Turkish Journal of Electrical Engineering & Computer Sciences
http://journals.tubitak.gov.tr/elektrik/
Research Article
Turk J Elec Eng & Comp Sci
(2016) 24: 3240 – 3252

doi:10.3906/elk-1405-198

Bayesian compressive sensing framework for spectrum reconstruction in Rayleigh fading channels

Nadia IQBAL[1], **Asad MAHMOOD**[2], **Sajjad HUSSAIN**[3], **Abdul GHAFOOR**[4,*]
[1]Department of Electrical Engineering, Military College of Signals, National University of Sciences and Technology, Islamabad, Pakistan
[2]School of Computer Science, University of the Witwatersrand, Johannesburg, South Africa
[3]Department of Electronics Engineering, Mohammad Ali Jinnah University, Islamabad, Pakistan
[4]Department of Electrical Engineering, Military College of Signals,National University of Sciences and Technology, Islamabad, Pakistan

Received: 30.05.2014 • **Accepted/Published Online:** 01.04.2015 • **Final Version:** 15.04.2016

Abstract: Compressive sensing (CS) is a novel digital signal processing technique that has found great interest in many applications including communication theory and wireless communications. In wireless communications, CS is particularly suitable for its application in the area of spectrum sensing for cognitive radios, where the complete spectrum under observation, with many spectral holes, can be modeled as a sparse wide-band signal in the frequency domain. Considering the initial works performed to exploit the benefits of Bayesian CS in spectrum sensing, the fading characteristic of wireless communications has not been considered yet to a great extent, although it is an inherent feature for all sorts of wireless communications and it must be considered for the design of any practically viable wireless system. In this paper, we extend the Bayesian CS framework for the recovery of a sparse signal, whose nonzero coefficients follow a Rayleigh distribution. It is then demonstrated via simulations that mean square error significantly improves when appropriate prior distribution is used for the faded signal coefficients and thus, in turns, the spectrum reconstruction improves. Different parameters of the system model, e.g., sparsity level and number of measurements, are then varied to show the consistency of the results for different cases.

Key words: Rayleigh fading, Bayesian compressive sensing, belief propagation, mean square error performance

1. Introduction

In mobile communications, satellite communications, and other wireless communication systems, radio waves transmitted from the transmitter undergo scattering, reflection, and diffraction, thus creating a multipath effect. In consequence, the received signal level undergoes fluctuations and changes in amplitudes and phases of the signal occur. This phenomenon is called fading and, due to fading, study of practical channels like Rayleigh fading channels is required, along with that of classical additive white Gaussian noise (AWGN) channels.

In the context of wide-band spectrum sensing, cognitive radio (CR) [1] promises the sensing and detection of such wide-band signals propagating in the Rayleigh fading channels. However, conventional channel-by-channel sensing techniques like energy detector-based sensing [2], waveform-based sensing [3], and cyclostationary-based sensing [4] become costly and complex in this regard. To facilitate the wide-band spectrum sensing and detection for CRs, compressive sensing (CS) [5] [6] has emerged as a fascinating method

*Correspondence: abdulghafoor-mcs@nust.edu.pk

of acquisition of the wide-band signals at sampling rates that are significantly lower than the Nyquist rates. The information contained in the few large coefficients can be encoded by few random linear projections, while throwing away the rest of the coefficients that might be useless in further signal processing applications. Those encoded random projections can then be used to reconstruct the wide-band spectrum, which will contain the useful information mainly in the wide-band signals [5] [6].

The CS method can be further simplified by using Bayesian inference, provided that the system model supports the Bayesian approach. Since Bayesian inference provides solutions that are conditional on the observed data, it estimates a full probability model, where probability distributions are associated with the parameters or hypotheses and the process of decoding the signal via belief propagation (BP) is considered as a maximum a posteriori (MAP) estimation problem. Bayesian compressive sensing (BCS) facilitated with Bayesian inference provides precise estimation of the wide-band signal and reduces the number of measurements for the CS decoding process [7] [8]. In the BP decoding process, prior distribution plays the key role for the estimation of the signal coefficients. In our work, we have emphasized the fact that the prior distribution must be selected according to the system model.

In the CS via BP (CSBP) algorithm [7] [8], a sparse CS sensing matrix similar to low density parity check (LDPC) codes in the channel coding [9] [10] is employed as an encoding matrix for the recovery of the signal. A two-state Gaussian mixture prior distribution is used to determine the posterior density of the desired signal iteratively in the decoding process via BP, where the conditional probability density messages are passed between the signal elements and the measurement vector elements.

Another algorithm known as sparse Bayesian learning via BP (SBL-BP) [11] employs the BP approach to the SBL framework, where the posterior density of the signal is determined iteratively on the basis of a three-layer hierarchical prior model. To estimate the sparse transform coefficients in large-scale CS problems, the density messages in the BP decoding process are treated as Gaussian probability density functions for a hierarchical Bayesian model.

In [12], the authors developed an algorithm similar to SBL-BP, but its framework is based on a Gaussian scale mixture [9], in which the authors introduced low density frames sensing matrices and a special type of prior called the Jeffreys prior for CS signal recovery. However, larger numbers of operations are needed, with the higher degree of distribution for low density frames, which increases the number of iterations required for the convergence of the algorithm.

In the CS via Bayesian support detection (CS-BSD) and Bayesian hypothesis test via BP (BHT-BP) algorithms [13] [14], the authors introduced the detection directed (DD) estimation approach in the BCS framework for noisy CS using spike and slab prior [15]. The posterior density used for Bayesian hypothesis testing is estimated and updated iteratively. The posterior density of signal is used to estimate the signal support at each iteration, and then signal value is estimated based on the DD estimation structure for the estimated support. The performance of the algorithms is enhanced as the noise statistics are used in measurement message calculations.

In the Bayesian sparse reconstruction (BSR) algorithm [16], the authors considered the reconstruction of Gaussian sparse signals via BCS using a fast relevance vector machine (FRVM) in the presence of impulsive noise. However, the computational complexity of the algorithm is high compared to BCS via BP. Moreover, the algorithm is effective for signal reconstruction in an AWGN environment, in which the fading factor is not considered.

The authors of [17] implemented an enhanced multitask CS algorithm using Laplace priors. The correlated signals under consideration are assumed to be of same group and zero-mean Gaussian random variables with unit variance, which are recovered using a Laplace-based multitask CS (LMCS) algorithm that follows a Bayesian approach. The performance degrades if the signals under consideration belong to different groups, for which a minimum description length (MDL)-based LMCS algorithm is proposed. However, in these algorithms, the signals are considered to follow Gaussian distribution, while, in practice, the signals undergo fading. This factor was not considered.

In the above-mentioned CS-based signal reconstruction algorithms, signal propagation has been considered in an AWGN environment, or the signals are considered to be Gaussian sparse signals. The probability density messages in the BP decoding process are treated as Gaussian probability density functions. However, in practical scenarios, the signals undergo fading, for which CS signal reconstruction algorithms should be developed such that the density messages are chosen according to the signal model.

In this paper, we propose the usage of the BCS framework for spectrum sensing in a practical wireless channel like a Rayleigh fading channel. The novelty of our work is that we have formulated the prior distribution for the estimation of wide-band signals propagated in Rayleigh flat fading channels. When the signal undergoes fading due to the multipath affect, the performance of the system will be different for the classic AWGN channel and for the practical Rayleigh channel as well. The BCS framework is modified with appropriate priors for the decoding of the Rayleigh faded signal via BP. Our results show that when an appropriate prior is used according to the system model, the wide-band spectrum can be recovered such that the information bands are retained while noisy bands are suppressed. When energy detection is applied on the recovered spectrum, the occupied bands can be identified more efficiently. We estimate that when BCS is performed to reconstruct the wide-band spectrum, the bands containing useful information are not compressed. The spectrum estimation is verified by observing the mean square error (MSE) of the recovered spectrum compared to the original spectrum. MSE is an indicator of the error of the overall spectrum sensing and detection process. When MSE performance of estimation of the spectrum via CS improves, the overall error in the detection of the occupied bands is reduced.

We demonstrate that varying different parameters like sparsity level and number of measurements has an impact on the MSE performance of the BCS algorithm. When the prior is chosen wisely according to the system model, then MSE can be further reduced by varying these parameters.

The remainder of this paper is organized as follows. In Section 2 we formulate our problem. The appropriate prior models based on Rayleigh distribution are presented in Section 3. In Section 4, the BCS framework for sparse signal recovery in a Rayleigh fading channel using appropriate priors is discussed. Simulation results are then presented in Section 5, while Section 6 concludes the paper. Regarding notations, we have used lowercase and uppercase letters for scalars, lowercase bold letters for vectors, and uppercase bold letters for matrices.

2. Problem formulation

Practically in wireless communication, the wide-band signal from the transmitter travels over multiple reflective paths towards the receiver. Due to multipath fading, the envelope of the signal is statistically described by the Rayleigh probability density function (PDF). The classic BCS frame-work where a Gaussian prior is used for decoding via BP will not be sufficient for BCS recovery of wide-band signals. Thus, practical Rayleigh fading channels suggest the modification of the BCS framework accordingly.

In this paper, we have explored the implementation of the BCS framework in Rayleigh fading channels, where the desired unknown signal follows the Rayleigh distribution. To estimate the wide-band signal in Rayleigh fading channels, consider a wide-band signal $\mathbf{k}(t)$, such that for a time interval T, the signal is sampled at a sampling period of T_0, i.e.

$$\mathbf{k_t} = \{\mathbf{k}(t)\}|_{t=nT_0}, \quad n = 1, 2, 3, \ldots, N \tag{1}$$

Each signal sample in $\mathbf{k_t}$, i.e. $\mathbf{k_t}(n)$, undergoes a Rayleigh flat fading channel such that $\mathbf{h}$ vector consisting of i.i.d. channel coefficients is given by

$$\mathbf{h} = \{\mathbf{h}(i)\}, \quad i = 1, 2, 3, \ldots, N \tag{2}$$

where each coefficient $h_i \in \mathbf{h}$ takes on some constant value independent of other channel coefficients. Thus, a Rayleigh faded signal vector $\mathbf{x_t} \in \mathbb{R}^N$ is obtained:

$$\mathbf{x_t} = \mathbf{h} * \mathbf{k_t}, \tag{3}$$

where $*$ represents the convolution operation. The wide-band spectrum for this wide-band signal in the Rayleigh flat fading channel will be sparse, given as:

$$\mathbf{x_f} = \mathcal{F}(\mathbf{x_t}) = \mathbf{H}\mathbf{k_f} \tag{4}$$

where $\mathbf{H}$ contains the frequency domain samples of each Rayleigh fading channel coefficient and $\mathbf{k_f}$ is the sparse wide-band spectrum of the signal $\mathbf{k_t}$ that undergoes Rayleigh fading.

The sparsity level of the signal $\mathbf{x_f}$ is given by some vector $\mathbf{r}$, i.e.

$$\|\mathbf{r}\|_0 = K \ll N, \tag{5}$$

where $\|\mathbf{r}\|_0$ represents the number of nonzero coefficients in $\mathbf{x_f}$. Thus, the sparse Rayleigh faded signal $\mathbf{x_f}$ contains only K nonzero coefficients out of N samples.

Sensing/sampling of such wide-band signals would be costly and complex as well, if the Nyquist sampling method is used. However, this problem can be solved using the BCS approach. The sparse nature of this Rayleigh faded signal allows us to estimate the signal via BCS with smaller numbers of measurements. A sensing matrix $\mathbf{\Phi} \in \mathbb{R}^{M \times N}$ provides us with the measurement vector $\mathbf{y}$ such that

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x_f} + \mathbf{n}, \tag{6}$$

where the additive noise vector $\mathbf{n} \in \mathbb{R}^M$ is drawn from i.i.d. zero-mean Gaussian noise distribution $\mathcal{N}(0, \sigma_n^2 \mathbf{I}_M)$, where σ_n^2 represents the variance of noise distribution and $\mathbf{I}_M$ is an identity matrix of order $M \times M$ and the number of measurements $M \ll N$. Our goal is to estimate the sparse Rayleigh distributed signal $\mathbf{x_f}$ via BCS. The Bayesian inference improves the performance of BCS algorithms such that the BCS recovery of the signal is considered as a MAP estimation problem [7] [8], given by

$$\hat{\mathbf{x_f}} = \arg\max_{\mathbf{x_f}} f_{\mathbf{X_f}}(\mathbf{x_f}|\mathbf{y}) \quad s.t.\ \mathbf{E}\|\mathbf{\Phi}\mathbf{x_f} - \mathbf{y}\| \leq \epsilon, \tag{7}$$

where $\hat{\mathbf{x_f}}$ is the estimated signal via BCS, $f_{\mathbf{X_f}}(\mathbf{x_f}|\mathbf{y})$ is the conditional PDF of $\mathbf{x_f}$, $\mathbf{E}[.]$ is the expectation operator, and ϵ is a user defined parameter for tolerance in error. The BCS approach offers solutions with low computational cost for the recovery of the signal. To ensure accurate and successful recovery of the signal, the minimum number of measurements required is $M \geq cK\log(N/K)$, where c is some constant. This condition also shows that recovery of the signal also depends on the sparsity level K of the signal.

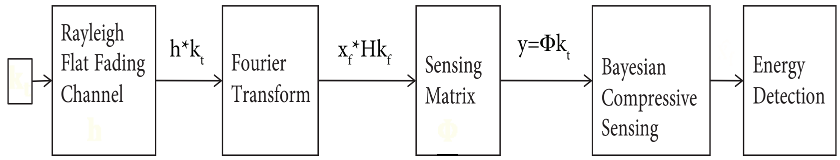

Figure 1. Wide-band spectrum estimation via BCS in Rayleigh fading channel.

The overall flow of the BCS approach for signal propagation in a Rayleigh fading channel is shown in Figure 1.

In the above figure, we estimate the spectrum via BCS such that MSE is minimum. This reduces the overall error in the detection of occupied bands in the wide-band spectrum. In the following sections, we proposed our system model for the BCS framework and then we discussed the solution approach for the recovery of the Rayleigh faded signal $\mathbf{x_f}$ via BCS.

3. Proposed system model based on Rayleigh prior distribution

The system model and the prior model should be chosen wisely according to the sparse Rayleigh faded signal under consideration. In this section, we discuss the graphical representation of the sensing matrix $\mathbf{\Phi}$ and the prior model to be used in the BCS framework for Rayleigh fading channels.

3.1. Sparse sensing matrix

To encode the sparse Rayleigh faded signal $\mathbf{x_f}$, we consider a sparse-binary sensing matrix $\mathbf{\Phi} \in \{0,1\}^{M\times N}$, where $M \ll N$. The sparsity of the sensing matrix is kept low and it is classified as fixed row weight R matrix, where each row of $\mathbf{\Phi}$ contains exactly R nonzero entries. Thus, maximum signal elements are sensed for smaller numbers of measurements. The sensing matrix for the Bayesian framework can be represented as a bipartite graph. We assume that a bipartite graph $\mathcal{G} = (\mathcal{U}, \mathcal{V}, \mathcal{E})$ represents the neighboring relations in $\mathbf{y} = \mathbf{\Phi x_f}$, where $\mathcal{U} = \{1,2,\ldots,N\}$ represents the set of indices corresponding to each element in $\mathbf{x_f}$, $\mathcal{V} = \{1,2,\ldots,M\}$ represents the set of indices corresponding to each measurement vector $\mathbf{y}$ element, and the set of edges connecting $\mathcal{U}$ and $\mathcal{V}$ can be defined as $\mathcal{E} = \{(j,i) \in (\mathcal{U}\times\mathcal{V}) | \phi_{ji} = 1\}$, where ϕ_{ji} represents the $(j,i)th$ element in $\mathbf{\Phi}$. The neighbor set of $\mathcal{U}$ will be $\mathcal{T}_{\mathcal{U}}(i) = \{j \in \mathcal{V} | (j,i) \in \mathcal{E}\}$ and the neighbor set of $\mathcal{V}$ will be $\mathcal{T}_{\mathcal{V}}(j) = \{i \in \mathcal{U} | (j,i) \in \mathcal{E}\}$.

3.2. Prior model

In the Bayesian framework, the selection of prior distribution according to the signal model plays a key role in the estimation of the sparsest solution of an underdetermined system. The classical AWGN channel suggests priors like the Gaussian–Spike prior or two-state Gaussian mixture prior for the recovery of the sparse signal. Practically, however, the signal undergoes fading and such sparse Rayleigh faded signals cannot be recovered accurately via BCS using the above-mentioned priors. For the recovery of the signal propagated in practical Rayleigh fading channels, the priors used in BP should be modified according to the Rayleigh fading system model. In this section, we have modified the BCS approach for signal propagation in Rayleigh fading channels in terms of prior distributions used for BP decoding of the wide-band signal. Here we have considered two types of system models, i.e. a noiseless system model and a noisy system model.

In the noiseless system model, the Rayleigh faded signal $\mathbf{x_f}$ is encoded without considering the noise factor, i.e. $\mathbf{y} = \mathbf{\Phi x_f}$, while in the noisy system model, AWGN also accounts for computation of the measurement vector, i.e. $\mathbf{y} = \mathbf{\Phi x_f} + \mathbf{n}$, where the additive noise vector $\mathbf{n} \in \mathbb{R}^M$ is drawn from i.i.d. zero-mean Gaussian noise distribution $\mathcal{N}(0, \sigma_n^2 \mathbf{I}_M)$, where σ_n^2 represents the variance of noise distribution and $\mathbf{I}_M$ is an identity matrix of order $M \times M$.

We have shown that different features of each system model require a different prior model for optimal employment in BP for the BCS decoding algorithm.

3.2.1. Prior model for noiseless system model

If $\mathbf{s}$ represents the state vector for $\mathbf{x_f}$ elements, then each signal coefficient $i \in \mathcal{U}$ in $\mathbf{x_f}$ can be associated either with high state, i.e. $s(i) = 1$, or with low state, i.e. $s(i) = 0$, where $s(i) \in \mathbf{s}$. For Rayleigh faded signal $\mathbf{x_f}$, the high state element is related to the Rayleigh distribution as

$$f_{\mathbf{X_f}}(\mathbf{x_f}|\mathbf{s}=1) = \left(\frac{\mathbf{x_f}}{\sigma_1^2}\right) e^{\left(\frac{-\mathbf{x_f}^2}{2\sigma_1^2}\right)}, \tag{8}$$

where σ_1^2 represents the variance of the Rayleigh distribution for signal coefficients associated with a high state. Similarly, the low state element in $\mathbf{x_f}$ can be related to a Spike distribution model as

$$f_{\mathbf{X_f}}(\mathbf{x_f}|\mathbf{s}=0) = \delta(\mathbf{x_f}), \tag{9}$$

where $\delta(\mathbf{x_f})$ represents a Dirac distribution having a nonzero value in the range $\mathbf{x_f} \in (0-, 0+)$, such that $\int_0^\infty \delta(\mathbf{x_f})\, \mathbf{dx_f} = 1$. The Rayleigh–Spike prior density associated with noiseless system model $\mathbf{y} = \mathbf{\Phi x_f}$ is given by

$$f_{\mathbf{X_f}}(\mathbf{x_f}) = q(\frac{\mathbf{x_f}}{\sigma_1^2}) e^{\left(\frac{-\mathbf{x_f}^2}{2\sigma_1^2}\right)} + (1-q)\delta(\mathbf{x_f}). \tag{10}$$

Such types of priors have been used for strictly sparse signal models [13] [14] [18], as the prior easily characterizes the signals by σ_1 and q factors. Since the elements in $\mathbf{x_f}$ are assumed to be i.i.d. Rayleigh distributed, the prior density represents each ith signal coefficient independently.

3.2.2. Prior model for noisy system model

For the noisy system model $\mathbf{y} = \mathbf{\Phi x_f} + \mathbf{n}$, the prior model should account for the Gaussian distribution associated with additive noise vector $\mathbf{n}$ along with Rayleigh distribution and Dirac distribution. For high state signal elements, when i.i.d. Rayleigh distributed coefficient $\mathbf{x_f}$ adds up with i.i.d. Gaussian noise coefficients n, then the resultant vector $\mathbf{z} = \mathbf{x_f} + \mathbf{n}$ would follow the probability distribution that is obtained after the convolution of probability distribution of $\mathbf{x_f}$, i.e. $f_{\mathbf{X_f}}(\mathbf{x_f}) = (\frac{\mathbf{x_f}}{\sigma_1^2}) e^{\left(\frac{-\mathbf{x_f}^2}{2\sigma_1^2}\right)}$, and the probability distribution of noise coefficient $\mathbf{n}$, i.e. $f_{\mathbf{N}}(\mathbf{n}) = (\frac{1}{\sqrt{2\pi\sigma_n^2}}) e^{\left(\frac{-\mathbf{n}^2}{2\sigma_n^2}\right)}$. The probability distribution for $\mathbf{z}$ is given by

$$f_{\mathbf{Z}}(\mathbf{z}|\mathbf{s}=1) = \int_0^\infty f_{\mathbf{X_f}}(\mathbf{x_f}) f_{\mathbf{X_f}}(\mathbf{z}-\mathbf{x_f})\, \mathbf{dx_f}, \tag{11}$$

where $\mathbf{n} = \mathbf{z} - \mathbf{x_f}$,

$$f_{\mathbf{Z}}(\mathbf{z}|\mathbf{s}=1) = \frac{1}{\sqrt{2\pi}\sigma_n\sigma_1^2}\int_0^{\infty} \mathbf{x_f} e^{\frac{-\mathbf{x_f}^2}{2\sigma_1^2}} e^{\left(\frac{-(\mathbf{z}-\mathbf{x_f})^2}{2\sigma_n^2}\right)} \mathbf{dx_f}. \tag{12}$$

Evaluating this integral via integration by parts yields

$$f_{\mathbf{Z}}(\mathbf{z}|\mathbf{s}=1) = \frac{\sigma_1 z}{(\sigma_1^2+\sigma_n^2)^{\frac{3}{2}}}\phi\left(\frac{\sigma_1}{\sigma_n}\frac{\mathbf{z}}{\sqrt{\sigma_1^2+\sigma_n^2}}\right) + \frac{\sigma_n}{\sqrt{2\pi}\sigma_1^2+\sigma_n^2} e^{\left(\frac{-\mathbf{z}^2}{2\sigma_n^2}\right)}, \tag{13}$$

where $\phi\left(\frac{\sigma_1}{\sigma_n}\frac{\mathbf{z}}{\sqrt{\sigma_1^2+\sigma_n^2}}\right)$ is the cumulative distribution function (CDF) of a standard normal random variable.

The low state signal coefficient in the noisy system model is represented by Dirac distribution, as it is special form of the Gaussian density when $\sigma_0 \to 0$ [15]. The prior density for noisy system model would be

$$f_{\mathbf{X_f}}(\mathbf{x_f}) = q\frac{\sigma_1 \mathbf{z}}{(\sigma_1^2+\sigma_n^2)^{\frac{3}{2}}}\phi\left(\frac{\sigma_1}{\sigma_n}\frac{\mathbf{z}}{\sqrt{\sigma_1^2+\sigma_n^2}}\right) + \frac{\sigma_n}{\sqrt{2\pi}\sigma_1^2+\sigma_n^2} e^{\left(\frac{-z^2}{2\sigma_n^2}\right)} + (1-q)\delta(\mathbf{x_f}). \tag{14}$$

In the following section, the BCS approach to estimate the signal is discussed. The proposed prior density models have been used for MAP estimation of each signal coefficient signal $\mathbf{x_f}_i$.

4. Bayesian compressed sensing framework based on appropriate priors

To determine the solution to underdetermined system models discussed above and estimate the Rayleigh faded signal $\mathbf{x_f}$ via BCS, we propose BCS decoding of the signal via BP using the proposed prior density.

As discussed earlier, the signal is estimated via BP as a MAP estimation problem. This can be represented as

$$\hat{\mathbf{x}}_{\mathbf{f}MAP} = \arg\max f_{\mathbf{X_f}}(\mathbf{x_f}|\mathbf{Y}=\mathbf{y}). \tag{15}$$

According to the Bayesian theorem,

$$f_{\mathbf{X_f}}(\mathbf{x_f}|\mathbf{Y}=\mathbf{y}) = \frac{f_{\mathbf{Y}}(\mathbf{y}|\mathbf{X_f}=\mathbf{x_f}) f_{\mathbf{X_f}}(\mathbf{x_f})}{f_{\mathbf{Y}}(\mathbf{y})}, \tag{16}$$

where $f_{\mathbf{X_f}}(\mathbf{x_f}|\mathbf{Y}=\mathbf{y})$ is the posterior distribution for $\mathbf{x_f}$, $f_{\mathbf{Y}}(\mathbf{y}|\mathbf{X_f}=\mathbf{x_f})$ is the likelihood of the estimate, and $f_{\mathbf{Y}}(\mathbf{y})$ is the marginal distribution. In Bayesian inference, marginal distribution plays no significant role in estimating posterior distribution. It only marginalizes the posterior distribution. Thus, MAP estimation becomes

$$\hat{\mathbf{x}}_{\mathbf{f}MAP} = \arg\max f_{\mathbf{Y}}(\mathbf{y}|\mathbf{X_f}=\mathbf{x_f}) f_{\mathbf{X_f}}(\mathbf{x_f}) \tag{17}$$

where $\hat{\mathbf{x}}_{\mathbf{f}MAP}$ represents the estimated sparse signal $\mathbf{x_f}$ via MAP estimation. It is obvious from the above equation that an accurate MAP estimation for $\mathbf{x_f}$ can be done when the sparsifying prior $f_{\mathbf{X_f}}(\mathbf{x_f})$ is selected according to the system model.

BP provides the posterior density for MAP estimation at every iteration. The prior density initializes the BP process. Afterwards, the posterior density of each signal element $\mathbf{x_f}_i$ is obtained and updated, while messages are being passed between the edges in factor graphs. Here sampled messages are considered to implement BP, where each message consists of the samples of the PDF [7] [8]. The sampled message approach is adaptive in the sense that various system models could be retrieved using it. When sampling step size is sufficient enough to store the PDF of the message under consideration, it also ensures faster convergence compared to the parametric-message approach [19] [20] [21]. In CS decoding via BP we follow the Bayesian rule, where the posterior density of each signal element $\mathbf{x_f}_i$ is represented in form of $Posterior\ Density = \frac{likelihood}{evidence} \times Prior\ Density$. The marginal posterior density $f_{\mathbf{X_f}_i}(\mathbf{x_f}|\mathbf{Y}=\mathbf{y})$ is given by

$$f_{\mathbf{X_f}_i}(\mathbf{x_f}|\mathbf{Y}=\mathbf{y}) = \frac{f_{\mathbf{Y}}(\mathbf{y}|\mathbf{X_f}_i = \mathbf{x_f}_i)}{f_{\mathbf{Y}}(\mathbf{y})} \times f_{\mathbf{X_f}}(\mathbf{x_f}), \tag{18}$$

where $f_{\mathbf{Y}}(\mathbf{y}|\mathbf{X_f}_i = \mathbf{x_f}_i)$ is the likelihood, $f_{\mathbf{X_f}}(\mathbf{x_f})$ is the prior density for signal $\mathbf{x_f}$, and the evidence $f_{\mathbf{Y}}(\mathbf{y})$ only marginalizes the posterior density and does not enter into determining the relative properties. The measurements associated with $\mathbf{X_f}_i$ according to the tree-like property of $\mathbf{\Phi}$, i.e. $\{\mathbf{Y}_k : k \in \mathcal{T}_{\mathcal{U}}(i)\}$, are statistically independent for $\mathbf{X_f}_i = \mathbf{x_f}_i$. Thus, Eq. (18) can be rewritten as

$$f_{\mathbf{X_f}_i}(\mathbf{x_f}|\mathbf{Y}=\mathbf{y}) \propto \prod_{j \in \mathcal{T}_{\mathcal{U}}(i)} f_{\mathbf{Y}_j}(\mathbf{y}|\mathbf{X_f}_i = \mathbf{x_f}_i) \times f_{\mathbf{X_f}}(\mathbf{x_f}). \tag{19}$$

Here each decomposition of likelihood $f_{\mathbf{Y}}(\mathbf{y}|\mathbf{X_f}_i = \mathbf{x_f}_i)$ is the measurement density, which is associated with the signal elements distributions and also the noise distribution for the noisy case. For the noiseless case, the measurement density is given by

$$f_{\mathbf{Y}_j}(\mathbf{y}|\mathbf{X_f}_i = \mathbf{x_f}_i) = \delta_{\mathbf{y}_j} \otimes \left(\bigotimes_{k \in \mathcal{T}_{\mathcal{V}}(j) \setminus \{i\}} f_{\mathbf{X_f}_k}(\mathbf{x_f}) \right), \tag{20}$$

where $\delta_{\mathbf{y}_j}$ is the probability density associated with each measurement vector element $\mathbf{y}_j$ and $\mathcal{T}_{\mathcal{V}}(j) \setminus \{i\} f_{\mathbf{x_f}_k}$ $(\mathbf{x_f})$ represent all the neighboring signal elements of $\mathbf{y}_j$ excluding $\mathbf{x_f}_i$. The $\otimes$ and $\bigotimes$ represent the operations of linear convolution and the linear convolution of a sequence of functions, respectively. For the noisy case, the noise distribution $f_{\mathbf{N}_j}(\mathbf{n})$ also accounts for the measurement density, i.e.

$$f_{\mathbf{Y}_j}(\mathbf{y}|\mathbf{X_f}_i = \mathbf{x_f}_i) = \delta_{\mathbf{y}_j} \otimes f_{\mathbf{N}_j}(\mathbf{n}) \otimes \left(\bigotimes_{k \in \mathcal{T}_{\mathcal{V}}(j) \setminus \{i\}} f_{\mathbf{X_f}_k}(\mathbf{x_f}) \right). \tag{21}$$

BP involves the process of exchanging and updating the probability density messages between the signal and the measurement coefficients that relate to each other according to the edges in the bipartite graphs [13] [14]. It is mainly accomplished in two steps known as multiplication and convolution. The marginal posterior for each $\mathbf{x_f}_i$ is updated in every iteration. The message passed from the ith signal coefficient to the jth measurement vector coefficient is called the signal message and is denoted as $m_{i \to j}$, and similarly the message passed from the jth measurement vector coefficient to the ith signal coefficient is called the measurement message, denoted as $m_{j \to i}$.

The signal message is the approximated density message of each signal element $\mathbf{x_f}_i$, i.e. $m_{i \to j} \approx f_{\mathbf{X_f}_i}(\mathbf{x_f}|\mathbf{y})$, and it is obtained from Eq. (19) by the multiplication of all the density messages associated with the neighboring measurement vector coefficients that are updated in the previous iteration, i.e.

$$m^{\ell}_{i \to j} = \eta \left[\prod_{k \in \mathcal{T}_{\mathcal{U}}(i) \setminus \{j\}} m^{\ell-1}_{k \to i} \times f_{\mathbf{X_f}}(\mathbf{x_f}) \right], \tag{22}$$

where $m^{\ell}_{i\to j}$ denotes the signal message at the ℓth iteration, $\eta[.]$ is a normalizing function such that $\int m^{\ell}_{i\to j}\,dx = 1$, and $m^{\ell-1}_{k\to i}$ denotes the neighboring measurement message updated in the previous iteration excluding that of $\mathbf{y}_j$.

Similarly, the measurement message is the approximated density message of each measurement vector coefficient $\mathbf{y}_j$, i.e. $m_{j\to i} \approx f_{\mathbf{y}_j}(\mathbf{y}|\mathbf{x_f})$. The measurement message is updated using Eq. (21) by the convolution of all the updated neighboring signal messages obtained, i.e.

$$m^{\ell}_{j\to i} = \delta_{\mathbf{y}_j} \otimes f_{\mathbf{N}_j}(\mathbf{n}) \otimes \left(\bigotimes_{k\in\mathcal{T}_{\mathcal{V}}(j)\setminus\{i\}} m^{\ell}_{k\to j} \right), \tag{23}$$

where $m^{\ell}_{j\to i}$ is the measurement message at the ℓth iteration and $m^{\ell-1}_{k\to j}$ denote the neighboring signal messages updated excluding that of $\mathbf{x_f}_i$ in the previous iteration, $\delta_{\mathbf{y}_j}$ is the probability density associated with each measurement vector element $\mathbf{y}_j$, and $f_{\mathbf{N}_j}(\mathbf{n})$ is the noise distribution for the noisy case. The $\otimes$ and $\bigotimes$ represent the operations of linear convolution and the linear convolution of a sequence of functions, respectively.

In BP, while the signal and measurement messages are being exchanged and updated, the posterior density for each signal element $\mathbf{x_f}_i$ is being computed at every iteration as

$$f^{\ell}_{\mathbf{X_f}_i}(\mathbf{x_f}|\mathbf{y}) = \eta\left[\prod_{j\in\mathcal{T}_{\mathcal{U}}(i)} m^{\ell}_{j\to i} \times f_{\mathbf{X_f}}(\mathbf{x_f})\right], \tag{24}$$

where $f^{\ell}_{\mathbf{X_f}_i}(\mathbf{x_f}|\mathbf{y})$ is the posterior density for $\mathbf{x_f}_i$ computed at the ℓth iteration.

The maximum value of the density estimated for each signal element $\mathbf{x_f}_i$ after T number of iterations determines the recovered value of each signal element, i.e. $\hat{\mathbf{x_f}}_i$. This BP approach for estimation of Rayleigh faded signals via CS is implemented in the following section.

5. Simulation results

We evaluated the performance of BCS for Rayleigh fading channels using MATLAB simulation results. For a simple case, we generated a strictly sparse CR signal $\mathbf{x_f}$ of length N, such that the signal elements are i.i.d. Rayleigh distributed. To sense this sparse signal and get enough number of measurements M for further recovery of the signal, we used a CS matrix $\mathbf{\Phi}$ ensuring $M \geq cK\log(N/K)$. To estimate the Rayleigh faded signal via BP for the noiseless case, we considered the Rayleigh–Spike prior given in Eq. (10). The sparse nature of the signal and the absence of noise results in almost accurate recovery of the signal. However, in the noisy case, when AWGN noise adds up in the measurement vector, the BCS recovery is not accurate using the Rayleigh–Spike prior due to the noise distribution. The noise distribution affects the MAP estimation for BCS recovery of the Rayleigh faded signal. To overcome this problem, we modified the prior and used a prior that is composed of convolution of Rayleigh–Gaussian distribution for nonzero coefficients and Spike distribution for zero coefficients as shown in Eq. (14). We performed BCS recovery for the noisy case using the modified prior model and varied different parameters of the system model like sparsity level K and number of measurements M to see the performance of the algorithm.

According to the condition $M \geq cK\log(N/K)$ for the BCS approach, a sufficient number of measurements are required for better BCS recovery of the signal. If the number of measurements is quite small, then the BCS recovery process does not perform well. In Figure 2, MSE has been plotted as a function of number of measurements M by varying row weight R for sensing matrix $\mathbf{\Phi}$. When BP converges in $t = 5$ iterations and the signal is recovered, MSE is calculated using the equation $MSE = \|\mathbf{\Phi}\hat{\mathbf{x}}_{\mathbf{f}} - \mathbf{y}\|$. MSE performance of the algorithm observed after $T = 10$ number of iterations using the modified prior suggests that as M increases for signal length of $N = 1000$, the MSE error is reduced. When row weight is small, i.e. $R = 10$, it misses some coefficients and results in poor recovery. MSE performance is improved when R is large and there are enough measurements M according to the signal length. MSE performance for $R = 30$ is better compared to that for $R = 20$, especially when $M \geq 400$ for $N = 1000$.

The above condition for enough number of measurements also shows that the length of signal N and the sparsity level K of the signal affects the performance of BCS recovery of the signal as well. In Figure 3, MSE has been plotted as a function of different values of K for a given N using the modified prior for BCS recovery of a Rayleigh faded signal. When the signal is more sparse, MSE is negligible. As the sparsity level K increases for given N, MSE becomes significant. For $N = 200$, MSE performance degrades when $K \geq 20$. Similarly, for $N = 500$, MSE is significant when $K \geq 30$. However, for $N = 1000$ the algorithm performs better even for $K = 50$, which shows that the performance of the BCS algorithm improves when the signal is more sparse.

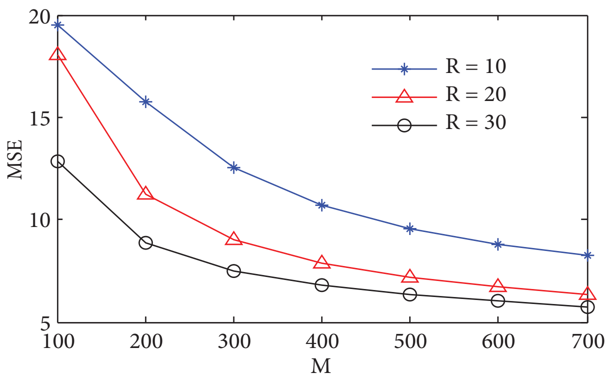

Figure 2. MSE as a function of number of measurements M using different matrix row weights R for BCS recovery of sparse Rayleigh faded signal using prior composed of convolution of Rayleigh–Gaussian distribution and Spike distribution ($N = 1000$, $K = 20$, $\sigma_n^2 = 1$, $R = 10$, 20, 30).

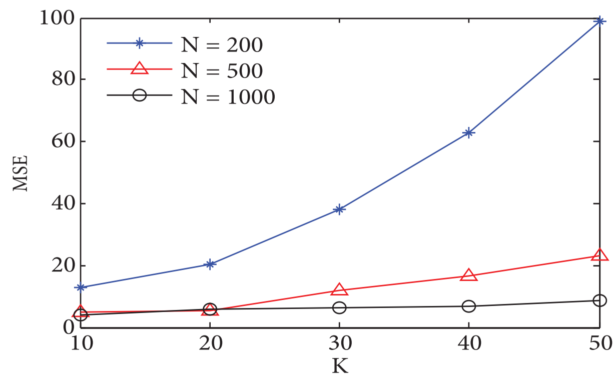

Figure 3. MSE as a function of sparsity level K for BCS recovery of sparse Rayleigh faded signal using prior composed of convolution of Rayleigh–Gaussian distribution and Spike distribution ($N = 200$, 500, 1000, $M = N/2$, $\sigma_n^2 = 1$).

As the sampled message approach has been used for the BCS recovery of the Rayleigh faded signal, the sampling of the prior distribution also affects the outcome of BP. The convolution operation in BP to evaluate the density message can be efficiently computed by using fast Fourier transform (FFT). For efficient use of the FFT, the sampling step should be chosen appropriately such that number of samples p is power of two. In Figure 4, the MSE performance has been investigated for different numbers of samples of prior distribution. As the number of samples p increases, the MSE performance of the algorithm improves. When the number of measurements is lower, the MSE performance of the algorithm is poor for all three cases.

In Figure 5, BCS recovery of the strictly sparse signal of length $N = 1000$ is performed for the noisy case using different priors and MSE is plotted as a function of different values of variance of noise σ_n^2. Figure 5 shows that MSE is reduced when decoding is performed for the noisy system model using the modified prior that is composed of convolution of Rayleigh–Gaussian distribution for nonzero coefficients and Spike distribution for zero coefficients. However, when a Gaussian prior is used for BCS recovery of the Rayleigh faded signal, MSE is significant. When the level of noise that is added to the measurement vector increases, MSE further increases in the case of Gaussian–Spike prior compared to the increase in MSE for the modified prior composed of convolution of Rayleigh–Gaussian distribution and Spike distribution. This shows that an appropriate prior should be used according to the signal model for BCS recovery process. The performance of the algorithm degrades when properties of the prior are not according to the signal model.

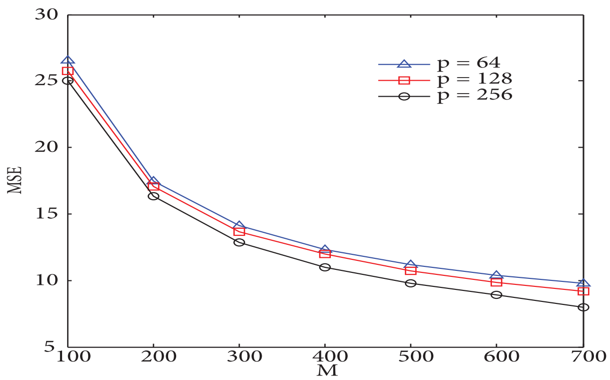

Figure 4. MSE as a function of number of measurements M using different number of samples for prior distribution p for estimation of Rayleigh faded signal via BCS using prior composed of convolution of Rayleigh–Gaussian distribution and Spike distribution ($N = 1000$, $K = 20$, $M = N/2$).

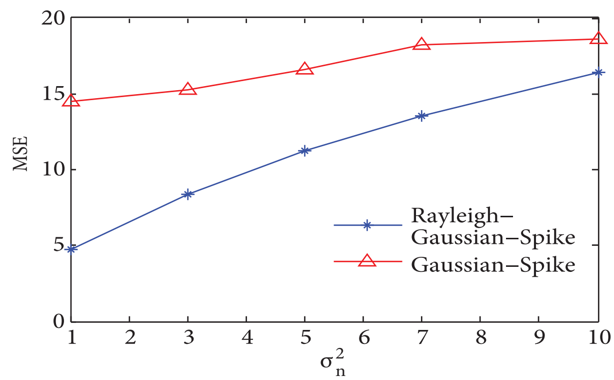

Figure 5. MSE as a function of noise variance σ_n^2 where BCS recovery of sparse Rayleigh faded signal is done for noisy model using two types of priors, i.e. prior composed of convolution of Rayleigh–Gaussian distribution and Spike distribution (Rayleigh-Gaussian-Spike), and Gaussian–Spike prior (Gaussian–Spike) ($N = 1000$, $K = 20$, $M = N/2$).

To further emphasize the appropriate choice of prior according to the system model, we varied the parameters for the system model and recovered the signal using the two above-mentioned priors. In Figure 6, a sparse Rayleigh faded signal of length $N = 500$ is decoded using our modified prior composed of convolution of Rayleigh–Gaussian distribution and Spike distribution as well as using the classical Gaussian–Spike prior. In Figure 6 MSE is plotted as a function of number of measurements M, which shows that as M increases, MSE is reduced when our modified prior is used, while it is significant when the Gaussian–Spike prior is used.

Similarly, in Figure 7, MSE is plotted as a function of sparsity level K for the two priors for a signal length of $N = 200$. It shows that when K increases, MSE is significant for the Gaussian–Spike prior compared to our modified prior. These results show that when the prior model is in accordance with the system model, the BP algorithm performs well and MSE is reduced. The wide-band spectrum sensing via BCS can become more efficient when practical channels like Rayleigh fading channels are considered in the algorithm. The selection of the prior model according to the system model under consideration improves the spectrum sensing process.

6. Conclusion

CS has shown great potential in the recent past in a large number of applications. Spectrum sensing in CRs is one of the fields that have found applications of CS and studies have been performed in the past to explore the application of CS in spectrum sensing. This paper extends this application to a more practical scenario, i.e. where the signal of interest has undergone fading. By appropriately modeling the prior distribution according to the fading distribution, an improvement in the MSE can be sought. An appropriate prior for a Rayleigh faded sparse signal was derived and then used in the BCS framework. Simulation results show that the usage of appropriate prior distribution for faded sparse signals can result in an improvement in MSE.

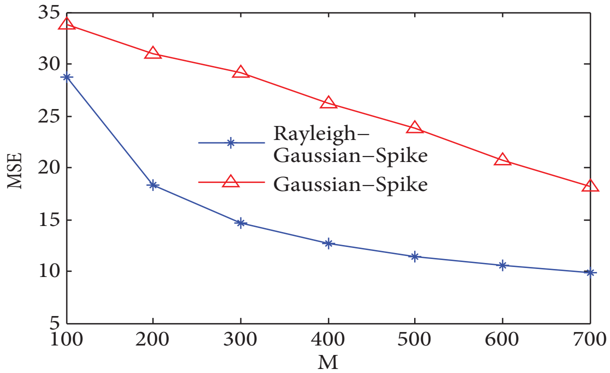

Figure 6. MSE as a function of number of measurements M where BCS recovery of sparse Rayleigh faded signal is done for noisy model using two types of priors, i.e. prior composed of convolution of Rayleigh–Gaussian distribution and Spike distribution (Rayleigh-Gaussian-Spike), and Gaussian–Spike prior (Gaussian–Spike) ($N = 500$, $K = 20$, $\sigma_n^2 = 1$, $R = 20$).

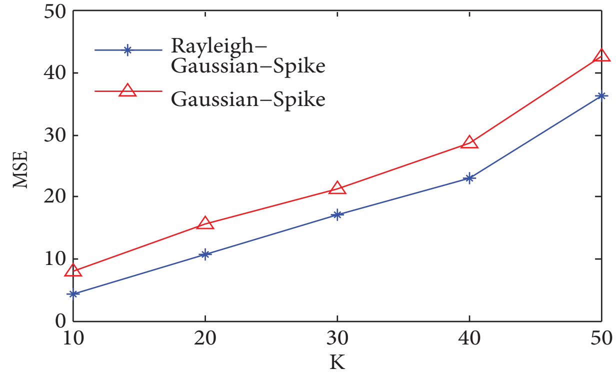

Figure 7. 5pt MSE as a function of sparsity level K where BCS recovery of sparse Rayleigh faded signal is done for noisy model using two types of priors, i.e. prior composed of convolution of Rayleigh–Gaussian distribution and Spike distribution (Rayleigh-Gaussian-Spike), and Gaussian–Spike prior (Gaussian–Spike) ($N = 200$, $M = N/2$, $\sigma_n^2 = 1$).

References

[1] Mitola J, Maguire GQ. Cognitive radio: making software radios more personal. IEEE Pers Commun 1999; 6: 13-18.

[2] Shankar NS, Cordeiro C, Challapali K. Spectrum agile radios: utilization and sensing architectures. In: 2005 First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks; 8–11 November 2005; Baltimore, MD, USA. pp. 160-169.

[3] Tang H. Some physical layer issues of wideband cognitive radio systems. In: 2005 First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks; 8–11 November 2005; Baltimore, MD, USA. pp. 151-159.

[4] Cabric D, Mishra SM, Brodersen RW. Implementation issues in spectrum sensing for cognitive radios. In: Conference Record of the Thirty-Eighth Asilomar Conference on on Signals, Systems and Computers; 7–10 November 2004; Pacific Grove, CA, USA. pp. 772-776.

[5] Candes EJ, Romberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE T Inform Theory 2006; 52: 489-509.

[6] Donoho D. Compressed sensing. IEEE T Inform Theory 2006; 52: 1289-1306.

[7] Sarvotham S, Baron D, Baraniuk R. Compressed Sensing Reconstruction Via Belief Propagation. Technical Report. Houston, TX, USA: Rice University, 2006.

[8] Baron D, Sarvotham S, Baraniuk R. Bayesian compressive sensing via belief propagation. IEEE T Signal Proces 2010; 58: 269-280.

[9] Gallager RG. Low-density parity check codes. IRE T Inform Theor 1962; 8: 21-28.

[10] MacKay DJC. Good error-correcting codes based on very sparse matrices. In: IEEE International Symposium on Information Theory; 29 June-4 July 1997; Ulm, Germany. p. 113

[11] Tan X, Li J. Computationally efficient sparse Bayesian learning via belief propagation. IEEE T Signal Proces 2010; 58: 2010-2021.

[12] Akcakaya M, Park J, Tarokh V. A coding theory approach to noisy compressive sensing using low density frame. IEEE T Signal Proces 2011; 59: 5369-5379.

[13] Kang J, Lee HN, Kim K. On detection-directed estimation approach for noisy compressive sensing. 2012 [Online]. arXiv: 121 1.1250 [cs.IT].

[14] Kang J, Lee HN, Kim K. Detection directed sparse estimation using bayesian hypothesis test and belief propagation. 2012 [Online]. arXiv: 121 1.1250 [cs.IT].

[15] Ishwaran H, Rao JS. Spike and slab variable selection: frequentist and Bayesian strategies. Ann Stat 2005; 33: 730-773.

[16] Ji Y, Yang Z, Li W. Bayesian sparse reconstruction method of compressed sensing in the presence of impulsive noise. Circ Syst Signal Pr 2013; 32: 2971-2998.

[17] Wang YG, Yang L, Tang L, Liu Z, Jiang WL. Enhanced multi-task compressive sensing using Laplace priors and MDL-based task classification. EURASIP J Adv Sig Pr 2013; 160: 1-17.

[18] He L, Carin L. Exploiting structure in wavelet-based Bayesian compressive sensing. IEEE T Signal Proces 2009; 57: 3488-3497.

[19] Guo D, Wang CC. Asymptotic mean-square optimality of belief propagation for sparse linear system. In: IEEE Information Theory Workshop; 22–26 October 2006; Chengdu, China: pp. 194-198.

[20] Wainwright MJ, Jaakkola T, Willsky A. Tree-based reparameterization framework for analysis of sum-product and related algorithms. IEEE T Inform Theory 2003; 49: 1120-1146.

[21] Rangan S. Estimation with random linear mixing, belief propagation and compressed sensing. In: 44th Annual Conference on Information Sciences and Systems; 17–19 March 2010; Princeton, NJ, USA. pp. 1-6.