

Imaging Applications with Variable Projections

Victor Pereyra¹, Godela Scherer²

¹Smart Fields, Stanford University, Stanford, CA, USA

²Department of Mathematics, University of Reading, Reading, GB

Email: vpereyra@yahoo.com, godela@gmail.com

How to cite this paper: Pereyra, V. and Scherer, G. (2019) Imaging Applications with Variable Projections. *American Journal of Computational Mathematics*, 9, 261-281. <https://doi.org/10.4236/ajcm.2019.94020>

Received: September 26, 2019

Accepted: December 13, 2019

Published: December 16, 2019

Copyright © 2019 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

We review two important fields of application for the method of Variable Projections (VARPRO): Seismic Prospecting and Medical Imaging. We cover the period since 2002 when a first review was published. Variable Projections is a method for the solution of nonlinear least squares problems where the model is a linear combination of nonlinear functions. Its success is based on the fact that the reduced functional, where the linear parameters are eliminated converges always faster than if the same method of solution is applied to the original problem. More importantly, many of these problems are very hard to solve in their original format and VARPRO has shown its value in many different applications.

Keywords

Medical Imaging, Geophysics, Variable Projections, Nonlinear Least Squares

1. Introduction

We consider nonlinear data fitting problems that have as their underlying model a linear combination of nonlinear functions. Models of this type are very common inasmuch as many inverse problems can be viewed as nonlinear data fitting problems [1].

Given a set of observations $\{y_i\}$, a separable nonlinear least squares problem is defined in [1] where the i th component of the residual vector is written as

$$r_i(\mathbf{a}, \alpha) = y_i - \sum_{j=1}^n a_j \phi_j(\alpha; t_i).$$

Here the t_i are independent variables associated with the observations y_i , while the a_j , and the k -dimensional vector α contains the parameters to be determined by minimizing the functional $\|\mathbf{r}(\mathbf{a}, \alpha)\|_2^2$, where $\mathbf{r}(\mathbf{a}, \alpha) = \sum_{i=1}^m r_i^2(\mathbf{a}, \alpha)$, and $\|\cdot\|_2$ stands for the l_2 vector norm. We can write

this functional using matrix notation as

$$\|\mathbf{r}(\mathbf{a}, \alpha)\|_2^2 = \|\mathbf{y} - \Phi(\alpha)\mathbf{a}\|_2^2,$$

where the columns of the matrix $\Phi(\alpha)$ correspond to the nonlinear functions $\phi_j(\alpha; t_i)$ of the k parameters α evaluated at all the t_i values, and the vectors \mathbf{a} and \mathbf{y} represent the linear parameters and the observations respectively.

Now it is easy to see that if we knew the nonlinear parameters α , then the linear parameters \mathbf{a} could be obtained by solving the linear least squares problem:

$$\mathbf{a} = \Phi(\alpha)^+ \mathbf{y}, \quad (1)$$

which stands for the minimum-norm solution of the linear least squares problem for fixed α , where $\Phi(\alpha)^+$ is the Moore-Penrose generalized inverse of $\Phi(\alpha)$. By replacing this \mathbf{a} in the original functional the minimization problem takes the form

$$\min_{\alpha} \frac{1}{2} \left\| (\mathbf{I} - \Phi(\alpha)\Phi(\alpha)^+) \mathbf{y} \right\|_2^2, \quad (2)$$

where the linear parameters have been eliminated.

We define $\mathbf{r}_2(\alpha) = (\mathbf{I} - \Phi(\alpha)\Phi(\alpha)^+) \mathbf{y}$, which will be called the Variable Projection (VP) of \mathbf{y} . Its name stems from the fact that the matrix in parentheses is the projector on the orthogonal complement of the column space of $\Phi(\alpha)$, which we will denote in what follows by $P_{\Phi(\alpha)}^{\perp}$. We will also refer to $\frac{1}{2} \|\mathbf{r}_2(\alpha)\|_2^2$ as the Variable Projection functional.

This is a more powerful paradigm than the simple idea of alternating between minimization of the two sets of variables (such as the NIPALS algorithm of Wold and Lyttkens [2]), which can be proven theoretically and practically not to result, in general, in the same enhanced performance.

In [3] you can find a full survey of applications of VARPRO in many different fields. The purpose of the current excerpt is to tell the story of these developments and applications since the first review in 2003 [4] in two important fields: geophysics and medical imaging. The common thread for the majority of the papers is the use of variable projections for separable models in a least squares context, while a small percentage uses only the results on the derivatives of pseudoinverses and projectors.

2. Geophysics, Petroleum Engineering

This is an area of particular interest to one of the authors, who has worked many years in exploration geophysics, mainly on geological modeling, seismic ray tracing and inverse problems [5] [6] [7]. In the past few years VARPRO has been discovered as one of the keys for solving the fundamental earth imaging problem using full waveform inversion [8]. This is a notoriously expensive procedure (it requires many solutions of the wave equation in 3D) and it leads also to multi-

modal optimization problems, where local optimization algorithms have difficulties in avoiding sub-optimal minima. It turns out that several different applications within this area can be stated as separable problems, making them amenable to the use of VARPRO, which tends to deliver much better behaved optimization problems in the reduced space of the nonlinear variables. It turns out that Bill Symes had foreseen this early on with his related method of differential semblance [9].

For instance, in [10] the authors tackle the well-known global convergence issue associated to any full waveform inversion (FWI) approach by solving an extended-image space least-squares migration problem to remove any local minima present in the FWI objective function. They discuss the connection between the reflectivity and migration velocity inversion and show the importance of combining the two problems using one objective function. Moreover, they show the full separability of the two inverse problems by using the variable projection method.

Furthermore, in [11], the same authors indicate that the main issue inherent to full waveform inversion (FWI) is its inability to correctly recover the Earth's subsurface seismic parameters from inaccurate starting models. This behavior is due to the presence of local minima in the FWI objective function. To overcome this problem, they propose a new objective function in which they modify the nonlinear modeling operator of the FWI problem by adding a correcting term that ensures phase matching between predicted and observed data. This additional term is computed by demigrating an extended model variable and its contribution is gradually removed during the optimization process while ensuring convergence to the true solution. Since the proposed objective function is quadratic with respect to the extended model variable, they make use of the variable projection method. They refer to this technique as full waveform inversion by model extension (FWIME) and illustrate its potential on two synthetic examples for which FWI fails to retrieve the correct solution.

In [12] the authors consider planar waves events recorded in a seismic array that can be represented as lines in the Fourier domain. However, in the real world, seismic events usually have curvature or amplitude variability, which means that their Fourier transforms are no longer strictly linear but rather occupy conic regions of the Fourier domain that are narrow at low frequencies but broaden at high frequencies, where the effect of curvature becomes more pronounced. One can consider these regions as localized "signal cones". In this work, the authors consider a space-time variable signal cone to model the seismic data. The variability of the signal cone is obtained through scaling, slanting, and translation of the kernel for cone-limited (C-limited) functions (functions whose Fourier transform lives within a cone) or C-Gaussian function (a multivariate function whose Fourier transform decays exponentially with respect to slowness and frequency), which constitutes their dictionary. The authors find a discrete number of scaling, slanting, and translation parameters from a continuum by optimally

matching the data. This is a non-linear optimization problem, which is solved by a fixed-point method that utilizes a variable projection method with l_1 constraints on the linear parameters and bound constraints on the non-linear parameters.

3. Medical and Biological

Many applications to imaging and spectroscopy in biological and medical systems can be found in this section. Again, identification problems are common. A software tool which appears in several articles is the open source package, TIMP [13]. TIMP is a problem solving environment for fitting separable nonlinear models to measurements arising in Physics and Chemistry written in R. It has been extensively tested for time-resolved spectroscopy and FLIM-FRET data (FLIM: Fluorescent Lifetime Imaging Microscopy; FRET: Forster resonance energy transfer).

In [14], a tool for the processing and automatic quality grading in the fish industry is developed based on diffuse reflectance imaging and the subsequent unmixing of the absorption spectra using a constrained least squares model to detect hemoglobin concentration. It is common for the absorption lines to have a Gaussian or Lorentzian distribution shape, so VARPRO is used in the decomposition of the spectra.

In [15], a dynamical model, described by the equations of motion of a spring-mass-damper system, was set up to estimate the impedance (force-position) for the human elbow. Then, a system identification technique based on prediction-error minimization (PEM) was developed of this non-parametric time-domain model augmented with a parametric noise model. VARPRO is used in the approximation.

The paper [16] describes a methodology to determine the Forster resonance energy transfer in live cells as measured using fluorescence imaging microscopy. The parameters of the nonlinear quantitative model can be computed using a variable projection type of algorithm. In [17] VARPRO is used to fit NMR spectra to depict small and highly oblique nerves of the lumbosacral plexus. The aim is to use the methodology for diagnostics of pathologies.

The articles [18] [19] [20] [21] are all about compartmental analysis, a mathematical modeling tool that originated in pharmacokinetics and is now used widely in medical and biological applications. The compartment model is formed by separate homogeneous compounds, the so called compartments. These may represent a certain space (blood, brain, etc) or a compound in a specific form (for example in a different chemical binding). The important point is that each compartment is assumed to be homogeneous. They interact with each other by exchanging material and for most medical and biological systems this exchange is assumed to obey a linear differential equation with constant transfer coefficients, so that the system behavior in time is modeled by a linear system of differential equations. Although the equations are linear the solution is not. It can

be fitted by a linear combination of nonlinear functions (basis functions), and the parameters can then be obtained through a variable projection algorithm.

For medical imaging applications, a radioligand (a radioactive biochemical substance, in particular a radiolabelled substance) or tracer is introduced, usually intravenously. The transport and the binding rate of the tracer are assumed to depend linearly of the difference of the tracer concentration between two compartments, so defining the system of ODE in the tracer concentrations. Often the data measured are sums of the concentrations.

3.1. Spectroscopy

After obtaining expressions for the identification of relaxation times associated with kinetic fluorescence decay and those associated with the dynamic evolution of fluorophores (chemical compounds that can re-emit light upon light excitation), the author of [22] suggests the use of variable projection algorithms in the evaluation of photochemical bioimaging, when the fluorophores are used as the probe molecules. In these studies a multi-exponential decay surface can be ascribed to each pixel, where the fluorescence decay times and the corresponding emission or excitation wavelength dependent amplitudes can be recovered by the VARPRO algorithm.

The first part of [23] is a survey of the adaptation of the variable projection algorithm to the case of matrix data and of constraints on the linear parameters. This form of least squares approximation to fit linear combination of nonlinear functions is common in the applications considered in the paper, namely, spectroscopy, microscopy and mass spectrometry. The authors emphasize the importance of forming the residual vector in a particular manner to avoid storing and operating with tensor products, and describe a way to compute the covariance for the linear parameters using less memory resources.

The application to spectroscopy involves determining the kinetics of a compartmental model of a photo-system for conversion of photons into chemical energy, using the time-resolved fluorescence measurements. This is a separable nonlinear least squares problem with matrix data Ψ and unknown kinetic model $C(z)E^T$, where z are the non-linear parameters and the spectra is represented by the (non-negative) linear coefficient matrix E^T : $\Psi \simeq C(z)E^T$. It is solved using the software package TIMP.

In the application to microscopy the authors consider the detection of a protein-protein interaction by the simultaneous analysis of multiple FLIM images. FLIM counts the photons detected at several time intervals and over many locations. The data analysis gives rise to a separable nonlinear problem with the same nonlinear parameters but different linear ones, and multiple right hand sides (also termed global analysis problem) and with a matrix data:

$$\min_{z \in \mathbb{R}^n} \left\| \Psi - C(z)E^T \right\|_2.$$

The kinetic processes of the fluorescence decays are exponential and are represented in the columns of $C(z)$ convolved with an instrumental response

function (IRF). The rows of E are the amplitudes corresponding to each kinetic process.

An additional point to consider here is the distribution of errors in the FLIM data, a count of number of photons detected at a given position and time. If the count is large, then to assume that the data errors have a Gaussian distribution with mean 0 and variance σ^2 is valid and the least squares criteria acceptable. For smaller counts, the errors have a Poisson distribution and the least squares estimates are not very good. One possible correction is to weight the data points $\Psi[i, j]$ with the factor $\frac{1}{\sqrt{\Psi[i, j]}}$.

The third area of application considered is mass spectrometry (MS), in conjunction with gas (GC/MS) or liquid (LC/MS) chromatography. These are analytical chemistry techniques that combine the physical separation of the different molecules in the chromatography column with the separation of the ions according to their mass-to-charge ratio in the MS step. In the GC or LC step the sample molecules passing through a column, elute or come off at different times according to their affinity with the chemical in the column. In the MS step the sample is ionized and the resulting ions separated according to mass-to-charge ratio by deflection due to an electric or magnetic field.

The GC/MS or LC/MS measurements of a sample can be modeled by $\Psi \approx CE^T$. Here C are the elution profiles and E the mass spectra resolved with respect to the mass-to-charge ratio. When several samples of the same compound are considered, the elution profiles are different but the mass spectra is considered the same. Often C can be well represented by choosing columns of exponentially modified Gaussians with 4 parameters, width, location, decay rate and amplitude. The problems can be solved with variable projections algorithms.

The performance of three gradient type algorithms: alternating least squares, VARPRO and the Kaufman simplification are compared in the case of a multi-exponential model of a photo-physical system [24]. Corroborating results of other authors, alternating least squares where the linear and nonlinear parameters are alternatively fixed and the problem minimized over the complementary set, is found to be the least efficient. The Fisher information matrix computation enables the authors to determine a lower bound for the covariance estimate of the precision of the nonlinear parameters, and conclude that in the present case the Kaufman simplification is the most cost efficient.

MEG or magneto-encephalogram is an imaging tool that can measure changes in the neural activity on a very small time scale (of the order of milliseconds). In [25] the authors compare several algorithms that solve the inverse problem: given magnetic field values at a number of measurement points, reconstruct the sources, *i.e.*, compute the location and moment parameters of the set of dipoles whose fields best approximates the data in the least squares sense. This is a large nonlinear optimization problem with a complex objective function and many local minima. However, the model is a separable function, *i.e.*, a linear combina-

tion, with coefficients depending on the dipole moment parameters, of nonlinear functions of the location parameters, amenable to variable projection. The authors compare several gradient-free algorithms for the reduced nonlinear problem: simulated annealing, genetic algorithms and a tabu search, and conclude that for the given problem, the best algorithm is a local genetic algorithm.

In 2D spectroscopy, contrary to conventional 1D, the third order optic signal at given population times is a function of two frequencies. In [26] the authors propose a method to analyze the 2D signals using a global analysis method based on the variable projection algorithm. To reduce the dimensionality, the 3D complex valued data arrays containing the signal as a function of the excitation and emission frequency and time, are reorganized into a matrix Y , each of its columns representing the evolution in time of a specific pair of frequencies. A multi-exponential model is then defined by $M = EA$, where each column of E contains a complex exponential function $E_{nk} = e^{b_n t_k}$ and the amplitudes are found in the matrix A .

The data analysis can also be extended to fit rephasing and non-rephasing data simultaneously by building appending blocks of data to form the matrix Y . As this problem is computationally challenging due to its size, the authors suggest using in the minimization step a subsample of the data in the frequency dimensions, reporting that even using only 5% of the data gives satisfactory results.

3.2. Tomography

The next articles concern tomography: the imaging by sectioning of objects, or more generally, 2 and 3D imaging of the inside of objects using different signals and the reconstruction techniques employed to recover the information about these objects from the imaging data. For some of the medical imaging methods, a radiation source is used and the data acquired are multiple 2D images or projections taken from different angles. In transmission tomography, such as computerized tomography (CT), the radiation source is outside the object: a rotating X-ray tube generates X-rays that travel through the object. The detector measures the line integral of the beam intensity and the quantity to be reconstructed is the attenuation coefficient of the medium. In emission tomography (ECT), like PET and SPECT, the radiation source is inside: a radioactive tracer is introduced into the body. In the case of SPECT, the tracers emit gamma radiation that is measured. PET tracers emit positrons that when colliding with nearby electrons produce two gamma photons. The gamma radiation is detected by a rotating gamma camera that acquires multiple 2D images from different angles. The quantity to be reconstructed in both cases is the distribution and concentration of the radioactive tracer in the different parts of the body. Finally, photo-acoustic tomography is based on the PA effect, the formation of sound waves following light absorption. The PA signals are acquired at several locations around the object using a transducer array, while the goal of the photo-acoustic imaging reconstruction is to retrieve the local pressure rise inside the tissue.

Article [27] evaluates reconstruction techniques applied to data obtained from CT scans. Incidentally, these techniques could also be used in other applications, for example seismic tomography. For medical tomography, there are two different reconstruction procedures: the direct inversion methods that reverse the Radon transform (the line integral of the beam intensity) by, for example, filtered back projection, and the algebraic reconstruction methods, more adequate when the data have not been regularly sampled.

The data produced by a CT scan of an unknown object $\mathbf{u} \in \mathbb{R}^3$ are a finite number of samples $p_i(s_i, \eta_i)$ of the intensity integral of a X-ray defined by its so called acquisition parameters, origin \mathbf{s} and direction $\boldsymbol{\eta}$:

$$p(\mathbf{s}, \boldsymbol{\eta}) = \int_{\mathbb{R}} \mathbf{u}(\mathbf{s} + t\boldsymbol{\eta}) dt.$$

The reconstruction problem is then reduced to solving the linear rectangular system $W\mathbf{u} = \mathbf{p}$, where W is the $m \times n$ projection matrix, depending on the parameters (s_i, t_i) . This can be computed by least squares methods. In practice though, often the geometry of the imaging system, *i.e.*, the acquisition parameters $(\mathbf{s}, \boldsymbol{\eta})$ are not known accurately, for example due to faulty calibration, and this produces alignment errors. One option would be to consider the problem as a Total Least Squares problem, *i.e.*, a linear problem where the errors are not only restricted to the observations \mathbf{p} but also the matrix W is not known exactly.

Instead the authors suggest a method belonging to the class of projection matching methods. Basically they solve the nonlinear model,

$$W(\mathbf{a})\mathbf{u} = \mathbf{p},$$

where the unknown $\mathbf{a} \in \mathbb{R}^l$ contains the parametrization of possible rigid motions, three shifts and three rotations, for each projection image. The authors note that this is a possible severely ill-conditioned problem and since its parameters separate, into the nonlinear \mathbf{a} and the linear \mathbf{u} , they suggest an alternating optimization procedure. After solving analytically for the linear parameters \mathbf{u} , *i.e.*, eliminating the reconstruction part of the problem, they describe several gradient-descent algorithms to solve the resulting nonlinear optimization problem, and state their convergence behavior, considering when approximate derivatives are used and in case of constraints on the alignment parameters.

The articles [27] [28] [29] consider the case of SPECT tomography. SPECT imaging methods are governed by the photon transport equation and the reconstruction involves the attenuated Radon transform. Two unknown quantities must be estimated simultaneously: the radioactive emission source (the distribution of the radionuclides) $f(x, E)$ and the photon attenuation coefficient of the tissue $\mu(x)$, from the acquired data $p(s, \theta)$ on the line defined by (s, θ) . The variables x and E are position and energy. The attenuated Radon transform is:

$$R_{\mu} f(s, \theta) = \int_{-\infty}^{\infty} f(s\theta + t\theta^{\perp}) \exp\left(-\int_t^{\infty} \mu(s\theta + \tau\theta^{\perp}) d\tau\right) dt = p(s, \theta).$$

Under the assumption that the emission data follow a Gaussian distribution, a nonlinear least squares problem can be defined including a regularization term to avoid irregular distributions of (f, μ) ,

$$\min_{f, \mu} \|R_\mu f - p\|_2^2 + \alpha I[f, \mu].$$

Bronnikov exploits the fact that the variables f and μ are respectively linear and nonlinear to design a variable projection algorithm along the lines of VARPRO that works well. On the other hand, Gourion *et al.* use nonlinear optimization methods directly. They also consider that a Poisson distribution is more appropriate for the data.

Finally, articles [30] and [31] study the newest field in biomedical imaging: photo-acoustic computed tomography (PACT). In this imaging technique, short laser pulses are directed at the object. The absorption of the optical energy produces local heating causing expansion of the tissue and consequent photo-acoustic wave-fields that can be measured outside the body using piezoelectric ultrasonic transducers. The signals that they receive are convolved with their acoustic-electric impulse response (EIR). In [31] the authors choose to incorporate the effect of EIR into the reconstruction. This results in an inverse model with separable linear and nonlinear parameters:

$$\mathbf{u} = \mathbf{H}(\mathbf{h})\boldsymbol{\theta}.$$

Here, the matrix \mathbf{H} contains the approximations to the function $\mathbf{A}(\mathbf{r}) = \sum \theta_j \phi_j(\mathbf{r})$ which represents the absorbed optical density of the object under study. \mathbf{H} also depends on the sampled EIR values represented by \mathbf{h} . The image reconstruction is now formulated as an optimization problem in \mathbf{h} and $\boldsymbol{\theta}$ with regularization terms:

$$\min \|\mathbf{u} - \mathbf{H}(\mathbf{h})\boldsymbol{\theta}\|^2 + \lambda R_1(\boldsymbol{\theta}) + \alpha R_2(\mathbf{h}).$$

To estimate the parameters, the authors alternate the minimization between the linear and the nonlinear parameters.

3.3. Nuclear Magnetic Resonance (NMR) Spectroscopy and Imaging

This is one of the most important applications of VARPRO as explained in [4] and we keep seen strong use as many more specific and new results come out for practical applications.

The physical phenomenon associated with NMR involves a sample that is placed in a magnetic field and is irradiated with a radio-frequency (RF) pulse of a determined resonant frequency. The nuclei in the sample emit a signal that can be recorded and interpreted. NMR can be used to identify what molecules are present because the resonant frequency (Larmor frequency) depends among other factors, on the molecules. In imaging applications, hydrogen is typically the molecule of interest.

In more detail, the $\frac{1}{2}$ -protons in the hydrogen nuclei have two eigenstates

($m = \pm \frac{1}{2}$), which in the absence of a magnetic field have the same energy. After a strong external static magnetic field is applied, most of the protons fall into the lower of the two states (for most isotopes used in NMR: $m = +1/2$). When the RF pulse is applied some protons are excited back into higher energy state ($m = -1/2$) and when they decay back to the lower state an electromagnetic radiation is emitted that can be measured. In the literature, T_1 or relaxation time, defines the equilibrium recovery time needed by the sample after the RF excitations. To produce images of the interior of the sample, *i.e.*, to localize the NMR signals, spatial variations in the magnetic field strength across the sample are generated, the gradient fields.

Already in our 2003 review article we listed a number of articles with applications in which the data obtained through NMR was evaluated using the variable projection algorithm. An important area of application is still *in vivo* MR spectroscopy. In [32] contrast enhanced MR imaging is used to obtain a time-series of the contrast concentration in the blood plasma and the extra-vascular, extra-cellular space (EES) of prostate tissue, both cancerous and non-cancerous. The perfusion model, *i.e.*, the pharmacokinetic model of the passage of fluid between the capillaries and the capillary bed (EES) used is a two-compartment model (Tofts), with the contrast concentration in each voxel given (after discretization) by $\mathbf{p}_i = \mathbf{a}_i v_p + \mathbf{A} \mathbf{h}_i (k_{ep}) K^{trans}$. Here, the linear parameters are the transfer constant K^{trans} and the vascular fraction v_p , while the nonlinear parameter is the rate constant k_{ep} . The arterial input function \mathbf{a} is a known function of time. This model can be fitted voxel-wise to the MR image data in the least squares sense.

The authors have compared two different algorithms to obtain the perfusion parameters: using Levenberg-Marquardt to estimate all the parameters and applying the variable projection separation of variables idea to define a nonlinear LSQ problem in the parameter k_p , followed by the solution of the linear LSQ problem in the two linear parameters. To solve the nonlinear optimization problem in one variable they use Golden Section Search. Their accuracy and noise sensitivity comparisons included numerical simulations using parameter values in the range reported both for normal and malignant tissue to generate simulated MR signals that were later converted back into concentration functions after the determination of the parameters by the two optimization algorithms mentioned above. The results were comparable, but the VP based technique was three times as fast as LM for all the variables. More important for the medical application were the clinical trials with 20 patients. Here the LM failed to converge in approximately 15% of the tissues, including normal and cancerous, whereas the VP technique converged in 100% of cases.

In article [33], global and target analysis of the time-resolved spectra obtained in bioenergetic applications are reviewed. Spectroscopy is used here as a tool to investigate the dynamic properties of complex biological systems through the application of a short pulse of high energy that produces reactions like absorp-

tion or fluorescence. Often the data collected depend on two parameters, one variable is the wave length and the second is the time after excitation.

To analyze the measurements and estimate the physicochemical parameters, both the kinetics (the compartmental model determined by transitions between the states) and the spectra must be modeled. One assumption is that the system is separable, meaning that the spectroscopic data of a complex compound is the superposition of the spectroscopic properties of the components weighted by their concentration. The simplest unidirectional kinetic model is used in global analysis, when the response is considered to be a sum of a few (2 - 4) exponential decays convolved with the instrument response function (IRF), *i.e.*, the data at different wavelengths are all approximated with a single set of exponentials. Target analysis is used when the problem requires a more complicated kinetic model involving for example forward and reversible reactions, independent decays, etc.

In both cases the parameters in the final model (both the kinetic and the spectra) must be fitted. Assuming normally distributed noise, a nonlinear least squares fit gives the maximum likelihood estimator. The nonlinear least squares approximation leads to a separable nonlinear problem that is solved using the variable projection algorithm. Results for the application to the study of ultrafast dynamics of the photoactive yellow protein are presented.

In [34] a complementary approach to global analysis is described, the so called lifetime density analysis (LDA). It is a technique used in ultrafast (femto- and pico-second) spectroscopy, when investigating energy and charge transfer in complex photosystems. Instead of using a small set of exponential decays to fit the data, one assumes that they are better represented by the integral of a continuous distribution of decays. This integral is discretized using a sum of a large (~100) lifetimes distributed along the time of the experiment. The resulting approximation problem in matrix notation is:

$$\min_{x,\tau} \|D(\tau)x - A_n\|_2^2.$$

Here, A_n is the $(m \times n)$ measurement matrix at the m time delays and n wavelengths, D is the matrix of the IRF convoluted with the exponential delays and x contains the amplitudes associated with the lifetimes. Again, it is a separable nonlinear least squares problem and it is solved by the variable projection algorithm.

Contrary to the global analysis strategy, the LDA approach has a large number of parameters and overfitting must be considered. The authors suggest the use of the well-known regularization techniques, truncated SVD and Tikhonov regularization. They also discuss briefly several other methods which they have included into the software, an open source Python package with GUI, that can be downloaded from the web.

Article [35] describes a technique to obtain more accurate estimates of spectral parameters when evaluating MR spectroscopic imaging (MRSI). One of the difficulties in spectral quantification is that the problem is not well conditioned.

A way to avoid the large uncertainties is to incorporate the spatial smoothness information contained in the tissue properties. So, instead of computing the spectral parameters for each voxel independently, the idea is to work with a joint formulation:

$$(\hat{\mathbf{a}}, \hat{\boldsymbol{\theta}}) = \arg \min_{\mathbf{a}, \boldsymbol{\theta}} \|\mathbf{d} - \mathbf{K}(\boldsymbol{\theta})\mathbf{a}\|_2^2 + \mathbf{R}(\mathbf{a}, \boldsymbol{\theta}). \quad (3)$$

Here $\mathbf{a}, \boldsymbol{\theta}$ are the linear and nonlinear parameters, \mathbf{d} the data, and $\mathbf{R}(\mathbf{a}, \boldsymbol{\theta})$ is a two-term penalty function added for regularization:

$$\mathbf{R}(\mathbf{a}, \boldsymbol{\theta}) = \lambda \|W_{\boldsymbol{\theta}}\|_2^2 + \eta \|\mathcal{W}_{\mathbf{a}}\|_1.$$

The terms in $W_{\boldsymbol{\theta}}$ and $\mathcal{W}_{\mathbf{a}}$ are designed to impose smoothness of all the parameters across the voxels. The proposed algorithm reformulates the optimization problem (3) as two consecutive subproblems, the first is the nonlinear least squares problem (3) but only with the l_2 penalty term of \mathbf{R} , while the second subproblem is again equation (3), now only with the l_1 penalty term of \mathbf{R} . The first subproblem is solved using the variable projection strategy to compute an updated value for the nonlinear parameters $\boldsymbol{\theta}$, which is then fed into the second subproblem. This is then a linear least squares problem in the parameters \mathbf{a} with an l_1 -norm regularization term and is solved by an alternating direction method of multipliers.

The technique was tested using several data sets, both from simulated and from *in vivo* experiments, and its performance was compared with QUEST, another method. The conclusion is that the accuracy is considerably improved, which was also confirmed by a Cramer-Rao bound analysis. The drawback is the cost of the computations, $\mathcal{O}(P^2N^2)$, with P the number of voxels and N the parameters for each voxel. This compares unfavorably with a voxel by voxel approach, where the cost is $\mathcal{O}(PN^2)$.

For long *echo-time* MRS signal modeling, VARPRO is used to fit a sum of complex damped exponentials. In [36] and [37] data quantification of metabolites in the case of *short echo-time* is considered. The *in vivo* short *echo-time* MRS are richer and therefore a more efficient solution is to create first a database of *in vitro* spectra measurements v_k of the individual metabolites. The *in vivo* signals $y(t)$ are then fitted with a combination of these v_k , corrected by parameters α_k , ς_k , and η_k to be determined in order to allow for specific scans. A baseline term $b(t)$ that globally represents the signals of other non-dominant, non-specified metabolites possible present, may also be added. This term is approximated by splines. The approximation problem considered is a non-linear least squares problem with a regularizing term:

$$\min_{\mathbf{a}, \mathbf{c}, \boldsymbol{\varsigma}, \boldsymbol{\eta}} \sum_{t=t_0}^{t_m} \left| y(t) - \sum_{k=1}^K \alpha_k (\varsigma_k)^t (\eta_k)^{t^2} v_k(t) - (\mathbf{Ac})(t) \right|^2 + \lambda^2 \mathbf{c}^H D^H D \mathbf{c}.$$

Here, the matrix D is a discrete differential operator acting as a regularization matrix. The parameters α_k have the form $\alpha_k = a_k \exp(i\phi_k)$ with a_k the real amplitudes and ϕ_k the phase shifts. The ς_k and η_k are nonlinear in the

damping correction parameters. The coefficients of the spline approximation of the baseline \mathbf{c} are always linear. If the amplitude parameters α_k are linear, the problem can be computed with VARPRO for complex problems. This is also the case even when constraints on the nonlinear parameters are imposed.

However, if there are constraints on the linear parameters α_k there is a difficulty because it is not possible, as is done in the variable projection technique, to write a closed-form solution for the linear problem part. One common occurrence for the approximation in MRS data quantification is that it requires equal phase corrections ϕ_k and non-negative real amplitudes a_k .

The algorithm described in the articles approaches this case by optimizing the nonlinear problem with the nonlinear parameters augmented by a common phase correction ϕ_0 . At each iteration of the nonlinear optimization an approximate value for the linear parameters is computed from the constrained linear least squares problem with the fixed nonlinear parameters of the last iteration step.

It is shown in [37], by numerical tests, that this algorithm is more efficient and more accurate than a “full” LS solution, where the linear and nonlinear parameters are optimized together. An open source software package AQSES based on the above techniques is described and tested. See [36]. In article [38] this software is further improved by taking into account the effect of inhomogeneities in the applied magnetic field and heterogeneities of the tissue. These cause distortions of the NMR signals, for example a broadening of the frequency-domain line with a consequent possible spectral overlap, thereby impeding a correct metabolite quantification.

The components of an NMR signal in the time-domain are the resonance frequencies multiplied by the natural damping function (Lorentzian, Gaussian or Voigt), in turn multiplied by an instrumental broadening function and all with an added noise. The line-shape (actually damping function because it is done in the time-domain) correction algorithm described in [38] is an iterative procedure: A first step is a nonlinear least squares approximation via VARPRO of the signals $\mathbf{y}(t)$ with a model $\bar{\mathbf{y}}(t)$, where line-shape distortions and baseline are ignored:

$$\bar{\mathbf{y}}(t) = \sum_{k=1}^K \alpha_k \exp(-d_k t - g_k t^2 + 2\pi i f_k t) v_k(t).$$

Using the spectral parameters thus obtained, an undisturbed signal $\hat{\mathbf{y}}(t)$ without the damping part is constructed. Then, from the quotient $\frac{\mathbf{y}(t)}{\hat{\mathbf{y}}(t)}$ a correction of the damping function $g(t)$ is constructed. After smoothing $g(t)$, eliminating any numerical instability and noise, a nonlinear least squares approximations step via VARPRO is again performed of $[\hat{\mathbf{y}}(t) \cdot g(t)]_{\text{denoised}}$. This process is continued until convergence of the spectral parameters or the damping function. The authors have validated the efficiency of this technique in improving the quantitation results using Monte Carlo simulations.

The articles [39] [40] [41] consider how to estimate the T_1 relaxation time parameter using the variable projection technique. The T_1 parameter is an intrinsic magnetic property of tissue and of importance in many clinical applications. For *in vivo* determination of T_1 , multiple datasets of signal intensity at different timings are obtained. The signal intensity can be modeled by the rational function:

$$S(t_i, A, B, C, T_1) = \frac{A + B \exp\left(-\frac{t_i}{T_1}\right)}{1 - C \exp\left(-\frac{t_i}{T_1}\right)}, \quad i = 1, \dots, N.$$

If the radio-frequency pulse flip angles are almost perfect (Ernst angle), then the parameter C in the above rational model is close to zero and the model is a very simple one where the linear parameters can easily be separated from the nonlinear one. This model is used in [39] and [41]. The rational model is considered in [40] and the parameters are again determined using VARPRO.

The two articles [42] and [43] describe a software for the quantification of brain metabolites using data obtained from 2D J-resolved magnetic resonance spectroscopy, a technique developed to reduce the spectral overlap common when using clinical strength 3T NMR. The ProFit tool described in the articles implements a variable projection algorithm that also takes into account possible linear relations between parameters of different metabolites, differences between parameters of different spins in the same metabolite and the fixing of specific parameters.

The Dixon technique is a chemical shift imaging method that allows to create fat only or water only images, and can therefore be used when fat or water conceal the signal of interest. In [44], in order to reconstruct the images from a dual-echo Dixon, a voxel-wise cost function is defined in the linear parameters, water and fat magnitudes W_v , F_v and the nonlinear ones, initial phase shared by water and fat Φ_{0_v} and ω_v , the phase induced by the inhomogeneity of the static magnetic field. To determine the maximum likelihood estimates of these parameters, a variable projection technique is used to solve a simplified nonlinear least squares problem. The method is applied to *in vivo* studies of foot/ankle and CE-MRA of thighs.

The following study [45], on the effect of intra-or-extracellular water accumulation and intracellular acidification in muscles on the rate of transverse relaxation was performed using spectroscopy before and after exercise. The resulting imaging data was modeled by a sum of exponentials and the maximum likelihood fit, a nonlinear least squares approximation, was solved using VARPRO.

The subject of article [46], diffusion MRI (dMRI) uses the diffusion of water molecules in the generation of MRI tissue images. Molecular diffusion in tissues is constrained by interactions with obstacles such as macromolecules, fibers or membranes. In this paper, MIX, a method to characterize the tissue microstruc-

ture of white matter fibers is developed. It uses a multi-compartmental model (intra-axonal, extra-axonal, isotropic) and allows for multiple fiber orientations. The data are fitted in the least squares sense to a combination of exponentials with constraints on the linear coefficients. In a first step, the algorithm constructs good initial values both of the linear and the nonlinear parameters. This is accomplished by applying the variable projection technique to separate linear from nonlinear parameters disregarding for the time being the fact that there are constraints on the linear parameters. The reduced nonlinear problem is solved using a stochastic method, a genetic algorithm, and the linear problem with constrained parameters is solved by CVX (Convex programming Matlab program designed by S. Boyd). In a second step, a trust region method is used to estimate all the parameters. The method was tested on synthetic and *ex-vivo* and *in-vivo* brain data.

An interesting technique for quantitative MRI inspired by Compressed Sensing and developed recently, Magnetic Resonance Fingerprinting (MRF), is the subject of the article [47]. The aim is to acquire enough magnetic resonance signal information in a reasonable short time to be able to deduce multiple tissue-specific parameters such as T_1 and T_2 times and spin density $\rho(\mathbf{x})$. Given the time restriction, and based on the assumption that the image can be approximated by a low-dimensional model, MRF opts for an under-sampled k -space using incoherent data acquisition schemes.

The data model has the form: $\mathbf{d} = \mathbf{FS}\Phi(\mathbf{T}_1, \mathbf{T}_2)\rho$, with Φ and ρ including the parameters that are of interest, \mathbf{S} contains data-acquisition information (coil sensitivities) and Φ is the Fourier encoding matrix. If the noise is Gaussian, a maximum likelihood estimation leads to a nonlinear least squares problem to determine the optimum times $\mathbf{T}_1, \mathbf{T}_2$ and spin density ρ . One of the complications of using optimization methods to solve the problem is the fact that, since there is no analytic expression for the elements of Φ , then Bloch equation simulations have to be performed at each iteration. The algorithm proposed by the authors is an iterative method of a reformulated problem obtained by the introduction of auxiliary variables and the subsequent splitting into three subproblems, two of them linear least squares problems. The third is a nonlinear separable variables least squares problem solved using variable projection techniques. The time consuming Bloch equation solutions needed here are substituted by dictionary values computed in advance.

Wilson [48] considers a new method for spectral registration. In contrast to previous approaches, the registration problem is formulated as a variable-projection (VARPRO) in the frequency domain. The use of VARPRO allows the incorporation of baseline modeling, whilst also reducing the iterative optimization complexity from two parameters (phase and frequency) to one (frequency). The approach is compared with TDSR (time-domain spectral registration), and found to be more robust to large frequency shifts (>7 Hz), baseline distortions and edited-MRS frequency misalignment. In his Ph. D. Thesis, Zhou [49] discusses and compares various methods for exponential fitting associated with

NMRI, including VARPRO. In [50] the authors consider an interesting application to the assessment of myofascial trigger point via MRI for patients with migraine.

3.4. Brain Imaging

This is an interesting area with great future potential as more brain research with electro-magnetic methods is performed.

In [51] the authors consider diffusion MRI to map the brain microstructure. For this purpose they introduce a novel data fitting procedure based on multi-compartment models. Their model is separable and they use VARPRO and a global search algorithm in the nonlinear phase. They show that the algorithm is faster than the commonly used linear search without separation. The accuracy and robustness is demonstrated on synthetic and real data.

Data-acquisition time length is a factor in the applicability of *in-vivo* spectroscopic imaging. Short data-acquisition time is possible when applying echo-planar spectroscopy imaging (EPSI) but with the disadvantage of a poor signal-to-noise ratio. The approach taken in [52] to obtain a high spatiotemporal resolution is to use a hybrid technique using a first step of double-echo chemical shift imaging (CSI) followed by an EPSI step. In the CSI step, the data sets D_{1s} and D_{1L} with a limited number of spatial values but with high temporal resolution are acquired. In the second EPSI stage, a data set D_2 with extended space coverage but limited temporal sampling is obtained. The algorithm uses the union-of-spaces idea, namely that the measured signals are a sum of nuisance, baseline (macromolecules) and metabolite signals. It also assumes partial separability of each sub-signal, *i.e.*, that it can be modeled by a linear combination of temporal basis functions $\varphi_i(t)$ with spatial components $c_i(\mathbf{k})$. In a first preliminary step of the data processing, the nuisance signals are removed from the measured data. In a second step, the temporal basis functions of the baseline and the metabolite signals are determined from the information in the data sets obtained from CSI. Finally, the spatial components are obtained from the signals measured with EPSI. The two last steps involve the solutions of a nonlinear least squares problems and are computed using VARPRO.

In his Ph. D. Thesis [53], J-B Poulet is concerned with the quantification and classification of MRS data for brain tumor diagnosis. MR Spectroscopy is a complementary technique to MR Imaging for brain cancer diagnosis. It provides metabolic information that is not available with MRI. One of the steps in the long and complicated overall process involves AQSES, a time domain quantification method designed for short echo time MR spectra. The model involves a linear combination of corrected *in vitro* metabolic profiles and therefore VARPRO is a natural tool that is used with advantage. Although the use of VARPRO in MRS is not new, the author points out that the previous work was restricted to long echo time MRS signals and the use of complex damped exponentials. He found that the advantages of VARPRO are even larger than in the previous applications,

since now there are a large number of linear parameters that get eliminated. Also, initializing the nonlinear variable to zero provides a good starting point for the optimization process.

4. Conclusion

We have reviewed in detail the application of the Variable Projection method for two important areas: Geophysical and Medical Imaging. The common thread among these applications is the solution of separable nonlinear least squares problems. These are problems in which the model is a linear combination of nonlinear functions. The Variable Projection method proceeds to eliminate analytically the linear variables from the problem. This leads to a much improved behavior that has been analyzed quantitatively in [54]. The bottom line is that problems that hitherto have been very difficult to solve, become much tamer under the VP transformation. The reason for this behavior is an area that deserves more scrutiny in the future.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Golub, G.H. and Pereyra, V. (1973) The Differentiation of Pseudo-Inverses and Nonlinear Least Squares Problems Whose Variables Separate. *SIAM Journal on Numerical Analysis*, **10**, 413-432. <https://doi.org/10.1137/0710036>
- [2] Wold, H. and Lyttkens, E. (1969) Nonlinear Iterative Partial Least Squares (NIPALS) Estimation Procedures. *Bulletin of the International Statistical Institute*, **43**, 29-51.
- [3] Pereyra, V. and Scherer, G. (2019) Variable Projections for Separable Nonlinear Least Squares Problems since 2002. <https://doi.org/10.1155/2019/4861708>
<http://vpereyra.com/wp-content/uploads/2019/08/Surveypaper2019.pdf>
- [4] Golub, G.H. and Pereyra, V. (2003) Separable Nonlinear Least Squares: The Variable Projection Method and Its Applications. *Inverse Problems*, **19**, R1-R26. <https://doi.org/10.1088/0266-5611/19/2/201>
- [5] Pereyra, V., Keller, H.B. and Lee, W.H.K. (1980) Computational Methods for Inverse Problems in Geophysics: Inversion of Travel Time Observations. *Physics of the Earth and Planetary Interiors*, **21**, 120-125. [https://doi.org/10.1016/0031-9201\(80\)90063-1](https://doi.org/10.1016/0031-9201(80)90063-1)
- [6] Pereyra, V. (1996) 3D Block Travel Time Inversion. SEG 66 Annual Meeting, Extended Abstracts. <https://doi.org/10.1190/1.1826565>
- [7] Pereyra, V. and Kaelin, B. (2008) Fast Wave Propagation by Model Order Reduction. *Electronic Transactions on Numerical Analysis*, **30**, 406-419.
- [8] Rickett, J. (2013) The Variable Projection Method for Waveform Inversion with an Unknown Source Function. *Geophysical Prospecting*, **61**, 874-881. <https://doi.org/10.1111/1365-2478.12008>
- [9] Symes, W.W. and Kern, M. (1994) Inversion of Reflection Seismograms by Differential Semblance Analysis: Algorithm Structure and Synthetic Examples. *Geophys-*

- ical Prospecting*, **42**, 565-614. <https://doi.org/10.1111/j.1365-2478.1994.tb00231.x>
- [10] Barnier, G.B., Biondi, E.B. and Biondi, B.B. (2018) Tomographic Full Waveform Inversion via Iterative Least-Squares Migration by Variable Projection Method. *1st EAGE/SBGf Workshop on Least-Squares Migration*, Rio de Janeiro, 27-28 November 2018, 1-5. <https://doi.org/10.3997/2214-4609.201803066>
- [11] Barnier, G.B., Biondi, E.B. and Biondi, B.B. (2018) Full Waveform Inversion by Model Extension. SEG Technical Program Expanded Abstracts. <https://doi.org/10.1190/segam2018-2998613.1>
- [12] Yarman, C.E., Kumar, R. and Rickett, J. (2018) A Model-Based Data-Driven Dictionary Learning for Seismic Data Representation. *Geophysical Prospecting*, **66**, 98-123. <https://doi.org/10.1111/1365-2478.12533>
- [13] Mullen, K.M. and van Stokkum, I.H.M. (2007) TIMP: An R Package for Modeling Multi-Way Spectroscopic Measurements. *Journal of Statistic Software*, **18**, 1-46. <https://doi.org/10.18637/jss.v018.i03>
- [14] Skjelvareid, M.H., Heia, K., Olsen, S.H. and Kristian, S. (2017) Detection of Blood in Fish Muscle by Constrained Spectral Unmixing of Hyper-Spectral Images. *Journal of Food Engineering*, **212**, 252-261. <https://doi.org/10.1016/j.jfoodeng.2017.05.029>
- [15] Westwick, D.T. and Perreault, E.J. (2011) Closed-Loop Identification: Application to the Estimation of Limb Impedance in a Compliant Environment. *IEEE Transactions on Biomedical Engineering*, **58**, 521-530. <https://doi.org/10.1109/TBME.2010.2096424>
- [16] Laptinok, S.P., Borst, J.B., Mullen, K.M., van Stokkum, I.H.M., Visser, A.J.W. and van Anerongen, H. (2010) Global Analysis of Forster Resonance Energy Transfer in Live Cells Measured by Fluorescence Lifetime Imaging Microscopy Exploiting the Rise Time of Acceptor Fluorescence. *Physical Chemistry Chemical Physics*, **12**, 7593-7602. <https://doi.org/10.1039/b919700a>
- [17] Sollmann, N., Weidlich, D., Cervantes, B., Klupp, E., Ganter, C., Kooijman, H., Rummeny, E.J., Zimmer, C., Kirschke, J.S. and Karampinos, D.C. (2019) High Isotropic Resolution T2 Mapping of the Lumbosacral Plexus with T2-Prepared 3D Turbo Spin Echo. *Clinical Neuroradiology*, **29**, 223-230. <https://doi.org/10.1007/s00062-017-0658-9>
- [18] Kadrmas, D.J. and Oktay, M.B. (2013) Generalized Separable Parameter Space Techniques for Fitting 1K-5K Serial Compartment Models. *Medical Physics*, **40**, Article ID: 072502. <https://doi.org/10.1118/1.4810937>
- [19] Riabkov, D.Y. and Di Bella, E.V.R. (2004) Blind Identification of the Kinetic Parameters in Three-Compartment Models. *Physics in Medicine and Biology*, **49**, 639-664. <https://doi.org/10.1088/0031-9155/49/5/001>
- [20] Riabkov, D.Y. and Di Bella, E.V.R. (2002) Estimation of Kinetic Parameters without Input Functions: Analysis of Three Methods for Multichannel Blind Identification. *IEEE Transactions on Biomedical Engineering*, **49**, 1318-1327. <https://doi.org/10.1109/TBME.2002.804588>
- [21] Wu, F.X. and Mu, L. (2009) Separable Parameter Estimation Method for Nonlinear Biological Systems. *3rd International Conference on Bioinformatics and Biomedical Engineering*, Beijing, 11-16 June 2009, 1-4. <https://doi.org/10.1109/ICBBE.2009.5163379>
- [22] Fisz, J.J. (2007) Fluorescence Polarization Spectroscopy at Combined High-Aperture Excitation and Detection: Application to One-Photon-Excitation Fluorescence Microscopy. *The Journal of Physical Chemistry A*, **111**, 8606-8621. <https://doi.org/10.1021/jp072113b>

- [23] Mullen, K.M. and Van Stokkum, I.H.M. (2009) The Variable Projection Algorithm in Time-Resolved Spectroscopy, Microscopy and Mass Spectrometry Applications. *Numerical Algorithms*, **51**, 319-340. <https://doi.org/10.1007/s11075-008-9235-2>
- [24] Mullen, K.M., Vengris, M. and van Stokkum, H.M. (2007) Algorithms for Separable Nonlinear Least Squares with Application to Modeling Time-Resolved Spectra. *Journal of Global Optimization*, **38**, 201-213. <https://doi.org/10.1007/s10898-006-9071-7>
- [25] Jiang, T., Luo, A., Li, X. and Kruggel, F. (2003) A Comparative Study of Global Optimization Approaches to MEG Source Localization. *International Journal of Computer Mathematics*, **80**, 305-324. <https://doi.org/10.1080/002071602200009255>
- [26] Volpato, A., Bolzonello, L., Meneghin, E. and Collini, E. (2016) Global Analysis of Coherence and Population Dynamics in 2D Electronic Spectroscopy. *Optics Express*, **24**, 24773-24785. <https://doi.org/10.1364/OE.24.024773>
- [27] van Leeuwen, T., Maretzke, S. and Batenburg, K.J. (2018) Automatic Alignment for Three-Dimensional Tomographic Reconstruction. *Inverse Problems*, **34**, Article ID: 024004. <https://doi.org/10.1088/1361-6420/aaa0f8>
- [28] Bronnikov, A.V. (1999) Numerical Solution of the Identification Problem for the Attenuated Radon Transform. *Inverse Problems*, **15**, 1315-1324. <https://doi.org/10.1088/0266-5611/15/5/314>
- [29] Gourion, D. and Noll, D. (2002) The Inverse Problem of Emission Tomography. *Inverse Problems*, **18**, 1435-1460. <https://doi.org/10.1088/0266-5611/18/5/315>
- [30] Chung, J. and Nguyen, L. (2016) Motion Estimation and Correction in Photo-Acoustic Tomographic Reconstruction. *SIAM Journal on Imaging Sciences*, **10**, 216-224.
- [31] Sheng, Q., Wang, K., Matthews, T.P., Xia, J., Zhu, L., Wang, L.V. and Anastasio, M.A. (2015) A Constrained Variable Projection Reconstruction Method for Photo-Acoustic Computed Tomography without Accurate Knowledge of Transducer Responses. *IEEE Transactions on Medical Imaging*, **34**, 2443-2458. <https://doi.org/10.1109/TMI.2015.2437356>
- [32] Kargar, S., Borisch, E.A., Froemming, A.T., Kawashima, A., Mynderse, L.A., Stinson, E.G., Trzasko, J.D. and Riederer, S.J. (2018) Robust and Efficient Pharmacokinetic Parameter Non-Linear Least Squares Estimation for Dynamic Contrast Enhanced MRI of the Prostate. *Magnetic Resonance Imaging*, **48**, 50-61. <https://doi.org/10.1016/j.mri.2017.12.021>
- [33] van Stokkum, I.H.M., Larsen, D.S. and van Grondelle, R. (2004) Global and Target Analysis of Time-Resolved Spectra. *Biochimica et Biophysica Acta*, **1657**, 82-104. <https://doi.org/10.1016/j.bbabi.2004.04.011>
- [34] Dorlhiac, G.F., Fare, C. and van Thor, J.J. (2017) PyLDM—An Open Source Package for Lifetime Density Analysis of Time-Resolved Spectroscopic Data. *PLoS Computational Biology*, **13**, e1005528. <https://doi.org/10.1371/journal.pcbi.1005528>
- [35] Ning, Q., Ma, C., Lam, F. and Liang, Z.P. (2017) Spectral Quantification for High-Resolution MR Spectroscopic Imaging with Spatiospectral Constraints. *IEEE Transactions on Biomedical Engineering*, **64**, 1178-1186. <https://doi.org/10.1109/TBME.2016.2594583>
- [36] Pouillet, J.B., Sima, D.M., Simonetti, A.W., De Neuter, B., Vanhamme, L., Lemmerling, P. and Van Huffel, S. (2007) An Automated Quantitation of Short Echo Time MRS Spectra in an Open Source Software Environment: AQSES. *NMR in Biomedicine*, **20**, 493-504. <https://doi.org/10.1002/nbm.1112>
- [37] Sima, D.M. and Van Huffel, S. (2007) Separable Nonlinear Least Squares Fitting with Linear Bound Constraints and Its Application in Magnetic Resonance Spec-

- troscopy Data Quantification. *Journal of Computational and Applied Mathematics*, **203**, 264-278. <https://doi.org/10.1016/j.cam.2006.03.025>
- [38] Sima, D.M., Garcia, M.I.O., Pouillet, J., Suvichakorn, A., Antoine, J.-P., Van Huffel, S. and van Ormondt, D. (2009) Line-Shape Estimation for Magnetic Resonance Spectroscopy (MRS) Signals: Self-Deconvolution Revisited. *Measurement Science and Technology*, **20**, Article ID: 104031. <https://doi.org/10.1088/0957-0233/20/10/104031>
- [39] Barral, J.K., Gudmundson, E., Stikov, N., Etezadi-Amoli, M., Stoica, P. and Nishimura, D.G. (2010) A Robust Methodology for *in Vivo* T1 Mapping. *Magnetic Resonance in Medicine*, **64**, 1057-1067. <https://doi.org/10.1002/mrm.22497>
- [40] Haldar, J.P., Anderson, J. and Sun, S.W. (2007) Maximum Likelihood Estimation of T1 Relaxation Parameters Using VARPRO. *Proceedings of the 15th Annual Meeting of ISMRM*, Berlin, 41.
- [41] Trzasko, J.D., Mostardi, P.M., Riederer, S.J. and Manduca, A. (2013) Estimating T1 from Multichannel Variable Flip Angle SPGR Sequences. *Magnetic Resonance in Medicine*, **69**, 1787-1794. <https://doi.org/10.1002/mrm.24401>
- [42] Fuchs, A., Boesiger, P., Schulte, R.F. and Henning, A. (2014) ProFit Revisited. *Magnetic Resonance in Medicine*, **71**, 458-468. <https://doi.org/10.1002/mrm.24703>
- [43] Schulte, F. and Boesiger, P. (2006) ProFit: Two-Dimensional Prior-Knowledge Fitting of J-Resolved Spectra. *NMR in Biomedicine*, **19**, 255-263. <https://doi.org/10.1002/nbm.1026>
- [44] Stinson, E.G., Trzasko, J.D., Fletcher, J.G. and Riederer, S.J. (2017) Dual Echo Dixon Imaging with a Constrained Phase Signal Model and Graph Cuts Reconstruction. *Magnetic Resonance in Medicine*, **78**, 2203-2215. <https://doi.org/10.1002/mrm.26620>
- [45] Damon, B.M., Gregory, C.D., Hall, K.L., Stark, H.J., Gulani, V. and Dawson, M.J. (2002) Intracellular Acidification and Volume Increases Explain R2 Decreases in Exercising Muscle. *Magnetic Resonance in Medicine*, **47**, 14-23. <https://doi.org/10.1002/mrm.10043>
- [46] Farooq, H., Xu, J., Nam, J.W., Keefe, D.F., Yacoub, E., Georgiou, T. and Lenglet, C. (2016) Microstructure Imaging of Crossing (MIX) White Matter Fibers from Diffusion MRI. *Scientific Reports*, **6**, Article No. 38927. <https://doi.org/10.1038/srep38927>
- [47] Zhao, B., Setsompop, K., Ye, H., Cauley, S.F. and Wald, L.L. (2016) Maximum Likelihood Reconstruction for Magnetic Resonance Fingerprinting. *IEEE Transactions on Medical Imaging*, **35**, 1812-1823. <https://doi.org/10.1109/TMI.2016.2531640>
- [48] Wilson, M. (2018) Robust Retrospective Frequency and Phase Correction for Single-Voxel MR Spectroscopy. *Magnetic Resonance in Medicine*, **81**, 2878-2886. <https://doi.org/10.1002/mrm.27605>
- [49] Zhou, Y. (2018) Estimation and Inference in Problems from Imaging and Biophysics. Ph. D. Thesis, University of Delaware, Newark.
- [50] Sollmann, N., Mathonia, N., Weidlich, D., Bonfert, M., Schroeder, S.A., Badura, K.A., Renner, T., Trepte-Freisleder, F., Ganter, C., Krieg, S.M., Zimmer, C., Rummeny, E.J., Karampinos, D.C., Baum, T., Landgraf Mirjam, N. and Heinen, F. (2019) Quantitative Magnetic Resonance Imaging, of the Upper Trapezius Muscles—Assessment of Myofascial Trigger Points in Patients with Migraine. *The Journal of Headache and Pain*, **20**, 8. <https://doi.org/10.1186/s10194-019-0960-9>
- [51] Farooq, H., Xu, J., Yacoub, E., Georgiou, T. and Lenglet, C. (2015) Brain Micro-

structure Mapping from Diffusion MRI Using Least Squares Variable Separation. CDMRI (MICCAI). https://doi.org/10.1007/978-3-319-28588-7_5

- [52] Ma, C., Lam, F., Ning, Q., Johnson, C.L. and Liang, Z.P. (2017) High-Resolution 1H-MRSI of the Brain Using Short-TE SPICE. *Magnetic Resonance in Medicine*, **77**, 467-479. <https://doi.org/10.1002/mrm.26130>
- [53] Pouillet, J.B. (2008) Quantification and Classification of Magnetic Resonance Spectroscopic Data for Brain Tumor Diagnosis. Ph. D. Thesis, Catholic University of Leuven, Leuven.
- [54] Ruhe, A. and Wedin, P.A. (1980) Algorithms for Separable Nonlinear Least Squares. *SIAM Review*, **22**, 318-337. <https://doi.org/10.1137/1022057>