Estimates for Green's functions of Schrödinger operators; also, a pure mathematician's adventures in wavelet applications

Michael Frazier
Mathematics Department
University of Tennessee

Talk at Oak Ridge National Laboratories on 11/6/07

Consider the operator $\mathcal{L} = \mathcal{L}_\alpha = (-\triangle)^{\alpha/2} - q$, and the equation

$$\mathcal{L}u = (-\triangle)^{\alpha/2}u - qu = \varphi \quad \text{in } \Omega \subseteq \mathbb{R}^n,$$

$$u = 0 \quad \text{on } \partial\Omega,$$

where $q \in L^1_{loc}(\Omega)$. We could have $\Omega = \mathbb{R}^n, 0 < \alpha < n$, $\Omega =$ half-plane, or e.g., $\Omega =$ bounded Lipschitz domain, with $0 < \alpha \leq 2$.

Remark: Kalton and Verbitsky studied the existence of solutions $u \geq 0$ to

$$(-\triangle)^{\alpha/2}u - qu^s = \varphi,$$

with $s > 1$. Curiously, their results did not apply to the linear case $s = 1$.

Main Equation:

$$\mathcal{L}u = (-\triangle)^{\alpha/2}u - qu = \varphi$$

If $\alpha = 2$, then $\mathcal{L}$ is the time-independent Schrödinger operator $-\triangle - q$.

Remark: Case $\alpha \neq 2$ is of interest in probability, where $(-\triangle)^{\alpha/2}$ corresponds to α-stable Lévy processes in the same way that $-\triangle$ corresponds to Brownian motion.

Joint work with Igor Verbitsky

Main Equation

$$\mathcal{L}u = (-\triangle)^{\alpha/2}u - qu = \varphi \quad \text{in } \Omega \subseteq \mathbb{R}^n$$

Let $G = G^{(\alpha)}$ be the Green's operator for $(-\triangle)^{\alpha/2}$ on Ω,

$$G^{(\alpha)}(f)(x) = \int_\Omega G^{(\alpha)}(x,y)f(y)\,dy,$$

and we assume $G^{(\alpha)}(x,y) \geq 0$. If $\Omega = \mathbb{R}^n$ then $G^{(\alpha)}$ is the Riesz potential $I^{(\alpha)}$ with kernel $c_n|x-y|^{\alpha-n}$; for $\alpha = 2$ and $n \geq 3$, G is the Newtonian potential.

We apply G to both sides of the Main Equation to obtain

$$u - G(qu) = G(\varphi) \equiv f$$

Supposing $q \geq 0$, let

$$T(u)(x) = G(qu)(x) = \int_\Omega G(x,y)u(y)q(y)\, dy$$

(if q is not non-negative, we let $T(u) = G(|q|u)$ and use this to obtain upper bounds). Then we have $u - T(u) = f$, or $(I - T)(u) = f$, so the formal solution is

$$u = (I - T)^{-1}(f) = \sum_{j=0}^{\infty} T^j(f).$$

If $f \in L^2(|q(y)|dy)$ and

$$\|T\|_{L^2(|q(y)|dy) \to L^2(|q(y)|dy)} < 1,$$

then $\sum_{j=0}^{\infty} T^j f$ converges in $L^2(\omega)$ to a solution. Our goal is pointwise estimates.

So

$$Tf(x) = \int_\Omega G(x,y) f(y) |q(y)|\, dy$$

$$= \int_\Omega G(x,y) f(y) d\omega(y),$$

for $d\omega(y) = |q(y)|\, dy$. Then

$$T^j f(x) = \int_\Omega G_j(x,y) f(y)\, d\omega(y),$$

where

$$G_1(x,y) = G(x,y),$$

and, for $j > 1$, inductively define

$$G_j(x,y) = \int_\Omega G(x,z)G_{j-1}(z,y)\,d\omega(z).$$

Let $V(x,y) = \sum_{j=1}^\infty G_j(x,y)$, so that

$$u(x) = f(x) + \sum_{j=1}^\infty T^j f(x)$$

$$= f(x) + \int_\Omega V(x,y)f(y)\,d\omega(y).$$

Our goal is to estimate V, the minimal Green's function for $\mathcal{L}$.

Theorem A (lower bound). Let $q \geq 0$. Then there exist $c_1, c_2 > 0$ such that

$$V(x,y) \geq c_1 G(x,y) e^{c_2 G_2(x,y)/G(x,y)}.$$

Remark: Theorem A is relatively easy. The more interesting result of the paper is that under a certain smallness condition on q, we obtain upper bounds of the same form.

Remark: In pretty general circumstances, there is a formula

$$V(x,y) = G(x,y)\mathbb{E}_y^x\left[e^{\frac{1}{2}\int_0^\zeta q(X_s)\,ds}\right],$$

where X_t is Brownian motion, if $\alpha = 2$, or an α-stable symmetric process, if $0 < \alpha < 2$, conditioned to start at x and end at y, and ζ is its lifetime. The quantity $\mathbb{E}_y^x\left[e^{\frac{1}{2}\int_0^\zeta q(X_s)\,ds}\right]$ is called the conditional Feynman-Kac gauge. Our results give estimates for the conditional gauge.

Theorem A follows from a general result about "quasi-metric kernels." Let (Ω, ω) be a σ-finite measure space. A function

$$K : \Omega \times \Omega \longrightarrow (0, \infty]$$

is a quasi-metric kernel on Ω if $K(x, y) = K(y, x)$ for all $x, y \in \Omega$, $K(x, y) < \infty$ if $x \neq y$, and there exists $\kappa > 1/2$ such that $d(x, y) = 1/K(x, y)$ satisfies

$$d(x, y) \leq \kappa(d(x, z) + d(z, y)), \qquad x, y, z \in \Omega,$$

for some $\kappa \geq 1/2$ (we don't require $d(x, x) = 0$).

Theorem A': Let K be a quasi-metric kernel on a σ-finite measure space (Ω, ω). Let $K_1 = K$ and inductively define

$$K_j(x,y) = \int_\Omega K(x,z) K_{j-1}(z,y)\, d\omega(z).$$

Then there exists c_2 depending only on κ such that

$$V(x,y) \equiv \sum_{j=1}^{\infty} K_j(x,y) \geq K(x,y) e^{c_2 K_2(x,y)/K(x,y)}$$

Sometimes Theorem A' implies Theorem A directly, with $K = G$. E.g., for $\Omega = \mathbb{R}^n$, $0 < \alpha < n$, then $G(x, y) = c_n |x - y|^{\alpha - n}$ is quasi-metric. However, for domains Ω, G may not be quasi-metric. But, for very general domains (including all bounded Lipschitz domains), there exists a function $m > 0$ on Ω such that

$$H(x, y) = \frac{G(x, y)}{m(x) m(y)}$$

is a quasi-metric kernel. To get Theorem A, apply Theorem A' with $K = H$ and the measure $d\nu = m^2 d\omega$, noting that $H_2 / H = G_2 / G$.

Now let's consider the upper estimate for $V(x,y)$. To see what is appropriate, recall the equation

$$u = T(u) + f,$$

where $T(u)(x) = \int_\Omega G(x,y)u(y)d\omega(y)$. If there exists $f \geq 0$ such that there is a solution $u > 0$ to $u = T(u) + f$, then $T(u)(x) \leq u(x)$ for all x.

Then, by Schur's Lemma,

$$\|T\|_{L^2(\omega)\to L^2(\omega)} \leq 1.$$

If we test the norm on χ_E, we obtain

$$\int_{E\times E} G(x,y)d\omega(x)d\omega(y) \leq \omega(E)$$

for all measurable $E \subset \Omega$.

The first condition is invariant: if $H(x,y) = G(x,y)/(m(x)m(y)), d\nu = m^2 d\omega$, and $S(u)(x) = \int_\Omega H(x,y)u(y)d\nu(y)$, then

$$\|S\|_{L^2(\nu)\to L^2(\nu)} = \|T\|_{L^2(\omega)\to L^2(\omega)}.$$

The second condition is not invariant. Define $\|\omega\|$ to be the smallest constant C such that

$$\int_{E\times E} G(x,y)m(x)m(y)\,d\omega(y) \le C \int_E m^2(z)\,d\omega(z)$$

for all measurable $E \subseteq \Omega$, where m is as above.

We define $\|\omega\|_*$ similarly, except that we only consider balls B instead of general sets E.

Theorem B (upper bound) Let Ω and α be as above. Then there exists $\epsilon > 0$ such that if either

(i) $\|T\|_{L^2(\omega)\to L^2(\omega)} < \epsilon$,

(ii) $\|\omega\| < \epsilon$,

or

(iii) ω is a doubling measure and $\|\omega\|_* < \epsilon$,

then there exist $c_3, c_4 > 0$ such that

$$V(x,y) \leq c_3 G(x,y) e^{c_4 G_2(x,y)/G(x,y)}.$$

As in Theorem A, the result follows from an abstract result for quasi-metric kernels.

Remark: There is interest in when q is sufficiently mild that $V(x, y) \approx G(x, y)$. We always have $V(x, y) \geq G(x, y)$, and by above, when (i), (ii), or (iii) holds, we obtain $V \approx G$ if $G_2 \leq cG$. This holds e.g., under the Kato condition.

Probably the main interest in our results is in the case where V is not equivalent to G.

Example: Let $\Omega = \mathbb{R}^n$, $0 < \alpha < n$, and let $q = \frac{A}{|x|^\alpha}$, for some constant A, so

$$\mathcal{L} = (-\triangle)^{\alpha/2} - \frac{A}{|x|^\alpha}.$$

Then there exists $\epsilon > 0$ such that for $0 \leq A < \epsilon$,

$$c_1 \frac{\left(\max\left\{\left|\frac{x}{y}\right|, \left|\frac{y}{x}\right|\right\}\right)^{c_2}}{|x-y|^{n-\alpha}} \leq V(x,y) \leq c_3 \frac{\left(\max\left\{\left|\frac{x}{y}\right|, \left|\frac{y}{x}\right|\right\}\right)^{c_4}}{|x-y|^{n-\alpha}},$$

for some $c_1, c_2, c_3, c_4 > 0$.

Proof: After some elementary computations, obtain

$$G_2(x,y)/G(x,y) \approx 1 + \log\left(\max\left\{\left|\frac{x}{y}\right|, \left|\frac{y}{x}\right|\right\}\right).$$

Apply Theorems A and B. Here ω is doubling, and one can check the condition $\|\omega\|_* \leq CA^2$.

Note that we don't get sharp powers, and we require the smallness condition on q. However, our results work for very general Ω and for a range of α.

In the literature, q is often assumed bounded, or very nice. From our results, we can see what singularities of q are feasible, and the general form of V.

Of course, estimates on V yield solvability results for the original equation $\mathcal{L}_\alpha u = \varphi$, because $u(x) = G(\varphi)(x) + \int_\Omega V(x,y) G(\varphi)(x) q(y)\, dy$.

Comments on proof of abstract theorem on quasi-metric kernels:

Actually show

$$V(x,y) \approx \int_{d(x,y)}^{\infty} \frac{e^{c(G_t(x)+G_t(y))}}{t^2}\, dt \approx K e^{c_1 K_2 / K},$$

where

$$G_t(x) = \int_0^t \frac{\omega(B_r(x))}{r^2}\, dr.$$

$$V(x,y) \approx \int_{d(x,y)}^{\infty} \frac{e^{c(G_t(x)+G_t(y))}}{t^2}\, dt$$

For the lower estimate, prove inductively that

$$K_j(x,y) \geq c^{j-1} \int_{d(x,y)}^{\infty} \frac{G_t(y)^{j-1}}{(j-1)!t^2}\, dt,$$

using an integration by parts. Then sum on j, obtaining $e^{cG_t(y)}$, use symmetry to get $e^{cG_t(x)}$, average, and use inequality between arithmetic and geometric means.

$$V(x,y) \approx \int_{d(x,y)}^{\infty} \frac{e^{c(G_t(x)+G_t(y))}}{t^2}\,dt$$

For the upper estimate, inductively prove

$$K_j(x,y) \le c_1 \left(\frac{c_2}{\beta}\right)^{j-1} \int_{d(x,y)}^{\infty} \frac{e^{\beta(G_t(x)+G_t(y))}}{t^2}\,dt.$$

Then for β large enough, sum on j. For the induction, need an estimate of the form

$$\int_{B_t(x)} e^{\beta G_t}\,d\omega \le c\omega(B_{2t}(x)).$$

We get this from

$$\int_{B_t(x)} G_t^m\,d\omega \le m!C^m\|\omega\|^m\omega(B_{2t}(x)).$$

We need $\|\omega\| < \epsilon$ to sum on m.

Transient Signal Detection

(Daniel Wagner Associates, 1990-1991)

Have noisy environment generating random discrete noise signals

$$n = (n(0), n(1), n(2), ..., n(N-1)).$$

Assume n is a jointly Gaussian random variable.

Want to detect a given prototype signal

$$s = (s(0), s(1), s(2), ..., s(N-1)).$$

Suppose we receive signal

$$x = (x(0), x(1), x(2), ..., x(N-1)).$$

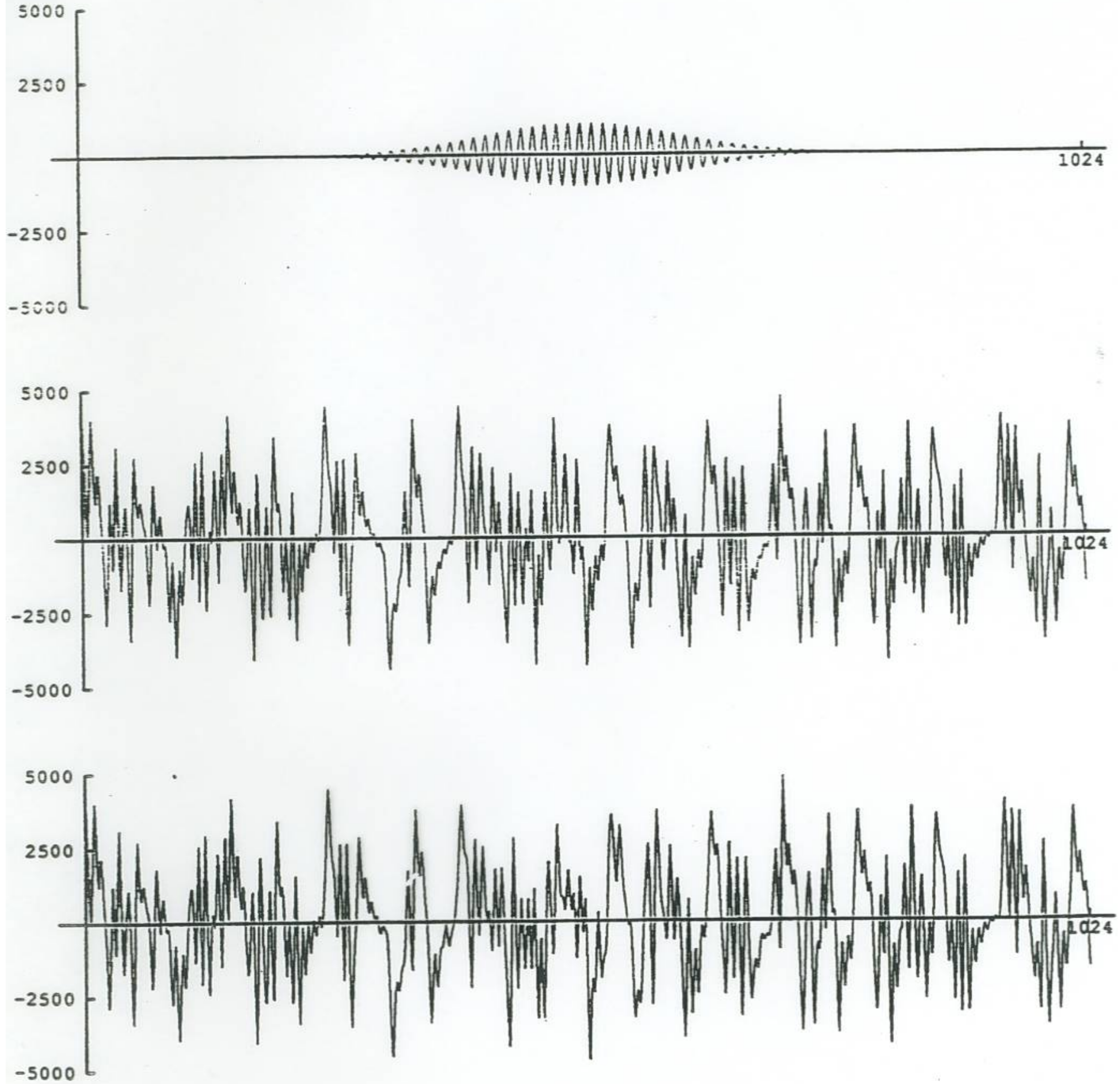
There are two possibilities:

$H_0 : x = n$ (null hypothesis: s not present in x)

$H_1 : x = s + n$ (s is present in x).

Need: hypothesis test.

Signal Plus Noise 9.1



A "decision criterion" assigns to each x a conclusion, either H_0 or H_1.

There are two types of mistakes:

False alarm: Accept H_1 when H_0 is true

Detection failure: Accept H_0 when H_1 is true

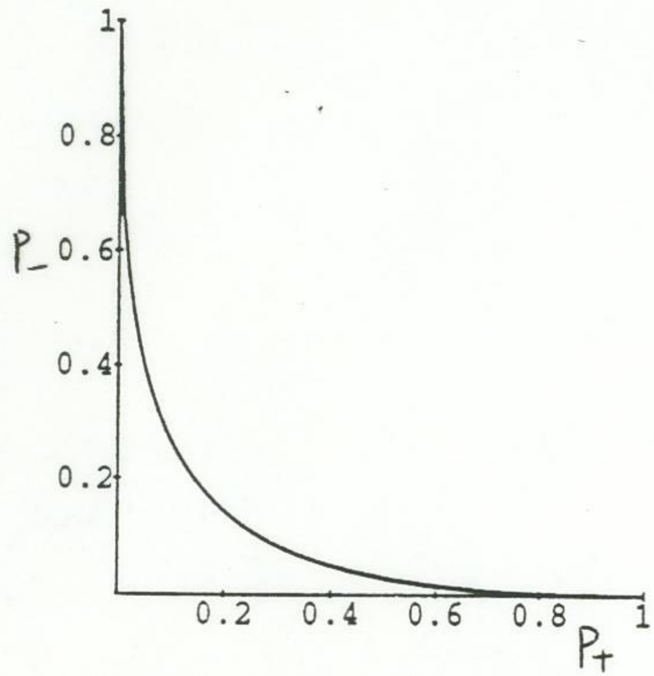
Let $P_+ = \text{Prob(False alarm)}$

and $P_- = \text{Prob(Detection failure)}$.

Want to be able to assign P_+.

A "test" assigns to each $P_+ \in [0, 1]$ a decision criterion having that value of P_+.

To evaluate a test, plot P_- as a function of P_+:



This is called the ROC (Receiver Operating Characteristic) curve for the test. One test is clearly superior to another if its ROC curve is everywhere lower.

A well-known, relatively elementary statistics result states that there is an optimal test for this problem, i.e., a test with lower ROC curve than any other test. This test is usually called the "matched filter" test.

Matched Filter Test

Let $R = (R_{i,j})_{i,j=0,1,2,3,\ldots,N-1}$ be the noise correlation matrix, i.e.,

$$R_{i,j} = E(n(i)n(j)).$$

Let $\lambda = \langle R^{-1}s, x\rangle = \sum_{i=0}^{N-1}(R^{-1}s)(i)x(i)$.

Matched filter test: Accept $H_1 \iff \lambda \geq \beta$.

Choice of parameter β determines P_+, hence a point on the ROC curve.

Problem: May have very large number of prototype signals s we want to test for. May not be able to compute all the test statistics

$$\lambda = \langle R^{-1}s, x \rangle = \sum_{i=0}^{N-1} (R^{-1}s)(i)x(i).$$

in real time.

Plan: Compress the test. Select a subset M of $0, 1, 2, ..., N-1$, and compute

$$\lambda_M = \sum_{i \in M} (R^{-1}s)(i)x(i).$$

If most of the information is contained in a small number of terms, and those terms are included in the sum by the choice of M, we can get a good approximation to the optimal test with much less computation. But the information may not be concentrated in a few terms in the standard basis.

Idea: change basis for the compressed test. Apply any orthogonal change of basis to $R, s,$ and x, to get $\widehat{R}, \widehat{s},$ and $\widehat{x}$. Then can compute λ in new basis:

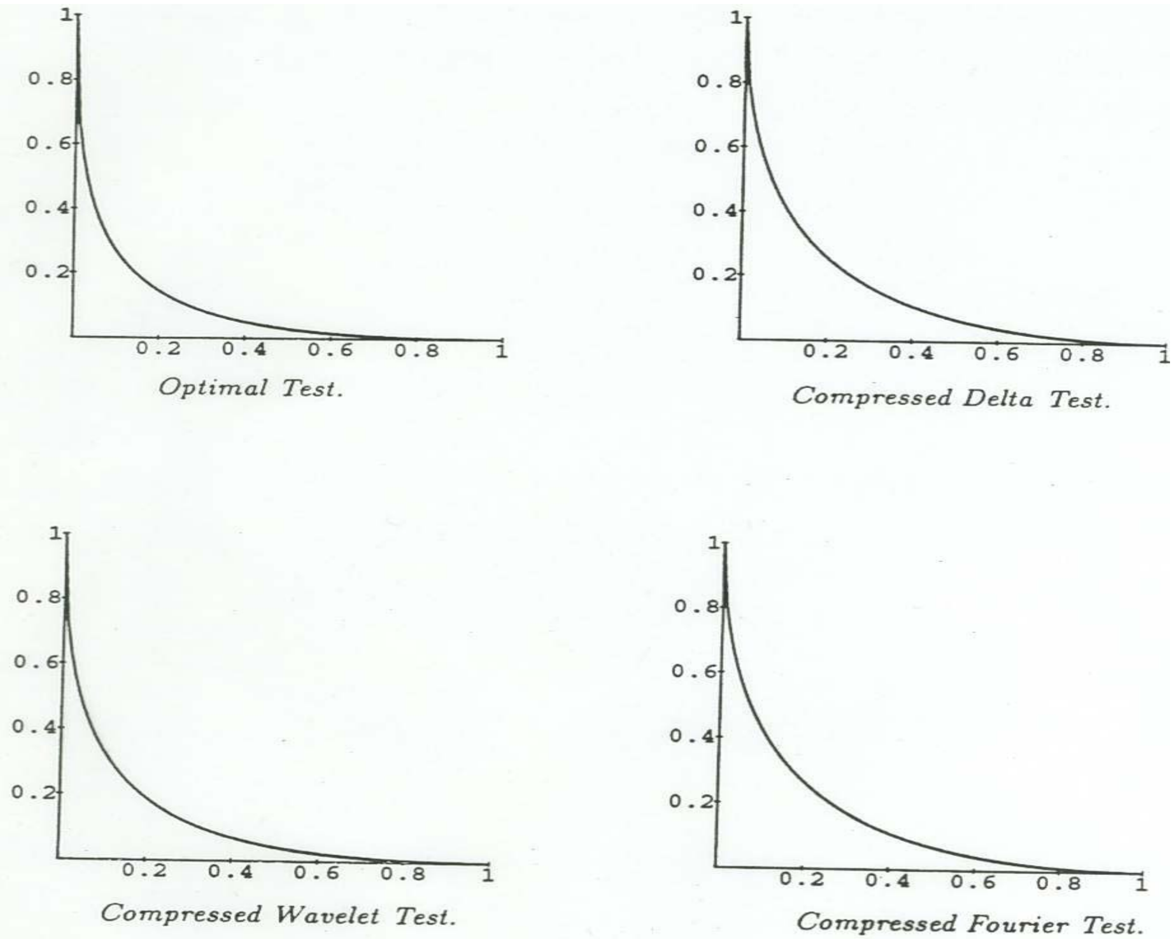
$$\lambda = \langle R^{-1} s, x \rangle = \langle \widehat{R}^{-1} \widehat{s}, \widehat{x} \rangle.$$

So full test is basis invariant (must be, since optimal). But compression of optimal test is not basis invariant.

Heuristic: If prototype signal is localized in space (a transient signal) and has definite frequency characteristics, and/or if the noise has definite frequency characteristics, then a wavelet basis should do a better job of compressing the optimal test.

Examples: $N = 64$, Cardinality of $M = 4$

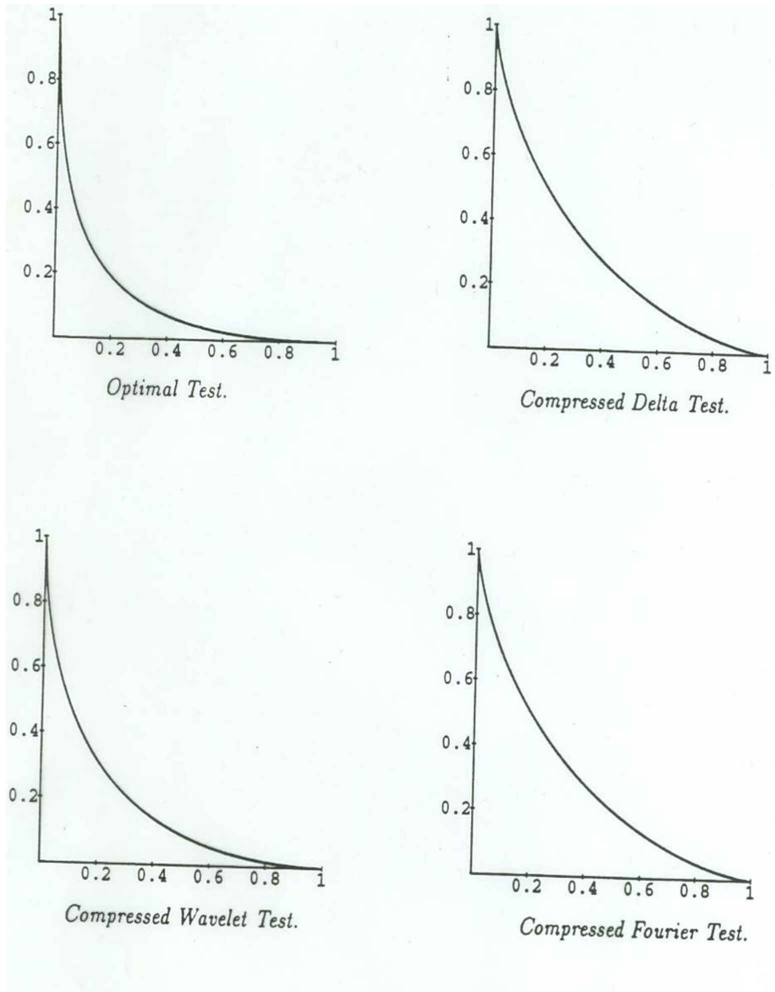
1.) $s(n) = .8\cos(n\pi/6)e^{-(n-31)^2/32}$, **white noise**



Example 1 ROC Curves.

Examples: $N = 64$, Cardinality of $M = 4$

2.) $s(n) = .8e^{-(n-31)^2/4}$, correlated noise



Example 2 ROC Curves.

Car Rattles and the Shift-Invariant Discrete Wavelet Transform

2001 Michigan State University

Ford Representative: David Scholl

Masters in Industrial Mathematics Team:

Joerg Enders

Weihua Geng

Peijun Li

Faculty Advisor: Michael Frazier

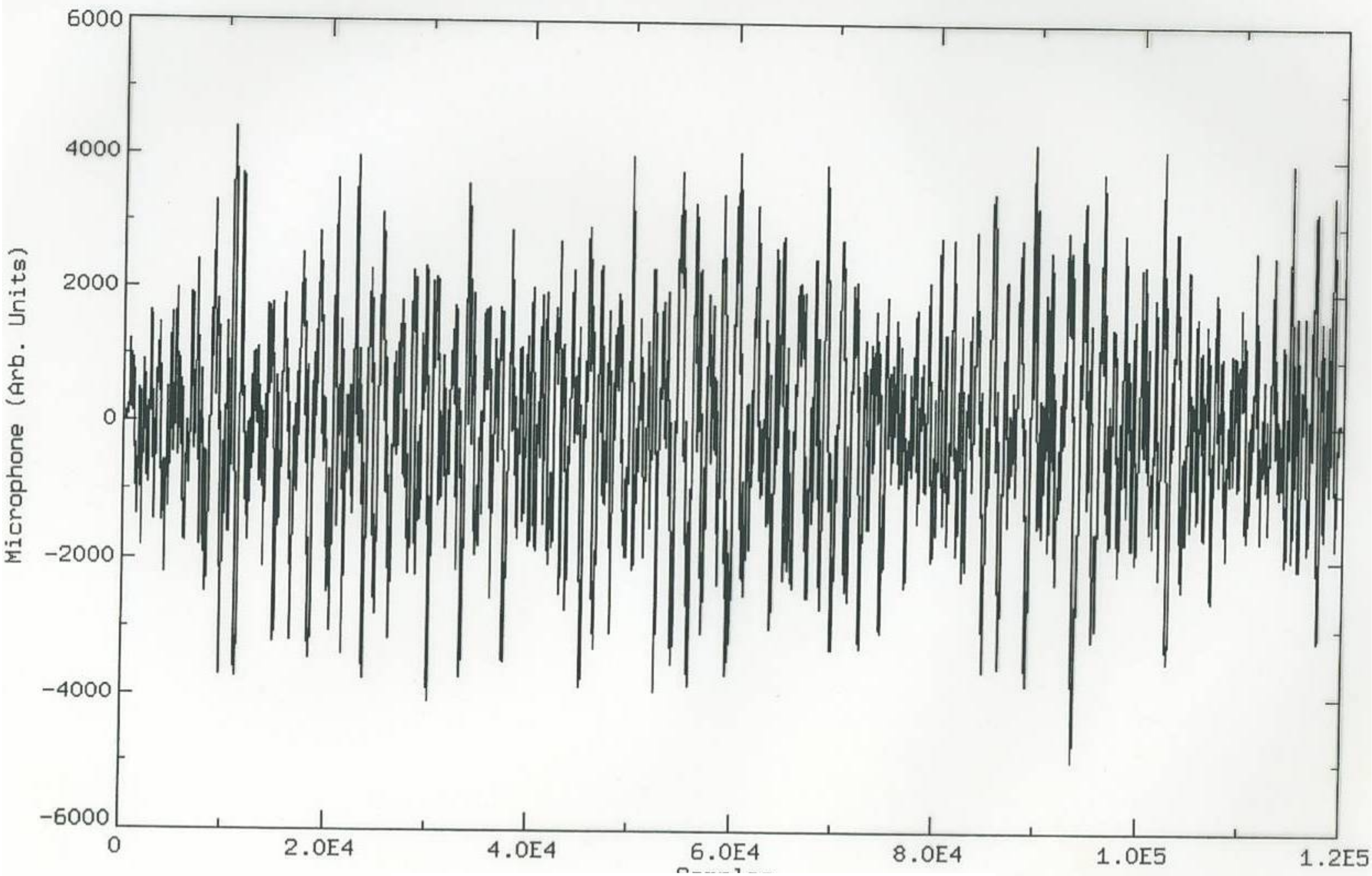
Project: Analyze and diagnose car rattles electronically

Example: Creak

Example: Clatter

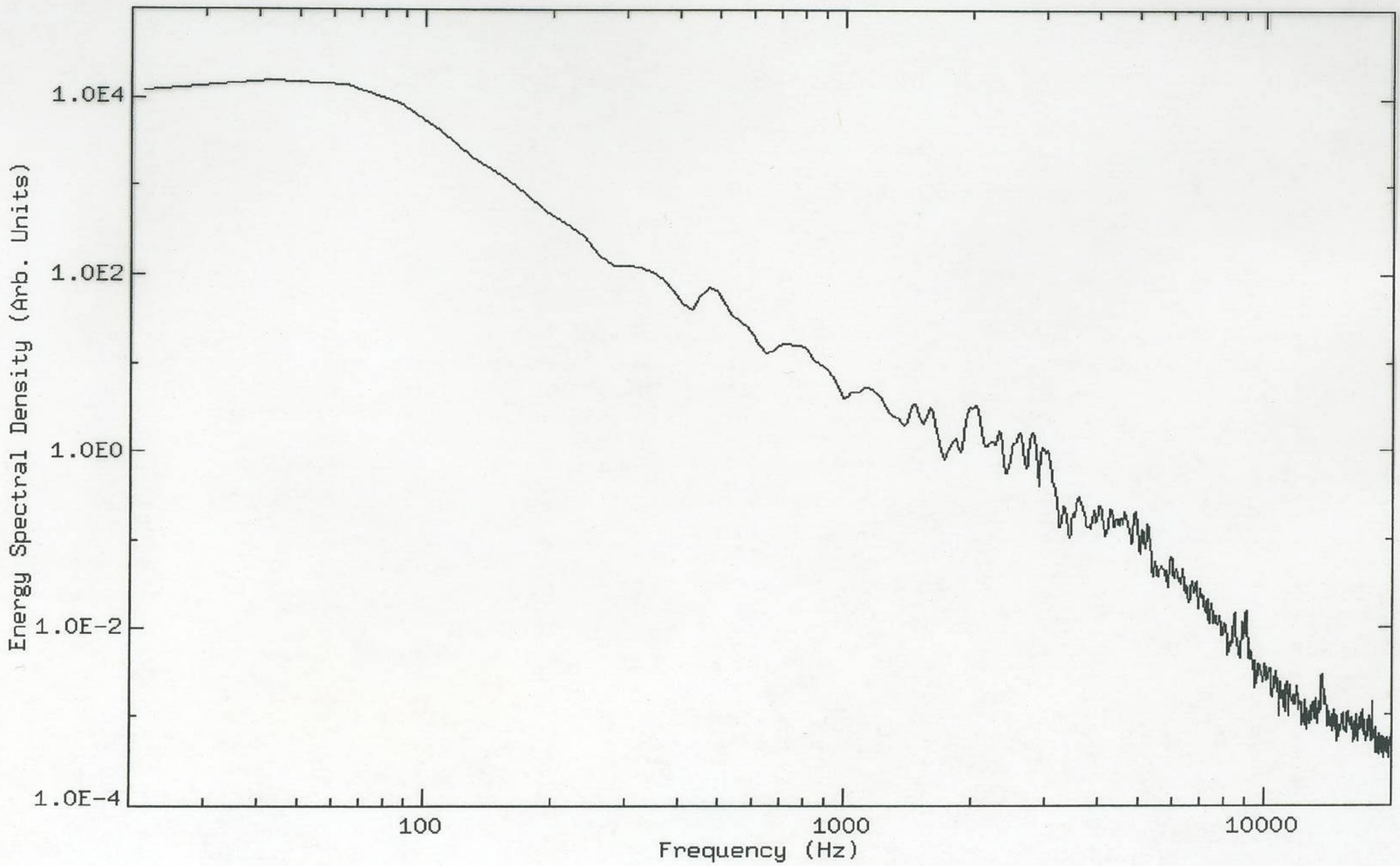
What is right representation to allow extraction of main features from the noise?

B Presented as a Time Series



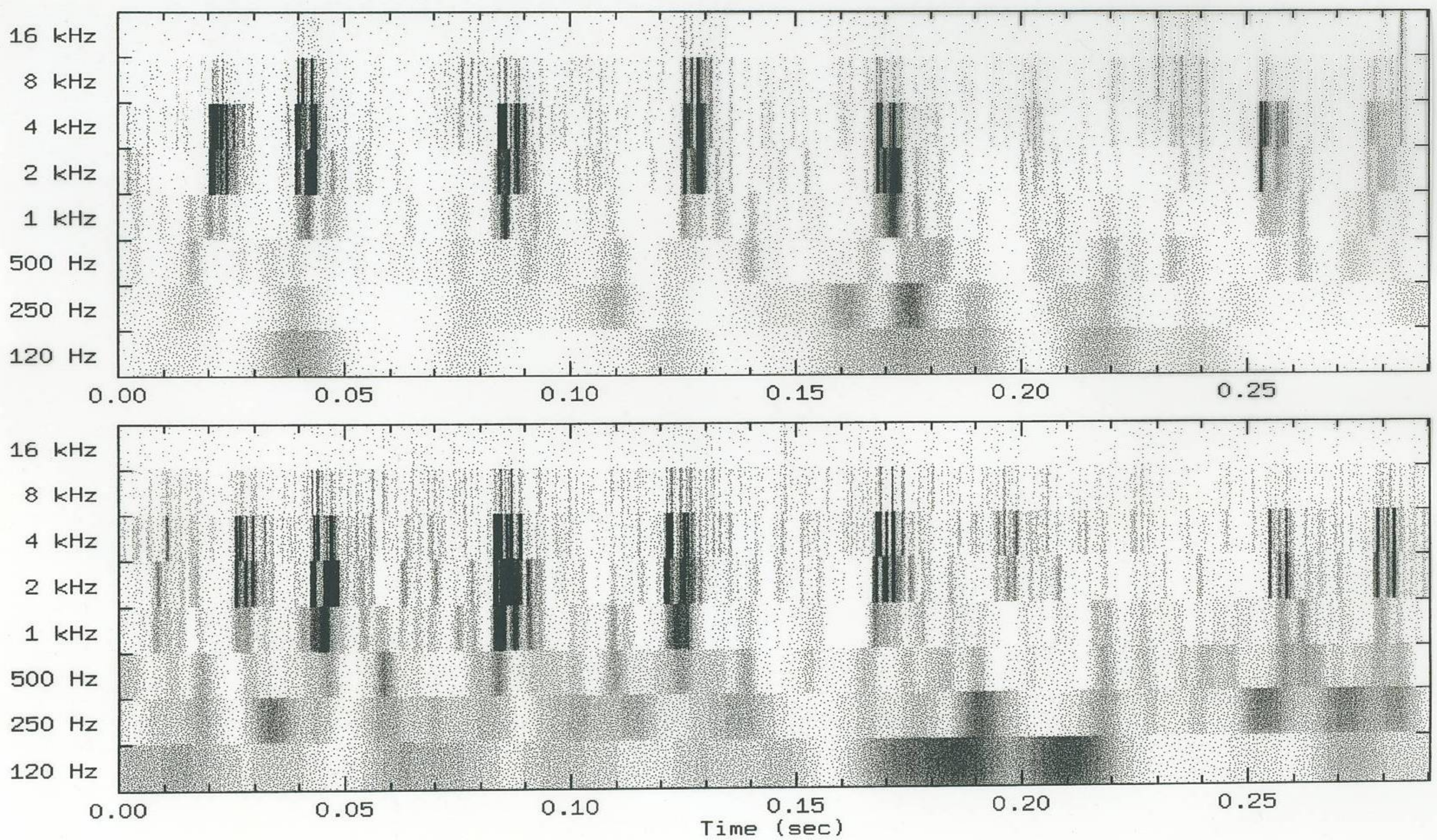
Time series: unclear

B Presented as a Spectral Density



Frequency Representation (spectrogram): unclear

B Presented as a Scalogram



Discrete Wavelet Representation (Top): Better

Shift-Invariant Discrete Wavelet Representation (Bottom): Even Better

If z is a vector of length N, wavelet transform of z is a vector of length N, and the wavelet transform is an invertible linear map. In fact, the inverse can be computed rapidly via convolutions, in $O(N)$ steps.

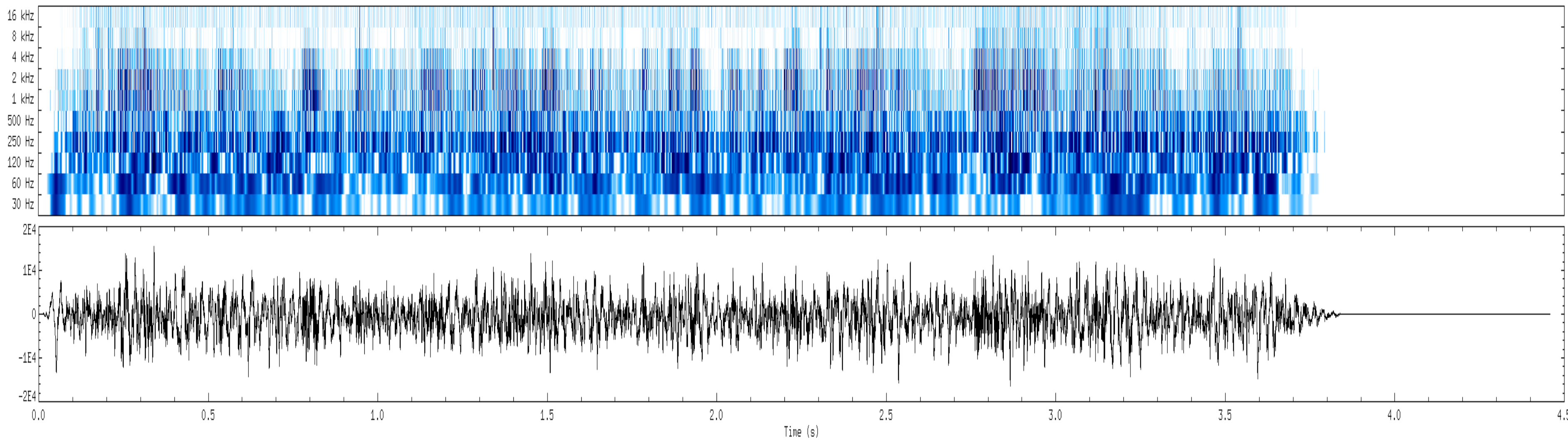
The Shift-Invariant Discrete Wavelet Transform (SIDWT) of z is obtained (roughly) by averaging the wavelet transform of z over all N translations of z. The SIDWT can be computed in $O(N \log^2 N)$ steps. However, the output is a vector of length $N \log_2 N$. Thus the SIDWT is a linear map from a lower-dimensional space into a much larger dimensional space. Hence it is not invertible.

In linear algebra, in this situation, one learns to use the pseudoinverse. If $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear and $1 - 1$, where $n < m$, define the pseudoinverse $S : \mathbb{R}^m \to \mathbb{R}^n$ as follows. For a point w in $\mathbb{R}^m$, find its orthogonal projection u on the range of T, and define $S(w)$ to be the unique $z \in \mathbb{R}^n$ such that $T(z) = u$.

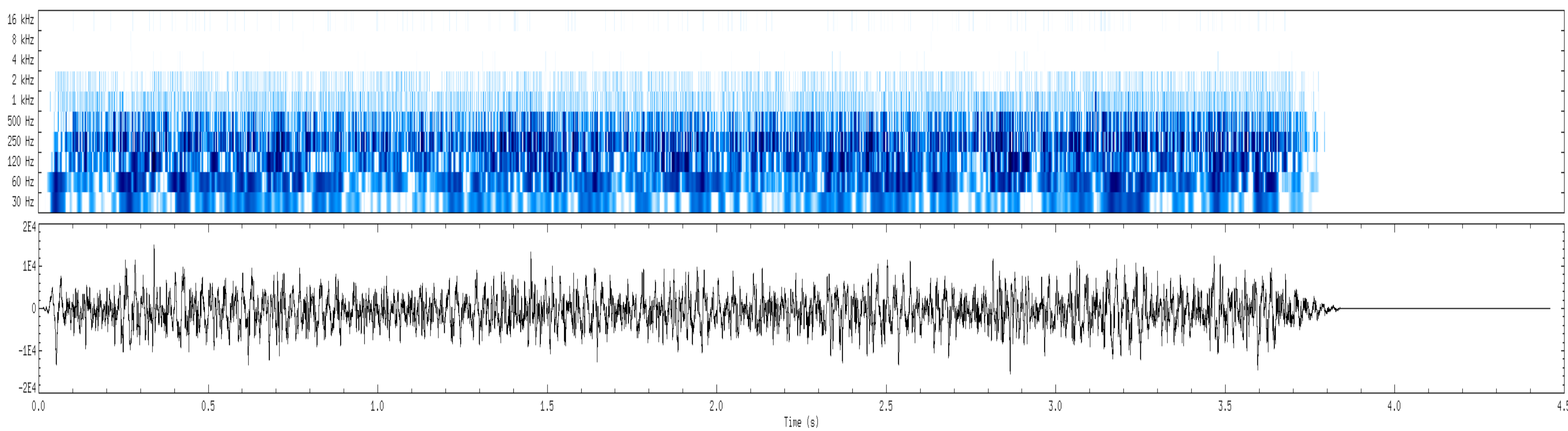
Problem: Formula for S is $S = (T^*T)^{-1}T^*$. Here T is a matrix of size $N \log_2 N \times N$, so T^*T is $N \times N$, where here typically $N \approx 10^5$ or 10^6. So the matrix is too large to invert rapidly and accurately.

My group figured out that the pseudoinverse of the SIDWT is also computable fast via convolutions, in fact in $O(N \log^2 N)$ steps. This allowed Dave Scholl to continue his examples.

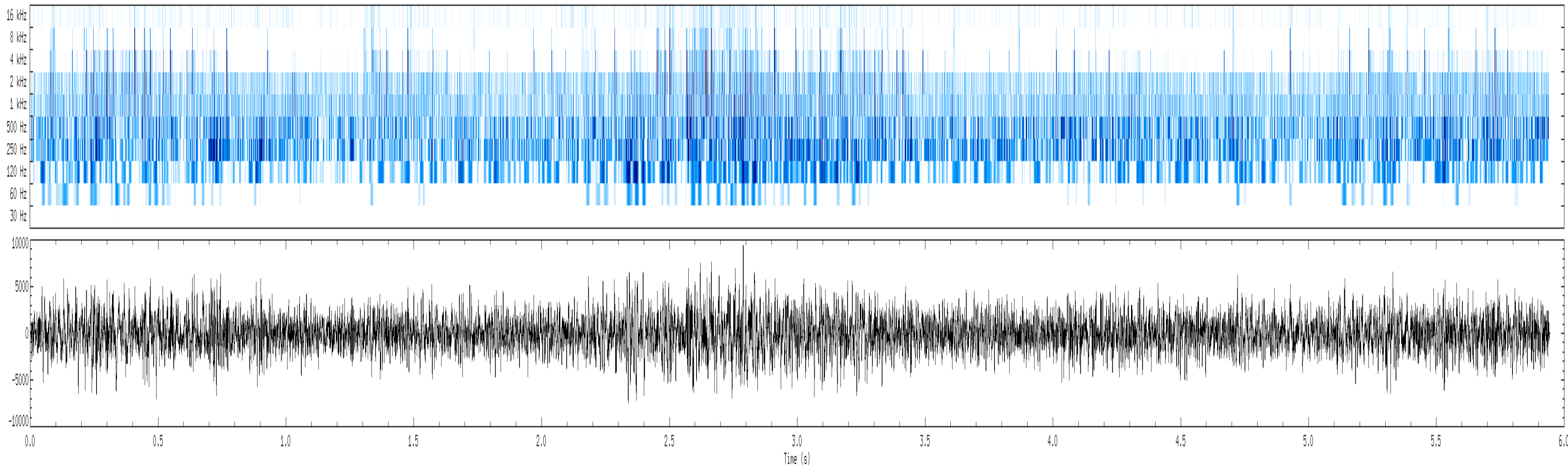
Creak: Original



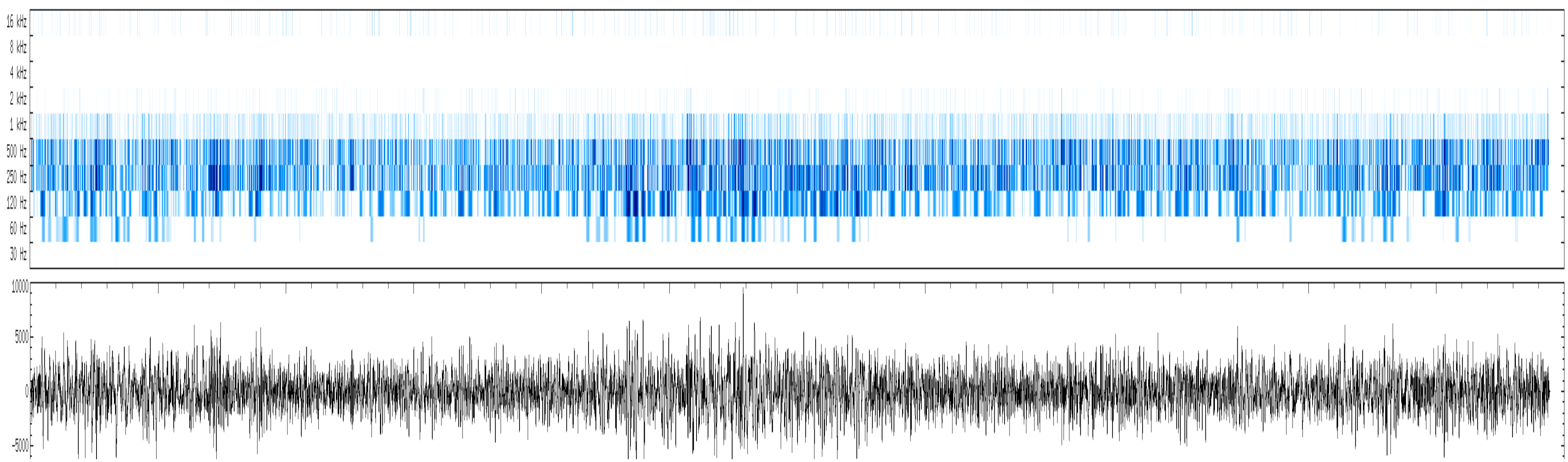
Creak: Background



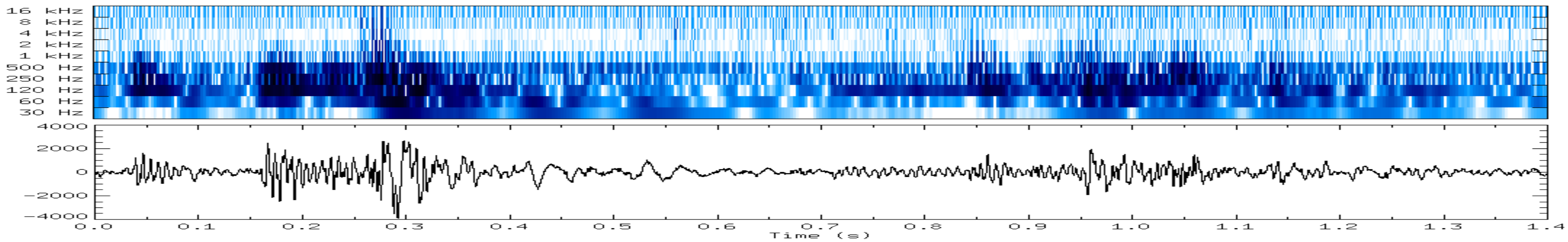
Clatter: Original



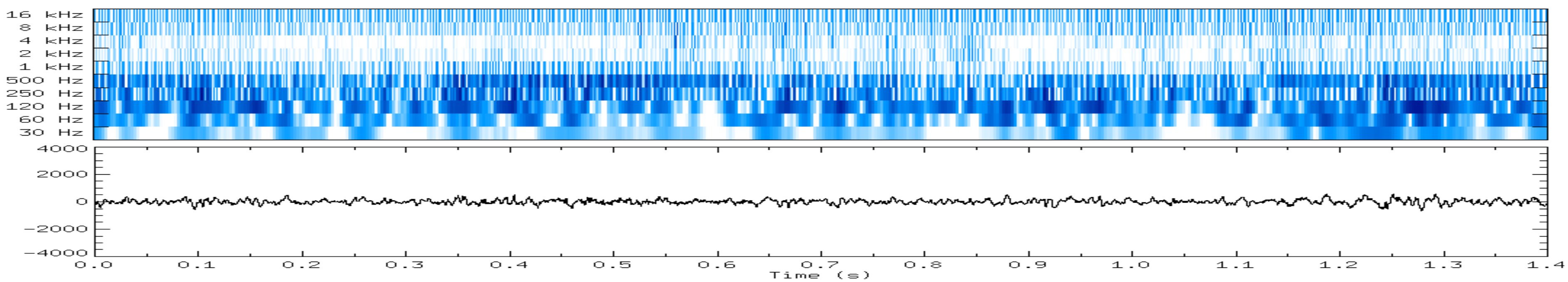
Clatter: Background



Windshield Wiper Reversal Thud: Original



Windshield Wiper Reversal Thud: Background



Windshield Wiper Reversal Thud: Thud