Deconvolution for DCE-MRI Using an Exponential Approximation Basis

Stephen L. Keeling[1], Thomas Kogler[2] and Rudolf Stollberger[3]

Abstract. A deconvolution approach for dynamic contrast enhanced magnetic resonance imaging using an approximation basis of exponential functions constrained to be non-negative and non-increasing is developed and compared with widely used methods. Monotonicity in an exponential basis is implemented in terms of a newly derived condition which is a considerable generalization over a previous condition that implies complete monotonicity. Since the constraints imply a bound on the total variation, a well known staircasing effect may result with other approximation bases, but an exponential basis is shown to resist staircasing. In addition to the choice of approximation basis, further regularization is implemented in terms of the numbers of basis functions and the distribution of their parameters. The exponential approach is applied to dynamic contrast enhanced magnetic resonance imaging data to determine physiological parameters pixelwise to visualize a cerebral tumor, and the results are compared favorably with those of the standard Truncated Singular Value Decomposition approach. In particular, kernels estimated with constrained exponentials are free of oscillations and staircasing, and the images of estimated kernel parameters are sharper than those obtained by Truncated Singular Value Decomposition.

1 Introduction

A major goal in functional magnetic resonance imaging is to evaluate tissue perfusion, or fluid flow, and the exchange between tissues and blood, or vessel permeability. For instance, ischemic tissues are characterized by a low oxygen state due to inadequate blood flow. On the other hand, tumors are characterized by excessive vascularization and a corresponding increase in blood flow. Further, the transport mechanisms involved may be convection or diffusion dominated when the volumetric flow rate is large or small, respectively, in relation to vessel permeability for instance. To evaluate the transport states of tissues, a bolus of contrast agent is injected into a patient, and rapid magnetic resonance imaging techniques are used to follow the concentration time course of the contrast agent at a given tissue site. The purpose of the present work was to develop estimation techniques by which dynamic contrast enhanced magnetic resonance imaging (DCE-MRI) data may be used to quantify the transport states described here. While attempting to clarify precisely what can be determined from these data, it was discovered that the theory of tracer transport and the convolution model of DCE-MRI requires revision, and the results of that investigation are reported separately in [11]. In the present paper, the authors present their deconvolution techniques whose applicability is naturally constrained by the theoretical results reported in [11].

In Section 2 aspects of the measurement process and the physiological modeling which are relevant to the present work are summarized briefly. Specifically, in the convolution model, an unknown kernel or impulse response for an identified tissue region is convolved with a (measured) arterial input concentration of contrast agent to obtain a (measured) contrast agent concentration for the identified tissue region. Then identifiable physiological parameters are

---

[1] Institut für Mathematik und Wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, Heinrichstraße 36, 8010 Graz, Austria; email: stephen.keeling@uni-graz.at; tel: +43–316–380–5156; fax: +43–316–380–9815. Supported by the Fonds zur Förderung der wissenschaftlichen Forschung under SFB 03, “Optimierung und Kontrolle”.

[2] Abteilung für Klinische und Experimentelle Magnetresonanzforschung, Universitätsklinik für Radiologie, Medizinische Universität Graz, Auenbruggerplatz 9, 8036 Graz, Austria; e-mail: th.kogler@meduni-graz.at; tel: +43-316-385-3707; fax: +43-316-385-3848; Supported by the Fonds zur Förderung der wissenschaftlichen Forschung under SFB 03, “Optimierung und Kontrolle”.

[3] Institut für Medizintechnik, Technische Universität Graz, Kronesgasse 5/II, 8010 Graz, Austria; e-mail: rudolf.stollberger@tugraz.at; tel: +43–316–873–5370; fax: +43–316–873–7890. Supported by the Fonds zur Förderung der wissenschaftlichen Forschung under SFB 03, “Optimierung und Kontrolle”.

given in terms of the kernel function [11], i.e., the volumetric flow rate per unit tissue volume, the mean transit time, and the effective volume fraction. Note that parameterized models have been used alternatively, e.g., in [12] [22], particularly in an effort to identify the tissue states described above [7], [8], [13]. It is shown in [11] that these tissue states cannot be decided on the basis of a single convolution equation alone. Nevertheless, in the present work regularized deconvolution methods are presented for the estimation of the kernel function and the physiological parameters which can be deduced from it.

While many regularization methods have been used for deconvolution [4], the methods popularized for transport kernel estimation in particular include filtered Discrete Fourier Transform [16], Truncated Singular Value Decomposition [2] [16], Tikhonov Regularization [2], penalized monotonicity [16], and constrained monotonicity [7]. In Section 3 the ill-posed nature of the deconvolution problem at hand is discussed in the present physiological context. Specifically, it is shown that this deconvolution problem is progressively more sensitive to measurement noise as the number of injection-to-tissue compartments increases. The deconvolution is regularized in a natural way by constraining the estimated kernel to be non-negative and non-increasing. However, it is seen that these constraints imply a bound on the total variation (TV) of the estimated kernel. TV regularization is known for its excellent reconstruction of apparently piecewise constant or *blocky* data; however, TV regularization suffers from a well known *staircasing effect* in regions with gradual variations [9], as can be seen in the results of [7]. It is shown here that an exponential basis naturally resists staircasing while, for instance, a spline basis does not. In addition to the choice of approximation basis, further regularization is implemented in terms of the numbers of basis functions and the distribution of their parameters. Additional criteria for a function basis are that it permit a convenient formulation of the constraints and possess approximation properties while simultaneously offering a physiologically realistic representation of the kernel in a low-dimensional approximation space. All these conditions are met by an exponential approximation basis. See also [3] for estimation with exponential functions.

The implementation of exponential basis functions for transport kernel deconvolution is presented in Section 4. The general approximation properties of exponential functions follow from the stated Müntz Theorem [21]. On the other hand, the physiological significance of exponential functions is explained in [11]. Then a condition for monotonicity of a sum of exponentials is derived and compared computationally with alternatives. The newly derived monotonicity constraint is a considerable generalization over the non-negative coefficients constraint of [3] which implies complete monotonicity [23]. It is shown in Section 4 that the freedom afforded by flexibility in the exponential time scales can be destabilizing. Yet, harmonically distributed time scales satisfy the theoretical conditions of the Müntz Theorem, and computationally observed results verify that such time scales exhibit good approximation properties; therefore, time scales are chosen to be harmonically distributed. Then the number of time scales is chosen according to an L-curve criterion [4], i.e., from the turning point in a log-residual/log-derivative curve generated by varying the number of time scales. Finally, the discretized problem is solved by linearly constrained least squares. Computational results with constrained exponentials for artificial and measured data are compared with results obtained by Truncated Singular Value Decomposition, Tikhonov Regularization and constrained splines, and the constrained exponentials yield estimations which are more natural and better suited to typical physiological data. On the basis of the strongest contrast seen between kernels estimated by the exponential approach and by Truncated Singular Value Decomposition, which has become standard for MRI perfusion analysis, these methods are next compared by applying them to a challenging DCE-MRI sequence with low signal-to-noise ratio to determine physiological parameters pixelwise to visualize a cerebral tumor. It is found that kernels estimated with constrained exponentials are free of oscillations and staircasing, and the images of estimated kernel parameters are sharper than those obtained by Truncated Singular Value Decomposition.

2 Summary of the Convolution Formulation of DCE-MRI

The quantitative objective of DCE-MRI is to infer tissue transport properties from the temporal variation of intensities in a series of magnetic resonance images measured during the injection of contrast agent. In particular, the two images shown in Fig. 1 are samples from a sequence of T_1-weighted images and they were measured respectively before and after the appearance of a Gadolinium-DTPA based contrast agent which causes a change in intensity. The degree of intensity change is used to compute the local concentration C of the contrast agent according to an algebraic relation, which depends upon the scanning technique; see, e.g, [16]. This procedure is used to determine contrast agent concentration in an arbitrary region of interest Ω from imaging. Regions of interest Ω are typically pixels of tissues infiltrated by a capillary bed, and pixels corresponding to macro-vessels or image artifacts may be removed using independent component analysis as addressed in Section 4. The distribution of a contrast agent of interest in the present work is determined by the purely passive transport mechanisms of convection and diffusion.

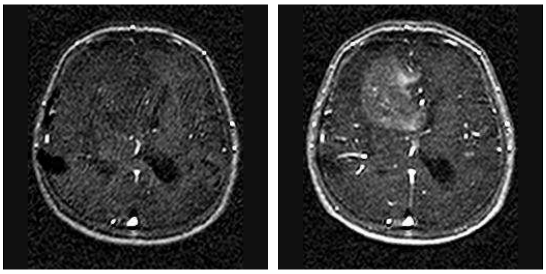

Figure 1: Magnetic resonance images taken from a series of images measured during the injection of a Gadolinium-DTPA based contrast agent. From left to right the images were measured respectively before and after the appearance of the contrast agent.

It is shown in [11] by integrating a convection-diffusion equation for the distribution of contrast agent concentration that the the concentration C_{T} within the tissue region Ω may be related to the concentration C_{AIF} on the inflow boundary of Ω (i.e., the arterial input function) by the following convolution:

$$C_{\mathrm{T}}(t) = \int_0^t K(t-s) C_{\mathrm{AIF}}(s) ds. \tag{2.1}$$

Here, K is an unknown convolution kernel equal to the tissue impulse response $C_{\mathrm{T}}(t) = K(t)$ for an impulsive arterial input $C_{\mathrm{AIF}}(t) = \delta(t)$. Thus, $K(t)$ is necessarily non-increasing and vanishes as the contrast agent clears from Ω [11]. In the manner described in Section 2, C_{T} is determined from intensities at the tissue region Ω and C_{AIF} is determined from intensities in a major artery nearby to Ω. Given measured C_{T} and C_{AIF}, the kernel $K(t)$ is determined by deconvolution methods such as those presented in this paper, and identifiable tissue transport properties are computed as follows; see [11]. The *volumetric flow rate per unit tissue volume* is given by:

$$\mathcal{F}_{\mathrm{T}} = K(0). \tag{2.2}$$

The *mean transit time* is given by:

$$\mathcal{T}_{\mathrm{T}} = \int_0^\infty \frac{K(t)}{K(0)} dt. \tag{2.3}$$

Finally, the *effective volume fraction* is given by:

$$\mathcal{V}_{\mathrm{T}} = \int_0^\infty K(t) dt. \tag{2.4}$$

The reader familiar with the literature concerning the association between kernel parameters and transport properties, e.g., [7], [8], [12], [13], [22], may inquire which types of examinations or tracers are characterized by (2.2), (2.3) and (2.4). It is shown in [11] that this characterization applies to all applications in which tracer is transported passively by convection and diffusion. In particular, it is shown in [11] that diffusive transport states are not identifiable from the convolution kernel since a purely convective transport state can be constructed which has a matching convolution kernel. In spite of the limitations imposed by the assumptions made to arrive at the convolution model, (2.1) is a useful formulation for DCE-MRI and it is applied below to estimate the identifiable physiological parameters given in (2.2), (2.3) and (2.4).

3 The Ill-Posedness and its Proposed Regularization

The deconvolution problem of solving for the kernel K in (2.1), from given data C_{T} and C_{AIF}, is well known to be ill-posed; see [4] and the references cited therein. Nevertheless, insight about the particular deconvolution problem at hand can be gained by considering the degree of ill-conditioning with respect to the physiological context. For this, suppose ideally that contrast agent is injected impulsively. Then the bolus may typically travel through at least three major compartments: from the injection site through the right heart, from the right heart through the lungs to the left heart, and from the left heart to the tissue. For simplicity, assume that each compartment is well mixed with the same mean transit time $1/\nu > 0$, and thus the same compartment kernel $\nu e^{-\nu t}$ [11]. When the impulsive input must travel through say $n+1$ such compartments, the arterial input at the tissue site is given through the $(n+1)$-fold convolution of the unit kernel:

$$C_{\mathrm{AIF}}(t) = \delta(t) * [\nu e^{-\nu t}]_1 * \cdots * [\nu e^{-\nu t}]_{n+1} = \nu \frac{(\nu t)^n}{n!} e^{-\nu t}. \tag{3.1}$$

For this exact arterial input function assume that K_{e} is the exact tissue kernel and that the exact tissue concentration is given by $C_{\mathrm{T}} = K_{\mathrm{e}} * C_{\mathrm{AIF}}$. Then ill-posedness follows from the existence of a noisy disturbance N_ε of the exact data C_{T} which generates an error E_ε in the kernel estimation which satisfies:

$$C_{\mathrm{T}}(t) + \underbrace{N_\varepsilon(t)}_{\mathrm{noise}\to 0, \varepsilon\to 0} = \int_0^t C_{\mathrm{AIF}}(t-s)[K_{\mathrm{e}}(s) + \underbrace{E_\varepsilon(s)}_{\mathrm{error}\to\infty, \varepsilon\to 0}]ds. \tag{3.2}$$

It is seen from the norm estimates in (3.3) below, which are proved in Appendix A, that the deconvolution problem is progressively more sensitive to noise as the number of injection-to-tissue compartments $n+1$ increases and thereby increases the initial smoothness of C_{AIF}.

Theorem 1 *For $n > 1$ there exists a noise function $N_\varepsilon(t)$ and an error function $E_\varepsilon(t)$ which satisfy (3.2) and:*

$$\max_{t\geq 0} |N_\varepsilon(t)| = \mathcal{O}(\varepsilon), \quad \max_{t\geq 0} |E_\varepsilon(t)| = \mathcal{O}(\varepsilon^{-n}). \tag{3.3}$$

In order to regularize this ill-posedness, the kernel K is computed in this work by solving a constrained minimization problem:

$$\min_{K\in B} \|C_{\mathrm{T}} - K * C_{\mathrm{AIF}}\|_{L_2[0,T]} \qquad \text{subject to:} \quad K' \leq 0, \quad K \geq 0 \tag{3.4}$$

where B is an appropriate function class. Also, $L_2[0,T]$ denotes the Lebesgue square integrable functions on the interval $[0,T]$ and $\mathcal{C}^k(I)$ below denotes the functions k-times continuously differentiable in an interval I; see [1] for further details on function spaces. The principal *a priori* information known of the kernel is that it is non-negative and non-increasing, and these constraints are incorporated in (3.4). Thus, when the data are positive on $(0,T]$, the minimizing

K in (3.4) can be infinite at most at $t = 0$, so K is bounded on every interval $[\epsilon, T]$, $\epsilon > 0$ and thus has bounded variation [1]. As a result, a kernel K estimated by solving (3.4) may manifest a staircasing effect, which is well known in the context of total variation regularization; see [9] and the references cited therein. This effect can be seen in the results of [7], and is illustrated here in Figs. 2 and 3 below. Specifically, it can be seen in Fig. 3 that an estimation under a bounded variation constraint provides a good reconstruction of a piecewise constant kernel; however, Fig. 2 shows that a kernel with gradual variations is estimated with a sequence of *stairsteps*.

Thus, further regularization is required to estimate gradual variations accurately, and that regularization is achieved here through the choice of basis functions in the finite dimensional set B as well as through the numbers of basis functions and the distribution of their parameters. The requirements for the function basis B are the following. First, it is required that a given kernel may be approximated with arbitrary accuracy for a sufficient number of basis functions; thus, the approach is nonparametric in this sense. On the other hand, while function bases such as step functions and global polynomials satisfy this technical property, they deliver physiologically unrealistic estimations when the dimension of the approximation space is chosen small for regularization purposes. Thus, the second criterion for basis functions is that low dimensional estimations be physically realistic. Finally, it is required that the inequality constraints in (3.4) be easily implemented in terms of linear conditions on the span coefficients, which do not rule out significant non-increasing elements in the basis span. These conditions are met by the exponential basis discussed below in Section 4.

4 Deconvolution with an Exponential Basis

In this section, the minimization problem (3.4) is solved over the set B composed of the span of decaying exponential functions. The deconvolution techniques are described and examples are presented to explain the methods chosen. As shown in [11], decaying exponentials have a certain physiological significance, and thus they constitute a meaningful basis for the kernel estimation problem. It will be shown here that the criteria set forth at the end of Section 3 are satisfied by exponential basis functions. Finally, based upon the apparently favorable performance of exponential functions seen in Fig. 9, the exponential deconvolution approach is compared qualitatively with the standard TSVD approach for the pixelwise determination of physiological parameters from the measured image sequence of Fig. 1.

4.1 Description of the Test Data

For the demonstrations shown here, artificial data are determined as follows. The input C_{AIF} is a piecewise linear interpolant of discrete samples from a sum of functions of the form shown in (3.1), where each subsequent summand has a delay to model recirculation through the blood stream. Two types of exact kernels K_{e} are used: either a step function or a bi-exponential function. The motivation for such test cases is as follows. Step function kernels are also tested in [16] and correspond to the extreme physiological case of plug flow [11] as well as the extreme mathematical case of a discontinuous kernel. Bi-exponential kernels correspond to the more benign physiological case of two-compartment transport, which may be either purely convective or (in fact indistinguishably) a combination of convection and diffusion [11]. For both types of exact kernels, C_{T} is the piecewise linear interpolant of discrete samples of $K_{\mathrm{e}} * C_{\mathrm{AIF}}$. Finally, the so constructed data C_{AIF} and C_{T} are corrupted by Gaussian distributed noise with a selected standard deviation. Although MRI intensity data are non-negative and Riccian distributed [14], the algebraic computation of concentrations, as explained in Section 2, changes the noise distribution of data; thus, a Gaussian noise model is standard [16].

The acquisition of measured data is described briefly in Section 2. Recall that C_{AIF} is approximated by measured concentrations in a major artery near the tissue of interest. All concentrations are taken from the same image sequence measured on a fixed regular time grid $\{t_i = iT/(N-1)\}_{i=0}^{N-1}$, but the time required for contrast agent to travel from the chosen major artery to the tissue site constitutes an unknown delay between the measured C_{AIF} and the actual arterial input at the tissue site. Furthermore, this delay is often shorter than the data time interval $t_1 - t_0$. The authors have performed computational experiments to estimate this delay by minimizing the constrained residual of (3.4) with respect to a temporal shift in C_{AIF}. However, estimating this delay accurately under typical noise levels is a challenging problem still under investigation, and it will not be addressed further in the present work.

4.2 Description of the Exponential Basis

For a given M and positive constants $\{\lambda_m\}_{m=1}^{M}$, let B denote the span of the exponential functions $\{e^{-\lambda_m t}\}_{m=1}^{M}$. Then the estimated kernel K takes the form:

$$K(t) = \sum_{m=1}^{M} k_m e^{-\lambda_m t} \tag{4.1}$$

where the constants $\{\lambda_m^{-1}\}_{m=1}^{M}$ in the terms $e^{-t/\lambda_m^{-1}}$ represent the time scales of the basis functions. The approximation properties of such sums of decaying exponential functions are given in the following theorem, which is proved in [21]. Because of these approximation properties, the exponential approach is nonparametric.

Theorem 2 *[Müntz] Suppose $\{\lambda_n\}_{n=1}^{\infty}$ are positive and satisfy $\lim_{n\to\infty} \lambda_n = +\infty$. Then the span of $\{e^{-\lambda_n t}\}$ is dense in $L_p[0,\infty)$ for $1 \le p < \infty$ if and only if $\sum_{n=1}^{\infty} 1/\lambda_n = +\infty$ holds.*

With the sum in (4.1), the residual in (3.4) can be computed exactly for piecewise linear data C_{AIF} and C_{T}. On the basis of results shown below in Subsections 4.4 and 4.5, the basis time scales are chosen to satisfy Theorem 2 according to the harmonic distribution $\lambda_m^{-1} = T/m$, $m = 1, \dots, M$, with a total number M determined by a balance between the constrained residual $\|C_{\mathrm{T}} - K * C_{\mathrm{AIF}}\|_{L_2[0,T]}$ and the kernel derivative $\|K'\|_{L_2[0,T]}$.

To enforce the constraints in (3.4), a non-negativity condition on (4.1), as used in [3], is sufficient:

$$k_m \ge 0. \tag{4.2}$$

However, such a condition implies that K is *completely monotone* and thus extremely smooth. As seen in the comparison between Fig. 4b and Fig. 5, the following condition is more flexible and is derived in Appendix B:

$$\boldsymbol{k}^{\mathrm{T}}[\Lambda D_1^{-1} \cdots D_{M-1}^{-1}] \ge 0 \tag{4.3}$$

where $\boldsymbol{k} = \langle k_1, \dots, k_M \rangle^{\mathrm{T}}$, $\Lambda = \mathrm{diag}\{\lambda_1, \dots, \lambda_M\}$ and:

$$D_m = \mathrm{tridiag}\left\{\begin{bmatrix} -\Delta_{m+1}^{1} & -\Delta_{m+2}^{2} & \cdots & -\Delta_{M}^{M-m} & 0 & \cdots & 0 & - \\ \Delta_{m+1}^{1} & \Delta_{m+2}^{2} & \cdots & \Delta_{M}^{M-m} & 1 & \cdots & 1 & 1 \\ - & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}^{\mathrm{T}}\right\}, \quad \Delta_i^j = (\lambda_i - \lambda_j)^{-1}. \tag{4.4}$$

That (4.3) is a sufficient condition for the constraints in (3.4) is established in Theorem 3, which is proved in Appendix B.

Theorem 3 *Assume that the values $\{\lambda_m\}_{m=1}^{M}$ are positive and distinct. Then under the condition (4.3), the function $K(t)$ in (4.1) is non-negative and non-increasing.*

The condition (4.3) on the coefficients $\{k_m\}$ is sufficient for the constraints of (3.4) but not necessary. An alternative pointwise constraint is also considered in [10] which is necessary but not sufficient for monotonicity, and the condition (4.3) is found in [10] to be superior.

The discretization of (3.4) using (4.1) and (4.3) leads to a linearly constrained least squares problem which can be solved readily using `lsqlin` in MATLAB; see [5]. The kernel so estimated and shown in Fig. 7a is representative of the results obtained in this work by deconvolution with an exponential basis. Note that in the experiments of Figs. 4 – 7, artificial data are used as discussed in Subsection 4.1, and the data are corrupted with Gaussian noise having a standard deviation of 10% of $\max\{C_{\mathrm{T}}\}$.

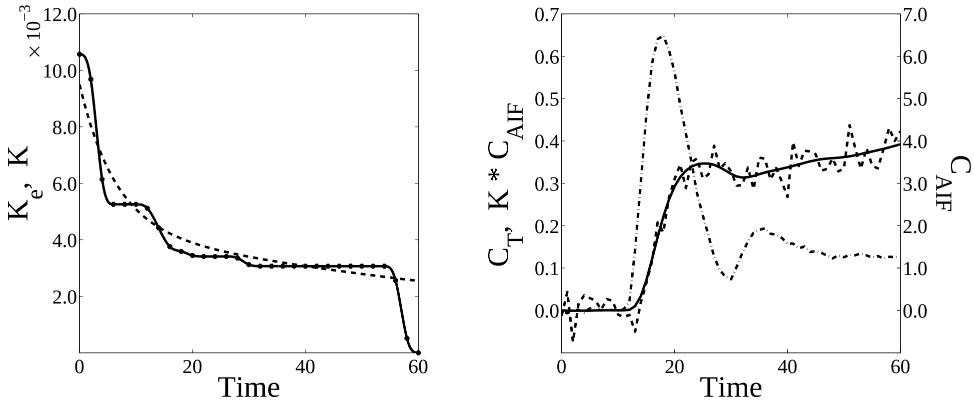

Figure 2: Results of solving (3.4) with $\mathcal{C}^2$ piecewise cubic splines. The dashed curve on the left is the exact kernel K_{e} which is bi-exponential. The solid curve on the left is the kernel estimation K. On the right the dash-dotted curve is C_{AIF} and the dashed curve is C_{T}. The solid curve on the right is the convolution $K * C_{\mathrm{AIF}}$. Note the stairsteps in the estimation kernel K.

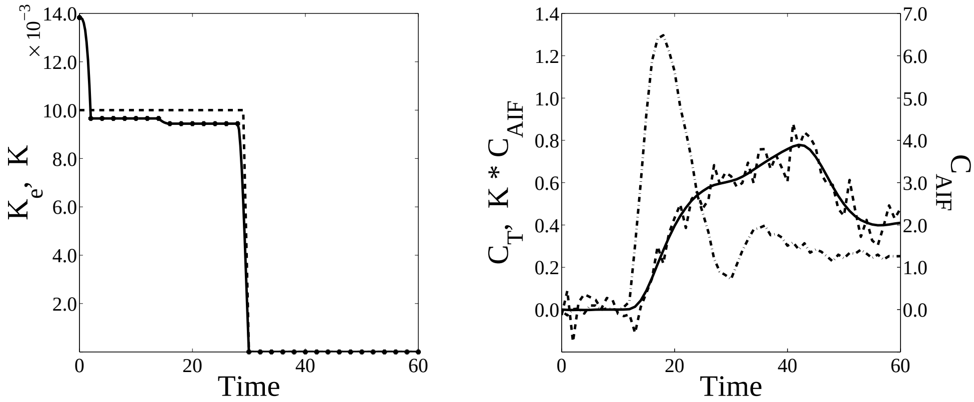

Figure 3: $\mathcal{C}^0$ Cubic splines are used to solve (3.4). The data format is the same as in Fig. 2, but here K_{e} is a step function. Note that the estimations K exhibit a sensitivity to initial noise but a relative absence of stairsteps in relation to Fig. 2.

4.3 The Staircasing Effect

Figure 2 shows the result of solving (3.4) with B given by piecewise $\mathcal{C}^2$ cubic spline basis functions [20] and the exact kernel is bi-exponential corresponding to two-compartment transport. Note that the estimated kernel in Fig. 2 manifests the staircasing effect in which a function with gradual variations is estimated with a sequence of stairsteps. On the other hand, Fig. 3 shows the result of solving (3.4) with B given by piecewise $\mathcal{C}^0$ cubic spline basis functions [20] and the exact kernel is a step function corresponding to a plug flow. Complete details of the calculations shown in Figs. 2 and 3 are given in [10]. Although the sensitivity of (2.1) to initial noise (cf. Section 3) is particularly evident in Fig. 3, the contrast between Figs. 2 and 3 demonstrates generally that estimations constrained by a bound on total variation achieve better reconstructions of apparently piecewise constant or *blocky* data than of functions with gradual variations.

In Fig. 4, it is seen that sums of exponentials approximate step functions poorly, and thus they resist manifesting the staircasing effect. It is also seen from Fig. 4 that exponential functions are not an appropriate basis for a kernel modeling plug flow; see also [16]. In each example shown in Fig. 4, $M = 10$. In addition, the dash-dotted curve shown in the left column for every example is $0.01 \cdot E_n(t)$ with $n = 100$, where $E_n(t) = e^{-2nt/T} \sum_{m=0}^{n} (2nt/T)^m/m!$ is the Erlang function, which is the residue function for the compartment network shown on the left in Fig. 6 with $\kappa = 2n/T$. Note that E_{100} also shows the approximating capacity of (4.1) with $M = 100$

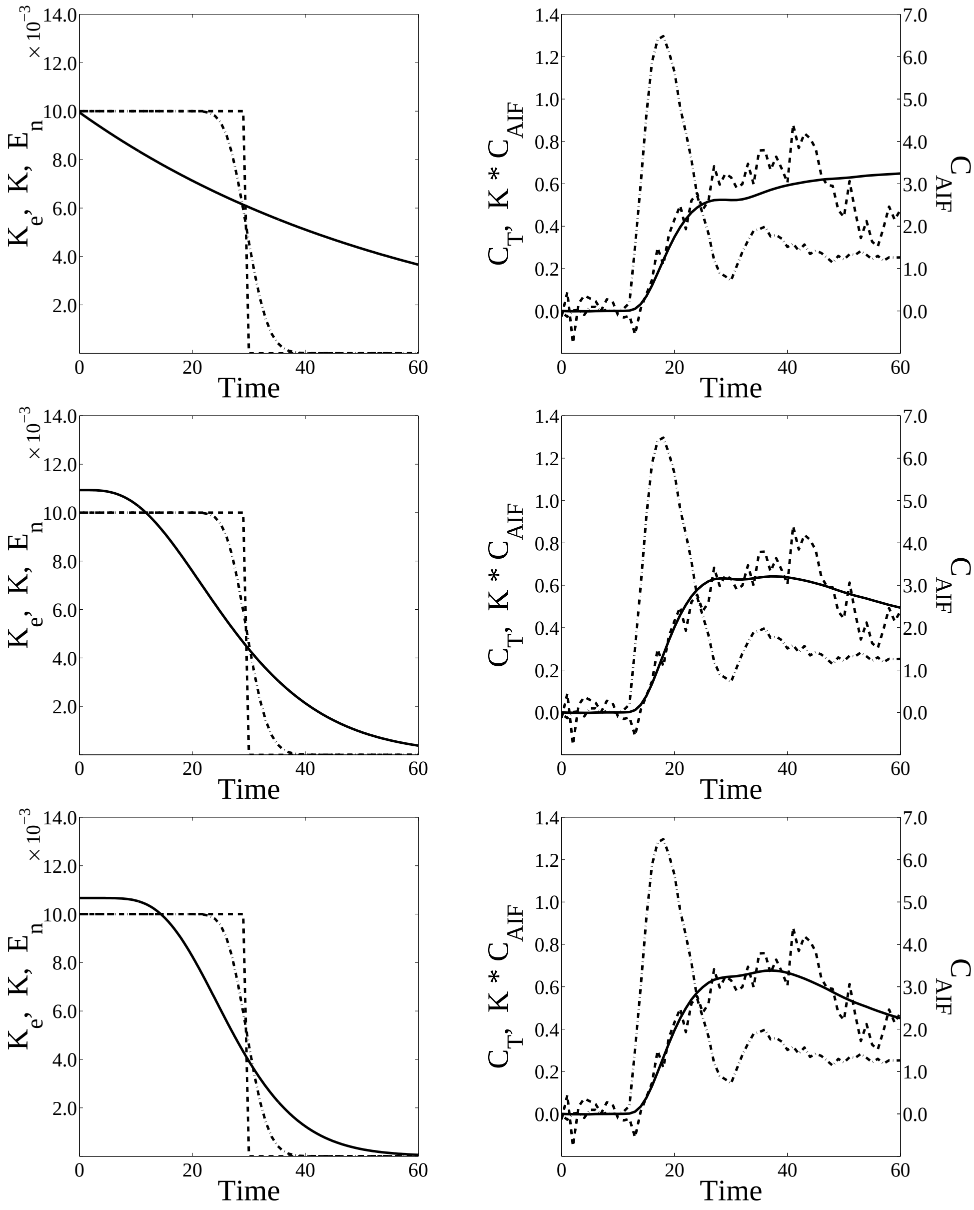

Figure 4: Results of solving (3.4) with $M = 10$ in (4.1) under the constraint (4.3). (a) In the first row, the time scales are distributed uniformly with $\lambda_m^{-1} = mT/M$. (b) In the second row, the time scales are distributed harmonically with $\lambda_m^{-1} = T/m$. (c) In the third row, the time scales are selected by a polytope search to minimize the constrained residual. The data format is the same as in Fig. 2, but here K_{e} is a step function. Also, the dash-dotted curve in the left column is the step approximating Erlang function $0.01 \cdot E_n$ with $n = 100$. Note that the harmonically distributed time scales have better approximation properties than the uniformly distributed time scales, and that the polytope search provides only a marginal improvement.

since a Trotter-Kato perturbation of the Erlang function can be made to form an arbitrarily near sum of exponentials with the same number of terms [17]. Clearly a great many exponential functions are required to approximate the step, whereas splines can approximate a step function much more readily.

Although the sums of exponentials shown in Fig. 4 are very smooth, Fig. 5 shows that when (4.1) is constrained by (4.2), the resulting completely monotone function is extremely smooth in relation to its more flexible counterpart constrained by (4.3). Specifically, for Fig. 5 $\|K_\mathrm{e} - K\|^2_{L_2[0,T]}$ is minimized over functions K in (4.1) under the constraints (4.2) and (4.3) for two different exact target functions K_e. In both cases $M = 10$ in (4.1) and the time scales are distributed harmonically with $\lambda_m^{-1} = T/m$. In Fig. 5a, K_e is a step function, which corresponds to the residue function for the compartment network shown on the left in Fig. 6, with $\kappa = 2n/T$ and $n \to \infty$. In Fig. 5b, K_e is the residue function for the compartment network shown on the right in Fig. 6, with $\kappa = 1/6$ and $\rho = 1/12$. Clearly, the approximations obtained using the newly derived monotonicity constraint in (4.3) are more accurate than those obtained using (4.2).

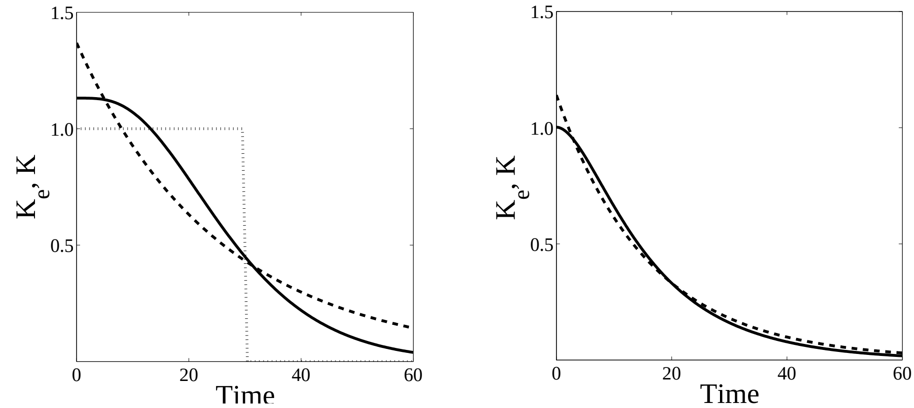

Figure 5: Results of minimizing $\|K_\mathrm{e} - K\|^2_{L_2[0,T]}$ under the constraint (4.2), shown with the dashed curve, and under the constraint (4.3), shown with the solid curve. In both cases, $M = 10$ in (4.1) and the time scales are distributed harmonically with $\lambda_m^{-1} = T/m$. The dotted curve is the exact target K_e which is a step function on the left (a) and lies beneath the solid curve on the right (b). Note that the completely monotone function constrained by (4.2) is extremely smooth in relation to its more flexible counterpart constrained by (4.3).

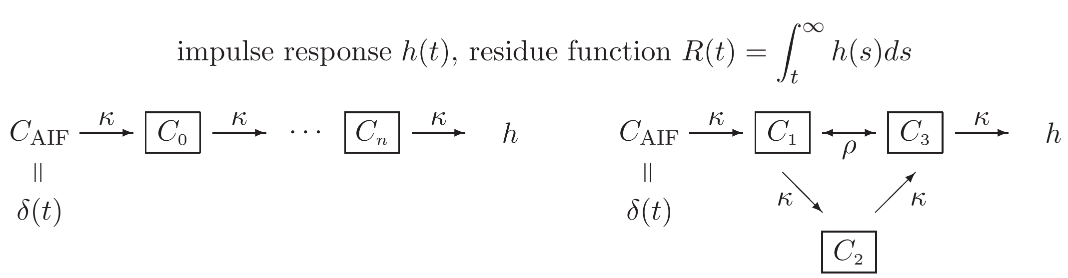

Figure 6: The compartment network on the left is a convective chain with concentrations satisfying $C_i' = \kappa(C_{i-1} - C_i)$, $C_{-1} = C_\mathrm{AIF}$. The concentrations of the convective and diffusive network on the right satisfy $C_1' = \kappa(C_\mathrm{AIF} - C_1) + \rho(C_3 - C_1)$, $C_2' = \kappa(C_1 - C_2)$, $C_3' = \kappa(C_2 - C_3) + \rho(C_1 - C_3)$.

4.4 Selecting the Placement of Time Scales

In Fig. 4a, the exponential time scales are chosen to be uniformly distributed according to $\lambda_m^{-1} = mT/M$. In Fig. 4b the time scales are distributed harmonically according to $\lambda_m^{-1} = T/m$.

In Fig. 4c the time scales are selected by a polytope search [5] to minimize the constrained residual $\|C_{\rm T} - K * C_{\rm AIF}\|_{L_2[0,T]}$ with respect to $\{\lambda_m\}$. Notice in Fig. 4a how poorly the shape and area (mean transit time) are estimated but how well the initial value (volumetric flow rate per unit tissue volume) is estimated. Notice too that the polytope search offers only marginal improvements over the harmonic time scales.

On the basis of the result shown in Fig. 4c, a gradient based optimization was also considered for the exponential basis. However, further simple tests with the polytope search reveal that the additional freedom afforded by time scale flexibility can be destabilizing. Specifically, Fig. 7 shows results of solving (3.4) for a bi-exponential kernel, where the effect of a harmonic distribution of time scales $\lambda_m^{-1} = T/m$ is shown in Fig. 7a and the effect of using the polytope search is shown in Fig. 7b. Note the over-estimation of the initial value (volumetric flow rate per unit tissue volume) in Fig. 7b. The improvement evident in Fig. 7a shows that the harmonic distribution is regularizing. At the same time, as explained in connection with Fig. 4, the harmonic distribution provides good approximation properties in relation to more uniformly distributed time scales. These computationally observed results are naturally a verification of the theoretical conditions of the Müntz Theorem 2, which the harmonically distributed time scales satisfy. Thus, time scales are henceforth assumed to be harmonically distributed.

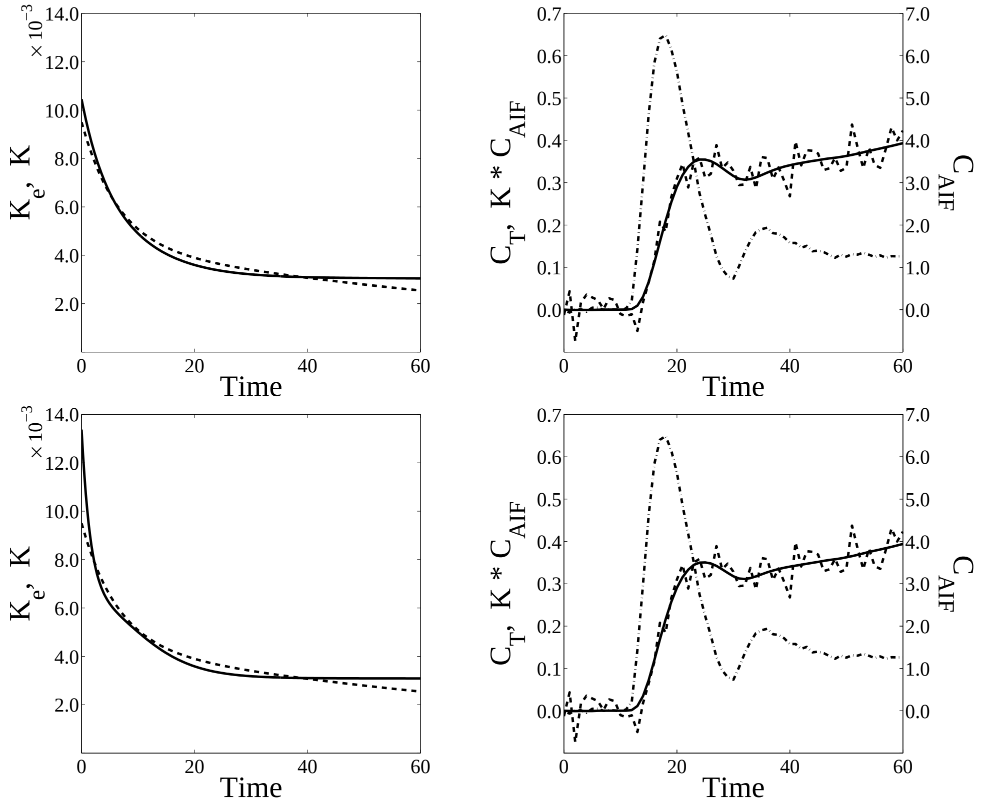

Figure 7: Results of solving (3.4) with (4.1), where $M = 12$ is chosen by the L-curve criterion in Fig. 8. (a) In the first row, the time scales are distributed harmonically with $\lambda_m^{-1} = T/m$. (b) In the second row, the time scales are selected by a polytope search to minimize the constrained residual. The data format is the same as in Fig. 2 and $K_{\rm e}$ is a bi-exponential. Note from the second row that the additional freedom afforded by time scale flexibility is destabilizing.

4.5 Selecting the Number of Time Scales

The over-estimation of the initial value in Fig. 7b is also typical of results achieved with a large number of times scales, even when they are harmonically distributed. To determine the appropriate number of time scales, the L-curve approach is adapted from the penalized least squares context [4]. Specifically, a plot of $\log\|C_\mathrm{T} - K * C_\mathrm{AIF}\|_{L_2[0,T]}$ versus $\log\|K'\|_{L_2[0,T]}$ is shown in Fig. 8 for varying numbers of harmonically distributed time scales. The chosen number of times scales is that associated with the first point where a local minimum is detected in the angle

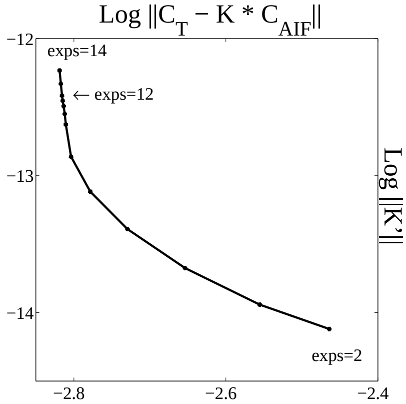

Figure 8: The L-curve is generated by varying the number of exponentials and plotting $\log\|K'\|_{L_2[0,T]}$ versus $\log\|C_\mathrm{T} - K * C_\mathrm{AIF}\|_{L_2[0,T]}$. The chosen number is that associated with the first point where a local minimum is detected in the angle (4.5) subtended at a vertex in the L-curve.

$$\alpha_k = \arccos\left[\frac{(\boldsymbol{\xi}_{k+1} - \boldsymbol{\xi}_k)\cdot(\boldsymbol{\xi}_{k-1} - \boldsymbol{\xi}_k)}{\|\boldsymbol{\xi}_{k+1} - \boldsymbol{\xi}_k\|\|\boldsymbol{\xi}_{k-1} - \boldsymbol{\xi}_k\|}\right], \quad \boldsymbol{\xi}_k = (\log\|C_\mathrm{T} - K * C_\mathrm{AIF}\|_{L_2[0,T]}, \log\|K'\|_{L_2[0,T]})_k. \tag{4.5}$$

subtended at a vertex in the L-curve. The L-curve shown in Fig. 8 was used to determine the number of time scales in Fig. 7. Note that for typical physiological data, the number of time scales required for an accurate representation of the kernel is apparently rather small, and therefore its calculation by the given heuristic rule is not expensive.

4.6 Application to Measured Data

The exponential deconvolution technique is now compared qualitatively with alternative methods in Fig. 9. The data used for the comparison were measured as discussed in Subsection 4.1, and thus there is no quantitative measure of accuracy. The kernel estimations shown here were obtained respectively by Truncated Singular Value Decomposition (TSVD), by Tikhonov regularization, and by constrained $\mathcal{C}^2$ cubic splines [10] as well as by the constrained exponentials. For TSVD, the singular values below 10% of the maximum were truncated, and this threshold was observed to reduce extreme oscillations without being overly smoothing, although comparable results are seen in [2] with a threshold of 20%. Nevertheless, the TSVD estimation still manifests significant oscillations while the residual $C_\mathrm{T} - K * C_\mathrm{AIF}$ is slightly smaller than the other residuals among the more regularized methods compared in Fig. 9. These other residuals are quite similar, and this similarity among residuals is also observed among Figs. 2 and 3. Such results are of course precisely the computational realization of the prediction represented in (3.2) that residuals will be indistinguishable for greatly different kernels, and more so without regularization. Note further that the regularization present in all methods prevents to match the steep ascent in the data, which could otherwise be achieved with unconstrained and sufficient freedom in (3.4). The regularization used for Tikhonov regularization was just enough to avoid the oscillations produced by TSVD, and yet there is significant smoothing in the solution at initial and later times; see also [2]. These difficulties do not appear in the constrained spline estimation or in the constrained exponential estimation. Since the underlying kernel in the data of Fig. 9 appears from these estimations to be roughly bi-exponential, the constrained exponential functions provide an apparently more natural kernel estimation. Note however that the

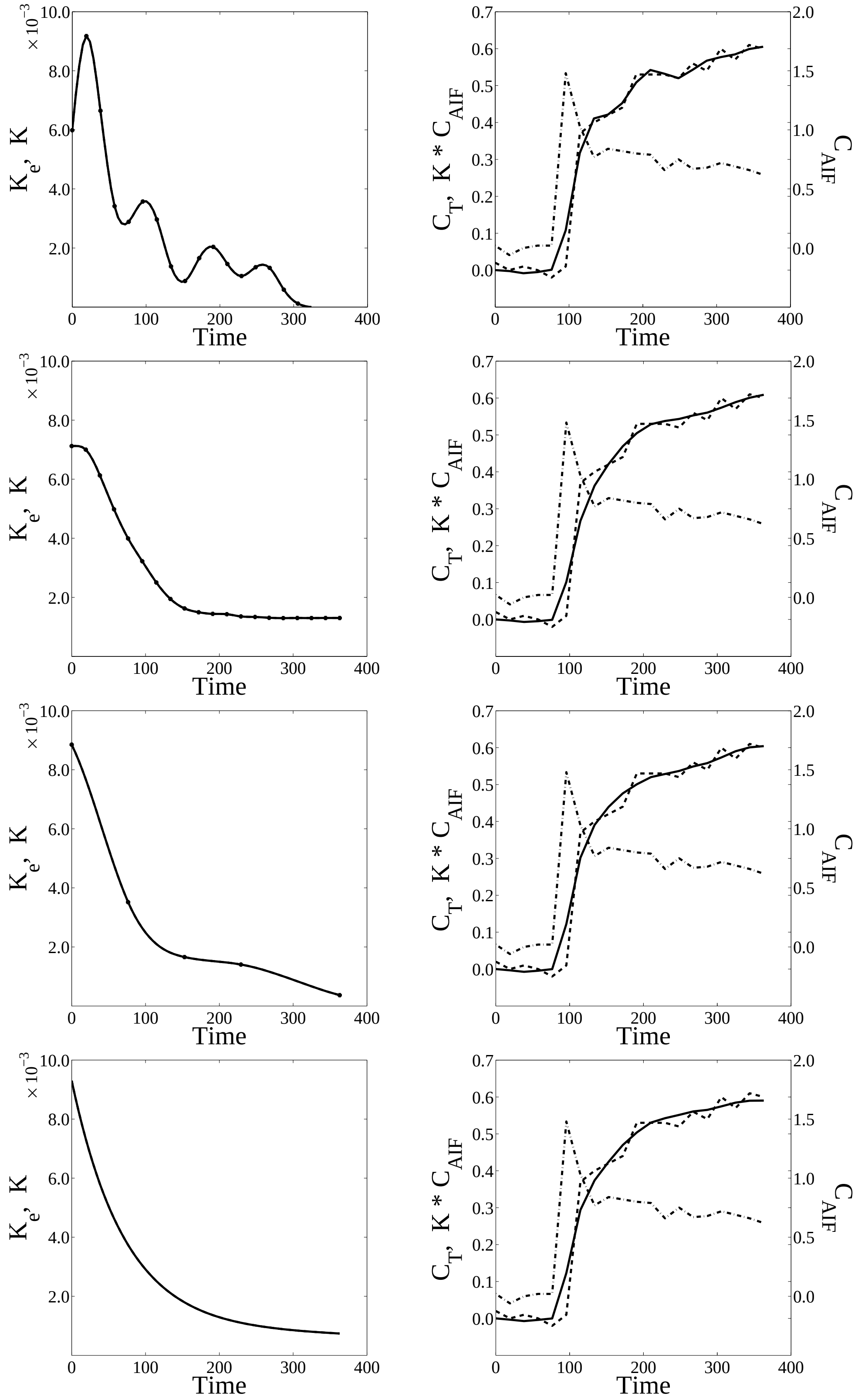

Figure 9: Results of solving (2.1) with Truncated Singular Value Decomposition (TSVD), Tikhonov regularization, constrained splines [10], and constrained exponentials are shown respectively in the (a) first, (b) second, (c) third and (d) fourth rows. The data format is the same as in Fig. 2. Note that TSVD exhibits significant oscillations and that Tikhonov regularization exhibits significant smoothing at initial and later times. The constrained splines and exponentials do not exhibit these difficulties, and the constrained exponentials apparently provide the most natural kernel estimation.

estimated sum in (4.1) is not strictly bi-exponential since the computed coefficients are given by $\boldsymbol{k} = 10^{-3}\langle 3.8, -7.6, 7.6, -3.8, 9.3\rangle^{\mathrm{T}}$. On the basis of the strongest contrast seen in Fig. 9 between the exponential approach and the standard TSVD approach, these methods are next compared by applying them to a challenging DCE-MRI sequence to determine physiological parameters pixelwise.

According to (4.1) the effective volume fraction in (2.4) is given by:

$$\mathcal{V}_{\mathrm{T}} = \int_0^\infty K(t)dt = \boldsymbol{k}^{\mathrm{T}}\boldsymbol{\lambda}^{-1} \tag{4.6}$$

the volumetric flow rate per unit tissue volume in (2.2) by:

$$\mathcal{F}_{\mathrm{T}} = K(0) = \boldsymbol{k}^{\mathrm{T}}\boldsymbol{e} \tag{4.7}$$

and the mean transit time in (2.3) by:

$$\mathcal{T}_{\mathrm{T}} = \int_0^\infty \frac{K(t)}{K(0)}dt = \frac{\boldsymbol{k}^{\mathrm{T}}\boldsymbol{\lambda}^{-1}}{\boldsymbol{k}^{\mathrm{T}}\boldsymbol{e}}. \tag{4.8}$$

The images in Fig. 10 show the pixelwise computed parameters derived from the image sequence of Fig. 1. For these calculations, the concentration in a major artery was used for C_{AIF}. Then each single pixel was chosen as a tissue region Ω, the L-curve procedure was used to determine the number of harmonically distributed time scales, and the exponential deconvolution method was used to estimate a kernel for the chosen pixel. In the top row of Fig. 10, the pixelwise computed parameters $\mathcal{V}_{\mathrm{T}}$, $\mathcal{F}_{\mathrm{T}}$, and $\mathcal{T}_{\mathrm{T}}$ are shown from left to right, as computed according to (4.6), (4.7), and (4.8), respectively. In the bottom row of Fig. 10, the same parameters are shown as computed according to the standard TSVD approach. In all images, a cerebral tumor is evident, particularly in the upper portion of the images. Note that this test with measured data is particularly challenging because of the low signal-to-noise ratio in the image sequence of Fig. 1. The visibility of affected tissues can be improved further by a preprocessing approach, as reported in [18] and [19], which reduces the influence of macro-vessels and image artifacts, and further such work will be reported separately. In the present results, it is remarkable that

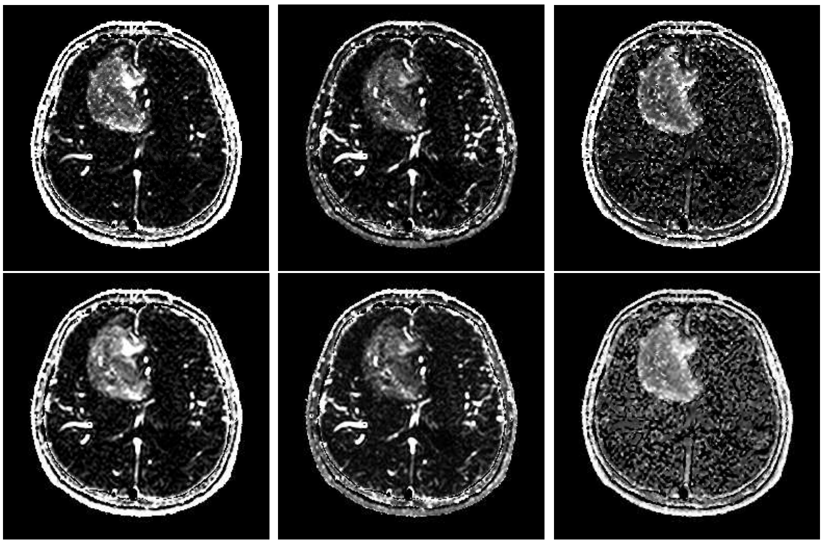

Figure 10: For the image sequence of Fig. 1, a convolution kernel and associated physiological parameters are computed for each pixel using constrained exponentials (top row) and Truncated Singular Value Decomposition (bottom row). The pixelwise computed parameters $\mathcal{V}_{\mathrm{T}}$, $\mathcal{F}_{\mathrm{T}}$, and $\mathcal{T}_{\mathrm{T}}$ are shown, respectively, in the left column, in the middle column, and in the right column. Note the cerebral tumor which is particularly evident in the upper portion of the images.

the obvious superiority of constrained exponentials in relation to TSVD, apparent in the curves of Fig. 9, is not more pronounced in the images of Fig. 10. Nevertheless, the images do manifest the following surprising spatial correlation, in spite of there being no explicit spatial coupling among the deconvolutions in either approach. The images obtained by the exponential approach manifest a sharpness typical of images regularized through spatial total variation, while the images obtained by TSVD manifest a slight blur typical of images spatially regularized through Gaussian filtering; see [9] for further details of such filtering comparisons. This apparent spatial correlation will be studied further in relation to the theory developed in [6].

5 Summary and Conclusions

Prior to the present work, transport kernel estimation had been investigated using filtered Discrete Fourier Transform [16], Truncated Singular Value Decomposition [2] [16], Tikhonov Regularization [2], penalized monotonicity [16], and constrained monotonicity [7]. In this paper a deconvolution approach for DCE-MRI was developed using an approximation basis of exponential functions constrained to be non-negative and non-increasing. Under these constraints it is seen that kernel estimates have bounded variation. Additional regularization was implemented in terms of the numbers of basis functions and the distribution of their parameters. Monotonicity in an exponential basis was implemented in terms of a newly derived condition which is a considerable generalization over the non-negative coefficients constraint of [3] which implies complete monotonicity [23].

The methods developed were compared directly with the widely used methods of Tikhonov Regularization [2], monotonicity constrained splines [10] and particularly Truncated Singular Value Decomposition [2] [16]. Also, the staircasing effect seen in [7] and in Fig. 2 was eliminated by the regularized exponential functions used for Fig. 7a. Fig. 4 shows that exponential functions naturally resist forming steps. Also, comparing Fig. 4 with Fig. 3 shows that splines are more appropriate for estimating kernels associated with plug flows. On the other hand, Figs. 2 and 7a as well as Fig. 9 indicate that exponential functions are better suited to typical physiological data. The direct comparison among TSVD, Tikhonov Regularization, constrained splines and constrained exponential functions shown in Fig. 9 reveals excessive oscillations in the TSVD estimation, excessive flattening in the Tikhonov estimation, a slightly unnatural concavity in the spline estimation, but an apparently natural estimation with exponential functions. On the basis of this comparison it was expected that the approach with constrained exponential functions would provide a considerable improvement over TSVD in Fig. 10. This improvement is as noticeable as that obtained from TV as opposed to Gaussian regularization for image reconstruction [9]; however, in light of Fig. 9, it is remarkable that the improvement from constrained exponentials in relation to TSVD seen in Fig. 10 is not more pronounced. Also remarkable is the finding in Fig. 10 that the images obtained by constrained exponentials manifest more spatial sharpness than the images obtained by TSVD, in spite of there being no explicit spatial coupling in either approach among the deconvolutions performed pixelwise. Further study of this implicit spatial correlation will be reported separately. Also, as discussed in more detail in [11], instead of integrating a convection-diffusion equation to obtain the convolution model, the transport properties may be estimated directly from the convection-diffusion equation, and this approach to DCE-MRI will be investigated separately.

A Proof of Theorem 1

It will be shown that for $n > 1$ (3.2) is satisfied by the following:

$$N_\varepsilon(t) = \varepsilon \Re \left[\nu^{n+1} \hat{\imath}^n \sum_{m=0}^{n} \begin{pmatrix} n \\ m \end{pmatrix} (-1)^m \frac{e^{\hat{\imath}(m+1)t/\varepsilon} - e^{-\nu t}}{\varepsilon\nu + \hat{\imath}(m+1)} \right], \quad \max_{t \geq 0} |N_\varepsilon(t)| = \mathcal{O}(\varepsilon), \tag{A.1}$$

$$E_\varepsilon(t) = \varepsilon^{-n}\Re\left[\sum_{m=0}^{n}\binom{n}{m}(-1)^m(\varepsilon\nu\hat{\imath}-(m+1))^n e^{\hat{\imath}(m+1)t/\varepsilon}\right], \quad \max_{t\geq 0}|E_\varepsilon(t)| = \mathcal{O}(\varepsilon^{-n}) \tag{A.2}$$

where $\Re$ denotes the real part and $\hat{\imath} = \sqrt{-1}$. Based on the partial integration relation,

$$\int_0^t f(s)D_s^n g(s)ds = \sum_{m=0}^{n-1} D_s^m f(s) D_s^{n-1-m} g(s)\Big|_{s=0}^{s=t} + (-1)^n \int_0^t D_s^n f(s) g(s) ds \tag{A.3}$$

the convolution,

$$\int_0^t C_{\mathrm{AIF}}(t-s)E_\varepsilon(s)ds = \nu^{n+1}(-1)^n e^{-\nu t}\int_0^t \frac{(s-t)^n}{n!}e^{\nu s}E_\varepsilon(s)ds \tag{A.4}$$

is simplified by defining functions of the form $f(s) = (s-t)^n/n!$ and $e^{\nu s}E_\varepsilon(s) = D_s^n g(s)$ for which:

$$\sum_{m=0}^{n-1} D_s^m f(s) D_s^{n-1-m} g(s) = 0, \qquad s = 0, t. \tag{A.5}$$

The sum in (A.5) vanishes for $s = t$ because $D_s^m f(s) = (s-t)^{(n-m)}/(n-m)!$ vanishes at $s = t$ for $0 \leq m \leq (n-1)$. For the sum in (A.5) to vanish for $s = 0$, $g(s)$ is constructed to have sufficient degrees of freedom:

$$g(s) = \Re\left[\sum_{m=0}^{n}\alpha_m e^{(\nu+\hat{\imath}\beta_m)s}\right], \quad E_\varepsilon(s) = e^{-\nu s}D_s^n g(s) = \Re\left[\sum_{m=0}^{n}\alpha_m(\nu+\hat{\imath}\beta_m s)^n e^{\hat{\imath}\beta_m s}\right] \tag{A.6}$$

so that each of the following conditions hold:

$$D_s^k g(s)|_{s=0} = \Re\left[\sum_{m=0}^{n}\alpha_m(\nu+\hat{\imath}\beta_m)^k\right] = 0, \quad k = 0,\ldots,n-1. \tag{A.7}$$

On the basis of the example cited after (3.2) with fundamental frequency $1/\varepsilon$, the modes β_m in E_ε are selected according to:

$$\beta_m = \frac{m+1}{\varepsilon}, \qquad \alpha_m = \binom{n}{m}(-1)^m\hat{\imath}^n, \qquad m = 0,\ldots,n-1 \tag{A.8}$$

where the coefficients $\{\alpha_m\}$ in (A.8) can be shown to solve (A.7) by expanding $(\nu + \hat{\imath}\beta_m)^k$ binomially and observing that:

$$0 = D_x^l(1-e^x)^n|_{x=0} = D_x^l\sum_{m=0}^{n}\binom{n}{m}(-e^x)^m|_{x=0} = \sum_{m=0}^{n}\binom{n}{m}(-1)^m m^l, \quad l = 0,\ldots,n-1. \tag{A.9}$$

Thus, E_ε is given through (A.8) and (A.6) by the formula in (A.2). Since (A.5) is satisfied, N_ε is given through (A.4) by (A.3):

$$\begin{aligned} N_\varepsilon(t) &= \int_0^t C_{\mathrm{AIF}}(t-s)E_\varepsilon(s)ds = \nu^{n+1}(-1)^n e^{-\nu t}(-1)^n\int_0^t D_s^n f(s)g(s)ds \\ &= \nu^{n+1}e^{-\nu t}\int_0^t \Re\left[\sum_{m=0}^{n}\alpha_m e^{(\nu+\hat{\imath}\beta_m)s}\right]ds = \nu^{n+1}\Re\left[\sum_{m=0}^{n}\alpha_m\frac{e^{\hat{\imath}\beta_m t}-e^{-\nu t}}{(\nu+\hat{\imath}\beta_m)}\right]. \end{aligned} \tag{A.10}$$

Thus, by (A.8) and (A.10), N_ε is given by the formula in (A.1).

The norm estimates shown in (A.1) and (A.2) are established as follows. First, consider N_ε. Since $\nu > 0$ and $(m+1) \geq 1$ hold, it follows for $\varepsilon > 0$ that $|(-1)^m(e^{\hat{\imath}(m+1)t/\varepsilon} - e^{\nu t})/(\varepsilon\nu + \hat{\imath}(m+1))| \leq 2$. The remaining binomial sum in (A.1) is $(1+1)^n$ and thus N_ε is majorized by $\varepsilon(2\nu)^{n+1}$, and the estimate follows. Finally, consider E_ε. For any integer l, $t_l = 2\pi l\varepsilon$ gives $e^{\hat{\imath}(m+1)t_l/\varepsilon} = 1$. Then, expanding $(\varepsilon\nu + \hat{\imath}(m+1))^n$ binomially and using (A.9) shows that the sum in (A.2) reduces to $E_\varepsilon(t_l) = \varepsilon^{-n}n!$, and the estimate follows. ■

B Proof of Theorem 3

Note that the kernel in (4.1) may be posed generally as the Laplace Transform of a signed measure [23]:

$$K(t) = \int_0^\infty e^{-xt} d\mu(x), \qquad \text{e.g.,} \quad \mu'(x) = \sum_{m=1}^{M} k_m \delta(x - \lambda_m). \tag{B.11}$$

To guarantee that K is non-increasing, the measure μ might be constrained to be non-decreasing so that the following holds:

$$-K'(t) = \int_0^\infty x e^{-xt} \underbrace{d\mu(x)}_{\geq 0} \geq 0, \qquad \text{e.g.,} \quad \mu'(x) = \sum_{m=1}^{M} \underbrace{k_m}_{\geq 0} \delta(x - \lambda_m) \geq 0. \tag{B.12}$$

Such a non-negative coefficient condition of (4.2) is used in [3]. However, (4.2) implies that K is *completely monotone* [23]:

$$(-1)^n K^{(n)}(t) = \int_0^\infty x^n e^{-xt} d\mu(x) \geq 0 \tag{B.13}$$

and thus extremely smooth. On the other hand, partial integration leads to an alternative condition:

$$-K'(t) = t^n \int_0^\infty dx_n \exp(-x_n t) \underbrace{\int_0^{x_n} dx_{n-1} \int_0^{x_{n-1}} dx_{n-2} \cdots \int_0^{x_1} d\mu(x_0) x_0}_{\geq 0} \geq 0 \tag{B.14}$$

and this observation is the basis for its discrete counterpart derived below.

Assume that μ in (B.11) is concentrated at discrete, positive values $\{\lambda_m\}_{m=1}^M$ which are distinct and numbered in ascending order so that (4.1) holds and $-K'(t) = \boldsymbol{k}^{\mathrm{T}} \Lambda \exp(-\boldsymbol{\lambda} t)$ where $\boldsymbol{k} = \langle k_1, \ldots, k_M \rangle^{\mathrm{T}}$, $\Lambda = \operatorname{diag}\{\lambda_1, \ldots, \lambda_M\}$, and $\exp(-\boldsymbol{\lambda} t) = \langle e^{-\lambda_1 t}, \ldots, e^{-\lambda_M t} \rangle^{\mathrm{T}}$. Analogous to (B.14), a condition to ensure $-K'(t) \geq 0$ is derived by employing discrete partial integration according to $-K'(t) = \boldsymbol{k}^{\mathrm{T}}[\Lambda D_1^{-1} \cdots D_{M-1}^{-1}][D_{M-1} \cdots D_1] \exp(-\boldsymbol{\lambda} t)$ where multiplication by the matrices in (4.4) gives the divided differences of $e_t(\lambda) = e^{-\lambda t}$,

$$e_t[\lambda_m] = e_t(\lambda_m), \qquad e_t[\lambda_m, \ldots, \lambda_{m+k}] = \frac{e_t[\lambda_{m+1}, \ldots, \lambda_{m+k}] - e_t[\lambda_m, \ldots, \lambda_{m+k-1}]}{\lambda_{m+k} - \lambda_m}, \tag{B.15}$$

as follows:

$$D_{M-1} \cdots D_1 \exp(-\boldsymbol{\lambda} t) = \{(-1)^{M-m} e_t[\lambda_m, \ldots, \lambda_M]\}_{m=1}^M. \tag{B.16}$$

Assuming that these discrete derivatives are non-negative leads to the discretization (4.3) of the constraints in (3.4) in terms of discrete integrations of the coefficients in (4.1). That (4.3) is a sufficient condition for the constraints in (3.4) is established as follows.

By the Mean Value Theorem for divided differences, the divided differences in (B.16) satisfy $(-1)^{M-m} e_t[\lambda_m, \ldots, \lambda_M] = (-1)^{M-m} e_t^{(M-m)}(\xi_m)/(M-m)! = t^{M-m} e^{-\xi_m t}/(M-m)! > 0$ with $\xi_m \in (\lambda_m, \lambda_M)$ [15], and thus each component of the right side of (B.16) is positive. Therefore, under the condition (4.3), $-K'(t) = \boldsymbol{k}^{\mathrm{T}}[\Lambda D_1^{-1} \cdots D_{M-1}^{-1}][D_{M-1} \cdots D_1] \exp(-\boldsymbol{\lambda} t) \geq 0$ holds, and $K(t)$ is non-increasing. Since the exponents $\{\lambda_m\}$ are assumed to be positive, it follows from (4.1) that $K(\infty) = 0$. Since $K(t)$ is non-increasing, $K(t) \geq K(\infty) = 0$ holds, and K is non-negative. ■

References

[1] R.B. Ash, *Real Analysis and Probability*, Academic Press, New York, 1972.

[2] F. Calamante, D.G. Gadian and A. Connelly, *Quantification of Bolus Tracking MRI: Improved Characterization of the Tissue Residue Function using Tikhonov Regularisation*, Magn. Reson. Med., Vol. 50, No. 6, pp. 1237–1247, 2003.

[3] V.J. Cunningham and T. Jones, *Spectral Analysis of Dynamic PET Studies*, J. Cereb. Blood Flow Metab, Vol. 13, No. 1, pp. 15–23, 1993.

[4] H.W. Engl, M. Hanke and A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.

[5] P.E. Gill, W. Murray and M.H. Wright, *Practical Optimization*, Academic Press, London, UK, 1981.

[6] F. Greensite, *The Temporal Prior in Bioelectromagnetic Source Imaging Problems*, IEEE Trans. on Biomed. Eng., Vol. 50, No. 10, pp. 1152-1159, 2003.

[7] J. Griebel, S. Pahernik, R. Lucht, A. DeVries, K.-H. Englmeier, M. Dellian and G. Brix, *Perfusion and Permeability: Can Both Parameters Be Evaluated Separately from Dynamic MR Data?*, Proc. Intl. Soc. Mag. Reson. Med. 9, 2001.

[8] E. Henderson, J.S. Sykes, D. Drost, H.-J. Weinmann, B.K. Rutt and T.-Y. Lee, *Simultaneous MRI Measurement of Blood Flow, Blood Volume, and Capillary Permeability in Mammary Tumors Using Two Different Contrast Agents*, J. Magn. Reson. Imaging, Vol. 12, pp. 991–1003, 2000.

[9] S.L. Keeling, *Total Variation Based Convex Filters for Medical Imaging*, Appl. Math. and Comp., Vol. 139, pp. 101–119, 2003.

[10] S.L. Keeling, R. Bammer, T. Kogler and R. Stollberger, *On the Convolution Model of Dynamic Contrast Enhanced Magnetic Resonance Imaging and Nonparametric Deconvolution Approaches*, Special Research Center Report 298, University of Graz, Austria, 2004.

[11] S.L. Keeling, R. Bammer and R. Stollberger, *Revision of the Theory of Tracer Transport and the Convolution Model of Dynamic Contrast Enhanced Magnetic Resonance Imaging*, J. Math. Biology, Vol. 55, pp. 389–411, 2007.

[12] S.S. Kety, *The Theory and Applications of the Exchange of Inert Gas at the Lungs and Tissues*, Pharmacol. Rev. Vol. 3, pp. 1–41, 1951.

[13] K.S. St. Lawrence and T.Y. Lee, *An Adiabatic Approximation to the Tissue Homogeneity Model for Water Exchange in the Brain. I. Theoretical Derivation*, J. Cereb. Blood Flow Metab., Vol. 18, pp. 1365–1377, 1998.

[14] Z. Liang and P.C. Lauterbur, *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective*, IEEE Press, Picsataway, 2000.

[15] L.M. Milne-Thomson, *The Calculus of Finite Differences*, Chelsea, New York, 1981.

[16] L. Østergaard, R.M. Weisskoff, D.A. Chesler, C. Glydensted and B.R. Rosen, *High Resolution Measurement of Cerebral Blood Flow using Intravascular Tracer Bolus Passages. Part I: Mathematical Approach and Statistical Analysis*, Magn. Reson. Med., Vol. 36, pp. 715–725, 1996.

[17] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983.

[18] G. Reishofer, C. Enzinger, F. Fazekas, S.L. Keeling and R. Stollberger, *Improved Visibility of Hemodynamic Parameter Maps Using ICA Corrected DCE-MRI Time Series*, ISMRM Workshop on Quantitative Cerebral Perfusion Imaging Using MRI: A Technical Perspective, Venice, Italy, March 21–23, 2004.

[19] G. Reishofer, F. Fazekas, S.L. Keeling, C. Enzinger, J. Simbrunner, R. Stollberger, *Minimizing Macro Vessel Signal in Cerebral Perfusion Imaging Using Independent Component Analysis*, Magnetic Resonance in Medicine, Vol. 57, No. 2, pp. 278 – 288, 2007.

[20] L. Schumaker, *Spline Functions: Basic Theory*, Wiley, New York, 1981.

[21] L. Schwartz, *Etude des sommes d'exponentielles*, Hermann, Paris, 1959.

[22] P.S. Tofts, *Modeling Tracer Kinetics in Dynamic Gd-DTPA MR Imaging*, J. Magn. Reson. Imaging, Vol. 7, pp. 91–101, 1997.

[23] D.V. Widder, *The Laplace Transform*, Princeton Univ. Press, Princeton, NJ, 1966.