and if (A.7) holds, return $z^{(t)} = z$ and exit the loop; if (A.7) does not hold, then *shrink the hyperrectangle*, as follows:

- if $x_{0,1} < x_{0,1}^{(t-1)}$, set $x_{0,1,\mathrm{L}} = x_{0,1}$; else if $x_{0,1} > x_{0,1}^{(t-1)}$, set $x_{0,1,\mathrm{U}} = x_{0,1}$;
- if $x_{0,2} < x_{0,2}^{(t-1)}$, set $x_{0,2,\mathrm{L}} = x_{0,2}$; else if $x_{0,2} > x_{0,2}^{(t-1)}$, set $x_{0,2,\mathrm{U}} = x_{0,2}$;
- if $d < d^{(t-1)}$, set $d_{\mathrm{L}} = d$; else if $d > d^{(t-1)}$, set $d_{\mathrm{U}} = d$;
- if $A < A^{(t-1)}$, set $A_{\mathrm{L}} = A$; else if $A > A^{(t-1)}$, set $A_{\mathrm{U}} = A$;
- if $\varphi < \varphi^{(t-1)}$, set $\varphi_{\mathrm{L}} = \varphi$; else if $\varphi > \varphi^{(t-1)}$, set $\varphi_{\mathrm{U}} = \varphi$;
- go back to 2).

Here, the hyperrectangles *shrink toward* $\boldsymbol{\phi}^{(t-1)} = [x_{0,1}^{(t-1)}, x_{0,2}^{(t-1)}, d^{(t-1)}, A^{(t-1)}, \varphi^{(t-1)}]^T$, which is clearly in the slice (see Step 1)).

Since the evaluation of $l(\boldsymbol{y}|\boldsymbol{z}, \boldsymbol{w})$ may cause a floating-point underflow, it is often safer to compute $\ln l(\boldsymbol{y}|\boldsymbol{z}, \boldsymbol{w})$ and modify the above algorithm accordingly (see [12, Sec. 4]).

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers and to Prof. R. B. Thompson from CNDE, Iowa State University, for their insightful comments and for bringing [1] and [6] to our attention.

REFERENCES

[1] P. J. Howard, D. C. Copley, and R. S. Gilmore, "The application of a dynamic threshold to C-scan images with variable noise," in *Rev. Progress Quantitative Nondestructive Evaluation*, D. O. Thompson and D. E. Chimenti, Eds. Melville, NY: Amer. Inst. Phys., 1998, vol. 17, pp. 2013–2019.

[2] A. Tsai, A. Yezzi, Jr., and A. S. Willsky, "Curve evolution implementation of the Mumford–Shah functional for image segmentation, denoising, interpolation, and magnification," *IEEE Trans. Image Process.*, vol. 10, pp. 1169–1186, Aug. 2001.

[3] D. L. Pham, C. Y. Xu, and J. L. Prince, "Current methods in medical image segmentation," *Annu. Rev. Biomed. Eng.*, vol. 2, pp. 315–337, 2000.

[4] W. E. Polakowski *et al.*, "Computer-aided breast cancer detection and diagnosis of masses using difference of Gaussians and derivative-based feature saliency," *IEEE Trans. Med. Imag.*, vol. 16, pp. 811–819, Sep. 1997.

[5] A. Dogandžić and B. Zhang, "Bayesian defect signal analysis," in *Rev. Progress Quantitative Nondestructive Evaluation*. Melville, NY: Amer. Inst. Phys., 2006, vol. 25, pp. 617–624.

[6] Aerospace Industries Association Rotor Integrity Sub-Committee, "The development of anomaly distributions for aircraft engine titanium disk alloys," in *Proc. 38th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, Materials Conf.*, Kissimmee, FL, Apr. 1997, pp. 2543–2553 [Online]. Available: http://www.darwin.swri.org/html_files/pdf_docs/pubs/aia1997.pdf

[7] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, 2nd ed. New York: Chapman & Hall, 2004.

[8] B. P. Carlin and T. A. Louis, *Bayes and Empirical Bayes Methods for Data Analysis*, 2nd ed. New York: Chapman & Hall, 2000.

[9] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed. New York: Springer-Verlag, 2004.

[10] J. von Neumann, "Various techniques used in connection with random digits," in *John von Neumann, Collected Works*. New York: Pergamon, 1961, vol. V, pp. 768–770.

[11] J. Geweke, "Efficient simulation from the multivariate normal and student-t distributions subject to linear constraints," in *Computing Science Statistics: Proc. 23rd Symp. Interface*, Seattle, WA, Apr. 1991, pp. 571–578.

[12] R. M. Neal, "Slice sampling," *Ann. Statist.*, vol. 31, pp. 705–741, Jun. 2003.

[13] C. Andrieu, A. Doucet, and C. P. Robert, "Computational advances for and from Bayesian analysis," *Stat. Sci.*, vol. 19, no. 1, pp. 118–127, Feb. 2004.

[14] D. J. Spiegelhalter, N. G. Best, B. R. Carlin, and A. van der Linde, "Bayesian measures of model complexity and fit," *J. R. Stat. Soc.*, ser. B, vol. 64, pp. 583–639, 2002.

Attenuation Estimation From Correlated Sequences

Tarek Medkour and Andrew T. Walden

***Abstract*—We calculate the frequency-dependent variance of the log spectral ratio for correlated time series. This is used to produce a weighted least-squares approach to attenuation estimation, with weights calculated from estimated coherence. Applications to synthetic and real data illustrate that, for correlated series, the method improves significantly on traditional unweighted least-squares attenuation estimates.**

***Index Terms*—Attenuation, coherence, correlation, spectral ratios.**

I. INTRODUCTION

Attenuation can be estimated from the change in frequency content observed between two sequences separated by two-way travel time ΔT (e.g., [14] and [16]).

Given power spectra $S_{11}(f)$ and $S_{22}(f)$ corresponding to the sequences $\{X_{1,t}\}$ and $\{X_{2,t}\}$, we define the attenuation parameters through the spectral ratio as follows, $S_{22}(f)/S_{11}(f) = c_0 \cdot \exp[-2\,\Delta T\,\alpha(f)]$, where $\alpha(f)$ is the attenuation coefficient and c_0 is a constant. The acoustic attenuation coefficient of soft biological tissue has been observed to have a linear-with-frequency characteristic [7]. Likewise a linear form has also been justified in seismology [4], [17], and the ubiquitous linear assumption for attenuation is made in this correspondence. Let $\alpha(f) = \mu f$ say, where μ is a constant, so that

$$\log\left[\frac{S_{22}(f)}{S_{11}(f)}\right] = c - 2\,\Delta T\,\mu f \tag{1}$$

where $c = \log c_0$ and the coefficient μ is called the logarithmic decrement and is measured in nepers (better known as the natural log of a voltage ratio). Since $\alpha(f) = \mu f$, by slight abuse of notation (we are using time not distance) $\alpha(f)$ has units of nepers/wavelength. In terms of the oft-used quality factor $Q(f)$, (1) can be written

$$\log\left[\frac{S_{22}(f)}{S_{11}(f)}\right] = c - \frac{2\pi\Delta T\, f}{Q} \tag{2}$$

so that $\mu = \pi/Q$ and $\alpha(f) = \mu f = \pi f/Q$. Since one neper is equivalent to $20\log_{10} e$ dB ≈ 8.686 dB, it is apparent that $\alpha(f) \approx (27.3/Q)f$ dB/wavelength.

While in attenuation studies via spectral ratios, it is invariably assumed that the sequences involved are independent, [3], [8], here we will consider estimation of the quality factor, Q, when the sequences are correlated.

When the log spectral ratio in (2) is estimated via multitapering, it is seen in Section II that at any frequency, this ratio can be viewed as a log variance ratio in complex Gaussian random variables. This is explored in Section III where the distribution of the estimated log spectral ratio (standardized by the true ratio) is developed, and, most important, its cumulant generating function, and hence variance, are derived. This frequency-dependent variance is a decreasing function of the ordinary coherence—which reflects sequence correlation—between the two sequences. Section IV sets up the regression model through which

Manuscript received October 3, 2005; revised April 4, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. A. Rahim Leyman. T. Medkour thanks the government of the People's Democratic Republic of Algeria for financial support.

The authors are with the Department of Mathematics, Imperial College London, London SW7 2BZ, U.K. (e-mail: tarek.medkour@imperial.ac.uk and a.walden@imperial.ac.uk).

Digital Object Identifier 10.1109/TSP.2006.885682

Q is estimated and shows how the variance of the estimated log spectral ratio enables a weighted least squares approach as an alternative to unweighted (ordinary) least squares. Estimation of Q using synthetic and real data is presented in Sections V and VI, respectively; it is demonstrated that taking account of sequence correlation through the weighted regression scheme leads to improved estimation.

II. Estimated Spectral Ratio

Consider two real-valued N-length sequences $\{X_{1,t}\}$ and $\{X_{2,t}\}$ which are taken to be jointly stationary, zero-mean, and Gaussian. We estimate the power spectra for sequences 1 and 2 using the multitaper method (e.g., [12]), which employs a set of K orthogonal tapers. We first form the product $u_{k,t}X_{l,t}$ of the kth real-valued taper, $k = 0, \ldots, K-1$, with the lth process, $l = 1, 2$, and then compute the (scaled) Fourier transform

$$J_{k,l}(f) \equiv (\Delta t)^{1/2} \sum_{t=0}^{N-1} u_{k,t} X_{l,t} e^{-i2\pi f t\, \Delta t} \tag{3}$$

where Δt is the sample interval.

Let $\mathbf{J}_k(f) = [J_{k,1}(f), J_{k,2}(f)]^T$. Then asymptotically, as $N \to \infty$

$$\mathbf{J}_k(f) \stackrel{\mathrm{d}}{=} \mathcal{N}_2^C(\mathbf{0}, \mathbf{S}(f)) \tag{4}$$

[15], i.e., $\mathbf{J}_k(f)$ has a complex Gaussian distribution of dimension 2 with mean $\mathbf{0}$ and covariance matrix $\mathbf{S}(f)$ for $0 < |f| < f_{\mathcal{N}}$, where $f_{\mathcal{N}} = 1/(2\Delta t)$, and

$$E\{\mathbf{J}_k(f)\mathbf{J}_k^H(f)\} = \begin{bmatrix} S_{11}(f) & S_{12}(f) \\ S_{12}^*(f) & S_{22}(f) \end{bmatrix} = \mathbf{S}(f). \tag{5}$$

The multitaper estimator of $\mathbf{S}(f)$ is

$$\widehat{\mathbf{S}}(f) = \frac{1}{K} \sum_{k=0}^{K-1} \mathbf{J}_k(f)\mathbf{J}_k^H(f) = \begin{bmatrix} \widehat{S}_{11}(f) & \widehat{S}_{12}(f) \\ \widehat{S}_{12}^*(f) & \widehat{S}_{22}(f) \end{bmatrix}. \tag{6}$$

Then

$$\frac{\hat{S}_{22}(f)}{\hat{S}_{11}(f)} = \frac{\frac{1}{K}\sum_{k=0}^{K-1} |J_{k,2}(f)|^2}{\frac{1}{K}\sum_{k=0}^{K-1} |J_{k,1}(f)|^2} = \frac{\sum_{k=0}^{K-1} |J_{k,2}(f)|^2}{\sum_{k=0}^{K-1} |J_{k,1}(f)|^2} \tag{7}$$

where each term of the form $|J_{k,l}(f)|^2$ is a tapered periodogram. From (2), we wish to understand the statistical distribution of the logarithm of this estimated spectral ratio. From (7), the ratio can be thought of as a (sample) variance ratio in complex Gaussian random variables. We next study the general case of variance ratios for two-dimensional complex Gaussian vectors.

III. Variance Ratios

Let $\mathbf{V}_k = [V_{k,1}, V_{k,2}]^T, k = 0, \ldots, K-1$, be a size K random sample of two-dimensional complex Gaussian vectors, each $\mathcal{N}_2^C(\mathbf{0}, \mathbf{\Sigma})$. Let $W_1 = \sum_{k=0}^{K-1} |V_{k,1}|^2$ and $W_2 = \sum_{k=0}^{K-1} |V_{k,2}|^2$, and define $\hat{R} = W_2/W_1$, the ratio of sample variances for the two complex vector components. The ratio in (7) has the same statistical properties as $\hat{R}$, with $\mathbf{S}(f)$ replacing $\mathbf{\Sigma}$.

Then [9, p.92] the probability density function (pdf) of $\hat{R}$ is given by, for $x \geq 0$

$$f_{\hat{R}}(x) = \frac{x^{K-1}(x+r)r^K(1-\gamma^2)^K}{B(K,K)[(x+r)^2 - 4rx\gamma^2]^{K+1/2}}. \tag{8}$$

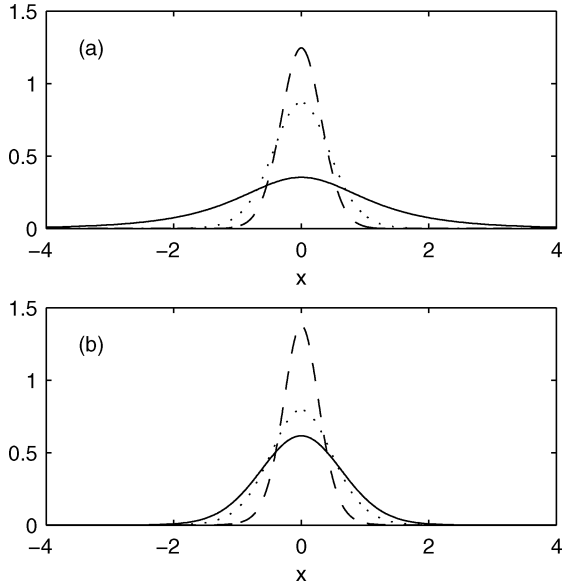

Fig. 1. pdf $\hat{P}$ (a) for $K = 1$ (solid line), 5 (dotted line), and 10 (dashed line) when $\gamma^2 = 0.5$, and (b) for $\gamma^2 = 0$ (solid line), 0.4 (dotted line), and 0.8 (dashed line) when $K = 5$.

Here, $B(K,K) = \Gamma^2(K)/\Gamma(2K)$ is the beta function, $r = \Sigma_{22}/\Sigma_{11}$ is the ratio of the true variances, and $\gamma^2 = |\Sigma_{12}|^2/[\Sigma_{11}\Sigma_{22}]$ is the modulus squared of the complex correlation coefficient, with $\gamma^2 < 1$ assumed. It then follows that $\hat{P} = \log(\hat{R}/r)$, has pdf

$$f_{\hat{P}}(x) = \frac{e^{Kx}(1+e^x)(1-\gamma^2)^K}{B(K,K)[(1+e^x)^2 - 4\gamma^2 e^x]^{K+1/2}} \tag{9}$$

$-\infty < x < \infty$, the complex version of a result in [2].

A. Uncorrelated Components

When the components of the vector are uncorrelated, $\hat{P}$ involves some familiar distributions. If $K = 1$, and $\Sigma_{12} = 0 \Rightarrow \gamma^2 = 0$, then $\hat{P} = \log\{[|V_{0,2}|^2/\Sigma_{22}]/[|V_{0,1}|^2/\Sigma_{11}]\}$, which is the log of the ratio of two independent chi-square random variables with two degrees of freedom (χ_2^2), i.e., the log of an $F_{2,2}$ random variable. In this case, the pdf in (9) takes the form $f_{\hat{P}}(x) = e^x(1+e^x)^{-2} = 1/[4\cosh^2(x/2)]$, agreeing with [3]. For general $K > 1$ and $\gamma^2 = 0$, $\hat{P}$ is the log of the ratio of two independent chi-square random variables each having $2K$ degrees of freedom, i.e., the log of an $F_{2K,2K}$ random variable, with pdf $f_{\hat{P}}(x) = e^{Kx}(1+e^x)^{-2K}/B(K,K)$; this pdf is shown for $K = 5$ by the solid curve in Fig. 1(b). Also, when $\gamma^2 = 0$, both W_1/Σ_{11} and W_2/Σ_{22} have the $\chi_{2K}^2/2K$ distribution and are independent, so $E\{\hat{P}\} = E\{\log(W_2/\Sigma_{22}) - \log(W_1/\Sigma_{11})\} = 0$. Further, $\mathrm{var}\{\hat{P}\} = \mathrm{var}\{\log(W_2/\Sigma_{22}) - \log(W_1/\Sigma_{11})\}$, and [1] $\mathrm{var}\{\log \chi_{2K}^2\} = \psi'(K)$, so that $\mathrm{var}\{\hat{P}\} = 2\psi'(K)$. Here, $\psi'(z) = \{\Gamma''(z)\Gamma(z) - [\Gamma'(z)]^2\}/\Gamma^2(z)$ denotes the trigamma function, the second derivative of the log of the gamma function.

B. Correlated Components: Mean and Variance

We are primarily interested in the case where the components of the random vector are correlated and the general pdf in (9) applies, illustrated in Fig. 1. From Fig. 1(a), we see that for a fixed γ^2, the effect of increasing K, the number of independent terms being averaged, is a reduction in the variance of $\hat{P}$. Likewise, we see from Fig. 1(b) that for a fixed K, increasing the correlation γ^2 also reduces the variance;

a high correlation between the elements $V_{k,1}$ and $V_{k,2}$ naturally makes the ratio formed from them less variable.

The moment-generating function (MGF), $M(t)$, corresponding to the pdf in (9), is derived in the Appendix . It is given by

$$M(t) = \frac{\Gamma(K+t)\Gamma(K-t)}{\Gamma^2(K)}(1-\gamma^2)^K {}_2F_1(K-t, K+t; K; \gamma^2). \tag{10}$$

Here, ${}_2F_1(a, b; c; y)$ is the hypergeometric function with 2 and 1 parameters and argument y. In fact [5, p. 1045], ${}_2F_1(K-t, K+t; K; \gamma^2)$ may be written explicitly as

$$\sum_{m=0}^{\infty} \frac{\Gamma(K-t+m)\Gamma(K+t+m)\Gamma(K)(\gamma^2)^m}{\Gamma(K-t)\Gamma(K+t)\Gamma(K+m)m!} \tag{11}$$

so from (10), $M(t)$ is given by

$$(1-\gamma^2)^K \sum_{m=0}^{\infty} \frac{\Gamma(K-t+m)\Gamma(K+t+m)(\gamma^2)^m}{\Gamma(K+m)\Gamma(K)m!}. \tag{12}$$

For calculation of the mean, and in particular the variance, it is convenient to use the cumulant-generating function given by $C(t) = \log M(t)$, for which the mean is $C'(0)$ and the variance is $C''(0)$. Clearly, $C(t)$ is given by

$$\log(1-\gamma^2)^K + \log \sum_{m=0}^{\infty} \frac{\Gamma(K-t+m)\Gamma(K+t+m)(\gamma^2)^m}{\Gamma(K+m)\Gamma(K)m!} \tag{13}$$

from which we obtain

$$E\{\hat{P}\} = C'(0) = 0 \tag{14}$$

and $\mathrm{var}\{\hat{P}\}$ is equal to

$$\frac{2}{D} \sum_{m=0}^{\infty} \frac{\left\{\Gamma''(K+m)\Gamma(K+m) - [\Gamma'(K+m)]^2\right\}(\gamma^2)^m}{\Gamma(K+m)\Gamma(K)m!} \tag{15}$$

where $D = \sum_{m=0}^{\infty}[\Gamma(K+m)(\gamma^2)^m]/[\Gamma(K)m!]$. We can write this variance in a simpler form. Since $\gamma^2 < 1$, we can write $D = (1-\gamma^2)^{-K}$ (using the negative binomial expansion). Hence, using the form of $\psi'(z)$ given above, we get

$$\mathrm{var}\{\hat{P}\} = 2(1-\gamma^2)^K \sum_{m=0}^{\infty} \frac{\Gamma(K+m)\psi'(K+m)(\gamma^2)^m}{\Gamma(K)m!}. \tag{16}$$

In the case when γ^2 is nonzero, its percentage effect on variance reduction can be calculated via

$$100 \cdot \left[\frac{\mathrm{var}\{\hat{P}|\gamma^2 = 0\} - \mathrm{var}\{\hat{P}|\gamma^2\}}{\mathrm{var}\{\hat{P}|\gamma^2 = 0\}}\right]. \tag{17}$$

This is illustrated in Fig. 2. We see that for $K \geq 3$, this variance improvement is approximately linear over most of the range of γ^2.

We now return to the estimation of Q and apply the statistical results derived above.

IV. Regression on Spectral Ratio Ordinates

A. Regression Model

Two common types of orthogonal tapers used in multitaper spectrum estimates are the Slepian tapers and sine tapers (see, e.g., [15] for details). In both cases, we can express the bandwidth of the estimator as

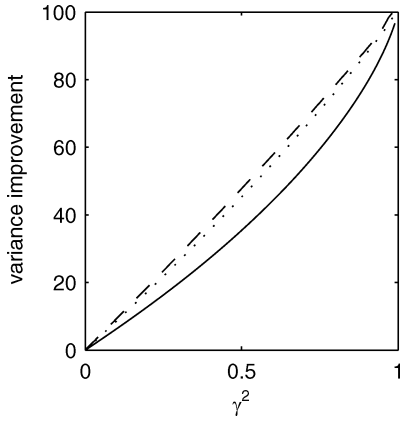

Fig. 2. Variance improvement percentage as a function of γ^2 for $K = 1$ (solid line), 3 (dotted line), and 6 (dashed line).

$B = l/(N\Delta t)$ Hz, for some small positive integer l. Hence, the multitaper spectrum estimators at Fourier frequencies $f_k = k/(N\Delta t)$ and f_{k+l} will be uncorrelated [12].

We choose, from the set of Fourier frequencies, a subset of m increasing frequencies $f'_j, j = 1, \ldots, m$, with $f'_{j+1} - f'_j \geq B, j = 1, \ldots, m-1$, within the signal band of the series. At these frequencies, the random variables $\eta_j = \hat{P}(f'_j) = \log\{[\hat{S}_{22}(f'_j)S_{11}(f'_j)]/[\hat{S}_{11}(f'_j)S_{22}(f'_j)]\}$ will be uncorrelated. From (1), $\log\{S_{22}(f'_j)/S_{11}(f'_j)\} = c - 2\Delta T\,\mu f'_j$. Adding η_j to both sides gives $\log\{[\hat{S}_{22}(f'_j)]/[\hat{S}_{11}(f'_j)]\} = c - 2\Delta T\,\mu f'_j + \eta_j$, for $j = 1, \ldots, m$. From (4) and (6) and the results of Section III, we know that $E\{\eta_j\} = 0$, (14), and $\mathrm{var}\{\eta_j\}$ is given by (16) with γ^2 replaced by $\gamma^2(f'_j) = |S_{12}(f'_j)|^2/[S_{11}(f'_j)S_{22}(f'_j)]$, the ordinary (magnitude squared) coherence. If we estimate β as $\hat{\beta}$ in the linear regression

$$Y_j = c + \beta x_j + \eta_j, \quad j = 1, \ldots, m \tag{18}$$

where $Y_j = \log[\hat{S}_{22}(f'_j)/\hat{S}_{11}(f'_j)]$, $x_j = 2\Delta T\, f'_j$, and $\beta = -\mu$, the estimate of Q follows as $\hat{Q} = -\pi/\hat{\beta}$.

We can write the m equations of (18) in the form

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{N} \tag{19}$$

where $\boldsymbol{Y}$ is the $m \times 1$ column vector of Y_j's, $\boldsymbol{X}$ is the $m \times 2$ matrix with first column all 1's and second column consisting of the x_j's, $\boldsymbol{\theta}$ is the 2×1 column vector of parameters c and β, and $\boldsymbol{N}$ is the $m \times 1$ column vector of the η_j's.

B. Unweighted Least Squares

As is well known [13, p. 592], the least-squares estimator corresponding to (19) is

$$\hat{\boldsymbol{\theta}}_U = \left[\boldsymbol{X}^T\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{Y} = \boldsymbol{\theta} + \left[\boldsymbol{X}^T\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{N} \tag{20}$$

where $\hat{\boldsymbol{\theta}}_U = [\hat{c}_U, \hat{\beta}_U]^T$. Since $\boldsymbol{N}$ has mean zero, $\hat{\boldsymbol{\theta}}$ is an unbiased estimator, and its covariance matrix, $\boldsymbol{\Theta}_U$ say, is

$$\boldsymbol{\Theta}_U = E\left\{\left[\hat{\boldsymbol{\theta}}_U - \boldsymbol{\theta}\right]\left[\hat{\boldsymbol{\theta}}_U - \boldsymbol{\theta}\right]^T\right\} = \left[\boldsymbol{X}^T\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{V}\boldsymbol{X}\left[\boldsymbol{X}^T\boldsymbol{X}\right]^{-1} \tag{21}$$

where $\boldsymbol{V}$ is the covariance matrix of $\boldsymbol{N}$. Since the frequencies f_j have been selected in order that the η_j are uncorrelated, we know that $\boldsymbol{V} = \mathrm{diag}(\mathrm{var}\{\eta_1\}, \ldots, \mathrm{var}\{\eta_m\})$.

C. Weighted Least Squares

The weighted least-squares estimator corresponding to (18) is [13, p. 592]

$$\hat{\boldsymbol{\theta}}_W = \left[\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{Y} = \boldsymbol{\theta}+\left[\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{N} \tag{22}$$

where $\hat{\boldsymbol{\theta}}_W = [\hat{c}_W, \hat{\beta}_W]^T$. $\hat{\boldsymbol{\theta}}_W$ is thus an unbiased estimator, and its covariance matrix, $\boldsymbol{\Theta}_W$ say, is given by

$$\boldsymbol{\Theta}_W = E\left\{\left[\hat{\boldsymbol{\theta}}_W - \boldsymbol{\theta}\right]\left[\hat{\boldsymbol{\theta}}_W - \boldsymbol{\theta}\right]^T\right\} = \left[\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X}\right]^{-1}. \tag{23}$$

D. Using Estimated Coherence

In practice, when forming up the matrix $\boldsymbol{V}$, we will only have an estimate of $\gamma^2(f_j')$ for use in $\mathrm{var}\{\eta_j\}$ in (16). We shall use the exactly unbiased estimator

$$\tilde{\gamma}^2(f_j') = 1 - (1-\hat{\gamma}^2(f_j'))\,{}_2F_1(1,1;K;1-\hat{\gamma}^2(f_j')) \tag{24}$$

where $\hat{\gamma}^2(f_j')$ is the simple raw estimate of coherence. This formula follows from [10, p. 211] by correct identification of degrees of freedom followed by a doubling to allow for the change from real to complex degrees of freedom; it is more accurate than other well-known formulas for relatively small degrees of freedom. The corresponding estimate of $\boldsymbol{V}$ will be denoted $\tilde{\boldsymbol{V}}$.

The actual weighted least-squares estimate used in practice, and referred to hereafter, is then given by (22) with $\boldsymbol{V}$ replaced by $\tilde{\boldsymbol{V}}$.

Since $Q = -\pi/\beta$, what we shall refer to as the unweighted and weighted least-squares estimates of Q are defined as $\hat{Q}_U = -\pi/\hat{\beta}_U$ and $\hat{Q}_W = -\pi/\hat{\beta}_W$, respectively.

V. Application to Synthetic Data

We first compare the unweighted and weighted estimates of Q under controlled conditions by simulation. We used

$$X_{1,t} = Y_t + \xi_t \quad \text{and} \quad X_{2,t} = b_t * Y_t \tag{25}$$

where $\{Y_t\}$ is a zero-mean stationary process, $\{\xi_t\}$ is a zero-mean white noise process with variance σ_ξ^2, independent of $\{Y_t\}$, and $\{b_t\}$ is a linear filter, with $*$ denoting convolution. Let us denote the power spectrum of $\{Y_t\}$ by $S_{YY}(f)$; a broadband spectrum was modeled having relatively less power around $f = 0.2$ and above 0.4. For model (25), the magnitude squared coherence may be readily calculated and is given by $\gamma^2(f) = S_{YY}(f)/[S_{YY}(f) + \sigma_\xi^2]$. Fig. 3(a) shows two coherence structures we used, coherence model 1 (solid line) and model 2 (dashed–dotted line), corresponding to $\sigma_\xi^2 = 0.1$ and 0.5, respectively. Given $S_{YY}(f)$ and σ_ξ^2, a particular log spectral ratio of the form (2) was accomplished by calculating $\{b_t\}$ appropriately.

A total of 500 independent simulations were used for all results shown. $K = 3$ tapers were used, and for computational speed, sine tapers were utilized. Fig. 3(b) shows median estimates of Q for coherence model 1. The upper part shows $\hat{Q}_U$ (solid line) and $\hat{Q}_W$ (dashed line) when the true value of Q was 100 for $N = 128, 256, 512$, and 1024. The lower part shows the same when the true value of Q was 50. (The horizontal dotted lines show the true Q values.) Convergence is seen to occur with increasing sample size, with the weighted estimates outperforming the unweighted.

Fig. 3(c) gives median estimates of Q for coherence model 2. Shown are $\hat{Q}_U$ (solid line) and $\hat{Q}_W$ (dashed line), for $Q = 50$. The coherence here is uniformly lower than for model 1. The result, as expected, is that although the weighted scheme still performs better, the difference is less marked.

Fig. 3(d) shows the average squared deviations from the true value of 50—i.e., the estimated mean-square error (MSE)—for $Q = 50$ and coherence model 1. The solid line is for the unweighted scheme and the dashed line for the weighted scheme. Again, convergence is seen to occur with increasing sample size, with the weighted estimates outperforming the unweighted.

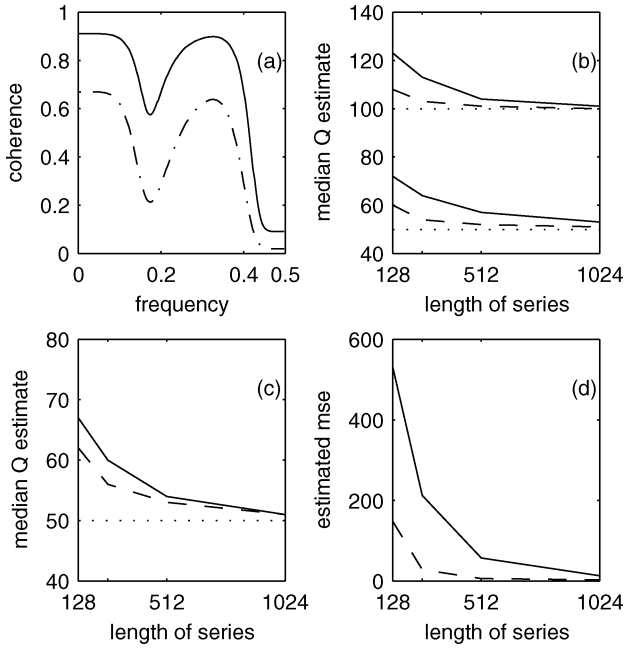

Fig. 3. (a) Coherence model 1 (solid) and model 2 (dashed–dotted). (b) Median estimates of Q for coherence model 1. Upper: $\hat{Q}_U$ (solid line), and $\hat{Q}_W$ (dashed line) for $Q = 100$. Lower: same for $Q = 50$. (c) Median estimates of Q for coherence model 2. $\hat{Q}_U$ (solid line), and $\hat{Q}_W$ (dashed line) for $Q = 50$. (d) Estimated MSE for unweighted (solid line) and weighted (dashed line) least squares for $Q = 50$ and coherence model 1.

VI. Application to Real Seismic Data

To compare the unweighted and weighted estimates of Q under uncontrolled conditions, we analyze 30 seismic series from a North Sea data set. The series were recorded at 25-m spacing along the sea surface. At each location, a seismic waveform created by an airgun array explosion travels into the sea and subsurface rock, reflects off the layers, and the amalgam of reflected events are recorded at the surface as the seismic series. Processing reduces multiple reflections, enhancing primary events. Absorption of seismic waves in the earth results in preferential attenuation of high frequencies with depth, and as stated earlier has a linear-with-frequency characteristic [17]. Processing of seismic data for attenuation estimation avoids any time-variant steps, meaning that no change in deconvolution operators or filters are allowed over the total time interval involved; the data discussed here was specially processed for attenuation estimation.

For each series, we extracted two 0.508-s segments ($N = 128$ samples, sample interval $\Delta t = 0.004$ s) with $\Delta T = 0.56$ s. These are shown in Fig. 4. For each of the series, the upper segment corresponds to the sequence $\{X_{1,t}\}$ and the lower segment to the sequence $\{X_{2,t}\}$. The lower segment results from reflections deeper in the earth and thus higher frequencies are more attenuated compared to the upper segment. The analysis bandwidth was chosen as $B = 4/(N\Delta t)$ Hz, i.e., B slightly less than 8 Hz. The f_j' were thus taken as $4, 12, 20, \ldots, 68$ Hz, since seismic frequencies are typically confined to the interval 4–70 Hz. Consequently, the number of Slepian tapers used, K, were chosen as $K = 3$ [12, p. 335].

The estimated covariance matrices of the form (21) and (23) using $\tilde{\boldsymbol{V}}$ will be denoted $\tilde{\boldsymbol{\Theta}}_U$ and $\tilde{\boldsymbol{\Theta}}_W$, respectively. The estimated variances are

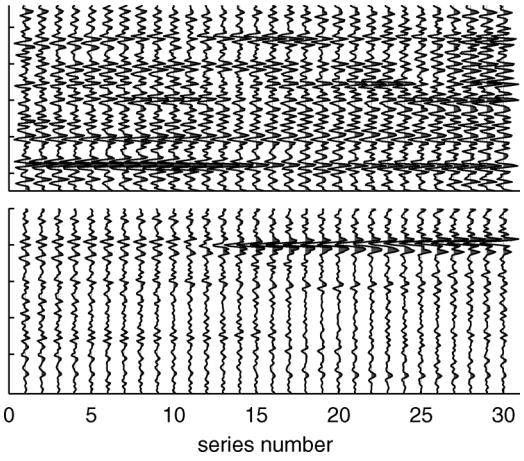

Fig. 4. Two segments of 30 seismic time series.

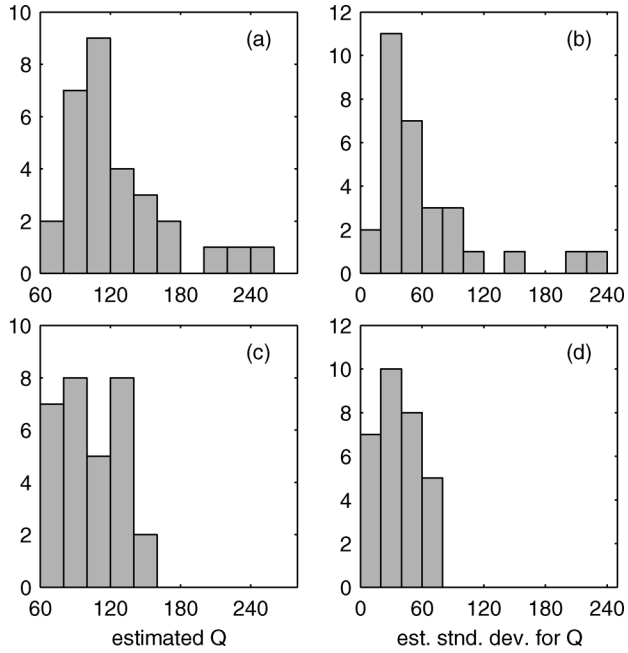

Fig. 5. Histograms over the 30 seismic series of: (a) the unweighted $\hat{Q}_U$ and (c) weighted $\hat{Q}_W$ least-squares estimates of Q; (b) and (d) the corresponding estimated standard deviations.

then $\widehat{\mathrm{var}}\{\hat{\beta}_U\} = \tilde{\Theta}_{U,22}$ and $\widehat{\mathrm{var}}\{\hat{\beta}_W\} = \tilde{\Theta}_{W,22}$, where the subscript "22" in $\tilde{\Theta}_{U,22}$ denotes the bottom right entry of the 2×2 matrix $\tilde{\boldsymbol{\Theta}}_U$.

Retaining only the first two terms of a Taylor expansion of $\hat{Q} = -\pi/\hat{\beta}$ and taking the variance [13, p. 90] gives the approximate result $\mathrm{var}\{\hat{Q}\} = (Q^4/\pi^2)\mathrm{var}\{\hat{\beta}\}$. The more useful estimated standard deviations of the estimator of Q for unweighted least squares and weighted least squares are given by, respectively

$$\widehat{\mathrm{sd}}\{\hat{Q}_U\} = \frac{\hat{Q}_U^2}{\pi}\widehat{\mathrm{sd}}\{\hat{\beta}_U\} = \frac{\hat{Q}_U^2}{\pi}[\tilde{\Theta}_{U,22}]^{1/2} \tag{26}$$

$$\widehat{\mathrm{sd}}\{\hat{Q}_W\} = \frac{\hat{Q}_W^2}{\pi}\widehat{\mathrm{sd}}\{\hat{\beta}_W\} = \frac{\hat{Q}_W^2}{\pi}[\tilde{\Theta}_{W,22}]^{1/2}. \tag{27}$$

Histograms over the 30 seismic series of the unweighted $\hat{Q}_U$ and weighted $\hat{Q}_W$ least-squares estimates of Q, and of the corresponding estimated standard deviations, are given in Fig. 5. Table I gives the minimum, average, and maximum of the values from which the histograms are formed. It can be seen that the weighted least-squares estimates of Q have a much smaller range and smaller average than achieved by unweighted least squares. Moreover, the average $\hat{Q}_W$ obtained, of about 100, agrees with typical North Sea results [14]. Moreover, the weighted least-squares estimates of the corresponding standard deviations have a much smaller range and much smaller average than achieved by unweighted least squares.

TABLE I
MINIMUM, AVERAGE AND MAXIMUM OVER 30 SEISMIC SERIES OF THE ESTIMATED Q VALUES, USING UNWEIGHTED, $\hat{Q}_U$, AND WEIGHTED, $\hat{Q}_W$ LEAST-SQUARES (TOP BLOCK) AND OF THE CORRESPONDING ESTIMATED STANDARD DEVIATIONS (BOTTOM BLOCK)

	Min	*Average*	*Max*
$\hat{Q}_U$	72	126	254
$\hat{Q}_W$	69	104	157
$\widehat{\mathrm{sd}}\{\hat{Q}_U\}$	17	63	237
$\widehat{\mathrm{sd}}\{\hat{Q}_W\}$	14	38	77

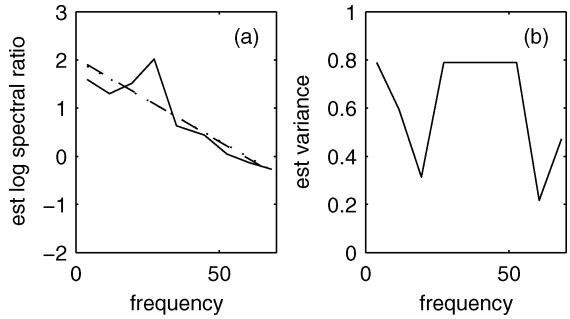

Fig. 6. (a) Estimated log spectral ratio (series 15) and unweighted (dotted line) and weighted (dashed line) least squares line fits. (b) Values of the diagonal matrix $\tilde{\boldsymbol{V}}$.

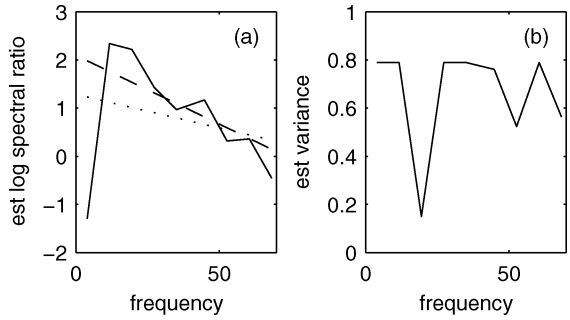

Fig. 7. As for Fig. 6, but for series 27.

Two particular analyses further illustrate the approach. Fig. 6(a) shows the estimated log spectral ratio for series 15 and regression fits using unweighted (dotted line) and weighted (dashed line) least squares. $\hat{Q}_U = 105$ and $\hat{Q}_W = 102$, i.e., both are very similar. Fig. 6(b) shows the values of the diagonal matrix $\tilde{\boldsymbol{V}}$. Note that $\mathrm{var}\{\hat{P}|\tilde{\gamma}^2(f_j') = 0\} = 2\psi'(3) = 0.7899$ when $K = 3$, as here. Reduced variance around 20 and 60 Hz is apparent (due to significant coherence), but the lack of any strong leverage points in the regression means that this information has little effect on the straight line fit. This can be contrasted with Fig. 7 resulting from series 27. Here, the lowest frequency point pulls the unweighted regression line downwards, resulting in $\hat{Q}_U = 254$, while for weighted regression the low variance around 20 Hz, due to strong coherence, causes significant compensation by upweighting the estimated log spectral ratio around 20 Hz. As a result, $\hat{Q}_W = 123$, a much more sensible value. These two figures illustrate extreme cases of none or large differences between the weighted and unweighted approaches, whereas Fig. 5 and Table I are an agglomeration of many more subtle effects.

$$M(t) = \frac{2^{K-1}c_1^K e^{p(K+1)/2}}{\Gamma(K)c_2^{K-1}} \int_0^\infty \int_0^\infty y^K u^{t+(K-1)/2} \times I_{K-1}(c_2 y e^{p/2} u^{1/2}) e^{-c_3[r+e^p u]y} du dy. \quad (33)$$

VII. Conclusion

In attenuation studies via spectral ratios, it is invariably assumed that the sequences involved are independent. We have shown that correlation between the sequences involved, as measured by ordinary coherence, can be exploited to adjust the frequency-dependent variance of estimated log spectral ratios; this variance was derived by explicitly calculating the cumulant generating function for the log spectral ratio. The frequency-dependent variance then informs a weighted least-squares approach to Q estimation. Studies using synthetic and real (seismic) data showed the efficacy of the method.

Appendix

The derivation of the MGF, $M(t)$, given in (10), begins by writing $f_{\hat{P}}(x) = e^{x+p} f_{\hat{R}}(e^{x+p})$, where $p = \log r$, and noting [11, p. 137] that

$$f_{\hat{R}}(x) = \int_0^\infty y f_{W_1,W_2}(y, xy) dy \quad (28)$$

where $f_{W_1,W_2}(\cdot,\cdot)$ is the joint pdf of W_1 and W_2 given by [6]

$$f_{W_1,W_2}(w_1,w_2) = \frac{2^{K-1}c_1^K}{\Gamma(K)c_2^{K-1}}(w_1 w_2)^{(K-1)/2} \times I_{K-1}(c_2[w_1 w_2]^{1/2}) e^{-c_3[r w_1 + w_2]} \quad (29)$$

where $I_\nu(\cdot)$ is the modified Bessel function of the first kind and order ν, $c_1 = 1/[\Sigma_{11}^4 r(1-\gamma^2)]$, $c_2 = 2\gamma/[\Sigma_{11}^2 r^{1/2}(1-\gamma^2)]$ and $c_3 = 1/[\Sigma_{11}^2 r(1-\gamma^2)]$.

So

$$f_{\hat{P}}(x) = e^{x+p} f_{\hat{R}}(e^{x+p}) = e^{x+p} \int_0^\infty y f_{W_1,W_2}(y, e^{x+p} y) dy \quad (30)$$

and therefore the MGF $M(t)$ is given by

$$E\{e^{\hat{P}t}\} = \int_{-\infty}^\infty e^{xt} f_{\hat{P}}(x) dx = \int_{-\infty}^\infty e^{xt} e^{x+p} \int_0^\infty y f_{W_1,W_2}(y, e^{x+p} y) dy dx. \quad (31)$$

Making the change of variable $u = e^x$, we obtain

$$M(t) = e^p \int_0^\infty u^t \int_0^\infty y f_{W_1,W_2}(y, e^p u y) dy du \quad (32)$$

which, using (29), becomes (33) at the top of the page. This may be conveniently written as

$$M(t) = \frac{2^{K-1}c_1^K e^{p(K+1)/2}}{\Gamma(K)c_2^{K-1}} \int_0^\infty y^K e^{-c_3 r y} G(y) dy \quad (34)$$

and $G(y) = \int_0^\infty u^{t+(K-1)/2} I_{K-1}(c_2 y e^{p/2} u^{1/2}) e^{-c_3 e^p u y} du$. Using [5, eq. 6.643.2], we obtain

$$G(y) = \frac{2\Gamma(K+t)}{c_2\Gamma(K)} y^{-[t+(K/2)+1]} c_3^{-[t+(K/2)]} e^{c_2^2 y/(8c_3)} \times M_{-[t+(K/2)],(K-1)/2}\left(\frac{c_2^2 y}{[4c_3]}\right) \quad (35)$$

for $t > -K$, where $M_{\kappa,\mu}(\cdot)$ is the Whittaker function. Substituting this back into (34) and making the substitution $w = c_2^2 y/(4c_3)$ gives

$$M(t) = \frac{2^{2(K-t)}\Gamma(K+t)c_1^K c_3^{-2t}}{\Gamma^2(K)c_2^{2(K-t)}e^{tp}} \cdot \int_0^\infty w^{K/2-t-1} e^{-\left[(4c_3^2 r/c_2^2)-1/2\right]w} \times M_{-[t+(K/2)],(K-1)/2}(w) dw. \quad (36)$$

Then, from [5, eq. 7.621.1], this can be written

$$M(t) = \frac{\Gamma(K+t)\Gamma(K-t)}{\Gamma^2(K)e^{tp}} \left[\frac{4c_1\gamma^2}{c_2^2}\right]^K \left[\frac{c_2}{2c_3\gamma}\right]^{2t} \times {}_2F_1(K-t, K+t; K; \gamma^2) \quad (37)$$

for $t < 1/4$ and $1/\gamma^2 > 1$; this latter condition always holds since $\gamma^2 < 1$. Once the forms of c_1, c_2 and c_3 are substituted, (10) is obtained, with validity $|t| < 1/4$.

Acknowledgment

The authors would like to thank the referees for constructive comments leading to an improved exposition.

References

[1] M. S. Bartlett and D. G. Kendall, "The statistical analysis of variance-heterogeneity and the logarithmic transformation," *Supp. J. Roy. Statist. Soc.*, vol. 8, pp. 128–138, 1946.

[2] S. Bose, "The distribution of the ratio of variances of samples drawn from a given normal bivariate correlated population," *Sankhya*, vol. 2, pp. 65–72, 1935.

[3] F. K. Forster, P. Guttorp, and E. L. Gow, "Variance reduction for ultrasonic attenuation measurements from backscatter in biological tissue," *IEEE Trans. Sonics Ultrason.*, vol. 32, no. 4, pp. 523–530, Jul. 1985.

[4] W. I. Futterman, "Dispersive body waves," *J. Geophys. Res.*, vol. 69, pp. 5279–5291, 1962.

[5] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products (Corrected and Enlarged Edition)*. New York: Academic, 1980.

[6] I. R. Joughin, D. P. Winebrenner, and D. B. Percival, "Probability density functions for multilook polarimetric signatures," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 3, pp. 562–574, May 1994.

[7] R. Kuc, "Estimating acoustic attenuation from reflected ultrasound signals: Comparison of spectral-shift and spectral difference approaches," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 1, pp. 1–6, Feb. 1984.

[8] R. Kuc and M. Schwartz, "Estimating the acoustic attenuation coefficient slope for liver from reflected ultrasound signals," *IEEE Trans. Son. Ultrason.*, vol. 26, no. 5, pp. 353–362, Sep. 1979.

[9] K. S. Miller, *Hypothesis Testing with Complex Distributions*. New York: Krieger, 1980.

[10] I. Olkin and J. W. Pratt, "Unbiased estimation of certain correlation coefficients," *Ann. Math. Stat.*, vol. 21, pp. 201–211, 1958.

[11] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 2nd ed. New York: McGraw-Hill, 1984.

[12] D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications*. Cambridge, U.K.: Cambridge Univ. Press, 1993.

[13] M. B. Priestley, *Spectral Analysis and Time Series*. New York: Academic, 1981.

[14] S. A. Raikes and R. E. White, "Measurements of earth attenuation from downhole and surface seismic recordings," *Geophys. Prospect.*, vol. 32, pp. 892–919, 1984.

[15] A. T. Walden, "A unified view of multitaper multivariate spectral estimation," *Biometrika*, vol. 87, pp. 767–787, 2000.

[16] Y. Wang, "Quantifying the effectiveness of stabilized Q filtering," *Geophys.*, vol. 68, pp. 337–345, 2003.

[17] R. E. White and P. N. S. O'Brien, "Estimation of the primary seismic pulse," *Geophys. Prospect.*, vol. 22, pp. 627–651, 1974.