Dictionary Learning for Reconstructing Measurements of Analog Wireless Sensor Nodes

Christopher Willuweit, Carsten Bockelmann and Armin Dekorsy
Department of Communications Engineering
University of Bremen
Bremen, Germany
Email: {willuweit, bockelmann, dekorsy}@ant.uni-bremen.de

***Abstract*—Wireless Sensor Nodes communicating measurements to a base station is one of the scenarios in the emerging field of Machine-Type-Communication. Those systems rely on low complexity of the nodes, due to cost and energy consumption. The main idea of this paper is to employ a low complexity analog modulation scheme in the node, and combine it with state of the art digital signal processing in the base station. Specifically, we focus on Amplitude Modulation in a point to point scenario facing noise and hardware offsets. We show that under certain assumptions this transmission can be described by a linear model. Subsequently we utilize payload (measurement) signal structure, namely sparsity, to estimate the payload signals as well as the hardware offsets using a dictionary learning algorithm. Numerical simulations show, that for realistic noise assumptions the algorithms are able to reconstruct payload signals and estimate hardware offsets.**

***Index Terms*—machine type communication, wireless sensor networks, analog sensor communication, amplitude modulation, hardware offsets, dictionary learning, K-SVD**

I. Introduction

Machine-type-communication (MTC) is expected to grow tremendously in the next years [1]. One important scenario of MTC is Wireless Sensor Networks (WSN) where physical measurements are to be transmitted from multiple sensor nodes to a base station for further processing. One important requirement for such systems is low complexity entailing low system cost and energy consumption which is especially important e.g. for battery-driven systems or those relying on *Energy Harvesting* [2].

Commercially available systems almost always employ digital signal processing and communication. An abstract block diagram of such a system consisting of a single sensor node transmitting to the base station is shown in Fig. 1a). Independent of the used protocols or standards it first converts the sensors analog signal $x(t)$ into the digital domain using Analog-to-Digital-Converters (ADC). The digital signal is then further processed, e.g. for source or channel coding and finally mapped onto discrete symbols. These are converted back to the analog domain using Digital-to-Analog-Converters (DAC), mixed to the used frequency band in the analog front end (FE) and emitted via antennas. In the base station the received signal is mixed to baseband, converted to digital domain and processed to estimate the payload signal. This signal path from node to base station is referred to as Uplink (UL). Usually sensor nodes do not operate continuously with the same settings but are controlled by base station via the Downlink (DL). The DL is also used to synchronize the node to the base station to deal e.g. with carrier-frequency offsets (CFO) and time offsets (TO). To establish the DL the node has to provide a receiver path, including FE and ADC, scaling up its technical complexity.

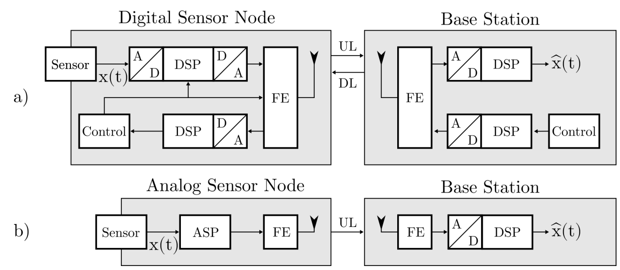

Fig. 1. a) Digital sensor node like in many commercially available systems. Payload signal $x(t)$ is converted into digital domain in which the communication signal is generated and transmitted via Uplink (UL). Those systems usually also provide a Downlink (DL) for system control and synchronization. In b) the proposed fully analog sensor node is depicted. Signal processing is done in analog domain (ASP) and the DL is spared.

A promising topic in current research is reducing system complexity using the *structure* of signals. One concept of describing signal structure is *sparsity*: the possibility of representing a signal by very few coefficients way below Nyquist rate [3].In [4] and [5] for example the sparsity of communication signals is used for Multiuser-Detection in WSN with a very large number of nodes. But also the concept of all-digital systems like in Fig.1a) is questioned. [6] for example uses low power analog signal processing (ASP) to compress the sensors signal $x(t)$ before it is sampled, reducing its bandwith and thus relieving the requirements for the following digital signal processing. This is only possible due to payload signals structured beyond bandlimitation. However, all these concepts still employ digital communication.

In this paper we propose a fully analog sensor node depicted in Fig 1b). We transmit using Analog Modulation whereas the base station is still digital trying to reconstruct the payload signal $x(t)$. Note that in this structure the downlink is omitted, further reducing complexity but imposing additional problems: Since no synchronization is possible, hardware offsets like the above mentioned CFO and TO have to be compensated by other means.

Among the various possibilities to realize a fully analog sensor node we focus on Amplitude Modulation (AM) of the payload signal. Further we only consider a point-to-point scenario where a single node is transmitting to the base station. It will be shown that the above mentioned concept of sparsity can help us finding reconstruction methods for the base station to recover the payload signal. Of further interest in this paper is the estimation of hardware offsets which we will provide algorithms for.

The remainder of this paper is organized as follows: First we will provide a system model in section II that breaks down to a linear model under certain assumptions. Signal reconstruction at the base station can then be interpreted as a 'dictionary learning problem' which we will tackle in section III. Also in this section we present methods to estimate the hardware offsets using the results generated by the dictionary learning. In section IV we show that the algorithms work fairly well under realistic noise assumptions using a numerical simulation for an exemplary setup. In sections V and VI we will conclude our results and present ideas for further research.

II. System Model

A. Analog modulation and digital reception

In this paper we focus on a specific realization of an analog sensor node based on Amplitude Modulation (AM) shown in Fig.1. The node is shown in Fig. 2 together with the assumed channel model and base station. AM in the node is realized

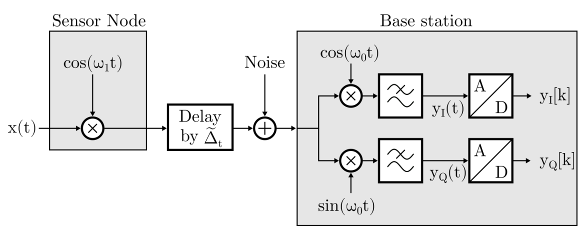

Fig. 2. Supposed structure of sensor node utilizing Amplitude Modulation, channel consisting of propagation delay and additive noise and the base station using IQ demodulation, anti aliasing lowpasses and analog-to-digital converters

by multiplying the payload signal $x(t)$ with a cosine carrier of angular frequency ω_1. The physical channel is assumed to impose a propagation delay of $\tilde{\Delta}_t$ and additive white gaussian noise (AWGN). Note that in this subsection noise will be neglected for the derivation of the system model. The base station applies a classical IQ demodulation using a carrier of angular frequency ω_0. The demodulated continuous signal in baseband representation

$$y(t) = y_I(t) + j \cdot y_Q(t) \tag{1}$$

is then sampled using standard Nyquist ADC generating a sequence

$$y[k] = y_I[k] + j \cdot y_Q[k]. \tag{2}$$

The task of the base station is the reconstruction of the nodes payload signal $x(t)$ applying DSP on the sampled receiver signal $y[k]$. Based the model shown in Fig. 2 we derive the connection between the payload signal $x(t)$ and the digital baseband representation $y[k]$. As a first step we formulate the connection between $x(t)$ and $y(t)$

$$y(t) = x(t - \tilde{\Delta}_t) \cdot \cos(\omega_1 t - \omega_1 \tilde{\Delta}_t) \cdot e^{-j\omega_0 t} \tag{3}$$

The IQ demodulation produces two frequency components: one at the difference of the carrier frequencies $\tilde{\Delta}_\omega := \omega_1 - \omega_0$, the other at the sum of the two frequencies. The latter part will be cancelled out by the anti aliasing filters in the front end resulting in the signal

$$y(t) = \frac{x(t - \tilde{\Delta}_t)}{2} \left(\cancelto{0}{e^{j(\omega_1 + \omega_0)t + j\omega_1 \tilde{\Delta}_t}} + e^{j\tilde{\Delta}_\omega t + j\omega_1 \tilde{\Delta}_t} \right). \tag{4}$$

It can be seen that $x(t)$ is multiplied with a complex exponential function with angular frequency $\tilde{\Delta}_\omega$ starting at phase $\omega_1 \tilde{\Delta}_t$. However, the phase at $t = 0$ in real systems is not only determined by the propagation delay $\tilde{\Delta}_t$ but also by the antennas and other phase perturbing elements in the FE. For this reason we assume the starting phase to be independent of the propagation delay and name it $\tilde{\Delta}_\phi$. After a proper scaling we get

$$y(t) = x(t - \tilde{\Delta}_t) \cdot e^{j(\tilde{\Delta}_\omega t + \tilde{\Delta}_\phi)} \tag{5}$$

at the ADCs input. To enable proper sampling of the signal, we assume a bandlimitation of $x(t)$ and consequently of $y(t)$. If $x(t)$ is limited to bandwidth B_{Tx} it can be written as a sum of delayed sinc functions also known as the *cardinal series* [7]. We additionally assume an approximate time limitation of $x(t)$, which leads to a *finite* cardinal series

$$x(t) = \sum_{n=0}^{N-1} x[n] \operatorname{sinc}\left(B_{\text{Tx}} t - n\right). \tag{6}$$

Note that this definition is strictly speaking not equal to approximately time- and bandlimited functions [8] but it is common practice to assume so [7]. So $x(t)$ has N degrees of freedom and is described by the finite sequence $x[n]$ with $n \in [0, N-1]$. From Eq. 5 follows that $y(t)$ has the same bandwith since its only shifted in frequency not scaled. Following the Shannon-Nyquist Sampling Theorem we have to sample $y(t)$ at rate $T_{\text{Rx}} := \frac{1}{B_{\text{Tx}}}$. If we now put Eq. 6 in Eq. 5 and evaluate it at $t = kT_{\text{Rx}}$ with $k \in \mathbb{N}$ we get a sequence describing our sampled signal

$$y[k] = e^{j(\tilde{\Delta}_\omega k T_{\text{Rx}} + \tilde{\Delta}_\phi)} \sum_{n=0}^{N-1} x[n] \operatorname{sinc}\left(k - n - \tilde{\Delta}_t B_{\text{Tx}}\right). \tag{7}$$

For reasons of better legibility we normalize $\tilde{\Delta}_\omega, \tilde{\Delta}_\phi$ and $\tilde{\Delta}_t$ and call them

$$\text{Carrier Frequency Offset (CFO):} \quad \Delta_\omega = \frac{\tilde{\Delta}_\omega}{2\pi B_{\text{Tx}}}, \tag{8}$$

$$\text{Phase Offset (PO):} \quad \Delta_\phi = \frac{\tilde{\Delta}_\phi}{2\pi}, \tag{9}$$

$$\text{Time Offset (TO):} \quad \Delta_t = B_{\text{Tx}} \tilde{\Delta}_t. \tag{10}$$

Because they are representing a complex phase, Δ_ω and Δ_ϕ are now in the restricted range from $-\frac{1}{2}$ to $\frac{1}{2}$. In this paper we focus on the case where the same holds for the TO Δ_t. In practice $|\Delta_t|$ can be greater than $\frac{1}{2}$ but this case is not considered here.

B. Formulation in linear algebra

So far we have derived the relation between the sampled receive signal and the payload signal. The task of the base station in Fig. 2 is to reconstruct $x(t)$ just with the knowledge of $y[k]$. To fulfill this task in finite time, k has to be in a finite range. Since $x(t)$ is assumed to have N degrees of freedom (Eq. 6) we will set $k = \{0, 1, ..., N-1\}$. Using this limitation, we can formulate a linear equation system which completely describes $y[k]$ in dependence of $x[n]$. For our derivation we define $\text{sinc}_{\Delta_t}(t) := \text{sinc}(t-\Delta_t)$: a time-shifted sinc function. Moreover we subdivide $y[k]$ in the sequences $c[k]$ and $z[k]$:

$$y[k] = \underbrace{e^{j2\pi(\Delta_\omega k+\Delta_\phi)}}_{=:c[k]} \underbrace{\sum_{n=0}^{N-1} x[n]\text{sinc}_{\Delta_t}(k-n)}_{=:z[k]}. \tag{11}$$

Since all the sequences are of length N, we can put them in vectors. For $x[k]$ we define

$$\mathbf{x} = \begin{bmatrix} x[0] & x[1] & \cdots & x[N-1] \end{bmatrix}^T \in \mathbb{C}^N \tag{12}$$

and do the same to get $\mathbf{y}, \mathbf{c}$ and $\mathbf{z}$, each in $\mathbb{C}^N$. The elements of the sum in Eq.11 only depend on $k-n$. Subsequently $\mathbf{z}$ can be described by a Toeplitz matrix $\mathbf{T}$ left multiplied with $\mathbf{x}$:

$$\mathbf{z} = \mathbf{T}\mathbf{x}. \tag{13}$$

Since $\mathbf{T} \in \mathbb{R}^{N\times N}$ is Toeplitz, its elements can be described by a vector $\mathbf{t} \in \mathbb{R}^{2N-1}$ using the *toep*-operator:

$$\mathbf{t} = [\text{sinc}_{\Delta_t}(1-N)\dots\text{sinc}_{\Delta_t}(0)\dots\text{sinc}_{\Delta_t}(N-1)], \tag{14}$$

$$\mathbf{T} = \text{toep}(\mathbf{t}) \Leftrightarrow [T]_{n,k} = [t]_{k-n+N}. \tag{15}$$

To get from $\mathbf{z}$ to $\mathbf{y}$ we have to perform an elementwise multiplication of $\mathbf{z}$ with $\mathbf{c}$. This can also be formulated as left-multiplying a diagonal matrix $\mathbf{C}$ with $\mathbf{z}$:

$$\mathbf{c} = e^{j2\pi\Delta_\phi} \begin{bmatrix} 1 & e^{j2\pi\Delta_\omega} & \dots & e^{j2\pi(N-1)\Delta_\omega} \end{bmatrix}, \tag{16}$$

$$\mathbf{C} = \text{diag}(\mathbf{c}) \in \mathbb{C}^{N\times N}. \tag{17}$$

At this point we introduce noise again. Usually it is assumed that AWGN in the high frequency domain can be represented as circular white gaussian noise in baseband representation. We take that into account by adding a vector $\mathbf{n} \in \mathbb{C}^N$ whose elements are i.i.d. taken from $\mathcal{CN}(0, \sigma^2)$:

$$\mathbf{y} = \mathbf{C}\mathbf{z} + \mathbf{n} = \underbrace{\mathbf{C}\mathbf{T}}_{\mathbf{D}}\mathbf{x} + \mathbf{n}. \tag{18}$$

Assuming time invariant offsets multiple transmit signals $\mathbf{x}_l$ and respective receive signals $\mathbf{y}_l$ with $l \in [0, L-1]$ can be summarized in a matrix equation

$$\mathbf{Y} = \mathbf{D}\mathbf{X} + \mathbf{N}, \tag{19}$$

where $\mathbf{D} := \mathbf{C}\mathbf{T}$ and $\mathbf{X}, \mathbf{Y} \in \mathbb{C}^{N\times L}$ containing receive signals $\mathbf{y}_l$ respectively payload signals $\mathbf{x}_l$ as columns. The elements of $\mathbf{N} \in \mathbb{C}^{N\times L}$ are again taken i.i.d. from $\mathcal{CN}(0, \sigma^2)$. For reasons clarified in the following sections, we call $\mathbf{D}$ a *dictionary*.

C. Sparsity

At this point we have a linear model of the relation between transmit and receive signals. The task of the base station is to recover the payload signals respectively their discrete representation $\mathbf{X}$ from the received samples $\mathbf{Y}$. This is only possible if we assume a certain structure of $\mathbf{X}$ which will be shown in Section III. Luckily many real world signals have a structure. One way to describe it is *sparsity*. Namely we assume that each payload signal can be described by

$$\mathbf{x}_l = \mathbf{\Psi}\mathbf{s}_l \quad \text{s.t.} \quad ||\mathbf{s}_l||_0 \le S \ll N, \tag{20}$$

where $\mathbf{\Psi} \in \mathbb{C}^{N\times N}$ is a basis usually called *sparse basis* and $\mathbf{s}_l \in \mathbb{C}^N$ is the *sparse representation* only having $S \in \mathbb{N}$ nonzero elements. An example for the corresponding *sparse bandlimited signals* $x(t)$ is given in [9]. In the next chapter it will be shown, that recovering $\mathbf{X}$ or $\mathbf{S}$ from $\mathbf{Y}$ is not solvable in general due to phase ambiguity. To alleviate this problem we make two last assumptions on $\mathbf{X}$. First we focus on the canonical sparse basis $\mathbf{\Psi} = \mathbf{I}$. So $\mathbf{X}$ itself is considered sparse. Second we assume that the elements of $\mathbf{X}$ are from $\mathbb{R}_+$. The generalization for basis elements and sparse coefficients to $\mathbb{R}$ or even $\mathbb{C}$ is left for further research.

III. Data and parameter estimation

Primary objective of the base station is to estimate $\mathbf{X}$ from $\mathbf{Y}$. This task is equivalent to factorize the matrix $\mathbf{Y}$ into $\mathbf{D}$ and $\mathbf{X}$. Hence it is called a *matrix factorization problem* and once we have an estimate $\hat{\mathbf{X}}$, we will also obtain an estimate $\hat{\mathbf{D}}$. The question arises if we can estimate the hardware offsets based on the estimated dictionary. This is important for two reasons: First because it will turn out that we need an estimate of the TO to process our estimate on $\mathbf{X}$ after the factorization. Second for multiple transmissions it might be useful to estimate the hardware offsets in the first transmission and afterwards use the corresponding dictionary to reconstruct all following transmissions.

The above mentioned matrix factorization problem has infinitely many solutions in general. In [10] the authors showed that for sparse $\mathbf{X}$ and especially if $\mathbf{D}$ is invertible, all sparse solutions are permuted and scaled versions of each other. This is called an *ambiguity*. The mathematical proof of invertibility of $\mathbf{D}$ is spared in this work due to lack of space, but will be available in future publications. We describe the ambiguity in scaling and permutation with the equivalence

$$\widehat{\mathbf{D}}\widehat{\mathbf{X}} = \underbrace{\overbrace{\widehat{\mathbf{D}}\mathbf{\Sigma}^{-1}}^{\widehat{\mathbf{D}}_N}\mathbf{P}^T}_{=:\widehat{\mathbf{D}}_P}\underbrace{\mathbf{P}\overbrace{\mathbf{\Sigma}\widehat{\mathbf{X}}}^{\widehat{\mathbf{X}}_N}}_{\widehat{\mathbf{X}}_P}, \tag{21}$$

where $\mathbf{\Sigma}$ is an invertible diagonal matrix characterizing the scaling and $\mathbf{P}$ is a permutation matrix. To fulfill Eq. 21,

$\mathbf{\Sigma}^{-1}\mathbf{P}^T\mathbf{P}\mathbf{\Sigma} = \mathbf{I}$ must apply. Additionally we assume the elements of $\mathbf{X}$ to be purely positive. This narrows $\mathbf{\Sigma}$ to purely positive elements as well. In simple words: If we find a $\widehat{\mathbf{X}}_P$ whose columns are S-sparse, it differs from the right solution $\widehat{\mathbf{X}}$ only in permuted and scaled rows. Same holds for the columns of $\widehat{\mathbf{D}}$.

In subsection III-A we describe how to find a sparse solution $\widehat{\mathbf{X}}_P$ and $\widehat{\mathbf{D}}_P$ using the well known dictionary learning algorithm K-SVD [11] which is slightly modified. Consequently in subsection III-B we use a sorting algorithm to remove the permutation ambiguity and get data and dictionary estimates that are just scaled by $\mathbf{\Sigma}$. To find $\mathbf{\Sigma}$ we need an estimate $\hat{\Delta}_t$ of the time offset. In subsection III-C the estimation of the hardware offsets and $\mathbf{\Sigma}$ is described. This estimate is used to scale the columns of $\widehat{\mathbf{D}}_N$ and rows of $\widehat{\mathbf{X}}_N$ to obtain the final estimates for dictionary and data

All these steps necessary to obtain estimates on payload signal and hardware offsets are depicted as a block diagram in Fig. 3.

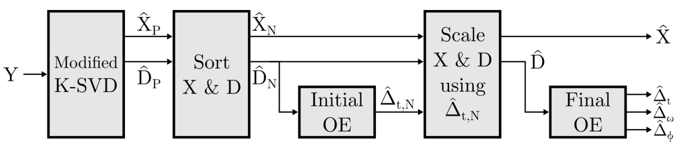

Fig. 3. The overall algorithm to estimate payload data and offsets from the received signal. It consists of a K-SVD (see III-A) sorting algorithm (see III-B) and a two stage offset estimation (see III-C and III-D).

A. Dictionary learning

To obtain a sparse representation of $\mathbf{Y}$ we use the well known K-SVD algorithm for dictionary learning. It tries to iteratively solve the minimization problem

$$\{\hat{\mathbf{X}}_P, \hat{\mathbf{D}}_P\} = \underset{\widetilde{\mathbf{X}}, \widetilde{\mathbf{D}}}{\text{argmin}} ||\mathbf{Y} - \widetilde{\mathbf{D}}\widetilde{\mathbf{X}}||_F^2 \tag{22}$$

with the columns of $\widetilde{\mathbf{X}}$ being sparse and the columns of $\widetilde{\mathbf{D}}$ having unit length. The first step in each iteration is sparse coding of $\mathbf{Y}$ with respect to the current dictionary. In the second step the dictionary is updated using a generalization of K-means algorithm. This procedure is repeated until a stopping criterion is met. For the first iteration an initial guess of the dictionary is needed. A detailed description of the algorithm can be found in [11].

In our work we use an implementation of K-SVD that works with complex dictionaries and data [12] and modified it to limit $\widehat{\mathbf{X}}_P$ to strictly positive elements. This is done by setting the phase of every nonzero element of $\widehat{\mathbf{X}}_P$ in each iteration to zero after sparse coding. For sparse coding the well known Orthogonal Matching Pursuit (OMP) is used. For OMP stopping criterion we assume known sparsity of the columns of $\widetilde{\mathbf{X}}$.

As initial dictionary we use a random complex matrix, each entry drawn from a circular gaussian distribution. Initial dictionaries with the expected structure of the true dictionary might seem like a better choice. However, we experienced that a random initial dictionary gives better results in terms of convergence speed.

It should be mentioned that there is no guarantee that K-SVD will reach the global optimum but in many cases it is a valid approach.

B. Sorting the estimated dictionary and data

As mentioned in the introduction of this section, we have to find the right permutation matrix $\mathbf{P}$ for sorting columns of $\widehat{\mathbf{D}}_P$ and subsequently rows of $\widehat{\mathbf{X}}_P$. We make use of the fact that the true dictionary $\mathbf{D}$ attains its per-column maxima at the main diagonal. This can easily be seen in Eq. 14 where $[t]_{k-n}$ attains its maximum at $k = n$ corresponding to the main diagonal of $\mathbf{T}$ defined in Eq. 15. To permute the columns of $\widehat{\mathbf{D}}_P$ in the way that the main diagonal contains maximum values we have to maximize the trace of its elementwise magnitude (called $\text{trace}_{\text{abs}}$ here):

$$\widehat{\mathbf{D}}_N = \max_{\mathbf{P}} \text{trace}_{\text{abs}} \left(\widehat{\mathbf{D}}_P \mathbf{P}^T \right). \tag{23}$$

This problem can be efficiently solved using the *hungarian method* described in [13]. After obtaining $\mathbf{P}$, we can use it to sort the Data and the dictionary:

$$\widehat{\mathbf{X}}_N = \mathbf{P}\hat{\mathbf{X}}_P \qquad \widehat{\mathbf{D}}_N = \hat{\mathbf{D}}_P\mathbf{P}^T \tag{24}$$

C. Offset estimation

As mentioned in the introduction of this section we need to know $\hat{\Delta}_t$ to scale dictionary and data. In the initial offset estimation we estimate the TO using $\widehat{\mathbf{D}}_N$. To estimate the TO, CFO and PO also have to be estimated. The proposed procedure for offset estimation is summarized in algorithm 1.

First we estimate Δ_ω and Δ_ϕ in steps 1 to 4. The basic approach is, that the left multiplication of $\mathbf{C}$ with $\mathbf{T}$ results in a phase rotation of the rows of $\mathbf{T}$. Also all elements of the main diagonal of $\mathbf{T}$ are given by $\text{sinc}_{\Delta_t}(0)$ which is positive. It follows that the main diagonal of $\mathbf{D}$ contains a sampled complex exponential function with frequency Δ_ω, phase offset Δ_ϕ and magnitude $\text{sinc}_{\Delta_t}(0)$. So to estimate CFO and PO we perform a spectral estimation on the main diagonal of $\widehat{\mathbf{D}}_N$. In this paper we used a zero padded periodogram (step 2 and 3) to find $\hat{\Delta}_\omega$ and subsequently calculate $\hat{\Delta}_\phi$ (step 4) as suggested in [14]. The total number of used samples N_{ZP} including zero padding determines the frequency resolution of the estimation. Using $\hat{\Delta}_\omega$ and $\hat{\Delta}_\phi$ we can compensate the CFO and PO related phase from $\widehat{\mathbf{D}}$ to get an estimate of $\mathbf{T}$ in steps 5 and 6. Because the elements of $\mathbf{T}$ are purely real, we discard the imaginary part of $\hat{\mathbf{C}}^{-1}\widehat{\mathbf{D}}$.

$\mathbf{T}$ as defined in Eq. 15 only contains values from the vector $\mathbf{t}$ which are samples of sinc_{Δ_t} function. From $\widehat{\mathbf{T}}$ an estimate on $\mathbf{t}$ is generated by taking the mean value of each of the $2N-1$ diagonals of $\widehat{\mathbf{T}}$ (step 7). In $\hat{\mathbf{t}}$ we now have estimates of $\text{sinc}_{\Delta_t}(m)$ with $m \in [1-N, N-1]$. At $m = 0$ the estimates are based on the N main diagonal elements of $\widehat{\mathbf{T}}$ whereas at $m = 1-N$ the estimate is just based on the single value $T_{1,N}$.

Now TO estimation is done by a Nonlinear Weighted Least Squares (NWLS) curve fitting [15]. The weighting specified

Algorithm 1: Proposed algorithm for estimating the offsets given an estimated dictionary. Parameter N_{ZP} determines the resolution for CFO-Estimation.

inputs : $\hat{\mathbf{D}} \in \mathbb{C}^{N\times N}$, $N_{\text{ZP}} \in \mathbb{N}$
outputs: Three real scalars $\hat{\Delta}_\omega, \hat{\Delta}_\phi, \hat{\Delta}_t$

// CFO and PO estimation
1 $\mathbf{d} \leftarrow \text{diag}(\hat{\mathbf{D}})$;
2 $\mathbf{d}_{\text{ZP}} = \begin{bmatrix}\mathbf{d} & 0 & \cdots & 0\end{bmatrix} \in \mathbb{C}^{N_{\text{ZP}}}$;
3 $\hat{\Delta}_\omega \leftarrow \max(\text{Periodogram}(\mathbf{d}_{\text{ZP}}))$; *// acc. to [14]*
4 $\hat{\Delta}_\phi \leftarrow \text{atan}\left(\frac{\text{Im}\left(\sum_{n=0}^{N-1} d_n e^{-j2\pi\hat{\Delta}_\omega n}\right)}{\text{Re}\left(\sum_{n=0}^{N-1} d_n e^{-j2\pi\hat{\Delta}_\omega n}\right)}\right)$;
// TO estimation
5 Generate $\hat{\mathbf{C}}$ using $\hat{\Delta}_\omega, \hat{\Delta}_\phi$ and Eq. 16 ;
6 $\hat{\mathbf{T}} \leftarrow \text{Re}\{\hat{\mathbf{C}}^{-1}\widehat{\mathbf{D}}\}$;
7 $[\hat{t}_{k+N}] \leftarrow \left[\frac{1}{N-k}\sum_{i=1}^{N-k}\hat{T}_{i,i+k}\right]$;
8 $\mathbf{w} = \begin{bmatrix}1 & 2 & \cdots & N & N-1 & \cdots & 1\end{bmatrix}^T$;
9 $\hat{\Delta}_t \leftarrow \underset{\substack{\Delta_t \in [-0.5, 0.5] \\ A \in \mathbb{R}_+}}{\text{argmin}} \left|\sum_{i=1}^{N} \hat{t}_i - A \cdot w_i \cdot \text{sinc}(i-1-\Delta_t)\right|^2$;

// using NWLS [15] with init. params. $\Delta_t = 0,\ A = 1$;

in vector $\mathbf{w}$ (step 8) is set according to the number of entries in $\hat{\mathbf{T}}$ that make up the corresponding entry of $\hat{\mathbf{t}}$.

D. Data and final offset estimation

With the previously estimated time offset $\hat{\Delta}_{t,N}$ we can generate a modeled $\widetilde{\mathbf{T}}$ according to Eq. 14 and 15. By determining the euclidian norms of the columns of $\widetilde{\mathbf{T}}$ we consequently calculate $\mathbf{\Sigma}$:

$$\Sigma_{k,k} = \left|\left|\begin{bmatrix}\tilde{T}_{1,k} & \tilde{T}_{2,k} & \cdots & \tilde{T}_{N,k}\end{bmatrix}\right|\right|_2, \tag{25}$$

which is used to scale the rows of $\widehat{\mathbf{X}}_N$ an the columns of $\widehat{\mathbf{D}}_N$ to obtain the final data and dictionary estimates:

$$\widehat{\mathbf{X}} = \mathbf{\Sigma}^{-1}\widehat{\mathbf{X}}_N \qquad \widehat{\mathbf{D}} = \widehat{\mathbf{D}}_N\mathbf{\Sigma}. \tag{26}$$

The latter we use to perform a final offset estimation by using the procedure described in sec. III-C. This time $\hat{\mathbf{D}}$ is used as input to generate the final offset estimates $\hat{\Delta}_t$, $\hat{\Delta}_\omega$ and $\hat{\Delta}_\phi$.

IV. Numerical Experiments

A. Defining error measures

The base stations main objective is to recover the payload signals respectively its discrete representation $\mathbf{X}$. To measure the reconstruction error we use the Normalized Mean Squared Error (NMSE)

$$\text{NMSE}(\hat{\mathbf{X}}) = \frac{1}{L}\sum_{l=1}^{L}\frac{||\mathbf{x}_l - \hat{\mathbf{x}}_l||_2^2}{||\mathbf{x}||_2^2}. \tag{27}$$

Secundary objective of the base station is to estimate the hardware offsets. Since the effects of the CFO and PO are cyclic, we also have to use a cyclic distance function to measure estimation errors:

$$E_{\Delta_{\omega,\phi}} = \left|\frac{\text{atan}\left(\frac{\sin(2\pi(\Delta_{\omega,\phi}-\hat{\Delta}_{\omega,\phi}))}{\cos(2\pi(\Delta_{\omega,\phi}-\hat{\Delta}_{\omega,\phi}))}\right)}{2\pi}\right|, \tag{28}$$

where $\Delta_{\omega,\phi}$ and $\hat{\Delta}_{\omega,\phi}$ represent CFO/PO and their estimates. The maximum error for CFO and PO is $\frac{1}{2}$. For TO we can measure the distance $E_{\Delta_t} = \left|\Delta_t - \hat{\Delta}_t\right|$ directly producing a maximum error of 1.

B. Simulation setup

In a toy example we will show that the proposed algorithm gives good results at least for the assumed model and algorithm parameters summarized in Table I. We use a total of

TABLE I
Used model and algorithm parameters.

Fixed parameter	Value	Random parameter	Range/Value
N	64	Δ_ω	$(-0.5, 0.5)$
S	6	Δ_ϕ	$(-0.5, 0.5)$
L	1024	Δ_t	$(-0.5, 0.5)$
SNR [dB]	$\{0, 6, \infty\}$	Entries of $\mathbf{N}$	$\mathcal{CN}(0, \sigma^2)$
N_{ZP}	1024		

1000 simulation runs per SNR. In each run an $\mathbf{X}$ is generated: nonzero elements are randomly placed on S of the N possible indices. The nonzero values are taken from a folded normal distribution to keep them purely positive. We also simulated using uniformly distributed values but the results did not show a qualitative difference. After $\mathbf{X}$ is generated, it is perturbed by randomly chosen offsets and noise according to section II. The proposed algorithm is then applied on the receive signals for 100 K-SVD iterations including offset estimation. The resulting error measures defined in the previous subsection are then averaged over the 1000 simulation runs. As a benchmark for NMSE we also apply the OMP on the true dictionaries, which we call *Oracle-OMP*. By using $N_{\text{ZP}} = 1024$ we have a resolution for $\hat{\Delta}_\omega$ of approximately 10^{-3}.

C. Results

The NMSE of the reconstructed payload signals for the chosen SNRs is depicted in Fig. 4. The errors are decreasing over the iterations seemingly converging to the oracle-OMP. As expected, the saturation level respectively the oracle-OMP error is smaller for higher SNRs. Not shown in Fig 4 is the oracle-OMP for the noiseless case which is at about 10^{-31}, which can only be explained by numerical effects.

Since CFO and PO error only show changes in the first three iterations in any SNR setting, we present them in Table II. It can be seen, that the estimation already gives good results in the first iteration. CFO estimation error quickly converges to the expected error given by $\frac{1}{4\cdot N_{ZP}} \approx 0.24 \cdot 10^{-3}$. E_{Δ_ω} and E_{Δ_ϕ} are slightly higher in the noiseless case. This somewhat unlikely result could not be explained so far. However, the absolute values of CFO and PO errors are nearly at the resolution of the periodogram and not expected to cause significant errors in $\widehat{\mathbf{X}}$.

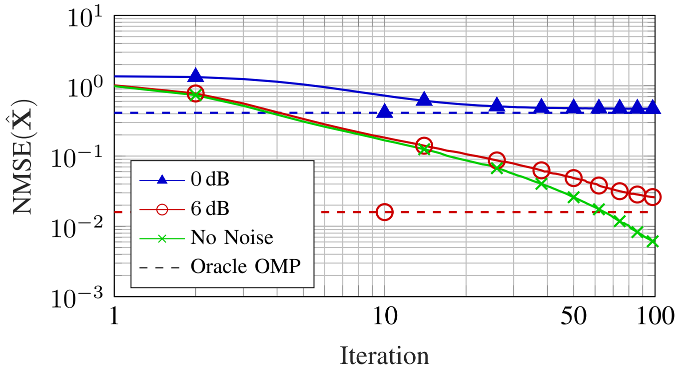

Fig. 4. Normalized mean squared error of the estimated payload matrix $\hat{\mathbf{X}}$ versus iterations for different noise levels.

TABLE II
ESTIMATION ERROR OF CFO AND PO FOR DIFFERENT SNRS.

	0dB SNR		6dB SNR		Noiseless	
Iteration	1	> 2	1	> 2	1	> 2
$E_{\Delta_\omega} \cdot 10^3$	0.40	0.24	0.26	0.24	0.29	0.25
$E_{\Delta_\phi} \cdot 10^3$	13.0	7.7	8.5	7.7	9.7	8.0

The time offset estimation error E_{Δ_t} is depicted in Fig. 5. It can be seen that in the first 20 iterations the TO estimation does not benefit from SNRs above 6 dB. After 20 iterations the curves become unsteady. This can be explained by the simulation runs where the TO estimation did not converge. This is the case for approximately 0.1 % of the simulations probably caused by a TO $|\Delta_t| \approx 0.5$.

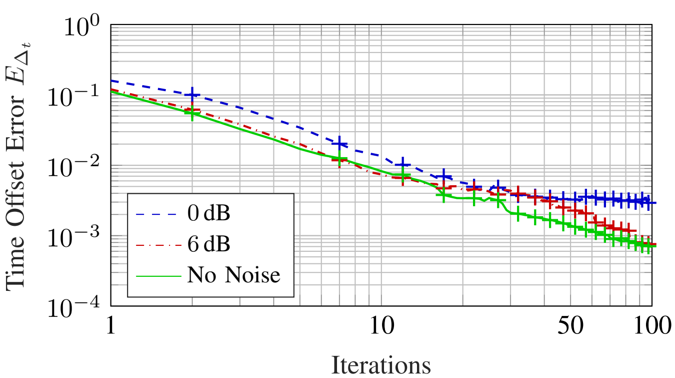

Fig. 5. Mean error of the Time Offset estimation $\hat{\Delta}_t$ for different SNRs versus iterations. Maximum is given by the cyclic error measure for fractional Time Offsets.

V. CONCLUSION

A model for a fully analog sensor node communicating to a digital base station using Amplitude Modulation was proposed. It was shown that for the assumptions of signal sparsity the base station can recover the payload signals and estimate the occuring hardware offsets between node and base station. The algorithms for signal reconstruction and hardware offset estimation based on dictionary learning were presented and evaluated in a numerical simulation using realistic assumptions on SNR. The results indicate that estimation of carrier frequency offset and phase offset is not a big issue for the algorithm since they already converge after three iterations. The estimation of the time offset needs about 20 Iteration to converge. In general, dictionary learning algorithms like K-SVD seem to be valuable for estimating payload signals transmitted using analog modulation if additional measures are taken.

VI. FUTURE WORK

To lead the proposed concepts to a more practical direction, the model restrictions have to be resolved. A major advantage would be the possibility to use arbitrary sparse bases. For this, an additional phase ambiguity in dictionary learning rather than just scaling and permutation has to be considered.

Another enhancement of the proposed model is to allow time variant hardware offsets. Whereas carrier frequency offsets in real life systems can often be considered time invariant the same does not hold for time offset and phase offset.

REFERENCES

[1] T. Taleb and A. Kunz. "Machine type communications in 3GPP networks: potential, challenges, and solutions." IEEE Communications Magazine Vol. 50, No. 3 (2012)

[2] X. Fafoutis, et al. "Energy-harvesting wireless sensor networks." in Proc., 9th European Conference on Wireless Sensor Networks (EWSN). 2012

[3] M. Verhelst and A. Bahai. "Where Analog Meets Digital: Analog to Information Conversion and Beyond." IEEE Solid-State Circuits Magazine, Vol.7,No.3, pp.67–80. 2015

[4] C. Bockelmann, H. Schepker and A. Dekorsy. "Compressive Sensing based Multi-User Detection for Machine-to-Machine Communication" Transactions on Emerging Telecommunications Technologies: Special Issue on Machine-to-Machine: An emerging communication paradigm, Vol. 24, June 2013, pp.389–400

[5] F. Monsees, M. Woltering, C. Bockelmann and A. Dekorsy. "A Potential Solution for MTC: Multi-Carrier Compressive Sensing Multi-User Detection" in Proc., The Asilomar Conference on Signals, Systems, and Computers. 2015, pp.18–22

[6] B. Rumberg, D.W. Graham, V. Kulathumani, und R. Fernandez: "Hibernets: Energy-Efficient Sensor Networks Using Analog Signal Processing", IEEE Journal on Emerging and Selected Topics in Circuits and Systems, Vol. 1, No. 3, 2011, pp.321–334

[7] A.I. Zayed, "Advances in Shannons Sampling Theory", London, CRC Press, 1993, p.9

[8] H.J.Landau and H.O.Pollak, "Prolate Spheroidal Wave Functions, Fourier analysis and uncertainty–II" Bell Labs Technical Journal, vol.40, pp.65–84, Jan.1961

[9] J.A. Tropp, et al. "Beyond Nyquist: Efficient sampling of sparse bandlimited signals." IEEE Transactions on Information Theory Vol.56,No.1, pp.520–544. 2010

[10] C. Hillar, and F. Sommer. "When can dictionary learning uniquely recover sparse data from subsamples?" IEEE Transactions on Information Theory Vol. 61, No. 11, 2015, pp.6290–6297.

[11] M. Aharon, M. Elad and AM. Bruckstein, "K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation" IEEE Transactions on signal processing, Vol.54 ,No.11, p.4311, 2006

[12] J. Turek, M. Elad, and I. Yavneh, "Clutter Mitigation in Echocardiography using Sparse Signal Separation", International Journal on Biomedical Imaging, 2015

[13] H.W. Kuhn, "The Hungarian Method for the Assignment Problem" Naval Research Logistics Quarterly 2, pp.83–97, 1955

[14] S.M.Kay, "Modern Spectral Estimation - Theory and Application" Englewood Cliffs, NJ, PTR Prentice Hall, 1988, p.409

[15] G.A.F. Seber, C. J. Wild. "Nonlinear Regression" Hoboken, NJ: Wiley-Interscience, 2003, pp.26