ROBUST MODELING OF A CLASS OF VIBRATION SIGNALS THROUGH BINARY NEURAL NETWORK-BASED SYMBOLIC REGRESSION

Nacer Yousfi [1,2], *Karim Abed-Meraim* [1], *Mohammed El-Badaoui* [2,3]

[1] PRISME Lab., Orléans Univ., France
[2] SAFRAN Tech, Dig. Sc. & Tech. Dep., Magny-Les-Hameaux, France
[3] Univ Lyon, UJM-St-Étienne, LASPI, EA3059, F-42023 Saint-Étienne, France

ABSTRACT

In this paper, we propose an automatic method to unveil the nonlinear multi-modulation model that describes vibration signal of certain complex systems. Firstly, we apply a harmonic estimation algorithm to the observed signal to identify all peaks in the spectral domain. Then, using the peaks' location information, a neural network-based symbolic regression is trained to determine a concise expression of the multi-modulation model underlying the considered signal. A regularized objective function is utilized to optimize the neural network weights and enforce their binarization (sparsity), providing a simple mathematical expression for the input signal model. The proposed algorithm improves the robustness of the determined model by eliminating any distortion in the input signal, such as transfer function and phase modulation distortion. This method could potentially be employed in the vibration analysis of rotating machines, such as planetary gearboxes, due to the structural spectral contents of the vibration signal, to reflect the interaction model between system elements.

Index Terms— Modeling, Symbolic regression, Multi-modulation, Neural network.

1. INTRODUCTION

The excitation responses of rotating systems are generated through an unknown nonlinear mixture function, termed a multi-modulation model [1], typically in a polynomial form. This model encapsulates the interactions among the system's moving parts. The identification of the model is challenged by the fact that the excitations are distorted with complex transfer function effects. Identifying this mixture function is crucial for interpreting system behavior and machine components interaction, particularly during phases of degradation [2][3], and for the implementation of individual signal separation and tracking algorithms within the identified model [4][5]. Removing transfer function effects is considered crucial for obtaining the excitation responses, thereby providing direct information about the health of rotating systems. Blind system identification (BSI) [6] has been explored in mechanical systems, where several preprocessing techniques aim to remove transfer function effects. These include the exponential liftering technique and OMA-based methods, which depend on system operating conditions, system degradation, and defect types [7]. On the other hand, a significant amount of research has been conducted to model rotating machine vibration, leveraging the fact that their signals have very structured spectral contents. This often involves manually inspecting the spectral contents of the vibration signal to establish a mathematical expression that accurately fits the spectrum [1]. However, a unified model for rotating machines with non-conventional architectures, such as planetary gearboxes, is still a challenging problem under study, e.g., [8]. This research aims to elucidate a simple and concise mathematical expression that provides a good fit for the observed signal. We extend the technique proposed in [8] to develop a simpler and more robust method for modeling harmonic signals originating from rotating machines, such as planetary gearboxes. [8] introduces a symbolic regression-based neural network approach to determine a compact formula that describes the observation and separation of individual contributions. but it does not include any processing for handling transfer function effects on system excitation.

In this paper, we propose to first estimate the spectral peak locations (i.e. frequencies) of the observed periodic signal via the high resolution method ESPRIT [9]. This information is then used as input of a symbolic regression network which unveils the underlying multi-modulation model of the signal. Note that our method's robustness, as will be shown in this paper, comes from the fact that spectral peak locations are unaffected by the transfer function or signal phase distortions. Once the modeling step is performed, the health state of the different machine's parts can be tracked and characterized using source separation algorithms.

2. PROBLEM FORMULATION

Consider a harmonic (vibration) signal y generated through a non-linear mixture (multi-modulation) $\mathcal{F}\colon \mathcal{R}^m \rightarrow \mathcal{R}$ of m periodic (source) signals. This mixture is distorted by the impulse response of the system's transfer function, h, during its transmission to the sensor location. The resulting signal, corrupted by an additive white Gaussian noise ϵ, is given by:

$$x(n) = h(n) \star y(n) + \epsilon(n) \quad (1)$$

where $\star$ denotes the convolution operator. y is obtained by applying the mapping $\mathcal{F}$, assumed to be of polynomial form[1], to a set of m ideally periodic signals $s_{1\leq i\leq m}(n)$ of fundamental frequencies $f_i,\ i = 1, \cdots, m$ as follows:

$$y(n) = \mathcal{F}(s_1(n), \ldots, s_m(n))$$

The fundamental frequencies f_i of the m source signals are assumed to be known (estimated from prior knowledge of the system). The source's Fourier series decomposition is then given by:

$$s_i(n) = \sum_{k=1}^{M_i} a_{i,k} sin(2\pi k f_i/f_s n + \phi_{i,k}) \quad (2)$$

where $\mathbf{a}_i = [\mathbf{a}_{i,k}, \cdots, \mathbf{a}_{M_i,k}]$ and $\phi_i = [\phi_{i,k}, \cdots, \phi_{M_i,k}]$ represent the unknown amplitudes and phases of the $i-$th source signal, respectively, with M_i being the number of its 'significant' harmonics. f_s represents the sampling frequency. Given the observed signal $x(n)$ and leveraging the essential information of the sources fundamental frequencies, we seek to identify the non-linear mixture $\mathcal{F}$ that characterizes the relationship between the source signals $s_i(n)$ and the harmonic input signal $x(n)$ without a need for an a priori estimation or knowledge of the unknown transfer function h.

3. SIGNAL MODELING

In vibration signal modeling for rotating machines, practitioners typically inspect the signal's spectral content manually by examining the locations of harmonics and the available fundamental frequencies of sources (related to system mobile components rotation frequencies, e.g., gears) to identify an analytical expression that provides a good fit for the spectrum lines, which can later be used for source separation and health monitoring [1]. In this work, we improve this process and automate it, as will be shown in the sequel.

3.1. Vibration data

In the Fourier domain, the spectral content $X(f)$ of the signal $x(n)$ can be represented as a sum of H complex sinusoids $S_k(n) = e^{j2\pi h_k n}$ and their complex conjugates $\bar{S}_k(n)$. h_k represents the unknown harmonic frequencies for $k = 1, \cdots, H$. We assume that N samples are available from a noisy measurement $x(n)$ expressed as:

$$x(n) = \sum_{k=1}^{H} \alpha_k S_k(n) + \bar{\alpha}_k \bar{S}_k(n) + \epsilon(n), n = 1, \ldots, N \quad (3)$$

[1]Indeed, the multi-modulation phenomenon results into a signal which can be expressed as a linear combination of products of source signals.

where $\alpha_{1\leq k\leq H}$ are the complex amplitudes of $\mathbf{S}_k$. We seek to regenerate a new signal $\tilde{x}(n)$ from $x(n)$ such that its spectral content $\tilde{X}(f)$ has a unit amplitude for all harmonic peaks. This approach eliminates the effects of the transfer function, which typically disrupt the amplitude symmetry of the sidebands of the carrier signal. Thus, we intend for estimating the harmonics frequencies h_k to construct the new signal defined as:

$$\tilde{x}(n) = \sum_{k=1}^{H} S_k(n) + \bar{S}_k(n), \ \ n = 1, \ldots, N \quad (4)$$

This is achieved by applying a classical frequency estimation technique to $x(n)$, namely ESPRIT (Estimation of Signal Parameters via Rotational Invariance) [9]. Following this, we propose a neural network-based symbolic regression method to approximate the nonlinear mapping $\mathcal{F}$ using the new signal $\tilde{\mathbf{x}}$.

3.2. Regression

Consider multivariate regression with a dataset defined as, $\{(\mathbf{D}_n, \tilde{\mathbf{x}}(n))\}_{1\leq n\leq T/N_b}$, where $\tilde{\mathbf{x}}(n) = [\tilde{x}((n-1)N_b + 1), \cdots, \tilde{x}(nN_b)]^T$ (N_b being a processing window and N is the total sample size). $\mathbf{D}_n$ is a dictionary given by $\mathbf{D}_n = [\mathbf{D}_{n,1}, \cdots, \mathbf{D}_{n,m}]^T$, where $\mathbf{D}_{n,1\leq i\leq m}$ is formed by vectors $\mathbf{d}_{k,i}$ of indices $k = (n-1)N_b + 1, \cdots, nN_b$, defined as:

$$\mathbf{d}_{k,i} = [e^{j2\pi f_i k}, e^{-j2\pi f_i k} \cdots e^{j2\pi H_{sup} f_i k}, e^{-2\pi H_{sup} f_i k}]^T$$

According to (2), $s_i(n)$ is a linear combination of the rows of $\mathbf{D}_{n,i}$, i.e. $\mathbf{s}_i(n) = \mathbf{w}_i^T \mathbf{D}_{n,i}$ where $\mathbf{s}_i(n) = [s_i((n-1)N_b + 1), \cdots, s_i(nN_b)]$, $\mathbf{w}_i$ is the vector of weight coefficients of source i, and H_{sup} is an overestimated harmonics number of $M_{1\leq i\leq m}$.

We seek to recover the sources (i.e. vectors $\mathbf{w}_i$) and an approximation ψ for the analytical function $\mathcal{F}$ that incorporates the selection of the active harmonics in each $\mathbf{D}_{n,i}$. This is achieved by minimizing the quadratic error defined as $\frac{1}{B}\sum_{b=1}^{B} \frac{1}{N_b}\sum_{k=1}^{N_b}(|\tilde{\mathbf{x}}(n) - \psi(\mathbf{w}_1^T\mathbf{D}_{n,1}, \cdots, \mathbf{w}_m^T\mathbf{D}_{n,m})|)^2$,where B denotes the number of data windows (batches). A Binary Neural network-based symbolic regression is proposed to achieve this objective.

Symbolic regression (SR) is a machine learning technique that identifies an analytical expression which describes a given dataset. SR is a difficult task that typically involves a two-step procedure: predicting the "skeleton" of the expression up to the choice of numerical constants, then fitting the constants by optimizing a non-convex loss function. It is often implemented through genetic programming (GP), which searches through the space of mathematical expressions while ensuring that the equation is viable through various heuristics [10]. However, GP do not scale well to high-dimensional problems. Neural networks have recently been successfully

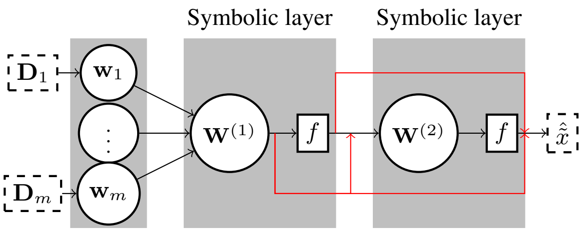

Fig. 1. Architecture of EQL network. The linear output layer is not shown here. Red lines represents skip connections.

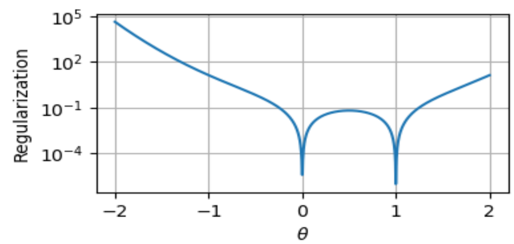

Fig. 2. Illustration of the proposed function to achieve binary neural network weights.

tasked with predicting the correct skeleton, motivated by the fact that neural networks are good at identifying qualitative patterns. Authors in [11] propose modifications to feed-forward networks to include interpretable components, i.e. replacing usual activation functions with operators such as multiplication, addition, identity, sine and cosine.

3.3. EQuation Learner (EQL) architecture

The EQL is a fully connected Neural Network (NN) that can perform symbolic regression by replacing the nonlinear activation functions with primitive functions [12], [11]. The output of the $i-th$ symbolic layer in Fig.1 can be described by

$$\mathbf{g}^{(i)} = \mathbf{W}^{(i)}\mathbf{h}^{(i-1)} + \mathbf{b}^{(i)} \tag{5}$$

$$\mathbf{h}^{(i)} = f(\mathbf{g}^{(i)}) \tag{6}$$

where $\mathbf{W}^i$ and $\mathbf{b}^{(i)}$ are the weight matrix for the $i-th$ symbolic layer and the bias vector, respectively. f represents the neural network primitive (activation) functions, and $\mathbf{h}^{(0)}$ is the output of the first layer. The activation function for the final layer is typically linear, so the output of the neural network with L hidden layers is $\hat{\tilde{x}} = \mathbf{W}^{(L+1)}\mathbf{h}^{(L)}$. In contrast, the EQL network utilizes a collection of primitive functions[2], with each component of $\mathbf{g}$ potentially undergoing a distinct primitive function and allowing for primitive functions to accept multiple inputs. In our case, the nodes in the first layer (Fig.1), are grouped into m weight vectors, $\mathbf{w}_{1\leq i\leq m}$, which are connected to each dictionary $\mathbf{D}_{1\leq i\leq m}$. This results in inputs for the first symbolic layer with a dimension equal to the number of sources, defined as $\mathbf{h}^{(0)}(n) = [\mathbf{w}_1^T\mathbf{D}_{n,1}, \cdots, \mathbf{w}_m^T\mathbf{D}_{n,m}]^T$. The training of the neural network of $L = 2$ (a hyperparameter) symbolic layers, and parameters

$$\theta = \{\mathbf{w}_1, \cdots, \mathbf{w}_m, \mathbf{W}^{(1)}, \cdots, \mathbf{W}^{(L+1)}, \mathbf{b}^{(1)}, \cdots, \mathbf{b}^{L+1}\}$$

is conducted using established methods akin to those used for conventional NNs, i.e. Adam [13].

[2] Since we know that the vibration signal has a polynomial form, we limit the primitive operators to multiplication ($\times$), addition ($+$), and the identity operator ($\mathbf{1}$).

3.4. Network interpretability

To ensure the interpretability of SR, we need to force the system to learn the simplest expression that describes the dataset. In genetic programming-based approaches, this is typically done by limiting the number of terms in the expression. For the EQL network, authors in [11] enforce this through sparsity regularization of the network weights, such that as many of the EQL weights are set to 0. Works [11, 12, 14] leverage a smoothed $\mathcal{L}_{0.5}$ or a relaxed $\mathcal{L}_0$, or simply $\mathcal{L}_1$ norm. In this work, we propose a new way to identify the simplest expression from the Neural network by constraining its weights, θ, to binary values from the set $\{0, 1\}$. To achieve this, we introduce a regularization term that minimizes a function with two possible solutions, 0 or 1. The latter, an elementwise function, is defined as:

$$\mathcal{L}(\theta) = \sinh^2(\theta(1-\theta)) \tag{7}$$

where θ represents the weight coefficients of the neural network, and $\sinh$ is the hyperbolic sine function defined as $\sinh(x) = \frac{\mathbf{e}^x - \mathbf{e}^{-x}}{2}$. Also, we add skip connections, similar to [11], to the EQL network to introduce an inductive bias towards simpler equations while simultaneously enabling the discovery of more complex equations. These skip connections enable us to bypass intermediate layers and connect directly to the output. We concatenate the output of the previous layer with that of the next layer (see Fig.1). ψ is a fully differentiable neural network with parameters θ, which allows us to train it in an end-to-end manner with back propagation. The objective is LASSO-like [15], defined as:

$$J(\theta) = \frac{1}{B}\sum_{n=1}^{B} ||\psi_\theta(\hat{\mathbf{s}}_1(n), \cdots, \hat{\mathbf{s}}_m(n)) - \tilde{\mathbf{x}}(n)||^2 + \lambda \sum_{l}^{L+1+m} |\mathcal{L}(\theta^{(l)})| \tag{8}$$

where λ is a penalty coefficient and $\hat{\mathbf{s}}_i(n) = \mathbf{w}_i^T\mathbf{D}_{n,i}$. In training, we start with non-regularized phase, $\lambda = 0$, to give the NN a good starting point, then the regularization is enabled with a nonzero λ value. The regularizer leads to a trade-off between minimizing the reconstruction loss and enforcing sparsity by binarizing the Neural network's weights (Fig.2).

The final model is selected on the basis of a balance between the accuracy of the reconstruction and the simplicity of the extracted expression.

4. RESULTS

We will now examine the case of widely used rotating machines in industries, such as planetary gearboxes. Vibration data collected from these machines exhibit highly structured spectral content. The vibration data of a planetary gearbox can be modeled as a multi-modulation process involving the meshing frequency and the rotation frequencies of the gearbox's mobile elements. This data is further distorted by transmission path effects, denoted as "h", which occur between the meshing points and the sensor location, as follows:

$$x(n) = h \star \sum_{p=1}^{P} s_{1,p}(n)(1 + s_{2,p}(n))(1 + s_{3,p}(n)) \quad (9)$$

where P represents the planets number of the planetary gearbox. We aim to identify the multi-modulation model in (9). For that, we generate the signal in (9) with $P = 1$, $f_1 = 99Hz$, $f_2 = 3Hz$, $f_3 = 8.48Hz$, and the signal harmonics are set to $M_1 = 2$, and $M_{2,3} = 4$. The sampling frequency is $fs = 1000Hz$, the overestimated harmonics number H_{sup} is set to 5, and the signal length is 300 $seconds$. The ESPRIT algorithm is applied with an autocorrelation matrix size of 1000. For comparison, we used the method in [8] where the NN is trained using the Adam optimizer with a learning rate of $lr = 10^{-3}$ and a batch size of $N_b = 256$. The training process utilizes the raw signal $x(n)$ over a 60-second duration, incorporating a smoothed sparsity regularization term $\mathcal{L}^*_{0.5}$. The objective function is given by $J_1(\theta) = \frac{1}{B}\sum_{n=1}^{B} ||\psi_\theta(\hat{\mathbf{s}}_1(n), \cdots, \hat{\mathbf{s}}_m(n)) - \mathbf{x}(n)||^2 + \lambda\mathcal{L}^*_{0.5}(\theta)$, where λ is the hyperparameter penalty coefficient. After selecting the hyperparameter penalty coefficient that achieves the best trade-off between reconstruction term and parameter θ sparsity, the signal $x(n)$ is reconstructed, as shown in Fig. 3. The figure shows that the neural network output does not perfectly fit all peaks, and their magnitudes are not accurately matched due to the transfer function effects on the raw signal. One of the expressions identified after thresholding neural network's weights is as follows:

$$x(n) = 6.52036 s_1 \cdot (0.085 + 5.22559 \cdot s_2 \cdot s_3)$$

The determined mathematical expression shows only some of the terms in model (9), while others related to modulation between s_1 and s_2 or s_3 are missing. The transfer function effects during the search for the data model leads in this case to an erroneous modelization. To avoid such issues, we now optimize the EQL network by minimizing the objective function in (8) with the new signal $\tilde{\mathbf{x}}$. Then, we apply a hard threshold of 0.001 to the neural network weights to obtain a binary neural network. For a Signal-to-Noise Ratio (SNR) of 20 dB, the accuracy of the estimated frequencies using the ESPRIT

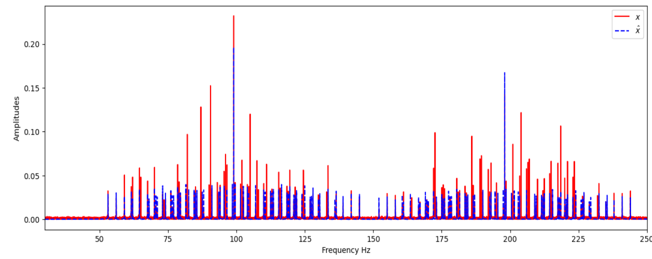

Fig. 3. Reconstruction of the raw signal x with a SNR of 10 dB.

method given by the normalized mean squares error are of the order of 10^{-7}. The resulting expression is:

$$\hat{\tilde{x}}(n) = 0.99 s_1(n) \cdot (0.99 + 0.99 s_2(n)(0.99 + 0.99 s_3(n))$$

The harmonic sources s_i for $1 \leq i \leq 3$ are well selected, and the reconstructed signal is shown in Fig. 4, demonstrating a good match between the power spectrum peaks. For a

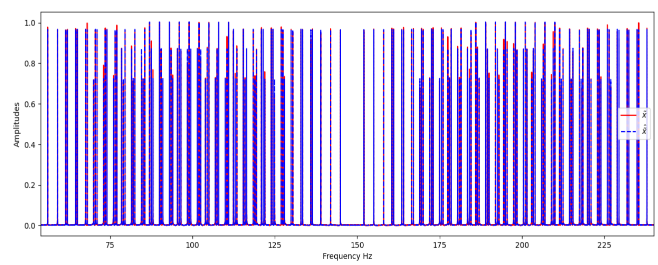

Fig. 4. Reconstruction of the new signal $\tilde{\mathbf{x}}$

lower SNR set to 5 dB, with an ESPRIT algorithm estimation accuracy of order 10^{-4}, the data model is determined as:

$$\hat{\tilde{x}}(n) = 0.93 s_1(n) \cdot (1.01 + 0.95 s_2(n)(1.01 + 0.94 s_3(n)).$$

The accuracy of the raw signal's harmonic frequencies does not affect the model selection, as the parameters θ still converge to the set $\{0, 1\}$. One of the original outcomes of our method is that, in addition to identifying the data model, we were able to estimate the number of harmonics of the elementary sources using the fit weights from the first layers. These outcomes could be vital for the initialization of source separation algorithms, and even tracking the amplification of the harmonic numbers for each source could provide a health indicator for the corresponding components.

5. CONCLUSION

In this paper, we introduce a framework for modeling vibration signals of rotating machines. Our approach improves upon previous work by eliminating distortions that affect the original input signal, such as transfer functions. We leverage a fully differentiable NN with symbolic layers, which is trained with a regularizer to enable the binarization of its weights for a compact model that describes the given data set. Through a synthesized planetary gearbox vibration signal, we demonstrate that our approach leads to a concise functional form that provides insights into the relationship between the entry sources.

6. REFERENCES

[1] Elisa Hubert, *Vibration monitoring of an aeronautic power transmission system*, Ph.D. thesis, Université de Lyon, 2019.

[2] RB Randall, "A new method of modeling gear faults," 1982.

[3] Murat Inalpolat and Ahmet Kahraman, "A theoretical and experimental investigation of modulation sidebands of planetary gear sets," *Journal of sound and vibration*, vol. 323, no. 3-5, pp. 677–696, 2009.

[4] Elisa Hubert, Axel Barrau, and Mohammed El Badaoui, "New multi-carrier demodulation method applied to gearbox vibration analysis," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 2141–2145.

[5] Yosra Marnissi, Yasmine Hawwari, Amadou Assoumane, Dany Abboud, and Mohamed El-Badaoui, "On the use of structured prior models for bayesian compressive sensing of modulated signals," *Applied Sciences*, vol. 11, no. 6, pp. 2626, 2021.

[6] Wanzhi Qiu K. Abed-Meraim and Yingbo Hua, "Blind system identification," *Proceedings of the IEEE*, vol. 85, no. 18, pp. 1310–1322, August, 1997.

[7] Runyu Lu, Md Rifat Shahriar, Pietro Borghesani, Robert B Randall, and Zhongxiao Peng, "Removal of transfer function effects from transmission error measurements using cepstrum-based operational modal analysis," *Mechanical Systems and Signal Processing*, vol. 165, pp. 108324, 2022.

[8] Nacer Yousfi, Karim Abed-Meraim, Yosra Marnissi, Maxime Leiber, and Mohammed El-Badaoui, "Neural network-based symbolic regression for empirical modeling of the behavior of a planetary gearbox," in *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2024, pp. 9866–9870.

[9] A. Paulraj, R. Roy, and T. Kailath, "Estimation of signal parameters via rotational invariance techniques- esprit," in *Nineteeth Asilomar Conference on Circuits, Systems and Computers, 1985.*, 1985, pp. 83–89.

[10] John R Koza, "Genetic programming as a means for programming computers by natural selection," *Statistics and computing*, vol. 4, pp. 87–112, 1994.

[11] Samuel Kim, Peter Y Lu, Srijon Mukherjee, Michael Gilbert, Li Jing, Vladimir Čeperić, and Marin Soljačić, "Integration of neural network-based symbolic regression in deep learning for scientific discovery," *IEEE transactions on neural networks and learning systems*, vol. 32, no. 9, pp. 4166–4177, 2020.

[12] Georg Martius and Christoph H Lampert, "Extrapolation and learning equations," *arXiv preprint arXiv:1610.02995*, 2016.

[13] P Kingma Diederik, "Adam: A method for stochastic optimization," *(No Title)*, 2014.

[14] Michael Zhang, Samuel Kim, Peter Y Lu, and Marin Soljačić, "Deep learning and symbolic regression for discovering parametric equations," *IEEE Transactions on Neural Networks and Learning Systems*, 2023.

[15] Robert Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 58, no. 1, pp. 267–288, 1996.