Estimating Faults Modes in Ball Bearing Machinery using a Sparse Reconstruction Framework*

Maria Juhlin*, Johan Swärd*, Marius Pesavento†, and Andreas Jakobsson*
*Div. of Mathematical Statistics, Lund University, Sweden,
†Communication Systems Group, Technische Universität Darmstadt, Germany

***Abstract*—In this work, we present a computationally efficient algorithm for estimating fault modes in ball bearing systems. The presented method generalizes and improves upon earlier developed sparse reconstruction techniques, allowing for detecting multiple fault modes. The measured signal is corrupted with additive and multiplicative noise, yielding a signal that is highly erratic. Fortunately, the damaged ball bearings give rise to strong periodical structures which may be exploited when forming the proposed detector. Numerical simulations illustrate the preferred performance of the proposed method.**

***Index Terms*—Ball bearing systems, sparse reconstruction, convex optimization, ADMM**

I. Introduction

Ball bearings are an essential part of many mechanical systems as they can carry high loads without significant energy being lost to friction. For optimal performance, minimizing the friction, the balls should not have any direct contact with the metal cage, running instead on a thin film of oil. It is important that such optimal conditions are preserved, as any disruptions, such as, for example, an imbalanced metal cage or a dirty oil film will subject the ball bearings to much wear and tear, and due to the rapid motion of the balls, the resulting degradation can occur swiftly. Malfunctions in the machinery can also have a large and costly business impact; as a result, it is important to identify and rectify malfunctions in the bearings at an early stage, before the errors start to affect the machinery, so that maintenance can be scheduled without unnecessary production halts. As ball bearings are often an integrated part of the machine, detecting malfunctions at an early stage is difficult, as direct inspection of the ball bearings is usually not feasible without stopping the machine. As a result, the development of non-invasive methods for error detection is important (see, e.g., [1]–[4]). Because of the rotational nature of ball bearings, disruptions will manifest as harmonically related ringings in the vibrational response of the ball bearing system. Although malfunctions can result for various reasons, these can roughly be divided into three groups of problems. As noted in [1], these groups of malfunctions give rise to distinct error patterns in the resulting vibrational signal. As a result, by estimating the fundamental frequencies and the harmonic structure in the vibrational response, one may detect and determine different kinds of problems with the ball bearing system. As ball bearings are an integrated part of the machine, the individual vibrational signals are generally not available from measurement, and instead a composite signal resulting from the entire machine is what may be measured, including various forms of engine vibrations and noise.

*This work was supported in part by the Swedish Research Council.

Due to the strong presence of noise and other forms of tonal-like interference, traditional techniques for estimating fundamental frequencies of harmonically related tones (see, e.g., [5]) will have difficulties in yielding reliable estimates, necessitating the development of techniques tailored to the expected structure of the vibrational signal. Reminiscent to the pitch estimation problem for audio signals (see, e.g., [6], [7]), one may reduce the problem of tonal interference as well as exploit the harmonic structure of the expected signal, by utilizing both the sparse structure of the signal and the group sparse structure resulting from the harmonics. Such an effort was made in [2], where some different approaches for identifying error patterns from ball bearing signals were developed and compared, including a group sparse pitch estimator. The here proposed method expands on this approach and first estimates the harmonically related sinusoids in the signal, and then, in a second stage, uses this knowledge to search for the other distinct error patterns that may be present in the signal. Furthermore, a computationally efficient Alternating Direction Method of Multipliers (ADMM) is derived which dramatically decreases the computational cost as compared to off-the-shelf convex solvers, such as [8]. Using simulated data, we examine the performance of the resulting estimator and compare with the results obtained using the group sparse estimator suggested in [2], clearly indicating the preferable performance of the proposed method.

II. The vibrational signal model

The vibrational response of a ball bearing system is not directly measurable, with the signal of interest being modulated with an unknown modulation that has to be removed before further processing [1]. For completeness, we begin by briefly reviewing how this is done. Typically, the demodulation of a signal with an unknown modulation is formed by computing the envelope of the corresponding discrete-time analytic signal [9]. This process causes the signal components to be mixed together, creating many spurious peaks, generally making it difficult to extract the signal of interest. However, as the vibrational response consists of harmonically related tones, this underlying structure will be preserved. In general, the noise-free envelope of the vibrational signal consisting of P harmonically related tones generated, e.g., by balls periodi-

ISBN 978-90-827970-1-5

cally running over a crack on the ring with frequency ω_0, may be well modeled as [1], [2]

$$x(t_n) = \sum_{\ell=1}^{P} \alpha_\ell \cos(\omega_0 \ell t_n + \phi_\ell) \quad (1)$$

where t_n denotes the nth time sample, α_ℓ, ω_0, and $\phi_\ell \in [0, 2\pi)$ the amplitude, frequency, and the phase of the rotational system, respectively. Due to the rotation of the bearing with frequency ω_c the ball bearing ring (if not fixed) periodically moves into and out of the pressure zone resulting in a modulation of the vibration signal $x(t_n)$ with a harmonic signal. The modulation of the signal implies that the measured signal can be expressed as

$$g(t_n) = \sum_{\ell=1}^{P} \alpha_\ell \alpha_c \cos(\omega_0 \ell t_n + \phi_\ell) \cos(\omega_c t_n + \phi_c) \quad (2)$$

where α_c, ω_c, and ϕ_c denote the unknown modulation gain, frequency, and phase, respectively. The envelope of the analytical signal of $g(t_n)$, denoted by $y(t_n) = |g_A(t_n)|^2$, can thus be written as

$$\begin{aligned} y(t_n) =& \frac{\alpha_c^2}{4} \Bigg(\alpha_i \alpha_j \cos\left(\omega_0 (j-i) t_n + (\phi_j - \phi_i)\right) + \\ & \sum_{i=1}^{P} \sum_{j=1}^{P} \alpha_i \alpha_j \cos\left(\omega_0 (i-j) t_n + (\phi_i - \phi_j)\right) + \\ & 2\alpha_i \alpha_j \cos\left(\omega_0 (j+i) t_n + \phi_j + \phi_i\right) \Bigg) \end{aligned} \quad (3)$$

The demodulation procedure ensures that the resulting signal is independent of the unknown modulation, but also causes the harmonic structure to be extended with $2P$ overtones, being formed as sums of these harmonics. The result emphasizes the importance to include the full harmonic structure when devising an algorithm aimed at retrieving the frequencies[1]. To simplify our notation, we proceed to express the demodulated noise-free ball bearing vibrational signal suffering from rotational imperfections as (see also [2])

$$\begin{aligned} y(t_n) =& \sum_{k=1}^{L} \alpha_k e^{i\omega_0 k t_n + ik\phi} + \sum_{k=1}^{L} \sum_{m=1}^{M_k} \beta_k e^{i\omega_0 k t_n \pm i\Delta m t_n + \phi k i} \\ & + \sum_{m=1}^{M_k} \beta_k e^{i\Delta m t_n + \phi k i} \end{aligned} \quad (4)$$

In this more general formulation, the response also allows for malfunctions that cause the presence of L harmonic overtones of the system frequency, each with a linearly increasing phase[2]. Some forms of imperfections, for instance due to a radial load [1], will also cause a distinct error pattern in the rotational information, causing the presence of the frequency modulation detailed in the second term of (4), where Δ denotes the frequency modulation offset caused by the rotational imperfections, with M_k being the number of such modulations for tone k. Some such modulation components may be missing, causing the corresponding modulation amplitudes, $\beta_k \in \mathbb{R}$, to be zero. Roughly, the error patterns resulting from malfunctions may be separated into two different groups; multiples of the fundamental frequency, and multiples of the fundamental frequency combined with sidebands around each harmonic [1]. As one is restricted to measuring a composite signal that can include, among other things, vibrations from the engine and other mechanical parts, the identification problem is twofold: firstly, one has to determine which of the signal components that belong to the rolling bearings, and then, secondly, identify the possible error pattern present. Figure 1 illustrates the periodogram of a typical demodulated ball-bearing signal. The signal, formed using (1) prior to modulation, contains a single harmonical vibration signal, consisting of a fundamental frequency with $L = 4$ harmonics and four sidebands ($M = M_k = 2, \forall k = 1, \ldots, L$), centered around each harmonics. The vibrational signal should thus contain $2LM = 16$ components. However, as can be seen from the figure, due to the demodulation step, the demodulated signal contains far more components, although the harmonic structure can still be seen, but now being extended to include $2L$ additional harmonics.

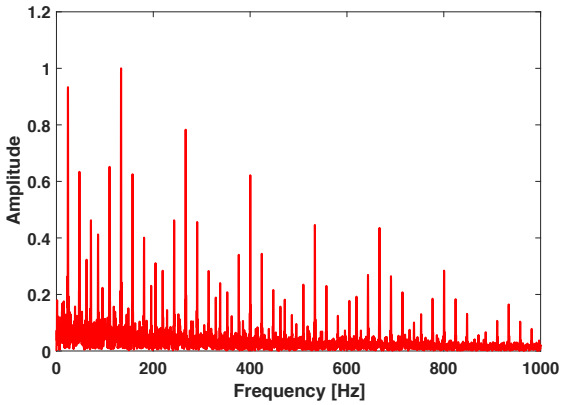

Fig. 1. The spectrum of a demodulated ball-bearing system. The figure shows the harmonically related groups of spectral peaks resulting from imperfections in the system, causing modulation components for each harmonic.

III. Proposed method

We proceed to introduce the proposed estimator for ball bearing signals. We start by separating the main problem in to two parts. The first part is to determine the harmonically related parts in the signal. In the second step, we utilize this information to determine the presence of any sidebands. In order to determine the harmonically related frequencies in (3), we begin by forming a grid over the fundamental frequencies,

[1]To simplify the presentation, the modulation signal has here been assumed to be sinuoidal; in real application, it is worth noting that this modulation is often better modelled using a low-order autoregressive model. This will have the effect of further spreading the signal power over the band, although the modulated frequencies will still appear shifted in the same way. This effect is also visible in Figure 1, where the low power noise component is seen to be shaped by the autoregressive modulation.

[2]It is worth noting that, occasionally, some of these overtones, including the actual fundamental frequency, may be missing, in which case the corresponding amplitude, $\alpha_k \in \mathbb{R}$, is zero.

ISBN 978-90-827970-1-5

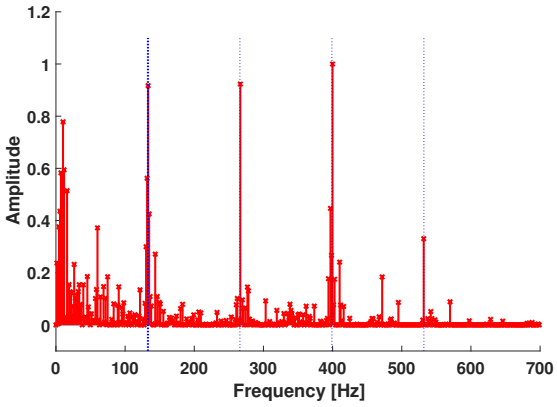

Fig. 2. The figure illustrates the activated components of a typical signal using the estimator in [2].

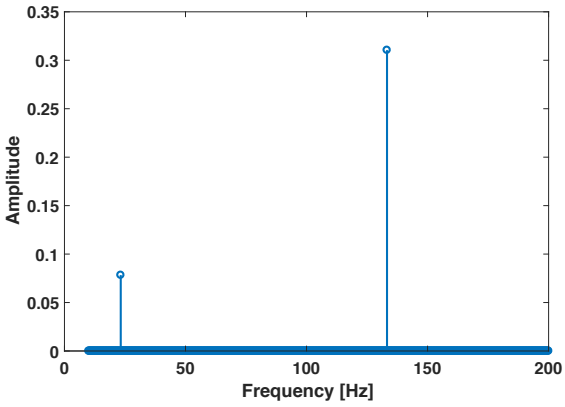

Fig. 3. The figure illustrates the activated pitch groups of a typical signal using the proposed estimator.

ranging from ω_{min} to ω_{max}, with length G. Furthermore, let $\mathbf{A}$ denote a dictionary matrix containing a large set of potential candidate fundamental frequencies and their corresponding harmonics, such that

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \dots & \mathbf{A}_G \end{bmatrix} \tag{5}$$

$$\mathbf{A}_g = \begin{bmatrix} \mathbf{a}_{g,1} & \dots & \mathbf{a}_{g,L_{\max}} \end{bmatrix} \tag{6}$$

$$\mathbf{a}_{g,\ell} = \begin{bmatrix} e^{i\omega_g \ell t_1} & \dots & e^{i\omega_g \ell t_N} \end{bmatrix}^T \tag{7}$$

where $(\cdot)^T$ denotes the transpose and $L_{\max}$ an upper limit on the expected number of harmonics. The problem of determining the fundamental frequencies may then be expressed as

$$\underset{\mathbf{x}}{\text{minimize}} \quad \frac{1}{2}||\mathbf{s} - \mathbf{A}\mathbf{x}||_2^2 + \lambda||\mathbf{x}||_1 + \gamma \sum_{g=1}^{G} ||\mathbf{x}_g||_2 \tag{8}$$

where $\mathbf{s} = \begin{bmatrix} s(t_1) & \dots & s(t_N) \end{bmatrix}^T$, with $\mathbf{x}$ denoting the amplitude vector for each candidate frequency, corresponding to each column in $\mathbf{A} \in \mathbb{C}^{N \times GL_{\max}}$, and $\mathbf{x}_g$ corresponds to all amplitudes for group g, whereas λ and γ are user determined regularization parameters governing the overall sparsity and the group sparsity of the solution, respectively. The latter penalty enforces the group sparse structure ensuring that only a few candidate fundamental frequencies are retained in the solution, whereas the former penalty strives to minimize the overall number of components in the solution. As (8) is a convex optimization problem, the solution may be found using a standard off-the-shelf solver such as SDP3 [10], although such solvers generally become computationally cumbersome as the problem dimension grows.

To alleviate this problem, we will instead exploit a solution formed using an ADMM (see, e.g., [11]). The ADMM works by separating the original problem into smaller and simpler problems which are solved separately. To this end, we split the original variable $\mathbf{x}$ in (8) into two separate variables, denoted $\mathbf{x}$ and $\mathbf{z}$, forming

$$\begin{aligned} \underset{\mathbf{x},\mathbf{z}}{\text{minimize}} \quad & \frac{1}{2}||\mathbf{s} - \mathbf{A}\mathbf{x}||_2^2 + \lambda||\mathbf{z}||_1 + \gamma \sum_{g=1}^{G} ||\mathbf{z}_g||_2 \\ \text{subject to} \quad & \mathbf{x} = \mathbf{z} \end{aligned} \tag{9}$$

To derive the steps in the ADMM method, one then forms the augmented Lagrangian of (9), i.e.,

$$\begin{aligned} L(\mathbf{x},\mathbf{z}) = & \frac{1}{2}||\mathbf{s} - \mathbf{A}\mathbf{x}||_2^2 + \lambda||\mathbf{z}||_1 + \\ & \gamma \sum_{g=1}^{G} ||\mathbf{z}_g||_2 + \frac{\rho}{2}||\mathbf{x} - \mathbf{z} + \mathbf{u}||_2^2 \end{aligned} \tag{10}$$

where ρ denotes the step size and $\mathbf{u}$ the (scaled) dual variable. In the first step of the ADMM algorithm, one solves (10) for the $\mathbf{x}$ variable, keeping the other variables fixed, i.e,.

$$\mathbf{x}^{(k+1)} = \left(\mathbf{A}^H\mathbf{A} + \rho\mathbf{I}\right)^{-1}\left(\mathbf{A}^H\mathbf{s} + \rho(\mathbf{z}^{(k)} - \mathbf{u}^{(k)})\right) \tag{11}$$

where $(\cdot)^{(k)}$ denotes the kth iteration. Since (10) is not differentible for all $\mathbf{z}$, one requires the use of subdifferential calculus to find the solution (see also [12])

$$\mathbf{z}_g^{(k+1)} = \mathscr{S}\left(S(\mathbf{x}_g^{(k+1)} + \mathbf{u}_g^{(k)}, \lambda/\rho), \gamma/\rho\right) \tag{12}$$

for $g = 1, \dots, G$, where

$$\mathscr{S}(\mathbf{v},\kappa) = \frac{\max(||\mathbf{v}||_2 - \kappa, 0)}{\max(||\mathbf{v}||_2 - \kappa, 0) + \kappa}\mathbf{v} \tag{13}$$

$$S(\mathbf{v},\kappa) = \frac{\max(|\mathbf{v}| - \kappa, 0)}{\max(|\mathbf{v}| - \kappa) + \kappa} \odot \mathbf{v} \tag{14}$$

with $\odot$ denoting the element-wise multiplication. Finally, one updates the (scaled) dual variable using

$$\mathbf{u}^{(k+1)} = \mathbf{x}^{(k+1)} - \mathbf{z}^{(k+1)} + \mathbf{u}^{(k)} \tag{15}$$

The steps in (11)-(15) are iterated until sufficient convergence has been achieved. The ADMM algorithm converges under very mild assumptions, such as the terms in (9) have to be closed, proper, and convex functions, and that the augmented

ISBN 978-90-827970-1-5

Algorithm 1 Identifying ω_0 and Δ

1: Let $\hat{G}$ be the number of estimated groups in the solution of (8) and $\hat{\mathbf{L}}$ be the number of harmonics for each group.
2: **if** $\hat{G} = 1$ **then**
3: Set $\hat{\omega}_0$ equal to the fundamental frequency of the found group. This will yield no prior information on Δ.
4: **else if** $\hat{G} = 2$ **then**
5: Form $\boldsymbol{\delta}$ and set $\hat{\Delta} = \min \boldsymbol{\delta}$ and $\hat{\omega}_0$ to the fundamental frequency of the remaining group.
6: **else**
7: Form $\boldsymbol{\delta}$. If there are any suitable candidates in δ, set $\hat{\Delta}$ to the δ with the most occurrences and set $\hat{\omega}_0$ primarily as the frequency with most δs equal to Δ or, secondarily, to the fundamental frequency corresponding to the group with corresponding to max $\hat{\mathbf{L}}$. If there are no candidates in δ that are suitable for Δ, then no prior information on Δ is available. Then, set $\hat{\omega}_0$ equal to the fundamental frequency corresponding to the group with max $\hat{\mathbf{L}}$ harmonics.
8: **end if**

Lagrangian has a saddle point, both being fulfilled in this case, and has been shown to produce solutions with high accuracy after only a few iterations [11], [13].

One ambiguity that arises in the estimation is the so-called halfling problem, which results from the fact that if the true fundamental frequency is ω_0, both ω_0 and $\omega_0/2$ will fit the frequency structure, and will therefore be valid solutions. However, this results in the selection of a too small fundamental frequency and thus needs some treatment. To this end, we exploit the idea presented in [14], and solve the ADMM multiple times, each time using weights to update the γ parameter. These weights are in each iteration i, defined as $w_g^{(i+1)} = 1/(|x_{g,1}^{(i)}| + \epsilon)$, where $x_{g,1}$ denotes the first element in the vector $\mathbf{x}_g$ and ϵ is a small number, typically around 10^{-4}, added to avoid numerical problems. Since the element in $\mathbf{x}$ corresponding to $\omega_0/2$ should be small, the corresponding group will be penalized; the algorithm will thus prefer to choose the group that has ω_0 as its fundamental frequency. Similarly, due to both the additive and the multiplicative noise, the spectrum is not really sparse. This makes the method sensitive to the choice of λ and, in particular, γ, which governs the group sparsity in the solution. To alleviate this, we here make use of the similar reweighting formulation described in [15]. To this end, we form a second weighting vector, $\tilde{w}_g = 1/(||\mathbf{x}_g||_2 + \epsilon)$, for $g = 1, \ldots, G$, which is applied to each group, respectively. The ADMM is then rerun, where in each iteration $\tilde{w}_g$ are updated, until convergence, which is reached after only a few iterations. To speed up the computation, one may warm start the ADMM in each iteration with the previously found solution.

The output from the ADMM determines which groups containing harmonically related sinusoids that are present in the signal, and is used to identify the main harmonic, ω_0. Since the first sideband, corresponding to the third sum in (4), also exhibits a harmonic structure, though often weaker since it contains less overtones than the main harmonic component, it often shows up in the solution to (8). When this is the case, one may use this information to form a prior on the sideband structure, although one still needs to identify any present noise components. To this end, one may utilize the fact that the sideband frequency Δ has an upper bound of half the frequency of ω_0, and the fact that potential sidebands are always formed at equal inter-peak distances from their corresponding ω_0 harmonic [2]. The algorithm for forming the prior on Δ and determining ω_0 can be divided into three cases depending on the number of frequency groups identified by the ADMM, and is summarized in Algorithm 1. The successive inter-peak distances, $\boldsymbol{\delta}$, are formed as the distances between successive peaks, with the first inter-peak distance, δ_1, being the distance between zero and the first peak, and the last one being $\delta_{last} = \infty$. If no such prior-information is attained from the solution to (8), one has to search a broader area around the found harmonic component. In this paper, we have defined this area to be $[2/3\omega_0, 4/3\omega_0]$.

To determine the presence of any sidebands, we propose to again utilize the ADMM algorithm, this time using a different group structure. Define the grid over possible sideband frequencies as $\boldsymbol{\Delta}$ with size Q, and the maximum allowed number of sideband harmonics as M. Furthermore, let $\hat{\omega}_0$ be the estimated fundamental frequency for the main harmonic component and let $\hat{L}$ be the estimated number of harmonics. One may then form a new dictionary for the sideband as

$$\begin{aligned}
\mathbf{D} &= \left[\begin{array}{ccc} \mathbf{D}_1 & \ldots & \mathbf{D}_Q \end{array} \right] \\
\mathbf{D}_q &= \left[\begin{array}{cccc} \mathbf{D}_{q,0} & \mathbf{D}_{q,1} & \ldots & \mathbf{D}_{q,\hat{L}} \end{array} \right] \\
\mathbf{D}_{q,\ell} &= \left[\begin{array}{ccccccc} \mathbf{d}_{q,\ell,-M} & \ldots & \mathbf{d}_{q,\ell,-1} & \mathbf{d}_{q,\ell,1} & \ldots & \mathbf{d}_{q,\ell,M} \end{array} \right] \\
\mathbf{d}_{q,\ell,m} &= \left[\begin{array}{ccc} e^{(\ell\hat{\omega}_0 + m\Delta)t_1} & \ldots & e^{(\ell\hat{\omega}_0 + m\Delta)t_n} \end{array} \right]^T
\end{aligned} \tag{16}$$

We again solve (8), this time with $\mathbf{D}$ as dictionary matrix. The estimated sideband is then selected as the group with the largest norm. Finally, we re-estimate the amplitude for all found sinusoids using the least squares

$$\hat{\mathbf{x}}_{\text{LS}} = \left(\tilde{\mathbf{D}}^H \tilde{\mathbf{D}}\right)^{-1} \left(\tilde{\mathbf{D}}^H \mathbf{s}\right) \tag{17}$$

where $\tilde{\mathbf{D}}$ denotes the dictionary with columns corresponding to the found frequencies. A sideband is deemed to be present if the corresponding group has at least half the energy compared to the main harmonic component. Thus, a sideband is deemed to be present if

$$||\mathbf{x}_\Delta||_1 > \tau ||\mathbf{x}_0||_1, \tag{18}$$

where $\mathbf{x}_\Delta$ and $\mathbf{x}_0$ denote the amplitudes for the group corresponding to the sideband and the main harmonic component, respectively, and τ is a predefined threshold, here set to $\tau = 0.5$. The resulting method, termed the MAD (MAlfunction Detection) algorithm, is summarized in Algorithm 2.

ISBN 978-90-827970-1-5

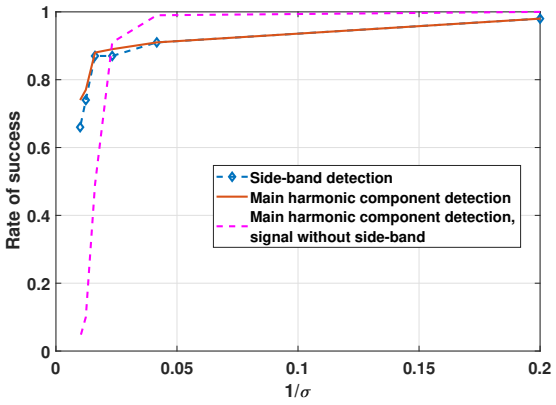

Fig. 4. The figure illustrates the rate of correctly identified error patterns for different values of σ^{-1}, computed using 100 Monte Carlo simulations. For a signal with sidebands and one without sidebands.

IV. Numerical examples

To illustrate the performance of the suggested method, we proceed to examine the method's performance using simulated ball bearing data, having a harmonic frequency at $133 + \epsilon_1$ Hz. The data is simulated both with and without a sideband disturbance at $23.24 + \epsilon_2$ Hz, where $\epsilon_1, \epsilon_2 \in U[0, 0.5]$ are additive noise components added to force the frequencies to lie off-grid. The signal is corrupted by first a multiplicative noise in the form of a second order autoregressive filter with coefficients $[1, -0.6119, -0.9801]$, and then by an additive zero-mean Gaussian noise with variance σ^2.

An example of the resulting signal spectrum is shown in Figure 1, illustrating the large number of spectral components resulting from the demodulation procedure. As is clear from the figure, it is non-trivial to extract the corresponding components from the resulting signal without exploiting the structure of the signal. Figures 2 and 3 illustrate the sparse representations of the harmonic frequency estimates resulting from the group sparse approach presented in [2] and from the here proposed method. As seen in the former figure, the group sparse approach yields a rather cluttered representation, making it difficult to determine which components that make up the harmonic structure, and which make up the demodulation components around the harmonic components. The method does not strive to group frequencies into harmonic groups, resulting in a non-sparse signal representation. On the other hand, the presented MAD estimator yields a sparse estimate of the activated frequency groups, allowing for a reliable subsequent error pattern identification.

Figure 4 shows the probability of correctly identifying the error pattern for varying noise powers for the signal with and without sideband. As can be seen in the figure, for the signal with a sideband the MAD algorithm is able to correctly identify the sideband in 66% of the cases and the main harmonic in 74% of the cases when $\sigma^{-1} = 0.01$, but only estimates the correct harmonic in 4% of the cases when there is no sideband present. However, the performance of the MAD algorithm improves rapidly for both signal types, and for $\sigma^{-1} = 0.02$ it is between 87-91%. The reason for this is that the energy content of the signals in the two cases differs, and the choices of λ and γ have not taken this into consideration. Here, we have selected $\lambda = \alpha_\lambda ||\mathbf{B}^H \mathbf{y}||_\infty$ and $\gamma = \alpha_\gamma ||\mathbf{B}^H \mathbf{y}||_\infty$, where $\mathbf{B}$ denotes the dictionary matrix and $\mathbf{y}$ the signal.

Algorithm 2 The MAD algorithm

1: Input: $\mathbf{x}^{(0)} = \mathbf{z}^{(0)} = \mathbf{u}^{(0)} = \mathbf{0}$, $\rho = 1, \lambda, \gamma > 0, k = 1$, $\alpha_\lambda = 0.3$, $\alpha_\gamma = 0.4$, and $\mathbf{A}$ formed from (5)-(7).
2: Using ADMM, solve (8) using reweighting w_g and $\tilde{w}_g$.
3: Form $\mathbf{D}$ from (16).
4: Using ADMM, solve (8) using reweighting $\tilde{w}_g$, and with $\alpha_\lambda = 0.05$ and $\alpha_\gamma = 0.1$, .
5: Reestimate the sinusoidal amplitudes using (17)
6: Decide if the sideband is present in the signal using (18).

References

[1] B. Geropp, *Schwingungsdiagnose an Wlzlagern mit Hilfe der Hllkurvenanalyse.*, Ph.D. thesis, Techn. Hochsch, Aachen, Aachen, Germany, 1995.

[2] M. Esswein, "Group sparsity in rolling bearing damage diagnosis," M.S. thesis, Darmstadt University of Technology, 64289 Darmstadt, Germany, 2015.

[3] J. Lin, M. J. Zuo, and K. R. Fyfe, "Mechanical fault detection based on the wavelet de-noising technique," *Journal of Vibration and Acoustics*, vol. 126, pp. 9–16, 2004.

[4] J. H. D. Guimarïaes, *Modelling the dynamic interactions of rolling bearings*, Ph.D. thesis, Techn. Hochsch, Aachen, Aachen, Germany, 2007.

[5] M. Christensen and A. Jakobsson, *Multi-Pitch Estimation*, Morgan & Claypool, San Rafael, Calif., 2009.

[6] S. I. Adalbjörnsson, A. Jakobsson, and M. G. Christensen, "Multi-Pitch Estimation Exploiting Block Sparsity," *Elsevier Signal Processing*, vol. 109, pp. 236–247, April 2015.

[7] C. Steffens, M. Pesavento, and M. E. Pfetsch, "A Compact Formulation for the $\ell_{2,1}$ Mixed-Norm Minimization Problem," *IEEE Trans. Signal Process.*, vol. 66, no. 6, pp. 1483–1497, 2018.

[8] Inc. CVX Research, "CVX: Matlab Software for Disciplined Convex Programming, version 2.0 beta," `http://cvxr.com/cvx`, Sept. 2012.

[9] S. L. Marple, "Computing the discrete-time "analytic" signal via FFT," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2600–2603, September 1999.

[10] R. H. Tutuncu, K. C. Toh, and M. J. Todd, "Solving semidefinite-quadratic-linear programs using SDPT3," *Mathematical Programming Ser. B*, vol. 95, pp. 189–217, 2003.

[11] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.

[12] R. Chartrand and B. Wohlberg, "A Nonconvex ADMM Algorithm for Group Sparsity with Sparse Groups," in *38th IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Vancouver, Canada, May 26-31 2013.

[13] J. Eckstein and D.P. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, pp. 293–318, April 1992.

[14] F. Elvander, J. Swärd, and A. Jakobsson, "Online Estimation of Multiple Harmonic Signals," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 2, pp. 273–284, February 2017.

[15] E. J. Candès, M. B. Wakin, and S. Boyd, "Enhancing Sparsity by Reweighted l_1 Minimization," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5, pp. 877–905, Dec. 2008.

ISBN 978-90-827970-1-5