J Med Syst (2016) 40:285
DOI 10.1007/s10916-016-0642-y

MOBILE & WIRELESS HEALTH

Energy Efficient Monitoring of Metered Dose Inhaler Usage

Aris S. Lalos[1] · John Lakoumentas[1] · Anastasios Dimas[1] · Konstantinos Moustakas[1]

Received: 26 February 2016 / Accepted: 17 October 2016

Abstract Life-long chronic inflammatory diseases of the airways, such as asthma and Chronic Obstructive Pulmonary Disease, are very common worldwide, affecting people of all ages, race and gender. One of the most important aspects for the effective management of asthma is medication adherence which is defined as the extent to which patients follow their prescribed action plan and use their inhaler correctly. Wireless telemonitoring of the medication adherence can facilitate early diagnosis and management of these diseases through the use of an accurate and energy efficient mHealth system. Therefore, low complexity audio compression schemes need to be integrated with high accuracy classification approaches for the assessment of adherence of patients that use of pressurized Metered Dose Inhalers (pMDIs). To this end, we propose a novel solution that enables the energy efficient monitoring of metered dose inhaler usage, by exploiting the specific characteristics of the reconstructed audio features at the receiver. Simulation studies, carried out with a large dataset of indoor & outdoor measurements have led to high levels of accuracy (98 %) utilizing only 2 % of the recorded audio samples at the receiver, demonstrating the potential of this method for the development of novel energy efficient inhalers and medical devices in the area of respiratory medicine.

Keywords Compressed sensing · AdaBoost · Time-frequency analysis · Support vector machines

Introduction

Asthma is a chronic disease of the airways that affects more than 235 million people worldwide [1]. In the region of Europe, 30 million adults suffer from asthma [2], while the number of children suffering from the disease is continuously rising [3]. This diversity of asthma prevalence is a global phenomenon [4] and reveals the inability of even developed countries to effectively support asthma patients [5]. Moreover, the socioeconomic consequences of asthma disease that reduce the quality of life of patients and the efficiency of the healthcare system [6], underline the need for novel healthcare approaches and innovative devices in support of patients and healthcare professionals.

One of the most important aspects for the effective management of asthma is medication adherence; the extent to which patients follow their prescribed action plan and use their inhaler correctly [7]. Reduced adherence has been associated with asthma attack incidents and patient hospitalizations [8]. Wireless telemonitoring of the medication adherence can facilitate early diagnosis and management of these diseases.

Several experts from the fields of information and communication technologies, respiratory medicine, and inhaler devices, focus on designing novel mHealth systems for monitoring medication adherence [9]. Most of these devices give a day-to-day measure of inhaler use but they do not assess inhaler technique. Detection of technique errors is traditionally carried out through a face to face process with a clinician [10]. However there is no way of assessing

This article is part of the Topical Collection on *Mobile & Wireless Health*

✉ Aris S. Lalos
aris.lalos@ece.upatras.gr

1 Electrical & Computer Engineering Department, University of Patras, Patras, Greece

Published online: 29 October 2016

technique performance once the patient returns home. Hence, the limitations of all of these methods suggest that there is a need for a technology to longitudinally and objectively monitor both temporal and technique adherence. One promising sensing approach adopted by electronic inhaler devices that can be used for overcoming the aforementioned limitation is the use of microphones.The recorded audio sounds can be processed and used to indicate not only inhaler actuation but also other critical events like inhalations and exhalations. Apart from the obvious advantages, wireless telemonitoring of inhalation technique adherence through processing of audio recordings requires new schemes and algorithms to be implemented in order to optimize important parameters, such as: (i) the energy consumption (ii) the total hardware cost and (iii) the accurate monitoring of events related to medication adherence. Low energy consumption significantly increases the battery lifetime of the audio sensor, while the hardware cost reduction makes the mHealth system economically viable for individual customers. Both requirements motivate the design and integration of compression/reconstruction schemes with high compression ratio capabilities, reduced computational requirements with state-of-the-art sophisticated classification methods.

The vast majority of audio compression schemes available in literature [11] charge the transmitter with most of the processing, thus not coping effectively with the aforementioned requirements. Compressed Sensing (CS) approaches for signal compression/reconstruction offers an affordable solution for audio compression in wireless sensor networks [12]. To the best of our knowledge, this is the first work that demonstrates the benefits of CS based compression/reconstruction schemes at the efficient telemonitoring of medication adherence. More specifically, the contributions of our work can be summarized as follows: i) We enhance the benefits of the conventional CS schemes proposed in [12], by taking into account specific characteristics (e.g., block sparsity, sample correlation) of the audio features, using a novel recovery algorithm named Decorrelated Group LASSO (DG LASSO), ii) we then integrate the DG LASSO, with state of the art classifiers allowing high levels of accuracy (98 %), from a very small number of linearly encoded samples increasing significantly the system energy efficiency, since then number of required encoded samples corresponds to the 2 % of the recorded ones.

The rest of the paper is outlined as follows: Section 2 includes a detailed summary of prior art. Section 3, presents preliminaries related to CS. The system model is discussed in section 4. Section 5, describes the proposed compression/reconstruction schemes. In Section 6, we present the adopted classification approaches. Section 7 presents the performance of the proposed schemes, highlighting the strengths and weaknesses. Finally, Section 8 concludes this paper.

Related works

One of the most important aspects for the efficient and effective management of asthma is the extent to which patients adhere to their prescribed action plan and use their medication correctly. Reduced adherence has been linked with significant indicators of health degradation [8]. More specifically, 24 % of the exacerbations and 60 % of hospitalizations can be credited to poor adherence [13]. A crucial step in this direction, is the formation of a sensing framework that can provide accurate information about the health of patients and help their doctors understand potential difficulties that prevent their patients from using their inhalers correctly [14]. Within this concept, a number of review studies have been recently published focusing on commercial products and their characteristics from the clinical point of view [9, 15–19].

The first of these studies has reviewed oral and nebulized medication monitors in addition to inhaler monitoring devices [16]. Two other studies provided a detailed review of the currently available devices focusing on the clinical point of view and producing a useful guide on how researchers and clinicians can select the most appropriate product and how to utilize the full spectrum of its capabilities [17, 18]. Finally, a recent work has provided a summary of the most common electronic monitors of inhaler adherence, but focused on measured dose inhalers (MDI) and mentioned some indicative devices for dry powder inhalers (DPI) [9]. The modern adherence monitoring environment has been also analyzed, dealing with the interpretation of results and the design of interventions that promote adherence [19].

The majority of devices presented above are based on electromechanical sensing capabilities, ranging from simple push buttons attached on the top of the inhaler's canister up to force sensing elements attached on the back of the inhaler's plastic casing. However, this approach is capable of identifying inhaler actuation, completely ignoring actions related to inhalation, exhalation prior to or subsequently to inhaler actuation [7]. One promising sensing approach adopted by electronic inhaler devices that can be used for overcoming the aforementioned limitation is the use of microphones. The recorded measurements can be locally processed and used to indicate not only inhaler actuation but also other critical events like inhalations and exhalations. A recent publication by Taylor et al. [20] introduced a fundamental and robust approach for the detection of MDI actuations, through context based audio classification in the laboratory environment. As an alternative approach, the authors in [21] utilized Convolutional Neural Networks aiming to produce more accurate results for real life environments. Although the aforementioned works focus on monitoring inhlaed medication using microphones, they mostly focus on identifying inhaler actuation without taking into account other critical events like inhalations and exhalations, while they completely ignore the energy required for transmitting these sounds and the robustness of the classification approaches to communication errors.[1]

Preliminaries on compressed sensing

CS provides a way of reconstructing a sparse signal $\mathbf{x} \in \mathbb{R}^N$ from a small number of linearly combined measurements [24]. The Random Linear Combinations (RLC) $\mathbf{y} \in \mathbb{R}^M$, $M < N$, are generated using a random matrix $\mathbf{A} \in \mathbb{R}^{M \times N}$ with i.i.d. elements as: $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}$, where $\mathbf{w}$ is a vector with noise samples..

Reconstruction by exploiting sample sparsity

In the noise-free case ($\mathbf{w} = \mathbf{0}_N$), vector $\mathbf{x}$ may be ideally recovered from $\mathbf{y}$ by solving the problem: $min_x \{\|\mathbf{x}\|_0 : \mathbf{y} = \mathbf{A}\mathbf{x}\}$, where $\|\cdot\|_0$ denotes the ℓ_0-norm that equals the number of nonzero entries in $\mathbf{x}$.

In order for the signal reconstruction to be robust in the presence of noise, the problem constraint is relaxed to: $min_x \left\{\|\mathbf{x}\|_0 : \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \leq \epsilon\right\}$, where ϵ is an error tolerance and $\| \cdot \|_2$ is the ℓ_2-norm of the input vector, respectively. The above optimization problem cannot be used for practical applications, since it is computationally intractable. CS suggests replacing the ℓ_0 quasi-norm by the convex ℓ_1-norm and solving the following problem: $min_x \left\{\|\mathbf{x}\|_1 : \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \leq \epsilon\right\}$, where the ℓ_1-norm is defined as $\|\mathbf{x}\|_1 = \sum_{i=1}^N |x_i|$. By employing Lagrangian relaxation, we are able to efficiently approximate the solution of the aforementioned problem by solving the ℓ_1 regularized least squares problem:

$$\hat{\mathbf{x}} := \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1, \tag{1}$$

where the parameter λ controls the balance between the two optimization objectives: (i) the noise level $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2$ and (ii) the sparsity of vector $\mathbf{x}$. Algorithmically, the convex optimization problem in Eq. 1, known as the the LASSO problem, can be tackled by any generic second-order cone program solver.

Reconstruction by exploiting block sparsity

A block sparse signal consists of clusters of zero and non - zero coefficients. To be more specific, vector $\mathbf{x}$ can be viewed as a concatenation of R blocks of length d:

$$\mathbf{x} = [\, \underbrace{x_1, \ldots, x_d}_{\mathbf{x}^T[1]}, \underbrace{x_{d+1}, \ldots, x_{2d}}_{\mathbf{x}^T[2]}, \ldots, \underbrace{x_{N-d+1}, \ldots, x_N}_{\mathbf{x}^T[R]} \,]^T, \tag{2}$$

where $\mathbf{x}[i]$ denotes the i^{th} block and $N = Rd$.

Similarly to Eq. 2, we can represent matrix $\mathbf{A}$ as a concatenation of sub-matrices $\mathbf{A}[i]$ of size $M \times d$: $\mathbf{A} = [\, \mathbf{A}[1]\,, \; \mathbf{A}[2]\,, \; \ldots \; \mathbf{A}[R] \,]$.

The block-sparse structure enables the signal recovery from a reduced number of samples, compared to sample sparse structures. To exploit block sparsity, we have to reconstruct vector $\mathbf{x}$ by solving:

$$\hat{\mathbf{x}} := \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{y} - \sum_{i=1}^{R} \mathbf{A}[i]\, \mathbf{x}[i]\, \|_2^2 + \sum_{i=1}^{R} \lambda_i \|\mathbf{x}[i]\, \|_2, \tag{3}$$

which is also known as the Group LASSO (GLASSO) problem [23].

Reconstruction by exploiting sparsity in a transform domain

In many applications, although the signal $\mathbf{x}$ is not sparse in the time domain, it can be sparse in other domains. Therefore, $\mathbf{x}$ can be expressed as $\mathbf{x} = \mathbf{W}\mathbf{s}$, where $\mathbf{W} \in \mathbb{R}^{N \times N}$ is an orthonormal basis matrix of a transformed domain and $\mathbf{s}$ is the representation coefficient vector, which is sparse. In such cases, in order to exploit either the sample or the block sparsity of $\mathbf{s}$ in the transformed domain, instead of Eq. 1 or Eq. 3, we can solve:

$$\hat{\mathbf{s}} := \underset{\mathbf{s}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{A}\mathbf{W}\mathbf{s}\|_2^2 + \lambda \|\mathbf{s}\|_1 \tag{4}$$

$$\hat{\mathbf{s}} := \underset{\mathbf{s}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{A}\mathbf{W}\mathbf{s}\|_2^2 + \sum_{i=1}^{R} \lambda_i \|\mathbf{s}[i]\, \|_2 \tag{5}$$

and then reconstruct $\hat{\mathbf{x}} = \mathbf{W}\hat{\mathbf{s}}$.

System model and audio features

Figure 1 illustrates the telemonitoring system under study. In particular, we consider two different settings, namely

[1] Most of the energy consumption of a biosensor comes from the radio frequency power amplifier [22]

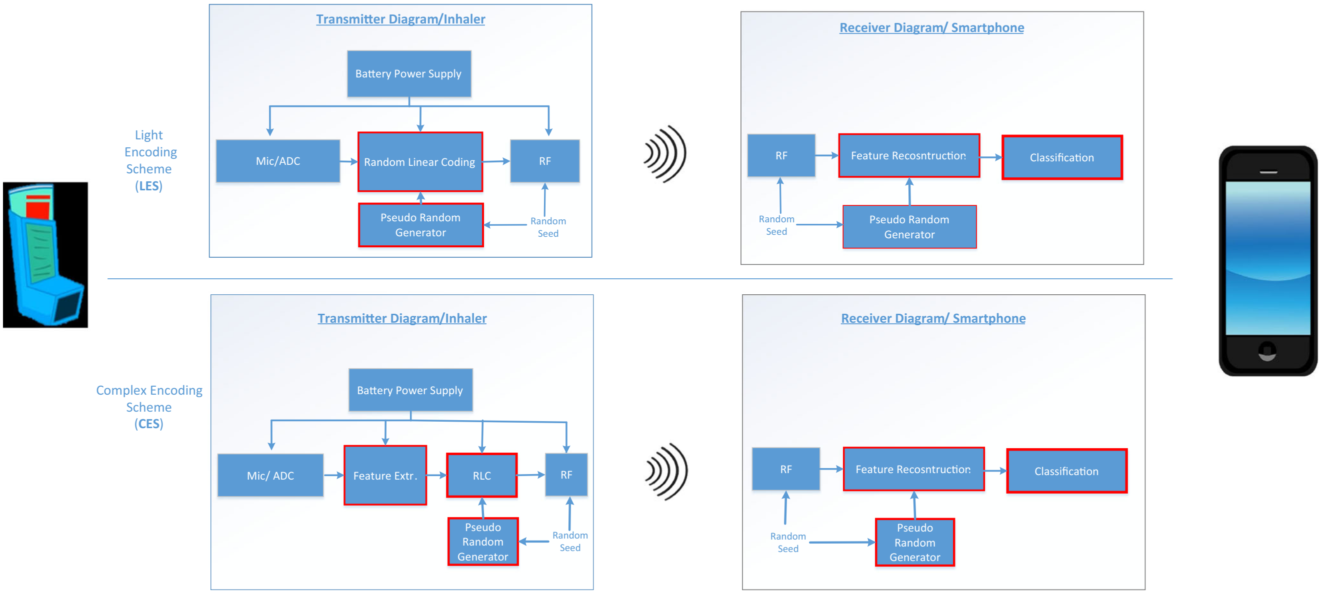

Fig. 1 Proposed architecture

the low encoding setup (LES) and the complex encoding setup (CES). In the LES, the acoustic sensor (source node), records a real time audio sound and transmits to the decoder (e.g., smartphone) RLC's of the recorded samples. At the decoder, the features are reconstructed from the encoded audio samples and then used for the audio classification problem. This significantly reduces the processing requirements at the transmitter, since it allows the extraction of features at the receiver side, and performs audio compression by simply generating random linear combinations of the digitized samples. In the CES approach, the sensor node evaluates the audio features and transmits random linear combinations of the features, allowing the decoder to directly execute the audio classifiers using the reconstructed features.

In both settings, the audio signal is recorded, digitized, and divided into segments of $N = f_s/2$ samples that correspond to 0.5 sec of audio, e.g., $\mathbf{x} = [x_1, \ldots, x_N]^T$, where $x_i \in \mathbb{R}$. In the LES, we assume that the transmitted samples contains noise and, as a result, may be written as $\mathbf{u} = \mathbf{x} + \mathbf{w}_s$, where $\mathbf{u} = [u_1, \ldots, u_N]^T$ are samples of the noisy signal and $\mathbf{w}_s = [w_1, \ldots, w_N]^T$ is the random noise. For each segment, the source generates M random linear combinations (see Fig. 1) by using a random matrix $\mathbf{A}$ of dimension $M \times N$ and performs quantization as follows:

$$\mathbf{y}_q = Q(\mathbf{Au}) = \mathbf{Ax} + \mathbf{w}_q, \tag{6}$$

where $Q : \Re \rightarrow \mathbf{Y_i}$ is a scalar quantization function, and $\mathbf{w}_q$ represents the combination of the sensing and quantization error. The encoded samples are then quantized, and $\mathbf{y}_q$ is transmitted to the decoder where the reconstruction of the features and the classification of the audio segments takes place.

In the CES, we extract the features at the encoder and then compress them using a random matrix $\mathbf{A}$. Motivated by the fact that existing acoustic approaches [20] suggest the use of time-frequency analysis to automatically detect pMDI actuations, we suggest deriving the audio features from the spectrogram of the samples that is generated by applying a short time Fourier Transform. To be more specific, the audio features $\mathbf{f}_i$ are represented by a one dimensional vector containing a summation of all high frequency content at a given time index $i = 1, \ldots, R$.

At this point, it should be mentioned that there are several ways of constructing matrix $\mathbf{A}$ for encoding the audio samples/features that directly affects i) the storage and processing requirements at the transmitter side, ii) the communication load and iii) the storage requirements at the receiver. Below we present the most widely used sensing matrices in the literature of CS [24]:

Gaussian random encoding

By selecting the coefficient $\mathbf{A}_{i,j} \sim \mathcal{N}(0, 1/\sqrt{N})$ as Gaussian i.i.d. elements the recovery conditions are satisfied. Even though the authors in [2]showed that the quantization of the Gaussian elements do not affect significantly the signal quality loss, the aforementioned choice requires i) the implementation of a Gaussian distributed random generator, ii) the multiplication of the audio samples/features with real values iii) the storage of a large matrix with real values.

Binary random encoding

An alternative approach is to select the i.i.d. entries of matrix $\mathbf{A}$ from the Bernoulli distribution, i.e. $\mathbf{A}_{i,j} = \pm 1/\sqrt{N}$ with probability 0.5. It has been shown that the aforementioned matrix reduces significantly the processing at the transmitter, while satisfying the recovery conditions. Note that the transmitted data are generated by simply adding/subtracting the original audio samples/features.

Sparse binary random encoding:

To reduce even more the required compression complexity, the entries Ai,jmay be selected according to: $\mathbf{A}_{i,j}$ are either $\{1, 0\}$ with probabilities $\{1/s, 1 - 1/s\}$ where s is a parameter that determines the degree of sparsity of the sensing matrix A. The authors in [26] showed that sparse random matrices still yields good recovery properties. The optimal choice of s depends on the structure of the audio signal/features and the decoding algorithm that is used at the receiver side.

In general, it is assumed that the encoding matrix $\mathbf{A}$ is known at the destination in order to perform reconstruction of the audio samples/features. To overcome this limitation, the random linear matrix A is constructed both at the encoder and decoder using a pseudo-random number generator (PRNG) that generates a sequence of numbers that approximate the properties of random numbers, as shown in Fig 1. The generated sequence, is completely determined by a random seed, represented by a single real number that is the only information that has to be transmitted at the receiver side. To reduce even more the communication requirements the same seed is used for encoding a large number of audio segments/features. Therefore even if we take into account the overhead for transmitting this seed, the calculated compression ratio at the decoder is not affected at all, since the same seed is used for encoding a large number of audio samples/features.

Efficient recovery of the audio features

The DCT coefficients of audio signals are usually modeled as multivariate Gaussian distributions with a specific correlation matrix [25]. Thus, we assume that both the representation of the audio samples in the DCT domain ($\mathbf{s} = \mathbf{Dx}$, where $\mathbf{D}$ is the DCT matrix) and the time-frequency features ($\mathbf{f}$) in the CES setup are, to a certain degree, temporally correlated. In the following section, we propose a novel recovery algorithm (DG LASSO) that efficiently exploits these temporal correlations, enabling the features reconstruction either the reconstruction of the DCT coefficients (LES case) or the reconstruction of spectrogram features (CES case), from less received measurements.

Temporal correlation-aware block sparse recovery

Let us assume that: (i) the matrix $\mathbf{R}_i \in \mathbb{R}^{d\times d}$ captures the correlation structure of the i-th block of $\mathbf{x}$, $\mathbf{x}[i]$, and (ii) the correlation between elements of different signal blocks is zero, e.g.,

$$E\left[\mathbf{x}[i]\,\mathbf{x}^T[j]\right] = \begin{cases} \mathbf{R}_i & if\ i = j \\ 0 & if\ i \neq j \end{cases} \tag{7}$$

Each block correlation matrix $\mathbf{R}_i$ can be approximated by a Toeplitz symmetric matrix:

$$\mathbf{R}_i = \begin{bmatrix} r_0 & r_1 & \dots & r_{d-1} \\ r_1 & r_0 & \dots & r_{d-2} \\ \vdots & \ddots & \ddots & \vdots \\ r_{d-1} & \dots & r_1 & r_0 \end{bmatrix}. \tag{8}$$

The values of r_k, $k = 0, \dots, d-1$, can be estimated by assuming that intra-block correlation follows an exponential correlation model by making the approximation $r_k = r^k$, $k = 0, \dots, d-1$ and selecting specific values for r that capture the degree of correlation between adjacent samples.

Based on the fact that the GLASSO become more efficient when the difference between the norms of non-zero blocks is small, we propose a practical way to achieve this (especially in highly correlated cases) by performing block-sparse reconstruction of the decorrelated segment $\mathbf{d}_\mathbf{x}$, written as:

$$\mathbf{d}_\mathbf{x} = \mathbf{R}^{-1/2}\mathbf{x}, \tag{9}$$

$$\mathbf{R}^{-1/2} = \begin{bmatrix} \mathbf{R}_1^{-1/2} & \mathbf{0}_d & \dots & \mathbf{0}_d \\ \dots & \ddots & \ddots & \dots \\ \mathbf{0}_d & \dots & \dots & \mathbf{R}_R^{-1/2} \end{bmatrix}. \tag{10}$$

Consequently, by solving the problem defined in Eq. 5 after selecting $\mathbf{W} = \mathbf{R}^{1/2}$:

$$\hat{\mathbf{d}}_\mathbf{x} := \underset{\mathbf{d}_\mathbf{x}}{\text{argmin}} \|\mathbf{y} - \mathbf{A}\mathbf{R}^{1/2}\mathbf{d}_\mathbf{x}\|_2^2 + \sum_{i=1}^{R} \lambda_i \|\mathbf{d}_\mathbf{x}[i]\|_2, \tag{11}$$

allows us to recover the coefficient of the original vector at the receiver, by eliminating any temporal correlations within signal blocks (intra-block correlation), therefore controlling the block sparsity of the decorrelated segment $\mathbf{d}_x$,.

In the LES, the recovery of the DCT coefficients after adopting the aforementioned approach, maybe written as:

$$\hat{\mathbf{d}}_{\mathbf{s}}:=\underset{\mathbf{d}_{\mathbf{s}}}{\operatorname{argmin}}\|\mathbf{y}-\mathbf{A}\mathbf{D}^{-1}\mathbf{R}^{1/2}\mathbf{d}_{\mathbf{s}}\|_2^2+\sum_{i=1}^{R}\lambda_i\|\mathbf{d}_{\mathbf{s}}[i]\|_2, \quad (12)$$

and $\hat{\mathbf{s}} = \mathbf{R}^{1/2}\hat{\mathbf{d}}_{\mathbf{s}}$. Then, the features used for the classification are evaluated as

$$\mathbf{f}^{LES} = \left[\left\|\hat{\mathbf{s}}[1]\right\|_{l_2}, \ldots, \left\|\hat{\mathbf{s}}[R]\right\|_{l_2}\right], \quad (13)$$

where $R = N/d$.

Similarly, in the CES case we reconstruct the features as:

$$\hat{\mathbf{d}}_{\mathbf{f}}:=\underset{\mathbf{d}_{\mathbf{f}}}{\operatorname{argmin}}\|\mathbf{y}-\mathbf{A}\mathbf{R}^{1/2}\mathbf{d}_{\mathbf{f}}\|_2^2+\sum_{i=1}^{R}\lambda_i\|\mathbf{d}_{\mathbf{f}}[i]\|_2, \quad (14)$$

and then we evaluate the features as: $\mathbf{f}^{CES} = \mathbf{R}^{1/2}\hat{\mathbf{d}}_{\mathbf{f}}$.

Feature classification for medication adherence

In order to differentiate inhaler actuations, exhalations and inhalations from noise, we performed pattern classification. Given a set of subjects described by features, including a special categorical attribute called class, classification aims to create a mapping between each class value and the combination of values of the rest features. Then, a classifier is able to predict the unknown classes of subjects, given their feature values (e.g., reconstructed features). We used the following types of state-of-the-art classifiers [31]:

Support vector machines (SVM)

The SVM is applied both to linearly and non-linearly separable data, with the use of kernel transformations. Specifically, it transforms the data to a higher dimension, from where it can identify a hyperplane that separates the data.

AdaBoost (AB)

AdaBoost is the most common boosting algorithm.[2] It uses decision trees as weak learners and treats them sequentially. Subsequent decision trees are tweaked in favor of those subjects misclassified by previous decision trees.

Random forests (R-F):

Each tree is constructed by a bootstrap sample from the data, using a small set of attributes selected from a random set. Once the forest is formed by the training data, test subjects are percolated down each decision tree and trees make their respective class predictions. The accuracy of random forest overall depends on the strength of each individual tree and the correlation between any two trees.

Simulation results

The focus of this study is to identify the strengths and weaknesses of the two different setups (LES and CES), and the energy efficiency (EE) of the mHealth system. We also evaluate the effects of the different reconstruction and classification approaches, with respect to the achieved classification accuracy and EE. The proposed telemonitoring schemes are studied by using audio segments recorded during the used of an inhaler according to the guidelines provided in [7]. In specific, the correct usage of a pMDI is defined according to clinical expert's suggestions by the next steps. a) Remove the cap. b) Breathe out (exhale), away from your inhaler. c) Bring the inhaler to your mouth. Place it in your mouth between your teeth and close your mouth around it. d) Start to breathe in (inhale) slowly. Press the top of your inhaler once and keep breathing in slowly until you have taken a full breath. e) Take away the inhaler from your mouth and hold your breath for about 10 seconds, then breathe out (exhale).

Those recordings are divided into 0.5 sec of audio segment that are classified into noise, inhaler actuations, inhlations and exhalations. The measurements are collected from indoor and outdoor environments using a recording device composed of a wireless Bluetooth microphone attached to the pMDI and a smartphone. Due to the lack of a single standard inhaler usage dataset, we constructed a database consisting of 500 inhaler activation samples, and

[2]Boosting is a classification ensemble meta algorithm that was built to answer the following question: can a set of weak learners create a single strong learner? A weak learner is defined to be a classifier which can classify subjects slightly better than random guessing. A strong learner is a classifier that is correlated with the true classification

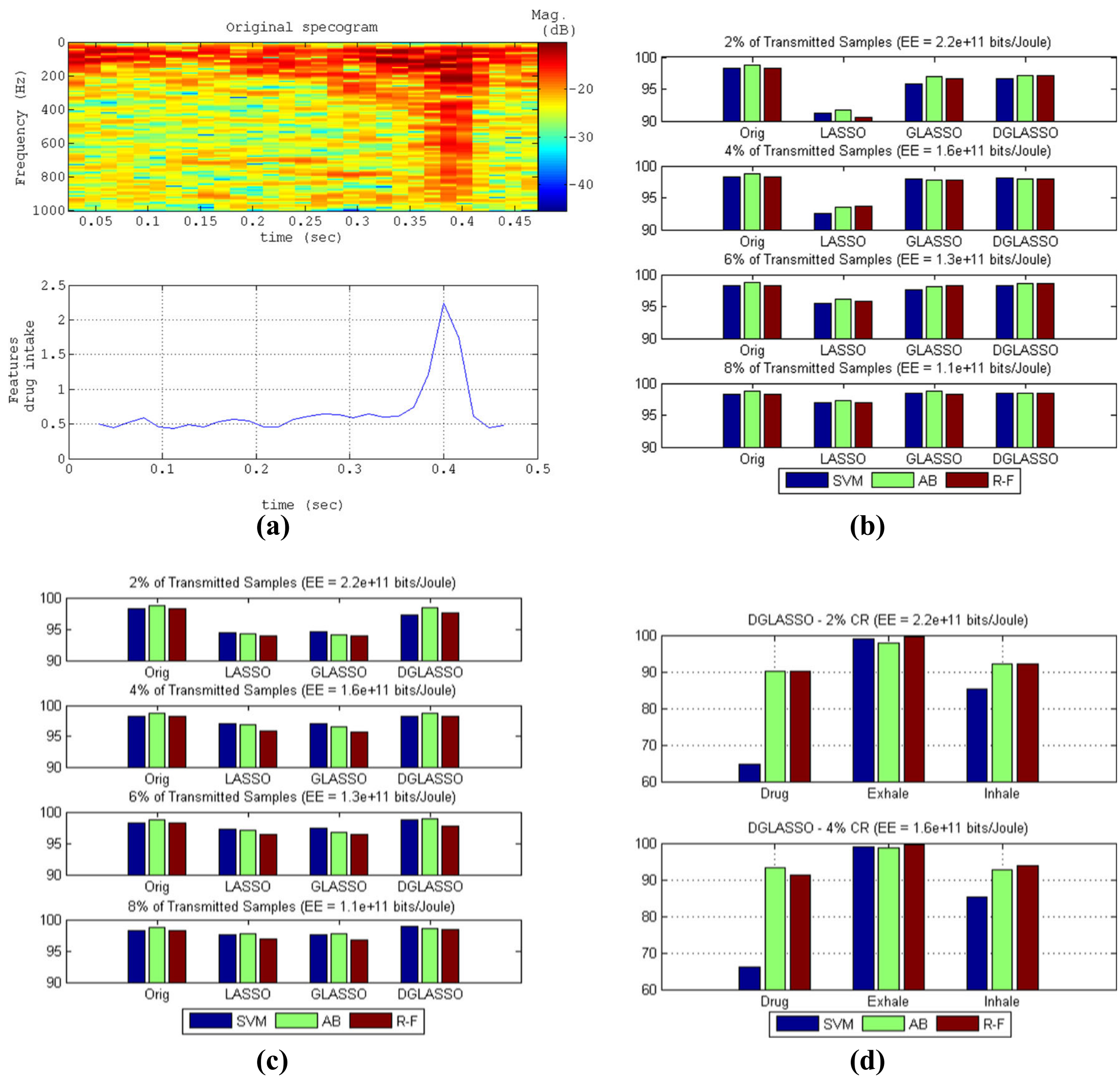

Fig. 2 **a** Time frequency analysis & features of original signal (**b**) Evaluation in term of energy efficiency - LES case (Dataset 1) (**c**) Evaluation in term of energy efficiency - CES case (Dataset 1) (**d**) Evaluation using (Dataset 2)

500 noise samples (Dataset-1) and a database that consist of 200 inhaler actuations, 200 inhlation, 200 exhalation and 200 noise samples (Dataset-2). Classification aims at distinguishing these types of audio sounds using the set of features that were extracted as described above. These sounds were recorded using 4 kHZ sampling rate and 4-bit

Table 1 LES: CR vs Classification accuracy using Dataset 1

CR	Decoder											
	Original			LASSO			GLASSO			DGLASSO		
	SVM	AB	R-F	SVM	AB	R-F	SVM	AB	R-F	SVM	AB	R-F
0.05	98.3	98.8	98.3	91.3	91.8	90.6	95.9	97	96.7	96.7	97.2	97.1
0.1	98.3	98.8	98.3	92.5	93.5	93.7	98	97.8	97.8	98.1	98	98
0.15	98.3	98.8	98.3	95.5	96.1	95.8	97.7	98.1	98.3	98.2	98.6	98.6
0.2	98.3	98.8	98.3	97	97.3	96.9	98.5	98.5	98.4	98.5	98.8	98.3

Table 2 CES: CR vs Classification accuracy using Dataset 1

CR	Decoder											
	Original			LASSO			GLASSO			DGLASSO		
	SVM	AB	R-F	SVM	AB	R-F	SVM	AB	R-F	SVM	AB	R-F
0.4	98.3	98.8	98.3	94.4	94.3	93.9	94.6	94.1	94	97.2	98.5	97.6
0.5	98.3	98.8	98.3	97	96.9	95.8	97	96.6	95.7	98.3	98.8	98.2
0.7	98.3	98.8	98.3	97.2	97.1	96.4	97.4	96.7	96.4	98.7	99	97.7
0.8	98.3	98.8	98.3	97.6	97.7	97	97.6	97.7	96.7	98.9	98.6	98.4

depth ($l_q = 4$). Two healthy individuals participated in the study and were used to produce the databases of the audio samples.

Simulation setup

We assume that each breath signal, is divided into segments of N = 2000 samples. In both the LES and CES, the encoding is performed using a binary matrix and we consider three types of decoders: (i) the conventional LASSO (ii) the GLASSO and iii) the DG LASSO presented in section 5. These algorithms were adopted to extract the features that were then used for classification. Then, we employed the different classifiers presented above (i) SVM, (ii) AB and (iii) R-F. The first dataset used for classification (Dataset-1) was composed of 1000 subjects (500 inhaler activation samples and 500 noise samples) and 29 features. The second dataset (Dataset-2) was composed of 800 subjects (200 inhaler activation samples, 200 exhalations, 200 inhalations, and 200 noise samples) and 29 features. We did not use a training or a test set; instead, in order to evaluate our algorithms, 10-fold cross validation was performed on the datasets that were generated for different compression rates. Training a support vector machine was performed by utilizing the radial (RBF) kernel, cost equal to 1 and gamma equal to 0.25. Training random forests was done by using 500 small decision trees without pruning and setting the number of features used by each tree as the floor of the squared root of the total number of features. Finally, the multiclass AdaBoost algorithm (SAMME) was executed by utilizing 100 small decision trees in sequence.

Table 3 DGLASSO in LES: CR vs Classification accuracy using Dataset 2

CR	Case studies								
	Drug vs. all			Exhale vs. all			Inhale vs. all		
	SVM	AB	R-F	SVM	AB	R-F	SVM	AB	R-F
0.1	64.8	89.5	90.1	99.6	97.9	100	85.4	92.1	92.1
0.2	66.0	93.2	91.4	99.6	98.7	100	85.4	92.7	93.9

The EE of the considered mHealth system, assuming that the transmit and receive power is equal to 3.8 mW and 4.6 mW respectively, that the encoded measurements have $l_q = 4$ bits and the duration of a packet transmission is $t_p \approx$ 2.94 ms [27][3] can be evaluated by:

$$EE = \frac{Segment\ Bits}{Total\ Energy} = \frac{8N}{M(P_T + P_R)t_p/10}\ bits/Joule. \quad (15)$$

Performance evaluation

In Fig.2a, the obtained features for the LES and CES are plotted. A block length $d = 16$ was selected and the scaling rules for the parameter λ in the LASSO and GLASSO approaches follow the results of [29]. In Fig. 2b and Table 1, the DG LASSO algorithm performs almost optimally; that is, compared to the results of the original data, despite the high CR. LASSO and GLASSO algorithms perform a bit worse, while all the decoding algorithms achieve accuracies greater than 90 %. Among the executed classifiers, AB performs better in most of the cases. As expected, when the compression rate increases, the classification accuracies increase as well, while it is clear that we achieve accuracies over 96 % even for 2 % compression ratios (CR).

Similar conclusions are drawn from Fig. 2c (CES) and Table 2. By comparing the two schemes, one can deduce that the LASSO algorithm performs better in the second case, the GLASSO algorithm performs better in the first case, while the DGLASSO algorithm has almost identical results. The robustness to quantization effects can be further increased by adopting the policies described in [30]. Finally, it should be noted that the integration of DG LASSO with AB allows us to adopt the LES scheme that achieves similar results with the CES scheme, without requiring the reconstruction of features at the transmitter.

Finally, in Fig. 2d and Table 3 we present experiments with the second database using the LES setup and the

[3] We have assumed packets with 14 bytes header and 80 bytes payload (10 audio samples/packet), and a data rate equal to 256 kbps [28].

DGLASSO decoder, in order to evaluate whether the proposed system is capable of identifying not only inhaler actuation but also, inhalations & exhalations. It should be noted that the increased accuracies (> 92 %), even when using a number of encoded samples equal to the 2 % of the recorded ones, demonstrate the potential of this method for the development of novel energy efficient inhalers that allow the wireless monitoring of medication adherence.

Conclusion

Wireless telemonitoring of inhaler medication adherence from acoustic sounds facilitate the early diagnosis and management of chronic inflammatory disease of the airways but introduce challenges related to the real-time compression, transmission and classification of the audio signals. To this end, we propose a novel Compressed Sensing framework that, when integrated with state-of-the-art classification solutions, offers significant gains in terms of both energy efficiency and accuracy of the considered mHealth system, by exploiting the benefits of the group LASSO approaches in the feature domain.

As future steps, the current study will be extended by evaluating the accuracy of the algorithm for different signal to noise ratios in order to better characterize the performance of the algorithm in a variety of environments and types of acoustic noises. Moreover, the compression efficiency of the proposed schemes will be investigated when using data dependent dictionaries based on 2D PCA and subspace tracking approaches instead of STFT. Finally, it should be underlined that the proposed solution is aiming to be integrated with a miniaturized smart blue tooth microphone that could be attached to traditional inhalers and extend their function with adherence monitoring capabilities.

Acknowledgments This work has been supported by the H2020-PHC-2014-2015 Project MyAirCoach (Grant Agreement No. 643607).

References

1. World Health Organization. Asthma. Fact sheet No 307. Available online:, http://www.who.int/mediacentre/factsheets/fs307/en/, [Online; accessed on 30-Sept-2015].
2. European Lung White Book. Adult asthma. Available online:, http://www.erswhitebook.org/chapters/adult-asthma/, [Online; accessed on 30-Sept-2015].
3. European Lung White Book. Childhood asthma. Available online:, http://www.erswhitebook.org/chapters/childhood-asthma/, [Online; accessed on 30-Sept-2015].
4. Innes Asher, M., Montefort, S., Bjrkstn, B., Lai, C. K. W., Strachan, D. P., Weiland, S. K., and Williams, H., Worldwide time trends in the prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and eczema in childhood: {ISAAC} phases one and three repeat multicountry cross-sectional surveys. *The Lancet* 368(9537):733–743, 2006.
5. Akinbami, O. J., and et al., Trends in asthma prevalence, health care use, and mortality in the united states, 2001-2010, 2012.
6. Masoli, M., Fabian, D., Holt, S., and Beasley, R., The global burden of asthma: executive summary of the gina dissemination committee report. *Allergy* 59(5):469–478, 2004.
7. National Asthma Education and Prevention Program: How to use a metered-dose inhaler. http://www.nhlbi.nih.gov/health/public/lung/asthma/asthma_tipsheets.pdf. [Online; accessed on 26-Feb-2016].
8. Murphy, A. C., Proeschal, A., Brightling, C. E., Wardlaw, A. J., Pavord, I., Bradding, P., and Green, R. H., The relationship between clinical outcomes and medication adherence in difficult-to-control asthma, Thorax, pp. thoraxjnl–2011, 2012.
9. Howard, S., Lang, A., Patel, M., Sharples, S., and Shaw, D., Electronic monitoring of adherence to inhaled medication in asthma. *Curr. Respir. Med. Rev.* 10(1):50–63, 2014.
10. Press, V. G., Arora, V. M., Shah, L. M., Lewis, S. L., Ivy, K., Chareneau, J., and et al., Misuse of respiratory inhalers in hospitalized patients with asthma and COPD. *J. Gen. Intern. Med.* 26(6):635–42, 2011.
11. Painter, T., and Spanias, A., Perceptual coding of digital audio. *Proc. IEEE* 88(4):451–515, 2000.
12. Griffin, A., Hirvonen, T., Tzagkarakis, C., Mouchtaris, A., and Tsakalides, P., Single-channel and multi-channel sinusoidal audio coding using compressed sensing. *IEEE Trans. Audio Speech Lang. Process.* 19(5):1382–1395, 2011.
13. Barnes, C. B., and Ulrik, C. S., Asthma and adherence to inhaled corticosteroids: Current status and future perspectives. *Respir. Care* 60(3):455–468, 2015.
14. Sumino, K., and Cabana, M. D., Medication adherence in asthma patients. *Curr. Opin. Pulm. Med.* 19:49–53, 2013.
15. Kikidis, D., and et al., The digital asthma patient: The history and future of inhaler based health monitoring devices. *J. Aerosol Med. Pulm. Drug Deliv.* 29(3):219–232, 2016.
16. Ingerski, L. M., Hente, E. A., Modi, A. C., and Hommel, K. A., Electronic measurement of medication adherence in pediatric chronic illness: a review of measures. *J. Pediatr.* 159:528–534, 2011.
17. Chan, A. H., Reddel, H. K., Apter, A., Eakin, M., Riekert, K., and Foster, J. M., Adherence monitoring and e-health: How clinicians and researchers can use technology to promote inhaler adherence for asthma. *J. Allergy Clin. Immunol. Pract.* 1:446–454, 2013.
18. Chan, A. H., Harrison, J., Black, P. N., Mitchell, E. A., and Foster, J. M., Using electronic monitoring devices to measure inhaler adherence: A practical guide for clinicians. *J. Allergy Clin. Immunol. Pract.* 3(3):335–349, 2015. e5.
19. Boulet, L. P., Vervloet, D., Magar, Y., and Foster, J. M., Adherence: The goal to control asthma. *Clin. Chest. Med.* 33:405–417.
20. Taylor, T. E., and et al., An acoustic method to automatically detect pressurized metered dose inhaler actuations. Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE. IEEE, 2014.
21. Kikidis, D., Votis, K., and Tzovaras, D., Utilizing convolution neural networks for the acoustic detection of inhaler actuations. IEEE E-Health and Bioengineering Conference (EHB), 2015.
22. Khan, J. Y., Yuce, M. R., Bulger, G., and Harding, B., Wireless body area network (wban) design techniques and performance evaluation. *J. Med. Syst.* 36(3):1441–1457, 2012.
23. Friedman, J., Hastie, T., Hfling, H., and Tibshirani, R., Pathwise coordinate optimization, Tech. Rep., Ann. Appl. Stat., 2007.
24. Donoho, D. L., Compressed sensing. *IEEE Trans. Inf. Theory* 52(4):1289–1306, 2006.

25. Rubinstein, R., Bruckstein, A. M., and Elad, M., Dictionaries for sparse representation modeling. *Proc. IEEE* 98(6):1045–1057, 2010.
26. Gilbert, A., and Indyk, P., Sparse recovery using sparse matrices. *Proc. IEEE* 98(6):937–947, 2010.
27. LAN/MAN Standards Committee of the IEEE Computer Society: IEEE STandard for Local and metropolitan area networks – Part 15.6: Wireless Body Area Networks, IEEE Std 802.15.6–2012, 2012.
28. Movassaghi, S., and et al., Wireless body area networks: a survey. *IEEE Commun. Surveys Tuts.* PP(99):1–29, 2014.
29. Angelosante, D., Bazerque, J. A., and Giannakis, G. B., Online adaptive estimation of sparse signals: Where rls meets the ℓ_1 -norm. *IEEE Trans. Signal Process* 58(7):3436–3447, 2010.
30. Laska, J. N.., Boufounos, P. T.., Davenport, M. A.., and Baraniuk, R. G.., Democracy in action: Quantization, saturation, and compressive sensing. *Appl. Comput. Harmon. Anal.* 31(3):429–443, 2011. ISSN 1063-5203. http://dx.doi.org/10.1016/j.acha.2011.02.002.
31. Hastie, T., Tibshirani, R., and Friedman, J., The elements of statistical learning: data mining, inference, and prediction, Springer Series in Statistics, 2nd ed, 2009.