

# Ensemble Kalman filters for reliability estimation in perfusion inference

P. Zaspel

Departement Mathematik und Informatik  
Fachbereich Mathematik  
Universität Basel  
CH-4051 Basel

Preprint No. 2017-04  
April 2017

[www.math.unibas.ch](http://www.math.unibas.ch)

# Ensemble Kalman filters for reliability estimation in perfusion inference \*

Peter Zaspel<sup>†</sup>

---

**Abstract.** We consider the solution of inverse problems in dynamic contrast-enhanced imaging by means of Ensemble Kalman filters. Our quantity of interest is blood perfusion, i.e. blood flow rates in tissue. While existing approaches to compute blood perfusion parameters for given time series of radiological measurements mainly rely on deterministic, deconvolution-based methods, we aim at recovering probabilistic solution information for given noisy measurements. To this end, we model radiological image capturing as sequential data assimilation process and solve it by an Ensemble Kalman filter. Thereby, we recover deterministic results as ensemble-based mean and are able to compute reliability information such as probabilities for the perfusion to be in a given range. Our target application is the inference of blood perfusion parameters in the human brain. A numerical study shows promising results for artificial measurements generated by a Digital Perfusion Phantom.

**Key words.** Dynamic Contrast-Enhanced Imaging, Perfusion, Ensemble Kalman Filter, Inference

**AMS subject classifications.** 62M40, 60G35, 62P10

**1. Introduction.** Medical imaging by x-rays, *magnetic resonance imaging* (MRI) and *computed tomography* (CT) has considerably changed medical diagnosis throughout the last decades. Often, *contrast agents*, i.e. specific liquid chemicals, are injected into the patients blood circulation during the imaging process. This leads to *contrast-enhanced images* of higher contrast in some regions of the human body. In this work, we study inverse problems for a specific MRI or CT imaging task. That is, we aim at recovering a quantity of interest in medical imaging, which is derived by *dynamic contrast-enhanced* (DCE) imaging. In dynamic contrast-enhanced imaging, a time-dependent series of radiological images of a part of the patients body (e.g. the brain) is taken immediately after injecting a contrast agent into the patient's blood circulation. By observing the time-dependent concentration evolution of the contrast agent inside the patient's tissue, it is possible to recover information about the blood flow rates, i.e. the *perfusion*.

The image acquisition process delivers a time-discrete series of three-dimensional (space-discrete) concentration images  $\mathbf{c}$  of a part of the patient body. The actual perfusion evaluation is a post-processing step, being preceded by image de-noising and motion compensation. Currently, blood perfusion is computed independently per discrete tissue volume element, i.e. *voxel*, thus spatial information is mostly neglected. Variants of the *indicator-dilution theory* [3, 6, 18] describe the concentration of a contrast agent in tissue at a given point in time as the result of a convolution in time of the (known) time-dependent arterial or blood circulation inflow concentration  $c_{art}$  with an unknown tissue-dependent kernel function  $\mathbf{k}$ . *Blood*

---

\*Submitted to the editors DATE.

**Funding:** This work is funded by the Swiss National Science Foundation (SNF) under project number 407540\_167186. Furthermore, this research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

<sup>†</sup>University of Basel, Basel, Switzerland ([peter.zaspel@unibas.ch](mailto:peter.zaspel@unibas.ch)).

*perfusion* is computed as a weighted maximum or point evaluation of the unknown kernel function.

Current state of the art methods aim at recovering the unknown time-dependent kernel function for given discrete measurements of the contrast agent in tissue. The kernel function is either modeled as a parametrized analytic function [21, 12, 6] or discretized as a fully unknown function [3, 11, 6]. Then, most approaches rely on a deterministic reconstruction of the kernel function, involving the solution of a deconvolution problem with regularization.

One drawback of the use of motion compensation, de-noising and deterministic deconvolution lies in the loss of information on the *quality* of a computed solution. That is, probabilistic information about the measurement accuracy and errors in space and time with their influence on the exactness on the computed quantity of interest are neglected or even lost.

In this work, we propose an approach to infer perfusion in the discussed application case while keeping the probabilistic information on the solution. Thereby, we overcome the discussed drawback of knowledge loss. To achieve this, we model the inference problem as a sequential data assimilation problem: First, the unknown kernel function  $\mathbf{k}$  is described as an unknown system state, for which a predictive time-discrete stochastic system state model is introduced. In this model, the kernel function is represented as a random variable. Then, a time-discrete stochastic observation model describes the relationship of the current approximation of  $\mathbf{k}$  and the noisy measurements delivered by medical imaging. Finally, the well-known Ensemble Kalman filter (EnKF) [5, 19, 7, 4] is used to compute an ensemble-based approximation of the posterior probability density function (PDF) of  $\mathbf{k}$  given the system state model and the (noisy) measurements. Based on this PDF, means, cumulative distribution functions, etc. can be computed. The whole methodology is applied to (noisy) artificial measurement data generated by a *Digital Perfusion Phantom* [17, 13, 9], i.e. a forward model describing the mapping of perfusion information to medical images.

While the use of the Ensemble Kalman filter is wide-spread, its predictive quality is sometimes in question [15] and many extensions and alternatives have been developed [1, 2, 20, 22, 14, 16]. On the one hand, it relies on having a linear observation model. This will be the case for a major part of our application. On the other hand, all random fields are required to be Gaussian guaranteeing that the posterior PDF is Gaussian, too. We will make sure that this requirement is fulfilled in our artificial test cases. However, in real world, this might not be the case. Therefore, the use of e.g. ensemble particle filters, overcoming this limitation, is future work.

To the best of our knowledge, we consider the discussed work to be a new contribution to the field. Nevertheless, there has been previous work on the use of Ensemble Kalman filters in the application scenario. In [10], the authors concentrate on the introduction of a tissue model that includes space-dependent information. To achieve this, a blood flow model is combined with an EnKF. Preliminary results for this approach are given. In contrast, we focus here directly on the mathematical setting based on the indicator-dilution theory that is well-known and, therefore, well accepted by radiologist. Hence, our methodology is considered as an extension to the existing standard methodology introducing the opportunity to derive statistical information on the computed solution. In addition to the different objective compared to [10], we also perform a large number of parameter studies and convergence tests, which are crucial to understand the properties of the method.

This article is organized as follows. In [Section 2](#), we give a mathematical model for the radiological imaging and perfusion extraction mechanism. [Section 3](#) outlines our numerical approach based on sequential data assimilation using EnKF. Numerical results are given in [Section 4](#) while [Section 5](#) summarizes the discussed work.

**2. Modeling radiological imaging and perfusion extraction.** In the following, we start by giving an abstract model for the transport of contrast agent. Then, measurements by e.g. MRI are abstractly modeled. A concrete model for the contrast agent distribution is given by the indicator-dilution theory. Finally, our quantity of interest, i.e. blood perfusion, is introduced and the deterministic inference problem is summarized.

**2.1. Abstract model for contrast agent transport.** Shall  $\mathcal{D}_{tiss} \subset \mathbb{R}^3$  be the tissue domain in the human body for which we want to derive information by dynamic contrast-enhanced imaging. We study contrast agent transport / concentrations in a time interval  $[0, T] \subset \mathbb{R}$  with  $T$  being the final time. The inflow concentration of the contrast agent (at some arterial inlet) is a function

$$c_{art} : [0, T] \rightarrow \mathbb{R}_{\geq 0}.$$

The time-continuous contrast agent concentration in tissue can be modeled as a function

$$c_{tiss} : \mathcal{D}_{tiss} \times [0, T] \rightarrow \mathbb{R}_{\geq 0}.$$

Both are related to each other by an (unknown) operator  $\mathcal{B}$ , with

$$(1) \quad c_{tiss}(\cdot, t) = \mathcal{B}[c_{art}](t)$$

that models the function of the human body with respect to contrast agent transport.

**2.2. Measuring contrast agent concentration in tissue.** Appropriate measurement devices (CT, MRI, ...) usually have a cuboidal measurement domain. Therefore, we start by limiting  $\mathcal{D}_{tiss}$  to

$$\mathcal{D}_{meas} = \prod_{d=1}^3 [0, a_d]$$

with  $\mathbf{a} = (a_1, a_2, a_3)^\top \in \mathbb{R}^3$  describing the size of the measurement domain. For simplicity, we assume the measurement domain and the area of interest to match exactly, i.e.  $\mathcal{D}_{meas} = \mathcal{D}_{tiss}$ , excluding cases in which some part of the measurement domain does not contain valid tissue. Moreover,  $\mathcal{D}_{meas}$  is simplified as being stationary in time, i.e. the measurement device (or the patient) does not move or movements are considered as measurement error.

The finite spatial resolution  $\mathbf{N}_{\mathcal{D}} \in \mathbb{N}^3$  of the measurement device leads to a decomposition of  $\mathcal{D}_{meas}$  into  $N_{voxel} = \prod_{d=1}^3 N_{\mathcal{D}}^{(d)}$  volume elements or *voxels* of volume  $V_{voxel} = \prod_{d=1}^3 a_d / N_{\mathcal{D}}^{(d)}$  for which we obtain averaged (constant) measurements. We introduce a measurement operator  $\Psi$  that gives for a given exact contrast agent concentration  $c_{tiss}$  and a chosen point in time  $t \in [0, T]$  a measurement vector  $\mathbf{c} \in \mathbb{R}_{\geq 0}^{N_{voxel}}$  as

$$\mathbf{c} = \Psi[c_{tiss}](t) := \Theta[c_{tiss}](t) + \mathcal{E}[c_{tiss}](t).$$

Here,  $\Theta$  is a noise-free measurement-operator and is usually a volumetric average over each voxel being equivalent to a piece-wise constant approximation in space.  $\mathcal{E}$  abstractly models a (potentially non-linear) additive error (noise, movements, technical problems, ...).

To reflect time-discrete measurements, we introduce  $N_{obs}$  ordered, pair-wise different discrete observation times  $t_i^{obs} \in [0, T]$ ,  $i \in \{1, \dots, N_{obs}\}$ , at which measurements or observations are done, giving observation vectors  $\mathbf{c}_i^{obs} = \left( c_{i,1}^{obs}, \dots, c_{i,N_{voxel}}^{obs} \right)^\top$  with

$$(2) \quad \mathbf{c}_i^{obs} = \Psi[\mathbf{c}_{tiss}](t_i^{obs}) := \Theta[\mathbf{c}_{tiss}](t_i^{obs}) + \mathcal{E}[\mathbf{c}_{tiss}](t_i^{obs}).$$

**2.3. Contrast agent transport model following the indicator-dilution theory.** The indicator-dilution theory (IDT) [3] provides a model for the time evolution of the contrast agent concentration in a reference voxel  $\mathcal{D}_{voxel} \subset \mathcal{D}_{tiss}$  with volume  $V_{voxel}$ , given the arterial inflow  $c_{art}$ . While in this standard model, the contrast agent's concentrations are assumed to be constant in each voxel, we first want to formulate the IDT as a space-continuous model and then move over to a discrete description as consequence of a measurement process. Our continuous version of the indicator-dilution theory-based transport model replaces  $\mathcal{B}$  in (1) with the model operator  $\mathcal{B}_{IDT}$  given via

$$(3) \quad \mathbf{c}_{tiss}(\mathbf{x}, t) = \mathcal{B}_{IDT}[c_{art}, \mathbf{k}_{tiss}](t) := \int_0^T c_{art}(\tau) \mathbf{k}_{tiss}(\mathbf{x}, t - \tau) d\tau, \quad (\mathbf{x}, t) \in \mathcal{D}_{tiss} \times [0, T].$$

Kernel  $\mathbf{k}_{tiss} : \mathcal{D}_{tiss} \times [0, T] \rightarrow \mathbb{R}$  fully characterizes the properties of the tissue at point  $\mathbf{x}$ . In order to have an well-defined integrand, we assume  $\mathbf{k}_{tiss}(\cdot, t) = 0$  for  $t < 0$ . Note that the model operator  $\mathcal{B}_{IDT}$  actually is independent of the spatial position.

We now apply the measurement operator  $\Psi$  to (3) obtaining

$$\begin{aligned} \Psi[\mathbf{c}_{tiss}](t) &= \Theta[\mathcal{B}_{IDT}[c_{art}, \mathbf{k}_{tiss}]](t) + \mathcal{E}[\mathcal{B}_{IDT}[c_{art}, \mathbf{k}_{tiss}]](t) \\ &= \int_0^T c_{art}(\tau) \mathbf{k}(t - \tau) d\tau + \mathcal{E}[\mathcal{B}_{IDT}[c_{art}, \mathbf{k}_{tiss}]](t), \end{aligned}$$

where  $\mathbf{k} = (k_1, \dots, k_{N_{voxel}})^\top$  is a vector of univariate kernel functions  $k_j : [0, T] \rightarrow \mathbb{R}$ . Since we are interested in time-discrete observations, we limit our discussion to observation times  $t_i^{obs}$  yielding

$$\mathbf{c}_i^{obs} = \Psi[\mathbf{c}_{tiss}](t_i^{obs}) = \int_0^T c_{art}(\tau) \mathbf{k}(t_i^{obs} - \tau) d\tau + \boldsymbol{\varepsilon}_i^{obs},$$

with the abbreviation  $\boldsymbol{\varepsilon}_i^{obs} := \mathcal{E}[\mathcal{B}_{IDT}[c_{art}, \mathbf{k}_{tiss}]](t_i^{obs})$ . For a single voxel  $j \in \{1, \dots, N_{voxel}\}$ , we obtain

$$c_{i,j}^{obs} = \int_0^T c_{art}(\tau) k_j(t_i^{obs} - \tau) d\tau + \varepsilon_{i,j}^{obs}.$$

In case of  $\varepsilon_{i,j}^{obs} = 0$ , this boils down to the classical indicator-dilution-theory model given on a reference voxel  $j$ . Obviously, this model is independent of the spatial position of the voxel  $j$ . The classical theory further introduces a mean density  $\rho_j \in \mathbb{R}_{\geq 0}$  in a voxel  $j$ , which becomes of interest in the following subsection.

**2.4. Perfusion.** The inference task discussed in this article is to compute a time-stationary perfusion (blood flow) information  $\mathbf{p} \in \mathbb{R}^{N_{\text{voxel}}}$  given the (assumed to be exactly known) inflow concentration  $c_{\text{art}}$  and the observations  $\mathbf{c}_i^{\text{obs}}$ . Formally, the blood perfusion in a given voxel  $j$  can be evaluated as quantity of interest of the computed response function  $k_j$  as

$$p_j := p(k_j) := \frac{1}{\rho_j} k_j(0).$$

From a mathematical point of view, this quantity has nice properties, since it is just a point evaluation of the response function. In practice [3], perfusion is however often evaluated as

$$\tilde{p}_j := \tilde{p}(k_j) := \frac{1}{\rho_j} \max_{t \in [0, T]} k(t).$$

For simplicity and since we use just artificial input data, we stick to the first version of this quantity of interest.

**2.5. Deterministic inference problem.** To summarize this section, we can now formulate the deterministic problem, which we aim to solve for: For given measurement time  $T \in \mathbb{R}$ , arterial inflow  $c_{\text{art}} : [0, T] \rightarrow \mathbb{R}$ , measurement/observation times  $t_1^{\text{obs}} < t_2^{\text{obs}} < \dots < t_{N_{\text{obs}}}^{\text{obs}}$  and observation vectors  $\mathbf{c}_i^{\text{obs}} \in \mathbb{R}^{N_{\text{voxel}}}$ ,  $i \in \{1, \dots, N_{\text{obs}}\}$ , we aim at computing a vector  $\mathbf{k} = (k_1, \dots, k_{N_{\text{voxel}}})^T$  of kernel functions  $k_j : [0, T] \rightarrow \mathbb{R}$  and the derived quantity of interest  $\mathbf{p}$  such that

$$(4) \quad c_{i,j}^{\text{obs}} \approx \int_0^T c_{\text{art}}(\tau) k_j(t_i^{\text{obs}} - \tau) d\tau + \varepsilon_{i,j}^{\text{obs}}, \quad i \in \{1, \dots, N_{\text{obs}}\}, j \in \{1, \dots, N_{\text{voxel}}\}.$$

Clearly, this problem is underdetermined with the given requirements. Furthermore, we have not specified the nature of the error term, yet. This is why we used the notion “ $\approx$ ”. A much clearer idea of the concept of a *solution* to this problem is given in the next section, where we reformulate the problem as Bayesian sequential data assimilation problem.

**3. Numerical approach by sequential data assimilation.** In this section, we first introduce a discretization for the model discussed in the last section. This is necessary, since we will use its discretized version in context of sequential data assimilation, afterwards. An approximation to the solution of the assimilation problem is derived by the Ensemble Kalman filter that is briefly introduced as final part of this section.

**3.1. Discretized observation model.** We start by discretizing (4) for fixed  $i \in \{1, \dots, N_{\text{obs}}\}$  and fixed  $j \in \{1, \dots, N_{\text{voxel}}\}$ . Numerical quadrature using a rectangular rule gives

$$c_{i,j}^{\text{obs}} \approx \Delta\tau \sum_{q=0}^{N_q-1} c_{\text{art}}(\tau_q) k_j(t_i^{\text{obs}} - \tau_q) + \varepsilon_{i,j}^{\text{obs}}.$$

Here, we have introduced  $N_q$  equidistant abscissas  $\tau_q := q \cdot \Delta\tau$  with  $\Delta\tau := \frac{T}{N_q}$ . In the original problem setting, the observation times  $t_i^{\text{obs}}$  can be chosen arbitrarily. However, we

here introduce a simplification, in which we assume the observation times to be given for a fixed time step size  $\Delta t_{obs}$ . Moreover, this time step size shall be a multiple of  $\Delta\tau$ , i.e.

$$t_i^{obs} := i \Delta t_{obs}, \quad \Delta t_{obs} := s \cdot \Delta\tau, \quad s \in \mathbb{N}.$$

Thereby, we obtain

$$\begin{aligned} c_{i,j}^{obs} &\approx \Delta\tau \sum_{q=0}^{N_q-1} c_{art}(\tau_q) k_j(t_i^{obs} - \tau_q) + \varepsilon_{i,j}^{obs} \\ &= \Delta\tau \sum_{q=0}^{N_q-1} c_{art}(q\Delta\tau) k_j((i s - q)\Delta\tau) + \varepsilon_{i,j}^{obs}. \end{aligned}$$

Since we now only need  $c_{art}$  and  $k_j$  being evaluated at multiples of  $\Delta\tau$ , we can replace them by vectors  $\mathbf{c}_{art} = (c_{art,0}, \dots, c_{art,N_q-1})^\top$ ,  $\mathbf{k}_j = (k_{j,0}, \dots, k_{j,N_q-1})^\top$ , with  $c_{art,q} := c_{art}(q\Delta\tau)$  and  $k_{j,q} := k_j(q\Delta\tau)$  giving

$$c_{i,j}^{obs} \approx \Delta\tau \sum_{q=0}^{N_q-1} c_{art,q} k_{j,(i s - q)} + \varepsilon_{i,j}^{obs}.$$

With the extension of  $k_j(t) = 0$  for  $t < 0$  and some index substitutions, we can finally find (for each  $j, i$ ) vectors  $\mathbf{h}_{i,j} \in \mathbb{R}^{N_q}$  such that

$$(5) \quad c_{i,j}^{obs} \approx \langle \mathbf{h}_{i,j}, \mathbf{k}_j \rangle + \varepsilon_{i,j}^{obs}.$$

Following the nomenclature of [15], we next reformulate the deterministic inference problem from [Subsection 2.5](#) as a sequential data assimilation problem. To this end, we first translate the involved quantities into random variables as in a Bayesian inference problem. Thereafter, we introduce the basic concepts of sequential data assimilation.

Since the problem decouples for all voxels  $j \in \{1, \dots, N_{voxel}\}$ , we keep  $j$  fixed for the rest of this section.

**3.2. Probabilistic view of inference.** Let be  $(\Omega, \mathcal{F}, \mathbb{P})$  a probability space. In Bayesian inference we want to gain information on a *system state variable* for given observation(s). In our context, the state variable is the time-continuous kernel function  $k_j$ . However, since we want to avoid the technical difficulties of Bayesian inference in infinite-dimensional (function) spaces, we infer the discrete  $\mathbf{k}_j \in \mathbb{R}^{N_q}$ , instead. Therefore, we introduce a new random variable

$$(6) \quad \mathbf{K}_j : \Omega \rightarrow \mathbb{R}^{N_q},$$

replacing the time-discrete deterministic solution vector  $\mathbf{k}_j$ . Moreover, we introduce a random variable

$$\mathcal{E}_{i,j}^{obs} : \Omega \rightarrow \mathbb{R},$$

replacing the error term used before. Note that we assume  $\mathbf{K}_j$  and  $\mathcal{E}_{i,j}^{obs}$  to be independent random variables. Finally, we also consider each observation  $c_{i,j}^{obs}$  as random variable

$$C_{i,j}^{obs} : \Omega \rightarrow \mathbb{R},$$

which is usually called *observed variable*. Using (5),  $C_{i,j}^{obs}$  is defined as

$$(7) \quad C_{i,j}^{obs} = H_{i,j} \mathbf{K}_j + \mathcal{E}_{i,j}^{obs},$$

with  $H_{i,j} \in \mathbb{R}^{1 \times N_q}$  such that  $H_{i,j} \mathbf{K}_j := \langle \mathbf{h}_{i,j}, \mathbf{K}_j \rangle$  is the (linear) forward map

$$H_{i,j} : \mathbb{R}^{N_q} \rightarrow \mathbb{R}.$$

The aim of inference is to find a reference *trajectory*  $\mathbf{k}_j^{ref}$ , being a realization of  $\mathbf{K}_j$  such that the (measured) observations fit to the observed variable.

**3.3. Sequential data assimilation.** Sequential data assimilation relies on an *evolution model* and a *forward model* to obtain  $\mathbf{k}_j^{ref}$ . The models run on different time scales. The evolution model is a stochastic difference equation implying a certain predicted evolution of the system state variable over many small time steps. The forward model defines a relationship between the reference trajectory (which is to be found) and the observed data at the observation times.

**3.3.1. Evolution model.** The choice of an appropriate evolution model is crucial for the quality of the solution generated by the sequential data assimilation problem. We take the evolution model

$$(8) \quad \mathbf{K}_j^{(n+1)} = \mathbf{K}_j^{(n)} + \sqrt{\Delta\tau} \Xi^{(n)}, \quad n \in \{0, \dots, N_q - 1\}$$

for our application, implying a time-constant evolution of the kernel-function with some noise. In (8),  $(\mathbf{K}_j^{(0)}, \dots, \mathbf{K}_j^{(N_q-1)})$  is a sequence of random variables of the type given in (6) for time steps  $n \Delta\tau$ . We assume  $\mathbf{K}_j^{(0)} \sim \mathcal{N}(\mathbf{0}, \sigma_0^2 \Sigma_{\Xi})$ , corresponding to a zero initial guess for the kernel function with Gaussian noise with a covariance matrix  $\Sigma_{\Xi}$ .  $\sigma_0 \in \mathbb{R}$  is a scaling coefficient. The  $\Xi^{(n)}$ s are a sequence of independent identically distributed random variables with  $\Xi^{(n)} : \Omega \rightarrow \mathbb{R}^{N_q}$  drawn as  $\Xi^{(n)} \sim \mathcal{N}(0, \Sigma_{\Xi})$ .  $\Sigma_{\Xi} \in \mathbb{R}^{N_q \times N_q}$  is the same covariance matrix as before. We define  $\Sigma_{\Xi}$  by a Gaussian covariance kernel as

$$\Sigma_{\Xi} := (\sigma_{n,n'})_{n,n'=0}^{N_q-1}, \quad \sigma_{n,n'} := \alpha e^{-\frac{\|\tau_n - \tau_{n'}\|_2^2}{2\ell^2}}$$

with a parametrization in the scale  $\alpha \in \mathbb{R}$  and the correlation length  $\ell \in \mathbb{R}$ . Parameters  $\sigma_0$ ,  $\alpha$  and  $\ell$  will be discussed in [Section 4](#).

Note, that our evolution model can be understood as Euler-Maruyama-based discretization of the system of stochastic ordinary differential equations

$$(9) \quad d\mathbf{K}_j = d\mathbf{W}_t,$$

where  $\mathbf{W}_t$  is a vector of correlated univariate Wiener processes.



**3.3.2. Forward model.** We choose (7) as our forward model, i.e. we get the forward model

$$(10) \quad C_{i,j}^{obs} = H_{i,j} k_{j,i_s}^{ref} + \mathcal{E}_{i,j}^{obs}, \quad i \in \{1, \dots, N_{obs}\}.$$

The  $\mathcal{E}_{i,j}^{obs}$  are sequences of i.i.d. random variables for growing observation time index  $i$  following

$$\mathcal{E}_{i,j}^{obs} \sim \mathcal{N}(0, \Sigma_{\mathcal{E}}), \quad i \in \{1, \dots, N_{obs}\}$$

with  $\Sigma_{\mathcal{E}} \in \mathbb{R}$  the observation error variance. Due to  $\Delta t_{obs} = s\Delta\tau$ ,  $k_{j,i_s}^{ref}$  is the unknown reference trajectory evaluated at observation time  $t_i^{obs}$ .

**3.3.3. Assimilation task.** To be concise, we here only briefly summarize the general idea of the actual assimilation task. Further details can be found e.g. in [15].

Let  $\pi_{\mathbf{K}_j^{(i_s)}}(\mathbf{k}_j)$  be the probability density function of the random variable  $\mathbf{K}_j^{(i_s)}$  at time  $t_i^{obs}$  for  $i \in \{1, \dots, N_{obs}\}$ . Then, sequential data assimilation computes posterior PDFs

$$\pi_{\mathbf{K}_j^{(i_s)}}(\mathbf{k}_j | c_{1:i,j}^{obs}), \quad i \in \{1, \dots, N_{obs}\},$$

i.e. probability density functions of the random variables  $\mathbf{K}_j^{(i_s)}$  conditioned to the observations  $c_{1,j}^{obs}, \dots, c_{i,j}^{obs}$ . This is done using an iterative approach. It is started with  $\pi_{\mathbf{K}_j}(\mathbf{k}_j | c_{1:0,j}^{obs})$  being the PDF of  $\mathbf{K}_j^{(0)}$ . Then, for a given PDF  $\pi_{\mathbf{K}_j^{((i-1)s)}}(\mathbf{k}_j | c_{1:i-1,j}^{obs})$ , it iteratively

1. computes the density  $\pi_{\mathbf{K}_j^{(i_s)}}(\mathbf{k}_j | c_{1:i-1,j}^{obs})$  and thereby solves a prediction problem for the given evolution model (8),
2. applies Bayes theorem

$$\pi_{\mathbf{K}_j^{(i_s)}}(\mathbf{k}_j | c_{1:i,j}^{obs}) = \frac{\pi_{C_{i,j}^{obs}}(c_{i,j}^{obs} | \mathbf{k}_j) \pi_{\mathbf{K}_j^{(i_s)}}(\mathbf{k}_j | c_{1:i-1,j}^{obs})}{\int_{\mathbb{R}^{N_q}} \pi_{C_{i,j}^{obs}}(c_{i,j}^{obs} | \mathbf{k}_j) \pi_{\mathbf{K}_j^{(i_s)}}(\mathbf{k}_j | c_{1:i-1,j}^{obs}) d\mathbf{k}_j}$$

in an update step to compute  $\pi_{\mathbf{K}_j^{(i_s)}}(\mathbf{k}_j | c_{1:i,j}^{obs})$ .

In other words, the idea is to start from knowledge (encoded in  $\pi_{\mathbf{K}_j^{((i-1)s)}}(\mathbf{k}_j | c_{1:i-1,j}^{obs})$ ) at an observation time step  $t_{i-1}^{obs}$ . Then, knowledge for a new observation time step is forecasted / predicted using only the evolution model (8). This forecast is finally corrected using the information given by observation  $c_{i,j}$ . The unknown reference trajectory is ultimately given as mean of the marginal PDF  $\pi_{\mathbf{K}_j^{(i_s)}}(\mathbf{k}_j | c_{1:N_{obs},j}^{obs})$ .

**3.4. Ensemble Kalman filter.** The EnKF is a Monte-Carlo-type implementation of the above discussed iterative data assimilation task. Instead of explicitly computing the posterior PDFs  $\pi_{\mathbf{K}_j^{(i_s)}}(\mathbf{k}_j | c_{1:i-1,j}^{obs})$  and  $\pi_{\mathbf{K}_j^{(i_s)}}(\mathbf{k}_j | c_{1:i,j}^{obs})$ , the EnKF constructs an *ensemble of realizations* of random variables representing these PDFs in an empirical sense. In that context, *forecast* and *analysis ensembles* are distinguished. As we will see, the computation of the forecast ensemble corresponds to approximating  $\pi_{\mathbf{K}_j^{(i_s)}}(\mathbf{k}_j | c_{1:i-1,j}^{obs})$ , while the computation of the analysis ensemble corresponds to the approximation of  $\pi_{\mathbf{K}_j^{(i_s)}}(\mathbf{k}_j | c_{1:i,j}^{obs})$ .

Shall  $N_e$  be the size of the ensembles. Then, the EnKF algorithm starts by drawing  $N_e$  samples  $\mathbf{k}_j^{(0),1}, \dots, \mathbf{k}_j^{(0),N_e}$  of the (initial) system state according to the PDF of  $\mathbf{K}_j^{(0)}$ . The algorithm consists of two main steps which are iteratively done for  $i \in \{1, \dots, N_{obs}\}$ .

**3.4.1. Forecast step.** In the forecast step, the ensemble is propagated over  $s$  steps of the evolution model in (8) to reach the next observation time step  $t_i^{obs} = i s \Delta\tau$ . To achieve this realizations  $\xi_{m,n}$  for  $m \in \{1, \dots, N_e\}$  are drawn i.i.d. from  $\Xi^{(n)}$  in each of the  $s$  steps. Then the propagation equation reads for  $n = 1, \dots, s$  as

$$\mathbf{k}_j^{(i(s-1)+n),m} = \mathbf{k}_j^{(i(s-1)+(n-1)),m} + \sqrt{\Delta\tau} \xi_{m,n}, \quad m \in \{1, \dots, N_e\}.$$

The newly generated ensemble is the *forecast ensemble*  $(\mathbf{k}_j^{f,m})_{m=1}^{N_e}$  with  $\mathbf{k}_j^{f,m} := \mathbf{k}_j^{i s, m}$ . We further compute the empirical forecast mean

$$(11) \quad \bar{\mathbf{k}}_j^f := \frac{1}{N_e} \sum_{m=1}^{N_e} \mathbf{k}_j^{f,m} \in \mathbb{R}^{N_q}$$

and the empirical forecast covariance (matrix)

$$(12) \quad \Sigma_{\mathbf{k}_j}^f := \frac{1}{N_e - 1} \sum_{m=1}^{N_e} (\mathbf{k}_j^{f,m} - \bar{\mathbf{k}}_j^f) (\mathbf{k}_j^{f,m} - \bar{\mathbf{k}}_j^f)^\top \in \mathbb{R}^{N_q \times N_q}.$$

**3.4.2. Analysis step.** In the analysis step, the Kalman filter [8, 15] is applied to the forecast ensemble to compute an analysis ensemble  $(\mathbf{k}_j^{a,m})_{m=1}^{N_e}$  representing the PDF  $\pi_{\mathbf{K}_j^{(i s)}}(\mathbf{k}_j | c_{1,i,j}^{obs})$ , which is conditioned to the new observation  $c_{i,j}^{obs}$ . As part of the Kalman filter, the forward model (10) with  $k_{j,i s}^{ref}$  being replaced by  $\mathbf{K}_j^{(i s)}$  is evaluated. The standard construction of the Kalman filter assumes to have a linear forward map  $H_{i,j}$ . Moreover all involved random variables shall be Gaussian. If these requirements are fulfilled, it can be shown that the analysis ensemble follows a Gaussian distribution, too. Therefore, it can be fully characterized by the empirical analysis mean  $\bar{\mathbf{k}}_j^a$  and the empirical analysis covariance  $\Sigma_{\mathbf{k}_j}^a$  (defined analogously to equations (11) and (12)).

Based on this observation, the core idea of the Kalman filter is to compute the empirical analysis mean as minimization problem

$$\bar{\mathbf{k}}_j^a = \arg \min_{\mathbf{k}_j \in \mathbb{R}^{N_q}} \frac{1}{2} \left( \left\| \mathbf{k}_j - \bar{\mathbf{k}}_j^f \right\|_{\left( \Sigma_{\mathbf{k}_j}^f \right)^{-1}}^2 + \left\| H_{i,j} \mathbf{k}_j - c_{i,j}^{obs} \right\|_{\Sigma_{\mathcal{E}}^{-1}}^2 \right).$$

Given the linearity of  $H_{i,j}$ , the minimum can be exactly computed as

$$\bar{\mathbf{k}}_j^a = \bar{\mathbf{k}}_j^f - U_{i,j} (H_{i,j} \bar{\mathbf{k}}_j^f - c_{i,j}^{obs}),$$

where  $U_{i,j}$  is the the Kalman (update) matrix

$$U_{i,j} = \Sigma_{\mathbf{k}_j}^f H_{i,j}^\top (H_{i,j} \Sigma_{\mathbf{k}_j}^f H_{i,j}^\top + \Sigma_{\mathcal{E}})^{-1}.$$

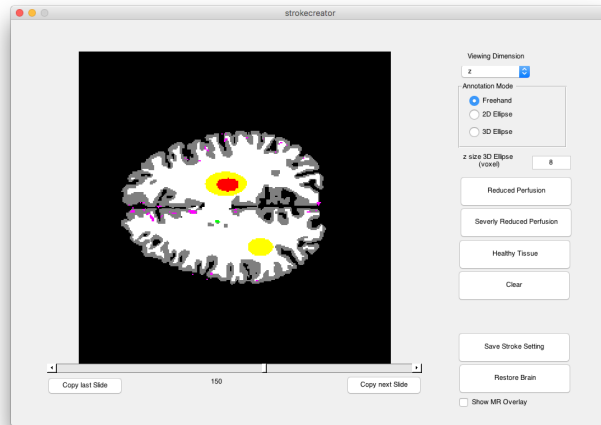
Instead of explicitly computing the empirical analysis mean and covariance (the latter by an analogous update idea), the analysis part of the Ensemble Kalman Filter (with perturbed observations) [15, Chapter 7] directly updates the forecast ensemble by

$$\mathbf{k}_j^{a,m} = \mathbf{k}_j^{f,m} - U_{i,j}(H_{i,j}\mathbf{k}_j^{f,m} + \varepsilon_{i,j,m} - c_{i,j}^{obs}), \quad m \in \{1, \dots, N_e\},$$

where  $\{\varepsilon_{i,j,m}\}_{m=1}^{N_e}$  are realizations of  $\mathcal{E}_{i,j}^{obs}$ . If required, empirical versions of the analysis mean and analysis covariance can be computed analogously to (11) and (12). Finally, the next forecast step is initialized with  $\mathbf{k}_j^{(i,s),m} = \mathbf{k}_j^{a,m}$ , that is, the analysis ensemble replaces the system state for  $t_i^{obs}$ .

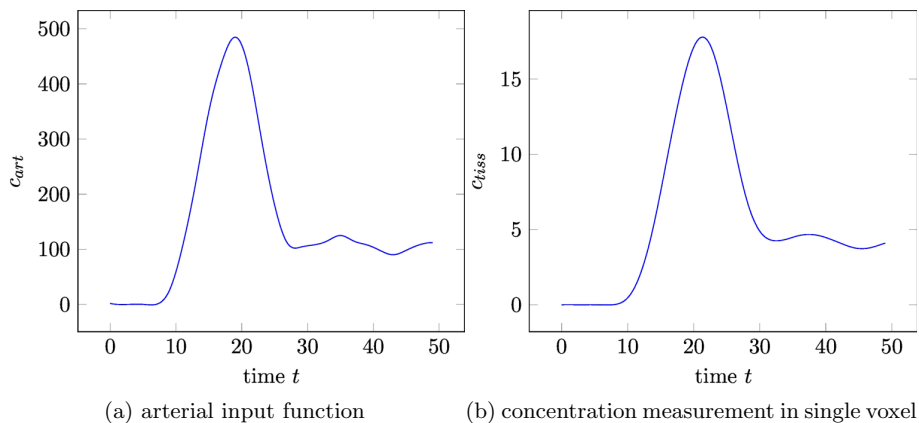
**3.4.3. Result.** For  $i = N_{obs}$  the algorithm terminates with an analysis ensemble, representing the posterior PDF  $\pi_{\mathbf{K}_j^{(N_{obs},s)}}(\mathbf{k}_j | C_{1:N_{obs},j}^{obs})$ . The reference trajectory is extracted as empirical mean  $\bar{\mathbf{k}}_j := \frac{1}{N_e} \sum_{m=1}^{N_e} \mathbf{k}_j^{(N_{obs},s),m}$ . The (mean) perfusion  $\bar{p}_j$  can be derived as  $\bar{p}_j = \frac{1}{\rho_j} \bar{\mathbf{k}}_j |_{t=0}$ . Empirical covariances are extracted as discussed before. Moreover, in case cumulative distribution functions or other probabilistic quantities shall be extracted, a kernel-density estimator (such as `ksdensity` in Matlab) is applied to the generated ensemble.

**4. Numerical results.** In this section, we demonstrate the beforehand introduced numerical method for artificial test data. To this end, we first introduce the source of this test data, which is a *Digital Perfusion Phantom*. Then, we study the numerical properties of our method in terms of convergence, parameter dependence and input dependence in a single-voxel scenario. Finally we solve the perfusion inference problem for a slice of a full (artificial) DCE imaging brain data set.



**Figure 1.** The source of our artificial measurements is the *Digital Brain Perfusion Phantom* package [9]. A Matlab implementation of this work is available. It allows to mark brain regions with reduced and severely reduced perfusion, here shown with the colors yellow and red. Given this data, artificial DCE imaging data is created.

**4.1. Digital Perfusion Phantom.** Digital Perfusion Phantoms (DPP) [17, 13, 9] allow to artificially generate DCE image data for perfusion analysis. Thereby new algorithms can be tested on such data without the additional constraints of true patient data. Perfusion Phantoms basically solve the forward problem, which involves to transform perfusion information into contrast agent concentrations. In our work, we use the *Digital Brain Perfusion Phantom* package [9], which is a Matlab implementation of the model introduced in [17]. The software provides a radiological image of a brain. A user interface, see Figure 1, allows to mark regions of reduced and strongly reduced perfusion. It is possible to control the observation snapshot time step size (i.e.  $\Delta t_{obs}$ ) of the artificial radiological imaging process. The measurement time is  $T = 49$ . The resolution of the artificially generated data is  $\mathbf{N}_D = (256, 256, 256)$ . The arterial input function is provided as discrete evaluations  $c_{art}(t_i^{input})$  with  $t_i^{input} = 2i$ . The Perfusion Phantom package uses a piecewise cubic spline interpolant through this data as exact  $c_{art}$ , see Figure 2a. During the artificial imaging process, each snapshot (i.e.  $c_i^{obs}$ ) is written in a separate file. A *baseline* for the radiological images is written, too. It contains the measurement data without contrast agent concentrations. In our examples, we always subtract this baseline data from the artificial measurements to obtain just the necessary concentration information.

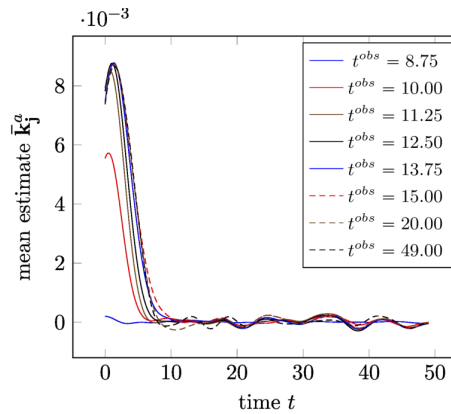


**Figure 2.** We use idealized concentration functions for one tissue voxel in order to test the implemented numerical method.

We perform a major part of our numerical tests on a single reference voxel which has been chosen arbitrarily as  $(100, 130, 150)$ . The observation data for that single voxel is stored with a time step size of  $\Delta t_{obs} = 0.25$ . This data is interpolated by a piecewise cubic spline to obtain measurement data at arbitrary points in time for our initial tests, cf. Figure 2b. Towards the end of this section, results for a full slice  $(\cdot, \cdot, 150)$  of the full data set are discussed.

We start by showing a series of numerical results obtained for given artificial input without noise. These results will give an insight into the choice of the different parameters of the method and into the convergence properties of the method. Noisy data is discussed afterwards.

**4.2. Data assimilation process.** Let us first have a look at the evolution of the analysis ensemble during the sequential data assimilation process. We have chosen an observation time step size of  $\Delta t_{obs} = 0.25$ , a quadrature step size of  $\Delta \tau = 0.0625$  (i.e.  $s = 4$ ) and an ensemble size of  $N_e = 5000$ . For a meaningful definition of the (co-)variances, we have to account for the scales of the involved quantities. By experiments, we found out that the kernel function  $k_j$  has a magnitude of about  $10^{-3}$ . Therefore, the scaling  $\alpha$  of the covariance matrix  $\Sigma_{\Xi}$  should be relative to a *standard deviation* of  $10^{-3}$ . With this in mind, we set  $\alpha = (10^{-3})^2 0.001$ . This corresponds to a *relative variance* of 0.001. The correlation length is set to  $\ell = 2$ . For the covariance of the initial state  $\mathbf{K}_j^{(0)}$ , we impose an additional scaling of  $\sigma_0 = 100$ , accounting for a much larger uncertainty in the initial state. The observation error variance also needs a problem-adapted scaling. Since the concentration measurements are in the range of 10, we shift the (co-)variance by a standard deviation of 10. Using a relative variance of 0.0001, we obtain  $\Sigma_{\mathcal{E}} = 10^2 0.0001$ .

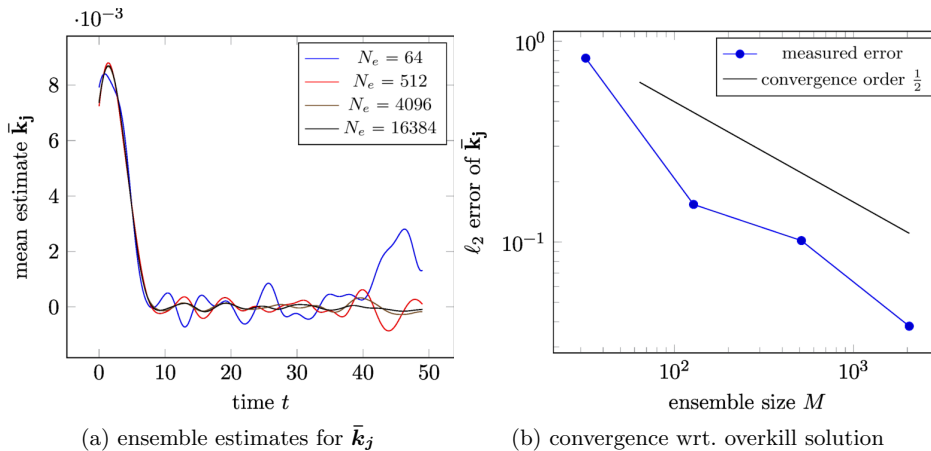


**Figure 3.** During the sequential data assimilation process, the analysis ensemble and thereby the empirical mean of the kernel function gets continuously updated, here shown for different update time steps.

In **Figure 3**, we show the evolution of the empirical mean  $\bar{k}_j^a$ , i.e. the prediction for the unknown kernel function, for different observation times during the operation of the EnKF. Note that a scaled evolution of  $\bar{k}_j^a$  at  $t = 0$  corresponds to the (scaled) unknown perfusion  $\bar{p}$ . Therefore, discussing numerical results for  $\bar{k}_j^a$  is equivalent to discussing results for  $\bar{p}$ . The major information gain for the predicted result is in time interval  $[10, 20]$ . This is the time interval in which the concentration at the arterial inlet grows. Afterwards, the data assimilation process only gains very little more information and converges towards the final result.

**4.3. Convergence in the ensemble size.** Next, we discuss the convergence of the empirical mean of  $k_j$  with respect to the ensemble size  $N_e$ . In the following, we will always concentrate on the last analysis ensemble obtained after assimilating the observation for  $t_{obs} = 49$ . To shorten notation, we skip additional indices, indicating this and call the empirical mean of this analysis ensemble  $\bar{k}_j$ .

Our convergence study with respect to the ensemble size uses the same parameters as in the

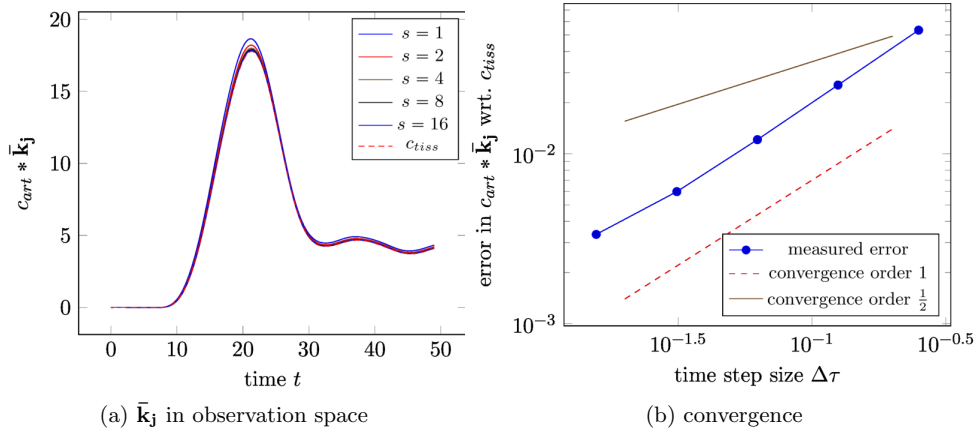


**Figure 4.** With growing ensemble size, the empirical estimate for the mean of the response / kernel function gets more accurate (left) and converges with order  $\frac{1}{2}$  (right).

previous paragraph. However, this time, we change the size of the ensemble. In Figure 4a, we show the empirical mean  $\bar{k}_j$  for ensemble sizes  $N_e \in \{64, 512, 4096, 16384\}$ . The convergence in the error of the empirical mean is shown in Figure 4b. Here, we define the solution for  $N_e = 16384$  as overkill solution and show convergence in the relative  $\ell_2$  error  $\frac{\|k_j - k_j^{\text{overkill}}\|_{\ell_2}}{\|k_j^{\text{overkill}}\|_{\ell_2}}$  towards this solution. The results indicate a convergence order of approximately  $\frac{1}{2}$ . This is the expected order of convergence, since we use a Monte Carlo-type estimator.

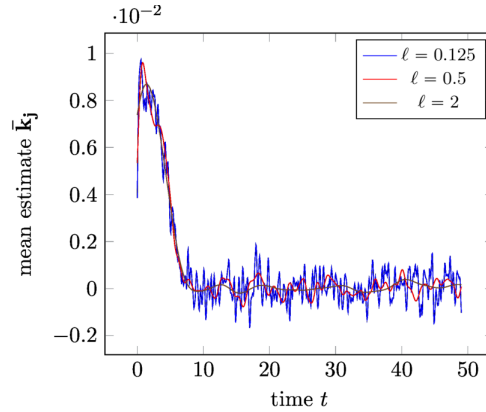
**4.4. Convergence in the time sub-steps  $\Delta\tau$ .** In the following, we have a look at convergence with respect to the quadrature and evolution model step size  $\Delta\tau$ . Here, we try to avoid the use of an overkill solution. To achieve this, we (discretely) fold the empirical mean  $\bar{k}_j$  againsts the arterial input function  $c_{art}$ , i.e. we transfer the prediction for  $k_j$  into observation space. In observation space, we have the analytically given artificial measurement result  $c_{tiss}$ , which we use to compare  $c_{art} * \bar{k}_j$  against. In our numerical study, we vary the sub-step number  $s$ , i.e. change  $\Delta\tau$ , while keeping all other parameters as in Subsection 4.2.

In Figure 5a, we visually compare the results obtained for an increasing number of sub-steps  $s$  (i.e. decreasing  $\Delta\tau$ ). The convergence plot in Figure 5b further shows the error reduction in the relative  $\ell_2$  norm for decreasing  $\Delta\tau$  if we compare the convolved mean estimate  $\bar{k}_j$  with the real observation data. The results indicate approximately first order convergence. In fact, parameter  $\Delta\tau$  influences the Euler-Maruyama approximation of the continuous stochastic differential equation (9) and the quadrature of the convolution integral. While the Euler-Maruyama method is known to have halve order convergence, the rectangular rule is convergent of second order for sufficiently smooth integrands. The observed convergence behavior strongly depends on the dominance of one of the errors (time-integration, quadrature). The observed first order seems to indicate that the quadrature error for the convolution integral is the dominant one. Nevertheless, full second order convergence is not achieved. This observation is clearly a pre-asymptotic and strongly problem-dependent result.



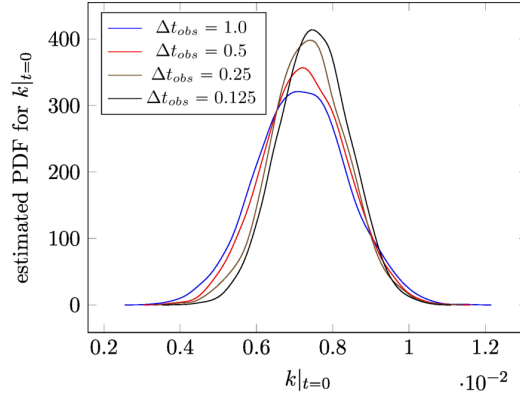
**Figure 5.** A smaller time step size for the quadrature / system state model leads to convergence of  $\bar{k}_j$  in observation space towards the measurement concentration  $c_{tiss}$ .

**4.5. Influence of the correlation length in the system state noise.** Our next study shall give an insight into the influence of the system state model, more specifically the influence of the correlation length  $\ell$  of the random variable  $\Xi$  on the inferred solution. To study the influence of the correlation length, we keep the parameters as in [Subsection 4.2](#) and apply different correlation lengths  $\ell \in \{0.125, 0.5, 2\}$ . The results of this numerical study are given in [Figure 6](#). Here, the inferred kernel function  $\bar{k}_j$  is shown for different correlation lengths. For growing correlation length the result gets less noisy. Hence, a larger correlation lengths has a smoothing effect on the solution. Since, in general, we seek for smooth solutions, we always choose  $\ell = 2$ .



**Figure 6.** Longer correlation lengths  $\ell$  impose a higher smoothness on the ensemble estimate.

**4.6. Influence of the number of observations.** Our final test with noise-free model data on a single voxel highlights the influence of a change of the observation time step size  $\Delta t_{obs}$ ,



**Figure 7.** The more observation samples are taken, the more reliable the estimate of the solution. Hence, the estimated PDF for  $k|_{t=0}$  shows a smaller variance for smaller observation time steps  $\Delta t_{obs}$ .

i.e. a change in the number of observations that are made during the imaging process. To test this, we take the same parameters as in [Subsection 4.2](#), but change the observation time step size as  $\Delta t_{obs} \in \{0.125, 0.25, 0.5, 1.0\}$  while keeping  $\Delta \tau$  constant. The quantity that we study is the computed probability density function for  $k|_{t=0}$ , hence a scaled version of  $\bar{p}$ . We use the kernel density estimator `ksdensity` in Matlab to reconstruct a continuous PDF for the ensemble data.

The results of this study can be seen in [Figure 7](#). Here, we make two observations. First, the mean of the PDF still changes for growing number of measurements, converging towards a true solution. Second, and more important, we observe a variance reduction if we increase the number of measurements. Clearly, this is the kind of information that we are interested in, since we thereby obtain confidence information for our solution.

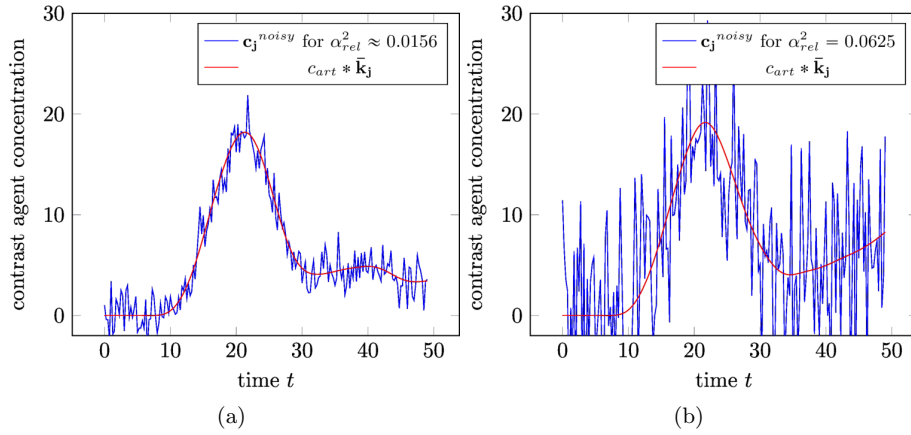
**4.7. Inference from noisy data.** Until now, we considered noise-free input data. Instead, we now discuss the same one-voxel input as before, but add artificial noise as

$$c_{i,j}^{noisy} = c_{i,j}^{obs} + \lambda_{i,j}, \quad i \in \{1, \dots, N_{obs}\},$$

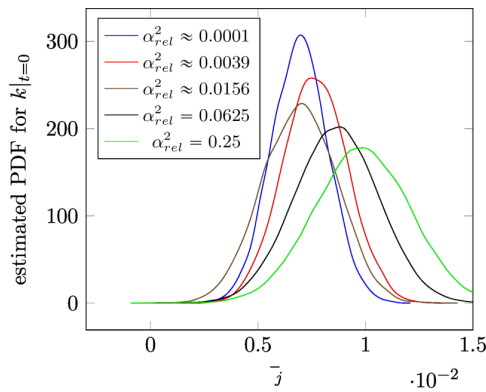
where the  $\lambda_{i,j}$  are realizations of i.i.d. random variables  $\Lambda_{i,j} : \Omega \rightarrow \mathbb{R}$ ,  $\Lambda_{i,j} \sim \mathcal{N}(0, \Sigma_{\Lambda})$  with  $\Sigma_{\Lambda} \in \mathbb{R}$  the variance of the noise.

We use a series of test cases with  $\Sigma_{\Lambda} = 10^2 \alpha_{rel}$  and  $\alpha_{rel} \in \{2^{-10}, 2^{-8}, 2^{-6}, 2^{-4}, 2^{-2}\}$ . Hence,  $\alpha_{rel}$  corresponds to the relative variance with respect to the magnitude of the measurements. We keep a major part of the parameters from [Subsection 4.2](#). However, we change the fixed observation error variance  $\Sigma_{\mathcal{E}}$ , to a problem-adapted one, namely, i.e.  $\Sigma_{\mathcal{E}} = \Sigma_{\Lambda}$ . Note that in practice, one would empirically *estimate* the noise in the measurement data and would set  $\Sigma_{\mathcal{E}}$  accordingly. Another change concerns the number of samples  $N_e$  in the EnKF. As our experiments showed, the size of the ensemble has to be increased for higher variances. This is well covered by classical Monte-Carlo theory. Therefore, we set  $N_e = 10000$  for  $\alpha_{rel} \in \{2^{-10}, 2^{-8}, 2^{-6}\}$  while we use  $N_e = 60000$  and  $N_e = 100000$  for  $\alpha_{rel} = 2^{-4}$  and  $\alpha_{rel} = 2^{-2}$ , respectively.





**Figure 8.** Even for stronger noise on the input data the inference of  $\bar{k}_j$  is acceptable, as long as the measurement variance  $\Sigma_{\mathcal{E}}$  is chosen appropriately. Here, we compare the noisy input  $\mathbf{c}_j^{\text{noisy}} \bar{k}_j$  in observation space for  $\alpha_{rel}^2 = 0.015625$  (left) and  $\alpha_{rel}^2 = 0.0625$  (right).



**Figure 9.** The proposed method is pretty robust with respect to noise. This can be seen, if we study the estimated probability density functions for  $k|_{t=0}$ . With growing noise variance, the empirical PDF estimate still recovers the mean appropriately. Extreme noise variances degenerate the result, as expected.

In **Figure 8**, we give two examples of noisy inputs for  $\alpha_{rel}^2 = 0.015625$  and  $\alpha_{rel}^2 = 0.0625$ . In the latter case, the original input signal is already severely degenerated. The predicted mean solutions in observation space  $c_{art} * \bar{k}_j$  are also given in **Figure 8**. In fact, the reconstructed solution is almost not influenced for  $\alpha_{rel}^2 = 0.015625$  and gets a little distorted for stronger noise. In **Figure 9**, we compare estimates of the PDF for  $k|_{t=0}$  for growing noise in the data. Since we appropriately account for the noise in the input, the mean is almost identical up to  $\alpha_{rel}^2 = 0.015625$ . For higher relative noise variances, the probability density functions still cover the general tendency of the results. Note that the variance in the solutions grows for larger noise in the input. This effect is not primarily caused by the noisy input, but by the imposed observation error  $\mathcal{E}_{i,j}$ , which acts here as a regularization for the noisy input.

Nonetheless, as long as the observation error variance is set in the range of the input noise variance, the variance in the solution correctly represents the variance coming from the noise in the input.

**4.8. Application problem.** We finally apply the beforehand studied method to a full application problem given by the Digital Brain Perfusion Phantom introduced in [Subsection 4.1](#). We use the slice  $(\cdot, \cdot, 150)$  with the choice of regions with reduced and severely reduced perfusion as in [Figure 1](#). We discuss a result for a large observation time-step size  $\Delta t_{obs} = 1.0$ , a highly resolved quadrature with  $\Delta\tau = 0.0625$  and a noise with variance  $\Sigma_{\Lambda} = 10^2 0.015625$ . The observation error variance is adapted as  $\Sigma_{\mathcal{E}} = \Sigma_{\Lambda}$ . All other parameters are kept as in [Subsection 4.2](#).

**4.8.1. Storage and performance considerations.** Storing and computing the ensembles for the discussed test cases is a rather challenging task. Just considering the analysis ensemble for a single slice, we need to store for each of the  $256 \times 256$  voxels 5000 realizations of discrete kernel functions  $k_j$  given via  $N_q = 785$  double precision values leading to a total storage requirement of

$$256 \times 256 \times 5000 \times 785 \times 8 \text{ Bytes} \approx 1834 \text{ GBytes}.$$

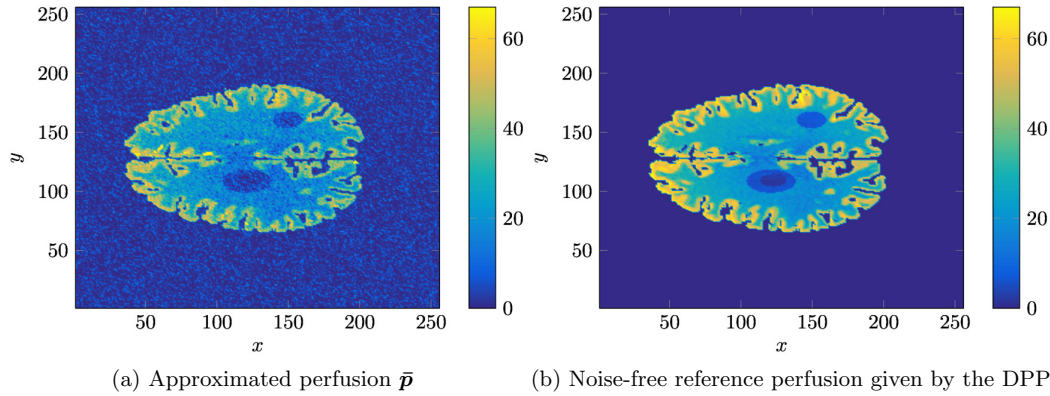
All our calculations are done in Matlab. We always compute 8 rows of the final  $256 \times 256$  slice at the same time and reuse the random input for each voxel in order to reduce the runtime. Note that especially sampling from  $\Xi^{(n)}$  is very computationally demanding. In order to do the calculations, we need constant access to way more than 64 GBytes of RAM. Due to storage and memory requirements, we to use nodes of the cluster *Rhea* at Oak Ridge National Lab to compute the full problem. Each node has 128 GBytes of RAM and a dual Intel® Xeon® E5-2650 CPU with 16 cores. To compute 8 lines, i.e. results for  $8 \times 256 = 2048$  voxels, we need about 3 hours and 15 minutes, noting that Matlab uses approximately 14 cores of the full machine. The total computing time (with respect to one node of *Rhea*) is thereby roughly 104 hours or about 4.3 days on a single machine.

Even though this amount of computing time seems to be rather prohibitive for the specific application case, it is clear that the discussed algorithm is extremely easy to parallelize. Especially, it seems to be very well suited to a parallelization on graphics processing units (GPUs) or other many-core hardware, as long as the results of the calculation are constantly streamed out to CPU memory. An appropriate parallel implementation is future work.

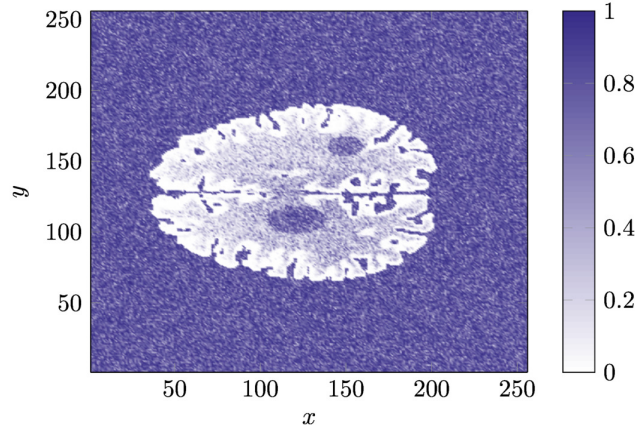
**4.8.2. Quantities of interest.** In our application examples, we consider the approximation of probabilistic quantities of interest in connection with the perfusion  $p_j := p(k_j) = \frac{1}{\rho_j} k(0)$ . Besides of the mean  $\bar{p}_j$  we are especially interested in probabilities for  $p_j$  to be in a given range. To be more specific, we compute the probabilities

$$\mathbb{P}(p_j < 10), \quad \mathbb{P}(20 \leq p_j < 40), \quad \mathbb{P}(p_j \geq 50),$$

noting that  $p_j \in [0, 70]$  in the case of the Digital Brain Perfusion Phantom data that we consider. These quantities give probabilities for low, medium and high perfusion in some region of the brain. Given the final analysis ensemble for  $k_j$ , it is easy to compute the above quantities by using the kernel density estimator `ksdensity`. The latter one can compute a cumulative distribution function (CDF) for each voxel, which is finally evaluated appropriately.



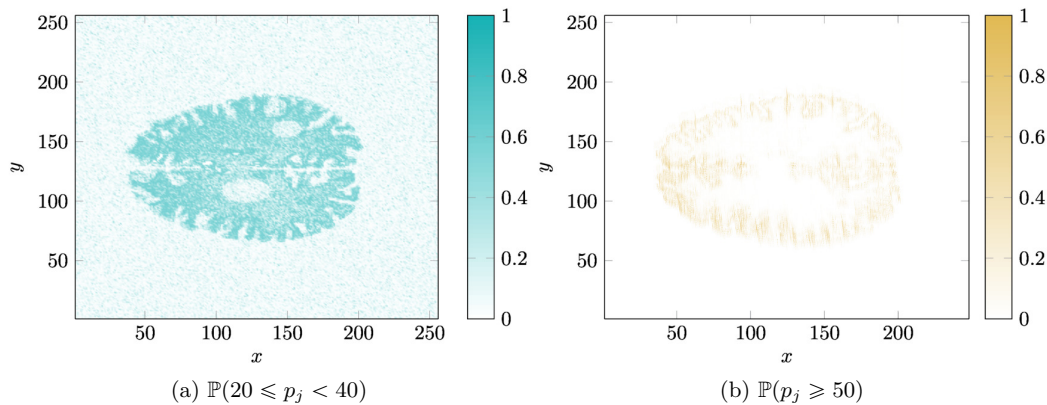
**Figure 10.** Our approximation method recovers the reference perfusion result (right) as mean of the ensemble in the Ensemble Kalman Filter. Both results match well, even though we introduced a considerable amount of artificial noise.



**Figure 11.** The advantage of the proposed method is that we are now also able to compute probabilistic information for the solution, here shown by plotting the probability  $\mathbb{P}(p_j < 10)$ . Hence, the depicted results give the space-dependent probability for low ( $< 10$ ) perfusion.

**4.8.3. Discussion of results.** An important advantage of the use of a Digital Perfusion Phantom is the existence of a reference solution to compare with. The DPP software that we use stores the reference solution together with the other generated data. In Figure 10b, we show the reference solution for our full application test case. The approximated result of our application example study, i.e.  $\bar{p}$ , is shown in Figure 10a. As expected from our single-voxel study, the inferred perfusion matches the exact perfusion result well. Note that this is the case even though we add a considerable amount of noise on the measurements.

As discussed before, we can use the ensemble-based estimate of the posterior probability density function to extract a wide range of probabilistic information on the inferred solution. This is the main result of this work. To exemplify this, we compute space-dependent prob-



**Figure 12.** Based on the results of the EnKF, it is easily possible to identify regions of high probability to have medium (left) and high (right) perfusion.

abilities for low (Figure 11), medium (Figure 12a) and high (Figure 12b) perfusion ranges, cf. Subsection 4.8.2. In case of Figure 11, we e.g. can now easily identify ranges of low perfusion and even can give a probability for this result.

In general, we claim that this probability information or derived probabilistic quantities (variance, percentiles, etc.) can give domain-experts in radiology a much clearer information on the reliability of the inferred estimates.

**5. Summary.** In this work, we have discussed the use of Ensemble Kalman filters for sequential data assimilation in order to infer probabilistic information on (blood) perfusion in tissue for given measurements from dynamic contrast-enhanced imaging. The deterministic inference of perfusion is well-known in the field of radiological imaging. However, to the best of the author’s knowledge, the new contribution is the approximation of PDFs for the perfusion given (noisy) measurements. EnKF are well-known in inference for dynamical systems and partial differential equations with stochastic coefficients. Hence, modeling the dynamic contrast-enhanced imaging process as sequential data assimilation in a Bayesian context was the main contribution of the work. Given the ensemble-based approximation of the PDF, we could compute probabilistic quantities such as probabilities for perfusion parameter ranges.

The new approach was first investigated for a single-voxel example with respect to convergence and parameter influence. Afterwards, it was applied to artificial application data generated by a Digital Perfusion Phantom, i.e. a model for deriving DCE image data for given perfusion data. Overall, the effectiveness of the method could be demonstrated, showing empirical convergence results and appropriate approximations of probabilistic information. The use of advanced filtering techniques and an efficient parallel implementation are future work.

**Acknowledgements.** The author likes to thank Wolfram Stiller and Christian Weis of the department of *Diagnostic and Interventional Radiology* of the *University Medical Center Heidelberg* and Holger Fröning of the *Institute of Computer Engineering at University of Heidelberg* for fruitful initial discussions on the application background.

## REFERENCES

- [1] J. L. ANDERSON, *An ensemble adjustment Kalman filter for data assimilation*, Monthly weather review, 129 (2001), pp. 2884–2903.
- [2] J. L. ANDERSON, *A local least squares framework for ensemble filtering*, Monthly Weather Review, 131 (2003), pp. 634–642.
- [3] G. BRIX, J. GRIEBEL, F. KIESSLING, AND F. WENZ, *Tracer kinetic modelling of tumour angiogenesis based on dynamic contrast-enhanced CT and MRI measurements*, European journal of nuclear medicine and molecular imaging, 37 (2010), pp. 30–51.
- [4] O. G. ERNST, B. SPRUNGK, AND H.-J. STARKLOFF, *Analysis of the ensemble and polynomial chaos Kalman filters in Bayesian inverse problems*, SIAM/ASA Journal on Uncertainty Quantification, 3 (2015), pp. 823–851.
- [5] G. EVENSEN, *Data assimilation: the ensemble Kalman filter*, Springer Science & Business Media, 2009.
- [6] A. FIESELMANN, M. KOWARSCHIK, A. GANGULY, J. HORNEGGER, AND R. FAHRIG, *Deconvolution-based CT and MR brain perfusion measurement: theoretical model revisited and practical implementation details*, Journal of Biomedical Imaging, 2011 (2011), p. 14.
- [7] M. A. IGLESIAS, K. J. LAW, AND A. M. STUART, *Ensemble Kalman methods for inverse problems*, Inverse Problems, 29 (2013), p. 045001.
- [8] R. E. KALMAN ET AL., *A new approach to linear filtering and prediction problems*, Journal of basic Engineering, 82 (1960), pp. 35–45.
- [9] M. MANHART, *Digital Brain Perfusion Phantom Documentation*, provided on the data web page of the pattern recognition lab at FAU Erlangen-Nürnberg, Germany (last check: Feb. 23, 2017), <http://www5.cs.fau.de/research/data/digital-brain-perfusion-phantom/>.
- [10] G. NÆVDAL, O. SEVAREID, AND R.-J. LORENTZEN, *Data assimilation using MRI data*, in Proceedings of the ECCOMAS Congress 2016, 2016.
- [11] L. ØSTERGAARD, R. M. WEISSKOFF, D. A. CHESLER, C. GYLDENSTED, AND B. R. ROSEN, *High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: Mathematical approach and statistical analysis*, Magnetic resonance in medicine, 36 (1996), pp. 715–725.
- [12] G. J. PARKER, C. ROBERTS, A. MACDONALD, G. A. BUONACCORSI, S. CHEUNG, D. L. BUCKLEY, A. JACKSON, Y. WATSON, K. DAVIES, AND G. C. JAYSON, *Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced mri*, Magnetic resonance in medicine, 56 (2006), pp. 993–1000.
- [13] O. S. PIANYKH, *Digital perfusion phantoms for visual perfusion validation*, American Journal of Roentgenology, 199 (2012), pp. 627–634.
- [14] S. REICH, *A nonparametric ensemble transform method for Bayesian inference*, SIAM Journal on Scientific Computing, 35 (2013), pp. A2013–A2024.
- [15] S. REICH AND C. COTTER, *Probabilistic forecasting and Bayesian data assimilation*, Cambridge University Press, 2015.
- [16] S. REICH AND C. J. COTTER, *Ensemble filter techniques for intermittent data assimilation*, Large Scale Inverse Problems. Computational Methods and Applications in the Earth Sciences, 13 (2013), pp. 91–134.
- [17] A. J. RIORDAN, M. PROKOP, M. A. VIERGEVER, J. W. DANKBAAR, E. J. SMIT, AND H. W. DE JONG, *Validation of CT brain perfusion methods using a realistic dynamic head phantom*, Medical physics, 38 (2011), pp. 3212–3221.
- [18] S. SOURBRON, *A tracer-kinetic field theory for medical imaging*, IEEE transactions on medical imaging, 33 (2014), pp. 935–946.
- [19] A. M. STUART, *Inverse problems: a Bayesian perspective*, Acta Numerica, 19 (2010), pp. 451–559.
- [20] M. K. TIPPETT, J. L. ANDERSON, C. H. BISHOP, T. M. HAMILL, AND J. S. WHITAKER, *Ensemble square root filters*, Monthly Weather Review, 131 (2003), pp. 1485–1490.
- [21] P. S. TOFTS, *Modeling tracer kinetics in dynamic Gd-DTPA MR imaging*, Journal of Magnetic Resonance Imaging, 7 (1997), pp. 91–101.
- [22] F. ZHANG, M. ZHANG, AND J. A. HANSEN, *Coupling ensemble Kalman filter with four-dimensional variational data assimilation*, Advances in Atmospheric Sciences, 26 (2009), pp. 1–8.

## LATEST PREPRINTS

No.	Author:	Title
2016-16	<b>A. Hyder</b>	<i>Conformally Euclidean metrics on <math>R^n</math> with arbitrary total Q-curvature</i>
2016-17	<b>G. Mancini, L. Martinazzi</b>	<i>The Moser-Trudinger inequality and its extremals on a disk via energy estimates</i>
2016-18	<b>R. N. Gantner, M. D. Peters</b>	<i>Higher order quasi-Monte Carlo for Bayesian shape inversion</i>
2016-19	<b>C. Urech</b>	<i>Remarks on the degree growth of birational transformations</i>
2016-20	<b>S. Dahlke, H. Harbrecht, M. Utzinger, M. Weimar</b>	<i>Adaptive wavelet BEM for boundary integral equations: Theory and numerical experiments</i>
2016-21	<b>A. Hyder, S. Iula, L. Martinazzi</b>	<i>Large blow-up sets for the prescribed Q-curvature equation in the Euclidean space</i>
2016-22	<b>P. Habegger</b>	<i>The norm of Gaussian periods</i>
2016-23	<b>P. Habegger</b>	<i>Diophantine approximations on definable sets</i>
2016-24	<b>F. Amoroso, D. Masser</b>	<i>Lower bounds for the height in Galois extensions</i>
2016-25	<b>W. D. Brownawell, D. W. Masser</b>	<i>Zero estimates with moving targets</i>
2016-26	<b>H. Derksen, D. Masser</b>	<i>Linear equations over multiplicative groups, recurrences, and mixing III</i>
2016-27	<b>D. Bertrand, D. Masser, A. Pillay, U. Zannier</b>	<i>Relative Manin-Mumford for semi-abelian surfaces</i>
2016-28	<b>L. Capuano, D. Masser, J. Pila, U. Zannier</b>	<i>Rational points on Grassmannians and unlikely intersections in tori</i>
2016-29	<b>C. Nobili, F. Otto</b>	<i>Limitations of the background field method applied to Rayleigh-Bénard convection</i>

## LATEST PREPRINTS

- | No.     | Author: Title  |
|---------|--|
| 2016-30 | <b>W. D. Brownawell, D. W. Masser</b><br><i>Unlikely intersections for curves in additive groups over positive characteristic</i>  |
| 2016-31 | <b>M. Dambrine, H. Harbrecht, M. D. Peters, B. Puig</b><br><i>On Bernoulli's free boundary problem with a random boundary</i>  |
| 2016-32 | <b>H. Harbrecht, J. Tausch</b><br><i>A fast sparse grid based space-time boundary element method for the nonstationary heat equation</i>   |
| 2016-33 | <b>S. Iula</b><br><i>A note on the Moser-Trudinger inequality in Sobolev-Slobodeckij spaces in dimension one</i>   |
| 2016-34 | <b>C. Bürli, H. Harbrecht, P. Odermatt, S. Sayasone, N. Chitnis</b><br><i>Mathematical analysis of the transmission dynamics of the liver fluke, <i>Opisthorchis viverrini</i></i> |
| 2017-01 | <b>J. Dölz and T. Gerig, M. Lüthi, H. Harbrecht and T. Vetter</b><br><i>Efficient computation of low-rank Gaussian process models for surface and image registration</i>           |
| 2017-02 | <b>M. J. Grote, M. Mehlin, S. A. Sauter</b><br><i>Convergence analysis of energy conserving explicit local time-stepping methods for the wave equation</i>                         |
| 2017-03 | <b>Collinear CM-points</b><br><i>Y. Bilu, F. Luca, D. Masser</i>   |
| 2017-04 | <b>P. Zaspel</b><br><i>Ensemble Kalman filters for reliability estimation in perfusion inference</i>   |