Energy Spreading and Data Compression Using the Prometheus Orthonormal Set

J. S. Byrnes *B. Saffari* *H. S. Shapiro*
Prometheus Inc., Newport, RI, USA
Proceedings of the IEEE DSP Conference, Loen, Norway, 1996

ABSTRACT

The Prometheus Orthonormal Set (PONS^{tm}) can be effectively used to compress all common digital audio signals. This compression method is effective because of two fundamental properties, computational simplicity and energy spreading. Although there exist other transform coding methods, such as Walsh-Hadamard, which give compression while limiting the computational burden, we believe that the energy spreading feature of PONS is unique. We express this feature in precise mathematical terms and indicate how it is helpful for compression.

1. INTRODUCTION

We defined in [1] a "Walsh-like" complete orthonormal sequence for $L^2(0, 2\pi)$ which satisfies several important properties. Each function in this *Prometheus Orthonormal Set* (PONS^{tm}) is piecewise ± 1, can change sign only at points of the form $\frac{k}{2^n}(2\pi), 1 \leq k \leq 2^n - 1$, and is easily computable using a straightforward and fast recursive algorithm. In addition to these features shared with the Walsh functions, PONS:

- Is optimal with respect to the global uncertainty principle described in [1]
- Yields the *uniform* crest factor $\sqrt{2}$
- Spreads energy almost equally among all transform domain bins

We discuss the utilization of these properties below, for applications to one-dimensional digital signals. Required theoretical properties for multidimensional PONS have been derived and proven. Next we state several mathematical results, which give a rigorous foundation to our energy-spreading concept. Finally, we indicate several possible advantages of energy spreading for signal processing.

2. ONE-DIMENSIONAL SIGNALS

2.1. Audio processing

Our first implementation of PONS [3] yields high quality data compression for audio. For 16-bit 44.1, 22.05, and 11.025 kHz monaural signals we achieve almost 4 to 1 compression with virtually no audible difference in sound. The compression and decompression algorithms operate in real-time on all modern PCs. For example, it takes less than 3 seconds, including file I/O, to compress or decompress a 30-second 22 kHz 16-bit sound snippet using a Power PC 8100 or a Pentium 100. These results are to be contrasted with a current standard, ADPCM, where roughly the same compression ratio is achieved, but at much higher computational cost and with somewhat lower quality.

2.2. Spread spectrum communications

A second application of one-dimensional PONS is in multi-user spread spectrum communications. The recently developed IS-95 standard for commercial code-division multiple access (CDMA) communications involves a two-stage direct sequence spreading process, first with a Walsh function and then with a longer pseudonoise (PN) sequence. The energy spreading property of PONS offers two important advantages over the Walsh functions in this application:

- The minimal crest factor property of the PONS sequences provides much more uniform spreading of the signal's energy across the frequency band. This increases robustness with respect to channel effects, such as narrowband fading and interference, and also reduces spectral features that can be exploited by an adversary in military scenarios.
- Spreading with PONS sequences rather than Walsh sequences yields signals requiring lower short-term ("peak") power to maintain a specified average transmission power. This offers potential for reducing overall transmitter power requirements.

Since PONS sequences can be as long as desired (any power of two), it is possible to use them in place of long PN sequences as spreading codes. PN sequences can have arbitrarily long constant subintervals (i.e., consecutive terms all of which are 0 or all of which are 1), whereas PONS sequences cannot have constant subintervals of length greater than 5.

3. ENERGY SPREADING

It is straightforward to describe, in heuristic terms, what we mean by "energy spreading." Namely, when a digital signal of any dimension is expanded in the PONS basis, each of the terms in the transform domain has approximately the same amount of energy. It turns out to be natural to mathematically describe and analyze this property in terms of the Beurling minimal extrapolation norm.

3.1. Norm estimates and energy spreading

We consider orthogonal transform coding, $y^T = Ax^T$, where A is an $N \times N$ unitary matrix, T denotes transposition of matrices, $x = [x_1, x_2, \ldots, x_N]$ is the input signal, and $y = [y_1, y_2, \ldots, y_N]$ is the output signal. Our main interest is for Hadamard matrices A, which means all entries a_{ij} satisfy $|a_{ij}| = 1/\sqrt{N}$.

Before defining the Beurling norm, standard norms of N-tuples x which are important for our discussion are:

- The *Euclidean norm*, $\|x\|_E$
- The l^1 *norm*, $\|x\|_1$, and
- The *supremum norm* (or sup norm), $\|x\|_\infty$

We claim that an inequality

$$\|Ax\|_\infty \leq c\|x\|_\infty \qquad (1)$$

with c small (certainly, small with respect to $\sqrt{N}$ and, preferably, small like an absolute constant) is a good index of energy spreading, and we shall now clarify and amplify this assertion.

Let S denote a certain class of test signals, to be defined more precisely below. If no restrictions were put on S, then all we could say is

$$\|Ax\|_\infty \leq |||A||| \cdot \|x\|_\infty$$

where $|||A|||$ denotes $\max_i \sum_{j=1}^N |a_{ij}|$, and this quantity is $\sqrt{N}$ for a Hadamard matrix. Thus, in place of (1) we have the universal (but trivial) bound

$$\|Ax\|_\infty \leq \sqrt{N}\|x\|_\infty,$$

which holds for all Hadamard matrices A and all input signals x.

For any signal (N-tuple) y, Cauchy's inequality yields

$$\|y\|_1 \leq \sqrt{N}\|y\|_E. \qquad (2)$$

Suppose x has "perfect energy spreading," say $|x_i| = 1$ for all i; then $\|y\|_E = \|Ax\|_E = \|x\|_E = \sqrt{N}$ so (2) says $\|y\|_1 \leq N$. For equality to hold here, it is necessary and sufficient that $|y_i| = 1$ for all i, i.e., that y has also "perfect energy spreading." Hence we have the following criterion:

Given a class of test signals S, *each of which has all entries of modulus* 1, A preserves perfect energy spreading for this class of inputs if and only if $\|Ax\|_1 = N$, for all $x \in S$, and A has good energy spreading qualities if $\inf_{x\in S} \|Ax\|_1/N$ is "large," that is, close to its theoretical maximum value 1.

Now, for any signal y,

$$\|y\|_E^2 = \sum |y_i|^2 \leq \|y\|_\infty \cdot \|y\|_1$$

hence for $y = Ax$:

$$\begin{aligned} \|Ax\|_1 &\geq \frac{\|Ax\|_E^2}{\|Ax\|_\infty} \\ &= \frac{N}{\|Ax\|_\infty} \qquad x \in S. \end{aligned}$$

Thus,

$$\frac{\|Ax\|_1}{N} \geq \frac{1}{\|Ax\|_\infty} = \frac{\|x\|_\infty}{\|Ax\|_\infty} \qquad x \in S. \qquad (3)$$

This shows our energy-spreading parameter $\inf_{x\in S} \|Ax\|_1/N$ is close to 1 if $\|Ax\|_\infty \leq c\|x\|_\infty$ holds for all $x \in S$, where c is a constant (bigger than, but) close to 1. Thus, we see the importance of estimates of type (1).

3.2. The Wiener norm

For a bounded complex measure μ on the real line $\Re$, its Fourier (or Fourier-Stieltjes) transform, denoted F_μ, is defined by

$$(F_\mu)(t) := \int e^{-it\omega}\, d\mu(\omega). \qquad (4)$$

Definition 1 *For a complex valued function $M(t)$ on $\Re$, its Wiener norm, denoted $\|M\|_W$, is the total variation $V(\mu)$ of the measure μ, if one exists, satisfying $F_\mu = M$. If no such μ exists, $\|M\|_W = +\infty$.*

It is well known that, if such μ exists for a given M, it is unique, so $\|M\|_W$ is well defined. In particular, if $\hat{f}$, the Fourier transform of an integrable function f on $\Re$, is absolutely integrable over $\Re$, then $\|f\|_W = \int_{-\infty}^{\infty} |\hat{f}(\omega)|\, d\omega$.

An important fact, obvious from (4) is:

$$\sup_{t\in\Re} |M(t)| \leq \|M\|_W. \qquad (5)$$

3.2.1. Some examples

For $M(t) = e^{iat}$, $a \in \Re$, M is the Fourier transform of a unit mass at $\omega = -a$, so the Wiener norm of a pure harmonic oscillation e^{iat} is 1. Likewise the W norm of a superposition of these, $M(t) = \sum_{k=1}^{l} \alpha_k e^{i a_k t}$, is $\sum |\alpha_k|$. In particular, the sinusoid $\sin at$ has Wiener norm 1.

Another interesting example is when M is positive definite (in the sense of Bochner), that is $M = F_\mu$ where μ is non-negative. The total variation of μ is $\int d\,\mu(\omega) = M(0)$ (from (4)), so $\|M\|_W = M(0)$ for such functions. For example, in all the cases $M(t) = e^{-t^2}$, $e^{-|t|}$, $(1+t^2)^{-1}$ or $M(t) =$"the triangle function":

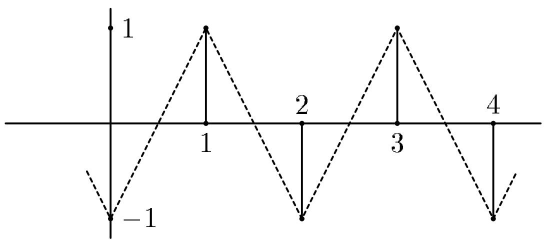

Figure 1. Graph of $M(t)$.

$$M(t) = (1 - |t|^+) = \begin{cases} 1 - |t|, & |t| \le 1 \\ 0, & |t| > 1 \end{cases}$$

we have $\|M\|_W = 1$. In particular, equality holds in (5) in these cases, although *generally* the W-norm is much larger than the supremum norm.

It is also easy to verify from the definition: *the W-norm does not change if we subject M to an affine change of variables:* $M(at+b)$, for fixed $a > 0$ and $b \in \Re$, has the same W-norm as M.

3.3. The Beurling minimal-extrapolation norm

Let $\varphi(t)$ be a complex-valued function defined on some nonempty subset E of $\Re$.

Definition 2 *The Beurling minimal extrapolation norm of φ, $\|\varphi\|_{\mathrm{ME}}$, is the infimum of $\|M\|_W$ over all functions M defined on $\Re$ which extrapolate φ, that is, satisfy $M(t) = \varphi(t)$, $t \in E$.*

Remarks 1 *A minimizing choice of M may or may not exist. Also, it can happen that no choice of M exists which is the Fourier transform of a bounded measure, and in this case $\|\varphi\|_{\mathrm{ME}} = +\infty$. Note that, in case $E = \Re$, $\|\varphi\|_{\mathrm{ME}} = \|\varphi\|_W$.*

Example 1 $E = (0, \infty)$, $\varphi(t) = e^{-t}$. *Here $\|\varphi\|_{\mathrm{ME}} = 1$. The extrapolation is $M(t) = e^{-|t|}$.*

Definition 3 *For an N-tuple $x = [x_1, x_2, \ldots, x_N]$ its* ME*-norm, denoted $\|x\|_{\mathrm{ME}}$, is $\|\varphi\|_{\mathrm{ME}}$ where φ is the function defined on the set $\{1, 2, \ldots, N\}$ by $\varphi(n) = x_n$.*

Example 2 $N = 4$, $x = [1, -1, 1, -1]$. *Here* $\|x\|_{\mathrm{ME}} = 1$.

Indeed, we can extrapolate φ by $M(t)$, a periodic "sawtooth" function as shown in Figure 1. Here $\|M\|_W = 1$, either by direct calculation of the Fourier coefficients of M, or by utilizing that $-M$ is positive definite.

Example 3 *If $\phi(t)$ is a linear function on $[a, b]$, then $\|\phi\|_{\mathrm{ME}} = \sup|\phi(t)|$ $(= max|\phi(a)|, |\phi(b)|)$.*

3.4. PONS, the Beurling norm, and energy spreading

As noted in 3.1,

$$\|Ax\|_\infty \le \sqrt{N}\|x\|_\infty, \tag{6}$$

which holds for all Hadamard matrices A and all input signals x. Moreover, it is easy to see that there are signals x for which equality holds in (6). However, *for a fairly broad class of signals x, a significant improvement of (6) is possible, in the case where A is the PONS matrix.* This improvement is embodied in the inequality

$$\|y\|_\infty \le \sqrt{2}\|x\|_{\mathrm{ME}} \qquad \text{for PONS} \tag{7}$$

Since, for a wide class of signals x, the ME norm is equal to the sup norm, or only exceeds it by a modest factor, it is clear that (7) represents in such cases a significant improvement of (6). Actually, (7) does not tell the whole story. A far-reaching extension of (7), based on a deeper property inherent in the structure of the PONS matrix, implies that, for a wide class of signals (for example, all x where $x_n = \pm 1$) *with only a small number of sign changes*, even though $\|x\|_{\mathrm{ME}}$ may be large, $\|y\|_\infty$ in fact remains bounded; thus (7) may be essentially strengthened for this kind of signal. The following theorems and examples illustrate this point.

Theorem 1 *If x is a signal consisting of ± 1 entries, and exactly $r - 1$ sign changes, its* PONS *transform satisfies*

$$\|y\|_\infty \le 3\sqrt{6/5}\sqrt{r}. \tag{8}$$

Theorem 2 *Let $x = [x_1, \ldots, x_N]$ be a signal, and split it into r blocks of consecutive terms whose successive lengths are $\lambda_1, \lambda_2, \ldots, \lambda_r$ (thus, $\lambda_j \ge 1$ and $\sum_{j=1}^r \lambda_j = N$). We may write this symbolically as*

$$x = [x^{(1)}|x^{(2)} \cdots |x^{(r)}]$$

where $x^{(i)}$ is the i^{th} block. (Thus, $x^{(1)} = [x_1, x_2, \cdots, x_{\lambda_1}]$, etc.)

Let K_i denote $\|x^{(i)}\|_{\mathrm{ME}}$. Then the PONS transform $y^T = Ax^T$ satisfies

$$\|y\|_\infty \le 3\sqrt{6/5}\left(\sum_{i=1}^r K_i^2\right)^{1/2}. \tag{9}$$

Example 4 *Suppose x consists of a block of sinusoidal terms at one frequency, followed by a second block of a different frequency, thus:*

$$x = [x'|x'']$$

where

$$\begin{aligned} x' &= [a\cos(\omega_1 + b), \ldots, a\cos(k\omega_1 + b)] \\ x'' &= [c\cos(\omega_2 + d), \ldots, c\cos(l\omega_2 + d)] \end{aligned}$$

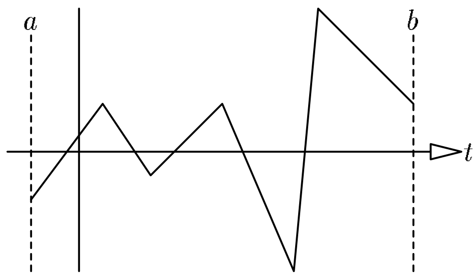

Figure 2. An example on $[a,b]$ of $f(t)$ consisting of 6 linear pieces ($r=6$).

where $a,b,c,d,\omega_1,\omega_2$ are arbitrary and $k+l=N$. As $\|x'\|_{\mathrm{ME}}=a$ and $\|x''\|_{\mathrm{ME}}=c$, (9) yields

$$\|y\|_\infty \le 3\sqrt{6/5}\sqrt{a^2+c^2}.$$

Further pursuit of this idea leads to an interesting and probably difficult new problem in harmonic analysis: to optimize the estimate (9) for a given signal x. In other words, *How does one break x into blocks of consecutive terms, such that the sum of squares of the* ME*-norms of these blocks is minimal? What is a good upper bound for this minimum, e.g., in terms of other, readily computable parameters?*

Example 5 *Suppose x is a signal obtained by equi-spaced sampling of a piecewise linear function $f(t)$ consisting of r (continuously) formed linear pieces (see Figure 2.) Then, the PONS transform y of x satisfies*

$$\|y\|_\infty \le 3\sqrt{6/5}\sqrt{r}\sup_t |f(t)| \qquad (10)$$

What is remarkable about (10) is that a sup-norm appears on the right (at the cost of the $\sqrt{r}$ factor). The ME-norm of the signal x (or the function f) will in general be much larger than the right side of (10), especially if r is small.

Estimates such as those given above are very finely tuned to specific structural features of the PONS transform, and no other forms of orthogonal coding known to us have such properties. In a rough qualitative way one can say that if x and x' are signals with which the PONS transform copes well (such as pure harmonic, purely linear, purely quadratic, etc.), then it copes well also with the signal consisting of some portion of x followed by some portion of x' (and likewise for a composite of $r \ll N$ "good" signals). The transient effect of abrupt switching, at an arbitrary instant, from x to x' does not elevate the output amplitude by much (e.g., switching from samples of $\cos\omega_1 t$ to samples of $\cos\omega_2 t$ can only bring in an amplitude factor of $\sqrt{2}$ in the output, regardless of the values of ω_1, ω_2).

3.5. Advantages of energy spreading

In addition to the advantages for spread spectrum communications given in 2.2, PONS energy spreading yields the following desirable properties:

- Added security, in that it makes a signal look like white noise
- Low crest factor array and optimal uncertainty principle bounds [1]
- Near optimum coding gain for sufficiently wideband input (independent of input statistics or spectrum)
- Quantizes all transform domain coefficients to the same precision
- Aliasing errors are canceled to the same precision in all transform domain subspaces
- Quantization errors appearing in the reconstruction are approximately Gaussian (as a linear sum of independent uniform errors)
- Eliminates bit allocation computations
- Good performance over unreliable media

To explain the final item above, suppose there is some kind of bursty noise which substantially degrades transmission at isolated and unpredictable times. No matter which PONS terms were lost because of such noise, there would be only very gradual signal degradation (i.e., slightly more degradation with each term that is lost). For all other known transforms (wavelet, DCT, Walsh-Hadamard, etc.), if low energy terms were lost because of such noise there would also be very little signal degradation, but if even one high energy term was lost the degradation would probably be quite substantial.

Also, PONS whitens any "signal," including "noise." Thus, if the noise energy is less than the energy of the desired signal, under quantization (of the PONS coefficients) the noise energy should be eliminated before the signal energy.

REFERENCES

[1] J.S. Byrnes, "Quadrature mirror filters, low crest factor arrays, functions achieving optimal uncertainty principle bounds, and complete orthonormal sequences – a unified approach," *Applied and Computational Harmonic Analysis*, vol. 2, pp. 261–266, 1994.

[2] J.S. Byrnes, "A low complexity energy spreading transform coder," *Proc. 1994 Conf. on Image Processing*, Haifa, Y. Zeevi, Ed., 1996

[3] James S. Byrnes, Michael A. Ramalho, Gerald Ostheimer and Izidor Gertner, Discrete one dimensional signal processing method and apparatus using energy spreading coding. U.S. Patent pending.

[4] H.S. Shapiro, Extremal problems for polynomials and power series. Sc.M. Thesis, Massachusetts Institute of Technology, 1951.