Optimal decay rate constant estimates from phased array data utilizing joint Bayesian analysis

James D. Quirk [a,*], Alexander L. Sukstanskii [a], G. Larry Bretthorst [a], Dmitriy A. Yablonskiy [a,b]

[a] Mallinckrodt Institute of Radiology, Washington University School of Medicine, 4525 Scott Avenue, Campus Box 8227, St. Louis, MO 63110, USA
[b] Department of Physics, Washington University, St. Louis, MO, USA

A B S T R A C T

Since their initial description, phased array coils have become increasingly popular due to their ease of customization for various applications. Numerous methods for combining data from individual channels have been proposed that attempt to optimize the SNR of the resultant images. One issue that has received comparatively little attention is how to apply these combination techniques to a series of images obtained from phased array coils that are then analyzed to produce quantitative estimates of tissue parameters. Herein, instead of the typical goal of maximizing the SNR in a single image, we are interested in maximizing the accuracy and precision of parameter estimates that are obtained from a series of such images. Our results demonstrate that a joint Bayesian analysis offers a "worry free" method for obtaining optimal parameter estimates from data generated by multiple coils (channels) from a single object (source). We also compare the properties of common channel combination techniques under different conditions to the results obtained from the joint Bayesian analysis. If the noise variance is constant for all channels, a sensitivity weighted average provides parameter estimates equivalent to the joint analysis. If both the noise variance and signal intensity are similar in all channels, a simple channel average gives an adequate result. However, if the noise variance differs between channels, an "ideal weighted" approach should be applied, where data are combined after weighting by the channel amplitude divided by the noise variance. Only this "ideal weighting" provides results similar to the automatic-weighting inherent in the joint Bayesian approach.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

Since their early description [1], phased array coils have become increasingly widespread due to their ease of customization for various applications. Their recent surge in popularity can be traced to improvements in coil technologies and the development of rapid imaging techniques that utilize the spatial information from the phased array coils to decrease acquisition times [2,3]. However, these coils are also used in more traditional imaging experiments simply for their flexibility and increased signal-to-noise ratio (SNR).

Numerous investigators [1,4–9] have attempted to optimize the SNR of images from phased array coils and proposed various techniques for combining such data. Roemer [1,5] and others [6,7,9] have suggested that the sum-of-squares (SOS) combination provides a near optimal signal-to-noise ratio in the reconstructed image, approaching that of a reconstruction using the "correct" channel sensitivity profile without requiring additional acquisitions. Others have refined this technique by weighting the channels using more complex factors that reduce the impact of local

signal and noise fluctuations and more accurately characterize the true coil response profiles [6,8,10–12]. These weighting factors are commonly obtained from either a smoothed version of the data itself or a separately acquired lower-resolution image.

One issue that has received relatively little attention is how the different channel combination techniques affect parameter estimates obtained from modeling the signal in a series of images [11,13]. Here, we are interested in maximizing the accuracy and precision of parameter estimates that are obtained from a series of array coil images. In addition to SNR optimization, this imposes the additional constraint of accurately preserving the relationships between the series images.

We also explore the effects of various channel combination methods on the accuracy and precision of parameter estimates and examine the case where the common assumption of equal noise power across channels is violated. In the simplest method for combining channels, the channel average or sum, all channels are treated equally. When signals of differing SNR are averaged, information from the high SNR channels are diluted by the lower SNR channels, resulting in less accurate and less precise parameter estimates [14]. Such variations in signal and noise power across channels are common in real imaging experiments as the array ele-

ments are seldom equidistant from a particular region of interest. If there are systematic effects in the data that are not properly modeled (e.g. the Rican noise profile induced by processing the magnitude images from each channel), they can coherently combine and further distort the parameter estimates. Channels may also experience different loading due to their placement on different parts of the sample or patient, and will have some degree of coupling. Thus, the simple averaging of channels is rarely advisable.

Sensitivity-weighted averaging attempts to mitigate these effects by accounting for spatial variations in signal intensity for each channel. However, these methods do not account for variations in noise power between channels and can still magnify systematic effects and artifacts in the data. If the channel weighting factors for each image in a series is derived from its own intensity, these weighting factors will differ across the series images, potentially biasing the parameter estimates. An extreme case of this occurs with the sum-of-squares (SOS) combination of images. While reported to provide a "near-optimal" SNR, the SOS combination also artificially distorts the relationships between images in a series by forcing all low SNR points to take positive values, introducing a DC offset that coherently combines across channels. For an exponential decay model, this increased noise floor produces a systematic underestimation of the decay rate constant and will affect the accuracy and precision of rate constant estimation, as previously described [6,15–20].

As an alternative to determining the optimal channel weighting factors for a given experimental setup, we could also analyze multi-channel data without signal combination. The signals from the various channels can be jointly analyzed with a model that allows the channel-specific properties (such as signal amplitude and noise power) to vary across channels while requiring the MRI properties inherent to the imaged object (such as a decay rate constant) to be identical for all channels. We have implemented this framework using Bayesian probability theory and demonstrate its benefits for modeling simulated multi-channel data compared to more traditional combination techniques. For simplicity, we consider here only the mono-exponential decay model prevalent in MR, but these general principles have an obvious extension to more complex estimation problems. We conclude that a joint Bayesian analysis offers a "worry free" method for obtaining optimal parameter estimates from multi-channel data.

2. Theory

A ubiquitous model in MRI experiments is the mono-exponential decay. For simplicity of the analysis below, we assume a simple mono-exponential analysis without a constant, such as in T_2 or diffusion measurements. For an array of M-channels used to acquire decay measurements at N different times, the measured signal can be expressed as:

$$S_m(t_n) = A_m \exp(-R\,t_n) + \eta_m(t_n), \quad (1)$$

where $S_m(t_n)$ is the signal measured on the mth channel at the nth sampling time (or b-value in the case of the diffusion experiment), $\eta_m(t_n)$ is the noise in the mth channel at the nth sampling time t_n, A_m is the signal amplitude in channel m, and R is a rate constant (e.g. R_2 or ADC). While we assume for simplicity that the channels are sampled simultaneously, which is typically the case, and that there is no coupling between coils, no other assumptions are made as to the distribution of data samples in time. The rate constant is treated as an inherent property of the sample, and is therefore independent of which channel is performing the measurement, whereas the signal amplitude and the noise are properties of each channel.

NMR/MRI scanners typically produce a complex signal (quadrature detection), and the real and imaginary components are com-

monly combined to produce a magnitude signal or image. For our purposes, the complex signal from each channel is assumed to have been "phased", i.e. the coherent signal moved entirely to the real channel, and only the real signals are analyzed [21,22]. This produces an improvement in SNR and removes the bias introduced by using magnitude images. An alternative would be the simultaneous analysis of the real and imaginary components from all channels. However, as this would complicate the model by introducing an additional amplitude (or phase) for each channel, we will assume for simplicity that the data have already been phased.

In Bayesian analysis, we are interested in calculating $p(AR|D\sigma I)$, the joint posterior probability of the model parameters A and R given the data, D, the standard deviation of the noise prior probability, σ, and the prior information, I. Using Bayes' theorem and the product rule, omitting constant terms that will cancel upon normalization, and assuming independence in our prior knowledge of A, R, and σ, this can be expressed as

$$p(AR|D\sigma I) \propto p(R|I)p(A|I)p(D|AR\sigma I). \quad (2)$$

In this equation, $p(A|I)$ and $p(R|I)$ are the prior probabilities of the parameters A and R, and represent what is known about the possible values of these parameters before acquiring the data; $p(D|AR\sigma I)$ is the direct probability of the data given the parameters and is proportional to a likelihood function.

Initially, we consider the signal generated by a single channel and calculate the expected uncertainty in the resultant parameter estimates. Using uniform and comparatively non-informative priors, the joint posterior probability of the model parameters in Eq. (2) can be expressed as [14,23,24]

$$p(AR|D\sigma I) \propto \exp\left(-\frac{Q}{2\sigma^2}\right); \quad Q = \sum_{n=0}^{N-1} [D(t_n) - A\exp(-Rt_n)]^2. \quad (3)$$

In the majority of exponential decay experiments, the actual value of the amplitude parameter is of little interest and we are primarily concerned with estimation of the rate constants. In such cases, the amplitudes can be removed from the analysis by calculating the marginal probability for the decay rate constant, R. This requires integrating Eq. (3) over all possible values of A:

$$p(R|D\sigma I) \propto \int dA\, p(AR|D\sigma I) \propto \int dA\, \exp\left(-\frac{Q}{2\sigma^2}\right). \quad (4)$$

Assuming high SNR, that the data are sampled at uniformly spaced times, and that the data are acquired until the exponential decays into the noise, the uncertainty in the decay rate constant estimate for a single channel was previously estimated as the standard deviation of the posterior probability distribution for parameter R, $\sqrt{8\hat{R}^3\Delta t}/\text{SNR}$, where Δt is the sampling interval between data points and $\hat{R}$ is the true value of the rate constant [25].

To broaden the applicability of this result to imaging experiments, here we relax the assumptions of uniform sampling and acquiring data until the signal decays into the noise. In the next section, we will also expand this result to the multi-channel case. To analytically evaluate Q in Eq. (4), the data are expressed in terms of $\hat{A}$, $\hat{R}$, and $\hat{\eta}(t_n)$, the true values of the amplitude, rate constant, and noise:

$$D(t_n) = \hat{A}\exp(-\hat{R}t_n) + \hat{\eta}(t_n). \quad (5)$$

In the high SNR approximation, which allows us to neglect the noise term from Eq. (5), evaluation of the amplitude integral in Eq. (4) and omitting constant terms produces the marginal probability for the rate constant in the form:

$$p(R|D\sigma I) \propto \exp\left(\frac{(\mathbf{d}\cdot\mathbf{G})^2}{2\sigma^2(\mathbf{G}\cdot\mathbf{G})}\right), \quad (6)$$

where $\mathbf{d}$ is a vector of data values acquired at the N sampling times and the elements of the model vector, $\mathbf{G} = \{G_n\}$, are defined as $G_n = \exp(-Rt_n)$.

Note that the specific time dependence of the signal determines only the form of the vectors $\mathbf{d}$ and $\mathbf{G}$. The basic structure of Eq. (6) will occur with any signal model containing a single marginalized amplitude, not just for the mono-exponential model considered here.

In the high SNR regime, the marginal probability for the decay rate constant R in Eq. (6) has a maximum at $R = \hat{R}$. Using a second-order Taylor series expansion about $\hat{R}$, we obtain:

$$p(R|D, \sigma, I) \propto \exp\left(-\frac{(R - \hat{R})^2}{2\sigma_R^2}\right), \tag{7}$$

where

$$\sigma_R = \frac{\hat{R}}{\mathrm{SNR}} F(\hat{R}, \mathbf{t}), \quad \mathbf{t} = \{t_1, t_2, \ldots, t_n\},$$
$$F(\hat{R}, \mathbf{t}) = \left[\frac{\Sigma_0}{\left(\Sigma_0 \Sigma_2 - \Sigma_1^2\right)}\right]^{1/2}, \quad \Sigma_j = \sum_{n=0}^{N-1} (\hat{R} t_n)^j \exp(-2\hat{R} t_n). \tag{8}$$

Here σ_R is the standard deviation of the posterior probability distribution for parameter R, effectively representing the predicted uncertainty in the parameter estimate, and $\mathrm{SNR} = \hat{A}/\sigma$ is the signal-to-noise ratio for the un-attenuated signal ($t = 0$).

As expressed in Eqs. (7) and (8), estimation of the decay rate constant and calculation of the uncertainty in its estimation, σ_R, require explicit knowledge of the noise standard deviation for each channel. This dependence can be removed by integrating over all possible values of the noise standard deviation, analogous to what was done with the signal amplitude above. While this approach is generally useful, if the number of sampling times is small then the data contains only minimal information about the standard deviation of the noise and this extra degree of freedom in the calculation would produce a significantly greater uncertainty in the estimate of the rate constant compared to when the noise standard deviation is explicitly included. For simplicity, we will utilize only two sampling times in the simulations below and therefore will assume that the noise standard deviation is accurately known for each channel and insert these values into Eqs. (7) and (8). Analysis of simulated data after marginalization of the noise standard deviation produced larger but qualitatively similar parameter uncertainties (data not shown).

2.1. Joint analysis

The uncertainty in the decay rate constant from the joint analysis of data simultaneously measured from M independent channels can be similarly derived. Starting from Eq. (2), we now allow $\mathbf{A}$, $\boldsymbol{\sigma}$, and $\mathbf{D}$ to symbolize the set of amplitudes, noise standard deviations, and data for all channels, $\mathbf{A} = \{A_m\}$, $\mathbf{D} = \{D_m\}$, and $\boldsymbol{\sigma} = \{\sigma_m\}$. Assuming that the amplitude and noise from the different channels are effectively independent, i.e. the noise is uncorrelated between channels, each term in Eq. (2) can be expanded as the product of the probabilities for the individual channels. If we again use uniform and comparatively non-informative priors, this produces:

$$p(\mathbf{A}R|\mathbf{D}\boldsymbol{\sigma}I) \equiv \prod_m P(A_m R|D_m \sigma_m I) \propto \exp\left(-\frac{1}{2}\sum_{m=1}^{M} \frac{Q_m}{\sigma_m^2}\right)$$
$$Q_m = \sum_{n=0}^{N-1} [D_m(t_n) - S_m(t_n)]^2, \tag{9}$$

analogous to Eq. (3). The marginal posterior probability for R is obtained by integrating the joint posterior probability in Eq. (9) over all of the amplitudes, $\mathbf{A}$, producing:

$$p(R|D\sigma I) \propto \prod_m \exp\left(\frac{(\mathbf{d} \cdot \mathbf{G})^2}{2\sigma^2(\mathbf{G} \cdot \mathbf{G})}\right). \tag{10}$$

To obtain an analytic expression for the rate constant uncertainty, we again utilize a Taylor series expansion about the rate constant, producing:

$$p(R|D\boldsymbol{\sigma}I) \propto \exp\left(-\frac{(R - \hat{R})^2}{2\left(\sigma_R^{\mathrm{joint}}\right)^2}\right), \tag{11}$$

where

$$\sigma_R^{\mathrm{joint}} = \frac{\hat{R}}{\mathrm{SNR}_{\mathrm{eff}}^{\mathrm{joint}}} F(\hat{R}, \mathbf{t}). \tag{12}$$

Here, the function $F(\hat{R}, \mathbf{t})$ is still defined by Eq. (8) and $\mathrm{SNR}_{\mathrm{eff}}^{\mathrm{joint}}$ is the effective signal-to-noise ratio for the jointly analyzed data, given by:

$$\mathrm{SNR}_{\mathrm{eff}}^{\mathrm{joint}} = \left(\sum_{m=1}^{M} \mathrm{SNR}_m^2\right)^{1/2}, \quad \mathrm{SNR}_m = \frac{\hat{A}_m}{\sigma_m}. \tag{13}$$

When the noise level is identical for all the channels, $\sigma_m = \sigma$, the effective SNR reduces to

$$\mathrm{SNR}_{\mathrm{eff}}^{\mathrm{joint}} = \frac{(\overline{A^2})^{1/2}}{\sigma} \cdot \sqrt{M}, \quad \overline{A^2} = \frac{1}{M}\sum_{m=1}^{M} \hat{A}_m^2. \tag{14}$$

Further, if the signals in all channels are characterized by the same SNR, then $\mathrm{SNR}_{\mathrm{eff}}^{\mathrm{joint}} = \mathrm{SNR} \cdot \sqrt{M}$ and Eq. (12) reduces to

$$\sigma_R^{\mathrm{joint}} = \frac{\hat{R}}{\mathrm{SNR} \cdot \sqrt{M}} \cdot F(\hat{R}, \mathbf{t}) \tag{15}$$

i.e. when all channels have identical SNR, the uncertainty in the parameter estimate for R decreases as the square root of the number of channels, as expected.

2.2. Weighted average analysis

For comparison, we explore an alternative where the signals from different channels are combined prior to analysis. The most general method is to "weight" the signal from each channel by a channel-specific factor, λ_m, and generate the average weighted signal,

$$D^{WAv}(t_n) = \frac{1}{M}\sum_{m=1}^{M} \lambda_m D_m(t_n). \tag{16}$$

For an un-weighted average, $\lambda_m = 1$; for sensitivity weighting, $\lambda_m = \hat{A}_m$, which is approximated here for each channel by its signal in the un-weighted ($t = 0$) image. The noise standard deviation in each channel is also weighted by the same factor: $\sigma_m \rightarrow \lambda_m \sigma_m$. If the noise is uncorrelated across channels, then the effective noise in the combined data is equal to $\sigma^{WAv} = (\sum_{m=1}^{M} \lambda_m^2 \sigma_m^2)^{1/2}/M$. Here, we assume that for each channel, all N images in the series are weighted using the same channel-specific factor.

With joint analysis, "weighting" the channels does not alter the effective SNR in Eq. (13) as the scaling factors for the amplitude and noise standard deviation cancel. However, if the signals are averaged after scaling, as in Eq. (16), then the uncertainty in Eq. (12) becomes

$$\sigma_R^{WAv} = \frac{\hat{R}}{SNR_{eff}^{WAv}} F(\hat{R}, \mathbf{t}), \tag{17}$$

where the effective SNR, SNR_{eff}^{WAv} is:

$$SNR_{eff}^{WAv} = \frac{\sum_{m=1}^{M} \lambda_m \hat{A}_m}{\left(\sum_{m=1}^{M} \lambda_m^2 \sigma_m^2\right)^{1/2}}. \tag{18}$$

To compare the effective SNR from the joint analysis, Eq. (13), to the effective SNR from weighted averaging, Eq. (18), it is convenient to introduce the variables $x_m = \hat{A}_m/\sigma_m$ and $y_m = \lambda_m \sigma_m$, allowing us to rewrite the ratio of the two effective SNRs as:

$$\frac{SNR_{eff}^{WAv}}{SNR_{eff}^{joint}} = \frac{\sum_{m=1}^{M} x_m y_m}{\left(\sum_{m=1}^{M} x_m^2 \cdot \sum_{m=1}^{M} y_m^2\right)^{1/2}}. \tag{19}$$

According to the Cauchy inequality, the ratio in the right-hand side of Eq. (19) is always less than or equal to 1. Therefore, SNR_{eff}^{WAv} from weighted averaging cannot exceed SNR_{eff}^{joint} from the joint analysis:

$$SNR_{eff}^{WAv} \leqslant SNR_{eff}^{joint}. \tag{20}$$

The SNRs in Eq. (20) coincide if and only if $y_m = \gamma x_m$ for all m, where γ is an arbitrary positive coefficient. Thus, when combining channels using weighted averaging, the maximum SNR_{eff}^{WAv} and, therefore the best estimate of the parameter $\hat{R}$, can be achieved using the weighting factors:

$$\lambda_m = \gamma \hat{A}_m/\sigma_m^2. \tag{21}$$

The coefficient γ is a common scaling factor which cancels from the expression for SNR_{eff}^{WAv} and can therefore be set equal to 1. It is easy to verify that SNR_{eff}^{WAv} in Eq. (18) with λ_m from Eq. (21) exactly coincides with SNR_{eff}^{joint} in Eq. (13) obtained by the joint analysis. Hence, Eq. (21) provides an optimal weighting algorithm.

It should be noted, however, that obtaining a pixel-wise estimate of the signal amplitude and noise levels for each channel is not always practical, especially for low SNR cases, and any errors in these estimates would propagate into bias and/or increased uncertainties in the parameter estimation. The commonly used sensitivity weighting, $\lambda_m = \hat{A}_m$, provides the optimal weighting, with effective SNR equivalent to the joint analysis, if and only if the noise standard deviation is the same in all the channels, $\sigma_m = \sigma$. The effective SNR using the un-weighted average, $\lambda_m = 1$, is always lower than the effective SNR from the joint analysis, except for the trivial case when the signal amplitudes and noise levels on each channel are identical.

2.3. Sum-of-squares

The most commonly utilized technique for processing multi-channel data (and the default option on many MRI scanners), involves combining the magnitude data from each channel by calculating the square root of the sum-of-squares (SOS):

$$D^{SOS}(t_n) = \left[\sum_{m=1}^{M} |D_m(t_n)|^2\right]^{1/2} \tag{22}$$

where $|D_m(t_n)| = \sqrt{(\mathbf{Re}\, D_m(t_n))^2 + (\mathbf{Im}\, D_m(t_n))^2}$ is the magnitude of the complex data.

It is important to note the significant differences between SOS and the other combination techniques outlined above. In the SOS analysis, the channel weighting factors vary for each image in a series, potentially distorting the decay behavior. If we consider the SOS as a special case of the weighted average, then each data point is effectively used as its own weighting factor (sensitivity map). Since the SNR of the images decreases as t_n increases, the

sensitivity map estimate is necessarily of lower quality at later time points. Thus, we would expect parameter estimates from this method to be less accurate than those from weighting the images by a constant sensitivity map derived from the highest SNR image (or a high SNR reference image). As the sum-of-squares forces all data values to be positive, it also introduces a well-known DC offset that can significantly distort the decay behavior and introduce a substantial, sampling time-dependent bias in the estimate of the rate constant. Although not commonly performed, the sum-of-squares across channels can also be calculated using only the real images from each channel, after phasing. We do so here, in order to maintain a fair comparison between combination methods, with the understanding that the additional use of magnitude data would be expected to worsen the accuracy of the parameter estimates.

At present, the above theoretical analysis of the parameter uncertainty cannot be directly applied to SOS data as the channel weighting factors, and therefore the noise power, change for each time point. Incorporation of these effects would require a substantial expansion of the above theory and the mathematical form of these expressions would be significantly less tractable. However, we can still examine the accuracy of the parameter estimates produced by the analysis of SOS data as the value of the standard deviation of the noise prior probability, σ, does not affect the location of the maximum in Eq. (4), though the actual noise values do.

While the bias inherent in obtaining parameter estimates from magnitude data in a single channel system is widely known, few have considered the impact of the SOS combination of array images on parameter estimation [11]. Below, we analyze simulated data to compare the bias in the decay rate constant obtained by the SOS approach with the alternate approaches developed in the current manuscript.

3. Computer simulations

The theoretical equations derived above provide a lower-bound estimate of the expected parameter uncertainties for experiments using arbitrary signal combination techniques. However, due to its approximations, this theory may lose accuracy with low SNR data. To validate the theory and establish its accuracy, we analyzed computer simulated data and compare the uncertainties in the parameter estimation to the theoretically predicted values. This approach also measures the bias of the resulting parameter estimates at various SNR levels.

Simulated data were generated and analyzed using MATLAB (R2007b, Mathworks, Natick, MA) on a Windows XP workstation. For every parameter combination below, 160,000 synthetic one- or two-channel datasets were generated using Eq. (1) at two measurement times ($t_0 = 0$, $t_1 = 1$ s) and a decay rate constant ($R_0 = 1$ s^{-1}). Independent Gaussian noise was added to each channel to the desired SNR.

3.1. One channel simulations

For each dataset in the one channel simulations, we determined the marginal probability distribution for the decay rate constant $\hat{R}$ using Eq. (6) and the maximum and standard deviation of each distribution was selected as the parameter estimate and its uncertainty. The median of these values across simulations was calculated, to minimize the effect of outlying datasets, and compared to the theoretically predicted uncertainty from Eq. (8). Data for this analysis were generated with SNR values of 50, 40, 30, 25, 20, 15, 10, and 5:1.

Fig. 1 compares σ_R, the percent uncertainty in the rate constant estimate for a one channel dataset predicted from Eq. (8) (solid line), and the median percent standard deviation (% SD) of $\hat{R}$ from
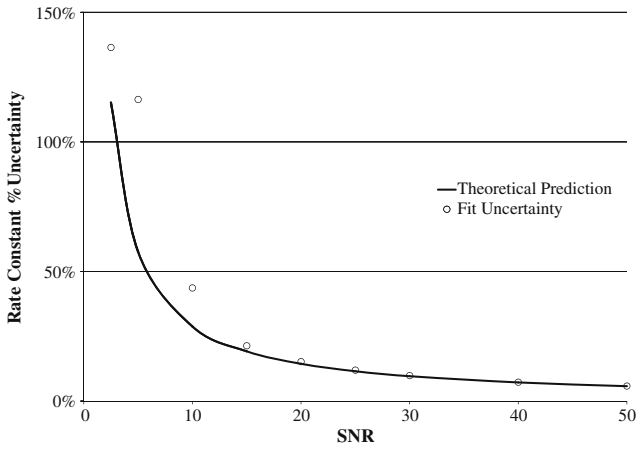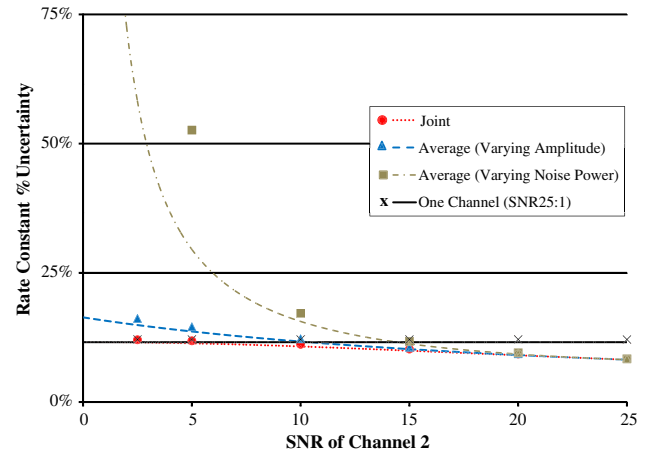
Fig. 1. The % uncertainty in estimation of the decay rate constant for one channel data as a function of SNR. The open circles represent the median % uncertainty from fitting the simulated one channel data and the lines show the theoretical % uncertainty predicted by Eq. (8).
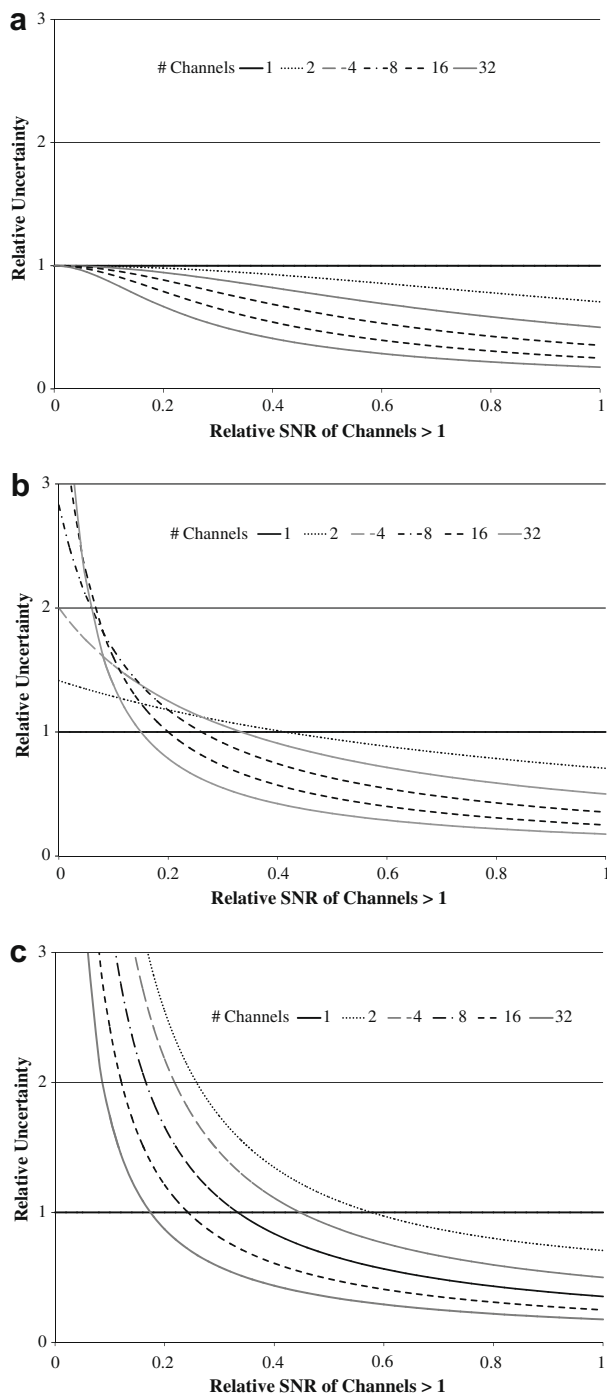

Fig. 2. The % uncertainty in estimation of the decay rate constant from two-channel data as a function of SNR in the second channel ($SNR_1 = 25:1$). Lines show the values predicted from theory and data points show the median % uncertainty in the decay rate constant obtained from modeling simulated data. Results are shown for the joint analysis, average analysis when the signal amplitude varies between channels, average analysis when the noise power varies between channels, and for analyzing the high SNR channel alone. The results for the sensitivity weighted average would overlap the joint analysis when the amplitude is varied and overlap the average analysis when the noise power is varied between channels. The optimally weighted data would overlap the results from the joint analysis under all conditions. For clarity, these overlapping results are omitted from the graph.

the Bayesian analysis (symbols), as functions of SNR. For one channel data at sufficiently high SNR (SNR $\geqslant$ 15), the theoretical values of the rate constant uncertainty, σ_R, are in excellent agreement with the % SD obtained from simulated data, whereas significant deviations appear at low SNR.

3.2. Two channel simulations

The computer simulations for a two-channel system were performed as follows. The first channel always had SNR of 25 and the second channel had SNR values of 25, 20, 15, 10, or 5:1. Every SNR value less than 25 was repeated twice; first by reducing the signal amplitude on the second channel and second by increasing the noise standard deviation, thereby producing a total of nine SNR combinations. Each simulated two-channel dataset was analyzed jointly and after channel combination using each of the following techniques: (i) average intensity – Eq. (16) with $\lambda_m = 1$, (ii) sensitivity weighted average – Eq. (16) with $\lambda_m = \hat{A}_m$, (iii) an "optimally weighted" average with $\lambda_m = \hat{A}_m/\sigma_m^2$, and (iv) square root of the sum of squares. For each dataset, the marginal probability distribution for the rate constant $\hat{R}$ for the averaged and joint analyses was determined using Eq. (6) or Eq. (10), and the maximum and standard deviation of each distribution was selected as the parameter estimate and its uncertainty. For the SOS analysis, the parameter estimates were also determined using Eq. (6), however the parameter uncertainty was not calculated due to the ambiguity in the value of the noise standard deviation, as described above. The maximum of the probability distributions was calculated with a resolution of 0.0025 s^{-1}.

The theoretically expected percent uncertainties in the decay rate constant estimates, σ_R, are shown in Fig. 2 for each combination method as functions of SNR_2, the SNR in the second channel. For comparison, the expected percent uncertainty in the decay rate constant estimates from analyzing the high SNR channel alone (denoted below as $\sigma_{R,1Ch}$) is also shown (horizontal line in Fig. 2). The symbols of each color represent the median percent uncertainty in the rate constant, over all 160,000 simulated datasets using that combination technique. The estimates of the rate constant uncertainty obtained by "optimal weighting" for both constant noise power and constant signal amplitude exactly coincide with the uncertainty obtained by the joint analysis, as does the uncertainty from "sensitivity map" averaging when SNR_2 is decreased by reducing signal amplitude. The uncertainty obtained by "sensitivity map" weighting when the noise standard deviation is increased

in the second channel exactly coincides with the uncertainty from the un-weighted average under the same conditions. For clarity, these overlapping results are not separately displayed in Fig. 2. There is an excellent agreement between the theoretical predictions for the rate constant uncertainty, σ_R, and the values obtained from simulated data for all the combination methods over the range of SNR studied.

When both the signal amplitude and noise standard deviation in the two channels are the same ($A_1 = A_2$, $\sigma_1 = \sigma_2$, Fig. 2 right side), the average, joint, and sensitivity weighted average analyses all produce equivalent uncertainties, equal to the $\sigma_{R,1Ch}/\sqrt{2}$, as expected, where $\sigma_{R,1Ch}$ is the uncertainty from analyzing the high SNR channel alone. As SNR_2 is reduced, σ_R produced by the joint analysis increases smoothly towards, but remains lower than $\sigma_{R,1Ch}$. In contrast, σ_R obtained from the average analysis can become higher than the one channel result as SNR_2 decreases.
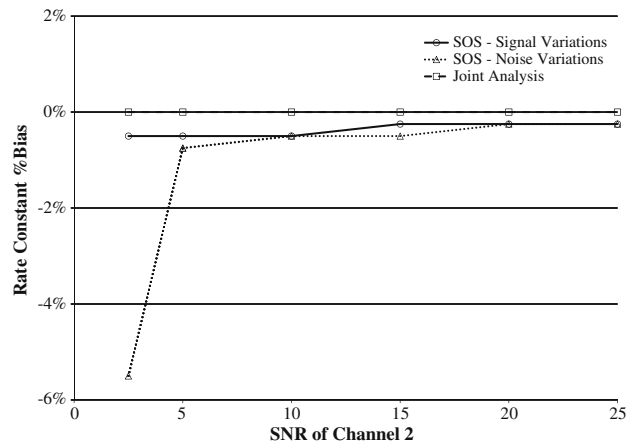
It is important to note that σ_R obtained by the average analysis (weighted and un-weighted) depends both upon the SNR of the second channel, SNR_2, and the whether the signal amplitude, A_2, or the noise standard deviation, σ_2 differs from the first channel. If the two channels have identical noise powers, as the SNR decreases by decreasing the amplitude in the second channel, σ_R from the sensitivity weighted average mimics the joint analysis and increases towards the single channel result, whereas the un-weighted average increases towards $\sigma_{R,1Ch} \cdot \sqrt{2}$. However, if the two channels have identical amplitudes, as the SNR decreases by increasing the noise power in the second channel, the uncertainty in the rate constant estimate for both the sensitivity-weighted and un-weighted averages rapidly increases and becomes proportional to the noise power in the second channel, σ_2. This is in sharp contrast to the joint Bayesian analysis, which is independent of the method of changing SNR in the regime considered here.

This theory was also utilized to generate Fig. 3, which show the expected rate constant uncertainty relative to the one channel result, $\sigma_R/\sigma_{R,1Ch}$, for an M-channel array where all channels other than the first have identical, lower SNRs, denoted as SNR_2. These results are plotted as functions of SNR_2/SNR_1 for different numbers of channels with the lower SNR. Fig. 3a corresponds to the joint

a

Fig. 4. The % bias in estimation of the decay rate constant as a function of SNR in the second channel calculated as the median bias over all two-channel simulated datasets for the sum-of-squares (SOS) and joint analyses. The results of the joint analysis are independent of whether the noise or amplitude of the second channel is varying.

channel result, $\sigma_{R,1Ch}$, when the lower SNR channels contain no signal and to $\sigma_{R,1Ch}/\sqrt{M}$ as the SNR in all channels becomes equal. The uncertainty from the averaged data also converges to $\sigma_{R,1Ch}/\sqrt{M}$ as the SNR in all channels becomes equal. However, as SNR_2 decreases, σ_R from averaged data can become larger than the one channel result. If all M channels have identical noise powers, the uncertainty from the un-weighted average converges to $\sigma_{R,1Ch} \cdot \sqrt{M}$ as the SNR of the additional channels decreases (Fig. 3b). If all M channels have identical amplitudes, the uncertainty from un-weighted averaging rapidly increases with decreasing SNR and becomes proportional to the noise power of the additional channels (Fig. 3c). As with the two channel case, the sensitivity weighted average would appear identical to Fig. 3a when only the signal amplitude changes between channels and Fig. 3c when only the noise power changes between channels. In real experiments where amplitude and noise power changes both contribute to SNR differences between channels would be expected to have a SNR dependence that lies between these two extremes.

The bias in estimation of the decay rate constant was calculated as the median bias over all two-channel simulated datasets. The % bias for the sum-of-squares, and joint analysis is shown in Fig. 4 as a function of SNR in the second channel. The sum-of-squares (SOS) combination of channels demonstrates the expected negative bias in the decay rate constant, with a greater bias when the noise power varies between channels. The joint analysis is effectively unbiased at all SNR values tested and is independent of whether the amplitude or noise power differs between channels. The optimally weighted average is also unbiased and the average and weighted average analyses exhibit a minimal bias only at the lowest SNR values (data not shown).

4. Discussion

The theoretical analysis and computer generated experiments above demonstrate that two approaches, joint analysis and the "optimal" weighted average, give the most accurate and precise results for parameter estimation from multi-channel data. While the "optimal" weighted average method requires knowledge of the noise power and amplitude for each pixel in each channel (information not always readily available), the joint Bayesian analysis "automatically" takes care of this problem by effectively weighting the contributions from the different channels based upon the strength of their projection onto the model.

b

c

Fig. 3. Theoretical uncertainties in estimation of a decay rate constant for different numbers of channels, as a function of the relative SNR in channels other than the first. The uncertainties are normalized to the uncertainty from analysis of the high SNR channel alone. (a) Joint channel analysis. (b) Analysis of the un-weighted average when the amplitude varies between channels. (c) Analysis of the un-weighted average when the noise power varies. As the channels become identical, all combination methods converge to the expected sqrt(M) lower than the one channel uncertainty, as expected. The results for a sensitivity weighted average would overlap the joint analysis when the amplitude is varied and the average analysis when the noise power is varied.

analysis, Fig. 3b to un-weighted averaging of channels with varying amplitudes and Fig. 3c to un-weighted averaging of channels with varying noise standard deviations. We see that the uncertainty in $\hat{R}$ from the joint analysis always improves with additional channels, even at low SNR. As expected, this uncertainty converges to the one

The estimates for parameter accuracy produced in this Bayesian analysis are conservative and consistent with the information explicitly stated in the model definition. If additional information is available, e.g. reliable sensitivity profiles for the array channel elements, this can be incorporated into the joint Bayesian analysis and may improve the quality of the parameter estimation. However, as these sensitivity profiles can also depend upon RF coil loading and the acquisition parameters used for sensitivity mapping, caution should be exercised when incorporating this information into this or other analyses.

Most previous treatments of multi-channel data have assumed that all channels are independent, but share the same noise power. As the noise in a given channel depends upon the coil construction, its loading, which may be different for array elements positioned about different regions of the subject's body, the degree of inter-channel coupling, and whether the noise is sample- or coil-limited, this assumption may not be valid in a particular experiment. Systematic effects in the data that are not fully modeled, such as noise correlations between channels or the Rican noise profile induced by processing the magnitude images from each channel, can add coherently upon channel combination and significantly distort the resulting parameter estimates. The joint Bayesian analysis of channels minimizes these effects, and provides a rigorous framework for incorporating factors such as inter-channel noise correlations.

As apparent from Eq. (13), the parameter uncertainty using the joint analysis is independent of whether the amplitude or the noise power is varying across channels and depends only upon the resultant signal-to-noise ratio. In contrast, the un-weighted or sensitivity-weighted average analyses, Eq. (18), are strongly dependent upon how these quantities individually vary.

As expected, the accuracy and precision of the parameter estimates improves with increasing SNR for all channel combinations. The joint analysis utilizes the additional information contained in low SNR channels without corrupting the high SNR channels and therefore produces parameter estimates that are always better than or equal to the one channel result. In contrast, the accuracy and precision of parameter estimates obtained after channel combination are sometimes degraded by the inclusion of a lower SNR channel, producing less accurate and precise results than if the high SNR channel was analyzed alone. This is especially true if the noise standard deviation varies between the two channels, as most combination techniques assume this is constant for all channels. The impact of non-optimal channel combination on the precision of parameter estimates increases with the number of low SNR channels and the number of sampling times (especially those with lower SNR).

Compensating for the reduced effective SNR of non-optimally combined data requires additional scanner time or improved hardware. For example, when using an eight channel array, if one channel has twice the SNR of the other channels for a given pixel or dataset, then a non-optimal channel combination can reduce the effective SNR by up to 10%, requiring a 14% increase in imaging time to compensate. If instead we assume that one channel has a four times greater SNR, the decrease in effective SNR could be as large as 64%, requiring a 167% increase in imaging time. In contrast, the joint analysis produces the maximum effective SNR with no additional scanner time and without distorting the signal relationships within a series of images. While a joint analysis is more computationally intensive than traditional methods, with modern computers this is less of a limitation.

The significance of these effects will depend upon SNR and the individual coil configuration. Laboratory-built arrays and arrays using heterogeneous elements, such as combined head and spine arrays, would be the most likely to display variations in the noise standard deviation. As the sensitivity of a surface coil array element is roughly proportional to the element's radius [26], large in-ter-coil differences in signal would be most prevalent when the physical extent of the array is much larger than the radius of the individual elements, such as in linearly aligned arrays and arrays with a large number of elements.

We have also assumed that the data from each channel are individually phased and that the analysis is performed only upon the "real" channel, which contains the entire signal while the noise is evenly distributed between the "real" and "imaginary" channels. Instead, if the magnitude images from each channel are analyzed, we distort the decay curve by introducing a DC offset for each channel, especially at values with low SNR. This offset will coherently add across the channels and produce a systematic underestimation of the decay rate constant if it is not incorporated into the data model. This would produce an effect analogous to the SOS data in Fig. 4. While proper modeling of the Rician noise distribution and the resulting DC offset can mitigate these effects [10,15,17–20], the increased model complexity and the uncertainty in the sign of low SNR data will still decrease the precision of the parameter estimates.

Implementing the joint analysis of the phased signals from each channel requires saving the raw, uncombined data from the individual channel elements. Saving this data is not the default behavior on many scanners and may not be possible with configurations where the individual coil signals are combined electronically prior to detection. However, even when only magnitude data are available, a joint analysis of the channels will still provide equal or superior parameter estimates to the coil combination methods described above (data not shown).

5. Conclusions

When analyzing data acquired using a phased array coil, the joint analysis of phased images provides the most robust parameter estimates. The accuracy and precision of these estimates are as good as or better than can be obtained from the channel combination methods considered here, without requiring reference scans or assumptions of constant noise power across channels. Of the possible channel combination methods, weighted averaging using the signal amplitude divided by the noise variance is an acceptable alternative provided that spatial profiles of the signal amplitude and noise standard deviation are known or can be accurately estimated for each channel.

The use of incorrect weighting factors decreases the effective SNR and can introduce a bias into the parameter estimates, sometimes producing less accurate and precise results than would be obtained from analyzing the single highest SNR channel alone. This effect is amplified at low SNR, for increased number of low SNR channels, increased number of low SNR points in the decay curve, and when noise power varies across channels. Sensitivity weighted images and sum-of-squares both assume that the noise power is identical across channels, which may not be valid for all cases. The sum-of-squares combination of all images, the default in many systems, introduces a known bias into the parameter estimation and therefore should generally be avoided.

References

[1] P.B. Roemer, W.A. Edelstein, C.E. Hayes, S.P. Souza, O.M. Mueller, The NMR phased array, Magn. Reson. Med. 16 (2) (1990) 192–225.
[2] K.P. Pruessmann, M. Weiger, M.B. Scheidegger, P. Boesiger, SENSE: sensitivity encoding for fast MRI, Magn. Reson. Med. 42 (5) (1999) 952–962.

[3] D.K. Sodickson, W.J. Manning, Simultaneous acquisition of spatial harmonics (SMASH): fast imaging with radiofrequency coil arrays, Magn. Reson. Med. 38 (4) (1997) 591–603.

[4] J.P. Debbins, J.P. Felmlee, S.J. Riederer, Phase alignment of multiple surface coil data for reduced bandwidth and reconstruction requirements, Magn. Reson. Med. 38 (6) (1997) 1003–1011.

[5] C.E. Hayes, P.B. Roemer, Noise correlations in data simultaneously acquired from multiple surface coil arrays, Magn. Reson. Med. 16 (2) (1990) 181–191.

[6] E.G. Larsson, D. Erdogmus, R. Yan, J.C. Principe, J.R. Fitzsimmons, SNR-optimality of sum-of-squares reconstruction for phased-array magnetic resonance imaging, J. Magn. Reson. 163 (1) (2003) 121–123.

[7] S.M. Wright, L.L. Wald, Theory and application of array coils in MR spectroscopy, NMR Biomed. 10 (8) (1997) 394–410.

[8] D.O. Walsh, A.F. Gmitro, M.W. Marcellin, Adaptive reconstruction of phased array MR imagery, Magn. Reson. Med. 43 (5) (2000) 682–690.

[9] C.D. Constantinides, E. Atalar, E.R. McVeigh, Signal-to-noise measurements in magnitude images from NMR phased arrays, Magn. Reson. Med. 38 (5) (1997) 852–857.

[10] M. Bydder, D.J. Larkman, J.V. Hajnal, Combination of signals from array coils using image-based estimation of coil sensitivity profiles, Magn. Reson. Med. 47 (3) (2002) 539–548.

[11] G. Gilbert, D. Simard, G. Beaudoin, Impact of an improved combination of signals from array coils in diffusion tensor imaging, IEEE Trans. Med. Imaging 26 (11) (2007) 1428–1435.

[12] M. Buehrer, K.P. Pruessmann, P. Boesiger, S. Kozerke, Array compression for MRI with large coil arrays, Magn. Reson. Med. 57 (6) (2007) 1131–1139.

[13] M.J. Graves, D. Emmens, H. Lejay, H. Hariharan, J. Polzin, D.J. Lomas, T2 and T2* quantification using optimal B1 image reconstruction for multicoil arrays, J. Magn. Reson. Imaging 28 (1) (2008) 278–281.

[14] G.L. Bretthorst, Bayesian spectrum analysis and parameter estimation, in: J. Berger, S. Fienberg, J. Gani, K. Krickeberg, B. Singer (Eds.), Springer-Verlag, New York, 1988, 209 p.

[15] O. Dietrich, S. Heiland, K. Sartor, Noise correction for the exact determination of apparent diffusion coefficients at low SNR, Magn. Reson. Med. 45 (3) (2001) 448–453.

[16] R.L. O'Halloran, J.H. Holmes, T.A. Altes, M. Salerno, S.B. Fain, The effects of SNR on ADC measurements in diffusion-weighted hyperpolarized He-3 MRI, J. Magn. Reson. 185 (1) (2007) 42–49.

[17] A. Kristoffersen, Optimal estimation of the diffusion coefficient from non-averaged and averaged noisy magnitude data, J. Magn. Reson. 187 (2) (2007) 293–305.

[18] J. Sijbers, A.J. den Dekker, E. Raman, D. Van Dyck, Parameter estimation from magnitude MR images, Int. J. Imaging Syst. Technol. 10 (1999) 109–114.

[19] O.T. Karlsen, R. Verhagen, W.M. Bovee, Parameter estimation from Rician-distributed data sets using a maximum likelihood estimator: application to T1 and perfusion measurements, Magn. Reson. Med. 41 (3) (1999) 614–623.

[20] C.G. Koay, P.J. Basser, Analytically exact correction scheme for signal extraction from noisy magnitude MR signals, J. Magn. Reson. 179 (2) (2006) 317–322.

[21] G.L. Bretthorst, Automatic phasing of MR images. I: linearly varying phase, J. Magn. Reason. 191 (2008) 184–192.

[22] G.L. Bretthorst, Automatic phasing of MR images. Part II: voxel-wise phase estimation, J. Magn. Reason. 191 (2008) 193–201.

[23] A.L. Sukstanskii, G.L. Bretthorst, Y.V. Chang, M.S. Conradi, D.A. Yablonskiy, How accurately can the parameters from a model of anisotropic 3He gas diffusion in lung acinar airways be estimated? Bayesian view, J. Magn. Reason. 184 (1) (2007) 62–71.

[24] G.L. Bretthorst, W.C. Hutton, J.R. Garbow, J.J.H. Ackerman, Exponential parameter estimation (in NMR) using Bayesian probability theory, Concepts Magn. Reason. A 27A (2) (2005) 55–63.

[25] G.L. Bretthorst, How accurately can parameters from exponential models be estimated? A Bayesian view, Concepts Magn. Reson. A 27A (2) (2005) 73–83.

[26] J.J. Ackerman, T.H. Grove, G.G. Wong, D.G. Gadian, G.K. Radda, Mapping of metabolites in whole animals by 31P NMR using surface coils, Nature 283 (5743) (1980) 167–170.