Control and systems theory problems in hyperpolarized carbon-13 MRI

John Maidens and Murat Arcak

Abstract Hyperpolarized carbon-13 magnetic resonance imaging (MRI) is an emerging technology for probing metabolic activity in living subjects, which promises to provide clinicians new insights into diseases such as cancer and heart failure. These experiments involve an injection of a hyperpolarized substrate, often [1-13C]pyruvate, which is imaged over time as it spreads throughout the subject's body and is transformed into various metabolic products. Designing these dynamic experiments and processing the resulting data requires the integration of noisy information across temporal, spatial, and chemical dimensions, and thus provides a wealth of interesting problems from an optimization and control perspective.

In this work we provide an introduction to the field of hyperpolarized carbon-13 MRI targeted toward researchers in control and systems theory. We then describe three challenge problems that arise in metabolic imaging with hyperpolarized substrates: the design of optimal substrate injection profiles, the design of optimal flip angle sequences, and the constrained estimation of metabolism maps from experimental data. We describe the current state of research on each of these problems, and comment on aspects that remain open. We hope that these challenge problems will serve to direct future research in control and systems theory.

John Maidens
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley
Cory Hall, Berkeley, CA, 94704, USA
e-mail: maidens@eecs.berkeley.edu

Murat Arcak
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley
Cory Hall, Berkeley, CA, 94704, USA
e-mail: arcak@eecs.berkeley.edu

1 Introduction

Carbon is arguably the most important element in biochemistry. It forms the basis of all organic molecules that make up the human body, yet only recently have we begun to be able to image carbon using magnetic resonance imaging (MRI). The emerging technology that makes this possible is known as hyperpolarized carbon-13 MRI, and it has enabled *in vivo* imaging with spatial, temporal and chemical specificity for the first time. This development is leading to new insights into the spatial distribution of metabolic activity through the analysis of dynamic image sequences.

The processes that are imaged in hyperpolarized carbon-13 MRI are inherently dynamic, resulting from blood flow, tissue perfusion, metabolic conversion, and polarization decay. Thus there is an opportunity for control and systems theory researchers to improve the dynamic models, excitation inputs and estimation algorithms used in hyperpolarized carbon-13 MRI.

The remainder of this paper is organized as follows. In Section 2 we present the basics of hyperpolarized carbon-13 MRI. In Section 3 we discuss how this technology enables the quantification of metabolism and the clinical significance of this advancement. In Section 4 we present a dynamic model of metabolic flux and discuss methods of estimating model parameters from experimental MRI data. In Section 5 we discuss formal formulations of optimal design for dynamic experiments. Finally, in Section 6 we present three control and systems theory problems that arise in metabolic MRI using hyperpolarized carbon-13 and discuss open questions.

2 Hyperpolarized carbon-13 MRI for imaging metabolism

The measurable signal in MRI arises from radio-frequency electromagnetic waves generated by oscillating atomic nuclei. Nuclei containing an odd number of protons and/or neutrons possess a nuclear spin angular momentum, each giving rise to a small magnetic moment. Thus nuclei such as carbon (^{12}C) and oxygen (^{16}O) are invisible to MRI, while hydrogen (^{1}H) and the carbon-13 isotope (^{13}C) exhibit magnetic resonance (MR). Hydrogen MR, sometimes known as proton MR, is currently the most commonly-used in clinical settings due to the high abundance of hydrogen atoms in the human body (largely in the form of H_2O) and its high sensitivity [20]. Hydrogen MRI is pervaisive for noninvasive imaging of anatomic structure, but provides little functional information. In this work, we focus on carbon-13 MR, which can be used to provide information about metabolic function (Figure 1).

2.1 Polarization

Each carbon-13 nucleus in a sample gives rise to a small magnetic moment, which we think of as a vector in three-dimensional space. In the absence of a magnetic

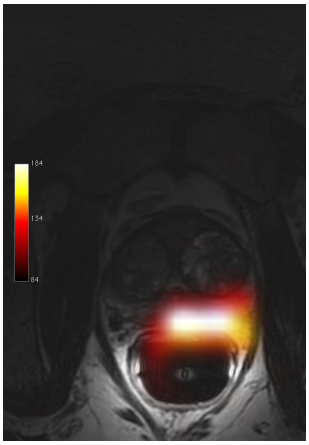
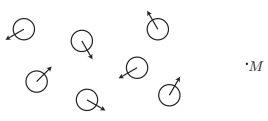
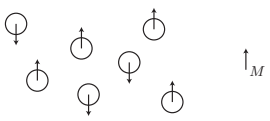


Fig. 1 Metabolic image collected in a clinical prostate cancer study. The color overlay corresponds to a metabolic map of a lactic acid fermentation pathway generated using hyperpolarized carbon-13 MRI. High metabolic activity coincides with the region of biopsy-proven prostate cancer. Image courtesy of Dr. Jeremy Gordon, Department of Radiology and Biomedical Imaging, UCSF School of Medicine.

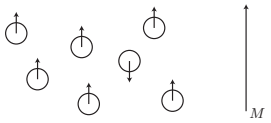
field these magnetic moments, or spins, are oriented randomly in space. In this state, where the spins are oriented randomly, the net magnetization $M = (M_x, M_y, M_z)$ of the sample (computed by summing the magnetic moments from all nuclei in the sample) is zero (Figure 2a).



(a) Spins in the absence of an external magnetic field are oriented randomly, leading to a net magnetization $M = 0$.



(b) In a magnetic field, spins orient themselves with the field, leading to a small net magnetization M.



(c) Hyperpolarizing the sample causes the spins to coordinate direction, leading to a greater net magnetization M.

Fig. 2: Polarization of a collection of spins leads to a net magnetization M. The magnitude of the vector M can be increased by hyperpolarization.

In the presence of a magnetic field, the spins will orient themselves to the magnetic field with some aligned parallel (n^+) with the magnetic field lines and some aligned anti-parallel (n^-) with the magnetic field lines. Due to a small energy difference between the two states, there is a slight bias toward the lower energy state aligned parallel to the magnetic field. This bias leads to a polarization of the sample, defined as the excess of spins in the lower energy state. At thermal equilibrium, the polarization P is given in terms of the applied magnetic field strength B_0, ambient temperature T, gyromagnetic ratio γ of the nucleus (10.705 MHz/T for ^{13}C), and the Boltzmann (k_B) and reduced Planck ($\hbar$) constants as

$$P = \frac{n^+ - n^-}{n^+ + n^-} = \tanh\left(\frac{\gamma\hbar B_0}{2k_B T}\right).$$

This slight bias leads to a small net magnetization of the sample at thermal equilibrium (Figure 2b) given by

$$M = \begin{bmatrix} 0 \\ 0 \\ M_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \frac{N\gamma^2\hbar^2 I_z(I_z+1)B_0}{3k_B T} \end{bmatrix}$$

where N is the number of spins in the sample and I_z is the spin operator. Note that we use the standard convention that the z axis is chosen such that it points in the direction of the applied B_0 field.

Due to the low natural abundance of carbon-13 in the body and its low gyromagnetic ratio, the thermal equilibrium magnetization is insufficient to achieve sufficient signal for imaging. Thus carbon-13-based imaging relies on hyperpolarization technology to increase the polarization beyond the equilibrium level (Figure 2c).

2.2 Hyperpolarization using DNP

Hyperpolarized carbon-13 MRI has been enabled by new technologies for hyperpolarizing carbon-13-containing substrates in liquid state, leading to a greater than $10000\times$ increase in signal-to-noise ratio (SNR) when imaging carbon-13. This technology relies on dissolution dynamic nuclear polarization (D-DNP) to achieve significant polarization gains [1].

Dynamic nuclear polarization relies on transferring polarization to carbon-13 nuclei from electrons using microwave radiation. In this procedure, a sample is doped with a small quantity of stable electron radical. The sample is then cooled to a temperature of 1.1 K and placed in a 3.35 T magnet. At this temperature and magnetic field strength, electrons become almost 100% polarized. Then by irradiating the sample with microwaves, polarization is transferred from the electrons to the carbon-13 nuclei in a biochemical substrate of interest. To prepare the sample for injection and *in vivo* imaging, it is then rapidly dissolved in warm water, neutralized to a safe pH and the electron radical is removed before injection [19].

2.3 Polarization decay in hyperpolarized substrates

Upon warming and removal from the magnet, the magnetization induced by hyperpolarization begins to decay over time toward the thermal equilibrium magnetization due to a phenomenon known as T_1 relaxation. The dynamics of the magnetization vector are governed by a system of state equations know as the rotating frame Bloch equations:

$$\frac{d}{dt}\begin{bmatrix} M_x \\ M_y \\ M_z \end{bmatrix} = \begin{bmatrix} -\frac{1}{T_2} & u_2 & 0 \\ -u_2 & -\frac{1}{T_2} & u_1 \\ 0 & -u_1 & \frac{1}{T_1} \end{bmatrix} \begin{bmatrix} M_x \\ M_y \\ M_z \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{M_0}{T_1} \end{bmatrix} \tag{1}$$

with initial condition $M(0) = (0,0,M_z(0))$. Here, the evolution of the state M is dependent on a sequence of control inputs u_1 and u_2 corresponding to the amplitude and frequency of the applied radio-frequency (RF) electromagnetic excitation pulse (known as the B_1 field) that rotates the vector M about the origin, and T_1 and T_2 parameters that govern the relaxation time in the longitudinal (z) and transverse (x,y) directions respectively.

When the sample is hyperpolarized we have $M_z(0) \gg M_0$, therefore the contribution of the affine term in (1) is negligible. Thus in the absence of RF excitation, the longitudinal magnetization exhibits exponential the decay

$$M_z(t) = M_z(0)e^{-t/T_1}.$$

In addition to T_1 relaxation, magnetization also decays due to repeated RF excitation. Throughout this paper we will assume that the RF pulse occurs on a time scale much faster than T_1 and T_2, therefore it can be modelled as an instantaneous state reset that rotates M to some angle α away from the z axis, known as the flip angle. We also assume that a spoiled gradient echo pulse sequence [2] is used, thus between RF pulses a strong magnetic field gradient is applied to dephase the transverse magnetization ensuring that $M_x = M_y = 0$. Thus at a time t^+ immediately after an RF pulse, the magnetization is given in terms of the magnetization at time t^- immediately before the RF pulse as

$$\begin{aligned} M_z(t^+) &= \cos(\alpha)M_z(t^-) \\ M_{xy}(t^+) &:= \sqrt{M_x(t^+)^2 + M_y(t^+)^2} = \sin(\alpha)M_z(t^-). \end{aligned}$$

It now follows that at a time t following a sequence of RF pulses with flip angles $\alpha_0, \ldots, \alpha_{N-1}$ the longitudinal magnetization remaining has decayed to

$$M_z(t) = M_z(0)e^{-t/T_1} \prod_{k=0}^{N-1} \cos(\alpha_k).$$

2.4 Chemical shift

The unique aspect of hyperpolarized carbon-13 MRI, when compared to competing metabolic imaging technologies such as positron emission tomography (PET), is that it is the only technique that provides chemical specificity. It is possible to infer chemical information from MRI data due to a phenomenon known as chemical shift.

Chemical shift results in a small change in the resonant frequency of spins. This change is caused by shielding of the nuclei from the main magnetic field B_0 due to

nearby electron orbitals [20]. The resulting frequency shift can be exploited to selectively excite specific metabolites [13], or distinguish between metabolites produced. This gives hyperpolarized carbon-13 MRI the unique ability to quantify metabolic flux in specific pathways.

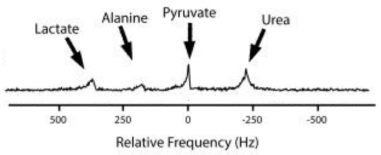


Fig. 3 MR spectrum illustrating chemical shift of ^{13}C biochemical compounds. Image courtesy of Dr. Jeremy Gordon, Department of Radiology and Biomedical Imaging, UCSF School of Medicine.

3 Metabolism and Disease

Hyperpolarized carbon-13 MRI is uniquely positioned to investigate metabolic disease biomarkers due to its ability to track the metabolism of injected substrates with spatial, temporal and chemical specificity. In practice, a commonly-used substrate is [1-^{13}C]pyruvate, which consists of pyruvate molecules labeled with a carbon-13 atom at the first carbon position (Figure 4).

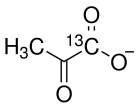


Fig. 4 [1-^{13}C] pyruvate molecule.

Pyruvate plays an important role in cellular metabolism, as the end product of glycolysis. Following glycolysis, under anaerobic conditions pyruvate is typically converted to lactate via lactic acid fermentation, which serves as a short-term means of producing energy. Activity in this pathway can be quantified using carbon-13 MRI via observed [1-^{13}C]lactate signal. Under aerobic conditions, pyruvate can also serve as an input to the citric acid cycle, which produces cellular energy in the form of adenosine triphosphate (ATP) through cellular respiration in the mitochondria. Activity in this pathway can be quantified via pyruvate flux through the pyruvate dehydrogenase complex (PDC), which is proportional to observed ^{13}C-bicarbonate signal in carbon-13 MRI studies [18].

3.1 Pyruvate metabolism changes in cancer

Cancer results in changes to metabolism that can be used to diagnose and monitor treatment response. In particular, a phenomenon known as the Warburg Effect results in increased lactic acid fermentation in place of cellular respiration in cancerous tissues even under aerobic conditions [7]. Thus the rate of production of [1-^{13}C]lactate from injected hyperpolarized [1-^{13}C]pyruvate can be used to monitor cancer metabolism *in vivo*. The ability of hyperpolarized carbon-13 MRI to distinguish cancerous tissue from healthy tissue been demonstrated in animal model [6] and clinical human prostate cancer [19] studies.

3.2 Pyruvate metabolism changes in heart failure

Metabolic changes in the heart have also been proposed as a significant contributing factor to congestive heart failure [24]. Hyperpolarized carbon-13 MRI studies are being used to study this hypothesis by quantifying metabolic fluxes in the heart. Hyperpolarized carbon-13 MRI has been used to demonstrate that the development of heart failure leads to decreased [1-^{13}C]pyruvate flux through the PDC in a porcine model of dilated cardiomyopathy [22]. Recent clinical studies have also demonstrated feasibility of measuring PDC flux in heart tissue using hyperpolarized [1-^{13}C]pyruvate in healthy human subjects [5].

4 Quantifying Metabolic Flux

Hyperpolarized carbon-13 MRI enables dynamic experiments that show metabolic activity with spatial, temporal and chemical specificity. This enables quantifying the spatial distribution of the activity of specific metabolic pathways. In this section, we discuss model-based methods of fusing this information into spatial maps of metabolic activity. This is done by estimating kinetic parameters in a model describing the evolution of the MR signal observed in each spatial volume element (voxel).

4.1 Kinetic models of hyperpolarized MRI signal in a single voxel

Hyperpolarized carbon-13 MRI researchers commonly rely on linear compartmental models for describing the evolution of signal in a voxel [4, 11, 12]. These models describe the magnetization exchange from the pool of injected hyperpolarized substrate to pools corresponding to various metabolic products. In the its simplest form, this amounts to the irreversible metabolic conversion of the substrate S to a single

product P performed at a characteristic kinetic rate k_{SP}:

$$S \xrightarrow{k_{SP}} P.$$

Throughout this article, we will focus on extremely simple pathways of this form, though extension to multiple products or bidirectional conversion is straightforward. Note that the clinically-relevant pathways discussed in Sections 3.1 and 3.2 can both be modelled in this simple form.

In the absence of external RF excitation, magnetization in a particular voxel i evolves via T_1 decay and label exchange according to the differential equations

$$\frac{d}{dt}\begin{bmatrix} M_{z,i,S}(t) \\ M_{z,i,P}(t) \end{bmatrix} = \begin{bmatrix} -R_{1,i,S} - k_{SP,i} & 0 \\ k_{SP,i} & -R_{1,i,P} \end{bmatrix} \begin{bmatrix} M_{z,i,S}(t) \\ M_{z,i,P}(t) \end{bmatrix} + \begin{bmatrix} k_{TRANS,i} \\ 0 \end{bmatrix} u(t) \tag{2}$$

where the states $M_{z,i,S}$ and $M_{z,i,P}$ represent the longitudinal magnetization in voxel i in the substrate and product compartments respectively, the input u model an arterial input function (AIF) describing the arrival of substrate from the circulatory system, and the parameters $k_{SP,i}$, $R_{1,i,S}$, $R_{1,i,P}$, and k_{TRANS} describe the metabolic rate, T_1 decay rate in the substrate pool, and T_1 decay rate in the product pool, and perfusion rate respectively.

When a constant flip angle excitation sequence and repetition time is used for imaging, decay due to RF excitation can be modelled by replacing $R_{1,i,X}$ by an effective decay rate

$$R_{1,i,X,\text{effective}} = R_{1,i,X} - \frac{\log(\cos\alpha)}{T_R}$$

where α is the flip angle and T_R is the repetition time, and X denotes an arbitrary compound (either S or P) [23]. However, when a variable flip angle sequence is used, signal decay due to RF excitation must be accounted for as in Section 2.3. This leads to a discrete time model for the transverse and longitudinal magnetization immediately preceding excitation k given by

$$\begin{bmatrix} M_{z,i,S}[k+1] \\ M_{z,i,P}[k+1] \end{bmatrix} = A_d \begin{bmatrix} \cos\alpha_S[k] & 0 \\ 0 & \cos\alpha_P[k] \end{bmatrix} \begin{bmatrix} M_{z,i,S}[k] \\ M_{z,i,P}[k] \end{bmatrix} + B_d u[k] \tag{3}$$

where A_d and B_d are computed from (2) assuming a zero order hold on the arterial input. A model for the transverse magnetization immediately following excitation k given by

$$M_{xy,i,X}[k] = \sin\alpha_X[k] M_{z,i,X}[k], \tag{4}$$

which describes the observable outputs from voxel i at time k.

4.2 Estimation of unknown model parameters

Estimating metabolic rate parameters θ_i from experimental data collected from voxel i involves minimizing a statistical loss function $L(\theta_i|Y_i)$ that describes how well a signal model fits the observed data Y_i. Using the model equations (3)–(4) as the basis of a signal model describing the predicted measurement

$$y_i(\theta_i) = \left[[M_{xy,i,S}[1]\ M_{xy,i,P}[1]\ \ldots\ M_{xy,i,S}[N]\ M_{xy,i,P}[N]\right]$$

in terms of the vector model parameters θ_i. Loss functions include:

- the least squares loss
$$L(\theta_i|Y_i) = \|Y_i - y_i(\theta)\|^2$$
which corresponds to a nonlinear least squares estimation problem and
- the negative log likelihood loss
$$L(\theta_i|Y_i) = -\log p_{\theta_i}(Y_i)$$
which corresponds to a maximum likelihood estimation problem. Unlike the least squares loss function, this loss requires a that a probability density function describing the joint distribution of Y_i be specified. Common choices are $Y_i \sim y_i + \varepsilon$ where ε is independent, identically-distributed (iid) Gaussian noise or independent Rician noise with location parameters given by y_i [10].

5 Optimal Design of Dynamic Experiments

Two of the three problems we will discuss in this paper address the design of optimized experiments for estimating the value of unknown parameters in a mathematical model of a dynamical system from noisy output data. Thus, in this section we provide background on optimal experiment design.

In dynamical systems with noisy outputs, the reliability of the parameter estimates depends on the choice of input used to excite the system, as some inputs provide much greater information about the parameters than others. Much work has been done on the optimal experiment design problem in the last 50 years [8, 9, 14, 21, 25]. Historically, a great deal of work on this problem has taken a frequency domain approach, where the input to the system is designed based on its power spectrum. Here, we will approach this problem in the time domain, hoping to be able to perform experiment design for systems with nonlinear dynamics.

5.1 Problem Description

We consider a discrete-time dynamical system with noisy observations

$$\begin{aligned} x_{t+1} &= f(t, x_t, u_t, \theta) \\ Y_t &\sim P_{x_t} \end{aligned} \tag{5}$$

where $x_t \in \mathbb{R}^n$ denotes the system's state, $u_t \in \mathbb{R}^m$ is a sequence of inputs to be designed and $\theta \in \mathbb{R}^p$ is a vector of unknown parameters that we wish to estimate. Observations are drawn independently from a known distribution that is parametrized by the system state x_t. We assume that for all $x_t \in \mathbb{R}^n$ the probability distribution P_{x_t} is absolutely continuous with respect to some measure μ and we denote its density with respect to μ by $p_{x_t}(y_t)$. We consider this system over a finite horizon $0 \leq t \leq N$. Our goal is to design a sequence u that provides a maximal amount of information about the unknown parameter vector θ. This problem can be addressed by maximizing the Fisher information about θ.

5.2 Fisher Information

An important notion in frequentist statistics is the Fisher information matrix for the vector of model parameters θ. The Fisher information is fundamental in the analysis of numerous statistical estimators from unbiased estimation to maximum-likelihood estimation. We begin with a definition.

Definition 1. Let $\mathscr{P} = \{P_\theta : \theta \in \Omega\}$ be a family of probability distributions parametrized by θ in an open set $\Omega \subseteq \mathbb{R}^p$ and dominated by some measure μ. Denote the probability densities with respect to μ by p_θ and assume that the densities are differentiable with respect to θ. We define the Fisher information matrix as the $p \times p$ matrix $\mathscr{I}(\theta)$ with (i, j)-th entry defined as

$$\mathscr{I}(\theta)_{i,j} = \mathbb{E}\left[\frac{\partial \log p_\theta(Y)}{\partial \theta_i} \frac{\partial \log p_\theta(Y)}{\partial \theta_j}\right]$$

where $Y \sim P_\theta$.

5.3 The Cramér-Rao Inequality for Unbiased Estimators

The Cramér-Rao inequality provides a lower bound on the achievable covariance of any unbiased estimator of θ in terms of the Fisher information matrix.

Theorem 1 (Cramér-Rao Bound). *Let $\delta(Y)$ be any unbiased estimator of the parameter θ. Under the conditions that*

- $\mathscr{I}(\theta)$ *exists and is nonsingular*
- $\frac{\partial}{\partial\theta}\left[\int \delta(y)p_\theta(y)dy\right] = \int \delta(y)\left[\frac{\partial}{\partial\theta}p_\theta(y)\right]dy$

the estimator δ *satisfies*

$$cov(\delta(Y)) \geq \mathscr{I}(\theta)^{-1}$$

where $A \geq B$ *indicates that* $A - B$ *is positive semidefinite.*

5.4 Asymptotic distribution of the Maximum-Likelihood Estimator

Let $Y^1, \ldots Y^n$ be n random variables drawn independently from the probability distribution P_θ. Given that some mild regularity conditions are satisfied, as $n \to \infty$ the maximum-likelihood estimator

$$\hat{\theta}_{MLE} = \arg\max_\theta \sum_{i=1}^{n} \log p_\theta(Y^i).$$

converges in distribution to a normal distribution with covariance $\mathscr{I}(\theta)^{-1}$.

Theorem 2 (Asymptotics of the MLE). *Suppose that*

- *the MLE is consistent*
- $p_\theta(Y)$ *is bounded and* C^2 *in* θ
- $\mathscr{I}(\theta)$ *exists and is nonsingular*
- $\mathbb{E}[\|\log H_\theta(Y)\|] < \infty$ *where* H *denotes the Hessian of* $p_\theta(Y)$ *with respect to* θ.

Then $\sqrt{n}(\hat{\theta}_{MLE} - \theta) \xrightarrow{d} N(0, \mathscr{I}(\theta)^{-1})$.

5.5 Scalar Measures of Positive Semidefinite Matrices

In order to use the Fisher information as an objective function for choosing an input sequence u, we must choose a method of summarizing $\mathscr{I}(\theta)$ by a scalar quantity. We survey some of the most common choices here.

- **Nonnegative linear functions**
  The most convenient class of objective functions are those that are linear functions of the information matrix. For any positive semidefinite matrix $K \in \mathbb{R}^{p\times p}$ we can define the objective function

  $$\phi(\mathscr{I}) = \mathrm{tr}(K\mathscr{I}).$$

  Particular cases are the "T-optimal" design criterion (when K is the identity matrix) and the c-optimal design criterion (when $K = cc^T$ has rank 1), which is optimal for the scalar parameter $c^T\theta$.

- **Other convex functions**
  A number of other functions of the Fisher information matrix are also commonly used as scalar measures of its size. Choices include:
  - $\phi_D(\mathscr{I}) = \log\det(\mathscr{I})$ know as "D-optimal design,"
  - $\phi_E(\mathscr{I}) = \lambda_{\min}(\mathscr{I})$ known as "E-optimal design", and
  - $\phi_A(\mathscr{I}) = 1/\operatorname{tr}(\mathscr{I}^{-1})$ known as "A-optimal design."

To understand the geometric meaning of each of these optimal design criteria, consider an asymptotic confidence region associated with the MLE of θ. Such a region is described by the ellipsoid $\mathscr{E} = \{\tilde{\theta} : n(\tilde{\theta} - \hat{\theta}_{MLE})^T \mathscr{I}(\tilde{\theta} - \hat{\theta}_{MLE}) \leq \varepsilon\}$ for some $\varepsilon > 0$. The volume of this ellipsoid is proportional to $\sqrt{\det(\mathscr{I}^{-1})}$ and therefore the D-optimal design criterion corresponds to minimizing the volume of the confidence ellipsoid $\mathscr{E}$. If the spectrum of $\mathscr{I}$ is denoted $\lambda_1, \ldots, \lambda_p$ then the length of the axes of $\mathscr{E}$ are $\frac{1}{\sqrt{\lambda_1}}, \ldots, \frac{1}{\sqrt{\lambda_p}}$. So the E-optimal design criterion corresponds to minimizing the length of the longest axis of the confidence ellipsoid and the A-optimal design criterion corresponds to minimizing the average squared length of the confidence ellipsoid axes.

5.6 Optimizing Under Parametric Uncertainty

Given a particular scalar-valued function ϕ with which we will measure the size of the Fisher information, the optimal experiment design problem is given by

$$\begin{aligned} &\textbf{maximize}_u \;\; \phi(\mathscr{I}(\theta)) \\ &\textbf{subject to} \;\; u \in \mathscr{U}. \end{aligned} \tag{6}$$

for some set $\mathscr{U}$ of admissible signals. In general the objective $\phi(\mathscr{I}(\theta))$ depends on the true value of the parameter $\theta \in \Omega$, which is unknown to the experimenter. Several approaches are commonly taken to address this problem.

The simplest approach, and the one that we will use, is to assume a nominal value $\theta_0 \in \Omega$ for the model parameters and to solve the problem

$$\begin{aligned} &\textbf{maximize}_u \;\; \phi(\mathscr{I}(\theta_0)) \\ &\textbf{subject to} \;\; u \in \mathscr{U}. \end{aligned} \tag{7}$$

about this nominal value. This is referred to by some authors as "local design" [25], though we will refer to it as "nominal design" and reserve the terms local and global to refer to local and global solutions of the optimization problem. If multiple experiments can be performed, this approach can also be extended to an iterative procedure where we start with some nominal value of the parameter vector and design an optimal experiment using this value. Data are then collected, a new value of the parameter vector is estimated from the data, and then the experiment design is performed again using the updated value of the parameter vector.

Another approach is to consider a minimax criterion, solving the worst-case optimization problem

$$\begin{aligned} \textbf{maximize}_u \;\; & \min_{\theta\in\Omega} \phi(\mathscr{I}(\theta)) \\ \textbf{subject to} \;\; & u \in \mathscr{U}. \end{aligned} \tag{8}$$

This approach ensures a minimum level of estimator performance regardless of the true value of θ.

Another alternative is to consider a prior distribution $\pi(\theta)$ on the values that the parameter vector may take. The experiment may then be designed with respect to the average criterion

$$\begin{aligned} \textbf{maximize}_u \;\; & \mathbb{E}_\pi[\phi(\mathscr{I}(\theta))] \\ \textbf{subject to} \;\; & u \in \mathscr{U}. \end{aligned} \tag{9}$$

6 Systems Theory Problems in Hyperpolarized Carbon-13 MRI

We now present three optimization problems that arise in the design of hyperpolarized carbon-13 MRI experiments and the subsequent data analysis. The first involves the design of substrate injection inputs to generate maximally informative data, a problem in which the control input enters linearly. The second involves the design of optimized flip angle sequences, again for generating maximally informative data. In contrast with Problem 1, this problem involves a nonlinear control system model, which is significantly more difficult to analyze globally. The third problem involves estimating the spatial distribution of metabolic flux parameters from the acquired data. Problem 3 completes the experimental sequence from experimental design to data acquisition to data analysis.

Problem 1: Substrate Injection Design

Data collected in MRI experiments is typically noisy due to thermal movement of electrons in the receiver coil and the object being imaged. This makes it challenging to estimate model parameters from dynamic data sets when the signal to noise ratio is small. This challenge can be addressed by designing experimental parameters with the goal of maximizing the information about unknown parameters contained in the data collected.

The first problem we consider is the optimal design of the injection input subject to constraints on the maximum injection rate and volume. This results in an dynamic optimal experiment design problem of the form discussed in Section 5. More formally, we consider the dynamic model defined in Equation (3) with an output defined in Equation (4) which is corrupted by iid additive Gaussian noise. Problem

1 is to design an injection input $u[k]$ to maximize the Fisher information about the parameter of interest k_{SP} contained in the data generated from a finite number of samples under this model. The input is constrained such that both the maximum injection rate $\|u\|_\infty$ and the maximum injection volume $\|u\|_1$ are upper bounded by some positive constant.

We first formulated this problem in [15], where it was shown that this problem can be reformulated as a nonconvex quadratic program (QP). We then developed a procedure for approximating the global solution of the QP using a semidefinite programming relaxation. This method allowed us to compute approximate solutions to realistic instances of this problem. We conjecture that all optimal solutions are of the form found in this paper: a bolus applied at the beginning of the experiment injected an the maximum rate until the volume budget is reached (Figure 5). However, this conjecture remains unproven.

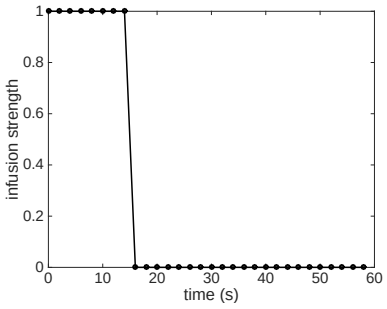


Fig. 5 Conjectured solution to a particular instance of Problem 1. The optimal input sequence $u[k]$ applies a bolus injection at the maximum allowable rate until the total input budget is reached.

Problem 2: Flip Angle Sequence Design

Similarly to the first problem, the second problem we consider involves designing experimental parameters to maximize the Fisher information about unknown rates in the model. Here we consider the problem of designing optimal RF flip angle excitation sequences.

Again we use the model defined in Equation (3) with an output defined in Equation (4) corrupted by iid noise. In Problem 2, we wish to select a sequence of flip angles $\alpha_S[k]$ and $\alpha_P[k]$ used to excite each of the chemical species. Here the choice of $\alpha_X[k]$ at each time is unconstrained. Since the flip angles enter the model in a nonlinear fashion, the resulting optimization problem is no longer a QP, so other optimization techniques must be used.

This problem is solved to local optimality under additional smoothness constraints in [16] using a nonlinear programming approach. However satisfactory methods for computing a global optimum remain to be found.

Problem 3: Constrained Parameter Mapping

The third problem involves computing maps of metabolic activity from the experimental data collected. Here we assume that we are given a statistical model for the data as well as a loss function, as described in Section 4.2. The challenge is to summarize the spatial, temporal and chemical information contained in the dynamic experimental data into a single spatial map of metabolic activity. We do so by estimating a value for the metabolic rate parameter $\theta_i = k_{SP,i}$ for each voxel i in space.

Since the objects imaged often contain spatial structure, this structure can be exploited to improve the quality of the estimated parameter maps. This can be achieved by adding regularization to the objective function that is optimized. We solve an optimization problem of the form

$$\textbf{minimize} \sum_i L(\theta_i|Y_i) + \lambda\, r(\theta)$$

where L is a loss function that depends on the data Y_i collected in each voxel i, and r is a regularization term that couples nearby voxels thereby enforcing spatial structure in the estimated maps. Possible choices of regularization used to enforce smoothness, sparsity and edge preservation include ℓ_2, ℓ_1 and total variation penalties. By including such penalties to exploit spatial correlations in the data, we have shown that better image quality can be achieved compared with independently fitting each voxel [17].

Both choices of loss function described in Section 4.2 are nonconvex. However, we have observed that despite the nonconvexity of the problem satisfactory solutions can be found using convex optimization algorithms such as ADMM [3]. Problem 3 is to better understand the convergence of this algorithm for estimating parameters in spatially-distributed dynamical system models. Why does this algorithm successfully converge to the same optimum for various initial conditions? And can we provide any formal guarantees of global convergence?

References

[1] Ardenkjær-Larsen JH, Fridlund B, Gram A, Hansson G, Hansson L, Lerche MH, Servin R, Thaning M, Golman K (2003) Increase in signal-to-noise ratio of $>$ 10,000 times in liquid-state NMR. Proceedings of the National Academy of Sciences 100(18):10,158–10,163

[2] Bernstein MA, King KF, Zhou XJ (2004) Basic pulse sequences. In: Handbook of MRI Pulse Sequences, Academic Press, pp 579 – 647

[3] Boyd S, Parikh N, Chu E, Peleato B, Eckstein J (2011) Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends in Machine Learning 3(1):1–122

[4] Brindle KM (1988) NMR methods for measuring enzyme kinetics in vivo. Progress in Nuclear Magnetic Resonance Spectroscopy 20(3):257–293

[5] Cunningham CH, Lau JY, Chen AP, Geraghty BJ, Perks WJ, Roifman I, Wright GA, Connelly KA (2016) Hyperpolarized ^{13}C metabolic MRI of the human heart: Novelty and significance. Circulation Research 119(11):1177–1182

[6] Day SE, Kettunen MI, Gallagher FA, Hu DE, Lerche M, Wolber J, Golman K, Ardenkjaer-Larsen JH, Brindle KM (2007) Detecting tumor response to treatment using hyperpolarized ^{13}C magnetic resonance imaging and spectroscopy. Nature Medicine 13(11):1382–1387

[7] Gatenby RA, Gillies RJ (2004) Why do cancers have high aerobic glycolysis? Nat Rev Cancer 4(11):891–899

[8] Gevers M, Bombois X, Hildebrand R, Solari G (2011) Optimal experiment design for open and closed-loop system identification. Communications in Information and Systems 11(3):197–224

[9] Goodwin G, Payne R (1977) Dynamic System Identification: Experiment Design and Data Analysis. Academic Press

[10] Gudbjartsson H, Patz S (1995) The Rician distribution of noisy MRI data. Magnetic Resonance in Medicine 34(6):910–914

[11] Harrison C, Yang C, Jindal A, Deberardinis R, Hooshyar M, Merritt M, Dean Sherry A, Malloy C (2012) Comparison of kinetic models for analysis of pyruvate-to-lactate exchange by hyperpolarized ^{13}C NMR. NMR in Biomedicine 25(11):1286–1294

[12] Kazan SM, Reynolds S, Kennerley A, Wholey E, Bluff JE, Berwick J, Cunningham VJ, Paley MN, Tozer GM (2013) Kinetic modeling of hyperpolarized ^{13}C pyruvate metabolism in tumors using a measured arterial input function. Magnetic Resonance in Medicine 70(4):943–953

[13] Larson PE, Kerr AB, Chen AP, Lustig MS, Zierhut ML, Hu S, Cunningham CH, Pauly JM, Kurhanewicz J, Vigneron DB (2008) Multiband excitation pulses for hyperpolarized ^{13}C dynamic chemical-shift imaging. Journal of Magnetic Resonance 194(1):121 – 127

[14] Ljung L (1999) System Identification: Theory for the User. Pearson Education

[15] Maidens J, Arcak M (2016) Semidefinite relaxations in optimal experiment design with application to substrate injection for hyperpolarized MRI. In: Proceedings of the American Control Conference (ACC), pp 2023–2028

[16] Maidens J, Gordon JW, Arcak M, Larson PEZ (2016) Optimizing flip angles for metabolic rate estimation in hyperpolarized carbon-13 MRI. IEEE Transactions on Medical Imaging 35:2403–2412

[17] Maidens J, Gordon JW, Arcak M, Chen HY, Park I, Criekinge MV, Milshteyn E, Bok R, Aggarwal R, Ferrone M, Slater JB, Kurhanewicz J, Vigneron DB, Larson PEZ (2017) Spatio-temporally constrained reconstruction for hyperpolarized carbon-13 MRI using kinetic models. In: Proceedings of the ISMRM Annual Meeting (under review), URL http://submissions.mirasmart.com/ISMRM2017/ViewSubmissionPublic.aspx?sei=GH0eaFQTF

[18] Merritt ME, Harrison C, Storey C, Jeffrey FM, Sherry AD, Malloy CR (2007) Hyperpolarized ^{13}C allows a direct measure of flux through a single enzyme-catalyzed step by NMR. Proceedings of the National Academy of Sciences 104(50):19,773–19,777

[19] Nelson SJ, Kurhanewicz J, Vigneron DB, Larson PEZ, Harzstark AL, Ferrone M, van Criekinge M, Chang JW, Bok R, Park I, Reed G, Carvajal L, Small EJ, Munster P, Weinberg VK, Ardenkjaer-Larsen JH, Chen AP, Hurd RE, Odegardstuen LI, Robb FJ, Tropp J, Murray JA (2013) Metabolic imaging of patients with prostate cancer using hyperpolarized [1-^{13}C]pyruvate. Science Translational Medicine 5(198):198ra108

[20] Nishimura DG (2010) Principles of Magnetic Resonance Imaging. Lulu

[21] Pukelsheim F (1993) Optimal design of experiments. Probability and mathematical statistics, Wiley

[22] Schroeder MA, Lau AZ, Chen AP, Gu Y, Nagendran J, Barry J, Hu X, Dyck JR, Tyler DJ, Clarke K, Connelly KA, Wright GA, Cunningham CH (2013) Hyperpolarized ^{13}C magnetic resonance reveals early- and late-onset changes to in vivo pyruvate metabolism in the failing heart. European Journal of Heart Failure 15(2):130–140

[23] Søgaard LV, Schilling F, Janich MA, Menzel MI, Ardenkjaer-Larsen JH (2014) In vivo measurement of apparent diffusion coefficients of hyperpolarized ^{13}C-labeled metabolites. NMR in Biomedicine 27(5):561–569

[24] Stanley WC, Recchia FA, Lopaschuk GD (2005) Myocardial substrate metabolism in the normal and failing heart. Physiological Reviews 85(3):1093–1129

[25] Walter É, Pronzato L (1997) Identification of parametric models from experimental data. Communications and control engineering, Springer