HIGH-RESOLUTION NON-PARAMETRIC SPECTRAL ESTIMATION USING THE HIRSCHMAN OPTIMAL TRANSFORM

Guifeng Liu and Victor DeBrunner
Department of Electrical and Computer Engineering
Florida State University
Tallahassee, FL, 32310
{liugu,victor.debrunner}@eng.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, its definition extends naturally from the case of infinitely supported continuous-time signals to the cases of both finitely and infinitely supported discrete-time signals, and, 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 separate or resolve two limiting cases of localization in frequency, viz pure tones and additive white noise. In this paper we implement a stationary spectral estimation method using an orthogonal matching pursuit method whose dictionary members are constructed from the combination of HOT-based and DFT atoms (elements) [3] in combination with the interpolating procedure developed in [4]. We call the resulting algorithm the smoothed HOT-DFT periodogram. We compare its performance (in terms of frequency resolution) to Quinn's smoothed periodogram. In particular, we compare the performance of the HOT-DFT with that of the DFT in resolving two close frequency components in additive white Gaussian noise (AWGN). We find the HOT-DFT to be superior to the DFT in frequency estimation, and ascribe the difference to the HOT's relationship to entropy.**

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

1. INTRODUCTION

IN earlier work, [5] introduced an entropy-based measure U_p that quantifies the compactness of a discrete-time signal in the sample-frequency phase plane that allowed us to overcome the limitations inherent to discretizing the Heisenberg uncertainty. A naïve discretization of the Heisenberg uncertainty leads to a discrete measure that fails to preserve translation invariance and is therefore not useful. The entropy-based measure showed 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 found that a discretized Gaussian pulse has an uncertainty U_p that 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 that currently use the DFT?* Using the HOT and DFT we first develop a smoothed HOT-DFT periodogram, and then compare its performance to that of the smoothed periodogram of Quinn that uses only the DFT. Our experiment is to distinguish two closely-spaced frequency components with different amplitude ratios embedded in AWGN. We observe that, after thresholding, the smoothed 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 orthogonal matching pursuit algorithm for estimating the power spectrum of a signal that uses Quinn's method [4], where the elements of the dictionary are derived using both the HOT and the DFT. We then compare the performances, and we find that using the HOT basis can significantly improve performance above that of current state-of-the-art methods.

2. THE HIRSCHMAN OPTIMAL TRANSFORM

Fix a finite set of nonnegative 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

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

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

$$X_D[k] = Fx[n] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] W_N^{nk}, \ k \in \mathcal{D} \qquad (1)$$

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

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

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. Hirschman Uncertainty uses entropy instead of energy. For $x \in \mathcal{H}_\mathcal{N}$ with $\|x\|_2 = 1$, the (Shannon) entropy is

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

Note that 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 $0 \leq p \leq 1$:

$$U_p(x) = pS(x) + (1-p)S(Fx), \quad x \in \mathcal{H}_\mathcal{N}, \|x\|_2 = 1. \quad (2)$$

In the special case where $p = \frac{1}{2}$, the measure (2) is called the *digital Hirschman uncertainty* [2]. Before describing the minimizers of (2), 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 Orthogonal Matching Pursuit and Quinn's Method

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 yield the orthogonal basis functions that define the HOT. This choice leads to an efficient computational structure (growing as $\sqrt{N} = K$) as compared to the N-point DFT. This DFT kernel can also produce transforms for other factorizations $N = KL, K \ne L$, but these possess an uncertainty U_p that varies as a function of p and are suboptimal in this sense [1]. In general [1], we have the analysis

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

and the synthesis

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

Note the similarities and differences of the HOT and DFT. One may think of the HOT as a "1-1/2 dimensional DFT" [2] in the sense that the equations for the HOT look like DFT's along the rows (or columns) of a data matrix.

Next we will show how we use the HOT to get a spectral estimate. To simplify the notation, define $X_D = Fx$, where x is the input signal, F is as in Eq. (1), and X_D denotes the DFT coefficients. Similarly, we define $X_H = Hx$, where H is the HOT. The HOT coefficients can be directly transformed into a set of frequencies by noting that the input signal can be expressed either as $x = F^{-1}X_D$ or as $x = H^{-1}X_H$. Thus, $X_D = FH^{-1}X_H$ and $X_H = HF^{-1}X_D$. Since both F and H are unitary, $F^{-1} = F^*$ and $H^{-1} = H^*$. Let the signal length be L and $B = FH^{-1}$. Then the periodogram is:

$$\begin{aligned} P_D &= \frac{1}{L}\text{diag}\left(|X_D|^2\right) \\ &= \frac{1}{L}\text{diag}\left(B|X_H|^2 B^{-1}\right) \end{aligned}$$

Now consider an M-point signal $x \triangleq [x_1, x_2, \cdots x_M]$ where each sub-sequence is L samples long. We build the average periodogram estimators:

$$\begin{aligned} \hat{P}_D &= \frac{1}{M}\sum_{m=1}^{M} P_D(x_m) \\ &= \frac{1}{ML}\sum_{m=1}^{M}\text{diag}\left(|X_D|^2\right) \end{aligned} \quad (3)$$

or

$$\hat{P}_H = \frac{1}{M}\left(\sum_{m=1}^{M}\frac{1}{L}\text{diag}\left(|BX_H|^2\right)\right) \quad (4)$$

From our derivation of P_D, one can see that $\hat{P}_D$ of Eq. (3) and $\hat{P}_H$ of Eq. (4) are different estimates of the same spectrum. *Our purpose here in this paper is to compare their relative performances.* The Periodogram estimator $\hat{P}_D$ is extremely well-known and studied. All of its statistical foibles are unaltered in this presentation. For purposes that will become evident slightly later, we define a dictionary D_1 that is a realization of the ideal power spectra of a pure tone (single sinusoid). D_1 realizes the periodogram estimate if every component is used. Similarly, from the definition of $\hat{P}_H$, we build D_H suitable for $\hat{P}_H$ that incorporates the sequency to frequency conversion, i.e.

$$D_H = \sqrt{L}B^{-1}$$

The $\sqrt{L}$ scale merely normalizes the energy in each element. For our analysis, we use the over-complete dictionary $D_2 = [D_1 \;\; D_H]$. For convenience, we use the acronym DFT to denote estimation using D_1 only (the periodogram), and HF to denote estimation using $D_2 = [D_1 \;\; D_H]$ in an orthogonal matching pursuit [3] algorithm. First, however, we must present a complete spectral analysis method.

Quinn studied the power spectrum problem using the DFT in the 1990's [4]. In that work, sinusoids in noise are modeled using an ARMA(p,q), where p and q are order of the AR (Autoregressive) part and MA (Moving Average) part respectively:

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

The parameters of the hybrid model are the a_m and b_m, $x[n]$ is the time series data for $n \in [0, N-1]$, and $\nu[n]$ is white noise with zero mean and finite variance. The signal is assumed to be (with discrete frequencies $\omega_k \in (0,\pi)$):

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

where the ρ_k and ϕ_k are the amplitude and initial phase of the k^{th} sinusoid. The special ARMA$(2p,2q)$ system that annihilates each sinusoidal component in $x[n]$ occurs when the parameters $a_m = b_m = \beta_m$ satisfy

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

and

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

Quinn's "smoothed periodogram" results when $x[n] = \rho\cos(\omega n + \phi) + \nu[n]$ is annihilated:

$$\kappa_N(\omega) = \int_{-\pi}^{\pi} P_x(\lambda)\mu_N(\omega - \lambda)d\lambda \quad (5)$$

Here, $P_x(\lambda)$ is the periodogram of $x[n]$ and $\mu_N(\omega) = \sum_{k=1}^{N-1} k^{-1}\cos(k\omega)$ is the kernel (or 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. (5) gives accurate frequency estimation without significant zero padding while simultaneously removing nearly all side lobes (refer to Fig. 2). Sometimes, this method is referred to as an interpolated periodogram.

Our spectral estimation methods are thus developed by first using the orthogonal matching pursuit using D_1 to approximate Eq. (3), followed by the smoothed periodogram of (5) and comparing it to the new HOT-based method using D_2 in the same approach. The first approach is the "DFT" approach to spectral estimation, the second is the "HOT" approach. More details can be found in [6].

4. Simulations

We consider two pure tone signals, $s_1[n] = A_1\cos[2\pi f_1/f_s n]$ and $s_2[n] = A_2\cos[2\pi f_2/f_s n]$. The length of the signal, N, is set to 256 for efficiency of calculation. The sampling frequency $f_s = 1000$ Hz The signal to be estimated using the orthogonal matching pursuit method is $y[n] = s_1[n] + s_2[n] + v[n]$, where $v[n]$ is ZMWG noise. Our results are averaged over $T = 100$ different noise realizations. In all applications of orthogonal matching pursuit, the dictionary is redundant. The stopping criterion, i.e. how many elements are selected for the representation, is critical. As we see in Fig. 1, the $\hat{P}_D$ performance is worse than $\hat{P}_H$ for all selections (<40) that include more than 6 elements, only achieving parity near the complete $L = 256$. In this case, $f_1 = 61$ Hz, $f_2 = 68.8$ Hz, the SNR is 15 dB, and the amplitude ratio A_1/A_2 is $1/2$. After applying orthogonal matching pursuit the performance of the algorithm is determined via the Relative Power of Error (RPE):

$$RPE = \frac{1}{T}\sum_{i=1}^{T}\frac{\sum_{n=0}^{N-1}|y_i[n] - \hat{y}_i[n]|^2}{\sum_{n=0}^{N-1}|y_i[n]|^2}$$

where y_i and $\hat{y}_i$ are the measured and reconstructed signals with i^{th} noise respectively.

For each reconstruction, we apply Quinn's method followed by a peak picking process to determine the frequencies f_1 and f_2. Also, we interpolate the spectral estimates via zero padding to a level sufficient to eliminate any "picket fence" in Quinn's method. Both the DFT and HF have bad resolution when less than 10 dictionary elements are used. We use 18 elements in the following figures for convenience.

Fig. 2 shows the spectral estimates for the case in Fig.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 elements are chosen, the lower frequency peak for the DFT is gradually shown, 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 HOT-DFT is more accurate than that from DFT. We have another example that shows an even more significant difference in [6]. The two PSD estimates are only identical when the two curves in Fig.1 coincide. Furthermore, the HF algorithm selects dictionary elements from D_H after the first few selections (that come from D_1). In fact, when 25 dictionary elements are used, over 40% of the selected dictionary elements are from D_H , and this increases to nearly 50% by the time 50 elements are used in the algorithm.

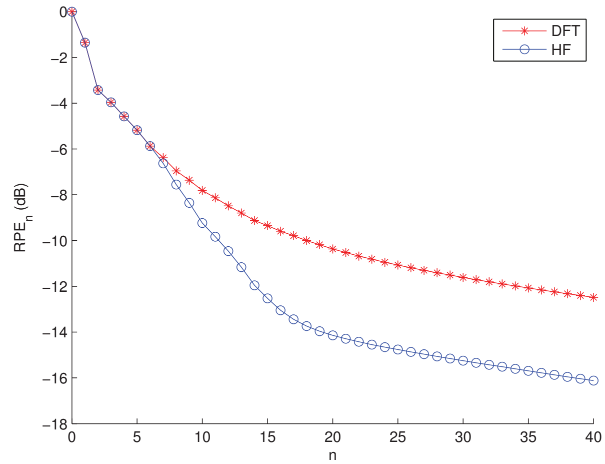

Figure 1. RPE (dB) vs. n (Number of dictionary elements)

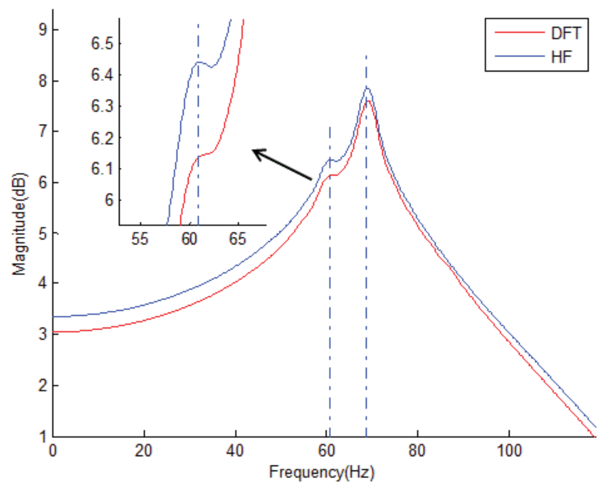

Figure 2. Power Spectrum Densities for $f_1 = 61$ Hz, $f_2 = 68.8$ Hz

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

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

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. $MSE_n = \frac{MSE}{f^t f}$.

The frequency separation $\triangle f$ is 7.8 Hz in Fig.2. If we fix $f_1 = 61$Hz, and vary $\triangle f$ between 6.5 Hz and 9 Hz, we observe the frequency resolution in a different way. Fig.3 shows that the MSE_n of the HF is much smaller than that of the DFT with SNR=15 dB when 18 elements are selected from the dictionary, especially $7.7 \leq \triangle f \leq 8$ Hz. When $\triangle f$ is less than 7.6 Hz, both methods degrade, though the DFT performance decrease is more severe. For frequency separations greater than 8 Hz, the methods behave similarly as expected though HF yields more accurate estimation. Note that when MSE_n is 0dB, the frequency is missed, i.e. there is no peak as in the case of the lower frequency 61Hz for the DFT case of Fig.2.

The MSE_n 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 perfom better in moderate and low SNR environment is very important in practical applications. Suppose that we change the SNR from -2 dB to 30 dB while keeping all other parameters unaltered; this comparison is shown in Fig. 4. We can clearly see that the performance of HF is always better than that of the DFT over the entire SNR range. We find that with increasing SNR, the MSE_n of the HF drops at a much lower SNR. Even when the $SNR \geq 17$ dB the HF still performs better than does the DFT. Changing the base (lower) frequency does not alter the relative performance substantially.

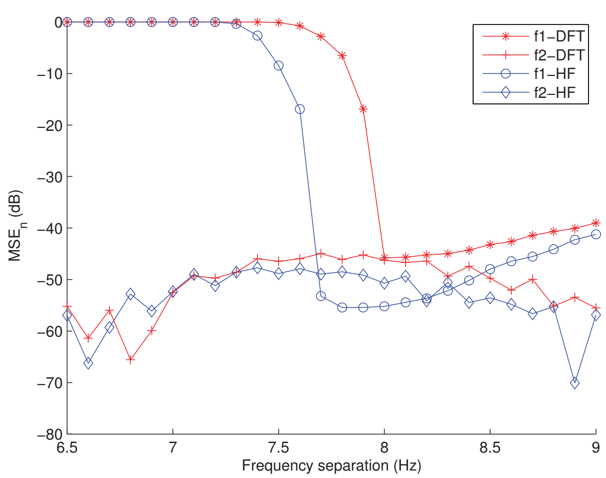

Figure 3. MSE_n of two frequency components with increasing frequency separation $\triangle f$

5. Conclusions

This paper introduces a method of nonparametric spectral estimation based on the HOT that uses orthogonal matching

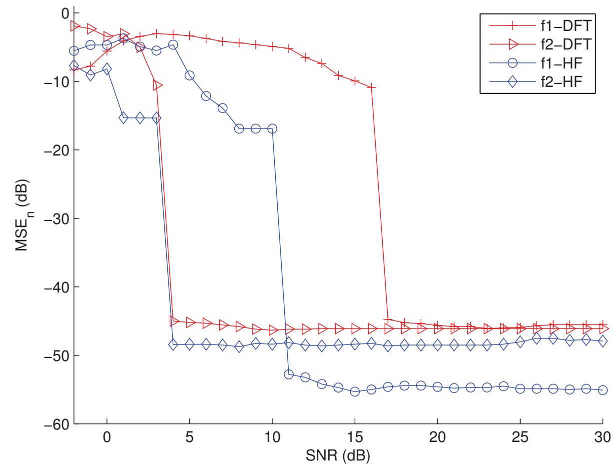

Figure 4. MSE_n of two frequency components with SNR

pursuit and Quinn's method. *Our results clearly show that the impact of our choice of transform is critical.* Specifically, we develop a dictionary generated with a combination of the HOT and DFT operations which we call the HF dictionary to calculate the smoothed periodogram with Quinn's method. *When compared to the DFT-only standard periodogram method, the power spectrum generated with our proposed method is superior over varying SNR, amplitude ratios, frequency separation, and number of elements used in the orthogonal matching pursuit algorithm.* We have answered our question in the introduction – using the entropy based HOT does improve the frequency resolution over the very similar DFT based spectral estimation techniques. Our main point is to compare the algorithm resolutions, not the computational complexity. Future work must be done to determine an automated method for stopping the orthogonal matching pursuit algorithm, as well as for the peak determination.

References

[1] T. Przebinda, V. E. DeBrunner, and M. Özaydın, "The optimal transform for the discrete hirschman uncertainty principle," *IEEE Trans. on Sig. Process.*, vol. 47, pp. 2086–2090, Jul. 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 $\ell^2(\mathbb{Z}/N\mathbb{Z})$," *IEEE Trans. Signal Process.*, vol. 53, pp. 2690–2699, Aug. 2005.

[3] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, pp. 3397–3415, 1993.

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

[5] V. DeBrunner, M. Özaydın, and T. Przebinda, "Resolution in time-frequency," *IEEE Trans. Sig. Process.*, vol. 47, pp. 783–788, Mar. 1999.

[6] G. Liu and V. E. DeBrunner, "Matching pursuits may yield superior results to orthogonal matching pursuits when secondary information is estimated from the signal model." Pacific Grove, CA: Proc. 44th Asilomar Conference on Signals, Systems and Computers, Nov. 2010.