

Submitted to MRM, 24th April 2019

## Spatially regularized estimation of the tissue homogeneity model parameters in DCE-MRI using proximal minimization

Michal Bartoš<sup>1</sup>, Pavel Rajmic<sup>2</sup>, Michal Šorel<sup>1</sup>, Marie Mangová<sup>2</sup>, Olivier Keunen<sup>3</sup>, Radovan Jiřík<sup>4</sup>

<sup>1</sup>The Czech Academy of Sciences, Institute of Information Theory and Automation, Prague, Czech Republic

<sup>2</sup>SPLab, Department of Telecommunications, FEEC, Brno University of Technology, Brno, Czech Republic

<sup>3</sup>Norlux Neuro-Oncology Laboratory, Department of Oncology, Luxembourg Institute of Health, Luxembourg, Luxembourg

<sup>4</sup>The Czech Academy of Sciences, Institute of Scientific Instruments, Brno, Czech Republic

**Correspondence:** Michal Bartoš, ÚTIA AV ČR, Pod Vodárenskou věží 4, CZ-182 08, Prague, Czech Republic, [bartos@utia.cas.cz](mailto:bartos@utia.cas.cz)

**Submitted to Magnetic Resonance in Medicine. Resubmission date: 24th April 2019**

## **Abstract**

**Purpose:** The Tofts and the extended Tofts models are the pharmacokinetic models commonly used in DCE-MRI perfusion analysis, although they do not provide two important biological markers, namely the plasma flow and the permeability-surface area product. Estimates of such markers are possible using advanced pharmacokinetic models describing the vascular distribution phase, such as the tissue homogeneity model. However, the disadvantage of the advanced models lies in biased and uncertain estimates, especially when the estimates are computed voxel-wise. The goal of this work is to improve the reliability of the estimates by including information from neighboring voxels.

**Theory and Methods:** Information from the neighboring voxels is incorporated in the estimation process through spatial regularization in the form of total variation. The spatial regularization is applied on five maps of perfusion parameters estimated using the tissue homogeneity model. Since the total variation is not differentiable, two proximal techniques of convex optimization are used to numerically solve the problem.

**Results:** The proposed algorithm helps to reduce noise in the estimated perfusion-parameter maps together with improved accuracy of the estimates. These conclusions are proved using a numerical phantom. In addition, experiments on real data show improved spatial consistency and readability of perfusion maps without considerable lowering the quality of fits.

**Conclusion:** The reliability of the DCE-MRI perfusion analysis using the tissue homogeneity model can be improved by employing spatial regularization. The proposed utilization of modern optimization techniques implies only slightly higher computational costs compared to the standard approach without spatial regularization.

## **Keywords**

DCE-MRI, perfusion parameter estimation, spatial regularization, tissue homogeneity model, proximal methods, total variation

# 1 INTRODUCTION

Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is becoming an established tool in obtaining information about tissue perfusion and capillary permeability. The trend today is to represent such information as a set of images/maps of physiological parameters related to the perfusion as for example blood plasma flow, permeability surface area product and plasma or extracellular-extravascular-space volumes. The type and number of the parameters represented by perfusion maps depend on the pharmacokinetic model used. To estimate the perfusion parameters, the respective model is typically fitted to the concentration-time curve of each voxel. These curves are calculated from a T1-weighted MR image sequence recorded before, during and after administration of a contrast agent bolus (1).

The problem of fitting the concentration-time curves is unfortunately a nontrivial optimization problem due to the model nonlinearity, insufficient temporal sampling, poor signal-to-noise ratio and uncertainties in the model and the measurement. This results in bias and uncertainty in the estimates or even wrong estimates because of the presence of local minima. The estimation problems are often categorized as a priori structural identifiability and a posteriori identifiability. The a priori identifiability is influenced by the nonlinear model structure itself (2–4) and by the experimental design – sampling and duration of the experiment (5–10). The a posteriori identifiability includes the errors in the measurement – the signal-to-noise ratio, arterial input function errors and the inaccuracy of conversion from the T1-weighted image sequence to the concentration-time curves (10–14). Additionally, the local minima may also be caused by an improper discretization of the model (15–17).

One way to reduce the uncertainties caused by errors is an incorporation of a spatial prior into the modeling. Such a means of regularization is based on the assumption that neighboring voxels in the parameter maps belonging to the same tissue should have similar values, i.e. the perfusion maps should be piecewise smooth. Although using spatial priors is usual in image reconstructions including MRI, they have been used only occasionally in the DCE-MRI analysis. To the authors' knowledge, it has been used only by few groups (18–21); (22); (23); (24). The priors in the mentioned papers are based on image gradients of perfusion maps except for (24) using a wavelet transform and (23) using the difference of the image from its denoised variant. The denoising is guided by the input DCE-MRI sequence. All three mentioned transformations generate so-called feature images on which a metric is computed to express, by one value, the spatial consistency of the voxels. The used metrics range from the smoothing  $\ell^2$  norm (18,20); (22); (23) to an approximation of the edge-preserving  $\ell^1$  norm (22); (24). The minimization of these spatial-regularization criterial functions is problematic because the voxels cannot be processed independently as in the case without the regularization.

Compared to (22,24) where an approximate  $\ell^1$  norm was used, in the present paper we derive a solution to the strict sparsity-inducing  $\ell^1$  norm regularizer operating on image gradients, i.e. the total variation regularization. This reduces smoothing of edges between different tissues. Since the strict  $\ell^1$  norm is not differentiable, we use proximal algorithms (25–27) to find the numerical solution. They are based on alternating minimization of the data and regularization terms and are the state of the art for solving image-processing problems such as image denoising, deconvolution or MRI reconstruction. However, the adaptation of these algorithms to DCE-MRI is not straightforward, because the data formation model in DCE-MRI is not linear. Motivated by good performance of the Levenberg-Marquardt method (LM) (28,29) in DCE-MRI perfusion parameter estimation (30,31), we decided to use the proximal Newton method (26) adapted to use the LM keeping good convergence and low computational costs. The derived algorithm iteratively performs one step of the LM method in each voxel followed by an image denoising step in all parameter maps.

All the above spatial regularization approaches are based on the Tofts (32) or extended Tofts (33) models. On the contrary, this paper proposes a spatial regularization for DCE-MRI based on the tissue homogeneity (TH) pharmacokinetic model (16,34). Use of the TH model provides estimates of additional perfusion parameters such as the plasma flow and the permeability surface area product, as opposed to the use of the Tofts models, but its use is not widespread because of its complexity (number of parameters, nonlinearity, a posteriori identifiability). In this paper we show that spatial regularization stabilizes the estimation procedure and makes use of such more complicated pharmacokinetic models feasible.

## 2 THEORY

### 2.1 PROBLEM DESCRIPTION

The goal of the DCE-MRI analysis is to estimate perfusion parameters from a sequence of  $N$  images capturing the distribution of the administered contrast agent bolus in time. The image intensity values are related to the actual concentration of the contrast agent in the voxel. The voxels from a region of interest are extracted and rearranged to form a matrix  $\mathbf{S}$  of size  $I \times N$  (number of selected voxels  $\times$  number of observations in time). The values in  $\mathbf{S}$  can be described by the model

$$S_{i,n} = \bar{S}_{i,n} + \epsilon_{i,n}, \quad i = 0, \dots, I - 1, n = 0, \dots, N - 1; \quad \epsilon_{i,n} \sim \text{Rice}(\bar{S}_{i,n}, \sigma_R), \quad [1]$$

where  $S_{i,n}$  is a single measured element of  $\mathbf{S}$  ( $i^{\text{th}}$  voxel,  $n^{\text{th}}$  time sample) and  $\bar{S}_{i,n}$  is the true value, not distorted by the noise  $\epsilon_{i,n}$ . The noise follows the signal-dependent Rice distribution (35), assuming Cartesian imaging for simplicity.

The intensity values  $\mathbf{S}$  are transformed to represent concentration of the contrast agent  $\mathbf{S}^\phi$ :

$$S_{i,n}^\phi = \phi_i(S_{i,n}) \approx \bar{S}_{i,n}^\phi + \epsilon_{i,n}^\phi, \forall i, n, \quad [2]$$

where  $\phi_i$  is a non-linear transformation function derived from the pre-contrast measurements (e.g. with several different flip angles) (1), thus the additive-noise model is not exact. The values  $\bar{S}_{i,n}^\phi$  can be modeled as samples of a nonlinear pharmacokinetic model  $c(n\Delta t, p_{i,\cdot})$ . It describes the contrast agent concentration within voxel  $i$  in time  $t = n\Delta t$ ,  $\Delta t$  is the sampling interval, and it is parametrized by the vector  $p_{i,\cdot}$  of length  $J$ , consisting of the perfusion parameters. A consequence of the transformation  $\phi_i$  is that the noise  $\epsilon_{i,n}^\phi$  now follows a complicated distribution. Thus, it is usually approximated by normal distribution, i.e.  $\epsilon_{i,n}^\phi \sim N(0, \sigma_i)$ ,  $\forall n$ ,  $\sigma_i$  is noise standard deviation in voxel  $i$ , which neglects variations in time and asymmetry (in case of low signal-to-noise ratios (SNR < 10 dB) (36,37) or for strongly nonlinear  $\phi_i$ ).

The goal of the DCE-MRI analysis is to estimate parameters  $p_{i,\cdot}$  of the pharmacokinetic model  $c$  by fitting it to the curve  $S_{i,\cdot}^\phi$  in each voxel  $i$ . Let  $\mathbf{p}$  denote the matrix gathering the perfusion parameters such that  $p_{i,\cdot}$  is the  $i^{th}$  row of  $\mathbf{p}$ . The maximum a posteriori probability (MAP) estimate  $\hat{\mathbf{p}}$  of the perfusion parameters  $\mathbf{p}$  of the size  $I \times J$  ( $I$  – number of voxels,  $J$  – number of perfusion parameters), assuming Gaussian noise, can be formulated as:

$$\hat{\mathbf{p}} = \arg \min_{\mathbf{p} \in P^I} \sum_{i=0}^{I-1} \sigma_i^{-2} \sum_{n=0}^{N-1} (S_{i,n}^\phi - c(n\Delta t, p_{i,\cdot}))^2 + \psi(\mathbf{p}) = \arg \min_{\mathbf{p} \in P^I} \sum_{i=0}^{I-1} f(p_{i,\cdot}) + \psi(\mathbf{p}) \quad [3]$$

where  $P$  is the feasible domain of the perfusion parameters. The regularization term  $\psi(\mathbf{p})$  describes the a priori knowledge about the problem and the first term, i.e. the data term, describes the fidelity of the fit. If no prior information is known or available in a suitable form,  $\psi$  is substituted by a zero function making the optimization problem to reduce to the standard nonlinear least squares problem, solvable by minimizing  $f(p_{i,\cdot})$  independently for each voxel (11,38).

## 2.2 PHARMACOKINETIC MODEL

The core of the functional [3] is a pharmacokinetic model  $c$  defined in general by a set of differential equations. This model can be transformed to the Laplace domain, where it has the form of multiplication of the Laplace spectra of two functions: the arterial input function (AIF) describing concentration of the contrast agent in voxel's arterial input and the impulse residue function (IRF) of the tissue voxel. If both functions have a closed-form expression in the time domain, the model can also be represented in the time domain, either as a convolution integral or even better in a closed form (15). If the closed forms are not available, as in our case of the TH model, the multiplication in the Laplace domain can be transformed to

the Fourier domain (16). In a discrete setting, the Laplace operator  $s$  is substituted by the sampling at discrete angular frequencies  $j\Delta\omega\mathbf{w}$ , where  $\Delta\omega$  is the sampling angular frequency,  $\mathbf{w}$  is a vector of indices; and the model is evaluated as:

$$c(\mathbf{n}\Delta t, p_{i,\cdot}) = \text{DFT}^{-1}\{C_a(j\Delta\omega\mathbf{w})H(j\Delta\omega\mathbf{w}, p_{i,\cdot})\} \quad [4]$$

where the multiplication of the Fourier spectra of the AIF and IRF, is element-wise and  $\text{DFT}^{-1}$  is the inverse discrete Fourier transform (DFT). The symbol  $\mathbf{n}$  represents a vector of time indices meaning that the DFT returns  $N$  values. The vector of frequency indices is  $\mathbf{w} = (0, 1, \dots, (N' - \text{mod}(N', 2))/2)$ , where  $\text{mod}$  is the modulo operation and  $N'$  is the number of time-domain samples that guarantee avoidance of the time-domain aliasing, see (16,39). The sampling angular frequency is then  $\Delta\omega = 2\pi / N'\Delta t$ . To use the defined representations of the functions, the  $\text{DFT}^{-1}$  includes complex conjugate symmetrization and final truncation to  $N$  samples. Additionally, we assume that the AIF is represented as a vector of time-domain samples  $c_a(\mathbf{n}\Delta t)$  which are transformed to the Fourier domain:  $C_a(j\Delta\omega\mathbf{w}) = \text{DFT}\{c_a(\mathbf{n}\Delta t)\}$  as in (16) with zero-padding to  $N'$  and the truncation to the half-spectrum. The IRF is represented by the TH model in the Laplace domain,  $H_{\text{TH}}(s, \mathbf{p})$ , with a modified parametrization compared to (16,40):  $\mathbf{p} = \{F_p, T_c, T_e, \alpha, \tau\}$  where we included  $\tau$  as the bolus arrival time (see Table 2 for descriptions). An additional modification is that we use a windowed version of the TH model,  $H_{\text{TH}}^w(s, \mathbf{p}, t_w)$ . This speeds up the evaluation keeping the number of samples low,  $N' = 3N$ , limited by the time-domain aliasing, see (39) for more details. The TH model is evaluated as:

$$H_{\text{TH}}(s, \mathbf{p}) = \frac{F_p(1 - e^{-(\alpha+T_e s)})(T_c + \alpha T_e + T_c T_e s)(\alpha + T_e s)}{\alpha(1 - e^{-(\alpha+T_e s)}) + s(T_c + \alpha T_e + T_c T_e s)(\alpha + T_e s)} \quad [5]$$

$$H_{\text{TH}}^w(s, \mathbf{p}, t_w) = h_{\text{TH}}(s, \mathbf{p}) - F_p E e^{-k_{\text{ep}}(t_w - \tau - T_c)} \frac{e^{-s t_w}}{s + k_{\text{ep}}}$$

where  $E = 1 - e^{-\alpha}$ ,  $k_{\text{ep}} = E/(T_e \alpha)$  and  $t_w = (N - 1)\Delta t$ .

### 2.3 REGULARIZATION FUNCTION

The stabilizing factor in [3] is the regularization function  $\psi(\mathbf{p})$ . Here, we describe the prior knowledge in the form of an image prior taking spatial relations of the voxels into account. The image prior in the form of a sparsifying  $\ell^1$  norm of a linearly transformed image is a widely used option in image and signal processing community (27,41,42). The transformation can be e.g. a wavelet transform or image gradients. It is in the form of a linear operator  $\mathbf{A}$  transforming the vector of parameters,  $p_{\cdot,j}$ , of the length  $I$  to a domain, where most of the values are close to zero. This so-called sparsifying transform  $\mathbf{A}$  can be represented by a set of  $V$  matrices  $\mathbf{A}_v, v \in \{1, \dots, V\}$  each corresponding to one sparse feature. We suppose the case of the image gradients, i.e. there are two matrices  $\mathbf{A} = (\nabla_r, \nabla_c)$  representing forward differences with respect to rows and columns and taking into account the spatial position of the voxels (43). Applying this operator, a

vector field is obtained. Before the application of  $\ell^1$  norm usually a magnitude of this vector field is computed (represented by  $|\cdot|$ ). This is repeated for  $J$  perfusion parameters leading to:

$$\psi(\mathbf{p}) = \sum_{j=0}^{J-1} \gamma_j \|\mathbf{A}p_{\cdot,j}\|_1 = \sum_{j=0}^{J-1} \gamma_j \sum_{i=0}^{I-1} \sqrt{\sum_v |\mathbf{A}_v p_{\cdot,j}|^2} = \sum_{j=0}^{J-1} \gamma_j \sum_{i=0}^{I-1} \sqrt{(\nabla_r p_{\cdot,j})^2 + (\nabla_c p_{\cdot,j})^2} \quad [6]$$

where  $\gamma_j$  is a regularization weight for particular parameter  $p_{\cdot,j}$ . The magnitude in the middle term reduces to elementwise absolute value and the square and the square root act elementwise. The last term represents the discrete isotropic total variation regularization used in this paper.

## 2.4 OPTIMIZATION METHODS

For the spatially regularized estimation of the perfusion maps, we use two proximal optimization methods, the proximal Newton method (26) and the primal-dual algorithm (27). Both methods are applicable to the class of problems of the form

$$\arg \min_{\mathbf{x}} d(\mathbf{x}) + r(\mathbf{A}\mathbf{x}) \quad [7]$$

where  $d, r$  are convex lower-semicontinuous, not necessarily differentiable functions. In our case,  $d$  and  $r$  represent the data and regularization terms, respectively.  $\mathbf{A}$  is an arbitrary linear operator possibly in matrix form.

### 2.4.1 PROXIMAL NEWTON METHOD

The proximal Newton method (26) is a variant of the Newton method (28,29) for non-differentiable functions applicable for functions in the form of [7] if  $d$  is twice differentiable. Similarly to the Newton method, it exploits the Hessian of  $d$ , which improves the convergence of the algorithm. Since  $r$  is typically not differentiable, its gradient is replaced by the so-called proximal operator, a useful tool proximal methods use to treat non-differentiable functions.

The proximal operator of an arbitrary convex function  $f(\mathbf{x})$  can be defined (44) as:

$$\text{prox}_{f}^{\mathbf{\Lambda}}(\mathbf{y}) = \arg \min_{\mathbf{x}} (f(\mathbf{x}) + (\mathbf{x} - \mathbf{y})^{\top} \mathbf{\Lambda}^{-1} (\mathbf{x} - \mathbf{y})) \quad [8]$$

where  $\mathbf{y}$  is a point at which the proximal operator is evaluated and  $\mathbf{\Lambda}$  is called the scaling matrix. Note that in majority of the literature,  $\mathbf{\Lambda}$  is the identity matrix. In the proximal Newton method,  $\mathbf{\Lambda}$  is necessary to compensate for the scaling implied by the Hessian in the Newton step.

The proximal Newton method iteratively performs two steps until convergence:

- 1) Do a Newton step in  $d$ , where  $\nabla d$  and  $\mathbf{H}$  are the gradient and Hessian of  $d$ :

$$\mathbf{y}^k = \mathbf{x}^k - \mathbf{H}^{-1} \nabla d(\mathbf{x}^k) \quad [9]$$

2) Evaluate the proximal operator of  $r(\mathbf{Ax})$  at the point  $\mathbf{y}^k$  scaled by  $\mathbf{H}^{-1}$ :

$$\mathbf{x}^{k+1} = \text{prox}_{r(\mathbf{A}\cdot)}^{\mathbf{H}^{-1}}(\mathbf{y}^k) = \arg \min_{\mathbf{x}} \left( r(\mathbf{Ax}) + \frac{1}{2}(\mathbf{x} - \mathbf{y}^k)^\top \mathbf{H}(\mathbf{x} - \mathbf{y}^k) \right) \quad [10]$$

Since this sub-problem 2) is solved in every iteration, it must be done efficiently. It resembles [7], but  $d$  is now much simpler – a quadratic function. This case can be again computed by a proximal method, e.g. by the nested primal-dual algorithm (27) described in the next section. Let us note that in case of a non-convex  $d$ , as in our case, the proximal Newton method converges to the closest minimum (45,46), if the minimization steps are not too large.

### 2.4.2 PRIMAL-DUAL ALGORITHM

Another proximal method for solving problems in the form [7] is the primal-dual algorithm (27), especially suitable for quadratic  $d$ . The algorithm is initialized by choosing constants influencing the convergence:  $\tau, \sigma > 0, \theta \in [0,1]$ , setting starting-point variables:  $\mathbf{u}^0 = \mathbf{y}^k, \mathbf{v}^0 = \mathbf{A}\mathbf{u}^0, \mathbf{x}^0 = \mathbf{u}^0$ , and proceeds by iterative updates until convergence (27):

$$\begin{aligned} \text{a) } & \mathbf{v}^{n+1} = \text{prox}_{r^*}^{\sigma}(\mathbf{v}^n + \sigma \mathbf{A}\mathbf{x}^n) \\ \text{b) } & \mathbf{u}^{n+1} = \text{prox}_d^{\tau}(\mathbf{u}^n - \tau \mathbf{A}^* \mathbf{v}^{n+1}) \\ \text{c) } & \mathbf{x}^{n+1} = \mathbf{u}^{n+1} + \theta(\mathbf{u}^{n+1} - \mathbf{u}^n) \\ \text{d) } & \text{Repeat a) – c) until convergence} \end{aligned} \quad [11]$$

Here,  $r^*$  is the convex conjugate (25) of  $r$  and  $\mathbf{A}^*$  is the adjoint operator to  $\mathbf{A}$ ; see [22] for details in case of the total variation.

## 2.5 ESTIMATION OF PERFUSION PARAMETERS

The final goal is to estimate perfusion parameter maps, i.e. to minimize [3] using total variation regularization [6]:

$$\hat{\mathbf{p}} = \arg \min_{\mathbf{p} \in \mathcal{P}^I} \sum_{i=0}^{I-1} f(p_{i,\cdot}) + \sum_{j=0}^{J-1} \gamma_j \| \mathbf{A} p_{\cdot,j} \|_1 \quad [12]$$

This problem is solved by the proximal Newton method (Sec. 2.4.1). Unfortunately, the computation of the Hessian matrices  $\mathbf{H}_i$  of the data term  $f(p_{i,\cdot})$  for each curve  $i$  is computationally demanding and their inversions are unstable. For this reason we replace them by the Levenberg-Marquardt modification (29):  $\mathbf{H}_i = 1/\lambda_i^k \mathbf{I} + 2\mathbf{J}_i^\top \sigma_i^{-2} \mathbf{J}_i$ , where  $\lambda_i^k$  is related to the step length in the iteration  $k$  and  $\mathbf{J}_i$  is the Jacobian matrix of  $c(\mathbf{n}T_s, \hat{p}_{i,\cdot}^k)$  of the size  $N \times J$ . Then, the proximal Newton method has the form:

1) Do a Newton (Levenberg-Marquardt) step [9] in terms of  $f(\hat{p}_{i,\cdot}^k)$  for each voxel:

$$y_{i,\cdot}^k = \hat{p}_{i,\cdot}^k - \mathbf{H}_i^{-1} \nabla f(\hat{p}_{i,\cdot}^k) \cong \hat{p}_{i,\cdot}^k - \lambda_i^k (\mathbf{I} + \lambda_i^k 2\mathbf{J}_i^\top \sigma_i^{-2} \mathbf{J}_i)^{-1} \nabla f(\hat{p}_{i,\cdot}^k), \forall i, \quad [13]$$

where  $\lambda_i^k$  is estimated by the LM scheme according to (29).

- 2) Evaluate the scaled proximal operator [10] of the regularization function in [12]:

$$\hat{\mathbf{p}}^{k+1} = \arg \min_{\mathbf{p} \in \mathcal{P}^I} \left( \sum_{j=0}^{J-1} \gamma_j \|\mathbf{A} \mathbf{p}_{\cdot,j}\|_1 + \frac{1}{2} \sum_{i=0}^{I-1} (p_{i,\cdot} - y_{i,\cdot}^k) \mathbf{H}_i (p_{i,\cdot} - y_{i,\cdot}^k)^\top \right), \text{ where} \quad [14]$$

$$\mathbf{H}_i = (\lambda_i^k)^{-1} (\mathbf{I} + \lambda_i^k 2\mathbf{J}_i^\top \sigma_i^{-2} \mathbf{J}_i), \forall i.$$

The resulting functional [14] is minimized using the primal-dual algorithm [11]. After initialization of the starting point  $\mathbf{u}^0 = \mathbf{y}^k$ ,  $v_{\cdot,j}^0 = \gamma_j \mathbf{A} \mathbf{u}_{\cdot,j}^0 \forall j$ ,  $\mathbf{x}^0 = \mathbf{u}^0$  and setting the constants  $\tau, \theta$ ;  $\sigma \leq \frac{1}{8\tau}$  (valid for the operator of forward differences [6]), the algorithm iterates through the following steps:

- a) Evaluate the proximal operator [11] of the convex conjugate of the regularization function in the form of  $r(\mathbf{x}) = \|\mathbf{A}\mathbf{x}\|_1$ . This is well known in image reconstruction problems and the required proximal operator of its convex conjugate gets the form of a projection onto unit ball (27):

$$v_{\cdot,j}^{n+1} = \frac{v_{\cdot,j}^n + \sigma \gamma_j \mathbf{A} \mathbf{x}_{\cdot,j}^n}{\max(1, |v_{\cdot,j}^n + \sigma \gamma_j \mathbf{A} \mathbf{x}_{\cdot,j}^n|)}, \forall j \quad [15]$$

i.e. the vector field is in each point divided by either its magnitude or by unity.

- b) Evaluate the proximal operator of a quadratic function [11]:

$$\mathbf{u}^{n+1} = \arg \min_{\mathbf{x}} \left( \frac{1}{2} \sum_{i=0}^{I-1} (x_{i,\cdot} - y_{i,\cdot}^k) \mathbf{H}_i (x_{i,\cdot} - y_{i,\cdot}^k)^\top + \sum_{i=0}^{I-1} \frac{1}{2\tau} (x_{i,\cdot} - u_{i,\cdot}^n + \tau(\gamma_j \mathbf{A}^* v_{\cdot,j}^{n+1})_{i,\cdot}) (x_{i,\cdot} - u_{i,\cdot}^n + \tau(\gamma_j \mathbf{A}^* v_{\cdot,j}^{n+1})_{i,\cdot})^\top \right), \forall j \quad [16]$$

The optimizer of this quadratic form was derived as:

$$u_{i,\cdot}^{n+1} = (\tau \mathbf{H}_i + \mathbf{I})^{-1} \left( \tau \mathbf{H}_i y_{i,\cdot}^k + u_{i,\cdot}^n - \tau(\gamma_j \mathbf{A}^* v_{\cdot,j}^{n+1})_{i,\cdot} \right), \forall j, i,$$

where  $(\cdot)_{i,\cdot}$  stands for the row  $i$  of the result of the parenthesized operation and  $\mathbf{A}^*$  is the adjoint operator [22].

- c) Update estimate [11]:

$$\mathbf{x}^{n+1} = \mathbf{u}^{n+1} + \theta(\mathbf{u}^{n+1} - \mathbf{u}^n) \quad [17]$$

- d) Repeat steps a), b), c) until convergence and then set:

$$\hat{\mathbf{p}}^{k+1} = \mathbf{x}^{n+1} \quad [18]$$

- 3) Repeat main steps 1), 2) until convergence.

## 3 METHODS

### 3.1 EXPERIMENTAL DATA

For the comparisons with the ground truth, we have designed a numerical DCE-MRI rat phantom. The phantom was based on a DCE-MRI examination of a real rat with implanted glioblastoma (47) (next paragraph). A high-resolution image ( $1024 \times 1024$  pixels) of an axial slice of head including the tumor was created based on manual segmentation of a real DCE-MRI image sequence into 41 homogeneous regions (Figure 1). Each region was described by a set of perfusion parameters based on literature and on perfusion analysis results from (47). For each set of parameters, a concentration-time curve was generated using the TH model and a pre-defined AIF (48) in high temporal resolution (sampling period equal to TR). The curves were converted to signal-intensity curves based on the acquisition model of the FLASH acquisition with no  $T_2^*$  effect and assuming a constant contrast-agent relaxivity  $r_1$  and spatially invariant native relaxation time  $T_{10}$  and proton density. The signal intensity curves were used to construct a high-spatial-resolution image at each TR. These images, multiplied by coil sensitivities (estimated from real measurements), were then Fourier transformed to the k-space and echoes extracted as k-space lines corresponding to the acquisition scheme (next paragraph). Complex uncorrelated zero-mean Gaussian noise was added to obtain SNR according to real conditions (Figure 1). In addition to the dynamic DCE-MRI scans, multi-flip-angle pre-contrast scans were simulated. The simulated dynamic sequence was converted to the contrast-agent concentration using the pre-contrast images according to (49).

A real DCE-MRI recording of a glioblastoma-bearing rat was used as a testing preclinical dataset (detailed description in (47)), acquired on a 7 T horizontal PharmaScan (Bruker Biospin, Germany) with a four-channel rat head surface coil using the FLASH acquisition: one 1mm slice, TR/TE 8/2.1 ms, FA  $17^\circ$ , acquisition matrix  $128 \times 128$ , temporal resolution 0.768 s, total scan time  $\sim 13$  min. The contrast agent (Omniscan - GE Healthcare, Norway) of 0.1 mmol/kg was injected intravenously after 25 s of recording. The pre-contrast scans were acquired using the same parameters except for the FA ( $5^\circ$ ,  $10^\circ$ ,  $15^\circ$ ,  $20^\circ$ ,  $25^\circ$ ,  $30^\circ$ ). These scans were used to convert the dynamic sequence to the contrast-agent concentration images (49). The AIF was derived using multichannel blind deconvolution (50) as stated in the original paper (47).

A testing clinical dataset of a renal-cell-carcinoma-metastasis patient (details in (51)) was acquired on a Magnetom Avanto 1.5 T MRI scanner (Siemens AG, Munich, Germany) using the T1-weighted 2D saturation-recovery prepared Turbo FLASH (nonselective SR pulse): TR/TE/TI 400/1.09/200 ms, FA  $16^\circ$ , acquisition matrix  $128 \times 128$ , temporal resolution 1.2 s, three coronal slices, total scan time 10 min. The contrast agent bolus of 7.5 ml (Gadovist – Bayer Schering Pharma, Berlin, Germany) was injected into antecubital vein. The pre-contrast scans acquisition preceded with the same parameters except TI (500,

1000, 3000 ms, five frames per each) to perform conversion to contrast-agent concentration as in (52). The AIF was estimated using multi-channel blind deconvolution according to (51).

### 3.2 INITIAL SETUP

To keep the regularization weights unchanged across measurements and AIFs with different energies, the measured perfusion curves and the AIF were normalized by scalar constants  $\alpha_{\text{TRF}}$ ,  $\alpha_{\text{AIF}}$  to always obtain similar ratio of the data and the regularization terms values. Additionally, we have separated relative weights  $\gamma'_j$  and the global weight  $\Gamma$  such that  $\gamma_j = \Gamma\gamma'_j$ . This extended the functional [12] to:

$$\hat{\mathbf{p}}' = \arg \min_{\mathbf{p}' \in P'} \sum_{i=0}^{I-1} \sigma_i^{-2} \sum_{n=0}^{N-1} \left( \alpha_{\text{TRF}}^{-1} S_{i,n}^\phi - \alpha_{\text{AIF}}^{-1} c(n\Delta t, p'_{i,\cdot}) \right)^2 + \Gamma N \sum_{j=0}^{J-1} \gamma'_j \|\mathbf{A}p'_{\cdot,j}\| \quad [19]$$

where  $\alpha_{\text{TRF}}$ ,  $\alpha_{\text{AIF}}$  are estimated from the measured data and AIF using:

$$\alpha_{\text{TRF}} = \text{quantile}_{0.75}(\text{median}(S_{i,\cdot}^\phi), i = 0, 1, \dots, I-1), \alpha_{\text{AIF}} = \text{median}(c_a(\mathbf{n}\Delta t)) \quad [20]$$

The estimates are finally computed as:  $\hat{\mathbf{p}} = \hat{\mathbf{p}}'$  except for  $\hat{p}_{\cdot,1} = \alpha_{\text{TRF}} \alpha_{\text{AIF}}^{-1} \hat{p}'_{\cdot,1}$ , since only the perfusion map  $F_p$  is influenced by the energy of the inputs. It is worth noting that the number of samples  $N$  was included in [19] to eliminate possibly different number of time samples in the measured data  $\mathbf{S}^\phi$ .

The relative weights were estimated based on the numerical phantom using the maximum likelihood approach (53), i.e.  $\gamma_j'' = I / \|\mathbf{A}p'_{\cdot,j}\|$ ,  $\forall j$ , transformed to  $\gamma'_j = \gamma_j'' / \sum \gamma''$ ,  $\forall j$ . The minimization was run only from one starting point based on the authors' experience and results in (16) using the TH model. The relative weights, the starting point as well as the constraints of the parameters are defined in Table 2. The estimates were projected onto the constraints  $P$  before any evaluation of the pharmacokinetic model throughout the iterative procedure (45). This was done to ensure the respective physiological ranges and numerical stability. The stopping criterion in the main loop of the algorithm (proximal Levenberg-Marquardt) was set experimentally to 50 iterations and each sub-problem (primal-dual) was stopped after 200 iterations. The estimates of noise standard deviations ( $\sigma_i, \forall i$ ) were computed using the median of the absolute deviation estimator with Daubechies wavelet (54). The source code of the algorithm is available at: [https://github.com/Bartolomejka/DCE-MRI\\_Regularization\\_MRM](https://github.com/Bartolomejka/DCE-MRI_Regularization_MRM).

## 4 RESULTS

### 4.1 STRENGTH OF REGULARIZATION

The goal of this numerical-phantom experiment was to analyze the influence of the global regularization weight  $\Gamma$  on the perfusion parameter estimates. The proposed algorithm was run for 16 values of  $\Gamma$  spread

logarithmically from  $10^{-3}$  to  $10^2$ . The results are shown in Figure 2 and Supporting Information Figures S1, S2 as perfusion-parameter maps and in Figure 3 as the mean absolute error (MAE) for each perfusion parameter and as the mean,  $\mu(\cdot)$ , of the reduced chi-squared statistic,  $\chi_{\text{red}}^2$ , computed per voxel, which is related to the data-term value in [12]. The metrics were defined as:

$$\begin{aligned} MAE_j &= \frac{1}{I} \sum_{i=0}^{I-1} |\hat{p}_{i,j} - p_{i,j}^*| \\ \chi_{\text{red}}^2 &= \frac{1}{(N-J)} (\sigma_i^*)^{-2} \sum_{n=0}^{N-1} \left( S_{i,n}^\phi - c(n\Delta t, \hat{p}_{i,\cdot}) \right)^2 \end{aligned} \quad [21]$$

where  $p_{i,j}^*, \forall j$  is the ground truth. The noise standard deviation  $\sigma_i^*, \forall i$  was estimated using the ground truth. Additional metrics showing the bias and precision are plotted in Supporting Information Figure S3. The areas where the model is invalid or unstable, i.e. bones (no signal), arteries (no extravascular space), brain tissue (no contrast-agent extravasation), and areas where  $\text{SNR} < 5$  dB (areas far from surface coils), were excluded from the analysis.

When  $\Gamma$  increased, the MAE values (Figure 3) decreased until their minimum, after which they started increasing exhibiting an over-regularization effect. The minima are achieved for  $\Gamma = 0.22$ , except for  $F_p, v_p$ , where the optimum is  $\Gamma = 4.64$ . The value  $\Gamma = 0.22$  also corresponded to the best perfusion-parameter maps visually (Figure 2). The optimal value  $\Gamma = 0.22$  was also consistent with the visual analysis of the bias and precision (Supporting Information Figure S3) for most perfusion parameters. Supporting Information Figure S3 shows that increasing  $\Gamma$  could further improve precision but at the cost of increased bias. MAE shapes similar to ours were reported also in (22); however, in (22) the shape of the curve equivalent to our  $\mu(\chi_{\text{red}}^2)$  (Figure 3) was similar to the MAE curves. This was not observed here. A probable explanation is that our pharmacokinetic model and minimizer are more robust in terms of local optima. Interestingly,  $\mu(\chi_{\text{red}}^2) \approx 1$  for any  $\Gamma$  meaning that the estimated concentration-time curves fitted the data well even for high regularization (cf. Supporting Information Figure S1 showing  $\chi_{\text{red}}^2$  per voxel).

## 4.2 BIAS AND PRECISION OF THE ESTIMATOR

The proposed method (with fixed  $\Gamma = 0.22$ ) was quantitatively compared to the standard non-regularized method (Table 3 – tumor tissues, Supporting Information Table S1 – all tissues). The metrics used were the mean and the standard deviation of perfusion-parameter estimates within each simulated tissue region, representing the bias and precision of the estimator, respectively. They were estimated from a single numerical rat phantom dataset (i.e. one realization of noise). The proposed method performed the best for large tissue areas with low SNR as expected. The non-regularized version led to slightly less biased or more precise estimates only in some small tissue areas or regions with high SNR.

We additionally tested the performance of the non-regularized and the proposed methods with the optimal setting (Section 4.1) on the numerical phantom for 50 noise realizations. To measure the bias and the precision of the estimators, the mean and the standard deviation of the perfusion-parameter estimates were computed for each voxel. Supporting Information Figure S4 (Mean) shows a distinct systematic difference between the methods in the low SNR areas, i.e. a clearly lower bias of the regularized estimation. In addition, the standard deviation of the proposed estimator is lower (i.e. higher precision), leading to an improved readability of perfusion maps (cf. Figure 2 with Supporting Information Figure S4 – standard deviation, parameter  $F_p$  in tumor outer region).

To compare the two estimators quantitatively, we have again estimated their bias and precision per tissue (Supporting Information Table S2) by averaging their mean and standard deviation in Supporting Information Figure S4 within each simulated tissue region. In agreement with the evaluation based on a single noise realization, the proposed method enjoyed consistently better precision and in majority of the cases also a lower bias (Supporting Information Table S2).

### 4.3 COMPUTATIONAL DEMANDS

To assess the time requirements of the proposed method, additional data from the experiment in Section 4.1 were analyzed. It included the number of evaluations of the pharmacokinetic model  $c(nT_s, p'_i)$  for each voxel  $i$  and the total duration of the minimization procedure measured on a 6-core Intel(R) Core(TM) i7-8700K CPU @ 3.70GHz, implementation in Matlab without parallelization. The same was analyzed for the standard method without regularization, i.e. the denoising step [14] was excluded. Additionally, the standard non-regularized method was accelerated by including a stopping criterion of a minimal change in its data term.

The proposed regularized perfusion analysis with  $\Gamma = 0.22$  needed 36 min to finish in comparison with 20 min in the non-regularized case. The data consisted of 9916 curves, with 1000 time points each. The evaluation of the LM step [13] took 80 % and the denoising step [14] took 15 % of the total time in the proposed algorithm.

As the calculation of the pharmacokinetic model [4] was the most time-demanding operation, the distribution of the number of model evaluations in each pixel and their total count were analyzed (Figure 4). Theoretically, the number of model evaluations in the regularized method is at least twice the number of iterations times the number of the concentration-time curves, since recalculation of the gradients after each denoising step is needed. Contrarily, the non-regularized method can perform almost an arbitrary number of model evaluations, since the number of iterations is not fixed for particular curves. However, the regularized method does not need twice the non-regularized method model evaluations (black points in

Figure 4). This means that the recomputations of the model and its derivatives inside one iteration to estimate the LM step length are reduced in case of the regularized method. This stabilizing effect of the regularization is also visible in the distribution of the model evaluations in pixels (boxplots in Figure 4). As the regularization grows, the numbers of model evaluations in each pixel reduce, i.e. the pixels with problematic convergence are now converging better.

#### 4.4 REAL DATASETS

The results of the experiments on real preclinical and clinical data are shown in Figure 5 and Figure 6, respectively; data were acquired under the approval of an Institutional Review and Ethics Boards. In both cases, the number of outlier perfusion-parameter estimates was clearly reduced and the spatial correspondence of the maps to the underlying anatomy was substantially improved.

## 5 DISCUSSION

The goal of this work was to improve the accuracy and precision of perfusion-parameter estimates in DCE-MRI. In the standard voxel-by-voxel approach, errors in the estimates are mainly caused by the presence of local minima of the curve-fitting problem and appear mainly in low SNR conditions starting at approximately 13 dB (cf. Figure 1 – SNR, Figure 2 – Non-regularized). However, it can differ based on the tissue type. Erroneous estimates reduce the readability of the perfusion maps and their usability in practice. The estimation error in perfusion-parameter estimates is not only random but the estimates are also biased, as was demonstrated in Supporting Information Figure S4 (Mean, Non-regularized).

Using spatial regularization, we were able to stabilize the estimates. The introduction of additional prior information led to a distinct improvement in the parameter estimates (Figure 2 or Supporting Information Figures S1, S2 for  $\Gamma = 0.22$ ). The improvements were of two types: our method has improved the precision (i.e. reduced the variance of the estimates in homogeneous regions) improving the readability of the parameter maps (e.g. Figure 2 –  $T_c$  map of the tumor) and it has reduced the bias of the estimates in areas with low SNR (see e.g. Figure 2 – bottom of the maps for  $F_p$ ,  $E$ ,  $T_c$ ). These observations were supported by the quantitative evaluation of the bias and precision per tissue (Table 3 and Supporting Information Tables S1, S2).

In comparison with previous publications (18–24); our work was based on the TH model as a natural successor of the previously used simpler models. Our study showed that the increased complexity of the TH model can still be handled using spatial prior. In addition, as opposed to the relaxed  $\ell^1$  norm or  $\ell^2$  norm used by other groups, our prior is the total variation using exact  $\ell^1$ -norm. This leads to a non-differentiable

criterion functional of the estimator. As was derived in the paper, application of the state-of-the art minimization techniques, i.e. proximal algorithms, solved this fundamental drawback efficiently.

In contrast to the previous spatially regularized DCE-MRI approaches, our spatial-regularization model [12] adjusts the strength of regularization automatically to the spatially varying noise level. Hence, voxels with high noise level are regularized more than those with low noise level. To utilize this property, the standard deviations of noise in each voxel were estimated using an independent estimator (54) prior to the minimization. Alternatively, estimates from non-regularized curve fits similar to (22) can be used, which is more precise but time-consuming. A more challenging approach would be estimation of noise standard deviations for each time instant or to formulate the problem using an exact noise distribution model. We assume a zero-mean Gaussian noise with spatial-variant standard deviation, which is not valid for low-SNR areas.

The presented method is an experimental method pushing the DCE-MRI limits utilizing the modern image processing techniques. The target application is pre-clinical DCE-MRI data processing, where strongly inhomogeneous coil sensitivities are causing dramatic reduction of SNR in areas distant from the coils. Similarly, the clinical datasets may suffer from insufficient SNR, e.g. in case of 3D imaging. The target application of our minimizer is defined by the model used in the core of the algorithm, i.e. the pharmacokinetic model. Its replacement by another pharmacokinetic model or a completely different model with a similar structure can solve non-linear minimization problems with regularization in other applications, e.g. arterial spin labeling (ASL) or diffusion tensor imaging (DTI).

Another aspect that deserves discussion is the used regularization function and its relation to the properties of the numerical rat phantom. We have used the total variation, which is a simple regularization function that favors piecewise-constant functions. In this way, we incorporated the assumption that tissues contain a small number of homogeneous regions with constant perfusion properties. To model real semi-homogeneous tissues better, we could choose one of more complex alternatives (by replacing the image gradients of the total variation ( $\mathbf{A}$  in [6]). These are for example the wavelet transform as in (24) favoring piecewise-polynomial functions or the total generalized variation (55) favoring piecewise linear functions. However, in our experience, the real-life difference between them is not large. What is important, is to apply any regularization with a preference of piecewise-smooth functions, even a simple one.

Our numerical rat phantom contains several dozens of regions, where all the voxels have the same perfusion parameters. We are aware of the fact that this simplistic choice favors in a sense the used regularization function. We cannot solve this problem by using another regularization function, because all common functions favor piecewise constant functions. On the other hand, we are not aware of any good model to

describe the change in perfusion parameters within the same tissue. Any ad-hoc noise introduced into the parameters could influence the results in an unpredictable way. These considerations lead us to the unideal but plausible choice of the total variation regularization and the simple numerical rat phantom with constant regions.

In general, to set up the regularization strength is challenging even for the standard problem when regularizing a single image. In the DCE-MRI, there are five perfusion-parameter maps to regularize simultaneously, i.e. five regularization weights have to be set. The weights are not independent, since each of them affects the others via the pharmacokinetic model. If, for example, one perfusion-parameter map is forced to be too smooth (by setting its weight high), the estimates of the remaining perfusion parameters compensate for this effect to preserve a good fit in the data term. Thus their maps become uneven (data not shown but similar behavior is visible in Figure 2 or Supporting Information Figure S1, right column, parameters  $F_p$  vs.  $E, T_c$ ). Therefore, the weights must not take only the values of the perfusion parameters into account but also their interconnection through the data-term fitting.

We have approached the problem of setting the relative regularization weights by the maximum likelihood estimation on a realistic numerical phantom. The optimal global weight  $\Gamma$  was chosen based on knowledge of the ground-truth and the estimated perfusion parameters. We have applied these weights to the real dataset and obtained a substantial improvement of the spatial consistency between the perfusion-parameter maps and the anatomical images when compared to the non-regularized version. However, the accuracy improvement gained from the proposed spatial regularization is difficult to quantify for the real data where no ground-truth is available. Since the proposed algorithm utilizes data normalization and an estimator of the noise variance based on the input data, it should not be necessary to change the global regularization weight  $\Gamma$ , nor the relative regularization weights  $\gamma$  for new datasets. However, this beneficial property is difficult to be thoroughly tested and thus a slight modification of  $\Gamma$  can improve the estimates in new scenarios.

A related question is, which perfusion parameters should be spatially regularized. In this paper, the same parameters as the parameters of the pharmacokinetic model are regularized:  $T_c, T_e, \alpha, \tau$  (parameters independent of AIF and concentration-time curves scaling) and  $F_p$  (dependent on scaling of the curves). This choice of the regularized parameters was also motivated by the fact that  $F_p, T_c, \alpha$  are hard to estimate in comparison with the derived perfusion parameters as e.g.  $v_p, v_e, K^{\text{trans}}$ . Despite the fact that the derived parameters were not regularized, their estimates were stable and robust to over-regularization contrary to the regularized parameters; see perfusion-parameter maps of  $v_p = F_p T_c$  and  $K^{\text{trans}} = F_p E$  in Supporting Information Figures S1, S2.

Although we have justified the set of parameters for regularization, it is not clear whether this set is equal to an optimal set of parameters to parametrize the model in the minimization, which influences convergence. In the non-regularized case, the optimal parametrization was studied in (56) for the extended Tofts model, but in our work, the parameterization was not studied and the selection of the parameters was driven by the regularization and practical reasons.

## 6 CONCLUSION

Incorporation of spatial prior information in terms of total variation helps to improve the estimates of perfusion parameters. This was clearly shown on realistically simulated data. Perfusion maps estimated from pre-clinical and clinical data showed a substantially better consistency with anatomical images than in case of the traditional estimation with no spatial prior.

Our implementation of the spatial prior incorporates an additional image denoising step applied to the perfusion maps after each iteration of the voxel-wise Levenberg-Marquardt algorithm. The time demands of the denoising step are negligible in comparison with the evaluation of the pharmacokinetic model and its gradients keeping the presented algorithm tractable. However, the algorithm needs more model evaluations than the non-regularized algorithm, because of the additional model evaluations after each denoising step and a missing smart stopping-criterion.

## ACKNOWLEDGMENT

This study was supported by the Czech Science Foundation (GA16-13830S), Ministry of Education, Youth and Sports of the Czech Republic (CZ.1.05/2.1.00/01.0017, LO1212, LO1401, CZ.02.1.01/0.0/0.0/16\_013/0001775) and the Czech Academy of Sciences (MSM100751802). Infrastructure of the SIX Center was used. We would like to thank Michal Standara for providing the clinical data.

## A. APPENDIX

### A.1 ADJOINT OPERATOR OF THE IMAGE GRADIENT

The adjoint operator  $\mathbf{A}^*$ , required in the primal-dual algorithm, has for the image gradient the form of a negative discrete divergence (27):

$$\mathbf{A}^* \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \nabla_r^T X_1 + \nabla_c^T X_2 = -\operatorname{div} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad [22]$$

where  $\mathbf{X}_1$ ,  $\mathbf{X}_2$  are respective components of the vector field with size equivalent to the result of the gradient operator in [6]. The divergence takes the position of the voxels in the image into account (43).

## REFERENCES

1. Schabel MC, Parker DL. Uncertainty and bias in contrast concentration measurements using spoiled gradient echo pulse sequences. *Phys. Med. Biol.* [Internet] 2008;53:2345–73. doi: 10.1088/0031-9155/53/9/010.
2. Koh TS. On the a priori identifiability of the two-compartment distributed parameter model from residual tracer data acquired by dynamic contrast-enhanced imaging. *IEEE Trans. Biomed. Eng.* [Internet] 2008;55:340–4. doi: 10.1109/TBME.2007.910682.
3. Brix G, Zwick S, Kiessling F, Griebel J. Pharmacokinetic analysis of tissue microcirculation using nested models: Multimodel inference and parameter identifiability. *Med. Phys.* [Internet] 2009;36:2923–2933. doi: 10.1118/1.3147145.
4. Brix G, Salehi Ravesh M, Griebel J. Two-compartment modeling of tissue microcirculation revisited. *Med. Phys.* [Internet] 2017;44:1809–1822. doi: 10.1002/mp.12196.
5. Lopata RGP, Backes WH, van den Bosch PPJ, van Riel NAW. On the identifiability of pharmacokinetic parameters in dynamic contrast-enhanced imaging. *Magn. Reson. Med.* [Internet] 2007;58:425–429. doi: 10.1002/mrm.21336.
6. Henderson E, Rutt BK, Lee T-Y. Temporal sampling requirements for the tracer kinetics modeling of breast disease. *Magn. Reson. Imaging* [Internet] 1998;16:1057–1073. doi: 10.1016/S0730-725X(98)00130-1.
7. Aerts HJWL, Jaspers K, Backes WH. The precision of pharmacokinetic parameters in dynamic contrast-enhanced magnetic resonance imaging: the effect of sampling frequency and duration. *Phys. Med. Biol.* [Internet] 2011;56:5665–78. doi: 10.1088/0031-9155/56/17/013.
8. Larsson C, Kleppestø M, Rasmussen I, Salo R, Vardal J, Brandal P, Bjørnerud A. Sampling requirements in DCE-MRI based analysis of high grade gliomas: simulations and clinical results. *J. Magn. Reson. Imaging* [Internet] 2013;37:818–29. doi: 10.1002/jmri.23866.
9. Evelhoch JL. Key Factors in the Acquisition of Contrast Kinetic Data for Oncology. *J. Magn. Reson. Imaging* [Internet] 1999;259:254–259. doi: 10.1002/(SICI)1522-2586(199909)10:3<254::AID-JMRI5>3.0.CO;2-9.

10. Kershaw LE, Cheng H-LM. Temporal resolution and SNR requirements for accurate DCE-MRI data analysis using the AATH model. *Magn. Reson. Med.* [Internet] 2010;64:1772–80. doi: 10.1002/mrm.22573.
11. Bartoš M, Jiřík R, Kratochvíla J, Standara M, Starčuk Z, Taxt T. The Precision of DCE-MRI Using the Tissue Homogeneity Model with Continuous Formulation of the Perfusion Parameters. *Magn. Reson. Imaging* [Internet] 2014. doi: 10.1016/j.mri.2014.02.003.
12. Garpebring A, Brynolfsson P, Yu J, Wirestam R, Johansson A, Asklund T, Karlsson M. Uncertainty estimation in dynamic contrast-enhanced MRI. *Magn. Reson. Med.* [Internet] 2013;69:992–1002. doi: 10.1002/mrm.24328.
13. Dale BM, Jesberger JA, Lewin JS, Hillenbrand CM, Duerk JL. Determining and optimizing the precision of quantitative measurements of perfusion from dynamic contrast enhanced MRI. *J. Magn. Reson. Imaging* [Internet] 2003;18:575–84. doi: 10.1002/jmri.10399.
14. De Naeyer D, De Deene Y, Ceelen WP, Segers P, Verdonck P. Precision analysis of kinetic modelling estimates in dynamic contrast enhanced MRI. *MAGMA* [Internet] 2011;24:51–66. doi: 10.1007/s10334-010-0235-6.
15. Koh TS, Cheong DLH, Hou Z. Issues of discontinuity in the impulse residue function for deconvolution analysis of dynamic contrast-enhanced MRI data. *Magn. Reson. Med.* [Internet] 2011;66:886–92. doi: 10.1002/mrm.22868.
16. Garpebring A, Östlund N, Karlsson M. A Novel Estimation Method for Physiological Parameters in Dynamic Contrast-Enhanced MRI: Application of a Distributed Parameter Model Using Fourier-Domain Calculations. 2009;28:1375–1383. doi: 10.1109/TMI.2009.2016212.
17. Bartoš M, Keunen O, Jiřík R, Bjerkvig R, Taxt T. Perfusion Analysis of Dynamic Contrast Enhanced Magnetic Resonance Images Using a Fully Continuous Tissue Homogeneity Model with Mean Transit Time Dispersion and Frequency Domain Estimation of the Signal Delay. In: *Analysis of Biomedical Signals and Images*. Vol. 20. Brno, Czech Republic: Brno University of Technology; 2010. pp. 269–274.
18. Sommer JC, Gertheiss J, Schmid VJ. Spatially regularized estimation for the analysis of dynamic contrast-enhanced magnetic resonance imaging data. *Stat. Med.* [Internet] 2014;33:1029–1041. doi: 10.1002/sim.5997.
19. Sommer JC, Schmid VJ. Spatial two-tissue compartment model for dynamic contrast-enhanced magnetic resonance imaging. *J. R. Stat. Soc. Ser. C (Applied Stat.* [Internet] 2014;63:695–713. doi: 10.1111/rssc.12057.

20. Schmid VJ, Whitcher B, Padhani AR, Taylor NJ, Yang G-Z. Bayesian methods for pharmacokinetic models in dynamic contrast-enhanced magnetic resonance imaging. *IEEE Trans. Med. Imaging* [Internet] 2006;25:1627–36. doi: 10.1109/TMI.2006.884210.
21. Feilke M, Bischl B, Schmid VJ, Gertheiss J. Boosting in nonlinear regression models with an application to DCE-MRI data. *Methods Inf. Med.* 2016;55:31–41. doi: 10.3414/ME14-01-0131.
22. Kelm BM, Menze BH, Nix O, Zechmann CM, Hamprecht FA. Estimating Kinetic Parameter Maps From Dynamic Contrast-Enhanced MRI Using Spatial Prior Knowledge. *IEEE Trans. Med. Imaging* [Internet] 2009;28:1534–1547. doi: 10.1109/TMI.2009.2019957.
23. Yang W, Yu L, Lu Z, Yun Z, Feng Y, Feng Q, Chen W. Estimating pharmacokinetic parameter maps from breast DCE-MRI with implicit regularization by guided image filtering. In: 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI). IEEE; 2014. pp. 878–881. doi: 10.1109/ISBI.2014.6868011.
24. Guo Y, Lingala SG, Zhu Y, Lebel RM, Nayak KS. Direct estimation of tracer-kinetic parameter maps from highly undersampled brain dynamic contrast enhanced MRI. *Magn. Reson. Med.* [Internet] 2017;78:1566–1578. doi: 10.1002/mrm.26540.
25. Parikh N, Boyd S. Proximal Algorithms. *Found. Trends Optim.* 2013;1:123–231. doi: 10.1561/24000000003.
26. Lee JD, Sun Y, Saunders MA. Proximal Newton-Type Methods for Minimizing Composite Functions. *SIAM J. Optim.* [Internet] 2014;24:1420–1443. doi: 10.1137/130921428.
27. Chambolle A, Pock T. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis.* 2011;40:120–145. doi: 10.1007/s10851-010-0251-1.
28. Seber GAF, Wild CJ. *Nonlinear Regression (Wiley Series in Probability and Statistics)*. Wiley-Interscience; 2003.
29. Madsen K, Nielsen HB, Tingleff O. *Methods for Non-Linear Least Squares Problems (2nd ed.)*. Richard Petersens Plads, Building 321, {DK-}2800 Kgs. Lyngby: Informatics and Mathematical Modelling, Technical University of Denmark, DTU; 2004.
30. Ahearn TS, Staff RTT, Redpath TWW, Semple SIKIK. The use of the Levenberg–Marquardt curve-fitting algorithm in pharmacokinetic modelling of DCE-MRI data. *Phys. Med. Biol.* [Internet] 2005;50:N85–N92. doi: 10.1088/0031-9155/50/9/N02.

31. Feilke M, Bischl B, Schmid VJ, Gertheiss J. Web-based Appendix for Boosting in nonlinear regression models with an application to DCE-MRI data: Details on the optimization in the boosting algorithm. *Methods Inf. Med.* 2016;55:31–41.
32. Tofts PS, Brix G, Buckley DL, et al. Estimating kinetic parameters from dynamic contrast-enhanced T1-weighted MRI of a diffusable tracer: Standardized quantities and symbols. *J. Magn. Reson. Imaging* [Internet] 1999;10:223–232. doi: 10.1002/(SICI)1522-2586(199909)10:3<223::AID-JMRI2>3.0.CO;2-S.
33. Sourbron SP, Buckley DL. Classic models for dynamic contrast-enhanced MRI. *NMR Biomed.* [Internet] 2013;26:1004–27. doi: 10.1002/nbm.2940.
34. Johnson JA, Wilson TA. A model for capillary exchange. *Am. J. Physiol. Heart Circ. Physiol.* [Internet] 1966;210:1299–1303. doi: 10.1152/ajplegacy.1966.210.6.1299.
35. Baselice F, Ferraioli G, Pascazio V. Relaxation Time Estimation from Complex Magnetic Resonance Images. *Sensors* [Internet] 2010;10:3611–3625. doi: 10.3390/s100403611.
36. Nowak RD. Wavelet-based Rician noise removal for magnetic resonance imaging. *IEEE Trans. Image Process.* 1999;8:1408–1419. doi: 10.1109/83.791966.
37. Gudbjartsson H, Patz S. The Rician distribution of noisy MRI data. *Magn. Reson. Med.* [Internet] 1995;34:910–4. doi: 10.1002/mrm.1910360222.
38. Garpebring A, Ostlund N, Karlsson M, Östlund N. A Novel Estimation Method for Physiological Parameters in Dynamic Contrast-Enhanced MRI: Application of a Distributed Parameter Model Using Fourier-Domain Calculations. *IEEE Trans. Med. Imaging* [Internet] 2009;28:1375–83. doi: 10.1109/TMI.2009.2016212.
39. Bartoš M, Šorel M, Jiřík R. Time-Efficient Fourier Domain Evaluation of Pharmacokinetic Model in Dynamic Contrast-Enhanced Magnetic Resonance Imaging. In: *IFMBE Proceedings*. Vol. 68. ; 2019. pp. 777–781. doi: 10.1007/978-981-10-9035-6\_143.
40. Sourbron SP, Buckley DL. Tracer kinetic modelling in MRI: estimating perfusion and capillary permeability. *Phys. Med. Biol.* [Internet] 2011;57:R1–R33. doi: 10.1088/0031-9155/57/2/R1.
41. Rudin LI, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms. *Phys. D Nonlinear Phenom.* [Internet] 1992;60:259–268. doi: 10.1016/0167-2789(92)90242-F.
42. Rajmic P, Novosadova M. On the limitation of convex optimization for sparse signal segmentation. 2016 39th Int. Conf. Telecommun. Signal Process. TSP 2016 2016:550–554. doi:

10.1109/TSP.2016.7760941.

43. Michel V, Gramfort a, Varoquaux G, Eger E, Thirion B. Total variation regularization for fMRI-based prediction of behaviour. *IEEE Trans. Med. Imaging* [Internet] 2011;30:1–12. doi: 10.1109/TMI.2011.2113378.

44. Rakotomamonjy A, Flamary R, Gasso G. DC Proximal Newton for Nonconvex Optimization Problems. *IEEE Trans. Neural Networks Learn. Syst.* 2016;27:636–647. doi: 10.1109/TNNLS.2015.2418224.

45. Salzo S, Villa S. Convergence analysis of a proximal Gauss-Newton method. *Comput. Optim. Appl.* 2012;53:557–589. doi: 10.1007/s10589-012-9476-9.

46. Maier O, Schoormans J, Schloegl M, Strijkers GJ, Lesch A, Benkert T, Block T, Coolen BF, Bredies K, Stollberger R. Rapid T<sub>1</sub> quantification from high resolution 3D data with model-based reconstruction. *Magn. Reson. Med.* [Internet] 2019;81:2072–2089. doi: 10.1002/mrm.27502.

47. Obad N, Espedal H, Jirik R, Sakariassen PO, Brekke Rygh C, Lund-Johansen M, Taxt T, Niclou SP, Bjerkvig R, Keunen O. Lack of functional normalisation of tumour vessels following anti-angiogenic therapy in glioblastoma. *J. Cereb. Blood Flow Metab.* 2017. doi: 10.1177/0271678X17714656.

48. Jirik R, Soucek K, Mezl M, et al. Blind deconvolution in dynamic contrast-enhanced MRI and ultrasound. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol. 2014. IEEE; 2014. pp. 4276–4279. doi: 10.1109/EMBC.2014.6944569.

49. Li KL, Zhu XP, Waterton J, Jackson A. Improved 3D quantitative mapping of blood volume and endothelial permeability in brain tumors. *J. Magn. Reson. Imaging* 2000;12:347–357. doi: 10.1002/1522-2586(200008)12:2<347::AID-JMRI19>3.0.CO;2-7.

50. Taxt T, Jiřík R, Rygh CB, Grüner R, Bartoš M, Andersen E, Curry FR, Reed RK. Single-channel blind estimation of arterial input function and tissue impulse response in DCE-MRI. *IEEE Trans. Biomed. Eng.* 2012;59:1012–1021. doi: 10.1109/TBME.2011.2182195.

51. Kratochvíla J, Jiřík R, Bartoš M, Standara M, Starčuk Z, Taxt T. Distributed capillary adiabatic tissue homogeneity model in parametric multi-channel blind AIF estimation using DCE-MRI. *Magn. Reson. Med.* [Internet] 2016;75:1355–1365. doi: 10.1002/mrm.25619.

52. Henderson E, Sykes J, Drost D, Weinmann H-J, Rutt BK, Lee T-Y. Simultaneous MRI measurement of blood flow, blood volume, and capillary permeability in mammary tumors using two different contrast agents. *J. Magn. Reson. Imaging* [Internet] 2000;12:991–1003. doi: 10.1002/1522-2586(200012)12:6<991::AID-JMRI26>3.0.CO;2-1.

Submitted to MRM, 24th April 2019

53. Šorel M, Bartoš M. Fast Bayesian JPEG Decompression and Denoising With Tight Frame Priors. *IEEE Trans. Image Process.* [Internet] 2017;26:490–501. doi: 10.1109/TIP.2016.2627802.

54. Foi A. Noise estimation and removal in MR imaging: The variance-stabilization approach. *Proc. - Int. Symp. Biomed. Imaging 2011*:1809–1814. doi: 10.1109/ISBI.2011.5872758.

55. Bredies K, Holler M. A TGV-Based Framework for Variational Image Decompression, Zooming, and Reconstruction. Part I: Analytics. *SIAM J. Imaging Sci.* [Internet] 2015;8:2814–2850. doi: 10.1137/15M1023865.

56. Fusco R, Sansone M, Petrillo M, Petrillo A. Influence of parameterization on tracer kinetic modeling in DCE-MRI. *J. Med. Biol. Eng.* 2014;34:157–163. doi: 10.5405/jmbe.1097.

Table 1: Mathematical notations

Symbol	Description
$\mathbf{A}, \mathbf{p}$	Matrices
$\mathbf{x}$	Vector
$x_j$	The $j^{th}$ element of $\mathbf{x}$
$A_{i,j}$	One element of $\mathbf{A}$ in the $i^{th}$ row and the $j^{th}$ column
$p_{i,}$	The $i^{th}$ row of $\mathbf{p}$
$p_{.,j}$	The $j^{th}$ column of $\mathbf{p}$
$\mathbf{H}_i$	The $i^{th}$ matrix from the array of matrices
$\mathbf{y}^k$	Vector $\mathbf{y}$ in the $k^{th}$ iteration
$\theta$	constant
$\hat{\mathbf{p}}, \mathbf{p}^*$	Estimate, ground truth of $\mathbf{p}$
$ \cdot $	Magnitude of a vector field or a number

Table 2: Definitions of perfusion parameters and related quantities

Perfusion parameter	Description	Units	Starting point, $\mathbf{p}^0$	Constraints, $P$	Relative weight, $\gamma'$
$F_p$	Plasma flow <sup>a</sup>	ml/min/ml	1	$[10^{-3}, 10^2]$	0.025
$T_c$	Mean capillary transit time	min	0.1	$[\Delta t, 3]$	0.283
$T_e$	Mean transit time of EES <sup>b</sup>	min	2.5	$[\Delta t, 10^2]$	0.024
$\alpha$	$= PS/F_p = -\ln(1 - E)$	–	0.4	$[10^{-4}, 3]$	0.103
$\tau$	Bolus arrival time	min	0	$[-0.5, 1]$	0.565
$PS$	Permeability-surface area product	ml/min/ml			
$E$	Extraction fraction	–			
$v_p$	Plasma volume	ml/ml			
$v_e$	EES <sup>b</sup> volume	ml/ml			
$k_{ep}$	EES <sup>b</sup> -to-plasma rate constant	1/min			
$K^{trans}$	Volume transfer constant	ml/min/ml			

<sup>a</sup>– values take the reweighting by  $\alpha_{TRF}$ ,  $\alpha_{AIF}$  into account

<sup>b</sup>– extravascular extracellular space

Table 3: Quantitative comparison of the proposed TV-regularized (TV) with  $\Gamma = 0.22$  and non-regularized voxel-wise (VW) estimations with the ground truth (GT) numerical rat phantom evaluated statistically per tissue in tumor.  $\mu$  stands for mean,  $\sigma$  for standard deviation. Bold values indicate lower bias or higher precision. Complete table can be found in supporting information (Supporting Information Table S1).

Tissue <sup>a</sup> [label]	Area [voxels]	SNR <sup>b</sup> [dB]	Method	$F_p$ [ml/ml/min]		$T_c$ [min]		$T_e$ [min]		$E$ [-]		$v_p$ [ml/ml]		$K^{trans}$ [ml/ml/min]	
				$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
41	1	23.0	GT	0.13		0.27		1.85		0.41		0.033		0.052	
			VW	0.12		<b>0.24</b>		<b>1.95</b>		<b>0.39</b>		0.029		<b>0.047</b>	
			TV	<b>0.13</b>		0.23		1.99		0.37		<b>0.029</b>		0.046	
39	5	22.1	GT	0.12		0.23		1.98		0.39		0.028		0.047	
			VW	<b>0.12</b>	0.01	0.27	0.03	<b>2.09</b>	0.09	<b>0.38</b>	0.02	0.032	0.001	0.046	0.002
			TV	0.13	<b>0.00</b>	<b>0.25</b>	<b>0.01</b>	2.12	<b>0.09</b>	0.36	<b>0.02</b>	<b>0.031</b>	<b>0.001</b>	<b>0.046</b>	<b>0.002</b>
37	7	20.9	GT	0.12		0.27		2.39		0.35		0.032		0.041	
			VW	0.11	0.02	<b>0.27</b>	0.04	<b>2.49</b>	<b>0.26</b>	<b>0.34</b>	0.04	<b>0.028</b>	0.004	0.037	0.007
			TV	<b>0.11</b>	<b>0.01</b>	0.24	<b>0.01</b>	2.52	0.27	0.32	<b>0.03</b>	0.028	<b>0.004</b>	<b>0.037</b>	<b>0.007</b>
40	4	18.4	GT	0.12		0.29		2.23		0.30		0.035		0.036	
			VW	<b>0.12</b>	0.01	<b>0.26</b>	0.04	<b>2.41</b>	0.20	<b>0.28</b>	0.03	<b>0.030</b>	0.002	<b>0.033</b>	0.003
			TV	0.12	<b>0.01</b>	0.24	<b>0.01</b>	2.45	<b>0.17</b>	0.26	<b>0.01</b>	0.030	<b>0.001</b>	0.033	<b>0.003</b>
38	8	17.2	GT	0.11		0.25		2.77		0.27		0.027		0.029	
			VW	<b>0.11</b>	0.02	<b>0.25</b>	0.04	<b>2.92</b>	<b>0.30</b>	<b>0.26</b>	0.02	<b>0.026</b>	0.002	0.028	0.003
			TV	0.11	<b>0.01</b>	0.23	<b>0.01</b>	2.95	0.31	0.25	<b>0.01</b>	0.026	<b>0.002</b>	<b>0.028</b>	<b>0.003</b>
36	57	15.4	GT	0.07		0.22		3.97		0.25		0.016		0.018	
			VW	<b>0.08</b>	0.02	<b>0.22</b>	0.05	3.95	0.54	<b>0.25</b>	0.04	0.017	0.004	<b>0.019</b>	0.004
			TV	0.08	<b>0.01</b>	0.20	<b>0.03</b>	<b>3.96</b>	<b>0.50</b>	0.24	<b>0.03</b>	<b>0.016</b>	<b>0.004</b>	0.019	<b>0.004</b>
35	116	8.7	GT	0.09		0.08		6.11		0.10		0.007		0.009	
			VW	<b>0.07</b>	0.06	0.16	0.09	<b>6.11</b>	1.52	0.18	0.09	0.008	0.003	<b>0.009</b>	0.002
			TV	0.07	<b>0.03</b>	<b>0.12</b>	<b>0.02</b>	5.88	<b>0.53</b>	<b>0.15</b>	<b>0.03</b>	<b>0.008</b>	<b>0.003</b>	0.009	<b>0.002</b>
34	97	7.2	GT	0.07		0.09		6.93		0.10		0.006		0.007	
			VW	<b>0.07</b>	0.07	0.16	0.08	<b>7.13</b>	1.91	0.16	0.09	0.007	0.003	0.007	0.002
			TV	0.06	<b>0.02</b>	<b>0.11</b>	<b>0.02</b>	6.37	<b>0.40</b>	<b>0.12</b>	<b>0.05</b>	<b>0.007</b>	<b>0.002</b>	<b>0.007</b>	<b>0.002</b>

<sup>a</sup> – position of the labels can be found in Figure 1

<sup>b</sup> – mean SNR in the tissue

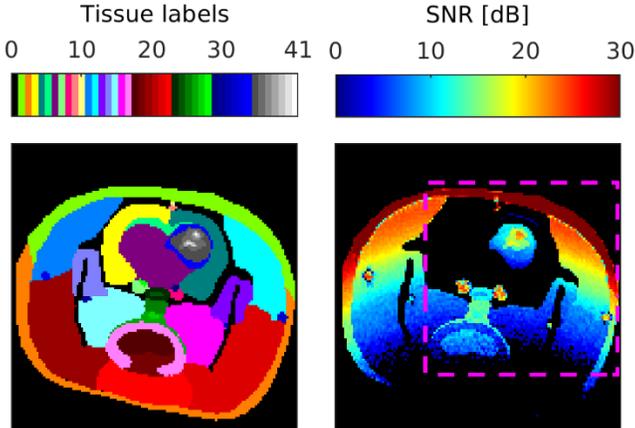


Figure 1 Numerical rat phantom – 41 different color-coded tissues (left) and the SNR induced to the phantom data by adding noise (right). The magenta rectangle shows the close-up used in Figure 2

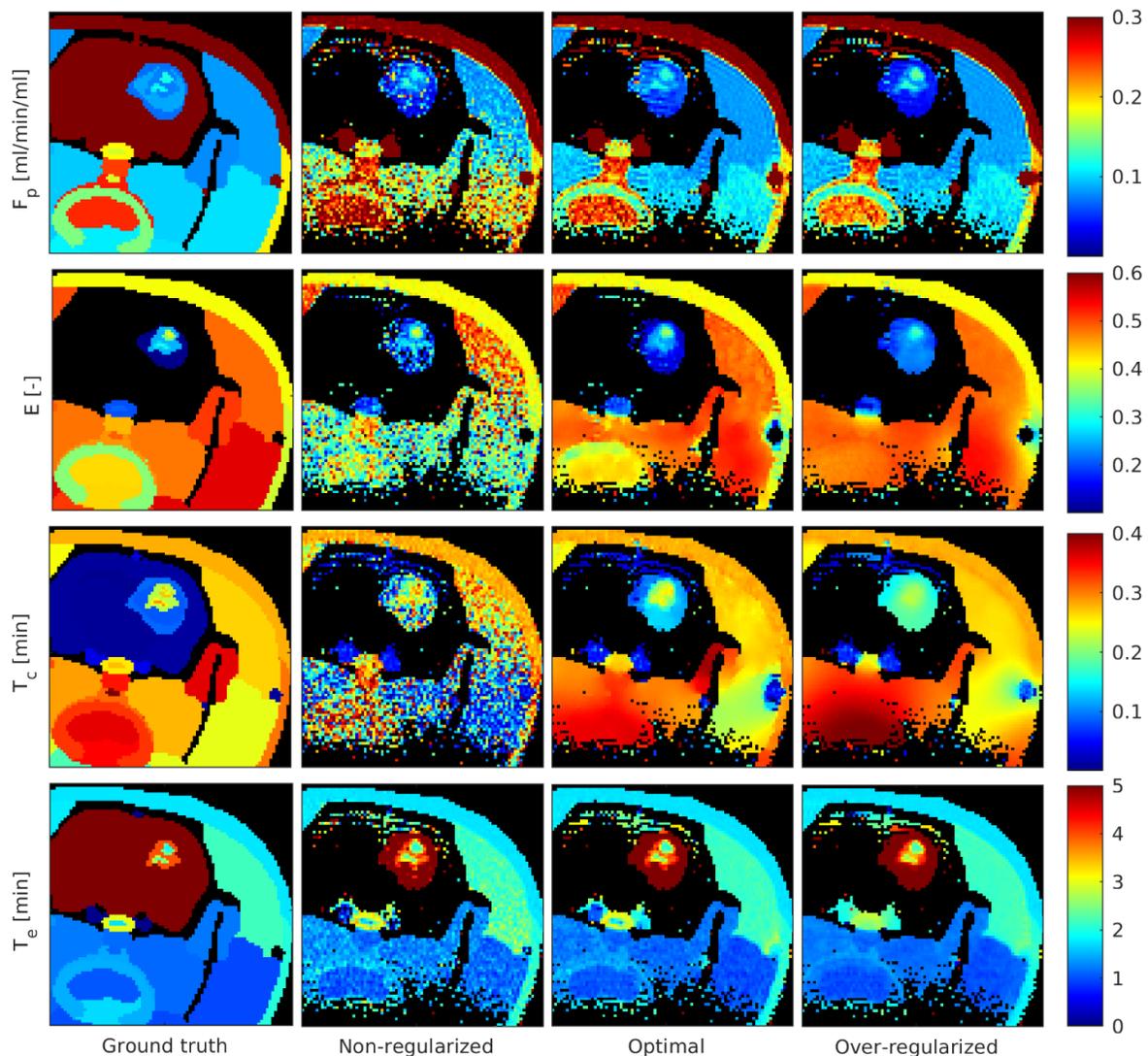


Figure 2 Effect of the regularization on perfusion parameters estimated from synthetic data in comparison with the ground truth. Only results for selected perfusion parameters and regularization weights are shown in a close-up defined in Figure 1. The black areas (e.g. brain tissue surrounding the tumor and areas far from the used surface coils) had  $\text{SNR} < 0$  dB (cf. Figure 1) and were excluded from the computation. The closest MAE distance to GT for most of the parameters is obtained for  $\Gamma = 0.22$  – column “Optimal” (cf. Figure 3). For  $\Gamma = 2.2$  – column “Over-regularized”, the maps are too smooth causing loss of details.

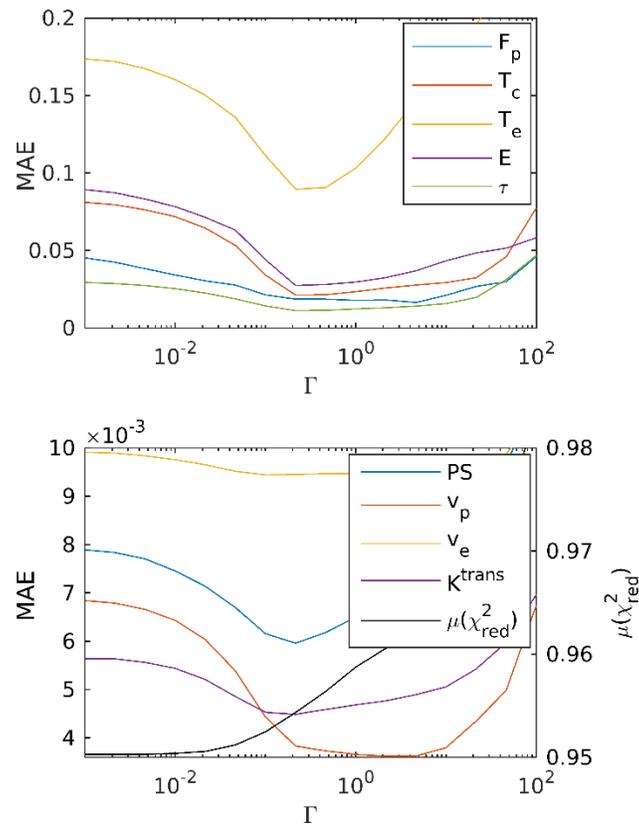


Figure 3 Effect of the regularization weight  $\Gamma$  on the distance of the parameter estimates from the ground truth (MAE) for the regularized parameters (top) and for the derived parameters (bottom – color curves). The effect of the regularization on the quality of fit represented by the mean of the reduced chi-squared statistic  $\mu(\chi_{red}^2)$  is also shown (bottom – black line).

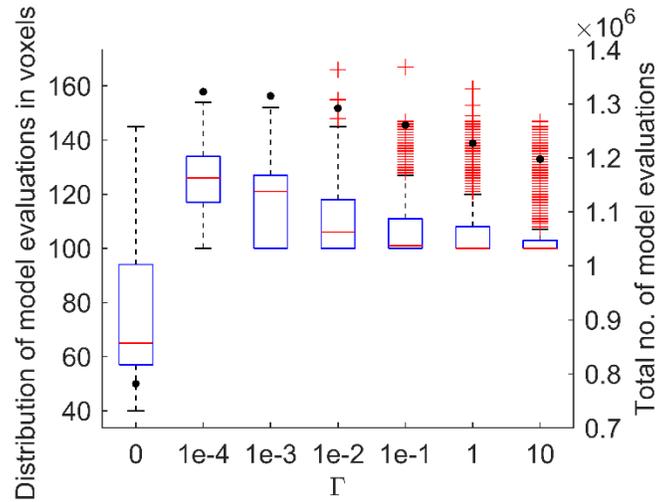


Figure 4 Distribution of the number of pharmacokinetic model evaluations in voxels shown in boxplots together with the total number of function evaluations (black points) as a function of the regularization strength.  $\Gamma = 0$  stands for the standard non-regularized algorithm. The number of voxels is  $I = 9916$ , the maximum number of iterations was 50.

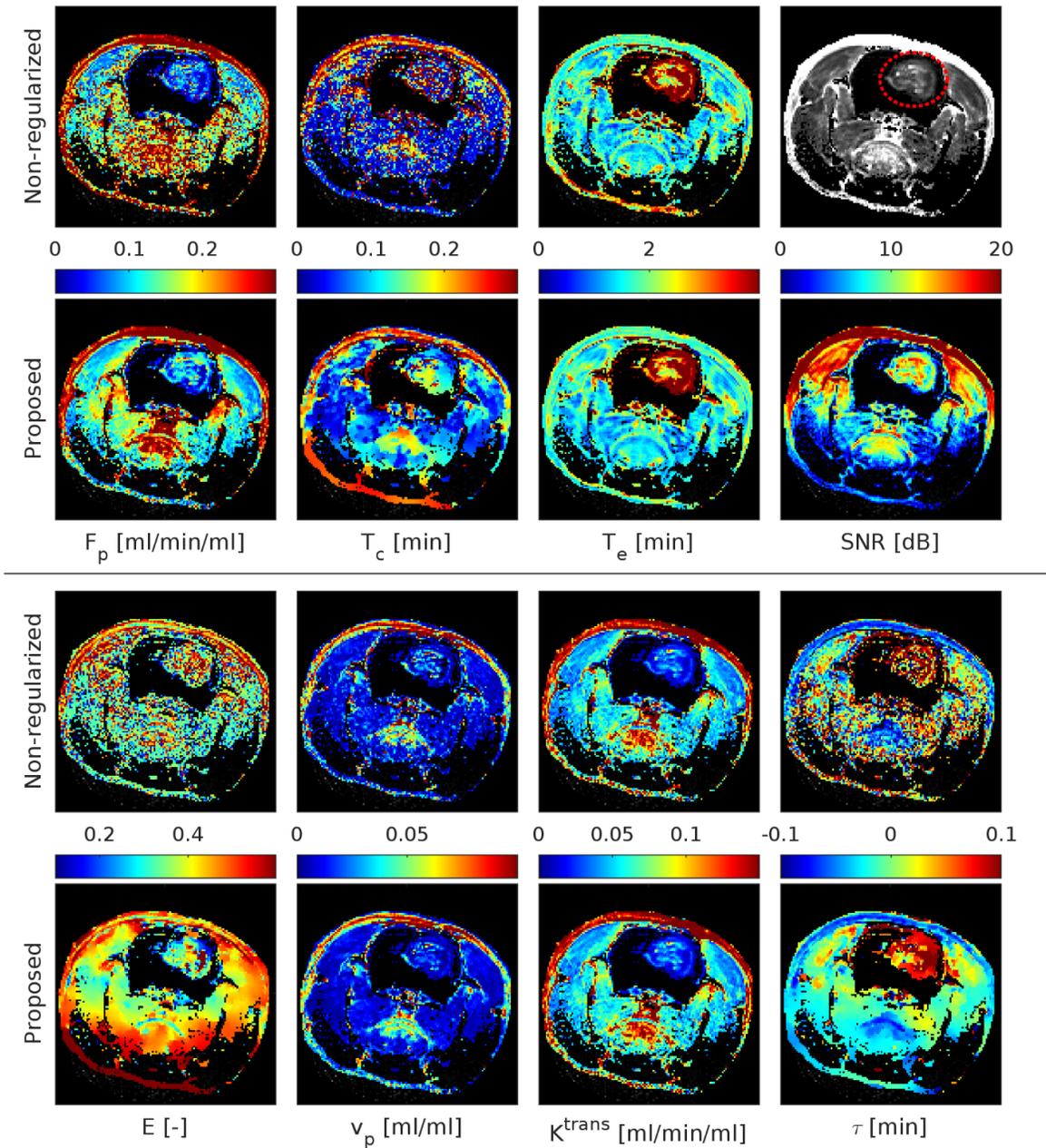


Figure 5 Preclinical data (rat brain with glioblastoma). Comparison of the perfusion parameter maps estimated without and with spatial regularization (top, bottom row in a group, respectively). The dotted circle in the anatomical image (top right) indicates the tumor. The final SNR map is shown only once since it was similar for both methods. The perfusion parameters were calculated only for curves satisfying  $\text{SNR} > 0$  dB.

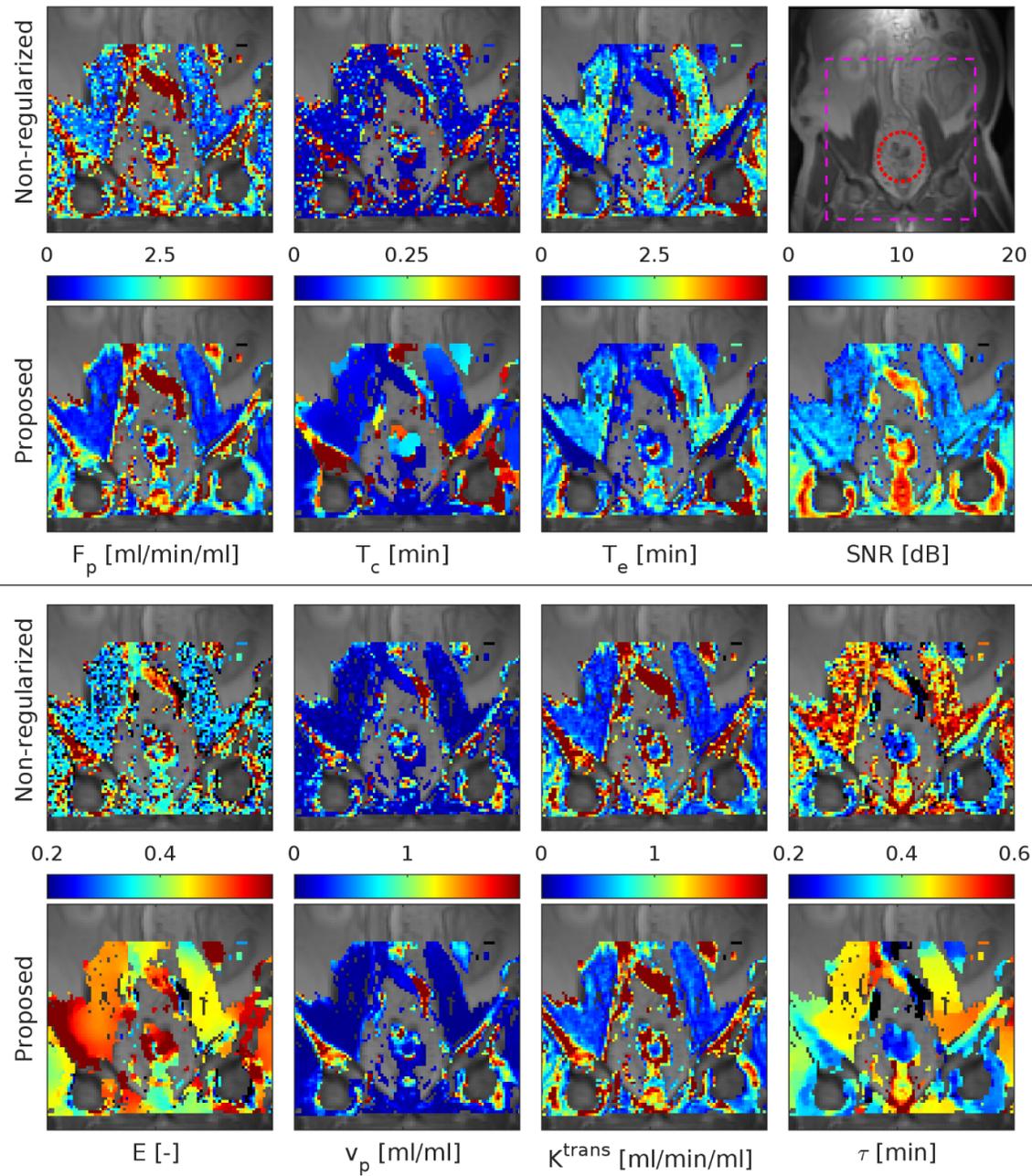


Figure 6 Clinical data (abdomen area with renal cell carcinoma metastasis). Comparison of the perfusion parameter maps estimated without and with spatial regularization (top, bottom row in a group, respectively). The dotted circle in the anatomical image (top right) indicates the tumor, the magenta rectangle indicates the closeup. The final SNR map is shown only once since it is similar for both methods. The perfusion parameters were calculated only for curves in a rectangular ROI satisfying  $SNR > 0$  dB.

Supporting Information Figure S1 Effect of the global regularization weight  $\Gamma$  on perfusion parameters estimated from the numerical rat phantom in comparison with the ground truth (GT). Estimates of the regularized perfusion parameters and the reduced  $\chi^2$  metric are shown for selected regularization weights. The location of the used close-up is defined in Figure 1.

Supporting Information Figure S2 Effect of the global regularization weight  $\Gamma$  on perfusion parameters estimated from the numerical rat phantom in comparison with the ground truth (GT). Estimates of the derived perfusion parameters and for selected regularization weights are shown in a close-up defined in Figure 1.

Supporting Information Figure S3 Simulated data, effect of the regularization weight  $\Gamma$  on the bias (estimated within a tissue) averaged using its absolute value over all tissues (left column) for the regularized parameters (top) and for the derived parameters (bottom). Similarly, the mean of tissue standard deviations is shown (measure of precision - right column). In the formulas,  $K, k$  relate to tissue indices and  $N, n$  relate to indices of voxels inside the tissues.

Supporting Information Figure S4 Effect of the regularization (Proposed,  $\Gamma = 0.22$ ) on the mean and standard deviation of the perfusion-parameter estimates from 50 noise realizations of the numerical rat phantom in comparison with the ground truth (GT).

Supporting Information Table S1: Quantitative comparison of the proposed TV-regularized (TV) and non-regularized voxel-wise (VW) estimations from a single noise realization with the ground truth of the numerical rat phantom. The mean,  $\mu$ , and standard deviation,  $\sigma$ , of the estimates are evaluated within a tissue region. Bold values indicate the lowest bias and the highest precision.

Supporting Information Table S2: Statistical quantitative comparison of the proposed TV-regularized (TV) and non-regularized voxel-wise (VW) estimations from 50 noise realizations with the ground truth of the numerical rat phantom. The estimation statistic is expressed as the mean and standard deviation shown in Figure S4 averaged within each tissue region, denoted as  $\mu$  and  $\sigma$ , respectively. Bold values indicate the lowest bias and the highest precision. ##### represents values out of range caused by an outlier.