Recovery of missing samples in Orthogonal Frequency Division Multiplexing signals with optimisation using data carriers

Received: 26 October 2023 | Revised: 13 February 2024 | Accepted: 10 March 2024

IET Radar, Sonar & Navigation

DOI: 10.1049/rsn2.12560

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

Recovery of missing samples in Orthogonal Frequency Division Multiplexing signals with optimisation using data carriers

Anders Haglund[1,2] | **Per-Olov Frölind**[1] | **Lars M. H. Ulander**[1,2]

[1]Swedish Defence Research Agency (FOI), Linköping, Sweden

[2]Department of Space, Earth and Environment, Chalmers University of Technology, Gothenburg, Sweden

Correspondence

Anders Haglund.
Email: anders.haglund@foi.se

Funding information

Totalförsvarets Forskningsinstitut; Chalmers Tekniska Högskola

Abstract

A method is proposed for reconstructing an Orthogonal Frequency Division Multiplexing (OFDM) signal that contains data gaps, with the aim to improve demodulation. The main objective is to use the method in a passive radar application with missing data samples and to improve target detection. The OFDM signal is assumed to comply with the Digital Video Broadcasting Terrestrial standard. The proposed recovery method is based on optimisation of a novel objective function, which consists of two parts. The first part is a function of the energy in the out-of-band frequencies, whereas the second, and novel part, uses the location of data carriers in the constellation diagram. The method is evaluated using both simulations and real data. The authors show that the proposed method significantly improves the OFDM signal in just a few iteration steps. The proposed method improved the condition number more than a factor ten thousand millions compared to using the least square method on the out-of-band frequencies only. The authors also decode the symbols with the Viterbi decoding algorithm and show how the required number of iterations with the proposed algorithm depends on the amount of missing samples and on the Signal-to-Noise Ratio in order to achieve a Bit Error Rate of less than one in one hundred thousand millions.

KEYWORDS

channel estimation, OFDM modulation, passive radar, radar signal processing, signal reconstruction

1 | INTRODUCTION

The aim of this paper is to propose and evaluate a method to facilitate demodulation of an Orthogonal Frequency Division Multiplexing (OFDM) signal sampled with a receiver that does not register all samples in a symbol but where some samples are missing.

In our application case, samples are lost in a radar system, originally built for transmitting pulses and receiving pulse echoes, when it operates as a passive radar. By demodulating the OFDM signal (Digital Video Broadcasting Terrestrial (DVB-T) [1] in our case), it is possible to reconstruct a time signal that closer resembles the originally transmitted signal (e.g. with less multipath effects) that can function as a reference signal with better quality than the originally received signal [2, 3].

The problem of missing samples was previously investigated [4] and it was shown that missing samples give rise to significant performance degradation. It was suggested to consider using algorithms that mitigate the effect of missing samples in order to improve the results.

In several different applications, some of the samples of a band-limited signal are lost due to different reasons [5, 6]. The problem of missing samples has given rise to the research and development of different methods for reconstructing the missing samples during the latest decades and several methods have been proposed [5, 6]. However, very little research was found about recovery of missing samples explicitly in OFDM signals. The algorithm proposed here takes advantage of the OFDM signal structure.

In passive radar application, no pulses are transmitted but instead, signals originating from other, unsynchronised

sources are received. Passive radar uses transmitters of opportunity and is a technology that has been under strong development in recent years [7–9]. In the processing part, the bistatic range is determined by measuring the time difference between the signal reflected by the target (surveillance channel) and the direct signal from the selected transmitter (reference channel). The time difference is measured by correlating the signals, and the cleaner the reference signal is from noise and multipath effects the better the quality of the results.

When demodulating the DVB-T OFDM signal [3], the discrete structure of the OFDM symbol makes it possible to make corrections to the received signal. A reference time signal can then be reconstructed by modulating the corrected signal.

This is especially desirable in the absence of a dedicated directional antenna as is the case for the LORA (LOw frequency RAdar) [10, 11] Synthetic Aperture Radar (SAR) system, where wide-beam antenna elements are mounted fixed on an airborne platform. The LORA system does not register measured samples continuously but about 1% of the samples are missing in periodically distributed groups.

After demodulation, it is also possible to decrease the strength of certain ambiguities in the DVB-T signal by manipulating for example the guard interval and the pilot carriers in the reconstructed reference signal. A constant false alarm rate (CFAR) detector using methods based on such manipulated reference signals can avoid false alarms of side peaks and improve target detection [12].

In this paper, an algorithm is proposed that recovers (estimates) the missing samples in the signal when demodulating the OFDM signal. The proposed method is based on minimising the data carriers' deviations from their respective correct constellation point.

Although not necessary for reconstructing a reference signal for application in passive radar, the signal is further decoded in this paper in order to demonstrate how the proposed algorithm improves the Bit Error Rate (BER) in applications where the information in the encoded data bits is actually used. It is also possible to encode and modulate the decoded (and error corrected) signal for use as a reference signal in passive radar, thereby possibly achieving better quality. However, it requires implementation of decoding and encoding algorithms and the execution work will increase.

A constellation diagram of the complex amplitude (in frequency domain) of all data carriers in a simulated DVB-T symbol, is shown in Figure 1 below. In the symbol data, about 1% of the samples in the time domain have been set to zero to represent the missing samples from LORA but there is no noise present. The data carriers' amplitudes deviate significantly from the ideal 8 × 8 points square grid constellation because the zeroed samples give rise to inter-carrier interference. Time samples that are replaced with values that differ from their true values impose errors on carriers across the frequency spectrum. Noise in the signal will give additional deviations.

2 | BACKGROUND

2.1 | The Digital Video Broadcasting Terrestrial signal

The channel coding and modulation for the DVB-T signal is described by the standard in ref. [1]. The signal is specified for 8 MHz, 7 MHz and 6 MHz channels, with the same main specification for these bandwidths apart from the choice of the elementary time period (e.g. 7/64 μs for 8 MHz channel spacing). Two modes of operation are defined, a "2K mode" and an "8K mode". Orthogonal Frequency Division Multiplexing transmission is used and the data carriers in an OFDM symbol are modulated with one of Quadrature Phase Shift Keying, 16-QAM (Quadrature Amplitude Modulation (QAM)) or 64-QAM (uniform or non-uniform). Before the data is mapped onto the signal constellation, the data is randomised.

In this paper, we focus on the 8K mode but the proposed method does not depend on this mode so we expect that the same method works for other modes. For the 8K mode, the signal has 6817 carriers, out of which there are 6048 data carriers. The data carriers are exploited in the proposed algorithm. The rest of the carriers are pilot carriers (continual or scattered) or Transmission Parameter Signaling (TPS) carriers [1] (not included in Figure 1). The pilot carriers are mapped onto two different constellation points with a "boosted" power on the real axis. The TPS carriers use Differential Binary Phase Shift Keying and are also mapped onto two different constellation points on the real axis but with lower power. Both pilot and TPS carriers are predictable over a number of symbols, and the pilots can be utilised for estimating the channel, for synchronisation in time and frequency, for determining the transmission mode, and for estimating the phase noise [1]. The time domain signal can then be created by applying the inverse discrete Fourier transform to the symbol's constellation. In the time domain, each symbol is composed of two parts. Before the useful part, a guard interval part consisting of a copy of the last part of the useful part is inserted. The guard interval has a length of 1/4, 1/8, 1/16 or 1/32 times the length of the useful part [1]. For 8K mode, the duration of the useful part of the symbol, T_U, is 896 μs. The symbols are also organised in frames and super-frames consisting of fixed numbers of symbols. The frame and super-frame structures are not exploited in this study. Before the data stream is mapped onto the signal constellation, the data stream will first be randomised for attaining energy dispersal, then subject to outer coding (Reed-Solomon code for error protection) and convolutional byte-wise interleaving. After that, the data undergo inner coding (punctured convolutional code), introducing redundancy for error correction. The amount of redundancy is described by the code rate. Then bit-wise interleaving is used followed by symbol interleaving. The interleaved data is then mapped onto the constellation using Gray coding [1].

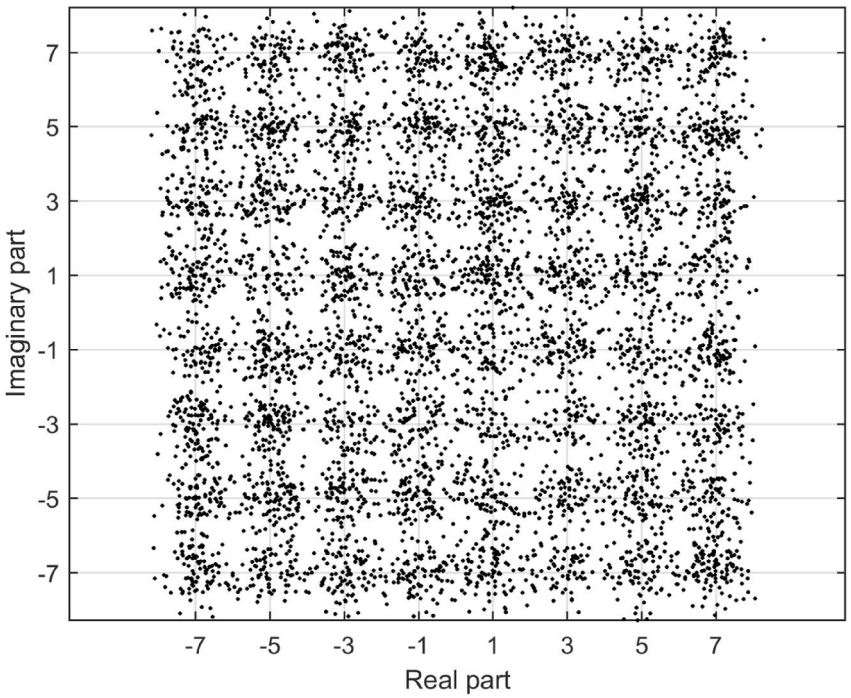

FIGURE 1 An example of a simulated uniform 64-QAM constellation diagram for a Digital Video Broadcasting Terrestrial (DVB-T) symbol containing 8192 complex samples in the time domain, where about 1% of the samples (4 data gaps of 20 samples each with a periodicity of 2048 samples) have been set to zero and no noise has been added. As can be seen, carriers deviate from the ideal 8 × 8 points square grid constellation.

2.2 | Demodulating the Digital Video Broadcasting Terrestrial signal

The standard [1] describes the signal processing on the transmitter side in detail, while "the processing at the receiver side is left open to different implementation solutions". The receiver processing is typically performed in several steps, in reverse order to the modulation steps described in the standard. Below, we follow the initial demodulation steps as described in ref. [3].

First, the time for the start of a received symbol and the frequency of the receiver need to be synchronised. In brief, time synchronisation is used for preventing inter-symbol interference and frequency synchronisation is used for preventing inter-carrier interference. Since a guard interval, consisting of the last part of the useful part of the symbol is inserted before the useful part, the symbol start time can be synchronised by correlation [13]. A piece of the signal (of length equal to the guard interval) is correlated with another piece of the signal separated by the length of the useful part. The frequency offset can be divided into an integer part and a fractional part. The fractional frequency offset can be found as a maximum likelihood estimate from using the phase from the same correlation [13]. The integer frequency offset can be found as a maximum likelihood estimate by calculating the discrete Fourier transform of the useful part of the symbol and correlating the complex coefficients of the continual pilot carriers of two adjacent symbols. After having compensated for both fractional and integer frequency offset, a similar correlation is used to find the symbol's scattered pilot pattern (out of four possible patterns). The channel is estimated in the frequency domain by using the continual and scattered pilot carriers as well as the TPS carriers in a least square channel estimation or a minimum mean square error estimation [3].

2.3 | Digital Video Broadcasting Terrestrial signal used in passive radar

Utilising DVB-T for passive radar has advantages like high transmitted power, large coverage areas, wide bandwidth (for resolution) and the possibility to use the structure in the digital signal to reconstruct the reference signal.

When used in passive radar signal processing, targets are detected by cross-correlating a reference signal with a surveillance signal. It is common to have an antenna and a receiver dedicated for the reference signal but some systems (like LORA) has no receiver dedicated for the reference signal but the measured signal instead contains both the direct path signal and the surveillance signal.

The radar ambiguity function for a DVB-T signal has ambiguities in both time (range) and frequency (Doppler) that depend on the structure of the DVB-T signal [7, 14, 15]. These ambiguities could mask actual targets or introduce false targets. The continual and scattered pilots and the guard interval give rise to both intra-symbol (from boosted power of pilots) and inter-symbol ambiguities.

The strength of the ambiguities can be decreased by correlating with a mismatched signal constructed by manipulating the received signal [7, 14, 15]. The mismatched signal could be constructed from the demodulated signal [14] or without demodulating the signal [7, 15], but then a dedicated receiver for the reference signal was used.

In order to suppress the guard interval ambiguities, a mismatched signal is constructed by setting the samples in the guard interval to zero in the received reference signal [7]. The intra-symbol peaks arising from the pilots are suppressed by correlating with a mismatched signal where the pilot carrier amplitudes have been multiplied by 9/16, while the inter-

symbol peaks arising from the pilots are suppressed by correlating with a mismatched signal where the pilot carrier amplitudes have been set to zero [7]. These two different mismatched signals can be combined. The demodulated signal is also used to estimate and take into account the relative drift between the local oscillators in the transmitter and in the receiver [14]. In the same reference, the Signal-to-Noise Ratio (SNR) of the reference signal is assumed to be sufficient so that decoding using error correction is not needed. Additionally, the demodulated signal could be used to suppress the direct path interference in the surveillance signal [14].

2.4 | The radar system LORA

LORA [10, 11, 16, 17] has mostly been used for SAR applications and each receiver channel samples a 10 MHz signal with a sampling rate of 25.6 MHz real samples. If a DVB-T 8 MHz channel and the 8K transmission mode signal is sampled with this sampling rate, then the signal will be 40% oversampled compared to the elementary time period [1]. The length of five OFDM symbols with or without guard interval corresponds to an integer number of sampling periods in 8K mode.

The LORA system was originally designed for pulsed radar operation but it registers an almost continuous signal in passive mode except that a minimum of 1% of the measured samples are missing since data is registered in data packets. Every data packet contains some descriptive information (header) together with the measured samples and LORA was designed so that samples cannot be registered at the same time as the header information is registered but simultaneous samples are thrown away.

In the LORA system, in standard passive mode, exactly 40 consecutive samples are missing every 4096 samples, otherwise the sampling rate is constant. Hence, the duration of the data gaps is $\frac{40}{25.6}$ μs and the maximum duration of contiguous samples is $\frac{4056}{25.6}$ μs.

2.5 | Recovery of missing samples

Recovery of missing samples in a signal where the signal is sampled uniformly (in time) but where some samples are missing can be seen as a special case of non-uniform sampling [5] and sparse signal processing [18]. The recovery of missing samples from an arbitrary signal is impossible [5], but the signal often satisfies known constraints that can make recovery possible. One constraint that can be used is that the signal is band-limited. The DVB-T signal is band-limited and this constraint is therefore used in this paper, as well as additional constraints as will be described in the following.

The missing samples can be calculated if the density of the known samples is above a minimum density (related to the Nyquist-Landau density [5]), although the stability of the method critically depends on the distribution of the missing samples. In general, for low-pass and high-pass signals, equidistantly distributed (scattered) missing samples lead to improved stability while contiguous missing samples (large gaps) lead to degraded stability. The proposed recovery method iterates one or a few times over the method of least squares (which itself is non-iterative) to solve the non-linear problem.

The magnitude of the error in the output of one iteration can be estimated by using the condition number of the numerical method. The condition number measures the maximal amount that a small relative numerical error in the input data is magnified by the method [19]. The condition number for the Euclidean norm (which we use throughout this paper) for a matrix is the quotient between the largest and smallest singular values of the matrix. The condition number of particularly the matrix inverse in the method of least squares can become very large so we direct attention to it.

3 | MODEL OF THE REGISTERED SIGNAL

Let $\boldsymbol{S^A}$ be the complex amplitudes of the carriers in a DVB-T symbol (and where the rest of the Fourier coefficients are set to zero) in the transmitter. Let $\boldsymbol{S^A}$ be converted to a continuous time signal, $s^A(t)$, after applying the inverse discrete Fourier transform and adding the guard interval. If $s^A(t)$ pass through a linear time-invariant system before it arrives at the receiver, the signal that arrives at the receiver can be modelled as

$$\left(s^A * h\right)(t) + u(t), \tag{1}$$

where $*$ denotes convolution, $h(t)$ is the channel impulse response and $u(t)$ is an unknown signal which is possibly present outside the DVB-T frequency band. After sampling in the receiver and throwing away the samples that will be missing, the registered signal can be modelled

$$s_R[t] = m[t](s_H[t] + u[t] + n[t]), \tag{2}$$

where $s_H[t] = (s^A * h)[t]$, and where $n[t]$ is the thermal noise in the receiver and $m[t]$ denotes the missing samples modelled as a signal that is equal to zero where the corresponding samples are missing, elsewhere $m[t]$ is equal to one.

If we restrict the time to the useful part of one symbol, (duration T_U), and assume that there is no inter-symbol interference, the corresponding samples of $s_H[t]$, $s_R[t]$, $u[t]$ and $n[t]$ can be written as vectors $\mathbf{s_H}$, $\mathbf{s_R}$, $\mathbf{u}$ and $\mathbf{n}$ respectively, where the lengths of the vectors equals the number of samples in the useful part of one symbol. Then (2) can be written as

$$\mathbf{s_R} = \mathbf{M}(\mathbf{s_H} + \mathbf{u} + \mathbf{n}) \tag{3}$$

where $\mathbf{M}$ is a diagonal matrix with elements either zero or one so that it inserts zeros in the positions of the missing samples.

For the special case where the channel impulse response $h(t)$ equals the unit impulse (delta function) we have $\boldsymbol{S^A} = \mathbf{Gs_H}$, where $\mathbf{G}$ equals the rows of the Fourier matrix that correspond to the carriers of the DVB-T band. Let $\mathbf{P}$ be a matrix with elements either equal to zero or one that selects the carriers from the vector of frequency components, so that

$$\boldsymbol{S^A} = \mathbf{Gs_H} = \mathbf{PFs_H}, \tag{4}$$

where $\mathbf{F}$ is the Fourier matrix. The complex amplitudes of the carriers in the frequency domain for the registered signal can then be written as a circular convolution

$$\begin{aligned}\mathbf{Gs_R} &= \mathbf{PFM}(\mathbf{s_H} + \mathbf{u} + \mathbf{n}) \\ &= \mathbf{P}\left(\frac{1}{N_s}(\mathbf{Fm}) * \left(\boldsymbol{S^A} + \mathbf{Fu} + \mathbf{Fn}\right)\right),\end{aligned} \tag{5}$$

where $\mathbf{m}$ is a vector consisting of the samples of $m[t]$ during one symbol and where N_s is the number of samples in one symbol. It is noted that the missing samples give a convolution in the frequency domain with the complex amplitudes of the carriers, the Fourier transform of the noise and the Fourier transform of unknown signals possibly present in the received data. The convolution spreads the signals' (and carriers') energies outside the Fourier coefficients of the frequency bands in which the signals were transmitted.

4 | PROPOSED METHOD

Since the DVB-T signal is band-limited, values for the missing samples can be estimated by optimisation so that the Fourier coefficients for frequencies outside the signal's band are minimised [5]. An objective function was chosen consisting of the sum of the squared absolute values of the complex Fourier coefficients for every frequency outside the frequency band within the sampled baseband signal, as these frequencies should only contain noise. In the same way as the missing samples make the DVB-T signal spread to Fourier coefficients outside the DVB-T band, the missing samples will make another strong signal (possibly present in the sampled data) in a different frequency band spread to Fourier coefficients outside its band. Therefore, only Fourier coefficients outside the frequency bands nominally occupied by other, possible, strong signals present in the sampled data were considered in the objective function. The method of least squares was used for the minimisation.

The bandwidth constraint was not sufficient to arrive at a stable solution, so in order to increase the accuracy (or to decrease the condition number w.r.t. matrix inversion of the matrix of the system of linear equations), additional information about the signal was used which resulted in more constraints and thus, extra equations. This is explained in the next paragraph.

We define the circumscribing square as the smallest square that circumscribes the ideal QAM constellation grid (pilot and TPS carriers excluded) with sides parallel with the real and imaginary axes and centred on the origin. We also define the concentric square as a second square of chosen size that is concentric to the circumscribing square (Figure 2). For data carriers whose complex amplitudes fall outside the concentric square, we let the objective function include their shortest distances (along the real or imaginary axes) to the circumscribing square. To the objective function, we thus add the squared distances (along the real or imaginary axes) to the circumscribing square for all data carrier points outside the concentric square. In this way, the circumscribing square indicates desired value(s) of the real or imaginary part of data carriers outside the concentric square.

The objective function is thereby defined to consist of a weighted sum of two different parts. The first part is the sum of the squared absolute values of the complex amplitude for every frequency outside the DVB-T band. The second part is the sum of the squared distances (along the real or imaginary axes) to the circumscribing square for all data carrier points that are outside the concentric square.

4.1 | Definition of the optimisation problem

In this section, first we derive the mathematical formulas for the objective function described in the previous two paragraphs and arrive in the norm in expression (13) below. Then, we derive the formulas for solving the optimisation problem in (13) with the method of least squares in (17) below. We show an overview of the proposed algorithm in Figure 3.

Let $\mathbf{s}$ be the (real- or complex-valued) sample column vector during one DVB-T symbol, ordered in time order with known, measured samples and with zeros representing the missing samples in the positions of the missing samples.

Let N_s be the number of samples in one DVB-T symbol and let N_f be the number of frequency components in the channel estimation. Let $\mathbf{G}_k$ be the $N_f \times N_s$ time-samples-to-carrier-amplitudes function matrix in iteration k+1 that transforms time samples of the sampled signal to DVB-T carriers and to other frequency components outside the DVB-T spectrum in the frequency domain. $\mathbf{G}_k$ ideally consists of only the Fourier transform and does not need to be updated from iteration to iteration. If the time-samples-to-carrier-amplitudes function is non-linear w.r.t. the time samples (e.g. if the least-squares channel estimator is used [20]), then $\mathbf{G}_k$ should be a linearised version of the time-samples-to-carrier-amplitudes function.

Let N_m be the number of missing samples in one DVB-T symbol. Let $\mathbf{x}_k$ be a column vector of length N_m of the unknown, missing samples to be estimated in iteration k. Choose the starting input value in the first iteration to be

$$\mathbf{x}_0 = \overline{0}, \tag{6}$$

where $\overline{0}$ is a column vector of appropriate length containing only zeros.

Let $\mathbf{P^M}$ (M is a superscript here) be a $N_s \times N_m$ matrix that inserts zeros in the positions of the known samples of $\mathbf{s}$ and

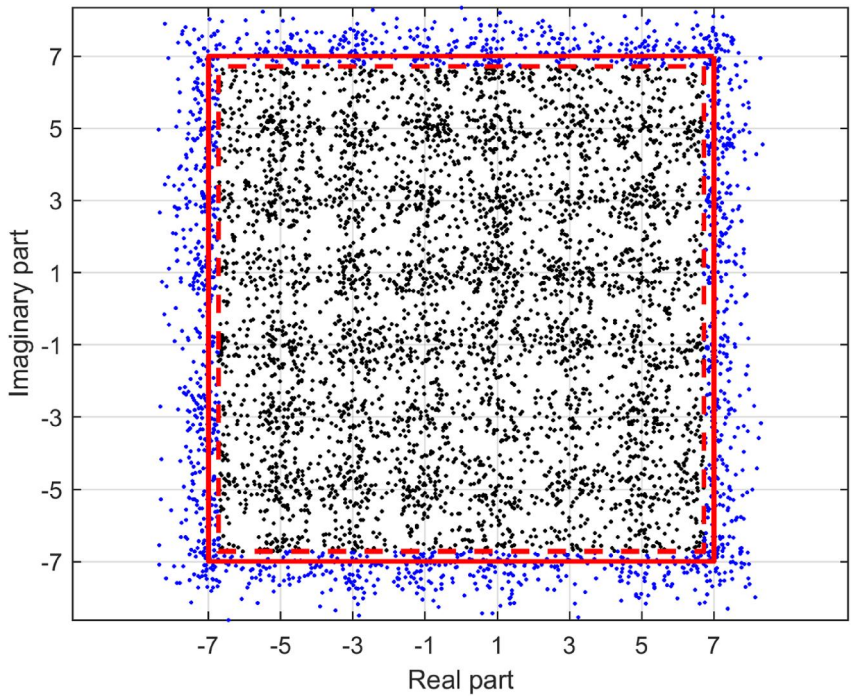

FIGURE 2 Simulated constellation diagram for a symbol containing 8192 complex samples with a Signal-to-Noise Ratio (SNR) of 25 dB and 80 samples (4 data gaps of 20 samples each) missing and set to zero. The circumscribing square is drawn with solid red line and the concentric square is drawn with dashed red line.

that inserts the samples of $\mathbf{x}_k$ into the positions of the missing samples of $\mathbf{s}$. Let the position indices of the missing samples in $\mathbf{s}$ be denoted p_μ^M for $\mu = 1, 2, \ldots, N_m$. The components of $\mathbf{P^M}$ are then given by

$$\left(\mathbf{P^M}\right)_{\tau,\mu} = \begin{cases} 1 & if\ \tau = p_\mu^M \\ 0 & else \end{cases}, \tag{7}$$

with matrix row index $\tau = 1, 2, \ldots, N_s$. Let the vector of DVB-T carriers and the other frequency components be denoted $\boldsymbol{S}_k$. Then

$$\boldsymbol{S}_k = \mathbf{G}_k\mathbf{s} + \mathbf{G}_k\mathbf{P^M}\mathbf{x}_k. \tag{8}$$

Let N_d be the number of data carriers (i.e. not pilot or TPS carriers) in the DVB-T symbol and let $\mathbf{P^B}$ be a $N_d \times N_f$ matrix that selects the data carriers from the vector of frequency components $\boldsymbol{S}_k$. Let the position indices of the data carriers in $\boldsymbol{S}_k$ be denoted p_δ^B for $\delta = 1, 2, \ldots, N_d$, and let $\varphi = 1, 2, \ldots, N_f$. The components of $\mathbf{P^B}$ are then given by

$$\left(\mathbf{P^B}\right)_{\delta,\varphi} = \begin{cases} 1 & if\ p_\delta^B = \varphi \\ 0 & else \end{cases}, \tag{9}$$

and the data carriers can be expressed $\mathbf{P^B}\boldsymbol{S}_k$. Let d denote half of the length of the side of the circumscribing square and let $\mathbf{D}_k^\mathbf{R}$ and $\mathbf{D}_k^\mathbf{I}$ denote two diagonal scaling $N_d \times N_d$ matrices in iteration k whose elements are all real and non-negative. The matrix elements in these matrices define the weights of the equations for the real and imaginary part, respectively, of each DVB-T data carrier point. The probability that a data carrier belongs to a constellation point along the left- or the right-hand side of the circumscribing square (i.e. the absolute value of the carrier's real part is equal to d) determines the weight in $\mathbf{D}_k^\mathbf{R}$ for the associated equation. In the same way, the probability that a data carrier belongs to a constellation point along the upper or the lower side of the circumscribing square (i.e. the absolute value of the carrier's imaginary part is equal to d) determines the weight in $\mathbf{D}_k^\mathbf{I}$ for the associated equation.

We used the weight 1 for equations associated with the real (or imaginary) part of a data carrier if the absolute value of the real part (or imaginary part, respectively) was larger than half of the length of the side of the chosen concentric square, otherwise we used the weight 0. The components of $\mathbf{D}_k^\mathbf{R}$ can be written

$$\left(\mathbf{D}_k^\mathbf{R}\right)_{\delta,\Delta} = \begin{cases} 1 & if\ \delta = \Delta\ and\ |(Re\{\mathbf{P^B}\boldsymbol{S}_k\})_\delta| \geq d \cdot r \\ 0 & else \end{cases}, \tag{10}$$

where r is a chosen real positive constant that determines the side of the concentric square. The components of $\mathbf{D}_k^\mathbf{I}$ can be written

$$\left(\mathbf{D}_k^\mathbf{I}\right)_{\delta,\Delta} = \begin{cases} 1 & if\ \delta = \Delta\ and\ |(Im\{\mathbf{P^B}\boldsymbol{S}_k\})_\delta| \geq d \cdot r \\ 0 & else \end{cases}. \tag{11}$$

As an option, equations given the weight 0 may be excluded from the matrices since the result does not depend on them.

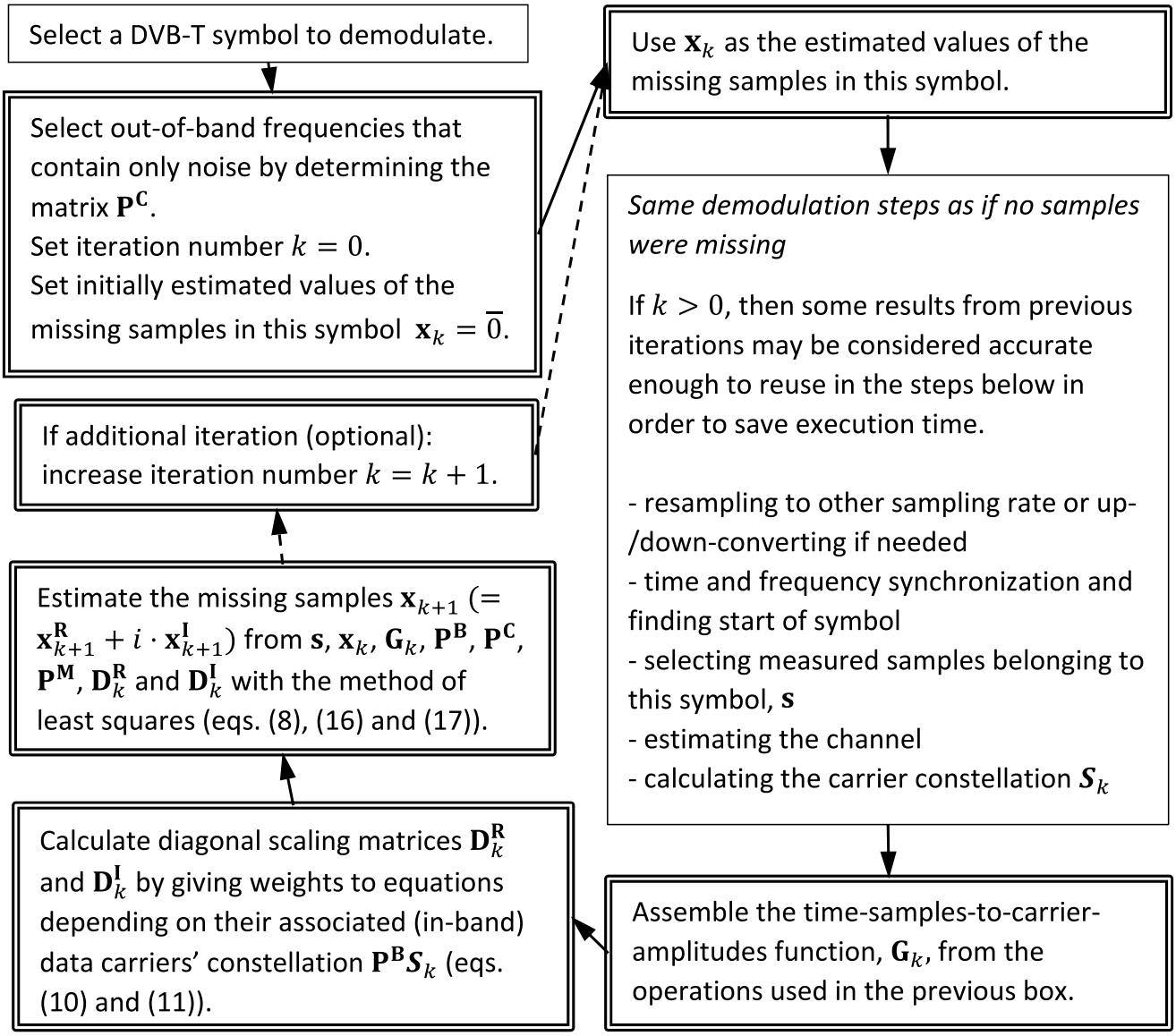

FIGURE 3 Overview of the proposed algorithm. Boxes with double lines indicate the steps introduced to estimate the missing samples. One or a few iterations may be sufficient.

Let N_o be the number of frequency components in $\boldsymbol{S}_k$ that should only contain noise and let $\mathbf{P^C}$ denote a $N_o \times N_f$ matrix whose elements are all real and positive. The elements in this matrix are used to give weights to the equations for the complex coefficients for frequencies outside the DVB-T frequency band in the sampled signal. We used the same weight 1 in all iterations for equations associated with frequencies that should only contain noise. $\mathbf{P^C}$ thereby weights these equations associated with frequencies outside the DVB-T frequency band against the equations associated with carriers inside the frequency band (with weights in $\mathbf{D}^{\mathbf{R}}_k$ and $\mathbf{D}^{\mathbf{I}}_k$). Let the position indices of the frequency components in $\boldsymbol{S}_k$ that should only contain noise be denoted p^C_ω for $\omega = 1, 2, \ldots, N_o$. The components of $\mathbf{P^C}$ are then given by

$$\left(\mathbf{P^C}\right)_{\omega,\varphi} = \begin{cases} 1 & \text{if } p^C_\omega = \varphi \\ 0 & \text{else} \end{cases}, \tag{12}$$

and the frequency components that should only contain noise can be expressed $\mathbf{P^C}\boldsymbol{S}_k$.

The missing samples can be estimated by minimising the objective function consisting of the square of the Euclidean norm of the estimated error in the selected frequency components in (13) where the *sign* function operates on each component of the vector in the argument. An output vector component of the *sign* function equals −1 if the value of the corresponding input vector component is negative and it equals 1 if the value of the component is positive. If the value of the corresponding input vector component is exactly 0, then

$$\min_{\mathbf{x}_{k+1}} \left\| d \cdot \begin{pmatrix} \mathbf{D}_k^{\mathbf{R}} sign\left(Re\{\mathbf{P}^{\mathbf{B}}\boldsymbol{S}_k\}\right) \\ \mathbf{D}_k^{\mathbf{I}} sign\left(Im\{\mathbf{P}^{\mathbf{B}}\boldsymbol{S}_k\}\right) \\ \overline{0} \\ \overline{0} \end{pmatrix} - \begin{pmatrix} Re\{\mathbf{D}_k^{\mathbf{R}}\mathbf{P}^{\mathbf{B}}\mathbf{G}_k\mathbf{s}\} \\ Im\{\mathbf{D}_k^{\mathbf{I}}\mathbf{P}^{\mathbf{B}}\mathbf{G}_k\mathbf{s}\} \\ Re\{\mathbf{P}^{\mathbf{C}}\mathbf{G}_k\mathbf{s}\} \\ Im\{\mathbf{P}^{\mathbf{C}}\mathbf{G}_k\mathbf{s}\} \end{pmatrix} - \begin{pmatrix} Re\{\mathbf{D}_k^{\mathbf{R}}\mathbf{P}^{\mathbf{B}}\mathbf{G}_k\mathbf{P}^{\mathbf{M}}\mathbf{x}_{k+1}\} \\ Im\{\mathbf{D}_k^{\mathbf{I}}\mathbf{P}^{\mathbf{B}}\mathbf{G}_k\mathbf{P}^{\mathbf{M}}\mathbf{x}_{k+1}\} \\ Re\{\mathbf{P}^{\mathbf{C}}\mathbf{G}_k\mathbf{P}^{\mathbf{M}}\mathbf{x}_{k+1}\} \\ Im\{\mathbf{P}^{\mathbf{C}}\mathbf{G}_k\mathbf{P}^{\mathbf{M}}\mathbf{x}_{k+1}\} \end{pmatrix} \right\|_2^2, \tag{13}$$

the corresponding weight in the diagonal matrix $\mathbf{D}_k^{\mathbf{R}}$ or $\mathbf{D}_k^{\mathbf{I}}$ will be set to zero anyway and the value of the *sign* function does not matter. By defining

$$\mathbf{x}_k^{\mathbf{R}} = Re\{\mathbf{x}_k\} \tag{14}$$

and

$$\mathbf{x}_k^{\mathbf{I}} = Im\{\mathbf{x}_k\} \tag{15}$$

for $k \geq 0$, the last term in the norm in expression (13) can be written $\mathbf{K}_k \begin{pmatrix} \mathbf{x}_{k+1}^{\mathbf{R}} \\ \mathbf{x}_{k+1}^{\mathbf{I}} \end{pmatrix}$ where

$$\mathbf{K}_k = \begin{pmatrix} Re\{\mathbf{D}_k^{\mathbf{R}}\mathbf{P}^{\mathbf{B}}\mathbf{G}_k\mathbf{P}^{\mathbf{M}}\} & -Im\{\mathbf{D}_k^{\mathbf{R}}\mathbf{P}^{\mathbf{B}}\mathbf{G}_k\mathbf{P}^{\mathbf{M}}\} \\ Im\{\mathbf{D}_k^{\mathbf{I}}\mathbf{P}^{\mathbf{B}}\mathbf{G}_k\mathbf{P}^{\mathbf{M}}\} & Re\{\mathbf{D}_k^{\mathbf{I}}\mathbf{P}^{\mathbf{B}}\mathbf{G}_k\mathbf{P}^{\mathbf{M}}\} \\ Re\{\mathbf{P}^{\mathbf{C}}\mathbf{G}_k\mathbf{P}^{\mathbf{M}}\} & -Im\{\mathbf{P}^{\mathbf{C}}\mathbf{G}_k\mathbf{P}^{\mathbf{M}}\} \\ Im\{\mathbf{P}^{\mathbf{C}}\mathbf{G}_k\mathbf{P}^{\mathbf{M}}\} & Re\{\mathbf{P}^{\mathbf{C}}\mathbf{G}_k\mathbf{P}^{\mathbf{M}}\} \end{pmatrix}. \tag{16}$$

Then $\begin{pmatrix} \mathbf{x}_{k+1}^{\mathbf{R}} \\ \mathbf{x}_{k+1}^{\mathbf{I}} \end{pmatrix}$ (and $\mathbf{x}_{k+1} = \mathbf{x}_{k+1}^{\mathbf{R}} + i \cdot \mathbf{x}_{k+1}^{\mathbf{I}}$ from Equations (14) and (15)) can be calculated with the method of least squares

$$\begin{pmatrix} \mathbf{x}_{k+1}^{\mathbf{R}} \\ \mathbf{x}_{k+1}^{\mathbf{I}} \end{pmatrix} = \left(\mathbf{K}_k^T\mathbf{K}_k\right)^{-1}\mathbf{K}_k^T \left(d \cdot \begin{pmatrix} \mathbf{D}_k^{\mathbf{R}} sign\left(Re\{\mathbf{P}^{\mathbf{B}}\boldsymbol{S}_k\}\right) \\ \mathbf{D}_k^{\mathbf{I}} sign\left(Im\{\mathbf{P}^{\mathbf{B}}\boldsymbol{S}_k\}\right) \\ \overline{0} \\ \overline{0} \end{pmatrix} - \begin{pmatrix} Re\{\mathbf{D}_k^{\mathbf{R}}\mathbf{P}^{\mathbf{B}}\mathbf{G}_k\mathbf{s}\} \\ Im\{\mathbf{D}_k^{\mathbf{I}}\mathbf{P}^{\mathbf{B}}\mathbf{G}_k\mathbf{s}\} \\ Re\{\mathbf{P}^{\mathbf{C}}\mathbf{G}_k\mathbf{s}\} \\ Im\{\mathbf{P}^{\mathbf{C}}\mathbf{G}_k\mathbf{s}\} \end{pmatrix} \right), \tag{17}$$

where the superscript T indicates transpose operation. The result from iteration $k+1$, that is, $\begin{pmatrix} \mathbf{x}_{k+1}^{\mathbf{R}} \\ \mathbf{x}_{k+1}^{\mathbf{I}} \end{pmatrix}$, depends on the result from previous iteration $\begin{pmatrix} \mathbf{x}_k^{\mathbf{R}} \\ \mathbf{x}_k^{\mathbf{I}} \end{pmatrix}$ through equations (14), (15), (8), (10) and (11). An overview of the proposed algorithm is shown in Figure 3.

For demodulation of real, measured signal with non-linear synchronisation and channel estimation, we carried out a second iteration by performing these operations once again but this time using the estimated values of the missing samples from the first iteration. See Figure 3.

4.2 | Ratio of incorrect extra equations

In this section, we derive expressions for a channel with only white Gaussian noise and no missing samples. In such a channel, there is a probability that an extra equation associates a data carrier point to an edge of the circumscribing square to which the data carrier's ideal constellation point does not belong, that is, the extra equation is incorrect. Below, we derive expressions for the expected ratio of incorrect equations. From detection theory (e.g. equation 6.12 in ref. [21]), the noise probability density function at the output of the matched filter for a received pulse can be written

$$f(n_R, n_I) = \frac{1}{2\pi\sigma^2} e^{-\left(n_R^2+n_I^2\right)/2\sigma^2}, \tag{18}$$

where σ^2 is the noise variance and n_R and n_I are random variables for the noise.

Assume that the radar pulse consists of one single OFDM carrier (i.e. a sinus signal) synchronised in time so that it corresponds to the useful part of one OFDM symbol. Assume a 64-QAM constellation is used and that all 64 points in the constellation are equally probable. By averaging the noise probability density function over the constellation points, the probability density for a complex amplitude of $Z_R + iZ_I$ at the output of the matched filter can be written

$$f_{QAM}(Z_R, Z_I, \sigma, k_0) = \frac{1}{64}\frac{1}{2\pi\sigma^2}\sum_{k=k_0}^{3} e^{-(Z_R-(k+1/2)\gamma)^2/2\sigma^2}\sum_{l=-4}^{3} e^{-(Z_I-(l+1/2)\gamma)^2/2\sigma^2} \tag{19}$$

with $k_0 = -4$, where γ is the QAM grid spacing at the output of the matched filter. For extra equations corresponding to the negative real part of data carriers, the expected ratio of incorrect equations, ρ_{R-}, can be written

$$\rho_{R-}(Z_R, Z_I, \sigma) = \frac{f_{QAM}(Z_R, Z_I, \sigma, -3)}{f_{QAM}(Z_R, Z_I, \sigma, -4)} = \frac{\sum_{k=-3}^{3} e^{-(Z_R-(k+1/2)\gamma)^2/2\sigma^2}}{\sum_{k=-4}^{3} e^{-(Z_R-(k+1/2)\gamma)^2/2\sigma^2}} \tag{20}$$

for a complex amplitude of $Z_R + iZ_I$ at the output of the matched filter. As is visible, the ratio of incorrect equations, ρ_{R-}, is independent of the imaginary part Z_I and of the carrier frequency. This means that the ratio of incorrect equations is constant along one side (and in its extension) of the concentric square. As can be shown with differentiation w.r.t. Z_R, ρ_{R-} increases along the real axis. It is visible that ρ_{R-} only depends on Z_R, on the noise variance σ^2 (which is inversely proportional to the SNR) and on γ that depends on the signal without noise at the output of the matched filter.

The ratios of incorrect equations for positive real part and negative and positive imaginary part can be calculated similarly. The same calculations can be carried out for other QAM-constellation schemes, for example, 16-QAM, by changing the summations. The shape of the concentric square (with straight edges parallel to the axes) is optimal in the sense that the expected ratio of incorrect extra equations is constant along its edges which thus allows a certain expected ratio to be specified and guaranteed along the edges. The ratio decreases if the concentric square is made larger and increases if the concentric square is made smaller.

5 | RESULTS FROM SIMULATIONS

Several simulations were performed where the proposed algorithm was applied to simulated complex DVB-T signals with varying number of missing samples and with varying noise. In the time domain, the useful part of each simulated symbol consisted of 8192 complex samples, corresponding to the 8K mode. A variable amount of white Gaussian noise was added to the complex DVB-T signal in order to simulate noise.

Groups of equal number of consecutive missing samples were evenly spaced every 2048 samples (in order to imitate the registration in the LORA system) in the DVB-T signal. This means that every symbol contained the same total number of missing samples, distributed among four or five groups. The missing samples were set to zero in the simulations. The guard interval of 1/32 made the position of the first missing sample change from symbol to symbol in the DVB-T signal containing many symbols.

Also, for all simulations, the 8 MHz channel with the 64-QAM constellation were used [1]. A DVB-T symbol has 6817 carriers in this mode and the simulation signal used low-pass complex baseband signal with Fourier frequencies from −3408 to 3408.

In order to simplify the simulations, a channel impulse response equal the unit impulse (delta function) was assumed (i.e. no multipath effects) and white Gaussian noise was added to the signal. When demodulating the signal, the transfer function matrix for the channel was assumed to be the identity matrix, so there was no need for estimating the channel and the pilot and TPS carriers were not used. In addition, the symbol start was known and no time synchronisation was needed. The guard interval was not used for the demodulation. However, in the simulation of a recovered signal in a passive radar application with correlations (described in section 5.4 below), the guard interval was included.

We varied the number of data carriers associated with extra equations (the equations that were given non-zero weights in either of the two diagonal scaling matrices in $\mathbf{D}_k^{\mathbf{R}}$ and $\mathbf{D}_k^{\mathbf{I}}$) in the simulations and studied how it affects the condition number of the system of equations to be solved and also how it is related to the side length of the concentric square. Results after one iteration and after two iterations of recovery are shown in sections 5.2 and 5.3 respectively. In section 5.4, it is shown how the proposed algorithm improves the BER. Results from a simulation of cross-correlations in a radar are shown in section 5.5. In section 6, the result from applying the algorithm to a real measured signal is shown.

5.1 | Before recovery

The condition number (w.r.t. matrix inversion of the system of linear equations) was improved several orders of magnitude when equations for a large number of data carrier points were used in the objective function, see Figure 4. The condition number is dependent on how many carrier points are used for

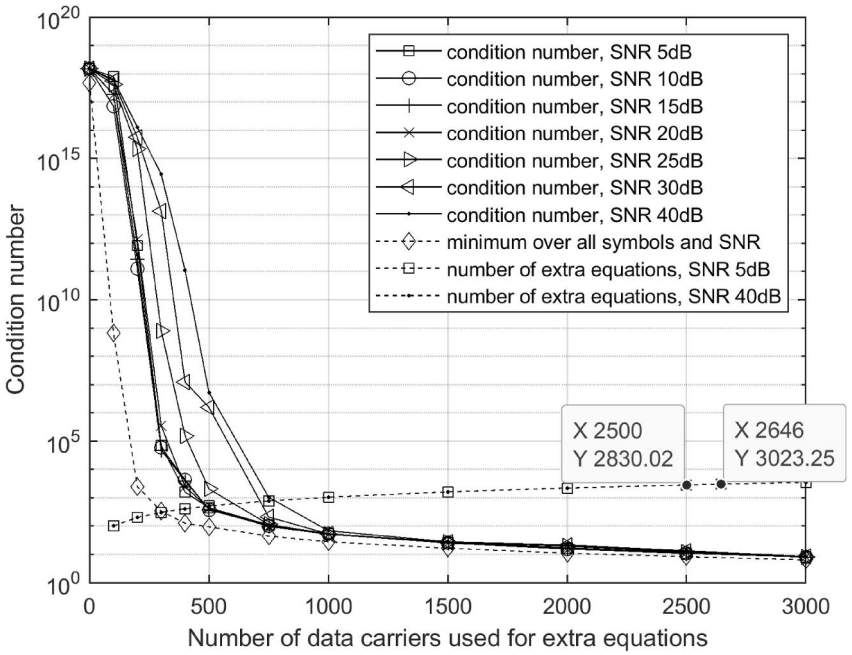

FIGURE 4 The condition number is dependent on how many carrier points are used for extra equations. The condition number was calculated with MATLAB's function "cond". The minimum of the condition number taken over all symbols and over all different values of Signal-to-Noise Ratio (SNR) in the figure is also shown. The number of extra equations are shown for SNRs of 5 and 40 dB and these two curves almost coincide. The unit on the y-axis for these two curves is "number".

extra equations in addition to the 2750 (=2·8192–2·6817) real-valued equations derived from the band limitation property of the complex-valued signal. A low condition number decreases the error in the estimation of the missing samples. The error in the constellation depends also on the portion of missing samples in the symbol. A signal with 800 symbols (which corresponds to a time signal of at least 0.73 s with the shortest guard interval of 1/32) and 80 missing samples (out of 8192) per symbol was used in the simulations.

The condition number varies between symbols depending on the signal and the distribution of the selected carriers (in addition to the distribution of missing samples).

For QAM constellation with square $a \times a$ grid, the number of extra real-valued equations per data carrier is about $4a/a^2 = 4/a$ in a single symbol on average if the data carrier coefficients are evenly distributed over all constellation points. There are 6048 data carriers in the 8K mode [1], so the number of extra equations is about 3024 (i.e. 6048/2) on average for 64-QAM.

In Figure 5, the relationship between the size of the concentric square and the number of data carriers outside it, is shown. For each of these simulations (and other simulations with number of data carriers used for extra equations on the x-axis), the data carriers were first sorted with respect to the largest absolute value of the real and imaginary parts (as this relationship changes from symbol to symbol). The square side ratio was then chosen to give the number of data carriers (used for extra equations) on the x-axis. Note that the sorting is not part of the proposed method and was used only to make these diagrams. Figure 5 was used to determine the size of the concentric square for successive simulations.

In the same figure, the ratio of incorrect extra equations is shown. The ratio of incorrect extra equations was calculated as the mean number (taken over all symbols) of incorrect extra equations divided by the mean number of extra equations. In general, errors in the estimated samples will increase if incorrect extra equations are included in the objective function as the desired values are for the outer points (i.e. data carrier points that belong to ideal constellation points lying on the circumscribing square).

5.2 | One iteration of recovery

Below are the results from simulations where a signal with 80 missing samples (out of 8192) per symbol was simulated over 800 symbols with the proposed method of recovery.

The data carriers' coefficients in the recovered signal were mapped to ideal constellation points by rounding the coefficient to the nearest constellation point. The data carriers that were mapped to the (decision region of) wrong constellation point were counted for each symbol and are shown as "number in wrong bin" in Figure 6. Also, the Root Mean Square (RMS) errors in the data carriers' coefficients (i.e. deviations in the constellation diagram) are shown in the same figure.

It is observed that the number of data carriers in wrong bin and the RMS error in the carriers' coefficients decreased when the number of data carriers used for extra equations was increased from zero up to about 1500. When the number of data carriers used for extra equations was further increased to 2000 or 2500, the results did not change significantly (except for curves that start from low levels). When this number was increased even further, the number of incorrect extra equations included in the objective function increased (shown in Figure 5) and the results became worse. Thus, it seems that the results are not sensitive to accidently using a small portion of incorrect extra equations among the correct extra equations.

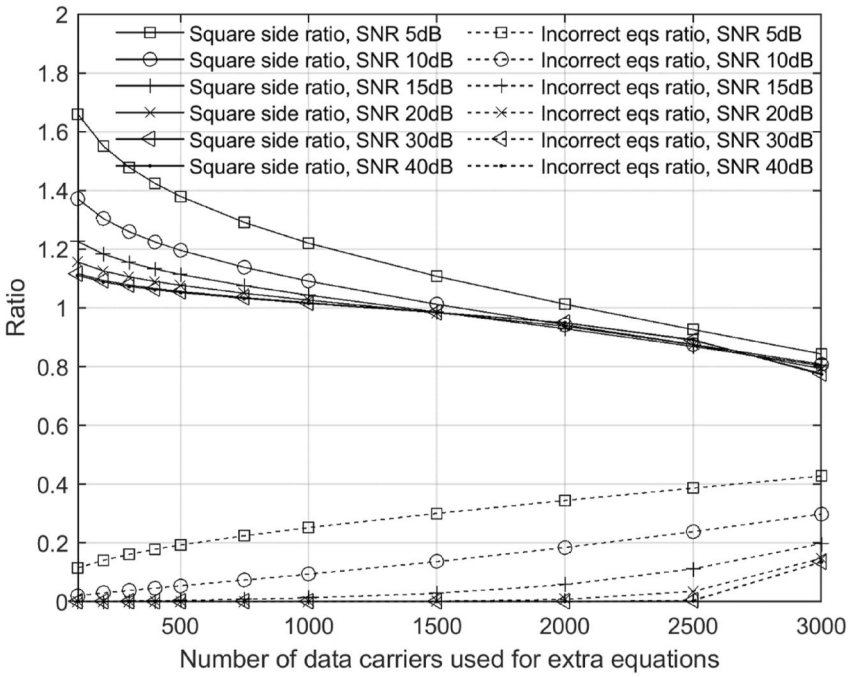

FIGURE 5 The ratio between the lengths of the sides of the second, concentric square to the circumscribing square is plotted against the number of data carriers outside the concentric square. The data carriers outside the concentric square were used for extra equations. In addition, the ratio of incorrect extra equations is shown. Note that the values on the *x*-axis starts from 100.

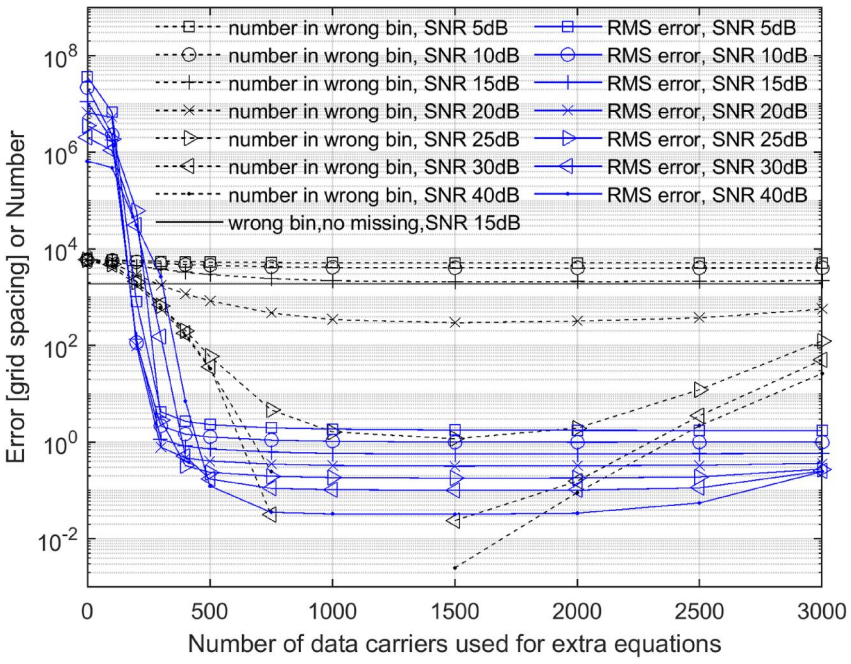

FIGURE 6 The errors in the carriers calculated using the recovered samples after one iteration of the proposed algorithm are plotted against the number of data carriers used for extra equations for different values of Signal-to-Noise Ratio (SNR). Errors are given relative to grid spacing. For 1000 data carriers used for extra equations and SNR equal to 30 and 40 dB respectively, no data carriers were mapped to wrong bin during the 800 simulated symbols so these points are not plotted due to the logarithmic scale.

The errors were very small for a SNR of 30 and 40 dB, but increased when incorrect extra equations were included in the objective function.

For the simulations using the least square method with no data carriers used for extra equations, that is, with only equations for the out-of-band frequencies, the results became much worse than setting the missing samples to zero. This is observed by comparing the number of data carriers in wrong bin and the RMS error in the carriers' coefficients in Figure 6 with the curves "RMS error before recovery" and "ratio in wrong bin before recovery" in Figure 7 below.

In Figure 7, results are shown from one iteration of the proposed algorithm with the concentric square side length set to be the length of the side of the circumscribing square times 0.98. Figures 5 and 6 were used for selecting 0.98 for the ratio between the square side lengths in order to give 1500 to 2000 data carriers used for extra equations (as Figure 6 shows good results for this interval of numbers of data carriers used for extra equations). Except for the simulations with SNR of 5 dB, the actual measured mean number of carriers was between 1540 and 1730.

Also in Figure 7, it is observed that the RMS errors in the data carrier coefficients in the recovered signal were

lower than the errors in the signal before recovery of missing samples. The RMS error curves for the recovered signal form straight lines since the logarithm of the RMS error is inversely proportional to the SNR if only Gaussian noise is present in the signal [22]. It is also observed that the number of data carriers in the wrong bin decreased compared to before the recovery. The errors in the recovered signal were slightly larger than the errors calculated with noise present but with no missing samples and without the use of the proposed algorithm.

For low values of SNR, it seems that the errors from noise dominate over errors from the 80 missing samples (set to zero) and the error values were about the same for the noisy signal whether samples were thrown away (set to zero) or not. For these low SNR values and the 80 missing samples, the proposed method does not improve the result.

The results from a simulation with constant SNR of 25 dB and different number of missing samples are shown in Figure 8 below. It can be seen that the errors increased with the number of missing samples as expected.

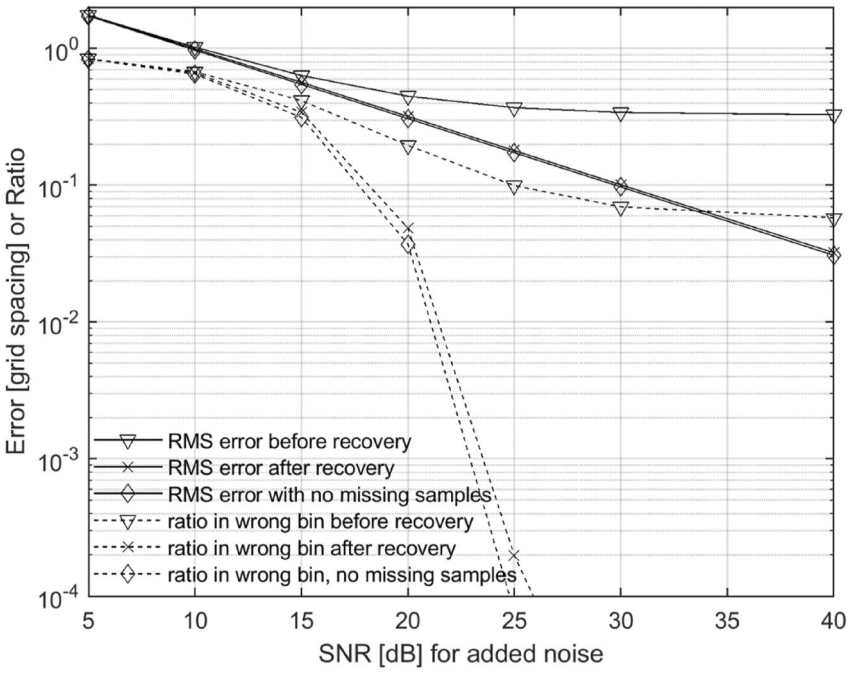

FIGURE 7 The Root Mean Square (RMS) error and ratio of data carriers in wrong bin for a signal with recovered samples (after one iteration) compared to the signal before recovery and before throwing away (setting to zero) samples. Errors are given relative to grid spacing. The concentric square side was 0.98 times the side of the circumscribing square.

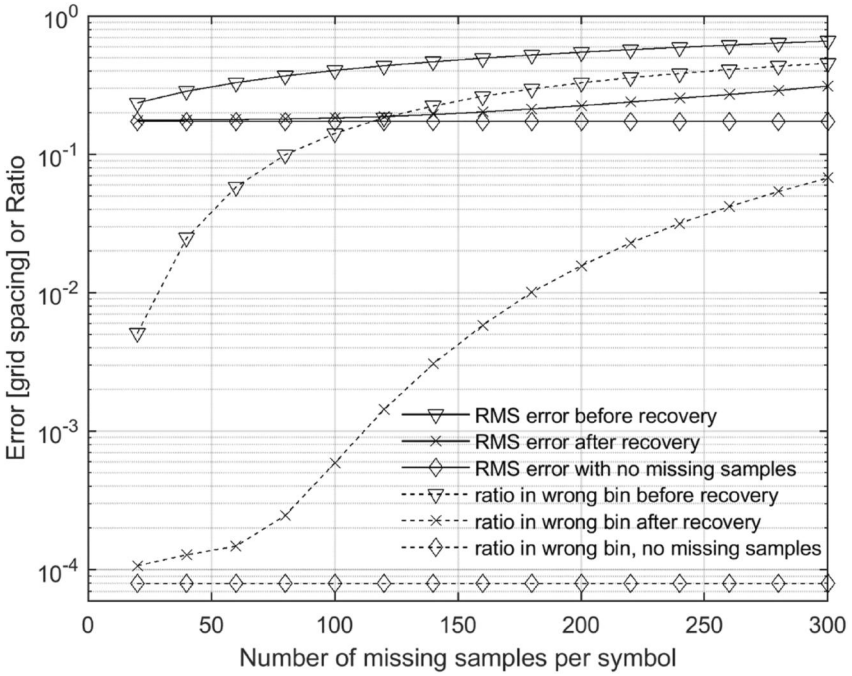

FIGURE 8 Errors after one iteration of the proposed algorithm are plotted against increasing number of missing samples. Errors are given relative to grid spacing. The ratio between the concentric square side length and the circumscribing square side length was set to 0.96 in this simulation.

For example, for 80 missing samples, the RMS error was about 0.37 before recovery, and about 0.18 after recovery, which is close to 0.17 for no missing samples. Also, for the same number of missing samples, the ratio of data carriers that were closer to another constellation grid point than the correct one, was about 0.10 before recovery, and about $2{\cdot}10^{-4}$ after recovery, which is closer to $8{\cdot}10^{-5}$ for no missing samples.

5.3 | Two iterations of recovery

Results from running two iterations of the proposed algorithm for 200 missing samples per symbol are shown in Figure 9 below. In both iterations, extra equations were added for data carrier points outside a square with side 0.96 times the side of the circumscribing square. There was a significant improvement after the first iteration of recovery and also from the first iteration to the second. These improvements were seen in both the RMS error and in the ratio of data carriers that were closer to another constellation grid point than the correct one after recovery. Again, it was shown that the values for the signal with noise but no missing samples were approached for the respective measures. However, for the lowest values of SNR, there was no significant improvement. For example, for an SNR of 25 dB, the RMS error was about 0.55 before recovery, and about 0.19 after two iterations of recovery, which is close to 0.17 for no missing samples. Also, for the same SNR, the ratio of data carriers that were closer to another constellation grid point than the correct one was about 0.33 before recovery, and about 0.002 after two iterations of recovery, which is closer to $8{\cdot}10^{-5}$ for no missing samples.

A contributing factor to the improvement in the second iteration was that the ratio of incorrect extra equations in the extra equations were lower in the second iteration because the first iteration moved the data carriers relative to the concentric square. In fact, since a channel impulse response equal the unit impulse was assumed in these simulations, and no equalisation was performed, the only difference in inputs between the first and the second iteration was the carriers selected for extra equations.

5.4 | Number of iterations required for one error event per hour after decoding

In Figure 10, it is shown how the number of iterations required for achieving a BER less than 10^{-11} depends on the SNR and on the amount of zeroed samples. A BER of about $2{\cdot}10^{-4}$ after convolutional decoding will give a BER of less than 10^{-11} after Reed-Solomon decoding which corresponds to an uncorrected error event rate of less than one per hour [1].

The DVB-T signal in Figure 10 uses 64-QAM and code rate 2/3, has white Gaussian noise added to it and zeroed samples that are grouped consecutively within each symbol. At least 1088 symbols (same as in 16 frames [1]) were simulated in each point. The Viterbi algorithm was used to decode the convolutional codes and the results for both hard decision decoding and unquantised soft decision decoding [23] are shown in the figure.

The square side ratio for the proposed algorithm was kept the same in all iterations and was chosen to be 0.98 for the hard decision decoding and 0.95 for the soft decision decoding, except for the curve marked "Soft decision after 6 iter. limit 0.98", where 0.98 was chosen. These values appeared to be good compromises for all choices of SNR, amount of zeroed samples and number of iterations in the figure.

As can be seen in the figure, both hard and soft decision decoding were improved with the proposed algorithm. For

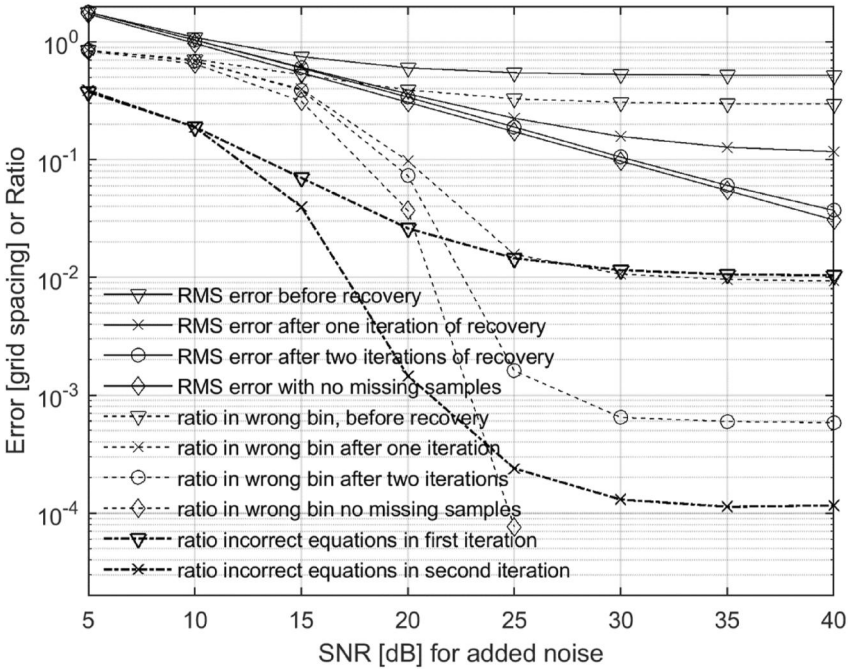

FIGURE 9 Root Mean Square (RMS) error relative to grid spacing and ratio of data carriers closer to another constellation grid point than the correct one are shown for a signal recovered with one and two iterations respectively. The signal was missing 200 samples per symbol (about 2.4%). Choosing this number of missing samples clearly showed that a second iteration gives a significant improvement. 800 symbols were used.

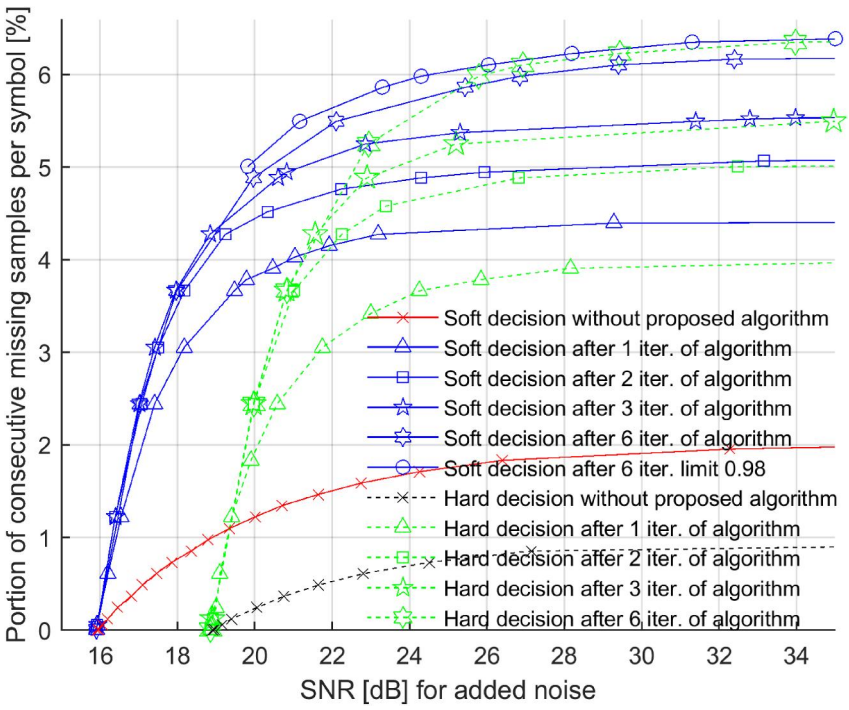

FIGURE 10 The number of iterations required for achieving a Bit Error Rate (BER) of about $2{\cdot}10^{-4}$ after convolutional decoding depends on Signal-to-Noise Ratio (SNR) and amount of zeroed samples. A BER of $2{\cdot}10^{-4}$ corresponds to a signal with less than one uncorrected error event per hour after Reed-Solomon decoding. The square side ratio was chosen to be 0.98 for hard decision and 0.95 for soft decision in all iterations, unless otherwise written.

example, the maximum amount of zeroed samples for achieving a BER of about $2{\cdot}10^{-4}$ after convolutional decoding without the proposed algorithm was increased roughly by a factor between 4 (for high SNR) and 10 (for low SNR, i.e. the end of the curves near the x-axis) for hard decision and by a factor between 2.2 and 4.7 for soft decision after only one iteration of the proposed algorithm.

After six iterations, the maximum amount of zeroed samples for achieving a BER of about $2{\cdot}10^{-4}$ after convolutional decoding was increased by a factor between 6.5 and 10 for hard decision and by a factor between 3.1 and 5.5 for soft decision.

Also, with about 5% missing consecutive samples and a SNR less than 20 dB, a signal with less than one uncorrected error event per hour could be accomplished with six iterations of the proposed algorithm. This can be compared to the oversampling of the signal by 20% (8192/6817-1) before missing samples.

In ref. [1] (Table A.1 in ref. [1]), the carrier-to-noise ratio (CNR) required for 64-QAM, code rate 2/3 and Gaussian Channel to achieve a BER of about $2{\cdot}10^{-4}$ after convolutional decoding with the Viterbi decoder, is about 16.7 dB. In Figure 10, for no (or a few) zeroed samples, the SNR required for a BER of about $2{\cdot}10^{-4}$ is about 15.9 dB for soft decision decoding with or without use of the proposed algorithm. This is in agreement with the CNR value of 16.7 dB given in ref. [1] since the value of SNR in the simulations is about 0.80 dB ($10 \cdot lg(8192/6817)$) lower than the value of the CNR for the same signal.

The time-samples-to-carrier-amplitudes function matrix, $\mathbf{G}_k$, needed only be calculated in the first iteration and was then reused in subsequent iterations saving execution work and time.

5.5 | Simulation of recovered signal in radar

In this section, results are presented from a simulation where the algorithm (without further decoding) was applied to a passive bistatic radar using the same received signal for both reference and surveillance and with a receiver with missing samples. A strong ambiguity peak in the cross-ambiguity function for a DVB-T signal appears at zero-Doppler and a delay of $T_U/3$ [14], where T_U is the time duration of the useful part of a symbol without the guard interval [1]. A simulation was performed with six point targets located around $T_U/3$ at different delays and with different strengths, see Figure 11. In terms of demodulating the received signal in order to create a mismatched reference signal from this only one received signal, the targets added some multipath effects to the received signal and since the guard interval $T_U/8$ was simulated, they also gave rise to some inter-symbol interference.

In the upper diagram in Figure 11 the autocorrelation of the received signal (with targets, noise and missing samples) is shown. The cross-correlation of the received signal with the signal directly after inverse Fourier transform and guard interval insertion (signal without distortion, without targets, with no noise and with no samples missing) is shown in the same diagram. It is visible that the ambiguity at zero-Doppler and delay of $T_U/3$ is stronger than the targets. The two correlation curves in the diagram almost coincide for the targets and the ambiguity.

The received signal was recovered in one iteration with the proposed algorithm with the concentric square side length equal to 0.96 times the length of the side of the circumscribing square. The recovered signal was then demodulated by rounding the data carrier points in the constellation to the nearest point. The coefficients of the pilot carriers were then attenuated by multiplying them with a factor of 9/16 in order to construct a mismatched

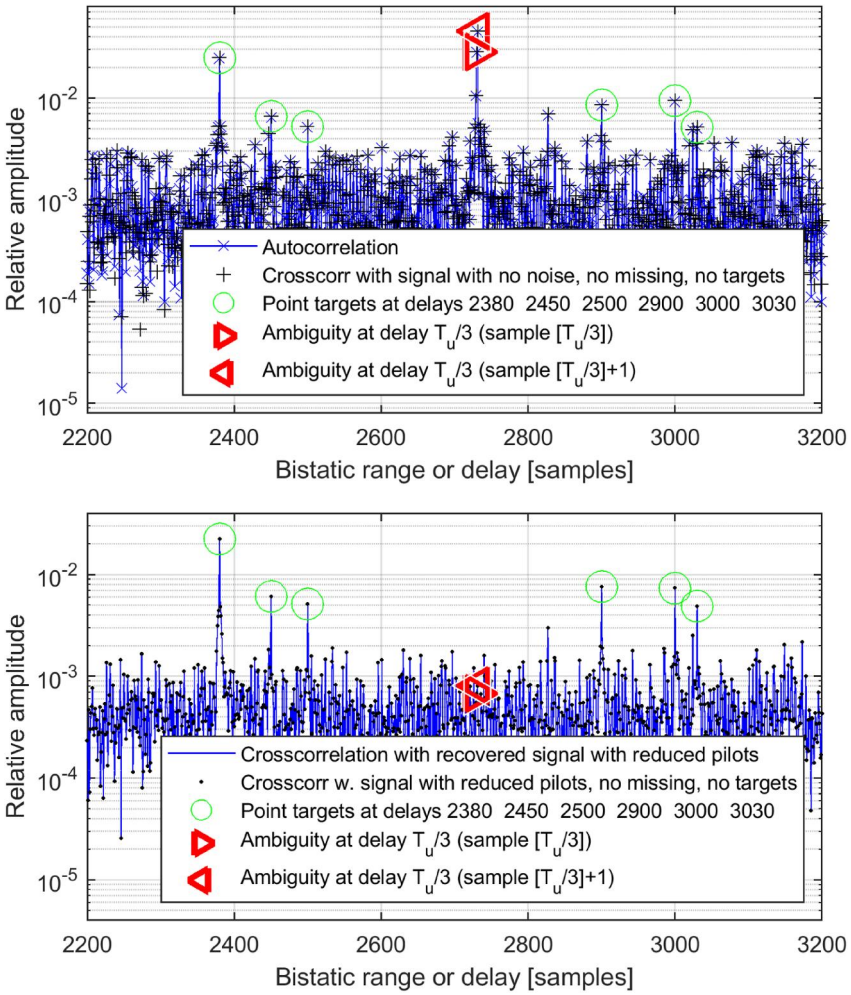

FIGURE 11 Cross-correlation amplitude relative (normalised) to the autocorrelation at zero delay. A Digital Video Broadcasting Terrestrial (DVB-T) signal with targets, noise and missing samples was correlated over 2000 symbols with the indicated signals. The signal was missing 4·20 samples out of 8192 and the Signal-to-Noise Ratio (SNR) was 25 dB. Six point targets were located at bistatic delays (ranges) of 2380, 2450, 2500, 2900, 3000 and 3030 times $T_U/8192$ respectively. The targets amplitudes were approximately the signal amplitude divided by 40, 160, 180, 120, 160 and 180 respectively. Doppler processing was not simulated.

filter to suppress the intra-symbol ambiguities [14]. The mismatched filter was then constructed by applying the inverse Fourier transform and guard interval insertion.

The received signal was cross-correlated with this mismatched filter and the result is shown in the lower diagram in Figure 11. Shown in the same diagram, is also the cross-correlation of the received signal with the mismatched filter instead based on the signal directly after inverse Fourier transform and guard interval insertion (signal without distortion, without targets, with no noise and with no samples missing). The two correlation curves in the diagram almost coincide for the targets. It is visible that the ambiguity at zero-Doppler and delay of $T_U/3$ has become weaker than the six targets by using the mismatched filter. It is visible that the amplitudes of the targets relative to the amplitude of the ambiguity improved about 16 dB, which is about 32 dB power. The proposed algorithm made it possible to use rounding to successfully demodulate the signal with missing samples so that the intra-symbol ambiguity at delay of $T_U/3$ was suppressed. The result of the cross-correlations with the signal recovered with the proposed method were very similar to the results of cross-correlating with the signal directly after inverse Fourier transform and guard interval insertion (in the transmitter).

6 | RESULTS FROM RECOVERY OF MISSING SAMPLES IN MEASURED SIGNAL

In order to show that the method works for a signal with multipath effects present, requiring use of pilot and TPS carriers and time and frequency synchronisation, the proposed method was also applied to real data from measurements performed with the LORA system in a stationary configuration.

In Figure 12, results are shown when two iterations of the proposed algorithm were run on a real-valued (not I/Q sampled) DVB-T signal measured with the LORA system in a stationary configuration using a commercial TV-antenna. With LORA's sampling rate of 25.6 MHz, the duration of the useful part of the symbol, T_U (896 μs), corresponds to 22,937.6 samples. This non-integer number was used as the denominator in the exponential functions of the Fourier transform inside $\mathbf{G}_k$ and the missing data was kept as an integer number of unknown samples. Between 200 and 240 samples are missing in a symbol measured with LORA. The measured signal was a band-pass (not low-pass) signal. Multipath effects were present in the signal, the channel was estimated using pilots and TPS carriers, and equalisation was performed.

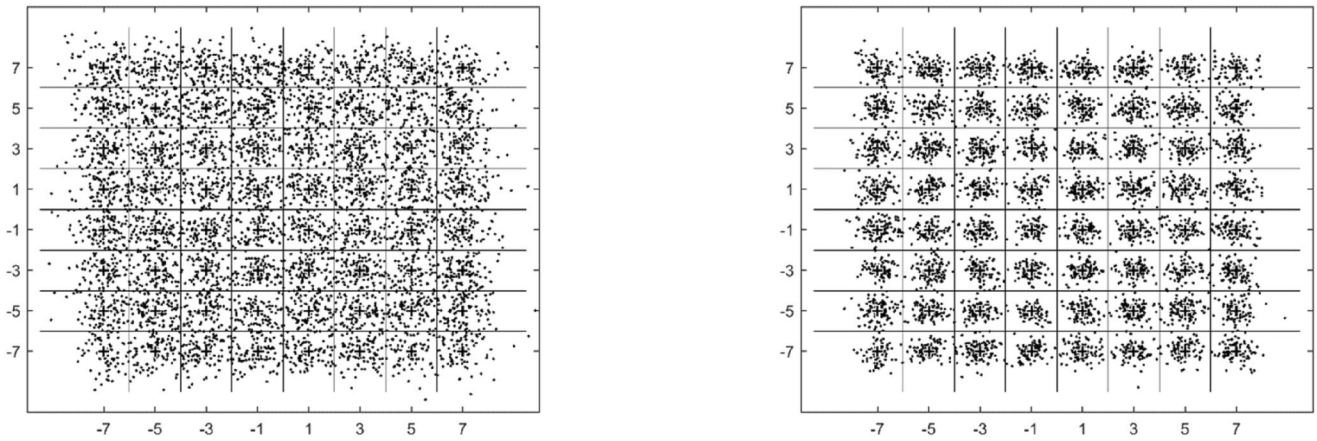

FIGURE 12 Constellation diagrams: (left diagram) real, measured data where 240 missing samples have been replaced with zeros; (right diagram) the same data where the 240 missing samples have been replaced with values calculated after two iterations of the proposed algorithm.

The RMS error for the data carriers before recovery (left diagram in Figure 12) was estimated to 0.38 relative to the grid spacing by measuring the error to the nearest ideal constellation point. Note that the ideal constellation point that a data carrier point belongs to, is not known, so there will be errors in the estimate. The estimate of the RMS error will be lower than it actually is when data carriers belong to an ideal constellation point that is not the nearest. The RMS error of 0.38 relative to the grid spacing is only slightly lower than the theoretical RMS value of about 0.41 for a constellation where the points are evenly distributed in the square area around an inner point.

After one iteration of recovery, the RMS error for the data carriers decreased to 0.28 relative to the grid spacing. The RMS error after two iterations of recovery (right diagram in Figure 12) was estimated to 0.25 relative to the grid spacing. If only the inner 6 × 6 data carrier points are considered, the RMS error before recovery was estimated to 0.36 relative to the grid spacing. After one iteration of recovery, the RMS error for the inner points decreased to 0.27 and after two iterations of recovery, the RMS error for the inner points decreased to 0.24 relative to the grid spacing. It is visible how the proposed algorithm concentrated the data carriers' coefficients in the constellation diagram.

7 | DISCUSSION

The equations (6)-(17) are written in matrix representation for simplification. A computer programme implementing them could be made more efficient (with respect to used memory and execution time) if for example, the coefficient equations are weighted without implementing the diagonal scaling matrices. If the size of the problem is increased then the number of multiplications per iteration required in the proposed method (except for computing the time-samples-to-carrier-amplitudes function matrix $\mathbf{G}_k$) is proportional to either the square of the number of samples in a symbol (N_s) or to the total number of in-band and out-of-band equations (i.e. the total number of non-zero elements in $\mathbf{D}_k^{\mathbf{R}}$, $\mathbf{D}_k^{\mathbf{I}}$ and $\mathbf{P}^{\mathbf{C}}$ (twice)) multiplied with the square of the number of missing samples (N_m).

In the proposed objective function, only the data carriers in the DVB-T signal are considered. However, the values of the pilots and TPS carriers are used for estimating the channel for calculation of the data carrier points and will thereby indirectly affect the value of the objective function.

Expression (20) for the ratio of incorrect extra equations and expression (19) integrated to give expected number of extra equations in one symbol can be used together with the SNR and the condition number to make the weights for the extra equations depend on the real part or imaginary part of the data carrier in a different way than proposed.

The proposed constraints associated with the known real or imaginary parts of selected outer data carriers could probably be utilised (e.g. together with optimisation as shown above) in other applications where the known (real and imaginary parts of) pilots and TPS carriers are used today, for example, in OFDM channel estimation.

The algorithm could also be used for interference that gives short bursts of unusable samples.

8 | CONCLUSIONS

We have proposed a method for reconstructing (or recovering) the OFDM signal in the DVB-T standard where data gaps appear and samples are missing. The proposed method is based on optimisation of an objective function. The proposed objective function consists of two parts. The first part is a function of the energy in the out-of-band frequencies, whereas the second, and novel part, uses the real or imaginary parts of the location of selectable high amplitude data carriers in the constellation diagram.

The proposed method was shown to significantly improve the OFDM signal in just one or two iteration steps. It was shown with simulations that the proposed method improved demodulation of a DVB-T signal so that it can be used for reconstructing a mismatched reference signal in a passive radar

missing a portion of samples without having to implement and perform decoding and coding of the OFDM signal. It was shown that the use of this mismatched reference signal suppressed intra-symbol ambiguities in the DVB-T signal, which can improve detection in a passive radar. We showed an example of a simulation with a passive radar with a receiver with about 1% missing samples, where the power of the targets relative to the power of a strong ambiguity in the DVB-T signal was improved about 32 dB.

We have also applied both hard and unquantised soft decision decoding to the signal after using the proposed algorithm. We have shown that the proposed algorithm significantly increases the maximum amount of missing samples to achieve less than one uncorrected error event per hour for both hard decision and soft decision decoding compared to without using the proposed algorithm. After six iterations of the proposed algorithm, the maximum number of missing samples for achieving less than one uncorrected error event per hour was increased by a factor between 6.5 and 10 for hard decision and by a factor between 3.1 and 5.5 for soft decision.

The proposed method was shown to improve the condition number more than a factor 10^{10} compared to using the least square method on only the out-of-band frequencies (the condition number estimates how much small relative numerical errors in the input data grow into the output). It was shown that for about 1% missing samples (distributed consecutively in four or five groups) per symbol, using the least square method with only the out-of-band frequencies, that is, without the proposed extra equations, made the results much worse than setting the missing samples to zero.

It was shown that for about 1% missing samples (distributed consecutively in four or five groups) per symbol and an SNR of 25 dB, the portion of data carriers that were rounded to wrong constellation point, was about 0.10 before recovery, and about $2{\cdot}10^{-4}$ after only one iteration of the proposed method.

It was shown that for about 2.4% missing samples (distributed consecutively in four or five groups) per symbol and an SNR of 25 dB, the portion of data carriers that were rounded to wrong constellation point, was about 0.33 before recovery, and about 0.002 after two iterations of recovery.

It was shown that the outer data carriers can be selected in a way that the results are not very sensitive to accidently using a portion of incorrect extra equations among the correct equations.

Finally, when the proposed method was applied to a symbol in a measured signal, the estimated RMS error in the constellation decreased from 0.38 to 0.25 relative to the grid spacing.

AUTHOR CONTRIBUTIONS

Anders Haglund: Conceptualization; formal analysis; investigation; methodology; visualization; software; writing – original draft; writing – review & editing; validation. **Per-Olov Frölind**: Conceptualization; project administration; supervision; writing – review & editing. **Lars M. H. Ulander**: Conceptualization; funding acquisition; project administration; supervision; writing – review & editing.

ACKNOWLEDGEMENTS

We would like to thank the Swedish Defence Research Agency (FOI) and the Department of Space, Earth and Environment at Chalmers University of Technology for funding, for providing resources and for making this work possible.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The datasets generated during and/or analysed during the current study are not publicly available (due to that the data was generated with publicly available tools and data) but are available from the corresponding author on reasonable request. The data (i.e. the simulated DVB-T signal) was generated with the open-source software GNU radio [1] together with additional software and video sample file from the gr-dvbt project [2]. [1] GNU Radio, GNU Radio Project, http://www.gnuradio.org, Accessed January 2019 [2] DVB-T implementation using GNU Radio, http://github.com/BogdanDIA/gr-dvbt. Accessed January 2019.

ORCID

Anders Haglund https://orcid.org/0009-0004-4546-4869
Lars M. H. Ulander https://orcid.org/0000-0001-5757-9517

REFERENCES

1. ETSI: Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television. In: ETSI EN 300 744 V1.6.2 (2015-10). European Telecommunications Standards Institute (2015)
2. Baczyk, M.K., Malanowski, M.: Decoding and reconstruction of reference DVB-T signal in passive radar systems. In: 11-th International Radar Symposium, pp. 1–4 (2010)
3. Mahfoudia, O.: Optimum Reference Signal Reconstruction for DVB-T Passive Radars. PhD Thesis. Royal Military Academy and Université Libre de Bruxelles, Belgium (2017)
4. Haglund, A., Frölind, P.-O., Ulander, L.M.H.: Simulation of effect of periodically missing samples on decoding in passive synthetic aperture radar system using OFDM. Inter. Geosci. Remote Sens. Sympos., 2949–2952 (2019). https://doi.org/10.1109/IGARSS.2019.8898806
5. Ferreira, PJSG: Iterative and noniterative recovery of missing samples for 1-D band-limited signals. In: Marvasti, F. (ed.) Nonuniform Sampling. Theory and Practice, pp. 235–281. Springer, Boston (2001)
6. Selva, J.: FFT interpolation from nonuniform samples lying in a regular grid. IEEE Trans. Signal Process. 63(11), 2826–2834 (2015). https://doi.org/10.1109/tsp.2015.2419178
7. Saini, R., Cherniakov, M.: DTV signal ambiguity function analysis for radar application. IEE Proc. Radar, Sonar Navig. 152(3), 133–142 (2005). https://doi.org/10.1049/ip-rsn:20045067
8. Griffiths, H.D., Baker, C.J.: An Introduction to Passive Radar, ed. Artech House, Norwood (2022)
9. Malanowski, M.: Signal Processing for Passive Bistatic Radar. Artech House, Norwood (2019)
10. Frölind, P.-O., et al.: Analysis of a ground target deployment in an airborne passive SAR experiment. In: 2017 IEEE Radar Conference, pp. 273–278 (2017). https://doi.org/10.1109/RADAR.2017.7944211
11. Ulander, L.M.H., et al.: Passive synthetic-aperture radar for detection of ground vehicles. Proc. IET Radar 2017, 4 (2017). https://doi.org/10.1049/cp.2017.0458

12. Tao, R., Gao, Z., Wang, Y.: Side peaks interference suppression in DVB-T based passive radar. IEEE Trans. Aero. Electron. Syst. 48(4), 3610–3619 (2012). https://doi.org/10.1109/taes.2012.6324746
13. van de Beek, J.J., Sandell, M., Börjesson, P.O.: ML estimation of time and frequency offset in OFDM systems. IEEE Trans. Signal Process. 45(7), 1800–1805 (1997). https://doi.org/10.1109/78.599949
14. Palmer, J.E., et al.: DVB-T passive radar signal processing. IEEE Trans. Signal Process. 61(8), 2116–2126 (2013). https://doi.org/10.1109/tsp.2012.2236324
15. Bongioanni, C., et al.: A new approach for DVB-T cross-ambiguity function evaluation. In: 2009 European Radar Conference (EuRAD), pp. 37–40 (2009)
16. Ulander, L., et al.: The VHF/UHF-band LORA SAR and GMTI system. In: Proceedings of SPIE, Algorithms for Synthetic Aperture Radar Imagery X, vol. 5095, pp. 206–215. SPIE, Orlando (2003). https://doi.org/10.1117/12.487851
17. Ulander, L.M.H., et al.: Change detection of vehicle-sized targets in forest concealment using VHF- and UHF-band SAR. IEEE AES Syst. Magaz. 26(7), 30–36 (2011). https://doi.org/10.1109/maes.2011.5958761
18. Marvasti, F., et al.: A unified approach to sparse signal processing. EURASIP J. Appl. Signal Process. 2012(1), 1–44 (2012). https://doi.org/10.1186/preaccept-1686979482577015
19. Dahlquist, G., Björck, Å.: Numerical Methods in Scientific Computing, vol. I. Society for Industrial and Applied Mathematics, Philadelphia (2008)
20. van de Beek, J.-J., et al.: On Channel estimation in OFDM systems. IEEE Vehic. Technol. Conf. 2, 815–819 (1995)
21. Budge, M.C., German, S.R.: Basic Radar Analysis, 2nd ed. Artech House (2020)
22. ETSI: Digital video broadcasting (DVB); measurement guidelines for DVB systems. In: ETSI TR 101 290 V1.3.1. Joint Technical Committee (JTC) Broadcast of the European Broadcasting Union (EBU), Comité Européen de Normalisation ELECtrotechnique (CENELEC) and European Telecommunications Standards Institute (ETSI) (2014-07)
23. Proakis, J.G., Salehi, M.: Digital Communications, 5th ed. McGraw-Hill (2008)

How to cite this article: Haglund, A., Frölind, P.-O., Ulander, L.M.H.: Recovery of missing samples in Orthogonal Frequency Division Multiplexing signals with optimisation using data carriers. IET Radar Sonar Navig. 1–18 (2024). https://doi.org/10.1049/rsn2.12560