SPECTRAL ESTIMATION WITH THE HIRSCHMAN OPTIMAL TRANSFORM FILTER BANK AND COMPRESSIVE SENSING

Guifeng Liu and Victor DeBrunner
Department of Electrical and Computer Engineering
Florida State University
Tallahassee, FL, 32310
gl07d@my.fsu.edu, victor.debrunner@fsu.edu

***Abstract*—The traditional Heisenberg-Weyl measure quantifies the joint localization, uncertainty, or concentration of a signal in the phase plane based on a product of energies expressed as signal variances in time and in frequency. Unlike the Heisenberg-Weyl measure, the Hirschman notion of joint uncertainty is based on the entropy rather than the energy [1]. Furthermore, as we noted in [2], the Hirschman optimal transform (HOT) is superior to the discrete Fourier transform (DFT) and discrete cosine transform (DCT) in terms of its ability to resolve two limiting cases of localization in frequency, viz pure tones and additive white noise. We found in [3] that the HOT has a superior resolution to the DFT when two pure tones are close in frequency. In this paper, we improve on that method to present a more complete spectral analysis tool. Here, we implement a stationary spectral estimation method using compressive sensing (in particular, Iterative Hard Thresholding) on HOT filterbanks. We compare its frequency resolution to that of a DFT filterbank using compressive sensing. In particular, we compare the performance of the HF with that of the DFT in resolving two close frequency components in additive white Gaussian noise (AWGN). We find the HF method to be superior to the DFT method in frequency estimation, and ascribe the difference to the HOT's relationship to entropy.**

***Index Terms*—Hirschman Optimal Transform, Orthogonal Matching Pursuits, Periodogram, Quinn's method**

1. INTRODUCTION

We introduced an entropy-based measure U_p [4] that quantifies the compactness of a discrete-time signal in the sample-frequency phase plane. Use of an entropy-based measure allowed us to overcome the limitations inherent to discretizing the Heisenberg uncertainty. Our entropy-based measure was used to show that discretized Gaussian pulses may not be the most compact basis with respect to joint time-frequency resolution. In [1], we found a basis (HOT transform) that is orthonormal and uniquely minimizes the discrete-time, discrete-frequency Hirschman uncertainty principle. For comparison, we considered a discretized Gaussian pulse, which is comparable to the HOT basis. We found the uncertainty U_ρ realized by the discretized Gaussian pulse is greater than that of the HOT basis functions [2].

The question we ask is: Can this improved localization of the HOT be used to improve spectral estimation techniques? Using the HOT and DFT, we examine the power spectrum experimentally, and compare the performance of our developed HOT-DFT periodogram to that of the classical periodogram using the DFT. Our experiment is to distinguish two closely-spaced frequency components with different amplitude ratios embedded in AWGN. The composite signal passes through the filter banks, then we reconstruct the selected channel signals using an energy criterion and apply the compressive sensing, i.e. Iterative Hard Thresholding algorithm, before applying classical Quinn's smoothing kernel [5] to get the power spectrum. We observe that, after thresholding, the HOT-DFT estimated spectrum is superior to the DFT when the signal-to-noise ratio (SNR) is as low as 0 dB.

In this paper, we briefly review the HOT, then we develop an filter bank method and apply the Iterative Hard Thresholding algorithm to estimate the power spectrum of a signal with the use of Quinn's method, where the channels of the filter bank are derived using both the HOT and the DFT. We carefully derive these filter banks, because they show the utility of the entropy-based spectral estimation method.

2. THE HIRSCHMAN OPTIMAL TRANSFORM

In this section, we consider discrete 1–D signals on a finite domain. Fix a finite set of non negative integers $\mathcal{D} = 0, 1, 2 \ldots, N-1$. Let $\mathcal{H}_\mathcal{N}$ denote the Hilbert space of sequences $x : \mathcal{D} \to \mathbb{C}$ with squared-norm

$$\|x\|_2^2 = \sum_{n=0}^{N-1} |x[n]|^2 \tag{1}$$

Using the twiddle factor notation $W_N = e^{-j(2\pi/N)}$, the DFT is

$$X[k] = Fx[n] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] W_N^{nk},\ k \in \mathcal{D} \tag{2}$$

This defines an isometry on $\mathcal{H}_\mathcal{N}$ with inverse given by

$$x[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X[k] W_N^{-nk} \tag{3}$$

By the digital phase plane, we mean the set of all points $(n, k) \in \mathcal{D} \times \mathcal{D}$. The translation and modulation operators (see [2] for details) allow us to view the entire digital phase plane. The uncertainty measure we use is based on entropy instead of energy. Consider the following definition: For $x \in \mathcal{H}_{\mathcal{N}}$ with $\|x\|_2 = 1$, the (Shannon) entropy is defined as

$$S(x) = -\sum_{n=0}^{N-1} |x[n]|^2 \ln\left(|x[n]|^2\right) \tag{4}$$

This entropy is defined on the pseudo-density determined from the normalized-energy signal, and not from any statistical definition. Using this entropy, we define a general class of digital uncertainty measures for $0 \le \rho \le 1$:

$$U_\rho(x) = \rho S(x) + (1-\rho)S(Fx), \quad x \in \mathcal{H}_{\mathcal{N}}, \|x\|_2 = 1. \tag{5}$$

In the special case where $\rho = \frac{1}{2}$, the measure (5) is called the *digital Hirschman uncertainty* [2]. In general, ρ allows for a tradeoff between concentration in time and in frequency. In the extreme where $\rho = 1$, the measure (5) ignores frequency localization, and the minimizing signals are those concentrated at single points. Similarly, if $\rho = 0$, the minimizing signals are those for which all the sample magnitudes $|x(n)|$ are equal. Intermediate values of ρ give a weighted measure of joint time-frequency localization of the signal. Before describing the minimizers of (5), we define periodization:

Definition 1. For $N = KL$, the periodization of $v \in \mathbb{C}^k$ is defined as $x[sK+n] = (1/\sqrt{L})v[n]$ for $0 \le s \le L-1$ and $0 \le n \le K-1$. We refer to the sequence $v \in \mathbb{C}^k$ given by $v[0] = 1$, $v[1] = 0, \cdots v[K-1] = 0$, as the Kronecker delta or impulse (unit sample) sequence, without specifying the signal length K. We proved the following theorem in [1]:

Theorem 2. *The only sequences $x \in \mathbb{C}^k$, with $\|x\|_2 = 1$, for which $U_{\frac{1}{2}}(x)$ is minimal are obtained from the Kronecker delta sequence by applying any composition of periodization, translation, modulation, the DFT, and multiplication by a complex number of unit magnitude.*

3. HOT/DFT Spectral Estimation Using Filter Bank

We use the K-dimensional DFT kernel as the originator signals for our $N = K^2$-length HOT basis. Each of these basis functions must then be shifted and up-sampled to produce the sufficient number of orthogonal basis functions that define the HOT. While other choices are possible, this one leads to an efficient computational structure with a complexity less than that of the N-point DFT. Note that the DFT kernel could also be used in a similar manner to produce transforms for other factorizations $N = KL, K \neq L$, but these possess an uncertainty H_p that varies as a function of p and are suboptimal in this sense [1].

The HOT is unitary to a scale, and so the inverse transform results by taking the conjugate transpose and scaling by $\sqrt{K}$. In general [1]:

$$H[Kr+l] = \frac{1}{\sqrt{K}} \sum_{n=0}^{K-1} x[Kn+l]\, e^{-j\frac{2\pi}{K}nr}, 0 \le r, l \le K-1. \tag{6}$$

and its inverse

$$x[Kn+l] = \frac{1}{\sqrt{K}} \sum_{r=0}^{K-1} H[Kr+l]\, e^{j\frac{2\pi}{K}nr}, 0 \le n, l \le K-1. \tag{7}$$

Note that while similar, the HOT and DFT are different. The N-point HOT is computationally more efficient than the N-point DFT – its complexity is equivalent to that of a $\sqrt{N}$-point DFT – and increasingly more efficient as $N \to \infty$.

Next we will show how we set up DFT and HOT filter bank frame. Let F be the DFT filter bank of complex FIR filters, and denote $F^{-1} \equiv IF$. The HOT filterbank is similarly constructed. Note that the HOT is built using the K-dimensional DFT kernel, where $N = K^2$. Denote $H^{-1} \equiv IH$. We build a HOT filter bank which is called FIH by applying F on the IH. Since F, IH and H are unitary, FIH will be unitary. We have two operators F and FIH. To compare their performance, we combine F and FIH to form the hybrid matrix $HF = [FIH;\, F]$. Let the signal pass through the HF and F filter banks, respectively.

4. Compressive Sensing Algorithm

Compressive Sensing (CS) is an efficient acquisition framework for signals $x \in \mathbb{C}^{\mathbb{N}}$ which are s-sparse in a known orthonormal basis matrix $\Psi \in \mathbb{C}^{N \times N}$. We write $x = \Psi\theta$, where only $s \ll N$ out of N signal coefficients θ are nonzero. The CS observation vector y is given by $y = \Phi x$, where Φ is an $M \times N$ measurement matrix. So the number of measurements required to ensure that y retains all of the information in x is $M = O(s \log(N/s))$ [6], [7], [8]. Furthermore, x can be recovered from its compressive measurements y via a convex optimization or some other iterative method.

The compressed sensing observation vector y is given by

$$y = \Phi x = (\Phi\Psi)\,\theta \triangleq A\theta$$

Here we define $A = \Phi\Psi_0 \in \mathbb{C}^{M \times N}$ as the compressive sensing measurement matrix with respect to θ. Since A has reduced dimension, many vectors x' exist which may yield the same y. However, if A satisfies the Restricted Isometry Property (RIP) [6], the s-sparse signals can be found by standard sparse approximation algorithms.

Definition 3. The $s-$restricted isometry constant for matrix A, denoted by δ_s, is the smallest non negative number such that, for all $\theta \in \mathbb{R}^N$ with $\|\theta\|_0 = s$,

$$(1-\delta_s)\|\theta\|_2^2 \le \|A\theta\|_2^2 \le (1+\delta_s)\|\theta\|_2^2$$

Matrix A is called RIP if $\delta_s < 1$; using i.i.d. Gaussian or Bernoulli entries ensures A is RIP.

Iterative methods such as the iterative hard thresholding (IHT) algorithm [9] are simple to implement and works well in the presence of noise on compressible signals. Thus, we choose the IHT algorithm to recover our compressible signal ($s = 2$) in this paper. The ℓ_0 regularized optimization problem is defined as follows: for given x and A , find coefficients θ minimizing the cost function:

$$C_{\ell_0}(\theta) = \|x - A\theta\|_2^2 + \lambda \|\theta\|_0$$

where $\|\theta\|_0$ is defined as the number of non-zero coefficients. To minimize the cost function, IHT is derived in [10]:

$$\theta_{i+1} = H_s\left(\theta_i + A^T(x - A\theta_i)\right)$$

where $\theta_0 = 0$, and $H_s(\mathbf{a})$ is the non-linear operator that sets all but the largest (in magnitude) s elements of $\mathbf{a}$ to zero. If there is no unique such set, a set can be selected either randomly or based on a predefined ordering of the elements. The IHT algorithm will perfectly recover $s-$sparse signals when $\delta_{3s} \leq 1/\sqrt{32}$.

As mentioned in Section 3, the compressive sensing theory can extend to noisy signals (or compressible signals) that are not exactly sparse but can be approximated as sparse signals. The sorted coefficients θ decay according to the power law: $|\theta[i]| = C\, i^{-1/p}$ for some $p \leq 1$ [11]. Our signals are power spectrum coefficients obtained from the reconstructed channel signals derived from the filter banks.

We need to note that there is a "mismatch" between the assumed basis for sparsity and actual the basis in which the θ is sparse[12]. The DFT coefficients will be sparse only when sinusoids in x have integral frequencies of the form $2\pi n/N$, where n is an integer. Otherwise the DFT coefficients will be compressible (less sparse) because of the spectral leakage introduced by windowing. It is intuitive to employ a redundant DFT frame to reduce the leakage caused by non-integral frequencies. However, standard sparse approximation algorithms for x in redundant DFT frames do not perform well when the redundancy factor ($c \in \mathbb{N}$) increases even though the frequency sampling interval decreases from $2\pi/N$ to $2\pi/(cN)$. The reason is the *coherence* between the frame vectors becomes high for large values of c and the performance of CS deteriorates. In this paper, we mainly focus on the performance comparison between HF and DFT. The redundancy factor does not change the relative performance much. So, we set $c = 1$ for computational convenience. The coherence of frame Ψ is defined as:

$$\mu(\Psi) = \arg \max_{1 \leq i,j \leq N} |\langle \psi_i, \psi_j \rangle|$$

where ψ_i denotes the i^{th} column of Ψ. The problem we target in this paper is to resolve two close frequency components. To observe the effect of the HOT, we push the frequency distance limit even smaller, so the coherence value in this case will be higher than 1, while in [13], the authors focus on coherence inhibited structured sparse estimation.

5. Classic Quinn's method

Quinn studied the power spectrum problem with DFT in 1990's [5]. He started from ARMA(p, q) model, where p and q are order of the AR (Autoregressive) part and MA (Moving Average) part respectively:

$$\sum_{m=0}^{p} a_m x[n-m] = \sum_{m=0}^{q} b_m \nu[n-m] \qquad (8)$$

a_n and b_n are parameters of the model, $x[n]$ is time series data, $n \in [0, N-1]$, and $\nu[n]$ is the white noise with zero mean and finite variance. For signal with frequencies $\omega_k \in (0, \pi)$:

$$x[n] = \sum_{k=1}^{p} \rho_k \cos(\omega_k n + \phi_k) + \nu[n] \qquad (9)$$

where ρ_k are the amplitude of the k_{th} sinusoidal signal and ϕ_k are the corresponding initial phase.

A special ARMA$(2p, 2q)$ annihilates all the sinusoidal components in $x[m]$, if the parameters $a_m = b_m = \beta_m$, and satisfy:

$$\sum_{m=0}^{2p} \beta_m z^m = \prod_{m=1}^{p} \left(1 - 2z\cos\omega_m + z^2\right)$$

The parameters β need to satisfy:

$$\beta_{2p-m} = \beta_m (m = 0, \ldots, p-1), \quad \beta_0 = 1$$

For example, set one sinusoidal signal $x[n] = \rho\cos(\omega n + \phi) + \nu[n]$, it can be annihilated by $ARMA(2,2)$. From this special ARMA model, Quinn derived the so called smoothed periodogram:

$$\kappa_N(\omega) = \int_{-\pi}^{\pi} P_x(\lambda)\mu_N(\omega - \lambda)d\lambda \qquad (10)$$

where $P_x(\lambda)$ is the periodogram of $x[n]$, $\mu_N(\omega) = \sum_{k=1}^{N-1} k^{-1}\cos(k\omega)$ is the kernel or we can say window function. The smoothed periodogram is the convolution of $P_x(\lambda)$ and the kernel. The key point of Quinn's method is: $\kappa_N(\omega)$ in Eq. (10) can give the accurate frequency estimation without much more zero padding.

6. Simulations

We consider two signals, $x_1[n] = A_1\cos[2\pi f_1/f_s n + \phi_1]$ and $x_2[n] = A_2\cos[2\pi f_2/f_s n + \phi_2]$, where ϕ_1 and ϕ_2 are the initial phases (we set both to zero). The length of the signal, N_o, is set to 256. The sampling frequency is $f_s = 1000$ Hz. The signal to be estimated is $x[n] = x_1[n] + x_2[n] + v[n]$, where $v[n]$ is ZMWG noise. In our simulations, the input signal $x[n]$ is normalized to unit energy. Our results are based on $t = 100$ different noise realizations. The length of the HF and DFT filter bank is $L = 64$, so the HF has 128 different channels while the DFT filterbank has 64 channels. The signal length increases to $N = N_0 + L - 1$ in the filtering. During the compressive sensing step, for safety, the number of measurements is 32 ($>M = O(s\log(N/s)) \approx 25$) . The maximum number of iterations is 10. If more channels are selected, more noises are added, which will deteriorate the performance. For each reconstruction by IHT, we apply Quinn's method followed by a peak picking process to estimate the frequencies f_1 and f_2. Both filterbanks have poor resolution when less than 4 channels are used. When using 7 channels with $SNR = 10$ dB, and the amplitude ratio $AR = \log_{10}(A_1/A_2) = -0.1$, the minimum frequency interval (i.e. the Bin Width) is $\triangle = f_s/N = 3.13$ Hz. Since we are using $f_1 = 19\triangle$Hz and $f_2 = 20.2\triangle$Hz, the frequency separation is 1.2 times the bin width. The power spectrum is shown in Fig.1, where we find that the performance of $\hat{P}_{HF}$ is superior to that of $\hat{P}_{DFT}$. In general, we find this to be a very common occurence for different base (lower) frequencies.

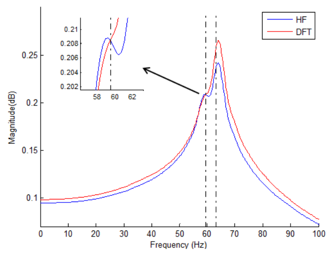

Figure 1. Power spectrum for $f_1 = 19\triangle$ Hz and $f_2 = 20.2\triangle$ Hz with AR=-0.1.

Note that the DFT estimate shows no low frequency peak, though the zoomed view does show that "something" is happening. In fact, when more channels are chosen, the lower frequency peak for the DFT is gradually revealed, but it is still only showing a small peak when compared to the HF peak. Another thing to note is: the measured f_2 peak value from HF is more accurate than that obtained from the DFT.

To compare the two methods, we use the Normalized Mean Square Error $(NMSE)$ of the peak positions of f_1 and f_2:

$$NMSE = \frac{1}{T-1}\sum_{i=1}^{T}(x_i - \bar{x})^2 + (\bar{x} - f)^2$$

where x_i are the estimated frequency values with the i^{th} noise, f is the vector of true frequencies, and $\bar{x} = \frac{1}{T}\sum x_i$. We normalize by the true frequencies, i.e. $NMSE = \frac{MSE}{f^t f}$.

If we keep the same base frequency f_1, and vary the frequency separation from $0.7\triangle$Hz to $1.7\triangle$Hz, we observe the problem in a different way. Fig.2 shows that the $NMSE$ of the HF is much smaller than that of the DFT with SNR=10 dB when 7 channels are selected from the filter bank, When frequency separation is less than $1.18\triangle$ Hz, both methods degrade, though the DFT performance decrease is more severe. For frequency separations greater than $1.18\triangle$ Hz, the DFT performance somewhat improves, though relatively its performance remains poor. Note that when $NMSE$ is 0dB, the frequency is missed, i.e. there is no peak as in the case of the lower frequency f_1 for the DFT case of Fig.1.

The $NMSE$ performance of the HF is superior and the difference is more pronounced with small SNR, which is consistent with the prediction in [2]. That the HOT can perform better in moderate and low SNR environments is very important in practical applications. Suppose that we change the SNR from -2 dB to 30 dB with 7 channels selected, while keeping all other parameters unaltered; this comparison is shown in Fig.2. We can clearly see that even though the f_1 estimation by HF is not so stable, the performance of HF is always better than that of the DFT over the entire SNR range. We find that with increasing SNR, the $NMSE$ of the HF drops at a much lower rate. Changing the base frequency does not alter the relative performance substantially.

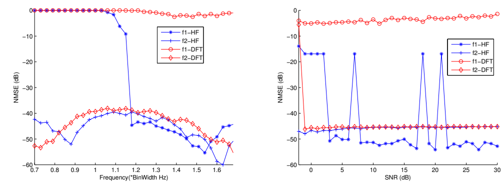

Figure 2. a) $NMSE$ of two frequency components with different frequency position . b) $NMSE$ of two frequency components with different SNR .

7. Conclusions

We introduce a method of nonparametric spectral estimation based on the HOT filter bank using the Iterative Hard Thresholding method of Compressive Sensing. *From our results, it should be clear that the impact of our choice of transform is important.* Specifically, we develop a filter bank generated with a combination of the HOT and DFT operations which we call the HF filter bank to preprocess the signal, and apply the compressive sensing method to their reconstructed signals. We calculate the smoothed periodogram with Quinn's method. *When compared to the DFT-only standard filter bank method, the power spectrum generated with our proposed method is superior over varying SNR, amplitude ratios, frequency separation, and number of channels used in the compressive sensing algorithm.* Future work must be done to determine an automated method for choosing the right number of channels, as well as for the peak determination. Also, we need to analyze the frequency resolution of the HF method.

References

[1] T. Przebinda, V. E. DeBrunner, and M. Özaydın, "The optimal transform for the discrete hirschman uncertainty principle," *Information Theory, IEEE Transactions on*, vol. 47, no. 5, pp. 2086–2090, July 2001.

[2] V. E. DeBrunner, J. P. Havlicek, T. Przebinda, and M. Özaydın, "Entropy-based uncertainty measures for $\ell^2(\mathbb{R}^n)$, $\ell^2(\mathbb{Z})$, and $\ell^2(\mathbb{Z}/N\mathbb{Z})$ with a hirschman optimal transform for $\ell^2(\mathbb{Z}/N\mathbb{Z})$," *Signal Processing, IEEE Transactions on*, vol. 53, pp. 2690–2699, Aug. 2005.

[3] G. Liu and V. E. DeBrunner, "High-resolution non-parametric spectral estimation using the hirschman optimal transform," vol. 10. Kyoto, Japan: Acoustics, Speech, and Signal Processing, IEEE International Conference on, Mar. 2012, p. Accepted.

[4] V. DeBrunner, M. Özaydın, and T. Przebinda, "Resolution in time-frequency," *Signal Processing, IEEE Transactions on*, vol. 47, pp. 783–788, Mar. 1999.

[5] B. G. Quinn and E. J. Hannan, *The Estimation and Tracking of Frequency*. Cambridge,UK: Cambridge University Press, 2001.

[6] E. J. Candés, "Compressive sampling," vol. 3. Madrid, Spain: Proc. International Congress of Mathematicians, 2006, pp. 1433–1452.

[7] D. L. Donoho, "Compressed sensing," *IEEE Trans. Info. Theory*, vol. 52, no. 4, pp. 1289–1306, Sep. 2006.

[8] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Processing Mag.*, vol. 24, no. 4, pp. 118–120, 2007.

[9] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265 – 274, 2009.

[10] ——, "Iterative thresholding for sparse approximations," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5, pp. 629–654, 2008.

[11] M. Duarte and R. G. Baraniuk, "Recovery of frequency-sparse signals from compressive measurements." Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on, 2010, pp. 599–606.

[12] Y. Chi, L. Scharf, A. Pezeshki, and A. Calderbank, "Sensitivity to basis mismatch in compressed sensing," *Signal Processing, IEEE Transactions on*, vol. 59, no. 5, pp. 2182–2195, may 2011.

[13] M. F. Duarte and R. G. Baraniuk, "Spectral compressive sensing," *unpublished*, 2011.