# $T_2$-Relaxometry for Myelin Water Fraction Extraction Using Wald Distribution and Extended Phase Graph

Alireza Akhondi-Asl, Onur Afacan, Robert V. Mulkern, Simon K. Warfield

Computational Radiology Laboratory, Boston Children's Hospital, and Harvard Medical School 300 Longwood Ave. Boston MA 02115 USA

**Abstract.** Quantitative assessment of myelin density in the white matter is an emerging tool for neurodegenerative disease related studies such as multiple sclerosis and Schizophrenia. For the last two decades, $T_2$ relaxometry based on multi-exponential fitting to a single slice multi-echo sequence has been the most common MRI technique for myelin water fraction (MWF) mapping, where the short $T_2$ is associated with myelin water. However, modeling the spectrum of the relaxations as the sum of large number of impulse functions with unknown amplitudes makes the accuracy and robustness of the estimated MWF's questionable. In this paper, we introduce a novel model with small number of parameters to simultaneously characterize transverse relaxation rate spectrum and $B_1$ inhomogeneity at each voxel. We use mixture of three Wald distributions with unknown mixture weights, mean and shape parameters to represent the distribution of the relative amount of water in between myelin sheets, tissue water, and cerebrospinal fluid. The parameters of the model are estimated using the variable projection method and are used to extract the MWF at each voxel. In addition, we use Extended Phase Graph (EPG) method to compensate for the stimulated echoes caused by $B_1$ inhomogeneity. To validate our model, synthetic and real brain experiments were conducted where we have compared our novel algorithm with the non-negative least squares (NNLS) as the state-of-the-art technique in the literature. Our results indicate that we can estimate MWF map with substantially higher accuracy as compared to the NNLS method.

**Keywords:** $T_2$ Relaxometry, MWF, EPG, Wald, Variable projection.

## 1 Introduction

Myelin is a layer of dielectric material derived mainly from lipids that form a sheath around neuronal axons and is well known to be crucial to support brain function [1]. Myelin-related disorders affect an estimated 3 million people around the world where this number is increasing every year. As such, the development of myelin imaging holds out the potential of providing pathologically specific quantitative information about myelin content. Among myelin imaging techniques, $T_2$

relaxometry is the most advantageous and effective non-invasive MRI approach for measuring alterations in myelin water content. The rationale is that the water molecules bound between myelin sheets, tissue water, and cerebrospinal fluid (CSF) have short, medium, and long $T_2$ relaxation times, respectively. Consequently, the fraction of water molecules with fast decay corresponds directly to the density of the myelin at each voxel. Therefore, by measuring the MR signal for multiple echo times and forming an estimate of the distribution of relaxation rates at each voxel, the fraction of water molecules characterized by fast decay can be estimated. The most well established approach for imaging of this $T_2$ decay is a Carr-Purcell-Meiboom-Gill (CPMG) sequence that collects many spin echo samples of the $T_2$ decay curve [2, 3]. The standard CPMG sequence can be extended to a multi-slice CPMG sequence by changing the 180 degrees RF pulse to a slice selective RF pulses. Multi-slice $T_2$ CPMG acquires several slices simultaneously and allow dramatic acceleration of the acquisition [4].

Conventionally, the myelin-bound water and free water fractions are identified by fitting a discrete mixture of impulse functions, each centered at prespecified $T_2$ values across the range of anticipated $T_2$ values. Each one of the impulse functions represents a single relaxation rate. A linear weight for each impulse function is fit to the multi-echo $T_2$ data via non-negative least squares (NNLS). The fraction of myelin-bound water is computed by summing the weights for all of the short components (below 50ms), and dividing by the total weight for all of the components [5, 6]. Recently, an extension of the NNLS approach is introduced where Extended Phase Graph (EPG) method is used to model the imperfect refocusing in CPMG based sequences. EPG method can be used for the precise calculation of observed echoes as the function of flip-angle, $T_1$, $T_2$, and echo time [7]. The authors optimize the flip angle and weights in a two-step optimization process where in the first step they estimate the flip angle and in the second optimization stage the estimated flip angle is used to estimate the weights of the EPG functions. However, this approach of fitting a discrete mixture of impulse basis functions fails to exploit the continuity of the true distribution of $T_2$ in the tissue.

In this paper, we have developed an alternative representation in which we use a finite mixture of continuous distributions to describe the complete $T_2$ spectrum. The fraction of the myelin-bound water is the area under the fast component curve divided by the total area of each component curve. This representation has the specific advantage that the number of parameters that must be estimated from the data is much smaller. We simultaneously estimate 3 parameters per component and the flip angle, for a total of ten parameters, where the NNLS approaches estimates more than 40 parameters from 32 spin echoes.

This approach, which uses a more physically realistic model of the signal, is also easier to estimate and leads to less noisy MWF estimates. The model we have developed for the $T_2$ distribution of each component is the Wald distribution, which has parameters of volume of occupancy, mean and shape that completely characterize the distribution [8]. The Wald distribution has a Gaussian-like distribution with positive support which makes it suitable for the representation

of transverse relaxation rate distribution. Robust and reliable estimation of the parameters of a mixture of Wald distributions can be achieved with a well-known technique called the variable projection method, which allows us to rapidly solve this nonlinear estimation problem [9, 10]. We have compared our algorithm with a well-known approach in the literature using both synthetic and real brain images and have shown the superiority of our approach.

## 2 Methods

### 2.1 Problem Definition

In the most general form, the MR signal observed at a voxel as a function of echo time is the sum of the signals from a population of spins where each one of them contributes to the observed signal as a function of $R_2$. Therefore, the $i$-th observed signal at the echo time $t_i = i \times TE$ can be expressed as:

$$S_i = S_0 \int_0^\infty f(R_2) EPG(R_2, \theta, TE, i) dR_2 \quad R_2 > 0 \tag{1}$$

where $f(R_2)$ is the probability density function (pdf) of the relaxivity rates,$R_2 = \frac{1}{T_2}$, $S_0$ is a constant, $EPG(R_2, \theta, TE, i)$ is the $i$-th stimulated echo for the spins with the relaxation rate of $R_2$ ,$\theta$ is the flip angle, and $TE$ is the interecho spacing. Since the impact of the $T_1$ value on the stimulated echoes is negligible in CPMG based sequences, we use a fixed $T_1 = 1s$ in all of the experiments. Without loosing generality, we can assume that the density function $f(R_2)$ can be expressed as a mixture of distributions of $n$ components:

$$f(R_2) = \sum_{j=1}^{n} a_j f_j(R_2) \quad \sum_{j=1}^{n} a_j = 1 \tag{2}$$

where $f_j(R_2)$ is the pdf of the $j$-th component and $a_j \leftarrow S_0 a_j$ for simplicity. It is known that the spectrum of relaxivity rate has $n \leq 3$ Gaussian-like components [3]. There are variety of pdf's with positive support which can be used to model the distribution of the components such as truncated Gaussian, log-normal, and Wald distributions. Since the pdf should satisfy $f(R_2 = 0) = 0$, truncated Gaussian is not appropriate distributions in the general case. Here, we use Wald distribution to model $f_j(R_2)$:

$$f_j(R_2) = \left(\frac{\lambda_j}{2\pi R_2^3}\right)^{\frac{1}{2}} \exp\left(\frac{-\lambda_j}{2\mu_j^2 R_2} (R_2 - \mu_j)^2\right) \quad R_2 > 0 \tag{3}$$

where $\mu_j > 0$ and $\lambda_j > 0$ are mean and shape parameter of the distribution, respectively and $\frac{\mu_j^3}{\lambda_j}$ is the variance of the distribution. The Wald distribution has several properties similar to the normal distribution. In addition, for small standard deviations, it becomes very similar to the Gaussian distribution.

### 2.2 Optimization

We are interested to estimate the parameters of the Wald distributions, their mixture weights, and flip angle using the observed signals at different echo times. However, in practice, we observe $y_i$, a noisy version of the signal $S_i$. We assume that zero mean, additive white Gaussian noise is added to the signal $S_i$.

Let, $\{\boldsymbol{\Phi}(\boldsymbol{\alpha})\}_{i,j} = \phi_j(\mu_j, \lambda_j, \theta; t_i)$ be a matrix of size $m \times n$ where

$$\phi_j(\mu_j, \lambda_j, \theta; t_i) = \int_0^\infty \left(\frac{\lambda_j}{2\pi R_2^3}\right)^{\frac{1}{2}} \exp\left(\frac{-\lambda_j}{2\mu_j^2 R_2}(R_2 - \mu_j)^2\right) EPG(R_2, \theta, TE, i) dR_2 \tag{4}$$

and $\boldsymbol{\alpha} = (\mu_1, \lambda_1, \ldots, \mu_n, \lambda_n, \theta) \in \mathcal{R}^{2n+1}$ be the vector of the mean and shape parameters of $n$ Wald distributions, and the flip angle. Given data $(t_i, y_i), i = 1, \ldots, m \geq 3n + 1$, we want to find set of parameters $\hat{\mathbf{a}} = (\hat{a}_1, \ldots, \hat{a}_n)$, $\hat{\boldsymbol{\alpha}} = (\hat{\mu}_1, \hat{\lambda}_1 \ldots, \hat{\mu}_k, \hat{\lambda}_k, \hat{\theta})$ which minimize the following functional:

$$r(\mathbf{a}, \boldsymbol{\alpha}) = ||\mathbf{y} - \boldsymbol{\Phi}(\boldsymbol{\alpha})\mathbf{a}||^2 = \sum_{i=1}^{m} \left(y_i - \sum_{j=1}^{n} a_j \phi_j(\mu_j, \lambda_j, \theta; t_i)\right)^2 \tag{5}$$

where $\mathbf{a} = (a_1, \ldots, a_n) \in \mathcal{R}^n$ and $\mathbf{y} = (y_1, \ldots, y_m) \in \mathcal{R}^m$ are the vectors of mixture weights and noisy observations, respectively.

This functional can be optimized using any Non-linear least squares (NLLS) optimization algorithm. However, since, it has separable NLLS formulation, it is possible to use a smart approach to improve the performance of the optimization [10]. Let us assume that the nonlinear parameters $\boldsymbol{\alpha}$ are known. Therefore, the linear parameters which satisfies the minimal least square solution can be written as $\hat{\mathbf{a}} = \boldsymbol{\Phi}(\boldsymbol{\alpha})^+ \mathbf{y}$ where the matrix $\boldsymbol{\Phi}(\boldsymbol{\alpha})^+$ is the Moore-Penrose generalized inverse of $\boldsymbol{\Phi}(\boldsymbol{\alpha})$. By replacing $\hat{\mathbf{a}} = \boldsymbol{\Phi}(\boldsymbol{\alpha})^+ \mathbf{y}$, the variable projection functional can be constructed:

$$\min_{\mathbf{a}} r(\mathbf{a}, \boldsymbol{\alpha}) = r(\hat{\mathbf{a}}, \boldsymbol{\alpha}) = ||\left(\mathcal{I} - \boldsymbol{\Phi}(\boldsymbol{\alpha})\boldsymbol{\Phi}^+(\boldsymbol{\alpha})\right)\mathbf{y}||^2 = ||\mathcal{P}^{\perp}_{\boldsymbol{\Phi}(\boldsymbol{\alpha})}\mathbf{y}||^2 = r_2(\boldsymbol{\alpha}) \tag{6}$$

where the matrix $\mathcal{P}^{\perp}_{\boldsymbol{\Phi}(\boldsymbol{\alpha})} = \mathcal{I} - \boldsymbol{\Phi}(\boldsymbol{\alpha})\boldsymbol{\Phi}^+(\boldsymbol{\alpha})$ is the projector on the orthogonal complement of the column space of $\boldsymbol{\Phi}(\boldsymbol{\alpha})$.

This indicates that we can first optimize nonlinear parameters by eliminating the linear parameters from the optimization problem. Then, we can use the obtained non-linear parameters to estimate the linear ones using the minimal least square solution. This approach has shown to be very effective in cases where the number of linear parameters is substantial. To optimize the variable projection functional without analytic derivatives, one needs to iteratively compute $r_2(\boldsymbol{\alpha})$. However, to improve the performance of the algorithm, it is possible to use the analytical derivatives in a Levenberg-Marquardt type NNLS solver. Let $\mathbf{D}_k$ be a matrix of size $m \times n$ where $\{\mathbf{D}_k\}_{i,j} = \frac{\partial \phi_j(\mu_j, \lambda_j; t_i)}{\partial \alpha_k} \quad k = 1 \ldots 2n + 1$. It is

known that $\mathbf{J}_k$, $k$-th column of $m \times 2n+1$ Jacobian matrix $\mathbf{J} = \frac{\partial \mathcal{P}^{\perp}_{\boldsymbol{\Phi}(\boldsymbol{\alpha})}\mathbf{y}}{\partial \alpha_k}$, can be computed as [10, 9]:

$$\mathbf{J}_k = -\mathcal{P}^{\perp}_{\boldsymbol{\Phi}(\boldsymbol{\alpha})}\mathbf{D}_k\boldsymbol{\Phi}^{+}\mathbf{y} - (\boldsymbol{\Phi}^{+})^T\mathbf{D}_k^T\mathcal{P}^{\perp}_{\boldsymbol{\Phi}(\boldsymbol{\alpha})}\mathbf{y} \tag{7}$$

Therefore, to derive Jacobian matrix, we only need to provide $\frac{\partial \phi_j(\mu_j,\lambda_j,\theta;t_i)}{\partial \alpha_k}$ which can be computed using the following relation:

$$\frac{\partial \phi_j(\mu_j,\lambda_j,\theta;t_i)}{\partial \alpha_k} = \tag{8}$$

$$\begin{cases} \int_0^{\infty} \left(\frac{\lambda_j}{2\pi R_2^3}\right)^{\frac{1}{2}} \exp\left(\frac{-\lambda_j(R_2-\mu_j)^2}{2\mu_j^2 R_2}\right) \frac{\lambda_j R_2 (R_2-\mu_j)}{\mu_j^3 R_2} EPG(R_2,\theta,TE,i)dR_2 & k = 2j-1 \\ \int_0^{\infty} \exp\left(\frac{-\lambda_j(R_2-\mu_j)^2}{2\mu_j^2 R_2}\right) \frac{1-\frac{\lambda_j(R_2-\mu_j)^2}{\mu_j^2 R_2}}{2(\lambda_j 2\pi R_2^3)^{\frac{1}{2}}} EPG(R_2,\theta,TE,i)dR_2 & k = 2j \\ \int_0^{\infty} \left(\frac{\lambda_j}{2\pi R_2^3}\right)^{\frac{1}{2}} \exp\left(\frac{-\lambda_j}{2\mu_j^2 R_2}(R_2-\mu_j)^2\right) \frac{\partial EPG(R_2,\theta,TE,i)}{\partial \theta} dR_2 & k = 2J+j \\ 0 & otherwise \end{cases}$$

where $\frac{\partial EPG(R_2,\theta,TE,i)}{\partial \theta}$ can be computed recursively.

## 3 Results

### 3.1 Synthetic Data

We use synthetic data with the known ground truth to evaluate our developed method and the NNLS algorithm. To demonstrate the impact of modeling $R_2$ spectrum with a mixture of continuous distribution, we also evaluate a modified version of our model where the Wald distribution is replaced with the impulse function. Mixture of three Wald distributions is considered as the ground truth with the peaks at 50Hz, 10Hz, and 1Hz, the shape parameters of $600Hz$, $400Hz$, and $300Hz$, and weights of 0.2,0.6, and 0.1, respectively.

Figure 1.a shows the normalized Cramer-Rao lower bound (CRLB) of MWF estimated using our method where we normalized the CRLB by the true MWF value. We have computed CRLB for all combinations of 5 equally spaced SNRs between $30dB$ and $50dB$ and 5 equally spaced flip angles between $120^o$ and $200^o$. This figure indicates that MWF can be estimated with very high accuracy in a clinical imaging scenario (SNR=$40dB$ and flip angle larger than $120^o$). Next, we evaluate our optimization algorithm for the same range of SNRs and flip angles. We initialize our method with three Wald distributions with means at $30ms$, $90ms$, and $1500ms$ and with shape parameters of $500Hz$. To have a more robust optimization, we use constraints on the mean and shape parameters of our model. We assume that the mean of the components are in the range of $15-40ms$, $60-120ms$, $200-2000ms$ and the shape parameters are between $0.01-10kHz$. These are reasonable numbers without any strong assumption about the $R_2$ spectrum. We observe 32 echoes at the range of $8-256ms$ and we

repeat each experiment 1000 times at each SNR. Figure 1.b shows the relative mean absolute error (MAE) of our method at different SNRs and flip angles. As seen, for the practical flip angles and SNRs, our method estimates the MWF with very small error. Figure 1.c shows a stimulated multicomponent decay

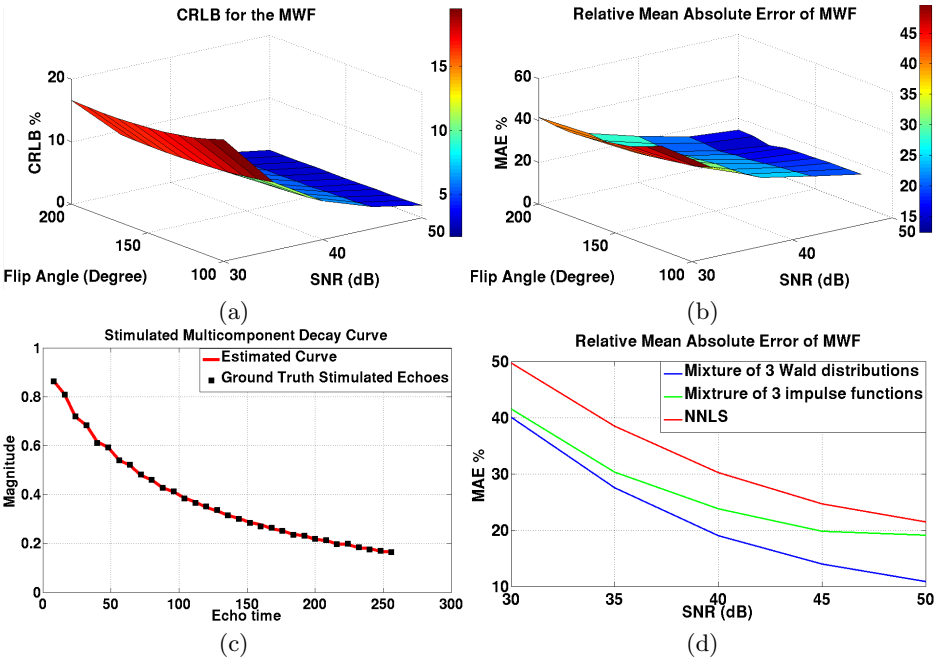
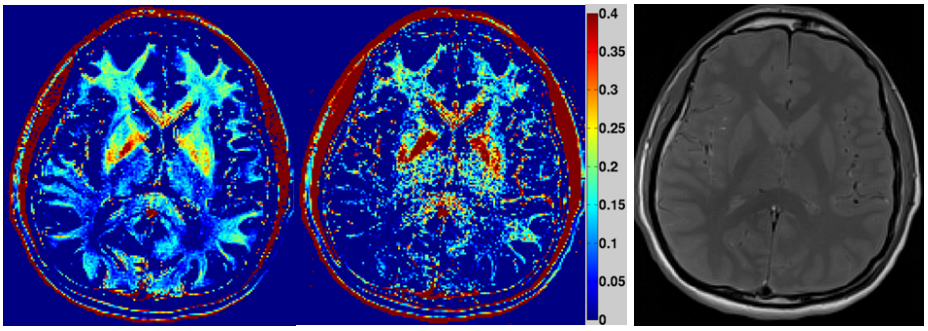



**Fig. 1.** Quantitative evaluation of mixture of Wald distributions. (a) CRLB of MWF estimation using mixture of Wald distributions where the standard deviation is normalized by the true MWF value. (b) Relative MAE of Wald distribution for a range of SNRs and flip angles. (c) Estimated stimulated echoes using the mixture of Wald distributions for $SNR = 45dB$ and flip angle of $150^o$. (d) Relative MAE of the three methods for the range of SNRs and flip angle of $180^o$. The results indicate that we can estimate the MWF accurately for the practical flip angles and SNRs. It can also be seen that our method has lower MAE as compared to the other methods for all the SNRs in the range of $30 - 50dB$.

curve for $SNR = 45dB$ and flip angle of $150^o$ and the estimated curve using our algorithm. This figure shows that we can accurately estimate both the $R_2$ spectrum parameters and the flip angle, as the stimulated echoes are estimated with very high accuracy. Finally, we compare the performance of our method, NNLS, and our modified model with three impulse functions for a range of SNRs and flip angle of $180^o$. For this experiment, we use inverse gamma distribution to generate the ground truth with the same mean, standard deviation, and fractions of the previous experiments. In this way, we will have a fair comparison between

the methods, as a different distribution is utilized to generate the ground truth spectrum. For NNLS method we estimate the amplitude of 50 impulse functions logarithmically spaced within the range of $15ms$ and $2s$. For our model with three impulse functions, we use the initialization and constraints of the mean parameters of the Wald distributions. Figure 1.d shows the relative MAE of the estimation of MWF using three different methods. It can be seen that our method has substantially smaller error as compared to the other methods for the whole SNR range. This indicates that our approach can produce the performance of NNLS method in the substantially lower SNR.




(a) (b)

**Fig. 2.** Qualitative comparison of algorithms. Estimated MWF map using mixture of Wald distribution (a-left) and NNLS (a-right) indicates that the estimated map using our approach is less noisy as compared to the NNLS algorithm. (b) The image of the first echo .

### 3.2 Real Brain MRI

In addition, our method and the NNLS algorithm were tested on 10 volunteers. $T_2$ relaxation measurements were performed on a 3T Siemens TRIO scanner with a multi-echo CPMG sequence acquiring 32 echoes with an echo spacing of 9 ms. A 21cm FOV was used with a matrix size of 192x192 (in plane resolution of 1.1mm) and the total scan time was 9 minutes and 41 seconds for acquisition of 3mm thick slices. Parameter initialization for each of the estimation procedures was the same as the simulation experiment. For the NNLS algorithm the two-step optimization approach in [6] was used to correct for the stimulated echoes. Figure 2.a shows the MWF mapping of one slice of a subject using our method and NNLS. For both models the components with $T_2$ shorter than 50ms are used to estimate the MWF. The results show that our approach estimated the MWF more accurately as compared to the NNLS method, since the MWF map is smoother and less noisy. Figure 2.b shows the image of the first echo.

## 4 Conclusions

In this paper, we have introduced a novel model to represent the spectrum of the relaxation rate at each voxel. To this end, we have utilized a mixture of three Wald distributions with unknown mixture weights, means and shape parameters. We also used EPG method to model stimulated echoes. Finally, we have utilized variable projection method to optimize the unknown parameters and used the estimated mean and mixture weights to identify the MWF at each voxel. We have used both synthetic and real brain images for the validation of our method. In addition, we have compared our method with the state-of-the-art MWF extraction algorithm and showed the superiority of our method.

## 5 Acknowledgment

This research was supported in part by NIH grants R01 EB013248, R01 LM010033, R01 NS079788, R42 MH086984, P30 HD018655, R03 EB008680 and by a research grant from Boston Children's Hospital Translational Research Program.

## References

1. Nave, K.A.: Myelination and support of axonal integrity by glia. Nature **468**(7321) (2010) 244–252
2. Kolind, S.H., Madler, B., Fischer, S., Li, D.K., MacKay, A.L.: Myelin water imaging: implementation and development at 3.0 T and comparison to 1.5 T measurements. Magnetic Resonance in Medicine **62**(1) (2009) 106–115
3. Whittall, K.P., MacKay, A.L.: Quantitative interpretation of NMR relaxation data. Journal of Magnetic Resonance (1969) **84**(1) (1989) 134–152
4. Oh, J., Han, E.T., Lee, M.C., Nelson, S.J., Pelletier, D.: Multislice brain myelin water fractions at 3T in multiple sclerosis. Journal of Neuroimaging **17**(2) (2007) 156–163
5. Mackay, A., Whittall, K., Adler, J., Li, D., Paty, D., Graeb, D.: In vivo visualization of myelin water in brain by magnetic resonance. Magnetic Resonance in Medicine **31**(6) (1994) 673–677
6. Prasloski, T., Rauscher, A., MacKay, A.L., Hodgson, M., Vavasour, I.M., Laule, C., Mädler, B.: Rapid whole cerebrum myelin water imaging using a 3D GRASE sequence. Neuroimage (2012)
7. Hennig, J.: Echoes-how to generate, recognize, use or avoid them in MR-imaging sequences. part I: Fundamental and not so fundamental properties of spin echoes. Concepts in Magnetic Resonance **3**(3) (1991) 125–143
8. Seshadri, V.: The inverse Gaussian distribution: a case study in exponential families. Clarendon Press Oxford (1993)
9. O'Leary, D.P., Rust, B.W.: Variable projection for nonlinear least squares problems. Computational Optimization and Applications (2012) 1–15
10. Golub, G., Pereyra, V.: Separable nonlinear least squares: the variable projection method and its applications. Inverse Problems **19**(2) (2003) R1