

The Bias/Variance Trade-Off When Estimating the MR Signal Magnitude From the Complex Average of Repeated Measurements

M. Dylan Tisdall,^{1,2*} Richard A. Lockhart,³ and M. Stella Atkins⁴

The signal-dependent bias of MR images has been considered a hindrance to visual interpretation almost since the beginning of clinical MRI. Over time, a variety of procedures have been suggested to produce less-biased images from the complex average of repeated measurements. In this work, we re-evaluate these approaches using first a survey of previous estimators in the MRI literature, then a survey of the methods statisticians employ for our specific problem. Our conclusions are substantially different from much of the previous work: first, removing bias completely is impossible if we demand the estimator have bounded variance; second, reducing bias may not be beneficial to image quality. Magn Reson Med 000:000–000, 2011. © 2011 Wiley-Liss, Inc.

Key words: signal-to-noise ratio; image bias; maximum likelihood estimate; denoising; nuisance parameters

INTRODUCTION

MRI magnitude images are constructed from noisy measurements of some presumed “true” signal magnitude at each voxel in the subject. An estimator is a function which takes the noisy measurement as its arguments and returns an estimate of the “true” value. In the context of MRI, we use an estimator function to estimate the values of the “true” signal magnitude at each voxel and then render an image of these estimates. The quality of the estimator function used is then of great importance in determining the quality of the resulting image.

An estimator is biased if “true magnitude” s is, on average, mapped to an estimated value $\hat{s} \neq s$; the bias is then $b_{\hat{s}} = \hat{s} - s$. An estimator is said to have a signal-dependent bias if $b_{\hat{s}}$ varies with s . In the context of MRI images, this could, for example, reduce contrast by overestimating the magnitude in low-magnitude regions. This is, in fact, a known issue: interest in the signal-dependent bias of MRI magnitude images seems to have first appeared, along with

the earliest attempt to compensate for it, in work by Henkelman (1,2). However, the recognition that the magnitude of a quadrature MRI measurement has a Rician distribution, and this distribution’s link to the signal processing literature, was first noted by Bernstein et al. (3), who proposed that correcting the Rician’s bias would improve image contrast and in turn improve feature detectability. Several authors have since suggested techniques for correcting or reducing the bias of a Rician-distributed MRI magnitude measurement (4–10), in addition to related previous work on this problem in other disciplines (11–14).

The problem of bias correction can be thought of in statistical terms as the construction of an estimator function taking the measured data as its arguments and giving an estimate of the magnitude where the function’s bias varies slowly—ideally not at all—as the “true” signal varies. It is important to note that these estimators are only tasked with correcting for the measurement contribution due to thermal noise. Artifacts (e.g., ghosts) are not considered noise for the purpose of these estimators but are lumped in as part of the “true” signal. Although ghosts and other artifacts almost always have a structured, systematic influence on the measurements, the thermal noise is a stochastic contribution to the measurement and is thus particularly amenable to statistical estimation methods.

Several related estimation problems have arisen in the MRI literature from this basic setup, so we will begin by surveying them and carefully specifying the estimation problem that we are interested in for the remainder of this work. Our chosen estimation problem has been well-studied in the MRI literature, but we make three contributions here. First, we survey the previous work and contextualize it with the suggested methods from statistical estimation theory, along the way introducing a novel estimator [the mean Bayesian (MB) estimator]. Second, we introduce a novel bound on estimator performance for our chosen problem with Eqs. 32 and 33, and their extensions in the Appendix. This bound imposes a trade-off between bias and variance for any possible estimator in our problem, and any improvement beyond the bound must come from the inclusion of prior information (e.g., smoothness assumptions). Third, we analyze the estimators from the previous literature, and the new estimator we present, relative to this bound.

Before we describe the estimation problem under study, we need a model of the thermal noise. All of the estimators we are interested in rely on a common probabilistic model: measurements in quadrature MRI are assumed to

¹Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, Massachusetts, USA

²Department of Radiology, Harvard Medical School, Brookline, Massachusetts, USA

³Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, British Columbia, Canada

⁴School of Computing Science, Simon Fraser University, Burnaby, British Columbia, Canada

Grant sponsor: Natural Sciences and Engineering Research Council of Canada

*Correspondence to: M. Dylan Tisdall, Ph.D., Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, 149 13th Street, Charlestown, MA 02129. E-mail: tisdall@nmr.mgh.harvard.edu

Received 18 October 2010; revised 25 January 2011; accepted 15 February 2011.

DOI 10.1002/mrm.22910

Published online in Wiley Online Library (wileyonlinelibrary.com).

© 2011 Wiley-Liss, Inc.

consist of two independent channels, each corrupted with the addition of zero-mean gaussian noise (1,6,15–17):

$$p(a, b; s, \phi, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(a - s \cos \phi)^2 + (b - s \sin \phi)^2}{2\sigma^2}\right). \quad [1]$$

In our notation here, a and b are the random variables for the measurements in the real and imaginary channels respectively, s and ϕ are the true signal magnitude and phase (where “true” phase is due, for example, to B_0 inhomogeneity), and σ is the standard deviation of the thermal noise. If we represent our measurements in polar coordinates, we have

$$p(r, \theta; s, \phi, \sigma) = \frac{r}{2\pi\sigma^2} \exp\left(-\frac{s^2 + r^2 - 2sr \cos(\theta - \phi)}{2\sigma^2}\right), \quad [2]$$

where r and θ are the random variables for the measured signal magnitude and phase.

A family of similar-looking estimation problems can be generated from this model depending on what assumptions are made about the signal. In the following, we assume that s and ϕ are always unknown and σ is always known.

1. r and θ are measured one or more times at each voxel. We assume that s , ϕ , and σ are constant for all measurements at each voxel. We want to estimate the unknown s at each voxel and ignore the unknown ϕ . This form of the problem has been addressed by several groups (1,4–6,9). Under the additional assumption that the unknown ϕ is slowly varying spatially, this problem has been addressed by several groups (3,18–20); instead, with a spatial local-smoothness constraint on s this problem has been addressed by (21,22).
2. r is measured one or more times at each voxel (θ is not measured). We assume that s and σ are constant for all measurements at each voxel as we allow ϕ to vary between measurements. Estimation of the unknown s while ignoring the unknown ϕ has been addressed by Sijbers et al. (8,9).

There are many more permutations of this problem structure; others are addressed in (6,9,10,23–25).

In this work, we are interested solely in the problem of estimating the true signal magnitude from single-channel quadrature MRI measurements without assuming spatial smoothness—where only measurements from a single voxel are used in the estimation of the true signal magnitude at that voxel. We assume that we have taken one or more measurements of the pair (r, θ) for our voxel of interest, and that the unknown s , ϕ , and σ are constant for all measurements of our voxel. Our goal is to estimate s while ignoring ϕ and σ . We will represent our measurements via the measurement lists $\vec{r} = [r_0, \dots, r_n]$ and $\vec{\theta} = [\theta_0, \dots, \theta_n]$ or alternatively $\vec{a} = [a_0, \dots, a_n]$ and $\vec{b} = [b_0, \dots, b_n]$.

Additionally, although s and ϕ vary between voxels, we note that, for a given volume, σ is fixed for all measurements of all voxels as the thermal noise results from processes unrelated to the sampling location in k -space (15). This fact makes σ particularly easy to estimate. Any region of homogenous true signal magnitude (usually a region of air where it is known that $s = 0$) can be used to produce an effective estimate of σ . Alternatively, a series of noise-only

measurements can be quickly generated with the subject in the scanner by acquiring data without an excitation pulse or gradients. Based on this, many of the previous estimators have assumed that σ is effectively known, and remove it from their models’ parameters, treating it as a fixed, known value (1,4–6,8,9). In the remainder of this article, we will similarly treat σ as known.

We will use the notation $\hat{s}(\vec{r}, \vec{\theta})$ to represent an estimator for the unknown s at a voxel where the n -length list $(\vec{r}, \vec{\theta})$ of measurements was acquired. Two estimators are most-used for our problem. First,

$$\hat{s}_{\text{Mag}}(\vec{r}, \vec{\theta}) = \frac{1}{n} \sqrt{\left[\sum_{i=0}^{n-1} (r_i \cos \theta_i)\right]^2 + \left[\sum_{i=0}^{n-1} (r_i \sin \theta_i)\right]^2}, \quad [3]$$

is commonly called a “magnitude image.” This estimator averages the complex measurements and subsequently takes the magnitude of the average. An alternative formulation that also goes by the name “magnitude image” is given by first taking the magnitude of each measurement and then combining via sum-of-squares:

$$\hat{s}_{\text{MagAlt}}(\vec{r}) = \frac{1}{n} \sqrt{\sum_{i=0}^{n-1} r_i^2}. \quad [4]$$

It is important to note that these two estimators will not generate the same image and that the resulting “magnitude images” will have different noise distributions in general. Under the assumptions we have given for our problem, \hat{s}_{Mag} has a Rician distribution while \hat{s}_{MagAlt} has a noncentral- χ distribution with $2n$ degrees of freedom—these two estimators are therefore the same only when $n = 1$.

Most of the attention on bias-correction in previous MRI literature has been focused on the Rician distribution (i.e., images generated via \hat{s}_{Mag} or where $n = 1$). Henkelman suggested that the unbiased signal could be estimated by setting the expected-value function for the Rician distribution equal to the measured magnitude and solving for s (1,2). Using the notation $E[z]$ to represent the expected value of the random variable z , we can write this as

$$E[\hat{s}_{\text{Mag}}] = \hat{s}_{\text{Mag}}. \quad [5]$$

In statistical terminology, this is the simplest form of the method of moments. McGibney and Smith (4) and Miller and Joseph (5) independently also used the method of moments, but noted that the Rician’s second moment

$$E[(\hat{s}_{\text{Mag}})^2] = s^2 + 2\sigma^2 = (\hat{s}_{\text{Mag}})^2 \quad [6]$$

permits a simpler method of moments estimator:

$$\hat{s}_{\text{Mom2}}(\vec{r}, \vec{\theta}) = \sqrt{(\hat{s}_{\text{Mag}}(\vec{r}, \vec{\theta}))^2 - 2\sigma^2}. \quad [7]$$

This method has also previously been employed for the same statistical problem as it arises in other disciplines [e.g., (11,13,14)].

One complaint that has been leveled against \hat{s}_{Mom2} is that it can produce imaginary-valued estimates for the real-valued s (8). We have a different interpretation, suggested by estimation of s^2 in (11). We note that, as $(\hat{s}_{\text{Mag}})^2 - 2\sigma^2$ is an unbiased estimator of s^2 , when $s \simeq 0$ this equation must

sometimes take negative values. However, as we know that s^2 must be positive, there is another estimator that is guaranteed to have equal or smaller error: $\max[0, (\hat{s}_{\text{Mag}})^2 - 2\sigma]$, although it will now be biased. Taking the square root then gives us

$$\{\max[0, (\hat{s}_{\text{Mag}})^2 - 2\sigma]\}^{1/2} = \text{Re}\{[(\hat{s}_{\text{Mag}})^2 - 2\sigma]^{1/2}\}. \quad [8]$$

Thus, by simply taking the real part of $\hat{s}_{\text{Mom}2}$ (likely what the original authors intended), we are basing our result on an improved estimator of s^2 and producing a valid estimator for s .

A related estimator has been suggested by Gudbjartsson and Patz,

$$\hat{s}_{\text{CP}}(\vec{r}, \vec{\theta}) = \sqrt{|(\hat{s}_{\text{Mag}}(\vec{r}, \vec{\theta}))^2 - \sigma^2|}, \quad [9]$$

which is proposed to have a more gaussian-like distribution (6).

Sijbers et al. have considered several applications of the maximum likelihood method for deriving estimators in the context of MRI (8,9,26). Their conclusions, as relevant to our problem, can be summarized as: when given a pair of lists \vec{r} and $\vec{\theta}$ of quadrature MRI measurements from a single voxel meeting our stated assumptions, we should take \hat{s}_{Mag} as the estimate of s (9). Additional previous work on this problem has extended estimators to multiple channels (10), but for simplicity we will leave the channel-combination problem aside here.

In the ‘‘Theory’’ section, we consider the major methods that statisticians have suggested for deriving information-maximizing estimators (i.e., estimators that seek to get the most information about the true MRI signal magnitude based on the available measurements). Although maximizing the use of available information is theoretically appealing, it is also possible, as suggested by Bernstein et al. (3), that the signal-dependent, and thus spatially varying, bias is a significant perceptual issue in radiologists’ interpretation of MRI magnitude images. With this in mind, we continue the ‘‘Theory’’ section by studying the properties of estimators that seek explicitly to reduce bias. As part of this, we show that there is an unavoidable trade-off between bias and variance when estimating the true MR signal magnitude.

In the ‘‘Computing Estimator Metrics’’ section, we quantify all the previous estimators’ performance relative to this trade-off bound. In addition to quantifying mean squared error and bias [some of which we previously computed with a different method in (27)], we attempt to illuminate the bias/variance trade-off between these estimators via comparison with computed bounds on best-case bias or variance (28). The results of these comparisons are presented in the ‘‘Results’’ section. In the ‘‘Discussion’’ section, we summarize our conclusions generally and suggest measures other than bias that we believe are worth considering when evaluating image estimators.

THEORY

Choice of Distribution

Our approach to the problem of estimating the magnitude of the true MRI signal begins by selecting the distribution which we will use as the basis for our estimators. We begin

by collapsing the probability distribution for our multiple measurements (also called multiple excitations) into one distribution. We do this using the concept of a sufficient statistic. Informally, if we have a probability distribution $p(x, y; P)$, with x , and y being measurements and P being a parameter, the derived statistic $f(x, y)$ is sufficient for P if knowing just $f(x, y)$ tells us as much about P as knowing both x , and y [for further discussion of sufficiency see (29,30)].

For our problem, it is known that the optimal sufficient statistic is the complex average of the measurements (30). Thus, we define our sufficient statistics to be $A = \frac{1}{n} \sum_{i=1}^n a_i$, $B = \frac{1}{n} \sum_{i=1}^n b_i$, or equivalently $R = \sqrt{A^2 + B^2}$ and $\Theta = \arg(A, B)$, where \arg is the argument function (e.g., atan2 in the C programming language) with range from $-\pi$ to π . The distribution of A and B is then also binormal, but with an appropriately scaled standard deviation

$$p(A, B; s, \phi) = \frac{n}{2\pi\sigma^2} \exp\left(-\frac{n[(A - s \cos \phi)^2 + (B - s \sin \phi)^2]}{2\sigma^2}\right), \quad [10]$$

or in polar coordinates

$$p(R, \Theta; s, \phi) = \frac{Rn}{2\pi\sigma^2} \exp\left(-\frac{n[s^2 + R^2 - 2sR \cos(\Theta - \phi)]}{2\sigma^2}\right). \quad [11]$$

If we define $s' = \frac{s\sqrt{n}}{\sigma}$, $A' = \frac{A\sqrt{n}}{\sigma}$, and $B' = \frac{B\sqrt{n}}{\sigma}$, we can write

$$p(A', B'; s', \phi) = \frac{1}{2\pi} \exp\left(-\frac{(A' - s' \cos \phi)^2 + (B' - s' \sin \phi)^2}{2}\right), \quad [12]$$

and with $R' = \frac{R\sqrt{n}}{\sigma}$ we have

$$p(R', \Theta; s', \phi) = \frac{R'}{2\pi} \exp\left(-\frac{s'^2 + R'^2 - 2s'R' \cos(\Theta - \phi)}{2}\right). \quad [13]$$

This equation makes clear that the quantities marked with ‘‘’’ do not depend on any of the fixed parameters and so their estimation is the same regardless of the number of measurements or the standard deviation of the noise. For the remainder of this work we will operate on the standardized sufficient statistics A' , B' , R' , and Θ and their standardized binormal distributions in terms of s' and ϕ . It is useful to note that $\hat{s}'_{\text{Mag}} = R'$. More concretely, the remainder of this work will address the question of whether \hat{s}'_{Mag} can be improved by selecting a different function applied to complex-average of the measurements.

Information Maximizing Estimators

The majority of statistical prescriptions for deriving estimators proceed from the idea that the measurements contain information about the unknown parameters. Estimators are then suggested that optimally use the available information in the measurements to calculate an estimate of the

desired parameters. However, there are several theories in the statistics literature regarding how the informational relationship between the measurements and the unknown parameters should be turned into an estimator. We will consider several approaches from the statistical theory and derive the estimators that they consider optimal.

Maximum Likelihood Estimator

The basic mechanism of the maximum likelihood estimator (MLE) is to view the distribution as a function of the parameters with the measurements fixed (viewed this way, a distribution is renamed a likelihood), and maximize the resulting function. It is often the case that taking the logarithm of the likelihood function makes taking derivatives easier as having no effect on the location of maxima, and so the maximum log-likelihood is often substituted. In our problem, taking derivatives of the log-likelihood with respect to s' and ϕ and using the our sufficient statistics as our input data gives

$$\frac{\partial \ell(s', \phi; A', B')}{\partial s'} = A' \cos \phi + B' \sin \phi - s' \quad [14]$$

$$\frac{\partial \ell(s', \phi; A', B')}{\partial \phi} = s'(B' \cos \phi - A' \sin \phi), \quad [15]$$

where $\ell(s', \phi; A', B') = \log p(A', B'; s', \phi)$. Setting these equal to zero gives the maximum likelihood (ML) estimator, previously noted by Sijbers and den Dekker (9),

$$(\hat{s}'_{\text{ML}}, \hat{\phi}_{\text{ML}}) = (\sqrt{A'^2 + B'^2}, \arg(B'/A')) = (R', \Theta) \quad [16]$$

As our model has two parameters, the maximum likelihood estimator of our model is necessarily a pair of values. If we wish to take just the portion representing the estimate of s' alone, then we refer to ϕ as a nuisance parameter. Simply dropping the nuisance parameter in this way gives an estimator called the maximum profile likelihood estimate. More formally, it is justified by replacing the likelihood by the profile likelihood where $\hat{\phi}_{\text{ML}}$ is substituted for ϕ , giving (29).

$$\frac{d\ell_{\text{P}}(s'; A', B', \hat{\phi}_{\text{ML}})}{ds'} = \sqrt{A'^2 + B'^2} - s' \quad [17]$$

where we use ℓ_{P} to designate the log of the profile likelihood and move $\hat{\phi}_{\text{ML}}$ to the list of known parameters as necessitated by the profiling. Setting this equal to zero gives the maximum profile likelihood estimate

$$\hat{s}'_{\text{P}}(R') = R' \quad [18]$$

We can see from this that the maximum profile likelihood estimate of the true signal magnitude is the same as \hat{s}'_{Mag} .

Information-Based Nuisance Parameter Removal

Although profiling to remove nuisance parameters from the maximum likelihood estimator is a well-known practice, it is essentially ad hoc and several correction factors have been suggested by statisticians in attempts to improve the performance, in particular the bias, of the resulting estimators. We will briefly consider two famous examples of these corrections.

Although we omit details of the derivation here, the Bartlett-corrected profile likelihood is produced by introducing a correction factor Δ to the derivative of the profile likelihood function (31). We have previously shown (27) that for our problem, the Bartlett correction is given by

$$\begin{aligned} \Delta &= -\frac{1}{2s'^2} E \left[\frac{\partial}{\partial s'} \frac{\partial^2}{\partial \phi^2} \ell(s', \phi; A', B') \right] \\ &= \frac{1}{2s'}. \end{aligned} \quad [19]$$

Applying this correction gives us the equation

$$\frac{d\ell_{\text{P}}(s'; A', B')}{ds'} - \Delta = R' - s' - \frac{1}{2s'} = 0, \quad [20]$$

whose solution defines the Bartlett-corrected profile likelihood estimator for our problem

$$\hat{s}'_{\text{Corr}}(R') = \frac{R' + \sqrt{R'^2 - 2}}{2}. \quad [21]$$

We note that this estimator can produce complex-valued estimates of the real-valued s' when $R' < \sqrt{2}$. We also note that this estimator is simply the average of \hat{s}'_{Mag} and \hat{s}'_{Mom2} , and so we can use the same approach as with \hat{s}'_{Mom2} to force our estimates to be real-valued—in this case, the logic implies that when $R' < \sqrt{2}$, the estimator becomes $\hat{s}'_{\text{Corr}} = R'/2$.

\hat{s}'_{Corr} also results from application of another famous method of correction: the stably adjusted profile likelihood estimator proposes multiplying the profile likelihood by a factor $M(s')$ designed to approximate the likelihood function that would result if ϕ could be factored out of the probability density function (29,32). It can be shown that, for our problem, the estimator that results from maximizing the stably adjusted profile likelihood is identical to the one produced via Bartlett correction and so we will refer to the estimator in Eq. 21 simply as the maximum corrected profile likelihood estimator for the remainder of this article.

Maximum Marginal Likelihood Estimator

Noting that none of the previous estimators based on the profile likelihood involve the phase of the sufficient statistic Θ , it is tempting to discard the phase statistic Θ entirely from the model (called marginalizing the distribution) in Eq. 13 giving the Rician distribution

$$\begin{aligned} p(R'; s') &= \int_{-\pi}^{\pi} p(R', \Theta; s', \phi) d\Theta \\ &= R' \exp\left(-\frac{s'^2 + R'^2}{2}\right) I_0(s'R'), \end{aligned} \quad [22]$$

This marginalized model is especially more convenient for estimation as it also removes ϕ from the parameters; it leaves us with a single statistic R' and has a single parameter s' . Although it may be convenient, the critical issue is whether basing estimators on this model is correct. It has been a major point of debate in the statistics community exactly when such a marginalization is actually “information preserving,” and thus whether we should proceed to

use this model [for a survey of this debate, see (33)]. For the present, we leave aside arguments about when this reduction is justified. The maximum marginal likelihood estimator for our problem is the same as the maximum likelihood estimator for the Rician distribution. As noted above, this has been presented in several contexts (8,11,12). We will denote this estimate as \hat{s}'_{Marg} . Unfortunately, the estimator does not have a closed-form equation, but Sijbers et al. showed that $\hat{s}'_{\text{Marg}} = 0$ is the unique solution when $R'^2 \leq 2$ and there is a single positive solution that can be found numerically when $R'^2 > 2$, which allows us to certify that we have found the correct root using an efficient numerical algorithm (8).

Bayesian Estimators

There is a separate school of thought in statistics that, although also much-debated, provides a different set of rules for generating estimators beyond the likelihood-based ones. The Bayesian approach differs from likelihood-based methods in that Bayesian estimation requires us to specify our beliefs about values of the parameters before we see any measurement data; these beliefs are called prior distributions. In our problem, we would like to specify that the true magnitude and phase are unrelated, and all phases are equally likely to occur at any given point. Thus, we chose the prior $p(\phi) = 1/(2\pi)$ and

$$p(s') = \begin{cases} 1 & s' \geq 0 \\ 0 & \text{otherwise.} \end{cases} \quad [23]$$

The prior $p(s')$ is clearly not a well-defined probability density function, as it does not integrate to unity. However, it is common practice in Bayesian estimation to allow improper priors for parameters so that we can express equivalent probabilities of events over an infinite range, as in our case where R' is equally likely on the range $(0, \infty)$. Using these priors and Bayes' Theorem, we can write the posterior distribution

$$p(s', \phi; R', \Theta) = \frac{p(R', \Theta; s', \phi)}{p(R', \Theta)2\pi}. \quad [24]$$

If we desire to remove ϕ from the model, we can do this by simply marginalizing the posterior to remove ϕ , thus

$$p(s'; R', \Theta) = \int_{-\pi}^{\pi} p(s', \phi; R', \Theta) d\phi = \frac{p(R'; s')}{p(R', \Theta)2\pi}. \quad [25]$$

where $p(R'; s')$ is the Rician distribution as in Eq. 22.

Having derived the posterior distribution for our parameter of interest, we must now decide how to summarize it with a single estimate for s' . If we take the maximum of the posterior distribution (commonly called the maximum a posteriori or MAP estimate), we end up with the same result as the maximum marginalized likelihood estimator; formally, $\hat{s}'_{\text{MAP}} = \hat{s}'_{\text{Marg}}$.

As an alternative to the MAP, we could also consider the mean value of the parameter s' specified by this posterior. This choice is attractive because it minimizes the expected

squared error of the estimate given this distribution and data. We define this MB estimator as

$$\begin{aligned} \hat{s}'_{\text{MB}} &= \int_0^{\infty} s' p(s'; R', \Theta) ds' \\ &= \frac{\int_0^{\infty} s' p(R'; s') ds'}{p(R', \Theta)2\pi} \end{aligned} \quad [26]$$

Solving

$$\begin{aligned} p(R', \Theta) &= \int_0^{\infty} \int_{-\pi}^{\pi} p(R', \Theta; s', \phi) d\phi ds' \\ &= \frac{R'}{2\pi} \sqrt{\frac{\pi}{2}} \exp\left(-\frac{R'^2}{4}\right) I_0\left(\frac{R'^2}{4}\right), \end{aligned} \quad [27]$$

and

$$\int_0^{\infty} s' p(R'; s') ds' = R', \quad [28]$$

we can substitute these back into our estimator equation to get

$$\hat{s}'_{\text{MB}}(R') = \sqrt{\frac{2}{\pi}} \frac{\exp\left(\frac{R'^2}{4}\right)}{I_0\left(\frac{R'^2}{4}\right)}. \quad [29]$$

Estimators Minimizing Bias

As an alternative to focusing on maximizing the use of information in the measurements, it is often possible to construct estimators that optimize a chosen error metric. Given the focus in previous work on the detrimental effects of signal-dependent estimator bias, it seems natural to ask if we can construct an estimator of s' whose bias does not vary as s' varies. Let us call this constant-biased estimator \hat{s}'_{CB} . We will use $b_{s'}$ to denote the bias of the estimator \hat{s}' . Importantly, if we can produce a constant-bias estimator (with bias $b_{\hat{s}'_{\text{CB}}}$ everywhere), then we can produce an unbiased estimator by taking $\hat{s}'_{\text{CB}} - b_{\hat{s}'_{\text{CB}}}$. Thus, statements about the existence of a constant-bias estimator and an unbiased estimator are equivalent.

In particular, we will focus on estimators that ignore the phase of the sufficient statistic (i.e., radially symmetric estimators of the form $\hat{s}'(R')$) as all of the estimators derived previously for our problem ignore the sufficient statistic's phase—refer to Eqs. 3, 7, 9, 18, 21, and 29. A lengthier analysis, which can be found in the Appendix, is required for estimators of the form $\hat{s}'(R', \Theta)$, but we will leave this aside here as it is not necessary to understand the bias-variance trade-off for the cited estimators.

In general, we can study the relationship between an estimator's bias and variance using the Cramér-Rao bound. This a commonly used bound from statistics that links the variance of an estimator with the gradient of its bias and the second-derivative of the log-likelihood function, which is also known as the Fisher information. As R' is Rician-distributed, we are interested in the Fisher information for the Rician distribution, previously shown by Sijbers and den Dekker to be (9)

$$I(s') = Z(s') - s'^2, \quad [30]$$

where

$$Z(s') \equiv \int_0^\infty x^3 \frac{I_1(s'x)^2}{I_0(s'x)} \exp\left(-\frac{s'^2 + x^2}{2}\right) dx. \quad [31]$$

We then note that if an estimator $\hat{s}'(R')$ has bias gradient $\frac{d}{ds'} b_{\hat{s}'}(s') = \frac{dE[\hat{s}']}{ds'}(s') - 1$, we can write the Cramér-Rao bound for the estimator's variance (28)

$$\sigma_{s'}^2(s') \geq \left(1 + \frac{d}{ds'} b_{\hat{s}'}(s')\right)^2 I(s')^{-1}. \quad [32]$$

Studying Eq. 32, we can see that $\lim_{s' \rightarrow 0^+} I(s')^{-1} = \infty$, which implies that, if $\frac{d}{ds'} b_{\hat{s}'}(s') = 0$ for all $s' \neq 0$, we cannot put an upper bound on $\lim_{s' \rightarrow 0^+} \sigma_{s'}^2(s')$ in Eq. 32. Thus, there is no estimator for our problem with constant bias and bounded variance. This is a significant result as it explains why none of the estimators we show above are able to achieve an unbiased output: there is in fact no estimator of this form that can do so as maintaining a useful bound on variance.

Having shown the lack of a useful constant-bias estimator, we might naturally ask how much signal-dependent variation in the estimator's bias we must accept in order to bound the estimator's variance at some chosen v . Considering [32], we can have $\frac{d}{ds'} b_{\hat{s}'} = 0$ as $I(s')^{-1} \leq v$. Then, where $I(s')^{-1} > v$, we must decrease $\frac{d}{ds'} b_{\hat{s}'}$ proportionally to compensate for the growth of $I(s')^{-1}$. As $I(s')^{-1}$ is strictly decreasing with $\lim_{s' \rightarrow \infty} I(s')^{-1} = 1$, we know that

for $v > 1$, $I(s')^{-1} = v$ coincides with one value of s' , which we label s'_v . If $v \leq 1$, we will define $s'_v = 0$. We also know that the minimum absolute value of $\frac{d}{ds'} b_{\hat{s}'}$ is attained for $s' < s'_v$ by having equality in Eq. 32. Thus, we can write

$$\frac{d}{ds'} b_{\hat{s}'} = \begin{cases} 0 & s' > s'_v \\ \sqrt{\frac{v}{I(s')^{-1}}} - 1 & s' \leq s'_v \end{cases}. \quad [33]$$

This equation indicates that, irrespective of the value of v chosen, as $s' \rightarrow 0^+$ we must have $\frac{d}{ds'} b_{\hat{s}'} \rightarrow -1$. However, it also indicates that changes to v affect the rate at which $\frac{d}{ds'} b_{\hat{s}'} \rightarrow -1$ as $s' \rightarrow 0^+$; the effect is $O(\sqrt{v})$. Thus, if we want to halve the rate at which $\frac{d}{ds'} b_{\hat{s}'} \rightarrow -1$ as $s' \rightarrow 0^+$, we must pay for this with a quadrupling of the bound on variance v (a variation of this result is extended to all estimators $\hat{s}'(R', \Theta)$ in the Appendix). We will return to this bound when we evaluate our estimators in the ‘‘Results’’ section.

Visualization of Estimators

As these estimators are all functions from the real-valued R' to a real-valued estimate \hat{s}' , they are easy to visualize as filters by plotting the response (\hat{s}') as a function of the input (R'). In Fig. 1, we have performed such a plot to facilitate comparison of estimator behavior.

In Fig. 2, we show an illustrative example of the estimators' output given a synthetic input image corrupted with complex gaussian noise. We note that some of the

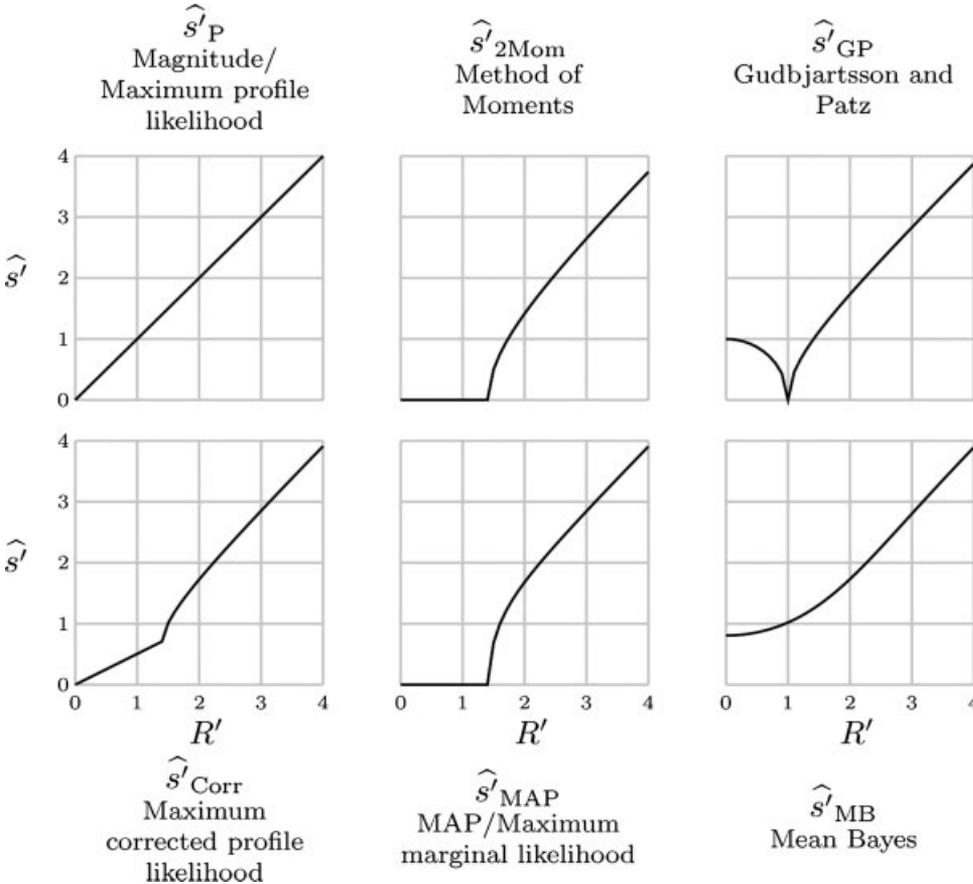


FIG. 1. Responses of the estimators. Each plot displays the estimate of s' produced by one estimator. The x-axis is the measured value R' , and the y-axis is the resulting estimate \hat{s}' .

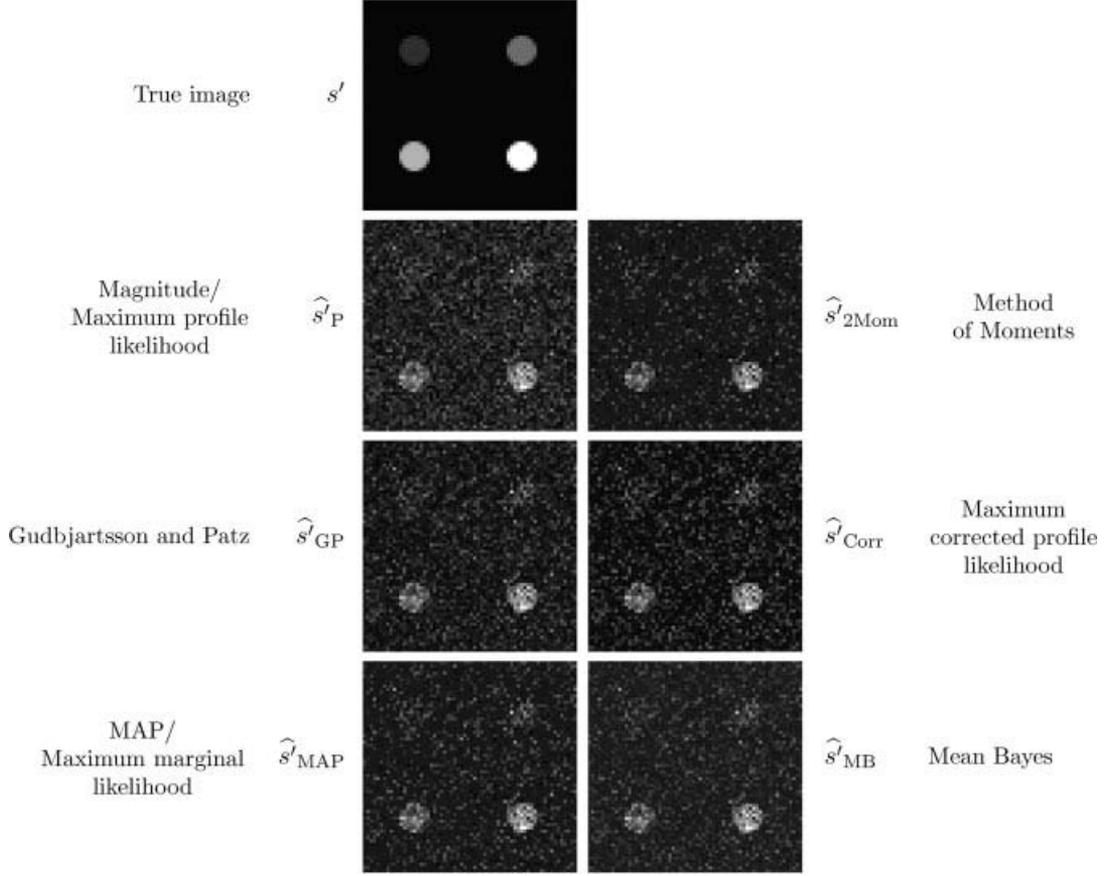


FIG. 2. Example of estimator output. Complex zero-mean gaussian noise with $\sigma = 1$ was added to the true image, and the resulting magnitude data was used to produce estimated images from the estimators discussed above. The four target circles have integer amplitudes from 1 to 4. The images are scaled independently so that each covers the entire dynamic range available for display.

estimators (e.g., the moment-based and MAP estimators) apply aggressive truncation at lower intensities of the magnitude signal, enhancing contrast between the brightest targets and the background, while simultaneously clipping regions of the fainter targets to zero.

COMPUTING ESTIMATOR METRICS

Computation of Estimator Bias and Variance

To compare these estimators, both to each other and to the bounds we outlined above, we computed the bias and variance of all the above estimators as a function of s' . The equations for the bias and variance of the maximum profile likelihood estimate are known (9,34):

$$b_{\hat{s}'_P}(s') = \sqrt{\frac{\pi}{2}} {}_1F_1\left(-\frac{1}{2}; 1; -\frac{s'^2}{2}\right) - s' \quad [34]$$

$$\sigma_{\hat{s}'_P}^2(s') = 2 + s'^2 - \frac{\pi}{2} {}_1F_1\left(-\frac{1}{2}; 1; -\frac{s'^2}{2}\right)^2, \quad [35]$$

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

For $\hat{s}'_{\text{Mom}2}$, we cannot find a closed form for the integral in the bias equation, and so simply write

$$b_{\hat{s}'_{\text{Mom}2}} = A(s') - s', \quad [36]$$

where we define $A(s') = \int_{\sqrt{2}}^{\infty} \sqrt{R'^2 - 2} R' \exp\left(-\frac{s'^2 + R'^2}{2}\right) I_0(s'R') dR'$. For the variance, we have

$$\sigma_{\hat{s}'_{\text{Mom}2}}^2(s') = \int_{\sqrt{2}}^{\infty} R'^3 \exp\left(-\frac{s'^2 + R'^2}{2}\right) I_0(s'R') dR' - 2Q(s', \sqrt{2}) - A(s')^2, \quad [37]$$

where $Q(s', \sqrt{2})$ is Marcum's Q -function (35). We solve these 1D integrals via numerical integration in Mathematica (36). We take a similar approach for the following estimators as well.

For the maximum corrected profile likelihood estimator, we find that the bias is

$$b_{\hat{s}'_{\text{Corr}}}(s') = b_{\hat{s}'_P}(s')/2 + b_{\hat{s}'_{\text{Mom}2}}(s')/2, \quad [38]$$

and the variance is given by

$$\begin{aligned} \sigma_{\hat{s}'_{\text{Corr}}}^2(s') &= \frac{\sigma_{\hat{s}'_P}^2(s') + \sigma_{\hat{s}'_{\text{Mom}2}}^2(s') - 2(b_{\hat{s}'_P}(s') + s')A(s')}{4} \\ &+ \frac{1}{2} \int_{\sqrt{2}}^{\infty} R'^2 \sqrt{R'^2 - 2} \exp\left(-\frac{s'^2 + R'^2}{2}\right) I_0(s'R') dR'. \end{aligned} \quad [39]$$

The Gudbjartsson and Patz estimator has bias

$$\begin{aligned} b_{\hat{s}'_{\text{GP}}}(s') &= \int_0^1 \sqrt{1-R'^2} R' \exp\left(-\frac{s'^2+R'^2}{2}\right) I_0(s'R') dR' \\ &\quad + \int_1^\infty \sqrt{R'^2-1} R' \exp\left(-\frac{s'^2+R'^2}{2}\right) I_0(s'R') dR' - s', \end{aligned} \quad [40]$$

and variance

$$\begin{aligned} \sigma_{\hat{s}'_{\text{GP}}}^2(s') &= 2 \int_0^1 (1-R'^2) R' \exp\left(-\frac{s'^2+R'^2}{2}\right) I_0(s'R') dR' \\ &\quad + s'^2 + 1 - [b_{\hat{s}'_{\text{GP}}}(s') + s']^2 \end{aligned} \quad [41]$$

The MB estimator's bias and variance are

$$\begin{aligned} b_{\hat{s}'_{\text{MB}}}(s') &= \sqrt{\frac{2}{\pi}} \int_0^\infty \frac{R'}{I_0\left(\frac{R'^2}{4}\right)} \exp\left(-\frac{R'^2+2s'^2}{4}\right) I_0(s'R') dR' - s' \end{aligned} \quad [42]$$

$$\begin{aligned} \sigma_{\hat{s}'_{\text{MB}}}^2(s') &= \frac{2}{\pi} \exp\left(-\frac{s'^2}{2}\right) \int_0^\infty \frac{R' I_0(s'R')}{\left[I_0\left(\frac{R'^2}{4}\right)\right]^2} dR' - [b_{\hat{s}'_{\text{MB}}}(s') + s']^2. \end{aligned} \quad [43]$$

Finally, we do not even have equations for the bias and variance of the MAP estimator. Instead, we rely on Monte Carlo estimates, using the sample mean and variance of the result to approximate the estimator bias and variance. We varied the underlying parameter s' from 0 to 4 at intervals of 0.1 and, for each choice of parameter value, used 5 million samples drawn from the distribution in Eq. 13.

Bias/Variance Trade-Off Analysis

The gap between the maximum profile likelihood estimator's bias gradient and variance and the computed bounds for these give some evidence of how much improvement to the estimator might be possible. In particular, a small gap indicates that switching to a different estimator is not likely an improvement, but instead representative of a different choice in the bias/variance trade-off. It is important to note that as the bounds we have given are not necessarily sharp, we should not draw the opposite conclusion—a large gap may not actually indicate that a better estimator is available, but simply that the bound is overly generous.

For all of our estimators except the MAP, we can compute the bias gradient from the bias equations given above. Those that contain integrals without closed forms can then also be solved numerically with satisfactory precision.

For the MAP estimator, we used the Monte Carlo method suggested by Hero et al. to estimate the gradient (28)

$$\begin{aligned} \nabla b_{\hat{s}'}(s') &\simeq \frac{1}{L-1} \sum_{i=1}^L \left(\hat{s}'(R_i) - \frac{1}{L} \sum_{j=1}^L \hat{s}'(R_j) \right) \\ &\quad \times \left(R_i \frac{I_1(R_i s')}{I_0(R_i s')} + s' \right) - 1, \end{aligned} \quad [44]$$

where L is the number of samples (5 million in our case) and R_i is the i th input sample generated for the Monte Carlo experiment.

RESULTS

To quantify the performance of these estimators, we plot their bias and mean squared error in Figs. 3 and 4, respectively. Considering these plots, we see that as the signal increases, the estimators' biases naturally tend toward zero. Similarly, as the estimators become increasingly unbiased, their mean squared error approaches the unbiased Cramér-Rao bound of 1. Thus, for large s' , all of these estimators are acceptable. The question, then, is if any of these estimators are preferable in low-SNR images or regions of images.

We can combine these results with our description of the bias/variance trade-off via the bounds we previously presented. In Fig. 5, we plot the difference between the true bias gradient at each point, and the bound on bias gradient implied by the true variance. This gap represents the most improvement in bias we could expect from any estimator with equivalent variance at a given value of s' . It is clear that the difference between estimators is, almost everywhere, larger than the gap between the observed bias and the bound. This indicates that choosing between the estimators is not simply a selection of the one with improved bias performance, but instead selecting different points along the bias/variance trade-off.

Similarly, in Fig. 6, we plot the difference between the variance at each point and the bound on variance implied by the true bias gradient. This allows us to visualize whether we could reduce our images' variance without increasing the signal-dependence of our bias by selecting any of these estimators. The smaller is this gap, the more confident we can be that we are operating close to the optimal variance implied by the observed bias gradient. It is clear from the mean squared error plots that the observed variance is far larger than the gap implied by the bound, thus we should be convinced that the choices between these estimators can be better described as selecting different bias/variance trade-offs, and not as a simultaneous improvement in all the relevant metrics.

DISCUSSION

We have surveyed the existing estimators for the problem of estimating the signal magnitude from the complex average of repeated measurements. We have made clear the links between previous results and the existing methods common in the statistical literature for these types of problems and have additionally introduced the MB estimator as

FIG. 3. Biases of the estimators. The x-axis is the true value s' ; the y-axis is the bias of the estimator computed directly or via Monte Carlo experiment.

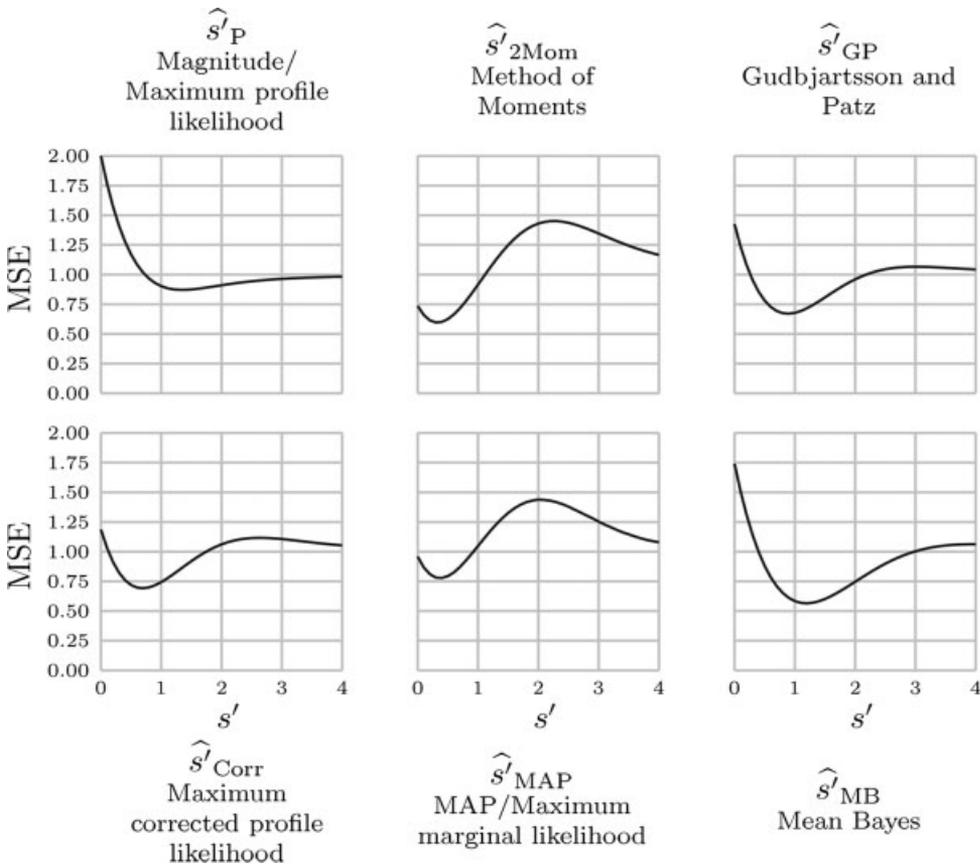
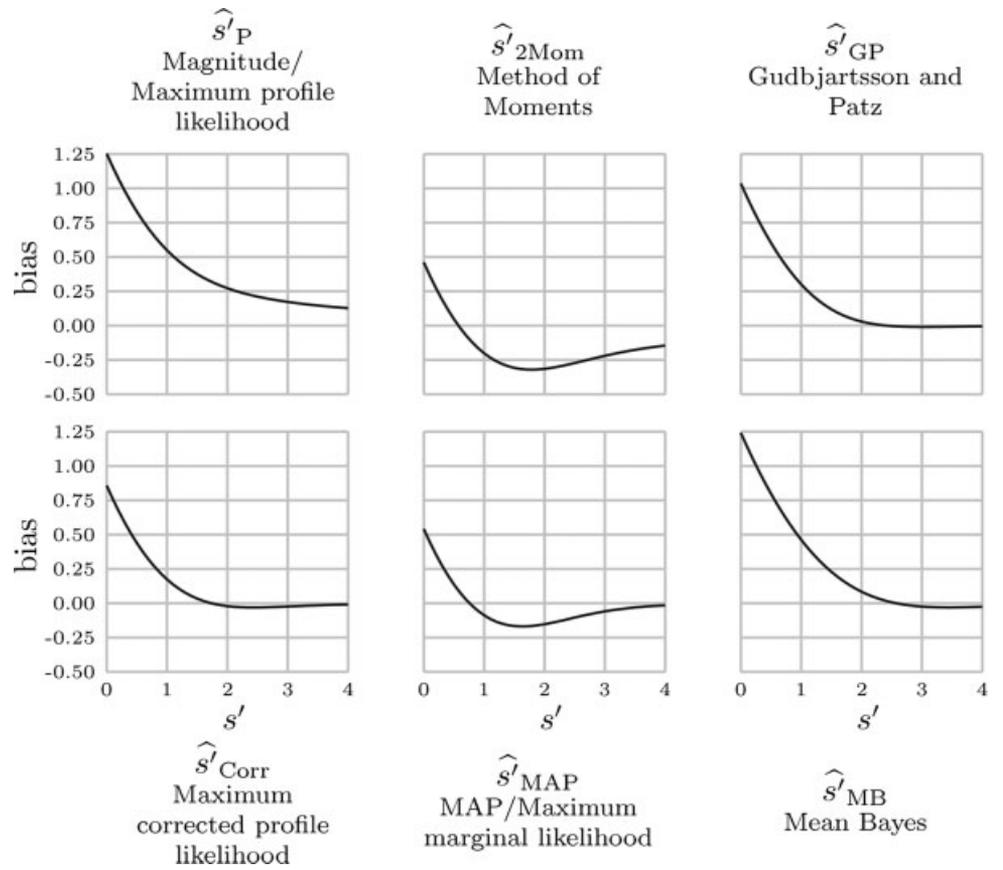


FIG. 4. Mean squared errors (MSEs) of the estimators. The x-axis is the true value s' ; the y-axis is the MSE of the estimator computed directly or via Monte Carlo experiment.

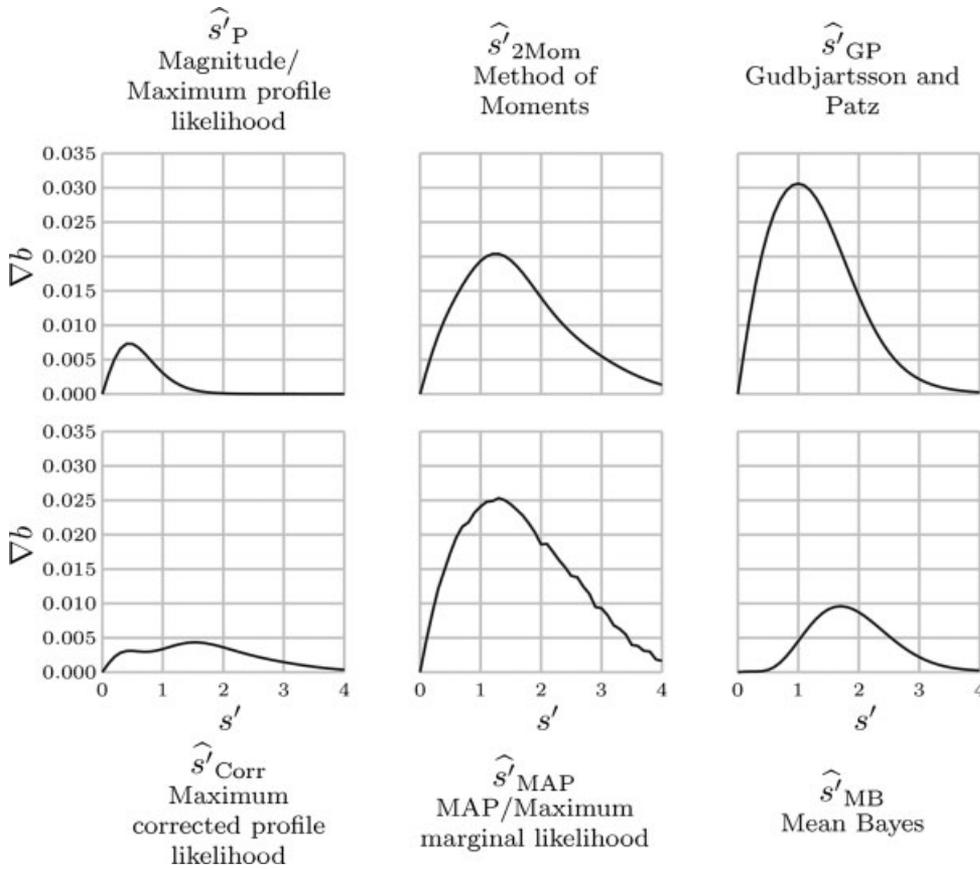


FIG. 5. Bounds on improvement of the estimators' bias gradients. The x-axis is the true value s' ; the y-axis is a bound on possible improvement of the bias gradient given the observed variance.

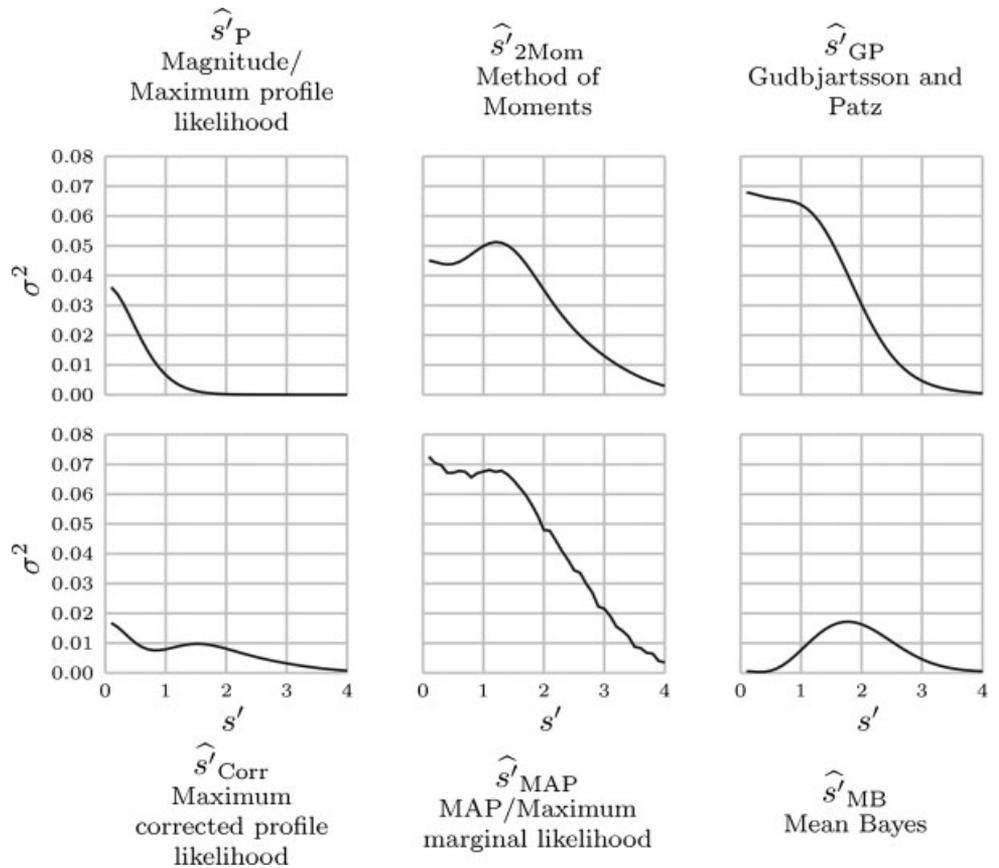


FIG. 6. Bounds on improvement of the estimators' variances. The x-axis is the true value s' ; the y-axis is a bound on possible improvement of the variance given the observed bias gradient.

yet another available estimator for this problem. We have demonstrated the important fact that, as an essential feature of this estimation problem, there is no unbiased estimator available with bounded variance. Finally, we have shown that there is an inevitable bias/variance trade-off that must be made in choosing estimators.

Our analysis of the estimators under various error metrics illustrate the bias/variance trade-off suggested by the theory we outlined above. Thus, of the available choices, we end up with the result that \hat{s}'_{Mag} represents a valid choice in the bias/variance trade-off. No other estimator, previously published or derived here, allows us to consistently improve both the bias and the variance across all values of s' , and so there is no obvious “free lunch” to be had by choosing a different estimator for image generation. Improvements beyond the trade-off must come, instead, from prior knowledge of the parameters (e.g., assuming local smoothness).

Having proved that there is no unbiased estimator with bounded variance available, and acknowledging an inevitable trade-off between bias and variance in any of the available estimators, we are left with the question of what trade-off should be preferred in practice. It is possible that a human observer might prefer the trade-off imposed by one of these estimators more than others for the distribution of s' commonly found in a clinical MRI. The answer to this question is likely complicated and task-dependent; we suggest that the literature on image perception provides a useful suite of methods for evaluating these estimators in the context of human observation (37–40). For quantitative analysis of MRI, we suggest that the methods outlined in this article be used to develop a range of estimators for the quantity of interest. The relative trade-off between bias and variance can then be determined for this new suite of estimators and a choice made based on the application domain. Additionally, care is required when performing quantitative analysis based on “magnitude images,” as the estimator used to generate these images can have a substantial effect on downstream analysis due to changes in the intensity distributions of the image voxels. Attempts to post-process images using estimators of the sort described here will normally require access to raw k -space measurement data, even for systems receiving with a single channel.

APPENDIX: EXTENSION OF THE CRAMÉR-RAO BOUND

We will now give an extension to the bound in [32] so that similar results exist for estimators taking the full (R', Θ) as arguments. In particular, we will show

$$\left(\frac{1}{2\pi} \int_0^{2\pi} \sigma_{\hat{s}'}(s', \phi) d\phi \right)^2 \geq \frac{\left(\frac{1}{2\pi} \int_0^{2\pi} \frac{d}{ds'} b_{\hat{s}'}(s', \phi) d\phi + 1 \right)^2}{I(s')}, \quad [\text{A1}]$$

where $\sigma_{\hat{s}'}(s', \phi)$ is the standard deviation and $b_{\hat{s}'}(s', \phi)$ the bias of an estimator $\hat{s}'(R', \Theta)$. This inequality shows that, for any choice of s' , if we replace the standard deviation and bias gradient from Eq. 32 with the means of each taken over ϕ , then the previous bound holds for all estimators $\hat{s}'(R', \Theta)$. As a result of this, the trade-off rate shown in Eq. 33 also holds for the means. This implies that an estimator can outperform the bound for certain choices of ϕ , but must pay

for this by underperforming at other ϕ so that it arrives at the bounded average result. When the estimator’s performance does not vary with ϕ at all, we recover the bound in Eq. 32.

For example, an estimator $\hat{s}'(R', \Theta) = R' \cos(\Theta)$ (i.e., the real component) is unbiased with bounded variance when $\phi = 0$, assuming negative estimates of magnitude are retained, but it substantially underperforms the bound when $\phi = \pi/2$. Practically, this is observed as intensity banding and drop-outs when an image is made using only the real component. The important result of this bound is that, as s' and ϕ are unknown and vary across the image volume, we cannot make an image that is unbiased for all voxels simultaneously without relying on methods which employ prior knowledge of s' or ϕ (e.g., via spatial smoothing).

We begin our proof of the bound with the Cauchy-Schwartz inequality, and simplify by noting that $E \left[\frac{d}{ds'} \ln p(R'; s') \right] = 0$,

$$\sigma_{\hat{s}'}(s', \phi) \geq \frac{E \left[\hat{s}'(R', \Theta) \frac{d}{ds'} \ln p(R'; s') \right]}{\sqrt{I(s')}} \geq -\sigma_{\hat{s}'}(s', \phi), \quad [\text{A2}]$$

where $I(s')$ is the Fisher information of the Rician distribution, shown in Eq. 30. Taking the mean over the range of ϕ gives

$$\begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} \sigma_{\hat{s}'}(s', \phi) d\phi &\geq \frac{\frac{1}{2\pi} \int_0^{2\pi} E \left[\hat{s}'(R', \Theta) \frac{d}{ds'} \ln p(R'; s') \right] d\phi}{\sqrt{I(s')}} \\ &\geq -\frac{1}{2\pi} \int_0^{2\pi} \sigma_{\hat{s}'}(s', \phi) d\phi. \end{aligned} \quad [\text{A3}]$$

Next, we simplify the middle of the inequality by noting that

$$\begin{aligned} &\frac{1}{2\pi} \int_0^{2\pi} \frac{d}{ds'} b_{\hat{s}'}(s', \phi) d\phi + 1 \\ &= \frac{1}{2\pi} \frac{d}{ds'} \int_0^{2\pi} E[\hat{s}'(R', \Theta)] d\phi \\ &= \frac{1}{2\pi} \frac{d}{ds'} \int_0^\infty \int_0^{2\pi} \left(\int_0^{2\pi} p(R', \Theta; s', \phi) d\phi \right) \\ &\quad \times \hat{s}'(R', \Theta) d\Theta dR'. \end{aligned} \quad [\text{A4}]$$

Then, using the identity $\int_0^{2\pi} p(R', \Theta; s, \phi) d\phi = p(R'; s')$ and simplifying via the chain rule we get

$$\begin{aligned} &\frac{1}{2\pi} \int_0^{2\pi} \frac{d}{ds'} b_{\hat{s}'}(s', \phi) d\phi + 1 \\ &= \frac{1}{2\pi} \int_0^{2\pi} E \left[\hat{s}'(R', \Theta) \frac{d}{ds'} \ln p(R'; s') \right] d\phi. \end{aligned} \quad [\text{A5}]$$

Substituting this back into [A3], we arrive at our final bound [A1] by squaring both sides.

REFERENCES

1. Henkelman RM. Measurement of signal intensities in the presence of noise in MR images. *Med Phys* 1985;12:232–233.
2. Henkelman RM. Erratum: measurement of signal intensities in the presence of noise in MR images [Med Phys 12, 232 (1985)]. *Med Phys* 1986;13:544.

3. Bernstein MA, Thomasson DM, Perman W. Improved detectability in low signal-to-noise ratio magnetic resonance images by means of a phase-corrected real reconstruction. *Med Phys* 1989;16:813–817.
4. McGibney G, Smith MR. An unbiased signal-to-noise ratio measure for magnetic resonance images. *Med Phys* 1993;20:1077–1078.
5. 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.
6. Gudbjartsson H, Patz S. The Rician distribution of noisy MRI data. *Magn Reson Med* 1995;36:910–914.
7. Gudbjartsson H, Patz S. Erratum: the Rician distribution of noisy MRI data. *Magn Reson Med* 1996;36:332.
8. Sijbers J, den Dekker AJ, Scheunders P, Van Dyck D. Maximum-likelihood estimation of Rician distribution parameters. *IEEE Trans Med Imaging* 1998;17:357–361.
9. Sijbers J, den Dekker AJ. Maximum likelihood estimation of signal amplitude and noise variance from MR data. *Magn Reson Med* 2004;51:586–594.
10. Koay CG, Basser PJ. Analytically exact correction scheme for signal extraction from noisy magnitude MR signals. *J Magn Reson* 2006;179:317–322.
11. Benedict TR, Soong TT. The joint estimation of signal and noise from the sum envelope. *IEEE Trans Inform Theory* 1967;13:447–454.
12. Stewart BG, Simmons JFL. Rician envelope estimation and confidence intervals in low signal/noise levels. *Electron Lett* 1987;23:832–834.
13. Talukdar KK, Lawing WD. Estimation of the parameters of the Rice distribution. *J Acoust Soc Am* 1991;89:1193–1197.
14. Tepedelenlioglu C, Abdi A, Giannakis GB. The Ricean K factor: estimation and performance analysis. *IEEE Trans Wireless Commun* 2003;2:799–810.
15. McVeigh ER, Henkelman RM, Bronskill MJ. Noise and filtration in magnetic resonance imaging. *Med Phys* 1985;12:586–591.
16. Jesmanowicz A, Hyde JS, Froncisz W, Kneeland JB. Noise correlation. *Magn Reson Med* 1991;20:36–47.
17. Macovski A. Noise in MRI. *Magn Reson Med* 1996;36:494–497.
18. Ahn CB, Cho ZH. A new phase correction method in NMR imaging based on autocorrelation and histogram analysis. *IEEE Trans Med Imaging* 1987;6:32–36.
19. Bernstein MA, Perman WH. Least-squares algorithm for phasing MR images. In: *Book of Abstracts Vol. 2, Sixth Annual Meeting of SMRM. Society of Magnetic Resonance in Medicine*; New York, NY; 1987. p 801.
20. Noll DC, Nishimura DG, Macovski A. Homodyne detection in magnetic resonance imaging. *IEEE Trans Med Imaging* 1991;10:154–163.
21. 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.
22. Bao P, Zhang L. Noise reduction for magnetic resonance images via adaptive multiscale products thresholding. *IEEE Trans Med Imaging* 2003;22:1089–1099.
23. Nowak RD. Wavelet-based Rician noise removal for magnetic resonance imaging. *IEEE Trans Image Process* 1999;8:1408–1418.
24. Cardenas-Blanco A, Nezamzadeh M, Footit C, Cameron IG. Accurate noise bias correction applied to individual pixels. In: *Proceedings of ISMRM, Berlin, Germany*; 2007.
25. Olariu E, Cardenas-Blanco A, Cameron IG. Noise bias correction for signal averaged images. In: *Proceedings of ISMRM, Honolulu, HI*; 2009.
26. Sijbers J, den Dekker AJ, Raman E, Van Dyck D. Parameter estimation for magnitude MR images. *Int J Imaging Syst Technol* 1999;10:109–114.
27. Tisdall MD, Atkins MS, Lockhart RA. Maximum likelihood estimators in magnetic resonance imaging. In: *Information Processing in Medical Imaging, Vol. 4584 Lecture Notes in Computer Science. Karssemeijer N, Lelieveldt B, eds.; Kerkrade, The Netherlands: Springer*; 2007. pp 434–445.
28. Hero AO, Fessler JA, Usman M. Exploring estimator bias-variance tradeoffs using the uniform CR bound. *IEEE Trans Signal Process* 1996;44:2026–2041.
29. Barndorff-Nielsen OE, Cox DR. *Inference and asymptotics. Monographs on statistics and probability.* Chapman and Hall: New York, NY; 1994.
30. Lehmann EL, Casella G. *Theory of point estimation*, 2nd ed. Springer: New York, NY; 1998.
31. Bartlett MS. Approximate confidence intervals. III. A bias correction. *Biometrika* 1955;42:201–204.
32. Barndorff-Nielsen OE. Adjusted versions of profile likelihood and directed likelihood, and extended likelihood. *J R Stat Soc B* 1994;56:125–140.
33. Basu D. On the elimination of nuisance parameters. *J Am Stat Assoc* 1977;72:355–366.
34. Rice SO. *Mathematical analysis of random noise.* In: *Selected papers on stochastic processes, Vol. 23.* Wax N, ed.; Dover: New York, NY; 1954. pp 133–294.
35. Marcum J. A statistical theory of target detection by pulsed radar. *IRE Trans Inf Theory* 1960;6:59–267.
36. Wolfram Research Inc. *Mathematica, Version 7.0.* Wolfram Research Inc.: Champaign, IL; 2008.
37. Salem KA, Lewin JS, Aschoff AJ, Duerk JL, Wilson DL. Validation of a human vision model for image quality evaluation of fast interventional magnetic resonance imaging. *J Electron Imaging* 2002;11:224–235.
38. Tisdall MD, Atkins MS. Using human and model performance to compare MRI reconstructions. *IEEE Trans Med Imaging* 2006;25:1510–1517.
39. Miao J, Huo D, Wilson D. Perceptual difference model (Case-PDM) for evaluation of MR images: validation and calibration. In: *Proceedings of the SPIE Medical Imaging 2007, Vol. 6515 SPIE*; San Diego, CA; 2007. p 15.
40. Tisdall MD, Atkins MS. Perception of dim targets on dark backgrounds in MRI. In: *Proceedings of the SPIE Medical Imaging 2007, Vol. 6515 SPIE*; San Diego, CA; 2007. p 13.