}$$

(NP-hard problem)

- To find an approximate solution, we first relax the integrality constraints to get a primal & a dual linear program:

primal LP: $$\begin{aligned} &\min \ \mathbf{c}^T\mathbf{x} \\ &\text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \end{aligned}$$

dual LP: $$\begin{aligned} &\max \ \mathbf{b}^T\mathbf{y} \\ &\text{s.t.} \quad \mathbf{A}^T\mathbf{y} \leq \mathbf{c} \end{aligned}$$

Duality

$$\begin{array}{lrcrcrl}
\text{min} & 3x_1 & + & 2x_2 & + & 4x_3 & \\
\text{subject to} & & & & & & \\
 & 2x_1 & + & x_2 & + & & \geq 2 \\
 & 2x_1 & - & x_2 & + & x_3 & \geq 5
\end{array}$$

$$\begin{array}{lrcrl}
\text{max} & 2y_1 & + & 5y_2 & \\
\text{subject to} & & & & \\
 & 2y_1 & + & 2y_2 & \leq 3 \\
 & y_1 & - & y_2 & \leq 2 \\
 & & & y_2 & \leq 4
\end{array}$$

Duality

$$
\begin{array}{llcl}
\min & 3x_1 + 2x_2 + 4x_3 & & \\
\text{subject to} & & & \\
& 2x_1 + x_2 \geq 2 & & y_1 \\
& 2x_1 - x_2 + x_3 \geq 5 & & y_2
\end{array}
$$

$$
\begin{array}{ll}
\max & 2y_1 + 5y_2 \\
\text{subject to} & \\
& 2y_1 + 2y_2 \leq 3 \\
& y_1 - y_2 \leq 2 \\
& y_2 \leq 4
\end{array}
$$

Duality

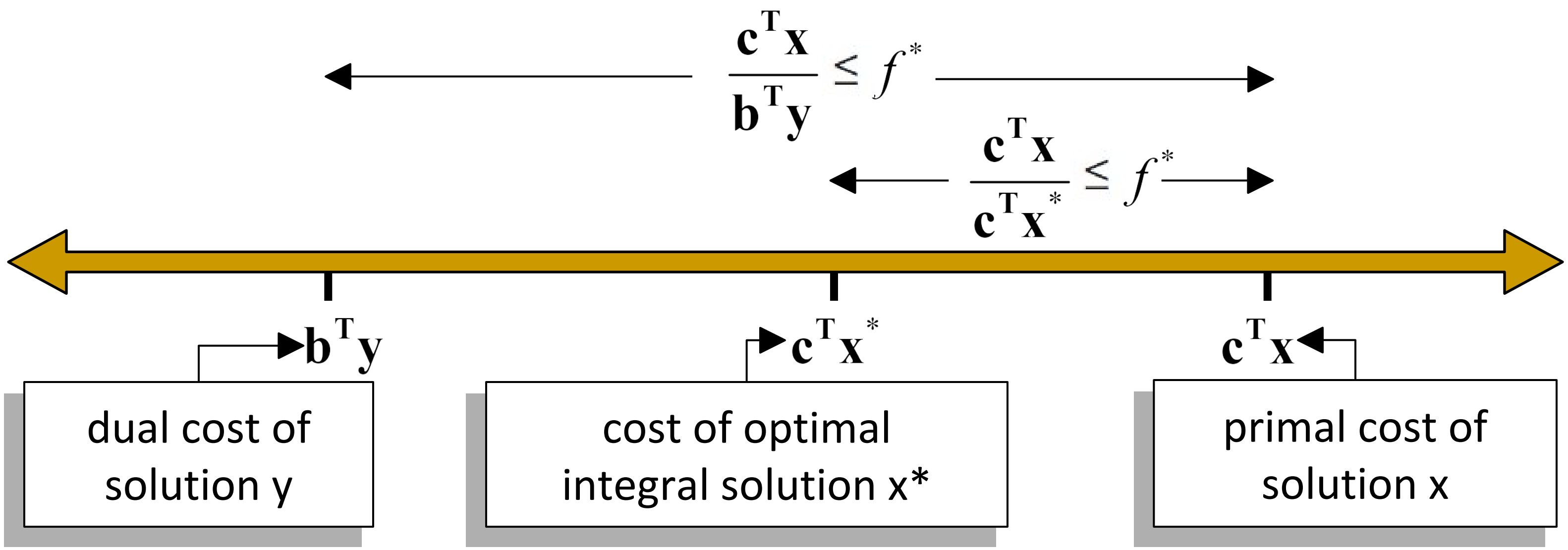
[redacted]

[redacted]

Theorem:
If the primal has an optimal solution,
the dual has an optimal solution with the same cost

The primal-dual schema

- <u>Goal:</u> find integral-primal solution x, feasible dual solution y such that their primal-dual costs are "close enough", e.g.,



Then x is an f^*-approximation to optimal solution x*

General form of the dual

$$
\begin{array}{llcll}
\min & c\,x & & & \\
\text{subject to} & & & & \\
& a_i\,x & = & b_i & (i \in E) \\
& a_i\,x & \geq & b_i & (i \in I) \\
& x_j & \geq & 0 & (j \in P) \\
& x_j & \in & \mathcal{R} & (j \in O)
\end{array}
$$

Primal

$$
\begin{array}{llcll}
\max & y\,b & & & \\
\text{subject to} & & & & \\
& y_i & \in & \mathcal{R} & (i \in E) \\
& y_i & \geq & 0 & (i \in I) \\
& yA_j & \leq & c_j & (j \in P) \\
& yA_j & = & c_j & (j \in O)
\end{array}
$$

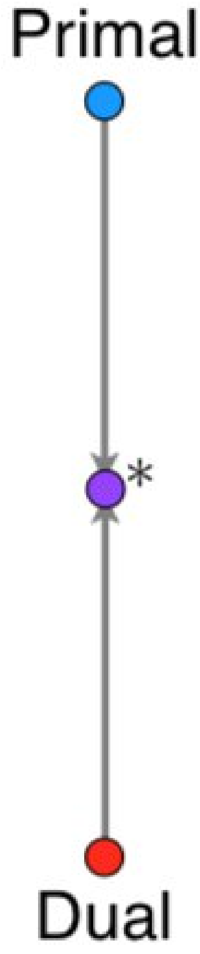
Dual

Properties of Duality

- The dual of the dual is the primal

	Finite Primal	Unbounded Primal	Infeasible Primal
Finite Dual	Yes	?	?
Unbounded Dual	?	?	?
Infeasible Dual	?	?	?

Primal and Dual

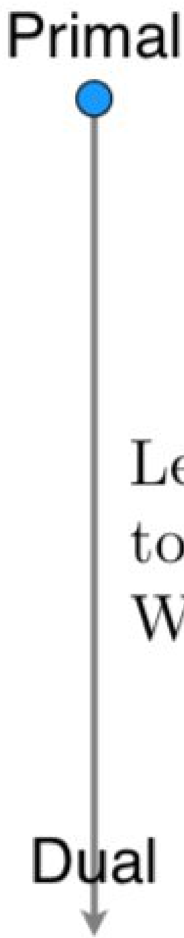


Properties of Duality

- The dual of the dual is the primal

	Finite Primal	Unbounded Primal	Infeasible Primal
Finite Dual	Yes	?	?
Unbounded Dual	?	?	?
Infeasible Dual	?	?	?

Primal and Dual



Let x and Π be feasible solutions to the primal and dual respectively. We have that $cx \geq \Pi A x \geq \Pi b$.

Properties of Duality

- The dual of the dual is primal

	Finite Primal	Unbounded Primal	Infeasible Primal
Finite Dual	Yes	?	?
Unbounded Dual	?	?	?
Infeasible Dual	?	?	?

Primal/Dual Relationships

$$
\begin{array}{ll}
\min & x_1 \\
\text{subject to} & \\
& \begin{array}{rcrcl} x_1 & + & x_2 & \geq & 1 \\ -x_1 & - & x_2 & \geq & 1 \end{array}
\end{array}
$$

infeasible primal

$$
\begin{array}{ll}
\max & y_1 + y_2 \\
\text{subject to} & \\
& \begin{array}{rcrcl} y_1 & - & y_2 & = & 1 \\ y_1 & - & y_2 & = & 0 \\ y_i & \geq & 0 & & \end{array}
\end{array}
$$

infeasible dual

Primal/Dual Relationships

$$\begin{array}{ll} \min & x_1 \\ \text{subject to} & \\ & \begin{array}{rcrcr} x_1 & + & x_2 & \geq & 1 \\ -x_1 & - & x_2 & \geq & 1 \\ x_j & \geq & 0 & & \end{array} \end{array}$$

infeasible primal

$$\begin{array}{ll} \max & y_1 + y_2 \\ \text{subject to} & \\ & \begin{array}{rcrcr} y_1 & - & y_2 & \leq & 1 \\ y_1 & - & y_2 & \leq & 0 \\ y_i & \geq & 0 & & \end{array} \end{array}$$

unbounded dual

Certificate of Optimality

- NP-complete problems
  - Certificate of feasibility
- Can you provide
  - A certificate of optimality?
- Consider now a linear program
  - Can you convince me that you have found an optimal solution?

Certificate of Optimality

primal

$$\begin{array}{ll} \text{min} & c\ x \\ \text{subject to} & \\ & Ax \geq b \\ & x_j \geq 0 \end{array}$$

dual

$$\begin{array}{ll} \text{max} & y\ b \\ \text{subject to} & \\ & yA \leq c \\ & y \geq 0 \end{array}$$

- Give me a x* that satisfies A x* ≥ b
- Give me a y* that satisfies y* A ≤ c
- Show me that c x* = y* b.

Bounding

$$
\begin{array}{lrcrcrcrcr}
\text{max} & 4x_1 & + & x_2 & + & 5x_3 & + & 3x_4 & & \\
\text{subject to} & & & & & & & & & \\
 & x_1 & - & x_2 & - & x_3 & + & 3x_4 & \leq & 1 \\
 & 5x_1 & + & x_2 & + & 3x_3 & + & 8x_4 & \leq & 55 \\
 & -x_1 & + & 2x_2 & + & 3x_3 & - & 5x_4 & \leq & 3
\end{array}
$$

- can we find an upper bound?

$$
10x_1 \quad + \quad 2x_2 \quad + \quad 6x_3 \quad + \quad 16x_4 \quad \leq \quad 110
$$

Bounding

$$
\begin{array}{lrcrcrcrcr}
\text{max} & 4x_1 & + & x_2 & + & 5x_3 & + & 3x_4 & & \\
\text{subject to} & & & & & & & & & \\
& x_1 & - & x_2 & - & x_3 & + & 3x_4 & \leq & 1 \\
& 5x_1 & + & x_2 & + & 3x_3 & + & 8x_4 & \leq & 55 \\
& -x_1 & + & 2x_2 & + & 3x_3 & - & 5x_4 & \leq & 3
\end{array}
$$

- can we find an upper bound?

$$
10x_1 + 2x_2 + 6x_3 + 16x_4 \leq 110
$$

Bounding

$$
\begin{array}{lrcrcrcrcr}
\max & 4x_1 & + & x_2 & + & 5x_3 & + & 3x_4 & & \\
\text{subject to} & & & & & & & & & \\
 & x_1 & - & x_2 & - & x_3 & + & 3x_4 & \leq & 1 \\
 & 5x_1 & + & x_2 & + & 3x_3 & + & 8x_4 & \leq & 55 \\
 & -x_1 & + & 2x_2 & + & 3x_3 & - & 5x_4 & \leq & 3
\end{array}
$$

- can we find an upper bound?

$$
\boxed{10}x_1 + \boxed{2}x_2 + \boxed{6}x_3 + \boxed{16}x_4 \leq \boxed{110}
$$

Bounding

$$
\begin{array}{lrcrcrcrcr}
\text{max} & \boxed{4}x_1 & + & \boxed{}x_2 & + & \boxed{5}x_3 & + & \boxed{3}x_4 & & \\
\text{subject to} & & & & & & & & & \\
 & x_1 & - & x_2 & - & x_3 & + & 3x_4 & \leq & 1 \\
 & 5x_1 & + & x_2 & + & 3x_3 & + & 8x_4 & \leq & 55 \\
 & -x_1 & + & 2x_2 & + & 3x_3 & - & 5x_4 & \leq & 3 \\
 & & & & & & & & & \\
 & \boxed{10}x_1 & + & \boxed{2}x_2 & + & \boxed{6}x_3 & + & \boxed{16}x_4 & \leq & \boxed{110}
\end{array}
$$

- can we find an upper bound?

Bounding

$$
\begin{array}{lrcrcrcrcr}
\text{max} & \boxed{4}x_1 & + & \boxed{}\,x_2 & + & \boxed{5}x_3 & + & \boxed{3}x_4 & & \\
\text{subject to} & & & & & & & & & \\
 & x_1 & - & x_2 & - & x_3 & + & 3x_4 & \leq & 1 \\
 & 5x_1 & + & x_2 & + & 3x_3 & + & 8x_4 & \leq & 55 \\
 & -x_1 & + & 2x_2 & + & 3x_3 & - & 5x_4 & \leq & 3
\end{array}
$$

- can we find an upper bound?

$$
\begin{array}{rcrcrcrcr}
\boxed{10}x_1 & + & \boxed{2}x_2 & + & \boxed{6}x_3 & + & \boxed{16}x_4 & \leq & \boxed{110} \\
4x_1 & + & 3x_2 & + & 6x_3 & + & 3x_4 & \leq & 58
\end{array}
$$

Bounding

$$
\begin{array}{lrcrcrcrcr}
\max & 4x_1 & + & x_2 & + & 5x_3 & + & 3x_4 & & \\
\text{subject to} & & & & & & & & & \\
 & x_1 & - & x_2 & - & x_3 & + & 3x_4 & \leq & 1 \\
 & 5x_1 & + & x_2 & + & 3x_3 & + & 8x_4 & \leq & 55 \\
 & -x_1 & + & 2x_2 & + & 3x_3 & - & 5x_4 & \leq & 3
\end{array}
$$

- can we find an upper bound?

$$
\begin{array}{rcrcrcrcr}
10x_1 & + & 2x_2 & + & 6x_3 & + & 16x_4 & \leq & 110 \\
4x_1 & + & 3x_2 & + & 6x_3 & + & 3x_4 & \leq & 58
\end{array}
$$

Bounding

$$
\begin{array}{lrcrcrcrcr}
\max & 4x_1 & + & x_2 & + & 5x_3 & + & 3x_4 & & \\
\text{subject to} & & & & & & & & & \\
& x_1 & - & x_2 & - & x_3 & + & 3x_4 & \leq & 1 \\
& 5x_1 & + & x_2 & + & 3x_3 & + & 8x_4 & \leq & 55 \\
& -x_1 & + & 2x_2 & + & 3x_3 & - & 5x_4 & \leq & 3
\end{array}
$$

- can we find an upper bound?

$$
\begin{array}{rcrcrcrcr}
10x_1 & + & 2x_2 & + & 6x_3 & + & 16x_4 & \leq & 110 \\
4x_1 & + & 3x_2 & + & 6x_3 & + & 3x_4 & \leq & 58
\end{array}
$$

Bounding

$$
\begin{array}{lrcrcrcrcr}
\text{max} & \boxed{4}x_1 & + & \boxed{}x_2 & + & \boxed{5}x_3 & + & \boxed{3}x_4 & & \\
\text{subject to} & & & & & & & & & \\
 & x_1 & - & x_2 & - & x_3 & + & 3x_4 & \leq & 1 \\
 & 5x_1 & + & x_2 & + & 3x_3 & + & 8x_4 & \leq & 55 \\
 & -x_1 & + & 2x_2 & + & 3x_3 & - & 5x_4 & \leq & 3
\end{array}
$$

- can we find an upper bound?

$$
\begin{array}{rcrcrcrcr}
10x_1 & + & 2x_2 & + & 6x_3 & + & 16x_4 & \leq & 110 \\
\boxed{4}x_1 & + & \boxed{3}x_2 & + & \boxed{6}x_3 & + & \boxed{3}x_4 & \leq & \boxed{58}
\end{array}
$$

Bounding

$$
\begin{array}{lrcrcrcrcr}
\text{max} & 4x_1 & + & x_2 & + & 5x_3 & + & 3x_4 & & \\
\text{subject to} & & & & & & & & & \\
 & x_1 & - & x_2 & - & x_3 & + & 3x_4 & \leq & 1 \\
 & 5x_1 & + & x_2 & + & 3x_3 & + & 8x_4 & \leq & 55 \\
 & -x_1 & + & 2x_2 & + & 3x_3 & - & 5x_4 & \leq & 3
\end{array}
$$

- positive combinations of the constraints

Bounding

$$
\begin{array}{lrcrcrcrcrl}
\text{max} & 4x_1 & + & x_2 & + & 5x_3 & + & 3x_4 & & & \\
\text{subject to} & & & & & & & & & & \\
 & x_1 & - & x_2 & - & x_3 & + & 3x_4 & \leq & 1 & y_1 \\
 & 5x_1 & + & x_2 & + & 3x_3 & + & 8x_4 & \leq & 55 & y_2 \\
 & -x_1 & + & 2x_2 & + & 3x_3 & - & 5x_4 & \leq & 3 & y_3
\end{array}
$$

- positive combinations of the constraints

Bounding

$$
\begin{array}{lrcrcrcrcrl}
\text{max} & 4x_1 & + & x_2 & + & 5x_3 & + & 3x_4 & & & \\
\text{subject to} & & & & & & & & & & \\
& x_1 & - & x_2 & - & x_3 & + & 3x_4 & \le & 1 & y_1 \\
& 5x_1 & + & x_2 & + & 3x_3 & + & 8x_4 & \le & 55 & y_2 \\
& -x_1 & + & 2x_2 & + & 3x_3 & - & 5x_4 & \le & 3 & y_3
\end{array}
$$

- positive combinations of the constraints

$$
\begin{array}{llrcrcrcrll}
y_1 & (& x_1 & - & x_2 & - & x_3 & + & 3x_4 &) & + \\
y_2 & (& 5x_1 & + & x_2 & + & 3x_3 & + & 8x_4 &) & + \\
y_3 & (& -x_1 & + & 2x_2 & + & 3x_3 & - & 5x_4 &) & \\
 & & & & \le & & & & & & \\
 & & y_1 & + & 55y_2 & + & 3y_3 & & & &
\end{array}
$$

Bounding

$$
\begin{array}{lrcrcrcrcrl}
\max & 4x_1 & + & x_2 & + & 5x_3 & + & 3x_4 & & & \\
\text{subject to} & & & & & & & & & & \\
& x_1 & - & x_2 & - & x_3 & + & 3x_4 & \leq & 1 & y_1 \\
& 5x_1 & + & x_2 & + & 3x_3 & + & 8x_4 & \leq & 55 & y_2 \\
& -x_1 & + & 2x_2 & + & 3x_3 & - & 5x_4 & \leq & 3 & y_3
\end{array}
$$

- positive combinations of the constraints

$$
\begin{array}{lcrcrcrcrcc}
y_1 & (& x_1 & - & x_2 & - & x_3 & + & 3x_4 &) & + \\
y_2 & (& 5x_1 & + & x_2 & + & 3x_3 & + & 8x_4 &) & + \\
y_3 & (& -x_1 & + & 2x_2 & + & 3x_3 & - & 5x_4 &) &
\end{array}
$$

$$\leq$$

minimize $\boxed{y_1 + 55y_2 + 3y_3}$

Complementarity slackness

- Let x* and y* be the optimal solutions to the primal and dual. The following conditions are necessary and sufficient for the optimality of x* and y*:

$$\sum_{j=1}^{n} a_{ij} x_j^* = b_i \ \vee y_i^* = 0 \quad (1 \leq i \leq m)$$

$$\sum_{i=1}^{n} a_{ij} y_i^* = c_j \ \vee x_j^* = 0 \quad (1 \leq j \leq n)$$

Economic Interpretation

Maximizing profit:

$$\max \quad \sum_{j=1}^{n} c_j x_j$$

subject to

Capacity constraints on your production:

$$\sum_{j=1}^{n} a_{ij} x_j \quad \leq \quad b_i \boxed{+ t_i} \quad (1 \leq i \leq m)$$

- for some small t_i, this linear program has an optimal solution

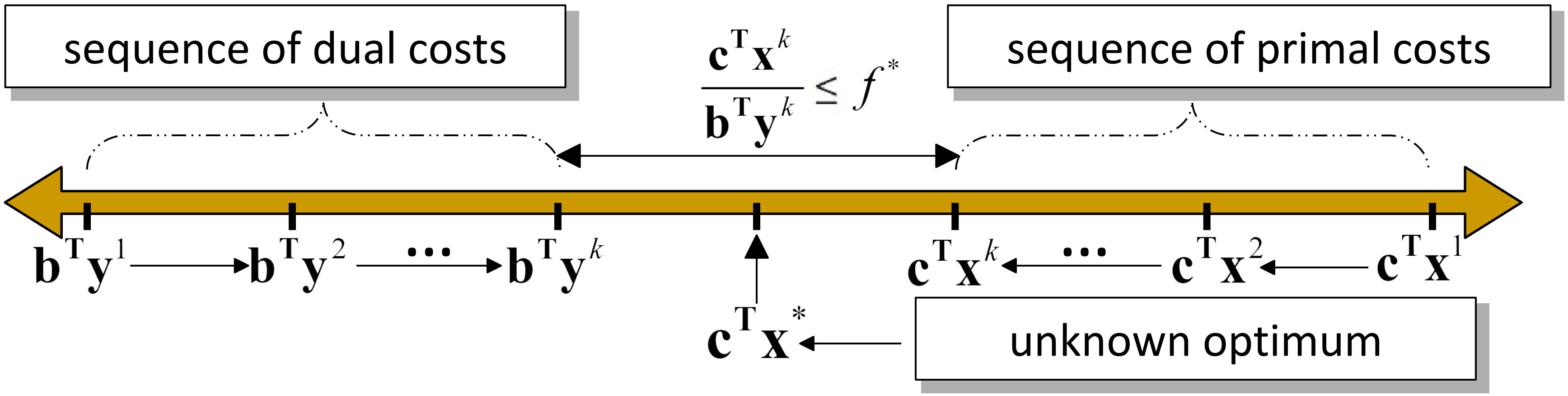
$$\underbrace{z^*}_{\text{optimal primal objective}} + \sum_{i=1}^{m} \underbrace{y_i^*}_{\text{dual solution}} \; t_i$$

Primal-Dual

- Why using the dual?
  - I have an optimal solution and I want to add a new constraint
  - The dual is still feasible (I am adding a new variable); the primal is not
  - Optimize the dual and the primal becomes feasible at optimality

The primal-dual schema

- The primal-dual schema works iteratively



- Global effects, through local improvements!

- Instead of working directly with costs (usually not easy), use <u>relaxed complementary slackness conditions</u> (easier)

- Different relaxations of complementary slackness
  Different approximation algorithms!!!

The primal-dual schema for MRFs

$$\min \left[\sum_{p \in G} \sum_{a \in L} V_p(a) x_{p,a} + \sum_{pq \in E} \sum_{a,b \in L} V_{pq}(a,b) x_{pq,ab} \right]$$

$$\text{s.t.} \sum\nolimits_{a \in L} x_{p,a} = 1$$

(only one label assigned per vertex)

$$\left.\begin{aligned} \sum\nolimits_{a \in L} x_{pq,ab} &= x_{q,b} \\ \sum\nolimits_{b \in L} x_{pq,ab} &= x_{p,a} \end{aligned}\right\}$$

enforce consistency between variables $x_{p,a}$, $x_{q,b}$ and variable $x_{pq,ab}$

$$x_{p,a} \geq 0, \; x_{pq,ab} \geq 0$$

Binary variables

$x_{p,a}=1$	$\Leftrightarrow$	label a is assigned to node p
$x_{pq,ab}=1$	$\Leftrightarrow$	labels a, b are assigned to nodes p, q

Complementary slackness

primal LP: $\min \mathbf{c}^T\mathbf{x}$

s.t. $\mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}$

dual LP: $\max \mathbf{b}^T\mathbf{y}$

s.t. $\mathbf{A}^T\mathbf{y} \leq \mathbf{c}$

Complementary slackness conditions:

$$\forall 1 \leq j \leq n: \quad x_j > 0 \Rightarrow \sum_{i=1}^{m} a_{ij} y_i = c_j$$

Theorem. tisfy the complementary slackness condition then they are both optimal.

Relaxed complementary slackness

primal LP: $\min \mathbf{c}^T\mathbf{x}$

s.t. $\mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}$

dual LP: $\max \mathbf{b}^T\mathbf{y}$

s.t. $\mathbf{A}^T\mathbf{y} \leq \mathbf{c}$

Exact CS: $$\forall 1 \leq j \leq n: \quad x_j > 0 \Rightarrow \sum_{i=1}^{m} a_{ij} y_i = c_j$$

Relaxed CS: $$\forall 1 \leq j \leq n: \quad x_j > 0 \Rightarrow \sum_{i=1}^{m} a_{ij} y_i \geq c_j / f_j$$

implies 'e

$f_j = 1 \forall j$

Theorem. If **x**, **y** primal/dual feasible and satisfy the **relaxed CS** condition then **x** is an f-approximation of the optimal integral solution, where f = max_j f_j.

Complementary slackness and the primal-dual schema

Theorem (previous slide). If **x**, **y** primal/dual feasible and satisfy the relaxed CS condition then **x** is an f-approximation of the optimal integral solution, where f = max_j f_j.

Goal of the primal dual schema: find a pair (**x**,**y**) that satisfies:
- Primal feasibility
- Dual feasibility
- (Relaxed) complementary slackness conditions.

FastPD: primal-dual schema for MRFs

- Regarding the PD schema for MRFs, it turns out that:

each update of primal and dual variables $\Longleftrightarrow$ solving max-flow in appropriately constructed graph

- Resulting flows tell us how to update both:
  - the dual variables, as well as
  - the primal variables

  for each iteration of primal-dual schema

- Max-flow graph defined from current primal-dual pair (x^k,y^k)
  - (x^k,y^k) defines connectivity of max-flow graph
  - (x^k,y^k) defines capacities of max-flow graph

- Max-flow graph is thus continuously updated

FastPD: primal-dual schema for MRFs

- Very general framework. Different PD-algorithms by RELAXING complementary slackness conditions differently.

- E.g., simply by using a particular relaxation of complementary slackness conditions (and assuming $V_{pq}(\cdot,\cdot)$ is a metric)
  <u>THEN</u> resulting algorithm shown equivalent to a-expansion!
  [Boykov,Veksler,Zabih]

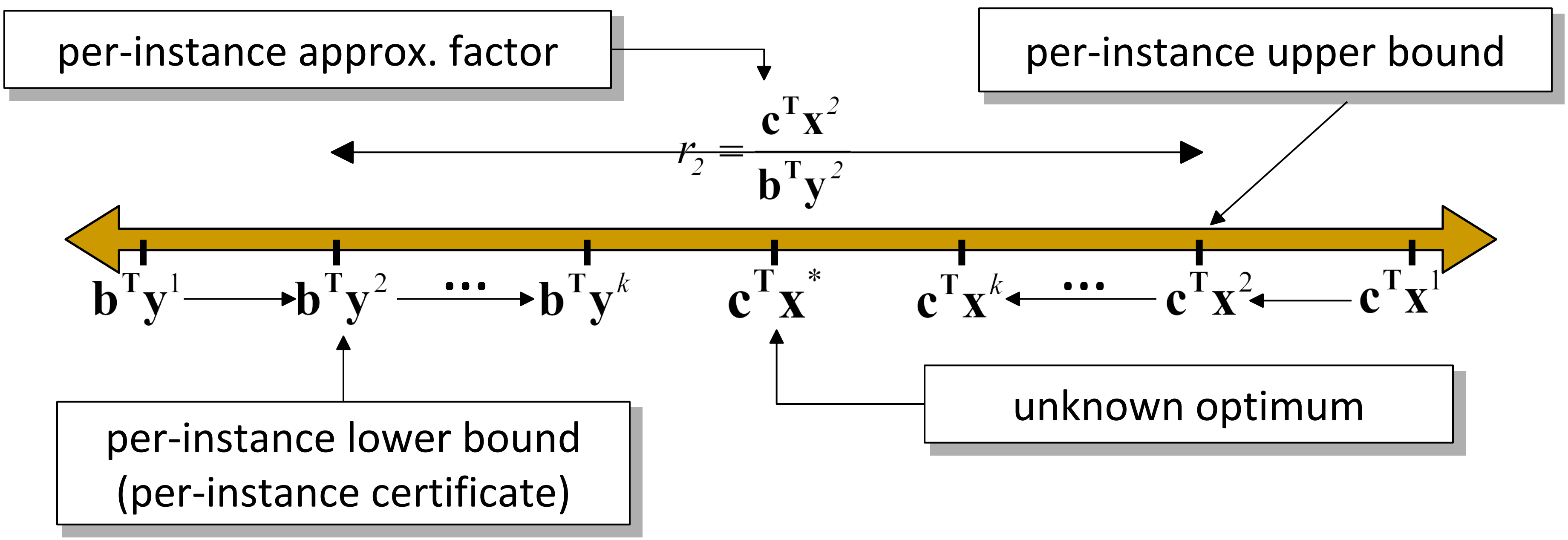
- PD-algorithms for non-metric potentials $V_{pq}(\cdot,\cdot)$ as well

- Theorem: All derived PD-algorithms shown to satisfy certain relaxed complementary slackness conditions

- <u>Worst-case</u> optimality properties are thus guaranteed

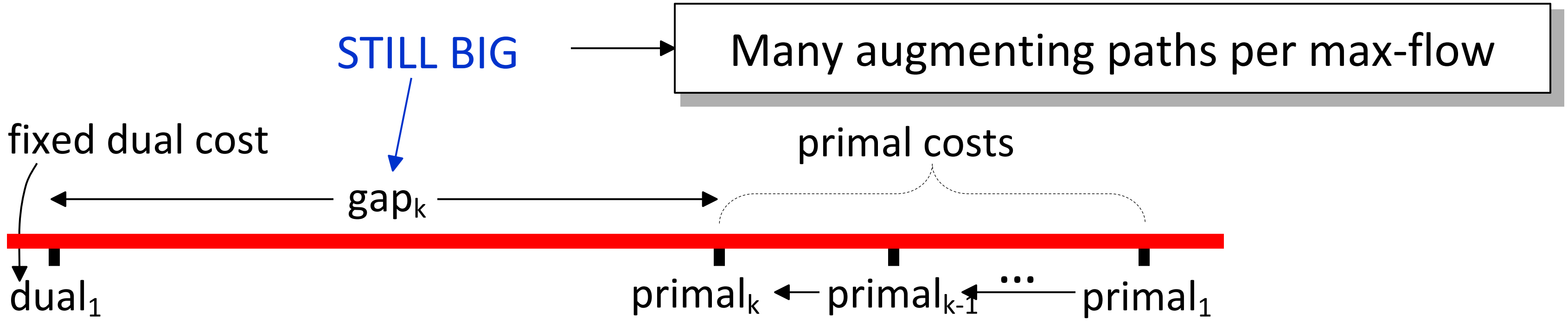
Per-instance optimality guarantees

- Primal-dual algorithms can always tell you **(for free)** how well they performed for a particular instance

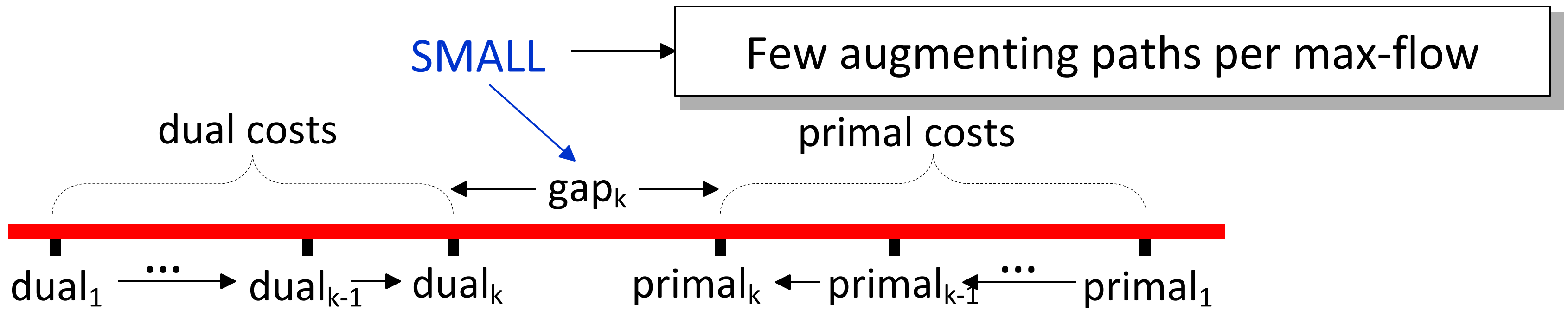


Computational efficiency (static MRFs)

- MRF algorithm only in the primal domain (e.g., a-expansion)



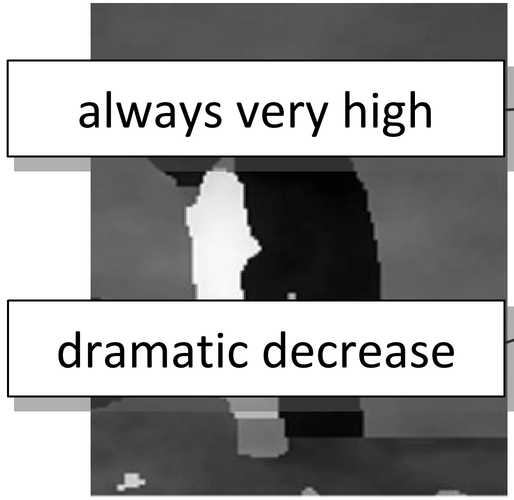
- MRF algorithm in the primal-dual domain (Fast-PD)



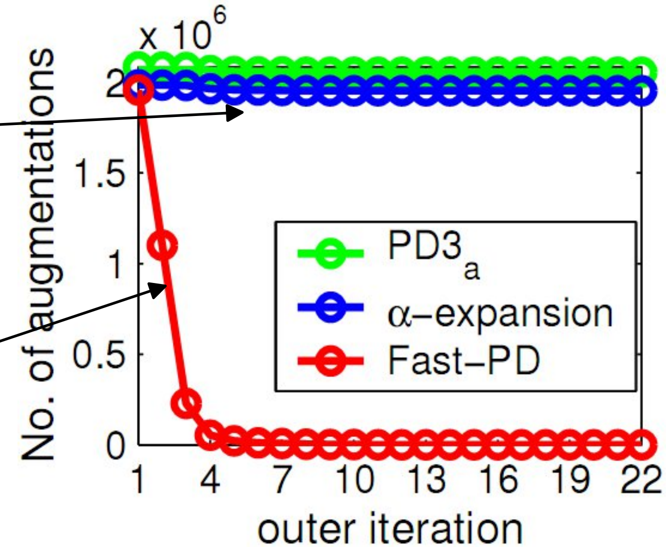
Theorem: primal-dual gap = upper-bound on #augmenting paths
(i.e., primal-dual gap indicative of time per max-flow)

Computational efficiency (static MRFs)

noisy image



denoised image

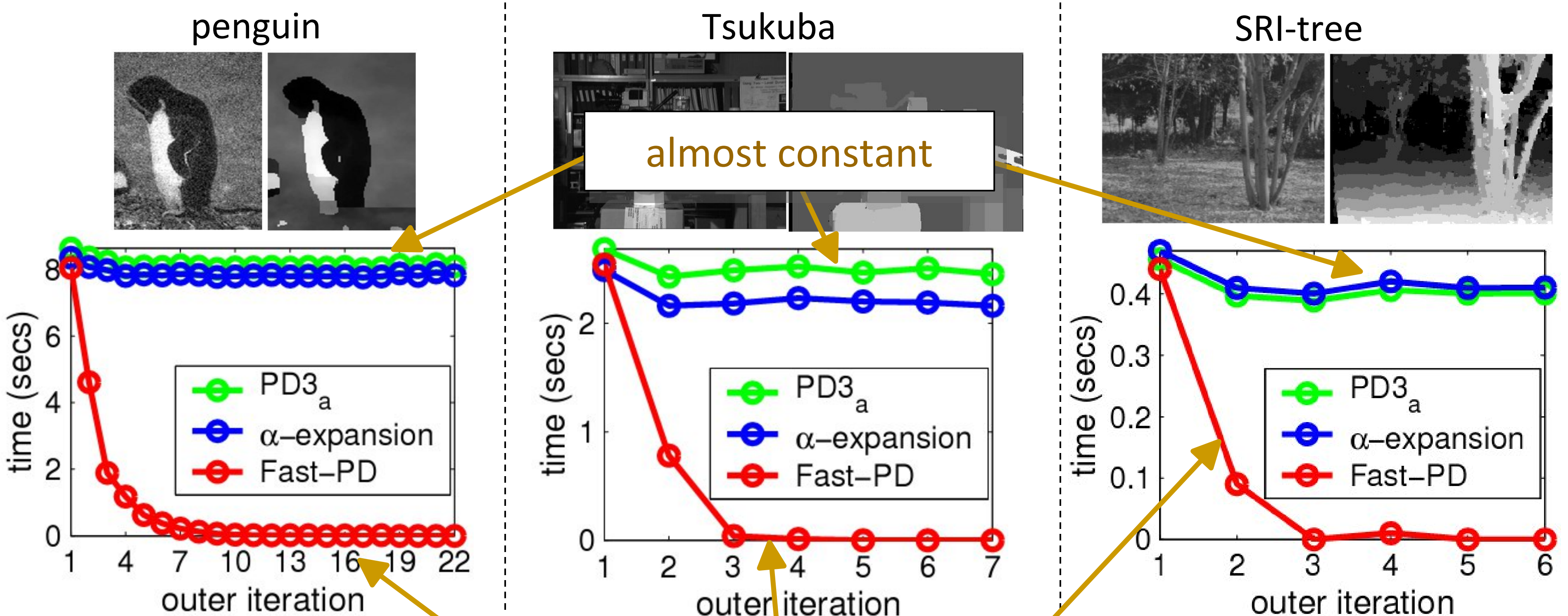


- Incremental construction of max-flow graphs
  (recall that max-flow graph changes per iteration)

- Possible because we keep <u>both</u> primal and dual information

- Principled way for doing this construction via the primal-dual framework

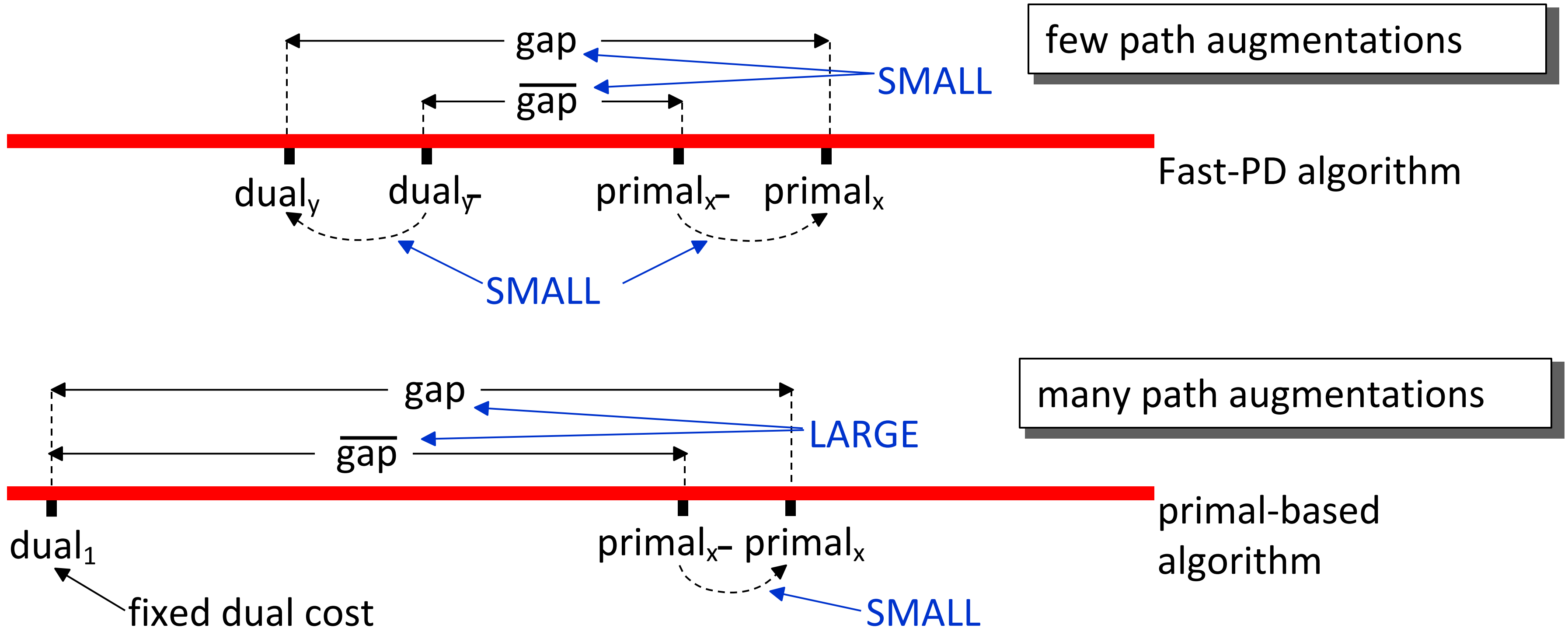
Computational efficiency (static MRFs)



total (secs)	Fast-PD	α-expansion
penguin	17.44	173.1
tsukuba	3.37	15.63
SRI tree	0.54	2.56

Computational efficiency (dynamic MRFs)

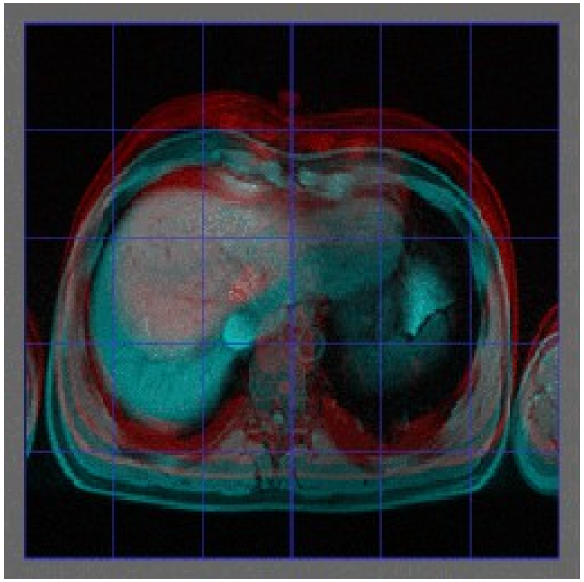
- Fast-PD can speed up dynamic MRFs [Kohli,Torr] as well (demonstrates the power and generality of this framework)

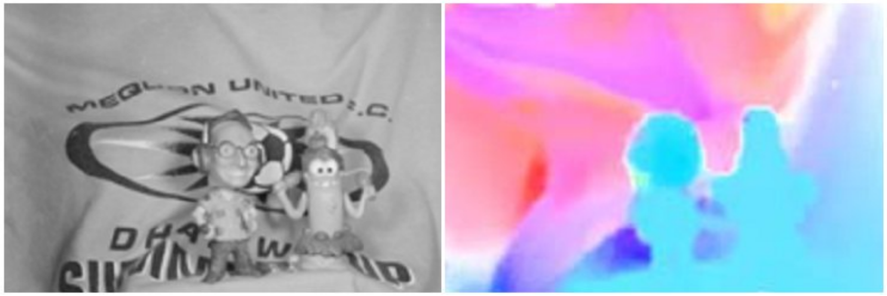


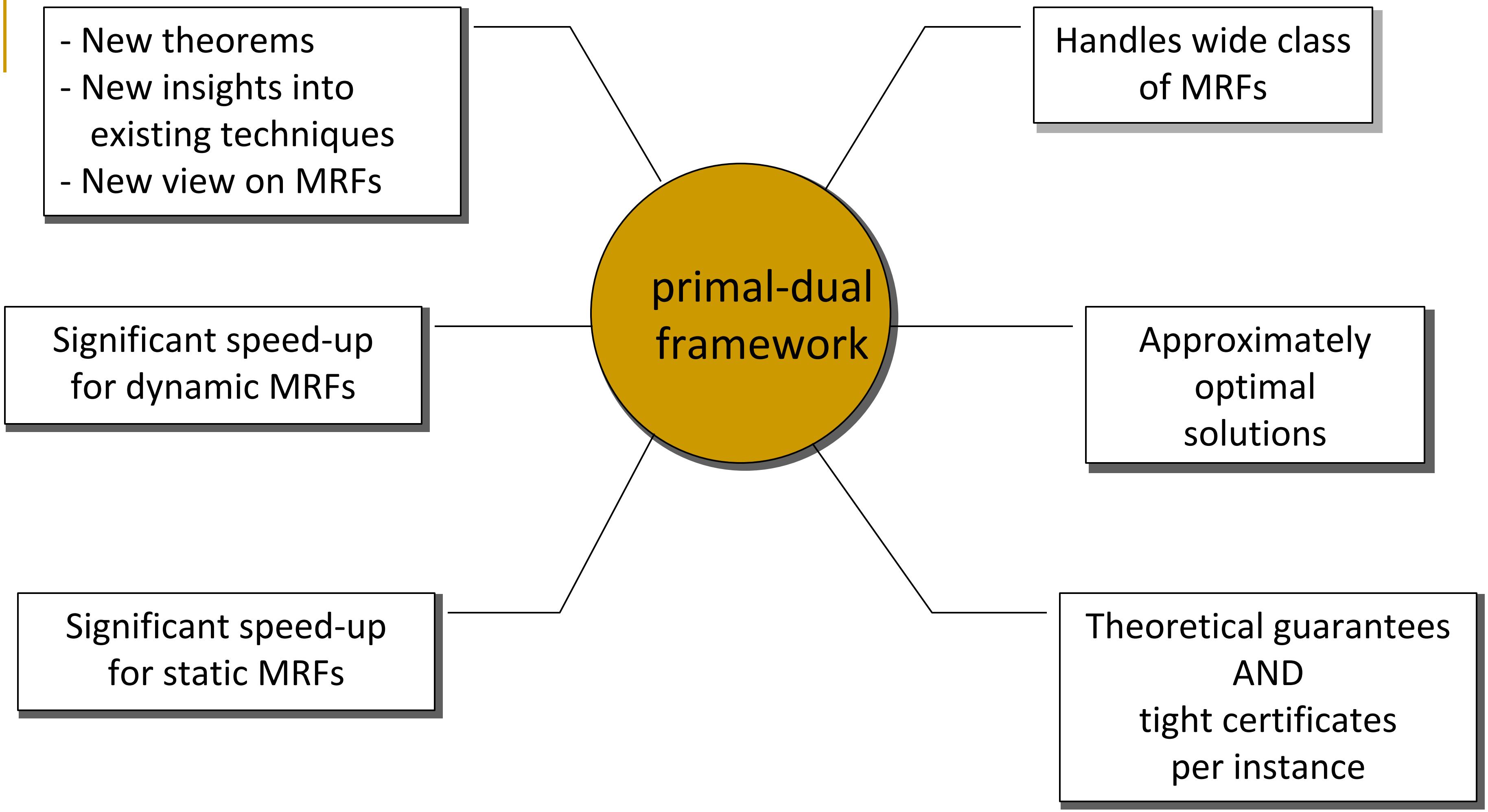
- Principled (and simple) way to update dual variables when switching between different MRFs

Drop: Deformable Registration using Discrete Optimization [Glocker et al. 07, 08]

- Easy to use GUI
- Main focus on medical imaging
- 2D-2D registration
- 3D-3D registration
- Publicly available:
  http://campar.in.tum.de/Main/Drop






primal-dual framework
- New theorems
- New insights into existing techniques
- New view on MRFs
Handles wide class of MRFs
Significant speed-up for dynamic MRFs
Approximately optimal solutions
Significant speed-up for static MRFs
Theoretical guarantees AND tight certificates per instance

Take home message:

LP and its duality theory provides:

Powerful framework for systematically tackling the MRF optimization problem

Unifying view for the state-of-the-art MRF optimization techniques