Enhanced Sparsity Order Estimation Techniques for Dynamic Compressed Sensing

A thesis submitted

in partial fulfillment for the award of the degree of

Doctor of Philosophy

by

THIRUPPATHIRAJAN S

Department of Avionics
Indian Institute of Space Science and Technology
Thiruvananthapuram, India

June 2023

Certificate

This is to certify that the thesis titled *Enhanced Sparsity Order Estimation Techniques for Dynamic Compressed Sensing* submitted by **THIRUPPATHIRAJAN S**, to the Indian Institute of Space Science and Technology, Thiruvananthapuram, in partial fulfillment for the award of the degree of **Doctor of Philosophy** is a bona fide record of the original work carried out by him under my supervision. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Dr. MANOJ B.S.
Professor,
Department of Avionics, IIST
(Supervisor)

Dr. LAKSMINARAYANAN R.
Associate Professor,
Department of Avionics, IIST
(Co-Supervisor)

Dr. SREELAL S.
Division Head,
VSSC, ISRO
(Supervisor, ISRO)

Dr. Selvaganesan N.,
Professor and Head,
Department of Avionics, IIST

Place: Thiruvananthapuram
Date: June 2023

Declaration

I declare that this thesis titled *Enhanced Sparsity Order Estimation Techniques for Dynamic Compressed Sensing* submitted in partial fulfillment for the award of the degree of **Doctor of Philosophy** is a record of the original work carried out by me under the supervision of **Dr. MANOJ B.S.**, **Dr. LAKSMINARAYANAN R.** and **Dr. SREELAL S.** and has not formed the basis for the award of any degree, diploma, associateship, fellowship, or other titles in this or any other Institution or University of higher learning. In keeping with the ethical practice in reporting scientific information, due acknowledgments have been made wherever the findings of others have been cited.

Place: Thiruvananthapuram   THIRUPPATHIRAJAN S
Date: June 2023   (SC13D010)

This thesis is dedicated to my parents, Sankaralingam V.M.(Late) and Gnanasoundari S, my beloved wife Babirekha T., and my children Poorna Sri T., and Karthik Kumar T., who have always loved me unconditionally and whose good examples have taught me to work hard for the things that I aspire to achieve.

Acknowledgements

Abstract

We present a composite Compressed Sensing (CS) system for the efficient acquisition and recovery of sparse and compressible time-varying signals. Sparsity Order Estimation (SOE) is critical in determining the efficiency of CS acquisition and recovery because the sparsity order parameter of an underlying signal plays an important role in determining the number of CS samples or measurements to be obtained from the underlying signal during acquisition and the number of iterations required to recover the support of the underlying signal during recovery. As a result, the proposed composite CS system is built by vertically stacking a sparse Binary Sensing Matrix (BSM) and a dense Gaussian Sensing Matrix (GSM), with the sparse BSM assisting the SOE during acquisition and recovery and the GSM assisting reconstruction during recovery. The proposed sparse BSM is deterministic and adapts to the sparsity order variations for an efficient SOE. The GSM is dense and random, and satisfies the Restricted Isometry Property (RIP) with an overwhelming probability for guaranteed recovery. Because of the BSM's weaker RIP, we limit the number of BSM-based measurements.

We propose the SOE based on two different Maximum Likelihood (ML) principles: (i) BSM-based SOE (BSOE), which takes advantage of the sparse structure of the BSM and the statistics of BSM-based measurements, and (ii) GSM-based SOE (GSOE), which takes advantage of the statistics of GSM-based measurements. The proposed BSOE method does not require any prior knowledge of the underlying signal, but it estimates the statistics of the underlying signal, and thus the sparsity order, with a limited number of measurements. To perform the SOE, the GSOE method requires statistical estimates obtained from the BSOE method. We demonstrate that both ML estimation methods produce unbiased estimates, and that the GSOE method meets the Cramer-Rao Lower Bound.

As the sparsity order varies over time owing to the continuous birth of newer supporting components and the death of existing supporting components, we characterize the sparsity order variation as a stochastic Markov birth-death process. A statistical measure, namely, survival time, is introduced here to statistically quantify the degree and duration of invariance of the sparsity order. We then refine the ML estimates of sparsity order using either of the two independent approaches, namely, Viterbi algorithm-based ML sequence estimation and Kalman filtering of ML estimation by exploiting the underlying discrete Markov process and the survival time, which characterizes the sparsity order variation.

We develop a BSM Aided Orthogonal Matching Pursuit (BAOMP) method for the faster recovery of sparse and compressible signals. Although the sparse BSM is limited owing to its weaker isometry bounds, each BSM-based measurement provides initial estimates of the probable support candidates based on the location of those in the corresponding BSM row. Because of these initial estimates, the number of iterations required for the recovery algorithm is subsequently reduced, and the speed of recovery is improved by at least 25% compared to existing recovery methods.

The proposed composite CS, ML sparsity order estimators, and BAOMP-based recovery algorithms are practically implementable and can be used in real-world applications such as (i) vibration signal acquisition and recovery, (ii) channel estimation, and (iii) electro cardiogram signal recovery. The proposed method's performance is then compared to existing methods using metrics such as SOE error, Normalized Recovery Error, and run-time complexities. The results on real-world and synthetic data show that the proposed methods work with better performances, even with a low signal-to-noise ratio.

Contents

Abbreviations

CS	Compressed Sensing
SOE	Sparsity Order Estimation
RIP	Restricted Isometry Property
ML	Maximum Likelihood
BSM	Binary Sensing Matrix
GSM	Gaussian Sensing Matrix
OMP	Orthogonal Matching Pursuit
BAOMP	BSM-Aided OMP
SAStOMP	Sparsity Adaptive segmented OMP
CoSaMP	Compressive Sampling Matching Pursuit
OAMP	Optimized Adaptive Matching Pursuit
CRLB	Cramer-Rao Lower Bound
TS-ACSS	Two-Step Adaptive Compressive Spectrum Sensing
MMV	Multiple Measurement Vector
SaMP	Sparsity Adaptive Matching Pursuit
ASaMP	Adaptive SaMP

Nomenclature

M	Measurements size
N	Signal dimension
k	Sparsity order
n	Discrete time index
$\|.\|_p$	ℓ_p norm operator
$\Pr[.]$	Probability operator
$\mathbb{E}\{.\}$	Expectation operator
$\text{VAR}\{.\}$	Variance operator
$\lfloor.\rceil$	Round the argument to nearest integer
$\lfloor.\rfloor$	Round the argument to greatest preceding integer
$\lceil.\rceil$	Round the argument to least succeeding integer
$\mathbb{R}$	Real number domain
$\mathcal{O}$	Order of complexity
$\binom{k}{L}$	Number of ways of choosing L out of k elements
$X \sim \mathcal{N}(\mu, \sigma^2)$	A Gaussian random variable X with mean μ and variance σ^2
$Y \sim \mathcal{B}(0, \lambda)$	A Bernoulli random variable Y such that $\Pr[Y = 0] = \lambda$

Chapter 1

Introduction

The conventional data acquisition system acquires any band-limited continuous time-varying signal $\mathbf{x}(t)$ of bandwidth B Hz by uniformly sampling at the Nyquist rate of at least $2B$ samples per second to generate N samples in every time interval T. These N samples are then transformed into N coefficients using an orthonormal sparse representation transform, or greater than N coefficients using an over-complete dictionary. Some of the orthonormal sparse representation transforms are the Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT), and Discrete Wavelet Transform (DWT), which are complete dictionaries as the transformation has the same dimension as that of the signal dimension. However, over-complete dictionaries are generally formed by the horizontal concatenation of two or more orthonormal transforms such that the transformation has a higher dimension than that of the signal dimension. The transformation is chosen such that the information of $\mathbf{x}(t)$ is concentrated over a very small number of coefficients, such as k coefficients and $k \ll N$. These k coefficients are termed significant coefficients because their amplitudes and locations are sufficient to reconstruct $\mathbf{x}(t)$. It should be noted that the number of significant components k also varies in every time interval T, i.e., $k \to k(nT)$ where nT denotes n^{th} time interval T and n is a positive integer. Thus, after transformation, these $k(nT)$ significant coefficients are retained and the remaining $N - k(nT)$ insignificant coefficients are discarded. For brevity, we omit the notation nT in this chapter.

The differentiation of significant and insignificant coefficients is user-defined based on some threshold τ, i.e., the coefficients having magnitudes above the threshold are termed significant coefficients, and the remaining coefficients are termed insignificant coefficients. In general, the magnitudes of the insignificant coefficients are mostly near zero. Thus, an $N-$dimensional signal $\mathbf{x}(t)$ is reduced to $2k$ information, i.e., k amplitudes and k positions in the orthonormal transform domain in a given time interval. In other words, the N time-domain samples obtained in a time interval are compressed into $2k$ information samples or

measurements. The number of significant coefficients in the transform domain represents the sparsity order or sparsity level k of the signal and the set representing the collection of locations or indices of such significant coefficients is the support $\mathcal{S}$ of the signal.

The aforementioned conventional data acquisition process is summarized in the following sequential steps.

1. Sampling of the signal at Nyquist rate to result in N time-domain samples.

2. Transformation of the acquired samples using an orthonormal transform or an over-complete dictionary to result in at least N transform-domain coefficients.

3. Retention of the significant transform domain coefficients and their locations to result in $2k$ compressed samples or measurements.

Considering the orthonormal transform, because N coefficients are computed from N Nyquist rate samples using N basis functions, the hardware complexity and the processing power requirements depend on the dimension N. However, in the end, only $2k \ll N$ information values are relevant, i.e., although the Nyquist rate results in N samples for a given time interval, the effective information rate results in $2k \ll N$ information values for the same time interval. Thus, there is a pressing need to acquire signals at the information rate rather than at the Nyquist rate using alternate methods. To address this need, Compressed Sensing (CS) [1, 2, 3, 4, 5, 6, 7] has emerged as an alternative method for acquiring a signal with sparse or concise representation. CS theory addresses the following aspects.

1. Sparse representation of the signal.

2. CS acquisition to obtain the compressed samples or measurements directly from the signal without applying the transformation.

3. CS recovery to reconstruct the signal using the obtained measurements.

1.1   Sparse representation

Most signals in nature are either sparse or compressible exhibiting conciseness when they are transformed or represented using orthonormal basis functions or over-complete dictionaries. The concise representation of a signal $\mathbf{x}(t)$ is identified by the least linear combination of vectors of a basis function or atoms of an over-complete dictionary $\{\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, ..., \boldsymbol{\psi}_N\}$

which resembles the original signal with minimal error, i.e.,

$$\mathbf{x}(t) = s_1(t)\boldsymbol{\psi}_1(t) + s_2(t)\boldsymbol{\psi}_2(t) + ... + s_k(t)\boldsymbol{\psi}_k(t) + ... + s_N(t)\boldsymbol{\psi}_N(t).$$

Here $s_i(t) : i = 1$ to N are descending order sorted coefficients. If $s_i(t) = 0$ for $i > k$, then $\mathbf{x}(t)$ is truly sparse signal and if $|s_i(t)| < \epsilon$(infinitesimal) for $i > k$, then $\mathbf{x}(t)$ is compressible signal.
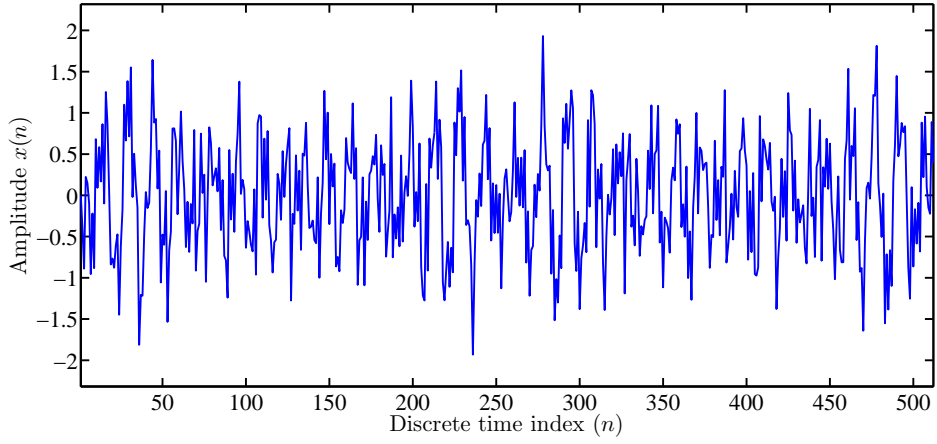
1.1.1   True sparse signal

A signal $\mathbf{x}(t)$ is truly sparse if it can be represented on a suitable orthonormal basis or dictionary $\boldsymbol{\Psi}$ such that $\mathbf{x}(t) = \boldsymbol{\Psi}\mathbf{s}(t)$, and the representation $\mathbf{s}(t)$ has very few nonzero components or coefficients compared to its dimension.

An example of a synthetic true sparse signal is shown in Figure 1.1. The discrete time-domain representation of the signal is shown in Figure 1.1 (a) and its frequency-domain representation using DFT and DCT are shown, respectively, in Figures 1.1 (b) and 1.1 (c). The dimension of the true sparse signal is $N = 512$. The frequency-domain representation using DFT has $k = 32$ nonzero coefficients whereas using DCT has $k = 200$ nonzero coefficients showing that DFT results in a more sparse representation than DCT for the given synthetic signal. It shows that generally the sparse representation differs for different orthonormal transformations and some of the transformations result in a more concise representation when the basis function or the dictionary matches with the signal characteristics.

Some real-world examples of true sparse signals are the Channel State Information (CSI) of OFDM channels [8, 9, 10] and the spectrum occupancy state of cognitive radio [11].

1.1.2   Compressible or sparse-approximated signal

A signal $\mathbf{x}(t)$ is compressible or sparse-approximated if its representation $\mathbf{s}(t)$ has all nonzero coefficients and the descending order sorted magnitudes of the coefficients obey the power-law decay [4], i.e., $|\tilde{s}_j| \leq Cj^{-r}$ where $|\tilde{s}_j|$ is the jth sorted magnitude and $C > 0$ and $r > 0$ are constants. If the signal decays rapidly, only a few coefficients have larger magnitudes to be considered significant and the remaining coefficients have near-zero magnitudes to be considered insignificant. Most real-world signals such as images, videos, and audio are examples of compressible signals. In other words, the cumulative energy of

(a) Sparse signal $\mathbf{x}(n)$ of dimension $N = 512$ in discrete time domain
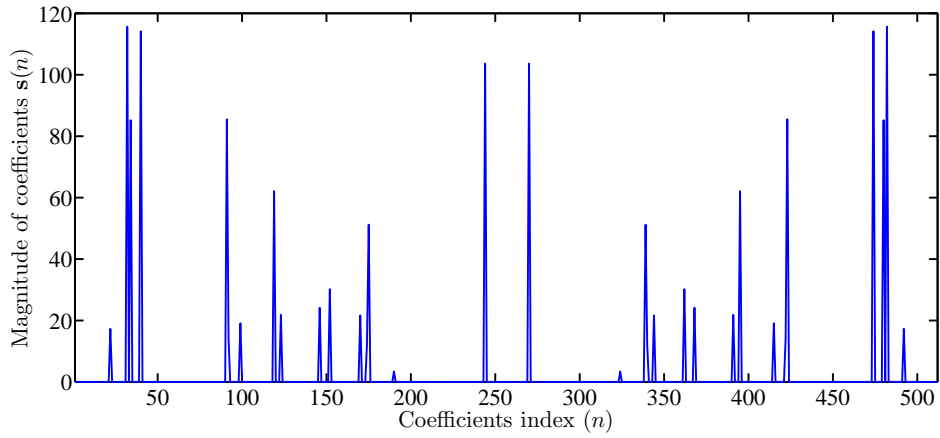
(b) Sparse representation $\mathbf{s}(n)$ of $\mathbf{x}(n)$ with sparsity order $k = 32$ using DFT
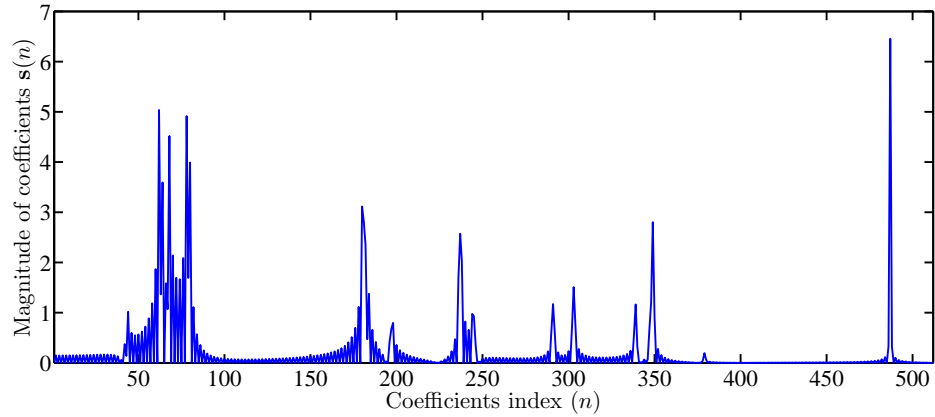
(c) Sparse representation $\mathbf{s}(n)$ of $\mathbf{x}(n)$ with sparsity order $k = 200$ using DCT

Figure 1.1: A synthetic true sparse signal in the time domain and frequency domain.
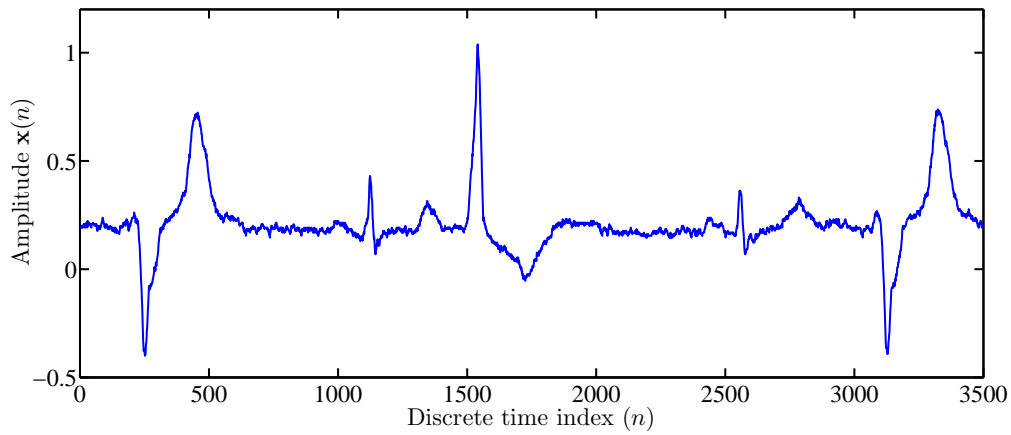
(a) Compressible signal $\mathbf{x}(n)$ of dimension $N = 3500$ in time domain
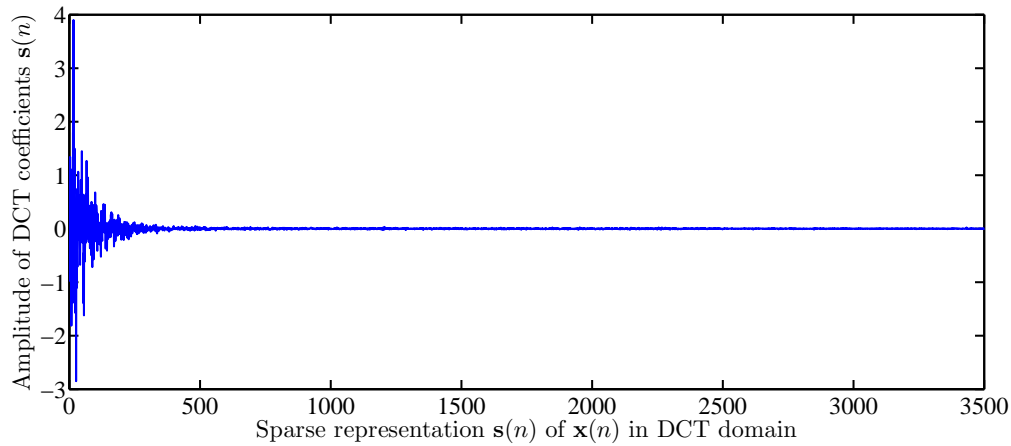
(b) Sparse representation $\mathbf{s}(n)$ of $\mathbf{x}(n)$ in DCT domain
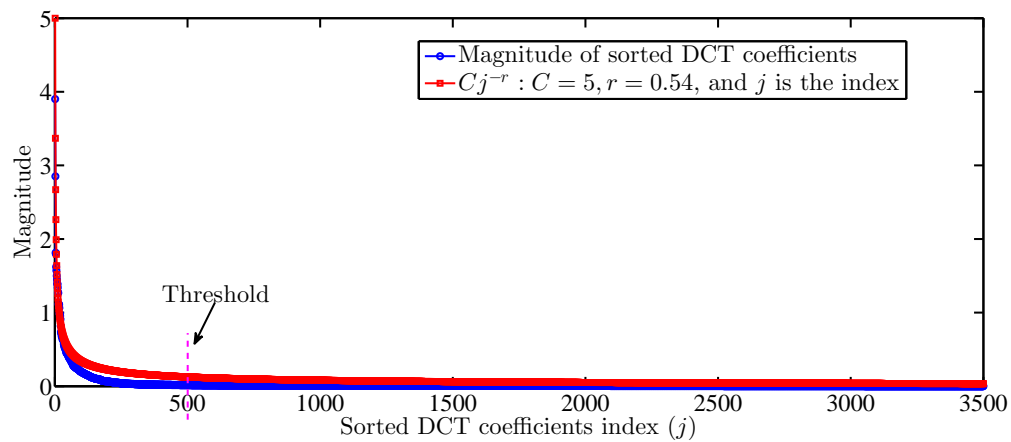
(c) Descending order sorted magnitudes of DCT coefficients

Figure 1.2: Compressible ECG signal.

the insignificant coefficients is very less compared to significant coefficients. For real-world compressible signals, the significant coefficients are identified such that they hold cumulatively 95 to 99% of energy of the signal and this percentage of cumulative energy concentration distinguishing the significant and insignificant coefficients is user defined.

An example of a real-world compressible Electrocardiogram (ECG) signal of dimension $N = 3500$ is shown in Figure 1.2. The time-domain representation of the signal is shown in Figure 1.2 (a). The frequency-domain representation of the signal using DCT is shown in Figure 1.2 (b). It is observed that the sparse representation has very few significant coefficients in the frequency domain. The descending order sorted magnitudes of the DCT coefficients are shown in Figure 1.2 (c) which reveals that the signal is highly compressible as the sorted magnitude decays rapidly obeying the power-law (exponential) decay (with the parameters $C = 5$ and $r = 0.54$) and the insignificant coefficients are in the tail region of the exponential decay. Here, 99% of total energy of the compressible signal is considered for distinguishing the significant and insignificant coefficients from the descending order sorted coefficients. It is observed that 99% of the energy is concentrated in the initial 500 sorted DCT coefficients. Thus, there are 500 significant coefficients and 3000 insignificant coefficients. Hence, the sparsity order is $k = 500$.

1.2   CS acquisition

CS acquisition in a discrete setting [1, 2, 3, 4] is a linear mapping of the $N-$dimensional $k-$sparse or compressible signal $\mathbf{x}$ using an $M \times N-$dimensional measurement or sensing matrix $\boldsymbol{\Phi}$ to obtain an $M-$dimensional measurement vector $\mathbf{y}$ as shown below.

$$
\begin{aligned}
\mathbf{y} &= \boldsymbol{\Phi}\mathbf{x} + \boldsymbol{\vartheta} & (1.1)\\
&= \boldsymbol{\Phi}\boldsymbol{\Psi}\mathbf{s} + \boldsymbol{\vartheta} & (1.2)\\
&= \boldsymbol{\Theta}\mathbf{s} + \boldsymbol{\vartheta} & (1.3)
\end{aligned}
$$

where $\boldsymbol{\Theta} = \boldsymbol{\Phi}\boldsymbol{\Psi}$, and $\boldsymbol{\vartheta}$ is the measurement noise. Here, $k < M$, $k \ll N$, and $M < N$.

The measurement matrix $\boldsymbol{\Phi}$ is designed such that it satisfies the Restricted Isometry Property (RIP) [1, 4], a property akin to the orthonormality of Fourier or wavelet matrices, to recover all sparse signals with sparsity order k. The RIP is given as,

$$
(1 - \delta_k) \leq \frac{\|\boldsymbol{\Theta}\mathbf{s}\|_2^2}{\|\mathbf{s}\|_2^2} \leq (1 + \delta_k) \tag{1.4}
$$

where $\delta_k \in (0,1)$ denotes the isometry constant of $\boldsymbol{\Theta}$. When $\delta_k = 0$, the RIP becomes $\|\boldsymbol{\Theta s}\|_2^2 = \|\mathbf{s}\|_2^2$ which is a necessary condition for the orthonormal transform matrices. However, for compressed sensing matrices $M < N$. Therefore, $\delta_k \neq 0$. If δ_k is very close to zero, i.e., if any subset of k columns of $\boldsymbol{\Theta}$ is nearly orthonormal, the recovery probability is very close to 1. In other words, the RIP guarantees that no two sparse signals of sparsity order k can be mapped to the same $\mathbf{y}$ through $\boldsymbol{\Theta}$, i.e., the RIP preserves the pairwise distance of any two $k-$sparse signals in the measurement space, as given below.

$$(1 - \delta_{2k}) \leq \frac{\|\boldsymbol{\Theta}(\mathbf{s}_1 - \mathbf{s}_2)\|_2^2}{\|(\mathbf{s}_1 - \mathbf{s}_2)\|_2^2} \leq (1 + \delta_{2k}). \tag{1.5}$$

When $\delta_{2k} \ll \sqrt{2} - 1$ for $\boldsymbol{\Theta}$, the recovery of any $k-$ sparse signal is exact from the measurement vector $\mathbf{y}$ using ℓ_1 norm minimization method [1]. However, identifying a sensing matrix that satisfies the RIP is a Non-deterministic Polynomial time (NP)-hard problem, as it requires $\binom{N}{k}$ searches. Using probabilistic approaches, it has been proven that all $M \times N$ independent and identically distributed (i.i.d) Gaussian matrices satisfy the RIP with a high probability when $M \geq c\, k \log(N/k)$ for a small constant c [3, 4].

Another parameter equivalent to the RIP in governing the construction of the sensing matrix is mutual coherence [3, 4, 12]. It measures the largest correlation factor between the rows of the sensing matrix $\boldsymbol{\Phi}$ and columns of the representation matrix $\boldsymbol{\Psi}$. Mutual coherence μ for $\boldsymbol{\Theta}$ is defined as follows.

$$\mu(\boldsymbol{\Theta}) = \mu(\boldsymbol{\Phi}, \boldsymbol{\Psi}) = \sqrt{N} \max |\langle \boldsymbol{\Phi}_i, \boldsymbol{\Psi}_j \rangle|; \quad 1 \leq i \leq M, 1 \leq j \leq N. \tag{1.6}$$

Using Equation (1.6), it can be observed that $1 \leq \mu \leq \sqrt{N}$. For CS, the mutual coherence μ should be as low as possible, i.e., $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ should be highly incoherent. This can be achieved when each row $\boldsymbol{\Phi}_i$ spreads out in $\boldsymbol{\Psi}$ and each column $\boldsymbol{\Psi}_j$ spreads out in $\boldsymbol{\Phi}$. This also implies that the rows of $\boldsymbol{\Theta}$ should be spread, rather than concentrated. If $\boldsymbol{\Phi}$ is constructed with i.i.d Gaussian entries, then the RIP and low mutual coherence conditions are satisfied with a high probability for any fixed orthonormal basis $\boldsymbol{\Psi}$. This is due to the fact that the columns of $\boldsymbol{\Theta}$ are mutually independent, and so are the rows of $\boldsymbol{\Theta}$ as shown below.

$$\mathbb{E}\{\boldsymbol{\Theta}^H \boldsymbol{\Theta}\} = \mathbb{E}\{\boldsymbol{\Psi}^H \boldsymbol{\Phi}^H \boldsymbol{\Phi} \boldsymbol{\Psi}\} = \boldsymbol{\Psi}^H \mathbb{E}\{\boldsymbol{\Phi}^H \boldsymbol{\Phi}\} \boldsymbol{\Psi} = \boldsymbol{\Psi}^H \mathrm{I}_{N \times N} \boldsymbol{\Psi} = \boldsymbol{\Psi}^H \boldsymbol{\Psi} = \mathrm{I}_{N \times N}$$
$$\mathbb{E}\{\boldsymbol{\Theta} \boldsymbol{\Theta}^H\} = \mathbb{E}\{\boldsymbol{\Phi} \boldsymbol{\Psi} \boldsymbol{\Psi}^H \boldsymbol{\Phi}^H\} = \mathbb{E}\{\boldsymbol{\Phi} \mathrm{I}_{N \times N} \boldsymbol{\Phi}^H\} = \mathbb{E}\{\boldsymbol{\Phi} \boldsymbol{\Phi}^H\} = (N/M)\, \mathrm{I}_{M \times M}$$
$$\tag{1.7}$$

The mutual coherence values for various pairs of $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ are listed in Table 1.1.

Table 1.1: Mutual coherence among different pairs

Φ	Ψ	$\mu(\Phi, \Psi)$
Identity	discrete Fourier	1
Identity	Hadamard-Walsh	1
Noiselets	Haar wavelets	$\sqrt{2}$
Noiselets	Daubechies D4	2.2
Noiselets	Daubechies D8	2.9
Orthobasis uniformly at random (basis with N orthonormalized vectors sampled independently and uniformly on a unit sphere)	Any fixed orthonormal basis	$\approx \sqrt{2 \log N}$
Random matrix with i.i.d Gaussian or Bernoulli entries	Any fixed orthonormal basis	≈ 1

Using RIP and mutual coherence, it can be observed that random sensing matrices with i.i.d Gaussian entries recover $k-$sparse signals with an overwhelming probability.

1.3 CS recovery

Because Equation (1.1) is a system of under-determined equations as $M < N$, there are infinite solutions on solving $\mathbf{s}$ from $\mathbf{y}$ and Θ. However, utilizing the fact that the solution is a sparse one, CS recovery solves Equation (1.1) using either convex relaxation techniques or greedy techniques.

1.3.1 Convex relaxation techniques

Consider the noiseless case, i.e., $\mathbf{y} = \Theta\mathbf{s}$. If $\hat{\mathbf{s}}$ is one of the solutions, then $\hat{\mathbf{s}} + \mathbf{h}$ is another possible solution for any vector $\mathbf{h}$ belonging to the null space of Θ, i.e., $\mathbf{h} \in \mathcal{N}(\Theta)$. Therefore, the CS recovery algorithm searches for a sparse signal in the $(N - M)$-dimensional null space, $\mathcal{H} = \mathcal{N}(\Theta) + \hat{\mathbf{s}}$.

One of the methods used to search and find a solution for the signal $\mathbf{s}$ from the measurements $\mathbf{y}$ in the translated null space with the least ℓ_2 norm, i.e., minimal energy is given by

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s}} \|\mathbf{s}\|_2 \quad \text{subject to} \quad \mathbf{y} = \Theta\mathbf{s}. \tag{1.8}$$

This technique has a unique closed-form least-squares solution $\hat{\mathbf{s}} = \Theta^H \left(\Theta\Theta^H\right)^{-1} \mathbf{y}$.

However, the solution is not sparse because ℓ_2 norm distributes energy among all the components of $\hat{s}$ such that the energy of the solution vector $\|\hat{s}\|_2^2$ is minimum resulting in a greater number of nonzeros in $\hat{s}$ which makes the solution non-sparse.

Considering ℓ_0 norm[1], which computes the number of nonzero elements, it is a suitable candidate for finding the sparse solution. Searching for and finding a vector in the translated Null space with the least number of nonzero components using ℓ_0 norm is given by

$$\hat{s} = \arg \min_s \|s\|_0 \quad \text{subject to} \quad y = \Theta s. \tag{1.9}$$

However, solving Equation (1.9) is a numerically unstable and combinatorial NP-hard problem, because the number of possible locations of k nonzero elements is $\binom{N}{k}$. Consider the following example to understand the difficulty of performing ℓ_0 norm minimization. Suppose $N = 2000$ and $k = 20$. Subsequently, $\binom{2000}{20} = 3.9 \times 10^{47}$ searches are required. If each search requires 10^{-9} s, completing the search requires 1.2×10^{31} years, which is impractical.

In general, when we move from ℓ_p norm minimization to ℓ_q norm minimization, such that $q < p$, sparsity is promoted. Thus, ℓ_1 norm minimization prefers sparse solutions to ℓ_2 norm minimization, as shown in Figure 1.3. Here a 2-dimensional sparse signal $s = \{s_1, s_2\}$ is considered, and a measurement $y = \phi_{11}s_1 + \phi_{12}s_2$ is obtained. The measurement is represented by the line $y = \Theta s$. The ℓ_1 norm minimization is performed by blowing a square ball from the origin, and its intersection with the measurement provides a sparse solution. By contrast, ℓ_2 norm minimization is performed by blowing a circular ball from the origin, and its intersection with the measurement provides the solution with the least energy instead of the sparse solution.

Minimizing ℓ_1 norm is a convex relaxation method and is given by

$$\hat{s} = \arg \min \|s\|_1 \quad \text{subject to} \quad y = \Theta s. \tag{1.10}$$

This convex relaxation problem can be reduced to a linear program by splitting s which has both positive and negative values, into s_1 and s_2 which can only take positive values, i.e., $s = s_1 - s_2$. Equation (1.10) can be reframed as a linear program, as in Equation (1.11),

---

[1] ℓ_0 norm is not properly a norm as it is non-differentiable and non-convex.
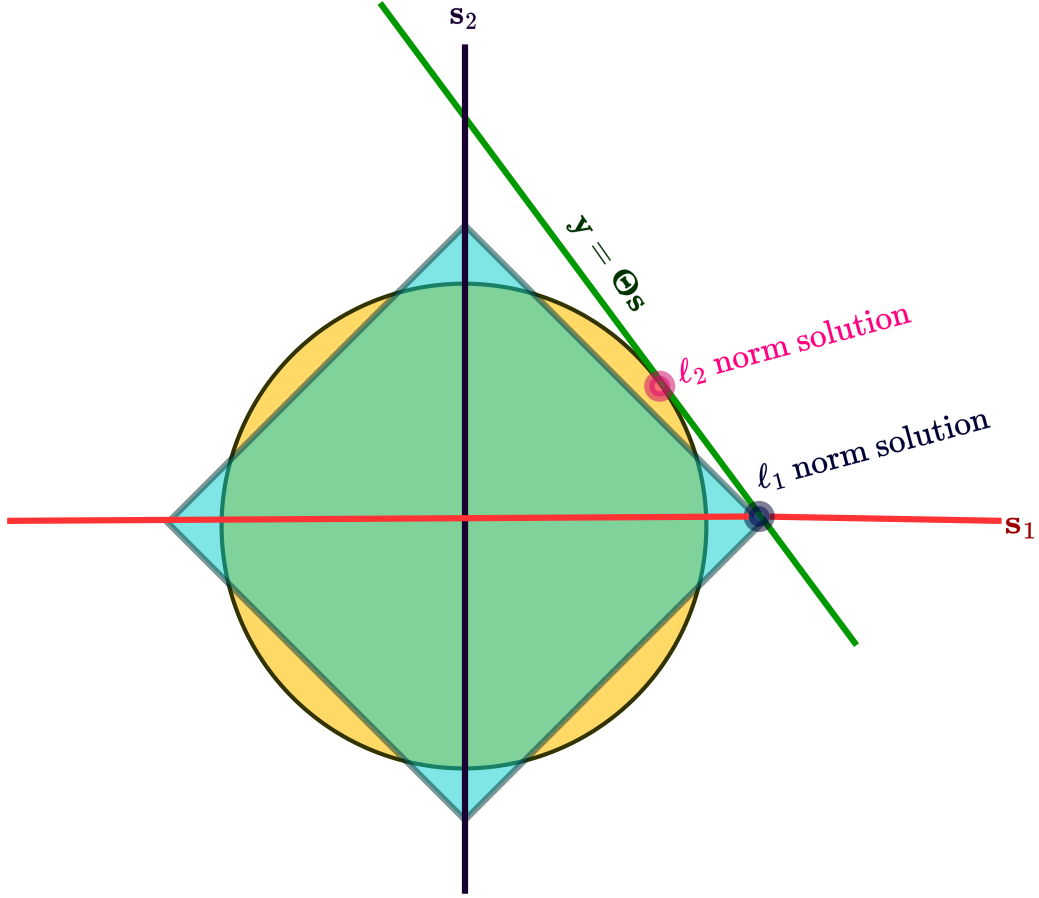
Figure 1.3: Comparison on ℓ_1 and ℓ_2 norm on finding the sparse solution.

which is known as basis pursuit [7].

$$
\begin{aligned}
\hat{\mathbf{s}} &= \arg\min \|\mathbf{s}\|_1 \\
&= \underset{\mathbf{s}_1,\mathbf{s}_2}{\arg\min} \sum_{i=1}^{n} \left(\mathbf{s}_1(i) - \mathbf{s}_2(i)\right) \quad \text{subject to } y = \boldsymbol{\Theta}\left(\mathbf{s}_1 - \mathbf{s}_2\right) \\
&\qquad\qquad\qquad\qquad\qquad\qquad \mathbf{s}_1(i) \geq 0, \mathbf{s}_2(i) \geq 0
\end{aligned}
\tag{1.11}
$$

When considering the noisy case, i.e., $\mathbf{y} = \boldsymbol{\Theta}\mathbf{s} + \boldsymbol{\vartheta}$, the following ℓ_1 norm minimization with relaxed constraints, as given in Equations (1.12) and (1.13), is used for the recovery.

$$
\hat{\mathbf{s}} = \min \|\mathbf{s}\|_1 \quad \text{subject to} \quad \|\boldsymbol{\Theta}\mathbf{s} - \mathbf{y}\|_2 \leq \epsilon
\tag{1.12}
$$

$$
\hat{\mathbf{s}} = \underset{\mathbf{s}}{\arg\min} \quad \frac{1}{2}\|\boldsymbol{\Theta}\hat{\mathbf{s}} - \mathbf{y}\|_2^2 + \nu\|\hat{\mathbf{s}}\|_1
\tag{1.13}
$$

10

Equations (1.12) and (1.13) represent Basis Pursuit Denoising (BPDN) and Least Absolute Shrinkage and Selection Operator (LASSO), respectively. Here, ϵ is the noise level and ν is the tuning parameter related to the sparsity order k and the noise level. These convex relaxation techniques are accurate, however, they generally have higher computational complexity and longer recovery time compared to greedy techniques which are discussed as follows.

1.3.2 Greedy techniques

The ℓ_1- minimization approach relies on linear programming, and the computational cost of linear programming is burdensome for many applications. Hence, the greedy approach has been widely used as a cost-effective and faster alternative to ℓ_1- minimization approach. Greedy techniques such as Orthogonal Matching Pursuit (OMP) [13], Compressive Sampling Matching Pursuit (CoSaMP) [14], etc., are generally used for faster recovery. These greedy techniques exploit the near-orthonormal properties of the column vectors of the measurement matrix to search for the support indices of the signal in an iterative manner. These techniques are usually initialized with a zero estimate of the sparse vector, i.e., $\hat{\mathbf{s}}^0 = \mathbf{0}$ and an empty support set, i.e, $\mathcal{S} = \varnothing$. With this initialization, the initial residual error is $\mathbf{r}^0 = \mathbf{y} - \boldsymbol{\Theta}\hat{\mathbf{s}}^0 = \mathbf{y}$. The proxy for the solution $\hat{\mathbf{s}}$ is $\boldsymbol{\Theta}^T\mathbf{r}^0$ and the support $\mathcal{S}$ of the sparse vector $\hat{\mathbf{s}}$ is updated by identifying the indices of the largest coefficients in the proxy based on a greedy rule. The solution $\hat{\mathbf{s}}^1$ is then updated such that $\hat{\mathbf{s}}^1_{\mathcal{S}} = \left(\boldsymbol{\Theta}^T_{\mathcal{S}}\boldsymbol{\Theta}_{\mathcal{S}}\right)^{-1}\boldsymbol{\Theta}^T_{\mathcal{S}}\mathbf{y}$ and $\hat{\mathbf{s}}^1_{\mathcal{S}^C} = \mathbf{0}$ where *(i)* $\hat{\mathbf{s}}^1_{\mathcal{S}}$ is the estimate of nonzero components in $\hat{\mathbf{s}}^1$ identified by the support $\mathcal{S}$, *(ii)* the submatrix $\boldsymbol{\Theta}_{\mathcal{S}}$ has column vectors of $\boldsymbol{\Theta}$ identified by the support $\mathcal{S}$ and *(iii)* $\hat{\mathbf{s}}^1_{\mathcal{S}^C}$ is the estimate of zero components in $\hat{\mathbf{s}}^1$ not identified by the support $\mathcal{S}$. With the updated solution, the residual is updated as $\mathbf{r}^1 = \mathbf{y} - \boldsymbol{\Theta}\hat{\mathbf{s}}^1$ and the iteration continues. Thus, each iteration updates the support $\mathcal{S}$, thereby, decreasing the residual error $\mathbf{r}$. The number of iterations required depends on the sparsity order. After completing the iterations, the signal is recovered as $\hat{\mathbf{x}} = \boldsymbol{\Psi}^{-1}\hat{\mathbf{s}}$.

The OMP and CoSaMP greedy algorithms, which are dependent on sparsity order k are explained in the following sections.

1.3.2.1 OMP

In the OMP algorithm, the estimate for $\mathbf{s}$ is updated in each iteration by projecting $\mathbf{y}$ orthogonally onto the columns of $\boldsymbol{\Theta}$ associated with the current support set $\mathcal{S}^i$. Consequently, OMP minimizes $\|\mathbf{y} - \boldsymbol{\Theta}_{\mathcal{S}^i}\mathbf{s}_{\mathcal{S}^i}\|_2^2$ over all $\mathbf{s}$ supported by $\mathcal{S}^i$. The OMP algorithm is refer-

enced in Algorithm 1. The columns in $\boldsymbol{\Theta}$ that are part of the support set $\mathcal{S}_i$ must be orthogonal to the residual $\mathbf{r}^i$. This implies that these columns will not be chosen again for the support set in the next iteration (and subsequent iterations), i.e., OMP never reselects an element and the residual at any iteration is always orthogonal to all of the current coefficients. The name of the algorithm originates from this orthogonalization.

Algorithm 1 Orthogonal Matching Pursuit

Input:
- Sensing matrix $\boldsymbol{\Theta} \in \mathbb{R}^{M \times N}$ such that $\|\boldsymbol{\Theta}_j\|_2 = 1 \ \forall j = 1$ to N
- Measurement vector $\mathbf{y} \in \mathbb{R}^{M \times 1}$
- Sparsity order k
- Iteration number $i = 0$
- Initial support $\mathcal{S}^i = \varnothing$
- Initial estimate of the sparse representation $\mathbf{s}^i = \mathbf{0}$

1: **while** $i < k$ **do**
2:      $i = i + 1$
3:      $\mathbf{r}^i = \mathbf{y} - \boldsymbol{\Theta}\hat{\mathbf{s}}^{i-1}$ ▷ Residual
4:      $\mathbf{z} = \boldsymbol{\Theta}^T \mathbf{r}^i$ ▷ Proxy of the sparse representation
5:      $\mathcal{J} = \arg \max_j |z_j|$ ▷ Identify the index of coefficient with the largest magnitude
6:      $\mathcal{S}^i = \mathcal{S}^{i-1} \cup \mathcal{J}$ ▷ Update the support
7:      $\hat{\mathbf{s}}^i_{\mathcal{S}^i} = \left(\boldsymbol{\Theta}_{\mathcal{S}^i}^T \boldsymbol{\Theta}_{\mathcal{S}^i}\right)^{-1} \boldsymbol{\Theta}_{\mathcal{S}^i}^T \mathbf{y}$ ▷ Update the significant coefficient
8: **end while**

Output: Estimate of the sparse representation: $\hat{\mathbf{s}}$.

. OMP can be terminated with a stopping rule based on the reconstruction error ϵ, i.e., $\|\mathbf{r}^i\|_2^2 \leq \epsilon$. However, it demands the noise structure *a priori*, i.e., the variance of noise ϵ as given in [15]. Fixing the reconstruction error based on the noise structure may lead to early termination resulting in poor reconstruction or late termination resulting in more execution time. Hence, the stopping rule based on reconstruction error is not suitable.

1.3.2.2 CoSaMP

A weakness of the OMP algorithm is that, once an incorrect index $\mathcal{J}$ corresponding to the column of $\boldsymbol{\Theta}$ is selected in the estimated support $\mathcal{S}^i$, it remains in all subsequent estimated support $\mathcal{S}^j : j > i$. As a result, if an incorrect index has been selected, k iterations of the OMP algorithm are insufficient to recover a vector with sparsity order k. One possible solution is to increase the number of iterations required. The CoSaMP algorithm [14]

proposes another strategy for estimating the sparsity order k. In contrast to OMP, in each iteration, CoSaMP finds the support $\mathcal{S}^i$ by selecting indices of $2k$ largest entries in the proxy $\mathbf{\Theta}^T \mathbf{r}^i$ and merges it with the support corresponding to k largest coefficients of the previous estimate of the sparse signal $\mathbf{s}^{i-1}$ as shown in steps 4 and 5 of Algorithm 2. Subsequently, it minimizes $\|\mathbf{y} - \mathbf{\Theta}\mathbf{b}\|_2$ among all $\mathbf{b}$ with the support $\mathcal{S}^i$ as shown in step 6. Then, the k largest entries of $\mathbf{b}$ become the k largest coefficients of the estimated sparse signal and the remaining $N - k$ coefficients are zeros. This iteration continues until the energy of the residual falls below a certain threshold, ϵ.

Algorithm 2 Compressive Sampling Matching Pursuit

Input:

- Sensing matrix $\mathbf{\Theta} \in \mathbb{R}^{M \times N}$ such that $\|\mathbf{\Theta}_j\|_2 = 1 \ \forall j = 1$ to N
- Measurement vector $\mathbf{y} \in \mathbb{R}^{M \times 1}$
- Sparsity order k
- Iteration number $i = 0$
- Initial estimate of the sparse representation $\mathbf{s}^i = \mathbf{0}$
- Initial residual $\mathbf{r}^i = \mathbf{0}$

1: **while** $\|\mathbf{r}^i\|_2^2 > \epsilon$ **do**
2:     $i = i + 1$
3:     $\mathbf{r}^i = \mathbf{y} - \mathbf{\Theta}\hat{\mathbf{s}}^{i-1}$ ▷ Residual
4:     $\mathcal{J}^i = \mathcal{L}_{2k}(\mathbf{\Theta}^T \mathbf{r}^i)$ ▷ indices of $2k$ largest coefficients in $\mathbf{\Theta}^T \mathbf{r}_i$
5:     $\mathcal{S}^i = \mathcal{S}^{i-1} \cup \mathcal{J}^i$ ▷ updated support
6:     $\mathbf{b} = \left(\mathbf{\Theta}_{\mathcal{S}^i}^T \mathbf{\Theta}_{\mathcal{S}^i}\right)^{-1} \mathbf{\Theta}_{\mathcal{S}^i}^T \mathbf{y}$ ▷ estimate of largest coefficients
7:     $\mathcal{S}^i = \mathcal{L}_k(\mathbf{b})$ ▷ indices of k largest coefficients in $\mathbf{b}$
8:     $\hat{\mathbf{s}}_{\mathcal{S}^i}^i = \mathbf{b}_{\mathcal{S}^i}$ and $\hat{\mathbf{s}}_{(\mathcal{S}^C)^i}^i = \mathbf{0}$ ▷ estimate of best $k-$sparse approximation
9: **end while**
Output: Estimate of the sparse representation: $\hat{\mathbf{s}}$.

In general, if the measurement matrix satisfies the RIP, then greedy techniques perform similarly to convex relaxation techniques with an overwhelming probability as shown in [16, 17]. Both the convex relaxation and greedy techniques exploit the properties of the underlying sensing matrix and sparsity order of the signal. Thus, the key components of the CS are the combination of a sensing matrix and the inherent sparsity of the signal in a suitable transform domain.

1.4 The purpose of sparsity order estimation

For a practical CS system, the knowledge of the exact sparsity order of a compressible signal is of great importance during CS acquisition and recovery. In the CS acquisition stage, the sparsity order dictates the minimum number of measurements, i.e., the size of the measurement vector to be acquired. In the CS recovery stage, the sparsity order controls the quality of estimation of the compressible signal. Therefore, the Sparsity Order Estimation (SOE) is crucial.

The number of CS measurements M required for the perfect recovery of an $N-$dimensional compressible signal using ℓ_1 norm based convex relaxation technique or greedy techniques depends on the sparsity order k of the underlying compressible signal, and is given in [3, 4] as

$$M \geq 2k \log \left(\frac{N}{k} \right) \tag{1.14}$$

when the sensing matrix Θ obeys RIP [12] and the components of the sensing matrix Θ are i.i.d random Gaussian variables. In many recovery algorithms, the optimal tuning of parameters requires knowledge of the sparsity order of the underlying signal. For example, in LASSO [5] techniques of convex relaxation-based CS recovery, as shown in Equation (1.13), the tuning parameter $\nu = \sigma_\vartheta \sqrt{2 \log k}$ is a function of the sparsity order k and the measurement noise level σ_ϑ for perfect recovery. Similarly, for greedy algorithms, such as OMP and CoSaMP, the recovery performance and number of iterations depend on the sparsity order k.

Most CS studies assume that the sparsity order is known in advance [18, 19, 20, 21]. This assumption makes the practical applicability of the CS theory difficult. An improper assumption of sparsity order during CS acquisition results in either an insufficient or an excess number of measurements, and affects the quality of the reconstructed signal during recovery. Similarly, improper assumptions during CS recovery result in either early or late termination of CS recovery algorithms, leading to either poor reconstruction or resource wastage.

1.4.1 Dynamic SOE

In many applications, the sparsity order, support, and amplitudes of the supporting components are not only unknown but also time varying [22, 23, 24, 25, 26, 27]. For example, in video imaging systems, the sparsity order and support vary between frames [27], whereas in wireless communication systems, the sparse CSI and tap delay locations vary

with time [26]. Therefore, in such practical systems, it is essential to estimate the sparsity order at each time instant of observation during the acquisition and recovery of sparse or compressible signals for better use of available resources. Therefore, the SOE in a dynamic environment has become an important topic for both CS acquisition and recovery.

1.5 Relevant and recent literature on SOE

There are direct and indirect SOE methods for estimating the sparsity order. Direct SOE methods are applicable for both CS acquisition and recovery, where the SOE is performed without signal reconstruction. However, indirect SOE methods are only applicable during CS recovery because the SOE is through the support estimation process of CS recovery. Recent and relevant studies on direct SOE methods are available in [27, 28, 29, 30, 31, 32, 33, 34, 35, 36] and indirect SOE methods are available in [37, 38, 39, 40, 41, 42, 43, 44].

1.5.1 Direct SOE methods

In the seminal work by Lan et al. [27], the direct SOE for an image was based on image complexity. Here, image complexity is estimated from the image texture and edge density using additional measurements that are not useful for image recovery. Other direct SOE methods exploit the design and characteristics of sensing matrices and the temporal correlation of the measurements, as discussed below.

1.5.1.1 SOE based on specially designed matrices

Direct SOE methods that use specially designed sensing matrices are available in [28, 30, 32]. Lopes [28] introduced the SOE for sparse signals from composite sensing matrix-based measurements, where the composite sensing matrix comprises a random Cauchy sensing matrix and a random Gaussian sensing matrix. Here, SOE is performed using the ratio of ℓ_1 norm to ℓ_2 norm of the sparse signal. As the random Cauchy sensing matrix preserves the ℓ_1 norm and the random Gaussian matrix preserves the ℓ_2 norm with overwhelming probability, the ratio of ℓ_1 norm to ℓ_2 norm of the measurements is used for the SOE. Although the measurements obtained using the random Cauchy sensing matrix are used for the SOE, they are not helpful for recovery [30]. Moreover, the distribution of the Cauchy sensing matrix entries depends on the knowledge of noise statistics.

The use of sparse random Gaussian sensing matrices for SOE of true sparse signals was proposed in [30]. Owing to the sparseness of the sensing, the obtained measurements are

distributed to a Gaussian mixture model parameterized by the sparsity order of the under-lying signal. This mixture model is approximated as a two-component Gaussian mixture model (2-GMM) so that the sparsity order can be easily estimated using the Expectation Maximization (EM) technique. However, the 2-GMM approximation is valid for a certain regime of signal's sparsity (which is unknown) and measurement's sparsity. In addition, the accuracy of this approximation and the efficiency of the SOE depend on *a priori* infor-mation about the unknown signal statistics, making the method practically infeasible. In addition, the signal recovery performance degrades because of the larger mutual coherence measure of the sparse Gaussian matrix.

A specially designed sensing matrix with a Khatri-Rao structure was recently presented for SOE [32]. Here, a single measurement vector is obtained from a specially designed matrix for the SOE. By suitable rearrangement of the obtained measurement vector into a matrix, the sparsity order is estimated from the rank of the matrix. However, the construc-tion of the specially designed matrix depends on the signal type and the SOE performance is limited to a lower sparsity order.

1.5.1.2 SOE by exploiting the characteristics of random sensing matrices

Direct SOE methods that exploit the characteristics of random sensing matrices are avail-able in [33, 34, 35, 36]. In [33], a spectrum sensing algorithm was used to solve an op-timization problem to remove the measurement noise effect, followed by an energy mini-mization problem using QR decomposition of the sensing matrix and applying a threshold to obtain the sparsity order. This algorithm requires the signal and noise power to be known *a priori* to tune the threshold parameter and is not suitable for estimating a higher sparsity order.

Exploiting the autocorrelation and cross-correlation properties of the column vectors of the Gaussian sensing matrix, a Two-Step Adaptive Compressive Spectrum Sensing (TS-ACSS) was proposed in [34]. In the first step, a coarse SOE is performed by identifying the slope change in the ordered arrangement of the inner product results of the column vectors of the sensing matrix with the obtained measurements. The SOE is then refined using the CS recovery of the spectrum and by comparing the estimates of the binary chan-nel occupation in multiple iterations in the second step. This method not only requires more measurements for SOE but also involves multiple iterations that make the method computationally intensive and slow.

Recently, SOE was performed iteratively by exploiting the RIP of a random sensing matrix with an isometry constant δ_k in [35]. Here, a Sparse Pre-estimated Adaptive Match-

ing Pursuit (SPAMP) algorithm is introduced using the criteria of sparsity underestimation and overestimation for the SOE. When a measurement vector $\mathbf{y}$ is obtained from the signal with sparsity order k using a sensing matrix $\boldsymbol{\Theta}$, then the sparsity underestimation criterion, i.e., $\|\boldsymbol{\Theta}_{\mathcal{S}}^T \mathbf{y}\|_2 \geq \frac{1-\delta_k}{\sqrt{1+\delta_k}} \|\mathbf{y}\|_2$ is satisfied if the estimated support $\mathcal{S}$ of the underlying signal has the cardinality $\hat{k} < k$. Otherwise, the sparsity overestimation criterion, i.e., $\|\boldsymbol{\Theta}_{\mathcal{S}}^T \mathbf{y}\|_2 \geq \frac{1+\delta_k}{\sqrt{1-\delta_k}} \|\mathbf{y}\|_2$ is satisfied if the estimated support $\mathcal{S}$ has the cardinality $\hat{k} \geq k$. Thus, iteratively, the underestimation and overestimation criteria are tested and the estimated sparsity order $\hat{k}$ is adjusted using two factors, i.e., a weak matching factor and an estimation factor until $\hat{k} = k$. Both these factors must be optimally chosen so that the trade-off between the speed of convergence and the accuracy of the estimation is minimized. However, their optimal values require knowledge of the sparsity order and are difficult to fix in practical scenarios, because the sparsity order is unknown and varies over time. In addition to this, computing δ_k for a given matrix $\boldsymbol{\Theta}$ and sparsity order k is NP-hard and not tractable as shown in [45].

In [36], a Sparsity Adaptive estimation Matching Pursuit algorithm based on a Sensing Dictionary (SAMP-SD) was presented for SOE. The method presented here is similar to that of [35], which considers the mutual correlation constant instead of the isometry constant. Here, a sensing dictionary $\boldsymbol{\Sigma}$ and a measurement matrix $\boldsymbol{\Theta}$ are constructed such that they have the same dimensions, and $\boldsymbol{\Sigma}$ has a weak mutual correlation with $\boldsymbol{\Theta}$, i.e., $\boldsymbol{\Sigma}^T \boldsymbol{\Theta} \approx \mathbb{I}$. Thus, $\boldsymbol{\Sigma}^T \mathbf{y}$ provides an estimate of sparsity order. However, this SOE method is not efficient for time-varying sparsity order as $\boldsymbol{\Sigma}$ and $\boldsymbol{\Theta}$ must be cooperatively constructed using iterative algorithms when the size of the measurement vector changes with the sparsity order.

1.5.1.3 SOE based on multiple observations

Within the Multiple Measurement Vector (MMV) framework, there exist direct SOE methods based on (i) the eigenvalues of the covariance matrix [46, 31] and (ii) the trace of the covariance matrix [29]. All these MMV-based SOE methods require multiple snapshots of the underlying signal and assume that the sparsity order is static during the period of multiple observations.

Using the random matrix theory, an eigenvalue-based SOE method is presented in [46]. A lookup table is maintained here, with entries for eigenvalues and the corresponding sparsity order. The maximum eigenvalue of the covariance matrix of the measurement vectors was calculated from multiple measurements, and the corresponding sparsity order was selected from a lookup table. However, the computation of eigenvalues in this method results

in high computational costs.

SOE was performed in [31] by identifying the slope change in the ordered eigenvalues of the covariance matrix of the measurements. However, this method is computationally intensive, similar to [46] for finding the eigenvalues, and is not suitable for estimating a higher sparsity order.

From the trace of the covariance matrix of the measurement vectors, SOE was performed in [29]. Although this method has comparatively less computational complexity than eigenvalue-based methods, its performance is affected under the scenarios of uncorrelated measurement vectors.

1.5.2 Indirect SOE methods

A Relative Threshold-based Sparsity Estimation method (RTSE) is proposed in [37] which performs SOE based on a reconstruction algorithm. Here, the threshold for finding the largest components is based on the training set and cannot be fixed *a priori* limiting its application.

The Sparsity adaptive Matching Pursuit (SaMP) algorithm [38] and its variants such as the Adaptive Step size-SaMP (AS-SaMP) algorithm [39], Modified CoSaMP (MCoSaMP) algorithm [44], and Sparsity Adaptive Segmented Orthogonal Matching Pursuit (SAStOMP) algorithm [43], estimate the sparsity indirectly through a variable step size and gradually increase the estimated support set to match the original. However, these SaMP algorithms require a step size, whose optimal value depends on an unknown sparsity order. An Optimized Adaptive Matching Pursuit (OAMP) algorithm was proposed in [40] which is similar to SaMP except for the energy-entropy-based order determination in updating the support.

Recently, Deterministic Binary Block Diagonal (DBBD) matrix-based sensing [41] and Kronecker-based recovery [42] have been proposed for acquiring and recovering compressible signals. Here, the DBBD sensing matrix reduces hardware complexity. However, its structure accumulates the energy of significant neighboring components, making support estimation difficult under noisy settings and leading to degraded recovery performance.

1.6 Literature on estimating time-varying signals

There exist literature on tracking of dynamic sparse signals [23, 26, 47] which assumes that sparse signals are slow-varying and the amplitude of each supporting component varies

according to a predefined Gauss-Markov model exhibiting temporal correlation. They recover sparse signals by estimating the support and do not involve in sparsity order estimation. Kalman filtering on CS recovery is proposed in [23] where support is estimated using convex optimization based CS recovery. Dynamic Compressed Sensing Approximate Message Passing (DCS-AMP) based tracking is proposed in [26], where in addition to amplitude modeling, the support is modeled using the Discrete Markov process. These support estimation techniques require a longer execution time and are suitable only during recovery and not during CS acquisition. An another method is Sparse Bayesian Learning (SBL) [47] method which has a Bayesian approach. SBL models the amplitude of components of sparse vector as Gaussian variables and their variances are hyper-parameters. From the CS measurements, the Bayesian approach estimates the posterior density functions for the values of non-zero components of signal. Since Bayesian approach provides full posterior density function about the components of the signal instead of a point estimate, a sense of confidence is possible for the estimates of non-zero components. In addition to this, it provides estimate for the posterior density function for the noise involved in the measurements. All these CS methods for estimating time-varying signals are useful during recovery and not optimal for determining the number of measurements during acquisition.

1.7 The need for composite sensing

Composite sensing is a method of sensing sparse or compressible signals using two or more CS measurement matrices. Inspired by SOE methods that perform estimation by exploiting the characteristics of the sensing matrix design, a composite CS measurement system for maximizing the SOE and recovery performance is presented in this research work. In general, sensing or measurement matrices can be classified into two categories: random and deterministic. The entries of the random sensing matrices are either i.i.d Gaussian or Bernoulli variates. These random sensing matrices satisfy the RIP required for the perfect recovery of the signal from the obtained measurements, and are nonadaptive and suitable for recovering all types of sparse or compressible signals. However, they are not optimal because they are primarily unstructured and not designed to exploit the structure of the signals. Compared to random sensing matrices, deterministic sensing matrices are fully structured and are signal- or application-specific. Chirp sensing matrices, Reed-Muller sensing matrices, Binary Sensing Matrices (BSM) such as Quasi-Cyclic Low-Density Party Check (QC-LDPC) matrices, and binary Bose-Chaudhuri-Hocquenghem (BCH) are examples of deterministic sensing matrices. Recently, sparse BSM [48, 49, 50, 51, 52, 53, 54]-based

CS measurement systems have been considered because they are multiplier-less (for time-domain sparse signals) and perform faster compression. Although these matrices are simple to implement, they possess the weaker RIP property[2] [55] and require additional measurements for recovery. However, the measurements obtained using these matrices have specific statistical properties that are suitable for SOE. Considering the guaranteed RIP and SOE using random and deterministic matrices, respectively, the sensing matrix can be a composite matrix comprising of both types of matrices for efficient CS acquisition and recovery.

1.8   Recent literature on composite sensing

Several studies have been conducted on composite sensing matrices [56, 28] that are used for SOE and support estimation. A Hybrid Compressed Sensing (HCS) method was proposed in [56], where two sub-matrices: a sparse complex-valued sub-matrix for support estimation and a random dense real-valued sub-matrix for lowering the number of measurements are used. Using this method, the success of signal recovery depends on the size of both the submatrices. However, the size of the submatrices is a function of the sparsity order that is being estimated. The Lopes method [28] uses a different composite CS measurement system comprising the Cauchy and Gaussian submatrices. This method is inefficient because the matrices are designed with *a priori* knowledge of noise statistics and the Cauchy matrix-based measurements are not helpful for recovery. Thus, there exists a need to develop a composite sensing matrix that is suitable for SOE and recovery.

1.9   Motivations

The problems associated with the above-discussed existing SOE methods and composite sensing systems are summarized as follows.

1. Existing SOE methods have the following assumptions and constraints that are not suitable for practical applications.

   - The signal and noise statistics are known *a priori* [28, 29, 30, 32].

   - The support is time-invariant [29, 34].

   - The sparsity order is restricted [31, 32, 33].

---

[2]In weaker RIP, the ℓ_2 norm is replaced with ℓ_p norm where $p \approx 1$

- The algorithm parameters are dependent on sparsity order [35, 36, 37, 38, 39, 40, 43, 44].

- They are computationally intensive [30, 31, 34].

- They require separate measurements just for SOE [27, 28].

2. The construction of existing composite sensing systems is based on the knowledge of sparsity order and statistics of noise and some of the obtained measurements are not helpful for recovery.

Thus, there is a strong need to develop a composite sensing system and an SOE algorithm suitable for both CS acquisition and recovery. Motivated by the purpose of SOE and to overcome the aforementioned limitations of the existing SOE methods, there exists a need to develop an efficient and practically implementable CS measurement system that performs a novel instantaneous SOE on the fly using the same set of measurements acquired for use during recovery. The SOE method should be capable of estimating the time-varying sparsity order and should not require any prior information regarding signal and noise statistics. Considering these requirements, the research problem statement of this thesis is formulated as follows.

1.10 Research problem statement

Problem 1.1. Model the CS acquisition system, CS recovery system, and time-varying parameters of dynamic sparse and compressible signals.

Problem 1.2. Estimate the time-varying parameters from noisy measurements and derive the performance bounds.

Problem 1.3. Apply the derived estimators for the recovery of real-world time-varying sparse and compressible signals.

1.11 Research contributions of this thesis

The research contributions of this thesis towards solving the aforementioned problem statement are summarized as follows.

1. A practical composite CS measurement system is developed with the help of a composite sensing matrix constituted by a deterministic sparse Binary Sensing Matrix (BSM) and a random dense Gaussian Sensing Matrix (GSM).

2. Time-varying parameters such as *(i)* amplitudes of significant and insignificant components are modeled using Gaussian random processes and *(ii)* sparsity order is modeled using discrete Markov birth-death random process. The composite sensing system, along with suitable statistical models for the time-varying parameters of sparse and compressible signals, motivates the development of novel SOE algorithms and maximises the recovery performance.

3. Two different approaches are devised for efficient SOE. The first approach performs Maximum Likelihood (ML) estimation of the sparsity order using the GSM, followed by ML Sequence (MLS) estimation using the Viterbi algorithm to refine the ML estimates. The second approach performs ML estimation using the BSM, followed by Kalman filtering to refine the ML estimates.

4. The proposed structure of BSM of composite sensing system provides initial estimates about the supporting significant components. The BSM-aided support estimation improves the speed of the greedy CS recovery algorithms by at least 25%.

5. The application of proposed composite sensing system, SOE algorithms and BSM-aided recovery algorithm is demonstrated on compressing the real-world vibration signals and estimation of CSI of communication systems.

Chapter 2

Modeling compressed sensing System

This chapter discusses a composite sensing system-based CS acquisition and its model. The proposed CS acquisition system is illustrated in Figure 2.1, which comprises the following elements: (A) compressible signal, (B) composite sensing system, (C) measurement vector, (D) SOE system, and (E) measurement vector size (number of measurements) estimator. A continuous-time compressible signal $\mathbf{x}(t)$ is acquired with the help of a composite sensing system Θ that includes Gaussian and impulse basis functions (both of which are modulated *a priori* using the inverse of the compressible signal's representation basis). The basis functions independently multiply the compressible signal and perform integration and dumping every $\mathcal{T}$ seconds to generate the measurement vector $\mathbf{y}(n\mathcal{T})$. Here, t represents the continuous time index, $n\mathcal{T}$ represents the sampling time index, and $n \in \mathbb{Z}$ represents the discrete time step. In the following, the sampling index $n\mathcal{T}$ is represented by n. Measurement vector $\mathbf{y}(n)$ contains $M(n)$ measurements. We know that an efficient CS system obtains $\mathbf{y}(n)$ in such a way that the measurement vector size $M(n)$ is the smallest necessary for recovery. Because the sparsity order determines the measurement vector size, it is estimated during the acquisition using BSM/GSM-based SOE techniques. An efficient CS system does not use separate measurements just for SOE. Instead, it estimates the sparsity order $k(n)$ by using the same set of measurements obtained for recovery. However, before obtaining the measurement vector $\mathbf{y}(n)$ using Equation (1.14), the measurement vector size $M(n)$ must be estimated based on the sparsity order $k(n)$. Thus, it is practically impossible to simultaneously estimate the sparsity order and the measurement vector size for the current time step n. Hence, the estimated sparsity order $\widehat{k}(n)$ for the current time step n is used to determine the number of measurements $M(n + 1)$ for the next time step $n + 1$ because naturally occurring time-varying compressible signals are quasi-static and exhibit a stronger temporal correlation.
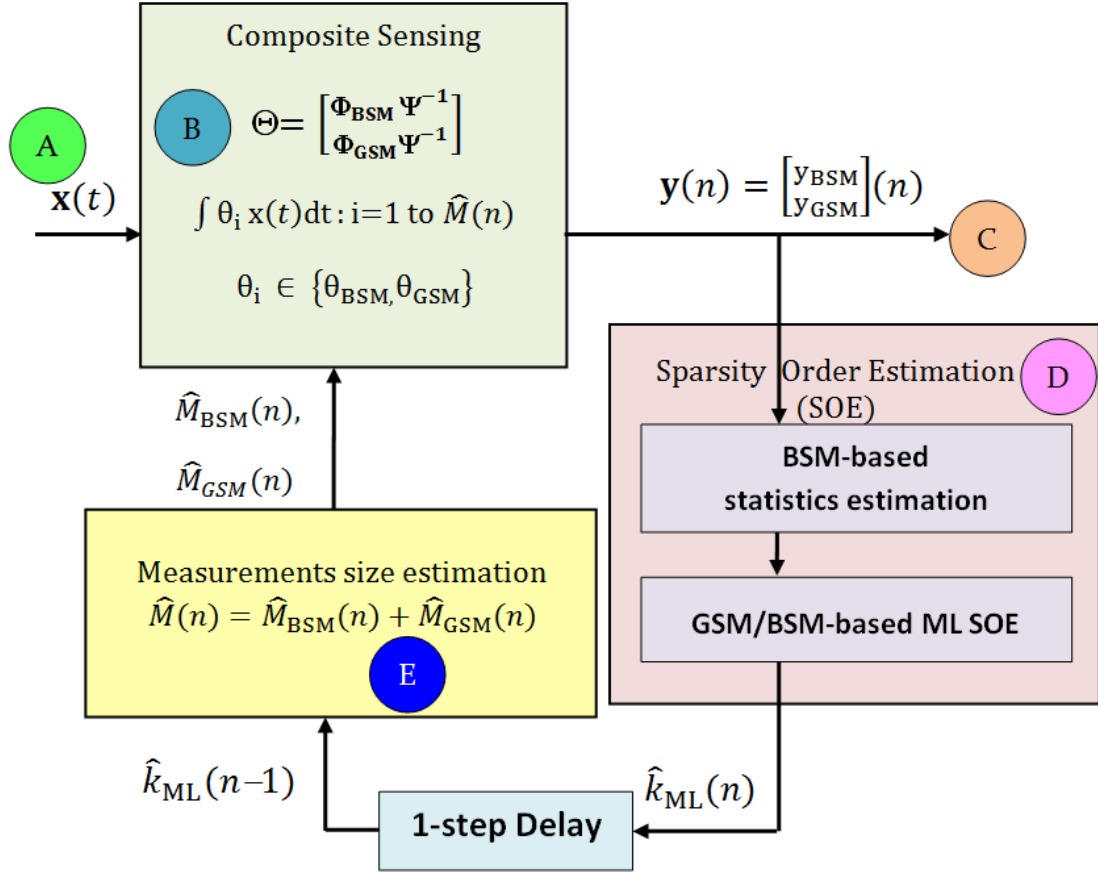
Figure 2.1: The block diagram of the proposed CS acquisition System.

2.1 Compressible signal model

Consider a continuous-time dynamic compressible signal $\mathbf{x}(t)$ with an $N-$dimensional representation $\mathbf{s}(t) = \{s_1(t), s_2(t), \ldots, s_N(t)\}$ on an orthonormal sparsifying basis $\mathbf{\Psi} = \{\psi_1(t), \psi_2(t), \ldots, \psi_N(t)\}$ at t^{th} time as follows:

$$\mathbf{x}(t) = \sum_{j=1}^{N} s_j(t)\psi_j(t) \tag{2.1}$$

where $s_j(t)$ denotes the coefficient of the jth basis, $\psi_j(t)$.

Using the basis $\mathbf{\Psi}$ given in Equation (2.1), the sparse representation $\mathbf{s}(t)$ has only $k(t)$ significant coefficients and the remaining are insignificant. The $k(t)$ significant coefficients have magnitudes well above a specified threshold and contain most of the energy of the compressible signal, thereby defining the sparsity order of the signal. The collec-

tion of indices of these $k(t)$ significant coefficients is the support set $\mathcal{S}_{\mathbf{s}(t)}$. Any classical compression technique retains these $k(t)$ significant coefficients and their support $\mathcal{S}_{\mathbf{s}(t)}$ and leaves the remaining $N - k(t)$ insignificant coefficients to obtain signal $\mathbf{s}_{\mathcal{S}}(t)$ which is a $k(t)-$sparse approximation of $\mathbf{s}(t)$. The error due to this approximation has energy $\mathbb{E}\{\|\mathbf{s(t)} - \mathbf{s}_{\mathcal{S}}(t)\|_2^2\} = (1 - E_s)\mathbb{E}\{\|\mathbf{s}(t)\|_2^2\}$, where E_s is user defined and is typically ≥ 0.95. This determines the threshold for distinguishing between significant and insignificant coefficients. Thus, the representation $\mathbf{s}(t)$ can be written as the sum of two disjoint signals, a $k(t)-$sparse signal $\mathbf{s}_{\mathcal{S}}(t)$, and an $(N - k(t))-$dense signal $\mathbf{s}_\epsilon(t)$, i.e.,

$$\mathbf{s}(t) = \mathbf{s}_{\mathcal{S}}(t) + \mathbf{s}_\epsilon(t) \tag{2.2}$$

where $\mathbf{s}_{\mathcal{S}}(t)$ contains $k(t)$ nonzero significant coefficients and $N - k(t)$ zeros, and $\mathbf{s}_\epsilon(t)$ contains $N - k(t)$ nonzero insignificant coefficients and $k(t)$ zeros.

Based on the compressible distributions given in [57], the insignificant coefficients are approximated as i.i.d Gaussian noise such that $s_j(t) \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ when the jth coefficient is insignificant. Considering that the entries of the significant components $(s_{\mathcal{S}})_{j:(j \in \mathcal{S}_{\mathbf{s}})}$ are the outcomes of a Gaussian random variable, the mean value of significant components is given by μ_s and the variance of the significant components is given by σ_s^2.

2.2 Composite CS acquisition model

The CS acquisition model acquires $\mathbf{x}(t)$ for every $\mathcal{T}$ seconds to obtain an $M(n\mathcal{T})-$dimensional discrete measurement vector $\mathbf{y}(n\mathcal{T})$ using $M(n\mathcal{T})$ composite sensing basis functions $\boldsymbol{\Theta} = \{\boldsymbol{\theta}_i(t)\}_{i=1}^M(n\mathcal{T})$. An ith measurement $y_i(n\mathcal{T})$ is obtained as given below:

$$y_i(n\mathcal{T}) = \int_{t=(n\text{-}1)\mathcal{T}}^{n\mathcal{T}} \boldsymbol{\theta}_i(t)\mathbf{x}(t)dt + \vartheta_i(n\mathcal{T}), 1 \leq i \leq M(n\mathcal{T}) \tag{2.3}$$

where a few of $\boldsymbol{\theta}_i(t)$ are generated using sparse impulse basis functions and the rest are with dense Gaussian basis functions and $\vartheta_i(n\mathcal{T}) \sim \mathcal{N}(0, \sigma_\vartheta^2)$ is the ith component of measurement noise $\boldsymbol{\vartheta}(t)$ and is generally modeled as i.i.d zero mean Gaussian noise [4].

The discrete versions of Equations (2.1), (2.2), and (2.3) are given as,

$$\mathbf{x}(n) = \boldsymbol{\Psi}\mathbf{s}(n), \quad \mathbf{x}(n) \in \mathbb{R}^{N \times 1} \tag{2.4}$$

$$\mathbf{s}(n) = \mathbf{s}_{\mathcal{S}}(n) + \mathbf{s}_\epsilon(n), \quad \mathbf{s}(n) \in \mathbb{R}^{N \times 1} \tag{2.5}$$

$$\mathbf{y}(n) = \boldsymbol{\Theta}\mathbf{x}(n) + \boldsymbol{\vartheta}(n), \quad \mathbf{y}(n) \in \mathbb{R}^{M \times 1} \tag{2.6}$$

where continuous time t is substituted with discrete time step n. The discrete compressible signal $\mathbf{x}(n) = \{x_j(n)\}_{j=1}^{N}$ has N samples in every time step of duration $\mathcal{T}$ seconds and has a representation $\mathbf{s}(n) = \{s_j(n)\}_{j=1}^{N}$ when projected onto an $N \times N$ orthonormal signal basis matrix $\boldsymbol{\Psi}$. The discrete version of sensing basis is the $M \times N-$dimensional sensing matrix $\boldsymbol{\Theta}$.

Substituting Equation (2.4) in Equation (2.6), the discrete CS acquisition model becomes,

$$\mathbf{y}(n) = \boldsymbol{\Theta}\boldsymbol{\Psi}\mathbf{s}(n) + \boldsymbol{\vartheta}(n). \tag{2.7}$$

The $\boldsymbol{\Theta}$ considered is a composite matrix such that

$$\boldsymbol{\Theta} = \begin{bmatrix} \boldsymbol{\Phi}_{BSM}\boldsymbol{\Psi}^{-1} \\ \boldsymbol{\Phi}_{GSM}\boldsymbol{\Psi}^{-1} \end{bmatrix} \tag{2.8}$$

where $\boldsymbol{\Phi}_{BSM}$ is the $M_{BSM} \times N-$dimensional deterministic sparse BSM, $\boldsymbol{\Phi}_{GSM}$ is the $M_{GSM} \times N-$dimensional random GSM, and $M = M_{BSM} + M_{GSM}$. The sparse BSM is deterministic such that it is constructed by horizontally stacking the identity matrices. For example, 4×12 BSM is constructed by stacking three 4×4 horizontally as given below.

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}}_{\boldsymbol{\Phi}_{BSM}}$$

The multiplication of BSM and GSM with inverse of sparsifying basis, i.e., $\boldsymbol{\Psi}^{-1}$ results in the direct acquisition of $\mathbf{s}(n)$. Thus Equation (2.7) becomes,

$$\mathbf{y}(n) = \begin{bmatrix} \mathbf{y}_{BSM} \\ \mathbf{y}_{GSM} \end{bmatrix} (n) = \begin{bmatrix} \boldsymbol{\Phi}_{BSM} \\ \boldsymbol{\Phi}_{GSM} \end{bmatrix} \mathbf{s}(n) + \boldsymbol{\vartheta}(n) \tag{2.9}$$

$$= \boldsymbol{\Phi}\mathbf{s}(n) + \boldsymbol{\vartheta}(n) \tag{2.10}$$

where $\boldsymbol{\Phi} = \begin{bmatrix} \boldsymbol{\Phi}_{BSM} \\ \boldsymbol{\Phi}_{GSM} \end{bmatrix}$.

Now substituting Equation (2.5) in Equation (2.10), the measurement vector model

becomes,

$$\mathbf{y}(n) = \mathbf{\Phi}\mathbf{s}_{\mathcal{S}}(n) + \mathbf{\Phi}\mathbf{s}_{\varepsilon}(n) + \boldsymbol{\vartheta}(n), \tag{2.11}$$

$$= \mathbf{y}_{\mathcal{S}}(n) + \mathbf{y}_{\epsilon}(n) + \boldsymbol{\vartheta}(n) \tag{2.12}$$

where $\mathbf{y}_{\mathcal{S}}(n)$ and $\mathbf{y}_{\epsilon}(n)$ are the components of the measurement vector corresponding to the significant and insignificant coefficients, respectively. The statistics of these components are given as follows. The discrete time notation n is dropped for the sake of brevity in the subsequent section.

2.3 Statistics of BSM measurements

Consider the sparse BSM having λ fraction of zeros in each row i.e.,

$$\lambda = \frac{\text{number of zeros in each row}}{N} = 1 - \frac{\sum_{j=1}^{N} \phi_{i,j}}{N}.$$

2.3.1 Statistics due to significant components

Because the significant components are randomly distributed, it may be noted from Equation (2.12) that the ith BSM measurement $(y_{\mathcal{S}})_i$ of $\mathbf{y}_{\mathcal{S}}$ is a random sum of significant coefficients, i.e.,

$$(y_{\mathcal{S}})_i = \sum_{j=1}^{N} \Phi_{i,j}(s_{\mathcal{S}})_j = \sum_{j \in \{\mathcal{S}_{\mathbf{\Phi}_i} \bigcap \mathcal{S}_{\mathbf{s}}\}} (s_{\mathcal{S}})_j \tag{2.13}$$

where $\mathcal{S}_{\mathbf{\Phi}_i}$ is the support set of ith row of $\mathbf{\Phi}_{BSM}$ and the set $\{\mathcal{S}_{\mathbf{\Phi}_i} \bigcap \mathcal{S}_{\mathbf{s}}\}$ contains the indices which are common to both the support sets $\mathcal{S}_{\mathbf{\Phi}_i}$ and $\mathcal{S}_{\mathbf{s}}$.

As each $(s_{\mathcal{S}})_{j:(j \in \mathcal{S}_{\mathbf{s}})} \sim \mathcal{N}(\mu_s, \sigma_s^2)$ and $(s_{\mathcal{S}})_{j:(j \notin \mathcal{S}_{\mathbf{s}})} = 0$, every BSM measurement $(y_{\mathcal{S}})_i$ is a random sum of Gaussians and is approximated by

$$(y_{\mathcal{S}})_i \sim \mathcal{N}(\ell_s \mu_s, \ell_s(\sigma_s^2 + \lambda \mu_s^2)) \tag{2.14}$$

where $\ell_s = k(1 - \lambda)$ is the average number of significant coefficients contributing to $(y_{\mathcal{S}})_i$.

The reason for $\ell_s = k(1 - \lambda)$, the mean of $(y_{\mathcal{S}})_i$ as $\ell_s \mu_s$ and the variance of $(y_{\mathcal{S}})_i$ as $\ell_s(\sigma_s^2 + \lambda \mu_s^2)$ are as follows.

The probability that an entry in each row of the BSM as 1 is $(1 - \lambda)$ and the probability that a coefficient in $\mathbf{s}$ as significant is k/N. Thus, the probability that an index to be

common to both support sets, $\mathcal{S}_{\Phi_i}$ and $\mathcal{S}_{\mathbf{s}}$ is $(k/N)(1-\lambda)$. Because there are N indices in each row of the BSM, the average number of indices common to both support sets $\mathcal{S}_{\Phi_i}$ and $\mathcal{S}_{\mathbf{s}}$, i.e., the average number of significant coefficients that contribute to $(y_{\mathcal{S}})_i$ is $\ell_s = N(k/N)(1-\lambda) = k(1-\lambda)$. In other words, because each measurement is a random sum of the significant components, the probability of selecting ℓ significant components from k significant components is binomially distributed with the probability of success in choosing a significant component as $1 - \lambda$. Hence, the mean of the binomial distribution is the average number of significant coefficients, i.e., $\mathbb{E}\{\ell\} = \ell_s = k(1-\lambda)$ and the variance of the binomial distribution is $k(1-\lambda)\lambda$.

The average number of significant coefficients is shown empirically by performing a simulation where the BSM was constructed with $\lambda = 0.95$, $M = 2500$, $N = 5000$ and a binary sparse signal was generated with sparsity order $k = 200$, i.e., there were 200 nonzero components and all of these components $s_{j:(j \in \mathcal{S}_{\mathbf{s}})} = 1$ with their support is randomly distributed. The histogram of $(y_{\mathcal{S}})_i$ corresponding to the significant components is shown in Figure 2.2 which shows that $\mathbb{E}\{(y_{\mathcal{S}})_i\} = \frac{1}{M}\sum_{i=1}^{M}(y_{\mathcal{S}})_i = \ell_s = k(1-\lambda) = 10$ is the average and has the highest probability.
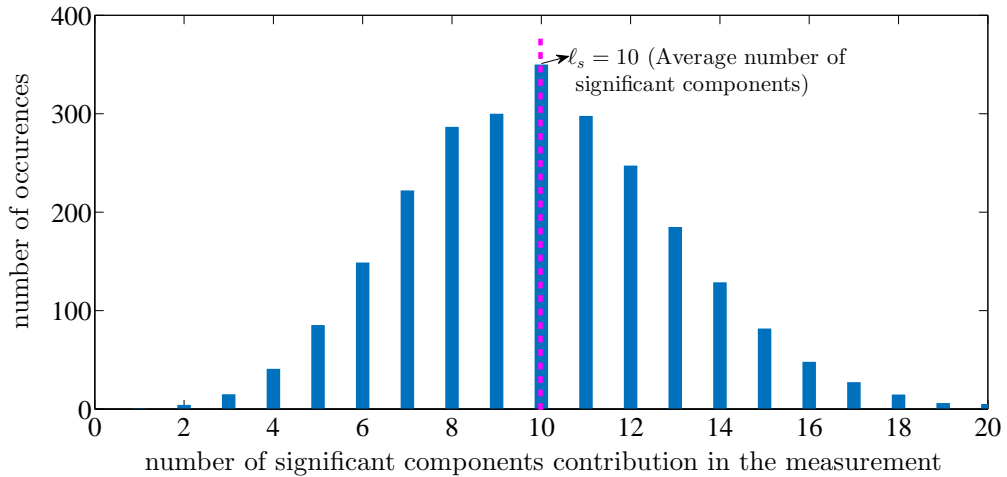

Figure 2.2: Empirical results on the number of significant coefficients contribution and its frequency of occurrences from the BSM-based measurements showing that each measurement contains an average number of significant coefficients as $\mathbb{E}\{(y_{\mathcal{S}})_i\} = \ell_s = k(1-\lambda) = 200(1 - 0.95) = 10$.

As there are $k(1 - \lambda)$ average significant coefficients and each significant coefficient's mean is μ_s, the mean of each measurement is $\mathbb{E}\{(y_{\mathcal{S}})_i\} = k(1 - \lambda)\mu_s$.

Equation (2.13) is the random sum of random variables. The random sum is due to the

randomness in ℓ out of k significant coefficients, and the random variables are significant coefficients. Hence, the variance of the random sum of random variables is given as

$$
\begin{aligned}
\text{Var}\{(y_\mathcal{S})_i\} &= \mathbb{E}\{\ell\}\text{Var}\{s_j\} + \text{Var}\{\ell\}(\mathbb{E}\{s_j\})^2 \\
&= k(1-\lambda)\sigma_s^2 + k(1-\lambda)\lambda\mu_s^2 \\
&= k(1-\lambda)(\sigma_s^2 + \lambda\mu_s^2) \\
&= \ell_s(\sigma_s^2 + \lambda\mu_s^2).
\end{aligned}
$$

2.3.2 Statistics due to insignificant components

Similar to $(y_\mathcal{S})_i$, the ith BSM measurement $(y_\epsilon)_i$ of $\mathbf{y}_\epsilon$ is a random sum of i.i.d insignificant coefficients, i.e.,

$$
(y_\epsilon)_i = \sum_{j=1}^{N} \Phi_{i,j}(s_\epsilon)_j = \sum_{j \in \{\mathcal{S}_{\Phi_i} \bigcap \mathcal{S}_\epsilon\}} (s_\epsilon)_j \tag{2.15}
$$

where $\mathcal{S}_\epsilon$ is the support set of insignificant coefficients and the set $\{\mathcal{S}_{\Phi_i} \bigcap \mathcal{S}_\epsilon\}$ contains the indices which are common to both the support sets $\mathcal{S}_{\Phi_i}$ and $\mathcal{S}_\epsilon$.

From Equation (2.15), the ith BSM measurement corresponding to insignificant coefficients is,

$$
(y_\epsilon)_i \sim \mathcal{N}(0, \ell_\epsilon \sigma_\epsilon^2) \tag{2.16}
$$

where $\ell_\epsilon = (N-k)(1-\lambda)$ is the average number of insignificant coefficients that contribute to $(y_\epsilon)_i$. The result for ℓ_ϵ was verified empirically by performing a simulation using the same setup as that discussed for the empirical verification of ℓ_s. Here, the binary sparse signal is a complement to the previous one, i.e., $(s_\epsilon)_{j:(j \in \mathcal{S}_\epsilon)} = 1$ and $(s_\epsilon)_{j:(j \notin \mathcal{S}_\epsilon)} = 0$. The histogram of $(y_\epsilon)_i$ corresponding to insignificant coefficients contribution is shown in Figure 2.3 which shows that $\mathbb{E}\{(y_\epsilon)_i\} = \frac{1}{M}\sum_{i=1}^{M}(y_\epsilon)_i = \ell_e = (N-k)(1-\lambda) = (5000-200)(1-0.95) = 240$ is the average and has the highest probability.

By combining Equations (2.14) and (2.16) with that of measurement noise, each component y_i of the BSM measurement vector $\mathbf{y}_{BSM}$ is

$$
y_i \sim \mathcal{N}(\ell_s\mu_s, \ell_s(\sigma_s^2 + \lambda\mu_s^2) + \ell_\epsilon\sigma_\epsilon^2 + \sigma_\vartheta^2). \tag{2.17}
$$

Here, the statistics μ_s and σ_s^2 are unknown *a priori*. Using the thresholding factor E_s to separate the significant and insignificant coefficients, an estimate of σ_ϵ^2 based on the
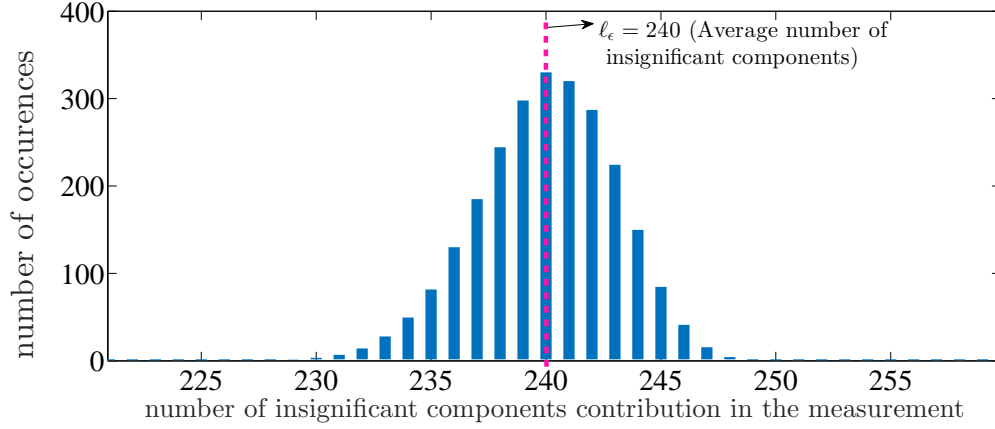
Figure 2.3: Empirical results on the number of insignificant coefficients contribution and its frequency of occurrences from the BSM-based measurements showing that each measurement contains an average number of insignificant coefficients as $\mathbb{E}\{(y_\epsilon)_i\} = \ell_\epsilon = (N - k)(1 - \lambda) = (5000 - 200)(1 - 0.95) = 240$.

concentration of measure is

$$M_{BSM}(N - k)(1 - \lambda)\sigma_\epsilon^2 \approx (1 - E_s)||\mathbf{\Phi}_{BSM}\mathbf{s}||_2^2.$$

$$\implies \sigma_\epsilon^2 \approx \frac{(1 - E_s)(\sum_{i=1}^{M_{BSM}} y_i^2 - M_{BSM}\widehat{\sigma_\vartheta^2})}{M_{BSM}(N - k)(1 - \lambda)}$$

$$\approx \frac{(1 - E_s)(\sum_{i=1}^{M_{BSM}} y_i^2 - M_{BSM}\widehat{\sigma_\vartheta^2})}{M_{BSM}N(1 - \lambda)} \quad \text{(as } N \gg k\text{)}.$$

It can be noted that an estimate $\widehat{\sigma_\vartheta^2}$ of measurement noise is available either during calibration by acquiring measurements in the absence of any signal or using Particle Swarm Optimisation (PSO) [58] method.

2.4 Statistics of GSM measurements

Each GSM measurement y_i is given by $y_i = \sum_j \phi_{i,j} s_j + \vartheta_i$. Since each $\phi_{i,j} \sim \mathcal{N}(0, \sigma_\phi^2) \in \mathbf{\Phi}_{GSM}$ is i.i.d Gaussian and independent of s_j, the mean $\mathbb{E}\{y_i\}$ and variance $\text{VAR}\{y_i\}$ of the GSM-based measurements are computed as follows.

$$
\begin{aligned}
\mathbb{E}\{y_i\} &= \mathbb{E}\{\sum_{j\in\mathcal{S}_\mathbf{s}}\phi_{i,j}(\mathbf{s}_\mathcal{S})_j\} + \mathbb{E}\{\sum_{j\in\mathcal{S}_\epsilon}\phi_{i,j}(\mathbf{s}_\epsilon)_j\} + \mathbb{E}\{\vartheta_i\} \\
&= \sum_{j\in\mathcal{S}_\mathbf{s}}\mathbb{E}\{\phi_{i,j}(\mathbf{s}_\mathcal{S})_j\} + \mathbb{E}\{\sum_{j\in\mathcal{S}_\epsilon}\phi_{i,j}(\mathbf{s}_\epsilon)_j\} \quad (\text{as } \mathbb{E}\{\vartheta_i\}=0) \\
&= \sum_{j\in\mathcal{S}_\mathbf{s}}\mathbb{E}\{\phi_{i,j}\}\mathbb{E}\{(\mathbf{s}_\mathcal{S})_j\} + \sum_{j\in\mathcal{S}_\epsilon}\mathbb{E}\{\phi_{i,j}\}\mathbb{E}\{(\mathbf{s}_\epsilon)_j\} \quad (\text{ applying the independence property}) \\
&= 0. \quad (\text{as } \mathbb{E}\{\phi_{i,j}\}=0)
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{VAR}\{y_i\} &= \sum_{j\in\mathcal{S}_\mathbf{s}}\mathrm{VAR}\{\phi_{i,j}(\mathbf{s}_\mathcal{S})_j\} + \sum_{j\in\mathcal{S}_\epsilon}\mathrm{VAR}\{\phi_{i,j}(\mathbf{s}_\epsilon)_j\} + \sigma_\vartheta^2 \\
&\qquad (\text{by applying independence property}) \\
&= \sum_{j\in\mathcal{S}_\mathbf{s}}(\mathbb{E}\{\phi_{i,j}\})^2\mathrm{VAR}\{(\mathbf{s}_\mathcal{S})_j\} + (\mathbb{E}\{(\mathbf{s}_\mathcal{S})_j\})^2\mathrm{VAR}\{\phi_{i,j}\} + \\
&\qquad \mathrm{VAR}\{(\mathbf{s}_\mathcal{S})_j\}\mathrm{VAR}\{\phi_{i,j}\} \\
&\quad + \sum_{j\in\mathcal{S}_\epsilon}(\mathbb{E}\{\phi_{i,j}\})^2\mathrm{VAR}\{(\mathbf{s}_\epsilon)_j\} + (\mathbb{E}\{(\mathbf{s}_\epsilon)_j\})^2\mathrm{VAR}\{\phi_{i,j}\} + \\
&\qquad \mathrm{VAR}\{(\mathbf{s}_\epsilon)_j\}\mathrm{VAR}\{\phi_{i,j}\} \\
&\quad + \sigma_\vartheta^2 \\
&= \sum_{j\in\mathcal{S}_\mathbf{s}}(0 + \mu_s^2\sigma_\phi^2 + \sigma_s^2\sigma_\phi^2) \\
&\quad + \sum_{j\in\mathcal{S}_\epsilon}(0 + 0 + \sigma_\epsilon^2\sigma_\phi^2) \\
&\quad + \sigma_\vartheta^2 \quad (\text{as } \mathbb{E}\{\phi_{i,j}\}=0 \text{ and } \mathbb{E}\{(\mathbf{s}_\epsilon)_j\}) \\
&= \sum_{j\in\mathcal{S}_\mathbf{s}}\sigma_\phi^2(\mu_s^2 + \sigma_s^2) + \sum_{j\in\mathcal{S}_\epsilon}\sigma_\phi^2\sigma_\epsilon^2 + \sigma_\vartheta^2 \\
&= k\sigma_\phi^2(\mu_s^2 + \sigma_s^2) + (N-k)\sigma_\phi^2\sigma_\epsilon^2 + \sigma_\vartheta^2 \\
&= k\sigma_\phi^2(\mu_s^2 + \sigma_s^2 - \sigma_\epsilon^2) + N\sigma_\phi^2\sigma_\epsilon^2 + \sigma_\vartheta^2.
\end{aligned}
$$

Thus, the mean and variance of each measurement are given as,

$$
\begin{aligned}
\mathbb{E}\{y_i\} &= 0 \quad \text{and} & (2.18) \\
\mathrm{VAR}\{y_i\} &= k\sigma_\phi^2(\mu_s^2 + \sigma_s^2 - \sigma_\epsilon^2) + N\sigma_\phi^2\sigma_\epsilon^2 + \sigma_\vartheta^2. & (2.19)
\end{aligned}
$$

2.5 Modeling time-varying sparsity order

Any real-world signal has time-varying support owing to the inclusions (births) and deletions (deaths) of the signal basis functions, as shown in Figure 2.4. The support spectrum of the vibration signal of an aircraft [59] is shown in this figure. The figure depicts a sparse representation of the vibration signal in the DCT domain. The figure captures the time evolution of the active cosine harmonics basis functions of the DCT, as shown in the different colored tracks. Here, it may be noted that there are (i) clusters of active basis functions centered on certain resonant frequencies, as shown in the highlighted portion of the plot, (ii) large spectral gaps between these clusters that reveal sparsity in the DCT domain, and (iii) inclusions and deletions of the basis functions with time, as revealed by the discontinuous lines. Thus, the birth and death of basis functions result in the variation of support and its cardinality, i.e., the sparsity order k.
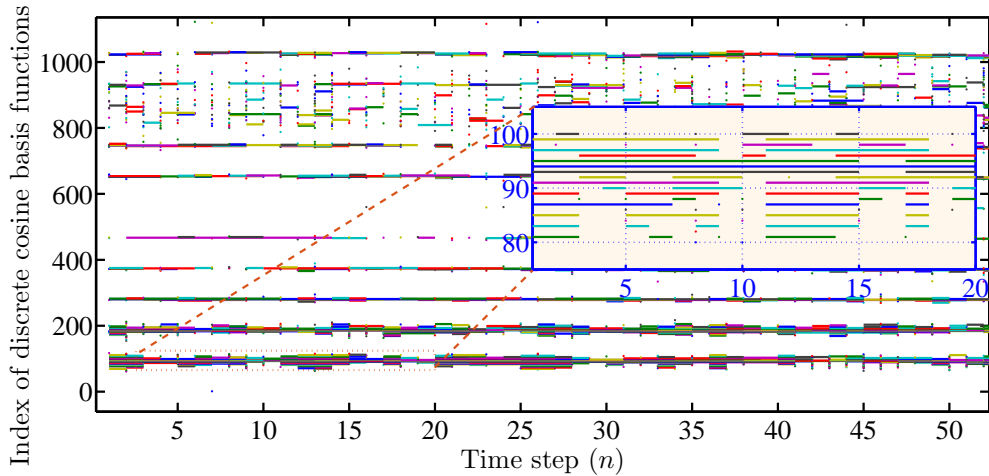
Figure 2.4: Time-varying support of a real-world vibration signal.

Let the dynamic signal $\mathbf{x}(n)$ be represented as

$$\mathbf{x}(n) = \mathbf{\Psi}\{\mathcal{K}(n)\mathbf{s}_{\mathcal{S}}(n) + \mathcal{K}^c(n)\mathbf{s}_{\epsilon}(n)\}$$

where

- $\mathcal{K}(n) = \text{diag}(\boldsymbol{\kappa}(n))$ is the support representation matrix for significant components,

- $\boldsymbol{\kappa}(n) = \{\kappa_j(n)\}_{j=1}^N$ is i.i.d $N-$dimensional vector,

- $\kappa_j(n) \sim$ Bernoulli-p, q such that $Pr[\kappa_j(n) = 1] = p$ and $Pr[\kappa_j(n) = 0] = q = 1-p$,

32

- $\mathcal{K}^c(n)$ =diag($\boldsymbol{\kappa}^c(n)$) is the support representation matrix for the insignificant components,

- $\boldsymbol{\kappa}^c(n) = \{\kappa_j^c(n)\}_{j=1}^N$ is the i.i.d $N-$ dimensional vector, and

- $\kappa_j^c(n)$ is complement to $\kappa_j(n)$ such that $Pr[\kappa_j^c(n) = 1] = q$ and $Pr[\kappa_j^c(n) = 0] = p = 1 - q$.

- There is a temporal structure imposed on the model, which will be discussed in detail below.

The variation in the support is modeled by appropriately changing the entries of the Bernoulli vector $\boldsymbol{\kappa}(n)$. The value of the difference $d = (\kappa_j(n) - \kappa_j(n\text{-}1))$ represents the birth, or death, or survival of the jth component, depending on whether it is $+1$, or -1, or 0, respectively. Owing to this difference, the sparsity order k at the nth time step is a generalized discrete Markov birth-death process and is given as

$$k(n) = k(n\text{-}1) + \underbrace{\sum_{j=1}^N (\kappa_j(n) - \kappa_j(n\text{-}1))}_{\text{sparsity order variation}}. \qquad (2.20)$$

The following parameters specify the discrete Markov birth-death process model of sparsity order variation.

- State space $\mathbb{S}$: $\mathbb{S}$ is a finite or countable set of states, that is, values that the sparsity order $k(n)$ may take on. As $1 \leq k(n) \leq N$, the state space $\mathbb{S} = \{1, 2, 3, .., N\}$.

- Probability transition rule P: This is specified by giving a transition matrix $\mathbf{P} = [p_{i,j}] : p_{i,j} = p_d = p_{j-i} = Pr\{k(n) = j | k(n\text{-}1) = i\}$. As S is the finite set, P is an $N \times N$ matrix. The interpretation of the number $p_{i,j}$ is the conditional probability, given that $k(n)$ is in state i at time step n-1 jumps to the state j at time step n.

In the probability $Pr\{k(n) = j | k(n\text{-}1) = i\}$, (i) when $i < j$, the probability indicates birth probability, (ii) when $i > j$ the probability indicates death probability, and (iii) when $i = j$, the probability indicates survival probability. In other words, the probabilities $Pr\{k(n) = j | k(n\text{-}1) = i\} = p_{i,j} = Pr[(k(n) - k(n\text{-}1)) = j - i = d] = p_d : d > 0, d < 0,$ and $d = 0$ are associated with the birth, death, and survival of supporting components. In particular, the probability that the sparsity order remains unchanged between two time steps, $p_{(d=0)} = p_0$ determines the degree of sparsity order variation. It shall be

noted that there are conditions such that though the sparsity order remains unchanged i.e., $k(n) = k(n-1)$, there can be complete change in the support. For example, suppose support is $\{1, 2\}$ at $(n-1)^{th}$ time step changes to $\{3, 4\}$ at n^{th} time step where sparsity order remains unchanged, i.e., $k(n-1) = k(n) = 2$. Such conditions are undesirable here.

The sparsity order variations can be slow, moderate, fast, or rapid. The sparse signals exhibit slow and gradual variations in the sparsity order under steady-state conditions and rapid variations during transients and bursty behavior. The idea of survival time $\widetilde{L}$ is introduced here to statistically quantify the degree and duration of invariance of the sparsity order in time akin to similar concepts in the context of modeling time-varying wireless channels in digital communications.

Definition 2.1. The survival time $\widetilde{L}$ is the statistical average number of time steps for which the sparsity order is practically time-invariant, and is given by Equation (2.21).

As the probability that the sparsity order remains unchanged for l time steps is $p_0^l(1 - p_0)$, the average time steps for which the sparsity order remains unchanged is $p_0/(1 - p_0)$. Thus, the survival time $\widetilde{L}$ is

$$\widetilde{L} = \left\lfloor \frac{p_0}{1 - p_0} \right\rfloor. \tag{2.21}$$

Intuitively, a longer survival time indicates slow variation and a shorter survival time indicates rapid variation in the sparsity order. For example, in this research work, the sparsity order variation is classified as (i) slow when $0.9 < p_0 < 1$ ($\widetilde{L} > 9$), (ii) moderate when $0.8 < p_0 \leq 0.9$ ($4 < \widetilde{L} \leq 9$), (iii) fast when $0.6 < p_0 \leq 0.8$ ($2 < \widetilde{L} \leq 5$), and (iv) rapid when $p_0 \leq 0.6$ ($\widetilde{L} \leq 2$), respectively. An example of a simulation of time-varying sparsity order $k(n)$ for four different values of p_0 is shown in Figure 2.5. It can be seen that, as p_0 decreases, the survival time $\widetilde{L}$ becomes smaller, and the rate of change of sparsity order increases, i.e., when (i) $p_0 = 0.98\{\widetilde{L} = 49\}$, (ii) $p_0 = 0.9\{\widetilde{L} = 9\}$, (iii) $p_0 = 0.75\{\widetilde{L} = 3\}$, and (iv) $p_0 = 0.5\{\widetilde{L} = 1\}$, the sparsity order variations are slow, moderate, fast, and rapid, respectively. This birth-death model is similar to the model given for time-varying sparse CSI for wireless channels in [60, 61, 62, 63].

2.5.1 Algorithm for simulating the time-varying sparse order

An algorithm used for simulating a discrete Markov birth-death process characterizing the time-varying sparsity order is given as follows.
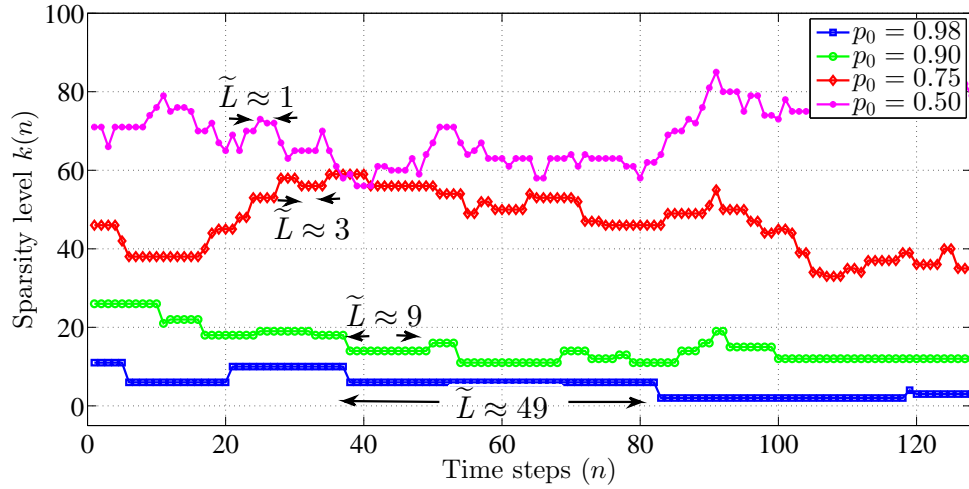
Figure 2.5: Simulation of time-varying sparsity order using Markov birth-death process.

Algorithm 3 Simulating the time-varying sparsity order

Input:

- Discrete time step $n = 0$

- Sparsity order variation probability matrix $\mathbf{P} = [p_{i,j}] : p_{i,j} = p_d = p_{j-i} = Pr\{k(n) = j | k(n-1) = i\}$

- Initial value of sparsity order $k(0)$

1: $n = n + 1$
2: $i = k(n-1)$
3: Take a sample u from an uniform distribution $\mathcal{U}(0,1)$
4: $k(n) = j$ if $\sum_{l=1}^{j} p_{i,l} < u \leq \sum_{z=1}^{j+1} p_{i,z}$
5: go to step 1
Output: Sparsity order: $k(n)$.

2.6 Summary

In this chapter, the models for (i) time-varying parameters: support, sparsity order, and amplitude of the compressible signal, (ii) sensing system, and (iii) the statistics of measurements are discussed.

Chapter 3

Maximum likelihood sparsity order estimation

In this chapter, estimators to estimate the instantaneous sparsity order on the fly are derived using the Maximum Likelihood (ML) principle when the measurements are obtained using the composite BSM and GSM. The derived estimators do not require any prior information on the compressible signal, and they perform SOE from the BSM and GSM-based measurements. The derivation of BSM-based SOE (BSOE) and GSM-based SOE (GSOE) and their properties are discussed here.

3.1 BSM-based SOE

In this section, the ML estimator for the sparsity order is derived from the statistical properties of the sparse representation $\mathbf{s}$ and the BSM-based measurement vector $\mathbf{y}_{BSM}$.

As each row of the sparse BSM has very few ones compared to the number of zeros, there exists a finite probability of obtaining measurements that have contributions only from insignificant coefficients and measurement noise. This finite probability is exploited here for estimating the sparsity order of the underlying compressible signal.

As discussed in Chapter 2, the compressible signal has a sparse representation $\mathbf{s}(n)$ and can be written as the sum of two disjoint signals, a $k(n)-$sparse signal $\mathbf{s}_{\mathcal{S}}(n)$, and an $(N - k(n))-$dense signal $\mathbf{s}_\epsilon(n)$, i.e.,

$$\mathbf{s}(n) = \mathbf{s}_{\mathcal{S}}(n) + \mathbf{s}_\epsilon(n) \tag{3.1}$$

where $\mathbf{s}_{\mathcal{S}}(n)$ is a $N-$dimensional vector containing $k(n)$ nonzero significant coefficients and $N - k(n)$ zeros, and $\mathbf{s}_\epsilon(n)$ is another $N-$dimensional vector containing $N - k(n)$

nonzero insignificant coefficients and $k(n)$ zeros. For the sake of brevity, the discrete-time notation n is omitted here after. Using the BSM, the measurement vector

$$
\begin{aligned}
\mathbf{y}_{BSM} &= \mathbf{\Phi}_{BSM}\mathbf{s}_{\mathcal{S}} + \mathbf{\Phi}_{BSM}\mathbf{s}_{\varepsilon} + \boldsymbol{\vartheta}_{BSM} \\
&= (\mathbf{y}_{BSM})_{\mathcal{S}} + (\mathbf{y}_{BSM})_{\epsilon} + \boldsymbol{\vartheta}_{BSM}.
\end{aligned}
$$

The vector $\mathbf{y}_{BSM} = \{y_i\}_{i=1}^{M_{BSM}}$ is composed of components $(\mathbf{y}_{BSM})_{\mathcal{S}} = \{(y_{\mathcal{S}})_i\}_{i=1}^{M_{BSM}}$ because of significant coefficients, $(\mathbf{y}_{BSM})_{\epsilon} = \{(y_{\epsilon})_i\}_{i=1}^{M_{BSM}}$ because of insignificant coefficients, and $\boldsymbol{\vartheta}_{BSM} = \{\vartheta_i\}_{i=1}^{M_{BSM}}$ because of the measurement noise. Thus, the ith measurement y_i is given as

$$
y_i = (y_{\mathcal{S}})_i + (y_{\epsilon})_i + \vartheta_i. \tag{3.2}
$$

An example illustrating the measurements obtained using BSM is shown below.

$$
\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}}_{\mathbf{y}_{BSM}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{\Phi}_{BSM}} \underbrace{\begin{bmatrix} (\mathbf{s}_{\mathcal{S}})_1 \\ (\mathbf{s}_{\varepsilon})_2 \\ (\mathbf{s}_{\varepsilon})_3 \\ (\mathbf{s}_{\mathcal{S}})_4 \\ (\mathbf{s}_{\mathcal{S}})_5 \\ (\mathbf{s}_{\varepsilon})_6 \\ (\mathbf{s}_{\varepsilon})_7 \\ (\mathbf{s}_{\varepsilon})_8 \end{bmatrix}}_{\text{sparse representation: } \mathbf{s}} + \underbrace{\begin{bmatrix} \vartheta_1 \\ \vartheta_2 \\ \vartheta_3 \\ \vartheta_4 \end{bmatrix}}_{\text{Measurement noise: } \boldsymbol{\vartheta}_{BSM}}
$$

$$
= \underbrace{\begin{bmatrix} (\mathbf{s}_{\mathcal{S}})_1 + (\mathbf{s}_{\mathcal{S}})_5 \\ 0 \\ 0 \\ (\mathbf{s}_{\mathcal{S}})_4 \end{bmatrix}}_{(\mathbf{y}_{BSM})_{\mathcal{S}}} + \underbrace{\begin{bmatrix} 0 \\ (\mathbf{s}_{\varepsilon})_2 + (\mathbf{s}_{\varepsilon})_6 \\ (\mathbf{s}_{\varepsilon})_3 + (\mathbf{s}_{\varepsilon})_7 \\ (\mathbf{s}_{\varepsilon})_8 \end{bmatrix}}_{(\mathbf{y}_{BSM})_{\epsilon}} + \underbrace{\begin{bmatrix} \vartheta_1 \\ \vartheta_2 \\ \vartheta_3 \\ \vartheta_4 \end{bmatrix}}_{\boldsymbol{\vartheta}_{BSM}}
$$

$$
= \begin{bmatrix} (\mathbf{s}_{\mathcal{S}})_1 + (\mathbf{s}_{\mathcal{S}})_5 + \vartheta_1 \\ (\mathbf{s}_{\varepsilon})_2 + (\mathbf{s}_{\varepsilon})_6 + \vartheta_2 \\ (\mathbf{s}_{\varepsilon})_3 + (\mathbf{s}_{\varepsilon})_7 + \vartheta_3 \\ (\mathbf{s}_{\mathcal{S}})_4 + (\mathbf{s}_{\varepsilon})_8 + \vartheta_4 \end{bmatrix}.
$$

In the above example, a compressible signal of dimension $N = 8$ has a sparse representation $\mathbf{s}$ with sparsity order $k = 3$ and support $\mathcal{S} = \{1, 4, 5\}$. The BSM $\mathbf{\Phi}_{BSM}$ of dimension $M \times N = 4 \times 8$ is constructed by horizontally stacking two identity matrices of

dimension 4×4. The measurement $\mathbf{y}_{BSM} = \{y_1, y_2, y_3, y_4\}$ has dimension $M = 4$. It can be observed that measurements y_1 and y_4 have contributions from significant coefficients, whereas measurements y_2 and y_3 are devoid of significant coefficients. Thus, each measurement is a random sum of the significant and insignificant coefficients along with the measurement noise.

Because the BSM is sparse, there exists a probability that y_i is devoid of significant coefficients, i.e., the probability that $\{(y_{\mathcal{S}})_i\} = 0$ resulting in $y_i = (y_\epsilon)_i + \vartheta_i$ can be used for the SOE. For a better understanding of the SOE using this probability, initially, true sparse signals under noiseless settings are considered, and then compressible signals with noise settings are considered.

3.1.1 Case 1: True sparse signals under noiseless settings

For true sparse signals, the insignificant coefficients are zeros. In other words, $\mathbf{s}_\epsilon = \mathbf{0}$ resulting in $\mathbf{s} = \mathbf{s}_{\mathcal{S}}$. Assuming zero measurement noise, any measurement y_i obtained using BSM is a random sum of the significant coefficients of $\mathbf{s}$. There exists a finite probability that $y_i = (\Phi_{BSM})_i \mathbf{s} = 0$, i.e., the zeros of both the ith BSM row $(\Phi_{BSM})_i$ and the sparse representation $\mathbf{s}$ mutually multiply with the nonzero entries of other vector resulting in zero-valued measurements. The probability P_0 of obtaining such zero-valued measurements can be used to estimate the sparsity order k for true sparse signals.

The following theorem derives the sparsity order estimator $\widehat{K}_{BSOE}$ for true sparse signals acquired under noiseless conditions.

Theorem 3.1. *The BSOE technique (True Sparse Signals): Consider a sparse random BSM $\Phi_{BSM} = [\phi_{i,j}] : \phi_{i,j} \sim \mathcal{B}(0, \lambda)$, which obtains measurements $\mathbf{y} = \{y_i\}_{i=1}^{M_{BSM}}$. Then, the sparsity order estimator $\widehat{K}_{BSOE}$ using the BSOE technique for any true $k-$sparse signal is given by $\widehat{K}_{BSOE} = \left\lfloor \log(\widehat{P_0})/\log(\lambda) \right\rceil$, where $\widehat{P_0} = \sum_{i=1}^{M_{BSM}} (\delta_0(y_i))/M_{BSM}$ is the estimated probability of obtaining a zero-valued measurement and $\delta_0(y_i) = \begin{cases} 1, \text{if } y_i(n) = 0; \\ 0, \text{otherwise}. \end{cases}$*

Proof. Let $\mathcal{S}_{\Phi_i}$ be the support of the ith row of a sparse BSM Φ_{BSM} and $\mathcal{S}_{\mathbf{s}}$ be the support of the sparse representation $\mathbf{s}$. If the number of elements in the support $\mathcal{S}_{\mathbf{s}}$ is k, then the probability P_ℓ that both support sets $\mathcal{S}_{\Phi_i}$ and $\mathcal{S}_{\mathbf{s}}$ to have ℓ common elements, is binomial distributed, and is given as

$$P_\ell = \binom{k}{\ell} (1 - \lambda)^\ell (\lambda)^{(k-\ell)} \tag{3.3}$$

where $0 \leq \ell \leq k$. As P_ℓ is a function of k, λ, and ℓ, an estimate of k can be obtained using the same. When the support sets $\mathcal{S}_{\Phi_i}$ and $\mathcal{S}_\mathbf{s}$ do not have a common element, i.e., $\ell = 0$, Equation (3.3) reduces to,

$$P_0 = \lambda^k \tag{3.4}$$

which provides the probability of having nonoverlapping sets $\mathcal{S}_{\Phi_i}$ and $\mathcal{S}_\mathbf{s}$, resulting in a zero-valued measurement. Using Equation (3.4), the sparsity order k can be derived as

$$k = \frac{\log(P_0)}{\log(\lambda)}. \tag{3.5}$$

Thus, a simple sparsity order estimator $\widehat{K}_{BSOE}$ is obtained by estimating the probability P_0. The probability P_0 is Maximum Likelihood (ML) estimated as the proportion of zero-valued measurements y_i in a total of M_{BSM} measurements as follows.

3.1.1.1 ML estimate of P_0

The probability for obtaining M_0 measurements devoid of significant coefficients out of M_{BSM} measurements is Binomial distributed, i.e., the likelihood function is given as

$$Pr[M_0] = \binom{M_{BSM}}{M_0} (P_0)^{M_0} (1 - P_0)^{M_{BSM} - M_0}. \tag{3.6}$$

Taking logarithm for Equation (3.6), the log-likelihood function is,

$$\log\{Pr[M_0]\} = \log\{\binom{M_{BSM}}{M_0}\} + M_0 \log\{(P_0)\} + (M_{BSM} - M_0) \log\{(1 - P_0)\} \tag{3.7}$$

The maximum likelihood of P_0 is obtained by taking first order derivative of Equation (3.7) with respect to P_0 and equating to 0, i.e.,

$$\frac{\partial \log\{Pr[M_0]\}}{\partial P_0} = \frac{M_0}{P_0} - \frac{M_{BSM} - M_0}{1 - P_0} = 0, \tag{3.8}$$

which provides

$$P_0 = \frac{M_0}{M_{BSM}}. \tag{3.9}$$

Here $M_0 = \sum_{i=1} (\delta_0(y_i))$. Thus the ML estimate of P_0 is given as

$$\widehat{P}_0 = \frac{\sum_{i=1}^{M_{BSM}} (\delta_0(y_i))}{M_{BSM}}. \tag{3.10}$$

39

Substituting $\widehat{P}_0$ into Equation (3.5) and rounding it to the nearest integer, the sparsity order estimator $\widehat{K}_{BSOE}$ is

$$\widehat{K}_{BSOE} = \left\lfloor \frac{\log(\widehat{P}_0)}{\log(\lambda)} \right\rceil = \left\lfloor \frac{\log\left(\frac{\sum_{i=1}^{M_{BSM}}(\delta_0(y_i))}{M_{BSM}}\right)}{\log(\lambda)} \right\rceil. \tag{3.11}$$

$\square$

3.1.2 Case 2: compressible signals under noise settings

The procedure for obtaining the sparsity order estimator for compressible signals is similar to the procedure discussed above for true sparse signals with certain changes in the computation of the probability P_0.

For compressible signals in the presence of additive noise, both $\mathbf{s}_\epsilon$ and $\boldsymbol{\vartheta}$ are nonzeroes. Thus the non-overlapping sets $\mathcal{S}_{\Phi_i}$ and $\mathcal{S}_s$ result in measurements without contributions from the significant coefficients, i.e., $y_i = (y_\epsilon)_i + \vartheta_i \neq 0$ and here P_0 is the probability of obtaining such a measurement devoid of contributions from the significant coefficients, which is the same as given in Equation (3.4), i.e.,

$$P_0 = Pr \underbrace{[y_i = (y_\epsilon)_i + \vartheta_i]}_{\text{devoid of significant components}} = \lambda^k. \tag{3.12}$$

As $(y_\epsilon)_i$ and ϑ_i are zero-mean Gaussian distributed, P_0 corresponds to the probability of obtaining measurements lying in a bounded interval from $-\tau$ to τ centered around the origin. When y_i is devoid of significant coefficients, i.e., when all the 1's in the ith row span only insignificant components, the variance of y_i is equal to the variance of $N(1-\lambda)$ insignificant components, i.e., $N(1-\lambda)\sigma_\epsilon^2$ added with measurement noise variance σ_ϑ^2. Hence, y_i has a Gaussian pdf whose variance is $N(1-\lambda)\sigma_\epsilon^2 + \sigma_\vartheta^2$. The 99% area of this pdf is bounded by y_i such that $-3\sqrt{N(1-\lambda)\sigma_\epsilon^2 + \sigma_\vartheta^2} \leq y_i \leq 3\sqrt{N(1-\lambda)\sigma_\epsilon^2 + \sigma_\vartheta^2}$. Hence, the bounding threshold τ is set to be, $\tau = 3\sqrt{N(1-\lambda)\sigma_\epsilon^2 + \sigma_\vartheta^2}$.

Thus, the probability P_0 is

$$P_0 = Pr\left[y_i = (y_\epsilon)_i + \vartheta_i\right] = Pr\left[|y_i| \leq \tau\right],$$

which is estimated as the proportion of measurements having magnitude less than the threshold τ, out of M_{BSM} measurements. However, there is a possibility that two or more

of the significant coefficients may also be present in the sum y_i, yet satisfying the above condition, i.e., $|y_i = (y_s)_i + (y_\epsilon)_i + \vartheta_i| \leq \tau$ and are akin to having false alarms in detection problems. Such conditions result in biased estimates of P_0 with an upward bias. Therefore, it is necessary to subtract the probability q of having two or more significant coefficients, whose sum is insignificant. Therefore,

$$Pr\left[|y_i| \leq \tau\right] = \underbrace{Pr\left[|(y_\epsilon)_i + \vartheta_i| \leq \tau\right]}_{P_0:\text{insignificant + noise}} + \underbrace{Pr\left[|(y_s)_i + (y_\epsilon)_i + \vartheta_i| \leq \tau\right]}_{q:\text{significant+insignificant+noise}}$$

where the estimate of $Pr\left[|y_i| \leq \tau\right]$ is $\frac{\sum_{i=1}^{M_{BSM}} \delta_\tau(y_i)}{M_{BSM}}$ and $\delta_\tau(y_i) = \begin{cases} 1, \text{if } |y_i(n)| \leq \tau \\ 0, \text{otherwise.} \end{cases}$

Thus,

$$\widehat{P_0} = \frac{\sum_{i=1}^{M_{BSM}} \delta_\tau(y_i)}{M_{BSM}} - \widehat{q}. \tag{3.13}$$

The BSOE technique estimates q on the fly by estimating the statistics of the significant coefficients from the measurement statistics. The estimate of q is given by Equation (3.17).

Extending Theorem 3.1, the BSOE technique for compressible signals is given by

Theorem 3.2. *The BSOE technique (Compressible Signals): The sparsity order estimator* $\widehat{K}_{BSOE}$ *for compressible signals is*

$$\widehat{K}_{BSOE} = \left\lceil \left| \frac{\log\left(\frac{\sum_{i=1}^{M_{BSM}} \delta_\tau(y_i)}{M_{BSM}} - \widehat{q}\right)}{\log(\lambda)} \right| \right\rceil. \tag{3.14}$$

The claim of the Theorem 3.2 is that the estimator $\widehat{K}_{BSOE}$ is obtained using the BSM measurements and the false alarm probability q. As the computation of false alarm probability depends on the knowledge of statistics of compressible signal, it is explained as follows.

Because each $\mathbf{y}$ is a random sum consisting of any $\ell < k$ significant coefficients, to determine q, first the probability $q_{\ell,k}$ for the presence of ℓ significant coefficients in the sum is considered. Then, the probability $q_{\ell,y}$ that such a sum is bounded by threshold τ is evaluated. Finally, combining the above probabilities and repeating the same for all values

of ℓ from 2 through $N(1 - \lambda)$ yields the desired probability q, i.e.,

$$q_{\ell,k} = \binom{k}{\ell}(1 - \lambda)^{\ell}(\lambda)^{(k-\ell)} \tag{3.15}$$

$$q_{\ell,y} = \int_{-\tau}^{\tau} p_{Y_\ell}(y_\ell) dy_\ell \tag{3.16}$$

$$q = \sum_{\ell=2}^{N(1-\lambda)} q_{\ell,k} q_{\ell,y} \tag{3.17}$$

where $p_{Y_\ell}(y_\ell)$ is the probability density function (pdf) of a measurement, given by

$$p_{Y_\ell}(y_\ell) \sim \mathcal{N}(\ell\mu_s, \ell\sigma_s^2 + (N(1 - \lambda) - \ell)\sigma_\epsilon^2 + \sigma_\vartheta^2). \tag{3.18}$$

Probability $q_{\ell,y}$ depends on the statistics μ_s and σ_s^2 which are functions of k. Thus, both $q_{\ell,k}$ and $q_{\ell,y}$ are functions of k whose estimation is the objective of this research work. That is, to estimate k, an estimate of q is required, but q depends on k. To address this yet another chicken and egg problem, $q_{\ell,k}$ can be precomputed for all values of ℓ and k from 0 to k_{max}. In the case of $q_{\ell,y}$, the evaluation of the integrals of the type given in Equation (3.16) can be performed using the Mclaurin expansion for the same and computing them to the desired level of accuracy.

Considering the pdf given in Equation (3.18), the problem of estimating q turns out to be an optimization problem such that, given the measurement vector $\mathbf{y}$, what would be the optimum pair of $\{k, \mu_s, \sigma_s^2\}$ which yields the probability P_0. Hence, a method for the joint estimation of the optimum pair from the obtained measurements is discussed as follows.

3.1.3 Joint estimation of statistics of significant coefficients and sparsity order

The estimates of μ_s and σ_s^2 of the significant coefficients in Equation (2.17) are given below.

$$\widehat{\mu_s} = \frac{\widehat{\mu_y}}{\widehat{k}(1 - \lambda)}$$

$$\widehat{\sigma_s^2} = \frac{\widehat{\sigma_y^2} - \left((N - \widehat{k})(1 - \lambda)\widehat{\sigma_\epsilon^2} + \widehat{\sigma_\vartheta^2}\right)}{\widehat{k}(1 - \lambda)} + \lambda\widehat{\mu_s^2}$$

where $\widehat{k} = \widehat{k}_{BSOE}$ is the estimated sparsity order obtained from the BSOE estimator, and

$$
\begin{aligned}
\widehat{\mu_y} &= \frac{\sum_{i=1}^{M_{BSM}} y_i}{M_{BSM}} \\
\widehat{\sigma_y^2} &= \frac{\sum_{i=1}^{M_{BSM}} (y_i - \widehat{\mu_y})^2}{M_{BSM}}
\end{aligned}
$$

are the sample mean and variance of the measurements, respectively.

A recursive feedback technique is proposed for the joint estimation of sparsity order and statistics of the significant coefficients from the obtained measurements. To start with, if one considers $\widehat{q} = 0$ in Equation (3.14), then $\widehat{k} < k$ resulting in $\widehat{\mu}_s > \mu_s$ as well as $\widehat{\sigma}_s^2 > \sigma_s^2$. Subsequently, $\widehat{q} < q$ and substituting it in Equation (3.14) decreases the estimate of P_0, leading to an increase in the estimate $\widehat{k}$ and a reduction in the estimate of statistics in the next iteration. This process continues until the false alarm probability estimate $\widehat{q}$ converges. Thus, in every iteration, the estimate $\widehat{k}$ is updated by approaching the true value. The BSOE procedure is summarized and given in Algorithm 4.

3.1.4 Experimental evaluation of convergence of recursive feedback technique

A simulation was performed to show the convergence of the estimates of the statistics $\widehat{\mu}_s$ and $\widehat{\sigma}_s^2$ and the sparsity order $\widehat{k}$. In the simulation, the sparsity order was kept constant at $k = 200$ and five realizations of a compressible signal of dimension $N = 2500$ were generated such that the average mean and variance of the significant coefficients were $\mu_s = 1$ and $\sigma_s^2 = 25$, respectively, and the mean and variance of each insignificant coefficient were $\mu_\epsilon = 0$ and $\sigma_\epsilon^2 = 0.2$, respectively. For each realization, different support was chosen. The sparsity order of BSM is $N(1-\lambda) = 20$. For each realization, $M = 130$ measurements were obtained. The measurement noise $\boldsymbol{\vartheta}_{BSM}$ was chosen such that the Signal to Noise Ratio (SNR) was 10 dB. The SNR is given as,

$$
SNR = 10 \log_{10} \frac{\|\boldsymbol{\Phi}_{BSM} \mathbf{s}\|_2^2}{\|\boldsymbol{\vartheta}_{BSM}\|_2^2}. \tag{3.19}
$$

The convergence of the mean, variance, and sparsity order estimates are illustrated in Figures 3.1 (a), 3.1 (b), and 3.1 (c), respectively. It is observed that the initial sparsity order estimate is less than the true value, and the initial estimate of the mean and variance of the significant coefficients is greater than the true value in all five realizations. In addi-

Algorithm 4 Estimation of sparsity order from BSM-based measurements

Input: Measurements $\mathbf{y} : \{y_1, y_2, .., y_{M_{BSM}}\}$, user-defined threshold E_s, measurement noise variance estimate $\widehat{\sigma_\vartheta^2}$.

1. Estimate the variance of insignificant coefficients:

- $\widehat{\sigma_\epsilon^2} = (1 - E_s) \sum_{i=1}^{M_{BSM}} \frac{y_i^2 - M_{BSM}\widehat{\sigma_\vartheta^2}}{M_{BSM} N(1-\lambda)}$, Threshold $\tau = 3\sqrt{N(1-\lambda)\sigma_\epsilon^2 + \sigma_\vartheta^2}$.

2. Estimate the sample mean and variance of the measurements:

$$\widehat{\mu_y} = \frac{\sum_{i=1}^{M_{BSM}} y_i}{M_{BSM}}; \quad \widehat{\sigma_y^2} = \frac{\sum_{i=1}^{M_{BSM}} (y_i - \widehat{\mu_y})^2}{M_{BSM} - 1}.$$

3. Estimate the probability P_0 by calculating the number of measurements such that $|y_i| \le \tau$ i.e., $\widehat{P_0} = \frac{\sum_{i=1}^{M_{BSM}} \delta_\tau(y_i)}{M_{BSM}}$ where $\delta_\tau(y_i) = \begin{cases} 1, \text{if } |y_i| \le \tau \\ 0, \text{otherwise.} \end{cases}$

4. Compute the initial estimate of sparsity order using BSM: $\widehat{K}_{BSOE} = \widehat{k} = \left\lfloor \log(\widehat{P_0})/\log(\lambda) \right\rceil$.

5. Iterative computation:

   5.1 Estimate the statistics of significant coefficients:

   $$\widehat{\mu_s} = \frac{\widehat{\mu_y}}{\widehat{k}(1 - \lambda)}; \quad \widehat{\sigma_s^2} = \frac{\widehat{\sigma_y^2} - \left((N - \widehat{k})(1 - \lambda)\widehat{\sigma_\epsilon^2} + \widehat{\sigma_\vartheta^2}\right)}{\widehat{k}(1 - \lambda)} + \lambda\widehat{\mu_s^2}.$$

   5.2 Estimate the probability q:

   - $\widehat{q} = \sum_{\ell=2}^{N(1-\lambda)} q_{\ell,k} q_{\ell,y}$ where $q_{\ell,k} = \binom{\widehat{k}(n)}{\ell}(1 - \lambda)^\ell(\lambda)^{(\widehat{k}-\ell)}$ and $q_{\ell,y} = \int_{-\tau}^{\tau} p_Y(y_i) dy_i$.
   - The probability $q_{\ell,k}$ can be precomputed and made available in a lookup table
   - The computation of $q_{\ell,y}$ can be done using the Mclaurin expansion in real time.

   5.3 Compute $\widehat{P_0} = \frac{\sum_{i=1}^{M_{BSM}} \delta_\tau(y_i)}{M_{BSM}} - \widehat{q}$.

   5.4 Update the sparsity order estimate, $\widehat{k}_{BSOE} = \widehat{k} = \left\lfloor \frac{\log(\widehat{P_0})}{\log(\lambda)} \right\rceil$;

   5.5 Go to 5.1 until $\widehat{k}_{BSOE}$ converges.

   End iteration.

Output: BSM-based estimate of sparsity order: $\widehat{k}_{BSOE}$.

tion, the convergence of the recursive feedback technique occurred after a maximum of five iterations.

3.1.5 Estimation performance of BSOE

A simulation was run to assess the BSOE's performance in terms of estimation accuracy when estimating the statistics of the significant coefficients and the sparsity order. The following are the estimation accuracy measures:

- Relative Mean Estimation Error (RMEE) = $\frac{|\mu_s - \widehat{\mu_s}|}{|\mu_s|}$

- Relative Variance Estimation Error (RVEE) = $\frac{|\sigma_s^2 - \widehat{\sigma_s^2}|}{\sigma_s^2}$

- Sparsity Order Estimation Error (SOEE) = $\frac{|k - \widehat{k}_{BSOE}|}{k}$

The true mean, variance, and sparsity order of the compressible signal dimension $N = 2500$ were set to $\mu_s = 1$, $\sigma_s^2 = 25$, and $k = 200$. The SNR was varied in 2 dB steps from 0 to 16 dB. For each SNR value, there were 500 realizations. In each realization, different support was chosen, and the parameters μ_s, σ_s^2, and k were estimated. The RMEE, RVEE, and SOEE were computed for each realization, and the averaged results for 500 realizations are shown in Figure 3.2. It is observed that the relative error is less than 15% for SNR values above 5 dB, and the performance of BSOE improves with increasing SNR values.

3.1.6 Properties of the estimator $\widehat{k}_{BSOE}$

Using Equation (3.9),

$$
\begin{aligned}
\mathbb{E}\{\widehat{P}_0\} = \mu_{\widehat{P}_0} &= \frac{\mathbb{E}\{M_0\}}{M_{BSM}} \\
&= \frac{M_{BSM} P_0}{M_{BSM}} \quad (\text{since } \mathbb{E}\{\widehat{M}_0\} = M_{BSM} P_0) \\
&= P_0.
\end{aligned}
$$

Thus the ML estimate $\widehat{P}_0$ is an unbiased estimate. Now applying functional invariance property, the transformation to $\left\lfloor \log(\widehat{P}_0)/\log(\lambda) \right\rfloor = \widehat{k}$ is also an ML estimate. As M_{BSM} approaches infinity, the estimate $\widehat{k} = k$ and the estimator $\widehat{K}_{BSOE}$ is asymptotically consistent.
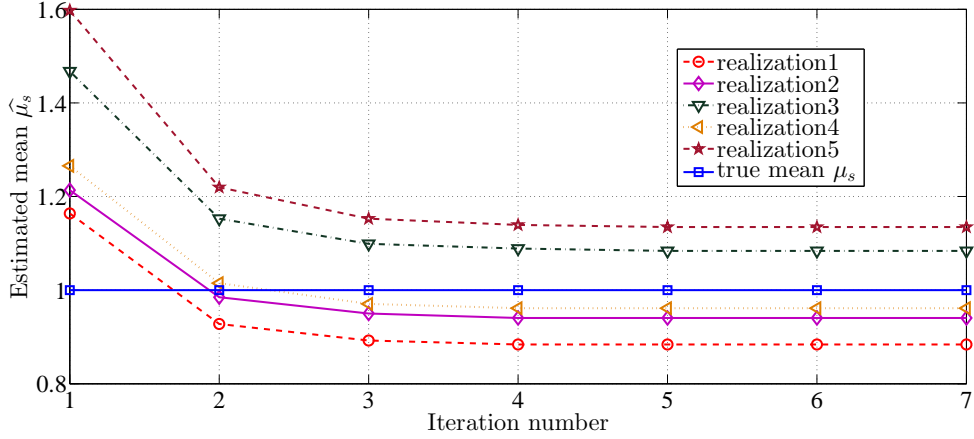
(a) Convergence of estimate of the mean μ_s
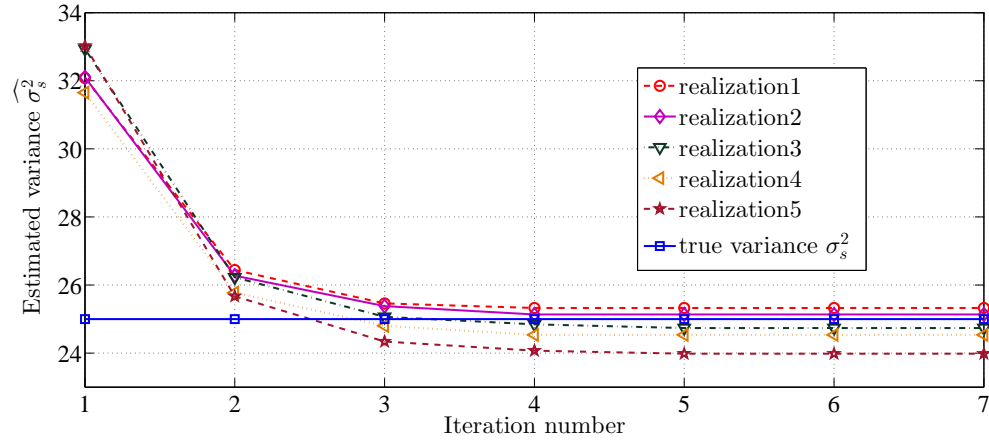
(b) Convergence of estimate of the variance σ_s^2
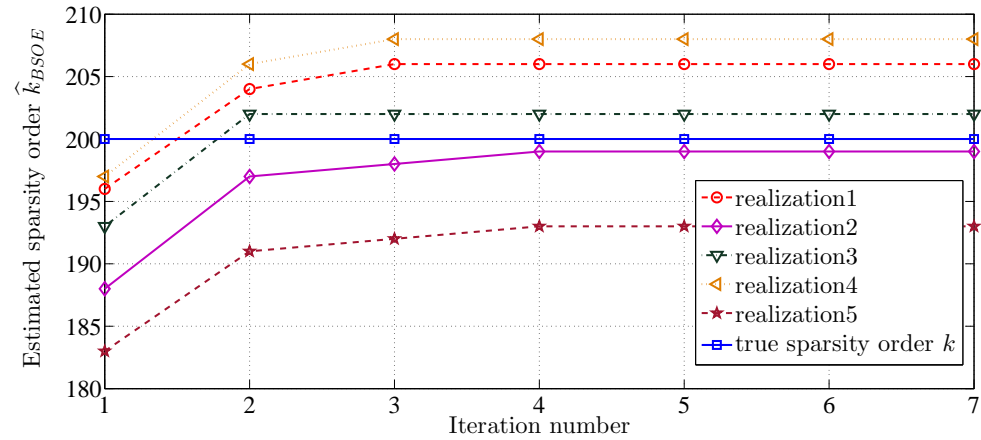
(c) Convergence of estimate of the sparsity order $\widehat{k}$

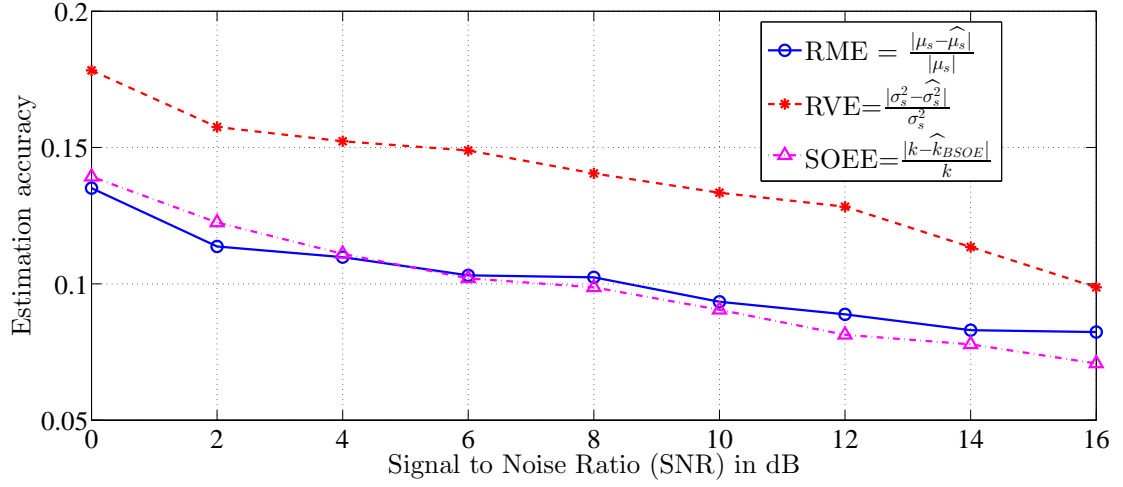**Figure 3.1:** Convergence of recursive feedback technique of BSOE.

Figure 3.2: Estimation performance of BSOE.

Using Taylor series approximation,

$$\text{VAR}\{\log(\widehat{P}_0)\} \approx \frac{\sigma_{\widehat{P}_0}^2}{\mu_{\widehat{P}_0}^2}{}^1, \tag{3.20}$$

$$\text{VAR}\{\widehat{K}_{BSOE}\} \approx \frac{1}{(\log(\lambda))^2} \frac{\sigma_{\widehat{P}_0}^2}{\mu_{\widehat{P}_0}^2}. \tag{3.21}$$

The mean $\mu_{\widehat{P}_0}$ and variance $\sigma_{\widehat{P}_0}^2$ of $\widehat{P}_0$ are computed using Equation (3.6) and are given by,

$$\mu_{\widehat{P}_0} = \lambda^k, \quad \sigma_{\widehat{P}_0}^2 = \frac{\lambda^k(1 - \lambda^k)}{M_{BSM}}. \tag{3.22}$$

Substituting $\mu_{\widehat{P}_0}$ and $\sigma_{\widehat{P}_0}^2$ in Equation (3.21), the variance of $\widehat{K}_{BSOE}$ is,

$$\text{VAR}\{\widehat{K}_{BSOE}\} \approx \frac{1 - \lambda^k}{M_{BSM}\lambda^k(\log(\lambda))^2}. \tag{3.23}$$

---

[1]For some continuous and differentiable function $g(\widehat{P}_0)$, Taylor series approximation to the variance of $g(\widehat{P}_0)$ is, $\text{VAR}\{g(\widehat{P}_0)\} \approx \left(\frac{\partial g}{\partial \widehat{P}_0}(\mu_{\widehat{P}_0})\right)^2 \sigma_{\widehat{P}_0}^2$ where $\mu_{\widehat{P}_0}$ and $\sigma_{\widehat{P}_0}^2$ are the mean and variance of $\widehat{P}_0$, respectively.

3.2 Optimum value of λ for minimizing the variance of sparsity order estimator

Parameter λ defines the sparsity order of BSM. This is crucial because it determines the SOE and recovery performances. From Equation (3.23), it is observed that $\mathrm{VAR}\{\widehat{K}_{BSOE}\}$ is a function of λ and the optimum value of λ to minimize $\mathrm{VAR}\{\widehat{K}_{BSOE}\}$ is computed as follows:

$$\frac{\partial \mathrm{VAR}\{\widehat{K}_{BSOE}\}}{\partial \lambda} = \frac{\lambda^k - \frac{k}{2}(\log(\lambda)) - 1}{\frac{M_{BSM}}{2}\lambda^{k+1}(\log(\lambda))^3} = 0$$

which implies

$$\lambda^k - \frac{k}{2}(\log(\lambda)) - 1 = 0. \tag{3.24}$$

Upon simplification, using the fact that $\lambda \approx 1$, Equation (3.24) becomes

$$\lambda \approx \exp(-1.6/k). \tag{3.25}$$

From Equation (3.25), it is observed that the optimum value of λ differs for different $k(n)$. Table 3.1 provides the optimum value of λ and the minimum variance value of $\widehat{K}_{BSOE}$ for different k.

k	λ	$\mathrm{VAR}\{\widehat{K}_{BSOE}\}$
100	0.984	38.607
150	0.989	57.962
200	0.992	77.210
250	0.994	96.693
300	0.995	116.036
350	0.995	135.887
400	0.996	154.417

Table 3.1: Optimum value of λ in minimizing the variance of the sparsity order estimate $\widehat{K}_{BSOE}$. It is observed that as the sparsity order k increases, the BSM's optimal sparsity parameter λ increases making the BSM too sparse to reduce the variance of SOE.

It is observed that λ must be increased as k increases, i.e., the BSM must be made very sparse to achieve a minimum variance of $\widehat{K}_{BSOE}$. However, as the parameter λ increases, the recovery performance of the BSM degrades as the number of measurements spanning the significant coefficients decreases. Observe from Equation (2.12) that each measurement $y_i(n)$ is a sum of the random sums of the significant and insignificant coefficients, and

noise. As λ approaches one, the number of ones in each row is reduced to zero, resulting in more measurements with contributions from insignificant coefficients and noise alone with a high probability, satisfying $|y_i(n)| \leq 3\sqrt{N(1-\lambda)\sigma_\epsilon^2 + \sigma_\vartheta^2}$ which is a condition required for a better SOE. However, at the same time, such measurements are not helpful for recovery as they do not have any contributions from significant coefficients. The remaining measurements with contributions from significant coefficients are helpful only for recovery. Thus, there is a trade-off between SOE and recovery, i.e., the sparser the BSM, the better the SOE; however, the poorer the recovery, as the BSM is poor in satisfying the RIP. Hence, with a limited number of BSM measurements, recovery of compressible signals is not possible. Hence, composite sensing matrices composed of GSM are used along with BSM for better SOE and recovery. The SOE technique using GSM is discussed in the next section.

3.3 GSM-based SOE

From the statistical properties of sparse representation and GSM-based measurement vector, the following theorem defines the GSM-based estimator, which is an ML estimator for the sparsity order.

Theorem 3.3. *The GSOE technique: Let $\mathbf{y} = \mathbf{\Phi}_{GSM}\mathbf{s} + \vartheta$ be the GSM-based CS acquisition model that obtains M_{GSM} measurements, where*

- $\mathbf{\Phi}_{GSM} = \{\phi_{i,j}\}_{i=1}^{M_{GSM}}{}_{,j=1}^{N}$ *is the CS acquisition matrix containing i.i.d entries such that $\phi_{i,j} \sim \mathcal{N}(0, \sigma_\phi^2)$,*

- $\mathbf{s} = \{s_j\}_{j=1}^{N}$ *is the $k-$sparse representation of the compressible signal such that the insignificant coefficients are approximated as zero mean i.i.d. Gaussian noise with variance σ_ε^2) and the mean and variance of the significant coefficients are given as μ_s and σ_s^2, respectively, and*

- $\vartheta = \{\vartheta_i\}_{i=1}^{M_{GSM}}$ *is the measurement noise containing M_{GSM} i.i.d entries such that $\vartheta_i \sim \mathcal{N}(0, \sigma_\vartheta^2)$.*

Then, a real-valued ML estimator $\widehat{K}_r$ of sparsity order k that maximizes the joint pdf of measurements parameterized by k and provides the estimate $\widehat{k}_r$ is given by

$$\widehat{K}_r = \frac{\frac{1}{M_{GSM}}\sum_{i=1}^{M_{GSM}} y_i^2 - (N\sigma_\phi^2\sigma_\epsilon^2 + \sigma_\vartheta^2)}{\sigma_\phi^2(\mu_s^2 + \sigma_s^2 - \sigma_\epsilon^2)}. \tag{3.26}$$

An integer-valued ML estimate $\widehat{k}_{GSOE}$ *for the sparsity order* k *is one of the two integer-valued estimates* $\lfloor \widehat{k}_r \rfloor$ *and* $\lceil \widehat{k}_r \rceil$ *which maximizes the pdf as follows.*

$$
p_Y(y;k) = \frac{1}{(2\pi(k\sigma_\phi^2(\mu_s^2+\sigma_s^2-\sigma_\epsilon^2)+N\sigma_\phi^2\sigma_\epsilon^2+\sigma_\vartheta^2))^{\frac{M_{GSM}}{2}}} \times
$$
$$
\exp\left[\frac{-1}{2(k\sigma_\phi^2(\mu_s^2+\sigma_s^2-\sigma_\epsilon^2)+N\sigma_\phi^2\sigma_\epsilon^2+\sigma_\vartheta^2)}\sum_{i=1}^{M_{GSM}}y_i^2\right],
$$

i.e.,

$$
\widehat{k}_{GSOE} = \underset{k\in\{\lfloor \widehat{k}_r \rfloor,\lceil \widehat{k}_r \rceil\}}{\arg\max} \; p_Y(y;k). \tag{3.27}
$$

Proof. Using Equations (2.18) and (2.19) given in Chapter 2, the mean and variance of each GSM measurement are given as,

$$
\mathbb{E}\{y_i\} = 0 \quad \text{and}
$$
$$
\mathrm{VAR}\{y_i\} = k\sigma_\phi^2(\mu_s^2+\sigma_s^2-\sigma_\epsilon^2)+N\sigma_\phi^2\sigma_\epsilon^2+\sigma_\vartheta^2.
$$

Now using central limit theorem and independence property, for any random variable s_j, i.e., for s_j need not be a Gaussian, the central limit theorem makes the pdf of measurements to be Gaussian, i.e., $y_i = (\sum_{j=1}^{N}\phi_{i,j}s_j+\vartheta_i) \sim \mathcal{N}(0, k\sigma_\phi^2(\mu_s^2+\sigma_s^2-\sigma_\epsilon^2)+N\sigma_\phi^2\sigma_\epsilon^2+\sigma_\vartheta^2)$. Thus, the joint pdf of M_{GSM} random variables $\{y_i\}_{i=1}^{M_{GSM}}$ is given as

$$
p_Y(y;k) = \frac{1}{(2\pi(k\sigma_\phi^2(\mu_s^2+\sigma_s^2-\sigma_\epsilon^2)+N\sigma_\phi^2\sigma_\epsilon^2+\sigma_\vartheta^2))^{\frac{M_{GSM}}{2}}} \times
$$
$$
\exp\left[\frac{-1}{2(k\sigma_\phi^2(\mu_s^2+\sigma_s^2-\sigma_\epsilon^2)+N\sigma_\phi^2\sigma_\epsilon^2+\sigma_\vartheta^2)}\sum_{i=1}^{M_{GSM}}y_i^2\right].
$$

Let $a = \sigma_\phi^2(\mu_s^2+\sigma_s^2-\sigma_\epsilon^2)$ and $b = N\sigma_\phi^2\sigma_\epsilon^2+\sigma_\vartheta^2$. Then, the joint probability distribution of M_{GSM} random variables, $\{y_i\}_{i=1}^{M_{GSM}}$ parameterized by k is given as

$$
p_Y(y;k) = \frac{1}{(2\pi(ak+b))^{\frac{M_{GSM}}{2}}}\exp\left[\frac{-1}{2(ak+b)}\sum_{i=1}^{M_{GSM}}y_i^2\right]. \tag{3.28}
$$

Equation (3.28) is the likelihood objective function for the GSM measurements. Taking logarithm, the log-likelihood function is,

$$
\log\{p_Y(y;k)\} = \log\{(2\pi(ak+b))^{\frac{-M_{GSM}}{2}}\} - \left[\frac{1}{2(ak+b)}\sum_{j=1}^{M_{GSM}}y_j^2\right] \tag{3.29}
$$
$$
= -\frac{M_{GSM}}{2}\log\{2\pi\} - \frac{M_{GSM}}{2}\log\{ak+b\} - \left[\frac{1}{2(ak+b)}\sum_{j=1}^{M_{GSM}}y_j^2\right] \tag{3.30}
$$

Since ML estimation is obtained by maximizing the Equation (3.30), the objective function becomes,

$$\log\{p_Y(y;k)\} = -\frac{M_{GSM}}{2}\log\{(ak+b)\} - \left[\frac{1}{2(ak+b)}\sum_{j=1}^{M_{GSM}} y_j^2\right] \quad (3.31)$$

as $-\frac{M_{GSM}}{2}\log\{2\pi\}$ is a constant. Setting the first order partial derivative of $\log(p_Y(y;k))$ with respect to k to 0 gives the real-valued ML estimator,

$$\begin{aligned}
\widehat{K}_r &= \frac{\frac{1}{M_{GSM}}\sum_{i=1}^{M_{GSM}} y_i^2 - b}{a} \\
&= \frac{\frac{1}{M_{GSM}}\sum_{i=1}^{M_{GSM}} y_i^2 - (N\sigma_\phi^2\sigma_\epsilon^2 + \sigma_\vartheta^2)}{\sigma_\phi^2(\mu_s^2 + \sigma_s^2 - \sigma_\epsilon^2)}.
\end{aligned}$$

Because sparsity order k is an integer, the integer-valued estimator $\widehat{K}_{GSOE}$ is obtained by finding an integer that maximizes the pdf $p_Y(y;k)$. Hence, the integer-valued estimate $\widehat{k}_{GSOE}$ can be obtained from the real-valued ML estimate $\widehat{k}_r$ and it is one among the two integer values $\lfloor\widehat{k}_r\rfloor$ and $\lceil\widehat{k}_r\rceil$ where $\lfloor\widehat{k}_r\rfloor = \max_{k\in\mathbb{Z}} k \leq \widehat{k}_r$ and $\lceil\widehat{k}_r\rceil = \min_{k\in\mathbb{Z}} k \geq \widehat{k}_r$ such that

$$\widehat{k}_{GSOE} = \arg\max_{k\in\{\lfloor\widehat{k}_r\rfloor,\lceil\widehat{k}_r\rceil\}} p_Y(y;k).$$

$\square$

3.3.1 Properties of $\widehat{K}_{GSOE}$ Estimator

1. The $\widehat{K}_{GSOE}$ estimator is an unbiased estimator.

2. The variance of the $\widehat{K}_{GSOE}$ estimator is,

$$\text{VAR}\{\widehat{k}_{GSOE}\} = \frac{2}{M_{GSM}}\frac{(\sigma_\phi^2(\mu_s^2 + \sigma_s^2 - \sigma_\epsilon^2)k + N\sigma_\phi^2\sigma_\epsilon^2 + \sigma_\vartheta^2)^2}{(\sigma_\phi^2(\mu_s^2 + \sigma_s^2 - \sigma_\epsilon^2))} + \frac{1}{12}. \quad (3.32)$$

3. The Cramer Rao Lower Bound (CRLB) for the integer parameter k is given by,

$$\text{CRLB}(\widehat{k}) = \frac{2}{M_{GSM}}\frac{(\sigma_\phi^2(\mu_s^2 + \sigma_s^2 - \sigma_\epsilon^2)k + N\sigma_\phi^2\sigma_\epsilon^2 + \sigma_\vartheta^2)^2}{(\sigma_\phi^2(\mu_s^2 + \sigma_s^2 - \sigma_\epsilon^2))}. \quad (3.33)$$

3.3.2 Proof for the properties of GSOE

The unbiasedness of GSOE is shown by proving $\mathbb{E}\{\widehat{k}_{GSOE}\} = k$. It is also shown that the variance of GSOE is closer and approximately equal to CRLB.

3.3.2.1 Proof for unbiasedness

Since $\widehat{k}_{GSOE}$ is the round-off result of $\widehat{k}_r$, it can be written as,

$$\widehat{k}_{GSOE} = \widehat{k}_r + k_\delta \tag{3.34}$$

where $k_\delta \sim \mathbb{U}\{-0.5, 0.5\}$ is the uniformly distributed round-off error.

From Equation (3.34)

$$
\begin{aligned}
\mathbb{E}\{\widehat{k}_{GSOE}\} &= \mathbb{E}\{\widehat{k}_r\} + \mathbb{E}\{k_\delta\} \\
&= \mathbb{E}\{\widehat{k}_r\} \quad \{\text{since } \mathbb{E}\{k_\delta\} = 0.\} \\
&= \frac{\frac{1}{M_{GSM}} \sum_{i=1}^{M_{GSM}} \mathbb{E}\{y_i^2\} - b}{a} \\
&= \frac{\frac{M_{GSM}(ak+b)}{M_{GSM}} - b}{a} \\
&= k.
\end{aligned}
$$

Since, $\mathbb{E}\{\widehat{k}_{GSOE}\} = k$, the estimator $\widehat{k}_{GSOE}$ is an unbiased estimator.

3.3.2.2 Variance of GSOE

The variance of $\widehat{k}_{GSOE}$ is derived from the variance of $\widehat{k}_r$ and is given by,

$$\text{VAR}\{\widehat{k}_{GSOE}\} = \quad \text{VAR}\{\widehat{k}_r\} + \text{VAR}\{k_\delta\} - 2\text{COVAR}\{\widehat{k}_r, k_\delta\} \tag{3.35}$$

$$= \quad \text{VAR}\{\widehat{k}_r\} + \frac{1}{12} - 2\text{COVAR}\{\widehat{k}_r, k_\delta\} \tag{3.36}$$

$$= \quad \text{VAR}\{\widehat{k}_r\} + \frac{1}{12}\{\text{since COVAR}\{\widehat{k}_r, k_\delta\} \approx 0\}. \tag{3.37}$$

The VAR$\{\widehat{k}_r\}$ is computed as follows.

$$
\begin{aligned}
\text{VAR}\{\widehat{k}_r\} &= \text{VAR}\left\{ \frac{\frac{1}{M_{GSM}}\sum_{i=1}^{M_{GSM}} y_i^2 - b}{a} \right\} \\
&= \frac{1}{(aM_{GSM})^2}\sum_{i=1}^{M_{GSM}} \text{VAR}\{y_i^2\} \\
&= \frac{1}{(aM_{GSM})^2}\sum_{i=1}^{M_{GSM}} (\mathbb{E}\{y_i^4\} - (\mathbb{E}\{y_i^2\})^2) \\
&= \frac{1}{(aM_{GSM})^2}\sum_{i=1}^{M_{GSM}} (3(ak+b)^2 - (ak+b)^2) \\
&= \frac{1}{(aM_{GSM})^2}2M_{GSM}(ak+b)^2 \\
&= \frac{2}{M_{GSM}}\frac{(ak+b)^2}{a^2} \\
&= \frac{2}{M_{GSM}}\frac{(\sigma_\phi^2(\mu_s^2+\sigma_s^2-\sigma_\epsilon^2)k + N\sigma_\phi^2\sigma_\epsilon^2 + \sigma_\vartheta^2)^2}{(\sigma_\phi^2(\mu_s^2+\sigma_s^2-\sigma_\epsilon^2))^2}.
\end{aligned}
$$

Thus,

$$
\text{VAR}\{\widehat{k}_{GSOE}\} = \frac{2}{M_{GSM}}\frac{(\sigma_\phi^2(\mu_s^2+\sigma_s^2-\sigma_\epsilon^2)k + N\sigma_\phi^2\sigma_\epsilon^2 + \sigma_\vartheta^2)^2}{(\sigma_\phi^2(\mu_s^2+\sigma_s^2-\sigma_\epsilon^2))^2} + \frac{1}{12}.
$$

3.3.2.3 CRLB for the SOE

For any integer parameter, the CRLB can be derived by computing either the second forward or second backward difference as given in [64]. Thus for the integer parameter k, the CRLB can be derived by computing the second forward difference for the $\ln(p_Y(y;k))$ of likelihood given in Equation (3.30).

The second forward difference of log-likelihood $\ln(p_Y(y; k))$ is given by,

$$\ln(p_Y(y; k+2)) - 2\ln(p_Y(y; k+1)) + \ln(p_Y(y; k))$$

$$=\frac{-M_{GSM}}{2}\ln\left(\frac{(ak+b)(a(k+2)+b)}{(a(k+1)+b)^2}\right)$$

$$-\frac{1}{2}\sum_{i=1}^{M_{GSM}}y_i^2\left(\frac{1}{ak+b}+\frac{-2}{a(k+1)+b}+\frac{1}{a(k+2)+b}\right)$$

$$\approx\frac{-M_{GSM}}{2}\left(\frac{-a^2}{(a(k+1)+b)^2}\right)$$

$$-\frac{1}{2}\sum_{i=1}^{M_{GSM}}y_i^2\left(\frac{2a^2}{(ak+b)(a(k+1)+b)(a(k+2)+b)}\right).$$

Using the second forward difference of log-likelihood,

$$\frac{-1}{\mathbb{E}\{\ln(p_Y(y; k+2)) - 2\ln(p_Y(y; k+1)) + \ln(p_Y(y; k))\}}$$

$$= -\left(\frac{-M_{GSM}}{2}\left(\frac{-a^2}{(a(k+1)+b)^2}\right)+\right.$$

$$\left.\frac{1}{2}\sum_{i=1}^{M_{GSM}}\mathbb{E}\{y_i^2\}\left(\frac{2a^2}{(ak+b)(a(k+1)+b)(a(k+2)+b)}\right)\right)^{-1}$$

$$\approx\frac{2}{M_{GSM}a^2}(ak+b)^2; (k \approx k+1 \approx k+2 \text{ when } k \text{ is large})$$

$$=\frac{2}{M_{GSM}}\frac{(\sigma_\phi^2(\mu_s^2+\sigma_s^2-\sigma_\epsilon^2)k+N\sigma_\phi^2\sigma_\epsilon^2+\sigma_\vartheta^2)^2}{(\sigma_\phi^2(\mu_s^2+\sigma_s^2-\sigma_\epsilon^2))}.$$

3.4 Estimation performance of GSOE

From Equation (3.26), it can be observed that GSOE requires knowledge of the statistics of significant coefficients. Because composite sensing was used, the estimates of the statistics of the significant coefficients were obtained using BSOE and substituted into Equation (3.26) to obtain the GSOE-based estimate. Hence, this section presents the estimation performance of GSOE using simulation with the help of the composite sensing matrices GSM and BSM.

Here, the simulation setup is similar to that used for evaluating the performance of BSOE, i.e., a compressible signal of dimension $N = 2500$ with sparsity order $k = 200$ was

generated with $\mu_s = 1, \sigma_s^2 = 25$ and $\sigma_\epsilon^2 = 0.2$. A total of 800 GSM-based measurements were obtained. The measurement noise was added such that the SNR values were 0 dB, 5 dB, and 10 dB. Here, the SNR is computed as $10 \log_{10} \frac{\|\mathbf{\Phi}_{GSM}\mathbf{s}\|_2^2}{\|\boldsymbol{\vartheta}_{GSM}\|_2^2}$. The GSM is fixed and not changed for different SNR values in the simulation setup. Initially, the statistics were estimated using $M_{BSM} = 130$ measurements obtained from the BSM. Then, for every GSM measurement, GSOE was performed to evaluate the efficiency of the GSOE with the number of measurements.

Figure 3.3 shows the GSOE-based estimates obtained under different SNR conditions. For SNR values of 5 dB and 10 dB, the average sparsity order estimate converged to 200 ± 10 with nearly $M_{GSM} = 200$ measurements. However, for an SNR value of 0 dB, the effect of measurement noise is significant, and the estimator requires nearly $M_{GSM} > 300$ measurements for convergence.
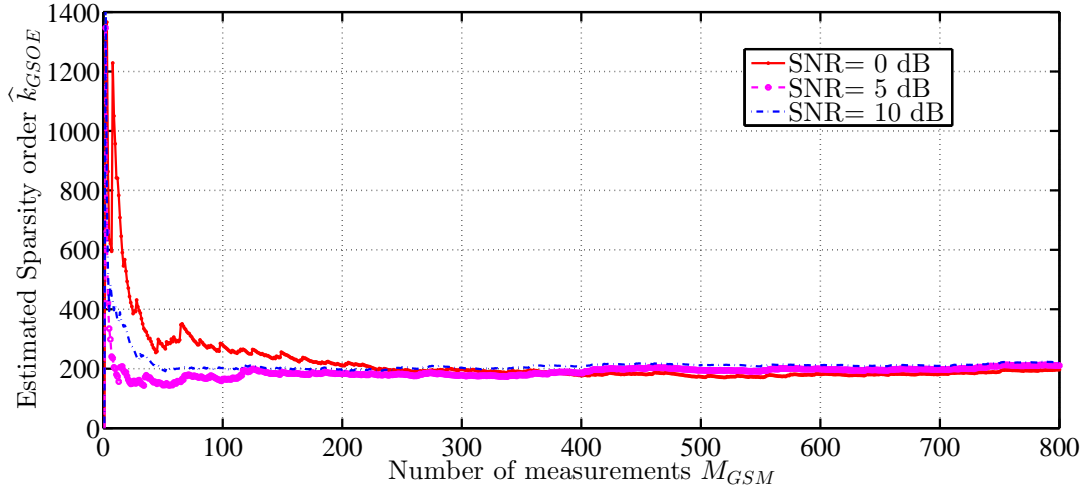
Figure 3.3: GSOE performance on estimating the sparsity order $k = 200$ for the different number of measurements under different SNR values.

3.5 Comparing BSOE and GSOE

Both the BSOE and GSOE are ML estimation techniques. The BSOE technique estimates the statistics and sparsity order with a limited number of measurements, and its estimation performance is shown in Figure 3.2. The GSOE technique requires estimates of statistics obtained from the BSOE to perform SOEE. Thus, GSOE depends on BSOE, and its accuracy depends on the estimation accuracy of BSOE. The performance results are presented in Fig. 3.3.

To compare the SOE performances of BSOE and GSOE, a simulation was performed for a compressible signal with dimension $N = 2500$ whose sparsity order k was varied from 50 to 500 in steps of 25, and the SNR was maintained at 10 dB. The statistics for the significant and insignificant coefficients were $\mu_s = 1, \sigma_s^2 = 25$ and $\sigma_\epsilon^2 = 0.2$, respectively. There were 100 realizations for each sparsity order value. In each realization, the support was randomly chosen, and the BSOE and GSOE estimates were obtained and averaged for 100 realizations. Considering $\widehat{k}_{BSOE}$ and $\widehat{k}_{GSOE}$ are the average BSOE and GSOE estimates, respectively, the weighted average of these two estimates is computed as

$$\widehat{k}_{avg} = \frac{M_{BSM}\widehat{k}_{BSOE} + M_{GSM}\widehat{k}_{GSOE}}{M_{BSM} + M_{GSM}}. \tag{3.38}$$

The results of these estimates are presented in Fig. 3.4. The BSOE and GSOE estimates perform similarly, and their weighted average $\widehat{k}_{avg}$ shows a better accuracy. As both BSOE and GSOE are ML estimation techniques, the weighted average of their estimates, i.e., $\widehat{k}_{avg}$ is denoted as $\widehat{k}_{ML}$ from now.
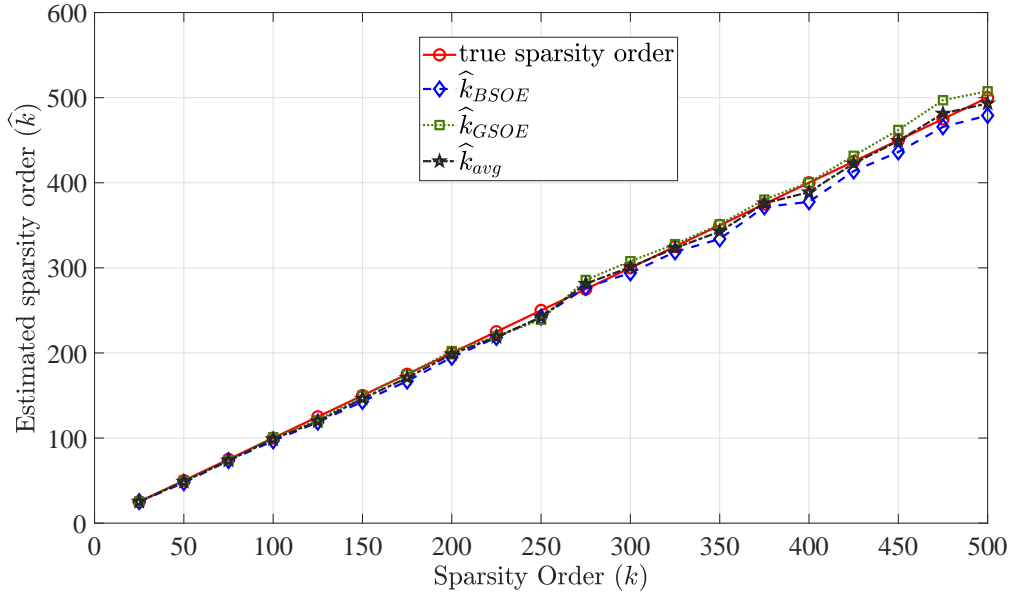
Figure 3.4: Performances of BSOE and GSOE on estimating the sparsity orders $k = 25$ to $k = 500$ under the SNR value of 10 dB. It is observed that both BSOE and GSOE perform similarly and the weighted average estimate of BSOE and GSOE i.e., $\widehat{k}_{avg}$ has better accuracy.

The SOEE performance of $\widehat{k}_{ML}$ is assessed for different SNR values, as shown in Figure 3.5. Here, for each SNR value, the SOEE was computed for different sparsity order

values $k = 25$ to $k = 500$ and the result was averaged as

$$\text{Average SOEE (ASOEE)} = \frac{\sum_k SOEE(k)}{\text{total number of sparsity order values}}. \qquad (3.39)$$

It was observed that as the SNR increased, the SOEE performance improved and the ASOEE decreased. This result is similar to the SOEE performance results for BSOE shown in Fig. 3.2.
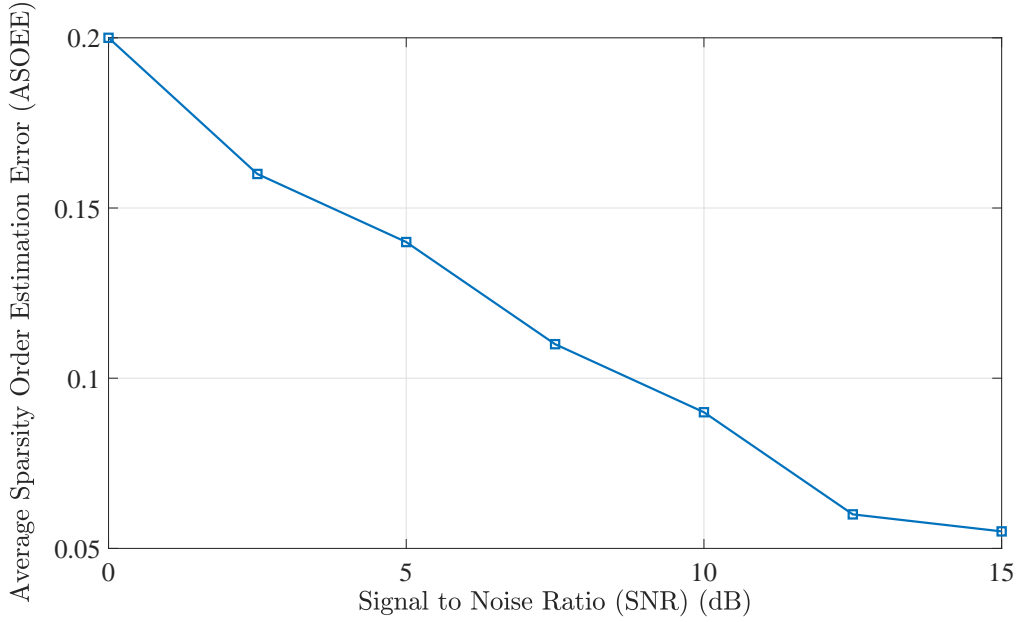
Figure 3.5: Average SOEE performance of $\widehat{k}_{ML}$ for different SNR values.

3.6 Summary

In this chapter, the ML estimators BSOE and GSOE using BSM and GSM matrices are derived. The statistical properties of both estimators are presented. Neither of these estimators requires any *a priori* information regarding the statistics of the compressible signal. These are unbiased estimators, and their performances are similar. The weighted average of both estimators is used as the ML estimator in subsequent chapters.

Chapter 4

Improving the maximum likelihood estimates

In the previous chapter, the ML estimators GSOE and BSOE were derived, and their properties were discussed. From Equation (3.23), it is observed that the variance of $\widehat{K}_{BSOE}$ increases with sparsity order k. In fact, a small perturbation in the value of $\widehat{p}_0$ would result in a large variation in the value of $\widehat{K}_{BSOE}$ as observed from the derivative of the function $\widehat{K}_{BSOE}$ with respect to the variable $\widehat{p}_0$ i.e.,

$$\frac{\partial \widehat{K}_{BSOE}}{\partial \widehat{p}_0} = \frac{1}{\widehat{p}_0 \log \lambda}.$$

Hence, for any fixed λ, the estimator $\widehat{K}_{BSOE}$ must be stabilized against its inherent sensitivity to randomness in the estimate $\widehat{p}_0$.

It is also observed from Equation (3.33) that in the absence of measurement noise and insignificant coefficients, the variance of $\widehat{K}_{GSOE}$ approaches the lower bound $2k^2/M_{GSM}$ which determines the spread in the estimate $\widehat{k}_{GSOE}$. The spread of the estimate decreases when the measurement size M_{GSM} increases. However, because the lower bound is directly proportional to k^2, the measurement size M_{GSM} required to reduce the spread increases quadratically with sparsity order k which defeats the purpose of CS.

The robustness of the ML estimators can be improved either by increasing the number of measurements or by utilizing an appropriate model for the time-varying sparsity order $k(n)$. For efficient CS acquisition, the measurement size should not be increased beyond the minimum value required for perfect CS recovery. Thus, by exploiting the underlying discrete Markov model for time-varying sparsity order, the ML estimates are improved. Two approaches are considered here to refine the ML estimates. The first approach employs an ML Sequence (MLS) estimator using the Viterbi algorithm, and the second employs a

Kalman filter, which is the optimal Linear Minimum Mean-Squared Estimator (LMMSE).

4.1 Approach 1: ML sequence estimation of ML estimates

The ML estimates can be refined further by exploiting the Markovian property of sparsity order variation. This property fits the framework of the MLS estimation implemented using the classical Viterbi algorithm.

The MLS estimates $\{\hat{k}_{MLS}(n)\}_{n=1}^{L}$ are obtained by identifying an optimal L-length sequence of sparsity orders that has the maximum probability of occurrence under the condition that the ML estimates of sparsity orders $\{\hat{k}_{ML}(n)\}_{n=1}^{L}$ are available, as given in Equation (4.1).

$$\{\hat{k}_{MLS}(n)\}_{n=1}^{L} = \underset{\{k_{MLS}(n)\}_{n=1}^{L}}{\arg\max} \ Pr\left[\{k_{MLS}(n)\}_{n=1}^{L} \mid (\{\hat{k}_{ML}(n)\}_{n=1}^{L}, \ p_d)\right]. \tag{4.1}$$

Before implementing Equation (4.1), the ML estimates must be filtered to remove outliers, if any, and to reduce the spread of values in the sequence of the ML estimates $\{\widehat{k}_{ML}(n)\}_{n=1}^{L}$. Because the sparsity order remains constant on an average of $\widetilde{\boldsymbol{L}}$ survival time steps, a simple $\widetilde{\boldsymbol{L}}$−tap moving average (MA) filter is considered to prefilter the ML estimates. The filtered estimate is rounded to the nearest integer. The effect of filtering the ML estimates using MA is shown in Figures 4.1(a) and (b) for two different sequences of sparsity orders with slow and fast variations, respectively. The ML estimates of the slow- and fast-varying sparsity orders (sparsity levels) are filtered using $\widetilde{\boldsymbol{L}} = 10$ and $\widetilde{\boldsymbol{L}} = 3$ taps, respectively. These filtered ML estimates are then refined using the Viterbi algorithm. To maintain the brevity of the notation, we denote the filtered estimates as $\widehat{k}_{ML}$ in the following sections.

4.1.1 Viterbi Algorithm

MLS estimation using the Viterbi algorithm is similar to estimating the most probable sequence of L hidden states from the sequence of L observations emitted by those hidden states in the hidden Markov problem. Here, the MLS and filtered ML estimates represent hidden states and observations, respectively. Because hidden states are refined versions of observations in this hidden Markov problem, the transition probability values of the hidden states and the emission probability values of observations from those hidden states are the same.
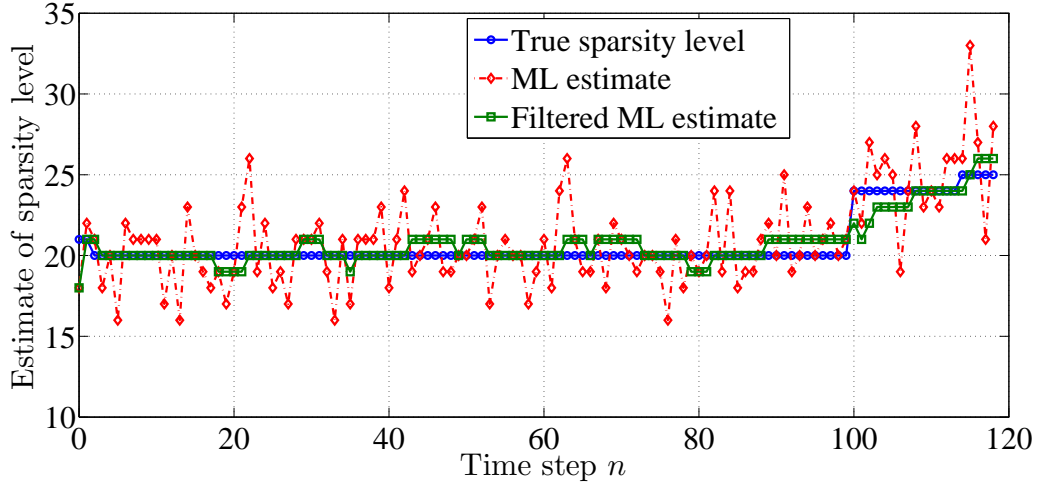
(a) Slow varying: $p_0 = 0.98$ and SNR=10 dB
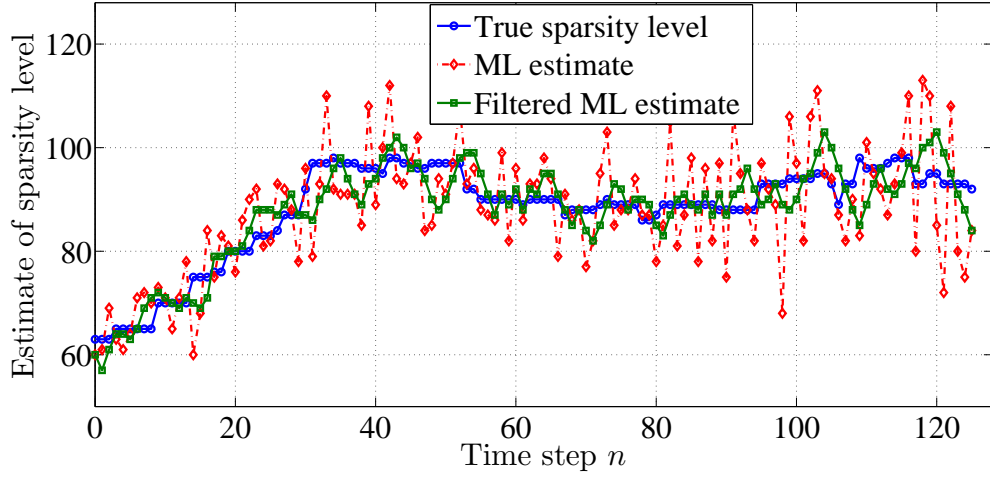
(b) Fast varying: $p_0 = 0.75$ and SNR=10 dB

Figure 4.1: Effect of MA filtering of ML estimates of slow and fast varying sparsity orders (sparsity levels).

The MLS estimate $\hat{k}_{MLS}(L)$ at L^{th} instant is obtained by maximizing the probability q_L^v, which is the probability that the sequence $\{\hat{k}_{MLS}(n)\}_{n=1}^{L-1}$ accounts for first $L-1$ refined estimates of $\{\hat{k}_{ML}(n)\}_{n=1}^{L}$ and $\hat{k}_{MLS}(L) = v$ at L^{th} instant given the knowledge of sparsity order variation probability p_d, i.e.,

$$q_L^v = \max_{\{\hat{k}_{MLS}(n)\}_{n=1}^{L-1}} Pr[\{\hat{k}_{MLS}(n)\}_{n=1}^{L-1}, \hat{k}_{MLS}(L) = v,$$
$$\{\hat{k}_{ML}(n)\}_{n=1}^{L} \mid p_d]. \tag{4.2}$$

Using the induction principle, the highest probability for selecting the sub-sequence $\{\hat{k}_{MLS}(n)_{n=1}^{L-2}$ such that the sequence $\{\hat{k}_{MLS}(n)_{n=1}^{L-2}, \hat{k}_{MLS}(L-1) = u, \hat{k}_{MLS}(L) = v\}$ ending at the sparsity order v can be computed as,

$$q_L^v = \max_u \left[q_{L-1}^u p_d \right]. \tag{4.3}$$

Considering $\Gamma_L(v) = \arg\max_u \left[q_{L-1}^u p_d \right]$, the MLS estimates at time steps $n = L-1, L-2, ..., 1$ can be computed using the trace-back technique, i.e., $\hat{k}_{MLS}(n) = \Gamma_{n+1}(\hat{k}_{MLS}(n+1))$.

4.1.2 ML Estimation of Markov Model Parameters

As the probability p_d is not available *a priori* for the MLS estimation, it is estimated using the ML principle [65] from the filtered ML estimates. Considering the sequence $\{\hat{k}_{ML}(n)\}_{n=1}^L$, the ML estimated transition probability $\hat{p}_d$ is given as,

$$\hat{p}_d = \frac{\text{number of } d \text{ transitions in } \{\hat{k}_{ML}(n)\}_{n=1}^L}{L}. \tag{4.4}$$

4.1.3 Sensitivity of MLS Estimation

The errors in the input sequence $\{\hat{k}_{ML}(n)\}_{n=1}^L$ to the MLS estimator affect the estimate of p_d, leading to a model mismatch. To demonstrate the sensitivity of the MLS estimator, we simulated a discrete Markov model in which the birth and death transition probabilities were considered equal and kept constant throughout the simulation. For each p_0, a sequence of 128 sparsity orders was generated and perturbed by introducing errors at random locations to mimic the errors in the ML estimate after filtering. The probabilities $p_{<0} = \sum_{d<0} p_d$, p_0, and $p_{>0} = \sum_{d>0} p_d$ were estimated using Equation (4.4), and the Viterbi estimator was applied to the perturbed sequence. For each p_0, the numbers of errors introduced were 10, 20, 30, 40, and 50. The estimates of the transition probabilities for each p_0 after the introduction of the errors are shown in Figure 4.2. The robustness of the MLS estimation method in terms of the number of sparsity order errors before and after MLS estimation is shown in Figure 4.3 for different values of model parameters.
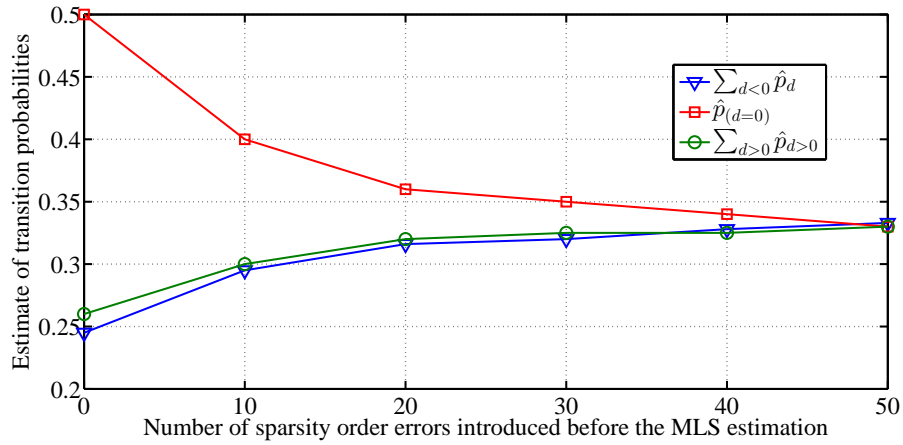
   Observations:

   1. For the slow varying case ($p_0 = 0.98$, $p_{<0} = 0.01$, and $p_{>0} = 0.01$), when the errors in ML estimates are less, the errors in the estimated transition probabilities ($\hat{p}_{<0}$,
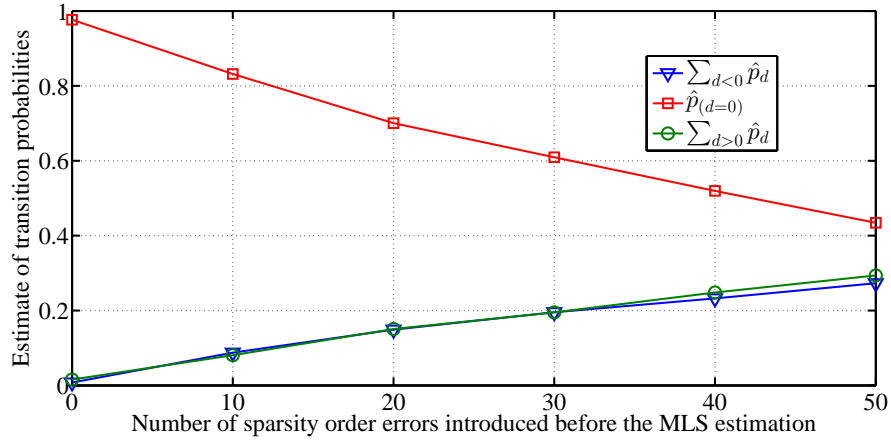
(a) Case1: Slow varying: $p_0 = 0.98$
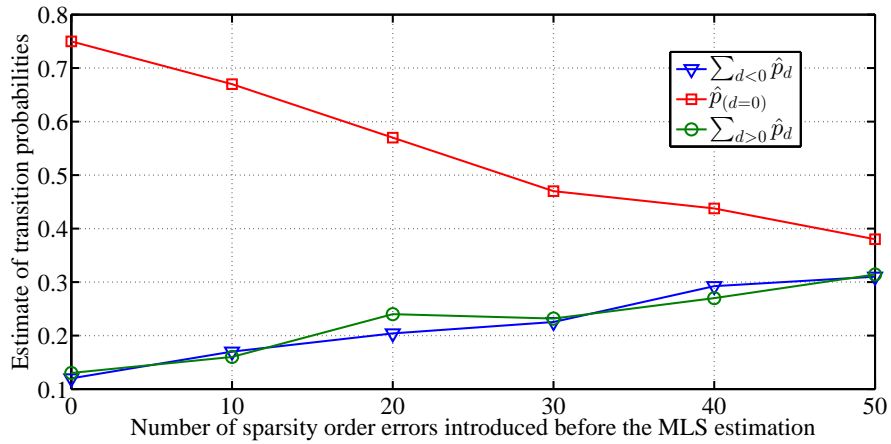
(b) Case2: Fast varying: $p_0 = 0.75$

(c) Case3: Rapid varying: $p_0 = 0.33$

Figure 4.2: Effect of ML estimation error on estimating the transition probabilities.
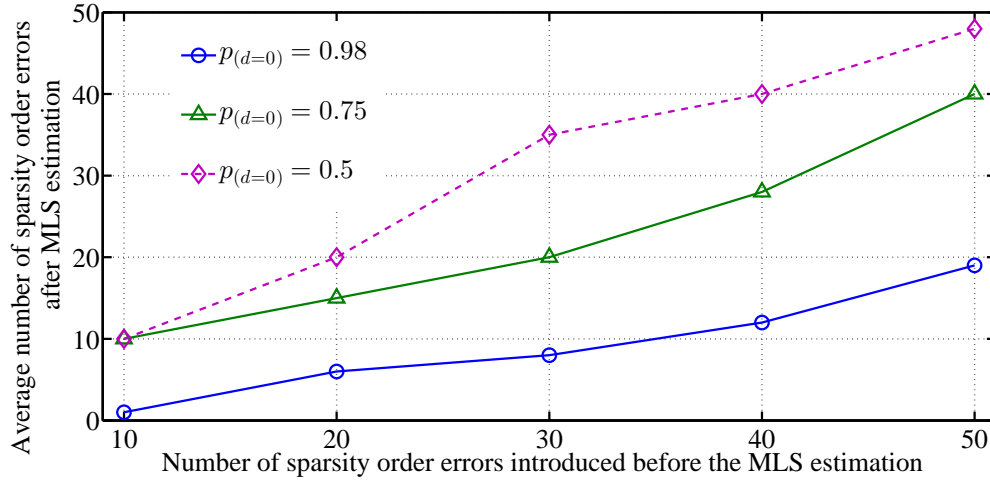
Figure 4.3: Robustness of MLS estimation method on correcting the sparsity order errors (errors in the $\hat{k}_{ML}$ estimates) for different values of Markov model parameters.

$\hat{p}_{>0}$, and $\hat{p}_0$) are also less, as shown in Figure 4.2(a), and the MLS estimation was sufficiently robust to correct most of the errors, as shown in Figure 4.3. However, when the errors in the ML estimates are large and bursty in nature, the errors in the estimated transition probabilities are large, as shown in Figure 4.2(a), and the performance of the Viterbi algorithm deteriorated and exhibited poor performance, as shown in Figure 4.3.

2. The worst-case scenario is for the rapid varying case i.e., when $p_0 = 0.33$, $p_{<0} = 0.33$, and $p_{>0} = 0.34$. Here, the errors in the ML estimates of the sparsity order are indistinguishable from those of the natural variation in the sparsity order, and the ML estimated transition probabilities are independent of the number of errors introduced, as shown in Figure 4.2(c). Therefore, the Viterbi estimator cannot recognize these errors. Hence, the correction was reduced, as shown in Figure 4.3.

Thus, the error in the estimate of transition probabilities harms the MLS output only if many ML estimates are in error or if the rate of change of the sparsity order is very high. It can be observed in Figures 4.2(a), (b), and (c) when no errors are introduced in the sparsity order, the estimate of the transition probabilities converges to the true value of transition probabilities. When a larger number of errors is introduced in the sparsity order, the perturbed sequence of the sparsity orders exhibits uniform variations. Thus, the estimates of the total birth, death, and survival probabilities converged to 0.33.

The MLS estimation technique is summarized in Algorithm 5.

Algorithm 5 MLS estimation of sparsity order

Input: $\mathbf{y}(n)$, Known values: $\widehat{\sigma_\vartheta^2}, \widetilde{\mathbf{L}}, L$

1. Estimate the statistics of the coefficients by using the BSOE algorithm given in Algorithm 4.

2. Estimate the sparsity order using BSOE.

   $\widehat{k}_{BSOE} = \frac{\log(\widehat{P}_0)}{\log(\lambda)}.$

3. Estimate the sparsity order using the GSOE.

   $\widehat{k}_{GSOE} = \frac{\frac{1}{M_{GSM}}\sum_{i=1}^{M_{GSM}} y_i{}^2 - (N\sigma_\phi^2\sigma_\epsilon^2 + \sigma_\vartheta^2)}{\sigma_\phi^2(\mu_s^2 + \sigma_s^2 - \sigma_\epsilon^2)}.$

4. Compute the ML estimate $\widehat{k}_{ML} = \left\lfloor \frac{M_{BSM}\widehat{k}_{BSOE} + M_{GSM}\widehat{k}_{GSOE}}{M_{BSM} + M_{GSM}} \right\rfloor.$

5. Filter the ML estimate using a MA filter of length $\widetilde{L}$.

   $\hat{k}_{ML}(n) = \lfloor \frac{\sum_{I=n-\widetilde{L}+1}^{n} \hat{k}_{ML}(I)}{\widetilde{L}} \rceil$

6. Estimate the transition probability values of the Markov model:

   $$\hat{p}_d = \frac{\text{number of } d \text{ transitions in } \{\hat{k}_{ML}(I)\}_{I=n-L+1}^{n}}{L}.$$

7. Apply the Viterbi algorithm to the sequence $\{\hat{k}_{ML}(I)\}_{I=n-L+1}^{n}$ using the estimated transition probability values to obtain the MLS estimates $\{\hat{k}_{MLS}(I)\}_{I=n-L+1}^{n}$.

8. $n = n + 1$ and repeat from Step 1.

Output: MLS estimates: $\hat{k}_{MLS}(n)$

4.2   Approach 2: Kalman filtering of ML estimates

Kalman filtering of BSOE-based ML estimates (KML) uses the discrete Markov model characterization of time-varying sparsity order $k(n)$ given in Equation (2.20) as the state model. This state model can be represented as,

$$k(n) = k(n-1) + w(n) : \text{VAR}\{w(n)\} = Q_w, \tag{4.5}$$

where $w(n)$ is an integer-valued random process with $w(n) > 0$ indicating birth, $w(n) < 0$ indicating death, and $w(n) = 0$ indicating survival.

The observation for the Kalman filter is the estimate of the probability $P_0(n)$ given as $\widehat{P_0}(n) = \frac{\sum_{i=1}^{M_{BSM}} \delta_\tau(y_i)}{M_{BSM}}$. Due to the inherent statistical variations in computing the probability $P_0(n)$, its estimate $\widehat{P_0}(n)$ can be written as,

$$\widehat{P_0}(n) = P_0(n) + v(n) = \lambda^{k(n)} + v(n), \tag{4.6}$$

where $v(n)$ denotes zero-mean random statistical fluctuation. Observe that the nonlinear relation in Equation (4.6) prevents one from applying the Kalman filter to obtain an estimate of $k(n)$. To overcome this problem, we apply the logarithm and arrive at a modified form of the observation model, as given below, with the state-dependent noise $\bar{v}(n)$.

$$
\begin{aligned}
\log(\widehat{P_0}(n)) &= \log(\lambda^{k(n)} + v(n)) && (4.7) \\
&= k(n)\log(\lambda) + \log\left(\frac{v(n)}{\lambda^{k(n)}} + 1\right) && (4.8) \\
&= k(n)\log(\lambda) + \bar{v}(n). && (4.9)
\end{aligned}
$$

Now let us consider the modified noise

$$\bar{v}(n) = \log\left(\frac{v(n)}{\lambda^{k(n)}} + 1\right),$$

which has a zero mean and variance $\sigma_{\bar{v}}^2 \approx \frac{1 - \lambda^{k(n)}}{M_{BSM}(n)\lambda^{k(n)}}$. Then, the linear Kalman filter using Equations (4.5) and (4.7) is as follows:

$$\widehat{k}(n|n\text{-}1) = \widehat{k}(n\text{-}1|n\text{-}1) \tag{4.10}$$

$$C(n|n\text{-}1) = C(n\text{-}1|n\text{-}1) + Q_w \tag{4.11}$$

$$K(n) = \frac{C(n|n\text{-}1)\log(\lambda)}{C(n|n\text{-}1)(\log(\lambda))^2 + \sigma_{\tilde{v}}^2} \tag{4.12}$$

$$r(n) = \log(\widehat{P}_0(n)) - \log(\lambda)\widehat{k}(n|n\text{-}1) \tag{4.13}$$

$$\widehat{k}(n|n) = \widehat{k}(n|n\text{-}1) + K(n)r(n); \tag{4.14}$$

$$C(n|n) = (1 - \log(\lambda)K(n))C(n|n\text{-}1) \tag{4.15}$$

$$\widehat{k}(n|n) = \widehat{K}_{KML}(n) = \left\lfloor \widehat{k}(n|n) \right\rceil. \tag{4.16}$$

where $C(n|n\text{-}1)$ and $C(n|n)$ are the *priori* and *posteriori* estimation error covariances, respectively, $K(n)$ is the Kalman gain, $r(n)$ is the residual, $\widehat{k}(n|n\text{-}1)$ and $\widehat{k}(n|n)$ are the *priori* and *posteriori* estimates of $k(n)$, respectively, and the Kalman filtered ML estimate is $\widehat{K}_{KML}(n)$.

The procedure for estimating the sparsity order using BSOE and GSOE followed by Kalman filtering is summarized and presented in Algorithm 6.

4.2.1 Effect of Kalman filtering with optimum value of λ

The effect of Kalman filtering on static and time-varying sparsity order is discussed below.

4.2.1.1 Case 1: Time-invariant sparsity order

The reduced variance of the Kalman-filtered estimate $\widehat{k}_{KML}$ was verified for different time-invariant sparsity order values k, as shown in Table 4.1. Here, at every time step, for a given sparsity order k, the support $\mathcal{S}_{\mathbf{s}}$ of the significant coefficients alone varies.

k	VAR$\{\widehat{k}_{ML}\}$	VAR$\{\widehat{k}_{KML}\}$
100	45.071	0.010
150	61.026	0.108
200	89.371	0.157
250	116.247	0.167
300	125.036	0.172

Table 4.1: Illustration of reduction in variance using Kalman filtering. The table shows that the Kalman filter reduces the variance of the ML estimate drastically.

Algorithm 6 Kalman Filtering of ML estimates

Input: $\mathbf{y}(n)$, Known values: $\widehat{\sigma_{\vartheta}^2}$

1. Estimate the statistics of the coefficients by using the BSOE algorithm given in Algorithm 4.

2. Estimate the sparsity order using BSOE.

   $\widehat{k}_{BSOE} = \frac{\log(\widehat{P}_0)}{\log(\lambda)}.$

3. Estimate the sparsity order using the GSOE.

   $\widehat{k}_{GSOE} = \frac{\frac{1}{M_{GSM}}\sum_{i=1}^{M_{GSM}} y_i{}^2 - (N\sigma_\phi^2\sigma_\epsilon^2 + \sigma_{\vartheta}^2)}{\sigma_\phi^2(\mu_s^2 + \sigma_s^2 - \sigma_\epsilon^2)}.$

4. Compute the ML estimate $\widehat{k}_{ML} = \left\lfloor \frac{M_{BSM}\widehat{k}_{BSOE} + M_{GSM}\widehat{k}_{GSOE}}{M_{BSM} + M_{GSM}} \right\rceil.$

5. Estimate the variance of the state noise.

   $$\sigma_{\overline{v}}^2 = \frac{1 - \lambda^{\widehat{k}_{ML}(n)}}{M_{BSM}(n)\lambda^{\widehat{k}_{ML}(n)}}.$$

6. Estimate the variance Q_w from the sparsity order variations.

   $$Q_w = VAR\{\widehat{k}_{ML}(n\text{-}1) - \widehat{k}_{ML}(n\text{-}2)\}.$$

7. Kalman filtering of $\widehat{P}_0(n)$:

   $$\begin{aligned}
   \widehat{k}_{KML}(n|n\text{-}1) &= \widehat{k}_{KML}(n\text{-}1|n\text{-}1) & (4.17) \\
   C(n|n\text{-}1) &= C(n\text{-}1|n\text{-}1) + Q_w & (4.18) \\
   K(n) &= \frac{C(n|n\text{-}1)\log(\lambda)}{C(n|n\text{-}1)(\log(\lambda))^2 + \sigma_{\overline{v}}^2} & (4.19) \\
   r(n) &= \log(\widehat{P}_0(n)) - \log(\lambda)\widehat{k}_{KML}(n|n\text{-}1) & (4.20) \\
   \widehat{k}_{KML}(n|n) &= \widehat{k}_{KML}(n|n\text{-}1) + K(n)r(n); & (4.21) \\
   C(n|n) &= (1 - \log(\lambda)K(n))C(n|n\text{-}1) & (4.22) \\
   \widehat{k}_{KML}(n|n) &= \widehat{k}_{KML}(n) = \left\lfloor \widehat{k}_{KML}(n|n) \right\rceil. & (4.23)
   \end{aligned}$$

Output: Kalman filtered estimate of sparsity order: $\widehat{k}_{KML}(n)$
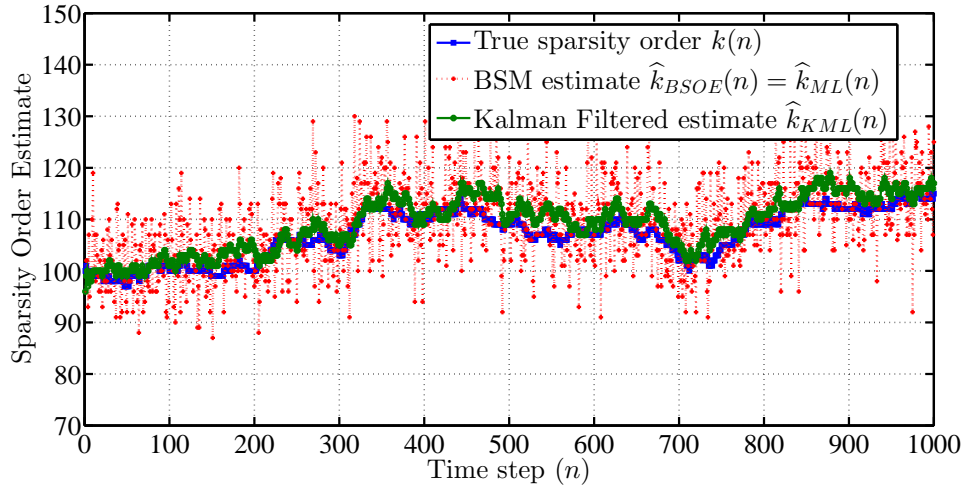
Figure 4.4: Comparision of sparity order estimates. The BSOE estimates have fluctuations around the true value and the Kalman filter reduces those fluctucations and improves the SOE.

4.2.1.2 Case 2: Time-varying sparsity order

A simulation example is presented in Figure 4.4 to establish an improvement by Kalman filtering for the time-variant sparsity order $k(n)$. In the simulation setup, the sparsity order was varied using a Markov process controlled by probabilities $\Pr[w(n) = 0] = 0.8$, $\Pr[w(n) < 0] = 0.1$, and $\Pr[w(n) > 0] = 0.1$. A time-varying compressible signal with $k(0) = 100$ significant components (above a certain threshold) and the rest with insignificant components (below the threshold), obeying power-law decay, is generated. The number of significant components is allowed to vary, i.e., $k(n) : n > 0$ is generated using the Markov process and is represented as true sparsity order. Figure 4.4 shows that the Kalman filtered estimates $\widehat{k}_{KML}(n)$ track the true sparsity order and have a reduced error compared with the estimates $\widehat{k}_{KML}(n)$ obtained using BSOE and GSOE.

Chapter 5

Proposed CS acquisition and recovery system

This chapter covers the practical implementation of the CS acquisition and recovery systems. The design of the composite sensing matrix and the procedure for determining the number of measurements are described. A method for coarse support estimation through BSM-based measurements that accelerates the recovery of compressible signals is also put forth.

5.1 CS acquisition system

The practical real-time composite sensing hardware for the proposed CS acquisition system is shown in Figure 5.1. There are M identical and independent modulator circuits working in parallel. The hardware components of a modulator circuit are (i) a multiplexer to select between the sensing basis $\mathbf{g}_i(t)$ (continuous-time version of rows of GSM multiplied by Ψ^{-1}) and the sensing basis $\mathbf{b}_i(t)$ (continuous-time version of rows of BSM multiplied by Ψ^{-1}) and (ii) a multiplier and an Integrate and Dump (I&D) circuit to multiply and integrate the compressible signal with the sensing basis for a duration of $\mathcal{T}$ seconds to output a measurement.

The select signal $\beta_i(t)$ of the multiplexer takes the value 0 for $i \leq M_{BSM}$ to select the basis $\mathbf{b}_i(t)$ and 1 for $M_{BSM} < i \leq M$ to select the basis $\mathbf{g}_i(t)$. From the obtained measurements, the proposed KML method estimates the sparsity order and then decides $\beta_i(t)$ and M for the next $\mathcal{T}$ seconds.

The proposed architecture is similar to the practical Random Modulation Pre-Integration (RMPI) hardware [66] with a difference in the multiplexer and select signal for multiplexing the sensing basis $\mathbf{b}_i(t)$ and $\mathbf{g}_i(t)$.
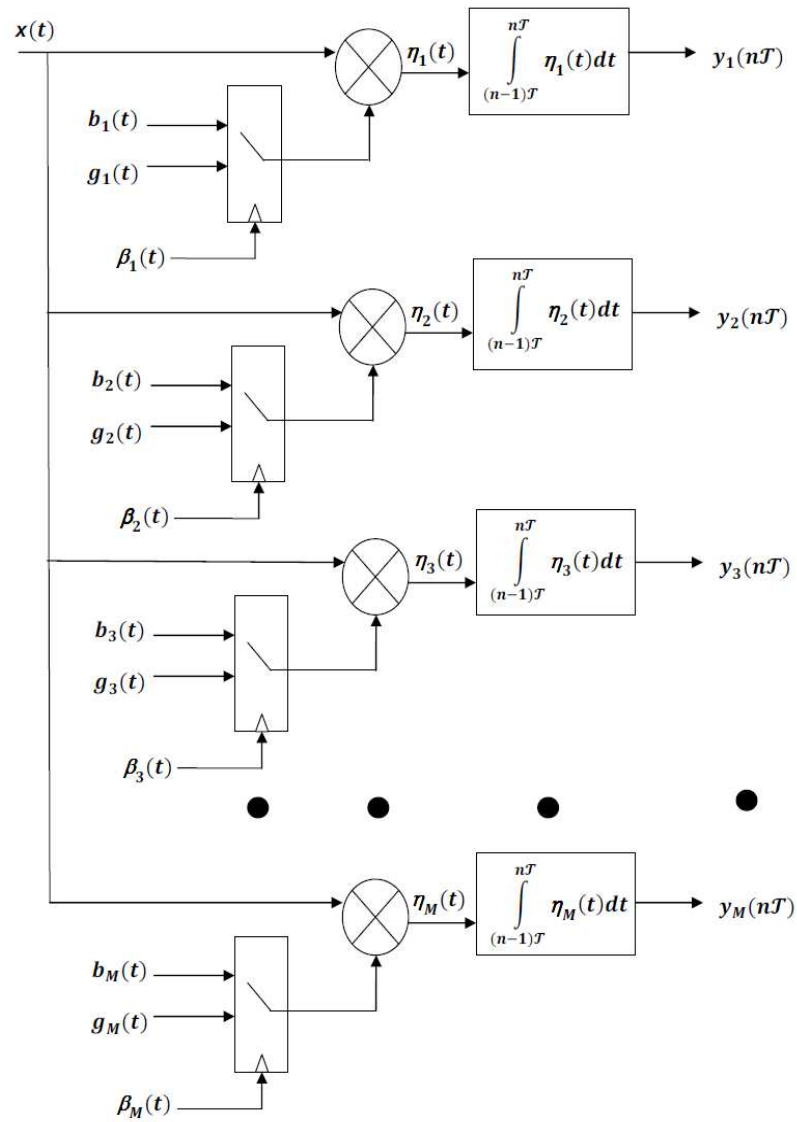
Figure 5.1: The block diagram of CS acquisition hardware. This is similar to RMPI hardware with the difference of multiplexer switching between GSM and BSM equivalent signals.

The real-time determination of the total number of measurements based on the estimated sparsity order is explained as follows.

5.1.1 Determining the number of measurements for the composite sensing system

It is known that the sparsity order $k(n)$ must be estimated from the obtained measurements $\mathbf{y}(n)$. However, the measurements size $M(n)$ of $\mathbf{y}(n)$ has to be estimated based on the sparsity order $k(n)$ before obtaining $\mathbf{y}(n)$ using Equation (1.14) to result in a Chicken-Egg problem. Thus, it is practically impossible to simultaneously estimate the sparsity order and the size of the measurements for the current time step n. Hence, the estimated sparsity order $\widehat{k}(n)$ for the current time step n is used to determine the number of measurements $M(n+1)$ for the next time step $n+1$. This determination is acceptable because naturally occurring time-varying compressible signals are quasi-static and exhibit a stronger temporal correlation resulting in quasi-static sparsity order. Thus after obtaining the measurement vector $\mathbf{y}(n)$, the sparsity order $k(n)$ is estimated to determine the number of measurements $M(n+1)$ for the next time step. As the composite sensing system is built using BSM and GSM, the number of measurements required for each is computed as follows.

5.1.1.1 Number of BSM-based measurements

The number of BSM measurements depends on the parameter λ which is the probability of an entry being zero in the BSM. The value of parameter λ is chosen according to the sparsity order $k(n)$ of the signal as $\lambda = 1 - \exp(-1.6/k(n))$. As the sparsity order $k(n)$ for the current time step is not known, λ is computed using the previous estimate $\widehat{k}(n-1)$ i.e., $\lambda \approx \exp(-1.6/\widehat{k}(n-1))$.

The BSM is designed by stacking $\lfloor N(1-\lambda) \rfloor$ numbers of $\lfloor \frac{1}{1-\lambda} \rfloor \times \lfloor \frac{1}{1-\lambda} \rfloor$ Identity matrices horizontally. Thus there are $\lfloor \frac{1}{1-\lambda} \rfloor$ rows. Hence the number of BSM-based measurements is,

$$
\begin{align}
M_{BSM}(n) &= \left\lfloor \frac{1}{1-\lambda} \right\rfloor \tag{5.1}\\
&= \left\lfloor \frac{1}{(1-\exp(-1.6/k(n-1)))} \right\rfloor \tag{5.2}\\
&= \left\lceil 0.63k(n-1) \right\rceil. \tag{5.3}
\end{align}
$$

5.1.1.2 Number of GSM-based measurements

The number of GSM-based measurements $M_{GSM}(n)$ is computed using Equation 1.14, and is given as,

$$M_{GSM}(n) = \left\lceil 2k(n-1)\log\left(\frac{N}{k(n-1)}\right)\right\rceil. \tag{5.4}$$

Each component $\phi_{i,j}$ of the GSM Φ_{GSM} is i.i.d. Gaussian, such that $\phi_{i,j} \sim \mathcal{N}(0, 1/M_{GSM})$.

Combing Equations (5.3) and (5.4), the total number of measurements acquired by the composite CS acquisition system is,

$$
\begin{aligned}
M(n) &= M_{BSM}(n) + M_{GSM}(n) \tag{5.5}\\
&= \left\lceil k(n-1)\left(0.63 + 2\log\left(\frac{N}{k(n-1)}\right)\right)\right\rceil \tag{5.6}
\end{aligned}
$$

5.1.2 Procedure for determining the number of measurements

When starting the acquisition at $n = 0$, the previous estimate $\widehat{k}(-1)$ is not available. Hence it is assumed that $\widehat{k}(-1) = k_{\max}$, where $k_{\max}$ is the maximum possible sparsity order for a compressible signal which is known *a priori*. Thus the number of measurements obtained initially is,

$$M(0) = \left\lceil k_{\max}\left(0.63 + 2\log\left(\frac{N}{k_{\max}}\right)\right)\right\rceil.$$

Once $M(0)$ measurements are obtained, the proposed SOE method provides the estimate $\widehat{k}(0)$ which in turn determines $M(1)$, and this process continues. Thus both the number of measurements $M(n)$ and the sparsity order $k(n)$ are estimated sequentially, i.e., $\widehat{k}(n-1)$ determines $M(n)$ and these $M(n)$ measurements provide the estimate $\widehat{k}(n)$.

5.1.3 Construction of BSM

The BSM is constructed such that it spans all the significant coefficients distributed across N positions to estimate the statistics μ_s and σ_s^2. As each row of BSM has $N(1 - \lambda)$ ones, the BSM matrix is designed by stacking $\lfloor N(1 - \lambda)\rfloor$ numbers of $\lfloor\frac{1}{1-\lambda}\rfloor \times \lfloor\frac{1}{1-\lambda}\rfloor$ Identity matrices horizontally. If $\lfloor N(1 - \lambda)\rfloor \lfloor\frac{1}{1-\lambda}\rfloor \neq N$, then the remaining $N - \lfloor N(1 - \lambda)\rfloor \lfloor\frac{1}{1-\lambda}\rfloor$ numbers of 1-sparse binary column vectors are added at the end to obtain the $\lfloor\frac{1}{1-\lambda}\rfloor \times N-$dimensional BSM.

As the structure of the BSM changes according to $k(n)$, the BSM must be transmitted to

the recovery process at every time step, along with the obtained measurements for recovery. However, the deterministic construction of the BSM avoids transmitting such overhead as the recovery process performs instantaneous sparsity order estimation from the obtained measurements as done during acquisition.

5.2 CS recovery system

Two approaches to CS recovery systems are presented in this section. The first approach uses the Viterbi algorithm, whereas the second approach uses a Kalman filter, as illustrated in Figure 5.2 and Figure 5.3, respectively. The approach based on Viterbi algorithm performs trace back-based sequence estimation and hence there is a delay of ν time-steps in providing the sparsity order estimates.
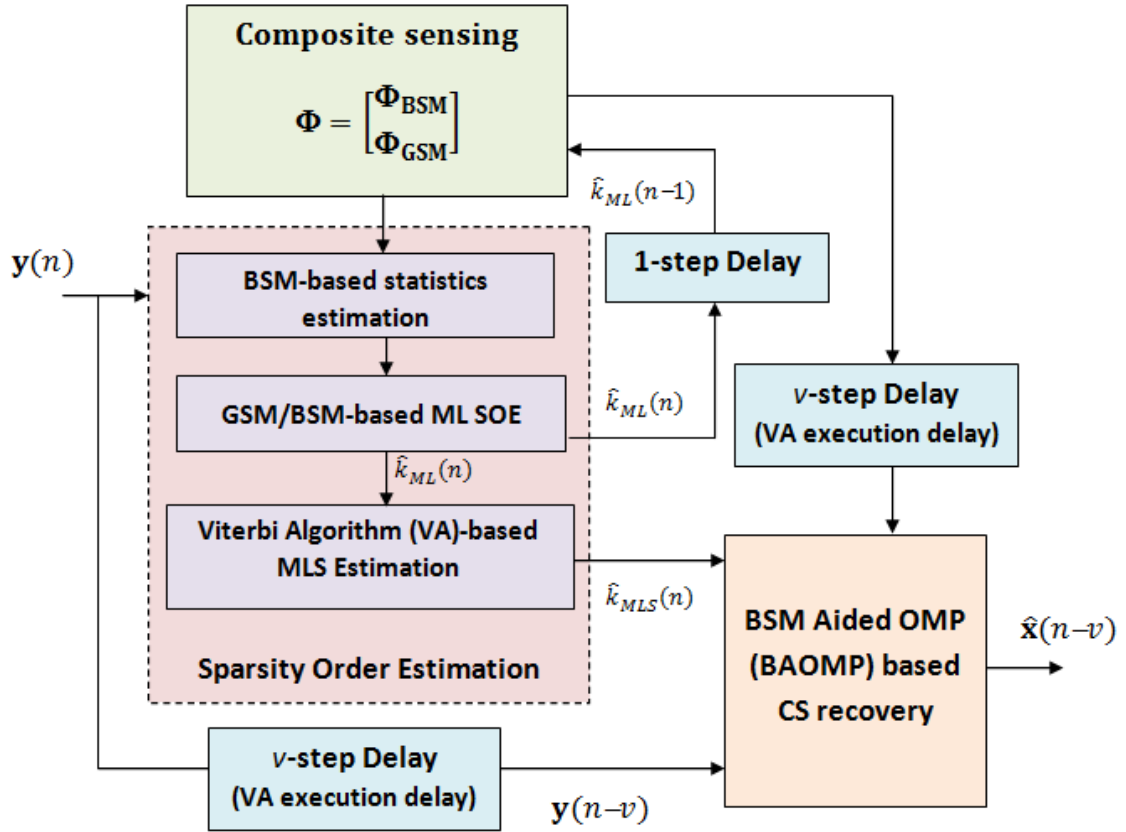
Figure 5.2: Viterbi algorithm and BSM aided CS recovery system. The system comprises of the major blocks: the input measurements, the composite sensing circuit, the Viterbi algorithm-based SOE block, the BAOMP-based recovery block, and the recovered signal.

73
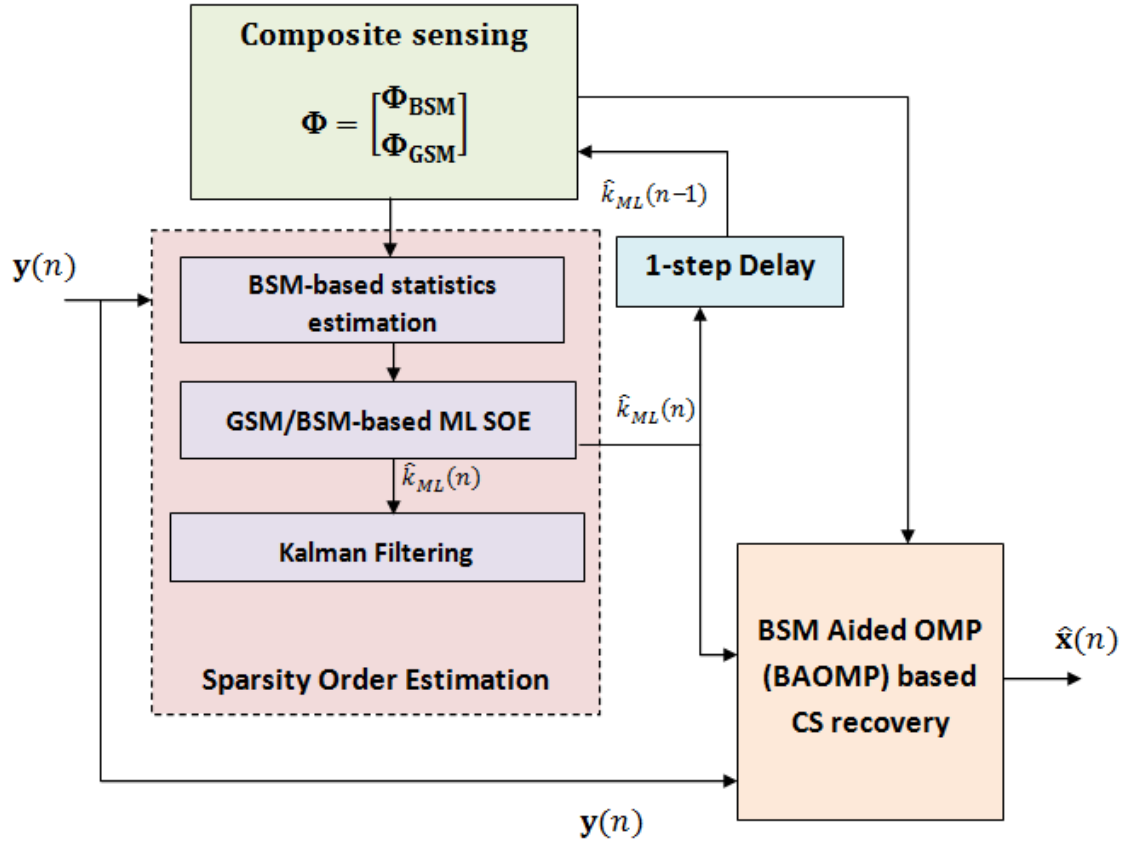
Figure 5.3: Kalman filter and BSM aided CS recovery system. The system comprises of the major blocks: the input measurements, the composite sensing circuit, the Kalman filter-based SOE block, the BAOMP-based recovery block, and the recovered signal.

5.2.1 Implementation of Viterbi algorithm for MLS estimation

Here sliding window-based Viterbi algorithm is considered. There exist two parallel sliding windows and each window requires a Viterbi estimator. The L-length sequence of ML estimates of the sparsity order is decomposed into blocks of M ML estimates. The value M is five times the trace-back length T where T is the maximum sparsity order transition possible in a time step. Each sliding window size is kept similar to the block size, M. Both the Viterbi estimators do not require the knowledge of initial state probabilities and the trellis terminated state.

From Figure 5.4, it can be observed that after receiving Block-1 at time $n = M$, the first Viterbi estimator starts the training of trace-back from sub-block TB. After T time steps i.e., at $n = (M + T)^{th}$ time step, the first Viterbi estimator starts providing the MLS estimates of sparsity order received during time steps $n = T$ to $n = M - T - 1$. Meanwhile,
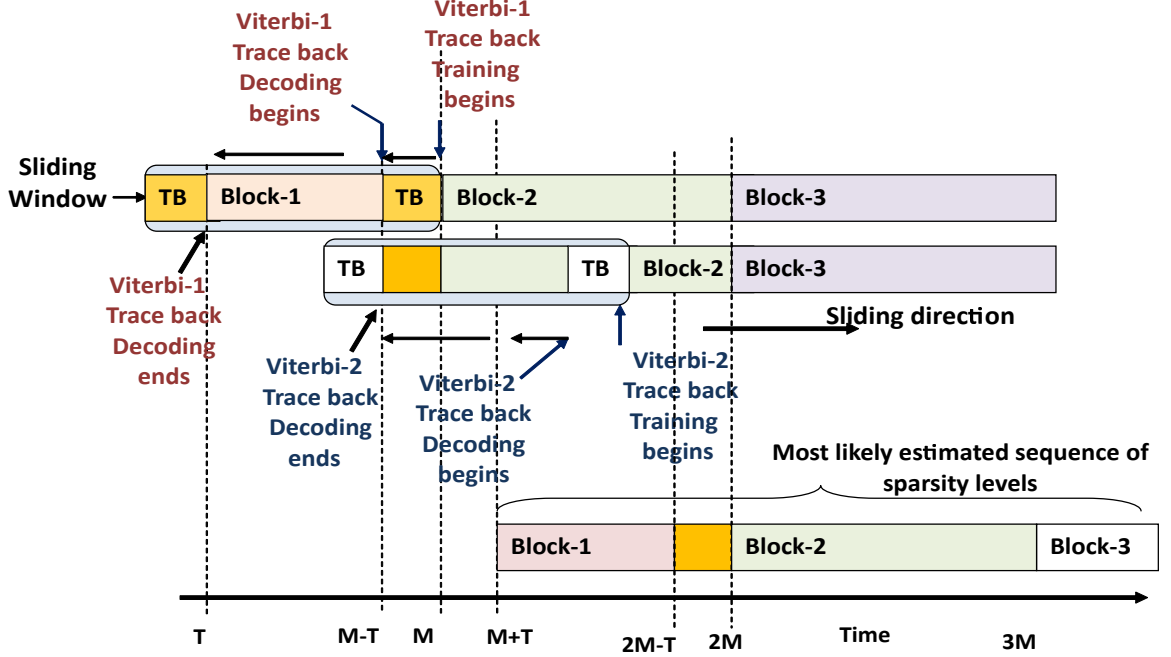
Figure 5.4: Illustration of sliding window Viterbi algorithm.

the second Viterbi estimator starts the trace-back training from the time step $n = 2M - 2T$. After T time steps i.e., at $n = (2M - T)^{th}$ time step, the second Viterbi estimator starts providing the MLS estimates of sparsity order received during time steps $n = M - T$ to $n = 2M - 3T - 1$. Thus with a group delay of M time steps, the sliding window-based Viterbi algorithm provides MLS estimates of sparsity order without knowing the initial state probabilities and trellis termination.

It should be noted that for fast and rapidly varying sparsity order, the Viterbi algorithm exhibits poor performance in correcting the errors introduced by the filtered estimates $\hat{k}_{ML}$ as shown in Figure 4.3. Thus, the proposed Viterbi algorithm-based MLS estimation technique is suitable for slow and moderately varying sparsity order to provide better MLS estimates with lower computational complexity. If the rate of sparsity order variation is higher than that of the estimation speed, then the filtered ML estimates are retained as the best estimates without executing the Viterbi algorithm.

5.2.2 The proposed BSM Aided OMP method

The composite sensing matrices used during CS recovery are the same as those used during CS acquisition. Our proposed recovery system uses BSM-based measurements along with GSM-based measurements. It differs from the conventional CS recovery system that

uses GSM-based measurements alone. The BSM-based measurements are used for SOE as a first step in CS recovery. The SOE techniques used during recovery are the same as those used during acquisition. The estimated sparsity order $\widehat{k}(n)$ is the input for the proposed BSM Aided OMP (BAOMP) recovery algorithm. We choose OMP for recovery as it is simple to implement and has robust recovery performance [5]. In OMP, the probable support indices are identified one by one in each iteration. Because there are k support indices for the $k-$sparse compressible signal, there are k iterations. As the BSM-based measurements provide a few initial estimates of the support indices, the OMP algorithm must estimate only the remaining support indices.

Some BSM measurements may have no contributions from the significant coefficients and have magnitudes less than the threshold $\tau = 3\sqrt{N(1-\lambda)\sigma_\epsilon^2 + \sigma_\vartheta^2}$ as discussed in Chapter 3. On contrary the measurements having magnitudes greater than the threshold, i.e., $|(y_i)_{BSM}| > \tau$ indicates that each of such measurements y_i has at least one significant component's contribution. The corresponding rows of such measurements y_i provide information about the probable support indices based on the locations of 1's in those rows. For example, suppose the dimension of the BSM is 10×250. We construct the BSM by horizontally stacking 25 identity matrices of dimensions 10×10. If the first BSM measurement has a magnitude greater than the threshold τ, then the compressible signal may have significant components probably located at either 1st location, 11th location, 21st location, or so on, up to 241st location. Thus, if $|(y_i)_{BSM}| > \tau$, then $\{i, i+M_{BSM}, i+2M_{BSM}, \ldots, i+((N/M_{BSM})-1)M_{BSM}\}$ are the support indices of the BSM-based support estimate $\widehat{\mathcal{S}}_{BSM}$ as shown in Algorithm 7.

As the GSM obeys the RIP property, the proxy $\mathbf{z} = \mathbf{\Phi}_{GSM}^T \mathbf{y}_{GSM}$ for the compressible signal $\mathbf{x}$ provides the GSM-based support estimate $\widehat{\mathcal{S}}_{GSM}$ by choosing indices of $1 \leq L \leq \widehat{k}$ largest coefficients in $\mathbf{z}$. Now, the common indices between $\widehat{\mathcal{S}}_{BSM}$ and $\widehat{\mathcal{S}}_{GSM}$ are selected as the support indices for the updated support estimate $\widehat{\mathcal{S}}$. Then $\widehat{k} - |\widehat{\mathcal{S}}|$ iterations are performed using the conventional OMP algorithm as shown in Algorithm 8 for the proposed BAOMP method. The value $L = \lfloor \widehat{k}/4 \rfloor$ is chosen experimentally such that the number of iterations is reduced without compromising the quality of the recovered signal $\widehat{\mathbf{x}}$ for all sparsity order values.

Figure 5.5 shows the performance of the BAOMP method for various L values during CS recovery of an $N = 2500-$dimensional compressible signal with a sparsity order of $k = 100$. The performance is measured using Normalised Recovery Error (NRE), Probability
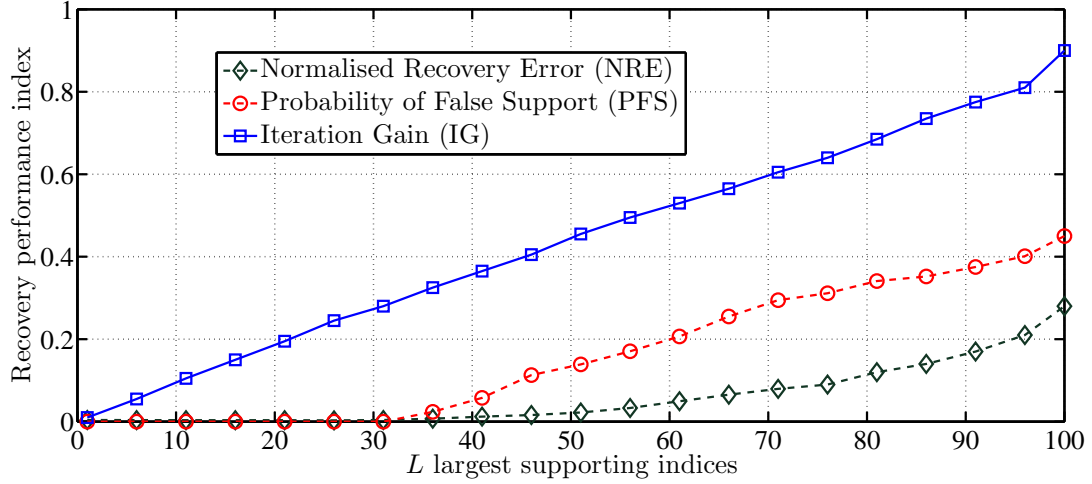
Figure 5.5: Performance of BAOMP algorithm for different L values.

of False Support (PFS), and Iteration Gain (IG) as given below.

$$
\begin{aligned}
\text{NRE} &= \frac{\|\widehat{\mathbf{x}} - \mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2}, \\
\text{PFS} &= \frac{|\widehat{\mathcal{S}}_{BSM} \cap \widehat{\mathcal{S}}_{GSM} \cap \mathcal{S}^C|}{|\mathcal{S}|}, \\
\text{IG} &= \frac{|\widehat{\mathcal{S}}_{BSM} \cap \widehat{\mathcal{S}}_{GSM}|}{\widehat{k}},
\end{aligned}
$$

where $\mathcal{S}$ is the original support of the compressible signal, $\mathcal{S}^C$ is the set containing indices that do not belong to $\mathcal{S}$, i.e., $\mathcal{S}^C$ contains the indices of insignificant coefficients, and $|\widehat{\mathcal{S}}_{BSM} \cap \widehat{\mathcal{S}}_{GSM}|$ represents the number of common indices between the BSM and GSM-based support estimates. Here, NRE measures the quality of the recovered signal, PFS measures the quantity of false support selection, and IG measures how fast BAMOP is compared to OMP. From Figure 5.5, it is observed that for $L = \lfloor \widehat{k}/4 \rfloor = 25$, the performance measures are PFS=0, NRE ≈ 0, and IG ≈ 0.25, that is, the BAOMP algorithm is 25% faster than the OMP algorithm.

Algorithm 7 Estimation of the support using BSM

Input: Threshold τ, BSM-based measurement vector $\mathbf{y}_{BSM}$, and iteration number $i = 1$.

1. An initial estimate of the support is an empty set, i.e., $\widehat{\mathcal{S}}_{BSM} = \varnothing$.

2. Iteration begins.

    do while $(i \leq M_{BSM})$

      if $(|y_{BSM}(i)| > \tau)$ $\widehat{\mathcal{S}}_{BSM} = \widehat{\mathcal{S}}_{BSM} \cup \{i, i + M_{BSM}, i + 2M_{BSM}, ..., i + ((N/M_{BSM}) - 1)M_{BSM}\}$.

      $i = i + 1$.

    end

3. Iteration ends.

Output: The estimated support: $\widehat{\mathcal{S}}_{BSM}$.

5.3 Performance comparison of proposed ML estimators and BAOMP methods with other existing methods

The proposed ML method, followed by the Viterbi algorithm or Kalman filtering method, is used for SOE during CS acquisition and recovery. The proposed BAOMP method is used for CS recovery. In this section, performance measures such as the SOEE and NRE of the proposed methods are compared with existing methods using synthetic signals.

The SOEE is given as,

$$\text{SOEE} = \frac{|\widehat{k} - k|}{k}.$$

The Signal to Noise Ratio (SNR) setting in evaluating the SOEE and NRE performances is calculated at the acquisition side, and it is given as,

$$\text{SNR} = 10 \log_{10} \left(\frac{\|\boldsymbol{\Phi}\mathbf{x}\|_2^2}{\|\vartheta\|_2^2} \right). \tag{5.7}$$

For all the simulations shown here, a Windows 7 Operating System-based PC with a processor running at 3 GHz clock speed and 4 GB RAM is used.

5.3.1 SOEE performance comparison using a synthetic signal

A simulation was performed to compare the SOEE performance of the KML method with other SOE methods such as Lopes [28], 2-GMM [30], eigenvalue [46], trace [29], SPAMP [35], and TS-ACSS [34]-based methods.

Algorithm 8 Estimation of the support using the BAOMP method and recovery of the compressible signal

Input: GSM $\boldsymbol{\Phi}_{GSM}$, Measurement vector $\mathbf{y}_{GSM}$, BSM estimated support $\widehat{\mathcal{S}}_{BSM}$, Sparsity order estimate $\widehat{k}$, and $i = 1$.

1. Initial residual: $\mathbf{r} = \mathbf{y}_{GSM}$.

2. Initial estimate of the support of the compressible signal: $\widehat{\mathcal{S}} = \varnothing$.

3. $\mathbf{z} = \boldsymbol{\Phi}_{GSM}^T \mathbf{r}$ {$\boldsymbol{\Phi}_{GSM}^T$ denotes the transpose of $\boldsymbol{\Phi}_{GSM}$}.

4. GSM-based initial estimate of the support: $\widehat{\mathcal{S}}_{GSM}$ = support of L largest components (in terms of magnitude) in the vector $\mathbf{z}$.

   {We choose $L = \lfloor \widehat{k}/4 \rfloor$ so that the error in the initial estimate of the support is less.}

5. Find the common support among the BSM- and GSM-based measurements: $\widehat{\mathcal{S}} = \widehat{\mathcal{S}}_{BSM} \cap \widehat{\mathcal{S}}_{GSM}$.

6. Update the residual $\mathbf{r} = \mathbf{y}_{GSM} - (\boldsymbol{\Phi}_{GSM})_{|\widehat{\mathcal{S}}} (\boldsymbol{\Phi}_{GSM})_{|\widehat{\mathcal{S}}}^\dagger \mathbf{y}_{GSM}$,

   {$(\boldsymbol{\Phi}_{GSM})_{|\widehat{\mathcal{S}}}$ is the sub-GSM matrix formed by selecting the columns indexed by the support $\widehat{\mathcal{S}}$ and $(\boldsymbol{\Phi}_{GSM})_{|\widehat{\mathcal{S}}}^\dagger$ is the pseudo-inverse of $(\boldsymbol{\Phi}_{GSM})_{|\widehat{\mathcal{S}}}$}.

7. Find the size of the common support: $|\widehat{\mathcal{S}}|$.

8. Iteration begins.

   do while $(i \leq (\widehat{k} - |\widehat{\mathcal{S}}|))$

   $\mathbf{z} = \boldsymbol{\Phi}_{GSM}^T \mathbf{r}$.

   Select the index of the largest component (in terms of magnitude) in the vector $\mathbf{z}$.

   Update the support estimate $\widehat{\mathcal{S}}$ by adding the selected index to it.

   $\mathbf{r} = \mathbf{y}_{GSM} - (\boldsymbol{\Phi}_{GSM})_{|\widehat{\mathcal{S}}} (\boldsymbol{\Phi}_{GSM})_{|\widehat{\mathcal{S}}}^\dagger \mathbf{y}_{GSM}$.

   $i = i + 1$.

   end

9. Iteration ends.

10. The recovered signal $\widehat{\mathbf{x}} = (\boldsymbol{\Phi}_{GSM})_{|\widehat{\mathcal{S}}}^\dagger \mathbf{y}_{GSM}$.

Output: The recovered compressible signal: $\widehat{\mathbf{x}}$.
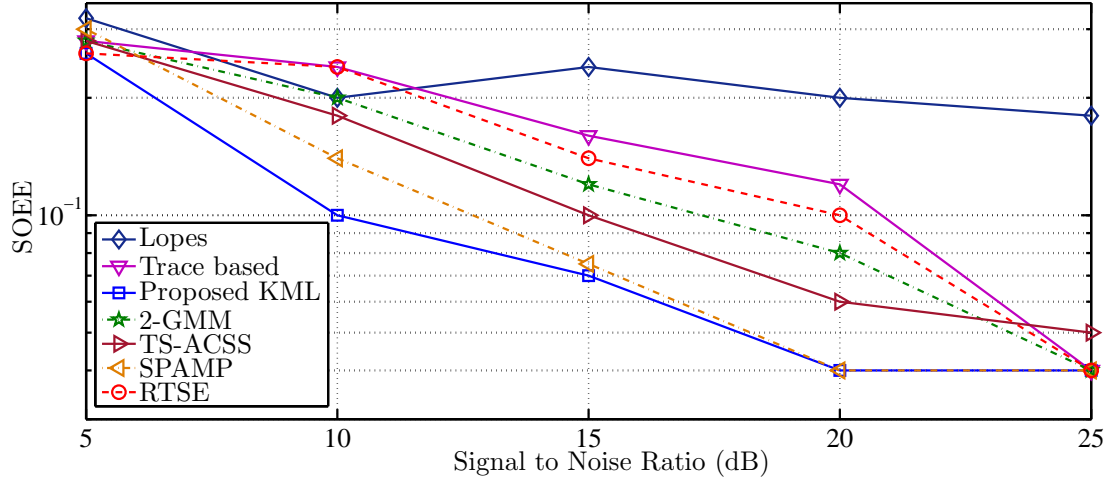
Figure 5.6: SOEE performance comparison for different SNR values.

In the simulation, an $N = 2500-$dimensional synthetic compressible signal was generated, whose sparsity order k was kept constant at $k = 250$. At every time step, the number of BSM measurements taken was $\lfloor \frac{1}{1-\lambda} \rfloor \approx 157$ for an optimal value of $\lambda = 0.9936$. For the Lopes method, $2k$ Cauchy sensing matrix-based measurements are obtained in every time step to compute the ℓ_1 norm for SOE. For eigenvalue- and trace-based methods, $2k$ GSM-based measurements were obtained. The total number of time steps considered is $k/2$. Thus, the total number of measurements obtained by the KML method is $78k$ compared with k^2 measurements obtained by other methods. Throughout the simulation, the support set remained the same, whereas the amplitudes of the significant coefficients varied according to normal distribution. The estimated sparsity order in every step is averaged for the Lopes, 2-GMM, TS-ACSS, SPAMP, and proposed KML methods. For the eigenvalue- and trace-based methods, only a single estimate is available after k^2 measurements. For the SPAMP method, the weak matching parameter is chosen as 0.5, and the estimation factor is kept at 0.2 for better results.

The performance was evaluated in terms of SOEE for different SNR values, as shown in Figure 5.6. The simulation results show that the KML method has better SOEE performance than the other methods. It is observed that the performance of the Lopes method is inferior and substantially invariant to the SNR owing to the use of the random Cauchy sensing matrix for which the variance is infinite. The 2-GMM SOE method requires knowledge of the energy of the significant coefficients for constructing the sparse Gaussian matrix, which is seldom known a priori. It also has an Expectation Maximization (EM) algorithm that adds to the complexity. The eigenvalue- and trace-based methods require at

80

least k^2 measurements, which are very expensive compared to other existing methods. The TS-ACSS performs a two-step SOE in which the first step performs coarse SOE, whose performance deteriorates with larger sparsity order values, and the second step refines the coarse SOE with the help of signal recovery, which is a time-consuming process. In addition, the first step was accurate only with additional measurements. Compared with other existing methods, the proposed KML method has the advantage of requiring three times fewer measurements with better performance.

5.3.2 NRE performance comparison using synthetic signals

The NRE performance of the proposed method (Composite sensing-KML) is compared with (i) other similar BSM methods for different sparsity order values, and (ii) 2-GMM [67], DBBD-Kronecker [41, 42], AS-SaMP [39], OAMP [40], and traditional GSM methods for different SNR values.

5.3.2.1 Performance comparison with other BSMs for different sparsity order values

A set of synthetic compressible signals of dimension $N = 5000$ with various sparsity order values $k = 50$ to $k = 400$ was generated. The generated signals are acquired and recovered using (i) the proposed composite sensing-KML-BAOMP method, (ii) random sparse BSM sensing followed by the BAOMP method, and (iii) DBBD matrix sensing [41] followed by modified Kronecker-based CS recovery [42]. For a given sparsity order k of a compressible signal, the proposed composite sensing matrix is designed using Equations (5.3) and (5.4). The random sparse BSM was designed with the same λ value as that of the proposed deterministic BSM. However, the ones in each row are randomly distributed. Both the random sparse BSM and DBBD methods obtain fixed $M = 1800$ measurements for $k < 300$ and fixed $M = 2500$ measurements for $300 \leq k \leq 400$, which is greater than the number of measurements obtained by composite sensing for a given k. The NRE performance for the 10 dB SNR settings is shown in Figure 5.7.

The simulation results show that the proposed composite sensing matrix outperforms both the random sparse BSM and DBBD sensing matrices with fewer measurements for all given sparsity order values. The probability of missing a significant component during acquisition was higher for random sparse BSM, resulting in poor NRE performance. Similarly, the high probability of incorrect support selection during recovery results in inferior NRE performance for the DBBD matrix-based method. The better performance of the composite sensing matrix is due to the use of GSM, which has better RIP property.
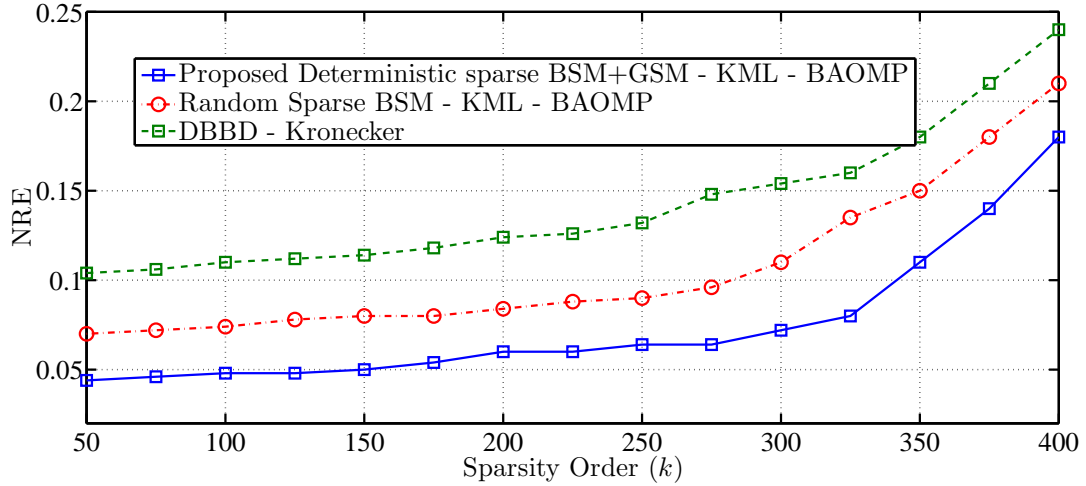
Figure 5.7: NRE performance comparison for different sparsity order values.

When $k < 300$, a smaller variance in KML results in minimal and invariant NRE. For the random sparse BSM and DBBD methods, the NRE remains invariant to k as the number of measurements $M = 1800$ is adequate for $k < 300$. When $k \geq 300$, the NRE performance degrades as the variance of the KML increases for the composite sensing matrix. However, even if $k \geq 300$, the NRE performance of KML is better than that of other BSM-based methods. For $k \geq 300$, the sensing matrix becomes too sparse for the random sparse BSM and DBBD methods, resulting in degraded NRE performance, and they require more measurements ($M > 2500$) to improve the NRE performance.

5.3.2.2 Performance comparison: For different SNR values

A time-varying synthetic compressible signal of dimension N=2500 is simulated for different SNR settings. Throughout the simulation, the sparsity order k was kept constant at $k = 250$ with varying support and amplitude. After acquisition and recovery using different CS methods, the NRE performance was compared.

Simulation results show the (i) improved performance of the proposed KML-based SOE followed by BAOMP-based recovery compared to the 2-GMM-based SOE method followed by OMP-based recovery, AS-SaMP-based recovery, OAMP-based recovery, and DBBD-Kronecker-based CS recovery, and (ii) comparable performance with GSM followed by basis pursuit, especially for low SNR conditions, as shown in Figure 5.8. The high probability of support estimation errors in the DBBD-Kronecker method results in poor NRE performance. The inaccuracy in estimating the statistics of the signal for the 2-GMM method affects NRE performance. The AS-SaMP and OAMP methods were fed
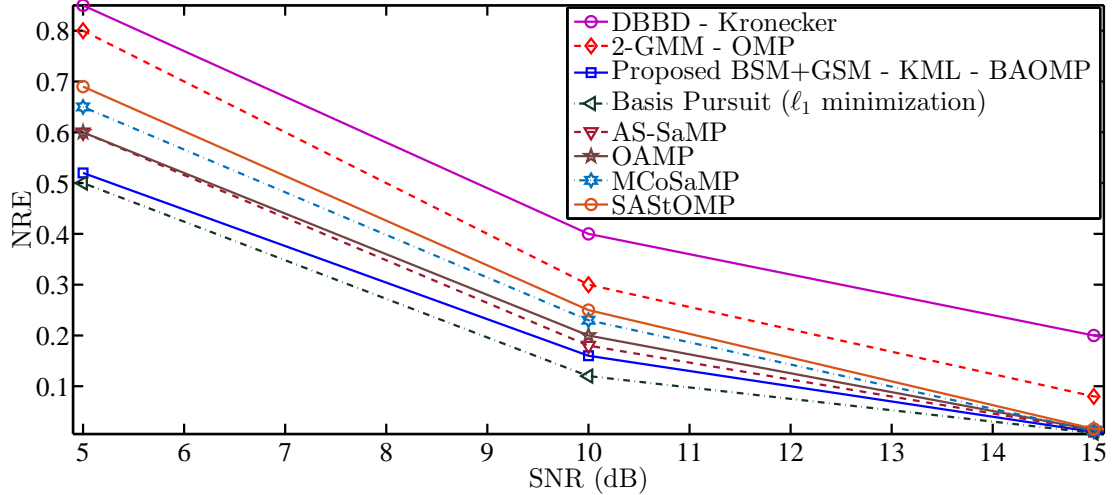
Figure 5.8: NRE performance comparison for different SNR values.

with the optimal parameters for the given sparsity order $k = 250$. Thus, their NRE performance is similar to that of the proposed method. However, tuning the parameters in real time is a challenging task for time-varying sparsity orders. Although basis pursuit has a better NRE performance, its higher computational complexity and longer recovery time are unsuitable for real-time recovery.

5.3.3 Performance comparison using recovery running time

A set of synthetic compressible signals of dimension $N = 2500$ with various sparsity order values $k = 50$ to $k = 400$ was generated, and the SNR was maintained at 10 dB. The generated signals were acquired and recovered using (i) the proposed composite sensing followed by the KML-BAOMP method, (ii) GSM sensing followed by the sparsity-aware OMP method, (iii) GSM sensing followed by AS-SaMP recovery, and (iv) GSM sensing followed by OAMP recovery. The running time for recovering the compressible signal is shown in Figure 5.9 and it is observed that KML-BAOMP is faster and outperforms all the existing methods for all sparsity order values.
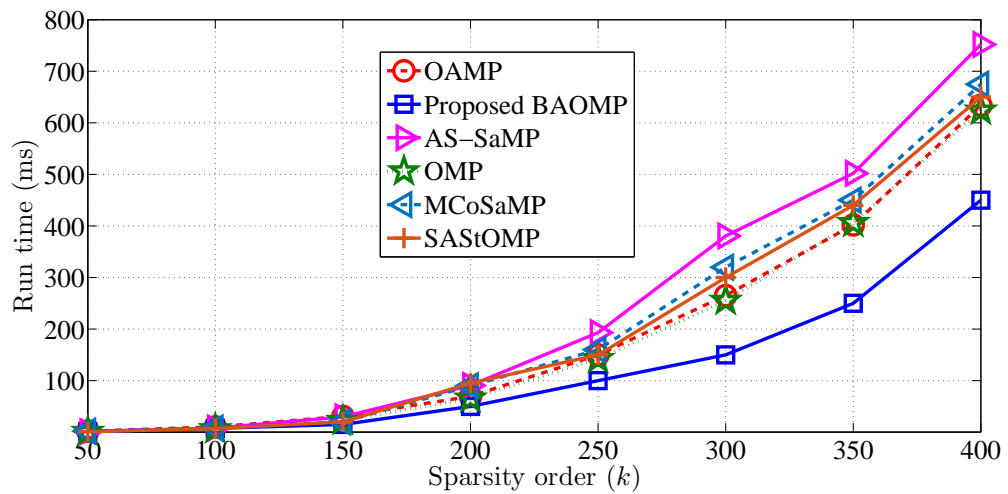
Figure 5.9: CS recovery run-time performance comparison for different sparsity order values. It is observed that the proposed BAOMP method is at least 1.5 times faster than the existing methods for a given sparsity order.

Chapter 6

Real-world applications

In this chapter, the proposed ML sparsity order estimators and BAOMP-based recovery algorithms are applied to real-world applications such as (i) acquisition and recovery of vibration signals, (ii) channel estimation, and (iii) recovery of ECG signals. Then, the performance of the proposed methods is compared with that of existing approaches.

6.1   Acquisition and recovery of vibration signals

Structural health monitoring (SHM) [68, 69] is generally performed to assess the integrity of structures in aerospace and related industries. SHM entails analysis of the time-frequency domain of the vibration signals that emanate from these structures for maintenance, safety, and reliability checks. For this purpose, a large set of vibration signals is acquired for identifying the excited frequencies during different operating conditions of the concerned structure. It would be preferable if these vibration signals could be compressed without sacrificing information in order to reduce data storage, processing power, and execution time.

CS-based SHM was recently investigated in [70, 71, 72, 73, 74, 75] for compressing vibration and its related signals, and it has been shown that vibration signals are acquired efficiently using the CS method. Bao et al. [70] and [71] studied the CS technique to acquire acceleration signals to derive the vibration characteristics of a bridge. The reconstruction of missing wireless SHM measurements was dealt with using CS recovery in [73] assuming that the sparsity order is known beforehand. A study on shock data acquisition using CS was presented in [74], which showed that shock signals are better compressible on wavelet bases. Recently, the CS of vibration signals from wireless sensor nodes exploiting group sparseness was analyzed in [75], which provided promising results for applying CS to vibration signals. All available CS-based SHM methods assume that the sparsity order is

known beforehand and does not vary with time. However, these assumptions were invalid. These circumstances indicate that SOE has become a vital problem for the better compression and recovery of vibration signals. Hence, the proposed BSOE and GSOE-based SOE estimator methods are applied to real-world vibration signals, and their performance is evaluated and compared with existing SOE and recovery methods.

6.1.1 Compressibility analysis of vibration signals

Real-world vibration signals are available from Mide Technologies [59] and are analyzed for performance evaluation. In this section, the vibration signals acquired from an aircraft and semi-trailer truck are analyzed.

6.1.1.1 Case 1: Vibration signal acquired from an aircraft during its climb

A snapshot of the smoothly varying vibration signal acquired using an accelerometer mounted on the surface of an aircraft during its climb is shown in Figure 6.1(a). The signal was sampled at 2500 s, and its time-frequency spectrogram is shown in Figure 6.1(b). The spectrogram reveals that the vibration signal is compressible in the frequency domain and that the number of significant frequency coefficients varies with time.

As compressibility depends on the number of significant frequency coefficients, it is essential to know $a\ priori$ that which orthonormal representation bases preserve the energy of the vibration signal in fewer significant frequency coefficients and provide the minimum sparsity order. Hence, the vibration signal is analyzed using the DCT, DFT, and discrete wavelet transform (DWT). To avoid spectral leakage, a Hamming window was applied to the vibration signal during DFT analysis. In every analysis, a segment of $N = 2500$ samples was transformed into 2500 frequency or scaling coefficients. The absolute values of these coefficients are sorted in descending order as functions of the new indices for each transform. The sorted values of one of the analysis segments are shown in Figure 6.2. The sorted values obey the power-law decay property of compressible signals. For a smoothly varying signal, DCT and DFT result in a faster decay of the sorted values and thus have better compressibility than DWTs. Because the original vibration signal is contaminated by sensor noise, all the frequency coefficients have a non-zero magnitude. This necessitates the identification of significant coefficients for the computation of sparsity order. Hence, the index at which the cumulative sum of the squared sorted values meets a threshold value is measured as sparsity order k. The threshold value is the energy of the uncontaminated vibration signal, and is assumed to be 95% of the energy of the contaminated signal. It is
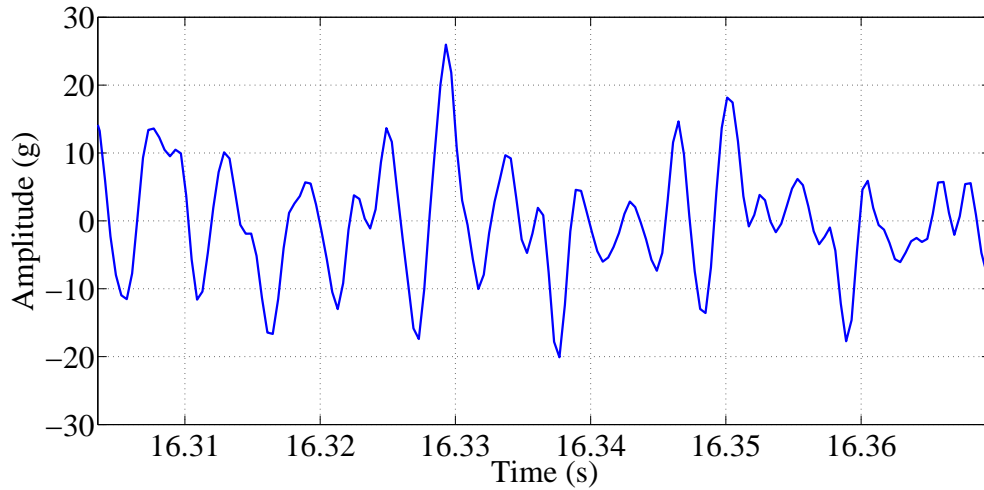
**(a)** A segment of vibration signal sampled at 2500 Hz
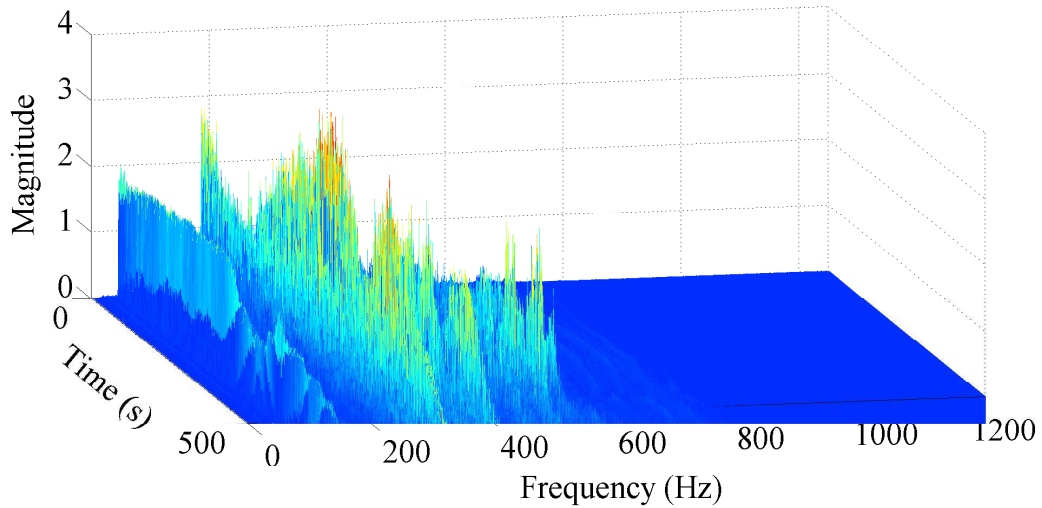
(b) Spectrogram of the entire vibration signal

Figure 6.1: Vibration signal measured outside of an aircraft during its climb and its Time-Frequency spectrogram. The spectrogram shows that only a few coefficients are significant having distinguishable magnitudes.

evident from Figure 6.3 that DCT provides better compaction than DFT, and thus provides a better sparse representation for a given NRE of 5%.

6.1.1.2   Case 2: Vibration signal acquired from a semi-trailer truck during its transit

An entire vibration signal with transients due to shock acquired from a semi-trailer truck during its transit and its frequency spectrum are shown in Figures 6.4(a) and (b), respectively. As the signal is having transients, DWT and DCT provide a lower sparsity order.
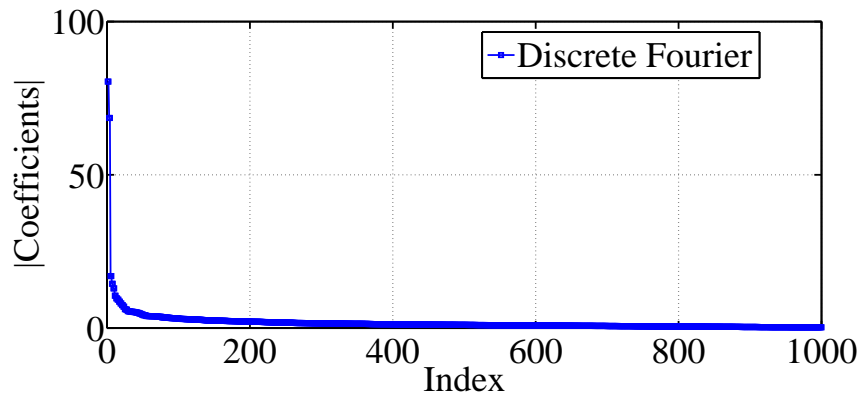
(a) Discrete Fourier bases
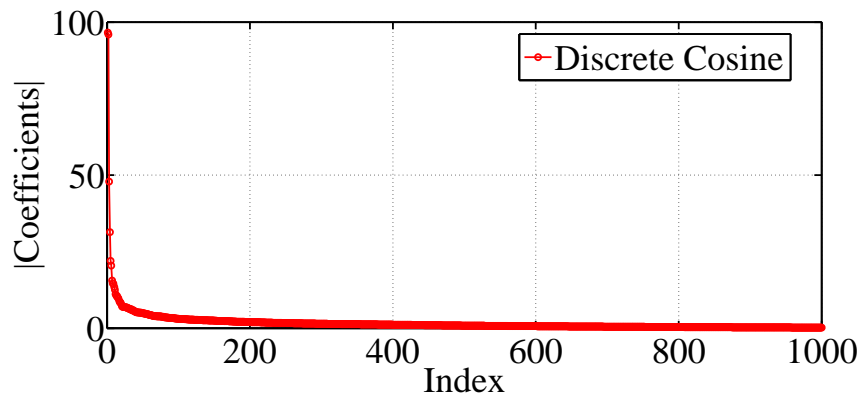
(b) Discrete Cosine bases
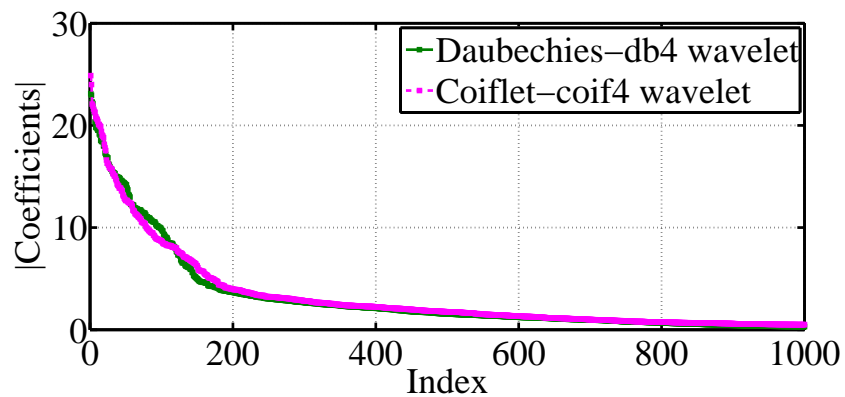
(c) Discrete wavelet bases

Figure 6.2: Illustration of compressibility using the power-decay property on applying different bases. The absolute values of 2500 coefficients are sorted and the largest 1000 values are shown. It is seen that discrete wavelets have slower decay compared to other bases and thus have poor compressibility.
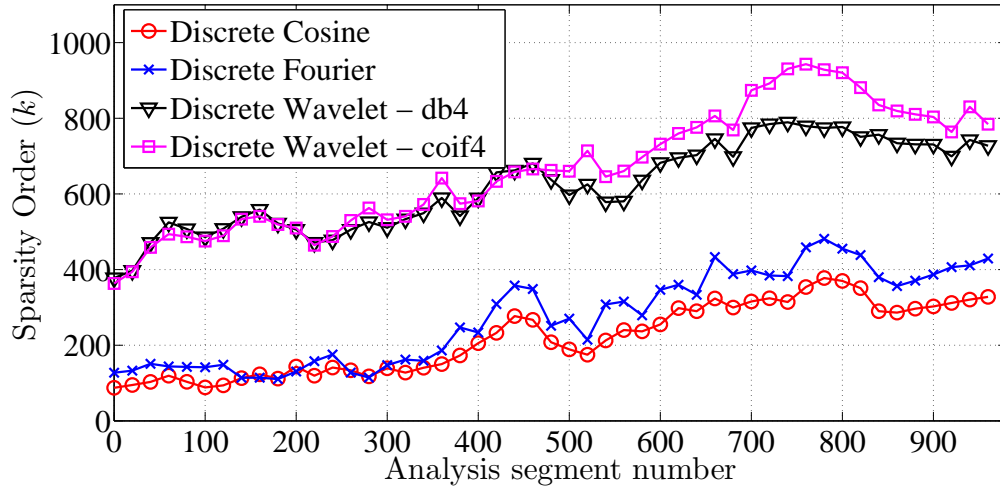
Figure 6.3: Time-varying sparsity order on different bases for every analysis segment of 2500 samples. It is evident that discrete cosine results in sparsity order values lesser than Fourier and wavelet counterparts throughout the analysis.

6.1.1.3 Case 3: Vibration signal acquired from the aircraft during its cruise

A snapshot of the vibration signal acquired from the aircraft during its cruise and its spectrum over the entire duration are shown in Figures 6.5(a) and (b), respectively. During the cruise, the vibration is benign, the dynamic range of the vibration signal is very low, and amplitude variations are not smooth. This causes the energy of the signal to spread evenly among all frequency components, and the signal is not compressible using any transform.

The performance metrics, such as the average value of the sparsity order and the compression ratio (CR) obtained for all the analysis segments using each transform function for all the vibration signals discussed above, are listed in Tables 6.1 and 6.2, respectively. The analysis shows that (i) DCT provides better CR for smoothly varying vibration signals, and (ii) DWTs provide better CR when vibration signals have transients owing to shock. In addition, it was observed that benign vibration signals with a lower dynamic range were not compressible. Considering the above, the DCT is considered here as the sparse representation matrix $\mathbf{\Psi}$ for analyzing the vibration signal.

6.1.2 Markov birth-death modeling of time-varying sparsity order of vibration signals

The sparsity order of vibration signals vary according to the Markov birth-death process model as shown in Fig 6.6. It is observed that sparsity order varies rapidly.
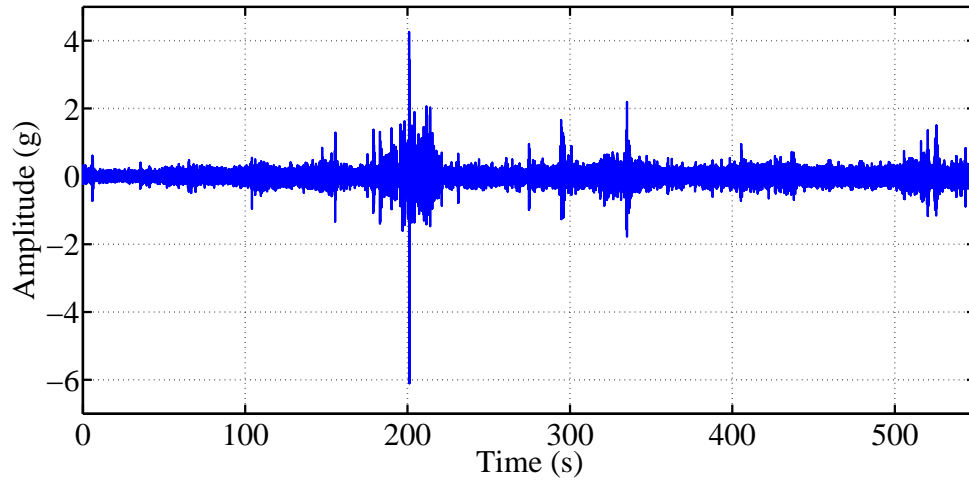
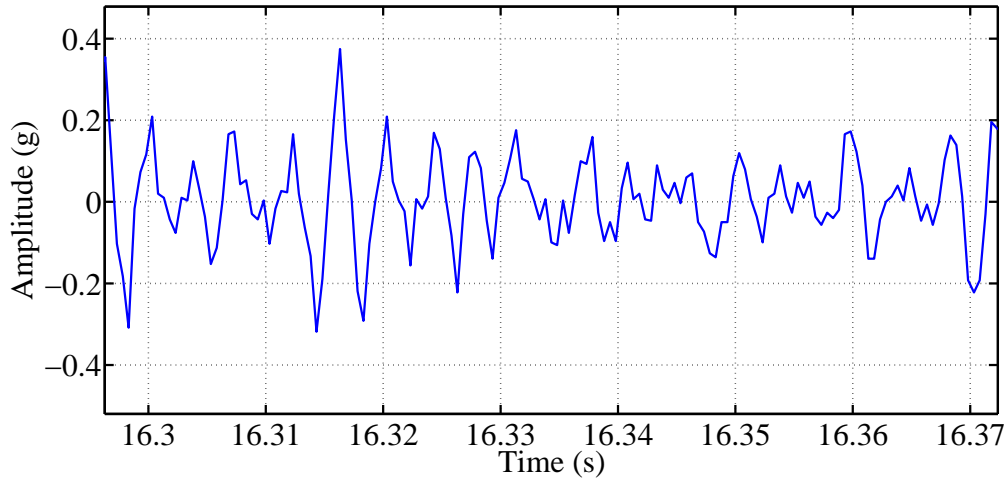**(a)** Entire segment of vibration signal sampled at 5000 Hz
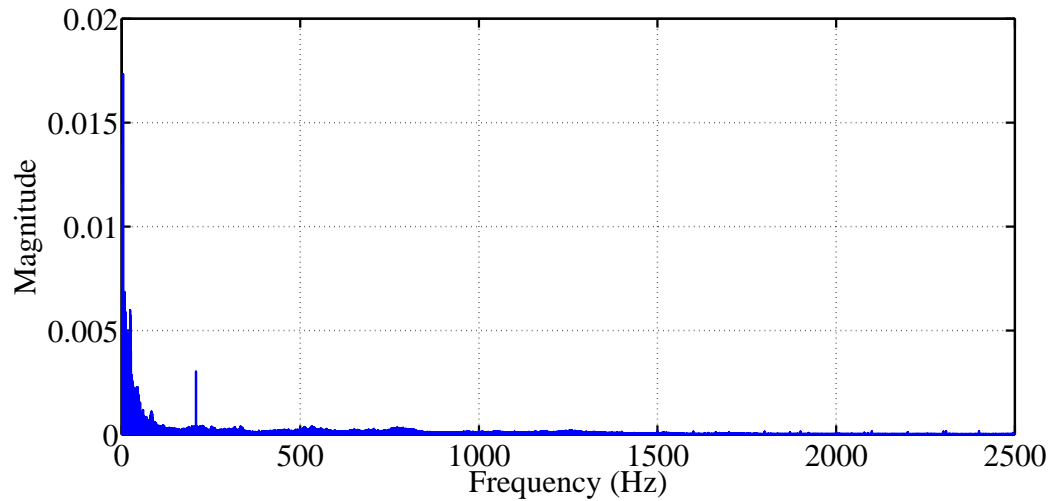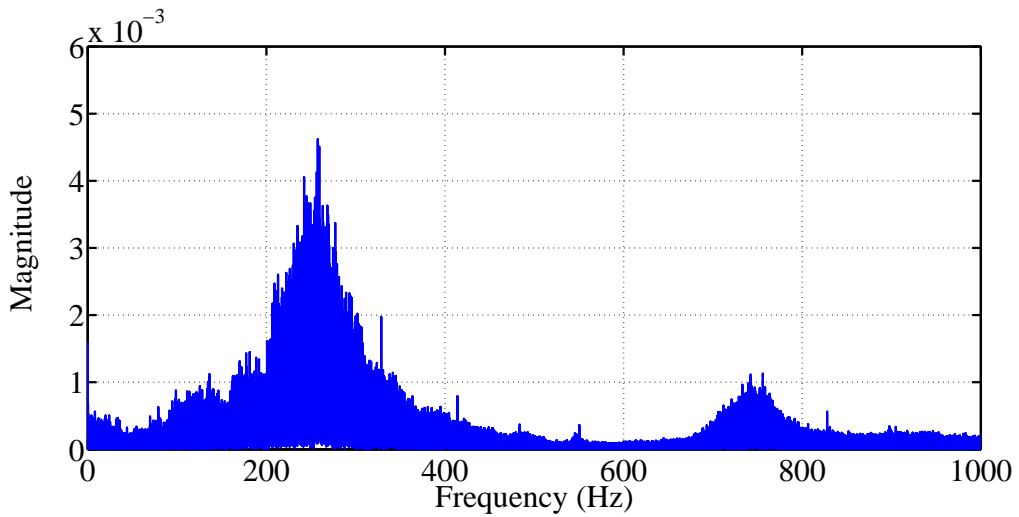

(b) DFT spectrum of entire vibration signal

Figure 6.4: Vibration signal measured from a semi-trailer truck during transit and its DFT spectrum.

Vibration signal	Sampling rate	Average Sparsity Order			
(acquired from)	(sps)	DFT	DCT	DWT-db4	DWT-coif4
aircraft during climb	2500	373	298	524	649
semi-trailer truck	5000	428	354	368	356
aircraft during cruise	2000	828	708	1343	1459

Table 6.1: Average sparsity order analysis using different transform

(a) A segment of vibration signal sampled at 2000 Hz

(b) DFT spectrum of entire vibration signal

Figure 6.5: Vibration signal measured from an aircraft during the cruise and its DFT spectrum.

Vibration signal	Sampling rate	Average compression ratio			
(acquired from)	(sps)	DFT	DCT	DWT-db4	DWT-coif4
aircraft during climb	2500	3.35	4.20	2.38	1.99
semi-trailer truck	5000	5.84	7.06	6.79	7.02
aircraft during cruise	2000	1.20	1.41	0.74	0.68

Table 6.2: Average compression ratio analysis using different transform
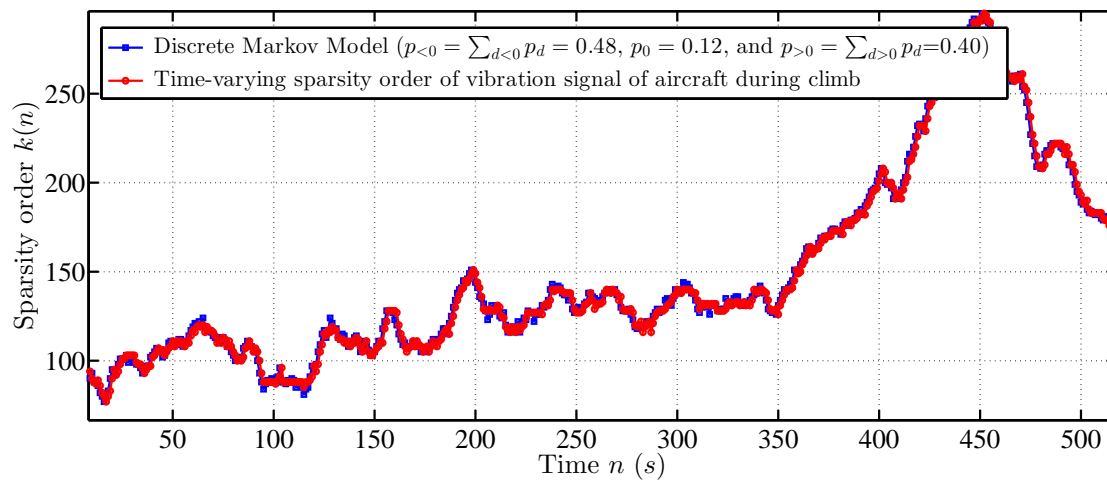
(a) Time-varying sparsity order of aircraft's vibration signal during its climb
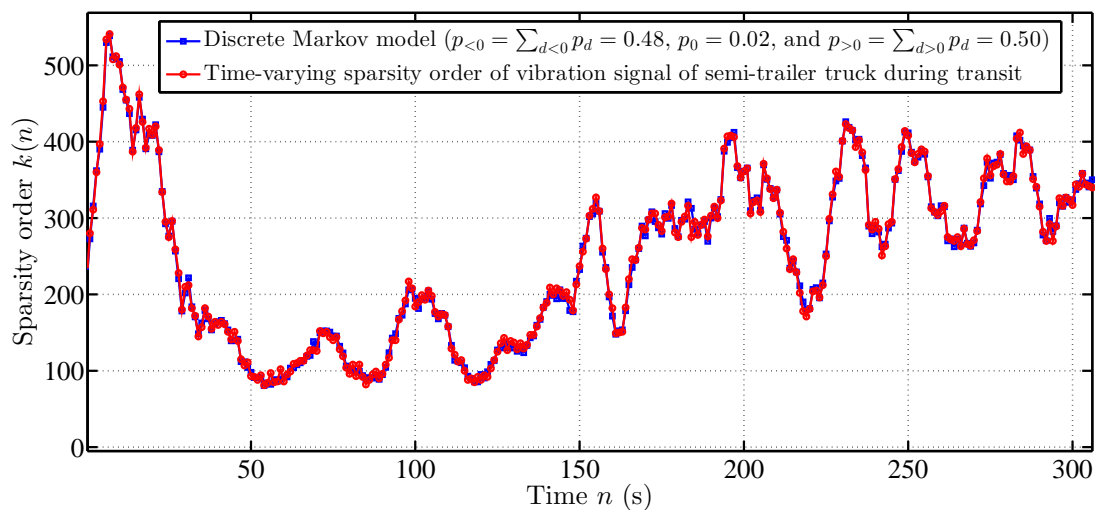
(b) Time-varying sparsity order of semi-trailer truck's vibration signal during transit

Figure 6.6: Markov birth-death modeling of time-varying sparsity order of vibration signals

6.1.3 Performance evaluation of proposed estimators and recovery method

The vibration signals acquired from a semi-trailer truck and aircraft were analyzed using the methods listed in Table 6.3. The 2-GMM method is provided with the knowledge of the energy of significant coefficients for the construction of a sparse GSM. The estimated sparsity order is the input for the CS recovery of the vibration signals for the proposed BAOMP method and the OMP algorithm for the 2-GMM method. In the traditional CS method, the OMP algorithm is provided with the original sparsity order value. For the SPAMP method, the weak matching parameter is chosen as 0.5 and the estimation factor is kept as 0.2 for better results. For the SBL method the noise variance and the thresholding parameters are kept as 0.02 and 10^{-4}, respectively for the optimal performance.

| Methods	Acquisition method	Recovery method
Proposed method	BSM+GSM+KML	KML+BAOMP
Traditional CS	GSM satisfying RIP	Basis Pursuit
2-GMM	sparse GSM	EM+OMP
DBBD-Kronecker	DBBD	Kronecker
SPAMP	GSM satisfying RIP	SAMP
SBL	GSM satisfying RIP	Bayesian learning
KFCS	GSM satisfying RIP	Dantzig selector + KF
DCT compression	DCT	Inverse DCT

Table 6.3: CS acquisition and recovery methods used for NRE comparison.

For the vibration signal measured from the aircraft, each analysis segment contained 2500 samples. For the vibration signal measured from a semi-trailer truck, each analysis segment had 5000 samples. The analysis segment was normalized to have unit energy. For KML, 2-GMM, and SPAMP methods, the sparsity order is estimated for each analysis segment. Based on the estimated sparsity order $\widehat{k}(n)$, measurements were obtained for recovery at the $(n + 1)^{th}$ time step. Using these measurements, the vibration signal was reconstructed using the respective recovery methods. Then, the NRE performance measure was compared with the traditional random GSM-based CS method, DBBD-Kronecker-based CS method, SBL method, KFCS method, and DCT-based compression method. For the traditional CS, DBBD, SBL, and KFCS methods, $M = 4k_{\max}$ measurements are obtained. The results of the NRE for both vibration signals for every 20 analysis segments are plotted in Figures 6.7(a) and (b). A snapshot of the reconstructed vibration signal obtained using the KML method followed by the BAOMP method for the vibration signal acquired outside the aircraft during its climb is shown in Figure 6.8 along with the original signal.

During recovery, the BAOMP method denoises the insignificant coefficients. Hence, the reconstructed signal is smooth over time compared with the original signal.

Although the proposed KML-based CS method has a slight degradation in NRE compared with the classical DCT method, the hardware complexity of acquiring a compressible signal is $\mathcal{O}(MN)$ (owing to the $M \times N$ sensing matrix), which is less than that of the DCT method's $\mathcal{O}(N^2)$ (owing to the $N \times N$ DCT matrix). Thus, it requires fewer hardware resources, storage, and power for the acquisition of time-varying compressible signals with a slight tolerable degradation in the recovery performance compared to the DCT method. It provides a good CR as it obtains a minimal number of measurements based on the estimated sparsity order compared to other CS methods. The existing SOE methods are not optimal during acquisition and result in less CR as they either (i) obtain a fixed number of measurements based on the conservative assumption of having maximum sparsity order $k_{\max}$ or (ii) obtain an additional set of measurements for the SOE.

It can be observed from Figures 6.7(a) and (b), the CS-based compression methods resulted in a slightly higher NRE than the traditional DCT-based compression method. The reasons for this are described in the following. The DCT method knows the support and amplitude of the significant coefficients among the N coefficients and approximates the insignificant coefficients to zeros. Hence, its NRE is the energy of insignificant coefficients, which is 0.05. The CS recovery methods do not have *a priori* knowledge of the significant coefficients, and they estimate the support and amplitude from the available $M < N$ measurements. There are some of the weakest significant DCT coefficients whose amplitudes are very close to the threshold value, distinguishing significant and insignificant coefficients. The CS recovery algorithms detect stronger information-bearing significant coefficients well above the threshold without any failure. However, sometimes they may not select the weakest significant coefficients very near the threshold, as the CS sensing matrices are not perfect orthonormal matrices for identifying the support of such coefficients. This effect is illustrated in Figure 6.9 which represents the DCT spectrum plot of one of the analysis segments of the vibration signal measured outside the aircraft. Here, the DCT coefficients before and after the sparse approximations and the estimated DCT coefficients using the proposed KML-BAOMP-based CS recovery are shown. It was observed that the DCT coefficients with indices from 83 to 87 and 100 to 102 had magnitudes slightly below and above the threshold, respectively, and were not detected during CS recovery, contributing to the slightly higher NRE compared to the DCT-based compression and recovery. However, among the CS methods, the proposed KML-BAOMP method performs better than the 2-GMM, DBBD-Kronecker, SPAMP, and AS-SaMP methods and is
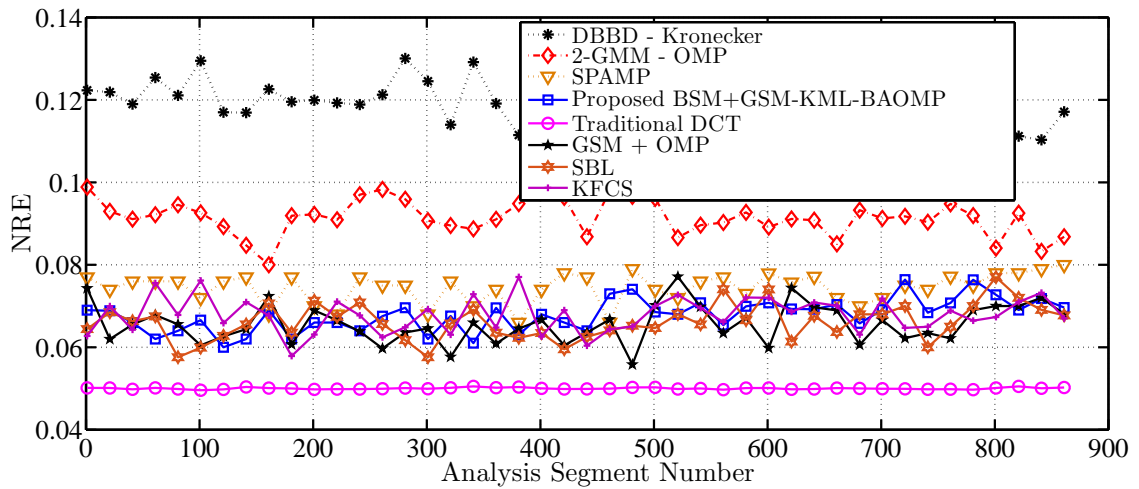
(a) NRE performance comparison on compressing vibration signal measured outside an aircraft during its climb
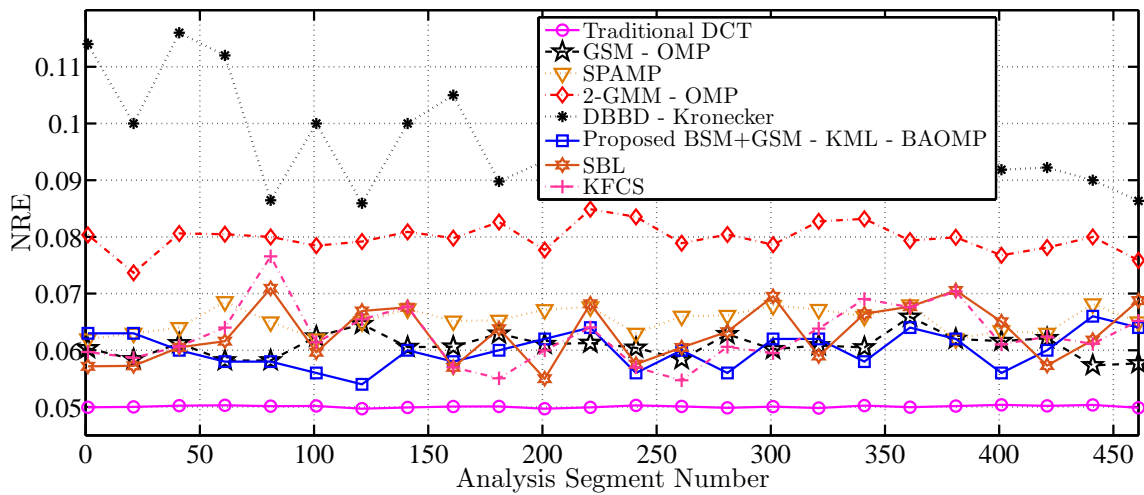
(b) NRE performance comparison on compressing vibration signal measured on a semi-trailer truck

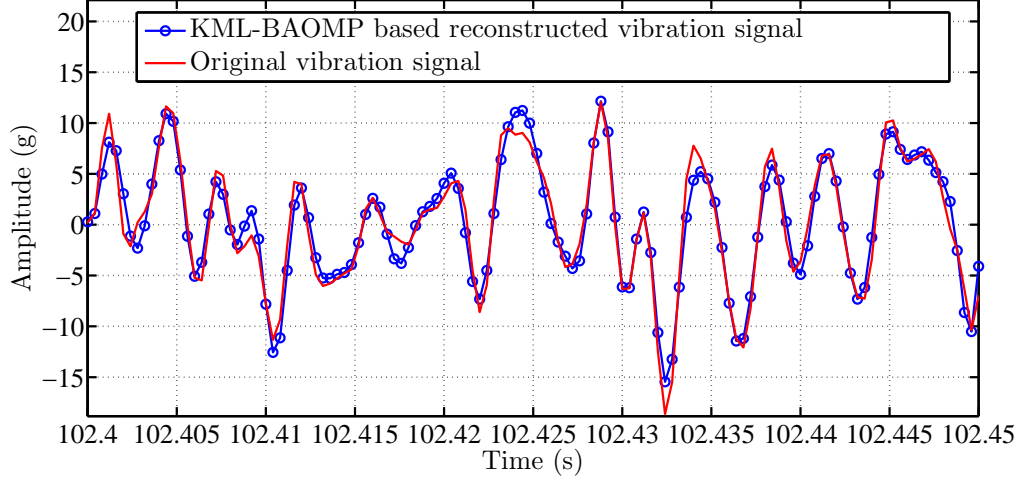**Figure 6.7:** NRE performance comparison.

Figure 6.8: Recovery performance of the KML method followed by the BAOMP method on reconstructing the vibration signal acquired on the skin of an aircraft during its climb.

comparable to the traditional GSM-based method, SBL and KFCS methods.

6.2 Estimation of channel impulse response using MLS estimator

Consider the problem of estimating the wireless channel impulse response (CIR) from the model,

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \boldsymbol{\vartheta},$$

where $\mathbf{y}, \mathbf{X}, \mathbf{h}$ and $\boldsymbol{\vartheta}$ are the received measurements, Toeplitz matrix of the known pilots, CIR, and measurement noise, respectively. The above model becomes the CS acquisition model as the CIR $\mathbf{h}$ is sparse and $\mathbf{X}$ is carefully chosen as a sensing matrix. As the minimum number of measurements obtained is dependent on the sparsity order of CIR, the sparsity order estimation is essential for reducing the number of pilots. Thus, the performance of the proposed sparsity order estimator in terms of estimation accuracy was evaluated using a real-world measured CIR obtained from the CRAWDAD dataset [76]. CIR was recorded using a 44-node wireless network. Multiple CIR measurements were performed for each of the $44 \times 43 = 1892$ pairwise links between the nodes in a standard office area by moving the transmitter and receiver between the nodes. A sample real and imaginary realization of the CIR between nodes 1 and 2 is shown in Figure 6.11(a). The time-varying sparsity order is shown in Figure 6.11(b) by stacking several realizations of
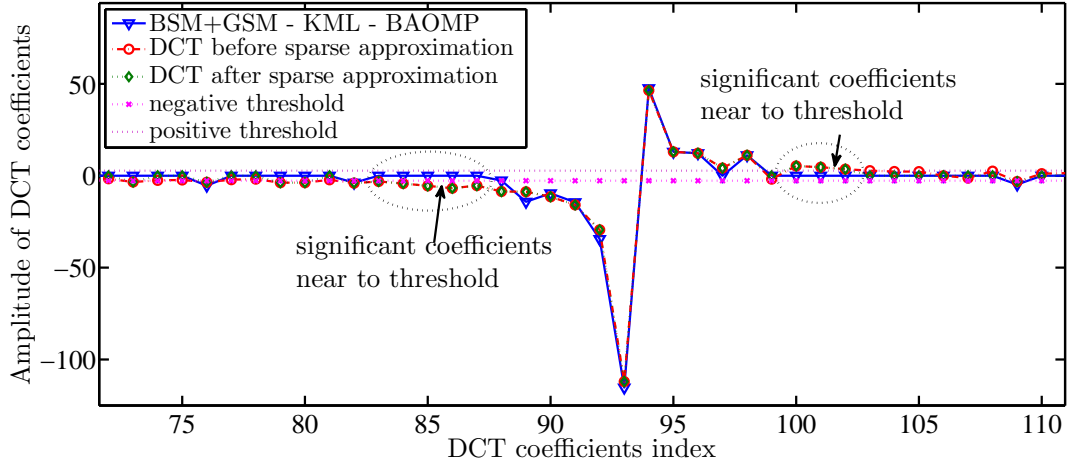
Figure 6.9: DCT spectrum plot comparing the recovery performances of DCT and KML-BAOMP methods. The components with magnitudes very near the threshold may or may not get recovered by the CS methods resulting that DCT based recovery method performing better than any CS methods.

CIR measurements. It is observed that sparsity order variation follows a discrete Markov process with transition probabilities $\sum_{d>0} p_d = 0.16$, $\sum_{d<0} p_d = 0.20$, and $p_0 = 0.64$ resulting in fast sparsity order variation as shown in Figure 6.10. As the sparsity order varies fast, the ML estimates of transition probabilities are computed as equally distributed i.e., $\sum_{d>0} \hat{p}_d = 0.33$, $\sum_{d<0} \hat{p}_d = 0.33$, and $\hat{p}_0 = 0.34$. The MLS estimation was performed using the Viterbi algorithm. The unfiltered ML and refined MLS estimates are presented in Figure 6.11(b), along with the actual sparsity order, and it is verified that the MLS estimation refines the unfiltered ML estimates.

6.3 Using real-world electrocardiogram signal

Real-world electrocardiogram (ECG) signals are available in Physiobank ATM [77]. An ECG signal sampled at 720 sps for a duration of 60 s was analyzed to evaluate the performance of the proposed ML-MLS algorithm. The ECG signal was divided into 60 segments, and each segment had a duration of 1 s. Each segment was analyzed using a DCT matrix, and it was observed that the ECG signal was sparse. The sparsity order varies rapidly in time, following a discrete Markov model with transition probabilities $\sum_{d>0} p_d = 0.17$, $\sum_{d<0} p_d = 0.17$, and $p_0 = 0.66$ as shown in Figure 6.12. The maximum variation in the sparsity order is ± 5 at a given time step. The transition probabilities are estimated
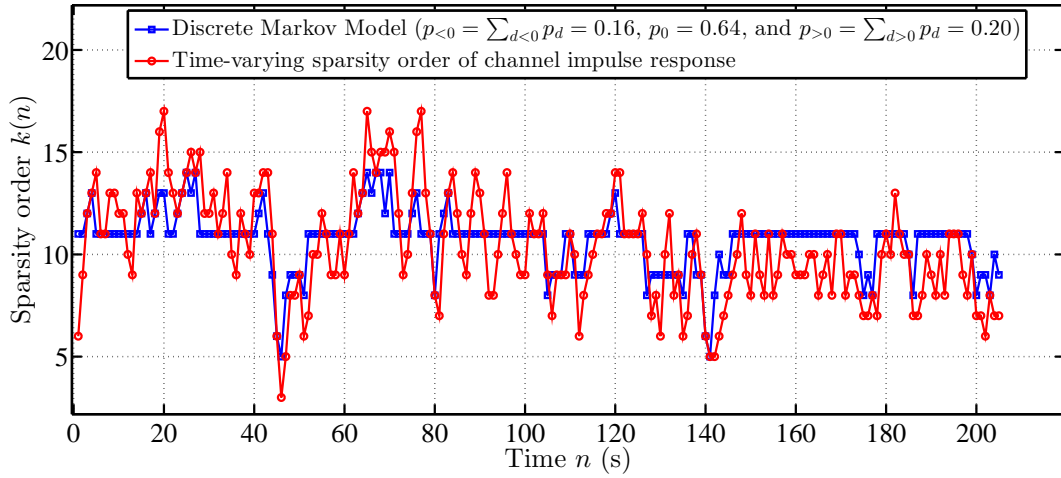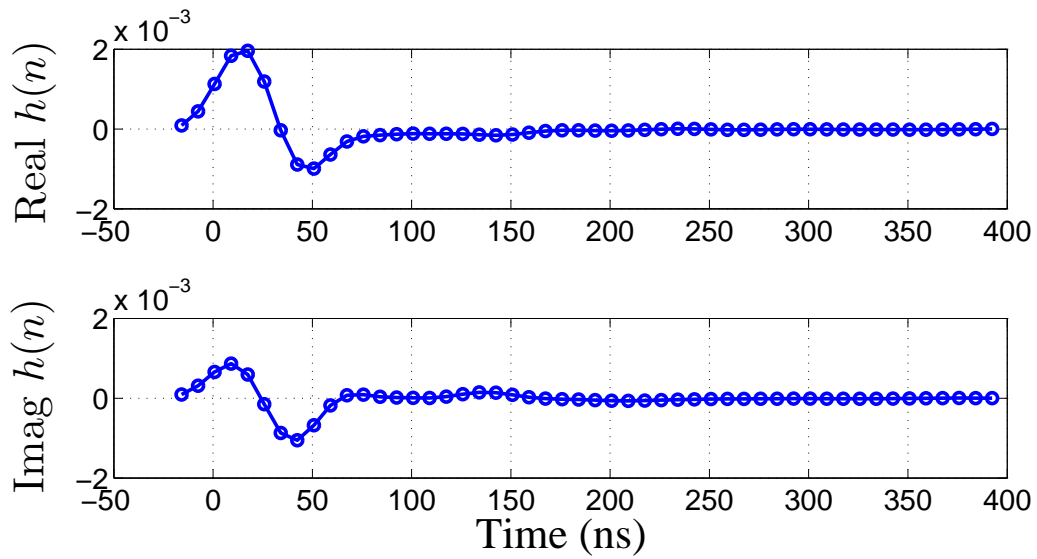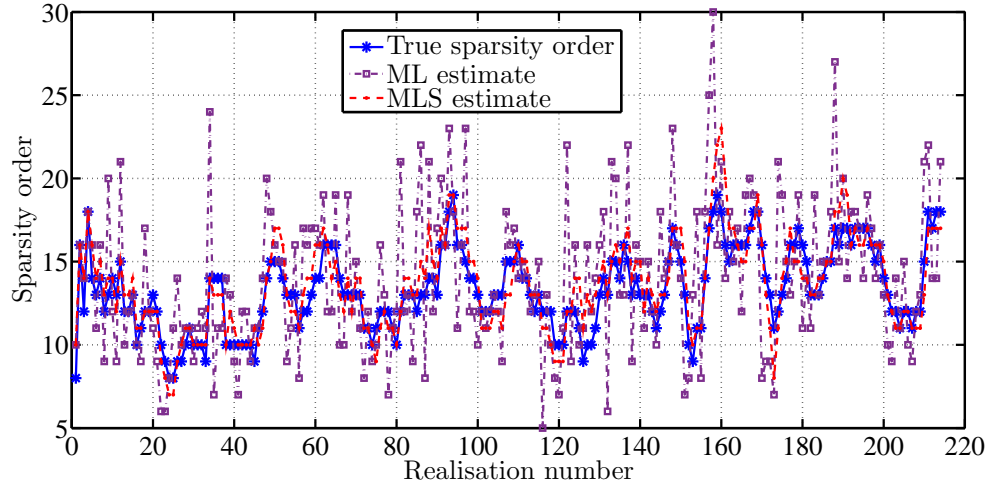
Figure 6.10: Time-varying sparsity order of channel impulse response

as $\sum_{d>0} \hat{p}_d = 0.34$, $\sum_{d<0} \hat{p}_d = 0.20$, and $\hat{p}_0 = 0.46$ from the sequence of the filtered ML estimates. The time-varying sparsity order and its ML and MLS estimates are shown in Figure 6.13 and it can be seen that the MLS algorithm provides refined estimates of the sparsity order. Using the estimated sparsity order, the ECG signal is recovered using the OMP algorithm. The recovered signal matches the original ECG signal as shown in Figure 6.14. The recovered signal has NRE $\approx 5\%$ compared to the original signal.

(a) Channel impulse response measured between transmitter 1 and reciever 2

(b) Performance of ML and MLS estimation techniques

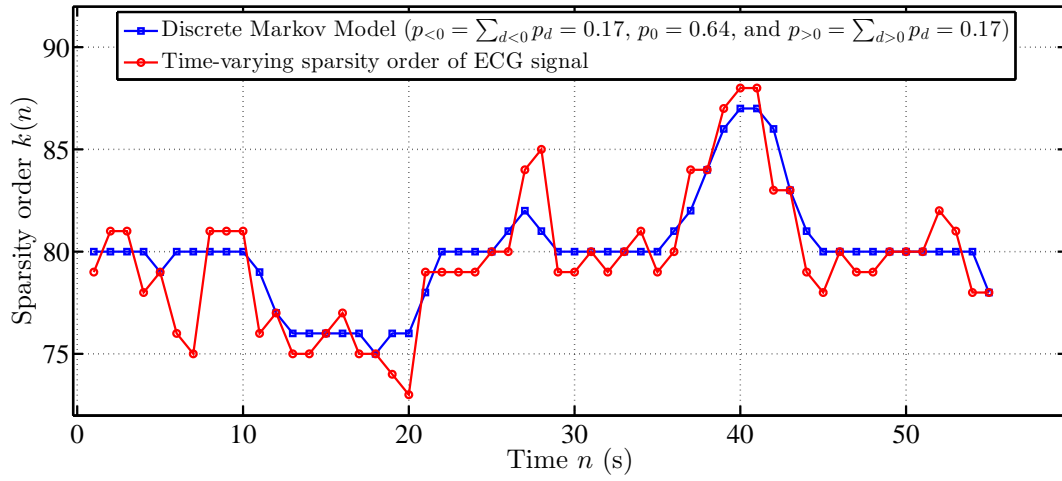**Figure 6.11:** Estimation of channel impulse response.

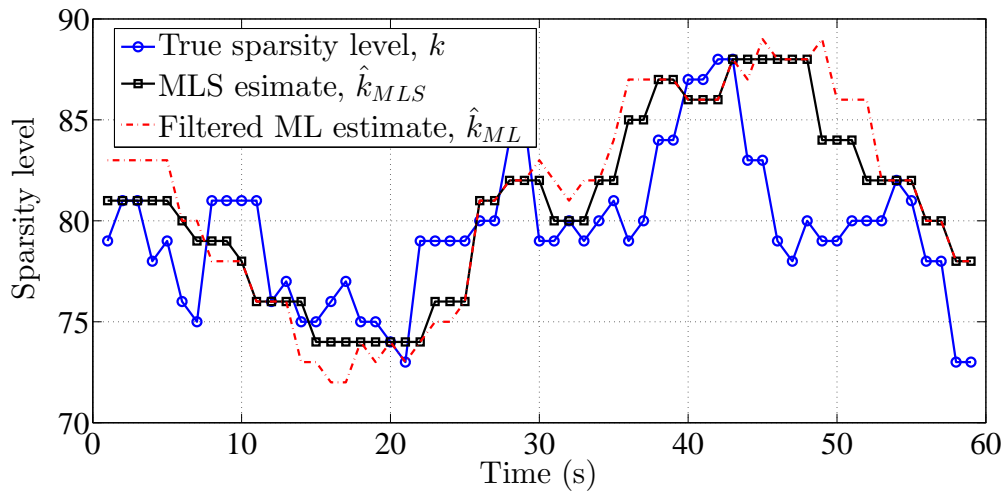**Figure 6.12:** Time-varying sparsity order of an ECG signal

Figure 6.13: Performance of ML and MLS estimation of sparsity order of ECG signal.
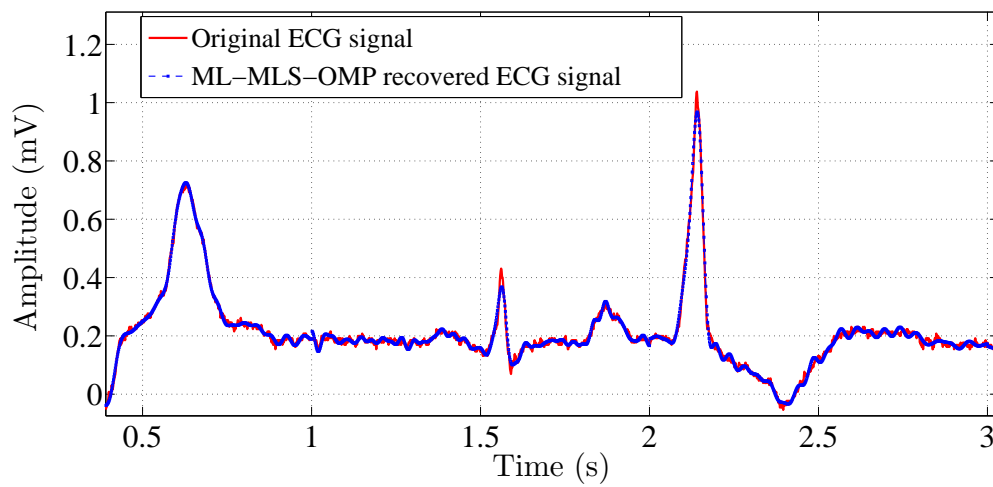
Figure 6.14: A segment of ML-MLS-OMP based recovered ECG signal compared with that of the original ECG signal.

Chapter 7

Complexity comparison

The performance of the ML-MLS estimation technique is compared in terms of computational complexity with the Lopes method [78], 2-GMM method [30], trace-based method [29], DCS-AMP method [26], and KFCS method [23] for a given M number of measurements, as follows.

The proposed estimator requires $M = M_{BSM} + M_{GSM}$ measurements per time step, where M_{BSM} measurements are used to estimate the mean energy of the supporting components and M_{GSM} measurements are used to estimate the sparsity order. The computational complexity for estimating the instantaneous sparsity order using m measurements was $\mathcal{O}(m)$.

At any time step, the Viterbi algorithm performs MLS estimation over a block of L length whose computational complexity is $\mathcal{O}(D_R^2 L)$, where $D_R = k_{max} - k_{min} + 1$ is the dynamic range of the sparsity order, k_{min} and k_{max} are the minimum and maximum sparsity orders, respectively, in the L length block. Thus, the overall computational complexity of the proposed ML-MLS method was $\mathcal{O}(M)+\mathcal{O}(D_R^2 L)$. The computational complexity of the other methods is listed in Table 7.1, where the remarks briefly explain their computational complexity. The proposed ML-MLS estimation technique demands an additional computational complexity of $\mathcal{O}(D_R^2 L)$ which is less than $\mathcal{O}(Mk^3)$ of the KFCS and $\mathcal{O}(MN)$ of the DCS-AMP methods. Except for the estimation of the mean energy of the supporting components, the ML-MLS technique does not involve iterations.

The computational complexity of the existing SOE and recovery methods was compared with the proposed KML and BAOMP methods. The computational complexity of the SOE method is based on the number of measurements and iterations. The KML method estimates the sparsity order k directly from $M_{BSM} < M$ measurements. The computational complexity of the proposed KML is dependent on (i) the probability p_0 computation, (ii) the iterations involved in the joint estimation of the statistics of significant coefficients

Method	Computational complexity	Remarks
Lopes	$\mathcal{O}(2M)$	M Cauchy and M Gaussian basis-based measurements
2-GMM	$\mathcal{O}(kM)$	k iteration per measurement for EM algorithm
Trace	$\mathcal{O}(M^3)$	Computation of Covariance of measurements
KF-CS	$\mathcal{O}(Mk^3)$	Complexity of Dantzig selector
DCS-AMP	$\mathcal{O}(MN)$	AMP's M iterative updation of N variables

Table 7.1: Computational Complexity Comparison

and sparsity order, and (iii) the Kalman filtering (KF) process. Probability p_0 is estimated by identifying the measurements from M_{BSM} measurements that are devoid of significant coefficients. This identification requires a computational complexity of $\mathcal{O}(M_{BSM})$. The joint estimation of probability q, statistics of significant coefficients, and sparsity order using the BSOE require $N(1 - \lambda)$ computations per iteration. The simulations show that a maximum of ten iterations are required for the convergence of the joint estimation. Because the number of iterations is fixed and independent of M_{BSM} measurements, the total computational complexity is dependent only on $N(1 - \lambda)$, and it is $\mathcal{O}(N(1 - \lambda))$ for the joint estimation step. The KF used for refining the BSOE estimate is scalar, and the computational complexity is $\mathcal{O}(1)$, which is less than the complexity of the BSOE. Thus, the overall computational complexity of KML is $\mathcal{O}(M_{BSM} + N(1 - \lambda))$.

Other existing SOE methods use all the M measurements for SOE. Given M measurements, the eigenvalue and trace-based methods require complex Matrix operations with a computational complexity $\mathcal{O}(M^3)$. The computational complexity of the EM algorithm in the 2-GMM method is $\mathcal{O}(kM)$. The Lopes method computes the median of the Cauchy sensed measurements and the mean of the energy of Gaussian sensed measurements with a computational complexity $\mathcal{O}(M \log M)$ to estimate the ℓ_1 and ℓ_2 norms for the SOE. Since $M_{BSM} + N(1 - \lambda) \ll M$, the proposed KML method is very advantageous for SOE in terms of less computational complexity compared to other existing methods.

The computational complexity for the OMP and SPAMP-based CS recovery methods is $\mathcal{O}(\widehat{k}MN)$, whereas for the proposed BAOMP method is $\mathcal{O}(\widehat{k} - |\widehat{\mathcal{S}}|MN)$. As $|\widehat{k} - \widehat{\mathcal{S}}| < \widehat{k}$, the BAOMP method requires fewer computations.

The run-time complexity of the proposed KML and BAOMP methods is evaluated and compared with other existing SOE methods for the real-world aircraft vibration signal anal-

ysis using a Windows 7 Operating System-based PC with a processor running at 3 GHz clock speed and 4 GB RAM. The values of average elapsed time for the SOE and recovery on analyzing each segment of the vibration signal are shown in Table 7.2 and Table 7.3, respectively.

SOE methods	Algorithms	Elapsed time
KML	BSOE + Kalman filter	40 ms +2 ms=42 ms
Lopes	median + mean computation	64 ms
2-GMM	EM algorithm	132 ms
SPAMP	Pre-estimation algorithm	84 ms

Table 7.2: Elapsed time comparison among SOE methods

Recovery methods	Algorithms	Elapsed time
KML+BAOMP	KML + BAOMP	42+202 ms=244 ms
SPAMP	pre-estimation + SAMP	84+384 ms = 468 ms
2-GMM+OMP	EM algorithm + OMP	132+398 ms = 530 ms

Table 7.3: Elapsed time comparison among recovery methods

The time elapsed by the proposed KML and BAOMP methods is significantly less than that of other existing methods. Considering the acquisition of less number of measurements, reduced computational and run-time complexity, and better recovery performance, the proposed methods are the best candidates for efficient acquisition, compression, and recovery of vibration and similar compressible signals.

Chapter 8

Discussions and conclusions

The advantages of the proposed CS acquisition and recovery systems are addressed in this chapter. A study of the effects of the ML estimation and the composite sensing system on resource usage and execution speed is presented. The applicability of the proposed SOE methods is also discussed here.

8.1 Discussions on results

The following discussion shows how the proposed Kalman filtered ML (KML) estimator enhances the CS acquisition and how the BAOMP method increases the execution speed of CS recovery.

8.1.1 KML estimation method during CS acquisition

The major problem in CS is estimating the sparsity order k of the compressible signal to reduce the measurements size M. However, real-time applications demand that sparsity order needs to be estimated from the measurements itself. Our proposed method of composite sensing followed by KML provides a solution to this chicken-egg problem of sparsity order determining the measurements size, and the measurements determining sparsity order.

We combined sparse BSM and dense GSM for composite sensing of compressible signals during CS acquisition. As the sparse BSM has very few ones in each of its columns, it exhibits a weak RIP and is not suitable for the perfect recovery of the compressed signal with limited measurements. However, we exploited the weak RIP for SOE. The challenge in the design of the BSM is to cater to the requirement of estimating the time-varying sparsity order of the compressible signal with limited measurements. Our solution for this challenge resulted in a BSM adjusting its dimensions and entries according to the time-varying

sparsity order. The BSM is deterministic, which also suits the practical implementation of the CS acquisition and recovery systems.

We derived a blind BSOE method that does not require any a priori knowledge of signal and noise statistics. These statistics were estimated from the statistics of the measurements and BSM entries. We proposed KML to reduce the variance of the BSOE and improve the SOE performance. Thus, the proposed KML results in an optimal number of CS measurements, which determines the efficient use of CS acquisition hardware. A simulation of the SOEE performance (Figure 5.6) shows that the proposed KML method performs better for all SNR conditions, even with three times fewer measurements than other SOE methods.

8.1.2 KML-BAOMP during CS recovery

The KML method performs SOE from CS measurements obtained during CS recovery. The SOE during recovery was the same as that during acquisition. The better SOEE performance of KML resulted in better NRE performance compared with other existing SOE and support estimation methods. The better NRE performance owing to the composite sensing and KML methods is demonstrated by analyzing synthetic and real-world signals, as shown in Figures 5.7, 5.8, 6.7a, and 6.7b.

8.1.3 Impact of KML on CS acquisition and recovery

Because the sparsity order determines the measurement size, the SOE must be accurate for efficient CS acquisition and recovery in terms of the optimal use of hardware resources and the quality of the recovered signal. Suppose the SOE error $\widehat{k} - k = \delta_k$, the error in the number of measurements is,

$$
\begin{aligned}
M_k - M_{k+\delta_k} &= \left\lceil k\left(0.63 + 2\log\left(\frac{N}{k}\right)\right)\right\rceil - \left\lceil (k+\delta_k)\left(0.63 + 2\log\left(\frac{N}{k+\delta_k}\right)\right)\right\rceil \\
&= \left\lceil \log\left(\frac{k+\delta_k}{k}\right) - \delta_k\left(0.63 + 2\log\left(\frac{N}{k+\delta_k}\right)\right)\right\rceil \\
&\approx \delta_k\left(0.63 + 2\log\left(\frac{N}{k+\delta_k}\right)\right).
\end{aligned}
$$

where M_k is the number of measurements required for the true k value and $M_{k+\delta_k}$ is the number of measurements required for the estimate $\widehat{k} = k + \delta_k$. Thus, the difference in the number of measurements is linearly proportional to the SOE error δ_k.

- Case 1: If $\delta_k > 0$, then $M_{k+\delta_k} > M_k$ resulting in excessive measurements and the inefficient use of hardware resources. However, the recovery performance was not

affected.

- Case 2: If $\delta_k < 0$, then $M_{k+\delta_k} < M_k$ resulting in a fewer number of measurements than the required degrading the recovery quality.

It has been shown in Chapter 4 that Kalman filtering reduces the variance of BSOE, thereby reducing the estimation error $\delta_k \ll k$, resulting in a minimal error in M and negligible impact on CS acquisition and recovery.

8.1.4 Impact of KML on the execution speed of CS

The KML method has an impact on both the CS acquisition and recovery processes. The time complexity analysis for the KML method shows that the acquisition rate is not affected during CS acquisition and execution speed is increased during recovery as given below.

8.1.4.1 Impact on CS acquisition

During the acquisition, the CS results in $M \ll N$ compressed samples for a given time interval $\mathcal{T}$. Because the proposed KML method determines M based on the previously estimated sparsity order $\widehat{k}(n-1)$, the time required for the SOE should be $< \mathcal{T}$. The time complexity of the KML method depends on the number of iterations involved in estimating probability q as given in Step 5 of Algorithm 1. Typically ten iterations are sufficient for the convergence of the KML method, and the time taken for convergence is significantly less than $< \mathcal{T}$ s. For example, Chapter 7 shows that the KML method requires 42 ms compared to $\mathcal{T} = 1s$ when analyzing real-world vibration signals. Thus, until the time required for the SOE is less than a fixed interval $\mathcal{T}$, the KML method does not reduce the CS acquisition rate.

8.1.4.2 Impact on CS recovery

For greedy recovery algorithms, the sparsity order k is an input. If k is unknown, the greedy algorithms iterate $k_{\max}$ times or, in the worst case, M times. For real-time applications, when k is unknown and assumed to be k_{max}, the iteration places the constraint that $k_{max}\rho < \mathcal{T}$, where ρ is the time elapsed per iteration. By using the proposed KML and BAOMP methods, the constraint becomes $(\widehat{k} - |\widehat{\mathcal{S}}|)\rho < \mathcal{T}$. Because $(\widehat{k} - |\widehat{\mathcal{S}}|)\rho < k_{max}\rho$, the KML method is fast, and the real-time recovery of fast-varying compressible signals is possible. For applications requiring offline recovery and analysis, CS execution increases by $\frac{k_{max}}{\widehat{k} - |\widehat{\mathcal{S}}|}$ using the proposed methods.

8.1.4.3   Impact on mismatches on the prior model for the SOE

The proposed KML method relies on the prior models for amplitude and sparsity order. These models are advantageous as they represent a wide range of information, including knowledge gained a priori about the process that created the signal as well as information about measurement system, noise characterization, underlying probabilistic structure, etc. The greater the amount of prior knowledge we can incorporate into the algorithm, the lower the resulting error variance. However, the approach of model-based estimation would perform poorly if the models or their parameters weren't well known enough.

In the proposed KML method, the amplitude of significant components is assumed to be Gaussian with mean μ_s and the variance σ_s^2. The ML estimator for the sparsity order is derived from the measurements statistics. Though the estimator does not demand Gaussian assumption for the amplitude of significant components, it assumes that the amplitude's mean is μ_s and the variance is σ_s^2 and it provides ML estimates every time step. Then, using the Discrete Markov model, these ML estimates are later refined by Maximum Likelihood Sequence (MLS) estimation method or Kalman filtering method. Though the error in the estimates of mean and variance of amplitude of significant components result in error in instantaneous sparsity order estimates, the discrete Markov model based MLS or Kalman filter method would correct those errors. However, as the Markov model parameters are estimated from the instantaneous estimates, the correction would be to an extent when the estimation of Discrete Markov model parameters have lesser variance. Thus under such model mismatches i.e., parametric and statistical uncertainties to certain extent,, the proposed method with good history of past measurements exhibit good amount of resilience and robustness in terms of their mean squared error variation.

8.1.5   Applicability of the KML method

The KML method applies to sparse and compressible signals. It should be noted that some signals such as noise are neither sparse nor compressible. Sometimes compressible signals become non-sparse due to disturbances. For example, the vibration signal of a rocket is almost a random signal during the transonic regime with rich in significant components making the signal non-sparse. Here, we do not consider such non-sparse signals or conditions. These signals or conditions will be understood while performing SOE using the KML method. When the sparsity order k of the compressible signal is increased, the number of measurements less than the threshold τ will decrease, i.e., the probability $p_0 = \lambda^k$ will decrease. Suppose the maximum sparsity order for a signal to be considered as compressible

is k_{max} which leads to the minimum probability $p_{0min} = \lambda^{k_{max}}$. If the estimated probability $\widehat{p}_0 < p_{0min}$, the acquired signal is neither sparse nor compressible. In such cases, the sensing matrix dimension matches with the signal's dimension satisfying Nyquist sampling conditions.

8.2 Conclusions

Novel practical implementable CS acquisition and recovery systems are proposed for dynamic sparse and compressible signals in this research work. The importance of sparsity order estimation is emphasized for the CS acquisition and recovery. Suitable models for the compressible signals, sparsity order, and composite sensing system are defined. We combined sparse BSM and dense GSM for composite sensing of compressible signals during CS acquisition. Though sparse BSM exhibits a weak RIP, we exploited the weak RIP for SOE. The challenge in the design of the BSM adapting to the time-varying sparsity order resulted in a BSM adjusting its dimensions and entries according to the time-varying sparsity order. The BSM proposed is deterministic, which also suits the practical implementation of the CS acquisition and recovery systems.

Using the ML principles, two unique innovative SOE methodologies based on BSM and GSM-based measurements have been developed. The BSM-based SOE, i.e., the BSOE technique estimates sparsity order without the *a priori* knowledge on statistics of compressible signals. Whereas, the GSM-based SOE, i.e, GSOE technique uses the estimate of statistics of compressible signal obtained from BSOE for the ML estimation of sparsity order. In the acquisition process, the ML estimators estimate sparsity order with a minimum number of measurements without any additional cost or resources compared to any other estimator. The instantaneous estimates are then refined using either of the two different proposed approaches namely MLS estimation using Viterbi algorithm and Kalman filtering. It is shown that the Viterbi algorithm-based MLS estimation technique is suitable for estimating slow and moderately varying sparsity orders with lower computational complexity. As Viterbi algorithm-based approach performs trace-back-based sequence estimation, there is a delay in providing the sparsity order estimates. However, Kalman filtering-based approach is instantaneous and has similar performance with respect to Viterbi algorithm-based approach.

In the recovery process, a novel BAOMP method is proposed for estimating the support of the compressible signals using the BSM-based measurements. It is shown that BAOMP reduces the recovery time by at least 25% compared to other existing methods.

The proposed ML sparsity order estimation methods and BAOMP-based recovery method show better performance in terms of various metrics in comparison with other methods published thus far while analyzing real-world signals and synthetic signals.

Chapter 9

Future works

The present work focuses only on the problem of estimating the sparsity order, and uses it for the acquisition and recovery of time-varying sparse and compressible signals. However, the problem of estimating and tracking time-varying sparse signals is three-dimensional, where time-varying (i) sparsity order, (ii) support, and (iii) amplitude must be estimated simultaneously. Hence, in future work, the characterization of time-varying sparsity order and support together as a two-dimensional discrete Markov birth-death model and the derivation of corresponding ML algorithms can be considered.

The amplitude of each of its supporting components shall be modeled using an independent autoregressive (AR) process, and independent scalar Kalman filters (KF) shall be employed to estimate and track the amplitude of the supporting components.

A simulation was carried out to see the performance of amplitude estimation using KF. The support was kept time-invariant and thereby the sparsity order of the compressible signal was kept constant at $k = 50$ for the signal dimension $N = 2500$ for 50 time steps. The amplitude of each those 50 components is varied using an AR process. In each time step, the sparsity order was estimated using the KML method. Then an iterative BAOMP-based CS reconstruction algorithm that works on thresholding operators was used to estimate the probable k supporting components. Although the BAOMP algorithm is faster in execution and easier to implement, it suffers from performance degradation similar to any greedy-based CS recovery algorithms in the presence of strong measurement noise. Hence each supporting component's amplitude was filtered using independent scalar KFs knowing the AR process model parameters. The improved performance by the KF in terms of recovery error is shown in Figure 9.1.

The above idea of employing KF for refining the time-varying amplitude needs to address the time-varying support and sparsity order i.e., the number of KFs required to run depends on time-varying sparsity order. This problem can be tackled using two different
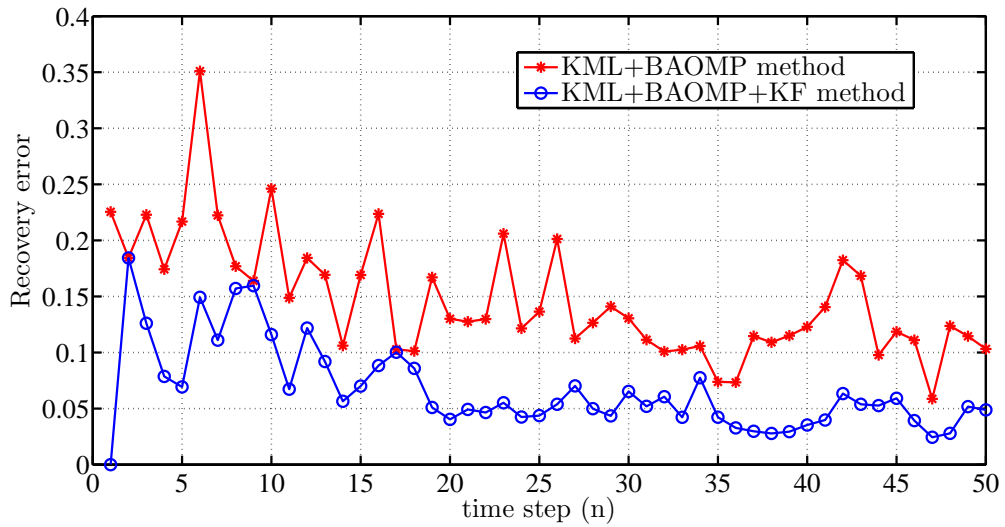
Figure 9.1: Improved recovery performance by employing Kalman filtering of the BAOMP estimate.

approaches. The first approach is discussed in Algorithm 9 which uses the coarse support estimates of BAOMP for KF-based amplitude tracking and the innovation error of KF as a feedback to BAOMP for refining the support.

The second approach is shown in Figure 9.2. Here, sparsity order is estimated initially and then BAOMP estimated supporting components' amplitudes are refined using KFs. Whenever the current estimated sparsity order does not differ from the previously estimated, then the BAOMP algorithm shall be bypassed, and previously estimated Kalman filtered amplitude itself can be taken as the prior estimate for the KF. Then KF amplitudes shall be used for updating the estimates of statistics of significant coefficients.

Algorithm 9 Estimation of the time-varying sparse signal using KML-BAOMP-KF algorithm

---

Input: Measurements $\mathbf{y}(n) : \{y_1(n), y_2(n), .., y_m(n)\}$;

1. Estimate the statistics of the coefficients by using the BSOE algorithm given in Algorithm 4.

2. Estimate the KML estimate of sparsity order.

3. Execute BAOMP algorithm to obtain an estimate of the support $\mathcal{S}(n)$ and a coarse estimate of amplitude $\widehat{\mathbf{s}}(n)$.

4. Kalman filtering:

   4.1 Prior amplitude estimate: $\widehat{s}(n|n\text{-}1) = \widehat{s}(n\text{-}1|n\text{-}1)$;

   4.2 Prior error covariance: $\mathbf{P}(n|n\text{-}1) = \mathbf{P}(n\text{-}1|n\text{-}1) + 1/k(\widehat{\mu_s^2 + \sigma_s^2})\mathbf{I}$;

   4.3 Innovation error: $\mathbf{y}_{e,n} = \mathbf{y}(n|n)\text{-}\boldsymbol{\Phi}\widehat{\mathbf{s}}(n|n\text{-}1)$;

   4.4 Kalman Gain: $\mathbf{K} = (\mathbf{P}(n\text{-}1|n\text{-}1)\boldsymbol{\Phi}^T(\boldsymbol{\Phi}\mathbf{P}(n\text{-}1|n\text{-}1)\boldsymbol{\Phi}^T + \boldsymbol{\sigma_\vartheta^2}\mathbf{I})^{-1})$;

   4.5 Posterior error covariance: $\mathbf{P}(n|n) = (\mathbf{I} - \mathbf{K}\boldsymbol{\Phi})\mathbf{P}(n|n\text{-}1)$;

   4.6 Posterior amplitude estimate: $\mathbf{s}(n|n) = \mathbf{s}(n|n\text{-}1)\text{+}\mathbf{K}\mathbf{y}_{e,n}$;

5. Apply BAOMP to the innovation error to update support $\mathcal{S}(n)$ and amplitude $\widehat{\mathbf{s}}(n)$.

6. Iterate steps [3]-[5] until $\mathbf{y}_{e,n}^T\mathbf{y}_{e,n} \leq 0.01$.

7. $\widehat{\mathbf{x}}(n) = \boldsymbol{\Psi}\widehat{\mathbf{s}}(n)$.

Output: Estimate of time-varying sparse signal: $\widehat{\mathbf{x}}(n)$.
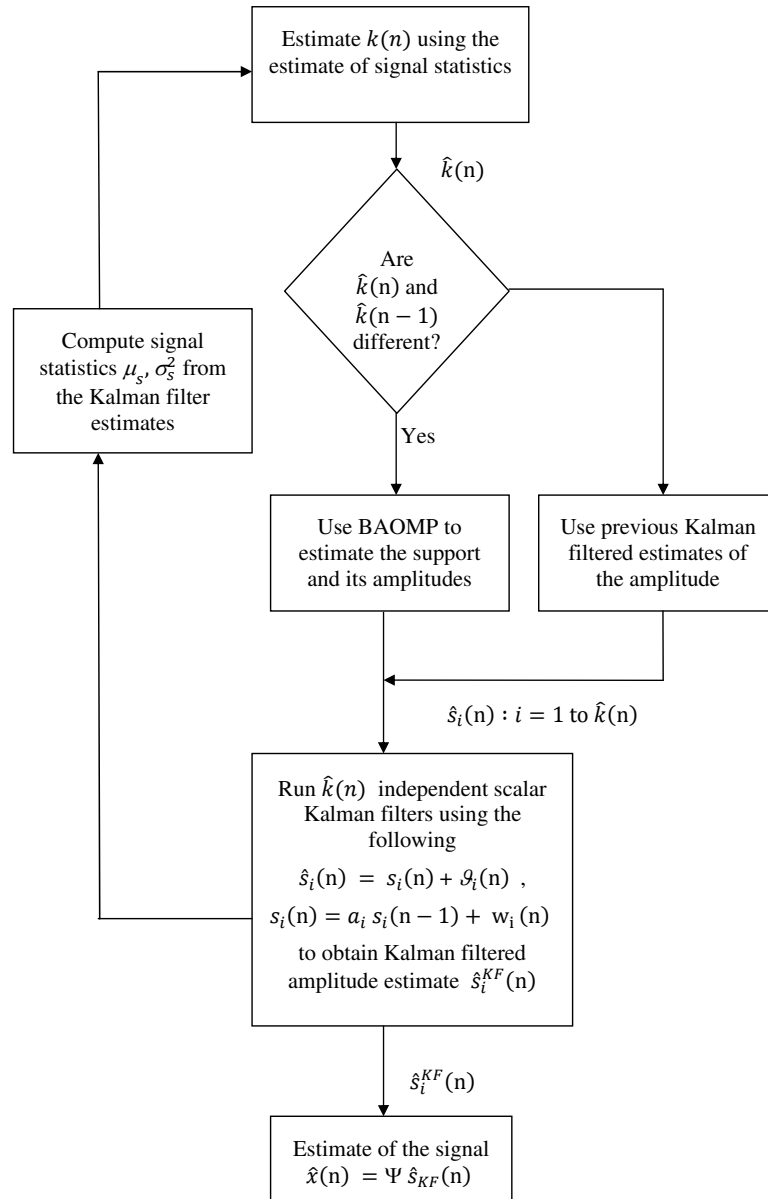

Figure 9.2: Employing Kalman filtering of the BAOMP estimate.

Bibliography

[1] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[2] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[3] M. A. Davenport, M. F. Duarte, Y. C. Eldar, and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge: Cambridge University Press, 2012.

[4] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*. Birkhäuser Basel, 2013.

[5] M. Rani, S. B. Dhok, and R. B. Deshmukh, "A systematic review of compressive sensing: Concepts, implementations and applications," *IEEE Access*, vol. 6, pp. 4875–4894, 2018.

[6] R. Arie, A. Brand, and S. Engelberg, "Compressive sensing and sub-nyquist sampling," *IEEE Instrumentation Measurement Magazine*, vol. 23, no. 2, pp. 94–101, 2020.

[7] A. Mousavi, M. Rezaee, and R. Ayanzadeh, "A survey on compressive sensing: classical results and recent advancements," *Journal of Mathematical Modeling*, vol. 8, no. 3, pp. 309–344, 2020.

[8] G. Taubock and F. Hlawatsch, "A compressed sensing technique for ofdm channel estimation in mobile environments: Exploiting channel sparsity for reducing pilots," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008, pp. 2885–2888.

[9] R. Mohammadian, A. Amini, and B. H. Khalaj, "Compressive sensing-based pilot design for sparse channel estimation in ofdm systems," *IEEE Communications Letters*, vol. 21, no. 1, pp. 4–7, 2017.

[10] N. R. Butt and A. Jakobsson, "Efficient sparse spectrum estimation for cognitive radios," in *2012 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2012, pp. 1–4.

[11] C. Qi, L. Wu, Y. Huang, and A. Nallanathan, "Joint design of pilot power and pilot pattern for sparse cognitive radio systems," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 11, pp. 5384–5390, 2015.

[12] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[13] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[14] D. Needell and J. A. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Commun. ACM*, vol. 53, no. 12, pp. 93–100, Dec. 2010.

[15] T. T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4680–4688, 2011.

[16] M. A. Davenport and M. B. Wakin, "Analysis of orthogonal matching pursuit using the restricted isometry property," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4395–4401, 2010.

[17] R. Giryes and M. Elad, "Rip-based near-oracle performance guarantees for sp, cosamp, and iht," *IEEE Transactions on Signal Processing*, vol. 60, no. 3, pp. 1465–1468, 2012.

[18] Z. Gao, L. Dai, S. Han, C. I, Z. Wang, and L. Hanzo, "Compressive sensing techniques for next-generation wireless communications," *IEEE Wireless Communications*, vol. 25, no. 3, pp. 144–153, JUNE 2018.

[19] S. Lu, I. A. Hemadeh, M. El-Hajjar, and L. Hanzo, "Compressed-sensing-aided space-time frequency index modulation," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 7, pp. 6259–6271, July 2018.

[20] V. S. Amin, Y. D. Zhang, and B. Himed, "Sparsity-based time-frequency representation of fm signals with burst missing samples," *Signal Processing*, vol. 155, pp. 25 – 43, 2019.

[21] E. Sejdi, I. Orovi, and S. Stankovi, "Compressive sensing meets time-frequency: An overview of recent advances in time-frequency processing of sparse signals," *Digital Signal Processing*, vol. 77, pp. 22 – 35, 2018.

[22] T. Wimalajeewa and P. K. Varshney, "Sparse signal detection with compressive measurements via partial support set estimation," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 3, no. 1, pp. 46–60, 2017.

[23] N. Vaswani, "Kalman filtered compressed sensing," in *Proceedings of the 15th IEEE International Conference on Image Processing*, San Diego, CA, USA, 2008, pp. 893–896.

[24] D. Angelosante, G. B. Giannakis, and E. Grossi, "Compressed sensing of time-varying signals," in *Proceedings of the 16th International Conference on Digital Signal Processing*, Santorini-Hellas, Greece, 2009, pp. 1–8.

[25] L. Gan, "Block compressed sensing of natural images," in *Proceedings of the 15th International Conference on Digital Signal Processing*, Cardiff, UK, 2007, pp. 403–406.

[26] J. Ziniel and P. Schniter, "Dynamic compressive sensing of time-varying signals via approximate message passing," *IEEE Transactions on Signal Processing*, vol. 61, no. 21, pp. 5270–5284, 2013.

[27] S. Lan, Q. Zhang, X. Zhang, and Z. Guo, "Sparsity estimation in image compressive sensing," in *Proceedings of the IEEE International Symposium on Circuits and Systems*, Seoul, South Korea, 2012, pp. 2669–2672.

[28] M. E. Lopes, "Unknown sparsity in compressed sensing: Denoising and inference," *IEEE Transactions on Information Theory*, vol. 62, no. 9, pp. 5145–5166, Sep. 2016.

[29] Z. Zhang, F. Liu, R. Du, C. Huang, and J. Sheng, "Trace-based sparsity order estimation with sparsely sampled random matrices," *IEEE Transactions on Communications*, vol. 64, no. 11, pp. 4790–4799, 2016.

[30] C. Ravazzi, S. M. Fosson, T. Bianchi, and E. Magli, "Sparsity estimation from compressive projections via sparse random matrices," *EURASIP Journal on Advances in Signal Processing*, vol. 2018, no. 56, 2018.

[31] Y. Gao, Y. Si, B. Zhu, and Y. Wei, "Sparsity order estimation algorithm in compressed sensing by exploiting slope analysis," in *2018 14th International Wireless Communications Mobile Computing Conference (IWCMC)*, 2018, pp. 753–756.

[32] S. Semper, F. Romer, T. Hotz, and G. DelGaldo, "Sparsity order estimation from a single compressed observation vector," *IEEE Transactions on Signal Processing*, vol. 66, no. 15, pp. 3958–3971, 2018.

[33] J. E. Shouhdy, B. Abdelhamid, and S. H. El Ramly, "Spectrum sensing with sparsity estimation for cognitive radio systems," in *2019 2nd IEEE Middle East and North Africa COMMunications Conference (MENACOMM)*, 2019, pp. 1–6.

[34] Y. Luo, J. Dang, and Z. Song, "Optimal compressive spectrum sensing based on sparsity order estimation in wideband cognitive radios," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 12, pp. 12 094–12 106, 2019.

[35] Zhang, L. Xinhe, W. Yufeng, and Xin, "A sparsity preestimated adaptive matching pursuit algorithm," *Journal of Electrical and Computer Engineering*, vol. 2021, 2021.

[36] S. Zhiyuan, W. Qianqian, and C. Xinmiao, "A sparsity adaptive compressed signal reconstruction based on sensing dictionary," *Journal of Systems Engineering and Electronics*, vol. 32, no. 6, pp. 1345–1353, 2021.

[37] X. Xu, J. Chen, N. Wan, D. Chen, and J. Wan, "Sparsity estimation method in compressed data gathering of wireless sensor networks," in *2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, 2019, pp. 833–836.

[38] Y. Wei, Z. Lu, G. Yuan, Z. Fang, and Y. Huang, "Sparsity adaptive matching pursuit detection algorithm based on compressed sensing for radar signals," *Sensors*, vol. 17, no. 5, 2017. [Online]. Available: https://www.mdpi.com/1424-8220/17/5/1120

[39] Y. Zhang, R. Venkatesan, O. A. Dobre, and C. Li, "Efficient estimation and prediction for sparse time-varying underwater acoustic channels," *IEEE Journal of Oceanic Engineering*, vol. 45, no. 3, pp. 1112–1125, 2020.

[40] Y. Huang, Y. He, L. Shi, T. Cheng, Y. Sui, and W. He, "A sparsity-based adaptive channel estimation algorithm for massive mimo wireless powered communication networks," *IEEE Access*, vol. 7, pp. 124 106–124 115, 2019.

[41] A. Ravelomanantsoa, H. Rabah, and A. Rouane, "Compressed sensing: A simple deterministic measurement matrix and a fast recovery algorithm," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 12, pp. 3405–3413, 2015.

[42] D. Mitra, H. Zanddizari, and S. Rajan, "Investigation of kronecker-based recovery of compressed ecg signal," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 6, pp. 3642–3653, 2020.

[43] J. Wu, D. Feng, J. Wang, and X. Huang, "Sar imaging from azimuth missing raw data via sparsity adaptive stomp," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[44] B. Wang and L. Li, "Sparsity adaptive channel estimation algorithm based on compressed sensing," in *2021 2nd International Symposium on Computer Engineering and Intelligent Communications (ISCEIC)*, 2021, pp. 30–33.

[45] A. M. Tillmann and M. E. Pfetsch, "The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing," *IEEE Transactions on Information Theory*, vol. 60, no. 2, pp. 1248–1259, 2014.

[46] S. K. Sharma, S. Chatzinotas, and B. Ottersten, "Compressive sparsity order estimation for wideband cognitive radio receiver," *IEEE Transactions on Signal Processing*, vol. 62, no. 19, pp. 4984–4996, 2014.

[47] D. P. Wipf and B. D. Rao, "An empirical bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3704–3716, 2007.

[48] R. R. Naidu and C. R. Murthy, "Construction of binary sensing matrices using extremal set theory," *IEEE Signal Processing Letters*, vol. 24, no. 2, pp. 211–215, Feb 2017.

[49] Y. Vasavada and C. Prakash, "Sub-nyquist spectrum sensing of sparse wideband signals using low-density measurement matrices," *IEEE Transactions on Signal Processing*, vol. 68, pp. 3723–3737, 2020.

[50] W. Zhao, B. Sun, T. Wu, and Z. Yang, "On-chip neural data compression based on compressed sensing with sparse sensing matrices," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 12, no. 1, pp. 242–254, 2018.

[51] K. Luo, Z. Wang, J. Li, R. Yanakieva, and A. Cuschieri, "Information-enhanced sparse binary matrix in compressed sensing for ecg," *Electronics Letters*, vol. 50, no. 18, pp. 1271–1273, 2014.

[52] W. Chen and C. Lu, "Recovery guarantee of sparse binary sensing matrices under greedy algorithm," in *2018 International Symposium on Information Theory and Its Applications (ISITA)*, 2018, pp. 423–426.

[53] H. Mamaghanian, N. Khaled, D. Atienza, and P. Vandergheynst, "Compressed sensing for real-time energy-efficient ecg compression on wireless body sensor nodes," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 9, pp. 2456–2466, 2011.

[54] Z. He, T. Ogawa, and M. Haseyama, "The simplest measurement matrix for compressed sensing of natural images," in *2010 IEEE International Conference on Image Processing*, 2010, pp. 4301–4304.

[55] A. Gilbert and P. Indyk, "Sparse recovery using sparse matrices," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 937–947, 2010.

[56] A. A. Moghadam and H. Radha, "Hybrid compressed sensing of images," in *2010 IEEE International Workshop on Multimedia Signal Processing*, 2010, pp. 99–104.

[57] R. Gribonval, V. Cevher, and M. E. Davies, "Compressible distributions for high-dimensional statistics," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5016–5034, 2012.

[58] M. R. Manesh, A. Quadri, S. Subramaniam, and N. Kaabouch, "An optimized snr estimation technique using particle swarm optimization algorithm," in *2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC)*, 2017, pp. 1–6.

[59] S. Hanly, "Vibration analysis on the exterior of an aircraft during climb," https://blog.mide.com/vibration-analysis-fft-psd-and-spectrogram, accessed on 2021-06-01.

[60] R. Zhang, H. Zhao, and J. Zhang, "Distributed compressed sensing aided sparse channel estimation in fdd massive mimo system," *IEEE Access*, vol. 6, pp. 18 383–18 397, 2018.

[61] S. Khademi, S. P. Chepuri, Z. Irahhauten, G. J. M. Janssen, and A. van der Veen, "Channel measurements and modeling for a 60 ghz wireless link within a metal cabinet," *IEEE Transactions on Wireless Communications*, vol. 14, no. 9, pp. 5098–5110, 2015.

[62] M. L. Jakobsen, T. Pedersen, and B. H. Fleury, "Analysis of stochastic radio channels with temporal birth-death dynamics: A marked spatial point process perspective," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 7, pp. 3761–3775, 2014.

[63] Z. Lian, L. Jiang, C. He, and Q. Xi, "A novel multiuser hap-mimo channel model based on birth-death process," in *2016 IEEE International Conference on Communications (ICC)*, 2016, pp. 1–5.

[64] T. Nishiyama, "Cramér-rao-type bound and stam's inequality for discrete random variables," 2019.

[65] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[66] V. M. L. Silva, B. W. S. Arruda, C. P. de Souza, E. C. GurjÃčo, V. L. Reis, and R. C. S. Freire, "A testing approach for a configurable rmpi-based analog-to-information converter," in *2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, 2018, pp. 1–5.

[67] V. Bioglio, T. Bianchi, and E. Magli, "On the fly estimation of the sparsity degree in compressed sensing using sparse sensing matrices," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Brisbane, Australia, 2015, pp. 3801–3805.

[68] T. L. Gafka, J. H. Kerr, K. C. Yap, J. Macias, and M. Kaouk, "Structural health monitoring and risk management of a reusable launch vehicle," *Journal of Spacecraft and Rockets*, vol. 49, no. 6, pp. 1099–1108, 2012.

[69] H. Herranen, A. Kuusik, T. Saar, M. Reidla, R. Land, O. MÃd'rtens, and J. Majak, "Acceleration data acquisition and processing system for structural health monitoring," in *2014 IEEE Metrology for Aerospace (MetroAeroSpace)*, 2014, pp. 244–248.

[70] Y. Bao, J. L. Beck, and H. Li, "Compressive sampling for accelerometer signals in structural health monitoring," *Structural Health Monitoring*, vol. 10, no. 3, pp. 235–246, 2011.

[71] Y. Q. Bao, H. Li, and J. P. Ou, "Applications of compressive sensing technique in structural health monitoring," in *Structural Health Monitoring: Research and Applications*, ser. Key Engineering Materials, vol. 558.   Trans Tech Publications, 8 2013, pp. 561–566.

[72] H. Yong, B. J. L., W. Stephen, and L. Hui, "Robust bayesian compressive sensing for signals in structural health monitoring," *Computer-Aided Civil and Infrastructure Engineering*, vol. 29, no. 3, pp. 160–179, 2014.

[73] S. Ji, Y. Sun, and J. Shen, "A method of data recovery based on compressive sensing in wireless structural health monitoring," *Mathematical Problems in Engineering*, vol. 2014, no. Article ID 546478, p. 9 pages, 2014.

[74] J. T. Panachakel and K. Finitha, "Energy efficient compression of shock data using compressed sensing," in *Intelligent Systems Technologies and Applications. Advances in Intelligent Systems and Computing*, B. S., T. S., and S. P., Eds.   Springer, Cham, 2008, pp. 273–281.

[75] Y. Bao, Z. Shi, X. Wang, and H. Li, "Compressive sensing of wireless sensors based on group sparse optimization for structural health monitoring," *Structural Health Monitoring*, vol. 17, no. 4, pp. 823–836, 2018.

[76] N. Patwari, "CRAWDAD dataset utah/cir (v. 2007-09-10)," Downloaded from https://crawdad.org/utah/CIR/20070910, Sep. 2007.

[77] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000 (June 13), circulation Electronic Pages: http://circ.ahajournals.org/content/101/23/e215.full PMID:1085218; doi: 10.1161/01.CIR.101.23.e215.

[78] M. E. Lopes, "Estimating unknown sparsity in compressed sensing," in *Proceedings of the 30th International Conference on Machine Learning*, ser. ICML'13, vol. 28, Atlanta, USA, 2013, pp. 217–225.

List of Publications

Refereed Journals

1. Thiruppathirajan S., Lakshmi Narayanan R., Sreelal S., and Manoj B. S., "Maximum Likelihood estimation of time-varying sparsity order for dynamic sparse signals," in IEEE Access, Sept, 2021, doi: 10.1109/ACCESS.2021.3113871.

2. Thiruppathirajan S., Lakshmi Narayanan R., Sreelal S., and Manoj B. S., "Sparsity Order Estimation for Compressed Sensing system using sparse Binary Sensing Matrix," in IEEE Access, Mar, 2022, 10.1109/ACCESS.2022.3161523.