

False positives detection in connectomics through hierarchical sparsity

matteo.frigo@epfl.ch

by

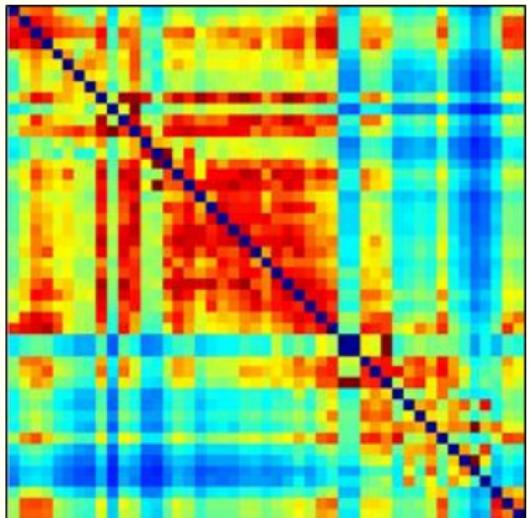
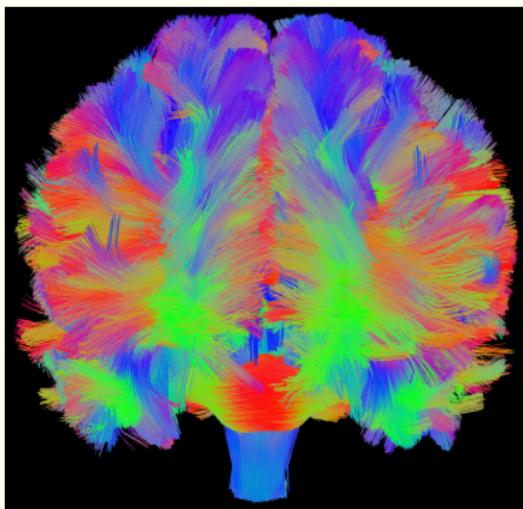
Matteo Frigo

Joint work with...

- Prof. Giandomenico Orlandi
- Prof. Jean-Philippe Thiran
- Prof. Alessandro Daducci
- Muhamed Barakovic



The goal: structural connectivity



Quantify the connections within the brain

The problem: false positives

Tractography-based connectomes are dominated by
false-positive connections

Maier-Hein et al., 2016 (bioarxiv)

- 96 distinct tractography pipelines
- “*most algorithms routinely extracted many false positive bundles*”
- “*Tractography identifies more invalid than valid bundles*”
- “***Tractography is fundamentally ill-posed***”

The problem: false positives

Connectome sensitivity or specificity: which is more important?

- “the impact of FPs and FNs on empirical connectomes indicate that specificity is at least twice as important as sensitivity”

Zalesky, A. et al., 2016 (*NeuroImage*)

Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited

- “The methods that show the highest sensitivity show the lowest specificity, and vice versa”

Thomas, C. et al., 2014 (*PNAS*)

Forward model

$$y = Ax + \varepsilon$$

Given

x : vector of weights

A : dictionary for the fibres

ε : systematic+random error

Obtain

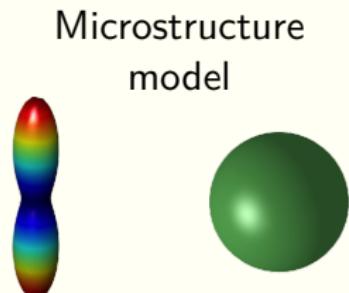
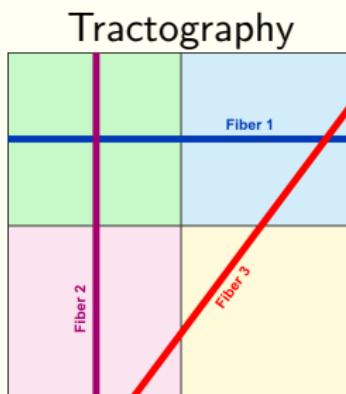
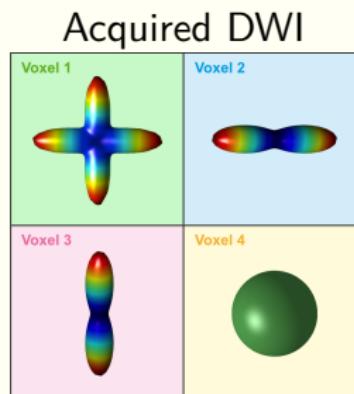
y : acquired dMRI data

Forward model: COMMIT

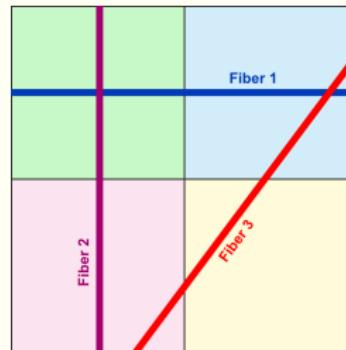
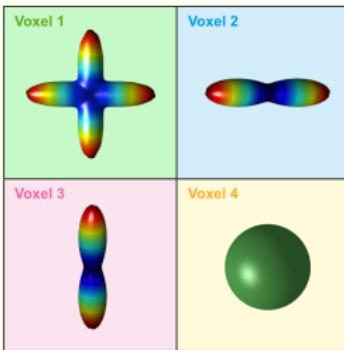
Convex Optimisation Modelling for Microstructure Informed Tractography

Daducci et al., 2015 (IEEE TMI)

Input:



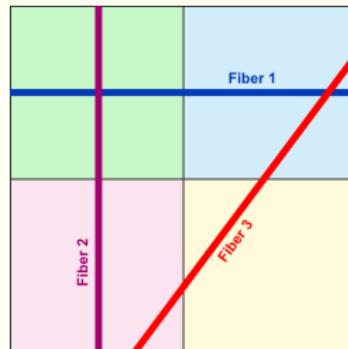
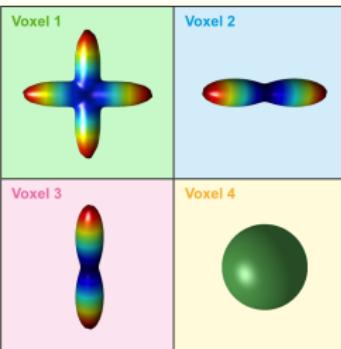
Forward model: COMMIT



$$y =$$

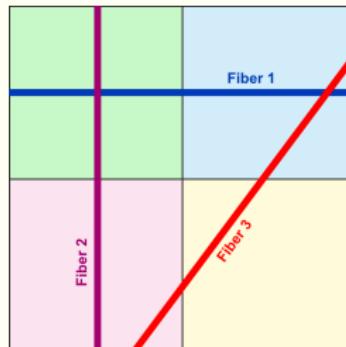
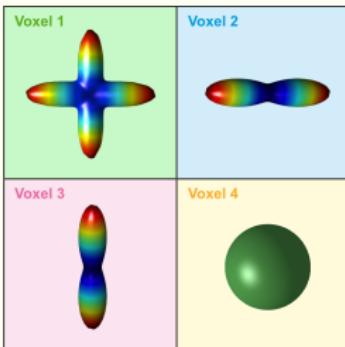
$$X = \begin{matrix} ? & ? & ? & ? & ? & ? & ? \end{matrix}$$

Forward model: COMMIT



$$X = \begin{bmatrix} ? & ? & ? & ? & ? & ? & ? \end{bmatrix}$$

Forward model: COMMIT



$$y = A \cdot x$$

Y-axis labels: Voxel 1, Voxel 2, Voxel 3, Voxel 4

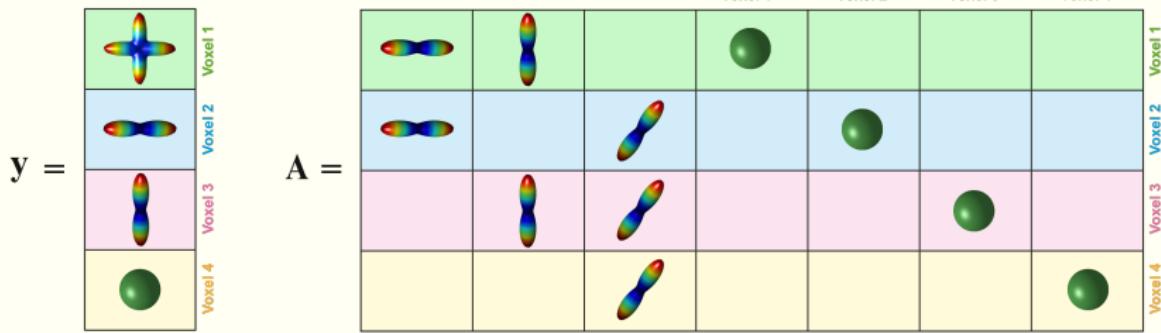
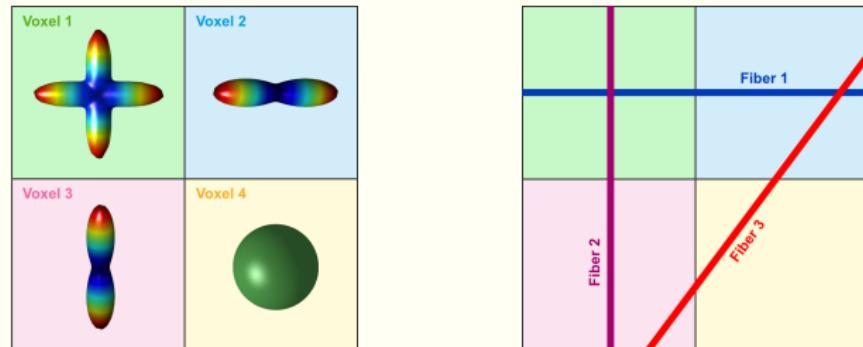
X-axis labels: Fiber 1, Fiber 2, Fiber 3

Matrix A (Forward Model Matrix):

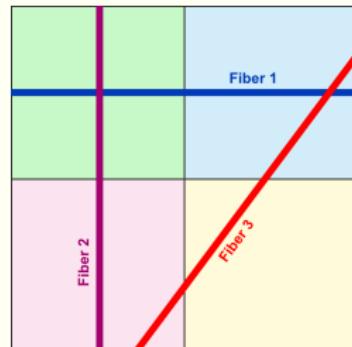
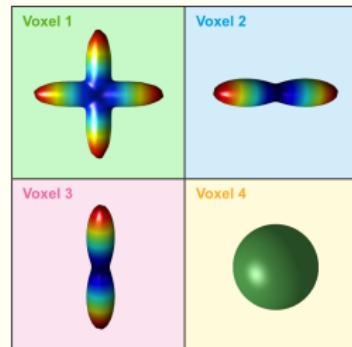
	Fiber 1	Fiber 2	Fiber 3						
Voxel 1	Tensor Field								
Voxel 2		Tensor Field							
Voxel 3			Tensor Field						
Voxel 4				Tensor Field					

$$x = [? ? ? ? ? ? ?]$$

Forward model: COMMIT



Forward model: COMMIT



$$y = Ax$$

where y is the measured signal vector, A is the system matrix, and x is the parameter vector.

$y =$

Voxel 1	Voxel 2	Voxel 3	Voxel 4

$A =$

Fiber 1	Fiber 2	Fiber 3	CSF in voxel 1	CSF in voxel 2	CSF in voxel 3	CSF in voxel 4

$x =$

<input checked="" type="checkbox"/>							
-------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------

Inverse problem

$$y = Ax + \varepsilon$$

Recover x from the acquired data y

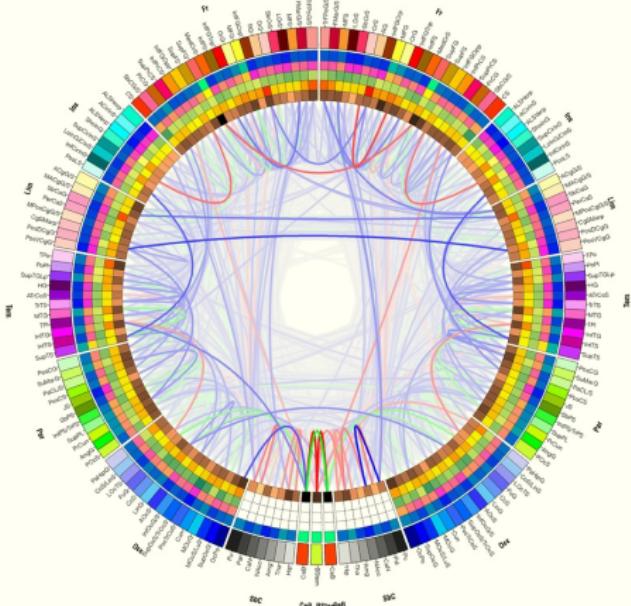
$$x^* = \underset{x \in \mathbb{R}_+^c}{\operatorname{argmin}} \frac{1}{2} \|Ax - y\|_2^2 + \lambda \Omega(x)$$

$\Omega : \mathbb{R}^c \rightarrow \mathbb{R}$ anatomy-based penalty

$\lambda \geq 0$ regularisation parameter

Anatomical prior knowledge

Brain hierarchy

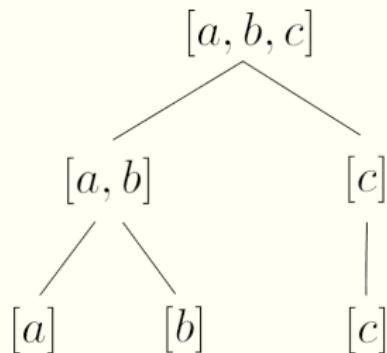
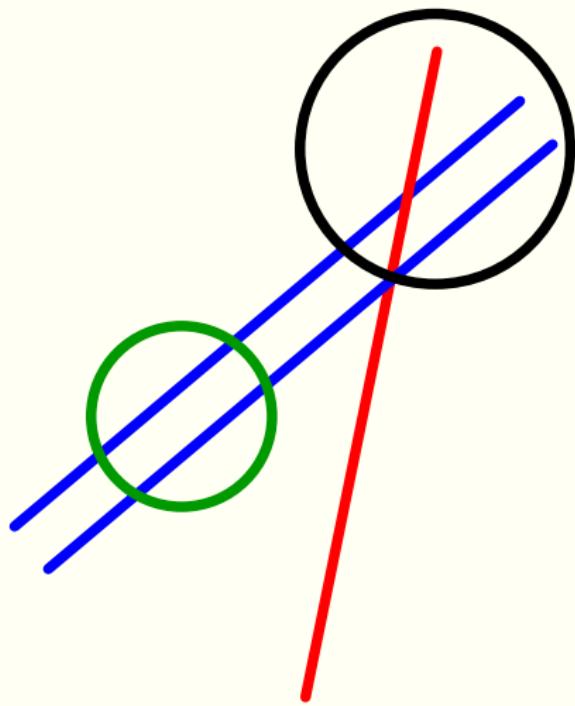


Connections within the brain are endowed with a hidden hierarchical pattern which should be exploited.

- ▢ Zhou et al., 2006
- ▢ Duarte-Carvajalino et al., 2012
- ▢ Moreno-Dominguez et al., 2012

© John Darrell Van Horn

Abstract model



$$\mathcal{G} = \left([[a], [b], [a, b], [c], [a, b, c]], \preceq \right)$$

E.g.:

$$[a] \preceq [a, b] \preceq [c]$$

$$[a, b, c] \not\preceq [c]$$

Nomenclature

- ❖ Valid/Invalid Bundle

- VB : fascicle belonging to the ground truth
 - IB : fascicle made of false positives

- ❖ Valid/Invalid Connection

- VC : fibre belonging to a VB
 - IC : fibre belonging to an IB

Formalisation

The problem

Find $x^* \in \mathbb{R}^c$ such that

$$x^* = \underset{x \in \mathbb{R}_+^c}{\operatorname{argmin}} \frac{1}{2} \|Ax - y\|_2^2 + \lambda \Omega(x)$$

RLS: Regularised Least Squares

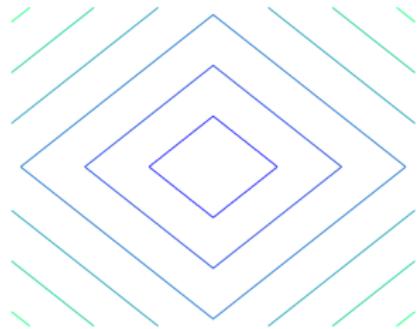
Penalty term Ω

Objective:

Cancel invalid bundles

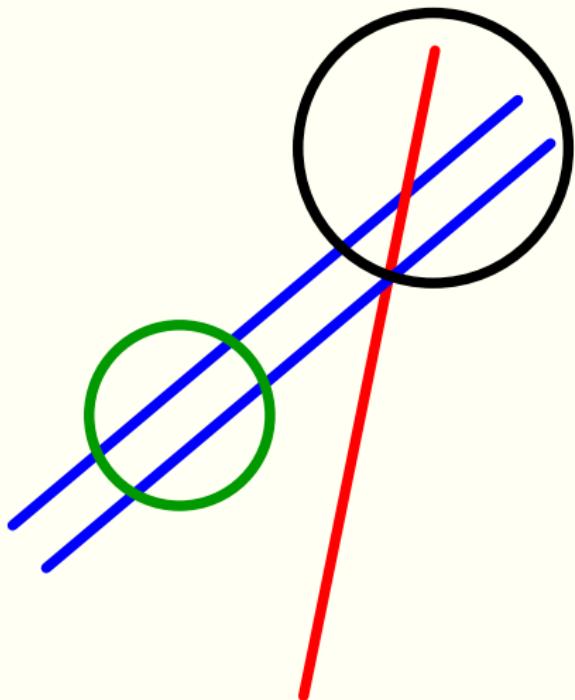


Sparsity in the space of fibres



Classical sparsity:
 $\Omega(x) = \|x\|_1$

Hierarchical sparsity



$$\Omega(x) = \|X_{\mathcal{G}}\|_1 = \sum_{g \in \mathcal{G}} \|x_{|g}\|_2$$

$$g = [a, b], \quad x = [x_a, x_b, x_c] \\ \Rightarrow x_{|g} = [x_a, x_b, 0]$$

$$\Omega(x) = \|[x_a, 0, 0]\|_2 + \|[0, x_b, 0]\|_2 + \\ \|[x_a, x_b, 0]\|_2 + \|[0, 0, x_c]\|_2 + \\ \|[x_a, x_b, x_c]\|_2$$

The problem

Find $x^* \in \mathbb{R}^c$ such that

$$x^* = \underset{x \in \mathbb{R}^c}{\operatorname{argmin}} \underbrace{\frac{1}{2} \|Ax - y\|_2^2}_{\text{smooth}} + \underbrace{\lambda \Omega(x) + \iota_{\geq 0}(x)}_{\text{non smooth}}$$

HNNLS: Hierarchical Non Negative Least Squares

Numerical optimisation

Proximal operator

Proximal operator of f :

$$\text{prox}_f(x) = \underset{y \in \mathbb{R}^c}{\operatorname{argmin}} \frac{1}{2} \|x - y\|_2^2 + f(y)$$

If $f(x) = \iota_S(x)$ with S convex set:

$$\text{prox}_f(x) = \Pi_S(x)$$

$$\text{prox}_{\iota_{\geq 0}}(x) = \Pi_{\geq 0}(x)$$

Proximal of Ω

Let \mathcal{G} tree structure with order \preceq

$$\Omega(x) = \sum_{g \in \mathcal{G}} \omega_g \|x_{|g}\|$$

Compute $\text{prox}_\Omega(x)$

1. Set $v = x$.
2. For $g \in \mathcal{G}$ following \preceq do

$$v_{|g} \longleftarrow v_{|g} - \Pi_{\|\cdot\|_* \leq \omega_g}(v_{|g}).$$

Jenatton et al., "Proximal methods for sparse hierarchical dictionary learning",
2010, ICML-10

Davis-Yin splitting scheme(2015)

Objective:

$$x^* = \operatorname{argmin}_{y \in \mathbb{R}^c} \overbrace{f(y)}^{\text{s}} + \overbrace{g(y) + h(y)}^{\text{ns}}$$

Algorithm:

$$y = \mathbf{prox}_{\gamma g}(x)$$

$$z = \mathbf{prox}_{\gamma h}(2x - y - \gamma \nabla f(y))$$

$$x^+ = x - y + z$$

Rate of convergence:

$$\mathcal{O}(1/k)$$

Non negativity

Definition (Absolute norm)

A norm $\Omega : X \rightarrow \mathbb{R}$ is called *absolute* if $\forall u, v \in \mathbb{R}^N$ such that $|u_j| \leq |v_j|$ for all j implies $\Omega(u) \leq \Omega(v)$.

Theorem (Proximal operator of absolute norms)

Let $w \in \mathbb{R}^n$ and $\lambda > 0$. Consider an absolute norm Ω . We have

$$\operatorname{argmin}_{z \in \mathbb{R}_+^n} \left[\frac{1}{2} \|w - z\|_2^2 + \lambda \Omega(z) \right] = \operatorname{argmin}_{z \in \mathbb{R}^n} \left[\frac{1}{2} \| [w]_+ - z \|_2^2 + \lambda \Omega(z) \right].$$

$$\begin{aligned}\operatorname{prox}_{\lambda \Omega + \iota_{\geq 0}}(w) &= \operatorname{prox}_{\lambda \Omega}([w]_+) \\ &= \operatorname{prox}_{\lambda \Omega} \left(\operatorname{prox}_{\iota_{\geq 0}}(w) \right)\end{aligned}$$

Jenatton et al., "Proximal methods for hierarchical sparse coding", 2011, JMLR

FISTA

Fast Iterative Shrinkage Thresholding Algorithm (2009)

Objective:

$$x^* = \underset{x \in \mathbb{R}^c}{\operatorname{argmin}} \overbrace{f(x)}^{\text{s}} + \overbrace{g(x)}^{\text{ns}}$$

Algorithm ($t_0 = 1$):

$$x_k = \operatorname{prox}_{t_k g}(x_{k-1} - t_k \nabla f(x_{k-1}))$$

$$t_{k+1} = \frac{1 + \sqrt{1 + 4t_k^2}}{2}$$

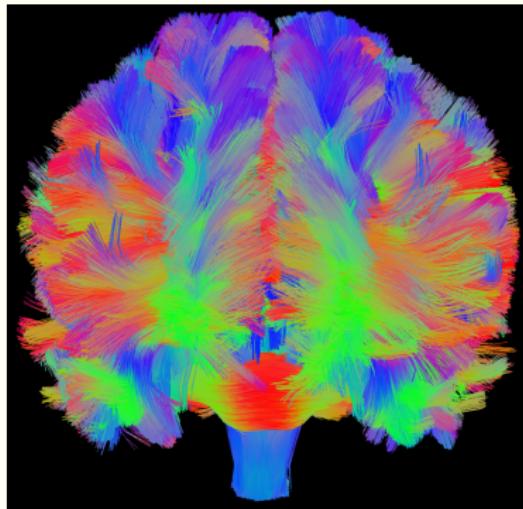
$$y_{k+1} = x_k + \left(\frac{t_k - 1}{t_{k+1}} \right) (x_k - x_{k-1})$$

Rate of convergence:

$$\mathcal{O}(1/k^2)$$

Results

Dataset



$$\|Ax - y\|_2^2$$

- A Tractography:
Particle Filtering¹
- y TDI map given by the
fibres belonging to VBs

¹ Girard and Descoteaux. *Online filtering tractography: tracking with anatomical priors*, ISMRM 2013

Competitors

SIFT2

Smith, Robert E., et al. "*SIFT2: Enabling dense quantitative assessment of brain white matter connectivity using streamlines tractography.*" *Neuroimage* 119 (2015): 338-351.

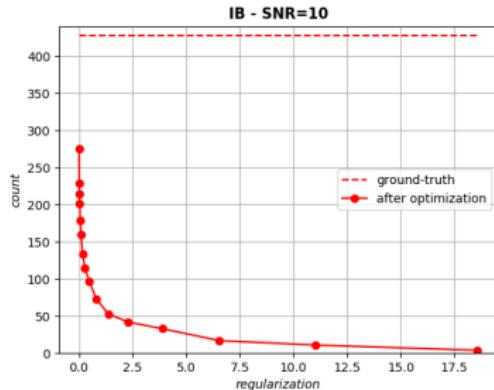
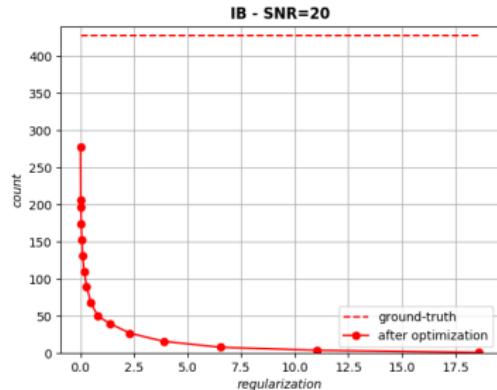
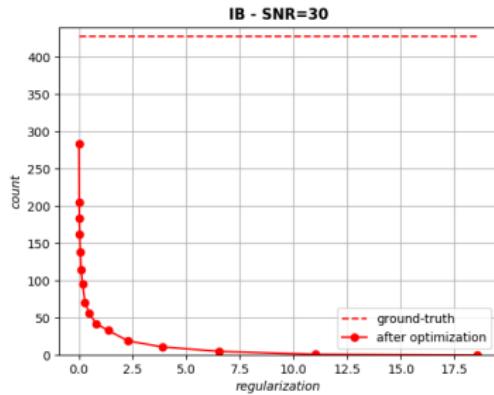
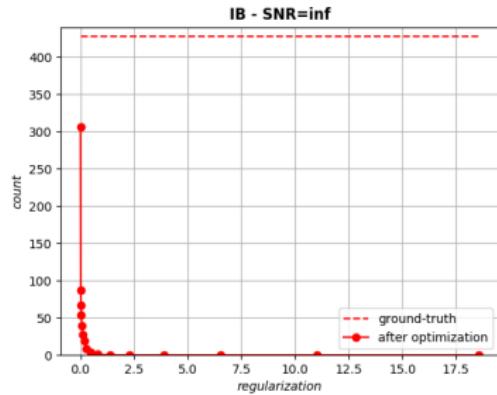
LiFE

Pestilli, Franco, et al. "*Evaluation and statistical inference for human connectomes.*" *Nature methods* 11.10 (2014): 1058-1063.

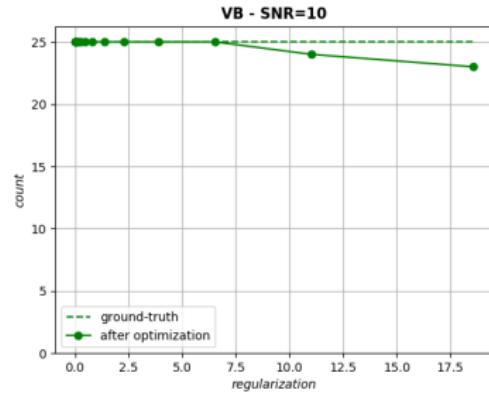
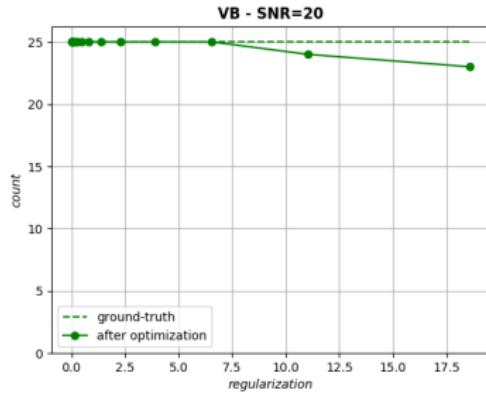
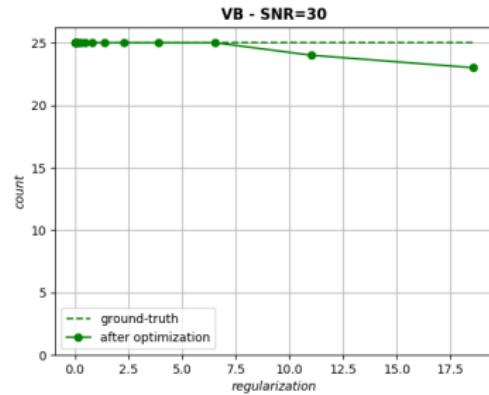
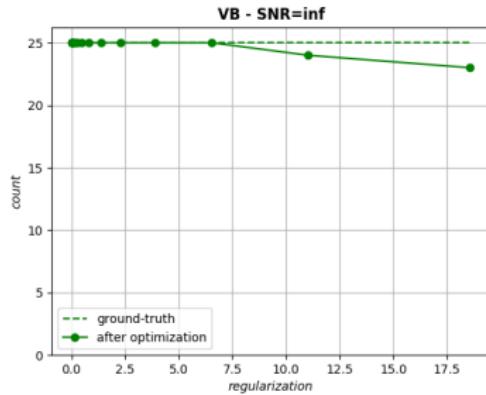
Same goal of HNNLS

No convexity, huge computational effort,
no microstructure information, ...

Results: Invalid Bundles (IB)



Results: Valid Bundles (VB)

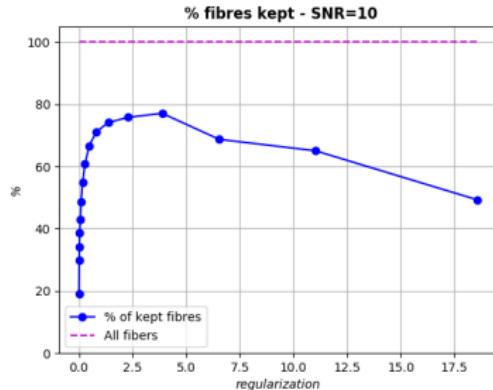
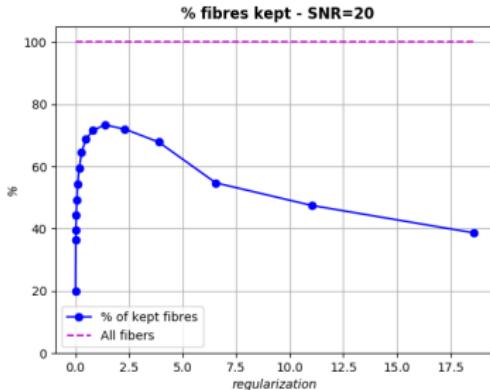
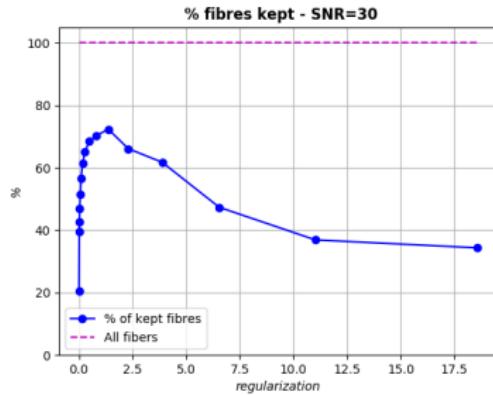
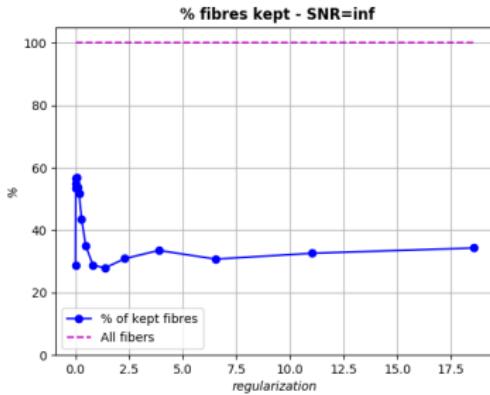


Results

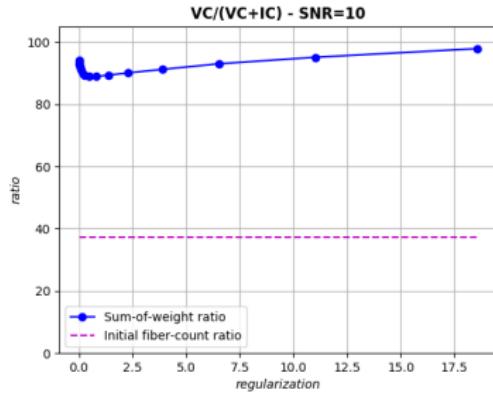
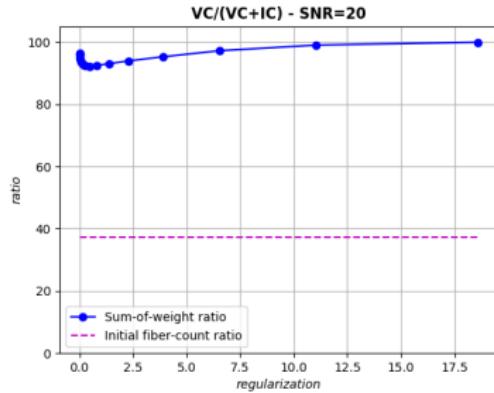
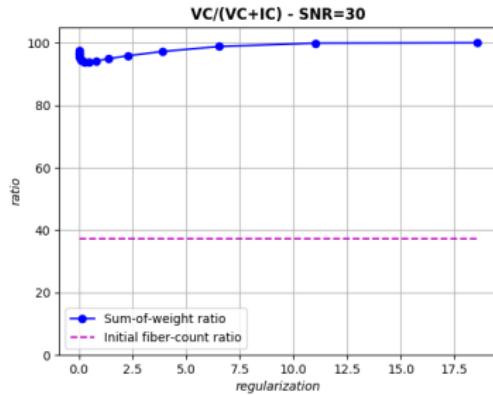
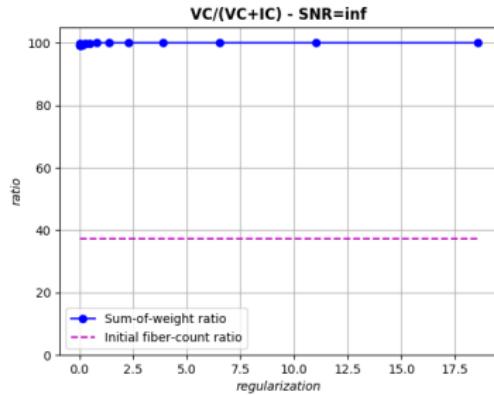
$$\lambda \in \{ 0 \quad 0.012 \quad 0.020 \quad 0.035 \quad 0.059 \quad 0.100 \quad 0.168 \quad 0.284 \\ 0.480 \quad 0.809 \quad 1.364 \quad 2.300 \quad 3.879 \quad 6.540 \quad 11.02 \quad 18.59 \}$$

	SIFT2	LiFE	$\lambda = 0$	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
VB	25	25	25	25	25	25	25	25	25
IB	427	423	284	205	184	162	138	115	95
	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}	λ_{13}	λ_{14}	λ_{15}
VB	25	25	25	25	25	25	25	24	23
IB	71	56	42	33	19	11	5	1	0

Results: Kept fibres



Results: Valid Connections



Future work

Still to be done...

- ☛ Sensitivity analysis w.r.t. tractogram
- ☛ Sensitivity analysis w.r.t. clustering thresholds
- ☛ Set up a strategy for selecting λ
- ☛ Exploit more sophisticated solvers (Canales, 2015)

Thank you

regularisation restricted optimisation norm
sparsity model connections mri estimated
results

algorithm tractography zero iteration connectivity group image

defined sparse order
matrix brain
structure tree
linear argmin
recover solution acquired solve voxel local gradient invalid
bundles diffusion minimisation
imaging hierarchical formulation neuroimaging

remark fibre convergence data

false positives signal microstructure

Backup

Reweighted ℓ_1

Objective:

Smart choice of w_g

Iteratively define

$$w_g = \frac{1}{\|x_g\|_1 + \varepsilon}$$

and restart the optimisation procedure.

Candes, Emmanuel J., Michael B. Wakin, and Stephen P. Boyd. "Enhancing sparsity by reweighted ℓ_1 minimization." *Journal of Fourier analysis and applications*, 2008

RUMBA

Robust and Unbiased Model-Based Spherical Deconvolution

$$\mathbb{P}(S_i | \bar{S}_i, \sigma^2, n) = \frac{\bar{S}_i}{\sigma^2} \left(\frac{S_i}{\bar{S}_i} \right)^n \exp \left\{ -\frac{1}{2\sigma^2} [S_i^2 + \bar{S}_i^2] \right\} I_{n-1} \left(\frac{S_i \bar{S}_i}{\sigma^2} \right) u(S_i),$$
$$f = f \circ \frac{H^t \left[S \circ \frac{I_n(S \circ Hf / \sigma^2)}{I_{n-1}(S \circ Hf / \sigma^2)} \right]}{H^t H f}$$

where $S_i = (Hf)_{i \cdot}$

Canales-Rodríguez, Erick J., et al. *Spherical deconvolution of multichannel diffusion MRI data with non-Gaussian noise models and spatial regularization*. PloS one 10.10 (2015).

Software

- ❖ Microstructure Informed Tractography
 - ❖ COMMIT *
 - ❖ HNNLS ‡
- ❖ Clustering
 - ❖ Tract Querier†
 - ❖ QuickBundlesX‡◊

* github.com/daducci/COMMIT

† github.com/demianw/tract_querier

‡ github.com/matteofrigo/COMMIT_demo

◊ github.com/MarcCote/dipy/