Analysis of Noisy Coefficients in the Discrete Hermite Transform Domain with Application in Signal Denoising and Sparse Signal Reconstruction

Miloš Brajović[1], Srdjan Stanković, and Irena Orović

University of Montenegro/Faculty of Electrical Engineering
20 000, Podgorica, Montenegro
phone: + (382) 20 245 839, fax: + (382) 20 245 873

Abstract

The influence of an additive Gaussian noise on the discrete Hermite transform based signal representation is analyzed. The Hermite coefficients of noisy signals are random Gaussian variables. Based on the derived respective mean values and the variances, an efficient nonlinear threshold for a simple signal denoising approach is introduced, suitable for signals well concentrated in this transform domain. Moreover, the results are easily incorporated into a coefficient thresholding based compressed sensing algorithm for the reconstruction of noisy signals with missing samples. These approaches and the theory behind are motivated by the signals concentrated in the Hermite transform domain, such as the QRS complexes and UWB signals. Numerical examples validate the presented theory.

Keywords: Hermite functions, Hermite transform, noise modelling, denoising, QRS complexes, sparse signal reconstruction.

1. Introduction

Recently, the Hermite transform (HT), in both continuous and discrete forms, has drawn a significant research interest in various practical applications [1]-[9], showing a better performance compared to the commonly used Fourier-based approaches. For instance, the HT is extensively applied in the processing and compression of ECG signals, as well as in the automatic recognition and classification of QRS complexes [1]-[9]. Other important applications include biomedicine [4], [5], image processing and computer tomography [2], [9], molecular biology [2], radar signal processing [10], etc. In the light of popular compressive sensing scenarios, the HT was also considered as a domain of signal sparsity [11], [12]. Therein, it has been shown that signals, exhibiting sparsity in the HT domain, can be efficiently reconstructed from a small set of random measurements. Besides many application examples, Hermite functions (HF) have been the subject of extensive research in the context of the discrete fractional Fourier transform [13]- [30]. These Hermite-Gaussian like functions, being closed-form Discrete Fourier Transform (DFT) eigenvectors used to define the discrete fractional Fourier transform, can be also used to define the HT. Various approaches have been reported in this context [13]: methods based on nearly tridiagonal matrices [14]-[23], methods based on orthogonal projections [24]-[29] as well as methods based on closed-form vectors, [9], [30].

[1] Corresponding author, phone: +38269486639, Email: milosb@ac.me

An interesting example of signals with suitable representation in the HT domain can be found in communications and remote sensing applications, namely, the ultra-wideband (UWB) signals received at antennas. Signals from this class typically have the Gaussian waveforms, such as Gaussian doublets or other shapes resembling the derivatives of the Gaussian function [31]-[34], that exhibit a compact support in the HT domain. The UWB signals are characterized by an inherent fine resolution in time as well as by good penetration into many common materials and therefore, these signals are widely used in the remote sensing applications [33]-[37]. However, the presence of noise, that is common in real scenarios, often leads to the performance degradation of algorithms dealing with UWB signals [35], [36], thus making the denoising techniques especially important. Having in mind the waveform similarity between the UWB signals and the Hermite basis, the UWB applications, including the denoising, may benefit from the HT [33]. HT has been continuously related to QRS complexes, particularly important parts of ECG signals. Recently, it was shown that the compressed sensing reconstruction of these signals is possible exploiting the Hermite transform [11]. Moreover, HT based compression of QRS complexes has been a widely studied topic, and the literature suggests the amenability of this particular transform [1], [8]. However, neither the HT based compression nor the compressed sensing has been previously studied for signals affected by common white Gaussian noise.

In this paper, the influence of an additive white Gaussian noise (AWGN) on the discrete HT is considered. This is particularly important, as the discrete HT form ordinarily assumes the Gauss-Hermite quadrature applied in the accurate numerical calculation of integrals defining Hermite coefficients, which inevitably leads to a different behavior of this transform in noisy environments when compared with common linear transforms [2], [10], [11]. We also discuss an alternative form of discrete Hermite transform, recently studied in [5], [6], [32]. The influence of AWGN is modeled through the expressions for the mean value and variance of noisy Hermite coefficients. Based on these expressions, a nonlinear threshold which separates signal components from the noise is derived. It enables an efficient signal denoising. Moreover, we incorporate the results into the compressed sensing algorithm based on the detection of signal support ([38], [39]), leading to the successful reconstruction of noisy signals with missing samples.

The paper is organized as follows. A basic theoretical background about Hermite expansion method using the Gauss-Hermite quadrature is presented in Section 2. The matrix form of the discrete Hermite transform is presented in Section 3. The influence of AWGN on the HT domain is analyzed in Section 4, while the thresholding method and denoising applications are presented in Section 5. The presented theory is put into the compressive sensing and sparse signal processing context in Section 6. The concluding remarks are provided in Section 7.

2. Hermite expansion

The p-th order Hermite function is defined as:

$$\begin{aligned}\psi_p(t,\sigma) &= \left(\sigma 2^p\, p!\sqrt{\pi}\right)^{-\frac{1}{2}} e^{\frac{-t^2}{2\sigma^2}} H_p(t/\sigma) \\ &= \left(\sigma 2^p\, p!\sqrt{\pi}\right)^{-\frac{1}{2}} e^{\frac{-t^2}{2\sigma^2}} (-1)^p e^{\frac{t^2}{\sigma^2}} \frac{d^p(e^{-\frac{t^2}{\sigma^2}})}{dt^p},\end{aligned} \tag{1}$$

with $H_p(t)$ being the p-th order Hermite polynomial (HP). A scaling factor σ which stretches or compresses HFs is often used in the definition (1) in order to match the analyzed signal [1], [5], [31]. Without loss of generality, it is assumed that it has a value $\sigma = 1$ [1], [11], and can be omitted in further notations. Hermite basis functions can be recursively calculated [1]-[3]:

$$\begin{aligned}&\psi_0(t) = \frac{1}{\sqrt[4]{\pi}} e^{-t^2/2}, \qquad \psi_1(t) = \frac{\sqrt{2}t}{\sqrt[4]{\pi}} e^{-t^2/2}, \\ &\psi_p(t) = t\sqrt{\frac{2}{p}}\,\psi_{p-1}(t) - \sqrt{\frac{p-1}{p}}\,\psi_{p-2}(t).\end{aligned} \tag{2}$$

The signal representation using the Hermite basis is referenced as the Hermite expansion:

$$s(t) = \sum_{p=0}^{\infty} C(p)\psi_p(t)\,. \tag{3}$$

For a continuous signal $s(t)$, an infinite number $M \to \infty$ of HFs is needed for an accurate signal representation. Otherwise, (3) is just an approximation of the analyzed signal. The p-th order Hermite coefficient $C(p)$ is calculated by:

$$C(p) = \int_{-\infty}^{\infty} s(t)\psi_p(t)dt\,. \tag{4}$$

2.1 Discrete Hermite transform based on the Gauss-Hermite quadrature approximation

The discrete Hermite transform can be considered as a discretized version of the continuous-time Hermite expansion (3). Namely, if HFs are sampled at zeros of the M-th order HP, then the summation (3) becomes the inverse form of the Hermite transform, [1]. In that case, any discrete signal of length M can be uniquely represented by the expansion (3), with a complete set of M discrete basis functions. The integral of the form (4) can be accurately calculated by the Gauss-Hermite (GH) quadrature [3], [8], [10]:

$$C(p) = \frac{1}{M}\sum_{m=1}^{M} \frac{\psi_p(t_m)}{\left[\psi_{M-1}(t_m)\right]^2} s(t_m), \tag{5}$$

where $t_m, 1 \le m \le M$ is used to denote zeros of the M-th order HP [1]-[11]. Functions $\psi_p(t_m),\ 1 \le m \le M,\ 0 \le p \le M-1$ obtained sampling the continuous HFs at points t_m are orthogonal [1]. When the sampling points are proportional to t_m, the transform (5) is a complete signal representation [1]-[2], [11]. This form of the discrete Hermite transform will be further denoted as DGHmT.

Note that a proper calculation of (5) requires a specific form of sampling. However, as indicated in our previous research [8], when the continuous signal $s(t)$, with assumption of a compact time support, $s(t) = 0$ for $t \notin [-T, T]$, is sampled

uniformly to obtain the corresponding finite duration discrete-time signal $s(n) = s(n\Delta t)$, then the discrete signal values at desired points t_m (or at proportional points λt_m where constant λ is directly related with the scaling factor σ (dilation parameter)), are obtained applying a resampling procedure. Without loss of generality, assume an odd signal length $M = 2K+1$, $n = -K,\ldots, K$, with Δt being the sampling period. The continuous-time signal can be reconstructed and resampled at the desired points $\lambda t_1, \lambda t_2,\ldots, \lambda t_M$ using [8]:

$$s(\lambda t_m) \approx \sum_{n=-K}^{K} s(n\Delta t) \frac{\sin\left(\pi(\lambda t_m - n\Delta t)/\Delta t\right)}{\pi(\lambda t_m - n\Delta t)/\Delta t}, \tag{6}$$

where $m = 1,\ldots, M$, $n = -K,\ldots, K$.

The truncation error [40], [41] using sinc interpolation is largest for time instants near the edges of the considered discrete grid. However, in the case of compact time-support signals, the truncation error will be significantly reduced. Finite signals interpolation problem is additionally discussed from the perspective of FIR filter-based sinc interpolation in [40], where it is emphasized that the truncation effects could be alleviated by multiplying the interpolation kernel $\sin\left(\pi(t-n\Delta t)/\Delta t\right)/(\pi(t-n\Delta t)/\Delta t)$ with a window function. The interpolation error has been derived in [42]. Recently, it has been confirmed that this error is negligible for signals with finite-time support, [8].

2.2 Discrete Hermite transform based on symmetric tridiagonal matrix that commutes with a centered Fourier matrix

The previous form of the discrete Hermite transform has been considered in many different application contexts [8], [10], [11], [12]. Recently, it has been successfully applied in the compression of QRS complexes [8], and in the reconstruction of compressively sensed UWB signals [51]. The highly successful applicability is related to the strong resemblance of these basis functions with analyzed signals. However, as it will be shown in next Section, the DGHmT matrix based on (5) is not orthogonal. As it is generally known that the sampling of continuous HFs does not lead to a compatible discrete orthogonal basis, significant research efforts have been made to determine other approaches for discrete Hermite functions definition [5], [6], [32]. For example, it has been shown that discrete HFs can be generated as eigenvectors of a centered or shifted Fourier matrix [5], [6]. These alternative HFs will be further denoted as $\tilde{\psi}_p(n,\sigma)$ when $\sigma = 1$ and $\tilde{\psi}_p(n)$ in the case when time-axis scaling factor $\sigma \geq 1$ (dilation parameter) is assumed. Note that n denotes discrete time index at a uniform time grid related with the sampling theorem. Other alternative approaches for the discrete HFs calculation can be also found in [13]-[30].

Observe a discrete-time signal $s(n)$ of length N, sampled according to the sampling theorem, with $0 \leq n \leq M-1$. For such signal, there exist a set of orthogonal digital functions $\tilde{\psi}_p(n,\sigma),\ p = 0, ..., M-1,$ being highly related to continuous HFs $\psi_p(t)$. It was recently shown that such functions are obtained as eigenvectors of the centered Fourier matrix $\boldsymbol{F}_C$ [5], satisfying

$$\mathbf{F}_C\tilde{\psi}_p(n,\sigma) = j^n\tilde{\psi}_p(n,\sigma),$$

where $j=\sqrt{-1}$. The set of functions is obtained based on eigenvalue decomposition

$$\mathbf{F}_C = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T \tag{7}$$

as columns of matrix $\mathbf{Q}=[\tilde{\boldsymbol{\psi}}_0, \ \tilde{\boldsymbol{\psi}}_1, \ \ldots, \ \tilde{\boldsymbol{\psi}}_M]$. Column vectors $\tilde{\boldsymbol{\psi}}_p$, $p=0,...,M-1$ contain values of discrete HFs $\tilde{\psi}_p(n,\sigma)$, $n=0,...,M-1$ and $\mathbf{\Lambda}$ is a diagonal matrix containing eigenvalues of matrix $\mathbf{F}_C$. It was shown that these discrete HFs can be generated in an computationally efficient manner as the set of eigenvectors of a symmetric sparse tridiagonal matrix defined by [5], [6], [32]:

$$\mathbf{T}=\begin{bmatrix} \varphi_0(0) & \varphi_1(1) & 0 & \cdots & 0 \\ \varphi_1(1) & \varphi_0(1) & \varphi_1(2) & \cdots & 0 \\ 0 & \varphi_1(2) & \varphi_0(2) & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \varphi_1(M-1) \\ 0 & 0 & 0 & \varphi_1(M-1) & \varphi_0(M-1) \end{bmatrix}, \tag{8}$$

with

$$\varphi_0(n) = -2\cos\left(\frac{\pi}{\sigma^2}\right)\sin\left(\frac{\pi n}{M\sigma^2}\right)\sin\left(\frac{\pi}{M\sigma^2}(M-1-n)\right),$$

for $0 \le n \le M-1$, and

$$\varphi_1(n) = \sin\left(\frac{\pi n}{M\sigma^2}\right)\sin\left(\frac{\pi}{M\sigma^2}(M-n)\right),$$

for $0 \le n \le M-1$. Eigenvalue decomposition of the form

$$\mathbf{T} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T. \tag{9}$$

leads to the same matrix of eigenvectors as in (7), in a computationally efficient manner [5]. Functions $\{\tilde{\psi}_0(n,\sigma), \tilde{\psi}_1(n,\sigma),..., \tilde{\psi}_{M-1}(n,\sigma)\}$ form an M-dimensional basis of the alternative discrete Hermite transform, further denoted as DHmT. Discrete HFs obtained in this manner are visually very similar to the continuous-time analogous functions. Similar to the continuous case, these functions are non-zero near the interval of the definition, and they are odd or even depending on the value of the index p (which also counts the number of zero-crossings of the signal), just as in the continuous case [5]. As previously noted [6], the difference between the continuous-time and discrete HFs produced by the presented method increases as the index p (order of the HF) increases.

This issue is illustrated in Fig. 1, where some of the DHmT basis functions (left column) are compared with corresponding basis functions of DGHmT (right column), obtained by sampling the continuous HFs at the roots of the Hermite polynomial of order M = 201. For the better comparability, functions are normalized with respect to their maximal amplitudes. For functions with p = 0, 2 and 5 there is a high similarity between $\tilde{\psi}_p(n)$ and $\psi_p(n)$, as presented in Fig 1., first and second rows. For p = 39, the difference becomes more obvious (third row), and it increases, as illustrated for p = 88 and p = 99 (fourth and fifth rows). Fig. 1 emphasizes the DGHmT relevancy for applications requiring a strong similarity with

continuous HFs. The RMSE between DGHmT and DHmT basis functions is shown in Fig. 1 (m), versus HF order p. It is important to note that calculation of the DHmT is numerically more efficient than the calculation of the DGHmT. For example, when M = 201, the average computation time needed for generation of the set of DHmT basis functions is 0.0025 [s], compared to 0.0060 [s] needed to generate the set of DGHmT basis functions (tested on the same computer with Intel(R) Core(TM) i7-6700HQ CPU @ 2.60GHz and 8GB of RAM). For M = 500, these computation times are 0.012 [s] in the first case, and 0.016 [s] in the DGHmT case. However, it should be also noted that DGHmT requires an additional computational burden if the analyzed signal needs to be resampled, according to (6), as well as for the Hermite coefficients calculation based on (5). The accuracy of DHmT basis with respect to continuous-time HFs can be improved by producing a larger set of functions and properly adapting the dilation (time-axis scaling) parameter, [5].

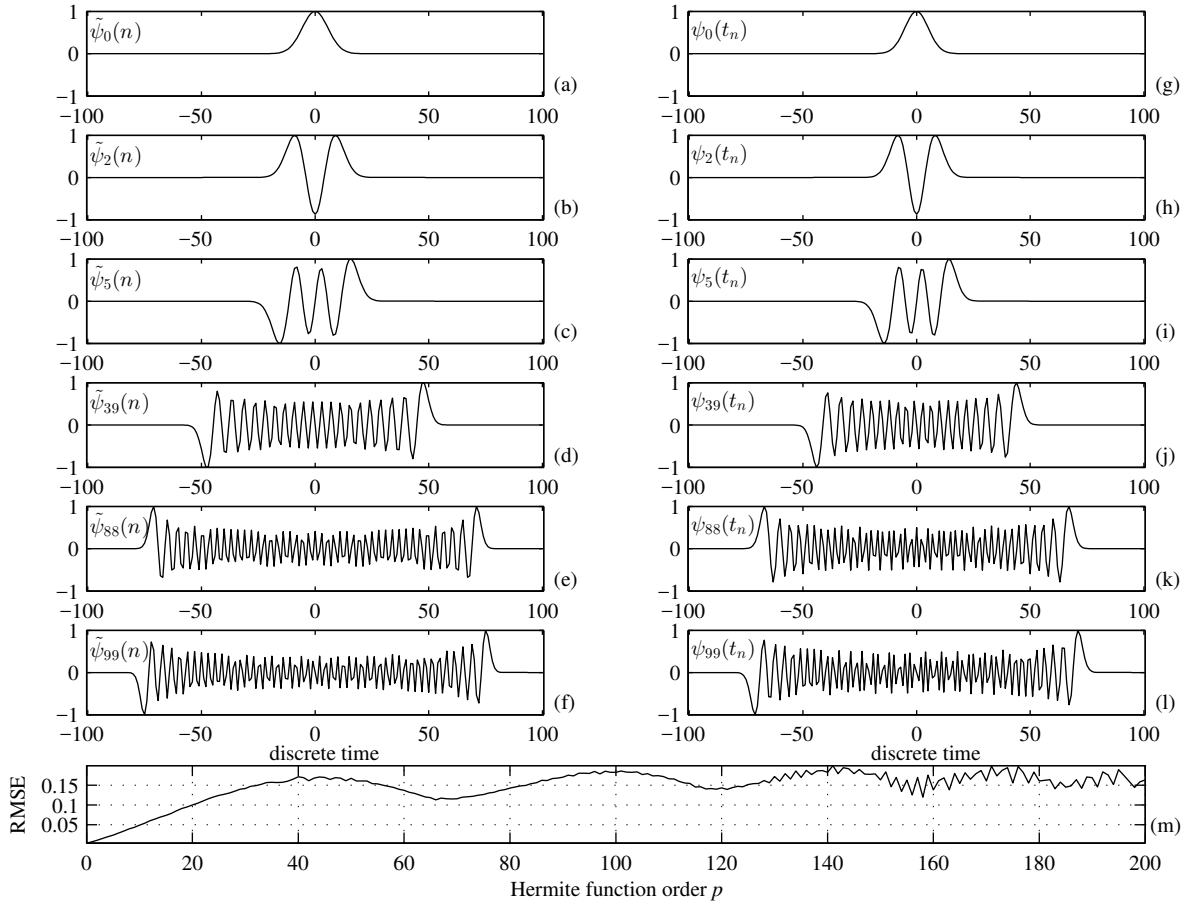

Fig. 1. Examples of: (a)-(f) DHmT and (g)-(l) DGHmT basis functions for M = 201; (m) RMSE between functions

3. Discrete Hermite transforms in matrix forms

The direct and the inverse DGHmT, given by (5) and (3), respectively, can be written in a matrix form. First, we define HT matrix $\mathbf{T}_H$:

$$\mathbf{T}_H = \frac{1}{M}\begin{bmatrix} \frac{\psi_0(t_1)}{(\psi_{M-1}(t_1))^2} & \frac{\psi_0(t_2)}{(\psi_{M-1}(t_2))^2} & \cdots & \frac{\psi_0(t_M)}{(\psi_{M-1}(t_M))^2} \\ \frac{\psi_1(t_1)}{(\psi_{M-1}(t_1))^2} & \frac{\psi_1(t_2)}{(\psi_{M-1}(t_2))^2} & \cdots & \frac{\psi_1(t_M)}{(\psi_{M-1}(t_M))^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\psi_{M-1}(t_1)}{(\psi_{M-1}(t_1))^2} & \frac{\psi_{M-1}(t_2)}{(\psi_{M-1}(t_2))^2} & \cdots & \frac{\psi_{M-1}(t_M)}{(\psi_{M-1}(t_M))^2} \end{bmatrix}, \tag{10}$$

and the corresponding inverse DGHmT matrix $\mathbf{T}_H^{-1}$ (of size $M \times M$)

$$\mathbf{T}_H^{-1} = \begin{bmatrix} \psi_0(t_1) & \psi_1(t_1) & \cdots & \psi_{M-1}(t_1) \\ \psi_0(t_2) & \psi_1(t_2) & \cdots & \psi_{M-1}(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_0(t_M) & \psi_1(t_M) & \cdots & \psi_{M-1}(t_M) \end{bmatrix}. \tag{11}$$

If $\mathbf{C} = [C(0), C(1), \ldots, C(M-1)]^T$ denotes the vector of Hermite coefficients and $\mathbf{s} = [s(t_1), s(t_2), \ldots, s(t_M)]^T$ is a vector of M signal samples, then:

$$\mathbf{C} = \mathbf{T}_H \mathbf{s}, \tag{12}$$

represents the matrix form of the DGHmT. The inverse discrete DGHmT has the following matrix form:

$$\mathbf{s} = \mathbf{T}_H^{-1} \mathbf{C}. \tag{13}$$

The inverse DGHmT matrix can be written as a product

$$\mathbf{T}_H^{-1} = \mathbf{T}_H^T \mathbf{D}, \tag{14}$$

where the matrix **D** is a diagonal matrix whose form is presented in [1], confirming that the discrete HT matrix is not orthogonal. The standard QR decomposition of the DGHmT matrix $\mathbf{T}_H$ leads to the product $\mathbf{T}_H = \mathbf{QR}$ with **Q** being the orthogonal matrix, satisfying $\mathbf{QQ}^T = \mathbf{I}$, with **I** being the identity matrix and the matrix **R** being a diagonal matrix with elements:

$$d_m = (-1)^{m-1}\left[\sqrt{M}\psi_{M-1}(t_m)\right]^{-1},\ m = 1, 2, \ldots, M. \tag{15}$$

Orthogonal matrix **Q**, taking into account (10) has the following form:

$$\mathbf{Q} = \frac{1}{\sqrt{M}}\begin{bmatrix} \frac{\psi_0(t_1)}{\psi_{M-1}(t_1)} & \frac{\psi_0(t_2)}{\psi_{M-1}(t_2)} & \cdots & \frac{\psi_0(t_M)}{\psi_{M-1}(t_M)} \\ \frac{\psi_1(t_1)}{\psi_{M-1}(t_1)} & \frac{\psi_1(t_2)}{\psi_{M-1}(t_2)} & \cdots & \frac{\psi_1(t_M)}{\psi_{M-1}(t_M)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\psi_{M-1}(t_1)}{\psi_{M-1}(t_1)} & \frac{\psi_{M-1}(t_2)}{\psi_{M-1}(t_2)} & \cdots & \frac{\psi_{M-1}(t_M)}{\psi_{M-1}(t_M)} \end{bmatrix}. \tag{16}$$

Due to the form of (14) and (15) where $d_m = 1$ does not hold for every $m = 1, 2, \ldots, M$, we may expect that the common AWGN will influence the DGHmT in a different manner, compared to standard orthogonal transforms, e.g. DFT.

In the case of DHmT, the transform matrix has the following form

$$\tilde{\mathbf{T}}_H = \begin{bmatrix} \tilde{\psi}_0(1) & \tilde{\psi}_0(2) & \dots & \tilde{\psi}_0(M) \\ \tilde{\psi}_1(1) & \tilde{\psi}_1(2) & \cdots & \tilde{\psi}_1(M) \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\psi}_{M-1}(1) & \tilde{\psi}_{M-1}(2) & \cdots & \tilde{\psi}_{M-1}(M) \end{bmatrix}, \tag{17}$$

Let us denote with $\tilde{\mathbf{C}} = [\tilde{C}(0), \tilde{C}(1), \dots, \tilde{C}(M-1)]^T$ the vector of DHmT coefficients and with $\mathbf{s} = [s(0), s(1), \dots, s(M-1)]^T$ a vector of M signal samples obtained according to the sampling theorem. The DHmT can be written as

$$\tilde{\mathbf{C}} = \tilde{\mathbf{T}}_H \mathbf{s}. \tag{18}$$

As it is formed based on eigenvectors of symmetric matrix (9), matrix $\tilde{\mathbf{T}}_H$ is orthogonal, implying that $\tilde{\mathbf{T}}_H^{-1} = \tilde{\mathbf{T}}_H^T$ and $\tilde{\mathbf{T}}_H \tilde{\mathbf{T}}_H^T = \mathbf{I}$. The inverse DHmT has the following form

$$\mathbf{s} = \tilde{\mathbf{T}}_H^T \tilde{\mathbf{C}}. \tag{19}$$

4. Additive Gaussian noise influence

Observe the DGHmT of zero-mean AWGN vector $\boldsymbol{\eta}$. Without loss of generality, it is inherently assumed that the signal and noise samples are available at the discrete points t_m, $m = 1, \dots, M$ corresponding to the roots of the M-th order HP. The discrete DGHmT of noise reads:

$$\boldsymbol{\Xi} = \mathbf{T}_H \boldsymbol{\eta} = \mathbf{Q}\mathbf{R}\boldsymbol{\eta} = \mathbf{Q} \begin{bmatrix} d_1\eta(t_1) & 0 & \cdots & 0 \\ 0 & d_2\eta(t_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & d_M\eta(t_M) \end{bmatrix}. \tag{20}$$

Thus, scaled noise samples $d_m\eta(t_m)$, $m = 1, 2..., M$ are expanded on the orthogonal vector space basis consisted of rows of matrix $\mathbf{Q}$. Note that in the case of DFT matrix, non-scaled noise samples $\eta(t_m)$ appear in $\mathbf{R}\boldsymbol{\eta}$, which means that AWGN differently influences the HT compared to the DFT.

Next, we derive the statistical properties of HT coefficients corresponding to the signal $s(t_m)$ affected by AWGN $\eta(t_m)$:

$$x(t_m) = s(t_m) + \eta(t_m). \tag{21}$$

Since the HT is linear, (21) leads to $X(p) = S(p) + \Xi(p)$, with $S(p)$ and $\Xi(p)$ being the DGHmT of the signal $s(t_m)$ and noise $\eta(t_m)$, respectively. Since the noise $\eta(t_m)$ is a random Gaussian process, due to the central limit theorem and the definition (5), it can be concluded that $\Xi(p)$ and $X(p)$ are also Gaussian random variables. We further analyze statistical properties of the random variable $X(p)$. The mean value can be expressed as:

$$\mu_X(p) = E\{X(p)\} = S(p) + E\{\Xi(p)\}. \tag{22}$$

For a zero-mean additive noise $E\{\eta(n)\} = 0$ we have:

$$\mu_X(p) = S(p) + \frac{1}{M} \sum_{m=1}^{M} \frac{\psi_p(t_m)}{[\psi_{M-1}(t_m)]^2} E\{\eta(t_m)\} = S(p). \tag{23}$$

Let us now calculate the variance of the random variable $X(p)$ with the assumption that the $\eta(t_m)$ is a zero-mean noise:

$$\sigma_X^2(p) = E\left\{\left|X(p)-\mu_X(p)\right|^2\right\} = E\left\{\left|S(p)+\Xi(p)\right|^2\right\} - \mu_X^2(p)$$

$$= \frac{1}{M^2}\sum_{m_1=1}^{M}\sum_{m_2=1}^{M}\frac{\psi_p(t_{m_1})}{\left[\psi_{M-1}(t_{m_1})\right]^2}\frac{\psi_p(t_{m_2})}{\left[\psi_{M-1}(t_{m_2})\right]^2}E\left\{\eta(t_{m_1})\eta(t_{m_2})\right\}.$$

For a white noise with variance σ_η^2 the autocorrelation function is: $r_{\eta\eta}(t_{m_1},t_{m_2}) = E\{\eta(t_{m_1})\eta(t_{m_2})\} = \sigma_\eta^2(t_{m_1}-t_{m_2})$. Hence, $\sigma_X^2(p)$ can be further written in the form:

$$\sigma_X^2(p) = \frac{\sigma_\eta^2}{M^2}\sum_{m=1}^{M}\frac{\psi_p^2(t_m)}{\left[\psi_{M-1}(t_m)\right]^4} = \gamma(p,M)\sigma_\eta^2. \tag{24}$$

This result indicates that the variance of the DGHmT coefficients depends on the coefficient index p. The mean value of the expression (24) can be now derived following the HFs orhonormality [11]:

$$\begin{aligned}\bar{\sigma}_X^2 &= \frac{1}{M}\sum_{p=0}^{M-1}\sigma_X^2(p) = \frac{\sigma_\eta^2}{M^2}\sum_{m=1}^{M}\frac{1}{\left[\psi_{M-1}(t_m)\right]^2}\frac{1}{M}\sum_{p=0}^{M-1}\frac{\psi_p(t_m)\psi_p(t_m)}{\left[\psi_{M-1}(t_m)\right]^2} \\ &= \frac{\sigma_\eta^2}{M^2}\sum_{m=1}^{M}\left[\psi_{M-1}(t_m)\right]^{-2} = \xi(M)\sigma_\eta^2.\end{aligned} \tag{25}$$

The function $\gamma(p,M)$ for M = 100 is shown in Fig. 2a and the function $\xi(M)$ is shown in Fig 1b. The mean value of the function $\gamma(p,M)$ in Fig.1a is 0.2753. It means that first 60 coefficients have the variance (scaled by $\gamma(p,M)$) below this mean value, while the remaining coefficients are scaled by larger values and are more sensitive to the noise. Also, the observed sparse signals are mainly concentrated on the first few DGHmT coefficients [1], [8], indicating that a proper threshold can be used to distinguish between signal and noise components. Fig 2b shows that the mean noise variance decreases as the number of M increases. Using these results, a polynomial fitting can be applied to approximate these functions and to define a general DGHmT-based denoising threshold.

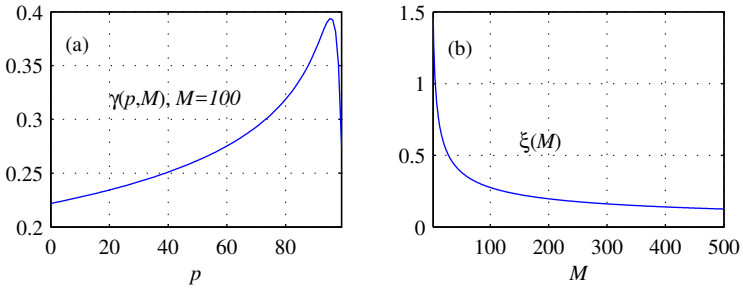

Fig. 2: Functions appearing in variance definitions: a) $\gamma(p,M)$, b) $\xi(M)$

In the case of DHmT, based on the matrix (17) orthogonality, and as the rows are normalized (being eigenvectors of matrix **T**), we conclude that the coefficient variance is

$$\tilde{\sigma}_X^2(p) = \sigma_\eta^2,$$

assuming that noisy samples are available at the uniform discrete time grid in (21).

5. The Hermite transform based denoising

The presented discussion can be applied in signal denoising. The considered signal (21) is affected by a zero-mean AWGN of variance σ_η^2. Note that the noise-only DGHmT coefficient at the position p is a random variable described by the Gaussian distribution $\mathcal{N}(0, \sigma_X^2(p))$. The denoising is done using hard-thresholding procedure [45]:

$$C_{den}(p) = \begin{cases} C(p), & |C(p)| > T(p) \\ 0, & |C(p)| \le T(p). \end{cases} \tag{26}$$

Observe that the threshold is dependent on p, due to the form of variance given by (24). Namely:

$$T(p) = l\sigma_X(p) = l\sqrt{\gamma(p, M)}\sigma_\eta, \tag{27}$$

where l is a constant ensuring that the noise-only coefficients are below the threshold level (27). For many real signals such as UWB signals or QRS complexes, the lower coefficients in the HT are the most significant ones [1], [11]. The non-linear form of the threshold (27) increases the probability for the successful separation of noise-free signal and noise-only Hermite coefficients in the denoising process (Fig. 2a). For $l = 3$, according to the well-known 3-sigma rule, noise-only coefficients are below the threshold with probability of 99.73%. In the case of DHmT, standard threshold $T = l\sigma_\eta$ should be used.

As the presented denoising approach may serve as an alternative to the standard DFT-based denoising, the additional calculation burden is analyzed here. Based on the fast algorithms for the HT calculation [1], [2], the discrete HT requires approximately $O(M \log_2^2 M)$ operations with real values, comparing with $O(M \log_2 M)$ operations with complex values in the case of DFT. The increase of the complexity by $\log_2 M$ times asymptotically is moderate even for a large M. For the threshold calculation in (27), additional $O(M)$ operations are needed.

6.1 Numerical examples and the discussion

Example 1: Consider the part of ECG signal known as QRS complex $qrs(t_m)$, sampled at the points proportional to the HP roots and sparsified according to the procedure in [1], with $M = 51$ samples. It is obtained from the MIT-BIH ECG database [43] (originally available as uniformly sampled in accordance with the sampling theorem) and corrupted with artificial zero-mean AWGN, such that SNR is 6dB. The results for the QRS complex denoising based on DGHmT are shown in Fig. 3a-d, illustrating a significant noise reduction. Parameter $l = 3$ was used in the example. Interestingly, the signal coefficient at $p = 3$, although smaller than the noisy coefficient at $p = 47$, is properly selected by the threshold (being above the threshold), due to the non-linear characteristic of the proposed threshold. Only a few signal coefficients much weaker than the noise remain under the threshold, at $p = 6$, $p = 7$ and $p = 8$. The MSE between the original noise-free and noisy signal of -28.71dB is reduced to -35.83dB after the denoising procedure is applied.

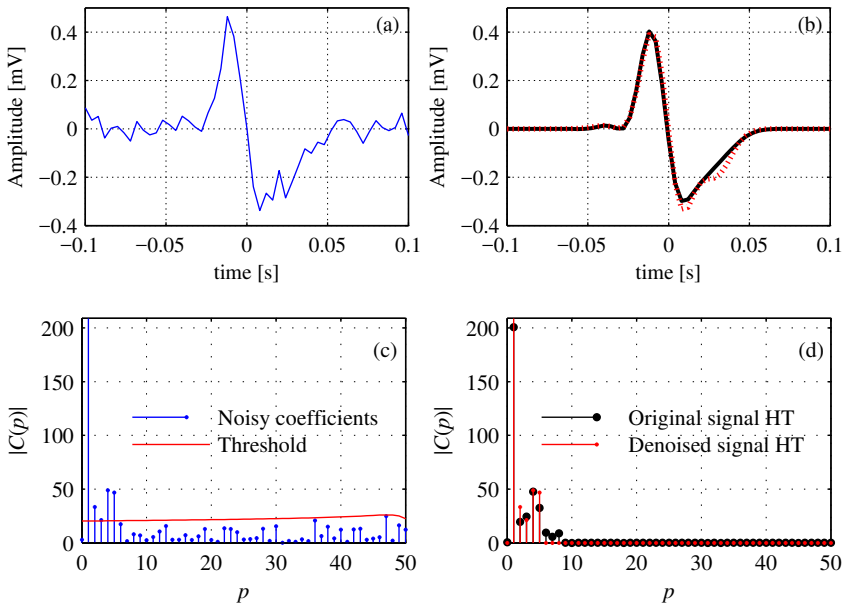

Fig. 3: Denoising a QRS complex using the hard-thresholding procedure: a) QRS complex embedded in AWGN with SNR = 6dB, b) Original QRS complex (solid) and denoised QRS complex (dots), c) DGHmT of noisy QRS complex and the threshold (27) with l = 3, d) DGHmT of original noise-free (circles) and denoised (dots) QRS complex

Now consider the shifted version of the considered noisy signal, $qrs_1(t_m) = qrs(t_m - c)$, where c = 0.028[s]. In this case signal will not be optimally concentrated in the DGHmT domain, as illustrated in Fig. 4a-d. Degradation of the denoising performance is illustrated in Fig. 4 (the MSE of the denoised signal is increased for about 2dB). In this case, prior to the hard-thresholding procedure (26), optimally concentrated DGHmT can be found solving

$$\nu_{opt} = \min_{\nu} \left\| \mathbf{C}_{\nu} \right\|_1 = \sum_{p=0}^{M-1} \left| \frac{1}{M} \sum_{m=1}^{M} \frac{\psi_p(t_m)}{\left[\psi_{M-1}(t_m)\right]^2} s(t_m - \nu) \right|, \tag{28}$$

and shifting the analyzed signal for value $-\nu_{opt}$. Problem (28) can be solved by performing a one-dimensional search over possible values of ν. Upon finding ν_{opt}, the hard-thresholding (26) is applied on the DGHmT of signal shifted in opposite direction $-\nu_{opt}$. We calculated the concentration measure in (28) for values of ν in range [-0.04s, 0.088s] with step 0.004. Results are shown in Fig. 4e. The global minimum of this function corresponds to shift value c = 0.028[s]. DGHmT of the signal $qrs_1(t_m + \nu_{opt})$ leads to the results presented in Fig. 3. Other signal examples suitable for the presented denoising procedure can be also found in [8].

Example 2: In this example, we observe a real 1.3GHz UWB signal transmitted between two UWB antennas 1m apart, through an indoor environment, in experiment described in [34]. First M = 165 samples of the signal 'ACW7FD45.dat' from the database [44] are observed. The artificial AWGN is added to this signal, such that SNR = 3dB. The noisy UWB signal is shown in Fig. 5a, while the corresponding DGHmT with the threshold (27) and l = 4 is shown in Fig. 5c. The original signal

and the denoised signal are shown in Fig. 5b, while the corresponding HTs are shown in Fig. 5d. It can be seen that the noise is significantly reduced. It is assumed that the noise variance σ_η^2 is known since it can be estimated as described in [45].

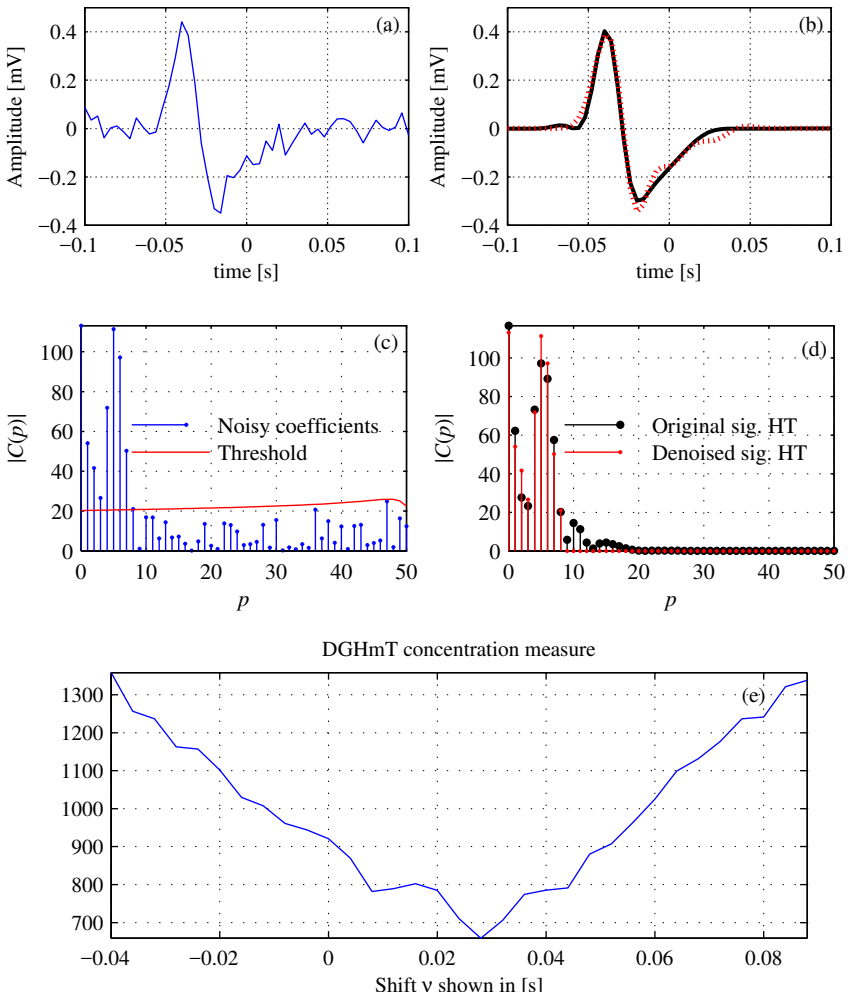

Fig. 4: Denoising a non-centered (shifted) QRS complex using the hard-thresholding procedure: a) QRS complex embedded in AWGN with SNR = 6dB, b) Original QRS complex (solid) and denoised QRS complex (dots), c) DGHmT of noisy QRS complex and the threshold (27) with l = 3, d) DGHmT of original noise-free (circles) and denoised (dots) QRS complex, e) concentration measure in (28) having a minimum at the optimal shift value 0.028 [s]

The non-linear threshold is able to select signal coefficients at p = 6 although having approximately the same value as the noise-only coefficient at p = 154 (that remains under the threshold). Only the smallest coefficients will remain below the threshold, but these are much weaker than the noise and do not contribute significantly to the resulting signal.

For the comparison, the signal denoising using a DFT-based hard thresholding approach is considered (with threshold $T_{\mathrm{DFT}} = lN\sigma_\eta^2$, and the same value l = 4). The results are presented in Fig. 6, showing a significantly degraded denoising performance. The resulting MSE between the original (non-noisy) and denoised signal, for the case of DGHmT-based procedure is -24.23 dB, whereas for the DFT-based procedure the achieved MSE is -16.07dB. The MSE reduction in the first case is approximately 11.8dB, and only 3.63dB in the DFT case.

Example 3: Additionally, previous experiment was conducted for a range of SNR values: from -10 to 15 dB, varied with step 1. For each SNR value, the MSE between the denoised and original signal was calculated based on 500 independent

realizations of artificial AWGN added to the signal. The results for the DGHmT and DFT hard threshold based denoising of the UWB signal from Example 2 are shown in Fig. 7a. The results confirm a significant MSE improvement when using DGHmT and the proposed non-linear threshold.

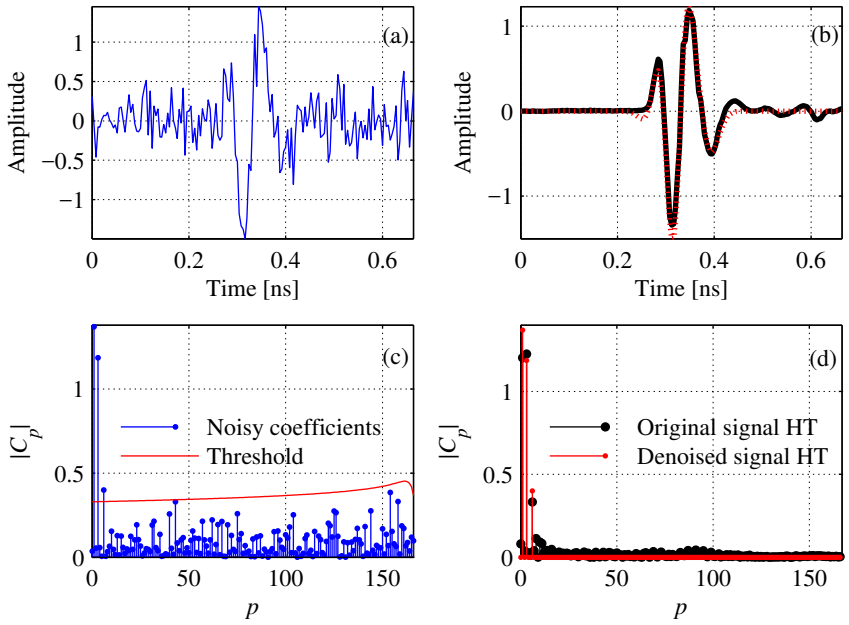

Fig. 5: Denoising a real UWB pulse using the hard-thresholding procedure based on HT: a) UWB signal embedded in AWGN with SNR = 3dB, b) Original (solid) and denoised UWB signal (dots), c) DGHmT of noisy UWB signal and the threshold (27), d) DGHmT of original (circles) and denoised (dots) signal

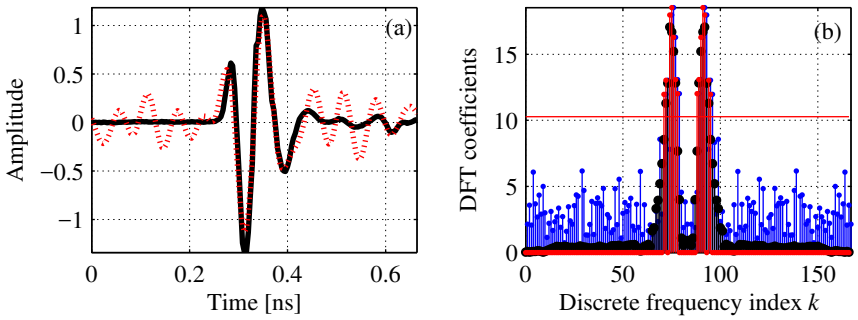

Fig. 6 Denoising a real UWB pulse using DFT-based hard-thresholding: a) original and denoised signal (black and red, respectively), b) DFT of noisy, denoised and original signal (blue, red and black, respectively)

Moreover, the experiment is repeated also with the Discrete wavelet transform (DWT) serving as a basis for signal denoising. We consider denoising based on Symlet 8 (sym8) and Daubechies 8 (db8) wavelets. In both cases decomposition of level 5 is considered. These particular wavelet types are chosen due to their visual similarity with the considered signal. Denoising is performed using two different threshold selection rules to the wavelet coefficients: Stein's Unbiased Risk Estimate (SURE) for db8 wavelet case, and Donoho and Johnstone's universal threshold with level-dependent estimation of the noise for sym8 wavelet [52]-[55]. In each case, the hard thresholding is applied. For this experiment we use MATLAB `wden` implementation from the Wavelet Toolbox. The threshold rescaling is done using a single estimation of level noise based on the first-level coefficients. As seen in Fig. 7a, two considered methods performed similarly. This experiment confirms that DGHmT's ability to represent signal sparsely can be crucial in the denoising, even exploited in very simple

methods such as the hard-thresholding. We additionally test the performance of the Discrete cosine transform (DCT) based hard-thresholding. The threshold level is set to $T_{\mathrm{DCT}} = l\sigma_\eta^2$, and the same value $l = 4$ is used as in the case of the DGHmT. The DGHmT-based method outperformed the DCT-based hard-thresholding (Fig. 7a).

Finally, we include the DHmT-based denoising based on hard-thresholding in our comparative analysis, with $T(p) = l\sigma_\eta$. The same value $l = 4$ is used as for the DGHmT. In both versions of the discrete Hermite transform, dilation (scaling) factor $\sigma = 1$ was used. As shown in Fig. 7a, DGHmT slightly outperforms DHmT. Interestingly, visual similarity of MSE curves in in accordance with the visual similarity of two sets of basis functions.

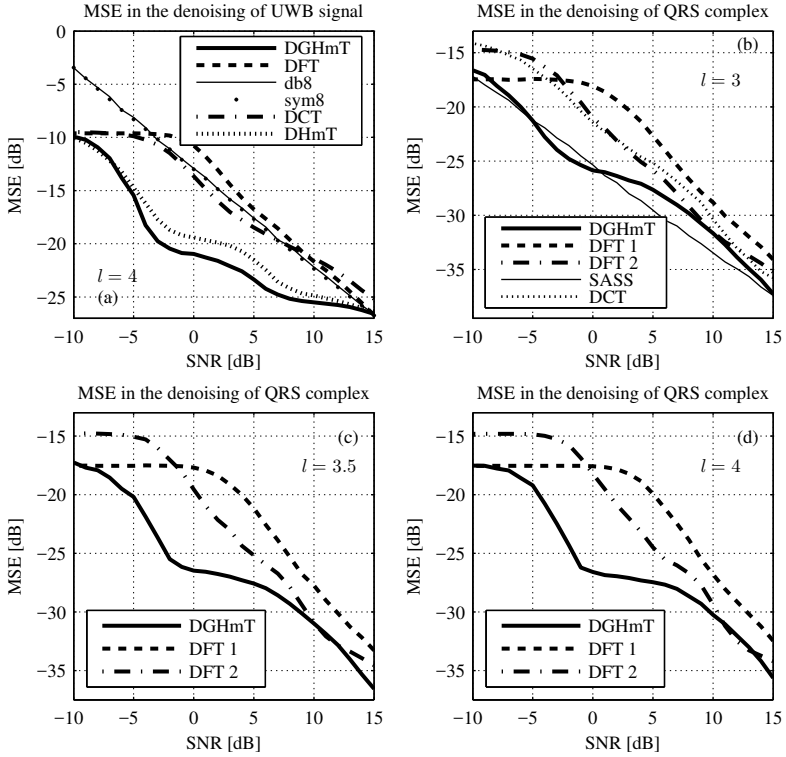

Fig. 7: The MSE between original (noise-free) and the signals denoised using various approaches, shown for various SNRs: (a) UWB signal case; (b)-(d) QRS complex denoising using various parameter l values.

We hypothesize that the performance of DHmT can be improved by setting a proper dilation parameter (time axis scaling factor). Namely, let us denote with $\tilde{\mathbf{T}}_H(\sigma)$ the DHmT matrix, with basis functions $\tilde{\psi}_p(n,\sigma)$, n, $p = 0,...,M-1$. This matrix retains orthogonality [5]. The improved DHmT concentration can be obtained by solving

$$\sigma_{opt} = \min_{\sigma} \left\| \tilde{\mathbf{T}}_H(\sigma)\mathbf{s} \right\|_1 = \min_{\sigma} \sum_{p=0}^{M-1} \left| \sum_{n=0}^{M-1} s(n)\tilde{\psi}_p(n,\sigma) \right|, \tag{29}$$

being a standard ℓ_1-norm minimization from the compressive sensing and sparse signal processing framework. Solution of (29) corresponds to the dilation parameter producing the best possible concentration of DHmT coefficients. It can be found by a direct search in the given range of possible σ values. The same procedure can be performed for DGHmT [8]. However, the further development of this concept and a comprehensive comparative analysis of the performance in the case of these

transforms with optimized parameters is outside the scope of this paper, and it is a topic of our further research. Note that in all presented experiments, for each considered SNR value, the MSE was calculated based on 500 independent realizations of the artificial AWGN added to the signal.

An experiment with a range of SNR values is also performed for the QRS complex from Example 1. Results are presented for various values of parameter l: $l = 3$ (Fig. 7b), $l = 3.5$ (Fig. 7c) and $l = 4$ (Fig. 7d). For a fair comparison, DFT based hard-thresholding was performed for both the original uniformly sampled signal case available in [43] (in Fig. 7b-d denoted as DFT 2) as well as for the non-uniform sampling case (in Fig. 7b-d denoted as DFT 1). In both cases, the same noise was added to signal samples. Results indicate that the DGHmT based denoising produces dominantly lower MSE values in all cases.

For $l = 3$ we compare the denoising performance with the DCT-based hard-thresholding method, where the threshold is set according to the DCT noisy coefficients variance, $T_{\mathrm{DCT}} = l\sigma_\eta^2$. The results shown in Fig 7b indicate the lower MSE obtained using the DGHmT-based method. We also compare the results with an advanced state-of-the-art ECG denoising algorithm, Sparsity-Assisted Signal Smoothing (SASS) [46]. We use the implementation available online, with the original parameter set [47]. The results are presented in Fig. 7b. The presented MSE curve for the SASS method is obtained in the following manner. As this algorithm is originally set to work with entire ECG signals, instead of proceeding as an input only the selected QRS complex, the original full length ECG signal from the MIT-BIH ECG database [43] (part of which is the observed QRS complex) was corrupted with AWGN with same variances as the QRS complex used for other methods evaluation. The denoising of this whole ECG signal was performed using SASS algorithm, with the default set of parameters. Then the QRS complex of interest was extracted from the denoised signal, and compared with the original QRS complex. As for other methods, the experiment was repeated 500 times for each SNR value, and averaged MSE is shown in Fig. 7b. In this way, we were able to use the algorithm without any parameter changes. Results shown in Fig. 7b indicate that this method moderately outperforms the DGHmT-based denoising, in the SNR range from 0dB to 10dB. However, the cost of this improvement is a significantly higher numerical complexity of the SASS algorithm. Nevertheless, the simple hard-thresholding procedure results are comparable with the results of this advanced denoising technique (in the context of QRS complex denoising).

6.2 Statistical validation of the variance

In order to statistically confirm the derived variance (24) and its average value (25), we perform two experiments using noisy UWB signal from Example 2. In the first test, the SNR level of 5 dB is observed (the noise variance $\sigma_\eta^2 = 0.030$). The variance of DGHmT coefficients is numerically evaluated based on 10000 independent realizations of the noise. High match with the theoretical expression (24) is achieved and shown in Fig. 8 (a). The average variance (25) is also shown by dots. In

the second experiment, the input noise variance σ_η^2 is varied from 10^{-6} to 1, with the step 10^{-2}, and the average variance of DGHmT coefficients is calculated. Observe that the results show high matching ratio with the variance (25), Fig 8 (b).

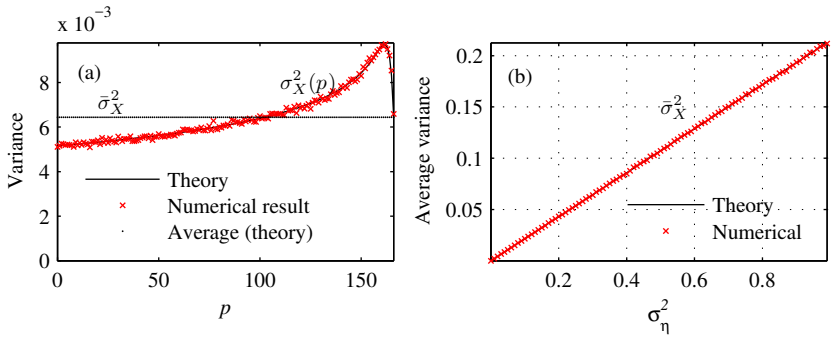

Fig. 8: HT coefficients variance for signal embedded in AWGN: a) Variance of coefficients at position p, SNR = 5 dB, b) Average variance in terms of the input noise variance.

6. Application in sparse signal reconstruction

In this Section, we present an overview of the presented theory applicability in the reconstruction of signals from a random subset of available samples (measurements, observations). To this end, observe a signal exhibiting sparsity in the DGHmT domain. Such signal can be represented with a small number K of non-zero Hermite coefficients, with positions belonging to the set $\mathbf{P} = \{p_1, p_2, ..., p_K\}$. It has the following form:

$$s(t_n) = \sum_{l=1}^{K} \alpha_l \psi_{p_l}(t_n). \tag{30}$$

Sparse signals can be reconstructed from a reduced set of observations [38], [39], [48], [49]. Assume that only M_A out of M randomly positioned samples are available. The available samples have random positions denoted by

$$\hat{t}_m \in \mathbf{M}_A = \{\hat{t}_1, \hat{t}_2, ..., \hat{t}_{M_A}\} \subseteq \mathbf{M} = \{t_1, t_2, ..., t_M\}.$$

Set $\mathbf{M}$ contains the sampling points of the full-length signal, and these points correspond to the roots of the M-th order Hermite polynomial. $\mathbf{M}_A$ is the random subset of $\mathbf{M}$ containing the sampling positions of available measurements. The common mathematical model for the compressive sensing procedure based on the randomly selected/acquired signal values involves the random measurement matrix $\mathbf{\Phi}$:

$$\mathbf{s_{cs}} = \mathbf{\Phi T}_H^{-1}\mathbf{C} = \mathbf{AC}, \tag{31}$$

with $\mathbf{s_{cs}}$ denoting the vector of available samples of the analyzed signal, and $\mathbf{C} = \mathbf{T}_H\mathbf{s}$ is the DGHmT coefficients vector of signal (30) containing all samples. The matrix $\mathbf{A}$ is obtained from the inverse DGHmT matrix $\mathbf{T}_H^{-1}$, by omitting the rows corresponding to the positions of missing samples. Set of M_A linear equations (31) with M unknowns can be solved if the additional sparsity condition is assumed for the solution. Therefore, sparse signal reconstruction problem is formulated as

$$\min \|\mathbf{C}\|_0 \text{ subject to } \mathbf{s_{cs}} = \mathbf{AC}, \tag{32}$$

where the so called ℓ_0-norm corresponds to the number of non-zero coefficients in **C**. It is known that the ℓ_0-norm cannot be used in the direct minimization and thus the problem (32) is usually reformulated using ℓ_1-norm, exploited in the application of efficient linear programming and iterative approaches. Indirectly, problem (32) can be solved by properly estimating the positions of non-zero DGHmT coefficients in the solution. If the signal support is known or appropriately estimated within a set $\hat{\mathbf{P}}$ containing $K \le \hat{P} \le M$ elements such that $\mathbf{P} \subseteq \hat{\mathbf{P}}$, the reconstruction is done using the pseudo-inversion [39]:

$$\mathbf{C}_K = \left(\mathbf{A}_{\hat{\mathbf{P}}}^T \mathbf{A}_{\hat{\mathbf{P}}}\right)^{-1} \mathbf{A}_{\hat{\mathbf{P}}}^T \mathbf{s}_{cs}, \tag{33}$$

often used in standard matching pursuit approaches, for example, in the OMP algorithm. The matrix $\mathbf{A}_{\hat{\mathbf{P}}}$ is the sub-matrix of matrix **A**, with omitted columns corresponding to positions $p \notin \hat{\mathbf{P}}$.

If samples are omitted from the signal, it produces the same result as if these samples assume zero values [38], [39], [51]. Consequently, a reduced number of signal samples can be considered as a complete set of samples, where some of them are affected by an additive noise. When missing samples assume zero values, the initial discrete DGHmT of such signal is

$$C_0(p) = \sum_{i=1}^{M_A} \sum_{l=1}^{K} \frac{\alpha_l}{M} \frac{\psi_p(\hat{t}_i)\psi_{p_l}(\hat{t}_i)}{(\psi_{M-1}(\hat{t}_i))^2}, \; p = 0,...,M-1.$$

Hermite coefficients $C_0(p)$ are random variables. Two categories of these coefficients can be identified. Coefficients corresponding to signal components, that is, $p = p_l, \quad l = 1,...., K$ are random variables with Gaussian distribution and mean values $\alpha_l \frac{M_A}{M}$ ([38], [50]) and variance $\sigma_s^2(p_l)$: $\mathcal{N}(\alpha_l \frac{M_A}{M}, \sigma_s^2(p_l)), \quad l = 1,..., K$. Coefficients that are not placed at positions corresponding to signal components, that is, $p \neq p_l$, are zero-mean random variables with variance [38], [50], [51]

$$\sigma_{cs}^2 = \frac{M_A M - M_A^2}{M^2(M-1)} \sum_{l=1}^{K} \alpha_l^2 \tag{34}$$

not depending on coefficient position as long as $p \neq p_l$ holds, that is, $\mathcal{N}(0, \sigma_{cs}^2)$.

The probability that $M-K$, $M \gg K$ independent noise alone coefficients are smaller than a value κ is

$$P_S(\kappa) = \operatorname{erf}\left(\frac{\kappa}{\sqrt{2}\sigma_{cs}}\right)^{M-K}, \tag{35}$$

where the CDF of the Half-normal distribution being $g(\kappa, \sigma_N) = \operatorname{erf}\left(\frac{\kappa}{\sqrt{2}\sigma_{cs}}\right)$ is used. If we set this probability to a fixed value, for example $P_S(\kappa) = 0.99,$ then, the threshold

$$\kappa = \sqrt{2}\sigma_{cs} \operatorname{erf}^{-1}\left(\left(P_S(\kappa)\right)^{\frac{1}{M-K}}\right) \approx \sqrt{2}\sigma_{cs} \operatorname{erf}^{-1}\left(\left(P_S(\kappa)\right)^{\frac{1}{M}}\right) \tag{36}$$

will, with given probability, separate noise-only coefficients from signal coefficients. If amplitudes α_l have close values, and a sufficiently large number of measurements M exists, then this threshold will help to detect positions of all coefficients corresponding to signal components, that is, $\hat{\mathbf{P}} = \arg\{|C_0(p)| > \kappa\}$. In that case, the reconstruction is simply achieved in one iteration, directly exploiting the pseudo-inversion (33).

Now assume that the available measurements are affected by a white Gaussian noise with variance σ_η^2. Both the external additive noise and the noise caused by missing samples are now sources of disturbances in the considered discrete Hermite domain. These disturbances are uncorrelated and Gaussian, and therefore the noise-only coefficients have the following variance

$$\sigma_T^2(p) = \sigma_{cs}^2 + \sigma_X^2(p) = \frac{M_A M - M_A^2}{M^2(M-1)} \sum_{l=1}^{K} \alpha_l^2 + \gamma(p, M)\sigma_\eta^2 \tag{37}$$

with constant σ_{cs}^2, whereas $\sigma_X^2(p)$ is dependent on coefficients positions. Therefore, we introduce non-linear position-dependent threshold for the detection of Hermite coefficients corresponding to signal components:

$$\kappa(p) = \sqrt{2}\sigma_T(p)\,\mathrm{erf}^{-1}\left(\left(P_S(\kappa)\right)^{\frac{1}{M}}\right). \tag{38}$$

Threshold (38) is incorporated in the efficient reconstruction algorithm presented in [39]. If amplitudes α_l do not have close values, then the variant of CoSaMP (OMP) based reconstruction can be easily developed based on the presented theory [39]. In this algorithm, based on initial coefficient vector $\mathbf{C}_0$, the first set of component positions is estimated as

$$\hat{\mathbf{P}} = \arg\{|C_0(p)| > \kappa(p)\}. \tag{39}$$

In the next step, partial sensing matrix $\mathbf{A}_1 = \mathbf{A}_{\hat{\mathbf{P}}}$ is formed from the sensing matrix $\mathbf{A}$, using only columns with indices $\hat{\mathbf{P}}$. First components are obtained by solving the system of measurement equations, using the well-known pseudo-inversion: $\mathbf{C}_K = \left(\mathbf{A}_1^T\mathbf{A}_1\right)^{-1}\mathbf{A}_1^T\mathbf{s}_{cs}$. Subsequently, the signal $\mathbf{s}_1 = \mathbf{A}_1\mathbf{C}_K$ is calculated. In case that $\mathbf{e} = \mathbf{s}_1$ holds, vector $\mathbf{C}_K$ is the problem solution and algorithm terminates (single iteration variant). If this is not the case, then the estimated component is removed from $\mathbf{s}_{cs}$ and the signal $\mathbf{e} = \mathbf{s}_{cs} - \mathbf{s}_1$ is formed.

Subsequently, the procedure is repeated for signal $\mathbf{e}$, where the set $\hat{\mathbf{P}}$ is expanded with new detected positions. The process iteratively continues until the solution is found, or until a required precision ε is acquired. A reasonable choice of parameter ε is the mean variance of noisy DGHmT coefficients (25), being the new stopping criterion for the reconstruction of signals in the presence of AWGN. Detailed description and analysis of presented reconstruction concepts and algorithms (single-iteration and iterative forms) can be found in [39].

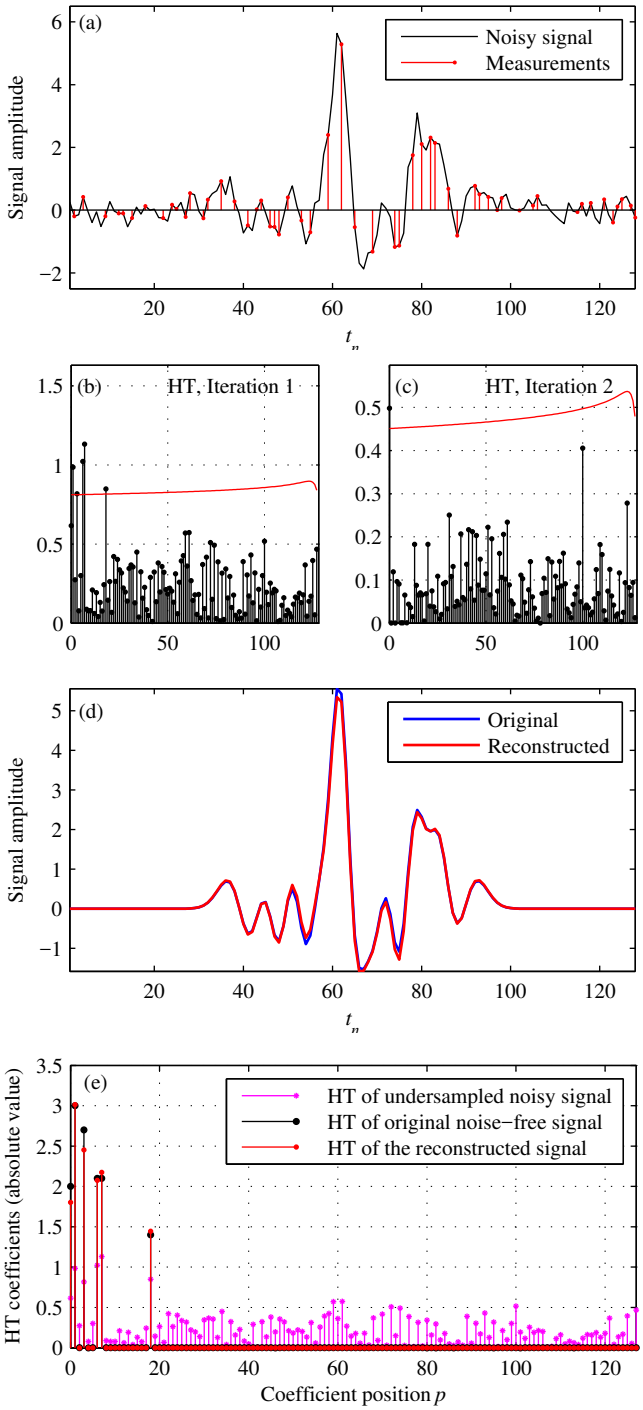

Fig. 9. Reconstruction of noisy signal sparse in the DGHmT domain with M_A = 54 out of M = 128 available samples: (a) noisy signal (black) and available measurements (red); (b) HT of available measurements (initial transform) and the first threshold (red); (c) HT calculated after removal of detected components, and the new threshold (red); (d) Original noise-free signal (blue) and the reconstructed signal (red); (e) HT of the undersampled noisy signal, original noise-free signal and the reconstructed signal

Example 4: Consider signal of the form (30), sparse in the DGHmT domain, with amplitudes $\alpha_1 = 2$, $\alpha_2 = -3$, $\alpha_3 = 2.7$, $\alpha_4 = 2.1$, $\alpha_5 = 2.1$, $\alpha_6 = 1.4$, and corresponding non-zero coefficients positions $p_1 = 0$, $p_2 = 1$, $p_3 = 3$, $p_4 = 6$, $p_5 = 7$, and $p_6 = 18$. It is affected by an additive white Gaussian noise; such that SNR is 7dB. Then only M_A = 54 out of M = 128 noisy samples (42.19%) are taken at random positions. The reconstruction is performed in only two iterations of the algorithm [39]

with the modified threshold (38). Results are shown in Fig. 9. The initial MSE between the original (noise-free) and noisy signal of -11.67dB dropped for ~12dB after the CS reconstruction. The resulting MSE between the original (noise-free) and the reconstructed signal is -23.66dB.

7. Conclusion

The discrete Hermite transform of signals affected by zero-mean AWGN is analyzed. The particular scope was on the discrete Hermite transform calculated based on the Gauss-Hermite quadrature. The mean values and variances of corresponding noisy Hermite coefficients are derived and statistically verified. Based on the modeled noise influence, a nonlinear threshold for the separation of signal and noise-alone components in the discrete Hermite transform domain is derived. Consequently, a method for signal denoising is proposed. It is based on a hard thresholding approach. The signal denoising is tested using real UWB signal and QRS complex, showing the benefits of using the derived nonlinear threshold. Moreover, the results are easily incorporated into the sparse signal reconstruction framework.

Acknowledgment

This work is supported by the Montenegrin Ministry of Science, project grant: CS-ICT “New ICT Compressive sensing based trends applied to: multimedia, biomedicine and communications”.

References

[1] A. Sandryhaila, S. Saba, M. Puschel, J. Kovacevic, "Efficient compression of QRS complexes using Hermite expansion," *IEEE Transactions on Signal Proccessing*, vol. 60, no. 2, pp .947-955, 2012, doi: 10.1109/TSP.2011.2173336

[2] G. Leibon, D. N. Rockmore, W. Park, R. Taintor, G. S. Chirikjian, "A fast Hermite transform," *Theoretical Computer Science*, vol. 409, issue 2, pp. 211-228, 2008, doi: 10.1016/j.tcs.2008.09.010

[3] A. Krylov, D. Kortchagine, “Fast Hermite projection method,” *Int. Conf. on Image Analysis and Recognition*, Portugal: 329-338, 2006

[4] I. Orović, S. Stanković, T. Chau, C. M. Steele, and E. Sejdić, “Time-frequency analysis and Hermite projection method applied to swallowing accelerometry signals,” *EURASIP Journal on Advances in Signal Processing*, Vol. 2010, Article ID 323125, 7 pages, 2010.

[5] A. Mahadevan, S. Acharya, D. B. Sheffer and D. H. Mugler, "Ballistocardiogram Artifact Removal in EEG-fMRI Signals Using Discrete Hermite Transforms," *in IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 6, pp. 839-853, Dec. 2008. doi: 10.1109/JSTSP.2008.2008367

[6] D. H. Mugler, S. Clary and Yan Wu, "Discrete Hermite expansion of digital signals: applications to ECG signals," *Proceedings of 2002 IEEE 10th Digital Signal Processing Workshop, 2002 and the 2nd Signal Processing Education Workshop.*, 2002, pp. 262-267.

[7] X. Ning, I. W. Selesnick, "ECG Enhancement and QRS Detection Based on Sparse Derivatives," *Biom. Sig. Proc. and Control,* vol. 8, no. 6, pp.713-723, 2013, doi: 10.1016/j.bspc.2013.06.005

[8] M. Brajović, I. Orović, M. Daković, and S. Stanković, “On the Parameterization of Hermite Transform with Application to the Compression of QRS Complexes,” *Signal Processing*, vol. 131, pp. 113-119, February 2017

[9] A. I. Rasiah, R. Togneri, Y. Attikiouzel, "Modelling 1-D signals using Hermite basis functions," *Vision, Image and Signal Processing, IEE Proceedings*, vol. 144, no.6, pp. 345-354, 1997, doi: 10.1049/ip-vis:19971613

[10] S. Stanković, I. Orović, A. Krylov, "The two-dimensional Hermite S-method for high resolution inverse synthetic aperture radar imaging applications," *IET Signal Processing*, vol. 4, no. 4, pp. 352-362, 2010, doi:10.1049/iet-spr.2009.0060

[11] M. Brajović, I. Orović, M. Daković, and S. Stanković, "Gradient-based signal reconstruction algorithm in the Hermite transform domain," *Electronics letters*, vol. 52, no. 1, pp. 41-43, 2016, doi: 10.1049/el.2015.1700

[12] S. Stanković, I. Orović, and LJ. Stanković, "Compressive Sensing approach in Hermite transform domain," *Mathematical Problems in Engineering*, Volume 2015 (2015), Article ID 286590, 9 pages

[13] J. R. de Oliveira Neto, J. B. Lima, "Discrete fractional Fourier transforms based on closed-form Hermite-Gaussian-like DFT eigenvectors," *IEEE Transactions on Signal Processing*, vol. 65, no. 23, pp. 6171-6184, December 2017. doi:10.1109/TSP.2017.2750105

[14] C. Candan, M. A. Kutay and H. M. Ozaktas, "The discrete fractional Fourier transform," in *IEEE Transactions on Signal Processing*, vol. 48, no. 5, pp. 1329-1337, May 2000.

[15] S. C. Pei, W. L. Hsue and J. J. Ding, "Discrete Fractional Fourier Transform Based on New Nearly Tridiagonal Commuting Matrices," in *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3815-3828, Oct. 2006.

[16] B. Santhanam and T. S. Santhanam, "Discrete Gauss-Hermite Functions and Eigenvectors of the Centered Discrete Fourier Transform," *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07*, Honolulu, HI, 2007, pp. 1385-1388.

[17] M. T. Hanna, N. P. A. Seif, and W. A. E. M. Ahmed, "Discrete fractional Fourier transform based on the eigenvectors of tridiagonal and nearly tridiagonal matrices," *Digital Signal Processing*, vol. 18, no. 5, pp. 709–727, Sep 2008.

[18] S. C. Pei, J. J. Ding, W. L. Hsue and K. W. Chang, "Generalized Commuting Matrices and Their Eigenvectors for DFTs, Offset DFTs, and Other Periodic Operations," in *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3891-3904, Aug. 2008.

[19] A, Serbes, L. Durak-Ata, "Efficient computation of DFT commuting matrices by a closed-form infinite order approximation to the second differentiation matrix," *Signal Processing*, vol. 91, Issue 3, 2011, pp. 582-589

[20] D. Y. Wei, Q. W. Ran, Y. M. Li, J. Ma and L. Y. Tan, "Fractionalisation of an odd time odd frequency DFT matrix based on the eigenvectors of a novel nearly tridiagonal commuting matrix," in *IET Signal Processing*, vol. 5, no. 2, pp. 150-156, April 2011.

[21] D. Wei and Y. Li, "Novel Tridiagonal Commuting Matrices for Types I, IV, V, VIII DCT and DST Matrices," in *IEEE Signal Processing Letters*, vol. 21, no. 4, pp. 483-487, April 2014.

[22] I. Bhatta and B. Santhanam, "A comparative study of commuting matrix approaches for the discrete fractional fourier transform," *2015 IEEE Signal Processing and Signal Processing Education Workshop (SP/SPE)*, Salt Lake City, UT, 2015, pp. 1-6.

[23] Ç. Candan, "On Higher Order Approximations for Hermite–Gaussian Functions and Discrete Fractional Fourier Transforms," in *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 699-702, Oct. 2007.

[24] M. T. Hanna, N. P. A. Seif and W. A. E. M. Ahmed, "Hermite-Gaussian-like eigenvectors of the discrete Fourier transform matrix based on the singular-value decomposition of its orthogonal projection matrices," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 51, no. 11, pp. 2245-2254, Nov. 2004.

[25] M. T. Hanna, N. P. A. Seif and W. A. E. M. Ahmed, "Hermite-Gaussian-like eigenvectors of the discrete Fourier transform matrix based on the direct utilization of the orthogonal projection matrices on its eigenspaces," in *IEEE Trans. on Signal Processing*, vol. 54, no. 7, pp. 2815-2819, July 2006.

[26] M. T. Hanna, "Direct Batch Evaluation of Optimal Orthonormal Eigenvectors of the DFT Matrix," in *IEEE Transactions on Signal Processing*, vol. 56, no. 5, pp. 2138-2143, May 2008.

[27] M. T. Hanna, "Direct sequential evaluation of optimal orthonormal eigenvectors of the discrete Fourier transform matrix by constrained optimization," *Digital Signal Processing*, vol. 22, Issue 4, 2012, pp. 681-689

[28] A. Serbes and L. Durak-Ata, "The discrete fractional Fourier transform based on the DFT matrix," *Signal Processing*, vol. 91, no. 3, pp. 571–581, March 2011.

[29] S.-C. Pei, M.-H. Yeh and C.-C. Tseng, "Discrete fractional Fourier transform based on orthogonal projections," in *IEEE Transactions on Signal Processing*, vol. 47, no. 5, pp. 1335-1348, May 1999.

[30] A. Kuznetsov, “Explicit Hermite-type eigenvectors of the discrete Fourier transform,” *SIAM Journal on Matrix Analysis and Applications*, vol. 36, no. 4, pp. 1443–1464, 2015.

[31] S. C. Pei and Y. C. Lai, "Signal Scaling by Centered Discrete Dilated Hermite Functions," in *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 498-503, Jan. 2012.doi: 10.1109/TSP.2011.2171687

[32] S. Clary, and D. H. Mugler. "Shifted Fourier matrices and their tridiagonal commutors." *SIAM Journal on Matrix Analysis and Applications* 24, no. 3 (2003): 809-821.

[33] M. Ghavami, L.B. Michael, R. Kohno, *Ultra Wideband Signals and Systems in Communication Engineering*, 2nd Edition, Wiley, 2007.

[34] R. J. M. Cramer, R. A. Scholtz, and M. Z. Win , "Evaluation of an Ultra-Wide Band Propagation Channel*", IEEE Transactions on Antennas and Propagation*, May 2002, doi: 10.1109/TAP.2002.1011221

[35] S. Gezici, H. Kobayashi, H. V. Poor and A. F. Molisch, "Performance evaluation of impulse radio UWB systems with pulse-based polarity randomization," in *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2537-2549, July 2005.

[36] F. Ramirez-Mireles, "On the performance of ultra-wide-band signals in Gaussian noise and dense multipath," in *IEEE Transactions on Vehicular Technology*, vol. 50, no. 1, pp. 244-249, Jan 2001.

[37] Y. Kopsinis and S. McLaughlin, "Development of EMD-Based Denoising Methods Inspired by Wavelet Thresholding," in *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1351-1362, April 2009, doi: 10.1109/TSP.2009.2013885

[38] LJ. Stanković, S. Stanković, and M. Amin, "Missing Samples Analysis in Signals for Applications to L-estimation and Compressive Sensing, " *Signal Processing*, vol. 94, pp. 401-408, Jan 2014.

[39] S. Stanković, I. Orović, and LJ. Stanković, "An Automated Signal Reconstruction Method based on Analysis of Compressive Sensed Signals in Noisy Environment, "*Signal Processing*, vol. 104, pp. 43 - 50, 2014.

[40] Z. Xu, B. Huang, K. Li, On Fourier Interpolation Error for Band-Limited Signals, *IEEE Trans. on Signal Process*. 57 (6) (2009) 2412–2416, doi: 10.1109/TSP.2009.2016263.

[41] A. Papoulis, *Signal analysis*, McGraw-Hill, New York, 1977.

[42] JP. Berrut, “A formula for the error of finite sinc-interpolation over a finite interval,” *Numerical Algorithms*, Vol. 45, Issue 1–4, pp 369–374, Aug. 2007

[43] PhysioNet: MIT-BIH ECG Compression Test Database, http://www.physionet.org/physiobank/database/cdb, accessed May 2017

[44] UltRa Lab UWB data: http://ultra.usc.edu/uwb_database/tdcindoor.htm, accessed May 2017

[45] Y. Kopsinis and S. McLaughlin, "Development of EMD-Based Denoising Methods Inspired by Wavelet Thresholding," in *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1351-1362, 2009

[46] I. Selesnick, "Sparsity-assisted signal smoothing (revisited)," *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA, 2017, pp. 4546-4550.

[47] I. W. Selesnick, Sparsity-assisted signal smoothing (SASS) [available online]: http://eeweb.poly.edu/iselesni/sass/, accessed January 2017.

[48] E. Candes, J. Romberg, and T.Tao, "Robust Uncertainty Principles: Exact Signal Reconstruction From Highly Incomplete Frequency Information, " *IEEE Transactions on Information Theory*, 52(2), pp. 489-509, 2006.

[49] E. Sejdić, A. Can, L. F. Chaparro, C. M. Steele, and T. Chau, “Compressive sampling of swallowing accelerometry signals using time-frequency dictionaries based on modulated discrete prolate spheroidal sequences,” *Eurasip Journal on Advances in Sig. Process*., vol. 2012, article 101, 2012.

[50] M. Brajović, I. Orović, M. Daković, and S. Stanković, “The Analysis of Missing Samples in Signals Sparse in the Hermite Transform Domain,” *23rd Telecommunications Forum TELFOR*, Belgrade, 2015

[51] M. Brajović, I. Orović, M. Daković, S. Stanković, “Compressive Sensing of Sparse Signals in the Hermite Transform Basis: Analysis and Algorithm for Signal Reconstruction”, *IEEE Transactions on Aerospace and Electronic Systems,* accepted for publication

[52] Antoniadis, A.; G. Oppenheim, Eds. (1995), *Wavelets and statistics*, 103, Lecture Notes in Statistics, Springer Verlag.

[53] Donoho, D.L. (1993), “Progress in wavelet analysis and WVD: a ten minute tour,” in *Progress in wavelet analysis and applications*, Y. Meyer, S. Roques, pp. 109–128. Frontières Ed.

[54] Donoho, D.L.; I.M. Johnstone (1994), “Ideal spatial adaptation by wavelet shrinkage,” *Biometrika*, Vol. 81, pp. 425–455.

[55] Donoho, D.L. (1995), “De-noising by soft-thresholding,” *IEEE Trans. on Inf. Theory*, 42 3, pp. 613– 627.