

A Low-Complexity Biosignal Recovery Algorithm using Compressed Sensing and Independent Component Analysis

Yahia Alghorani, *Member, IEEE*, and Salama S. Ikki, *Senior Member, IEEE*

Abstract—The aim of this study is to propose a low-complexity algorithm that can be used for the joint sparse recovery of biosignals. The framework of the proposed algorithm supports real-time patient monitoring systems that enhance the detection, tracking, and monitoring of vital signs via wearable biosensors. Specifically, we address the problem of sparse signal recovery and acquisition in wearable biosensor networks, where we develop an efficient computational framework using compressed sensing (CS) and independent component analysis (ICA) to reduce and eliminate artifacts and interference in sparse biosignals. Our analysis and examples indicate that the CS-ICA algorithm helps to develop low-cost, low-power wearable biosensors while improving data quality and accuracy for a given measurement. We also show that, under noisy measurement conditions, the CS-ICA algorithm can outperform the standard CS method, where a biosignal can be retrieved in only a few measurements. By implementing the sensing framework, the error in reconstructing biosignals is reduced, and a digital-to-analog converter operates at low-speed and low-resolution.

Index Terms—Patient monitoring systems, joint sparse recovery for biosignals, wearable biosensor networks.

1 INTRODUCTION

RECENT developments in patient monitoring systems have led to important applications in the field of biomedical engineering. Wearable devices integrated with cyber-physical systems (CPS) make the medical monitoring of patients with chronic conditions easier and more consistent, allowing clinicians and caregivers to supervise their patients remotely and providing feedback to help maintain an optimal health status, regardless of patient location. These systems are being increasingly used in hospitals and clinics to provide continuous high-quality care to patients in complex clinical scenarios [1].

Recent advances in wearable sensors make CPS a powerful candidate for real-time e-health monitoring, extending to different monitoring areas, including homes, buildings, means of transport, etc. Unlike traditional embedded systems, CPS is typically designed as a network of real-time embedded computing interaction with physical elements that can ensure the adaptability, autonomy, reliability and functionality of wearable biosensor networks. An essential part of CPS is the internet of things (IoT) edge computing platforms, which can enable smart mobile healthcare services, through which patient data is collected by wearable biosensors and clinical data aggregators (e.g., smart watches and mobile devices) to be transferred to ambient sensors and then stored in medical servers for monitoring health status. The IoT edge computing platform consists of a large number of real-time data aggregators and ambient sensors that collect large amounts of data from patients in different locations and make them accessible to clinicians at any time for analysis (see, e.g., Fig.1).

Big data analytics for CPS help clinicians predict illness, prepare diagnosis, and plan treatment, resulting in overall improved quality of care and reduced cost. By exchanging medical records for patients between public and private hospitals, doctors and specialists would be able to predict where the patient is located on the spectrum of disease progression more accurately and efficiently. Managing and monitoring vital signs [2], such as heart and respiratory rates, blood pressure, blood flow, blood glucose, body temperature, oxygen saturation, electroencephalogram (EEG), electrocardiogram (ECG), requires wearable sensing platforms that can capture vital signs and biometrics, and deliver data from the patient to IoT edge computing devices (e.g., data aggregators, ambient sensors), as well as to the CPS cloud, for medical analysis.

The main challenge is to implement remote monitoring and tracking of patients in sensors and data acquisition/detection, i.e., when vital signs (biosignals) contain noise and artifacts. For example, ECG signals are affected by patient motion and they often suffer from low signal-to-noise ratio (SNR) due to motion artifacts and interference effects. Artifacts can be defined as distorted signals caused by internal or external sources [e.g., muscle movement, overlapping of data transmission where inter-biosensor interference occurs within the same wearable biosensor network] or inter-network interference where the wearable biosensor network interferes with other nearby wireless sensor networks operating in the industrial, scientific, and medical (ISM) radio bands [3].

Due to ECG artifacts, high data acquisition is invisible to wearable biosensors, resulting in inadequate diagnosis and treatment. With a complete set of n discrete-time samples of a biosignal, the design of physical sampling devices such as digital-to-analog converters (DACs) and analog-to-digital (ADCs) becomes more complicated for wearable biosensors.

• Y. Alghorani and S. Ikki are with the Department of Electrical Engineering, Lakehead University, Thunder Bay, On, Canada (e-mail: yahia.alghorani@ieee.org; sikki@lakeheadu.ca).

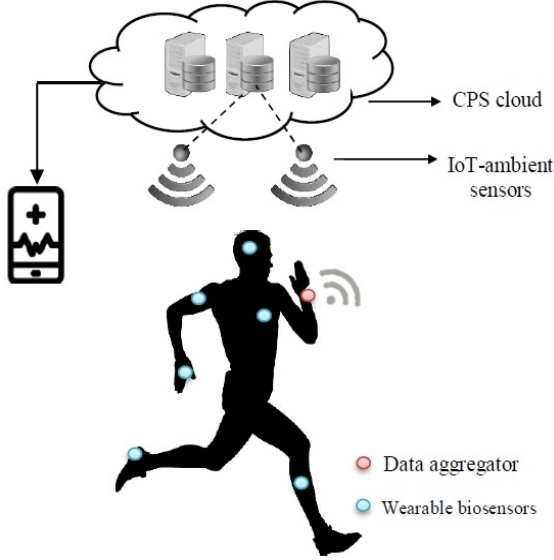


Fig. 1: IoT edge computing platform for continuous patient monitoring, where vital signs are collected through wearable biosensors and then send to the data aggregator (e.g., smart watch/mobile device) which in turn transmits data flows to ambient sensors and medical servers for analysis

Another challenge presented by patient monitoring systems is related to the large energy consumption of wearable biosensors and data aggregators due to continuous monitoring; therefore, the goal of this work is to design and implement a low-complexity computation algorithm that can eliminate noise and artifacts in wearable biosensor networks at a low-cost hardware implementation and power consumption.

The mutual interference between wearable biosensors (e.g., biosensor-to-biosensor interference) due to the overlapping of physiological data transmission can reduce the received signal strength, which may result in significant degradation of signal detection. In other words, the presence of noise and interference in the sparse biosignals requires an increase in the number of measurements needed by compressed sensing (CS) to improve the quality of the reconstructed image/signal, making the resolution of the sampling devices high, i.e., high-complexity acquisition/detection systems [4], [5].

In order to reduce the number of noisy biosignal measurements and obtain a high-resolution biosignal, we perform independent component analysis (ICA) [6]. ICA is a computational method that separates a multivariate signal into independent subcomponents, and is mainly used to remove artifacts from EEG recordings. The primary driving force behind the use of ICA in patient monitoring is that solutions are required to address the data sparsity¹ problem in the presence of noise and interference. Due to biosignal artifacts, high-resolution and high bit-depth DACs/ADCs are needed to restore biosignals.

In this paper, we propose an innovative approach to address the problem of sparse signal recovery and acquisition in wearable biosensor networks. The approach proposed in the block diagram is summarized in Fig.2. Since ICA recovery algorithms are powerful tools even in noisy environ-

ments, we aim to use these algorithms to remove noise and artifacts from sparse biosignals, so that the CS can reduce the number of measurements (i.e., sampling rate) needed to retrieve viable biosignals and use low-speed and low-resolution sampling devices (i.e., low-cost implementation and power consumption of patient monitoring systems). Our analysis shows that CS-ICA can perform better than conventional CS when the sparse biosignal contains noise and artifacts, improving the quality of the reconstructed biosignal.

2 RELATED WORK

2.1 CS

In recent years, there has been a research focus on CS applications in wireless body sensor networks and telemonitoring purposes (see, for example [7-10]), where the CS hardware architecture has been divided into two main methods: analog and digital CS. For instance, in [7], the digital CS method (where the linear CS compression is applied after the ADC) was used to recover fetal ECG signals. In [11] and [12], the analog CS method has been applied to ECG, where the compression occurs in the analog sensor read-out electronics prior to the ADC. In spite of the fact that the analog CS method reduces the cost and power consumption of sampling devices compared with the digital CS method, its demonstration still requires extensive work on the analog sensor read-out electronics. Hence, in the current work, we propose to use the digital CS (as described in Fig.2).

2.2 ICA

In medical applications, it is common practice to use the digital ICA algorithm to restore EEG signals (where the ADCs are located at the receiver and data is processed in the digital domain [13], [14]). This type of algorithm requires high-speed sampling rate (due to the use of full measurements: $m \geq n$) to find the independent components/EEG signals, resulting in high-cost and high-power devices. On the other hand, few studies have discussed the analog VLSI implementation of ICA algorithms [see [15] for intelligent hearing aid applications] where the low-energy ICA architecture has been proposed in a noisy and reverberant environment, and experiments have shown a clear separation and precise localization of two speech sources.

However, both CS and ICA methods have some limitations when used alone, and pre/post-processing techniques must be adopted. Some studies have recently proposed new frameworks for the joint CS-ICA recovery to remove artifacts from EEG signals; see, e.g., [16], [17], where the CS method is preceded by the ICA data processing method to reduce power consumption of the sensing unit (wearable biosensor network). It is noteworthy that the implementation of ICA prior to CS can lead to a power consumption problem because full measurements are used by the digital ICA algorithm to reconstruct biosignals, giving rise to additional power-hungry components in the sensing unit [18].

Another recent study discussed this issue (see, e.g., [19]), where the ICA is performed after the CS. There are drawbacks to this method because two important factors have been neglected: 1) The use of a Gaussian random

1. Sparsity is defined as having a small number of coefficients that capture most of the information contained in a biosignal.

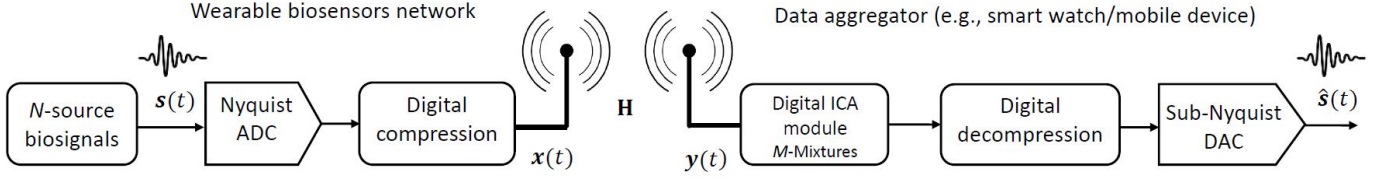


Fig. 2: Proposed Framework

matrix $\Phi \in R^{m \times n}$ in the CS-based sensing unit to measure source biosignals usually following non-Gaussian distributions (e.g., sub-Gaussian, super-Gaussian, or mixed super-, sub-, and Gaussian distributions) [20], and 2) Increased power consumption of the sensing unit as the number of noisy measurements (m) increases in the measurement matrix Φ due to artifacts and interference that exist in source biosignals. In order to bridge the gap between the CS and ICA methods in inter-biosensor interference scenarios, we provide a new co-design platform for CS-ICA algorithms that offers significant energy savings compared to other state-of-the-art methods. The proposed framework enables efficient implementation in CS hardware and provides valuable guidelines for communications engineers/researchers working on the design of sparse recovery algorithms for wearable biosensor networks. Specifically, the paper's contributions are summarized as follows:

- Developing low-cost, low-power wearable biosensors through digital CS-ICA algorithms, where we address the sparse signal recovery problem for inter-biosensor interference scenarios, and reconstruct a biosignal with only a handful of samples or measurements needed ($m \geq 2 \log_2(\frac{n}{K}) / \log_2(M \cdot SNR)$ samples) compared to conventional CS methods.
- Unlike conventional CS recovery methods, the CS-ICA algorithms reduce biosignal reconstruction errors and provide a low mean squared error compared to standard CS methods (where the noise power in the collected data is reduced to up $\frac{\epsilon}{M}$ when reconstructing a biosignal).

3 PROBLEM FORMULATION

In order to develop an energy-efficient sensing framework for patient monitoring systems, we outline the steps of CS used in noisy measurements, where source biosignals are sparse in the time domain (i.e., the K -source biosignal vector s_K contains K non-zero elements and satisfies $\|s\|_{l_0} \leq K \ll n$). If the source biosignal $s \in R^n$ is not sparse, we can make it so through CS, using an inverse discrete cosine transform (IDCT) matrix to produce a sparse vector; that is, $x = \Psi s$, where $\Psi_{n \times n}$ is a unitary matrix that can discard the small coefficients of s (i.e., many coefficients are set to zero via Ψ after adding a quantization step to the IDCT). With CS, we can employ low-speed DACs (i.e., sub-Nyquist sampling rates) to recover the biosignals and reduce the power consumption and cost of wearable biosensors/data aggregators.

3.1 CS Acquisition

In e-health monitoring applications, where interference can occur between N wearable biosensors (see, e.g., Fig.3), the

TABLE I
NOTATION

| | |
|---------------------|---|
| K | Small number of non-zero elements (coefficients) that capture most of the information contained in a biosignal. |
| n | Number of coefficients/samples in a biosignal |
| m | Number of measurements/samples for a biosignal |
| N | Number of wearable biosensors |
| M | Number of sensors for the data aggregator, $M \geq N$. |
| s_r | $s_r = [s_1, \dots, s_n]$ denotes the r -th source biosignal of size n |
| $\Psi_{n \times n}$ | $\Psi_{n \times n}$ is a unitary matrix that can discard the small coefficients of s_r |
| x_r | x_r denotes the r -th sparse biosignal of size n , defined as $x_r = \Psi_r s_r$ |
| Φ_r | The measurement matrix, $\Phi_r \in R^{m \times n}$, $n \gg m \gg K$ |
| A_r | The sensing matrix, defined as $A_r = \Phi_r \Psi_r$, $A_r \in R^{m \times n}$ |
| h_r | Constant channel vector for the r -th source biosignal, $h_r \in R^{M \times 1}$ |
| w_r | The digital combiner, $w_r \in R^{1 \times M}$ |
| n_r | Additive white Gaussian noise for the r -th sparse biosignal |
| y_r | Receiving signal for the r -th biosensor |
| \hat{x}_r | The estimated sparse biosignal, defined as $\hat{x}_r = w_r y_r$ |
| H | The unknown mixing matrix of source biosignals, $H \in R^{M \times N}$ |
| W | The unmixing matrix, defined as $W = H^T$, $W \in R^{N \times M}$ |
| $\ s_r\ _{l_1}$ | Standard l_1 norm of s_r , $\ s_r\ = \sum_{k=1}^n s_{r,k} $ |
| $\ s_r\ _{l_2}$ | Standard l_2 norm of s_r , $\ s_r\ ^2 = \sum_{k=1}^n s_{r,k} ^2$ |
| $H(\hat{x}_i)$ | Differential entropy of the i -th sparse biosignal \hat{x}_i , defined in (8). |

receiving sparse signal for each biosensor, $y_r \in R^{M \times n}$, $r \in \{1, 2, \dots, N\}$, at the M -sensor array (mixtures) of data aggregator (where $M \geq N$), is expressed as

$$y_r = \sum_{i=1}^N h_r x_i^T + n_r, \quad (1)$$

where $x_i = \Psi_i s_i$ represents n -pixel image/video/radio pulses, which are corrupted by additive white Gaussian noise (AWGN) $n_r \in R^{M \times n}$, $h_r \in R^{M \times 1}$ is a constant channel vector with array elements $h_{r,j}$, which depends on the distance between the r -th biosensor and the j -th sensor node of the data aggregator. The received signal is then processed by a digital combiner $w_r \in R^{1 \times M}$ to obtain $\hat{x}_r = w_r y_r$, which can be rewritten as

$$\hat{x}_r = \underbrace{w_r h_r x_r^T}_{\text{desired signal}} + \underbrace{\sum_{i \neq r}^N w_r h_r x_i^T}_{\text{interference}} + \underbrace{w_r n_r}_{\text{noise}}, \quad (2)$$

where $w_r = h_r^T$. After extracting the desired sparse biosignal, \hat{x}_r , and using the CS, the received signal is expressed as $z_r = \Phi_r \hat{x}_r$, namely

$$z_r = A_r s_r + v_r, \quad (3)$$

where $z_r \in R^m$ is the measurements vector for the biosensor r , $A_r = \Phi_r \Psi_r$ is the sensing matrix, $\Phi_r \in R^{m \times n}$ is the measurement matrix (where $n \gg m \gg K$), and $v_r \in R^m$ is the effective noise vector that contains the noise and interference component. Here, the sensing matrix $A_r \in R^{m \times n}$ obeys the restricted isometry property (RIP), i.e.,

$$(1 - \delta_K) \|s_r\|_{l_2}^2 \leq \|A_r s_r\|_{l_2}^2 \leq (1 + \delta_K) \|s_r\|_{l_2}^2, \delta_K \in (0, 1),$$

when δ_K is not too close to one and its entries are drawn from a suitable distribution, e.g. a Gaussian distribution [5]. So, if the RIP of order K is established, then it is sufficient to have $m = O(K \log_2 n / K)$ samples.

3.2 CS Recovery

In order to recover s_r from z_r , we consider the following optimization problem

$$\min_{s_r} \|s_r\|_{l_1} \text{ s.t. } \|A_r s_r - z_r\|_{l_2} \leq \epsilon, \quad (4)$$

where $\|v_r\|_{l_2} \leq \epsilon$ is the maximum noise power, $\|s_r\|_p$ stands for the standard l_p -norm and $\|s_r\|_{l_0}$ counts the number of nonzero elements in s_r . If we assume that $\delta_{2K} < \sqrt{2} - 1$, the solution to the convex problem above is satisfied by

$$\|\hat{s}_r - s_r\|_{l_2} \leq C_0 \cdot \|s_r - s_{r,K}\|_{l_1} / \sqrt{K} + C_1 \cdot \epsilon, \quad (5)$$

where the constants C_0 and C_1 are typically small [21, eq. (14)]. The result in (5) indicates that the noise power in the data can significantly increase the reconstruction errors, which means that we need more measurements (m -samples) to recover the source biosignal vector s_r . An equivalent expression has been determined by [22], $\|\hat{s}_r - s_r\|_{l_2} < \|v_r\|_{l_2} + \epsilon / (1 - \delta_{2K})$.

Note that when δ_{2K} approaches one, the reconstruction process becomes infeasible because of the high error rate. It should be noted that the RIP works well when there is no noise, where the reconstruction of the biosignal is more accurate (i.e., $\|\hat{s}_r - s_r\|_{l_1} \leq C_0 \cdot \|s_r - s_{r,K}\|_{l_1}$, [5, eq. (10)]). Although the CS technique resolves the sparse data acquisition problem, there are still some conditions and limitations that should be considered when dealing with scenarios of noise and interference (i.e., it is impossible to construct the sparse vector s_r when the effective noise level, v_r , is high). Therefore, the precise calculation of the digital combiner elements w_r is important to reduce the high noise level of v_r so that the biosignal is retrieved in relatively few measurements.

4 PROPOSED FRAMEWORK

ICA was originally developed as a computational method that can be used in various applications including medical signal and image processing (with a non-Gaussian input) [6], where a multivariate biosignal decomposes into independent additive subcomponents without prior knowledge of source biosignals or mixing coefficients. In other words, if we have a mixture of N independent source biosignals,

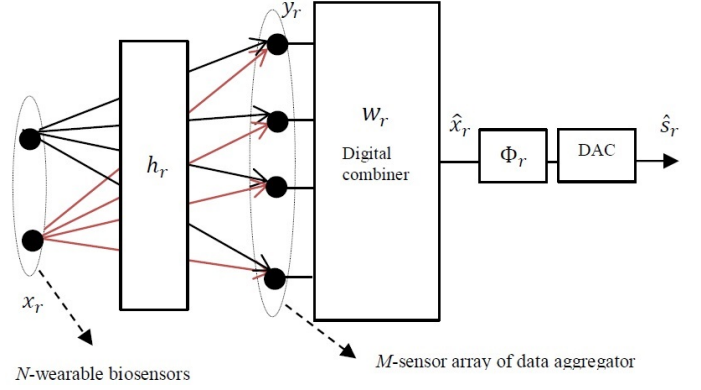


Fig. 3: Biosensor-to-biosensor interference scenarios where biosignals are retrieved through standard digital CS methods.

the ICA algorithm works to find the unmixing matrix $W \in R^{N \times M}$, to extract the source biosignal and remove artifacts from EEG recordings, see Fig.2, where the source biosignals are collected by wearable biosensors and compressed by the digital CS model to be transferred to the data aggregator (e.g., smart watch, mobile device). Here, the ICA scheme is performed to extract the sparse biosignals which are then decompressed by the digital CS model.

Using the joint CS-ICA recovery, we can employ low-speed and low-resolution DACs (i.e., sub-Nyquist sampling rates and low bit-depths) to restore biosignals and reduce the power consumption and cost of wearable biosensors/data aggregators. Linear receivers such as zero-forcing and minimum mean squared error equalizers can offer a fair trade-off between performance and complexity. The main drawbacks of these receivers are twofold: 1) Any defects in the estimation of channel and covariance matrices in a large network of biosensors result in residual interference in the data aggregator output, which in turn can lead to performance degradation. 2) There is an increase in the number of iterations needed for convergence. Therefore, the ICA algorithm is suggested as an alternative if the channel state information is unknown, and a fast convergence multi-biosensor detection scheme can be achieved by extracting all independent components directly. Suppose now that the data aggregator observes the mixture signal $y \in R^{M \times 1}$ in the time domain,

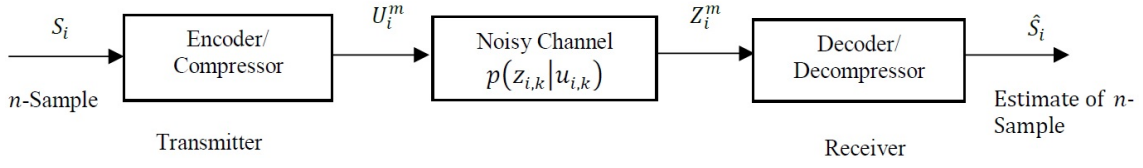
$$y(t) = Hx(t) + n(t), \quad (6)$$

where $n \in R^{M \times 1}$ is the Gaussian noise vector, $H \in R^{M \times N}$ is the unknown mixing matrix of source biosignals $s \in R^{N \times 1}$, and $x(t) = \Psi s(t)$ is the sparse vector, such that

$$x(t) = [x_1(t), \dots, x_N(t)]^T \\ s(t) = [s_1(t), \dots, s_N(t)]^T,$$

where the discrete-time representation of a continuous-time signal, which lies at the heart of analogue to digital conversion [23], defined as $x_i(t) = \Psi_i s_i(t)$, $s_i \in R^n$ is a source biosignal, and $\Psi_i \in R^{n \times n}$ is the unitary matrix generated by the IDCT. We define t as a time index, where $1 \leq t \leq T$.

The digital ICA algorithm estimates the sparse vector $x \in R^{N \times 1}$ using the unmixing matrix W , that is, $\hat{x}(t) = Wy(t)$, where $W = H^T$ is approached by maximum-likelihood estimation or FastICA algorithms [6]. Further-

Fig. 4: Block diagram for an n -sample compression

more, to eliminate the noise components $Wn(t)$, we use an adaptive filter (e.g., Least Mean Square (LMS) methods) to accomplish this task.

Assuming the independence of sparse biosignals x_i , the joint probability density function (PDF) of continuous random variables $P(X_1, \dots, X_N)$ is factorized as

$$p_{X_i}(x(t)) = \prod_i p_{X_i}(x_i(t)). \quad (7)$$

In order to maximize the statistical independence of extracted components (i.e., sparse biosignals), we need to minimize the mutual information of estimated signals \hat{x}_i and maximize non-Gaussianity, i.e., $I(\hat{X}_1; \dots; \hat{X}_N) = \sum H(\hat{X}_i) - H(\hat{X}_1, \dots, \hat{X}_N)$, where the differential entropy $H(\hat{X}_i)$ of a continuous random variable \hat{X}_i with a density function $p(\hat{x}_i)$, is defined as [24, Sec. 8.1]

$$H(\hat{X}_i) = - \int_S p(\hat{x}_i) \log p(\hat{x}_i) d\hat{x}_i, \quad (8)$$

where S is the support set of the variable \hat{X}_i , and the joint differential entropy of multiple continuous random variables $\hat{X}_1, \dots, \hat{X}_N$ is defined as [24, Sec. 8.4]

$$H(\hat{X}_1, \dots, \hat{X}_N) = - \int p(\hat{x}_1, \dots, \hat{x}_N) \times \log p(\hat{x}_1, \dots, \hat{x}_N) d\hat{x}_1 \dots d\hat{x}_N. \quad (9)$$

(9) can also be rewritten as $H(\hat{X}_1; \dots; \hat{X}_N) \leq \sum H(\hat{X}_i)$ with equality if $\hat{X}_1, \dots, \hat{X}_N$ are independent.

Now let $\hat{X} = WY$; using the differential entropy property: $H(WY) = H(Y) + \log |W|$, we obtain

$$I(\hat{X}_1; \dots; \hat{X}_N) = \sum H(\hat{X}_i) - H(Y_1, \dots, Y_M) - \log |W|. \quad (10)$$

In order to measure the difference in entropy between a given distribution and Gaussian distribution (containing the highest entropy) with the same mean and variance, we use negentropy, defined as $J(\hat{X}_i) = H(\hat{X}_g) - H(\hat{X}_i)$, where $H(\hat{X}_g)$ is the entropy of the Gaussian distribution with unit variance for all estimated sparse biosignals \hat{x}_i .

In information theory and statistics, minimization of mutual information between multiple random variables \hat{X}_i is achieved by minimizing entropy $H(\hat{X}_i)$ or maximizing negentropy $J(\hat{X}_i)$, which is also equivalent to minimizing Gaussianity. Therefore, finding the optimal unmixing matrix W will help to minimize the mutual information between the variables \hat{X}_i and make the extracted sparse biosignals uncorrelated (independent) and non-Gaussian. Once the

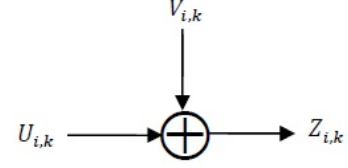


Fig. 5: Gaussian Channel

ICA extracts the sparse biosignals with improved SNRs, the CS senses and acquires each sparse biosignal, so the number of measurements m is minimized, and the biosignal is sampled by few measurements at a rate below Nyquist standards.

Suppose the sparse biosignals are independent and identically distributed (i.i.d) random variables (i.e., all sparse biosignals X_i have the same probability distribution and are mutually independent), then the sparse biosignal observed by the CS model can be expressed as

$$z_i(t_k) = u_i(t_k) + v_i(t_k), \quad (11)$$

where $u_i(t_k) = \Phi_i \hat{x}_i(t_k)$ is the received sparse biosignal vector, $\hat{x}_i(t_k)$ is the estimated sparse vector that corresponds to the i -th row of $Wy(t_k)$, and $v_i(t_k)$ is the i -th residual gaussian noise vector after filtering the noise component $Wn(t_k)$. Note that $t_k, k \in \{1, 2, \dots, m\}$ are the time instants when samples (m -snapshots) are taken.

In order to reconstruct the sparse biosignal x_i , we fix Ψ_i and pick up the input of the i -th sampling matrix Φ_i randomly from a non-Gaussian distribution with zero mean and variance $1/m$ (e.g., sub-Gaussian distribution [25]). Here is typically a flat PDF with a strong tail decay property, so that

$$E[e^{-\lambda u}] \leq e^{-\lambda^2 \rho^2 / 2}, \text{ for all } \lambda \geq 0, \quad (12)$$

where $E[U] = 0$, $\text{Var}[U] \leq \rho^2$, and the sensing matrix $[A_i = \Phi_i \Psi_i]$ satisfies the RIP condition (i.e., $m \geq c.K \cdot \log_2 n/K$, where c is a positive constant). Here, the digital CS-ICA algorithm takes advantage of the RIP property by providing minimum and maximum power for the samples, ensuring that the K -sparse vectors do not fall into the null space of the sampling matrix Φ_i , which in turn provides a stable recovery for the sparse biosignals.

5 GENERALIZED RATE OF CS-ICA ALGORITHMS

The sparse recovery algorithm works to reduce interference and noise, where the generalized rate R_i of independent sub-Gaussian random variables $U_{i,k} \sim N(0, \rho_{i,k}^2)$, is calculated in the following way:

Proposition. Suppose $U_i^m = (U_{i,1}, \dots, U_{i,m}) \in U_i^m$, and $Z_i^m = (Z_{i,1}, \dots, Z_{i,m}) \in Z_i^m$ are an m -tuple of the random

variables U_i and Z_i , $i \in \{1, \dots, N\}$, which are generated from the CS and data compression process. Now consider m independent Gaussian channels in parallel, in which we can send the non-Gaussian input u_i^m through m -AWGN channels v_i^m with power constraint $\frac{1}{m} \sum_{k=1}^m P_{i,k} \leq \frac{P}{N}$ and noise variance $\sigma_{i,k}^2$. The digital ICA combiner output is expressed as $z_i^m = u_i^m + v_i^m$ with power gain $\|h_i\|^2 = \sum_{j=1}^M |h_{ij,k}|^2$, h_i is the i -th column of the unknown mixing matrix H . Thus, if $U_{i,k}$ and $V_{i,k}$ are independent random variables, then

$$R_i = I(Z_i^m; U_i^m) \leq \sum_k \frac{1}{2} \log_2 \left(1 + \frac{P_{i,k}}{\sigma_{i,k}^2} \cdot \|h_i\|^2 \right), \quad (13)$$

with equality if $Z_{i,k}$ are independent and sub-Gaussian random variables (i.e., $Z_{i,k} \sim N(0, \rho_{i,k}^2 + \sigma_{i,k}^2)$ and $\sigma_{i,k}^2 = E[V_{i,k}^2]$), and the SNR is obtained for the k -th sample as $SNR_{i,k} = P_{i,k}/\sigma_{i,k}^2$.

Proof. By calling the definition of the information capacity of parallel Gaussian channels [24, Sec. 9.4], we obtain

$$I(Z_i^m; U_i^m) \leq \sum_k H(Z_{i,k}) - H(V_{i,k}). \quad (14)$$

Given $H(Z_{i,k}) \leq \frac{1}{2} \log 2\pi e (\rho_{i,k}^2 + \sigma_{i,k}^2)$, $H(V_{i,k}) = \frac{1}{2} \log 2\pi e \sigma_{i,k}^2$, and $\rho_{i,k}^2 = P_{i,k} \cdot \|h_i\|^2$, (13) can be proven.

Special cases:

- 1) Since the ICA algorithm mitigates interference between source biosignals, we assume that a) the received SNR is high enough so that we can distribute equal amounts of energy across channels using a *water-filling* solution (i.e., $P_{i,k} \leq P/N$), and b) with a coefficient magnitude of $|h_{ij,k}|^2 = 1$ (i.e., $\|h_i\|^2 = M$). Using Jensen's inequality, (13) can be expressed as

$$\begin{aligned} R &\leq \frac{1}{2} \log_2 \left(1 + \frac{1}{m} \sum_k \frac{P_{i,k}}{\sigma_{i,k}^2} \|h_i\|^2 \right) \\ &\leq \frac{1}{2} \log_2 (M \cdot SNR), \end{aligned}$$

where $SNR = P/N\sigma^2$.

- 2) On the other hand, assuming that the mixture signal y contains Gaussian and non-Gaussian random variables, we can maximize the entropies $H(\hat{X}_i)$ and minimize the negentropies $J(\hat{X}_i)$, hence finding an ideal unmixing matrix W becomes impossible due to Gaussian increase, leading to a low SNR (e.g., $\log(1+x) \approx x \log_2 e$ when x is sufficiently small), namely

$$R \approx \frac{1}{2} M \cdot SNR \cdot \log_2 e,$$

where $SNR = P/N\sigma^2$.

- 3) Assuming that the number of measurements is large enough (i.e., $m = n \rightarrow \infty$). By the strong law of

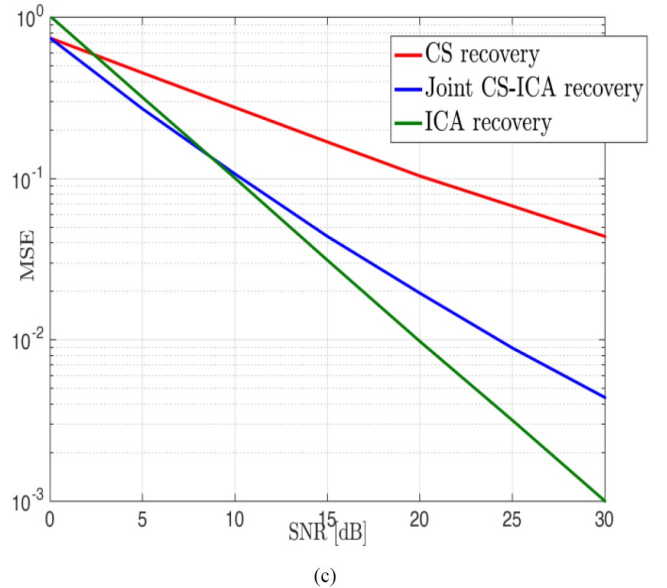
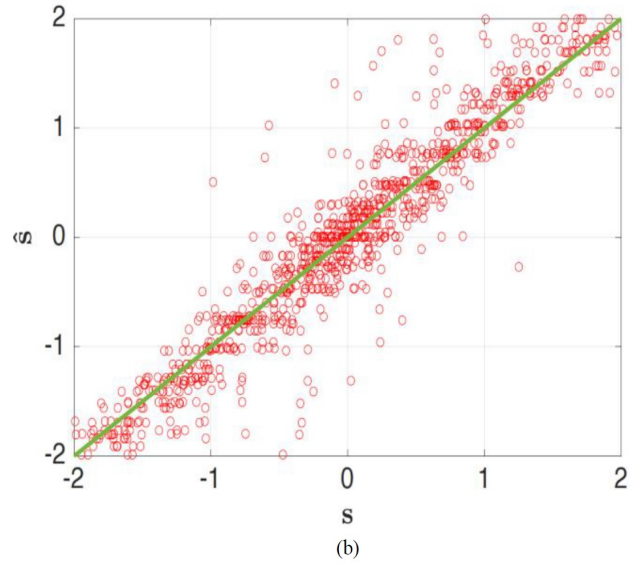
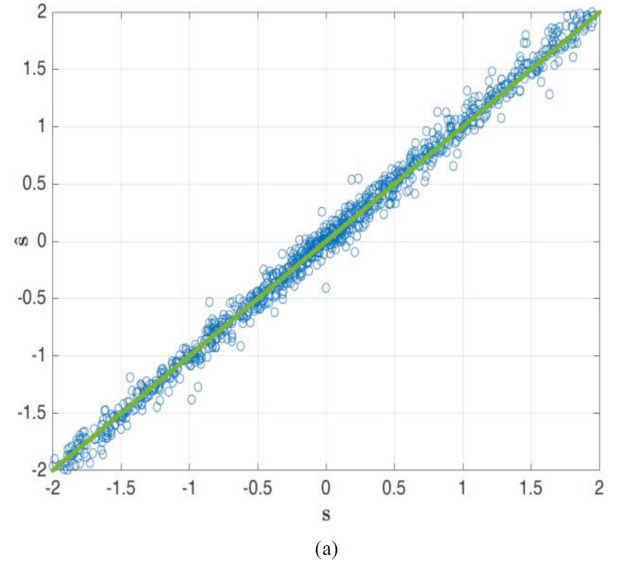


Fig. 6: The biosignal reconstruction errors for both CS-ICA (a) and CS (b) methods, where $N = 2$, $M = 2$. (c) Comparison of the mean squared error for various sparse biosignal recovery methods, where the biosignal is recovered at $n = 1000$, $m = 300$ and different SNR values.

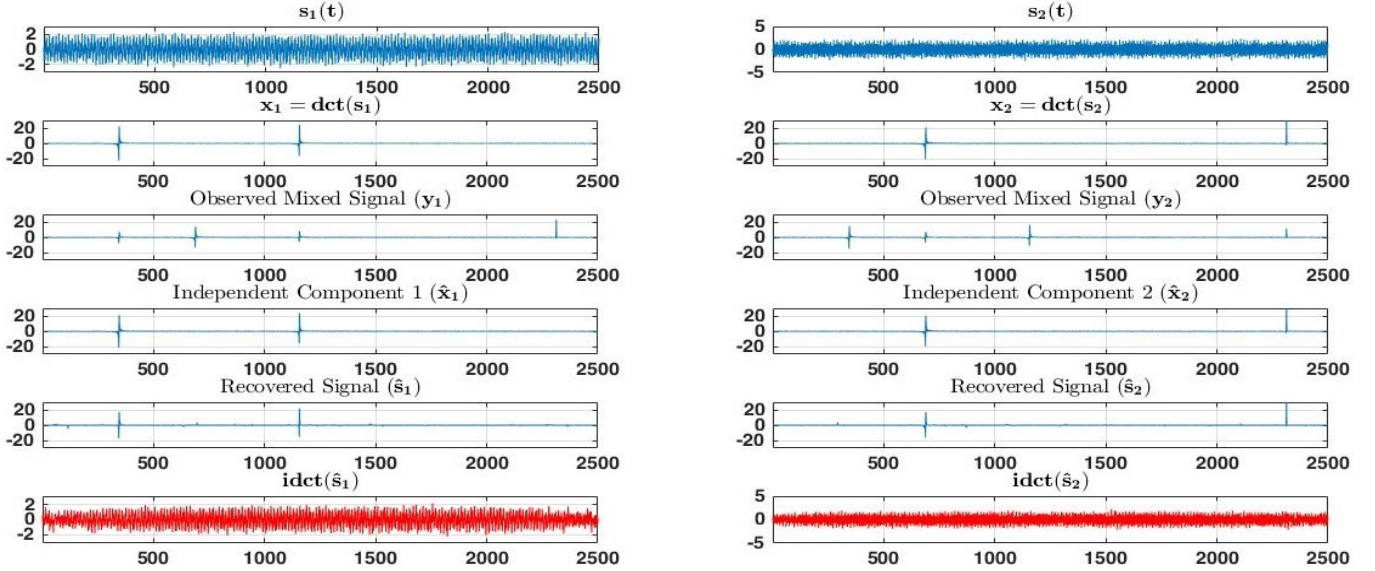


Fig. 7: The joint CS-ICA recovery for EEG $s_1(t)$ and EMG $s_2(t)$ signals. Both source biosignals are sparse in the frequency domain (e.g., the number of samples per sparse biosignal is $n = 2500$) and reconstructed in the time domain (e.g., $m = 170$) at an SNR level of 30dB.

large numbers, $\Pr \left[\lim_{m \rightarrow \infty} \frac{1}{m} \sum_k P_{i,k} = \frac{P}{N} \right] = 1$, (13) is calculated as

$$R \leq \frac{1}{2} \log_2 (1 + M \cdot SNR).$$

5.1 Channel Coding via CS-ICA Algorithms

To extract the performance gain of digital CS-ICA algorithms, we invoke the channel coding theorem for a discrete memoryless channel [24]. Suppose we connect through a noisy channel $(U_i^m, p(z_i^m | u_i^m), Z_i^m)$, where the input samples do not depend on the past output samples, that is to say, $p(z_{i,k} | u_i^k, z_i^{k-1}) = p(z_{i,k} | u_{i,k})$; $p(z_i^m | u_i^m) = \prod_{k=1}^m p(z_{i,k} | u_{i,k})$. The encoder (compressor) Ψ_i first takes a block of n samples and turns them into a K -sparse biosignal, where $\log_2(\frac{n}{K})$ bits are used to encode/compress the n -sample sequence $S_i = (S_{i,1}, \dots, S_{i,n})$, then the decoder (decompressor) does the inverse process. An $(2^{mR}, m)$ code for the channel $U_i^m, p(z_i^m | u_i^m), Z_i^m$ consists of a compressing function U^m and a decompressing function Z^m (see Fig.4). The rate of the code $(2^{mR}, m)$ is defined by $R = \log_2(\frac{n}{K})$, with a probability of error $P_e = \Pr(S_i \neq \hat{S}_i) < \epsilon$. So, if we assume that all entries of the source vector S_i are i.i.d random variables and the channel coefficients of the samples are defined as $|h_{i,k}|^2 = 1$, then the number of measurements for each source biosignal is calculated as

$$m \geq \frac{2 \log_2(\frac{n}{K})}{\log_2(1 + M \cdot SNR)}, \quad (15)$$

where $\log_2(\frac{n}{K}) = 2H(K/n) = 2K \log_2(\frac{n}{K})$ bits when $0 \leq K/n \leq 1$. Note that the ICA method brings an additional power gain to the CS method. The higher the SNR, the lower the number of measurements. We can also conclude that

$$m \geq 2 \log_2\left(\frac{n}{K}\right) / \log_2(M \cdot SNR),$$

when the number of ICA mixtures is very large i.e., $M \rightarrow \infty$, where $SNR = P/\sigma^2$. By assuming that the sensing matrix A_i satisfies the RIP property, $(1 - \delta_{i,K}) \|s_i\|_{l_2}^2 \leq \|u_i\|_{l_2}^2 \leq (1 + \delta_{i,K}) \|s_i\|_{l_2}^2$, where $u_i = \|h_i\|^2 A_i s_i$, then the receiver can reconstruct the n -sample sequence, $S_i = S_{i,1}, \dots, S_{i,n}$, in the digital domain by solving the convex problem

$$\hat{s}_i = \min_{s_i \in R^n} \|s_i\|_{l_1} \quad \text{s.t.} \quad \|h_i\|^2 A_i s_i - z_i\|_{l_2} \leq \epsilon. \quad (16)$$

By using the triangle inequality [21] and the theorem in (5) for $\|v_i\|_{l_2} \leq \epsilon = \sigma_i^2$, the solution to (16) obeys

$$\|\hat{s}_i - s_i\|_{l_2} \leq C_{i,0} \cdot \|s_i - s_{i,K}\|_{l_1} / \sqrt{K} + C_{i,1} \sigma_i^2 / \|h_i\|^2. \quad (17)$$

6 PERFORMANCE ANALYSIS

In order to gain further insight into the performance of digital CS-ICA algorithms, we display the error in retrieving source biosignals from noisy data as shown in Fig.6 (a) and (b), where the linear regression of modelling n data points is performed and each biosignal is restored from noisy measurements (e.g., $n = 1000$, $m = 250$, and $SNR = 25$ dB). Here, the simulation results show that the digital CS-ICA algorithms have fewer reconstruction errors than traditional digital CS methods, the reason for which is that the ICA algorithms can capture all n data points containing the K -coefficients and make the biosignal detectable at a high SNR, whereas traditional CS methods cannot guarantee whether the measured samples m contain all non-zero components K , which are randomly distributed across the sparse vector s_i and captured by increasing the number of measurements m .

Fig.6 (c) compares the performance of both methods in terms of the mean squared error (MSE) with standard (digital) ICA algorithms, where the CS-circuit components are omitted and the ICA scheme must perform full measurements to create W , i.e., $m = 1000$. The results indicate

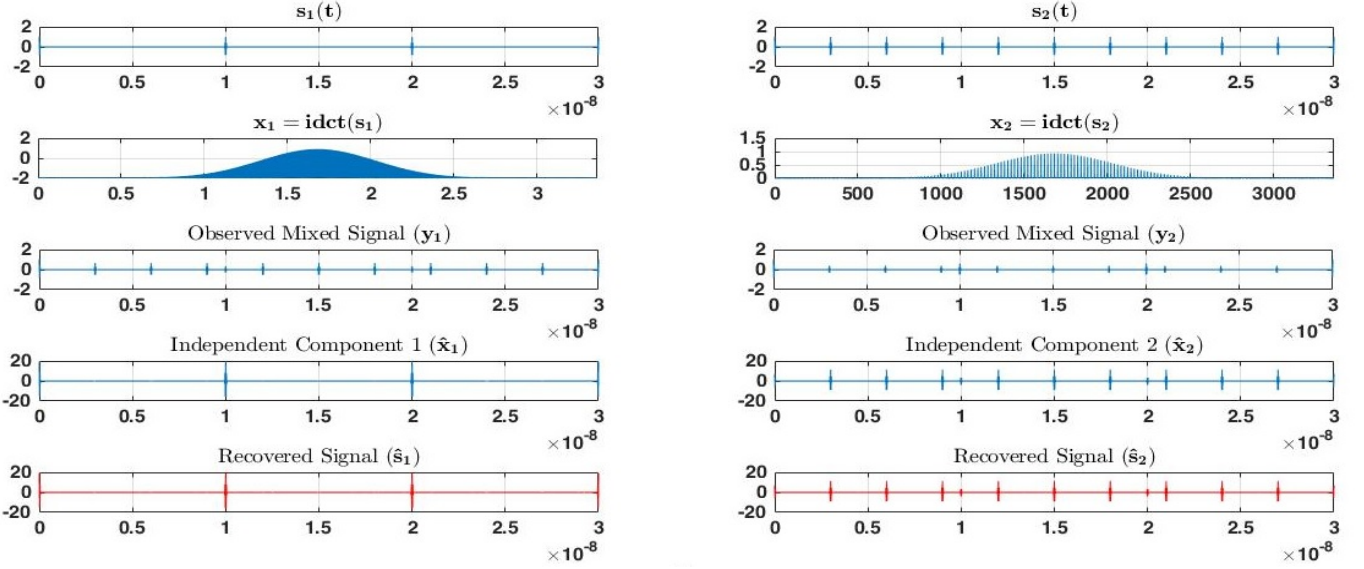


Fig. 8: Reconstruction of ECG signals at mm-wave, measured at SNR level of 20dB. The ECG signal is retrieved through the digital CS-ICA algorithms. In this example, the mm-wave pulses are sparse in the time domain for a certain period of 30 ns (i.e., $n = 3361$ samples after passing the source biosignal $s(t)$ through IDCT) and are recovered from just 300 snapshots.

that the performance of CS-ICA algorithms can be close to standard ICA algorithms for medium and high SNR regimes, both of which are roughly equivalent at low SNR values. We can consequently conclude that the digital CS-ICA algorithms can significantly reduce Gaussianity output because of the high level of additive Gaussian noise (i.e., $\|v_i\|_{l_2} \leq \sigma_i^2/\|h_i\|^2$).

According to the central limit theorem and the law of large numbers, when adding a large number of independent random variables (i.e., the number of wearable biosensors gets larger, $N \rightarrow \infty$), the total tends to a Gaussian distribution even if the original variables (source biosignals) are not normally distributed. Since the CS methods have inherently Gaussian characteristics, they can be used instead of the ICA algorithms to retrieve biosignals. Furthermore, the results demonstrate that the CS-ICA algorithms perform better than traditional CS methods when tested under the same conditions (e.g., $m = 300$).

In general, Fig.6 results can be treated as benchmarks for the CS-ICA algorithms. On one hand, the ICA can increase the sparse biosignal strength received (by reducing artifacts and interference) and use low-resolution DACs to restore the biosignal. On the other hand, the CS methods can reduce the sampling rate and make the DACs operate at low speed. This process is called an “exchange of interests” between CS and ICA methods.

7 EXPERIMENTAL RESULTS

Fig.7 shows the joint CS-ICA recovery for two source biosignals, e.g., EEG and electromyography (EMG) signals where $N = 2, M = 2$. The source biosignals are sparse in the frequency domain (e.g., the number of samples per source signal is $n = 2500$ and the number of spikes is $K = 2$) with various periods for each source biosignal (e.g., $T = 1$ second for the EEG signal $s_1(t)$ and $T = 2$ seconds for the EMG signal $s_2(t)$). The relatively few coefficients capture

most of the signal energy, e.g., $m = 170$, taken in the time domain. This example obviously shows the gain acquired by the digital CS-ICA algorithms, where the biosignals can be restored in only a few measurements.

Fig.8 compares the performance of CS-ICA algorithms and standard CS methods to two source ECG signals operating in the 28 GHz band [26]. The mm-wave pulse train is sparse in the time domain, e.g., the pulse repetition period (Δt) is 10 ns for the first ECG signal and 3 ns for the second ECG signal, $N = 2$, where each biosignal has the same frame size. For example, $T = 30$ ns for the ECG signals $s_1(t)$ and $s_2(t)$. The reason is due to the assumption that the mm-wave signals are sparse in the time domain rather than the frequency domain to simplify the construction of the IDCT matrix $\Psi_{n \times n}$ (where n is very large) required to produce the sparse source vectors $x_1(t)$ and $x_2(t)$. Here, the sampling rate is calculated as $f_s = 4 \times 28$ giga samples/sec, and the total number of samples (data points) taken to perform a Fourier transform, is $n = 3361$.

Moreover, in this example, assume that the wearable ECG sensors have line-of-sight (LOS) paths with the data aggregator sensors $M = 2$ (e.g., $SNR = 20$ dB), where each sparse ECG signal is retrieved from a few measurements that are captured in the frequency domain (e.g., $m = 300$). In this setup, the mixing matrix elements of $H \in R^{2 \times 2}$ are uniformly distributed and the unmixing matrix $W = H^T$ is calculated through the FastICA algorithms. Our experimental results show that the digital CS-ICA algorithms can provide a more accurate and rapid recovery of ECG signals compared to traditional CS methods. Another practical example of real-time patient monitoring systems can be found in wearable magnetoencephalography (MEG) scanners that are used for mapping brain activity.

Fig.9 illustrates a phantom recovery experiment that exactly mimics MEG images that can be contaminated with outside noise and artifacts. In this example, we restore a large image size of 256x256 pixels of noisy measurements

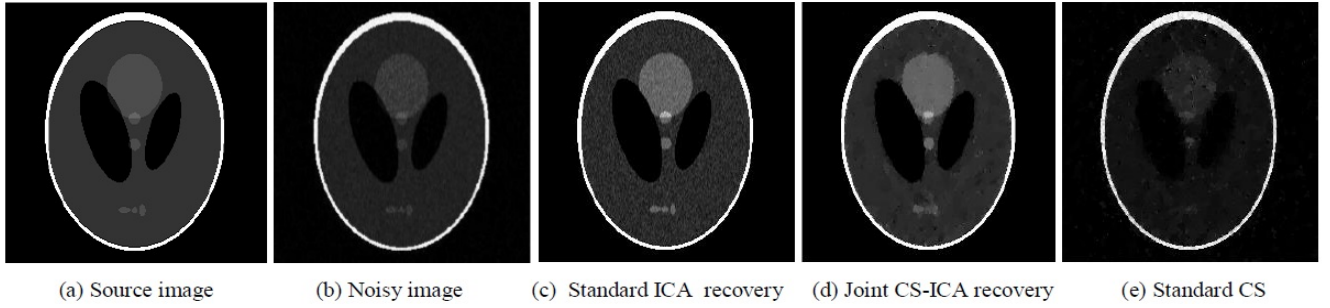


Fig. 9: Phantom recovery experiments at SNR 15dB, where the CS-ICA recovery and standard CS methods restore the MEG image at $n = 65,536$ samples and $m = 6213$ snapshots, while standard ICA methods perform accurate image recovery at $m = 65,536$ snapshots.

(e.g., $SNR = 15\text{dB}$), where the biosignal of interest s_1 is extracted from an unwanted biosignal s_2 (e.g., $N = 2, M = 2$). The sparse MEG image is recovered using the min-total variation (TV) ² method [27], with 25 non-zero frequency components distributed uniformly in the Fourier domain, where a few measurements are created to restore the desired MEG image (e.g., $m = 6213$). Compared with standard digital ICA methods, the CS-ICA recovery methods can reduce noise and number of measurements needed (e.g., from 65,536 to 6213 snapshots), allowing receivers to use low-speed/low-resolution DACs to retrieve the sparse source MEG signal. Furthermore, the experimental results show that the digital CS-ICA algorithm can achieve high-resolution of MEG images compared to standard digital CS method. However, it should be noted that standard digital ICA algorithms have poor resolution of highly correlated brain sources.

8 CONCLUSION

In this paper, we present a low-complexity algorithm that can address the sparse signal recovery and acquisition problem in wearable biosensor networks. Using digital CS-ICA implementation in wearable biosensor devices, we can reduce the number of measurements, increase the noise robustness, and improve the accuracy and efficiency of standard digital CS methods. Our results showed that the CS-ICA algorithms can perform better than standard CS methods when a biosignal contains noise and artifacts. In fact, we can recover K -sparse biosignals/images from just $m \geq 2 \log_2 \left(\frac{n}{K} \right) / \log_2 (M \cdot SNR)$ noisy measurements, thereby improving the quality of the reconstructed biosignals. Compared to standard ICA algorithms, the CS-ICA algorithms can reduce the sampling requirements for digital-to-analog converters as well as the computational complexity of recovery at mm-waves, so that we can reduce the power consumption of wearable biosensors and data aggregators, and retrieve biosignals in fewer measurements.

The proposed sensing framework will have a significant impact on the healthcare sector by improving the efficiency, reliability and accuracy of patients' continuous monitoring systems, resulting in better patient diagnosis and treatment options. The sensing method will generally provide significant environmental and economic benefits to

healthcare delivery systems, allowing clinicians to monitor patient conditions directly and lighten the burden of healthcare costs. For future work, it would be interesting to test out the proposed framework in real-life scenarios to analyze the objective computational advantages against the conventional approaches.

REFERENCES

- [1] S. K. Khaitan and J. D. McCalley, "Design techniques and applications of cyberphysical systems: A survey," *IEEE Syst. J.*, vol. 9, no. 2, pp. 350–365, Jun. 2015.
- [2] T.-P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, T. J. Sejnowski, "Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects," *Clin. Neurophysiol.*, vol. 111, pp. 1745–1758, 2000.
- [3] A. Lanata, E. P. Scilingo, E. Nardini, G. Loriga, R. Paradiso, and D. DeRossi, "Comparative evaluation of susceptibility to motion artifact in different wearable systems for monitoring respiratory rate," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 2, pp. 378–386, Mar. 2010.
- [4] Y. C. Eldar, "Sampling Theory: Beyond Bandlimited Systems," Cambridge University Press, Apr. 2015.
- [5] E. Candes and M. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
- [6] A. Hyvarinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural networks*, 13(4-5):411–430, 2000.
- [7] Z. Zhang *et al.*, "Compressed sensing for energy-efficient wireless telemonitoring of noninvasive fetal ECG via block sparse Bayesian learning," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 2, pp. 300–309, Feb. 2013.
- [8] H. Mamaghanian, N. Khaled, D. Atienza, and P. Vanderghenst, "Compressed sensing for real-time energy-efficient ECG compression on wireless body sensor nodes," *IEEE Trans. on Biomed. Eng.*, vol. 58, no. 9, pp. 2456–2466, 2011.
- [9] B. Liu, Z. Zhang, G. Xu, H. Fan, and Q. Fu, "Energy efficient telemonitoring of physiological signals via compressed sensing: A fast algorithm and power consumption evaluation," *Biomed. Signal Process.*, vol. 11, pp. 80–88, 2014.
- [10] F. Chen, A. P. Chandrakasan, and V. Stojanovic, "Design and analysis of a hardware-efficient compressed sensing architecture for data compression in wireless sensors," *IEEE J. Solid-State Circuits*, vol. 47, pp. 744–756, Mar. 2012.
- [11] D. Gangopadhyay, E. G. Allstot, A. M. R. Dixon, K. Natarajan, S. Gupta, and D. J. Allstot, "Compressed sensing analog front-end for bio-sensor applications," *IEEE J. Solid-State Circuits*, vol. 49, no. 2, pp. 426–438, Feb. 2014.
- [12] F. Pareschi, P. Albertini, G. Frattini, M. Mangia, R. Rovatti, and G. Setti, "Hardware-algorithms co-design and implementation of an analog-to-information converter for biosignals based on compressed sensing," *IEEE Trans. Biomed. Circuits Syst.*, vol. 10, no. 1, pp. 149–162, Feb. 2016.
- [13] K. K. Shyu, M. H. Lee, Y. T. Wu, and P. L. Lee, "Implementation of pipelined fastICA on FPGA for real-time blind source separation," *IEEE Trans. Neural Netw.*, vol. 19, no. 6, pp. 958–970, Jun. 2008.

2. The min-total variation (TV) solution is used for the large-scale 2D-structure instead of $\min -l_1$, which ensures stable recovery for sparse MEG images with high accuracy.

- [14] L. Van, D. Wu and C. Chen, "Energy-Efficient FastICA Implementation for Biomedical Signal Separation," *IEEE Trans. Neural Netw.*, vol. 22, no. 11, pp. 1809–1822, 2011.
- [15] M. Stanačević, S. Li, G. Cauwenberghs, "Micropower mixed-signal VLSI independent component analysis for gradient flow acoustic source separation," *IEEE Trans. Circuits Syst. I Reg. Papers*, vol. 63, no. 7, pp. 972–981, Jul. 2016.
- [16] S. Fauvel, A. Agarwal and R. Ward, "Compressed sensing and energy-aware independent component analysis for compression of EEG signals," 2013 *IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver, BC, 2013, pp. 973–977.
- [17] B. Zhou, X. Wu, Z. Lv, L. Zhang and C. Zhang, "Independent component analysis combined with compressed sensing for EEG compression in BCI," 2015 *10th International Conference on Information, Communications and Signal Processing (ICICSP)*, Singapore, 2015, pp. 1–4.
- [18] M. Stanačević, S. Li, and G. Cauwenberghs, "Micropower mixed-signal VLSI independent component analysis for gradient ow acoustic source separation," *IEEE Trans. Circuits Syst.*, vol. 63, no. 7, pp. 972–981, Jul. 2016.
- [19] D. Kanemoto, S. Katsumata, M. Aihara and M. Ohki, "Framework of Applying Independent Component Analysis After Compressed Sensing for Electroencephalogram Signals," 2018 *IEEE Biomedical Circuits and Systems Conference (BioCAS)*, Cleveland, OH, 2018, pp. 1–4.
- [20] J. Cao and N. Murata, "A stable and robust ICA algorithm based on t-distribution and generalized Gaussian distribution models," in *Proc. IEEE Signal Process. Soc. Workshop on Neural Netw. Signal Process.*, Madison, WI, Aug. 1999, pp. 283–292.
- [21] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Compte Rendus de l'Academie des Sciences*, vol. Serie I, no. 346, pp. 589–592, 2008.
- [22] F. Marvasti *et al.*, "A unified approach to sparse signal processing," *EURASIP J. Adv. Signal Process.*, vol. 2012, no. 44, pp. 1–45, 2012.
- [23] F. Renna, R. Calderbank, L. Carin, and M. R. D. Rodrigues, "Reconstruction of signals drawn from a Gaussian mixture via noisy compressive measurements," *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2265–2277, May. 2014.
- [24] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [25] O. Rivasplata, "Subgaussian random variables: An expository note," *Unpublished note*, 2012.
- [26] S. Suzuki, T. Matsui, B. Escalante, "Remote sensing for medical and health care applications," in *Remote Sensing—Applications*, Rijeka, Croatia: InTech, 2012.
- [27] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.



Yahia Alghorani (S'13-M'16) received his M.Sc. degree in computer engineering from University Kebangsaan Malaysia, Selangor, Malaysia, in 2007, and his Ph.D. degree in computer and software engineering from Ecole Polytechnique de Montreal, Montreal, QC, Canada, in 2015. From 2009 to 2010, he was an RF engineer with Batelco. From 2010 to 2012, he was a wireless networks engineer with Huawei. From 2017 to 2019, he was a research fellow at University College Dublin, Ireland. Dr. Alghorani is currently a research associate at Lakehead University, ON, Canada. His research interests include intelligent transportation systems, connected vehicles, artificial intelligence, machine learning, deep learning, medical wearables and e-health, statistical signal processing, and hardware/software co-design.



Salama S. Ikki received his Ph.D. degree in Electrical Engineering from Memorial University, St. Johns, NL, Canada, in 2009. From February 2009 to February 2010, he was a Postdoctoral Researcher at the University of Waterloo, ON, Canada. From February 2010 to December 2012, he was a Research Assistant with INRS, University of Quebec, Montreal, QC, Canada. He is currently an Associate Professor of wireless communications with Lakehead University, Thunder Bay, ON, Canada. He is the author of 100 journal and conference papers and has more than 4000 citations and an H-index of 30. His research interests include cooperative networks, multiple-input-multiple-output, spatial modulation, and wireless sensor networks. Dr. Ikki served on the Editorial Board of *IEEE Communications Letters* and *Institution of Engineering and Technology Communications*. He also served as a Technical Program Committee member for various conferences, including *IEEE International Conference on Communications*, *IEEE Global Communications Conference*, *IEEE Wireless Communications and Networking Conference*, *IEEE Spring/Fall Vehicular Technology Conference*, and *IEEE International Symposium on Personal, Indoor, and Mobile Communications*. Dr. Ikki received the Best Paper Award for his paper published in the *EURASIP Journal on Advanced Signal Processing*. He also received *IEEE Communications Letters*, *IEEE Wireless Communications Letters*, *IEEE Transactions on Vehicular Technology* and *IEEE Transactions on Communications* exemplary reviewer certificates for 2012, 2013 and 2014, respectively.