



MAX-PLANCK-GESELLSCHAFT

Unified Diffusion Tensor Estimation using Space-Varying Coefficient Models

Susanne Heim¹, Philipp G. Sämann², Ludwig Fahrmeir¹

¹ Department of Statistics, Ludwig-Maximilians-University, Munich

² Research Group NMR, Max-Planck-Institute of Psychiatry, Munich

susanne.heim@stat.uni-muenchen.de; <http://www.stat.uni-muenchen.de/~heim>



1. Review & Preview

To date, DTI raw data are processed in multiple steps risking error agglomeration:

First, there are various approaches to estimate the diffusion tensor, for example

- voxelwise regression of the linearized Stejskal-Tanner equation [1]
- non-Gaussian error regression [2]
- robust techniques [3, 4]
- variational framework [5].

Second, thermal noise contained in the raw data requires regularization of the tensor field, currently performed in a post-hoc step employing isotropic Gaussian [6] or anisotropic kernels [7] or other adaptive filtering schemes [8].

Third, due to the limited resolution in DTI several applications, in particular fiber tracking, need a subsequent interpolation, e. g. B-splines [9] or anisotropic sigmoid interpolation kernel [10].

GOAL: To provide a unified framework comprising estimation (*level1*), smoothing (*level2*) and interpolation (*level3*), thereby considering spatial correlation ab initio.

2. Methodology

Starting point: separate regression (*level1*) at each voxel, indexed by $s = 1, \dots, n$

$$y_i(s) = \mathbf{x}_i' \boldsymbol{\beta}(s) + \varepsilon_i(s), \quad i = 1, \dots, r,$$

$y_i(s)$ signal intensities transformed according to Stejskal-Tanner

\mathbf{x}_i space-invariant covariate vector constructed from the i th applied magnetic gradient

$\boldsymbol{\beta}(s)$ coefficient vector of unknown elements of the local tensor

$\varepsilon_i(s) \stackrel{iid}{\sim} N(0, \sigma^2)$ error term

CORE IDEA: Combine the $n = n_1 \times n_2 \times n_3$ separate regression models into **ONE** joint 3d space-varying coefficient model (SVCM).

1 SVCM – Tensor Product Variant:

Project $\boldsymbol{\beta}_j$ ($j = 1, \dots, p = 6$) onto penalized tensor product B-splines:

$$\boldsymbol{\beta}_j(s) = \sum_{v=1}^{KLM} \mathbf{B}(s, v) \boldsymbol{\gamma}_j(v) = (\mathbf{B}_3 \otimes \mathbf{B}_2 \otimes \mathbf{B}_1)(s, \cdot) \boldsymbol{\gamma}_j,$$

yielding the penalized objective:

$$\underbrace{\| \mathbf{y} - (\mathbf{B} \otimes \mathbf{X}) \boldsymbol{\gamma} \|^2}_{\text{least squares term}} + \underbrace{\lambda_1 \| \mathbf{P}_1 \boldsymbol{\gamma} \|^2 + \lambda_2 \| \mathbf{P}_2 \boldsymbol{\gamma} \|^2 + \lambda_3 \| \mathbf{P}_3 \boldsymbol{\gamma} \|^2}_{\text{penalty term}} \rightarrow \min_{\boldsymbol{\gamma}, \boldsymbol{\lambda}}$$

$\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)'$ tuning parameter for trade-off between smoothness and fidelity to the data
 $\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3$ dimension-specific difference penalties

Benefit of basis function approach:

straightforward increase of the resolution without additional interpolation.

Computational issues:

- \mathbf{B} is the same for all coefficient surfaces $\boldsymbol{\beta}_j$ and has to be calculated only once.
- Assuming $K \times L \times M = 32 \times 32 \times 8$ knots resp. basis functions for a region of interest sized $40 \times 40 \times 10$ voxels, \mathbf{B} has approx. 3×10^9 elements: if each floating point takes 8 bytes, $(\mathbf{B} \otimes \mathbf{X})'(\mathbf{B} \otimes \mathbf{X})$ would occupy more than 18 Gb.
- ➔ use sparse matrix classes
- ➔ switch to computationally efficient approximation (see 2)

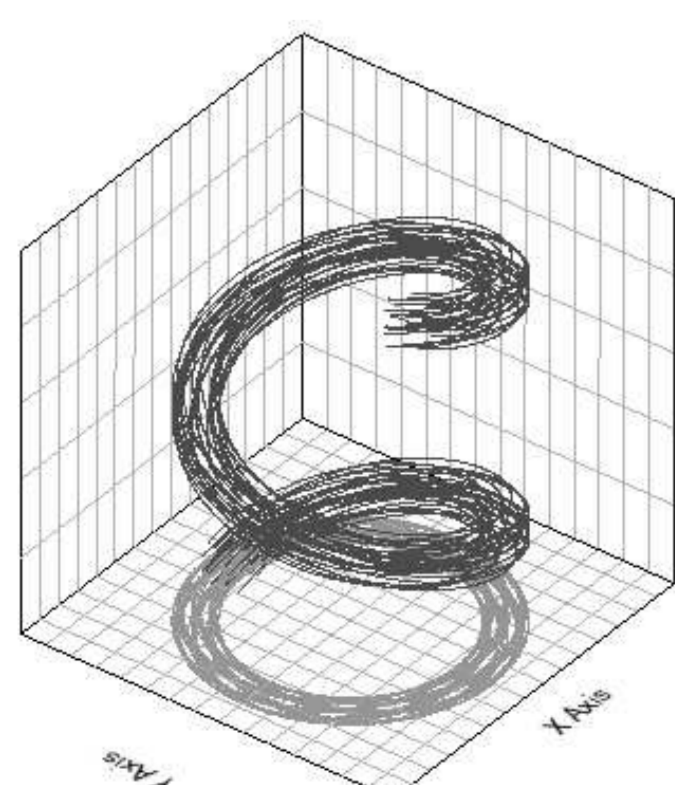
2 SVCM – Sequential Variant:

Drawing on 2d penalized smoothing [11], the coefficient vector $\boldsymbol{\gamma}$ results from consecutive univariate smoothing with the observational ($r \times n_1 \times n_2 \times n_3$)-data array \mathbf{Y} as input to the first smoothing cycle exclusively. The (preliminary) coefficients attained last serve each as input to further iterations. Implementation bases on array regression [12].

Extra benefit:

a larger number of basis functions are allowed ➔ detail preservation

3. Simulation Study



- 3d data grid of $\{15 \times 15 \times 5\} \subset \mathbb{R}^3$ voxels
- $2 \times 2 \times 4 \text{ mm}^3$ voxel size
- six 3d varying coefficient surfaces
- spiral tensors are anisotropic (2 : 1 : 1 eigenvalue ratio); background tensors are fully isotropic
- $N = 100$ noisy data sets generated with Gaussian error ($\sigma = 10$)
- acquisition schemes: 1 [repeat] $\times 6$ [directions], 3×6 , 1×15 , 1×31

Figure 1: Design of the simulation study and geometry of the underlying fiber bundle.

We compared:

voxelwise regression (*level1*)

voxelwise regression plus Gaussian kernel (*level2*)

voxelwise regression plus Gaussian kernel plus tri-linear interpolation (*level3*)

SVCM (sequential variant) with 1st order difference penalties, global tuning parameter, linear (*SVCMin*) vs quadratic (*SVCQuad*) B-spline basis functions, 1 knot per 1.25 voxels.

Outcome measures:

AMSE MSE averaged over fiber, non-fiber voxels, and all voxels

VMSE MSE per voxel and per tensor element averaged over the N simulated data sets

$\log(\text{VMSE}_{m1} / \text{VMSE}_{m2})$

e. g. a log ratio of 2 indicates seven times larger VMSE of method $m1$ than of $m2$.

4. Results

AMSE at original resolution: Independent of the acquisition scheme, both *SVCMin* and *SVCQuad* outperform standard *level2* but not *level1*. This benefit is most distinct in anisotropic voxels.

AMSE at double resolution: In anisotropic conditions, also both *SVCMin* and *SVCQuad* are superior to *level3*, however, both visual inspection and MSE analyses show background inhomogeneity. Neither increasing directions nor repeats could improve SVCM estimation quality overall voxels. Importantly, interpolation using *SVCMin* outperformed tri-linearly interpolated *level1* or *level3* for all acquisition schemes.

VMSE log ratio:

At original resolution, While *level2* tends to blur the spiral and to evoke additional shape artifacts (rose background blobs, Fig. 2), the *SVCM* catches the "height" better (rose ridges) but on the cost of inferior background smoothness and overshooting at edges (green echoes, in particular with *SVCQuad*). This pattern points to inadequate edge preservation.

At double resolution, VMSE log ratio maps expose the same method-specific deficiencies of either shape and intensity distortions (standard procedures) or wiggly background reproduction in combination with Gibbs phenomena (SVCM).

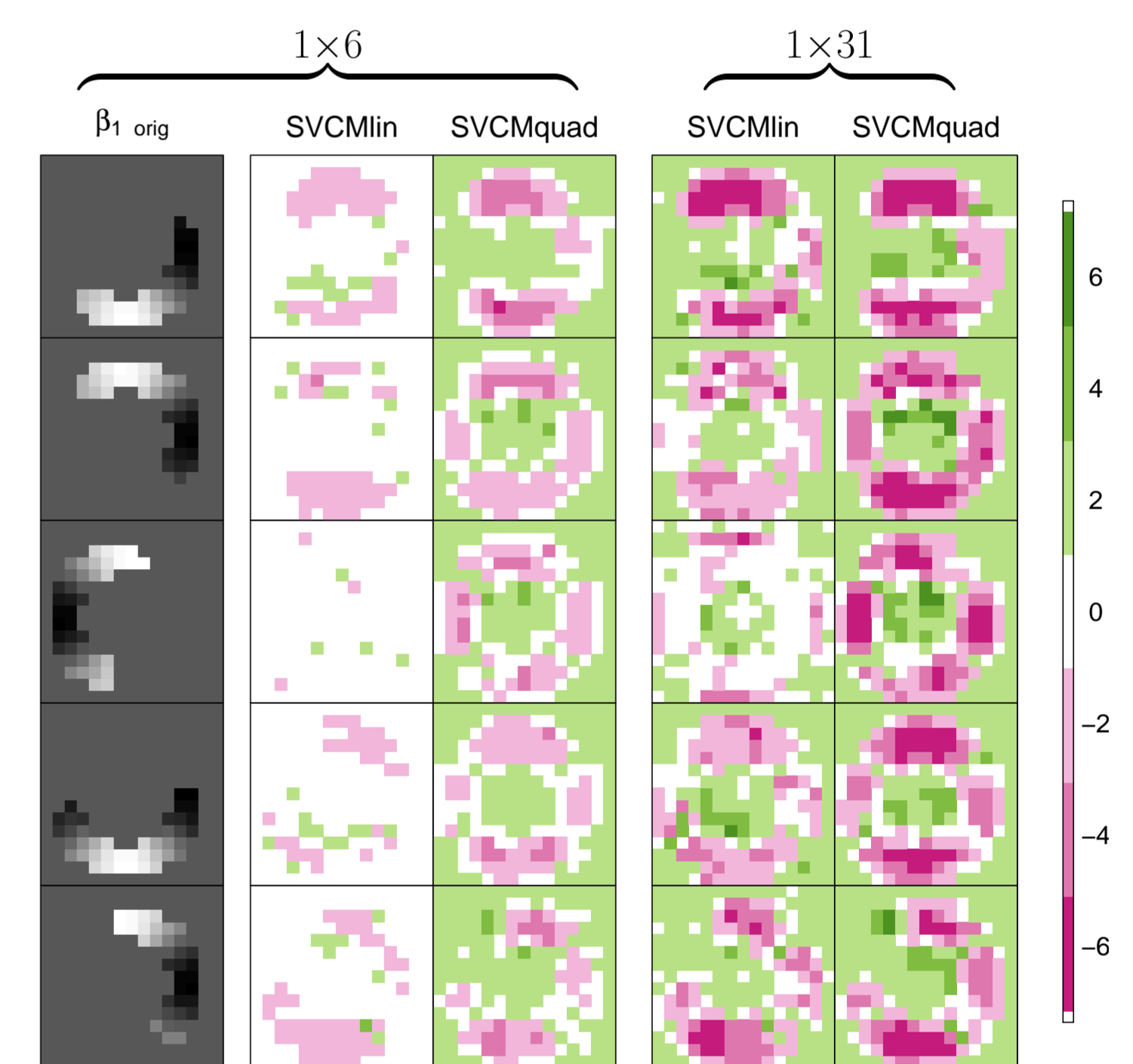


Figure 2: Exemplary maps of $\log(\text{VMSE}_{\text{SVCM}} / \text{VMSE}_{\text{level2}})$ when estimating the first tensor element. Rose (green) indicates higher accuracy of SVCM (*level2*).

5. Limitations and Future Directions

Limitation

insufficient local adaptiveness

nonlinearity of Stejskal-Tanner equation
uncertain or even missing data
positive definiteness constraint

Extension/Modification

- data-driven penalty weights (Fig. 3)
- radial basis functions or wavelets
- logarithmic link function
- observation weights
- e. g. by Cholesky decomposition

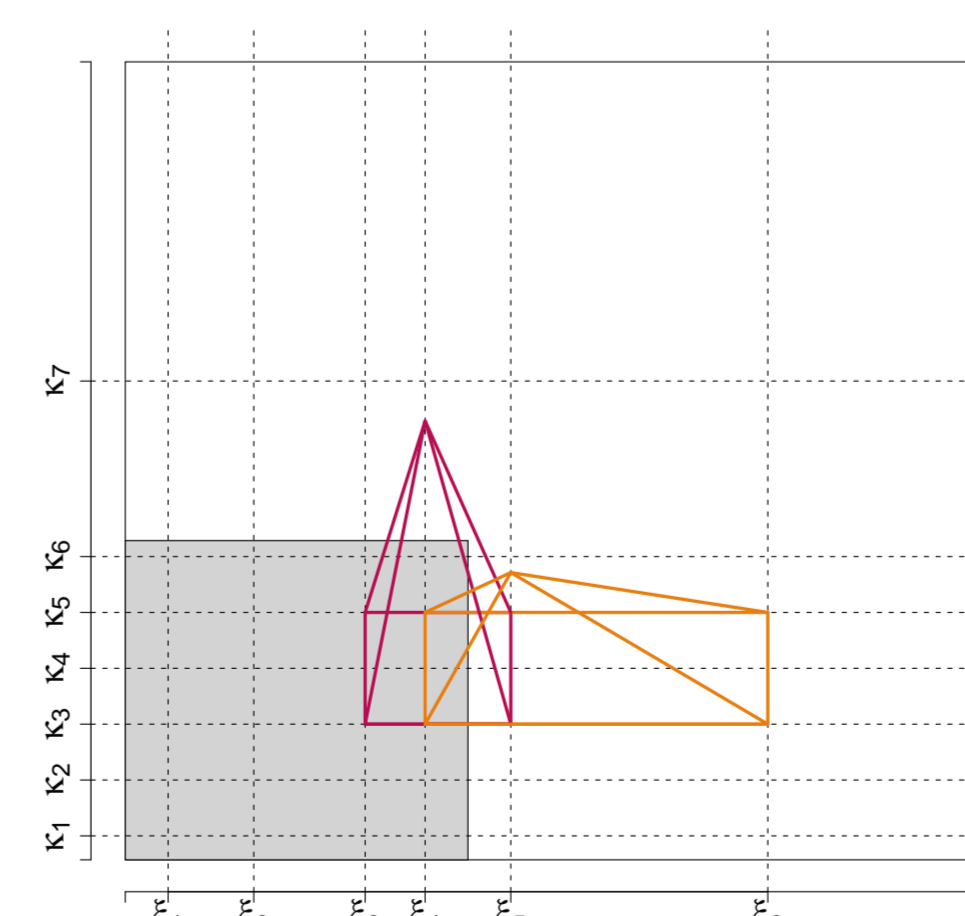


Figure 3: 2d example exhibiting sharp edges. The amplitudes of two neighbouring tensor product B-splines (orange, red) could be allowed to differ appropriately by relaxing the penalty.

References

- [1] Basser et al, J Magn Reson B 103, 1994 [2] Assaf et al, Magn Reson Med 52, 2004 [3] Mangin et al, Med Image Anal 6, 2002 [4] Chang et al, Magn Reson Med 53, 2005 [5] Tschumperlé et al, IEEE ICCV, 2003 [6] Gössel et al, Neuroimage 16, 2002 [7] Ding et al, Magn Reson Med 53, 2005 [8] Samsonov et al, Magn Reson Med 52, 2004 [9] Pajevic et al, J Magn Reson 154, 2002 [10] Mishra et al, Neuroimage, 2006 [11] Dierckx, SIAM J Num Anal 19, 1982 [12] Eilers et al, Comput Statist Data Anal 50, 2006