Computational approaches for parametric imaging of dynamic PET data

S Crisci[1], M Piana[2], V Ruggiero[1] and M Scussolini[3]

[1] Dipartimento di Matematica e Informatica, Università di Ferrara, Ferrara, Italy

[2] Dipartimento di Matematica, Università di Genova, and CNR-SPIN, Genova, Italy

[3] Dipartimento di Matematica, Università di Genova, Genova, Italy

E-mail: serena.crisci@unife.it

Abstract. Parametric imaging of nuclear medicine data exploits dynamic functional images in order to reconstruct maps of kinetic parameters related to the metabolism of a specific tracer injected in the biological tissue. From a computational viewpoint, the realization of parametric images requires the pixel-wise numerical solution of compartmental inverse problems that are typically ill-posed and nonlinear. In the present paper we introduce a fast numerical optimization scheme for parametric imaging relying on a regularized version of the standard affine-scaling Trust Region method. The validation of this approach is realized in a simulation framework for brain imaging and comparison of performances is made with respect to a regularized Gauss-Newton scheme and a standard nonlinear least-squares algorithm.

1. Introduction

Positron Emission Tomography (PET) [1] utilizes an isotope produced in a cyclotron to provide dynamical images of the metabolism-based isotope accumulation in the biological tissue. PET dynamic images of the tracer distribution are obtained by applying a reconstruction algorithm to the measured radioactivity and provide a reliable estimate of the tracer-related metabolism in the tissue [2, 3].

From a technical viewpoint, compartmental analysis [4, 5, 6] allows processing these dynamic PET data in order to estimate a set of physiological kinetic parameters that explain such metabolism in a quantitative manner (specifically, these parameters express the effectiveness of the tracer in changing its functional status within the tissue). Compartmental analysis requires, first, the formulation of a forward model for the tracer concentration represented by a Cauchy problem, in which the kinetic parameters are the coefficients of the differential equations; then, the numerical solution

of the corresponding nonlinear inverse problem, in which the kinetic parameters are the unknowns and the tracer concentrations in the tissue are the input data.

Relying on compartmental analysis, parametric imaging [7, 8, 9] allows the pixel-wise determination of the kinetic parameters by means of two possible alternative approaches. On the one hand, direct parametric imaging [10, 11] utilizes as input the PET raw sinograms and solves the inverse problem that relates them to the parameters; on the other hand, indirect parametric imaging [7, 8, 12, 13] is applied to the reconstructed PET images and solves pixel-wise the compartmental analysis problem. Direct approaches do not need the application of image reconstruction methods but have typically to deal with the intertwining of spatial and temporal correlations, which makes the optimization process more complex; this same optimization is more straightforward in indirect approaches but requires a higher computational burden, due to the need of solving a large number of nonlinear inverse problems.

The present paper deals with indirect parametric imaging and introduces a regularized optimization method for the solution of the nonlinear ill-posed inverse problem of compartmental analysis. The idea of the method is to introduce a regularizing strategy [14] in the standard affine-scaling Trust Region method [15, 16], which allows a better reduction of the numerical instabilities induced by the presence of the experimental noise in the measured data.

From a formal viewpoint, we prove a convergence result for the regularized algorithm, which enables a generalization to the non-negatively constrained case of the convergence analysis developed in [14] for the unconstrained problem. The numerical validation of the method is performed against synthetic data realized from an 'ad hoc' modification of the Hoffman Brain Phantom often used in PET and CT imaging (http://depts.washington.edu/petctdro/DROhoffman_main.html). Specifically, we mimicked a two-compartment experiment for the kinetics of [18F]fluoro-2-deoxy-D-glucose (FDG), which is the mostly utilized tracer in PET diagnostic and prognostic activities [17, 18, 19, 20, 21]. Using this simulation we could compare the computational effectiveness and reconstruction accuracy of the method with respect to the performances provided by two frequently used indirect parametric imaging methods.

The structure of the paper is as follows. Section 2 sets up the two-compartment problem for FDG kinetics. Section 3 describes in detail the nonlinear optimization method for the solution of this problem. Section 4 illustrates the validation experiment and its results. Our conclusions are offered in Section 5.

2. Compartmental analysis of dynamic PET data

Compartmental analysis of nuclear medicine data is the mathematical framework for the quantitative assessment of tracer kinetics in the biological tissue [22, 23, 24, 25, 26, 27]. The compartmental model of a specific organ comprises compartments representing the functional states of the tracer radioactive molecules (e.g.: physical location as
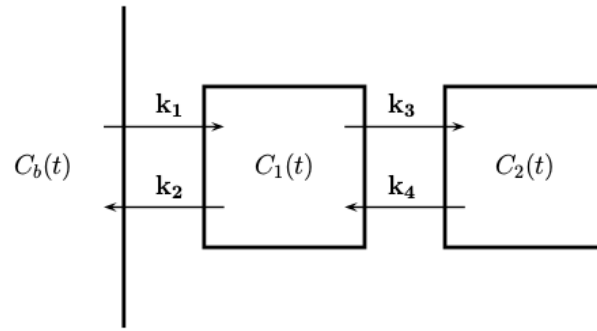
Figure 1: Compartmental model for FDG kinetics in a living tissue: external arterial blood compartment of concentration C_b, free tracer compartment of concentration C_f, metabolized tracer compartment of concentration C_m, and four kinetic parameters k_1, k_2, k_3, k_4.

intravascular space, extracellular space, intracellular space, etc., or chemical state as metabolic form, binding state, etc.), and kinetic parameters, which are the input/output tracer rates for each compartment. Figure 1 illustrates the standard two-compartmental model describing the FDG metabolism in the organ under consideration [28]. This model reproduces the main steps of the FDG path in a PET experiment. First, the tracer is injected into the blood with a concentration mathematically modelled by the Input Function (IF), here assumed as known and represented by the tracer concentration C_b in the arterial blood compartment. Then, the FDG metabolism within the tissue is characterized by two functional states: the free compartment with concentration C_f, associated to the tracer molecules outside the tissue cells, and the metabolized compartment with concentration C_m, associated to FDG molecules within the cytoplasm. Finally, the FDG kinetics is described by four rate constants connecting the model compartments: k_1 and k_2 describe the exchange rates between the input and free pools, and k_3 and k_4 describe the exchange rates at the basis of the phosphorylation/dephosphorylation process.

The system of Ordinary Differential Equations (ODEs) for the two-compartment model is

$$\frac{d\boldsymbol{C}}{dt}(t) = \dot{\boldsymbol{C}}(t) = \boldsymbol{M}\boldsymbol{C}(t) + k_1 C_b(t)\boldsymbol{e}_1, \tag{1}$$

where

$$\boldsymbol{C} = \begin{pmatrix} C_f \\ C_m \end{pmatrix}, \quad \boldsymbol{M} = \begin{pmatrix} -(k_2 + k_3) & k_4 \\ k_3 & -k_4 \end{pmatrix}, \quad \boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{2}$$

where t is the time variable, and $\pm k_i$, $i = 1, 2, 3, 4$, represent incoming and outgoing fluxes. In standard applications of compartmental analysis, the initial conditions are $C_f(0) = C_m(0) = 0$, meaning that the PET experiment starts at time $t = 0$ when there is no available tracer into the biological system. The analytical solution of (1) represents the forward model equation of determining the compartment concentrations given the

kinetic parameters, and takes the form

$$\boldsymbol{C}(t; \boldsymbol{k}) = k_1 \int_0^t \exp(\boldsymbol{M}(t - u)) C_b(u) \boldsymbol{e}_1 \, du, \tag{3}$$

where the entries of the vector $\boldsymbol{k} = (k_1, k_2, k_3, k_4)^T \in \mathbb{R}^4$ have to be non-negative real values.

The compartmental input function $C_b(t)$ can be obtained experimentally either from serial sampling of the arterial blood or reconstructed dynamic images [29], when a large arterial pool such as the left ventricle is in the field of view for many frames, or by using reference tissue methods [30, 31]. However, PET images cannot offer enough resolution power to provide information on $\boldsymbol{C}(t; \boldsymbol{k})$. Therefore the measurement equation

$$\tilde{C}_{(p,q)}(t; \boldsymbol{k}) = \boldsymbol{\alpha}^T \boldsymbol{C}_{(p,q)}(t; \boldsymbol{k}) + V_{(p,q)} C_b(t) \ , \quad \boldsymbol{\alpha} = \begin{pmatrix} 1 - V_{(p,q)} \\ 1 - V_{(p,q)} \end{pmatrix}, \tag{4}$$

should be added to equation (3) to connect the compartment model to the PET data. In this equation (p, q) represents a specific image pixel, $\tilde{C}_{(p,q)}$ denotes the measured tracer concentration at pixel (p, q) of the organ image, $\boldsymbol{C}_{(p,q)}$ is the formal analytic solution of (3), and $V_{(p,q)}$ is the fraction of tissue volume occupied by the blood. In general, the blood volume fraction depends on the pixel position, but within a homogeneous tissue it can be assumed as a known constant.

In equation (4) the unknown kinetic parameters are functions of (p, q) and therefore the inverse problem represented by this equation should be solved numerically and pixel-wise. Rather coarse approximations allow a linearization of this equation [32, 33]. However, the pixel-wise solution of the exact nonlinear equation requires the availability of an effective optimization scheme for the regularization of the ill-posed nonlinear compartmental inverse problem represented by the equation and eventually for the reconstruction of the four parametric images associated to k_1, k_2, k_3, and k_4.

3. Computational approaches for nonlinear ill-posed problems

The compartmental inverse problem described in the previous section is a special case of the following more general formulation. Given a set of measurements $\boldsymbol{y}^0$ of tracer concentration provided by PET images, corresponding to a finite sample of N time points $t_1, ..., t_N$, we have to determine the kinetic parameters $\boldsymbol{k} \in \mathbb{R}^n$, $n \le N$, by solving the non-negatively constrained nonlinear system

$$\boldsymbol{F}(\boldsymbol{k}) = \boldsymbol{y}^0, \qquad \text{subject to } \boldsymbol{k} \ge 0. \tag{5}$$

Here $\boldsymbol{y}^0 = (\tilde{C}_{(p,q)}(t_1, \boldsymbol{k}), ..., \tilde{C}_{(p,q)}(t_N, \boldsymbol{k}))^T$, and $\boldsymbol{F} \colon \mathbb{R}^n \to \mathbb{R}^N$ is the continuously differentiable function at the right hand side of (4). In real experiments a noisy version $\boldsymbol{y}^\delta$ of $\boldsymbol{y}^0$ is at disposal, where δ is a known bound on the measurement error, with $\delta \le \|\boldsymbol{y}^0\|$. A standard approach to address equation (5) [34, 35] consists in

Computational approaches for parametric imaging of dynamic PET data 5

approximating a solution $\boldsymbol{k}^{\dagger}$ of this nonlinear system by solving the following non-negatively nonlinear least squares problem via an iterative regularization technique with semiconvergent behaviour:

$$\min_{\boldsymbol{k} \geq 0} \ \Phi(\boldsymbol{k}) \equiv \frac{1}{2}\|\boldsymbol{y}^{\delta} - \boldsymbol{F}(\boldsymbol{k})\|^2. \tag{6}$$

In view of the discrepancy principle [34], the iterative method is stopped at the iteration $\bar{j}(\delta)$ satisfying the following condition

$$\|\boldsymbol{y}^{\delta} - \boldsymbol{F}(\boldsymbol{k}^{\bar{j}(\delta)})\| \leq \tau\delta < \|\boldsymbol{y}^{\delta} - \boldsymbol{F}(\boldsymbol{k}^{j})\| \quad 0 \leq j \leq \bar{j}(\delta), \tag{7}$$

for a suitable $\tau > 1$.

In this section, we describe a method for computing a regularized solution of problem (6); in particular, we combine the regularizing approach developed in [14] for unconstrained ill-posed problems with the affine scaling trust-region (TR) schemes for a box-constrained minimization problem [15, 36]. The key point to link these methods is represented by the following Proposition 1, which shows that possible projection steps do not prevent the convergence of the iterative scheme. Therefore, the main contribution of this section is to show that the theoretical framework developed for the unconstrained problem [14] still holds also in the non-negatively constrained case.

3.1. A regularizing affine scaling trust-region method for non-negatively nonlinear least-squares problems

For unconstrained nonlinear ill-posed least-squares problems, the state-of-the-art approaches are the regularized Levenberg-Marquadt (LM) method, proposed by Hanke [37], and its reformulation within a Trust-Region (TR) framework, proposed by Wang et al. [14] and, more recently, by Bellavia et al. [38]. As in the standard TR algorithm, the regularizing TR iteration requires to compute, at each iteration, a trial step $\boldsymbol{p}^{j}$, by minimizing the quadratic model $m_j(\boldsymbol{p})$ within a region around the current iterate $\boldsymbol{k}^{j}$:

$$\min_{\boldsymbol{p}} \ m_j(\boldsymbol{p}) \equiv \frac{1}{2}\boldsymbol{p}^T\boldsymbol{B}^j\boldsymbol{p} + \boldsymbol{p}^T\boldsymbol{g}^j$$
$$\text{s.t.} \ \ \|\boldsymbol{p}\| \leq \Delta_j \tag{8}$$

where $\boldsymbol{B}^j \equiv \boldsymbol{J}(\boldsymbol{k}^j)^T\boldsymbol{J}(\boldsymbol{k}^j)$ is the Gauss-Newton approximation of the Hessian of $\boldsymbol{F}$, $\boldsymbol{g}^j \equiv \nabla\Phi(\boldsymbol{k}^j) = \boldsymbol{J}(\boldsymbol{k}^j)^T(\boldsymbol{F}(\boldsymbol{k}^j)-\boldsymbol{y}^{\delta})$ and Δ_j denotes the TR radius; this can be expanded or reduced depending on whether a sufficient reduction of the model is achieved or not, i.e. if the ratio $\rho_j = \dfrac{\Phi(\boldsymbol{k}^j + \boldsymbol{p}^j) - \Phi(\boldsymbol{k}^j)}{m_j(\boldsymbol{p}^j)}$ between the actual reduction in the objective functional and the predicted reduction in the quadratic model is lower than some positive threshold $\beta \in (0, 1)$. The regularizing property is accomplished by requiring that the TR constraint is active at the solution, i.e., the solution $\boldsymbol{p}^j$ of (8) must be such that $\|\boldsymbol{p}^j\| = \Delta_j$ so that the associated Lagrange multiplier α_j plays the role of a penalization

parameter in a Tikhonov-like regularization. Indeed, given $\boldsymbol{k}^j$, the new iterate can be viewed as the solution of the penalized subproblem arising at the iteration of LM method:

$$\boldsymbol{k}^{j+1} = \boldsymbol{k}^j + \boldsymbol{p}^j = \arg\min_{\boldsymbol{k}}\{\|\boldsymbol{y}^\delta - \boldsymbol{F}(\boldsymbol{k}^j) - \boldsymbol{J}(\boldsymbol{k}^j)(\boldsymbol{k} - \boldsymbol{k}^j)\|^2 + \alpha_j\|\boldsymbol{k} - \boldsymbol{k}^j\|^2\}. \qquad (9)$$

This regularization technique for an unconstrained problem can be combined with the TR methods for box-constrained nonlinear least-squares problems. To this aim, we introduce a regularizing technique in the affine-scaling TR method [16, 15, 36] requiring that the TR constraint in the subproblem (8) is active at the solution. In particular, given $\boldsymbol{k}^j > 0$ and $\boldsymbol{g}^j \neq 0$, we find the solution $\alpha_j > 0$ of the nonlinear equation $\Delta_j - \|\boldsymbol{p}(\alpha)\| = 0$, where $\boldsymbol{p}(\alpha) = (\boldsymbol{J}(\boldsymbol{k}^j)^T\boldsymbol{J}(\boldsymbol{k}^j) + \alpha\boldsymbol{I}_n)^{-1}\boldsymbol{J}(\boldsymbol{k}^j)^T(\boldsymbol{y}^\delta - \boldsymbol{F}(\boldsymbol{k}^j))$. By setting $\boldsymbol{p}^j = \boldsymbol{p}(\alpha_j)$, in order to ensure the strict feasibility of a new iterate, the i-th entry of $\boldsymbol{k}^{j+1} = \boldsymbol{k}^j + \bar{\boldsymbol{p}}^j$ is computed in accordance with the following rule:

$$\bar{\boldsymbol{p}}_i^j = \begin{cases} \boldsymbol{p}_i^j & \text{if } (\boldsymbol{k}^j + \boldsymbol{p}^j)_i > 0 \\ t(\Pi(\boldsymbol{k}^j + \boldsymbol{p}^j) - \boldsymbol{k}^j)_i & \text{if } (\boldsymbol{k}^j + \boldsymbol{p}^j)_i \leq 0 \end{cases} \qquad (10)$$

where $\Pi(\cdot)$ denotes the Euclidean projection onto the non-negative orthant of $\mathbb{R}^n$ and $t \in (0,1)$. Clearly, in view of the properties of the projection operator, $\|\bar{\boldsymbol{p}}^j\| \leq \|\boldsymbol{p}^j\|$.

As emphasized in [15], a key point to assure the convergence of the affine-scaling TR method is that the new iterate $\boldsymbol{k}^j + \bar{\boldsymbol{p}}^j$ must be able to achieve at least as much reduction in the quadratic model as the one achieved by the generalized Cauchy point $\boldsymbol{p}_C^j = -\lambda_C^j \boldsymbol{D}(\boldsymbol{k}^j)\boldsymbol{g}^j$, where $\boldsymbol{D}(\boldsymbol{k})$ is a diagonal matrix such that

$$\boldsymbol{D}(\boldsymbol{k})_{i,i} = \begin{cases} |\boldsymbol{k}_i| & \text{if } \nabla\Phi(\boldsymbol{k})_i \geq 0 \\ 1 & \text{otherwise} \end{cases} \qquad (11)$$

and λ_C^j is defined as follows

$$\lambda_C^j = \begin{cases} \min\left\{\dfrac{\Delta_j}{\|\boldsymbol{D}(\boldsymbol{k}^j)\boldsymbol{g}^j\|}, \dfrac{\|\boldsymbol{D}^{1/2}(\boldsymbol{k}^j)\boldsymbol{g}^j\|^2}{\|\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{D}(\boldsymbol{k}^j)\boldsymbol{g}^j\|^2}\right\} & \text{if } \left(\boldsymbol{k}^j - \hat{\lambda}_C^j\boldsymbol{D}(\boldsymbol{k}^j)\boldsymbol{g}^j\right)_i > 0 \\ t\min_i\left\{\dfrac{\boldsymbol{k}_i^j}{(\boldsymbol{D}(\boldsymbol{k}^j)\boldsymbol{g}^j)_i} : (\boldsymbol{D}(\boldsymbol{k}^j)\boldsymbol{g}^j)_i > 0\right\} & \text{if } \left(\boldsymbol{k}^j - \hat{\lambda}_C^j\boldsymbol{D}(\boldsymbol{k}^j)\boldsymbol{g}^j\right)_i \leq 0 \end{cases} \qquad (12)$$

with $t \in (0,1)$. If $\rho_j^C = \dfrac{m_j(\bar{\boldsymbol{p}}^j)}{m_j(\boldsymbol{p}_C^j)} > \beta_C \in (0,1)$ and $\rho_j > \beta \in [0.25, 1)$, the current trial step is accepted and the next iterate is updated as $\boldsymbol{k}^{j+1} = \boldsymbol{k}^j + \bar{\boldsymbol{p}}^j$, otherwise the TR radius is reduced. In particular, if $\rho_j^C \leq \beta_C$, the unsatisfactory reduction of the quadratic model at $\bar{\boldsymbol{p}}^j$ with respect to the reduction obtained with the generalized Cauchy step highlights that we have to increase the effect of the regularization term by reducing the TR radius and computing a new reduced step; this vector tend to line up with $\boldsymbol{g}^j$ and the new generalized Cauchy step, so that the sufficient reduction of the quadratic model is obtained. Furthermore, when $\|\boldsymbol{g}^j\| \neq 0$, after a successful iteration of

Computational approaches for parametric imaging of dynamic PET data 7

the method, the TR radius can be further adjusted by increasing or reducing it within a prefixed range, accordingly to a strategy proposed in [38] (see Eq. (5.5)-(5.6)), as follows:

$$\Delta_{j+1} = \max\left(\mu_{j+1}\|\boldsymbol{F}(\boldsymbol{k}^{j+1}) - \boldsymbol{y}^{\delta}\|, 1.2\frac{(1-q)\|\boldsymbol{g}^{j+1}\|}{\|\boldsymbol{B}^{j+1}\|}\right), \qquad (13)$$

where

$$\mu_{j+1} = \begin{cases} \theta\mu_j & \text{if } q_j < q \\ \frac{\mu_j}{\eta} & \text{if } q_j > 1.1q \\ \mu_j & \text{otherwise} \end{cases} \qquad (14)$$

with $q \in (0,1)$, $q_j = \dfrac{\|\boldsymbol{y}^{\delta} - \boldsymbol{F}(\boldsymbol{k}^j) - \boldsymbol{J}(\boldsymbol{k}^j)\bar{\boldsymbol{p}}^j\|}{\|\boldsymbol{y}^{\delta} - \boldsymbol{F}(\boldsymbol{k}^j)\|}$ and $\theta, \eta \in (0,1)$. The regularizing affine-scaling TR method, called in the following reg-AS-TR, is summarized in Algorithm 1; for data affected by noise, the stopping criterion is based on the discrepancy principle (7).

The convergence analysis of reg-AS-TR requires to prove Proposition 1, which is analogous to Theorem 2.1 of [14], i.e., we need to prove that the distance between $\boldsymbol{k}^j$ and the exact solution $\boldsymbol{k}^{\dagger}$ decreases for $j \leq \bar{j}(\delta)$. To this aim, we give two essential assumptions on the local properties of the nonlinear system (5), very similar to the ones used in [37, 14] to handle ill-posed problems.

A1. Given an initial guess $\boldsymbol{k}^0 > 0$, there exist $\nu, c > 0$ such that $\boldsymbol{k}^{\dagger} \in B_{\nu}(\boldsymbol{k}^0) = \{\boldsymbol{k} \geq 0 : \|\boldsymbol{k} - \boldsymbol{k}^0\| \leq \nu\}$ and for all $\bar{\boldsymbol{k}}, \boldsymbol{k} \in B_{2\nu}(\boldsymbol{k}^0) = \{\boldsymbol{k} \geq 0 : \|\boldsymbol{k} - \boldsymbol{k}^0\| \leq 2\nu\}$ the following condition holds:

$$\|\boldsymbol{F}(\bar{\boldsymbol{k}}) - \boldsymbol{F}(\boldsymbol{k}) - \boldsymbol{J}(\boldsymbol{k})(\bar{\boldsymbol{k}} - \boldsymbol{k})\| \leq c\|\bar{\boldsymbol{k}} - \boldsymbol{k}\|\|\boldsymbol{F}(\bar{\boldsymbol{k}}) - \boldsymbol{F}(\boldsymbol{k})\| \qquad (15)$$

A2. $\|\boldsymbol{k}^0 - \boldsymbol{k}^{\dagger}\| < \min(\frac{q}{c}, \nu)$ for noisy-free data ($\delta = 0$) and $\|\boldsymbol{k}^0 - \boldsymbol{k}^{\dagger}\| < \min(\frac{q\tau - 1}{c(1+\tau)}, \nu)$ for noisy data ($\delta > 0$) with $\tau > 1/q$.

We highlight that, when at the first steps of the algorithm these assumptions are not verified, the initial iterations can enable to restrict the domain so that they hold from a certain j. Now, we are able to state the following key proposition (for the proof see the Appendix).

Proposition 1. *Let assume that* $\boldsymbol{J}(\boldsymbol{k}^j)^T\boldsymbol{J}(\boldsymbol{k}^j) + \alpha_j\boldsymbol{I}_n$ *is positive definite,* $\boldsymbol{g}^j \neq 0$ *and*

$$\|\boldsymbol{y}^{\delta} - \boldsymbol{F}(\boldsymbol{k}^j) - \boldsymbol{J}(\boldsymbol{k}^j)\bar{\boldsymbol{p}}^j\| \geq q\|\boldsymbol{y}^{\delta} - \boldsymbol{F}(\boldsymbol{k}^j)\| \qquad (16)$$

for a suitable $q \in (0,1)$*, with* $j \geq 0$ *and with* $j \leq \bar{j}(\delta)$ *when* $\delta > 0$*. Moreover, let assume that, for a suitable* $\gamma_{\delta} > 1$*, the following condition holds for* $\boldsymbol{k}^j > 0$*:*

$$\|\boldsymbol{y}^{\delta} - \boldsymbol{F}(\boldsymbol{k}^j) - \boldsymbol{J}(\boldsymbol{k}^j)(\boldsymbol{k}^{\dagger} - \boldsymbol{k}^j)\| \leq \frac{q}{\gamma_{\delta}}\|\boldsymbol{y}^{\delta} - \boldsymbol{F}(\boldsymbol{k}^j)\|. \qquad (17)$$

Thus we have

$$\|\boldsymbol{k}^{\dagger} - \boldsymbol{k}^j\|^2 - \|\boldsymbol{k}^{\dagger} - \boldsymbol{k}^{j+1}\|^2 > \frac{2t(\gamma_{\delta} - 1)q}{\gamma_{\delta}}\|\boldsymbol{y}^{\delta} - \boldsymbol{F}(\boldsymbol{k}^j)\|\|\boldsymbol{v}^j\| \qquad (18)$$

with $\boldsymbol{v}^j = (\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{J}(\boldsymbol{k}^j)^T + \alpha_j\boldsymbol{I}_N)^{-1}(\boldsymbol{y}^{\delta} - \boldsymbol{F}(\boldsymbol{k}^j))$.

We remark that condition (17) with $j = 0$ follows directly from the assumptions A1-A2 with $\gamma_0 \geq \dfrac{q}{c\|\boldsymbol{k}^\dagger - \boldsymbol{k}^0\|} > 1$ for noise-free data. For $\delta > 0$, condition (17) with $j = 0$ is obtained with $\gamma_\delta \geq \dfrac{q\tau}{1 + c\|\boldsymbol{k}^\dagger - \boldsymbol{k}^0\|(1 + \tau)} > 1$, combining the assumptions A1-A2 with the inequality $\dfrac{\|\boldsymbol{y}^\delta - \boldsymbol{F}(\boldsymbol{k}^j)\|}{\delta} > \tau$ which is satisfied for $j \leq \bar{j}(\delta)$ (see (7)). As a consequence of Proposition 1, $\boldsymbol{k}^1$ belongs to $B_{2\nu}(\boldsymbol{k}^0)$ and to $B_\nu(\boldsymbol{k}^\dagger)$. Therefore, for the same argument above, condition (17) holds by induction for $j \geq 0$ and for $j \leq \bar{j}(\delta)$ when $\delta > 0$; as a consequence, the sequence $\|\boldsymbol{k}^j - \boldsymbol{k}^\dagger\|$ is decreasing.

Based on the above proposition and the convergence results of the affine-scaling TR methods, the same properties of the regularizing TR method for an unconstrained nonlinear least-squares problem can be easily extended to the non-negatively constrained case. Under Assumptions A1-A2 on the exact solution $\boldsymbol{k}^\dagger$, reg-AS-TR terminates after $\bar{j}(\delta) < \infty$ iterations, where δ is the noise level on the data, whereas for $\delta = 0$ or $\delta \to 0$ the sequence $\{\boldsymbol{k}^j\}$ generated by Algorithm 1 converges to a solution of the original problem.

As a final remark, we point out that the ill-posedness and nonlinearity of the method, together with the local properties of reg-AS-TR imply that the effectiveness of our numerical scheme may be significantly influenced by the accuracy of both the initialization and the noise estimate. The reliability with which these two aspects are addressed is an essential requirement for the accuracy of the reconstruction results.

Algorithm 1 Regularizing affine-scaling Trust-Region (reg-AS-TR) method

---

      Initialize: Choose $\boldsymbol{k}^0 > 0, \beta \in [0.25, 1), \gamma, \beta_C \in (0, 1)$,
      $0 < \Delta_{min} < \Delta_{max}, \ q \in (0, 1), \ \mu_0 = 0.001$
      $j = 0$;

1: **while** the stopping rule is not satisfied **do**
2:      Evaluate $\boldsymbol{B}_j = \boldsymbol{J}(\boldsymbol{k}^j)^T J(\boldsymbol{k}^j)$ and $\boldsymbol{g}^j = \boldsymbol{J}^T(\boldsymbol{k}^j)(\boldsymbol{F}(\boldsymbol{k}^j) - \boldsymbol{y}^\delta)$
3:      $\Delta_j = \max\left(\mu_j\|\boldsymbol{F}(\boldsymbol{k}^j) - \boldsymbol{y}^\delta\|, \ 1.2\dfrac{(1-q)\|\boldsymbol{g}^j\|}{\|\boldsymbol{B}_j\|}\right) \in [\Delta_{min}, \Delta_{max}]$
4:      **repeat**
5:         Compute a feasible solution $\bar{\boldsymbol{p}}^j$ to the trust-region problem (8)
6:         Compute the Cauchy point $\boldsymbol{p}_C^j$
7:         Compute $\rho_j^C = \dfrac{m_j(\bar{\boldsymbol{p}}^j)}{m_j(\boldsymbol{p}_C^j)}$ and $\rho^j = \dfrac{\Phi(\boldsymbol{k}^j + \bar{\boldsymbol{p}}^j) - \Phi(\boldsymbol{k}^j)}{m_j(\bar{\boldsymbol{p}}^j)}$
8:         If $\rho_j^C \leq \beta_C$ or $\rho_j \leq \beta$ then set $\Delta_j = \gamma\Delta_j$
9:      **until** $\rho_j^C > \beta_C$ and $\rho_j > \beta$
10:     $\boldsymbol{k}^{j+1} = \boldsymbol{k}^j + \bar{\boldsymbol{p}}^j$
11:     j = j+1
12:     Update μ_{j+1} as specified in (13)
13: **end while**

---

4. Numerical experiments

The numerical validation of reg-AS-TR is performed using synthetic PET data generated by means of a digital phantom of the human brain. All simulations were realized on a

workstation equipped with an Intel Xeon QuadCore E5620 processor at 2,40 GHz and 18 Gb of RAM, by implementing the method in the Matlab® R2019a environment.

4.1. Simulation setting

The starting point was the 3D Hoffman Digital Reference Object, a digital representation of the Hoffman Brain Phantom used in PET and CT imaging studies, freely available from the Imaging Research Laboratory of the Department of Radiology at the Medical Center of the University of Washington (http://depts.washington.edu/petctdro/DROhoffman_main.html).

The 3D Hoffman brain phantom is composed of 250 slices, covering the entire head, consisting in black/white images of size 256×256. We reduced the image size to 128×128 to resemble typical PET acquisitions, preserving the shape and features of the original phantom. For sake of simplicity, we selected a middle slice including eight anatomical structures that can be subdivided into the four homogeneous functional regions in Figure 2(a): grey matter (region 1), white matter (region 2), basal ganglia (region 3), and thalamus (region 4). Then, for each region, we assigned a ground-truth set of rate constants of the two-compartment model for FDG kinetics (described in Section 2) and a specific blood volume fraction V. The numerical values of such parameters, as reported in Table 1, have been chosen in order to reproduce a realistic framework for the FDG uptake of a human brain [39, 40, 41, 42]. The ground-truth parametric images are shown in Figure 3.

In order to model the IF we implemented the following procedure [43]. We considered a mathematical function (see Eq. (2) in [43]) consisting of an increasing linear component followed by a tri-exponential decay; we fitted the free parameters of this function against measurements for 80 subjects; we selected the median estimated parameters computed over all 80 subjects (see Table 2 in [43]), a median initial distribution volume (12.7 L corresponding to 0.1683 L/kg body weight), and an Administered Activity (AA) of 350 MBq (typical of human PET acquisitions). The resulting simulated IF is shown in Figure 2(d).

The dynamic PET data were generated by solving the compartmental forward problem for each pixel of the processed Hoffman brain image. In particular, the two-compartment concentrations were evaluated by means of the integral equation (3) with the ground truth values of the compartmental parameters and the simulated IF, at 28 time frames (6×10 sec, 3×20 sec, 3×30 sec, 4×60 sec, 3×150 sec, 9×300 sec) with a time sampling typical of standard PET experiments, for a total time interval of 60 minutes. Then, the measurement equation (4) was computed to create the time concentration curves characteristic for each brain region (Figure 2(c)). The last frame of the obtained dynamic PET images is reported in Figure 2(b).

Once the noise-free dynamic PET images were obtained, we projected the images into the sinogram space by means of the Radon transform, and we added Poisson noise to the projected data through the Matlab function *poissrnd*. In this way, we created

Computational approaches for parametric imaging of dynamic PET data 10

ten independent identically-distributed noisy data. In addition to the noise-free IF case, we considered two further instances where the IF was perturbed by two Gaussian noise levels: $C_b^c = C_b(t_i)(1 + c \cdot r)$, for time points t_i, $i = 1, \ldots, N$, where r is randomly generated from a standard normal distribution of mean 0 and standard deviation 1, and $c = 0.10, \ 0.20$ (Figure 2(d)).

(a) Brain regions

(b) Last frame PET data

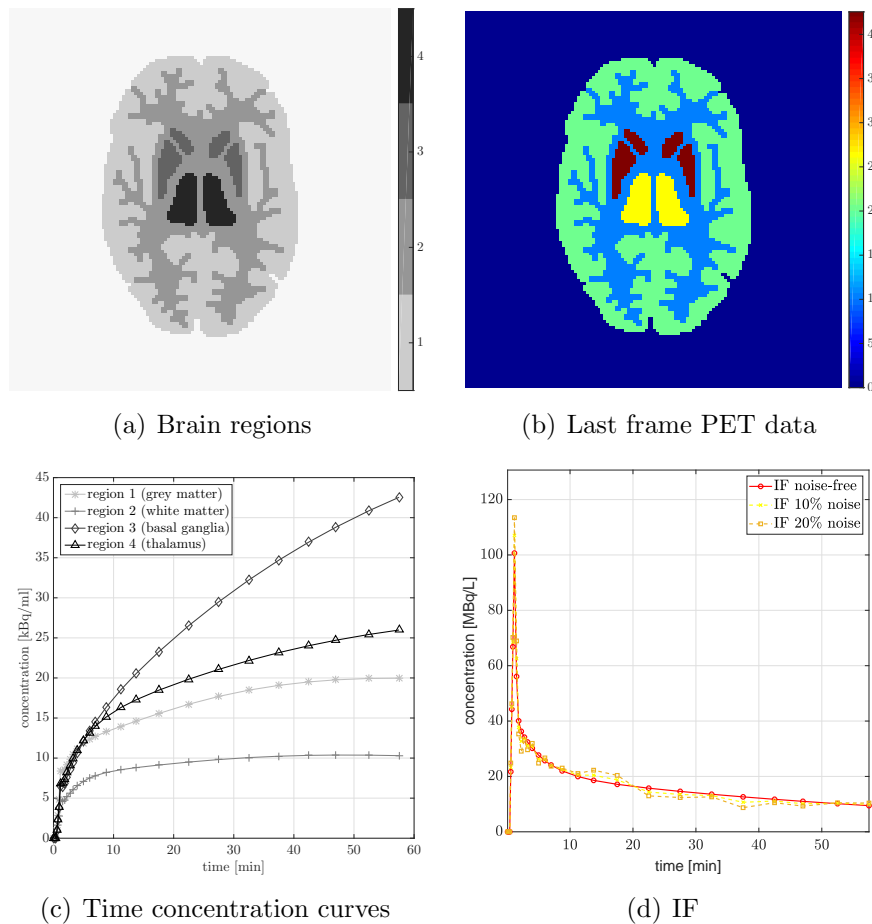(c) Time concentration curves

(d) IF

Figure 2: Simulation layout: (a) 128×128 Hoffman brain image subdivided into four homogeneous regions; (b) last frame of the simulated PET dynamic data ; (c) time-dependent concentration curves of all brain regions; (d) simulated IF for AA of 350 MBq as noise-free, and with 10%, and 20% of noise.

4.2. Setup of the algorithms

The parametric reconstruction by means of reg-AS-TR was performed as follows.

In order to remove blurring artifacts from the images of each dataset, we applied a well-known deblurring technique based on the minimization of the Kullback-Leibler divergence with a smooth total-variation regularization term referred to as *hypersurface potential*; this minimization is performed by means of the Scaled Gradient Projection

AUTHOR SUBMITTED MANUSCRIPT - IP-102361

Table 1: Ground truth values of the kinetic parameters k_1, k_2, k_3, k_4, (min^{-1}) and the blood volume fraction V, for each one of the four homogeneous regions.

	k_1	k_2	k_3	k_4	V
region 1	0.100	0.250	0.100	0.020	0.050
region 2	0.050	0.150	0.050	0.020	0.030
region 3	0.070	0.050	0.100	0.007	0.040
region 4	0.080	0.100	0.050	0.007	0.050

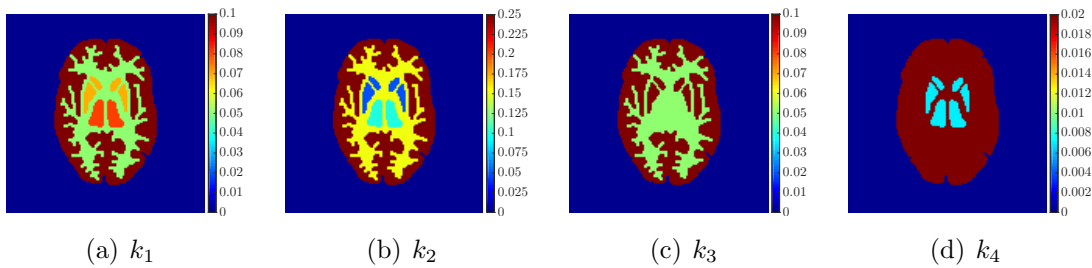

(a) k_1 (b) k_2 (c) k_3 (d) k_4

Figure 3: Ground truth parametric images.

(SGP) method proposed in [44] (see also [45]), starting from the inverse Radon transform of the noisy sinogram data. The deblurring procedure exploits the parallel toolbox of Matlab enabling the use of `GPUarray` and it requires about 7 minutes overall.

The stopping criterion of reg-AS-TR is the following:

$$\epsilon_j < \tau_1 \quad \text{or} \quad \left(\epsilon_j < \tau_2 \quad \text{and} \quad \left| 1 - \frac{\epsilon_{j-1}}{\epsilon_j} \right| < 10^{-2} \right) \tag{19}$$

where $\epsilon_j = \| \boldsymbol{y}^\delta - \boldsymbol{F}(\boldsymbol{k}^j) \|$, τ_1 is the sample standard deviation computed at the current pixel and τ_2 is a multiple of τ_1, which changes accordingly when the procedure switches between boundary ($\tau_2 = 10\tau_1$) and inner pixels ($\tau_2 = 3\tau_1$) of a region. In addition, if condition (19) is not satisfied, the execution terminates when stagnation or the maximum number of iterations are reached.

The stopping rule implemented allows to diversify the initialization procedure of reg-AS-TR. In general, the initial vector is randomly chosen in an interval determined by a priori knowledges on the physiology. However, when the current pixel is strictly inside a functional region and some neighboring pixels have been already successfully processed, the initialization value is the mean over the values obtained on these neighboring pixels.

The reconstruction accuracy of reg-AS-TR has been assessed by comparison with both the ground truth and the parametric images provided by a recently introduced regularized Gauss-Newton method (reg-GN) [13]. For sake of comparison, the setup of reg-GN is coherent with what is done in that paper, i.e.:

- *Deblurring.* The noise on the PET datasets was reduced by applying a Gaussian

Computational approaches for parametric imaging of dynamic PET data 12

smoothing filter (mean 0, standard deviation 1, window 3×3) directly to the noisy PET images.

- *Initialization.* The starting point of the kinetic parameters was chosen randomly in intervals determined by knowledge on the physiology.

- *Stopping criterion.* The iterative scheme is stopped when the relative error between the experimental dynamic concentration and the model-predicted one is less than an appropriate threshold, or the maximum number of iterations is reached.

4.3. Results

Figure 4, Figure 5, and Figure 6 show the mean images computed over the ten reconstructions obtained by the methods reg-AS-TR, reg-GN, and by the Matlab routine *lsqcurvefit* implementing a standard Trust-Region-Reflective least-squares algorithm [15, 46]. We used the noise-free IF and the perturbed IF with 10% and 20% of noise, respectively. Figure 7 contains mean and standard deviation values of the kinetic parameters computed over the ten reconstructions and over each one of the four homogeneous regions, for each one of the three noise levels on the IF.

Finally, Figure 8 represents the last frame of the dynamic PET data reconstructed with the mean parametric values returned by reg-AS-TR, reg-GN, and *lsqcurvefit*, with respect to the noise-free, 10%-noise, and 20%-noise IFs.

5. Comments and conclusions

In general, reg-AS-TR and *lsqcurvefit* seem to provide similar mean reconstructions, although uncertainties associated to *lsqcurvefit* are significantly bigger. On the other hand reg-GN seems to systematically underestimate the parameter values within region 1. Furthermore and as expected, for all methods the quality of the parametric reconstructions deteriorates with increasing noise levels; this is more clear from the k_3 and k_4 parametric images, probably due to the different sensitivities of the data with respect to the model parameters [47]. In reg-GN and some artifacts can be observed at the edges of the homogeneous regions, especially around region 1 and region 2, whereas the effect of regularization in reg-AS-TR results in a reduced presence of artifacts while the structure of the regions is preserved. This general trend is confirmed by the error-bar plots of Figure 7. Finally, the frames in Figure 8 corresponding to reg-AS-TR show a significant improvement of the image quality with respect to what is provided by the other two approaches.

The mean execution time for a single parametric reconstruction differs considerably between the reconstruction methods: reg-AS-TR requires about 20 minutes, reg-GN needs a computational time in the range $75 - 120$ minutes with run time increasing with noise level on IF (as a consequence of the stopping criterion implemented) and Matlab *lsqcurvefit* takes about 90 minutes. Therefore reg-AS-TR seems to be the most efficient approach in terms of both computational time and reconstruction accuracy.

Computational approaches for parametric imaging of dynamic PET data　　　13

Next steps for this piece of research activity will be the validation of reg-AS-TR against several experimental datasets in the case of both humans' and small animals' dynamic PET images. Further, we are going to generalize reg-AS-TR to the case of more complex compartmental models like the ones for the assessment of FDG kinetics in liver [26] and kidneys [27].
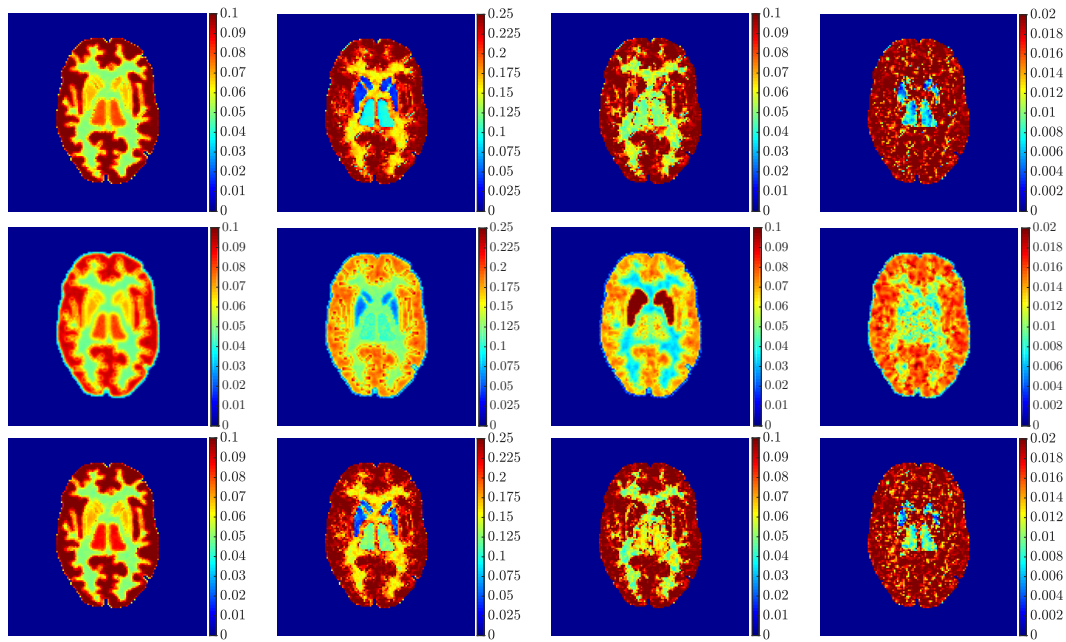

Figure 4: From left to right: mean parametric images corresponding to k_1, k_2, k_3, k_4, obtained by using reg-AS-TR (first row), reg-GN (second row), *lsqcurvefit* (third row). Case noise-free IF.

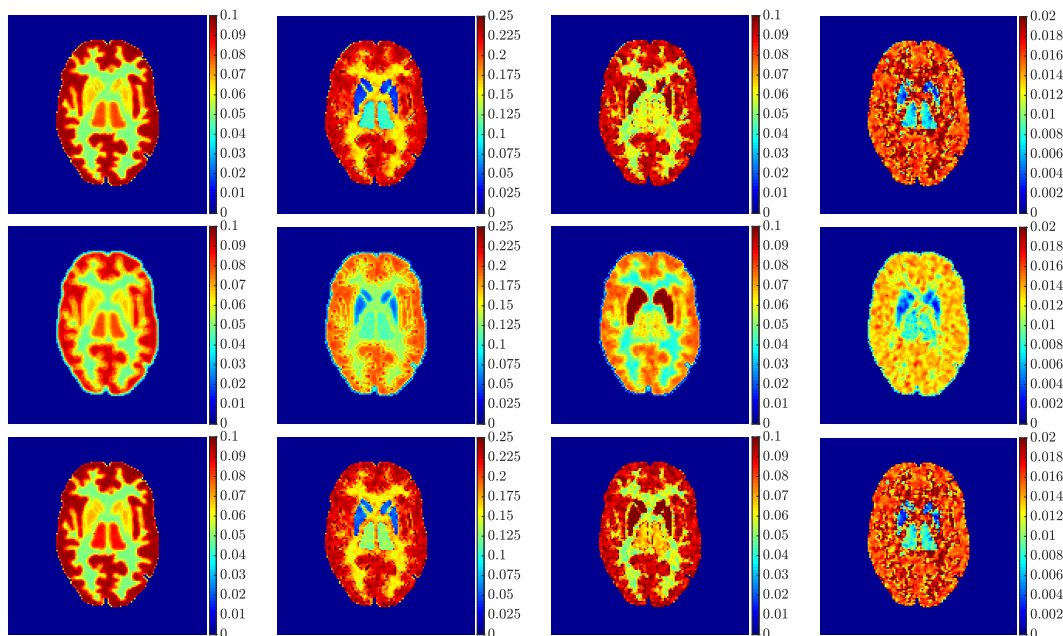*Computational approaches for parametric imaging of dynamic PET data*　　　14

Figure 5: From left to right: mean parametric images corresponding to k_1, k_2, k_3, k_4, obtained by using reg-AS-TR (first row), reg-GN (second row), *lsqcurvefit* (third row). Case 10%-noise IF.
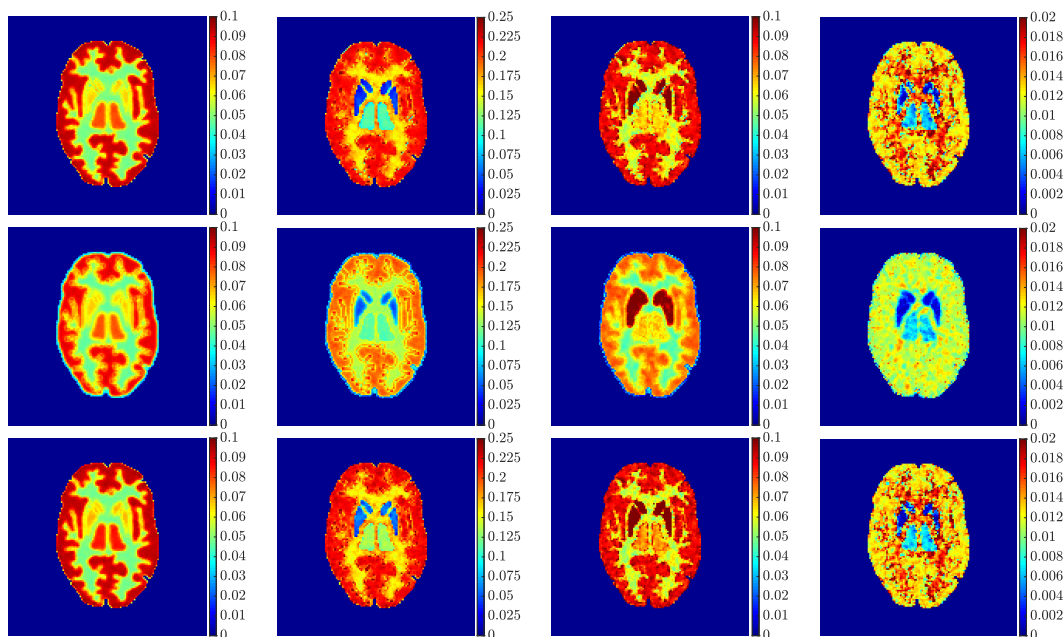

Figure 6: From left to right: mean parametric images corresponding to k_1, k_2, k_3, k_4, obtained by using reg-AS-TR (first row), reg-GN (second row), *lsqcurvefit* (third row). Case 20%-noise IF.

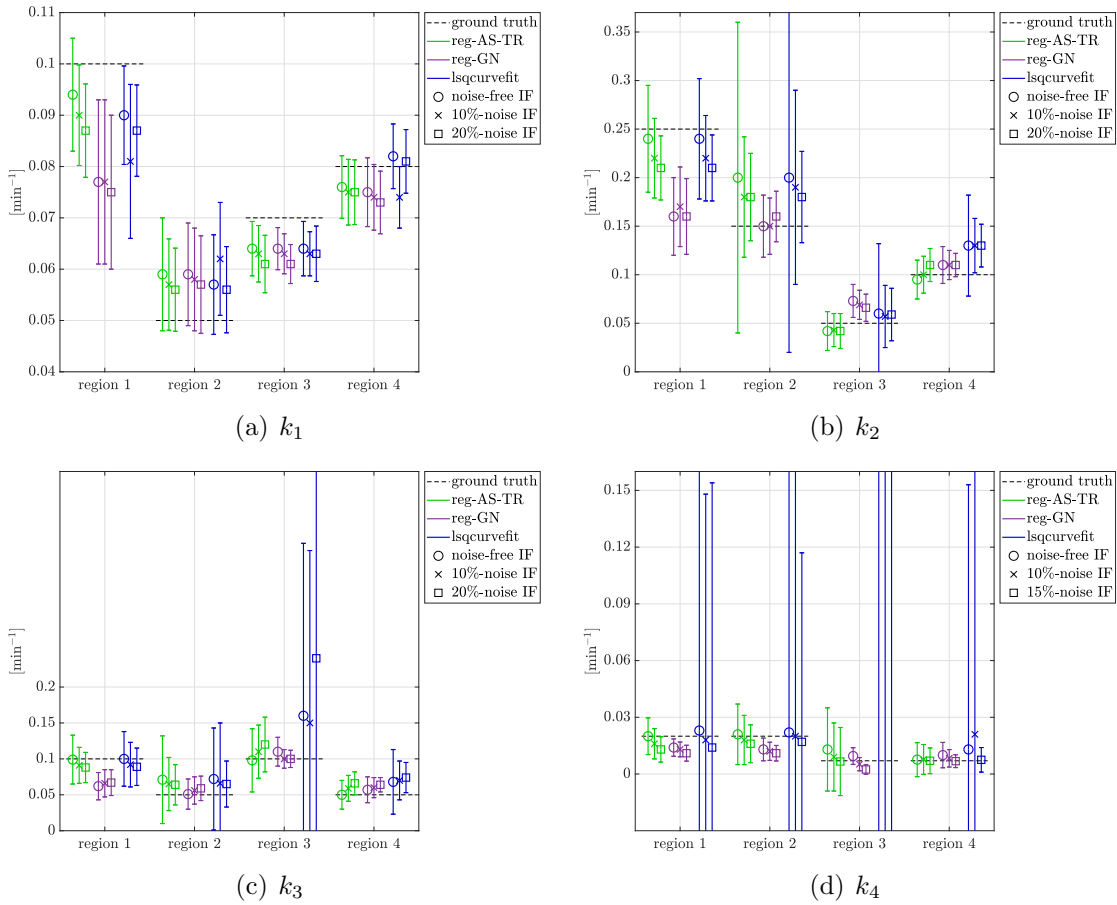*Computational approaches for parametric imaging of dynamic PET data* 15

Figure 7: Mean and standard deviation values of the kinetic parameters for the four homogeneous region as error bars over the reconstructions: reg-AS-TR (green bars), reg-GN (purple bars), *lsqcurvefit* (blue bars); noise-free IF (circle), 10%-noise IF (cross), 20%-noise IF (square).
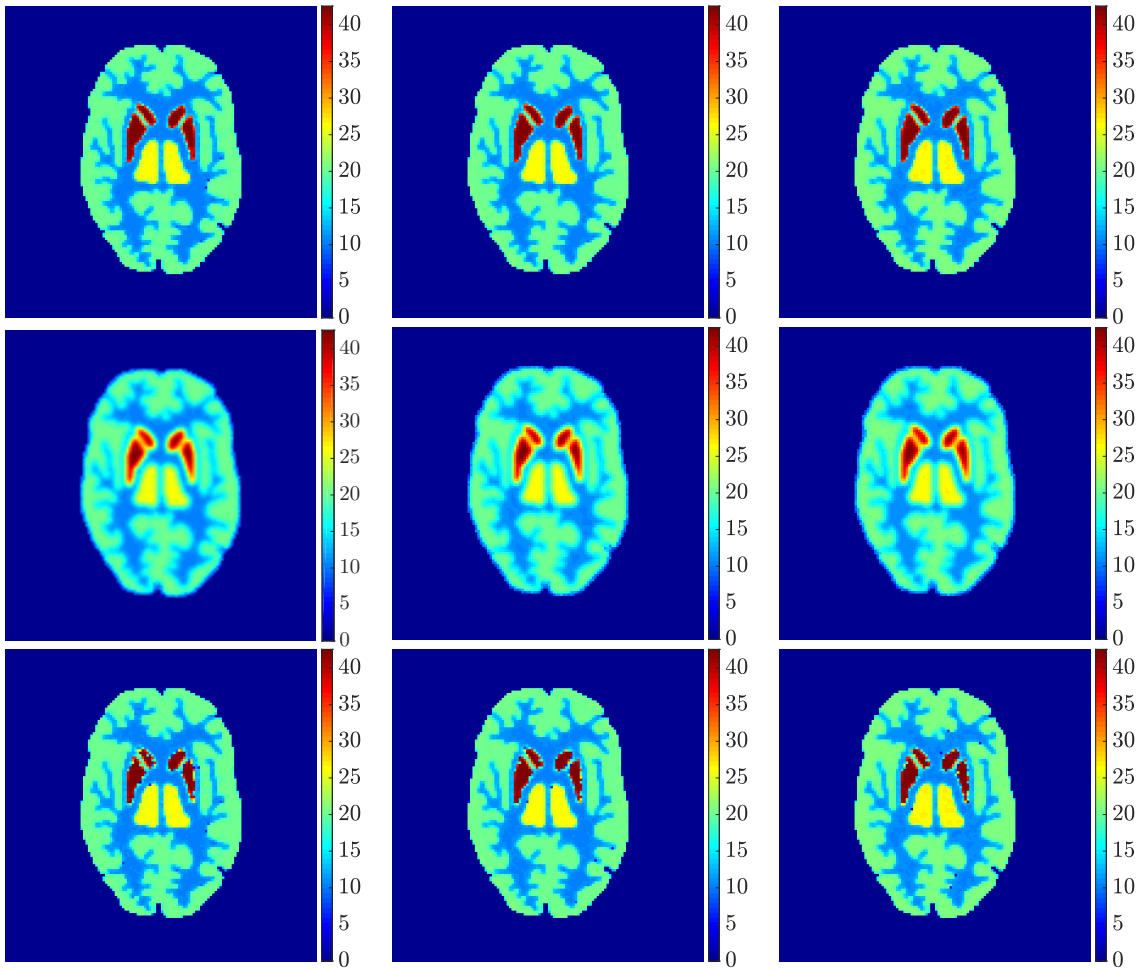
Figure 8: Last frame of the dynamic reconstruction obtained by using reg-AS-TR (first row), reg-GN (second row), and *lsqcurvefit* (third row). From left to right: noise-free IF, 10%-noise IF, 20%-noise IF.

Appendix

In this Appendix, we provide the proof of Proposition 1.

Proof. Let $\boldsymbol{p}^{j\dagger} = \boldsymbol{k}^{\dagger} - \boldsymbol{k}^{j}$ and $\boldsymbol{k}^{j+1} = \boldsymbol{k}^{j} + \bar{\boldsymbol{p}}^{j}$, with $\bar{\boldsymbol{p}}^{j}$ given in (10). If $\bar{\boldsymbol{p}}^{j} = \boldsymbol{p}^{j}$, we return back to the unconstrained case, for which Proposition 2.1 in [14] holds. Let assume for some j that

$$\bar{\boldsymbol{p}}^{j} = t(\Pi(\boldsymbol{k}^{j} + \boldsymbol{p}^{j}) - \boldsymbol{k^{j}}) \qquad 0 < t < 1.$$

Let $\bar{\boldsymbol{k}}^{j} = \Pi(\boldsymbol{k^{j}} + \boldsymbol{p^{j}})$. From the properties of the projection operator [48, Proposition 2.1.3], we have

$$(\bar{\boldsymbol{k}}^{j} - (\boldsymbol{k}^{j} + \boldsymbol{p}^{j}))^{T}(\boldsymbol{k} - \bar{\boldsymbol{k}}^{j}) \geq 0 \quad \forall \ \boldsymbol{k} \geq 0; \qquad (A.1)$$

in particular, Eq. (A.1) holds for $\boldsymbol{k}^{\dagger} \geq 0$ and, therefore, we obtain

$$(\bar{\boldsymbol{k}}^{j} - \boldsymbol{k^{j}})^{T}(\boldsymbol{k}^{\dagger} - \bar{\boldsymbol{k}}^{j}) \geq \boldsymbol{p}^{j^{T}}(\boldsymbol{k}^{\dagger} - \bar{\boldsymbol{k}}^{j}).$$

Computational approaches for parametric imaging of dynamic PET data 17

Setting $\hat{\boldsymbol{p}}^j = \bar{\boldsymbol{k}}^j - \boldsymbol{k}^j = \frac{1}{t}\bar{\boldsymbol{p}}^j$ and $\boldsymbol{k}^\dagger - \bar{\boldsymbol{k}}^j = \boldsymbol{p}^{j\dagger} - \hat{\boldsymbol{p}}^j$, we can write the previous inequality as follows:

$$\hat{\boldsymbol{p}}^{j^T}\boldsymbol{p}^{j\dagger} \geq \|\hat{\boldsymbol{p}}^j\|^2 + \boldsymbol{p}^{j^T}\boldsymbol{p}^{j\dagger} - \boldsymbol{p}^{j^T}\hat{\boldsymbol{p}}^j.$$

From the identity $\|\boldsymbol{p}^{j\dagger}\|^2 - \|\bar{\boldsymbol{p}}^j - \boldsymbol{p}^{j\dagger}\|^2 = 2\,\bar{\boldsymbol{p}}^{j^T}\boldsymbol{p}^{j\dagger} - \|\bar{\boldsymbol{p}}^j\|^2$, the definition $\bar{\boldsymbol{p}}^j = t\hat{\boldsymbol{p}}^j$, the previous inequality and $t \in (0,1)$, we have

$$
\begin{aligned}
\|\boldsymbol{p}^{j\dagger}\|^2 - \|\bar{\boldsymbol{p}}^j - \boldsymbol{p}^{j\dagger}\|^2 = 2\,\bar{\boldsymbol{p}}^{j^T}\boldsymbol{p}^{j\dagger} - \|\bar{\boldsymbol{p}}^j\|^2 = 2t\,\hat{\boldsymbol{p}}^{j^T}\boldsymbol{p}^{j\dagger} - t^2\|\hat{\boldsymbol{p}}^j\| &\geq \\
\geq 2t\,(\|\hat{\boldsymbol{p}}^j\|^2 + \boldsymbol{p}^{j^T}\boldsymbol{p}^{j\dagger} - \boldsymbol{p}^{j^T}\hat{\boldsymbol{p}}^j) - t^2\|\hat{\boldsymbol{p}}^j\| & \\
= t\left((2-t)\|\hat{\boldsymbol{p}}^j\|^2 + 2\boldsymbol{p}^{j^T}\boldsymbol{p}^{j\dagger} - 2\boldsymbol{p}^{j^T}\hat{\boldsymbol{p}}^j\right) & \\
> t(\|\hat{\boldsymbol{p}}^j\|^2 + 2\boldsymbol{p}^{j^T}\boldsymbol{p}^{j\dagger} + \|\boldsymbol{p}^j\|^2 - \|\boldsymbol{p}^j\|^2 - 2\boldsymbol{p}^{j^T}\hat{\boldsymbol{p}}^j) & \\
= t(\|\hat{\boldsymbol{p}}^j - \boldsymbol{p}^j\|^2 + 2\boldsymbol{p}^{j^T}\boldsymbol{p}^{j\dagger} - \|\boldsymbol{p}^j\|^2) & \\
> t(2\boldsymbol{p}^{j^T}\boldsymbol{p}^{j\dagger} - \|\boldsymbol{p}^j\|^2) & \\
> t(2\boldsymbol{p}^{j^T}\boldsymbol{p}^{j\dagger} - 2\|\boldsymbol{p}^j\|^2). & \quad (A.2)
\end{aligned}
$$

Now, we recall that, in view of positive definiteness of the matrix $(\boldsymbol{J}(\boldsymbol{k}^j)^T\boldsymbol{J}(\boldsymbol{k}^j) + \alpha_j\boldsymbol{I}_n)$, the following matrix identities hold:

$$
\begin{aligned}
(\boldsymbol{J}(\boldsymbol{k}^j)^T\boldsymbol{J}(\boldsymbol{k}^j) + \alpha_j\boldsymbol{I}_n)^{-1}\boldsymbol{J}(\boldsymbol{k}^j)^T &= \boldsymbol{J}(\boldsymbol{k}^j)^T(\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{J}(\boldsymbol{k}^j)^T + \alpha_j\boldsymbol{I}_N)^{-1} \\
\boldsymbol{I}_N - \boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{J}(\boldsymbol{k}^j)^T(\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{J}(\boldsymbol{k}^j)^T + \alpha_j\boldsymbol{I}_N)^{-1} &= \alpha_j(\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{J}(\boldsymbol{k}^j)^T + \alpha_j\boldsymbol{I}_N)^{-1}
\end{aligned}
$$

As a consequence, setting $\boldsymbol{r}^j = \boldsymbol{y}^\delta - \boldsymbol{F}(\boldsymbol{k}^j)$, we can write

$$\boldsymbol{r}^j - \boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{p}^j = \alpha_j(\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{J}(\boldsymbol{k}^j)^T + \alpha_j\boldsymbol{I}_N)^{-1}\boldsymbol{r}^j. \quad (A.3)$$

In view of inequality (A.2), the definition of $\boldsymbol{p}^j$ and the above identities, we have

$$
\begin{aligned}
&\|\boldsymbol{p}^{j\dagger}\|^2 - \|\bar{\boldsymbol{p}}^j - \boldsymbol{p}^{j\dagger}\|^2 \\
&> t(2\boldsymbol{r}^{j^T}(\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{J}(\boldsymbol{k}^j)^T + \alpha_j\boldsymbol{I}_N)^{-1}\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{p}^{j\dagger} - 2(\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{p}^j)^T(\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{J}(\boldsymbol{k}^j)^T + \alpha_j\boldsymbol{I}_N)^{-1}\boldsymbol{r}^j) \\
&= t(2\boldsymbol{r}^{j^T}(\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{J}(\boldsymbol{k}^j)^T + \alpha_j\boldsymbol{I}_N)^{-1}\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{p}^{j\dagger} - 2\boldsymbol{r}^{j^T}(\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{J}(\boldsymbol{k}^j)^T + \alpha_j\boldsymbol{I}_N)^{-1}\boldsymbol{r}^j + \\
&\quad + 2\alpha_j\boldsymbol{r}^{j^T}(\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{J}(\boldsymbol{k}^j)^T + \alpha_j\boldsymbol{I}_N)^{-2}\boldsymbol{r}^j) \\
&= t(2\alpha_j\|(\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{J}(\boldsymbol{k}^j)^T + \alpha_j\boldsymbol{I}_N)^{-1}\boldsymbol{r}^j\|^2 - 2\boldsymbol{r}^{j^T}(\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{J}(\boldsymbol{k}^j)^T + \alpha_j\boldsymbol{I}_N)^{-1}(\boldsymbol{r}^j - \boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{p}^{j\dagger})) \\
&> 2t(\alpha_j\|(\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{J}(\boldsymbol{k}^j)^T + \alpha_j\boldsymbol{I}_N)^{-1}\boldsymbol{r}^j\|^2 - \|(\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{J}(\boldsymbol{k}^j)^T + \alpha_j\boldsymbol{I}_N)^{-1}\boldsymbol{r}^j\|\|\boldsymbol{r}^j - \boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{p}^{j\dagger}\|) \\
&= 2t(\|\boldsymbol{r}^j - \boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{p}^j\| - \|\boldsymbol{r}^j - \boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{p}^{j\dagger}\|)\|(\boldsymbol{J}(\boldsymbol{k}^j)\boldsymbol{J}(\boldsymbol{k}^j)^T + \alpha_j\boldsymbol{I}_N)^{-1}\boldsymbol{r}^j\|,
\end{aligned}
$$

where the last inequality follows from the Cauchy-Schwarz inequality. Then, the q-condition and the assumption (17) yields

$$\|\boldsymbol{p}^{j\dagger}\|^2 - \|\bar{\boldsymbol{p}}^j - \boldsymbol{p}^{j\dagger}\|^2 > \frac{2t(\gamma_\delta - 1)q}{\gamma_\delta}\|\boldsymbol{y}^\delta - \boldsymbol{F}(\boldsymbol{k}^j)\|\|\boldsymbol{v}^j\|.$$

$\square$

References

[1] D. Bailey, D. Towsend, P. Valk, and M. Maisey. *Positron Emission Tomography - Basic Science*. Springer Verlag, 1st edition, 2005.

[2] F. Natterer and F. Wübbeling. *Mathematical Methods in Image Reconstruction*. SIAM, 2001.

AUTHOR SUBMITTED MANUSCRIPT - IP-102361

[3] L. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Trans. Med. Imag.*, 113:113–122, 1982.

[4] R. N. Gunn, S. R. Gunn, and V. J. Cunningham. Positron emission tomography compartmental models. *J. Cer. Blood Flow Metab.*, 32:635–652, 2001.

[5] K. C. Schmidt and F. E. Turkheimer. Kinetic modeling in positron emission tomography. *Q. J. Nucl. Med. Mol. Imaging*, 46:70–85, 2002.

[6] H. Watabe, Y. Ikoma, Y. Kimura, M. Naganowa, and M. Shidahara. PET kinetic analysis – compartmental models. *Annals of Nuclear Medicine*, 20:583, 2006.

[7] R. N. Gunn, A. A. Lammertsma, S. P. Hume, and V. J. Cunningham. Parametric imaging of ligand-receptor binding in PET using a simplified reference region model. *Neuroimage*, 6(4):279–287, 1997.

[8] N. A. Karakatsanis, M. A. Lodge, A. K. Tahari, Y. Zhou, R. L. Wahl, and A. Rahmim. Dynamic whole-body PET parametric imaging: I. Concept, acquisition protocol optimization and clinical application. *Phys. Med. Biol.*, 58(20):7391, 2013.

[9] A. J. Reader and J. Verhaeghe. 4D image reconstruction for emission tomography. *Phys. Med. Biol.*, 59(22):R371, 2014.

[10] M. E. Kamasak, C. A. Bouman, E. D. Morris, and K. Sauer. Direct reconstruction of kinetic parameter images from dynamic PET data. *IEEE Trans. Med. Im.*, 24(5):636–650, 2005.

[11] G. Wang and J. Qi. Direct estimation of kinetic parametric images for dynamic PET. *Theranostics*, 802:815, 2013.

[12] M. Castellaro, G. Rizzo, M. Tonietto, M. Veronese, F. E. Turkheimer, M. A. Chappell, and A. Bertoldo. A variational bayesian inference method for parametric imaging of PET data. *Neuroimage*, 150:136–149, 2017.

[13] M. Scussolini, S. Garbarino, G. Sambuceti, G. Caviglia, and M. Piana. A physiology-based parametric imaging method for FDG–PET data. *Inverse Problems*, 33(12):125010, 11 2017.

[14] Y. Wang and Y. Yuan. Convergence and regularity of trust region methods for nonlinear ill-posed problems. *Inverse Problems*, 21:821–838, 2005.

[15] T. F. Coleman and Y. Li. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 6(2):418–445, 1996.

[16] S. Bellavia and B. Morini. Subspace trust-region methods for large bound-constrained nonlinear equations. *SIAM Journal of Numerical Analysis*, 44:1535–1555, 2006.

[17] R. Boellaard, M. J. ODoherty, W. A. Weber, F. M. Mottaghy, M. N. Lonsdale, S. G. Stroobants, W. J. G. Oyen, J. Kotzerke, O. S. Hoekstra, J. Pruim, P. K. Marsden, K. Tatsch, Hoekstra. C. J., E. P. Visser, B. Arends, F. J. Verzijlbergen, J. M. Zijlstra, E. F. I. Comans, A. A. Lammertsma, A. M. Paans, A. T. Willemsen, T. Beyer, A. Bockisch, C. Schaefer-Prokop, D¿ Delbeke, R¿ P. Baum, A. Chiti, and B. J. Krause. FDG PET and PET/CT: Eanm procedure guidelines for tumour PET imaging: version 1.0. *Europ. J. Nucl. Med. Mol. Im.*, 37:181, 2010.

[18] J. W. Fletcher, B. Djulbegovic, H. P Soares, B. A. Siegel, V. J. Lowe, G. H. Lyman, R. E. Coleman, R. Wahl, J. C. Paschold, N. Avril, L. H. Einhorn, W. W. Suh, D. Samson, D. Delbeke, M. Gorman, and A. F. Shields. Recommendations on the use of 18F-FDG PET in oncology. *J. Nucl. Med.*, 49(3):480–508, 2001.

[19] T. Kato, Y. Inui, A. Nakamura, and K. Ito. Brain fluorodeoxyglucose (FDG) PET in dementia. *Ag. Res. Rev.*, 30:73–84, 2016.

[20] C. Marini, S. Morbelli, A. Cistaro, C Campi, C. Caponnetto, M. Bauckneht, A. Bellini, A. Buschiazzo, I Calamia, M. C. Beltrametti, S. Margotti, P. Fania, I. Poggi, C. Cabona, S. Capitanio, R. Piva, A. Calvo, C. Moglia, A. Canosa, A. M. Massone, F. Nobili, G. Mancardi, A. Chiò, M. Piana, and G. Sambuceti. Interplay between spinal cord and cerebral cortex metabolism in amyotrophic lateral sclerosis. *Brain*, 141(8):2272–2279, 2018.

[21] M. Massollo, C. Marini, M. Brignone, L. Emionite, B. Salani, M. Riondato, S. Capitanio, F. Fiz, A. Democrito, A. Amaro, S. Morbelli, M. Piana, D. Maggi, M Cilli, U. Pfeffer, and S. Sambuceti. Metformin temporal and localized effects on gut glucose metabolism assessed using 18F-FDG

PET in mice. *J. Nucl. Med.*, 54(2):259–266, 2013.

[22] R. E. Carson. *Tracer Kinetic Modeling in PET*, pages 127–159. Springer London, 2005.

[23] S. R. Cherry, J. A. Sorenson, and M. E. Phelps. *Physics in Nuclear Medicine.* Elsevier Health Sciences, 2012.

[24] E. D. Morris, C. J. Endres, K. C. Schmidt, B. T. Christian, R. F. Muzic, and R. E. Fisher. *Kinetic Modeling in Positron Emission Tomography*, pages 499–540. Academic Press, 11 2004.

[25] K. C. Schmidt and F. E. Turkheimer. Kinetic modeling in positron emission tomography. *The Quarterly Journal of Nuclear Medicine and Molecular Imaging*, 46(1):70–85, 3 2002.

[26] S. Garbarino, V. Vivaldi, F. Delbary, G. Caviglia, M. Piana, C. Marini, S. Capitanio, I. Calamia, Buschiazzo A., and G. Sambuceti. A new compartmental method for the analysis of liver FDG kinetics in small animal models. *Europ. J. Nucl. Med. Mol. Imag. Res.*, 5:35, 2015.

[27] S. Garbarino, G. Caviglia, G. Sambuceti, F. Benvenuto, and M. Piana. A novel description of FDG excretion in the renal system: application to metformin-treated models. *Phys. Med. Biol.*, 59(10):2469, 2014.

[28] L. Sokoloff, M. Reivich, C. Kennedy, M. H. Des Rosiers, C. S. Patlak, K. D. Pettigrew, O. Sakurada, and M. Shinohara. The [14C]deoxyglucose method for the measurement of local cerebral glucose utilization: theory, procedure, and normal values in the conscious and anesthetized albino rat. *Journal of Neurochemistry*, 28(5):897–916, 2 1977.

[29] Y. Su, T. M. Blazey, A. Z. Snyder, M. E. Raichle, R. C. Hornbeck, P. Aldea, J. C. Morris, and T. L. S. Benzinger. Quantitative amyloid imaging using image-derived arterial input function. *Plos One*, 10(4):0122920, 2015.

[30] C. A. Salinas, G. E. Searle, and R. N. Gunn. The simplified reference tissue model: Model assumption violations and their impact on binding potential. *J. Cer. Blood Flow Metab.*, 35(2):304–311, 2015.

[31] M. Scussolini, S. Garbarino, Piana M., G. Sambuceti, and G. Caviglia. Reference tissue models for FDG-PET data: Identifiability and solvability. *IEEE Trans. Rad. Pla. Med. Sci.*, 2(3):177–186, 2018.

[32] C. S. Patlak, R. G. Blasberg, and J. D. Fenstermacher. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. *J. Cer. Blood Flow Metab.*, 3(1):1–7, 1983.

[33] J. Logan. Graphical analysis of PET data applied to reversible and irreversible tracers. *Nuclear Medicine and Biology*, 27:661, 2000.

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

[35] B. Kaltenbacher, A. Neubauer, and O. Scherzer. *Iterative regularization methods for nonlinear ill-posed problems.* Berlin, Boston: De Gruyter, 2008.

[36] M. Macconi, B. Morini, and M. Porcelli. A Gauss-Newton method for solving bound-constrained undetermined nonlinear systems. *Optimization Methods and Software*, 24:219–235, 2009.

[37] M. Hanke. Regularizing Levenberg-Marquardt scheme, with applications to inverse groundwater filtration problems. *Inverse Problems*, 13:79–95, 1997.

[38] S. Bellavia, B. Morini, and E. Riccietti. On an adaptive regularization for ill-posed nonlinear systems and its trust-region implementation. *Computational Optimization and Applications*, 64:1–30, 2016.

[39] H. Hongbin Guo, R. A. Renaut, K. Chen, and E. Reiman. FDG–PET parametric imaging by total variation minimization. *Computerized Medical Imaging and Graphics*, 33(4):295–303, 2009.

[40] S. C. Huang, M. E. Phelps, E. J. Hoffman, K. Sideris, C. J. Selin, and D. E. Kuhl. Noninvasive determination of local cerebral metabolic rate of glucose in man. *American Journal of Physiology-Endocrinology and Metabolism*, 238(1):E69–E82, 1980.

[41] G. Wang and J. Qi. Generalized algorithms for direct reconstruction of parametric images from dynamic PET data. *IEEE Transactions on Medical Imaging*, 28(11):1717–1726, 2009.

[42] G. Wang and J. Qi. An optimization transfer algorithm for nonlinear parametric image reconstruction from dynamic PET data. *IEEE Transactions on Medical Imaging*, 31(10):1977–1988, 2012.

Computational approaches for parametric imaging of dynamic PET data 20

[43] D. Vriens, L. F. de Geus-Oei, W. J. Oyen, and E. P. Visser. A curve-fitting approach to estimate the arterial plasma input function for the assessment of glucose metabolic rate and response to treatment. *Journal of Nuclear Medicine*, 50(12):1933–1939, 2009.

[44] S. Bonettini, R. Zanella, and L. Zanni. A scaled gradient projection method for constrained image deblurring. *Inverse Probl.*, 25(1):015002, 2009.

[45] S. Crisci, V. Ruggiero, and L. Zanni. Steplength selection in gradient projection methods for box-constrained quadratic programs. *Applied Mathematics and Computation*, 356:312–327, 2019.

[46] T. F. Coleman and Y. Li. On the convergence of reflective Newton methods for large-scale nonlinear minimization subject to bounds. *Mathematical Programming*, 67(2):189–224, 1994.

[47] M. Scussolini, V. Cossu, C. Marini, Sambuceti. G, G. Caviglia, and M. Piana. The role of the endoplasmic reticulum in 'in vivo' cancer fdg kinetics. *bioarxiv*, http://dx.doi.org/10.1101/664417, 2019.

[48] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, 2nd edition, 1999.