Journal of the Royal Statistical Society Series C: Applied Statistics, 2025, **74**, 100–125
https://doi.org/10.1093/jrsssc/qlae046
Advance access publication 17 September 2024
Original Article

ROYAL STATISTICAL SOCIETY DATA | EVIDENCE | DECISIONS
Journal of the Statistics Society
Series C
Applied Statistics

A Bayesian complex-valued latent variable model applied to functional magnetic resonance imaging

Chase J. Sakitis[1], D. Andrew Brown[2] and Daniel B. Rowe[1]

[1]Department of Mathematical and Statistical Sciences, Marquette University, Milwaukee, WI, USA
[2]Department of Mathematical and Statistical Sciences, Clemson University, Clemson, SC, USA

Address for correspondence: Daniel B. Rowe, Department of Mathematical and Statistical Sciences, Marquette University, 313, Cudahy Hall, 1313 W. Wisconsin Ave., Milwaukee, WI 53233, USA. Email: daniel.rowe@marquette.edu

Abstract

In linear regression, the coefficients are simple to estimate using the least squares method with a known design matrix for the observed measurements. However, real-world applications may encounter complications such as an unknown design matrix and complex-valued parameters. The design matrix can be estimated from prior information but can potentially cause an inverse problem when multiplying by the transpose as it is generally ill-conditioned. This can be combat by adding regularizers to the model but does not always mitigate the issues. Here, we propose our Bayesian approach to a complex-valued latent variable linear model with an application to functional magnetic resonance imaging (fMRI) image reconstruction. The complex-valued linear model and our Bayesian model are evaluated through extensive simulations and applied to experimental fMRI data.

Keywords: Bayesian, fMRI, reconstruction, regression, SENSE

1 Introduction

1.1 Background

Linear regression is a common tool used for prediction analysis of one variable based on the value of another variable. The equation for linear regression is that of a line of best-fit with measurement error as expressed in equation (1.1)

$$y_j = \beta_0 + \beta_1 x_{j1} + \beta_2 x_{j2} + \cdots + \beta_p x_{jp} + \varepsilon_j, \quad j = 1, \ldots, n, \tag{1.1}$$

where p is the number of regression coefficients and n is the number of observations. Some applications, such as in functional magnetic resonance imaging (fMRI) image reconstruction, do not have a y-intercept in the model which would remove the β_0 from the model. For this article, we will focus on the regression model with no y-intercept which can be compactly written as

$$y = X\beta + \varepsilon, \tag{1.2}$$

where $y \in \mathbb{R}^{n\times 1}$ is the observed dependent variable, $X \in \mathbb{R}^{n\times p}$ is the design matrix, $\beta \in \mathbb{R}^{p\times 1}$ is the vector of regression coefficients, and $\varepsilon \in \mathbb{R}^{n\times 1}$ is the measurement error or residuals. With regression models, our goal is to solve for the coefficients β. If the variables of this model are

Received: October 11, 2023. **Revised:** May 18, 2024. **Accepted:** August 7, 2024

real-valued and the design matrix X is known, we can simply apply a least squares method to solve for β by using equation (1.3)

$$\beta = (X'X)^{-1}X'y. \tag{1.3}$$

This linear model can encounter complex values instead of real-valued variables which changes y to $y_c \in \mathbb{C}^{n\times 1}$, X to $X_c \in \mathbb{C}^{n\times p}$, β to $\beta_c \in \mathbb{C}^{p\times 1}$, and ε to $\varepsilon_c \in \mathbb{C}^{n\times 1}$. With complex-valued parameters, we can write the linear model using a real-valued isomorphic representation to essentially remove the nuisance of complex values. This isomorphic representation is shown as

$$\begin{bmatrix} y_R \\ y_I \end{bmatrix} = \begin{bmatrix} X_R & -X_I \\ X_I & X_R \end{bmatrix}\begin{bmatrix} \beta_R \\ \beta_I \end{bmatrix} + \begin{bmatrix} \varepsilon_R \\ \varepsilon_I \end{bmatrix}, \quad (\varepsilon_R, \varepsilon_I)' \sim N(0, \sigma^2 I_{2n}), \tag{1.4}$$

where $y_R \in \mathbb{R}^{n\times 1}$ and $y_I \in \mathbb{R}^{n\times 1}$ are the observed real and imaginary components, respectively, of y, $X_R \in \mathbb{R}^{n\times p}$ and $X_I \in \mathbb{R}^{n\times p}$ are the unobserved real and imaginary components of X, $\beta_R \in \mathbb{R}^{p\times 1}$ and $\beta_I \in \mathbb{R}^{p\times 1}$ are the unobserved real and imaginary components of β, and $\varepsilon_R \in \mathbb{R}^{n\times 1}$ while $\varepsilon_I \in \mathbb{R}^{n\times 1}$ are the real and imaginary components of ε. This isomorphic representation can be compactly written to be $y = X\beta + \varepsilon$, where $y \in \mathbb{R}^{2n\times 1}$, $X \in \mathbb{R}^{2n\times 2p}$, $\beta \in \mathbb{R}^{2p\times 1}$, and $\varepsilon \in \mathbb{R}^{2n\times 1}$.

1.2 Unknown design matrix

In real-world applications, we may not always have a known design matrix X. An example of this can be found in the blind source separation problem which has been studied by researchers in signal processing (Cardoso & Laheld, 1996; Comon, 1994; Yellin & Weinstein, 1996), identification of MA processing (Swami et al., 1994), and neural networks (Bell & Sejnowski, 1995; Cichocki et al., 1994; Roth & Baram, 1996). In source separation, measured signals are modelled using linear combinations of an operator matrix (design matrix) and the original source signals (regression coefficients). In this model, both the operator matrix and the source signals are unknown with the operator matrix not necessarily being full rank. The research performed in this field focuses on theoretical identification of the linear combinations through filtering or unsupervised learning algorithms to formally estimate the source signals (Cao & Liu, 1996; Choi et al., 2005; Lee et al., 1997). Despite this being a major area that consists of having an unknown design matrix in its linear model, even a small amount of *a priori* information is required to gain insight on the filtering processes to estimate the original source signals (Choi et al., 2005).

Aliased image reconstruction in functional magnetic resonance imaging (fMRI) is similar to a source separation problem. In this field, the design matrix is unknown but enough *a priori* information is obtained to estimate the design matrix and is treated as 'known'. This 'known' design matrix can then be used for ordinary least squares to estimate the regression coefficients. However, estimating the design matrix from prior information results in the matrix being generally ill-conditioned leading to an inverse problem. To address this issue, a common solution is to add regularizers such as ridge (Hoerl & Kennard, 1970) or lasso regression (Tibshirani, 1996) to the model. These regularizers, however, may not always mitigate the issues as they can introduce a bias-variance problem, be computational expensive, or produce subjective parameter estimates. This partially motivates our Bayesian approach to this latent variable linear regression problem.

1.3 Complex-valued applications

Also in real-world applications where a linear regression is modelled, the observed data can be complex-valued instead of real-valued. An example of a complex-valued, latent variable real-world application can be seen in speech enhancement. For speech enhancement, the goal is to improve the quality of noisy signals (Loizou, 2013). Most models in speech enhancement ignore the phase information yielding real-valued signals (Williamson et al., 2013, 2014) that can be modelled using linear regression. Chen et al. (2018a) incorporate the phase information in the reconstruction of the complex-valued short-time Fourier transformation using a nonlinear complex-valued Gaussian process model. This work is further improved by adding in locality-preserving and discriminative constraints (Chen et al., 2018b). Despite the use of non-linear models for the

complex-valued speech signal data, linear regression can be used on the complex-valued signals (Schreier & Schraf, 2010). With available prior information (Williamson et al., 2013, 2014) and a complex-valued linear model, our Bayesian approach can be applied to speech signal data. Also, Nguyen et al. (2017) address the possibility of under-determined systems in the complex-valued linear regression, in signals such as speech, by using a generalization of sparse filtering and K-hyperlines clustering. Even with under-determined systems, our Bayesian approach can still be applied without any alterations creating a fully automated process.

Similar to signal processing, the data in fMRI are also complex-valued and can be linearly modelled with an unknown design matrix. For this article, we introduce a Bayesian approach to a complex-valued latent variable linear model where the design matrix X along with the regression coefficients β and the noise variance σ^2 are treated as unknown parameters. Prior distributions are then placed on the unknown variables and combined with the likelihood to obtain the joint posterior distribution. This model can be applied to any complex-valued data that can be modelled using linear regression with (or without) an unknown design matrix. To demonstrate the utilization of our proposed isomorphic Bayesian complex-valued latent variable model, we applied the model to simulated and experimental fMRI data for image reconstruction.

1.4 Overview

The second section of this article will explain the model of the Bayesian complex-valued latent variable model. Section 3 of the article describes the fMRI application with Section 4 analysing the results of image reconstruction application. We will conclude in Section 5 with an overview of the important results of the article and a discussion of future work with this Bayesian model and its application to fMRI data.

2 Bayesian complex-valued model

For our Bayesian model, we use the isomorphic representation of the complex linear model as expressed in equation (1.4). In this work, two different representations of the design matrix will be used. The first representation is $X \in \mathbb{R}^{2n\times 2p}$ as shown in equation (1.4) which is necessary for the proper skew symmetric design matrix for complex-valued multiplication. The second is $G = [X_R, X_I]$, used in the prior distribution and ultimately for parameter estimation, since X_R and X_I uniquely determine X and do not need to be duplicated.

2.1 Data likelihood, prior, and posterior distributions

We assume that the residual error is normal and independent and identically distributed in the real and imaginary components (Macovski, 1996). The likelihood for the observed measurements for the n observations becomes

$$P(y \mid X, \beta, \sigma^2) \propto (\sigma^2)^{-\frac{2n}{2}} \exp\left[-\frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)\right]. \tag{2.1}$$

We can quantify available prior information about the regression coefficients β, the unobserved parameters of the design matrix X, and the residual variance σ^2 in the likelihood with assessed hyperparameters of prior distributions. The regression coefficients β are specified to have a normal prior distribution, expressed in equation (2.2). The design matrix, represented as G, is also specified to have a normal prior distribution (equation (2.3)) and the noise variance σ^2 is specified to have an inverse gamma prior distribution (equation (2.4)),

$$P(\beta \mid n_\beta, \beta_0, \sigma^2) \propto (\sigma^2)^{-\frac{2p}{2}} \exp\left[-\frac{n_\beta}{2\sigma^2}(\beta - \beta_0)'(\beta - \beta_0)\right], \tag{2.2}$$

$$P(G \mid n_X, G_0, \sigma^2) \propto (\sigma^2)^{-\frac{2np}{2}} \exp\left[-\frac{n_X}{2\sigma^2} tr(G - G_0)'(G - G_0)\right], \tag{2.3}$$

$$P(\sigma^2 \mid \alpha, \delta) \propto (\sigma^2)^{-(\alpha+1)} \exp\left[-\frac{\delta}{\sigma^2}\right], \tag{2.4}$$

where *tr* is the trace of the $(G - G_0)'(G - G_0)$ matrix. The hyperparameters n_X, G_0, n_β, β_0, α, and δ are assessed from available prior information, as discussed in the next subsection, but can also be determined using a fully subjective approach. The joint posterior distribution of the regression coefficients β, the design matrix X, and the noise variance σ^2 is

$$P(G, \beta, \sigma^2 \mid a) \propto P(y \mid X, \beta, \sigma^2)P(\beta \mid n_\beta, \beta_0, \sigma^2) \cdot P(G \mid n_X, G_0, \sigma^2)P(\sigma^2 \mid \alpha, \delta), \tag{2.5}$$

with the distributions specified from equations (2.1), (2.2), (2.3), and (2.4).

2.2 Hyperparameter determination

As mentioned in Section 1.2, in linear regression with an unknown design, prior information can be utilized to estimate the design matrix, treating it as a 'known' parameter. That same available data, $y_{0_c} \in \mathbb{C}^{n \times n_0}$, can be utilized to assess the hyperparameters (β_0, n_β, G_0, n_X, α, and δ) for the prior distributions of the unknown parameters for our Bayesian model. We utilize a logical straightforward objective process for hyperparameter assessment. For this, we average the n_0 prior data points (second dimension) for each y_{0_c} resulting in $y_{0\text{avg}} \in \mathbb{C}^{n \times 1}$ which can be utilized for hyperparameter assessment. Since our data are complex-valued, we can estimate an initial magnitude of our regression coefficients β_{0M} by computing the Euclidean norm of $y_{0\text{avg}}$. Then, the $y_{0\text{avg}}$ values can be pointwise divided by β_{0M} resulting in initial real and imaginary values for G_0. From this, we calculate the magnitude value by $\rho_G = \sqrt{R^2 + I^2}$ and the phase value by $\theta_G = \arctan(I/R)/2$ for G_0, where R and I are the initial real and imaginary components of G_0. By dividing the $arctan(I/R)$ by 2, both the real and imaginary components of the design matrix are incorporated into the estimation of the regression coefficients. Without dividing by 2, the least squares estimation zeroes out the imaginary component of the regression coefficients which directly removes the phase information from the complex-valued coefficients. The magnitude and phase estimates are then utilized to calculate the complex-valued prior means for G_0 using the equation $G_0 = \rho_G \exp(i\theta_G)$. These G_0 prior means, along with $y_{0\text{avg}}$, are used to estimate complex-valued prior means for the regression coefficients β_0 via least squares estimation.

The hyperparameters n_β and n_X, which are the scalar weights of the prior means for β and X respectively, are assessed to be the number of prior data points n_0. The average residual variance over the second dimension of our prior data points y_{0_c} is calculated to obtain a prior for the noise variance noted as σ_0^2. The hyperparameters α (shape parameter of the inverse gamma) and δ (scale parameter of the inverse gamma) are assessed to be $\alpha = n_0 - 1$ and $\delta = (n_0 - 1)\sigma_0^2$. This prior information is incorporated in estimating the p regression coefficients β for every jth data point of the observed measurements.

2.3 Posterior estimation

Using the posterior distribution in equation (2.5), two approaches are used to estimate the regression coefficients β, design matrix X, and residual variance σ^2. Maximum *a posteriori* (MAP) estimation using the Iterated Conditional Modes (ICM) optimization algorithm (Lindley & Smith, 1972; O'Hagan, 1994) to find the joint posterior mode, and marginal posterior mean (MPM) estimation via Markov chain Monte Carlo (MCMC) Gibbs sampling (Gelfand & Smith, 1990; Geman & Geman, 1984). It should be noted that with the current specifications with the likelihood and priors, the posterior conditional distributions are unimodal hence the joint posterior distribution is unimodal. Beginning with initial estimates of each parameter, ICM iterates over the parameters, calculating its posterior conditional mode until convergence at the joint posterior mode. The posterior conditional modes are

$$\hat{\beta} = (X'X + n_\beta I_{2p})^{-1}(X'y + n_\beta \beta_0), \tag{2.6}$$

$$\hat{G} = (B'B + n_X I_{2p})^{-1}(BY' + n_X G_0), \tag{2.7}$$

$$\hat{\sigma}^2 = \frac{\Theta}{2(2n + 2p + \alpha + 2np + 1)}, \tag{2.8}$$

where $\Theta = (y - X\beta)'(y - X\beta) + n_\beta(\beta - \beta_0)'(\beta - \beta_0) + \alpha\delta + n_X tr[(X - X_0)(X - X_0)']$, $Y = [y_R, y_I]$ and $B \in \mathbb{R}^{2p\times 2}$ is a skew symmetric matrix representation of the regression coefficients β as expressed by

$$B = \begin{bmatrix} \beta_R & \beta_I \\ -\beta_I & \beta_R \end{bmatrix}. \tag{2.9}$$

The full conditional distributions are given by

$$\beta \mid X, \sigma^2, y \sim N\{\hat{\beta}, \sigma^2(X'X + n_\beta I_p)\}, \tag{2.10}$$

$$G \mid \beta, \sigma^2, y \sim MN\{\hat{G}, \sigma^2(B'B + n_X I_{2p})\}, \tag{2.11}$$

$$\sigma^2 \mid \beta, X, y \sim IG(\alpha_*, \delta_*), \tag{2.12}$$

where $\alpha_* = np + n + p + \alpha$ and $\delta_* = [(y - X\beta)'(y - X\beta) + n_\beta(\beta - \beta_0)'(\beta - \beta_0) + n_X tr((G - G_0)(G - G_0)') + 2\delta]/2$. This process is completely objective providing a fully automated method without having to calculate a subjective penalty. Our Bayesian approach, however, is flexible enough to include subjective priors if desired. Because we are using available prior information, we expect the subsequent estimators to have smaller variance. Since we expect the expected mean of the of our regression coefficients to be equal to true mean, i.e. $E(\hat{\beta}) = \beta$, we can appropriately use the mean square error (MSE) estimate to quantify bias. This measure is also used to determine how accurate the regression coefficients are to the true values. To illustrate our Bayesian approach compared to the non-Bayesian approach, extensive realistic simulations are performed in Section 4.

3 fMRI application

3.1 fMRI background

Magnetic resonance imaging (MRI) is a type of medical imaging that creates images using magnetic fields. fMRI was developed in the early 1990's as a technique to noninvasively observe the human brain in action without exogenous contrast agents (Bandettini et al., 1993). This procedure examines brain activity by detecting changes in the brain using the blood-oxygen-level dependent (BOLD) contrast (Ogawa et al., 1990). When a neuron fires, the BOLD contrast increases in the proximity of the neuron and is a correlate for neuronal firing. The firing of neurons is a proxy for brain activity and is of interest when examining the brain in action in fMRI analysis. Measurements for images are arrays of complex-valued spatial frequencies in so called *k*-space (Kumar et al., 1975). These *k*-space arrays are then reconstructed into images using an inverse Fourier transform (IFT) producing brain images. The reconstructed brain images are made up of complex-valued voxels which contain the signal intensity for each pixel in the image. The magnitude and phase of the complex-valued reconstructed images can be utilized for analysis (Rowe, 2005; Rowe & Logan, 2004), but generally only the magnitude is used (Bandettini et al., 1993).

In fMRI, obtaining hundreds of volume images is necessary to detect activation in the brain. This series of observations are of the same underlying volume image taken over time. Measuring full arrays of data for all slices required to form volume images takes a considerable amount of time due to the size a dataset is from a single fMRI experiment. For example, the experimental data used in this article contain nine slices of 96×96 images with 510 time points yielding 41,472,000 complex-valued data points. Acquiring fully sampled *k*-space arrays where every value in the array is measured limits the temporal resolution of the reconstructed images which can diminish effectively capturing brain activity.

Historically, a single channel receiver coil has been utilized in fMRI to measure fully sampled *k*-space. The drawbacks of acquiring fully sampled *k*-space arrays directed fMRI research to increase the number of images acquired per unit of time. More recently, the focus of research has

been to acquire more images per unit of time by measuring less data without losing the ability to form a full image. To accomplish this, multiple receiver coils are utilized in parallel to each measure spatial frequencies. This would require the multiple coil images, after using the IFT, to be combined into a single, composite brain image. In 1999, Pruessmann et al. introduced a parallel imaging technique called SENSitivity Encoding (SENSE) which operates on the images after IFT.

The SENSE method uses the linear regression, as expressed in equation (1.2), with complex-valued parameters and a fixed design matrix. A complex-valued least squares solution (equation (1.3)) is used to estimate the unknown parameter, which would be the voxel values of the single, full brain image. This approach for parameter estimation can be difficult because the complex-valued design matrix, generally, is ill-conditioned. This can cause aliasing artefacts, low image quality, and signal-to-noise ratio (SNR) degradation in the final reconstructed image, which has led to variations of the traditional technique (King & Angelos, 2001; Liang et al., 2002; Lin et al., 2004; Liu et al., 2009; Ying et al., 2004). These modified regularization models have deficiencies that hardly mitigate the limitations of the traditional maximum likelihood SENSE procedure. These variations cause trade-off between SNR and aliasing artefacts (King & Angelos, 2001) or can lead to a significant increase in computational expense (Lin et al., 2004; Liu et al., 2009; Ying et al., 2004) due to selection of the regularizer which can render these techniques ineffective in practice. We can apply a Bayesian approach to this complex-valued linear model with an unknown design matrix called Bayesian SENSE (BSENSE). Our Bayesian model will incorporate prior information, which is assessed with complete automation and minimal computation time (<1 s) and does not use a single *a priori* fixed complex-valued sensitivity matrix. Through the extensive simulation study and application to experimental data, the results yield increased SNR, no aliasing artefacts, and increased image quality with improved task detection results.

For the fMRI application, the notation for the observed measurements (y), the design matrix (X), and the regression coefficients (β) in the linear model become a, S, and v, respectively. Also, n_C, the number of coils, will replace n and n_A, the acceleration factor, will replace p when discussing the dimensions of the parameters.

3.2 Research problem

As mentioned in the previous subsection, fMRI historically utilized a single channel receiver coil as illustrated in Figure 1. With a single channel coil, the height of the receiver is taller than the size of the subject's head, shown in the three-dimensional depiction in Figure 1a. Both parts a and b of Figure 1 show the single coil receiver wraps completely around the subject's head starting from posterior to anterior and connects back at the posterior.

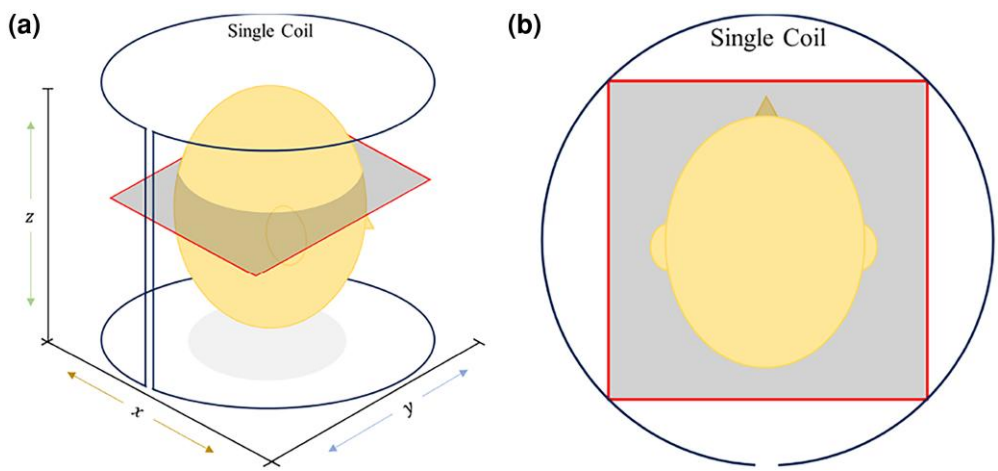

Figure 1. a) Illustration of a three-dimensional single coil channel along with b) the top-down view of the coil receiver.

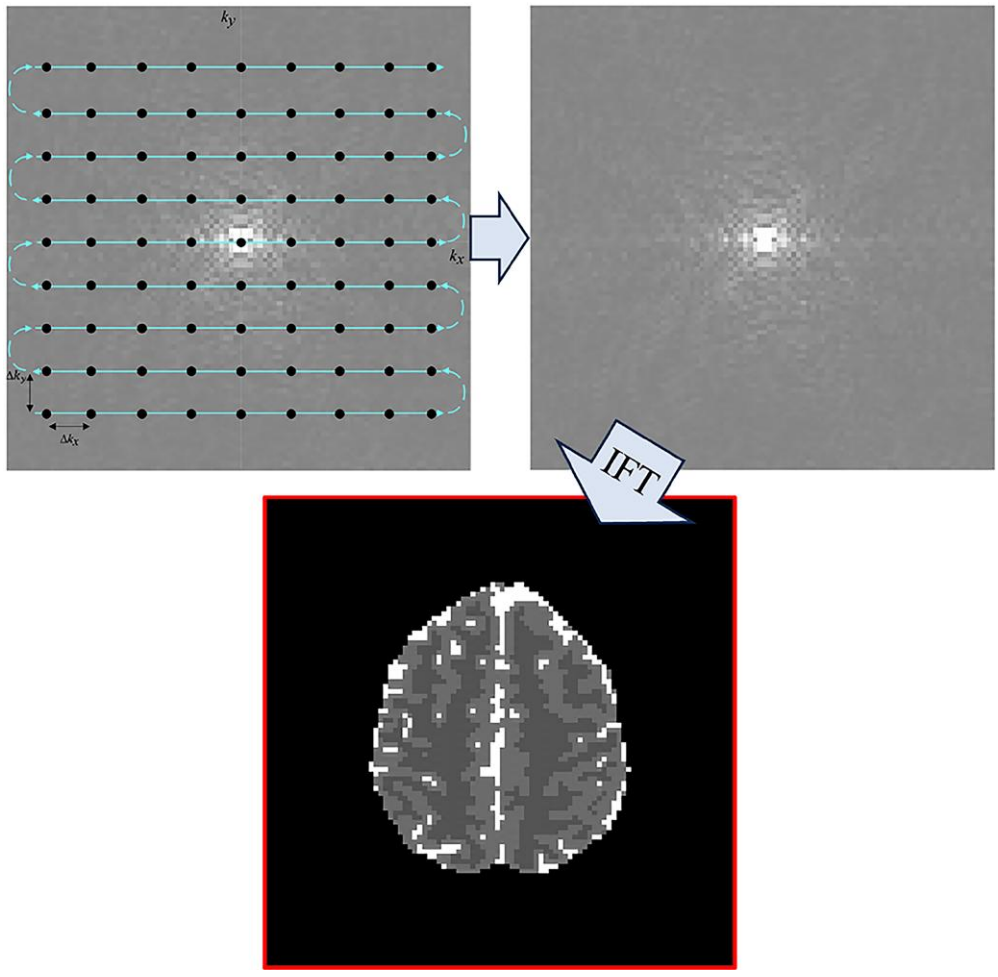

Figure 2. Fully sampled *k*-space zig-zag coverage (top left) with the finalized full *k*-space array after omitting the turn-around points (top right) and the reconstructed brain image using the IFT (bottom).

From the single channel coil, the *k*-space arrays are acquired along a trajectory as shown in Figure 2 (top left) where the machine starts in the bottom left corner and moves across the row measuring complex-valued spatial frequencies along the Cartesian grid. At the end of each row, you move up one line and the process is repeated in the opposite direction. This acquisition of complex-valued spatial frequencies is continued until all the rows of the *k*-space array is obtained, yielding fully sampled *k*-space depicted in Figure 2 (top right). These complex-valued spatial frequency arrays are then reconstructed into full field-of-view (FOV) magnitude and phase brain images using the IFT (bottom of Figure 2). The reconstructed phase image is not shown.

To acquire more images per unit of time, $n_C > 1$ receiver coils are utilized instead of a single channel coil. The number of coils $n_C > 1$ would be the n observations as described in Section 1.1. An example of a four-channel coil arrangement is illustrated in Figure 3. The three-dimensional depiction of the multi-coil arrays in Figure 3a show the height of the receiver coils being taller than the head of the subject. In Figure 3b, starting with coil 1 at the anterior of the subject, the coils increment clockwise with coil 2 on the right lateral, coil 3 on the posterior, and coil 4 on the left lateral of the subject's head. Each of the four coils can measure full sampled *k*-space arrays, as exhibited in Figure 2, in parallel which does not increase the acquisition time compared to the single channel coil array.

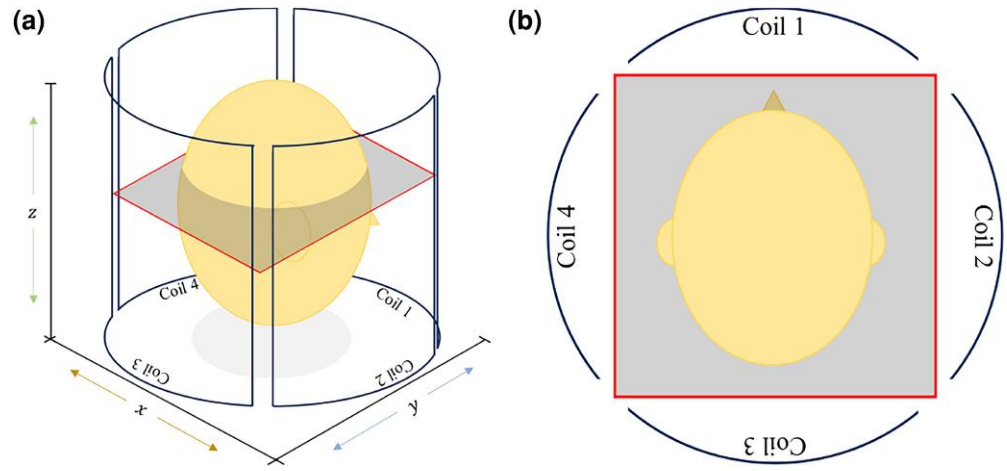

Figure 3. a) Illustration of a three-dimensional multi-coil channel with four receivers along with b) the top-down view of the multiple coils.

Each channel receiver coil possesses a depth sensitivity profile which depends on its size and location. This means that each coil can only 'see' parts of the object with a particular depth sensitivity that decreases as we move farther from the coil. The same four-channel coil configuration in Figure 3b is displayed in Figure 4 (centre image with four coils on each side) showing how the coils would look around a single slice brain image. Figure 4 gives an illustrative example of image slices with $n_C = 4$ coils (top, bottom, left, right) and their respective depth sensitivity to the true image slice (the four corners of the figure). The images for Figure 4 are magnitude images used to visualize the how the linear model is designed. In Figure 4, the top right corner image displays the true image point-wise multiplied by the depth sensitivity profile of coil 1 which is located at the front of brain. The resulting image shows that the signal intensity of the image decreases as you move farther from the coil location towards the back of the brain (bottom of the top right image). When examining a single complex-valued voxel in the weighted brain image for coil 1, the complex-valued voxel from the true image (centre) is multiplied by the complex-valued weighted sensitivity, S_{1c}, to get $a_{1c} = S_{1c}v_c$. The other three coils follow this same operation creating the system of equations $a_c = S_c v_c$ where $a_c = [a_{1c}, a_{2c}, a_{3c}, a_{4c}]'$ and $S_c = [S_{1c}, S_{2c}, S_{3c}, S_{4c}]'$. With this system of equations, a_c is the complex-valued coil measurements (the observed measurements y from equation (1.2)), S_c is the coil sensitivities (the design matrix X from equation (1.2)), and v_c is the unaliased, and coil combined, voxel values (the regression coefficients β from equation (1.2)). With this system of equations, a_c (the corner images in Figure 4) is the observed measurements, after applying the IFT, from the machine that need to be combined into a single, composite brain image. Since voxels are spatially discrete, this process is repeated for the rest of the voxels in the coil measurements.

As previously noted, the primary goal of parallel imaging is to increase the number of images acquired per unit of time which can be attained by measuring less data. This can be accomplished by skipping lines in the k-space array, i.e. subsampling, as displayed in Figure 5. Skipping lines in k-space introduces what is called an acceleration factor, n_A. The acceleration factor indicates the fraction of lines of data in k-space that are measured and how much sampling time is reduced for a volume image. For example, with an acceleration factor of $n_A = 2$, every other line horizontally in k-space is measured as exhibited on the left side of Figure 5. This would result in each slice of the volume k-space arrays to be 48×96 (top right of Figure 5) instead of the full 96×96. If it took 1 s to obtain a full volume k-space array, with $n_A = 2$, the subsampled volume image would take half a second, doubling the rate at which we can observe brain dynamics. If an acceleration factor of $n_A = 3$ is used, a third of the points along the horizontal lines of k-space are measured yielding each slice of the volume image to be 32×96 which means three subsampled volume images would be observed in the time it would take to observe one full sampled volume image.

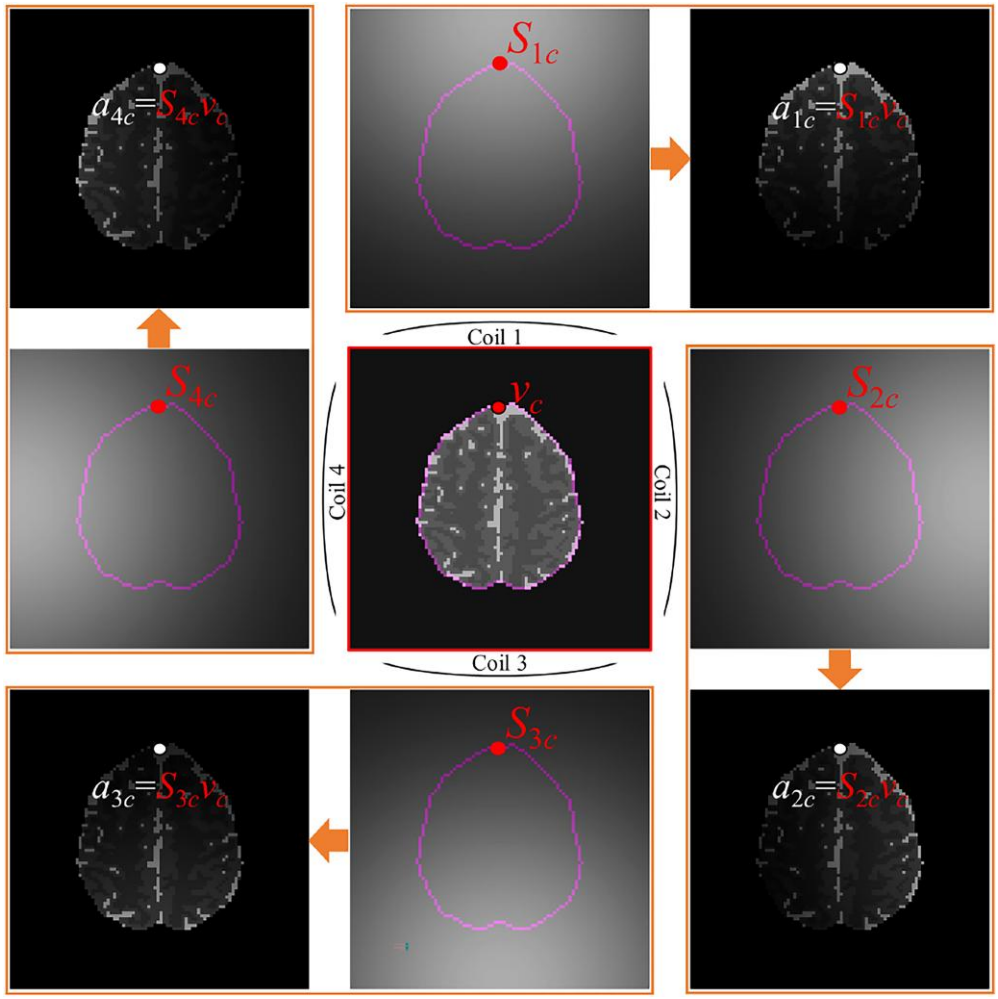

Figure 4. True slice image (centre) along with coil sensitivity profiles (top, bottom, left, right) and sensitivity weighted true images (the four corners). The coil sensitivity profiles are typically masked outside the brain but left here to show how the sensitivity decreases with voxels that are further from the coil.

However, skipping lines in k-space causes reconstructed coil-weighted brain images to appear folded over itself, or aliased, because the IFT cannot uniquely map the downsampled signals. We can see an example of this in Figure 5 where the IFT of the subsampled k-space (top right), with $n_A = 2$, causes the brain image to be aliased (bottom right). The depiction in Figure 5 only shows the aliasing for one of the coils, and since multiple coils are utilized in parallel imaging, a weighted aliased image transpires for each coil. It also only shows the magnitude images as the associated phase images are not shown.

Figure 6 shows a similar depiction of the full coil-weighted magnitude brain images to Figure 4, but introduces an acceleration factor of $n_A = 3$. The sequential subsampling pattern follows one similar to that shown in Figure 5 (left), but measuring every third line of k-space instead of every other line, resulting in aliased coil-weighted brain images. In Figure 6 (top right), the true aliased image is the point-wise multiplication of the given voxel by the sensitivity profile for coil 1 summed for the three strips, $a_{1c} = S_{11c}v_{1c} + S_{12c}v_{2c} + S_{13c}v_{3c}$. This process is repeated for a_{2c} in coil 2 (bottom right), a_{3c} in coil 3 (bottom left), and a_{4c} in coil 4 (top left). This depiction of four observed, complex-valued aliased images, a_c, along with the unobserved, complex-valued coil sensitivities, S_c, the unobserved, complex-valued unaliased voxel values, v_c, and the complex-valued measurement error, ε_c, create a linear system of complex-valued equations, shown in equation

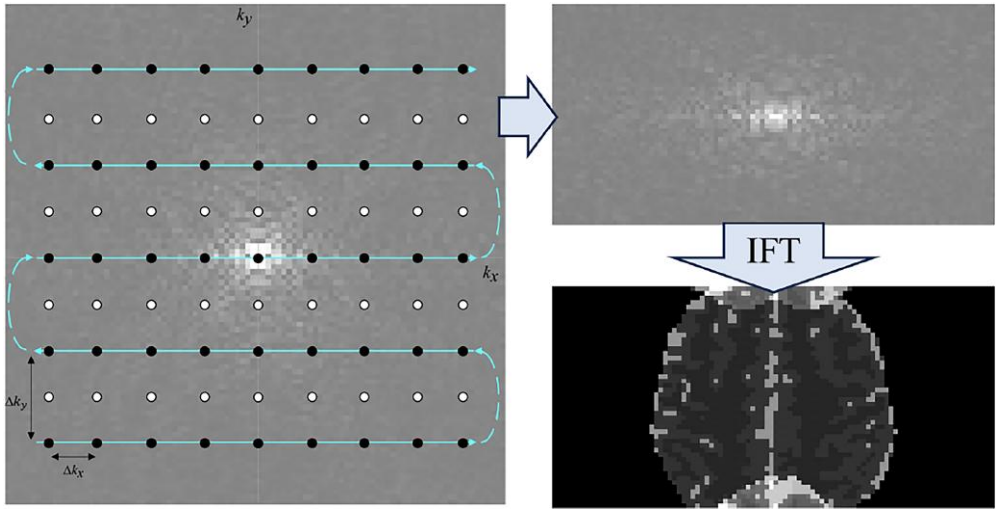

Figure 5. Subsampled *k*-space zig-zag coverage with $n_a = 2$ acceleration factor (left), the finalized subsampled *k*-space array after omitting the turn-around points (top right) and the aliased brain image after reconstruction using the IFT (bottom right).

(3.1). Since the unaliased voxel values, v_c, are the parameter of interest, SENSE estimates the coil sensitivities, S_c, treats it as a known parameter, and models the process as a complex-valued regression model,

$$\begin{bmatrix} a_{1c} \\ a_{2c} \\ a_{3c} \\ a_{4c} \end{bmatrix} = \begin{bmatrix} S_{11c} & S_{12c} & S_{13c} \\ S_{21c} & S_{22c} & S_{23c} \\ S_{31c} & S_{32c} & S_{33c} \\ S_{41c} & S_{42c} & S_{43c} \end{bmatrix} \begin{bmatrix} v_{1c} \\ v_{2c} \\ v_{3c} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1c} \\ \varepsilon_{2c} \\ \varepsilon_{3c} \\ \varepsilon_{4c} \end{bmatrix}. \tag{3.1}$$

BSENSE uses the isomorphic representation of equation (3.1), similar to equation (1.4) (Bruce et al., 2012). The likelihood, prior distributions, and posterior along with the parameter estimation are outlined in Section 2.

3.3 Prior assessment

The assessment for the hyperparameters is outlined in Section 2.2 and its application to the fMRI data is detailed in this subsection. The full pre-scan coil calibration images, which would be y_{0c} from Section 2.2, can be utilized to fully assess appropriate hyperparameters for the prior distributions in an automated way. For example, the n_{cal} (n_0) coil calibration images (top left of Figure 7) can be averaged together to give us full complex-valued coil images. An initial magnitude v_{0M} (β_{0M}) of the prior mean can be estimated for each voxel in the unaliased image by computing the Euclidean norm shown in the top right of Figure 7.

The n_C averaged coil calibration images can then be pointwise divided by v_{0M} to obtain a prior mean for the real and imaginary coil sensitivities, as displayed in the bottom of Figure 7. The phase of the coil sensitivities is estimated by $\arctan(I/R)/2$, where R and I are the real and imaginary components of the coil sensitivities, respectively. This phase is utilized to estimate complex-valued prior means for the coil sensitivities, H_0 (G_0). These coil sensitivity estimates, H_0, along with the full averaged calibration coil images are used to estimate complex-valued prior means for the voxel values, v_0 (β_0).

The hyperparameters n_S and n_v, which are the scalar weights of the prior means, are assessed to be the number of calibration images n_{cal}. The average residual variance over the voxels of the calibration images is calculated to obtain a prior for the noise variance noted as σ_0^2. The hyperparameters α (shape parameter of the inverse gamma) and β (scale parameter of the inverse

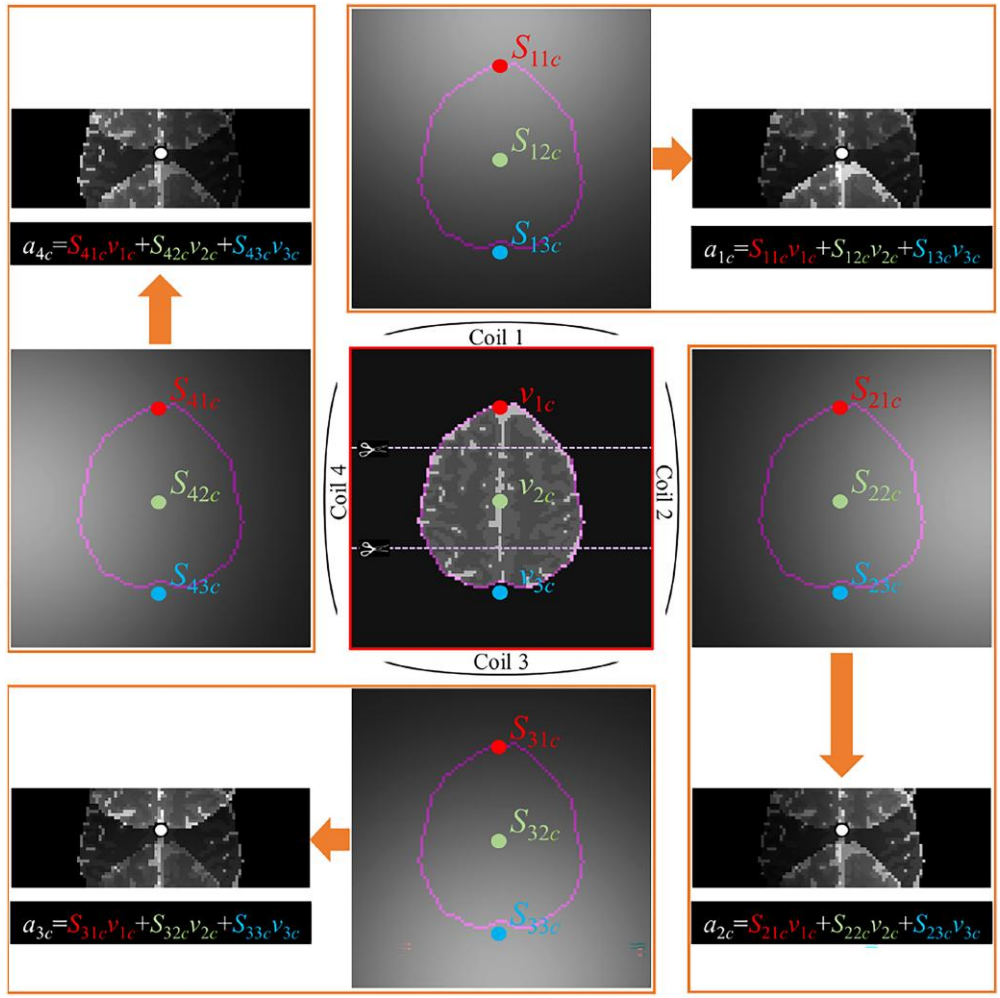

Figure 6. True slice image (centre) along with coil sensitivity profiles (top, bottom, left, right) and sensitivity weighted true aliased images (the four corners). The coil sensitivity profiles are typically masked outside the brain but left here to show how the sensitivity decreases with voxels that are further from the coil.

gamma) are assessed to be $\alpha = n_{\text{cal}} - 1$ and $\beta = (n_{\text{cal}} - 1)\sigma_0^2$. This prior information is incorporated to reconstruct each voxel measurement in the aliased coil image into the unaliased voxel values at every time in the fMRI series.

The software used for this research was MATLAB run on a 12th Gen Intel(R) Core(TM) i7-1255U laptop computer with 16 GB RAM, operating on Windows 11.

4 Simulation and experimental studies

4.1 Non-task data

A noiseless non-task image was used to create two series of 510 simulated full FOV coil images for one slice to mimic the experimental data shown in Section 4.6. The last n_{cal} time points of the first time series of non-task images served as calibration images that were utilized for hyperparameter assessment, and the second time series was used for a simulated non-task experiment. A complex-valued image was multiplied by a designed sensitivity map with $n_C = 8$ coils, similar to the four-channel coil shown in Figure 3 but with four additional coils in each corner as well, and then the series of images were Fourier transformed into full coil k-space arrays. In real-world MRI experiments, the first few images of the time series have increased signal as the magnetization reaches a

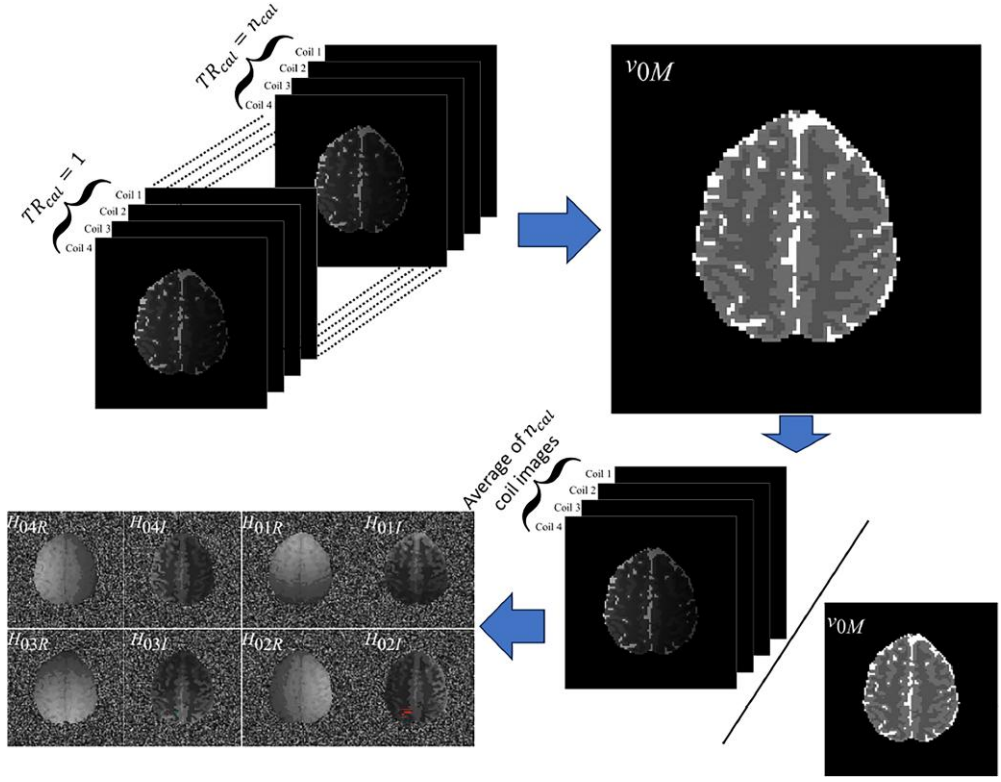

Figure 7. The n_{cal} calibration coil images (top left) are averaged through time and the Euclidean norm is taken yielding a prior mean for the magnitude unaliased voxel values v_{0M} (top right). The average of the coil calibration images are then pointwise divided by v_{0M} resulting in prior means for the real and imaginary parts of the coil sensitivities H_{0R} and H_{0I}, respectively.

steady state. The first three images in both the simulated series of non-task images are appropriately scaled, based on the experimental data, replicating the increased signal. These series were simulated by adding separate $N(0, 0.0036 n_x n_y)$ noise to both time series, where n_y and n_x are the number of rows and columns, respectively, in the full k-space array, to the real and imaginary parts of full coil k-space, corresponding to the noise in the fMRI experimental data used in Section 4. This data generation is following a general linear model with normally distributed noise and no spatial or temporal dependencies.

The arrays were then inverse Fourier transformed back into full coil images. To mimic the fMRI experiment shown in Section 4.6, the first 20 time points were discarded leaving 490 time points of non-task images for the single slice, though they could be used to estimate T_1 and magnetic field maps as described in Section 1.3 of the online supplementary material. The remaining 490 images in the time series were Fourier transformed and aliased by censoring lines in k-space according to the different acceleration factors used for the simulation, then back transforming the down-sampled data.

4.2 Reconstruction results

To analyse the reconstruction performance of BSENSE vs. SENSE, we first reconstructed aliased coil measurements at one time point, giving us a single unaliased image for both methods. For this, we used the first time point of the 490 simulated non-task time series with an acceleration factor of $n_A = 3$, shown in Figure 8.

The last $n_{\text{cal}} = 30$ time points, corresponding to 30 s, from the first 510 non-task full FOV calibration time series were utilized to assess the hyperparameters. The prior means from the calibration images for the unaliased voxels υ_0 and the sensitivity coils S_0 were used as initial values for H

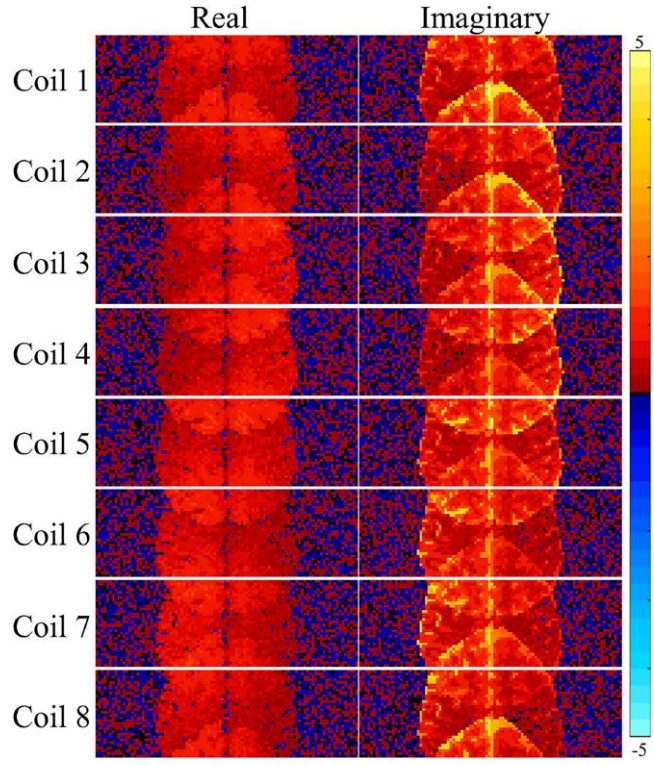

Figure 8. Simulated observed noisy aliased coil images for first time point in the non-task time series.

and ν. These initial values were used to generate a σ^2 value from the posterior conditional from equation (2.8), initializing the ICM algorithm and the Gibbs sampler. The simulated aliased coil images were reconstructed into a single, full brain images using the BSENSE MAP estimate from the ICM algorithm, the BSENSE MPM via MCMC, and traditional SENSE estimate. For the ICM algorithm, only three iterations were needed for estimating the parameters (computation time about 0.10 s per time point), and for the Gibbs sampling, 10,000 total iterations were run (computation time about 90 s per time point) with a burn of 2,500 leaving 7,500 iterations for estimation. For comparison, the computation time for SENSE is about 0.04 s per time points. Plots of the 10,000 iterations for a gray matter voxel at one single time point is displayed in Figure 9. The plot on the left of Figure 9 is the real part of the complex-valued voxel and the plot on the right is for the imaginary part. The red line in the plots show where the burn-in iterations end. Figure 9 shows that the Gibbs sampler converges relatively quickly.

Figure 10 displays the true simulated image (first column) along with the BSENSE MAP unaliased image (second column), the BSENSE posterior marginal mean unaliased image (third column), and the SENSE unaliased image (fourth column). We can see that the joint MAP estimate and the marginal posterior mean from BSENSE both produce magnitude and phase images that closely resemble the true non-aliased image in Figure 10 (left column). SENSE, on the other hand, produced an image with a higher noise level in the magnitude image resulting in less clear distinction between the different brain tissue when compared to our BSENSE approach and the true unaliased image. This is also evident by examining the noise level outside of the brain which is markedly higher in the SENSE reconstructed image. Typically, in fMRI studies, the voxels outside the brain are masked out, but here we leave them in to further show the spatial noise level of the reconstructed images for both techniques. Unlike the BSENSE and true phase images, SENSE also produced an unusable phase image with no anatomical structure. Activation using both magnitude and phase images has been shown to yield increased power of detection (Rowe, 2005; Rowe & Logan, 2004) and additional biological information (Petridou et al., 2006).

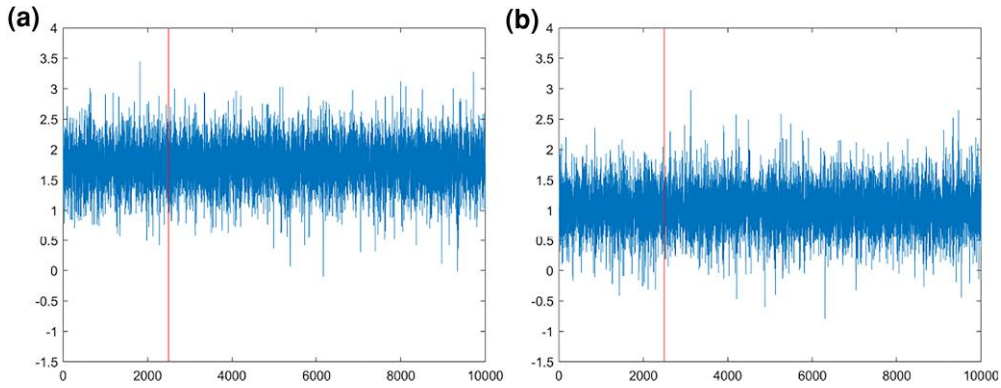

Figure 9. Time series of the 10,000 iterations from the Gibbs sampler with the real (left) and imaginary (right) components of a grey matter voxel. The line for both plots indicates the point where the burn-in iterations end.

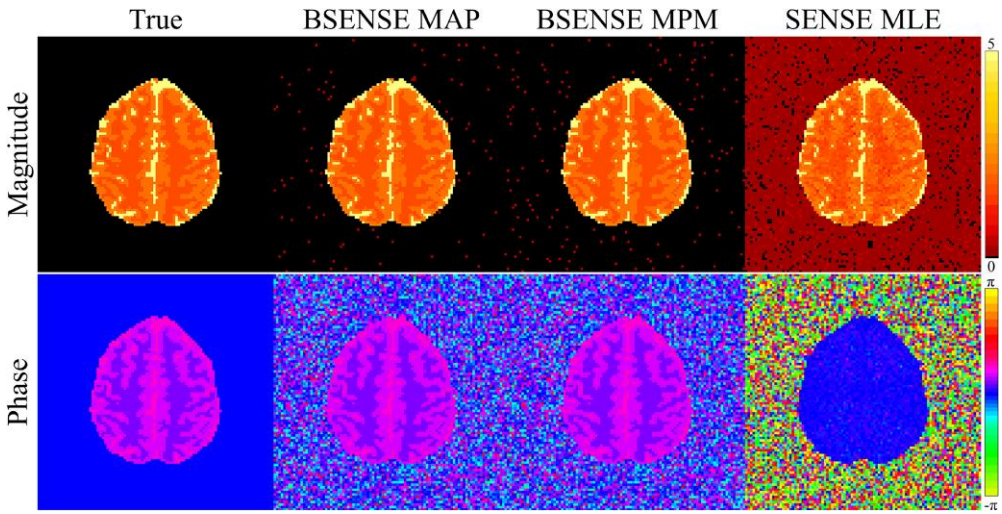

Figure 10. True non-task unaliased images (first column), BSENSE MAP unaliased non-task images (second column) using ICM, posterior mean BSENSE unaliased non-task images (third column) using Gibbs sampling, and SENSE non-task images (fourth column) with magnitude images in the first row and phase images in the second row.

To quantify the differences between the true and reconstructed magnitude images, we use the MSE to indicate the accuracy of a single reconstructed image compared to the true simulated image with lower MSE indicating a more accurate reconstructed image. The MSE is calculated by $MSE = \frac{1}{K}\sum_{j=1}^{K}(\nu_j - \overline{\nu}_j)^2$, where K is the number of voxels (either inside or outside the brain) in the full reconstructed image, ν_j is the reconstructed magnitude value of the jth voxel, and $\overline{\nu}_j$ is the true magnitude value of the jth voxel. The MSE for both BSENSE MAP and BSENSE MPM are <0.001 inside the brain and 0.001 outside the brain, respectively. For SENSE, the MSE was calculated to be 0.035 inside the brain and 0.03 outside the brain. This means that SENSE has a 26, 670% larger MSE inside the brain compared to BSENSE MAP and BSENSE MPM, and 2,940% and larger MSE outside the brain, respectively. The process illustrated here for reconstructing aliased coil images at a single time point can be replicated to reconstruct the rest of the series.

For the remaining results discussed in this article, only the BSENSE MAP estimate was used to reconstruct the time series of aliased coil. For the study covered in this article, we are only interested in a single estimate for each of the reconstructed images. From the Gibbs sampler, that is the posterior mean for each unaliased voxel, ν. Since the ν follows a normal distribution, the estimated

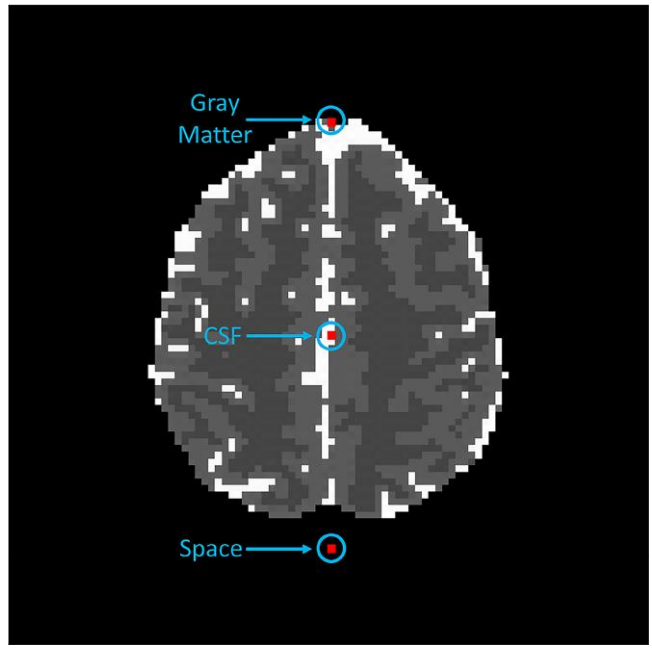

Figure 11. The dots indicate the three previously aliased voxels that are analysed in Table 1 and Figure 12. The top point is a grey matter voxel, the middle point is a cerebrospinal fluid voxel, and the bottom point is in space.

posterior mean and mode would be equivalent. This allows us to only need the MAP estimate for image reconstruction for this study, saving computation time.

Next, we reconstructed the 490 simulated time points with an acceleration factor of $n_A = 3$ using both BSENSE MAP and SENSE. Before evaluating the full reconstructed brain image results, we first analysed three reconstructed voxels that were previously aliased. The three previously aliased voxels are of different voxel types: cerebrospinal fluid (CSF), gray matter (GM), and outside the brain (Space). The location of these voxels are shown in Figure 11.

Table 1 shows the true magnitude of each voxel (row 1), the temporal magnitude means for BSENSE (row 2) and SENSE (row 3), and the MSE of the magnitude values for BSENSE (row 4) and SENSE (row 5). Phase analysis is also shown in Table 1 with the true phase value in row 6, the temporal mean phase values for BSENSE (row 7) and SENSE (row 8), and the MSE of the phase values for BSENSE (row 9) and SENSE (row 10). Along with being more accurate, the smaller the MSE the less bias the reconstructed voxel values are. With the MSE estimates relatively close to zero, this indicates that both methods have little to no bias. Note that the MSE estimates for BSENSE for both magnitude and phase is noticeably smaller than the MSE estimates for SENSE. This indicates that BSENSE more accurately reconstructs the unaliased voxel value compared to SENSE while having no bias from the true simulated magnitude and phases.

Figure 12 displays for the full time series for the three voxels analysed in Table 1 with the top plot showing the CSF voxel, the middle plot showing the gray matter voxel, and the bottom plot showing the voxel outside the brain. In each of the plots, the red time series is the SENSE reconstructed series, magenta is the temporal mean of the SENSE series, the black lines show the 95% confidence interval for the SENSE time series, The blue time series is the BSENSE reconstructed series, the purple line (not visible) is the temporal mean of the BSENSE series, the green lines show the 95% confidence interval for the BSENSE time series, and the light blue line with black dots shows the true value. The values in the plots in Figure 12 are the magnitude values. The plots for each of the voxels further illustrates how the BSENSE magnitude estimates are remarkably closer to the true simulated magnitude than SENSE is. This is exhibited by how close the true values (light blue line with black dots) are to the BSENSE values (blue time series) showing that there is no bias when using the Bayesian approach to the complex-valued latent variable model.

Table 1. Analysis of three previously aliased voxels comparing the true values to the reconstructed values from BSENSE and SENSE

	CSF voxel	GM voxel	Space voxel
True Mag.	5.6475	2.2704	0.0000
Mean BSENSE Mag.	5.6395	2.2601	0.0428
Mean SENSE Mag.	5.6149	2.2966	0.0332
BSENSE MSE Mag.	6.5e−05	0.0001	0.0018
SENSE MSE Mag.	0.0439	0.0470	0.0280
True Phase	0.7854	0.5236	0.0000
Mean BSENSE Phase	0.7847	0.5196	−0.8412
Mean SENSE Phase	−0.0045	0.0230	−0.7491
BSENSE MSE Phase	5.2e−07	1.6e−05	0.7088
SENSE MSE Phase	0.6253	0.2580	2.7581

Note. The aliased voxel contained a cerebrospinal fluid (CSF) voxel (column 1), a gray matter (GM) voxel (column 2), and a voxel outside the brain or out in space (column 3).

Then, we evaluated how the number of calibration images, n_{cal}, affected the reconstructed images. For the calibration image analysis, we fixed the acceleration factor to be $n_A = 3$ for the aliased coil measurements of the simulated non-task time series with $n_{TR} = 490$ time points and set the number of calibration images to be $n_{\text{cal}} = 5, 10, 15, 20, 25, 30$ for separate hyperparameter assessments. After assessing the hyperparameters using the different numbers of calibration images, the simulated non-task time series with the aliased coil measurements were reconstructed using BSENSE MAP and SENSE.

The results, displayed in Figure 13, indicate that increasing the number of calibration images decreases the noise level outside of the brain for BSENSE but has little effect inside the brain. Even the BSENSE MAP reconstruction with $n_{\text{cal}} = 5$ still produces an image with less noise than SENSE. The MSE of inside and outside the brain for both BSENSE and SENSE and the entropy for BSENSE and SENSE for the different number of calibration images were also calculated to quantify this result. Entropy analyses uncertainty and smoothness across a single image with lower entropy meaning less uncertainty throughout the image. The equation for entropy is given by $E = -\sum_{j=1}^{N} [\frac{\upsilon_j}{\upsilon_{\max}} ln(\frac{\upsilon_j}{\upsilon_{\max}})]$, where ln is the natural log, N is the number of voxels in the full reconstructed image, υ_j is the reconstructed magnitude value of the jth voxel, and $\upsilon_{\max}$ is the voxel intensity if all the image intensities were in one pixel given by $\upsilon_{\max} = \sqrt{\sum_{j=1}^{N} \upsilon_j^2}$ (Atkinson et al., 1997).

Shown in Figure 14a, the MSE for inside and outside the brain for the BSENSE MAP reconstructed magnitude images was immensely smaller than the SENSE reconstructed magnitude images. BSENSE also had much smaller entropy compared to SENSE, displayed in Figure 14b, as it decreased from 193.6 to 181.4 with the entropy for SENSE remaining around 218.5 as the number of calibration images increased. Lower MSE for BSENSE indicates a more precise reconstructed image while smaller entropy means less uncertainty with image reconstruction. Increasing the number of calibration images also decreases the temporal variance for BSENSE yielding increased SNR. For SENSE, the number of calibration images does not meaningfully affect the temporal variance, resulting in similar SNR for each n_{cal}. In all cases, the temporal variance for BSENSE is substantially lower than for SENSE. This demonstrates that BSENSE mitigates noise in the reconstructed image.

Along with analysis of the number of calibration images, we evaluated how different acceleration factors, n_A, affected the reconstructed images. Here, we fixed the number of calibration images to be $n_{\text{cal}} = 30$ for hyperparameter assessment and set the acceleration factors of the non-task time series to be $n_A = 2, 3, 4, 6, 8, 12$. For SENSE, the maximum acceleration factor was $n_A = 8$ since it cannot exceed the number of coils used as it yields a severely under-determined system of equations. These aliased coil measurements with separate acceleration factors were

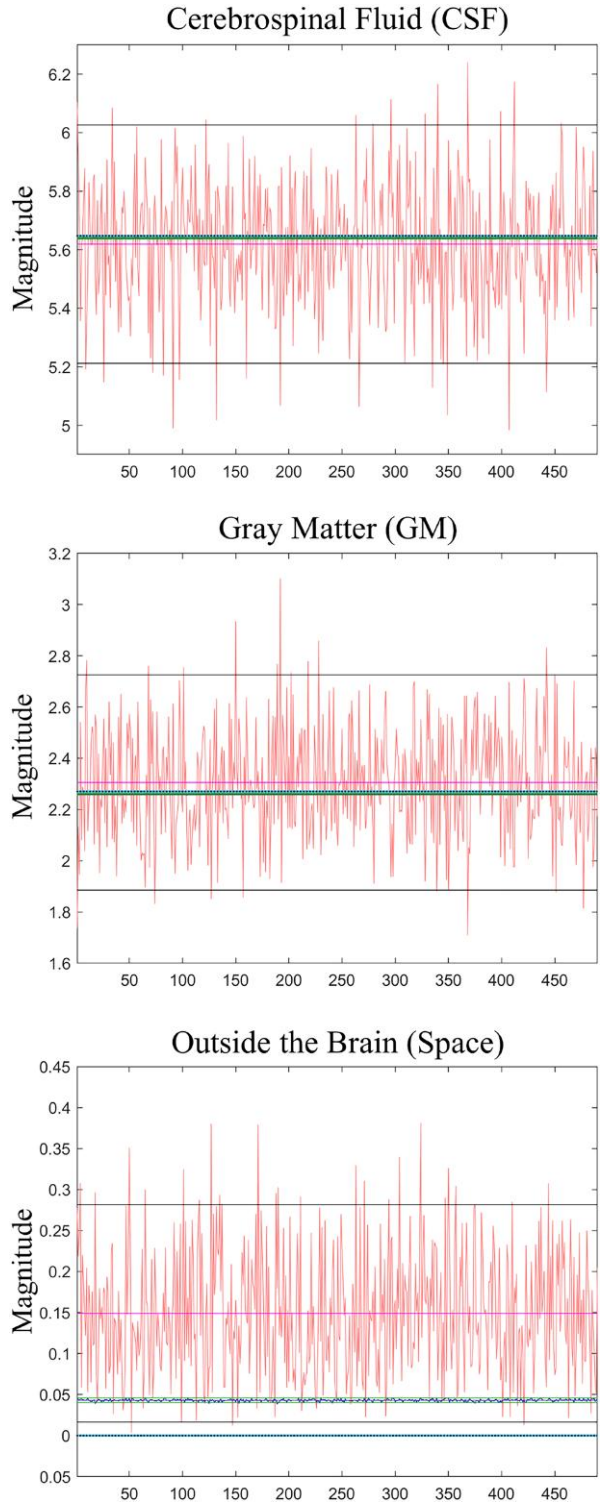

Figure 12. Time series of the magnitude of the same three previously aliased voxels from Table 1. For SENSitivity Encoding, the reconstruction over time is the highly varied series with the mean being at the centre of the time series and the lines 95% confidence interval being near the top and bottom of the plots. For Bayesian SENSE, the reconstruction over time series, the mean of the time series, and the 95% confidence interval are all close together the true magnitude (dotted line).

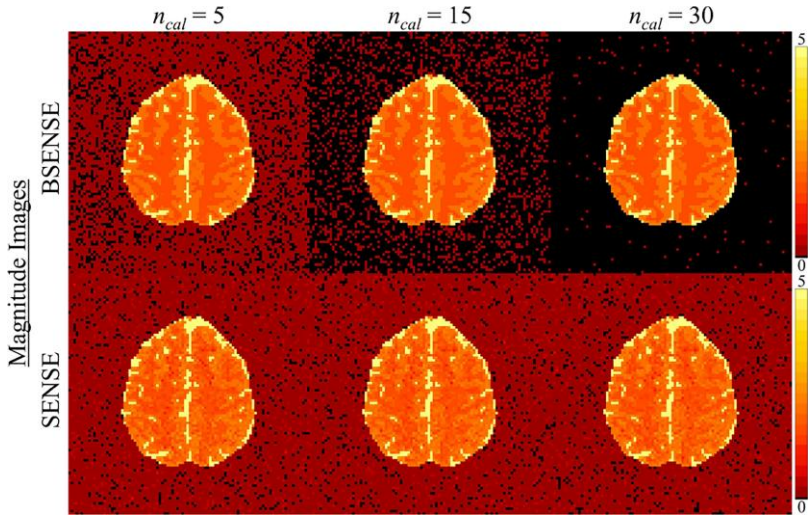

Figure 13. Reconstructed magnitude images for different number of calibration images using BSENSE MAP estimate (top row) and SENSE (second row).

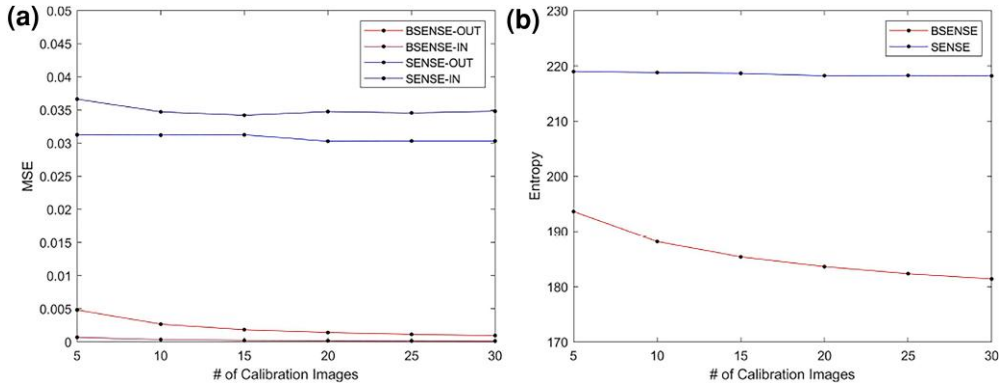

Figure 14. (a) Mean square error for inside and outside the brain for Bayesian SENSE (BSENSE) and SENSitivity Encoding (SENSE) comparing both method's reconstructed images to the true simulated magnitude image for each number of calibration images. (b) Entropy plot for BSENSE and SENSE for each number of calibration images.

reconstructed into full images using the BSENSE MAP estimate and SENSE, again comparing the results for both methods.

The results, exhibited in Figure 15, showed that the reconstructed magnitude images from BSENSE are negligibly affected by increasing the acceleration factor with SENSE being severely affected. The reconstructed phase images for BSENSE applying the different acceleration factors also closely resemble the true phase image while the reconstructed phase images for SENSE show zero phase inside the brain, rendering unusable phase information for anatomical and task analysis. These phase results are shown in Section 1.1 of the online supplementary material. The BSENSE temporal variance stays relatively the same (first row of Figure 16) with the increased acceleration factors, still producing high SNR (third row of Figure 16). SENSE was heavily influenced by the acceleration factor, as the reconstructed images with acceleration factors greater than $n_A = 3$ fail to produce usable images with distinct matter types throughout the brain as shown in Figure 15. The increased acceleration factor also markedly increases the temporal variance (second row of Figure 16) which substantially degrades the SNR of SENSE (fourth row of Figure 16).

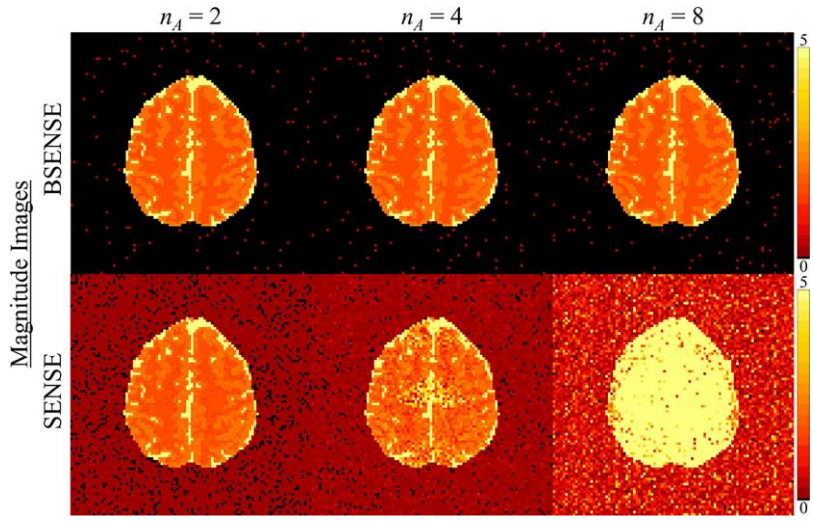

Figure 15. Reconstructed magnitude images for different acceleration factors using BSENSE MAP estimate (top row) and SENSE (second row).

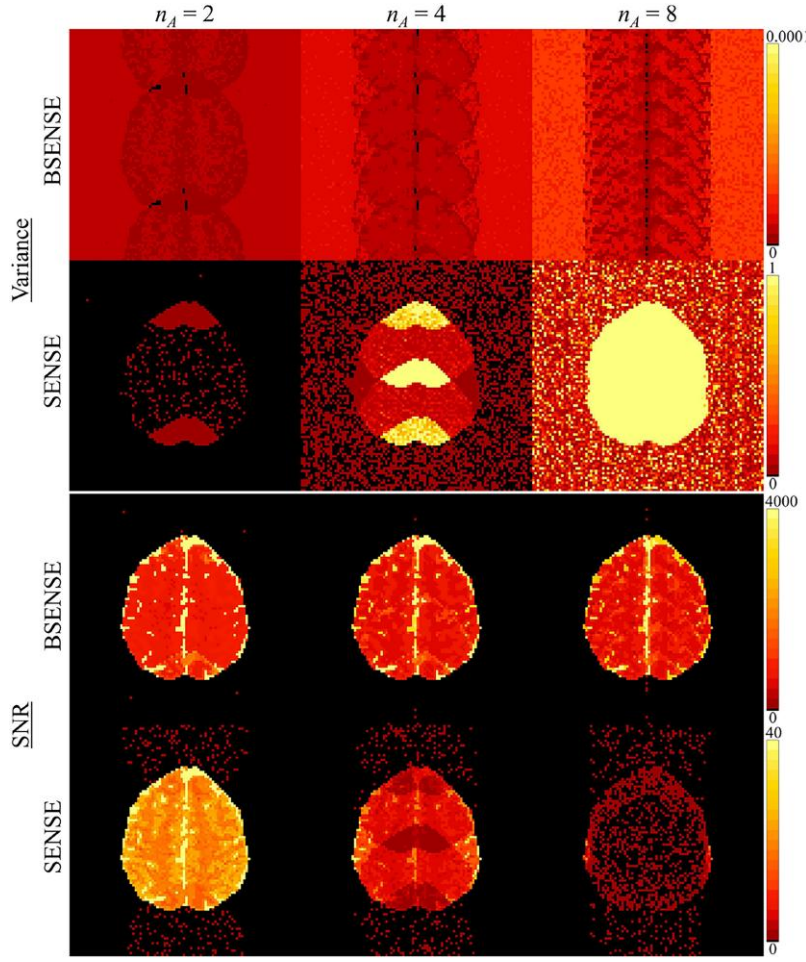

Figure 16. Temporal variance and SNR images for different acceleration factors using BSENSE MAP estimate (first row and third row, respectively) and SENSE (second row and fourth row, respectively).

We examined the reconstruction time of both methods for the different acceleration factors as well. The average time, in seconds, it took to reconstruct each image in the time series for BSENSE decreased from 0.1195 s to 0.0744 moving from an acceleration factor of 2 to 8 with a slight increase to 0.0855 with an acceleration factor of 12. SENSE does have a shorter reconstruction time with it decreasing from 0.0441 s to 0.0201 moving from an acceleration factor of 2 to 8. Despite this, our BSENSE approach still offers the potential for real time image reconstruction while producing remarkably better results in image reconstruction.

4.3 Task activation

In task-based fMRI, the non-task reconstructed images create a baseline value for each voxel giving us an intercept only simple linear regression $y = \beta_0 + \varepsilon$, where y is the unaliased voxel value. By adding in task activation to select images in the series of images, we have a simple linear regression $y = \beta_0 + x\beta_1 + \varepsilon$ for the unaliased voxel values. In this regression, β_0 is the baseline voxel value from the non-task reconstructed images determining the SNR $= \beta_0/\sigma$, and β_1 is the estimated task related increase from β_0 determining the contrast-to-noise ratio CNR $= \beta_1/\sigma$. The vector $x \in \{0, 1\}^{n_{\text{IMG}}}$, where n_{IMG} is the number of reconstructed images in the series, is a vector such that the zeros correspond to the images in the series without task activation and ones corresponding to the images with task activation. We can write this regression as $y = XB + \varepsilon$, where $X = [1,\ x] \in \mathbb{R}^{n_{\text{IMG}} \times 2}$ and $B = [\beta_0, \beta_1]'$.

Since the CNR is typically much lower than the SNR, the task is not usually visible on the reconstructed images. Instead, a right-tailed t-test is carried out with $\beta_1 \leq 0$ as the null hypothesis and $\beta_1 > 0$ as the alternative. The reason for the one-sided hypothesis test is because we anticipate an increased signal from the task activation. To simulate added task, a $\beta_1 = 0.045$ magnitude-only signal increase is added to the true noiseless non-task image with $\varepsilon \sim N(0, 0.0036)$ noise yielding a CNR of 0.75. A simulated phase task of $\pi/120$ was also added and analysed in Section 1.3 of the online supplementary material. This added task activation is located in the left motor cortex to resemble the region of interest (ROI) of brain activity from the fMRI unilateral right-hand finger-tapping experiment (Karaman et al., 2014). Ranges of tissue pixel intensities are 1.00–1.75 for white matter, 1.75–3.00 for grey matter, and 4.00–6.00 for the cerebral spinal fluid (CSF) in the simulation to mimic the experimental data discussed in Section 4.6.

4.4 fMRI time series data generation

A true noiseless task image along with a true noiseless non-task image were used to simulate a series of 510 full FOV coil images for one slice. The true images were multiplied by the same $n_C = 8$ coil sensitivity maps used for the non-task simulated time series, and then the series of images were Fourier transformed in full coil k-space arrays. This series was also generated by adding separate $N(0, 0.0036 n_x n_y)$ noise to the real and imaginary parts of the full coil k-space arrays and were then inverse Fourier transformed back into full coil images. To simulate the real-world fMRI experimental process, the series was generated by starting with 20 non-task time points. The scaling for the first few images in the fMRI simulated data was the same as the signal increases outlined in Section 4.1 for each of the tissue types. The initial 20 non-task time points are followed by 16 epochs alternating between 15 non-task and 15 task time points. An epoch is a stimulation period with time points of the subject at rest (non-task) and the subject performing an action or task. The series culminated with 10 non-task time points producing the simulated fMRI series of 510 images. To mimic the forthcoming fMRI experiment, the first 20 time points were discarded leaving 490 time points in the series. The last n_{cal} time points in the non-task time series from Section 4.1 were utilized as full FOV coil calibration images to assess the hyperparameters. For this simulation, we evaluate both BSENSE and SENSE using $n_C = 5, 10, 15, 20, 25, 30$ calibration images. The transformation and aliasing are the same as in Section 4.1. The different acceleration factors tested in this simulated fMRI experiment are $n_A = 2, 3, 4$.

4.5 fMRI time series reconstruction results

The hypothesis test described in Section 4.3 was utilized to determine voxels with a statistically significant signal increase. The statistically significant voxels for each number of calibration images were analysed for the BSENSE MAP reconstructed time series and the SENSE reconstructed

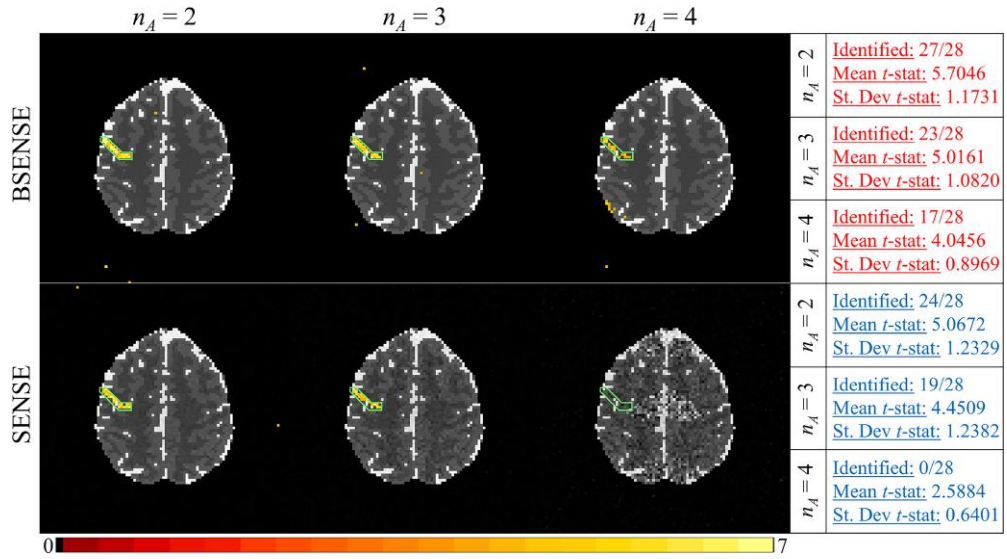

Figure 17. Statistically significant voxels in the region of interest (ROI) using false discovery rate (FDR) for Bayesian SENSE reconstructed images (first row), significant voxels in the ROI using FDR for SENSitivity Encoding (second row), and analysis of the *t*-statistics in the boxes on the right.

time series using the 5% false discovery rate (FDR) threshold procedure (Benjamini & Hochberg, 1995; Genovese et al., 2002; Logan & Rowe, 2004). The ROI here consists of 28 voxels located in the left motor cortex. Increasing the number of calibration images did not notably affect the detection of task for either BSENSE or SENSE, shown in Section 1.3 of the online supplementary material. The number of identified voxels with task activation and the mean value of the *t*-statistics was greater with BSENSE while having a smaller standard deviation compared to SENSE. These results indicate our BSENSE approach performs better with task detection compared to traditional SENSE, regardless of the number of calibration images that are utilized for hyperparameter assessment.

Figure 17 shows the statistically significant magnitude-only voxels from the BSENSE MAP reconstructed time series (first row) and the SENSE reconstructed time series (second row) for the different acceleration factors. Figure 17 also summarizes the *t*-statistics in the ROI for each acceleration factor. BSENSE identified more statistically significant voxels in the ROI for each acceleration factor while SENSE does not detect a single voxel when the acceleration factor was $n_A = 4$. The mean value for the *t*-statistics was again much higher for BSENSE with a lower standard deviation for the different acceleration factors compared to SENSE, demonstrating that BSENSE performs better when detecting task activation. Increasing the acceleration factor decreases the number of voxels identified and the mean of the *t*-statistics for both BSENSE and SENSE. The false positive rate for BSENSE for each of the acceleration factors were 0.033%, 0.033%, and 0.098%, respectively, while SENSE had 0.022%, 0%, and 0%, respectively.

4.6 Experimental data description

A 3.0 T General Electric Signa LX magnetic resonance imager was used to conduct an fMRI experiment on a single subject. The last $n_{\text{cal}} = 30$ full *k*-space arrays of a non-task series of 510 time points performed on the subject was inverse Fourier transformed into full coil brain images and used for hyperparameter assessment. A bilateral finger-tapping task was performed in a block design with an initial 20 s rest followed by 16 epochs with 15 s off and 15 s on. The experiment was concluded with 10 s of rest giving us a series of $n_{\text{IMG}} = 510$ repetitions with each repetition being 1 s, a flip angle of 90° and an acquisition bandwidth of 125 kHz. The data set consists of nine 2.5 mm thick axial slices with $n_C = 8$ receiver coils that have a 96×96 dimension for a 24 cm full FOV, with a posterior to anterior phase encoding direction. For this article, the time series

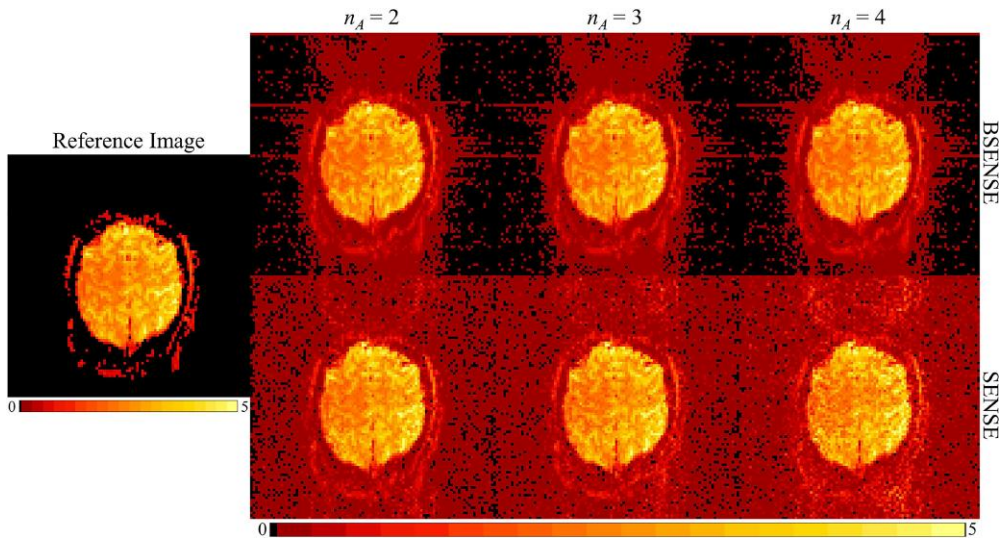

Figure 18. BSENSE MAP unaliased non-task magnitude images for each acceleration factor (first row) using the ICM algorithm, and SENSE unaliased non-task magnitude images for each acceleration factor (second row) with the magnitude reference image (left).

for all nine slices was used to analyse the effects of applying acceleration factors of $n_A = 2, 3, 4$ for both BSENSE and SENSE, but only the time series of the second slice is shown. Like the simulation study, the aliased coil images came from artificially skipping lines in the full coil k-space arrays, mimicking the effect of actually subsampling the coil k-space arrays. The first 20 images were discarded due to varying echo times and magnetization stability, leaving 490 time points for the fMRI experiment. The first 10 images not used for fMRI activation can be used to estimate a T_1 map (Karaman et al., 2014) as shown in Section 2.2 of the online supplementary material while the second 10 images could be used for static magnetic field mapping (Hahn et al., 2012), also discussed in the online supplementary material.

Before artificially aliasing the time series by omitting lines of k-space, a reference image (left image in Figure 18) was produced by taking the square norm between the $n_C = 8$ full FOV coil images at the first time point. This provides a magnitude image with which to compare to SENSE and our BSENSE. Rows of k-space in the fMRI experiment were omitted in each coil in accordance to the specified acceleration factors.

4.7 Experimental results

Similar to the process for the simulated data described in Section 4, each image in the entire time series of aliased coil measurements were simultaneously unaliased and combined using BSENSE and SENSE separately. Figure 18 displays the BSENSE MAP reconstructed images (top row) and the SENSE reconstructed images (bottom row) of the first time point of the 490 images using acceleration factors 2, 3, and 4. Just as the simulated results in Figure 15 demonstrated, the BSENSE reconstruction in Figure 18 produced clearer, less noisy images compared to SENSE reconstruction. It is noticeable inside of the brain where the signal is higher, but the distinction is strongest outside the brain.

MSE was again utilized to quantify the differences between the reference image and reconstructed images. The MSE for inside the brain for the BSENSE MAP was approximately 0.016 for each acceleration factor. For SENSE, the MSEs inside the brain were 0.030, 0.038, and 0.170 for acceleration factors 2, 3, 4, respectively. The MSE for outside the brain for BSENSE was 0.034, 0.033, and 0.033 for each of the acceleration factors while the MSE for SENSE was 0.061, 0.085, and 0.157. This means SENSE had an 88%, 140%, and 968% larger MSE inside the brain and an 81%, 160%, and 379% larger MSE outside the brain, respectively, for each acceleration factor. These results reflect the decreased noise from BSENSE vs. SENSE. The entropy

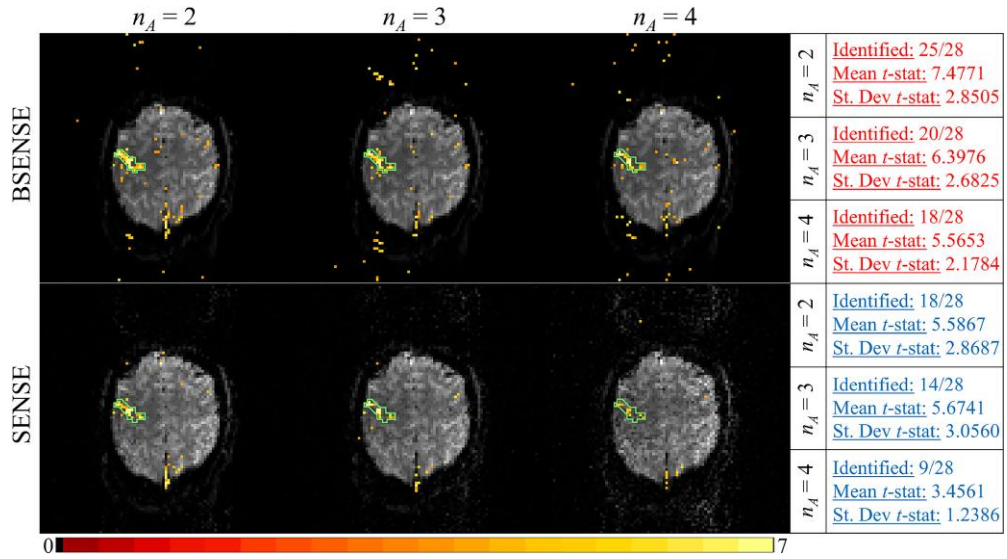

Figure 19. Statistically significant voxels in the region of interest (ROI) using false discovery rate (FDR) for Bayesian SENSE reconstructed images (first row) for three different acceleration factors, significant voxels in the ROI using FDR for SENSitivity Encoding (second row) for three different acceleration factors, and analysis of the *t*-statistics to the right of the images.

for BSENSE (191.0, 190.2, and 190.2, respectively) was also lower than the entropy for SENSE (214.9, 226.6, 243.3) indicating less uncertainty for each reconstructed image.

For the detection of task activation, the hypothesis test outlined in Section 4.3 was carried out. Figure 19 shows the statistically significant voxels under BSENSE (top row) and SENSE (bottom row) reconstruction. The images for the statistically significant voxels in Figure 19 for both methods use the 5% FDR threshold. Voxels outside the brain are usually masked out meaning the statistically significant voxels shown outside the brain in Figure 19 would typically not be there. Figure 19 also summarizes the *t*-statistics with BSENSE (red) and SENSE (blue). BSENSE correctly detected more voxels than SENSE as task activation in the ROI for all three acceleration factors. Our BSENSE approach also had a much higher mean *t*-statistic and lower standard deviation for all the acceleration factors. The false positive rate for BSENSE for each of the acceleration factors were 0.697%, 0.664%, and 0.642%, respectively, while SENSE had 0.283%, 0.163%, and 0.109%, respectively. We also evaluated BSENSE and SENSE task detection performance on the other eight slices which is outlined in Section 2.2 of the online supplementary material.

5 Discussion

In linear regression, having an unknown design matrix and complex-valued parameters can make parameter estimation of the regression coefficients more difficult. Here, we implement a Bayesian complex-valued latent variable linear model and applied it, along with the non-Bayesian model, to image reconstruction in fMRI. The results of the simulated and experimental studies showed that the Bayesian complex-valued latent variable model (BSENSE) outperformed the complex-valued non-Bayesian model (SENSE).

The BSENSE unaliased images were shown to more accurately reconstruct the truth compared to SENSE. The number of calibration images had minimal effect on the SENSE reconstructed images and its performance against BSENSE reconstructed images. Increasing the number of calibration images, however, did reduce the noise level in the BSENSE reconstructed images, leading to increased SNR. The results also indicated that the different acceleration factors had less influence on BSENSE than SENSE. BSENSE was able to successfully reconstruct images with an acceleration factor of up to 12, which was greater than the $n_C = 8$ coils used, without any aliasing artefacts or increasing the spatial variance but diminished activation. The SENSE reconstructed images

beyond acceleration factors of $n_A = 3$ were essentially unusable. Our BSENSE approach also had better performance when detecting the signal increase in the voxels that experienced task activation, as shown with both simulated and experimental data. The number of coils did not have a notable effect on our BSENSE approach which indicates that it works for any coil configuration as demonstrated in Section 1.1 of the online supplementary material. There was a noticeable effect on the SENSE image reconstruction. With four coils, the noise for the SENSE reconstructed images was higher compared to the 8, 12, and 16 simulated coil arrays as shown in Section 1.1 of the online supplementary material. This suggests that SENSE requires a deeper coil sensitivity map coverage to properly reconstruct images.

This article used the full posterior distribution for reconstructing images, meaning available prior information was quantified on all three parameters (ν, S, and σ^2) and utilized for parameter estimation. We have also analytically integrated out σ^2 yielding a marginal posterior where ν and S are the only two unknowns to be estimated. Integrating out σ^2 produces a joint Student-t posterior for S and ν from which we have obtained Gibbs sampling marginal estimates consistent with the three-parameter model.

For the task detection in this article, only the MAP estimate using the ICM algorithm was used to reconstruct the time series of non-task and task aliased images for capturing activation for both the simulated and experimental data. This is due to the Gibbs sampler being more computationally expensive when running a long series of images so it may not be as practical to use compared to evaluating the MAP estimate. This does not mean there is no value in running a Gibbs sampler, as it has the additional benefit of quantifying uncertainty. For instance, it could be utilized on a shorter series of images, provide us more statistical information about any voxel, or for hypothesis testing between two images. It is also possible to hybridize the ICM and Gibbs sampler with a couple of ICM steps followed by a short or no-burn Gibbs sampler. Our Bayesian approach allows for more options of how to run an fMRI experiment based on the objective of the scan compared to SENSE.

In this article, a magnitude-only activation model was utilized to detect task activation. In Section 1.3, phase-only activation for BSENSE is analysed with the results showing strong task detection power. Since the reconstructed images are complex-valued, our model is expected to be applicable for complex activation models for task detection (Rowe, 2005; Rowe & Logan, 2004). Further, an extension of our model would be to incorporate both covariance between the observations and covariance between the regression coefficients.

Acknowledgments

The authors thank the Wehr Foundation as this research was funded by the Computational Sciences Summer Research Fellowship (CSSRF) at Marquette University in the Department of Mathematical and Statistical Sciences.

Conflicts of interest: None declared.

Funding

This research was also funded by the Regular Research Grant (RRG) provided by the Committee on Research at Marquette University. A.B.'s work is supported by the National Science Foundation (DMS-2210686).

Data availability

The code and data are available per request and will be made publicly available via the arXiv immediately after the article has been accepted for publication.

Supplementary material

Supplementary material is available online at *Journal of the Royal Statistical Society: Series C.*

References

Atkinson D., Hill D. L. G., Stoyle P. N. R., Summers P. E., & Keevil S. F. (1997). Automatic correction of motion artifacts in magnetic resonance images using an entropy focus criterion. *IEEE Transactions on Medical Imaging*, *16*(6), 903–910. https://doi.org/10.1109/42.650886

Bandettini P., Jesmanowicz A., Wong E., & Hyde J. S. (1993). Processing strategies for time-course data sets in functional MRI of the human brain. *Magnetic Resonance in Medicine*, *30*(2), 161–173. https://doi.org/10.1002/mrm.v30:2

Bell A. J., & Sejnowski T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, *7*(6), 1129–1159. https://doi.org/10.1162/neco.1995.7.6.1129

Benjamini Y., & Hochberg Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bruce I. P., Karaman M. M., & Rowe D. B. (2012). The SENSE-isomorphism theoretical image voxel estimation (SENSE-ITIVE) model for reconstruction and observing statistical properties of reconstruction operators. *Magnetic Resonance Imaging*, *30*(8), 1143–1166. https://doi.org/10.1016/j.mri.2012.04.002

Cao X. R., & Liu R. W. (1996). General approach to blind source separation. *IEEE Transactions on Signal Processing*, *44*(3), 562–571. https://doi.org/10.1109/78.489029

Cardoso J. F., & Laheld B. H. (1996). Equivariant adaptive source separation. *IEEE Transactions on Signal Processing*, *44*(12), 3017–3030. https://doi.org/10.1109/78.553476

Chen S. H., Lee Y. S., & Wang J. C. (2018a). Complex-valued Gaussian process latent variable model for phase-incorporating speech enhancement. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 5439–5443). IEEE.

Chen S. H., Lee Y. S., & Wang J. C. (2018b). Locality preserving discriminative complex-valued latent variable model. In *24th International Conference on Pattern Recognition* (pp. 1169–1174). IEEE.

Choi, S., Cichocki, A., Park, H. M., & Lee, S. Y. (2005). Blind source separation and independent component analysis: A review. *Neural Information Processing - Letters and Reviews*, *6*(1), 1–57.

Cichocki A., Unbehauen R., & Rummert E. (1994). Robust learning algorithm for blind separation of signals. *Electronic Letters*, *30*(17), 1386–1387. https://doi.org/10.1049/el:19940956

Comon P. (1994). Independent component analysis, a new concept? *Signal Processing*, *36*(3), 287–314. https://doi.org/10.1016/0165-1684(94)90029-9

Gelfand A. E., & Smith A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of American Statistical Association*, *85*(410), 398–409. https://doi.org/10.1080/01621459.1990.10476213

Geman S., & Geman D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *6*(6), 721–741. https://doi.org/10.1109/TPAMI.1984.4767596

Genovese C. R., Lazar N. A., & Nichols T. E. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*, *15*(4), 870–878. https://doi.org/10.1006/nimg.2001.1037

Hahn A. D., Nencka A. S., & Rowe D. B. (2012). Enhancing the utility of complex-valued functional magnetic imaging detection of neurobiological processes through postacquisition estimation and correction of dynamic B_0 errors and motion. *Human Brain Mapping*, *33*(2), 288–306. https://doi.org/10.1002/hbm.v33.2

Hoerl A. E., & Kennard R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55–67. https://doi.org/10.1080/00401706.1970.10488634

Karaman M. M., Bruce I. P., & Rowe D. B. (2014). A statistical fMRI model for differential T_2* contrast incorporating T_1 and T_2* of gray matter. *Magnetic Resonance Imaging*, *32*(1), 9–27. https://doi.org/10.1016/j.mri.2013.09.001

King K. F., & Angelos L. (2001). SENSE image quality improvement using matrix regularization. In *Proceedings of the 9th Annual Meeting of ISMRM* (pp. 1771). ISMRM.

Kumar, A., Welti, D., & Ernst, R. R. (1975). NMR Fourier zeugmatography. *Journal of Magnetic Resonance Imaging*, *18*(1), 69–83. https://doi.org/10.1016/0022-2364(75)90224-3

Lee, T. W., Bell, A. J., & Orglmeister, R. (1997). Blind source separation of real world signals. In *Proceedings of International Conference on Neural Networks* (Vol. 4, pp. 2129–2134). IEEE.

Liang, Z. P., Bammer, R., Ji, J., Pelc, N. J., & Glover, G. H. (2002). Making better SENSE: wavelet denoising, Tikhonov regularization, and total least squares. In *Proceedings of the 10th Annual Meeting of ISMRM* (pp. 2388). ISMRM.

Lin F. H., Kwong K. K., Belliveau J. W., & Wald L. L. (2004). Parallel imaging reconstruction using automatic regularization. *Magnetic Resonance in Medicine*, *51*(3), 559–567. https://doi.org/10.1002/mrm.v51:3

Lindley D. V., & Smith A. F. M. (1972). Bayes estimates for the linear model. *Journal of the Royal Statistical Society, Series B*, *34*(1), 1–18. https://doi.org/10.1111/j.2517-6161.1972.tb00885.x

Liu B., King K., Steckner M., Xie J., Sheng J., & Ying L. (2009). Regularized sensitivity encoding (SENSE) reconstruction using Bregman iterations. *Magnetic Resonance in Medicine*, *61*(1), 145–152. https://doi.org/10.1002/mrm.v61:1

Logan B. R., & Rowe D. B. (2004). An evaluation of thresholding techniques in fMRI analysis. *Neuroimage*, *22*(1), 95–108. https://doi.org/10.1016/j.neuroimage.2003.12.047

Loizou P. C. (2013). *Speech enhancement*. CRC Press.

Macovski A. (1996). Noise in MRI. *Magnetic Resonance in Medicine*, *36*(3), 494–497. https://doi.org/10.1002/mrm.v36:3

Nguyen, A. H. T., Reju, V. G., Khong, A. W. H., & Soon, I. Y. (2017). Learning complex-valued latent filters with absolute cosine similarity. In *IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 2412–2416). IEEE.

Ogawa S., Lee T. M., Nayak A. S., & Glynn P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magnetic Resonance in Medicine*, *14*(1), 68–78. https://doi.org/10.1002/mrm.v14:1

O'Hagan A. (1994). *Kendall's advanced theory of statistics*. (Vol. *2B*). Wiley.

Petridou N., Dietmar P., Silva A. C., Loew M., Bodurka J., & Bandettini P. A. (2006). Direct magnetic resonance detection of neuronal electrical activity. *Proceedings of the National Academy of Sciences*, *103*(43), 16015–16020. https://doi.org/10.1073/pnas.0603219103

Pruessmann K. P., Weiger M., Scheidegger M. B., & Boesiger P. (1999). SENSE: Sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine*, *42*, 952–962. https://doi.org/10.1002/(ISSN)1522-2594

Roth Z., & Baram Y. (1996). Multidimensional density shaping by sigmoids. *IEEE Transactions on Neural Networks*, *7*(5), 1291–1298. https://doi.org/10.1109/TNN.72

Rowe D. B. (2005). Modeling both the magnitude and phase of complex-valued fMRI data. *Neuroimage*, *25*(4), 1310–1324. https://doi.org/10.1016/j.neuroimage.2005.01.034

Rowe D. B., & Logan B. R. (2004). A complex way to compute fMRI activation. *Neuroimage*, *23*, 1078–1092. https://doi.org/10.1016/j.neuroimage.2004.06.042

Schreier P. J., & Schraf L. L. (2010). *Statistical signal processing of complex-valued data: The theory of improper and noncircular signals*. Cambridge University Press.

Swami A., Giannakis G., & Shamsunder S. (1994). Multichannel ARMA processes. *IEEE Transactions on Signal Processing*, *42*(4), 898–913. https://doi.org/10.1109/78.285653

Tibshirani R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B*, *58*(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Williamson, D. S., Wang, Y., & Wang, D. (2013). A sparse representation approach for perceptual quality improvement of separated speech. In *IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 7015–7019). IEEE.

Williamson, D. S., Wang, Y., & Wang, D. (2014). A two-stage approach for improving the perceptual quality improvement of separated speech. In *IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 7034–7038). IEEE.

Yellin D., & Weinstein E. (1996). Multichannel signal separation: Methods and analysis. *IEEE Transactions on Signal Processing*, *44*(1), 106–118. https://doi.org/10.1109/78.482016

Ying, L., Xu, D., & Liang, Z. P. (2004). On Tikhonov regularization for image reconstruction in parallel MRI. In *26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 1056–1059). IEEE.