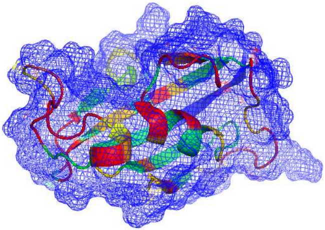
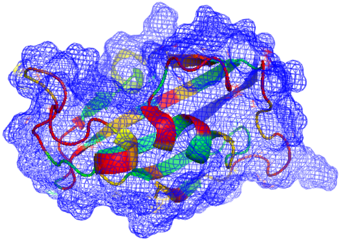
Improving the Resolution of Hydrogen/Deuterium Exchange Data

Stefan Canzar

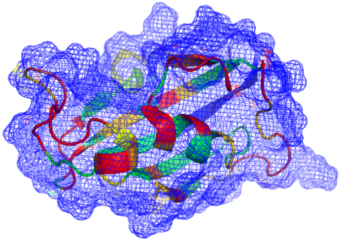
Centrum Wiskunde & Informatica



Motivation

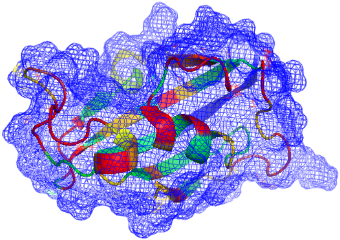


Motivation



- can be determined by X-ray crystal diffraction, NMR, ...
- but limitations! (concentration, conformational changes. . .)

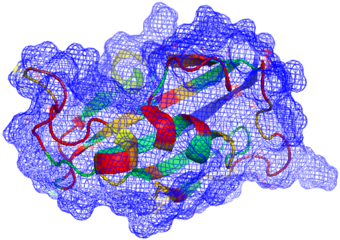
Motivation



- can be determined by X-ray crystal diffraction, NMR, ...
- but limitations! (concentration, conformational changes. . .)

⇒ Exchange of labile hydrogens for deuteriums (HDX)

Motivation

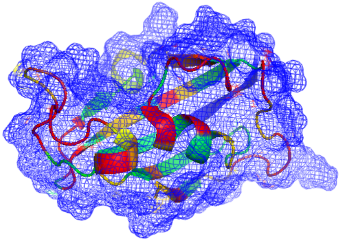


- can be determined by X-ray crystal diffraction, NMR, ...
- but limitations! (concentration, conformational changes...)

⇒ Exchange of labile hydrogens for deuteriums (HDX)

- exchange rates as measure of solvent accessibility → structure!

Motivation

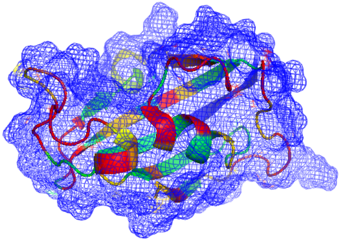


- can be determined by X-ray crystal diffraction, NMR, ...
- but limitations! (concentration, conformational changes. . .)

⇒ Exchange of labile hydrogens for deuteriums (HDX)

- exchange rates as measure of solvent accessibility → structure!
- monitored by mass spectrometry (MS) (NMR limited)

Motivation

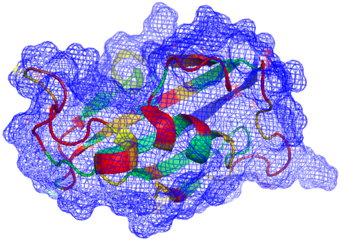


- can be determined by X-ray crystal diffraction, NMR, ...
- but limitations! (concentration, conformational changes. . .)

⇒ Exchange of labile hydrogens for deuteriums (HDX)

- exchange rates as measure of solvent accessibility → structure!
- monitored by mass spectrometry (MS) (NMR limited)
- output: aggregate exchange data for peptic fragments

Motivation



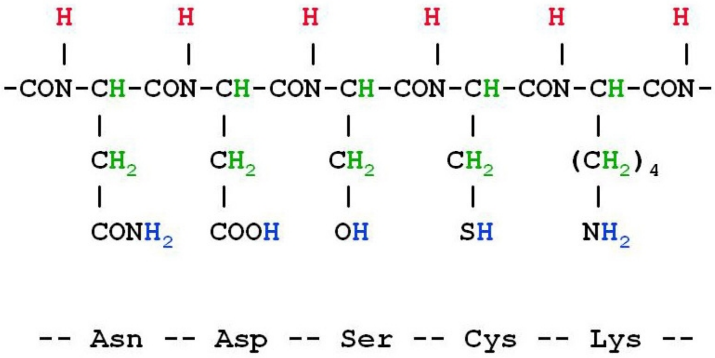
- can be determined by X-ray crystal diffraction, NMR, ...
- but limitations! (concentration, conformational changes...)

⇒ Exchange of labile hydrogens for deuteriums (HDX)

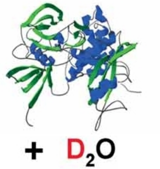
- exchange rates as measure of solvent accessibility → structure!
- monitored by mass spectrometry (MS) (NMR limited)
- output: aggregate exchange data for peptic fragments

⇒ Increase resolution from fragments to single residues!

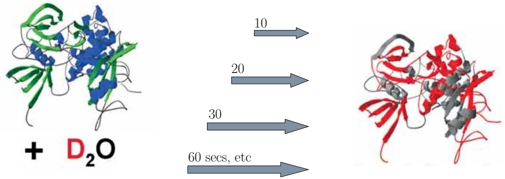
Hydrogens in Proteins



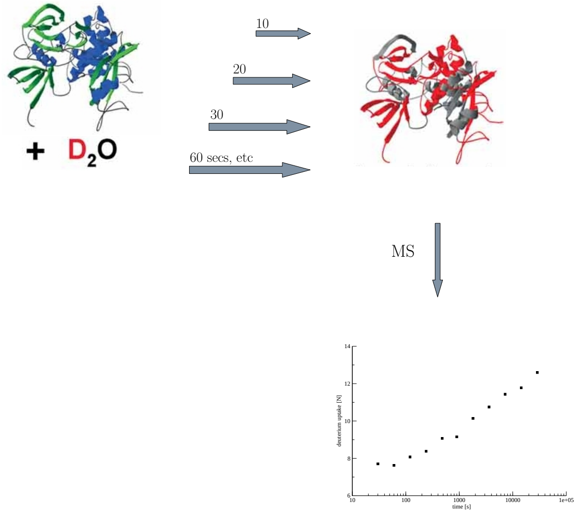
H/D Exchange + MS



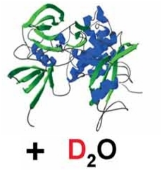
H/D Exchange + MS



H/D Exchange + MS



H/D Exchange + MS



10

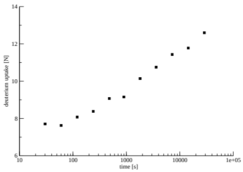
20

30

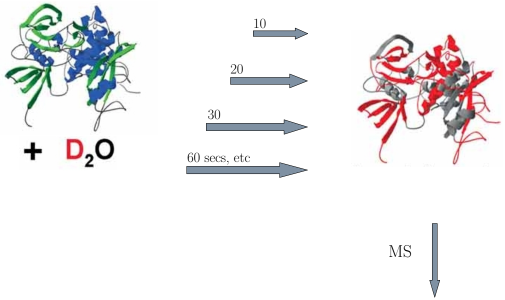
60 secs, etc

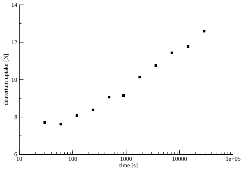
MS

Fit to 3-comp. model



H/D Exchange + MS

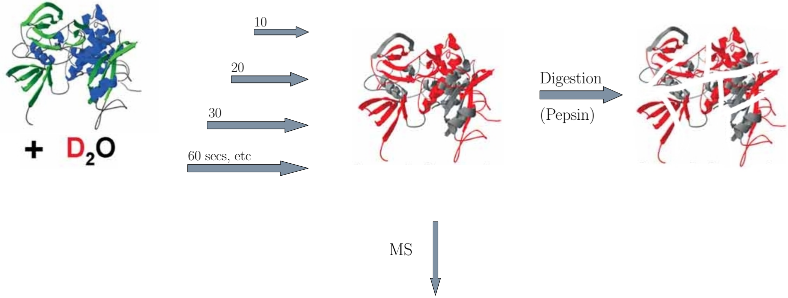




Fit to 3-comp. model

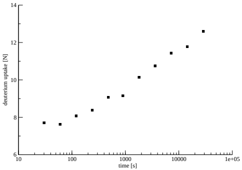
$\Rightarrow$ (F : 20, M : 14, S : 5)

H/D Exchange + MS

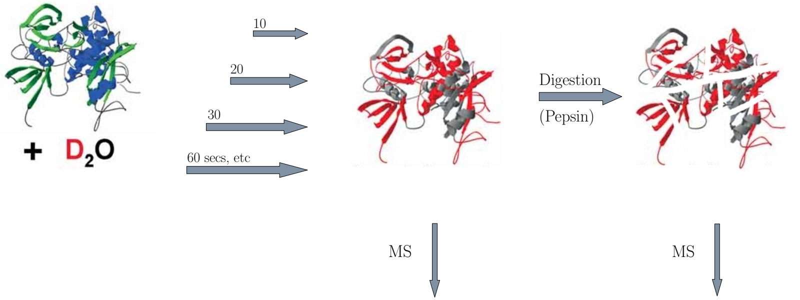


Fit to 3-comp. model

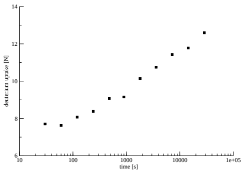
$\Rightarrow$ (F : 20, M : 14, S : 5)



H/D Exchange + MS



$(F_1 : 5, M_1 : 2, S_1 : 0)$
$(F_2 : 4, M_2 : 4, S_2 : 2)$
$\vdots$
$(F_k : 2, M_k : 5, S_k : 3)$



Fit to 3-comp. model

$\Rightarrow$ (F : 20, M : 14, S : 5)

Problem Illustration

G	R	Y	R	G	Y	R	R	Y
1	2	3	4	5	6	7	8	9

Problem Illustration

G	R	Y	R	G	Y	R	R	Y
1	2	3	4	5	6	7	8	9

Experiments

$[1,5]$ — — $(2,2,1)$

$[3,7]$ — — $(2,1,2)$

$[6,8]$ — — $(2,0,1)$

$[5,9]$ — — $(2,1,2)$

Problem Illustration

G	R	Y	R	G	Y	R	R	Y
1	2	3	4	5	6	7	8	9

Experiments ↓

$[1,5]$ — — $(2,2,1)$

$[3,7]$ — — $(2,1,2)$

$[6,8]$ — — $(2,0,1)$

$[5,9]$ — — $(2,1,2)$

↑ feasible assignment

Mathematical Abstraction

- Protein of n residues $\longmapsto$ Set $V = [n]$

Mathematical Abstraction

- Protein of n residues $\longmapsto$ Set $V = [n]$
- Peptic fragments $\longmapsto$ Set $\mathcal{I}$ of intervals defined on V

Mathematical Abstraction

- Protein of n residues $\longmapsto$ Set $V = [n]$
- Peptic fragments $\longmapsto$ Set $\mathcal{I}$ of intervals defined on V
- H/D exchange rates $\longmapsto$ Set of color classes $[k]$

Mathematical Abstraction

- Protein of n residues $\longmapsto$ Set $V = [n]$
- Peptic fragments $\longmapsto$ Set $\mathcal{I}$ of intervals defined on V
- H/D exchange rates $\longmapsto$ Set of color classes $[k]$
- Requirement function $r : \mathcal{I} \times [k] \mapsto \mathbb{N}$ where

$$r(I, 1) + \cdots + r(I, k) = |I|$$

Mathematical Abstraction

- Protein of n residues $\longmapsto$ Set $V = [n]$
- Peptic fragments $\longmapsto$ Set $\mathcal{I}$ of intervals defined on V
- H/D exchange rates $\longmapsto$ Set of color classes $[k]$
- Requirement function $r : \mathcal{I} \times [k] \mapsto \mathbb{N}$ where

$$r(I, 1) + \cdots + r(I, k) = |I|$$

IntervalConstrainedColoring

Is there a *feasible* coloring, i.e. a function $\chi : V \mapsto [k]$ such that for every $I \in \mathcal{I}$ and all $c \in [k]$ we have $|\{j \in I \mid \chi(j) = c\}| = r(I, c)$?

An ILP Formulation

Variables:

$$x_{i,j} := \begin{cases} 1 & \text{if residue } i \text{ has exchange rate } j \\ 0 & \text{otherwise.} \end{cases}$$

An ILP Formulation

Variables:

$$x_{i,j} := \begin{cases} 1 & \text{if residue } i \text{ has exchange rate } j \\ 0 & \text{otherwise.} \end{cases}$$

Constraints:

Each residue i gets exactly one exchange rate:

$$\sum_{j \in [k]} x_{i,j} = 1$$

An ILP Formulation

Variables:

$$x_{i,j} := \begin{cases} 1 & \text{if residue } i \text{ has exchange rate } j \\ 0 & \text{otherwise.} \end{cases}$$

Constraints:

Each residue i gets exactly one exchange rate:

$$\sum_{j \in [k]} x_{i,j} = 1$$

Each fragment I contains $r(I,j)$ residues exchanging at rate j:

$$\sum_{i \in I} x_{i,j} = r(I,j)$$

Modeling Measurement Errors

In Practice:

- experimental data contain noise
- A feasible coloring usually does not exist

Modeling Measurement Errors

In Practice:

- experimental data contain noise
- A feasible coloring usually does not exist

$\Rightarrow$ **Capture noise in model!**

Modeling Measurement Errors

In Practice:

- experimental data contain noise
- A feasible coloring usually does not exist

⇒ **Capture noise in model!**

(i) Drop constraint $\sum_{i \in I} x_{i,j} = r(I,j)$ and minimize

$$\sum_{I \in \mathcal{I}} \sum_{j \in [k]} |r(I,j) - \sum_{i \in I} x_{i,j}|$$

Modeling Measurement Errors

In Practice:

- experimental data contain noise
- A feasible coloring usually does not exist

$\Rightarrow$ **Capture noise in model!**

(i) Drop constraint $\sum_{i \in I} x_{i,j} = r(I,j)$ and minimize

$$\sum_{I \in \mathcal{I}} \sum_{j \in [k]} |r(I,j) - \sum_{i \in I} x_{i,j}|$$

(ii) Assign exchange rates fractionally to residues and round

Modeling Measurement Errors

In Practice:

- experimental data contain noise
- A feasible coloring usually does not exist

$\Rightarrow$ **Capture noise in model!**

(i) Drop constraint $\sum_{i \in I} x_{i,j} = r(I, j)$ and minimize

$$\sum_{I \in \mathcal{I}} \sum_{j \in [k]} |r(I, j) - \sum_{i \in I} x_{i,j}|$$

(ii) Assign exchange rates fractionally to residues and round

(iii) Find a maximum (weight) colorable subset $\mathcal{I}' \subseteq \mathcal{I}$

Modeling Measurement Errors

In Practice:

- experimental data contain noise
- A feasible coloring usually does not exist

$\Rightarrow$ **Capture noise in model!**

(i) Drop constraint $\sum_{i \in I} x_{i,j} = r(I,j)$ and minimize

$$\sum_{I \in \mathcal{I}} \sum_{j \in [k]} |r(I,j) - \sum_{i \in I} x_{i,j}|$$

(ii) Assign exchange rates fractionally to residues and round

(iii) Find a maximum (weight) colorable subset $\mathcal{I}' \subseteq \mathcal{I}$

Maximum Feasible Coloring

MaxFeasibleColoring (MFC)

Given non-negative weights $w : \mathcal{I} \mapsto \mathbb{R}_{+}$, find a maximum weight colorable subset $\mathcal{I}' \subseteq \mathcal{I}$.

Maximum Feasible Coloring

MaxFeasibleColoring (MFC)

Given non-negative weights $w : \mathcal{I} \mapsto \mathbb{R}_+$, find a maximum weight colorable subset $\mathcal{I}' \subseteq \mathcal{I}$.

- Mfc is $\mathcal{APX}$-hard for $k = 2$ (reduction from Max2SAT)

Maximum Feasible Coloring

MaxFeasibleColoring (MFC)

Given non-negative weights $w : \mathcal{I} \mapsto \mathbb{R}_{+}$, find a maximum weight colorable subset $\mathcal{I}' \subseteq \mathcal{I}$.

- MFC is $\mathcal{APX}$-hard for $k = 2$ (reduction from Max2SAT)
- $(1 + \epsilon)$ violation in quasi-polynomial time

Maximum Feasible Coloring

MaxFeasibleColoring (MFC)

Given non-negative weights $w : \mathcal{I} \mapsto \mathbb{R}_+$, find a maximum weight colorable subset $\mathcal{I}' \subseteq \mathcal{I}$.

- Mfc is $\mathcal{APX}$-hard for $k = 2$ (reduction from Max2SAT)
- $(1 + \epsilon)$ violation in quasi-polynomial time
- $\mathcal{O}(\sqrt{|OPT|})$-approximation algorithm (for constant k)

Maximum Feasible Coloring

MaxFeasibleColoring (MFC)

Given non-negative weights $w : \mathcal{I} \mapsto \mathbb{R}_+$, find a maximum weight colorable subset $\mathcal{I}' \subseteq \mathcal{I}$.

- Mfc is $\mathcal{APX}$-hard for $k = 2$ (reduction from Max2SAT)
- $(1 + \epsilon)$ violation in quasi-polynomial time
- $\mathcal{O}(\sqrt{|OPT|})$-approximation algorithm (for constant k)

Practically irrelevant, but related to

Maximum feasible subsystem problem with 0/1-coefficients!

Modeling Measurement Errors

In Practice:

- experimental data contain noise
- A feasible coloring usually does not exist

$\Rightarrow$ **Capture noise in model!**

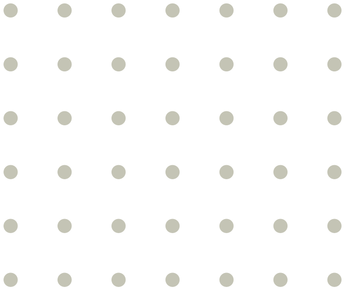
(i) Drop constraint $\sum_{i \in I} x_{i,j} = r(I,j)$ and minimize

$$\sum_{I \in \mathcal{I}} \sum_{j \in [k]} |r(I,j) - \sum_{i \in I} x_{i,j}|$$

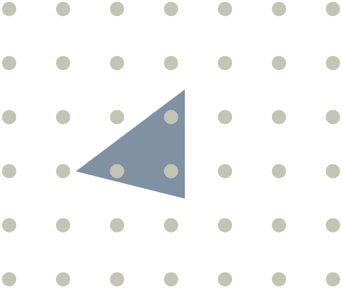
(ii) Assign exchange rates fractionally to residues and round

(iii) Find a maximum (weight) colorable subset $\mathcal{I}' \subseteq \mathcal{I}$

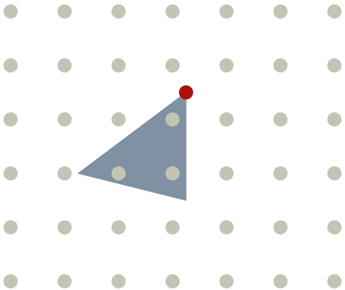
A ± 1 Guarantee



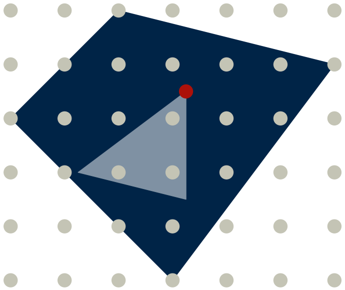
A ± 1 Guarantee



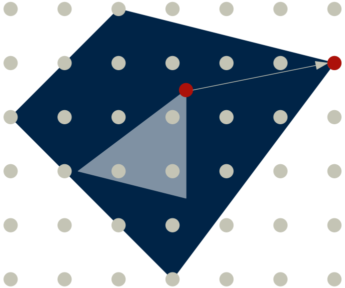
A ± 1 Guarantee



A ± 1 Guarantee

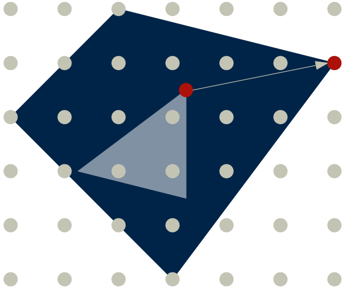


A ± 1 Guarantee



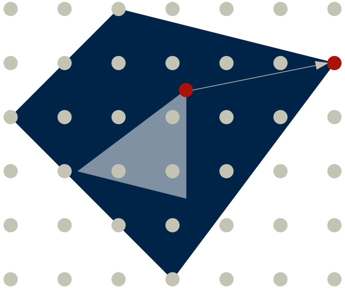
- # vertices in I colored j is $r(I, j) \pm 1$

A ± 1 Guarantee



- # vertices in I colored j is $r(I,j) \pm 1$
- $\Pr[\ I \text{ is satisfied }] \geq \frac{c(c+1-H_{c-1})}{(c+1)!}$

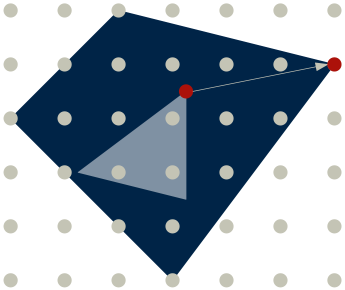
A ± 1 Guarantee



- # vertices in I colored j is $r(I, j) \pm 1$
- Pr[I is satisfied] $\geq \frac{c(c+1-H_{c-1})}{(c+1)!}$

- But: Feasible fractional solution must exist!

A ± 1 Guarantee



- # vertices in I colored j is $r(I, j) \pm 1$
- Pr[I is satisfied] $\geq \frac{c(c+1-H_{c-1})}{(c+1)!}$

- But: Feasible fractional solution must exist!
- Guarantees polynomial delay in enumeration

Modeling Measurement Errors

In Practice:

- experimental data contain noise
- A feasible coloring usually does not exist

$\Rightarrow$ **Capture noise in model!**

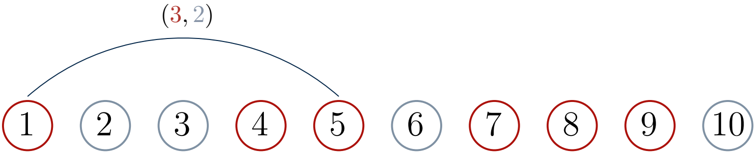
(i) Drop constraint $\sum_{i \in I} x_{i,j} = r(I,j)$ and minimize

$$\sum_{I \in \mathcal{I}} \sum_{j \in [k]} |r(I,j) - \sum_{i \in I} x_{i,j}|$$

(ii) Assign exchange rates fractionally to residues and round

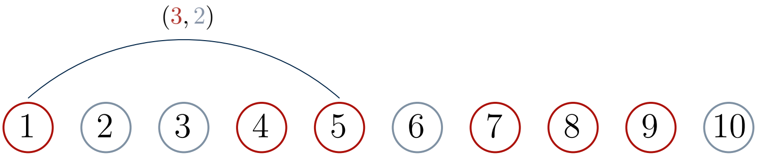
(iii) Find a maximum (weight) colorable subset $\mathcal{I}' \subseteq \mathcal{I}$

2 Colors without Errors



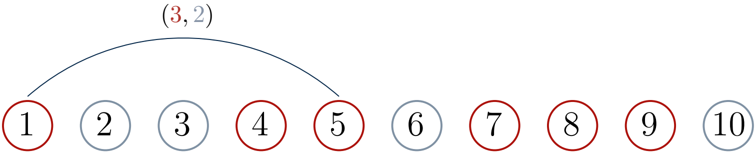
- Two colors 1 and 2 ⇒ Constraint matrix is totally unimodular

2 Colors without Errors



- Two colors 1 and 2 $\Rightarrow$ Constraint matrix is totally unimodular
- Assume $r(I, 1) + r(I, 2) = |I|$

2 Colors without Errors



- Two colors 1 and 2 $\Rightarrow$ Constraint matrix is totally unimodular
- Assume $r(I, 1) + r(I, 2) = |I|$
- Feasible coloring $\boldsymbol{x}_1$ suffices

Equivalence to Shortest Paths

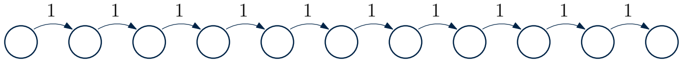
Let $d(i)$ denote the number of red vertices left of (and including) vertex i.

Equivalence to Shortest Paths

Let $d(i)$ denote the number of red vertices left of (and including) vertex i.

Idea: Add edges such that shortest path lengths from a new source vertex define a feasible d.

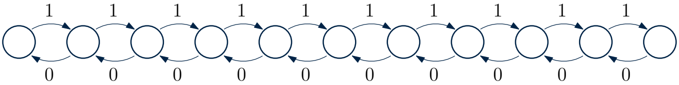
Constraint Graph Construction



1. Number of red vertices can increase by at most one:

$$d(i+1) \leq d(i) + 1$$

Constraint Graph Construction



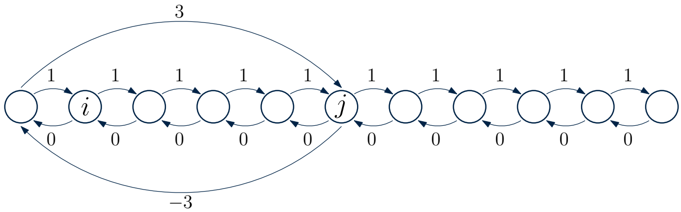
1. Number of red vertices can increase by at most one:

$$d(i+1) \leq d(i) + 1$$

2. Number of red vertices is monotonically increasing:

$$d(i) \leq d(i+1)$$

Constraint Graph Construction



1. Number of red vertices can increase by at most one:

$$d(i+1) \leq d(i) + 1$$

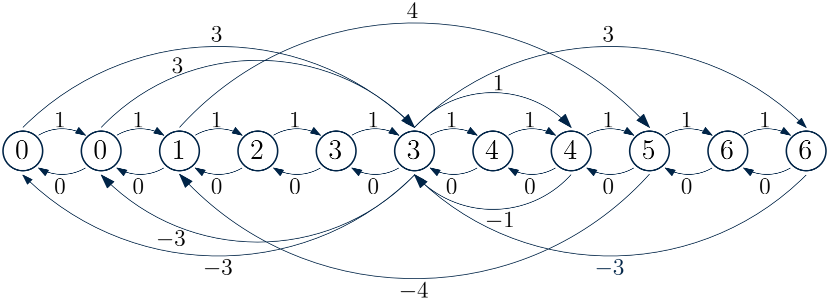
2. Number of red vertices is monotonically increasing:

$$d(i) \leq d(i+1)$$

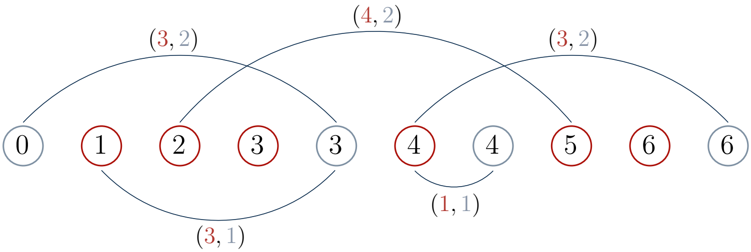
3. Fragment $I = [i, j]$ contains $r(I, 1)$ red vertices:

$$d(j) - d(i-1) = r(I, 1)$$

2 Colors by Shortest Paths



2 Colors by Shortest Paths



Extension to Real-World Instances

- 2 colors with error: *Minimum cost circulation problem* (MCS)

Extension to Real-World Instances

- 2 colors with error: *Minimum cost circulation problem* (MCS)
- $\mathcal{NP}$-hard for $k \geq 3$ [Byrka et al. LATIN'10]

Extension to Real-World Instances

- 2 colors with error: *Minimum cost circulation problem* (MCS)
- $\mathcal{NP}$-hard for $k \geq 3$ [Byrka et al. LATIN'10]
- Heuristic: Solve MCS and recurse on $k - 1$
- *Lagrangian Relaxation:* MCS per color!

Extension to Real-World Instances

- 2 colors with error: *Minimum cost circulation problem* (MCS)
- $\mathcal{NP}$-hard for $k \geq 3$ [Byrka et al. LATIN'10]
- Heuristic: Solve MCS and recurse on $k - 1$
- *Lagrangian Relaxation:* MCS per color!

Real world instances

- Cabin, Cytochrome P450, FK506 binding protein, myoglobin
- $74 \leq n \leq 152$, $18 \leq m \leq 49$, $k = 3$

Extension to Real-World Instances

- 2 colors with error: *Minimum cost circulation problem* (MCS)
- $\mathcal{NP}$-hard for $k \geq 3$ [Byrka et al. LATIN'10]
- Heuristic: Solve MCS and recurse on $k - 1$
- *Lagrangian Relaxation:* MCS per color!

Real world instances

- Cabin, Cytochrome P450, FK506 binding protein, myoglobin
- $74 \leq n \leq 152$, $18 \leq m \leq 49$, $k = 3$
- Optimal solution in < 0.1 second

Extension to Real-World Instances

- 2 colors with error: *Minimum cost circulation problem* (MCS)
- $\mathcal{NP}$-hard for $k \geq 3$ [Byrka et al. LATIN'10]
- Heuristic: Solve MCS and recurse on $k-1$
- *Lagrangian Relaxation:* MCS per color!

Real world instances

- Cabin, Cytochrome P450, FK506 binding protein, myoglobin
- $74 \leq n \leq 152$, $18 \leq m \leq 49$, $k = 3$
- Optimal solution in < 0.1 second
- All optimal solutions with minimal error in < 14 seconds

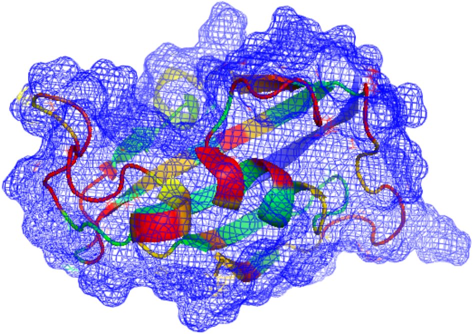
Extension to Real-World Instances

- 2 colors with error: *Minimum cost circulation problem* (MCS)
- $\mathcal{NP}$-hard for $k \geq 3$ [Byrka et al. LATIN'10]
- Heuristic: Solve MCS and recurse on $k - 1$
- *Lagrangian Relaxation:* MCS per color!

Real world instances

- Cabin, Cytochrome P450, FK506 binding protein, myoglobin
- $74 \leq n \leq 152$, $18 \leq m \leq 49$, $k = 3$
- Optimal solution in < 0.1 second
- All optimal solutions with minimal error in < 14 seconds
- 60%-75% agreement with NMR

Structural View of FKBP



red = fast
yellow = medium
green = slow

Acknowledgments

Experiments & Data:

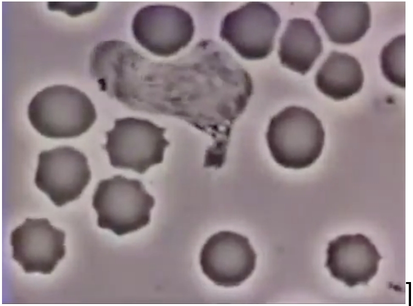
- M.R. Emmett - Department of Chemistry, FSU
- A. Marshall - Department of Chemistry, FSU
- A. Meyer-Baese - FAMU-FSU College of Engineering
- J. Tipton - Department of Chemistry, FSU
- H. Zhang - Institute of Molecular Biophysics, FSU

Theory & Implementation

- E. Althaus - Johannes Gutenberg-Universität Mainz
- C. Ehrler - Universität des Saarlandes
- K. Elbassioni - MPI Saarbrücken
- A. Karrenbauer - Institute of Mathematics, EPFL Lausanne
- J. Mestre - University of Sydney

Analysis of Live Cell Video

with K. Elbassioni, G. Klau, J. Mestre



Analysis of Life Cell Video

Given live cell video, we want to track individual cells

Analysis of Life Cell Video

Given live cell video, we want to track individual cells

Segmentation based methods:

Analysis of Life Cell Video

Given live cell video, we want to track individual cells

Segmentation based methods:

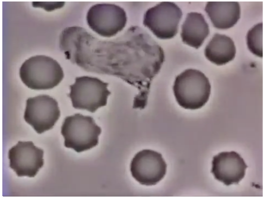
1. Perform image segmentation for each frame

Analysis of Life Cell Video

Given live cell video, we want to track individual cells

Segmentation based methods:

1. Perform image segmentation for each frame

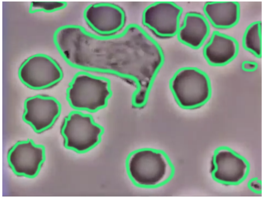


Analysis of Life Cell Video

Given live cell video, we want to track individual cells

Segmentation based methods:

1. Perform image segmentation for each frame

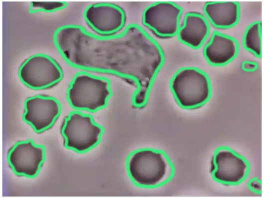


Analysis of Life Cell Video

Given live cell video, we want to track individual cells

Segmentation based methods:

1. Perform image segmentation for each frame
2. Match segments from adjacengt frames

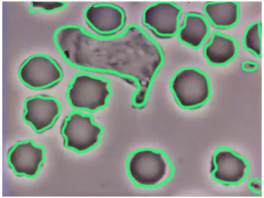


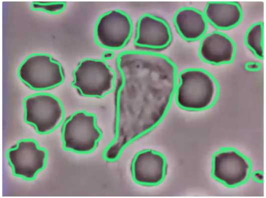
Analysis of Life Cell Video

Given live cell video, we want to track individual cells

Segmentation based methods:

1. Perform image segmentation for each frame
2. Match segments from adjacengt frames



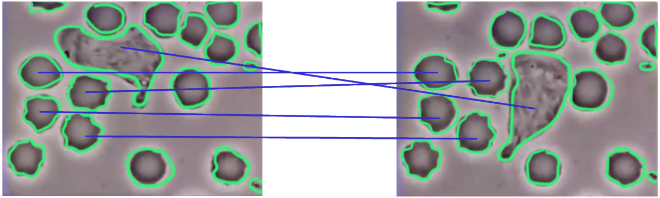


Analysis of Life Cell Video

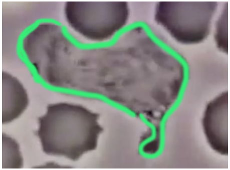
Given live cell video, we want to track individual cells

Segmentation based methods:

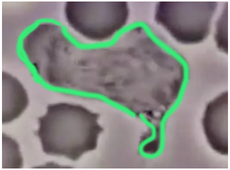
1. Perform image segmentation for each frame
2. Match segments from adjacengt frames

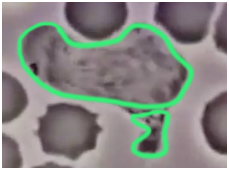


Over/under segmentation



Over/under segmentation





Over/under segmentation

Over/under segmentation

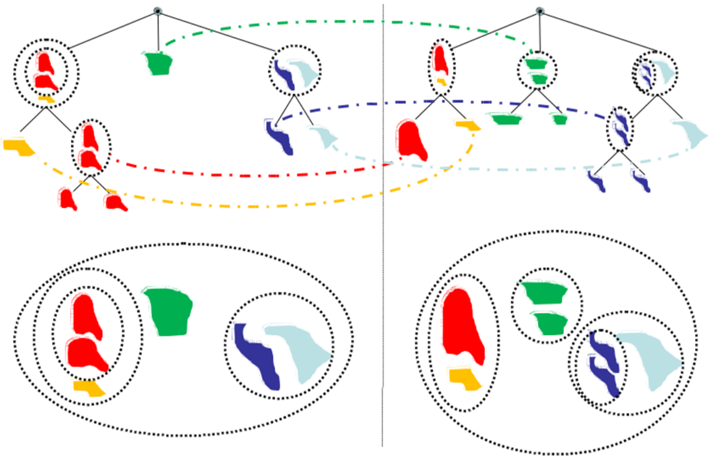
Challenge: Biological cell division vs. over-segmentation

Over/under segmentation

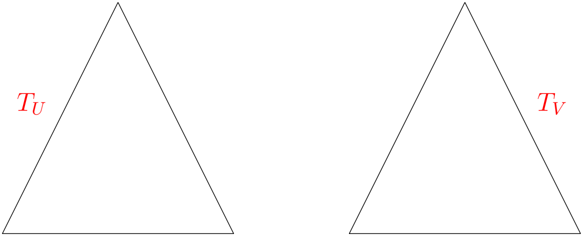
Challenge: Biological cell division vs. over-segmentation

⇒ [Mosig *et al.*, 2009]: Integrate identification and tracking steps!

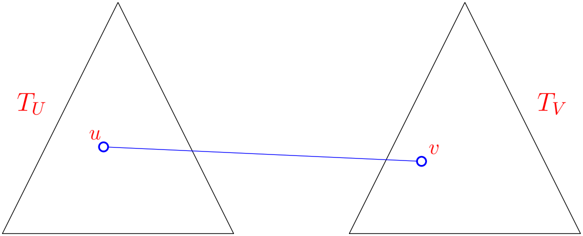
Cosegmentation



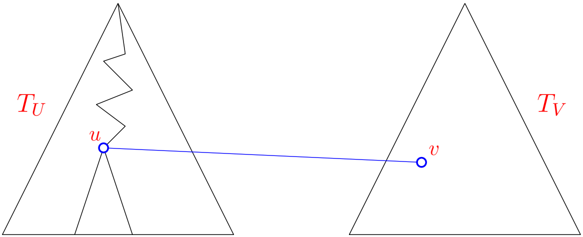
Tree-Constrained Matching



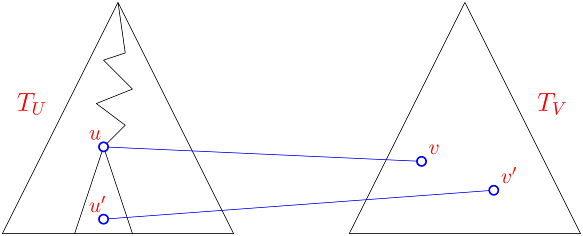
Tree-Constrained Matching



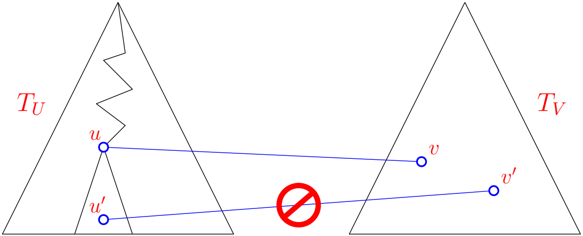
Tree-Constrained Matching



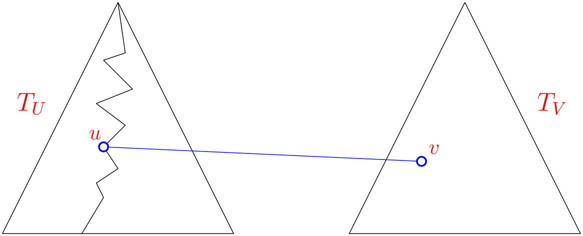
Tree-Constrained Matching



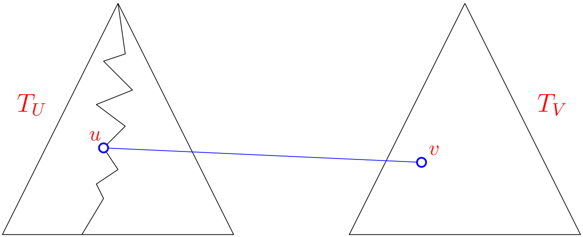
Tree-Constrained Matching



Tree-Constrained Matching

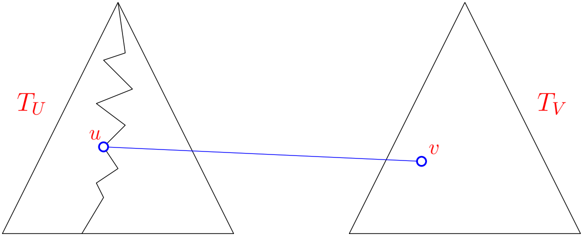


Tree-Constrained Matching



Given a weighted bipartite graph (U, V, E) and trees T_U and T_V over U and V, we want maximum weight matching $\mathcal{M}$ such that matched vertices in T_U and T_V are not comparable

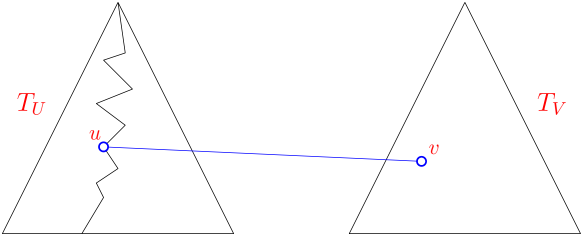
Tree-Constrained Matching



Given a weighted bipartite graph (U, V, E) and trees T_U and T_V over U and V, we want maximum weight matching $\mathcal{M}$ such that matched vertices in T_U and T_V are not comparable

Mosig *et al.* introduced TCM and gave a LP formulation, which they claimed was totally unimodular

Tree-Constrained Matching



Given a weighted bipartite graph (U, V, E) and trees T_U and T_V over U and V, we want maximum weight matching $\mathcal{M}$ such that matched vertices in T_U and T_V are not comparable

Mosig *et al.* introduced TCM and gave a LP formulation, which they claimed was totally unimodular

But...

Our Results

- *Tree-Constrained Matching* as special case of maximum independent set in 2-interval graphs

Our Results

- *Tree-Constrained Matching* as special case of maximum independent set in 2-interval graphs
- $\mathcal{APX}$-hard, 2-approximation

Our Results

- *Tree-Constrained Matching* as special case of maximum independent set in 2-interval graphs
- $\mathcal{APX}$-hard, 2-approximation
- fractional local ratio and properties of BFS

Our Results

- *Tree-Constrained Matching* as special case of maximum independent set in 2-interval graphs
- $\mathcal{APX}$-hard, 2-approximation
- fractional local ratio and properties of BFS
- Generalization to
  - posets (uncertainty in clustering): 4ρ-approximation

Our Results

- *Tree-Constrained Matching* as special case of maximum independent set in 2-interval graphs
- $\mathcal{APX}$-hard, 2-approximation
- fractional local ratio and properties of BFS
- Generalization to
  - posets (uncertainty in clustering): 4ρ-approximation
  - $k > 2$ frames: $2k\rho$-approximation

Our Results

- *Tree-Constrained Matching* as special case of maximum independent set in 2-interval graphs
- $\mathcal{APX}$-hard, 2-approximation
- fractional local ratio and properties of BFS
- Generalization to
  - posets (uncertainty in clustering): 4ρ-approximation
  - $k > 2$ frames: $2k\rho$-approximation
- dependence on ρ unavoidable

Plant Breeding

with M. El-Kebir

- *'Plant breeding is the art and science of changing the genetics of plants for the benefit of humankind'*

Plant Breeding

with M. El-Kebir

- ‘*Plant breeding* is the art and science of changing the genetics of plants for the benefit of humankind’
  - Common practice ever since mankind moved from hunting-gathering to farming

Plant Breeding

with M. El-Kebir

- *'Plant breeding is the art and science of changing the genetics of plants for the benefit of humankind'*
  - Common practice ever since mankind moved from hunting-gathering to farming
  - At first, simply selection for desirable traits

Plant Breeding

with M. El-Kebir

- *'Plant breeding is the art and science of changing the genetics of plants for the benefit of humankind'*
  - Common practice ever since mankind moved from hunting-gathering to farming
  - At first, simply selection for desirable traits
  - Now, we can plan more systematically

Powdery Mildew in Pepper

- Fungal disease incited by *Leveillula taurica*

Powdery Mildew in Pepper

- Fungal disease incited by *Leveillula taurica*
- Infestations result in sun-scalded fruit and crop loss

Powdery Mildew in Pepper

- Fungal disease incited by *Leveillula taurica*
- Infestations result in sun-scalded fruit and crop loss
- Pathogen is resistant to fungicides

Powdery Mildew in Pepper

- Fungal disease incited by *Leveillula taurica*
- Infestations result in sun-scalded fruit and crop loss
- Pathogen is resistant to fungicides

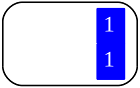
$\Rightarrow$ Host plant resistance is desired

Powdery Mildew in Pepper (cont'd)

- Consider only two traits: resistance and pungency

Powdery Mildew in Pepper (cont'd)

- Consider only two traits: resistance and pungency
  - Elite line is sweet but susceptible



- Pungency is monogenic
  - 0 : pungent
  - 1 : sweet

Powdery Mildew in Pepper (cont'd)

- Consider only two traits: resistance and pungency
  - Elite line is sweet but susceptible

0	0	0	1
0	0	0	1

  - Wildtype is resistant but pungent

1	1	1	0
1	1	1	0

- Pungency is monogenic
  - 0 : pungent
  - 1 : sweet
- Resistance is polygenic
  - 0 : susceptible
  - 1 : resistant

Powdery Mildew in Pepper (cont'd)

- Consider only two traits: resistance and pungency
  - Elite line is sweet but susceptible

0	0	0	1
0	0	0	1

  - Wildtype is resistant but pungent

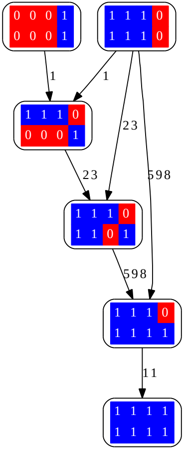
1	1	1	0
1	1	1	0

  - Desired is resistant and sweet

1	1	1	1
1	1	1	1

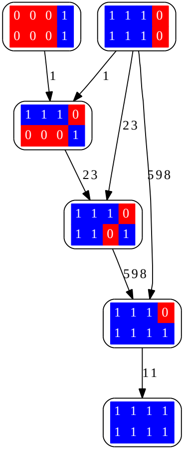
- Pungency is monogenic
  - 0 : pungent
  - 1 : sweet
- Resistance is polygenic
  - 0 : susceptible
  - 1 : resistant

Optimal Schedule



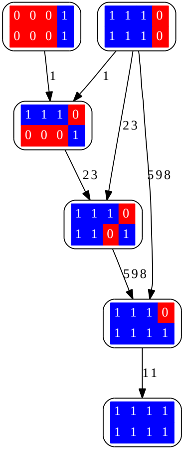
- 4 generations, 4 crossings, 634 individuals

Optimal Schedule



- 4 generations, 4 crossings, 634 individuals
- $\mathcal{NP}$-hard

Optimal Schedule



- 4 generations, 4 crossings, 634 individuals
- $\mathcal{NP}$-hard
- ingredients:
  - advanced mathematical programming techniques
  - combinatorial structure
  - implicit enumeration