D-optimal design of b-values for precise intra-voxel incoherent motion imaging

Mario Sansone¹, Roberta Fusco² and Antonella Petrillo²
¹ Department of Electrical Engineering and Information Technologies, University ‘Federico II’, Naples, Italy
² National Cancer Institute ‘Fondazione Pascale’, Naples, Italy
E-mail: msansone@unina.it

Keywords: diffusion weighted imaging, intravoxel incoherent motion, optimal design

Abstract
The aim of this paper is to optimally design the set of b-values for diffusion weighted MRI with the aim of precise estimation of intra-voxel incoherent motion (IVIM) parameters (f perfusion fraction, D_s slow diffusion, D_f fast diffusion) according to the model developed by Le Bihan. Previous studies have addressed the design in a Monte Carlo fashion; however, due to huge computation times, this approach is practical only for a limited number of values of the parameters (local design): however, as the parameters of a specific patient are not known a priori, it would be desirable to optimise b-values over a region of parameters. In order to address this issue, we propose to use a D-optimal design approach. Our study has two key results: first, under fairly general conditions, the optimal design does not depend on perfusion fraction: this allow to perform a search over a 2D parameter space instead of 3D; second, as an exhaustive search over all possible designs would still be time consuming, we proposed an algorithm to find an approximate solution very quickly.

1. Introduction
Optimal choice of b-values for acquisition of diffusion weighted MR (DW-MR) data is still under debate [1–4]. In general, for precise estimation of IntraVoxel Incoherent Motion (IVIM) parameters [5], appropriately chosen b-values should be used [6, 7]. The number of b-values should be compatible with clinical requirements such as the duration of the diffusion exam, the confort of the patient etc.. Depending on the specific field of view, in typical examinations, the acquisition of one b-value can take a few minutes to complete. Therefore, in clinical literature [1, 3, 4] a relatively small number of b-values have been used (e.g. 10 in breast and prostate cancer [8, 9], 8 in pancreatic cancer [10], 14 in cervical cancer [11]).

Previous attempts to design optimal b-values have been reported in [2–4, 12–20]. Mainly, they have tackled the problem using a Cramer–Rao-lower-bound approach by means of Monte Carlo simulations. It is important to highlight that optimal design in the case of nonlinear model requires an a priori approximate knowledge of the parameters to be estimated as the fisher-information-matrix is parameter-dependent. Therefore, previous studies typically achieved local optimal design for a few values of the parameters. However, it would be desirable to design optimal values for aregion in the parameter space (e.g. \( D_s \in (0, 2) \cdot 10^{-3} \text{mm}^2 \text{s}^{-1} \), \( D_f \in (0, 20) \cdot 10^{-3} \text{mm}^2 \text{s}^{-1} \), \( f \in (0, 0.5) \)) because the specific values for each patient are not known a priori, of course.

A full optimisation within the 4-Dimensional parameter space (see section 2.2) of the IVIM model is computationally intensive but computational load might be reduced using a D-optimal approach: in fact, we will show that under general hypothesis, optimisation of determinant of the fisher-information-matrix depends only on a small subset of parameters. To the best of our knowledge, the problem of b-values design has not yet been addressed using this kind of approach. This has a sounding mathematical basis and can lead, as we show in this manuscript, to a quick design of optimal b-values (among a set of predefined values) over a region of the parameters space.

The specific aim of this paper is to show how the design of the b-values can be simplified performing the search only in the space of the diffusion coefficients and to propose a fast algorithm for finding an (approximate) design over an entire region of this space. The optimal combination of b-values has been chosen...
from a set of predefined values taken from the literature. The reliability of the approximate design has been evaluated on the basis of the Cramer–Rao lower bound. Finally, we would like to underline that although diffusion kurtosis imaging (DKI) is a recent technique for measuring deviations from gaussian distribution at high b-values [18, 19, 21], in this study we focused only on IVIM modelling (see section 4 for further discussion on this point).

2. Methods

In the following, we briefly illustrate some background information concerning principles of optimal design, noise modelling and IVIM modelling. Then, we describe the specific approach for optimal design used in this study. Finally, validation via MonteCarlo simulations is presented.

2.1. Types of optimal design

We briefly introduce the main concepts concerning optimal experimental design; the interested reader is referred to [6, 22].

Let us assume that our measurement procedure provides data with statistical distribution \( f(S; \boldsymbol{\sigma}) \) where \( S = S(b, \boldsymbol{\theta}), \boldsymbol{\theta} = [\theta_1, \ldots, \theta_p]^T \in \Omega \) is a vector of parameters (\( \Omega \) is the parameter space), \( b \in \mathcal{X} \) is some control variable on which the experimenter can act upon (\( \mathcal{X} \) is the design space), and \( \boldsymbol{\sigma} \) are noise parameters. Our objective is to estimate \( \boldsymbol{\theta} \) on the basis of \( N \) independent measurements of \( S \) made with \( b = [b_1, \ldots, b_N]^T \). A common approach is based on maximisation of the log-likelihood of parameters \( \boldsymbol{\theta} \) given by:

\[
L(\boldsymbol{\theta}; \boldsymbol{b}) = \ln \prod_{n=1}^{N} f(S(b_n, \boldsymbol{\theta}), \boldsymbol{\sigma})
\]  

where we have emphasized the dependence upon \( \boldsymbol{b} \) (the noise parameters are usually included in \( \boldsymbol{\theta} \)). To simplify notation, from now on we drop the dependence on \( \boldsymbol{\sigma} \) which is implicitly understood. Maximum Likelihood (ML) estimates of \( \boldsymbol{\theta} \) are commonly chosen because they generally possess good asymptotic \((N \to \infty)\) properties such as unbiasedness and consistency [23]:

\[
\hat{\boldsymbol{\theta}}_{ML}(\boldsymbol{b}) = \text{arg max } L(\boldsymbol{\theta}; \boldsymbol{b}) \quad (2)
\]

in which we again highlight the dependence on \( \boldsymbol{b} \); in general, given two different sets of control variables \( \boldsymbol{b}' \) and \( \boldsymbol{b}'' \) the \( \hat{\boldsymbol{\theta}}_{ML} \) estimates and their asymptotic properties, such as precision, might be different. Moreover, it should be emphasized that, although the ML estimator relies on the asymptotic behaviour of log-likelihood, for a small number of measurements \((N \sim 1)\) it might be biased [23–25].

However, it is well known that, for a fixed \( \boldsymbol{b} \), there is a lower bound for the precision of parameters estimate whatever the estimator used: the Cramer–Rao lower bound (CRLB) per each parameter lies on the diagonal of the inverse of the Fisher Information Matrix (FIM) (equation (3)):

\[
\text{M}(\boldsymbol{\theta}, \boldsymbol{b}) = E \left[ \frac{\partial L}{\partial \theta} \frac{\partial L}{\partial \theta^T} \right] \quad (3)
\]

Therefore, when it comes to the problem of choosing the best \( \boldsymbol{b} \), one can optimise the precision designing the \( \boldsymbol{b} \) points leading to the lowest CRLB.

A geometrical interpretation of these facts in the parameters space is easily illustrated when data distribution is spherical normal \( S \sim N(S(\boldsymbol{\theta}), \sigma^2\mathbf{I}) \) and consequently the distribution of the estimates is also (at least locally, for small deviance from \( \hat{\boldsymbol{\theta}}_{ML} \)) normal \( \theta \sim N(\hat{\theta}_{ML}, \text{M}^{-1}(\hat{\theta}_{ML})) \): in this case the iso-level surfaces of \( \theta \) are ellipsoids whose principal axes are determined by FIM. In particular the square roots of the eigen-values of \( \text{M}^{-1} \) represent the size of the ellipsoids axes and the eigen-vectors represent the directions of the principal axes.

On the basis of previous considerations, different types of design criterion can be individuated: in the next subsections we briefly describe three criteria often used in practice [22] (see figure 1).

![Figure 1](image_url)
2.1.1. D-optimal design
In this criterion
\[
b_{\text{opt}} = \arg \min_{b \in \mathcal{X}} \det(\mathbf{M}^{-1}(\theta; b))
\]
(4)
where \(\det(\cdot)\) is the determinant of a matrix; this is equivalent to minimise the product of the eigen-values of \(\mathbf{M}^{-1}\) corresponding to the volume of the ellipsoide;

2.1.2. A-optimal design
In this criterion
\[
b_{\text{opt}} = \arg \min_{b \in \mathcal{X}} \text{Trace}(\mathbf{M}^{-1}(\theta; b))
\]
(5)
this is equivalent to minimise the sum of the eigen-values of \(\mathbf{M}^{-1}\) corresponding to the square of the diagonal of the rectangle enclosing the ellipsoide;

2.1.3. E-optimal design
In this criterion
\[
b_{\text{opt}} = \arg \min_{b \in \mathcal{X}} \lambda_{\max}(\mathbf{M}^{-1}(\theta; b))
\]
(6)
where \(\lambda_{\max}(\cdot)\) is the largest eigen-value of \(\mathbf{M}^{-1}\) corresponding to the longest axis of the ellipsoide.

For completeness we present another criterion which is interpreted in the measurements space \(\mathcal{W}\) instead of the parameter space \(\Omega\) (the measurement space is defined by \(\mathbf{S} = [S(b_1, \theta), \ldots, S(b_N, \theta)]^T \in \mathcal{W}\)).

2.1.4. G-optimal design
In this criterion
\[
b_{\text{opt}} = \arg \min_{b \in \mathcal{X}} \max_{\theta \in \Omega} \left| \frac{\partial S(b_{\text{opt}})}{\partial \theta^a} \mathbf{M}^{-1}(\theta; b) \frac{\partial S(b_{\text{opt}})}{\partial \theta^b} \right|
\]
(7)
i.e. the maximum variance of the measurements with respect to model expectation is minimised [22]. It can be shown that a G-optimal design is also D-optimal.

2.1.5. Optimisation over a region within \(\Omega\)
From previous sections it is clear that if we knew the specific \(\theta\) to be estimated, a corresponding optimal design can be found. Consider the simple example of a scalar parameter \(\theta\). Let \(c(\theta, b)\) the cost-function to be minimised. Given a certain \(\theta^*\) it is possible to find \(b_{\text{opt}}\) minimising \(c(\theta^*, b)\) in the design space \(\mathcal{X}\). This means that the \(b_{\text{opt}}\) will provide the lowest dispersion \(c(\theta^*, b_{\text{opt}})\) for that specific \(\theta^*\); in general, for the same design other \(\theta\) might have a higher dispersion \(c(\theta, b_{\text{opt}}) > c(\theta^*, b_{\text{opt}})\) (see figure 2).

However, in clinical practice, \(\theta^*\) is not known a priori (in fact, it is what we are looking for!); therefore, it would be auspicious to look for a design which can provide the lowest possible dispersion over a region within \(\Omega\) rather than a single value (see \(b_{\text{optimal}}\) in figure 2). With these considerations in mind, the design criterion might be rewritten as:
\[
b_{\text{opt}} = \arg \min_{b \in \mathcal{X}} \max_{\theta \in \Omega} c(\theta, b)
\]
(8)
in words: the maximum dispersion over \(\Omega\) is minimised by the design \(b_{\text{opt}}\).

It should be highlighted that such a region-optimised design might have, for some \(\theta\), higher dispersion with respect to a non-optimal design (see \(\theta_1\) in figure 2): what is guaranteed by a region-optimised design is that the dispersion will remain lower than a specified level (\(c_{\text{max}}\) in figure 2) over the specified region within \(\Omega\).

2.1.6. Choice of a design criterion
As seen in the previous section, the design criterion can be put in general in the form of minimisation of a cost-function of the type \(c(\theta, b)\). The three criteria
presented in the previous sections 2.1.1, 2.1.2, 2.1.3 do not provide, in general, the same design. From an inspection of figure 1 a good choice seems to be the E-optimal criterion which minimises the largest dispersion of the parameters. This might not be the best choice when parameters have different scales: minimising the dispersion of large scale parameters does not ensure to minimise small scale parameters dispersion. Moreover, this criterion has some disadvantages from a computational perspectives. In particular, eigenvalues computation is a particularly intensive task: exploration of a large design space requires the computation of eigenvalues for each design. Moreover, if we are interested in the optimisation over a region within \( \Omega \) as described in the previous section we would like to have a more quick approach in order to reduce computation times. Similar considerations apply to A-optimality.

From a computational perspective D-optimal design presents some advantages. In fact, given an optimal design with \( N \) design points, recursive formulas can be derived for easily calculating the determinant of the FIM with \( N + 1 \) design points (see later sections). This is particularly useful when one has the requirement to quickly compute the cost-function in order to explore a large design space in a reasonable time. This has been the criterion adopted in this work.

2.1.7. Designs with repetition

In principle, a specific design point could be used several times for data acquisition (repeated measures). Intuitively, this might lead to an improved estimate of noise level. However, computational time for optimisation with repetition is impractical in the case of exact design (see section 2.7.3): therefore, we computed exact design only without repetition. Approximate design (see section 2.7.4), instead, can be obtained very quickly and we computed both no-repetition and with-repetition designs.

2.2. IVIM modelling

The most used model separating the contribute of diffusion and perfusion in intra-voxel incoherent motion has been developed by Le Bihan [5]. According to this pioneering paper, the signal intensity of diffusion weighted MRI can be described by equation (2):

\[
S(b, S_0, f, D, D^*) = S_0[(1 - f)\exp(-bD) + f \exp(-b(D + D^*))]
\]  

where \( b \) is a factor depending on the gradient pulse sequence, \( D \) is the diffusion coefficient of water, \( D^* \) is a pseudo-diffusion coefficient describing blood microcirculation, \( f \) is the fraction of water flowing in perfused capillary, \( S_0 \) is the signal intensity when \( b = 0 \).

As suggested in [4], equation (9) can be rearranged as in equation (10):

\[
S(b, S_0, f, D, D_f) = S_0[(1 - f)\exp(-bD_f) + f \exp(-bD_f)]
\]  

where \( D_f = D \) represents the slow component of diffusion and \( D_f = D + D^* \) represents the fast component of diffusion. This form slightly simplifies formula manipulation and will be used in the following.

Units of the diffusion coefficients are \( \text{mm}^2 \text{s}^{-1} \) while \( b \) is measured in \( \text{mm}^2 \text{s}^{-1} \). Typically [4, 5], \( D_f \) is about \( 20 \cdot 10^{-3} \text{mm}^2 \text{s}^{-1} \) while \( D_f \) is about \( 1 \cdot 10^{-3} \text{mm}^2 \text{s}^{-1} \). In general \( D_f \approx 10D_f \). Values for \( b \), which is the focus of this paper, falls typically in the range \( 0 \rightarrow 1000 \text{ s mm}^{-2} \). Perfusion fraction \( f \) has no units and ranges \( 0 < f < 1 \); typically, in high perfused tissue \( f \leq 0.5 \) [1, 5, 26, 27].

2.3. Noise on diffusion weighted data

It is well known that, because of gaussian noise superimposed to receiving antennas, the measured signal intensity \( (S_m) \) of diffusion MRI data has a Rician distribution [23, 24, 28, 29] as in equation (11):

\[
p(S_m; S, \sigma) = \frac{S_m}{\sigma^2} \exp\left(\frac{S_m^2 + S^2}{2\sigma^2}\right) I_0\left(\frac{S_m S}{\sigma^2}\right)
\]  

where \( S \) is the signal intensity without noise (which should be given by the IVIM model equation (10)), \( \sigma \) is the noise level, \( I_0 \) is the modified Bessel function of the first kind.

It the limit of high Signal Noise Ratio (SNR, \( SS_m/\sigma^2 \rightarrow \infty \)) a very useful approximation is the Gaussian distribution [28–30]:

\[
p(S_m; S, \sigma) \approx \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(S_m - S)^2}{2\sigma^2}\right)
\]  

Therefore, within the limit of this approximation the measured signal intensity can be considered the superposition of zero-mean gaussian noise to the true value:

\[
S_m = S + \epsilon
\]  

with \( \epsilon \sim N(0, \sigma) \). In [25] it has been shown that gaussian approximation is sufficiently accurate when \( S / \sigma > 3 \).

2.3.1. IVIM model for b optimisation

It is a common approach in IVIM literature to normalise diffusion weighted data \( S_m(b) \) dividing by \( S_m(0) \) (in place of \( S_0 \) that is unknown) because in so doing the model simplifies to a three parameters model. This would corresponds to the following modelling:

\[
\frac{S_m(b)}{S_m(0)} = (1 - f)\exp(-bD_f) + f \exp(-bD_f)
\]  

However, as it has been observed in the previous section, while the noise superimposed on \( S \) can be well approximated (for \( S / \sigma > 3 \)) by a gaussian distribution, the distribution of \( S_m(b)/S_m(0) \) tends to be a Cauchy distribution instead [31]. As a consequence, the optimal design procedures, based on equation (13) become inapplicable.
With this in mind, for b design purposes (section 2.4) and for evaluation of attainable precision (section 2.6) we considered here a complete 4 parameters model \( S(b, S_0, f, D_s, D_f) \) including \( S_0 \) among the parameters, instead of the common approach to consider the normalised signal.

### 2.4. D-optimal design

As underlined in the previous sections, diffusion weighted signal intensities in real cases can be well approximated by a Gaussian distribution. Therefore, the measured signal \( S_m(b) \) is well represented by an IVIM component plus an additive noise component as in equation (13) that we rewrite here explicating parameters dependence:

\[
S_m(b; S_0, \theta) = S(b, S_0, \theta) + \epsilon
\]

with \( \epsilon \sim N(0, \sigma) \), and \( \theta = [f, D_s, D_f]^T \).

Under this hypothesis it is possible to use the theory for optimal design of experiments described, for example, in [8, 22]. Optimal design is based upon the computation of the Fisher information matrix of the IVIM model, which we perform in the following.

The Fisher information matrix corresponding to all parameters \( p = [S_0, f, D_s, D_f]^T \) is given by equation (16):

\[
M = [M_{ij}] = \sigma^{-2} \sum_{k=1}^{N} \frac{\partial S(b_k)}{\partial p_i} \frac{\partial S(b_k)}{\partial p_j}
\]

with \( \frac{\partial S(b_k)}{\partial p_i} \) defined as follows:

\[
\frac{\partial S(b)}{\partial p} = \frac{\partial S(b)}{\partial p_i}
\]

where the partial derivatives of \( S \) are evaluated at \( b_k \) with \( b = [b_1, b_2, \ldots, b_N]^T \). However, we note that the model in equation (10) is conditionally linear in the parameter \( S_0 \) thus it can be re-formulated as:

\[
S(b; S_0, \theta) = S_0 \cdot g(b, \theta)
\]

and therefore the Fisher Matrix can be rewritten as:

\[
M = \sigma^{-2} \begin{bmatrix} 1 & 0 \\ 0 & S_0 \end{bmatrix} f^T \begin{bmatrix} 1 & 0 \\ 0 & S_0 \end{bmatrix}
\]

where \( f \) is the identity matrix of size \( 3 \times 3 \), \( A(S_0) \) depends only on \( S_0 \) and

\[
J(b, \theta) = \begin{bmatrix} g(b_1, \theta) & \frac{\partial g(b_1, \theta)}{\partial \theta} & \frac{\partial g(b_1, \theta)}{\partial \theta} \\ g(b_2, \theta) & \frac{\partial g(b_2, \theta)}{\partial \theta} & \frac{\partial g(b_2, \theta)}{\partial \theta} \\ \vdots & \vdots & \vdots \\ g(b_N, \theta) & \frac{\partial g(b_N, \theta)}{\partial \theta} & \frac{\partial g(b_N, \theta)}{\partial \theta} \end{bmatrix}
\]

Further simplification can be obtained using the following notation:

\[
E_i = \exp(-bD_i)
\]

with \( \mathbf{g} = (1 - f)E_i + fE_f \) and \( \partial_f \mathbf{g} = E_f - E_i \). Moreover, we indicate with \( \odot \) the Hadamard product between two vectors e.g.: \( \mathbf{x} \odot \mathbf{y} = [x_1y_1, \ldots, x_Ny_N]^T \).

According to the D-optimal design approach (see section 2.1.1) we must find the \( b \) values maximising the determinant \( ||J|| \) of the Fisher matrix (see section 2.1.1):

\[
||M(b, S_0, \theta)|| = \sigma^{-2}||A(S_0)||^2||f||^2||H^T H(b, D_s, D_f)||
\]

where \( A \) is the design space (the set of all candidate \( b \)-values) and \( \Theta \) a region of interest within the parameters space (containing the expected values of the parameters).

It is well known [22] that the number of \( b \)-values (design points) must be greater (or equal) than the total number of parameters \( P \) (4 in this case) in order for \( M \) not to be singular. Moreover, in a continuous design [22] (i.e. when the number of measurements taken at the design points is very large) the number of design points is limited superiorly by \( P(P + 1)/2 + 1 \) (11 in the present case). In passing to a discrete design (i.e. a single measure is taken at each design point, as is the case for IVIM studies) the number could be greater than this limit. However, we used 20 as maximum because of the previous considerations (see introduction section) concerning the duration of the exam.

### 2.5. Justification for gaussian approximation on real data

In [25] it has been showed in detail that the validity of the gaussian approximation is fairly accurate for \( S/\sigma \geq 3 \). Moreover, they have shown that the Fisher information matrix for a generic model with Rician
noise is given by:
\[
\mathbf{M} = \sigma^{-2} \frac{\partial S(\mathbf{b}^\top)}{\partial \mathbf{b}} E[\mathbf{R}] \frac{\partial S(\mathbf{b})}{\partial \mathbf{b}^\top}
\]
(26)
where \(E[\mathbf{R}]\) is a diagonal matrix whose diagonal elements can be found in [25]; for \(S/\sigma \gtrsim 3\), \(E[\mathbf{R}]\) can be well approximated by the identity matrix \(\mathbf{I}\).

The validity of this approximation should be checked on the region of interest (ROI) under investigation. In our experience, based on diffusion images of the prostate at several \(b\)-values from 0 to 1000, the approximation \(S/\sigma \gtrsim 3\) is very well satisfied. In fact, for voxels within ROI the measured intensity \(S_0\) was in the range [50, 100] for \(b = 0\); the parameter evaluated outside the field of view, was in the range [2, 3] leading to \(S_0/\sigma \approx [25, 50]\). (At our institution images have been obtained using a Siemens scanner, 1.5 T, with echo planar (EP) pulse sequence, segmented k-space (SK), spoiled (SP), oversampling phase (OSP), TR = 7500 ms, TE = 91 ms, 3 averages, flip angle = 90 deg).

Moreover, over the regions of interest, the SNR measures reported in other studies such as e.g. [4] confirm that the noise level \(\sigma\) is typically very low with respect to the signal level \(S\).

2.6. Comparison of designs
To compare different designs we used the Cramer–Rao lower bound (CRLB) of the parameters. \cite{6, 22, 32, 33} Assuming the parameters estimates are not biased, according to the Cramer–Rao lower bound theorem \cite{33} the achievable precision is given by the diagonal elements of the inverse Fisher information matrix. The inverse of the Fisher matrix is given by:
\[
\mathbf{M}^{-1} = \sigma^2 \begin{pmatrix}
1 & 0 \\
0 & S_0^{-1}
\end{pmatrix} (\mathbf{f}^\top \mathbf{j})^{-1} \begin{pmatrix}
1 & 0 \\
0 & S_0^{-1}
\end{pmatrix}
\]
(27)

Calling \(j_k^{-1}\) with \(k = 1, \ldots, 4\), the diagonal elements of the matrix \((\mathbf{f}^\top \mathbf{j})^{-1}\) we have the following bounds for the estimated parameters:
\[
\begin{pmatrix}
\sigma_f^2 \\
\sigma_D^2
\end{pmatrix} \geq \begin{pmatrix}
\frac{1}{2} j_1^{-1} & \frac{1}{2} j_2^{-1} \\
\frac{1}{2} j_3^{-1} & \frac{1}{4} j_4^{-1}
\end{pmatrix}
\]
(28)
and \(\sigma_j^{2i_{11}}\) is the CRLB of the variance of \(S_0\) estimator, which is of no interest here. The square root of CRLB \(\left(\sigma_f, \sigma_D, \sigma_b\right)\) can be considered the achievable precision for the parameter estimate.

We have computed the square root of CRLB per each parameter at the \(\mathbf{b}_{\text{opt}}\) D-optimal values (equation (25)), over the whole parameters region used in the optimisation process.

2.7. Numerical optimisation
2.7.1. Parameters region
The region of the parameter space \((f, D_f, D_D) \in \Theta\) has been chosen on the basis of parameters values found in published literature (see table 1) and using the constraints illustrated in sections 2.3 and 2.5. In particular, by simple computation, it can be verified that for \(D_f \in [0, 2] \cdot 10^{-3} \text{ mm}^2 \text{s}^{-1}, D_D \in [10, 20] \cdot 10^{-3} \text{ mm}^2 \text{s}^{-1}, f \leq 0.5\), with \(S_0/\sigma \gtrsim 30\), the above mentioned conditions are satisfied. These are the range of parameters values expected on the basis on previous studies [2–4].

As explained in sections 2.1.5 and 2.1.6 the resulting region-optimised design guarantees the smallest determinant of inverse FIM over the whole region (see figure 2).

2.7.2. Range of candidate \(b\) values
Also the range of \(b\) values has been chosen after examination of literature [1, 3]. In particular, we have used the following set of \(b\)-values given as in a Matlab-like notation \(b = 0:10:1000\) (100 values).

2.7.3. Exact design
In principle, according to the max-min criterion (equation (25)), in order to find the optimal \(b\), it is required to test all the combinations of \(b\)-values taken from the series of candidate \(b\)-values. Per each combination of \(b\)-values, the determinant of equation (25) must be evaluated on a grid comprising all interesting values \(D_f, D_D\) within the parameter space \(\Theta\) and the minimum value over this region has to be found. Afterwards, the \(b\) giving maximum of all minima must be found. This guarantees that the Fisher determinant evaluated at \(\mathbf{b}_{\text{opt}}\) is the highest possible over the entire parameters region.

However, this approach implies large computational times for designs with a number of design points greater than the minimum (i.e. 4, see section 2.4); this makes also impractical the design based over a large region of the parameters space or the use of an large candidate set of \(b\)-values. For these reasons, in the next section we propose to use a fast although approximate design.

2.7.4. Approximate design
As observed the minimum number \(N\) of design points which gives non singular Fisher matrix for the IVIM model is \(N = 4\). As a starting point, we found the exact optimal 4 points design using the method described in the previous section: \(\mathbf{b}_{\text{opt}}(4) = [0, 180, 570, 1000]\). In particular, we evaluated the determinant of the Fisher matrix over all the parameters space \(\Theta\) and using as \(B\) the space of all possible combinations of 4 points chosen among the 101 candidate \(b\) values chosen. The

| Study | \(D_f \cdot 10^{-3}\) [\text{mm}^2 \text{s}^{-1}] | \(D_D \cdot 10^{-3}\) [\text{mm}^2 \text{s}^{-1}] |
|-------|-----------------|-----------------|
| [4]   | 0.23, 0.7       | 2.52, 2.9       |
| [2]   | 1.0, 5.6        | 2.1, 15.3, 81   |
| [3]   | 1.1, 3          | 11, 16.5, 61    |
| this  | 0.1:0.1:2, 2:1:2 | 0.1:0.1:2, 2:1:2 |
resulting number of combinations in this case is \( \binom{101}{4} = 4082925 \approx 4 \cdot 10^6 \).

However, as observed, with increasing number of design points, the number of possible combination increases quickly (e.g. \( \binom{101}{5} = 7920875 \approx 7 \cdot 10^7 \)), and the computational time increases as well. In order to reduce the computational time we proceeded as follows.

Let \( b(N) = [b_1, \ldots, b_N] \) a design set with \( N \) points. Starting from an optimal design \( b(N) \) we can construct a new design \( b(N+1) = [b(N), b_{\text{new}}] \) adding a new \( b_{\text{new}} \) value. The matrix \( \mathbf{H} \) corresponding to this new set of design points is (dropping down the dependence on \( D_i \) and \( D_f \) for simplicity):

\[
\mathbf{H}(b(N+1)) = \left[ \mathbf{H}(b(N)), \mathbf{r}(b_{\text{new}}) \right]
\]

where

\[
\mathbf{r}(b_{\text{new}}) = \begin{bmatrix}
\exp(-b_{\text{new}}D_i) \\
\exp(-b_{\text{new}}D_f) \\
b_{\text{new}}\exp(-b_{\text{new}}D_i) \\
b_{\text{new}}\exp(-b_{\text{new}}D_f)
\end{bmatrix}
\]

As we are interested in calculation of the determinant of the Fisher matrix we observe that \( \det(\mathbf{J}(b(N+1))) \) we drop the dependence from \( b_{\text{new}}, D_i, \) and \( D_f \) in order to simplify notation:

\[
||\mathbf{M}(b(N+1))|| = ||\mathbf{M}(b(N))|| \cdot (1 + \mathbf{r}^T(\mathbf{H}(b(N))\mathbf{H}(N)^{-1})^{-1})
\]

In order for the new design being optimal this determinant must be maximum. This implies that \( b_{\text{new}} \) must maximise equation (31) over the parameters region. Using equation (31) and starting from the optimal design \( b(4) \), designs with a fixed \( N \) can be constructed iteratively, evaluating one only matrix inverse and performing \( \sim N \) computations instead of examining all the possible combinations.

Unfortunately, this algorithm does not provide an absolutely optimal design, because there can be different combinations of \( N + 1 \) points extracted from the original set, giving a higher determinant. This algorithm provides therefore a sub-optimal solution. However, in our experience (we tested the case \( N = 5 \) and \( N = 6 \)), the precision attainable with the sub-optimal solution is not far away from the optimal solution (see figure 7).

2.8. Validation of optimality

As underlined in section 2.1.5 performances of the design, in terms of cost-function vary with the specific set of parameter: in practice no design can have optimal value of the cost function over the whole parameter region and we can only require that performance do not go below a certain threshold. In practice, there might exist a design having performance better than the optimal for a specific parameter; however, the optimal design is guaranteed not to go below a certain minimum of the cost-function. The cost-function used for the optimisation is the determinant of the inverse information matrix; the performance indices on the basis of which we compared different design is the standard deviation of the Monte Carlo simulation of each parameter.

With this in mind, in order to validate the optimality of the design proposed we attempted an exemplifying Monte Carlo Simulations using the optimal exact design with 5 b-values. In particular, we compared the optimal set \( b_{\text{opt}} = [0, 180, 570, 580, 1000] \) with respect to 3 alternative sets \( b_1 = [0, 100, 500, 900, 1000], b_2 = [0, 50, 100, 150, 1000], b_3 = [0, 700, 800, 900, 1000] \) that have been chosen in such a way that \( b_1 \) spans the whole range of \( b \)-values, \( b_2 \) spans mainly low \( b \)-values and \( b_3 \) spans mainly high \( b \)-values. As the performance can vary with the specific parameters set (see section 2.1.5) we have chosen a single set for exemplificative purposes: \( S_0 = 100, f = 0.3, D_i = 0.6 \cdot 10^{-3} \text{ mm}^2 \text{ s}^{-1}, D_f = 6 \cdot 10^{-3} \text{ mm}^2 \text{ s}^{-1} \). Specifically, we expect that the design \( b_1 \) should have performance roughly similar to \( b_{\text{opt}} \) while \( b_2 \) and \( b_3 \) should show lower standard deviations (SD) on \( D_i \) and \( D_f \), respectively. For each design, Monte Carlo simulations have been performed adding Rician noise to 5000 simulated curves sampled at the design points and Levenberg-Marquardt least-square fitting has been applied (using function lsqcurvefit in Matlab). The starting estimate was the same for all fitting and has been set to:

\( S_0 = 90, f = 0.1, D_i = 1 \cdot 10^{-3} \text{ mm}^2 \text{ s}^{-1}, D_f = 10 \cdot 10^{-3} \text{ mm}^2 \text{ s}^{-1} \). Standard deviations per each parameter have been computed.

3. Results

All computations have been performed in Matlab [34].

3.1. Exact designs with 4, 5 and 6 b-values

Table 2 reports the exact designs obtained for \( N = 4, 5 \) and 6 respectively. They have been obtained considering all the possible combinations (without repetitions) of \( N \) \( b \)-values chosen from the candidate set (see section 2.7.2). Exact design for higher values of \( N \) are impractical (huge computational time) because the number of combinations grows exponentially. For the same reason we did not attempt to search exact design with repetitions.

3.2. Approximate designs up to 20 b-values

Tables 3 and 4 reports the approximate designs obtained using the algorithm described in section 2.7.4 with and without repetitions, respectively. The first four entries of the table coincide with the exact 4 points design (table 2): starting from this one should add the successive entries in order to construct designs with more than 4 points. For example a 6 points design with repetitions includes 0, 180, 570, 1000, 170, 620.
In order to compare exact designs with approximate designs, we have conducted a study involving 20 simulated noisy curves. Parameters have been calculated using the exact algorithm described in section 2.4. For each number of design points (N), the combination of b-values chosen from the candidate set is reported.

Table 2. Optimal designs calculated using the exact algorithm described in section 2.4. Per each number of design points (N) the combination of b-values chosen from the candidate set is reported.

| N  | Exact design without repetitions |
|----|---------------------------------|
| 4  | 0 180 570 1000                  |
| 5  | 0 180 570 580 1000              |
| 6  | 0 170 180 570 580 1000          |

Table 3. Approximate optimal design with repeated measures. Starting from the exact 4 points design (the 4 first entries in the table) we derived additional b-values iteratively using the algorithm proposed in section 2.4. We report results up to N = 20 b-values. Values must be read left → right, top → down. Note that the values 0, 180, 550, 1000, are repeated.

| Approximate design with repetitions |
|------------------------------------|
| 0 180 570 1000 170 620 0 1000 550 |

3.3. Validation of optimality

The results of Monte Carlo simulations have been synthesised in figure 3. As a matter of fact, b_opt has a smaller SD for f and Df, with respect to other designs while for Dv and Sb, the SD is comparable to the others; in particular, b1 and b3 have higher SD for both f and Df while b2 has the smallest SD for Dv.

3.4. Comparison between exact and approximate design basing on CRLB

In order to compare exact designs with approximate ones the CRLB can be used: in particular, for illustrative purposes, the comparison between the exact 6 points design and the approximate one is reported in figure 7. The attainable precision (CRLB) has been computed over the whole parameter region. It is evident that, per each parameter, the CRLB with exact design is not very different from the corresponding precision attainable using the approximate design (without repetitions).

It is useful to compare the precision attainable with design using a small number of points (e.g. 11) with respect to the precision attainable using a very large N = 100 design (e.g. b = 0: 10: 1000, ideal design). The comparison is reported in figure 8. It is evident that no large difference is revealed for f and Df; on the contrary, ‘ideal’ design has a lower SD over the whole region, especially for Df, with respect to the 11 points design: however, in a clinical scenario this might not justify the extra time required. Further insight on this issue can be achieved observing figure 6. As a matter of fact, the use of 20 points produces similar results with respect to 11 points design.

In figures 7 and 8 a value of f = 0.1 has been used; other parameters were Sb = 100 and σ = 3 corresponding to an SNR of Sb/σ ≈ 3.33. These values can be considered to be representative of real values, within the ROIs of interest, from published literature.

In order to give an idea of the location of the design points, figure 4 shows an example of DW data vs b for two sets of parameters with the approximate 20 design (without repetitions) points superimposed.

3.5. Comparison between repeated and non-repeated designs

We computed CRLB both for with and without repetitions designs (5, 6, 7, 8). For repeated designs the results were closely similar to figures 6, 7 and 8 therefore we did not report them, with the interesting exception of CRLB for Dv with 11 points design which is reported in figure 5. It is seen that the CRLB is higher with respect to without-repetition design, over a region with higher values of Dv. This is probably due to the fact that, although repeated design allow a better estimation of noise variance, for small values of N it seems more important to spread the design points over the whole candidate set of b-values.

4. Discussion

The aim of this paper was to design the b-values for diffusion weighted MRI in an optimal manner suitable for precise estimation of intra-voxel incoherent motion parameters according to the model in equation (10).

The design has been conducted according to the principles of the D-optimal approach. Optimal combination of b-values has been chosen within a set of predefined values taken from the literature. As the design is affected by the parameters Dv, Df an exhaustive optimisation within a predefined parameter region has been made.

Our results are in line with other studies in literature [3, 4] which have been conducted via Monte Carlo simulation. Two main disadvantages of the approach via Monte Carlo simulation are connected with the computational burden: first, in order to have statistical precision a large number of simulation must be performed (typically 1000 or higher) and per each simulated noisy curve estimation of parameters must be performed (e.g. via least squares fitting); second, due to computational time only a small portion of the parameters region can be explored in reasonable time. In fact, least squares fitting of noisy simulated curves is time consuming and if Sb is neglected it might be not suitable to the noise structure on the data (see
Figure 3. Comparison of performances between optimal ($b_{opt}$) and non-optimal ($b_1, b_2, b_3$) designs. (a) design points on a curve having $S_0 = 100, f = 0.3, D_s = 0.0006 \text{ mm}^2 \text{s}^{-1}, D_f = 0.006 \text{ mm}^2 \text{s}^{-1}$; (b) performance of the designs per each parameter on a log scale: it can be seen that standard deviation (SD) of the design $b_{opt}$ is the smallest for both $f$ and $D_s$ and is slightly smaller for $S_0$; however, for $D_f$ the design $b_2$ has the smallest SD.

Figure 4. Example of normalised ($S(b)/S_0$) IVIM curve with various parameters. The solid line represents the theoretical curve; circles indicate the approximate $N = 20$ points design.

Figure 5. CRLB for $D_f$ parameter using an approximate design with 11 repeated measures.
section 2.3). Finally, as the parameters values \((f, D_s, D_f)\) are not known prior to the MR exam it would be desirable to have a set of \(b\) values optimised over a large portion of the parameter space.

The approach followed in this study might overcome the above mentioned disadvantages of the Monte Carlo approach. In particular, we showed (section 2.4) that the search for a D-optimal design can be addressed in the 2D space of the diffusion parameters only \((D_s, D_f)\) without considering \(S_0\) and \(f\): this dramatically reduces the computational load.

Moreover, we proposed a fast algorithm for finding an approximate design: on the basis of CRLB analysis we showed that the approximate design is comparable to the exact design (at least in the case of 5 and 6 points designs). In fact, inspection of table 3 and 2 reveals that the approximate designs with 5 and 6 points are only slightly different from the exact 5 and 6 points design. These similarities are confirmed analysing the CRLB in figure 7. From these similarities we infer that also for higher values of \(N\) the exact and approximate designs might be very similar (for \(N \to \infty\) the exact and approximate design should converge\([22]\)).

The comparison of 11 points design with the 100 points design (see figure 8) suggests that even with an ideal design the uncertainty over \(D_f\) is limited. Furthermore, use of 20 points only slightly improves the precision with respect of a 11 points design: this is a useful information in a clinical setting.

A few considerations referring to other DW models are worth here. As reported in\([20, 27]\) it seems that DKI is getting out of the research world to be used in the clinical routine especially for oncologic assessment. Although we recognize that DKI is a very important technique that has gained growing interest, in the following we will try to argue our choice to consider, in this study, only the IVIM effect (modelled as a combination of fast/slow diffusion) at low-intermediate \(b\)-values, neglecting the kurtosis effect (non-gaussian diffusion) at high \(b\)-values.

First, IVIM effect at low-intermediate \(b\)-values has been in use for a long time (the first paper by Le Bihan appeared in 1986), clinicians have become confident with it and the perfusion/diffusion information obtainable has been very well assessed in a number of studies for different clinical applications. However, precision of measurements and patient’s comfort (affecting reduction of movement artifacts) is an important trade-off that is still matter of research\([26]\).

Second, a direct interpretation of apparent kurtosis \((K)\) in terms of biophysical quantities has not been clearly established yet\([8, 9]\). On the contrary, perfusion fraction from IVIM model has been recognized to be linked with classical tissue perfusion\([8, 12]\).

Third, as observed in\([8, 9]\) there is still a debate concerning the measurement of both IVIM effect (perfusion) at low-intermediate \(b\)-values and kurtosis at intermediate-high \(b\)-values: in fact,\([9]\) suggests not to use DKI model at low \((<1000)\) \(b\)-values. Therefore, it seems that perfusion and kurtosis might not be effectively measured simultaneously: a fortiori, an optimal design for IVIM will help in reducing overall scan times when doing both investigations.

Fourth, an effective adoption of DKI by a clinical structure would involve some very strong hardware improvements such as high field (3T) MRI for improving signal-noise-ratio: in fact, DKI requires very high \(b\)-values \((>1000 \text{ s mm}^{-2})\), but because of the exponential decay SNR at high \(b\) might become very low on 1.5T systems. Therefore, in the short period, DKI might not be available to those clinical centres which are not yet up-to-date.

We would like to relate our results to other studies that investigated the influence of low \(b\)-values on estimates reliability\([1, 35–39]\). To this aim, we observe that looking at a logarithmic plot of \(S(b)/S_0\) it seems to be possible to put a threshold \(b_{th}\) separating a perfusion zone...
from a diffusion zone [5, 6]: this is because the pseudo-diffusion coefficient is higher than water diffusion coefficient. This observation is especially useful when a segmented approach to IVIM parameters estimation is used [36, 37]: for \( b > b_{th} \) the curve can be approximated with \( \log(S(b)/S0) = \log(1 - f) - bD_f \) from which an estimate of \( D_f \) and \( f \) can be obtained via simple linear regression. However, there is no clear consensus on which is the best threshold \( b_{th} \) [36–39]: some studies reported \( b_{th} \approx 100 \) [37] or lower [36, 39] others reported about 200 [38]. As a matter of fact, accuracy and precision of the segmented estimates relies on the goodness of the approximations which might depend on the specific parameters values to be investigated. Moreover, the
The number of low b-values to be used is still under debate [5, 6] because low b-values are more error prone and sensitive to signal-noise variation; further, not all currently commercially available MR scanners allow the user to set b-values under 100 [6]. Further, with the aim to interpret our resulting sets of b-values (table 4), we would like to highlight that the overall shape of the S(b) curve is determined simultaneously by all the parameters $S_0$, $f$, $D_f$, and $D_s$. Specifically, the aim of our study was to design the optimal b set for maximising estimates precision of all parameters simultaneously (without focussing on a single parameter), over a region of possible parameters without any a priori assumptions. Therefore, our algorithm neither assumed that the data will be analysed by means of a segmented approach, nor assumed specific parameters values: rather, the data should be fitted using a non-linear...
Least Squares or Maximum Likelihood approach because the $S(b)\text{curve}$ must be taken as a whole. As a further remark, we notice that the set of 11 $b$-values, from table 4, is 0, 10, 170, 190, 510, 570, 620, 960, 990, 1000: there you can see 5 values under 200, with a value lower than 50, which might be used for a segmented analysis.

As regards some practical aspects concerning noise level reduction such as the number of excitations (NEX), investigated e.g in [38], in our algorithm we did not explicitly considered them because, as it is seen in section 2.4, the noise level does not affect the design but only the precision of the estimate.

One final remark is the following. Our study has shown that it is possible to search for an optimal $N + 1$-point design starting from an (approximately) optimal design of $N$ points. This procedure is very fast. This might suggest an adaptive strategy for D-optimal design which could be performed during clinical exam: after a first scan using the minimum set of 4 $B$-values, an estimate $\theta_1$ of the parameters is performed over the region of interest (within the image), opportunely selected by the radiologist; a spatial average of the first estimate might be used for the design of the 5th point; the procedure can be repeated giving estimates $\theta_1, \theta_2$ and so on. The radiologist can stop after reaching the desired precision in the estimates or after a reasonable time. This adaptive procedure has not been investigated here but will be subject of future studies.

5. Conclusion

The design of the $b$-values for optimal estimation of IVIM parameters can be addressed using a D-optimal strategy. In this study we have shown that the optimal design does not depend on perfusion fraction $f$ and therefore the search can be performed in a 2-D space ($D_0, D_1$); moreover, as an exact exhaustive search is still time consuming, we have proposed an iterative algorithm for searching an approximate design starting from an optimal design formed by 4 points.

Acknowledgments

The authors are very grateful to the editor and referees for their comments which contributed to improve the quality of the study.

The authors would like to thank Dr Augusto Aubry at the Dept. Of Electrical Engineering and Information Technologies of University of Naples ‘Federico II’ for fruitful discussions.

This work has been partially funded by the project n. BioMatMRI-000010-ALTRI-2017-M-SANSONE_001_001 at the Department of Electrical Engineering and Information Technologies (DIETI) of the University ‘Federico II’, Naples Italy.

ORCID iDs

Mario Sansone https://orcid.org/0000-0003-3556-7771

References

[1] Koh D-M, Collins D J and Orton M R 2011 Intravoxel incoherent motion in body diffusion-weighted mri: reality and challenges AJR Am. J. Roentgenol. 196 1331–41
[2] Zhang Q, Wang Y-X, Ma H T and Yuan J 2013 Cramer-rao bound for intravoxel incoherent motion diffusion weighted imaging fitting Engineering in Medicine and Biology Society (EMBC), 2013 XXXV Annual International Conf. of the IEEE (Piscataway, NJ: IEEE) pp 511–4
[3] Lemke A, Stieltjes B, Schad L R and Laun F B 2011 Toward an optimal distribution of $b$ values for intravoxel incoherent motion imaging Magn. Reson. Imaging 29 766–76
[4] Jambor I, Merisari H, Aromen H J, Järvinen J, Saunavaara J, Kauko T, Borra R and Pesola M 2014 Optimization of $b$-value distribution for biexponential diffusion-weighted mri imaging of normal prostate Journal of Magnetic Resonance Imaging 39 1213–22
[5] Le Bihan D, Breton E, Lallemant D, Aubin M L, Vignaud J and Laval-Jeantet M 1988 Separation of diffusion and perfusion in intravoxel incoherent motion mri imaging Radiology 168 497–505
[6] Bates D M and Watts D G 1988 Nonlinear Regression: Iterative Estimation and Linear Approximations (New York, USA: John Wiley & Sons, Inc.)
[7] Seber G A F and Wild C J 1989 Nonlinear Regression (New York, USA: John Wiley & Sons, Inc.)
[8] Cho G Y, Moy L, Kim S G, Baete S H, Moccaldi M, Babb J S, Sodickson D K and Sigmund E E 2016 Evaluation of breast cancer using intravoxel incoherent motion (vim) histogram analysis: comparative analysis with malignant status, histological subtype, and molecular prognostic factors European Radiology 26 2547–58
[9] Pesapanee F, Patella F, Fumarola E M, Panella S, Ierardi A M, Pompili G G, Franceschelli G, Angleri S A, Biasina A M and Carrafiello G 2017 Intravoxel incoherent motion (vim) diffusion weighted imaging (dwi) in the periféric prostate cancer detection and stratification Medical Oncology 34 35
[10] Klauff M, Maier-Hein K, Tjaden C, Hackert T, Grenacher L and Stieltjes B 2016 Invivo dw-mri of autoimmune pancreatitis: therapy monitoring and differentiation from pancreatic cancer European Radiology 26 2099–106
[11] Zhu L, Wang H, Zhu L, Meng J, Xu Y, Liu B, Chen W, He J, Zhou Z and Yang X 2017 Predictive and prognostic value of intravoxel incoherent motion (vim) mri imaging in patients with advanced cervical cancers undergoing concurrent chemo-radiotherapy Sci. Rep. 7 11635
[12] Karki K, Hugo G D, Ford J C, Olsen K M, Saraiya S, Groves R and Weiss E 2015 Estimation of optimal $b$-value sets for obtaining apparent diffusion coefficient free from perfusion in non-small cell lung cancer Phys. Med. Biol. 60 7877–91
[13] Xing D, Papadakis N G, Huang C L, Lee V M, Carpenter T A and Hall L D 1997 Optimised diffusion-weighting for measurement of apparent diffusion coefficient (adc) in human brain Magn Reson Imaging 15 771–84
[14] Choi J and Ragunl L G 2010 Robust optimal design of diffusion-weighted magnetic resonance experiments for skin microcirculation J. Magn. Reson. 206 246–54
[15] Caruyer E, Lenglet C, Sapiro G and Deriche R 2013 Design of multishell sampling schemes with uniform coverage in diffusion mri Magn. Reson. Med. 69 1534–40
[16] Brihuega-Moreno O, Heese F P and Hall L D 2003 Optimization of diffusion measurements using cramér-rao lower bound theory and its application to articular cartilage Magn. Reson. Med. 50 1069–76
Le Bihan D, Poupon C, Amadon A and Lethimonnier F 2006 Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. Magn. Reson. Med. 53 1452–60

[19] Jensen J H and Helpern J A 2010 Mri quantification of non-gaussian water diffusion by kurtosis analysis NMR Biomed. 23 698–710

[20] Rosenkrantz A B, Padhani A R, Chenevert T L, Koh D-M, De Keyzer F, Taouli B and Le Bihan D 2015 Body diffusion kurtosis imaging: Basic principles, applications, and considerations for clinical practice J. Magn. Reson. Imaging 42 1190–202

[21] Poot D H J, den Dekker A J, Achten E, Verhooye M and Sijbers J 2010 Optimal experimental design for diffusion kurtosis imaging IEEE Trans. Med. Imaging 29 819–29

[22] Fedorov V V 1972 Theory of Optimal Experiments (Amsterdam: Elsevier)

[23] Sijbers J and den Dekker A J 2004 Maximum likelihood estimation of signal amplitude and noise variance from mr data Magn. Reson. Med. 51 586–94

[24] Sijbers J, den Dekker A J, Scheunders P and Van Dyck D 1998 Maximum-likelihood estimation of rician distribution parameters IEEE Trans. Med. Imaging 17 357–61

[25] Karlsen O T, Verhagen R and Bovée W M 1999 Parameter estimation from rician-distributed data sets using a maximum likelihood estimator: application to t1 and perfusion measurements Magn. Reson. Med. 41 614–23

[26] Le Bihan D, Poupon C, Amadon A and Lethimonnier F 2006 Artifacts and pitfalls in diffusion mri J. Magn. Reson. Imaging 24 478–88

[27] Le Bihan D 2017 What can we see with ivim mri? Neuroimage 187 56–67

[28] Gudbjartsson H and Patz S 1995 The rician distribution of noisy mri data Magn. Reson. Med. 34 910–4

[29] den Dekker A J and Sijbers J 2014 Data distributions in magnetic resonance images: a review Phys. Med. 30 725–41

[30] Kristoffersen A 2007 Optimal estimation of the diffusion coefficient from non-averaged and averaged noisy magnitude data J. Magn. Reson. 187 293–305

[31] Probability A P 1984 Random variables and stochastic processes McGraw Hill, New York, USA

[32] Sansone M, Fusco R and Petriillo A 2015 A geometrical perspective on the 3tp method in dce-mri Biomed. Signal Process. Control 16 32–9

[33] Smith ST 2005 Covariance, subspace, and intrinsic cramer-rao bounds Signal Processing, IEEE Transactions on 53 1610–30

[34] MATLAB 2017 version 9.3.0.713 579 (R2017b) (Natick, Massachusetts: The MathWorks Inc.)

[35] Lima M and Le Bihan D 2016 Clinical intravoxel incoherent motion and diffusion mri imaging: Past, present, and future Radiology 278 13–32

[36] Cohen A D, Schieke M C, Hohenwalter M D and Schmainda K M 2015 The effect of low b-values on the intravoxel incoherent motion derived pseudodiffusion parameter in liver Magn. Reson. Med. 73 306–11

[37] Jalnefjord O, Andersson M, Montelius M, Starck G, Elf A-K, Johanson V, Stensson J and Lungberg M 2018 Comparison of methods for estimation of the intravoxel incoherent motion (ivim) diffusion coefficient (d) and perfusion fraction (fi) Magnetic Resonance Materials in Physics, Biology and Medicine 31 715–23

[38] Liu C, Liang C, Liu Z, Zhang S and Huang B 2013 Intravoxel incoherent motion (ivim) in evaluation of breast lesions: comparison with conventional dvo European Journal of Radiology 82 678–9

[39] Ye C, Xu D, Qin Y, Wang L, Wang R, Li W, Kuai Z and Zhu Y 2019 Estimation of intravoxel incoherent motion parameters using low b-values PLOS ONE 14 1–16