

Maximum Likelihood estimation of signal amplitude and noise variance from MR data

J. Sijbers* and A. J. den Dekker

“Magnetic Resonance in Medicine, Vol. 51, Nr. 3, p. 586-594”

Abstract

In magnetic resonance imaging, the raw data, which are acquired in spatial frequency space, are intrinsically complex valued and corrupted by Gaussian distributed noise. After applying an inverse Fourier transform the data remain complex valued and Gaussian distributed. If the signal amplitude is to be estimated, one has two options. It can be estimated directly from the complex valued data set, or one can first perform a magnitude operation on this data set, which changes the distribution of the data from Gaussian to Rician, and estimate the signal amplitude from the thus obtained magnitude image. Similarly, the noise variance can be estimated from both the complex and magnitude data sets.

This paper addresses the question whether it is better to use complex valued data or magnitude data for the estimation of these parameters using the Maximum Likelihood method. As a performance criterion, the mean-squared error (MSE) is used.

1 INTRODUCTION

Data received from an MRI system are intrinsically complex valued, and represent the Fourier transformation (FT) of a magnetization distribution of a volume at a certain point in time (1). After an inverse FT, these complex MR data are generally transformed into magnitude and phase data as one is more interested in the magnitude and phase of the original magnetizations than in the real and imaginary components. This is because magnitude and phase data are more directly related to the physiological and anatomical quantities of interest.

Most of the current image processing applications applied to MR image data can be formulated as a parameter estimation problem. For example, in the case of noise filtering, the parameter to be estimated is given by the true signal component underlying the noise corrupted data, whereas in the construction of T_1 and T_2 maps the parameters to be estimated are given by the relaxation time constants (2–10).

Usually, the main parameter of interest is the magnitude signal component of the (generally complex valued) model. For this reason, the complex valued data are transformed into a set of magnitude data after which only the latter set is considered for parameter estimation. Nevertheless, one could as well estimate the magnitude model parameter directly from the complex data set. Hence, one could raise the question whether it is better to use the complex data set or the magnitude data set when estimating a specific model parameter.

*Jan Sijbers is a Postdoctoral Fellow of the F.W.O. (Fund for Scientific Research - Flanders, Belgium)

Consider for example the estimation of the underlying magnitude signal component (i.e., the signal amplitude) from a set of image data points (e.g., noise filtering). Thereby, for each image data point, belonging to the data set under concern, the underlying signal amplitude is assumed to be the same. Then, this amplitude can be estimated using either the magnitude MR data or the complex valued data. Indeed, both data sets contain the signal amplitude to be estimated. If N complex points are available, such a data set has in fact $2N$ observations (N real and N imaginary data points) and $N + 1$ unknowns (N true phase values and 1 true signal amplitude). If N magnitude data points are available, then only 1 unknown parameter (the underlying signal amplitude) needs to be estimated from N observations. Similarly, the noise variance can be estimated from both data sets.

Hence, questions may rise like “Should we use the complex data set or the magnitude data set when estimating the unknown signal amplitude or the noise variance?” and “Does it matter whether or not the true phase values of the complex data, from which the signal amplitude is estimated, are the same?”. In this paper, these questions are addressed.

In order to simplify the discussion, we will in the present work elaborate the estimation of the underlying signal amplitude from a set of data points of which this signal amplitude is assumed to be the same (i.e., a constant model). It is however clear that similar reasoning is valid for any other underlying (parametric) model of the data points. Furthermore, although in the past, several estimation methods have been proposed (11–13), we will restrict ourselves to Maximum Likelihood (ML) estimators, which are known to yield optimal (asymptotic) properties (14).

The paper is organized as follows. Section 2 reviews some results from statistical parameter estimation theory which are used in the remainder of the paper. Different performance measures for estimators, as well as the so-called Cramér-Rao Lower Bound and the ML estimator are discussed. In Section 3, ML estimation of the underlying signal amplitude from (complex as well as magnitude) MR data sets is discussed. Section 4 is devoted to ML estimation of the noise variance from these MR data sets. In section 5, the results obtained in sections 3 and 4 are discussed. Finally, in section 6 conclusions are drawn.

2 THEORY

In the following sections, several maximum likelihood estimators will be considered. In order to compare the performance of these estimators, we will discuss their *precision*, *accuracy*, and *mean-squared-error*. The *precision* of an estimator concerns the spread of the estimates when the experiment is repeated under identical conditions. It is represented by the standard deviation, or, equivalently, the variance of the estimator. The variance is thus a measure of the non-systematic error. The *accuracy* of an estimator can be described in terms of its *bias*, which is defined as the deviation of its expectation from the true value of the parameter (vector):

$$\mathbf{b}(\hat{\boldsymbol{\theta}}) = E[\hat{\boldsymbol{\theta}}] - \boldsymbol{\theta} \quad , \quad [1]$$

where $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \dots, \hat{\theta}_K)^T$ represents an estimator of the K -dimensional parameter vector $\boldsymbol{\theta} = (\theta_1, \dots, \theta_K)^T$.¹ Hence, the bias represents the systematic error. The *mean squared error (MSE)* is a measure incorporating both precision and accuracy. The MSE of the k^{th}

¹In Eq. [1] and in what follows, underlined characters are stochastic variables.

element of the estimator $\hat{\theta}$ is defined as:

$$\text{MSE}(\hat{\theta}_k) = E \left[(\hat{\theta}_k - \theta_k)^2 \right] . \quad [2]$$

Note that the MSE can also be written as the sum of the variance of the estimator and its bias squared:

$$\text{MSE}(\hat{\theta}_k) = b^2(\hat{\theta}_k) + \text{Var}(\hat{\theta}_k) . \quad [3]$$

2.1 Cramér-Rao Lower Bound

One and the same parameter can be estimated using different estimators. Generally, different estimators have different precisions. Then, one might ask what precision might be achieved, or, in other words, is there a lower bound on the attainable variance? The answer is yes. Such a lower bound exists. It can be computed from the joint PDF of the observations (i.e., data points) as follows (14).

Suppose that the joint PDF $p_{\underline{x}}(\underline{\mathbf{x}}; \boldsymbol{\theta})$ of a set of observations $\underline{\mathbf{x}} = (\underline{x}_1, \dots, \underline{x}_N)$ is determined by the parameter vector $\boldsymbol{\theta} = (\theta_1, \dots, \theta_K)^T$. In addition, define the matrix \mathbf{I} by

$$\mathbf{I}(\boldsymbol{\theta}) = -E \left[\frac{\partial^2 \log p_{\underline{x}}(\underline{\mathbf{x}}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right] . \quad [4]$$

In this expression, $\partial^2 \log p_{\underline{x}}(\underline{\mathbf{x}}; \boldsymbol{\theta}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T$ is the $K \times K$ Hessian matrix of $\log p_{\underline{x}}(\underline{\mathbf{x}}; \boldsymbol{\theta})$ defined by its (q, r) -th element $\partial^2 \log p_{\underline{x}}(\underline{\mathbf{x}}; \boldsymbol{\theta}) / \partial \theta_q \partial \theta_r$. The matrix \mathbf{I} is called the *Fisher information matrix*. Next, let $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \dots, \hat{\theta}_K)^T$ be any unbiased estimator of the parameter vector $\boldsymbol{\theta}$, that is, $E[\hat{\boldsymbol{\theta}}] = \boldsymbol{\theta}$. Then, under a number of not too restrictive conditions, the Cramér-Rao inequality states that (15):

$$\text{Cov}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}) \geq \mathbf{I}^{-1}(\boldsymbol{\theta}) \quad [5]$$

In this expression, $\text{Cov}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}})$ is the $K \times K$ covariance matrix of the estimator $\hat{\boldsymbol{\theta}}$. Therefore, the diagonal elements are the variances of $\hat{\theta}_1, \dots, \hat{\theta}_K$, respectively.

Inequality [5] expresses that the difference between the positive semi-definite left-hand and right hand members is positive semi-definite. The right-hand member defines the *Cramér-Rao Lower Bound (CRLB)* on the covariance of any unbiased estimator of $\boldsymbol{\theta}$. A property of positive semi-definite matrices is that their diagonal elements cannot be negative. Therefore, the diagonal elements of $\text{Cov}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}})$, that is, the variances of the elements of the estimator $\hat{\boldsymbol{\theta}}$, cannot be smaller than the corresponding elements of the CRLB. Consequently, the latter diagonal elements are a lower bound on the variances of the elements of the estimator $\hat{\boldsymbol{\theta}}$. The CRLB thus defines the highest attainable precision.

Finally, it is known that if there exists an unbiased estimator having the CRLB as covariance matrix, it is the Maximum Likelihood (ML) estimator (14). In the next subsection, the ML estimator will be discussed.

2.2 Maximum Likelihood estimation

Consider a set of N statistically independent observations $(\underline{x}_1, \dots, \underline{x}_N)$. The joint PDF of the observations is then given by:

$$p_{\underline{x}}(\underline{\mathbf{x}}|\boldsymbol{\theta}) = \prod_{n=1}^N p_{\underline{x}_n}(x_n|\boldsymbol{\theta}) , \quad [6]$$

where $\boldsymbol{\theta} = (\theta_1, \dots, \theta_K)^T$ and $p_{x_n}(x_n|\boldsymbol{\theta})$ is the PDF of x_n .

To construct the *Maximum Likelihood* (ML) estimator, we first substitute the available observations for the corresponding independent variables in Eq. [6]. Since these observations are numbers, the resulting expression depends only on the elements of the parameter vector $\boldsymbol{\theta}$. In a second step, we regard the fixed true parameters $\boldsymbol{\theta}$ as variables. The resulting function $L(x_1, \dots, x_N|\boldsymbol{\theta})$ is called the *Likelihood function* of the sample. The ML estimate $\hat{\boldsymbol{\theta}}$ of the parameters $\boldsymbol{\theta}$ is defined as the value of $\boldsymbol{\theta}$ that, within the admissible range of $\boldsymbol{\theta}$, maximizes the likelihood function (16, 17):

$$\hat{\boldsymbol{\theta}}_{\text{ML}} = \arg \left\{ \max_{\boldsymbol{\theta}} (\log L) \right\} \quad . \quad [7]$$

Under very general conditions, ML estimators are known to be consistent and asymptotically efficient (14). Moreover, as noted earlier, if there exists an unbiased estimator of which the variance attains the CRLB, it is given by the ML estimator.

3 SIGNAL AMPLITUDE ESTIMATION

In this section, ML estimation of the signal amplitude from complex as well as magnitude data is considered. Both options are evaluated in terms of accuracy and precision.

3.1 Signal amplitude estimation from complex data

We start by considering a set of N independent, Gaussian distributed, complex data points $\underline{\mathbf{c}} = \{(\underline{w}_{r,n}, \underline{w}_{i,n})\}$. The CRLB for unbiased estimation of the underlying amplitude signal as well as the ML estimator of this signal will be derived. This will be done for data with identical underlying phase values, as well as for data with different phase values.

3.1.1 Region of constant amplitude and phase

First, assume that the complex data $\underline{\mathbf{c}} = \{(\underline{w}_{r,n}, \underline{w}_{i,n})\}$ have an underlying signal amplitude A and identical phase values φ .² This means that $A \cos \varphi$ and $A \sin \varphi$ represent the true real and imaginary values, respectively. As the real and imaginary data are independent, the joint PDF of the complex data, $p_{\underline{\mathbf{c}}}$, is simply the product of the PDF's of the Gaussian distributed real and imaginary data points:

$$p_{\underline{\mathbf{c}}} = \left(\frac{1}{2\pi\sigma^2} \right)^N \prod_{n=1}^N e^{-\frac{(\omega_{r,n} - A \cos \varphi)^2}{2\sigma^2}} e^{-\frac{(\omega_{i,n} - A \sin \varphi)^2}{2\sigma^2}} \quad , \quad [8]$$

where σ^2 denotes the noise variance, and $\{(\omega_{r,n}, \omega_{i,n})\}$ are the real and imaginary variables corresponding with the complex data $\{(\underline{w}_{r,n}, \underline{w}_{i,n})\}$.

CRLB It follows from subsection 2.1 that the CRLB for unbiased estimation of A and φ can be computed from the Fisher information matrix \mathbf{I} :

$$\text{CRLB} = \mathbf{I}^{-1} = \begin{pmatrix} \frac{\sigma^2}{N} & 0 \\ 0 & \frac{\sigma^2}{NA^2} \end{pmatrix} \quad . \quad [9]$$

²Note that the parameter vector $\boldsymbol{\theta}$ in section 2 would in this case be given by (A, φ) .

ML estimation The Likelihood function L is obtained by substituting the available observations $\{(w_{r,n}, w_{i,n})\}$ for $\{(\omega_{r,n}, \omega_{i,n})\}$ into the joint PDF (cfr. Eq. [8]). Then, the ML estimates of A and φ are found by maximizing L , or equivalently $\log L$, with respect to A and φ . At the maximum, the first order derivative of $\log L$ with respect to A and φ are zero. From the resulting equations, the ML estimators of A and φ are found to be:

$$\hat{A}_{\text{ML}} = \frac{1}{N} \sqrt{\left(\sum_{n=1}^N w_{r,n}\right)^2 + \left(\sum_{n=1}^N w_{i,n}\right)^2} \quad [10]$$

$$\hat{\varphi}_{\text{ML}} = \arctan\left(\frac{\sum_{n=1}^N w_{i,n}}{\sum_{n=1}^N w_{r,n}}\right) \quad [11]$$

Notice that the estimator \hat{A}_{ML} is obtained by taking the square root of the quadratic sum of two Gaussian distributed variables. Hence, \hat{A}_{ML} is Rician distributed.

MSE As \hat{A}_{ML} is Rician distributed, we find for its MSE (cfr. Eq. [3]):

$$\text{MSE}\left(\hat{A}_{\text{ML}}\right) = \left[b\left(\hat{A}_{\text{ML}}\right)\right]^2 + \text{Var}\left(\hat{A}_{\text{ML}}\right) \quad [12]$$

$$= 2A\left(A - E\left[\hat{A}_{\text{ML}}\right]\right) + 2\sigma^2/N \quad [13]$$

with

$$E\left[\hat{A}_{\text{ML}}\right] = \frac{\sigma}{\sqrt{N}} \sqrt{\frac{\pi}{2}} {}_1F_1\left[-\frac{1}{2}; 1; -\frac{NA^2}{2\sigma^2}\right], \quad [14]$$

where ${}_1F_1$ is the confluent hypergeometric function of the first kind.

3.1.2 Region of constant amplitude and different phases

Now assume that the complex data $\underline{c} = \{(w_{r,n}, w_{i,n})\}$ have an underlying signal amplitude A and arbitrary phase values $\varphi_1, \dots, \varphi_N$. Then, the joint PDF of the complex data, $p_{\underline{c}}$, is given by:

$$p_{\underline{c}} = \left(\frac{1}{2\pi\sigma^2}\right)^N \prod_{n=1}^N e^{-\frac{(w_{r,n} - A \cos \varphi_n)^2}{2\sigma^2}} e^{-\frac{(w_{i,n} - A \sin \varphi_n)^2}{2\sigma^2}}. \quad [15]$$

CRLB From the Fisher information matrix \mathbf{I} of $(A, \varphi_1, \dots, \varphi_N)$, the CRLB for unbiased estimation of $(A, \varphi_1, \dots, \varphi_N)$ can be computed to be:

$$\text{CRLB} = \begin{pmatrix} \frac{\sigma^2}{N} & 0 & \dots & 0 \\ 0 & \frac{\sigma^2}{A^2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac{\sigma^2}{A^2} \end{pmatrix}. \quad [16]$$

ML estimation The ML estimator of A and φ_n are now found to be:

$$\hat{A}_{\text{ML}} = \frac{1}{N} \sum_{n=1}^N \sqrt{w_{r,n}^2 + w_{i,n}^2} \quad [17]$$

$$\hat{\varphi}_{n,\text{ML}} = \arctan\left(\frac{w_{i,n}}{w_{r,n}}\right) \quad [18]$$

MSE The ML estimator of A , given by Eq. [17], is distributed as the average of N independent, Rician distributed variables. Therefore, its mean value is simply given by the average of the mean values of the individual Rician distributed variables, whereas its variance is given by the sum of their variances, divided by N^2 . Hence, we have for the MSE:

$$\text{MSE}(\hat{A}_{\text{ML}}) = \left[b(\hat{A}_{\text{ML}}) \right]^2 + \text{Var}(\hat{A}_{\text{ML}}) \quad [19]$$

$$= \left(A - E[\hat{A}_{\text{ML}}] \right)^2 + \left(A^2 + 2\sigma^2 - E[\hat{A}_{\text{ML}}]^2 \right) / N \quad [20]$$

where $E[\hat{A}_{\text{ML}}]$ is now given by

$$E[\hat{A}_{\text{ML}}] = \sigma \sqrt{\frac{\pi}{2}} {}_1F_1 \left[-\frac{1}{2}; 1; -\frac{A^2}{2\sigma^2} \right] \quad . \quad [21]$$

Note that Eq. [21] is identical to Eq. [14] in case $N = 1$.

3.2 Signal amplitude estimation from magnitude data

In this subsection, we will discuss estimation of the signal amplitude from a set of N independent magnitude data. Unlike ML estimation of the signal amplitude from complex data, ML estimation of the signal amplitude from magnitude data requires either prior knowledge of the noise variance or simultaneous estimation of signal amplitude and noise variance.

- Prior knowledge of the noise variance may be obtained if a background region is available, i.e., a region in which the underlying signal is zero (see subsection 4.2.2). If large background regions are available, which is often the case, much more data points are available for the estimation of the noise variance than for the estimation of the signal amplitude. Then, the noise variance can be estimated with much higher precision. Hence, it might be a valid assumption to regard the noise variance as known (i.e., to regard the estimated noise variance as the true noise variance).
- If the noise variance cannot be estimated separately (with sufficient precision), it acts as a nuisance parameter that needs to be estimated simultaneously with the signal amplitude.

Both cases are discussed in this subsection. Thereby, we will consider the estimation of the underlying amplitude signal A from N (Rician distributed) magnitude data points $\underline{\mathbf{m}} = (m_1, \dots, m_N)$, where A is assumed to be constant. For this set, the joint PDF $p_{\underline{\mathbf{m}}}$ is given by:

$$p_{\underline{\mathbf{m}}} = \prod_{n=1}^N \frac{M_n}{\sigma^2} e^{-\frac{M_n^2 + A^2}{2\sigma^2}} I_0 \left(\frac{AM_n}{\sigma^2} \right) \quad , \quad [22]$$

where $\{M_n\}$ are the magnitude variables corresponding with the magnitude observations $\{m_n\}$; I_0 denotes the zeroth order modified Bessel function of the first kind.

3.2.1 Region of constant amplitude and known noise variance

CRLB The Fisher information matrix \mathbf{I} of A is given by (6):

$$\mathbf{I} = -E \left[\frac{\partial^2 \log p_{\mathbf{m}}}{\partial A^2} \right] = \frac{N}{\sigma^2} \left(Z - \frac{A^2}{\sigma^2} \right) , \quad [23]$$

with

$$Z = E \left[\frac{\frac{m^2}{\sigma^2} I_1^2 \left(\frac{Am}{\sigma^2} \right)}{I_0^2 \left(\frac{Am}{\sigma^2} \right)} \right] , \quad [24]$$

and m a Rician distributed random variable with true parameters (A, σ) . The expectation value in Eq. [24] can be evaluated numerically. Note that \mathbf{I} is in fact a scalar, from which the CRLB can easily be obtained by applying the inverse operator:

$$\text{CRLB} = \frac{\sigma^2}{N} \left(Z - \frac{A^2}{\sigma^2} \right)^{-1} . \quad [25]$$

ML estimation The ML estimate of A is constructed by substituting the available observations $\{m_n\}$ in Eq. [22] and maximizing the resulting function $L(A)$, or equivalently $\log L(A)$, with respect to A . Hence, it follows that:

$$\log L = \sum_{n=1}^N \log \left(\frac{m_n}{\sigma^2} \right) - \sum_{n=1}^N \frac{m_n^2 + A^2}{2\sigma^2} + \sum_{n=1}^N \log I_0 \left(\frac{Am_n}{\sigma^2} \right) , \quad [26]$$

The ML estimate is then found from the global maximum of $\log L$:

$$\hat{A}_{\text{ML}} = \arg \left\{ \max_A (\log L) \right\} . \quad [27]$$

Notice, that Eq. [27] cannot be solved analytically. Finding the maximum of the (log-)likelihood function is therefore a numerical optimization problem.

3.2.2 Region of constant amplitude and unknown noise variance

If the noise variance is unknown, the signal amplitude and the noise variance have to be estimated simultaneously (i.e., the noise variance is a nuisance parameter).

CRLB After some calculations, the elements of the Fisher information matrix of (A, σ^2) can be shown to be given by:

$$I(1, 1) = \frac{N}{\sigma^2} \left(Z - \frac{A^2}{\sigma^2} \right) \quad [28]$$

$$I(1, 2) = I(2, 1) = \frac{NA}{\sigma^4} \left(1 + \frac{A^2}{\sigma^2} - Z \right) \quad [29]$$

$$I(2, 2) = \frac{N}{\sigma^4} \left(1 + \frac{A^2}{\sigma^2} (Z - 1) - \frac{A^4}{\sigma^4} \right) \quad [30]$$

where $I(i, j)$ denotes the (i, j) th element of the matrix \mathbf{I} and Z is given by Eq. [24]. Then, the CRLB for unbiased estimation of (A, σ^2) becomes:

$$\text{CRLB} = \frac{1}{\det \mathbf{I}} \begin{pmatrix} I(2, 2) & -I(2, 1) \\ -I(1, 2) & I(1, 1) \end{pmatrix} . \quad [31]$$

ML estimation If a background region is not available for noise variance estimation, the signal A and variance σ^2 have to be estimated simultaneously from the N available data points by maximizing the log-likelihood function with respect to A and σ^2 :

$$\{\widehat{A}_{\text{ML}}, \widehat{\sigma}_{\text{ML}}^2\} = \arg \left\{ \max_{A, \sigma^2} (\log L) \right\} \quad , \quad [32]$$

where $\log L$ is given by Eq. [26].

4 NOISE VARIANCE ESTIMATION

In this section, we consider ML estimation of the noise variance from complex (subsection 4.1) as well as magnitude MR data (subsection 4.2) from a region in which the underlying signal amplitude is nonzero but constant as well as from a background region.

4.1 Estimation of the noise variance from complex data

Suppose the noise variance needs to be estimated from N complex valued observations $\underline{c} = \{\{\underline{w}_{r,n}, \underline{w}_{i,n}\}\}$. We will consider the case of identical underlying phase values, as well as the case of different underlying phase values.

4.1.1 Region of constant amplitude and phase

Let us first consider a region with a constant, nonzero underlying signal amplitude and identical underlying phase values.

CRLB From the Fisher information matrix of (A, φ, σ^2) , thereby using $p_{\underline{c}}$ given by Eq. [8], the CRLB for unbiased estimation of (A, φ, σ^2) can easily be found:

$$\text{CRLB} = \begin{pmatrix} \frac{\sigma^2}{N} & 0 & 0 \\ 0 & \frac{\sigma^2}{NA^2} & 0 \\ 0 & 0 & \frac{\sigma^4}{N} \end{pmatrix} \quad . \quad [33]$$

ML estimation For identical true phase values, the ML estimator of σ^2 is given by:

$$\widehat{\sigma}_{\text{ML}}^2 = \frac{1}{2N} \sum_{n=1}^N \left[\left(\widehat{A}_{\text{ML}} \cos \widehat{\varphi}_{\text{ML}} - \underline{w}_{r,n} \right)^2 + \left(\widehat{A}_{\text{ML}} \sin \widehat{\varphi}_{\text{ML}} - \underline{w}_{i,n} \right)^2 \right] \quad , \quad [34]$$

with \widehat{A}_{ML} and $\widehat{\varphi}_{\text{ML}}$ given by Eqs. [10] and [11], respectively.

MSE It can be shown that, for large N , the quantity $2N\widehat{\sigma}_{\text{ML}}^2/\sigma^2$ is approximately distributed as χ_{2N-2}^2 (i.e., chi-square distributed with $2N - 2$ degrees of freedom). Since the mean and variance of a chi-squared variable with λ degrees of freedom are given by λ and 2λ , respectively, we find for the bias and variance of $\widehat{\sigma}_{\text{ML}}^2$:

$$b \left(\widehat{\sigma}_{\text{ML}}^2 \right) \simeq \frac{\sigma^2}{2N} (2N - 2) - \sigma^2 = -\frac{\sigma^2}{N} \quad [35]$$

and

$$\text{Var} \left(\widehat{\sigma}_{\text{ML}}^2 \right) \simeq \frac{\sigma^4}{N} \left(1 - \frac{1}{N} \right) , \quad [36]$$

respectively. Then, the MSE of $\widehat{\sigma}_{\text{ML}}^2$ is given by:

$$\text{MSE} \left(\widehat{\sigma}_{\text{ML}}^2 \right) \simeq \frac{\sigma^4}{N} . \quad [37]$$

4.1.2 Region of constant amplitude and different phases

Next, consider a region with a constant nonzero underlying signal amplitude and different underlying phase values.

CRLB In that case, the CRLB for unbiased estimation of $(A, \varphi_1, \dots, \varphi_N, \sigma^2)$ is given by

$$\text{CRLB} = \begin{pmatrix} \frac{\sigma^2}{N} & 0 & \dots & 0 & 0 \\ 0 & \frac{\sigma^2}{A^2} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac{\sigma^2}{A^2} & 0 \\ 0 & 0 & \dots & 0 & \frac{\sigma^4}{N} \end{pmatrix} . \quad [38]$$

ML estimation In case of different phase values, we have:

$$\widehat{\sigma}_{\text{ML}}^2 = \frac{1}{2N} \sum_{n=1}^N \left[\left(\widehat{A}_{\text{ML}} \cos \widehat{\varphi}_{n,\text{ML}} - \underline{w}_{r,n} \right)^2 + \left(\widehat{A}_{\text{ML}} \sin \widehat{\varphi}_{n,\text{ML}} - \underline{w}_{i,n} \right)^2 \right] , \quad [39]$$

with \widehat{A}_{ML} and $\widehat{\varphi}_{n,\text{ML}}$ now given by Eqs. [17] and [18], respectively.

MSE It can be shown that, for large N , the quantity $2N\widehat{\sigma}_{\text{ML}}^2/\sigma^2$ is approximately distributed as $\chi_{2N-(N+1)}^2 = \chi_{N-1}^2$ (18). This means, that

$$b \left(\widehat{\sigma}_{\text{ML}}^2 \right) \simeq \frac{\sigma^2}{2N} (N-1) - \sigma^2 = -\frac{\sigma^2}{2} \left(1 + \frac{1}{N} \right) \quad [40]$$

and

$$\text{Var} \left(\widehat{\sigma}_{\text{ML}}^2 \right) \simeq \frac{\sigma^4}{2N} \left(1 - \frac{1}{N} \right) . \quad [41]$$

Hence, the MSE of $\widehat{\sigma}_{\text{ML}}^2$ is given by:

$$\text{MSE} \left(\widehat{\sigma}_{\text{ML}}^2 \right) \simeq \frac{\sigma^4}{4} \left(1 + \frac{4}{N} - \frac{1}{N^2} \right) . \quad [42]$$

4.1.3 Background region

Next, consider the case in which the noise variance is estimated from a background region.

CRLB It can easily be shown that the CRLB for unbiased estimation of σ^2 is given by:

$$\text{CRLB} = \frac{\sigma^4}{N} \quad , \quad [43]$$

independent of the underlying phase values.

ML estimation The ML estimator is given by

$$\widehat{\sigma^2}_{\text{ML}} = \frac{1}{2N} \sum_{n=1}^N (\underline{w}_{r,n}^2 + \underline{w}_{i,n}^2) \quad , \quad [44]$$

independent of the underlying phase values.

MSE It can easily be shown that the ML estimator in Eq. [44] is unbiased and that its variance equals the CRLB. Therefore, its MSE is simply given by:

$$\text{MSE} \left(\widehat{\sigma^2}_{\text{ML}} \right) = \frac{\sigma^4}{N} \quad . \quad [45]$$

4.2 Estimation of the noise variance from magnitude data

We will now describe estimation of the noise variance from a set of N magnitude MR data points $\{\underline{m}_n\}$. In subsection 4.2.1, we will consider estimation of the noise variance from a (non-zero) constant region. In this case, the noise variance has to be estimated simultaneously with the signal amplitude. In subsection 4.2.2, we will consider estimation of the noise variance from a background region.

4.2.1 Region of constant amplitude

If the data set $\{\underline{m}_n\}$ is available from a constant (non-zero) signal region, the CRLB for unbiased estimation of σ^2 is given by Eq. [31], in which the elements are defined by Eqs. [28-30].

Furthermore, the value of σ^2 can be estimated using the ML method as follows (cfr. Eq. [32]):

$$\left(\widehat{A}_{\text{ML}}, \widehat{\sigma^2}_{\text{ML}} \right) = \arg \left\{ \max_{A, \sigma^2} \left(-N \log \sigma^2 - \sum_{n=1}^N \frac{m_n^2 + A^2}{2\sigma^2} + \sum_{n=1}^N \log I_0 \left(\frac{Am_n}{\sigma^2} \right) \right) \right\} \quad [46]$$

Note that it requires the optimization of a 2D function, which cannot be solved analytically.

4.2.2 Background region

If the data set $\{\underline{m}_n\}$ is available from a background region, these data are known to be governed by a Rayleigh distribution and their joint PDF, $p_{\underline{m}}$, is given by:

$$p_{\underline{m}}(\{M_n\}) = \prod_{n=1}^N \frac{M_n}{\sigma^2} e^{-\frac{M_n^2}{2\sigma^2}} \quad , \quad [47]$$

where $\{M_n\}$ are the magnitude variables corresponding to the magnitude observations $\{\underline{m}_n\}$.

CRLB In that case, the CRLB for unbiased estimation of σ^2 is given by:

$$\text{CRLB} = \frac{\sigma^4}{N} \quad . \quad [48]$$

ML estimation The likelihood function is obtained by substituting the available background data points $\{m_n\}$ for the variables $\{M_n\}$ in Eq. [47]. Then, maximizing the log-likelihood function with respect to σ^2 , yields the ML estimator of σ^2 :

$$\widehat{\sigma^2}_{\text{ML}} = \frac{1}{2N} \sum_{n=1}^N m_n^2 \quad . \quad [49]$$

Notice, that the estimators given in Eqs. [44] and [49], are identical.

MSE It can be shown that Eq. [49] is an unbiased estimator for the noise variance, that is, its mean is equal to σ^2 . Furthermore, the variance of the ML estimator in Eq. [49] is equal to σ^4/N , which equals the CRLB given by Eq. [48] for all values of N . Hence,

$$\text{MSE} \left(\widehat{\sigma^2}_{\text{ML}} \right) = \frac{\sigma^4}{N} \quad . \quad [50]$$

5 DISCUSSION

5.1 Signal amplitude estimation

5.1.1 CRLB

The CRLB for unbiased estimation of the signal amplitude from *complex* MR data with identical and different phase values is given by Eq. [9] and [16], respectively. From these equations, it is clear that these lower bounds

- are inversely proportional to the number of data points used for the estimation.
- do not depend on the phase values.

The CRLB for unbiased estimation of the signal amplitude from *magnitude* MR data with known and unknown noise variance was obtained from Eq. [25] and [31], respectively. Thereby, the expectation values were evaluated numerically from Monte Carlo simulations. From these equations, it is clear that the lower bounds are inversely proportional to the number of data points used for the estimation. Fig. 1 shows these lower bounds as a function of the signal-to-noise ratio (SNR), defined as A/σ , for unbiased estimation of the signal amplitude from complex as well as from magnitude data for known as well as for unknown noise variance. From the figure, one can see that

- for low SNR (SNR<3), the CRLB for unbiased estimation of A is significantly smaller for estimation from complex data than for estimation from magnitude data.
- for low SNR (SNR<3), the CRLB for unbiased estimation of A from magnitude data with known noise variance is significantly smaller than for estimation from magnitude data with unknown noise variance.
- for high SNR (SNR>3), the CRLB's for unbiased estimation from magnitude data tend to the CRLB for unbiased estimation from complex data, which equals σ^2/N .

5.1.2 MSE

Recall that the MSE of an estimator is the sum of its bias squared and its variance. The bias, variance, and MSE of the ML estimators of A , as derived above, were computed, where the number of data points was set to $N = 25$ and the true variance was set to $\sigma^2 = 1$. For complex data with identical and different phase values, the bias, variance, and MSE of \hat{A}_{ML} were computed from Eq. [13-14] and [20-21], respectively. On the other hand, for magnitude data with known and unknown noise variance, the bias, variance, and MSE of \hat{A}_{ML} were obtained from a Monte-Carlo simulation experiment with sample size 10^5 . Thereby, \hat{A}_{ML} was obtained by maximizing the log-likelihood function [26] with respect to A and $\{A, \sigma^2\}$ using Eq. [27] and [32], respectively.

The bias of \hat{A}_{ML} has been plotted as a function of the SNR in Fig. 2. In Fig. 3, the MSE of \hat{A}_{ML} has been plotted as a function of the SNR. Both figures show the results obtained for complex data with identical and different phase values as well as for magnitude data with known and unknown noise variance. From these figures, one can observe that, in terms of the MSE,

- \hat{A}_{ML} for complex data with identical phases performs best, independent of the SNR,
- \hat{A}_{ML} for magnitude data with known noise variance is significantly better compared to \hat{A}_{ML} for magnitude data with unknown noise variance and \hat{A}_{ML} for complex data with different phase values,
- for high SNR (i.e., for $\text{SNR} > 5$), the performance of all ML estimators of A are approximately equal.

As in practice, the assumption of identical phases for complex data is generally invalid, it may be concluded that the signal amplitude is preferentially estimated from magnitude MR data for which the noise variance is known. The latter requisite is not too restrictive as often in practice the noise variance can be estimated with a much higher precision than the signal amplitude.

In addition, the dependence of both the bias and the variance of \hat{A}_{ML} on the number of data points has been investigated.

- The *bias* of \hat{A}_{ML} for complex data with identical phases as well as for magnitude data with known and unknown noise variance generally decreases with the number of data points used for the estimation. On the other hand, it turns out that for complex data with different phases, the bias of \hat{A}_{ML} does not decrease with the number of data points.
- The *variance*, as may be expected, turns out to be inversely proportional to the number of data points for all estimators.

5.2 Noise variance estimation

5.2.1 CRLB

The CRLB for unbiased estimation of the noise variance from *complex* data with identical and different phase values is given by Eq. [33] and [38], respectively. Note that, in both cases, and independent of the signal amplitude of the data points, the CRLB is equal to σ^4/N .

Furthermore, the CRLB for unbiased estimation of the noise variance from *magnitude* data has been computed for a background region as well as for a constant region. In Fig. 4, the CRLB is shown as a function of the SNR.

- If the noise variance is estimated from N magnitude data of a *background* region, the CRLB is equal to σ^4/N (cfr. Eq. [48]). Then, the CRLB is the same as for estimation from N complex data of a background region. This might be surprising as estimation from N complex data actually exploits $2N$ real valued (N real and N imaginary) observations while estimation from N magnitude data only exploits N real valued observations. However, this is compensated by the fact that the Rayleigh PDF has a smaller standard deviation.
- If the noise variance is estimated from N magnitude data of a *constant* region, the CRLB is given by Eq. [31]. It can numerically be shown that for magnitude data this CRLB tends to $2\sigma^4/N$ when the SNR increases, which is a factor 2 larger compared to estimation from complex data (cfr. Fig. 4). This is not surprising since the Rice PDF tends to a Gauss PDF for high SNR with the same variance as the PDF of the real or imaginary data. Hence, for high SNR, the difference in CRLB between magnitude and complex data can simply be explained by the number of observations available for the estimation of the noise variance.

5.2.2 MSE

For *complex* data from a region with constant amplitude with identical and different phase values, expressions for the bias, variance, and MSE of $\widehat{\sigma}_{\text{ML}}^2$ were derived. The bias of $\widehat{\sigma}_{\text{ML}}^2$ is given by Eq. [35] and [40], respectively. From these expressions, it is clear that:

- Both noise variance estimators are biased. Also, note that the bias of $\widehat{\sigma}_{\text{ML}}^2$ is independent of the true signal amplitude.
- For identical phases, the bias of $\widehat{\sigma}_{\text{ML}}^2$ decreases inversely proportionally with the number of observations (N). In contrast, the bias of $\widehat{\sigma}_{\text{ML}}^2$ for different phases does not; for large N , it converges to $\sigma^2/2$.

The variance of $\widehat{\sigma}_{\text{ML}}^2$ for *complex* data from a region with constant amplitude and identical and different phase values is given by Eq. [36] and [41], respectively. Note that only for complex data with identical phases, the variance of $\widehat{\sigma}_{\text{ML}}^2$ asymptotically attains the CRLB. This may be explained by the fact that for estimation from complex data with different phases the number of unknown parameters that need to be estimated simultaneously with σ^2 , is proportional to N .

The MSE of $\widehat{\sigma}_{\text{ML}}^2$ for *complex* data from a region with constant amplitude and identical and different phase values is given by Eq. [37] and [42], respectively. Both are shown in Fig. 5 as a function of N . Moreover, the MSE of $\widehat{\sigma}_{\text{ML}}^2$ as a function of the SNR is shown in Fig. 6.

The MSE of $\widehat{\sigma}_{\text{ML}}^2$ for *magnitude* data from a region with constant amplitude can be found numerically from Eq. [46]. The results for this estimator, as a function of the SNR, are also shown in Fig. 6.

Finally, the MSE of $\widehat{\sigma}_{\text{ML}}^2$ from background MR data is given by σ^4/N , for magnitude as well as for complex data, and independent of the phases. From this, it is clear that

the noise variance should be estimated from background data points, whenever possible. If a background region is not available, a similar reasoning for the estimation of σ^2 as for the estimation of A holds. That is, estimation of σ^2 from complex data with identical phases is then preferred to estimation from magnitude data, which, in turn, is preferred to estimation from complex data with different phases.

6 CONCLUSIONS

It has been shown that maximum likelihood estimation of the signal amplitude from complex data points with equal true phase values is generally better in terms of the mean-squared error compared to maximum likelihood estimation from magnitude data points. However, in practice, phase values usually vary within the data set from which the amplitude signal is estimated. In that case, estimation from magnitude data is significantly better in terms of the mean-squared error.

With respect to estimation of the noise variance, it is clear that estimation from a background region is preferred to estimation from a (non-zero) constant area. Furthermore, the mean-squared errors of the maximum likelihood estimators of the noise variance from complex and magnitude background data points are equal.

REFERENCES

1. Callaghan PT. Principles of Nuclear Magnetic Resonance Microscopy. Clarendon Press, Oxford, 2nd edition, 1995.
2. Sijbers J, den Dekker AJ, Verhoye M, Van Audekerke J, Van Dyck D. Estimation of noise from magnitude MR images. *Magn Reson Imaging* 1998;16:87–90.
3. Sijbers J, den Dekker AJ, Van der Linden A, Verhoye M, Van Dyck D. Adaptive, anisotropic noise filtering for magnitude MR data. *Magn Reson Imaging* 1999;17:1533–1539.
4. Nowak RD. Wavelet based Rician noise removal for magnetic resonance images. *IEEE Trans Image Processing* 1999;10:1408–1419.
5. Wood JC, Johnson KM. Wavelet packet denoising of magnetic resonance images: Importance of Rician noise at low SNR. *Magn Reson Med* 1999;41:631–635.
6. Karlsen OT, Verhagen R, Bovée WM. Parameter estimation from Rician-distributed data sets using maximum likelihood estimator: application to T_1 and perfusion measurements. *Magn Reson Med* 1999;41:614–623.
7. Sijbers J, den Dekker AJ, Raman E, Van Dyck D. Parameter estimation from magnitude MR images. *Int J Imag Syst Tech* 1999;10:109–114.
8. Borgia GC, Brown RJS, Fantazzini P. Uniform-penalty inversion of multiexponential decay data. *J Magn Reson* 2000;147:273–285.
9. van der Weerd L, Vergeldt FJ, de Jager PA, Van As H. Evaluation of algorithms for analysis of NMR relaxation decay curves. *Magn Reson Imaging* 2000;18:1151–1157.

10. Does MD, Gore JC. Compartmental study of T_1 and T_2 in rat brain and trigeminal nerve in vivo. *Magn Reson Med* 2002;47:274–283.
11. Miller AJ, Joseph PM. The use of power images to perform quantitative analysis on low SNR MR images. *Magn Reson Imaging* 1993;11:1051–1056.
12. McGibney G, Smith MR. An unbiased signal-to-noise ratio measure for magnetic resonance images. *Med Phys* 1993;20:1077–1078.
13. Sijbers J, den Dekker AJ, Scheunders P, Van Dyck D. Maximum likelihood estimation of Rician distribution parameters. *IEEE Trans Med Imag* 1998;17:357–361.
14. van den Bos A. *Handbook of Measurement Science*, Edited by P.H. Sydenham, Wiley, Chichester, England, volume 1. 1982; 331–377.
15. Zacks S. *The theory of Statistical Inference*. Wiley, New York, 1971.
16. Fisher RA. On the mathematical foundations of theoretical statistics. *Phil Trans R Soc Lond A* 1922;222:309–368.
17. Fisher RA. Theory of statistical estimation. *Proc Camb Phil Soc* 1925;22:700–725.
18. Seber GAF, Wild CJ. *Nonlinear regression*. John Wiley and Sons, New York, 1989.

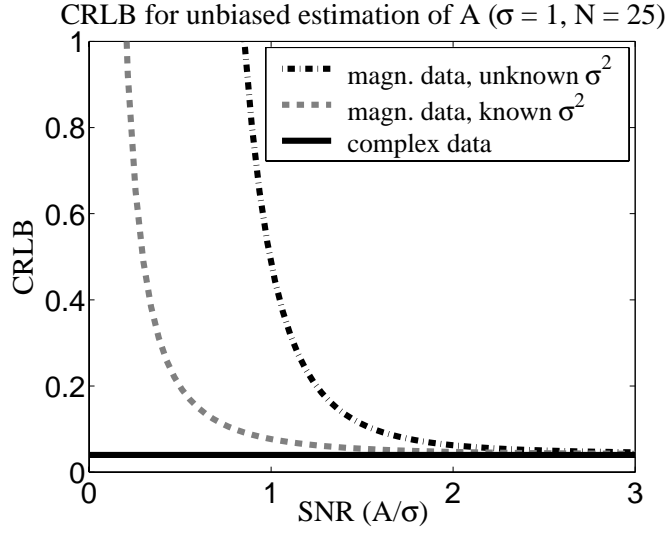


FIG. 1. CRLB for unbiased estimation of A from complex and magnitude data, with known and unknown noise variance.

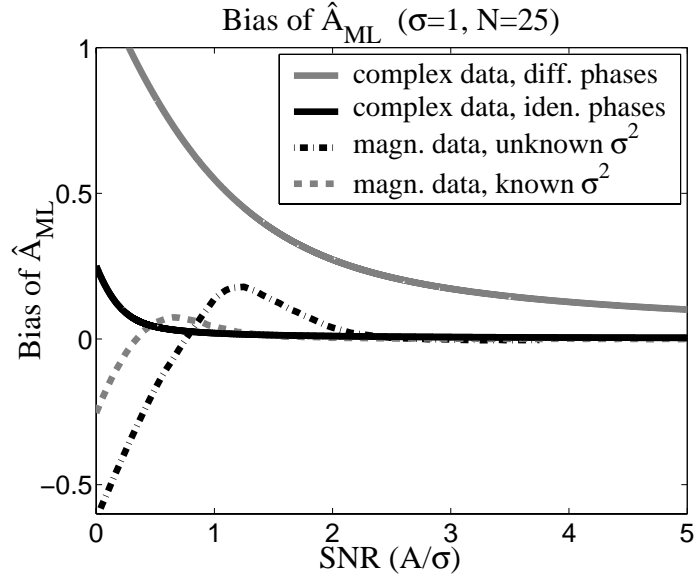


FIG. 2. Bias of \hat{A}_{ML} for complex data with identical and different phases and for magnitude data with unknown and known noise variance.

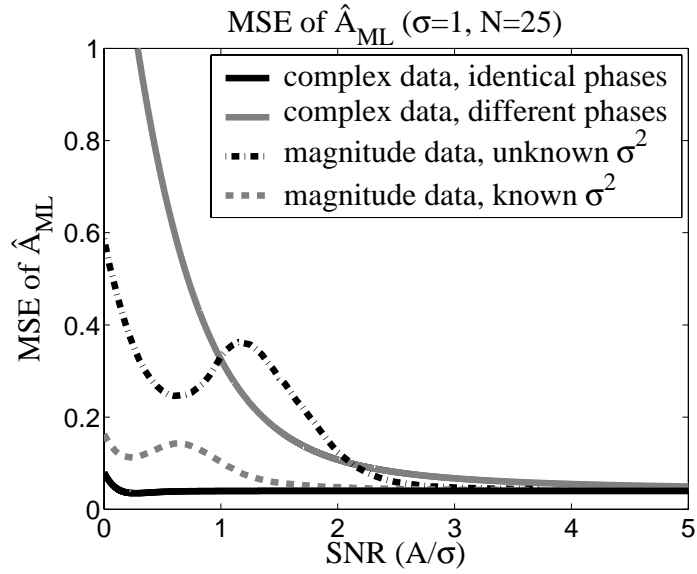


FIG. 3. MSE of \hat{A}_{ML} for complex data with identical and different phases and for magnitude data with unknown and known noise variance.

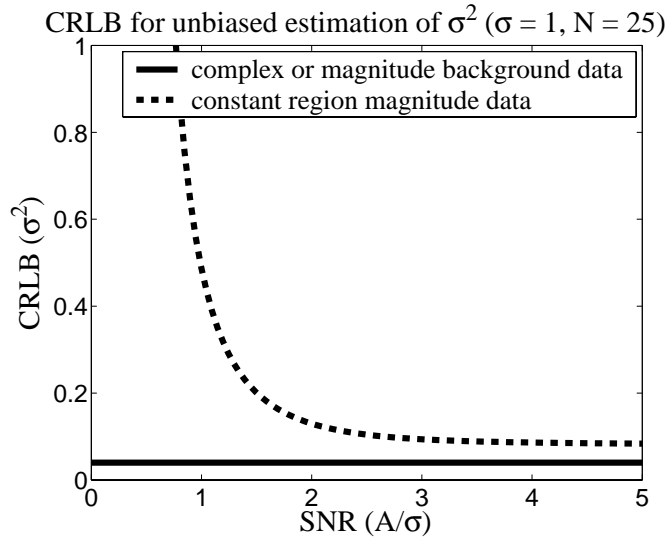


FIG. 4. CRLB for unbiased estimation of the noise variance for complex data and magnitude background data (full line) and for magnitude data from a constant region, that is, where A is estimated simultaneously (dashed line).

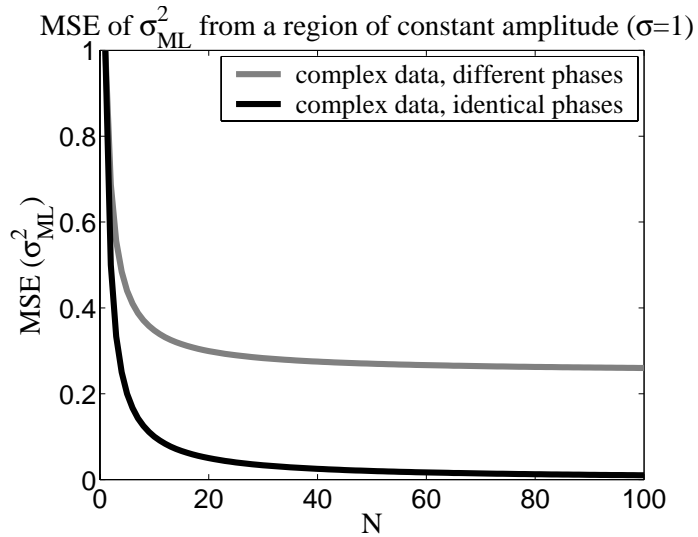


FIG. 5. MSE for unbiased estimation of the noise variance from a region of constant amplitude, for complex data with identical and different phases.

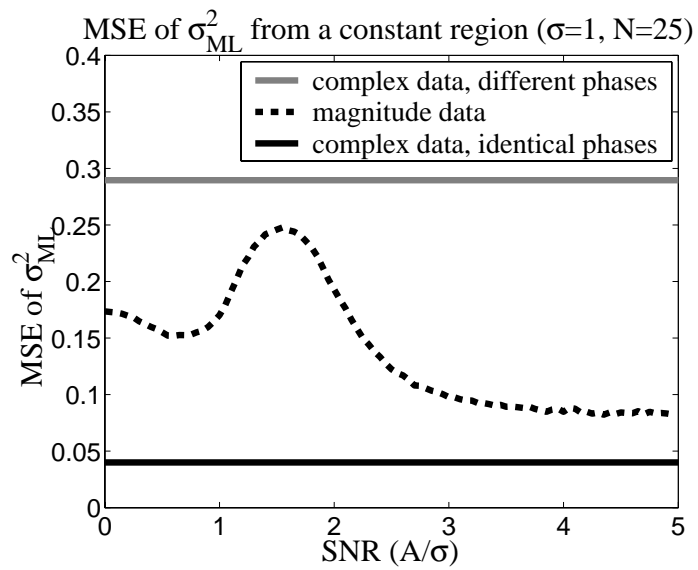


FIG. 6. MSE for unbiased estimation of the noise variance from a region of constant amplitude, for complex data with identical and different phases and for magnitude data.

LIST OF SYMBOLS

A	signal amplitude	4
\hat{A}	estimate of the signal amplitude	7
$\underline{\hat{A}}$	estimator of the signal amplitude	5
b	bias	2
$\underline{\mathbf{c}}$	set of complex observations (stochastic variables)	4
$E[.]$	expectation operator	2
${}_1F_1$	confluent hypergeometric function of the first kind	5
\mathbf{I}	Fisher information matrix	3
I_0	zeroth order modified Bessel function of the first kind	6
I_1	first order modified Bessel function of the first kind	7
K	number of parameters	2
L	likelihood function	4
m	magnitude observation (number)	7
\underline{m}	magnitude observation (stochastic variable)	6
M	magnitude variable corresponding to the observation m	6
N	number of observations	2
p	probability density function	3
φ	true phase value	4
$\underline{\hat{\varphi}}$	estimator of the phase	5
σ	standard deviation of the noise	4
$\underline{\sigma^2}$	variance of the noise	4
$\widehat{\sigma^2}$	estimate of the noise variance	8
$\underline{\widehat{\sigma^2}}$	estimator of the noise variance	8
$\boldsymbol{\theta}$	parameter vector	2
$\hat{\boldsymbol{\theta}}$	estimate of $\boldsymbol{\theta}$	2
$\underline{\hat{\boldsymbol{\theta}}}$	estimator of $\boldsymbol{\theta}$	3
$\omega_{i,n}$	imaginary variable corresponding to the observation $w_{i,n}$	4
$w_{i,n}$	imaginary observation (number)	5
$\underline{w_{i,n}}$	imaginary observation (stochastic variable)	4
$\omega_{r,n}$	real variable corresponding to the observation $w_{r,n}$	4
$w_{r,n}$	real observation (number)	5
$\underline{w_{r,n}}$	real observation (stochastic variable)	4