Adapting to changing signal statistics using compressed measurements

Sumeet Kumar Mishra
Department of Electrical Engineering, IIT Bombay
sumeet@iitb.ac.in
Guides: Prof. Ajit Rajwade & Prof. Nikhil Karamchandani

Abstract

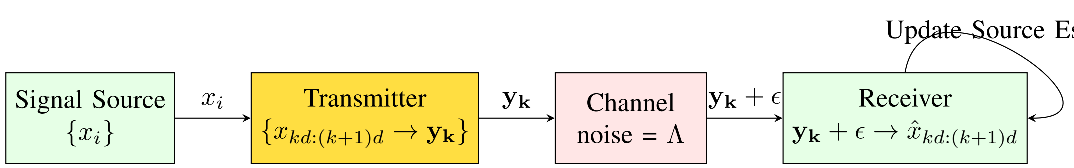
In this report, we present a method to compress the signals from a one-dimensional signal source by compressing a sequence of data points and using Statistical Compressed Sensing to reconstruct signals at the receiver. The sequence of data points is modelled as being drawn from a Gaussian Mixture Model **GMM**. We also propose a novel method to detect drift in the source distribution by considering the cross-validation error in reconstruction. On detecting drift, we update our estimate of the source GMM using compressed measurements. The proposed method can have various applications in monitoring systems, such as health monitoring or temperature monitoring systems, where we can have time-varying signal statistics.

I. Introduction

With the growth of transmission technologies (4G, 5G), big data, and IoT technologies, a huge amount of data is being transmitted constantly. But bandwidth limitations of channels specify limits on the data rate. To tackle this, compressed signals can be sent over the channels to increase the data rate. However, there needs to be a system in place to ensure reliable transmission. There has been a lot of work in the field of compressed sensing beginning with the work of Candès, Romberg and Tao[1], and Donoho[2] , which was followed by various developments and specific applications such as the work of Yu and Sapiro on Statistical Compressed Sensing of Gaussian Mixture Models[3], and the work of Jianbo Yang et al. on Compressive Sensing by Learning a Gaussian Mixture Model From Measurements[4]. We use the ideas in[3]-[4] in our model to reconstruct the compressed samples and to update our estimate of the source GMM in case a drift is detected in the source signal statistics.

The rest of the report is organised as follows. In Section 2, we describe the system and the problem statement. In Section 3, we present the method for reconstruction using SCS, and propose a method for drift detection. Updating the GMM estimate based on the compressed signals is described in Section 4, which is followed by experimental results in Section 5, and the conclusion.

II. System Model



Our system can be specified as follows

- **Signal Source**: The Signal Source is the source of information which generates a continuous signal or the sequence $\{x_i\}$
- **Transmitter**: The Transmitter collects these signals and groups them into partitions of size d (d = patch size). Using random sensing matrices, it compresses this sequence of d elements to a lower dimension (size($\mathbf{y_k}$) $< d$) and sends it over the channel. Note that we can use different sensing matrices for different samples, but the receiver must have the information about the sensing matrix corresponding to a particular compressed sample
- **Channel**: We assume that the channel adds Gaussian noise to the compressed signal of zero mean and a diagonal covariance matrix Λ
- **Receiver**: Finally, the receiver gets the noisy samples and reconstructs them using Statistical Compressed Sensing

III. Reconstruction & Drift Detection

A. Compressed Sensing

Before proceeding to Statistical Compressed Sensing, let us give a brief introduction to compressed sensing. We have a signal $\mathbf{x}$ that we compress using a sensing matrix $\boldsymbol{\phi}$ as $y = \boldsymbol{\phi}\mathbf{x} + \nu$, where ν is a noise vector. Our goal is to get a reconstruction $\hat{x}$ of x from y. However, this would require some assumption on the signals $\mathbf{x}$ since the number of measurements is less than the size of x. Suppose size of $\mathbf{x}$ is n and size of y is m, then $\boldsymbol{\phi}$ is a $m \times n$ matrix and $m < n$. For example, the assumptions on $\mathbf{x}$ can be that of sparsity. It has been shown in [5] that when the sensing matrix satisfies the RIP property we can get the optimal estimate $\hat{x}$ by solving the problem

$$\hat{x} = \arg\min_{x} \|x\|_1 \quad \text{s.t.} \quad \|y - \boldsymbol{\phi}x\|_2 < \epsilon$$

Where $\|x\|_1$ represents the $l1$ norm of x and ϵ depends on the noise variance. It has been proved

B. Statistical Compressed Sensing

If we assume that x is derived from a GMM, then closed-form solutions exist to the optimal $\hat{x}$, which minimises the Mean Square Error. In previous work, [7], a piecewise linear estimator was used to get the reconstruction, which is described in [6]. However, it has been shown in [3] that the MSE is minimised by using the weighted average of the piecewise linear estimates obtained from each component of the GMM. Let $\boldsymbol{\Theta} = \{\boldsymbol{\Lambda}\} \cup \{\boldsymbol{\pi}_k, \boldsymbol{\mu}_k, \mathbf{D}_k\}_{k=1}^{K}$ be the noise covariance matrix (assumed to be diagonal) and the parameters of the GMM and suppose the sensing model $\mathbf{y} = \boldsymbol{\phi}\mathbf{x} + \boldsymbol{\epsilon}$.

$$\begin{aligned}
\boldsymbol{\eta}_z(\mathbf{y}, \boldsymbol{\Theta}) &= \boldsymbol{\mu}_z + \mathbf{D}_z\boldsymbol{\phi}'(\boldsymbol{\Lambda} + \boldsymbol{\phi}\mathbf{D}_z\boldsymbol{\phi}')^{-1}(\mathbf{y} - \boldsymbol{\phi}\boldsymbol{\mu}_{\boldsymbol{z}}) \\
\mathbf{C}_z(\mathbf{y}, \boldsymbol{\Theta}) &= \mathbf{D}_z - \mathbf{D}_z\boldsymbol{\phi}'(\boldsymbol{\Lambda} + \boldsymbol{\phi}\mathbf{D}_z\boldsymbol{\phi}')^{-1}\boldsymbol{\phi}\mathbf{D}_z \\
\mathbf{R}_z(\boldsymbol{\Theta}) &= \boldsymbol{\Lambda} + \boldsymbol{\phi}\mathbf{D}_z\boldsymbol{\phi}'
\end{aligned}$$

$$\text{Let } \rho_z(\mathbf{y}, \boldsymbol{\Theta}) = \frac{\pi_z \mathcal{N}(\mathbf{y}; \boldsymbol{\phi}\boldsymbol{\mu}_{\boldsymbol{z}}, \mathbf{R}_{\mathbf{z}}\boldsymbol{\Theta})}{\sum_{l=1}^{K} \pi_l \mathcal{N}(\mathbf{y}; \boldsymbol{\phi}\boldsymbol{\mu}_{\boldsymbol{l}}, \mathbf{R}_{\mathbf{l}}\boldsymbol{\Theta})}$$

$$\hat{\mathbf{x}}_{\text{MMSE}}(\mathbf{y}, \boldsymbol{\Theta}) = \sum_{k=1}^{K} \rho_k(\mathbf{y}, \boldsymbol{\Theta})\boldsymbol{\eta}_k(\mathbf{y}, \boldsymbol{\Theta})$$

Here, $\hat{\mathbf{x}}_{\text{MMSE}}(\mathbf{y}, \boldsymbol{\Theta})$ is the optimal estimate in the mean square error sense. The mean square error is

$$\text{MSE}(\boldsymbol{\Theta}) = \int p(\mathbf{y}|\boldsymbol{\Theta}) \int \|\mathbf{x} - \hat{\mathbf{x}}(\mathbf{y}, \boldsymbol{\Theta})\|_2^2 p(\mathbf{x}|\mathbf{y}, \boldsymbol{\Theta}) d\mathbf{x} d\mathbf{y}$$

Note that $\boldsymbol{\eta}_z(\mathbf{y}, \boldsymbol{\Theta})$ can be thought of the piecewise linear estimate (reconstruction of the compressed sample based on the z^{th} component, $\rho_z(\mathbf{y}, \boldsymbol{\Theta})$ is the probability that the compressed sample $\mathbf{y}$ was drawn from the z^{th} component.

C. Drift Detection

Over a period, the source statistics may change and hence, there is a need to update our representation of the GMM (prior) since Statistical Compressed Sensing heavily relies on the prior. However, we do not have access to uncompressed signals. Hence, we use cross-validation as a measure to detect drift. Suppose the mean cross-validation error on the representative signals (on which our representation of the GMM was trained) was some $\overline{\boldsymbol{\epsilon_{cv}}}$ and the standard deviation of the cross-validation error was $\sigma_{\mathbf{cv}}$. Then we set a threshold $t = \overline{\boldsymbol{\epsilon_{cv}}} + \alpha \cdot \sigma_{\mathbf{cv}}$. Where $\alpha < 1$ is chosen based on experiment. We flag a drift when the cross-validation error of a continuous sequence of samples exceeds the threshold.

IV. Updating the Source GMM

Now suppose we have detected a drift in the source GMM. We want to update our prior based on the compressed samples received from the source. In previous work [7], the GMM was updated by requesting the source to send uncompressed samples. However, based on the algorithm in [4], the GMM can be learnt from compressed measurements. The method is based on the Maximum Marginal Likelihood. The updates are based on Expectation Maximisation. Suppose the GMM estimate at iteration t is $\boldsymbol{\Theta}^t$. The updates are given as follows

$$\pi_k^{(t)} = \frac{\sum_{i=1}^{N} \rho_{ik}^{(t-1)}}{\sum_{l=1}^{K}\sum_{i=1}^{N} \rho_{il}^{(t-1)}} \quad \mu_k^{(t)} = \frac{\sum_{i=1}^{N} \rho_{ik}^{(t-1)} \boldsymbol{\eta}_{ik}^{(t-1)}}{\sum_{i=1}^{N} \rho_{ik}^{(t-1)}}$$

$$\mathbf{D}_k^{(t)} = \frac{\sum_{i=1}^{N} \rho_{ik}^{(t-1)}[(\boldsymbol{\eta}_{ik}^{(t-1)} - \boldsymbol{\mu}_k^{(t)})(\boldsymbol{\eta}_{ik}^{(t-1)} - \boldsymbol{\mu}_k^{(t)})' + \mathbf{C}_{ik}^{(t-1)}]}{\sum_{i=1}^{N} \rho_{ik}^{(t-1)}}$$

Note that this algorithm is guaranteed to monotonically increase the marginal log-likelihood function

$$\sum_{i=1}^{N} \ln \sum_{k=1}^{K} \pi_k^{(t)} \mathcal{N}(\mathbf{y_i} | \boldsymbol{\phi} \boldsymbol{\mu}_k^{(t)}, \mathbf{R}_k(\Theta^{(t)})) \tag{1}$$

This method is beneficial for our purpose as we are already starting from a reasonable estimate of $\boldsymbol{\Theta}$ in the cases where the drift is not too significant. Otherwise, with random initialisation, the EM algorithm can get stuck at a local minimum.

V. Results

Note: The error metric in consideration is the Mean of the RMSE of each patch (patch size = d)

$$\text{Mean RMSE} = \frac{1}{n} \sum_{i=1}^{n} \sqrt{\frac{1}{d} \sum_{j=1}^{d} (x_{i,j} - \hat{x}_{i,j})^2}$$

Here, i is the index of the i^{th} patch and j in the index of the j^{th} element in the i^{th} patch

A. Synthetic GMM source with synthetic drift

In this experiment, we used a random GMM and sampled 10000 samples from this GMM. The GMM parameters were $n\_components = 5$, $n\_features = 20$. Using these uncompressed samples we trained an estimate on the source GMM with the assumption that the number of components was known ($n\_components = 5$). Following this, we compressed the samples using one of 10 random binomial matrices ($\{\phi_\mathbf{i}\}$) and tested the reconstruction methods and compared the Mean RMSE values. Note that the **DCT** reconstruction method is the same as retaining the first $n\_measurements$ number of DCT coefficients and setting the rest to zero. This implies that DCT would only preserve the low-frequency components of the signal, as can be seen in the image below. We also tried the reconstruction with the measurement matrices as random gaussian matrices but the performance of the random binomial matrices was superior. This was surprising as random gaussian matrices have better RIC (Restricted Isometric Constant) than random binomial matrices.

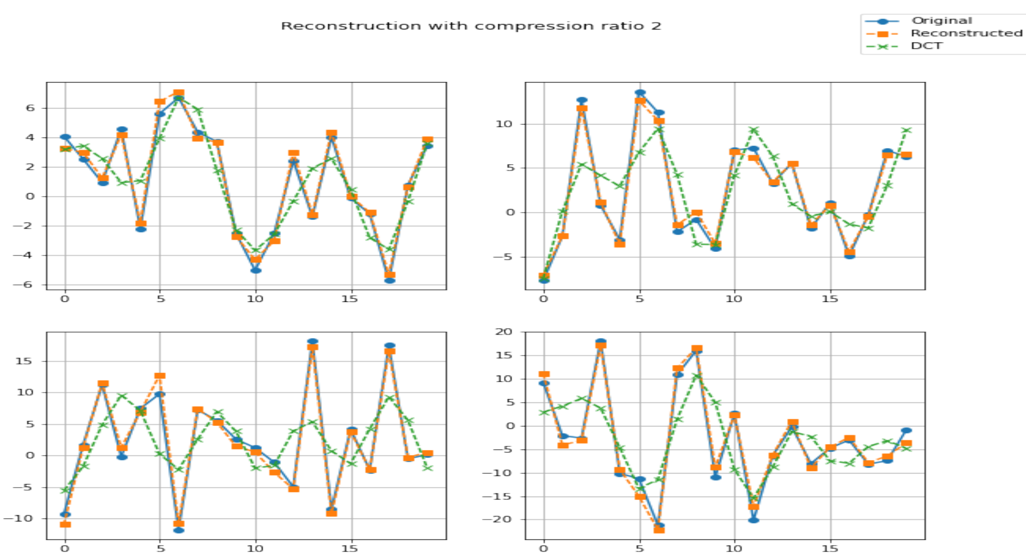


Fig. 1: Reconstruction at compression ratio = 2

Refer to (fig-2) for reconstructions at various compression levels using Statistical Compressed Sensing.

Now, to introduce drift synthetically, we only shifted the mean vectors of the source GMM by at most 20% of their norm in a random direction while keeping the covariance matrices the same. Here, $U(a, b) \sim Uniform(a, b)$.

$$G(\{\mu_i, \Sigma_i\}) \rightarrow G'(\{\mu_i + U(0, 0.2\|\mu_i\|)\delta, \Sigma_i\})$$

Following this, we sampled 2000 samples from the shifted GMM source and compressed them. Then, using the compressed measurements, updated our estimate of the GMM using the algorithm described in [4]. Note that we have a very high probability of converging to the shifted GMM as we already have a reasonable estimate of the source. However, starting from a random

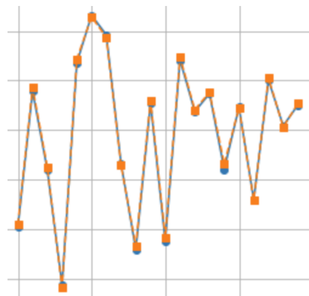


Fig. 2: Reconstruction at CR=1.25

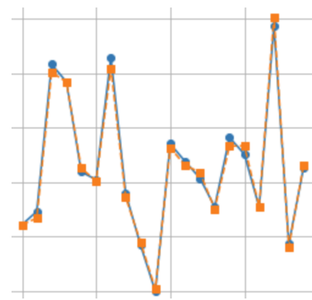


Fig. 3: Reconstruction at CR=1.67

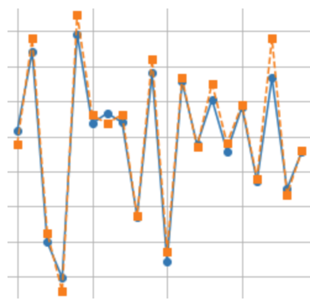


Fig. 4: Reconstruction at CR=2.5

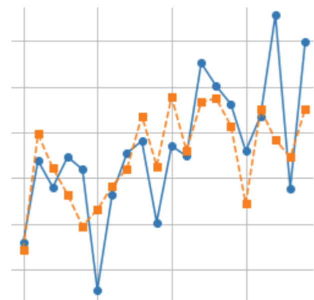


Fig. 5: Reconstruction at CR=5

Fig. 6: Reconstruction at various Compression Ratios using SCS

initialisation and using the update method can lead us to a local minimum. Refer to 7 to see the reconstructions of the compressed sampled after the drift is introduced, and RMSE-table to check the RMSE values of the reconstruction, pre and post update.

	Mean RMSE with method		
Compression	Min MSE	Max Max	DCT
5	5.265	7.770	7.189
2.5	2.278	2.677	6.203
2	1.527	1.579	5.798
1.67	1.124	1.128	4.939
1.25	0.549	0.549	3.48

	Mean RMSE (min MSE method)	
Compression	Updated Estimate	Old Estimate
5	5.59	5.60
2.5	2.58	2.61
2	1.89	1.92
1.67	1.48	1.53
1.25	0.74	1.07

B. Aluminium Panel Dataset

In this part, we use the data obtained by conducting experiments on an aluminium panel 8. An input pulse 9 was given at one point, and output was measured at the other points. Later, a rivet hole was introduced at the centre of the panel to model damage, and the experiment was repeated. We compare the response before and after the damage10. The experiment was conducted at various temperatures in a controlled environment to check the temperature dependence. However, there are only minor variations in the output response, and for our experiment, we concatenate data of temperatures close to each other to get more data points. (Temperatures in which experiment was conducted - 20, 25, . . . 70; we merge response from (25,30), (30,35). . .)

The response from the healthy Aluminium Panel was split into train and test sets. The training set was made using overlapping patches, however, and the test set was made using non-overlapping patches. We experimented with various patch sizes $d(50, 100, 150)$ and observed the standard trend of increasing PSNR with decreasing compression ratio. The reconstruction was done using 80% of the measurements, and the rest were used to calculate cross-validation errors. The average cross-validation error on the test set of the healthy panel was used to get an approximate value of the threshold for flagging drift. The PSNRs for different compression ratios were obtained as follows at different temperatures (1).

Suppose the mean of the cross-validation error on the test set of the healthy panel was $\overline{\boldsymbol{\epsilon_{cv}}}$ and the standard deviation was $\boldsymbol{\sigma}_{cv}$, then we set the threshold $t = \overline{\boldsymbol{\epsilon_{cv}}} + 0.5\boldsymbol{\sigma_{cv}}$. If the number of consecutive samples for which the cross-validation error is t is greater than 5, we flag a drift in the distribution.

Next, we built our anomalous dataset using non-overlapping patches from the rivet hole response and compressing them. We reconstructed these samples using the GMM trained on the healthy panel and observed the cross-validation error (reconstruction

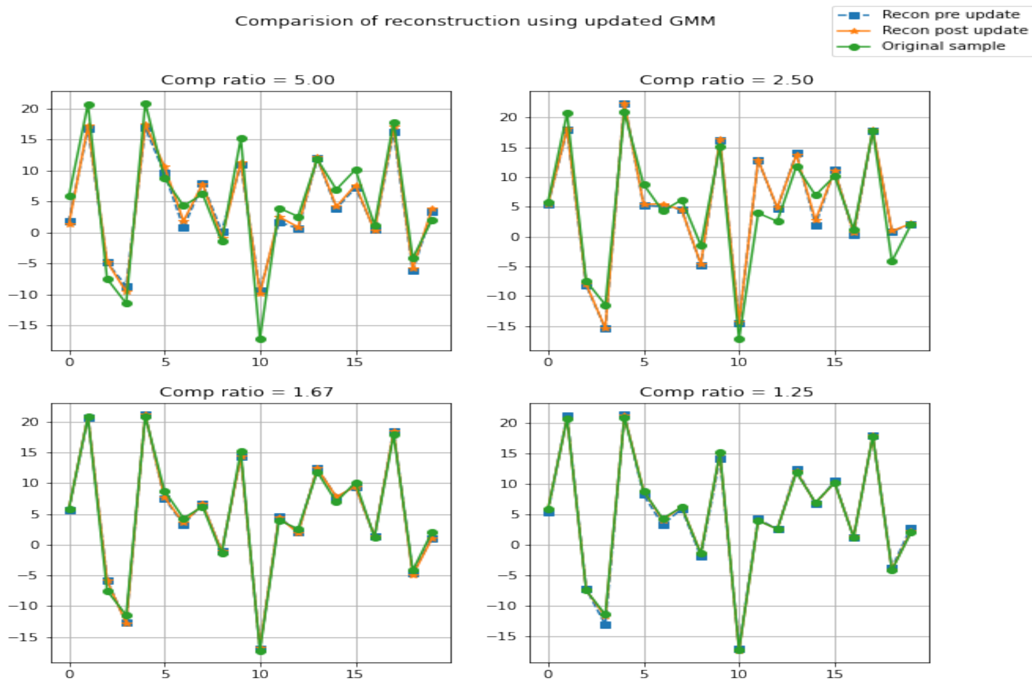


Fig. 7: Comparision of Mean RMSE after update

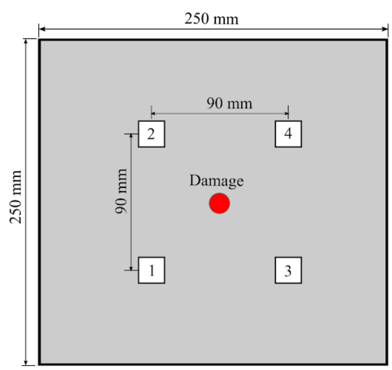


Fig. 8: Aluminium Panel

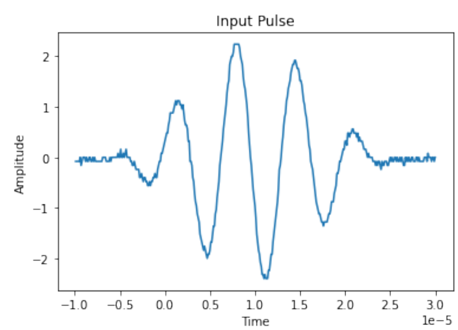


Fig. 9

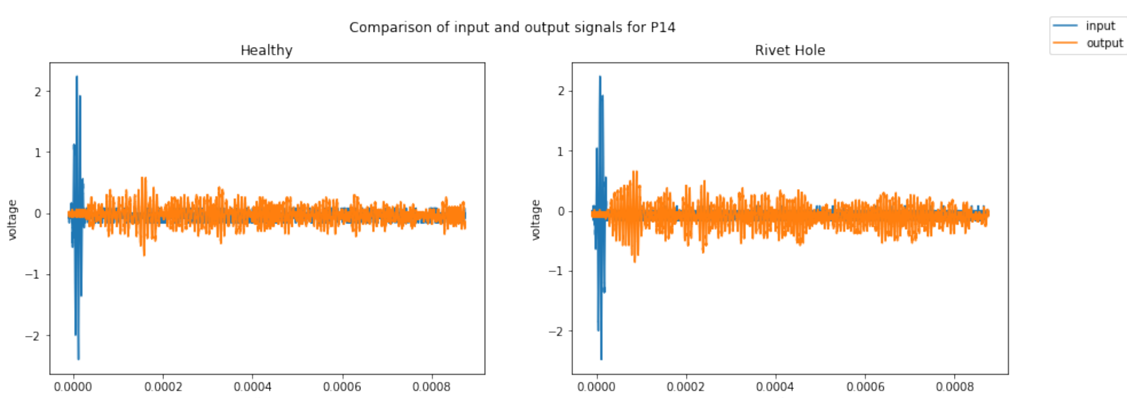


Fig. 10: Healty and Rivet Hole Panel response

	PSNR with Percent measurements									
Temp	10	20	30	40	50	60	70	80	90	100
20	29.37	31.204	32.40	33.13	33.717	34.425	35.193	36.083	37.291	38.755
30	29.823	33.52	34.49	34.969	35.58	36.40	37.10	38.07	39.195	40.65
40	29.28	31.86	32.504	33.185	33.82	34.544	35.25	36.07	37.13	38.689
50	30.044	32.14	33.23	33.829	34.460	35.143	35.83	36.70	37.78	39.25
60	29.15	32.243	33.25	33.973	34.609	35.15	35.940	36.89	38.08	39.735

TABLE I: PSNR on the reconstruction of healthy test set

	PSNR with Percent measurements									
Temp	10	20	30	40	50	60	70	80	90	100
20	32.89	35.481	36.727	37.385	38.002	38.682	39.436	40.283	41.401	42.844
30	30.621	34.685	35.574	36.134	36.765	37.428	38.139	39.009	40.146	41.784
40	22.47	32.567	33.60	34.403	35.341	36.140	36.897	37.874	39.076	40.719
50	30.323	32.508	33.880	34.471	35.055	35.750	36.50	37.365	38.43	39.96
60	26.63	31.515	32.628	33.289	33.880	34.48	35.37	36.343	37.420	38.85

TABLE II: PSNR on the reconstruction of damaged test set

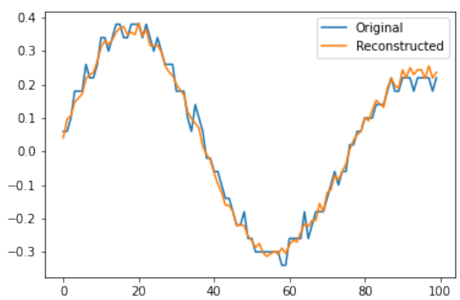


Fig. 11: Healthy panel with 50% measurements

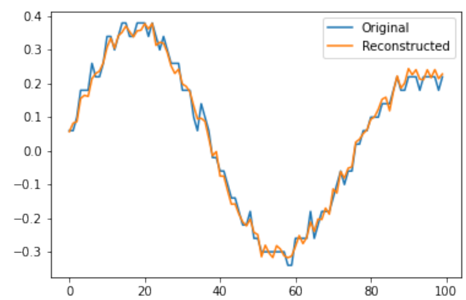


Fig. 12: Healthy panel with 80% measurements

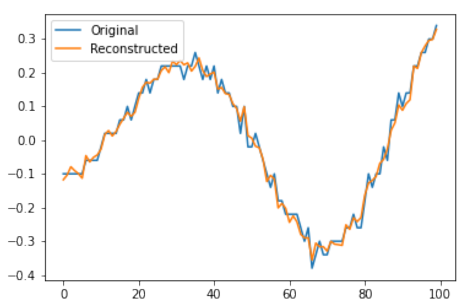


Fig. 13: Damaged panel with 50% measurements

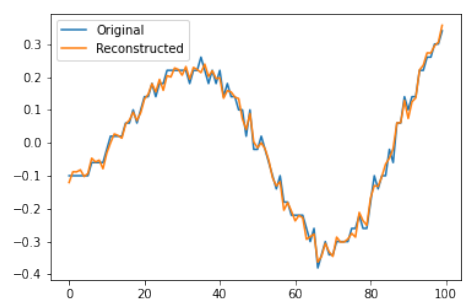


Fig. 14: Damaged panel with 80% measurements

using 80% of the measurements and cross-validation error using the rest). However, we observed that our GMM was modelling the river hole response very well and hence, the reconstruction error was low (comparable to the reconstruction error of the healthy panel response), and so was the cross-validation error. Hence, our method couldn't detect a drift in this experiment. However, the reconstructions error were low. The PSNR for reconstruction at different temperatures is at 2. We observe that the PSNR at each compression ratio is comparable (and better in some cases!) to that of the reconstructions of the healthy compressed samples. The following are some reconstructions of the compressed healthy 12 and the rivet hole 14 response. To try and detect the drift, we tried increasing the patch size to $d = 300$ to check for differences in a larger time scale. However, this resulted in some of the covariance matrices of the trained GMM being singular, and reconstruction is not possible in this case, as we need the covariance matrix to be invertible.

VI. Conclusion & Future Scope

In this report, we presented a model of a transmitter-receiver system which uses compressed sensing to reduce the bandwidth requirements. By modelling the consecutive samples of the signal as being drawn from a GMM, we could use Statistical Compressed Sensing to get accurate reconstructions of the signals. We also present the experiments done to validate the theoretical results. Finally, we looked at the method presented in [4] to update our source estimate using the compressed measurements.

There need to be improvements in the method to detect drift since the choice of threshold in our experiments is ad-hoc. Also, we need to prove that cross-validation is a good measure to detect drift.

ACKNOWLEDGMENT

I would like to thank Professor Ajit Rajwade and Professor Nikhil Karamchandani for their guidance and continuous feedback and hope to continue working under their guidance in the future. I would also like to thank Professor Siddharth Tallur for sharing the dataset of the experiments conducted on the Aluminium Panel. Finally, I would like to thank Amol Girish Shah, who is working with me on this project, for providing important insights on some topics.

REFERENCES

[1] E. Candès, J. Romberg, and T. Tao, ”Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information” *IEEE Trans. Inform. Theory, vol. 52, no. 2, pp. 489–509*, Feb. 2006

[2] D. L. Donoho, “Compressed sensing” *IEEE Trans. on Info. Theory, vol. 52, no. 4, pp. 1289–1306*, April 2006.

[3] Guoshen Yu and Guillermo Sapiro ”Statistical Compressive Sensing of Gaussian Mixture Models”, *arXiv:1010.4314 [cs.CV]*, 2011

[4] J. Yang et al., ”Compressive Sensing by Learning a Gaussian Mixture Model From Measurements,” *IEEE Transactions on Image Processing, vol. 24, no. 1, pp. 106-119*, Jan. 2015

[5] Vladislav Voroninski and Zhiqiang Xu. ”A strong restricted isometry property, with an application to phaseless compressed sensing”. *Applied and Computational Harmonic Analysis*, 40(2):386–395, 2016

[6] Guoshen Yu, Guillermo Sapiro, and Stéphane Mallat. ”Solving inverse problems with piecewise linear estimators: From gaussian mixture models to structured sparsity”. *IEEE Transactions on Image Processing*, 21(5):2481–2499, 2011

[7] Ajit Rajwade, Nimay Gupta. ”Signal compression using probability density function estimation”, BTP Report, 2022.

[8] S. Park and H. -N. Lee, ”On the derivation of RIP for random Gaussian matrices and binary sparse signals,” *ICTC 2011, 2011, pp. 120-124*, doi: 10.1109/ICTC.2011.6082562.

[9] Jonathan Monsalve, Juan Ramirez, Iñaki Esnaola, and Henry Arguello. ”Covariance esti- mation from compressive data partitions using a projected gradient-based algorithm”. *IEEE Transactions on Image Processing*, 31:4817–4827, 2022