Spatio-temporally constrained reconstruction for hyperpolarized carbon-13 MRI using kinetic models

John Maidens, Jeremy W. Gordon, Hsin-Yu Chen, Ilwoo Park, Mark Van Criekinge, Eugene Milshteyn, Robert Bok, Rahul Aggarwal, Marcus Ferrone, James B. Slater, John Kurhanewicz, Daniel B. Vigneron, Murat Arcak, and Peder E. Z. Larson

***Abstract*—We present a method of generating spatial maps of kinetic parameters from dynamic sequences of images collected in hyperpolarized carbon-13 MRI experiments. The technique exploits spatial correlations in the dynamic traces via regularization in the space of parameter maps. Similar techniques have proven successful in other dynamic imaging problems such as dynamic contrast enhanced MRI. In this paper we apply these techniques for the first time to hyperpolarized MRI problems, which are particularly challenging due to limited SNR. We formulate the reconstruction as an optimization problem and present an efficient iterative algorithm for solving it based on the alternation direction method of multipliers (ADMM). We demonstrate that this technique improves the qualitative appearance of parameter maps estimated from low SNR dynamic image sequences, first in simulation then on a number of data sets collected *in vivo*. The improvement this method provides is particularly pronounced at low SNR levels.**

I. Introduction

Magnetic resonance imaging (MRI) using hyperpolarized carbon-13 labeled substrates has made it possible to probe metabolism *in vivo* with chemical specificity [1], [2]. This technique is increasingly being applied in the clinic, allowing researchers to investigate metabolic conditions ranging from prostate cancer [3] to heart disease [4]. In particular, experiments studying the conversion of hyperpolarized [1-^{13}C]pyruvate to [1-^{13}C]lactate are common, as the rate of conversion is upregulated in many cancers, a phenomenon known as the Warburg effect.

MRI using hyperpolarized carbon-13 is challenging due to the dynamic nature of the data collected, the low signal-to-noise ratio (SNR), and the difficulty of presenting large data sets consisting of dynamic spectroscopic images in an interpretable manner. Metabolism mapping by estimating parameters in a kinetic model from hyperpolarized MRI data has been shown to be useful for overcoming a number of these challenges [5]. Constraining the time evolution of signal in a given voxel to follow a kinetic model has been shown to allow map reconstruction from noisy, undersampled dynamic images, and to reduce the number of signal-depleting excitations required to generate images. Parameter mapping also facilitates interpretation of dynamic image data by summarizing spatial, temporal and chemical (i.e. chemical shift spectrum) information in a single spatial map.

Parameter maps are naturally a form of constrained reconstruction, as they constrain the data to lie on a manifold of trajectories of the dynamical system parametrized by the system's parameters. This constrained reconstruction reduces the sequence of dynamic images to a single map by exploiting temporal correlations within the dynamic imaging data. In this paper, we demonstrate that we can exploit spatial correlations in addition to temporal correlations by integrating prior information about the parameter map through regularization. Similar approaches have proven useful recently in the context of pharmacokinetic parameter mapping in dynamic contrast enhanced and cardiac perfusion MRI [6]–[12]. To our knowledge, this is the first time this family of spatial regularization techniques have been used in hyperpolarized MRI, where they are particularly beneficial due to the challenges of working with low SNR images.

This paper is organized as follows. In Section II we introduce background on modelling hyperpolarized ^{13}C MRI data and existing approaches to parameter mapping. In Section III we introduce a framework for spatially-constrained parameter mapping to exploit spatial correlations in the data. In Section IV we present an algorithm for efficient inference in this framework. In Section V we present the results of simulation experiments where we demonstrate the effectiveness of the method. In Section VI we then apply the method to a collection of clinically-relevant data sets. Finally, Section VII concludes the paper and briefly discusses potential extensions of this work.

Preliminary results from this paper were presented at the 2017 Annual Meeting of the International Society of Magnetic Resonance in Medicine [13].

J. Maidens is with the Department of Mechanical and Industrial Engineering, Ryerson University, 350 Victoria Street, Toronto, ON, M5B 2K3, Canada. Corresponding author: johnmaidens@gmail.com

J. Gordon, H-Y Chen, M. Van Criekinge, E. Milshteyn, R. Bok, J. Slater, J. Kurhanewicz, D. Vigneron and P. Larson are with the Department of Radiology and Biomedical Imaging, University of California, San Francisco, 1700 4th Street, San Francisco, CA, 94158, USA.

I. Park is with the Department of Radiology, Chonnam National University Medical School, Jeabongro 42, Donggu, City of Gwangju, 61469, Korea.

R. Aggarwal is with the Department of Medicine, University of California, San Francisco, 1600 Divisadero, San Francisco, CA, 94115, USA.

M. Ferrone is with the Department of Clinical Pharmacy, University of California, San Francisco, 1700 4th Street, San Francisco, CA, 94158, USA.

M. Arcak is with the Department of Electrical Engineering & Computer Sciences, University of California, Berkeley, CA, 94720, USA.

This research was supported in part by the NSF under grant ECCS-1405413 and the NIH under grants R01EB017449, R01CA166655, and P41EB013598.

II. Background

A. Data Model

We model the dynamic evolution of the data Y_i collected from a single voxel i using a dynamic model for a two-dimensional state $x(t) = [x_1(t) \quad x_2(t)]^T$:

$$\frac{dx}{dt}(t) = \begin{bmatrix} -k_{PL} - R_{1P} & 0 \\ k_{PL} & -R_{1L} \end{bmatrix} x(t) + \begin{bmatrix} k_{TRANS} \\ 0 \end{bmatrix} u(t). \quad (1)$$

This system of ordinary differential equations (ODEs) has been widely used to model the uni-directional conversion of an injected substrate (pyruvate, in this case) to a metabolic product (lactate, in this case) [14]. The state $x_1(t)$ models the longitudinal magnetization in the substrate pool, and the state $x_2(t)$ models the longitudinal magnetization in the product pool. The parameter k_{PL} describes the rate at which the substrate is metabolized, the parameter k_{TRANS} describes the rate at which the substrate is taken up by the tissue, and the parameters R_{1P} and R_{1L} are lumped parameters that account for T_1 magnetization decay, metabolism of the substrate into unmeasured products and flow of substrate out of the voxel.

Measurements are collected at a sequence of times $\{t_1, \ldots, t_N\}$. Neglecting the effect of the input between t_k and t_{k+1}, integrating this continuous-time dynamic model and incorporating the effect of repeated radio-frequency (RF) excitation leads to a discrete-time model for the magnetization at acquisition times t_k of the form

$$\begin{aligned} \hat{L}(k+1) = e^{-R_{1L}\Delta t} \cos(\alpha_L(k)) \hat{L}(k) \\ - k_{PL} \frac{e^{-(R_{1P}+k_{PL})\Delta t} - e^{-R_{1L}\Delta t}}{R_{1P} - R_{1L} + k_{PL}} \cos(\alpha_P(k)) P(k). \end{aligned} \quad (2)$$

This gives a statistical model that describes the evolution of the predicted lactate signal $\hat{L}(k) = x_2(t_k)$ as a function of the measured pyruvate signal $P(k) = x_1(t_k)$ and the flip angles α_P and α_L applied to the pyruvate and lactate compartments. The predicted lactate is assumed to be $\hat{L}(0) = 0$ at the beginning of the experiment.

For the purpose of generating simulated data, the data measured at each time t_k are assumed to be independent and follow a bivariate normal distribution with mean $\delta_x \delta_y \delta_z x(t_k)$ and covariance $\sigma^2 I$ where I denotes the 2×2 identity matrix and δ_x, δ_y and δ_z describe the image resolution and slice thickness. We collect the time series data collected from voxel i into a matrix $Y_i = \begin{bmatrix} P(1) & \cdots & P(N) \\ L(1) & \cdots & L(N) \end{bmatrix}$ and denote the unknown parameters to be estimated from the data $\theta_i = k_{PL}$.

B. Voxel-Wise Parameter Estimation

Given a collection of data Y_i from a voxel i we wish to generate an estimate of the parameter θ_i that describes the tissue in that voxel. We assume that θ_i lies in a parameter space Θ. We consider the class of "M-estimators" [15] that minimize a loss function

$$\hat{\theta}_i \in \operatorname*{argmin}_{\theta \in \Theta} \ell(\theta_i | Y_i).$$

In the present paper, we consider the nonlinear least squares loss function

$$\ell(\theta_i | Y_i) = \|Y_i - \hat{Y}_i(\theta_i)\|_F \quad (3)$$

where $\hat{Y} = \begin{bmatrix} P(1) & \cdots & P(N) \\ \hat{L}(1) & \cdots & \hat{L}(N) \end{bmatrix}$ denotes the predicted signal given the pyruvate time series and $\|\cdot\|_F$ denotes the Frobenius norm (*i.e.* the ℓ_2 norm of the vectorized matrix). Under the assumption that the data collected are normally-distributed with mean proportional to $x(t_k)$, independent with identical variance, the minimum of this nonlinear least squares loss is also the maximum likelihood estimate of the parameter vector. While we consider only this loss in the present paper, the results are applicable generally to any computationally tractable loss function.

III. Constrained Parameter Mapping

In order to incorporate prior information about the spatial distribution of metabolic rates and exploit spatial correlations within the data, we constrain the maps to have a desired structure through regularization. This results in an optimization problem in Lagrangian form

$$\text{minimize} \sum_{i \in \mathcal{V}} \ell(\theta_i | Y_i) + \lambda r(\theta) \quad (4)$$

where $\theta = (\theta_i)_{i \in \mathcal{V}}$ denotes the map of parameters across all voxels, r is a regularization term, and λ denotes a Lagrange multiplier that can be tuned in order to achieve the desired regularization strength. The choice of an appropriate regularizer depends on the desired features of the parameter map. Common choices include Tikhonov (ℓ_2) regularization, ℓ_1 regularization, and total variation regularization. We briefly summarize these three methods below.

Tikhonov regularization, or ℓ_2 regularization penalizes the size of the parameters θ_i. It involves adding a quadratic penalty term

$$r(\theta) = \|\theta\|_2^2$$

where $\|\cdot\|_2$ denotes the ordinary Euclidean norm. For linear regression problems with orthogonal covariates, this regularization leads to uniform shrinkage of the estimates [16]. For the nonlinear parameter mapping problems we consider here, using Tikhonov regularization helps to suppress large parameter values in the unperfused "background" region.

ℓ_1 regularization is another shrinkage method that penalizes parameters based on their ℓ_1 norm

$$r(\theta) = \|\theta\|_1.$$

This method induces sparsity in the resulting parameter maps, and hence also helps to suppress parameter values in the background region. It is closely-related to basis pursuit denoising [17] and lasso regression [18].

Total variation (TV) regularization is another method commonly used for image denoising [19]. In this paper, we use an anisotropic total variation regularization term given by

$$r(\theta) = \|\nabla \theta\|_1 := \sum_{(i,j) \in \mathcal{N}} |\theta_i - \theta_j|$$

where ∇ denotes a discrete differencing operator and $\mathcal{N}$ denotes the set of all neighbouring voxels. As all applications we consider in this paper we consider three-dimensional images, the neighbourhood $\mathcal{N}$ consists of the six voxels j immediately

adjacent to the voxel i. Anisotropic total variation is chosen due to the availability of numerical packages for extremely fast computation of proximity operators via the proxTV package [20], [21]. TV regularization is known to preserve edges and large-scale structure in images while rejecting noise [22], resulting in natural-looking reconstructed images.

IV. Iterative Algorithms for Constrained Parameter Mapping

A naive algorithm for solving this optimization problem by directly optimizing the objective function (4) would be inefficient because it involves solving a joint optimization over all $\{\theta_i : i \in \mathcal{V}\}$. Thus the computation time required to directly solve the optimization problem increases dramatically with matrix size, making naive approaches inefficient even for the images of moderate resolution considered here. To solve the optimization problem more efficiently, we can take advantage of the particular structure of the problem using the ADMM algorithm.

The alternating direction method of multipliers (ADMM) is an iterative optimization algorithm that is well-suited to efficiently solving such problems that can be decomposed into a sum of two terms [23]. In contrast with other distributed optimization algorithms, the ADMM algorithm is particularly well-suited to the problem formulated in this paper as it splits the required optimization into the sum of a set of loss functions ℓ that are complex to optimize, but can be optimized independently for each voxel, and a regularization r that is relatively simple but high-dimensional as it couples a large number of neighboring voxels. By exploiting this decomposition, ADMM allows the optimization problem to be solved efficiently. The general problem that ADMM attempts to solve is an optimization problem of the form

$$\begin{aligned} \textbf{minimize} \quad & f(x) + g(z) \\ \textbf{subject to} \quad & Ax + Bz = c. \end{aligned} \tag{5}$$

The algorithm does so by iteratively applying the updates

$$\begin{aligned} x^{k+1} &= \operatorname*{argmin}_x \left(f(x) + \frac{\rho}{2}\|Ax - Bz^k - c + u^k\|_2^2 \right) \\ z^{k+1} &= \operatorname*{argmin}_z \left(g(z) + \frac{\rho}{2}\|Ax^{k+1} - Bz - c + u^k\|_2^2 \right) \\ u^{k+1} &= u^k + Ax^{k+1} + Bz^{k+1} - c. \end{aligned}$$

Under the assumption that f and g are closed, proper, convex functions and that the Lagrangian

$$L(x, z, \lambda) = f(x) + g(z) + \lambda^T(Ax + Bz - c)$$

has a saddle point, it can be shown [23] that the residuals $r^k = Ax^k + Bz^k - c$ converge to zero and the values $f(x^k) + g(z^k)$ converge to the optimal value of the problem (5).

A. ADMM for iterative parameter mapping

To solve (4) we transform the problem to a form amenable to the ADMM algorithm by introducing a new variable $z = \theta$ and solving

$$\begin{aligned} \textbf{minimize} \quad & \sum_{i \in \mathcal{V}} \ell(\theta_i | Y_i) + \lambda r(z) \\ \textbf{subject to} \quad & \theta - z = 0. \end{aligned} \tag{6}$$

The ADMM iteration is then given as

$$\begin{aligned} \theta^{k+1} &= \operatorname*{argmin}_\theta \sum_{i \in \mathcal{V}} \ell(\theta_i | Y_i) + \frac{\rho}{2}\|\theta - z^k + u^k\|_2^2 \\ z^{k+1} &= \operatorname*{argmin}_z \lambda r(z) + \frac{\rho}{2}\|\theta^{k+1} - z + u^k\|_2^2 \\ u^{k+1} &= u^k + \theta^{k+1} - z^{k+1}. \end{aligned}$$

This method is sometimes known as Douglas-Rachford splitting [24]. Note that the θ update is additively separable. Introducing the proximity operator

$$\operatorname{prox}_f(x) = \operatorname*{argmin}_u f(u) + \frac{1}{2}\|u - x\|_2^2$$

we can re-write this iteration as

$$\begin{aligned} \theta_i^{k+1} &= \operatorname{prox}_{\frac{1}{\rho}\ell(\cdot|Y_i)}(z_i^k - u_i^k) \quad i \in \mathcal{V} \\ z^{k+1} &= \operatorname{prox}_{\frac{\lambda}{\rho}r}(\theta^{k+1} + u^k) \\ u^{k+1} &= u^k + \theta^{k+1} - z^{k+1}. \end{aligned}$$

Here, the θ_i updates can be performed independently for each $i \in \mathcal{V}$, significantly decreasing time and memory required for computation and allowing the parallelization of this step.

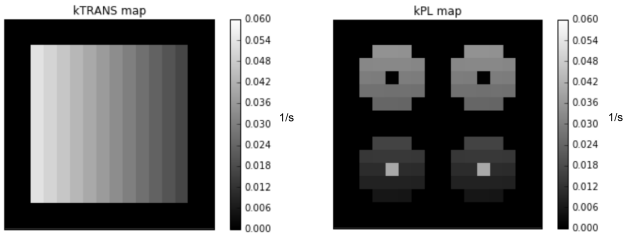
Note that for the particular choice of loss function given in Section III, $\ell(\cdot|Y_i)$ are nonconvex functions and thus the formal convergence guarantees do not apply. Despite this fact, we have seen in all the experimental instances of the problem we have considered that the algorithm converges to a sensible optimum robustly for a variety of initializations. In what follows, we use the modified Levenberg-Marquardt algorithm [25] implemented in MINPACK [26] to solve the nonlinear least squares problem corresponding to the θ update step in the ADMM iteration, and for the unregularized estimation.

V. Simulated Results and Discussion

To demonstrate the effectiveness of this method, we perform a sequence of experiments on simulated data. We begin with an experiment using a simple numerical phantom designed to test the robustness of metabolic parameter mapping methods to differences in perfusion, as well as their ability to reliably resolve large and small features.

A. Reconstruction at a variety of noise levels

To generate simulated data for validating our algorithm, we simulate trajectories for each voxel of the $16 \times 16 \times 16$ dynamic phantom described shown in Figure 1. This phantom describes maps of the k_{TRANS} and k_{PL} parameters and is designed to test an algorithm's ability to resolve both large and small features under high and low perfusion conditions. More details about the phantom can be found in Section 5.5



(a) k_{TRANS} map (b) k_{PL} map

Fig. 1: Slice through $z = 0$ of a $16 \times 16 \times 16$ voxel 3D dynamic phantom.

of [27]. The data are generated according to the model (1) with arterial input $u(t) = k_{TRANS} A_0 (t - t_0)^{\gamma} e^{(-(t-t_0)/\beta}$ added to the pyruvate compartment, and states scaled by $\cos(\alpha_{P/L}(k))$ and measured outputs scaled by $\sin(\alpha_{P/L}(k))$ each time that simulated data are collected, where $\alpha_{P/L}(k)$ is a spectrally-selective flip angle applied to spins in the P or L compartment during acquisition k. An optimized dynamic flip angle sequence based on the method of [28] is used for the simulation, and shown in Figure 2. This same flip angle sequence is also used for a majority of the *in vivo* experiments.

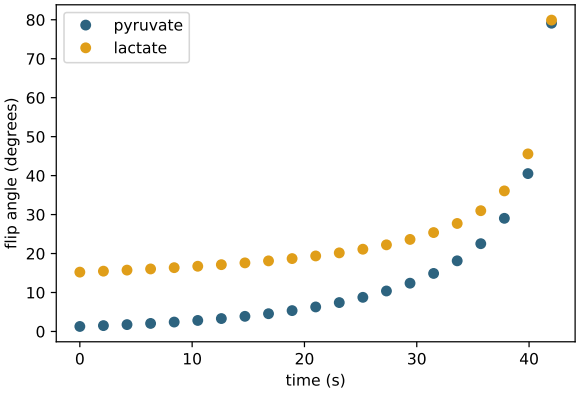
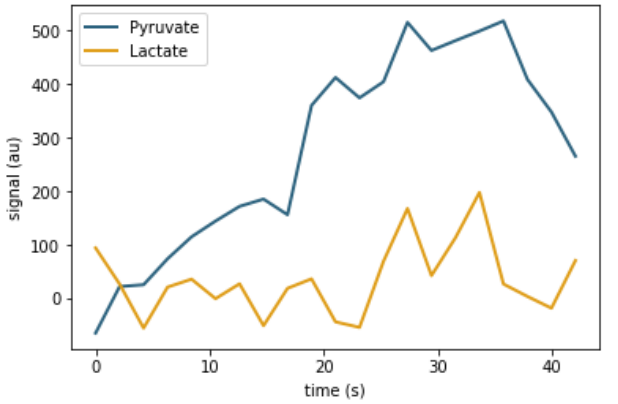


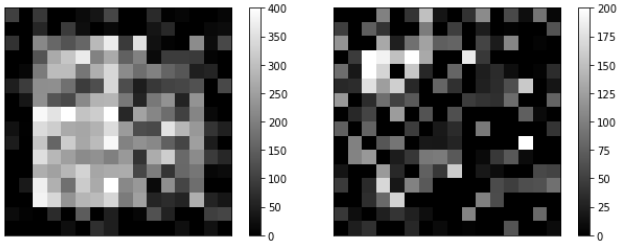
Fig. 2: Dynamic flip angle sequence used for experimental validation

We then add independent, identically-distributed (iid) Gaussian noise at a variety of SNR levels, measured based on the SNR in the lactate channel corresponding to the peak lactate level. Simulated time series and image data are shown in Figure 3.

For SNR levels of 8, 4, 2, and 1, we fit the model (2) to the data using the loss function (3) and the regularization $r(\theta) = \lambda_1 \|\nabla\theta\|_1 + \lambda_2 \|\theta\|_2^2$ with λ_1 =1e06 and λ_2 =1e08. A combination of ℓ_2 and TV regularization was chosen because the ℓ_2 penalty prevents estimation bias in the unperfused region while the TV penalty encourages smooth maps with well-defined tissue boundaries. The values of λ_1 and λ_2 are selected such that the total absolute error is minimized (see Section V-B). Before fitting, the simulated data are scaled by $1/\sin(\alpha_{P/L}(k))$ to counteract the effect of the time-varying flip angle sequence. In Figure 4 we compare the results of this



(a) Sample time series data from a high k_{TRANS}, high k_{PL} voxel



(b) Pyruvate image slice through $z = 0$ (c) Lactate image slice through $z = 0$

Fig. 3: Simulated data generated at a maximum lactate SNR level of 2.

constrained fit against two competing methods: independent voxel-wise fit (equivalent to our method with $\lambda_1 = \lambda_2 = 0$) and independent voxel-wise fit followed by anisotropic total variation denoising of the resulting parameter map. We see that the constrained reconstruction allows accurate parameter maps to be generated in high noise regimes where the competing methods have difficulty. In particular, the baseline method of unconstrained mapping followed by denoising performs poorly in unperfused areas where it is attempting to fit parameter values to pure noise. In contrast, the constrained fit is able to suppress noise in the unperfused region via ℓ_2 regularization.

B. *Quantitative Improvements*

In addition to the qualitative benefits of spatial regularization demonstrated in the previous section, regularization can also lead to quantitative improvements in the estimates of dynamic parameters. In simulation experiments where we have access to the ground truth values of the model parameters, we can quantify the improvement in estimates $\hat{\theta}$ of θ via the total absolute error

$$\|\hat{\theta} - \theta\|_1 = \sum_{i \in V} |\hat{k}_{PL_i} - k_{PL_i}|.$$

In Figure 5 we plot the total absolute error for various values of the regularization parameters λ_1 and λ_2. This experiment was performed using the $16 \times 16 \times 16$ phantom from Figure 1 with a maximum lactate SNR value of 2.0. We see that small values of λ_1 and λ_2 lead to larger quantitative errors in

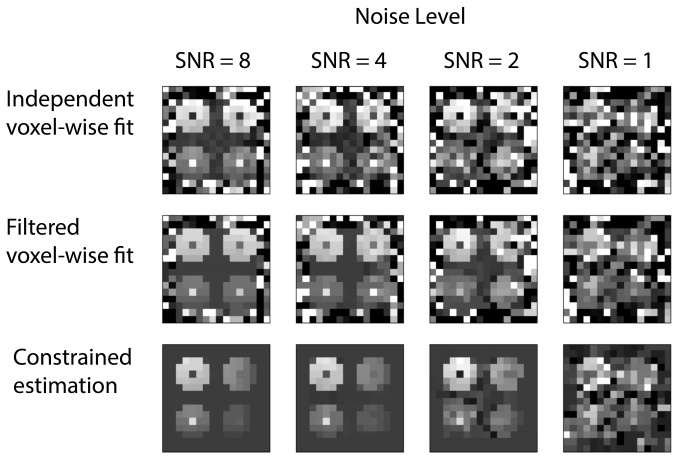


Fig. 4: Results of simulated k_{PL} mapping experiment for various values of the maximum lactate image SNR

the parameter maps than the optimized values λ_1 =1e06 and λ_2 =1e08 used in the previous section. Note that the optimal values will depend on a number of factors potentially including the geometry and sparsity of the phantom, and the noise distribution, SNR and signal amplitude in the dynamic images. Thus by appropriately choosing λ_1 and λ_2, we can achieve quantitative improvements in the parameter map in addition to the qualitative improvements we have already demonstrated.

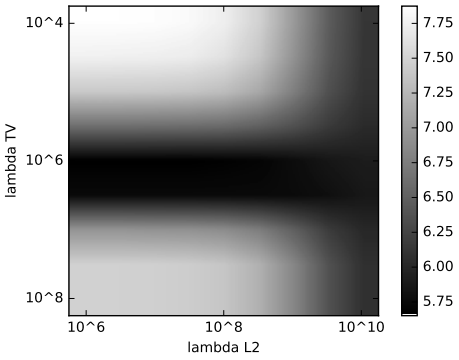


Fig. 5: Total absolute estimation error for k_{PL} for various values of the regularization parameters λ_1 and λ_2.

VI. In Vivo Results and Discussion

We now move on to experiments on a number of datasets collected *in vivo*. In contrast to the simulation experiments, we no longer have access to ground truth values of the model parameters to make quantitative comparisons. However, we will use the *in vivo* experiments to demonstrate that the spatially-constrained parameter mapping technique leads to qualitative improvements in the parameter maps.

We begin with an experiment in healthy rats where we can collect high SNR data. For these data, we add artificial noise to demonstrate how the spatially-constrained parameter mapping technique can be used to allow reconstruction in low SNR regimes, for realistic anatomies. We then apply this technique to the analysis of a number of low SNR clinical datasets collected in prostate cancer patients. These experiments demonstrate that spatio-temporally constrained kinetic modelling can be used to generate improved metabolic parameter maps from low SNR experimental data.

A. High SNR rat kidney data analysis

We begin by analyzing a metabolic dataset acquired in healthy Sprague-Dawley rats on a 3T MRI scanner (MR750, GE Healthcare). 2.5mL of 80mM hyperpolarized [1-13C]pyruvate was injected over 15s, and data acquisition coincided with the start of injection. Metabolites from a single slice were individually excited with a singleband spectral-spatial RF pulse and encoded with a single-shot EPI readout [29], an in-plane resolution of 3 x 3mm, a 15mm slice thickness centered on the kidneys, and a 2s sampling interval. The resulting dynamic image sequences are relatively high SNR with Rician noise resulting from magnitude images, are shown in Figure 6.

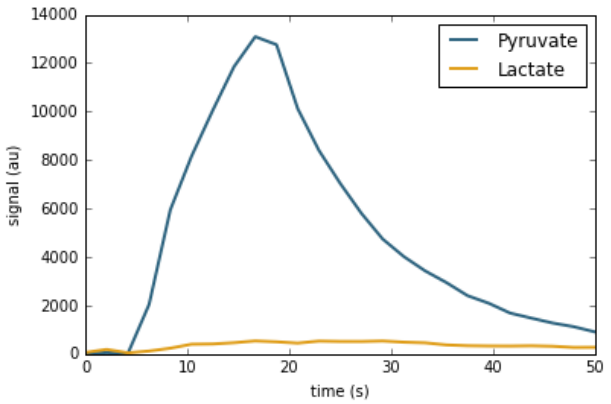
In Figure 7 we compare a spatially constrained fit of the data against an independent voxel-wise fit. The voxel-wise fit is masked to only show k_{PL} fit in the highly perfused regions where the total area under the pyruvate curve (AUC) is greater than 2e04. We see that the constrained fit leads to more smoothly-varying maps. Additionally, the Tikhonov regularization helps alleviate problems with artificially high k_{PL} estimates in the background region and tissues with low perfusion, a common problem with k_{PL} mapping from Rician-distributed data. This leads to more realistic k_{PL} values in the intestinal tissue proximal to the kidneys without significantly affecting the k_{PL} estimates in the kidney voxels, and also removes the need to mask the images to the high perfusion region.

To investigate the robustness of this technique to noise, we perform a sequence of experiments in which artificial iid Gaussian noise of varying strengths is added to the *in vivo* data using Python's `numpy.randn` random number generator before fitting k_{PL}. The random number generator is seeded explicitly using `numpy.random.seed(0)` to ensure reproducibility. This allows us to replicate the results of Figure 4 with more realistic anatomy. We see in Figure 8 that qualitatively, the spatially-constrained fit is more robust to strong noise than the independent fit. Further, we see in Figure 9 that spatially-constrained parameter mapping outperforms a baseline of simply downsampling the raw image sequence.

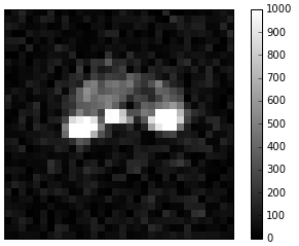
B. Human prostate cancer data analysis

To demonstrate feasibility of this technique on clinically-relevant data, we have analyzed two prostate cancer datasets collected during clinical experiments at UCSF. These datasets were chosen because they had relatively low SNR compared to our typical prostate cancer studies, and thus would potentially benefit the most from this approach.

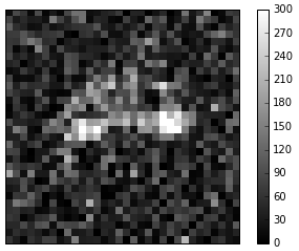
Imaging was performed using a 3T GE scanner using a abdominal clamshell ^{13}C transmission coil and an endo-rectal receive coil. The injected solution consisted of 220-260 mM [1-^{13}C]-pyruvate at a dose of 0.43 mL/kg. Dissolution DNP was performed using a 5T SpinLab polarizer (GE Healthcare).



(a) Sample time series data from high lactate SNR voxel

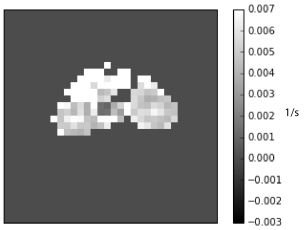


(b) Pyruvate image at time $t =$ 50s

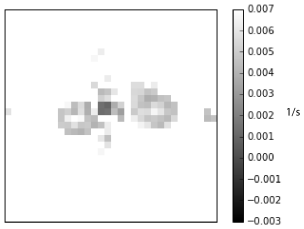


(c) Lactate image at time $t =$ 50s

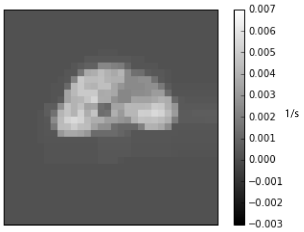
Fig. 6: Dynamic metabolite images collected in the healthy rat experiment. Maximum lactate SNR in these images is 21.1.



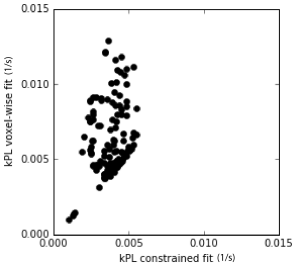
(a) Independent voxel-wise fit masked to region with pyruvate AUC $>$ 2e04.



(b) Independent voxel-wise fit without masking exhibits high k_{PL} values in the background region.

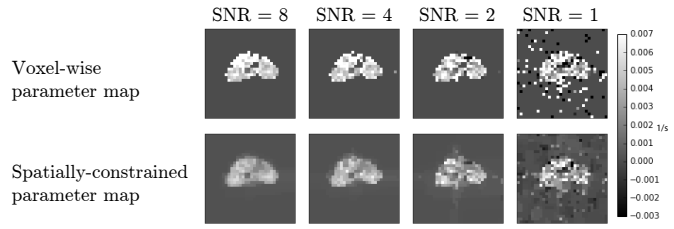


(c) Spatially-constrained fit with $\lambda_1 =$ 1e07 and $\lambda_2 =$ 1e10.

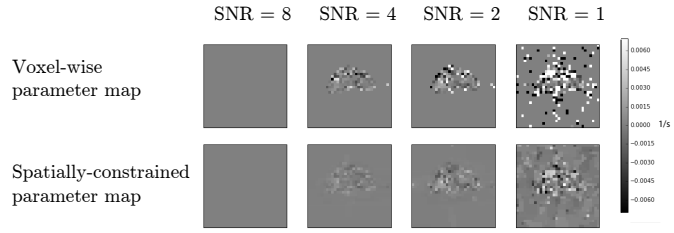


(d) Scatterplot of constrained and unconstrained k_{PL} fits.

Fig. 7: Comparison of unconstrained and constrained k_{PL} maps fit to the healthy rat dataset.



(a) Raw maps



(b) Difference maps using reconstruction without added noise as baseline

Fig. 8: Comparison of k_{PL} maps at various artificial noise levels. Noise level is measured based on maximum lactate SNR over the time and space dimensions in the dynamic images. Regularization parameters used for the constrained fits are chosen to be the same as in Figure 7.

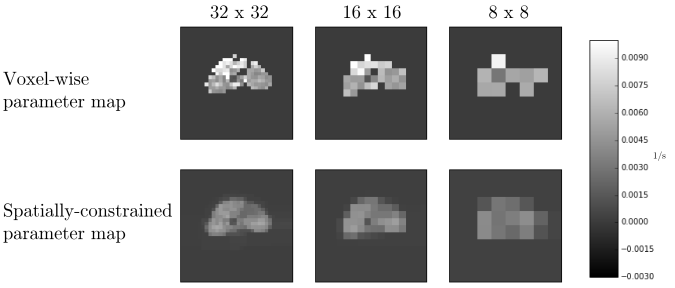


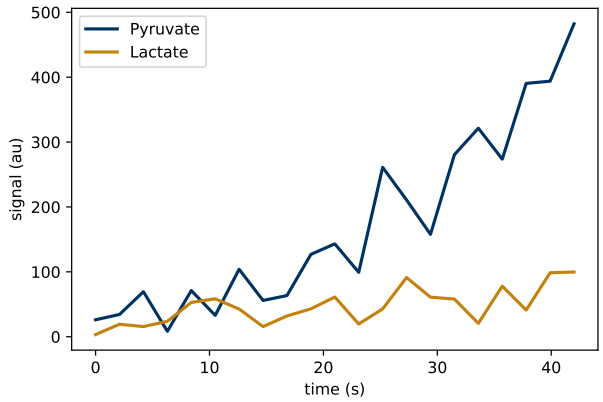
Fig. 9: Comparison of k_{PL} maps for varying spatial resolutions. Raw data is downsampled to the appropriate matrix size prior to fitting parameter maps for the independent voxel-wise and spatially-constrained fits.

Before injection the electron paramagnetic agent is filtered out, and automated pH, temperature, polarization, volume and EPA concentration tests were performed.

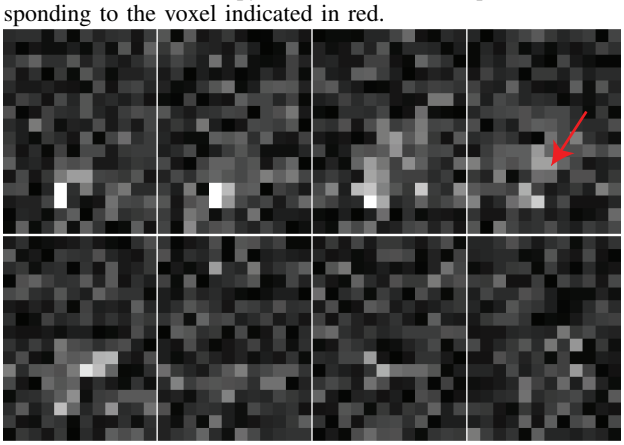
Images were encoded using two techniques. One set of images labeled "EPI" were collected using a spectrally-selective excitation with an echo-planar (EPI) readout [29]. The other set of images labelled "EPSI" was collected using a blipped EPSI acquisition with a compressed sensing reconstruction [30].

Raw space/time/chemical data reconstructed from the EPI acquisition are shown in Figure 10. The raw data are rather noisy and also difficult to interpret for metabolic activity due to 3D spatial, temporal and chemical dimensions.

We fit 3D k_{PL} parameter maps to the data using the constrained reconstruction method. Regularization strengths λ_1 and λ_2 are selected manually based on the qualitative



(a) Time series data at pyruvate and lactate frequencies corresponding to the voxel indicated in red.



(b) Lactate data from 8 of the 16 slices at the time of the final acquisition t=42 seconds from the start of injection.

Fig. 10: Sample of raw EPI data collected in a prostate cancer patient.

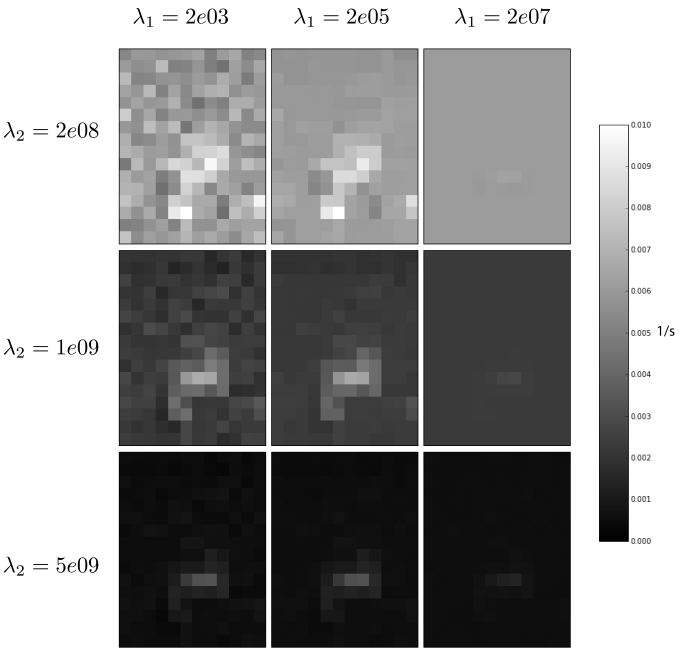
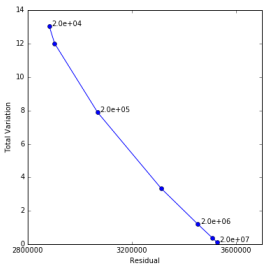
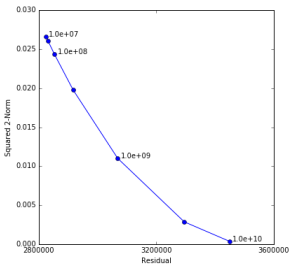


Fig. 11: Constrained estimates of the k_{PL} paramater with different regularization strengths compared on a single slice from the 3D EPI human prostate cancer dataset.



(a) L-curve for λ_1 for fixed $\lambda_2 = 1e09$.

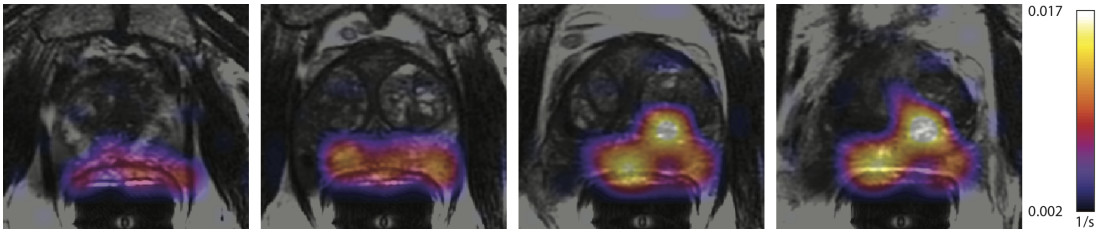


(b) L-curve for λ_2 for fixed $\lambda_1 = 2e05$.

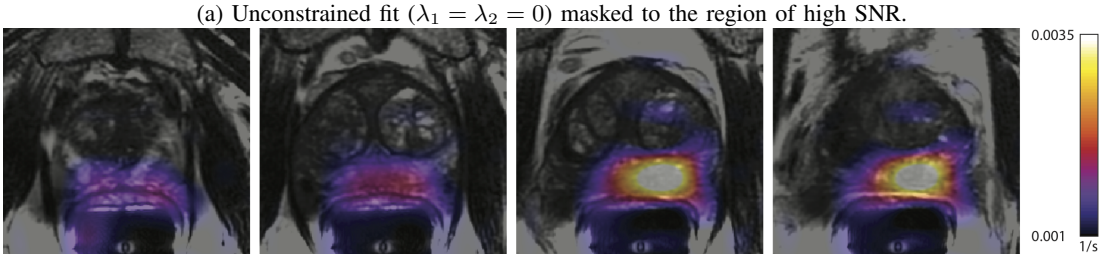
Fig. 12: L-curve analysis for the 3D EPI human prostate cancer dataset. The residual $\sum \ell(\theta_i|Y_i)$ is plotted against the regularizer $r(\theta)$ for various values of λ_1 and λ_2.

appearance of the parameter maps. Due to the quick parameter map estimation enabled by the parallelized ADMM iteration, it is possible to perform this hyperparameter exploration relatively efficiently. In Figure 11 we compare the resulting parameter maps for a variety of values for the regularization parameters λ_1 and λ_2. The results suggest that we should choose λ_1 large enough that the images do not appear noisy, but small enough that the signal does not disappear, and choose λ_2 large enough to suppress the bias in the unperfused region but small enough that it does not cause too much shrinkage in the perfused region. Figure 12 shows L-curves for the choice of λ_1 and λ_2, providing an alternative quantitative method of choosing parameters. We see that for very low or very high values of the regularization parameters, the regularization and residual terms cluster at the top left and bottom right of the figures respectively. Regularization parameter values approximately midway between the two clusters correspond to the qualitatively good parameter choices found in Figure 11. Additionally, in Figures 13 and 14 we compare unconstrained and constrained fits on the dataset from the EPI and EPSI acquisitions. The fits are overlaid on ^{1}H images of the anatomy using SIVIC [31]. The unconstrained fit is masked to voxels with a minimum pyruvate SNR due to fitting instability with low pyruvate signals, whereas this is not necessary for the constrained fit. We see that with an appropriate choice of regularization, we can recover qualitatively satisfying parameter maps for a variety of datasets. Note that the regularization parameters differ significantly between the EPI and EPSI acquisitions due mainly to the different amplitudes of the raw dynamic image data. Note that the strong regularization leads to significant quantitative shrinkage of the k_{PL} estimates. However, it improves the qualitative indication of the highly metabolically-active regions and removes noise-like characteristics of the fitting that is primarily due to low pyruvate SNR.

Figure 15 demonstrates how the constrained kPL maps could be integrated with the multi-parametric ^{1}H MRI into the clinical workflow to improve tumor localization and visualize treatment response. Elevated k_{PL} in the prostate regions

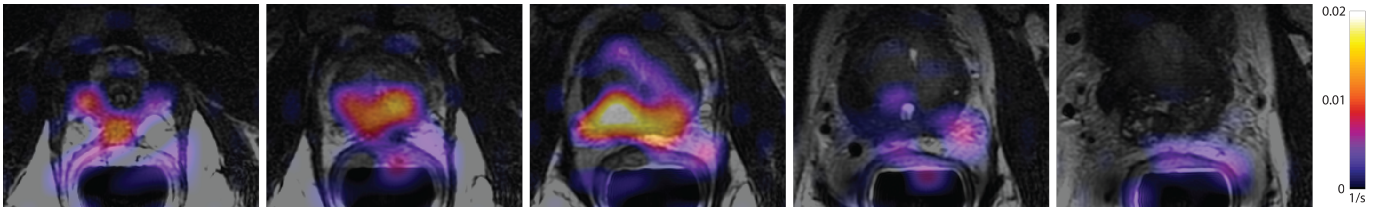


(a) Unconstrained fit ($\lambda_1 = \lambda_2 = 0$) masked to the region of high SNR.

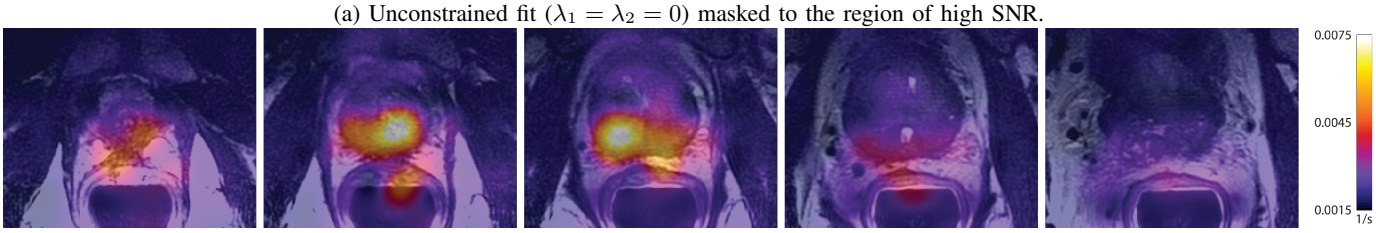


(b) Spatially-constrained fit ($\lambda_1 =$ 5e04 and $\lambda_2 =$ 1e09).

Fig. 13: Comparison of unconstrained and constrained k_{PL} maps fit to the 3D EPI data set overlaid on proton images of the prostate anatomy. Maps are plotted for four slices through the prostate with high lactate signal. This patient had biopsy proven cancer in the left base and midgland (Gleason 3+3 and 3+4), which is consistent with the results seen in the spatially-constrained k_{PL} fit.



(a) Unconstrained fit ($\lambda_1 = \lambda_2 = 0$) masked to the region of high SNR.



(b) Spatially-constrained fit ($\lambda_1 =$ 2e17 and $\lambda_2 =$ 1e14).

Fig. 14: Comparison of unconstrained and constrained k_{PL} maps fit to 3D EPSI data overlaid on prostate anatomy. Maps are plotted for five slices through the prostate with high lactate signal. This patient had extensive bilateral biopsy-proven prostate cancer (Gleason 4+4 and 4+3) involving the entire prostate. The spatially-constrained fit is consistent with significant bilateral disease, though the high k_{PL} region does not extend all the way to the prostate apex, likely due to its distance from the endo-rectal ^{13}C RF coil.

of Figures 13, 14 and 15 were consistent with biopsy and multiparametric (mp)-MRI [32] results. The patient studied in Figures 13 and 15A had biopsy proven cancer in the left base and midgland (Gleason 3+3 and 3+4). Their mp-MRI exam had an associated clear-cut region of reduced T_2 signal and water apparent diffusion coefficient (ADC), and enhanced uptake and washout on dynamic contrast enhanced (DCE) MRI in the left posterior peripheral zone of the midgland with extension across the midline. This is in strong agreement with the region of high k_{PL} shown with the constrained mapping in Figures 13 and 15A, which is in the left base and midgland with some extension across the midline. The patient studied in Figures 14 and 15B had extensive bilateral biopsy-proven prostate cancer (Gleason 4+4 and 4+3). mp-MRI demonstrated a large volume of prostate cancer involving the entire prostate, with right, posterior mid gland macroscopic extracapsular extension and bilateral seminal vesicle invasion. The k_{PL} fitting in Figures 14 and 15B also shows bilateral regions of high kPL, including the right, posterior midgland region identified by mp-MRI. The high k_{PL} does not, however, extend through the entire prostate, most likely due to low SNR further away from the endo-rectal ^{13}C RF coil sitting just below the prostate in the images. While further studies are required to fully evaluate the potential improvements in assessing cancer metabolism, this work demonstrates the feasibility and qualitative results of this approach on clinical

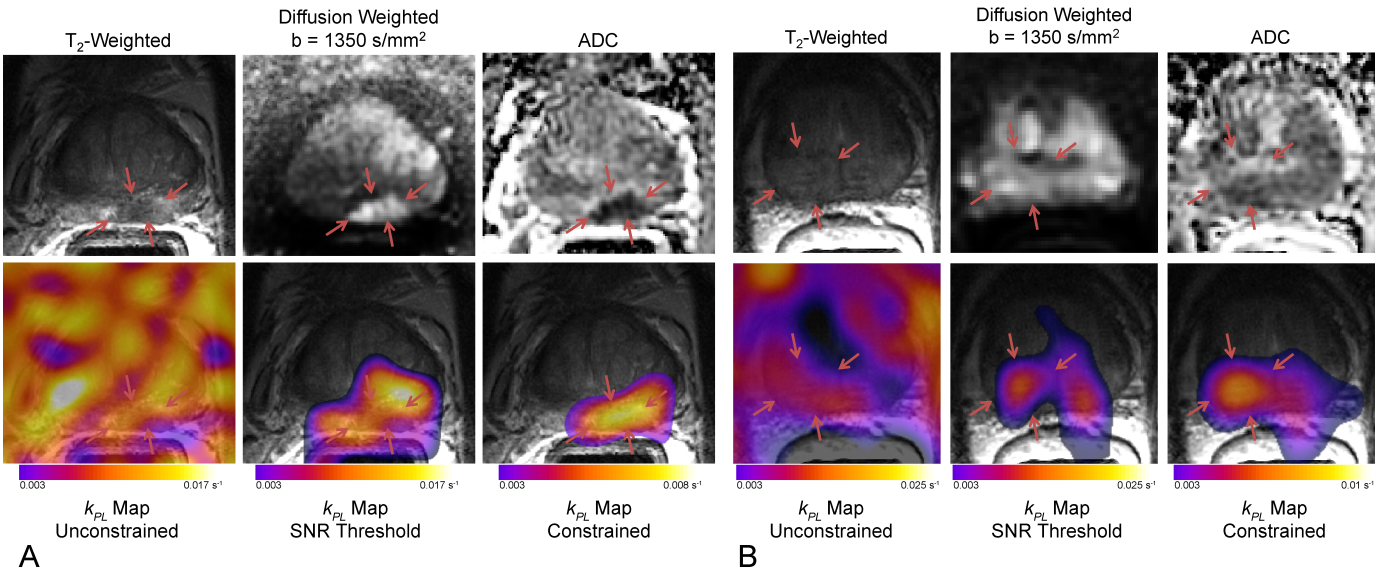


Fig. 15: Multi-parametric ^{1}H MRI and ^{13}C k_{PL} maps for the EPI (A) and EPSI (B) study showing the midgland prostate. Regions of high k_{PL} on the constrained reconstruction correlated well with biopsy proven aggressive cancer. It also agrees with lesions on multiparameteric MRI, including T_2-weighted, diffusion weighted, and ADC maps (red arrows). In contrast, the lesions are obfuscated by spurious noise on the unconstrained k_{PL} maps, or require an empirical hard threshold on the pyruvate signal to visualize.

datasets.

VII. Conclusion

We have demonstrated that constrained reconstruction of parameter maps via spatial regularization improves the qualitative performance of model-based parameter mapping. We have shown this first in simulated experiments where we can also demonstrate quantitative improvements in the parameter estimates. The results of the *in vivo* studies echo the qualitative benefits of constraining parameter maps through regularization, and validate that the ADMM-based algorithm we have presented enables efficient reconstruction of parameter maps for problems of practical interest by exploiting the objective function's structure.

Looking forward, the ability to exploit spatial and temporal correlations in the data for denoising could potentially help to overcome problems with low SNR in hyperpolarized 13C MRI, enabling the reconstruction of higher resolution k_{PL} maps. Also, developing methods to choose the regularization strength hyperparameters systematically may help to improve the quantitative bias seen in some of the *in vivo* experiments. In particular, methods based on Shure's unbiased risk estimate used for selecting hyperparameters in total variation denoising applications [33] can likely be adapted to this context. We suspect that the results of this paper could be further improved by replacing the ordinary least squares objective used by a weighted least squares objective where weights are chosen based on SNR, or based on an optimization problem based on maximizing Fisher information about the metabolic rate [34]. Finally, we would like to develop a better theoretical understanding of the ADMM algorithm's convergence on the non-convex optimization problem presented.

References

[1] K. Golman, J. H. Ardenkjær-Larsen, J. S. Petersson, S. Mansson, and I. Leunbach, "Molecular imaging with endogenous substances," Proceedings of the National Academy of Sciences, vol. 100, no. 18, pp. 10 435–10 439, 2003.

[2] S. E. Day, M. I. Kettunen, F. A. Gallagher, D.-E. Hu, M. Lerche, J. Wolber, K. Golman, J. H. Ardenkjær-Larsen, and K. M. Brindle, "Detecting tumor response to treatment using hyperpolarized ^{13}C magnetic resonance imaging and spectroscopy," Nature Medicine, no. 11, pp. 1382–1387, 2007.

[3] S. J. Nelson, J. Kurhanewicz, D. B. Vigneron, P. E. Z. Larson, A. L. Harzstark, M. Ferrone, M. van Criekinge, J. W. Chang, R. Bok, I. Park, G. Reed, L. Carvajal, E. J. Small, P. Munster, V. K. Weinberg, J. H. Ardenkjær-Larsen, A. P. Chen, R. E. Hurd, L.-I. Odegardstuen, F. J. Robb, J. Tropp, and J. A. Murray, "Metabolic imaging of patients with prostate cancer using hyperpolarized [1-^{13}C]pyruvate," Science Translational Medicine, vol. 5, no. 198, p. 198ra108, 2013.

[4] C. H. Cunningham, J. Y. Lau, A. P. Chen, B. J. Geraghty, W. J. Perks, I. Roifman, G. A. Wright, and K. A. Connelly, "Hyperpolarized ^{13}C metabolic MRI of the human heart: Novelty and significance," Circulation Research, vol. 119, no. 11, pp. 1177–1182, 2016.

[5] J. A. Bankson, C. M. Walker, M. S. Ramirez, W. Stefan, D. Fuentes, M. E. Merritt, J. Lee, V. C. Sandulache, Y. Chen, L. Phan, P.-C. Chou, A. Rao, S.-C. J. Yeung, M.-H. Lee, D. Schellingerhout, C. A. Conrad, C. Malloy, A. D. Sherry, S. Y. Lai, and J. D. Hazle, "Kinetic modeling and constrained reconstruction of hyperpolarized 1-13C-pyruvate offers improved metabolic imaging of tumors," Cancer Research, 2015.

[6] B. K. Felsted, R. T. Whitaker, M. Schabel, and E. V. DiBella, "Model-based reconstruction for undersampled dynamic contrast-enhanced MRI," in Medical Imaging 2009: Biomedical Applications in Molecular, Structural, and Functional Imaging, vol. 7262. International Society for Optics and Photonics, 2009, p. 72622S.

[7] S. G. Lingala, M. Nadar, C. Chefd'Hotel, L. Zhang, and M. Jacob, "Unified reconstruction and motion estimation in cardiac perfusion MRI," in Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on. IEEE, 2011, pp. 65–68.

[8] N. Dikaios, S. Arridge, V. Hamy, S. Punwani, and D. Atkinson, "Direct parametric reconstruction from undersampled (k, t)-space data in dynamic contrast enhanced MRI," Medical image analysis, vol. 18, no. 7, pp. 989–1001, 2014.

[9] J. C. Sommer, J. Gertheiss, and V. J. Schmid, "Spatially regularized estimation for the analysis of dynamic contrast-enhanced magnetic

resonance imaging data," Statistics in Medicine, vol. 33, no. 6, pp. 1029–1041, 2014.

[10] S. G. Lingala, Y. Guo, Y. Zhu, S. Barnes, R. M. Lebel, and K. S. Nayak, "Accelerated DCE MRI using constrained reconstruction based on pharmaco-kinetic model dictionaries," in Proceedings of the 23rd Annual Meeting of ISMRM, Toronto, Ontario, Abstract 1539, 2015.

[11] Y. Guo, Y. Zhu, S. G. Lingala, R. M. Lebel, and K. S. Nayak, "Highly accelerated brain DCE MRI with direct estimation of pharmacokinetic parameter maps," in Proceedings of the 23rd Annual Meeting of ISMRM, Toronto, Ontario, Abstract 1093, 2015.

[12] Y. Guo, R. M. Lebel, Y. Zhu, S. G. Lingala, M. S. Shiroishi, M. Law, and K. Nayak, "High-resolution whole-brain DCE-MRI using constrained reconstruction: Prospective clinical evaluation in brain tumor patients," Medical Physics, vol. 43, no. 5, pp. 2013–2023, 2016.

[13] J. Maidens, J. W. Gordon, M. Arcak, H.-Y. Chen, I. Park, M. V. Criekinge, E. Milshteyn, R. Bok, R. Aggarwal, M. Ferrone, J. B. Slater, J. Kurhanewicz, D. B. Vigneron, and P. E. Z. Larson, "Spatio-temporally constrained reconstruction for hyperpolarized carbon-13 MRI using kinetic models," in Proceedings of the ISMRM Annual Meeting, 2017. [Online]. Available: http://indexsmart.mirasmart.com/ISMRM2017/PDFfiles/3040.html

[14] C. Harrison, C. Yang, A. Jindal, R. J. DeBerardinis, M. A. Hooshyar, M. Merritt, A. Dean Sherry, and C. R. Malloy, "Comparison of kinetic models for analysis of pyruvate-to-lactate exchange by hyperpolarized ^{13}C NMR," NMR in Biomedicine, vol. 25, no. 11, pp. 1286–1294, 2012.

[15] P. J. Huber and E. M. Ronchetti, Robust Statistics. John Wiley & Sons, Inc., 2009.

[16] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, 2nd ed. Springer, 2009.

[17] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," SIAM review, vol. 43, no. 1, pp. 129–159, 2001.

[18] R. Tibshirani, "Regression shrinkage and selection via the lasso," Journal of the Royal Statistical Society. Series B (Methodological), pp. 267–288, 1996.

[19] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," Physica D: Nonlinear Phenomena, vol. 60, no. 1-4, pp. 259–268, 1992.

[20] A. Barbero and S. Sra, "Fast newton-type methods for total variation regularization." in International Conference on Machine Learning (ICML), L. Getoor and T. Scheffer, Eds. Omnipress, 2011, pp. 313–320.

[21] ——, "Modular proximal optimization for multidimensional total-variation regularization," 2014. [Online]. Available: http://arxiv.org/abs/1411.0589

[22] D. Strong and T. Chan, "Edge-preserving and scale-dependent properties of total variation regularization," Inverse problems, vol. 19, no. 6, p. S165, 2003.

[23] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," Foundations and Trends in Machine Learning, vol. 3, no. 1, pp. 1–122, 2011.

[24] L. Vandenberghe, "Lecture 13: Douglas-Rachford method and ADMM," 2016. [Online]. Available: https://web.archive.org/web/20170405001209/http://www.seas.ucla.edu/~vandenbe/236C/lectures/dr.pdf

[25] J. E. Dennis Jr and R. B. Schnabel, Numerical methods for unconstrained optimization and nonlinear equations. SIAM, 1996.

[26] J. J. Moré, B. S. Garbow, and K. E. Hillstrom, "User guide for MINPACK-1," Argonne National Laboratory, Tech. Rep. ANL-80-74, 1980.

[27] J. Maidens, "Optimal control for learning with applications in dynamic MRI," Ph.D. dissertation, EECS Department, University of California, Berkeley, Aug 2017. [Online]. Available: http://www2.eecs.berkeley.edu/Pubs/TechRpts/2017/EECS-2017-135.html

[28] Y. Xing, G. D. Reed, J. M. Pauly, A. B. Kerr, and P. E. Larson, "Optimal variable flip angle schemes for dynamic acquisition of exchanging hyperpolarized substrates," Journal of Magnetic Resonance, vol. 234, pp. 75–81, 2013.

[29] J. W. Gordon, D. B. Vigneron, and P. E. Larson, "Development of a symmetric echo planar imaging framework for clinical translation of rapid dynamic hyperpolarized 13C imaging," Magnetic Resonance in Medicine, vol. 77, no. 2, pp. 826–832, 2017.

[30] P. E. Z. Larson, S. Hu, M. Lustig, A. B. Kerr, S. J. Nelson, J. Kurhanewicz, J. M. Pauly, and D. B. Vigneron, "Fast dynamic 3D MR spectroscopic imaging with compressed sensing and multiband excitation pulses for hyperpolarized 13C studies," Magnetic Resonance in Medicine, vol. 65, no. 3, pp. 610–619, 2011.

[31] J. C. Crane, M. P. Olson, and S. J. Nelson, "SIVIC: Open-source, standards-based software for DICOM MR spectroscopy workflows," International Journal of Biomedical Imaging, vol. 2013, p. 169526, 2013.

[32] J. Kurhanewicz, D. Vigneron, P. Carroll, and F. Coakley, "Multiparametric magnetic resonance imaging in prostate cancer: present and future," Current Opinion in Urology, vol. 18, no. 1, pp. 71–77, 2008.

[33] V. Solo, "Selection of regularisation parameters for total variation denoising," in IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 3, 1999, pp. 1653 – 1655.

[34] J. Maidens, J. W. Gordon, M. Arcak, and P. E. Z. Larson, "Optimizing flip angles for metabolic rate estimation in hyperpolarized carbon-13 MRI," IEEE Transactions on Medical Imaging, vol. 35, no. 11, pp. 2403–2412, 2016.