Appendix A

Derivation of the Discrete Wavelet Transform

A.1 ANALYSIS ALGORITHM

The key to developing an efficient discrete-time implementation of the wavelet decomposition lies in recognizing a useful recursion. Because

$$\phi_0^0(t), \psi_0^0(t) \in \mathbf{V}_0 \subset \mathbf{V}_1,$$

there exists a pair of sequences $h[n]$ and $g[n]$ such that we can express these functions in terms of a basis for $\mathbf{V}_1$, i.e.,

$$\phi_0^0(t) = \sum_l h[l]\, \phi_l^1(t) \tag{A.1a}$$

$$\psi_0^0(t) = \sum_l g[l]\, \phi_l^1(t) \tag{A.1b}$$

where the coefficient $h[n]$ and $g[n]$ are given by the appropriate projections, viz., (2.20). Equivalently, we may express (A.1) in the frequency domain as

$$\Phi(\omega) = 2^{-1/2}\, H(\omega/2)\, \Phi(\omega/2) \tag{A.2a}$$

$$\Psi(\omega) = 2^{-1/2}\, G(\omega/2)\, \Phi(\omega/2). \tag{A.2b}$$

In any case, multiplying both sides of (A.1) by $2^{m/2}$, replacing t with $2^m t - n$, and effecting a change of variables we get, more generally,

$$\phi_n^m(t) = \sum_l h[l - 2n]\, \phi_l^{m+1}(t) \tag{A.3a}$$

$$\psi_n^m(t) = \sum_l g[l - 2n]\, \phi_l^{m+1}(t) \tag{A.3b}$$

where, in turn, we may rewrite (2.20) as

$$h[l-2n] = \int_{-\infty}^{\infty} \phi_n^m(t)\, \phi_l^{m+1}(t)\, dt, \tag{A.4a}$$

$$g[l-2n] = \int_{-\infty}^{\infty} \psi_n^m(t)\, \phi_l^{m+1}(t)\, dt. \tag{A.4b}$$

The discrete-time algorithm for the fine-to-coarse decomposition associated with the analysis follows readily. Specifically, substituting (A.3a) into (2.13) and (A.3b) into (2.5b), we get, for each m, the filter-downsample relations (2.21a) and (2.21b) defining the algorithm.

A.2 SYNTHESIS ALGORITHM

The coarse-to-fine refinement algorithm associated with the synthesis can be derived in a complementary manner. Since

$$\phi_n^{m+1}(t) \in \{\mathbf{V}_m \oplus \mathbf{O}_m\},$$

we can write

$$\begin{aligned} \phi_n^{m+1}(t) &= A_m\{\phi_n^{m+1}(t)\} + D_m\{\phi_n^{m+1}(t)\} \\ &= \sum_l \{h[n-2l]\, \phi_l^m(t) + g[n-2l]\, \psi_l^m(t)\} \end{aligned} \tag{A.5}$$

where the last equality follows by recognizing the projections in the respective expansions as (A.4). The upsample-filter-merge relation (2.21c) then follows immediately by substituting (A.5) into

$$a_n^{m+1} = \int_{-\infty}^{\infty} x(t)\, \phi_n^{m+1}(t)\, dt.$$

Appendix B

Proofs for Chapter 3

B.1 PROOF OF THEOREM 3.2

Let ω_0 and ω_1 be constants from Definition 3.1, and let $\lambda = \omega_1/\omega_0$. We first establish the following useful lemma.

Lemma B.1 *When a $1/f$ process $x(t)$ is passed through a filter with frequency response*

$$B_a(\omega) = \begin{cases} 1 & a\omega_0 < |\omega| \leq a\omega_1 \\ 0 & \textit{otherwise} \end{cases} \tag{B.1}$$

for any $a > 0$, the output $y_a(t)$ is wide-sense stationary, has finite variance, and has an autocorrelation satisfying

$$R_{y_a}(\tau) = E\left[y_a(t)\, y_a(t-\tau)\right] = a^{-2H} R_{y_1}(a\tau) \tag{B.2}$$

for all $a > 0$. Furthermore, for any distinct integers m and k, the processes $y_{\lambda^m}(t)$ and $y_{\lambda^k}(t)$ are jointly wide-sense stationary.

Proof:

First, from Definition 3.1 we have immediately that $y_1(t)$ is wide-sense stationary. More generally, consider the case $a > 0$. Let $b_a(t)$ be the impulse response of the filter with frequency response (B.1). To establish (B.2), it suffices to note that $y_a(t)$ has correlation function

$$R_{y_a}(t,s) \;=\; E\left[y_a(t)\, y_a(s)\right]$$

$$
\begin{aligned}
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} b_a(t-\alpha)\, b_a(s-\beta)\, R_x(\alpha,\beta)\, d\alpha\, d\beta \\
&= a^{-2H}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} b_1(at-\alpha)\, b_1(as-\beta)\, R_x(\alpha,\beta)\, d\alpha\, d\beta \\
&= a^{-2H} R_{y_1}(at, as) \qquad \text{(B.3)}
\end{aligned}
$$

where we have exploited the identities (3.2b) and

$$b_a(t) = a b_1(at).$$

However, since $y_1(t)$ is wide-sense stationary, the right side of (B.3) is a function only of $t-s$. Hence, $y_a(t)$ is wide-sense stationary and (B.2) follows. Furthermore, $y_a(t)$ has variance

$$R_{y_a}(0,0) = a^{-2H} R_{y_1}(0,0) < \infty$$

where the inequality is a consequence of Definition 3.1. To establish our final result, since $B_{\lambda^m}(\omega)$ and $B_{\lambda^k}(\omega)$ occupy disjoint frequency intervals for $m \neq k$, the spectra of $y_{\lambda^m}(t)$ and $y_{\lambda^k}(t)$ likewise occupy disjoint frequency intervals. Thus, $y_{\lambda^m}(t)$ and $y_{\lambda^k}(t)$ are uncorrelated, and, hence, jointly wide-sense stationary as well.

■

Proceeding now to a proof of our main theorem, let us establish that $y(t)$ is wide-sense stationary. Let M_L and M_U be any pair of integers such that

$$\lambda^{M_L}\omega_0 < \omega_L < \omega_U < \lambda^{M_U}\omega_1.$$

and consider preceding the filter (3.25) with a filter whose frequency response is

$$\tilde{B}(\omega) = \begin{cases} 1 & \lambda^{M_L}\omega_0 < |\omega| \leq \lambda^{M_U}\omega_1 \\ 0 & \text{otherwise} \end{cases} \qquad \text{(B.4)}$$

since this will not affect the output $y(t)$.

Let $\tilde{y}(t)$ be the output of the filter (B.4) when driven by $x(t)$. Then since

$$\tilde{B}(\omega) = \sum_{m=M_L}^{M_U} B_{\lambda^m}(\omega)$$

where $B_{\lambda^m}(\omega)$ is as defined in (B.1) of Lemma B.1, we can decompose $\tilde{y}(t)$ according to

$$\tilde{y}(t) = \sum_{m=M_L}^{M_U} y_{\lambda^m}(t) \qquad \text{(B.5)}$$

where $y_{\lambda^m}(t)$ is the response of the filter with frequency response $B_{\lambda^m}(\omega)$ to $x(t)$. Since, by Lemma B.1, all the terms comprising the summation (B.5) are jointly wide-sense stationary, $\tilde{y}(t)$ is wide-sense stationary. Then since $y(t)$ is obtained from $\tilde{y}(t)$ through the filter (3.25), the stationarity of $y(t)$ is an immediate consequence of the stationarity of $\tilde{y}(t)$ [40].

Let us now derive the form of the spectrum of $y(t)$, i.e., (3.26). We begin by rewriting (B.2) of Lemma B.1 in the frequency domain as

$$S_{y_1}(a\omega) = a^{-(2H+1)} S_{y_a}(\omega) \tag{B.6}$$

where $S_{y_a}(\omega)$ is the power spectrum associated with $y_a(t)$. For $1 < a < \lambda$, we observe that $S_{y_1}(\omega)$ and $S_{y_a}(\omega)$ have spectral overlap in the frequency range $a\omega_0 < |\omega| < \omega_1$, and can therefore conclude that the two spectra must be identical in this range. The reasoning is as follows. If we pass either $y_a(t)$ or $y_1(t)$ through the bandpass filter with frequency response

$$B^\dagger(\omega) = \begin{cases} 1 & a\omega_0 < |\omega| \leq \omega_1 \\ 0 & \text{otherwise} \end{cases}$$

whose impulse response is $b^\dagger(t)$, the outputs must identical, i.e.,

$$b^\dagger(t) * y_a(t) = b^\dagger(t) * y_1(t) = b^\dagger(t) * x(t).$$

Since $y_a(t)$ and $y_1(t)$ are jointly wide-sense stationary, we then conclude

$$S_{y_a}(\omega)|B^\dagger(\omega)|^2 = S_{y_1}(\omega)|B^\dagger(\omega)|^2$$

whence

$$S_{y_a}(\omega) = S_{y_1}(\omega), \qquad a\omega_0 < |\omega| < \omega_1. \tag{B.7}$$

Combining (B.7) with (B.6) we get

$$S_{y_1}(a\omega) = a^{-(2H+1)} S_{y_1}(\omega), \qquad a\omega_0 < |\omega| < \omega_1 \tag{B.8}$$

for any $1 < a < \lambda$. Differentiating (B.8) with respect to a and letting $a \to 1+$, we find that

$$\omega S'_{y_1}(\omega) = -(2H+1) S_{y_1}(\omega), \qquad \omega_0 < \omega < \omega_1,$$

and note that all positive, even, regular solutions to this equation are of the form

$$S_{y_1}(\omega) = \sigma_x^2/|\omega|^\gamma, \qquad \omega_0 < |\omega| \leq \omega_1 \tag{B.9}$$

for some $\sigma_x^2 > 0$ and $\gamma = 2H + 1$. Using (B.9) with (B.6) we find, further, that

$$S_{y_{\lambda^m}}(\omega) = \begin{cases} \sigma_x^2/|\omega|^\gamma & \lambda^m\omega_0 < |\omega| \leq \lambda^m\omega_1 \\ 0 & \text{otherwise} \end{cases}.$$

Via Lemma B.1, the $y_{\lambda^m}(t)$ are uncorrelated, so we deduce that $\tilde{y}(t)$ has spectrum

$$S_{\tilde{y}}(\omega) = \sum_{m=M_L}^{M_U} S_{y_{\lambda^m}}(\omega) = \begin{cases} \sigma_x^2/|\omega|^\gamma & \lambda^{M_L}\omega_0 < |\omega| \leq \lambda^{M_U}\omega_1 \\ 0 & \text{otherwise} \end{cases}.$$

Finally, since

$$S_y(\omega) = |B(\omega)|^2 S_{\tilde{y}}(\omega)$$

our desired result (3.26) follows.

■

B.2 PROOF OF THEOREM 3.3

To show that a fractional Brownian motion $x(t)$, for $0 < H < 1$, is a $1/f$ process according to Definition 3.1, it suffices to consider the effect on $x(t)$ of *any* LTI filter with a regular finite-energy impulse response $b(t)$ and frequency response $B(\omega)$ satisfying $B(\omega) = 0$. In particular, since $x(t)$ has correlation given by (3.16), the output of the filter

$$y(t) = \int_{-\infty}^{\infty} b(t-\tau)\, x(\tau)\, d\tau \tag{B.10}$$

has autocorrelation

$$\begin{aligned} R_y(t,s) &= E\left[y(t)y(s)\right] \\ &= \frac{\sigma_H^2}{2} \int_{-\infty}^{\infty} b(v)\, dv \int_{-\infty}^{\infty} |t-s+u-v|^{2H}\, b(u)\, du \end{aligned}$$

as first shown by Flandrin [44]. Since $R_y(t,s)$ is a function only of $t-s$, the process is stationary, and has spectrum

$$S_y(\omega) = |B(\omega)|^2 \cdot \frac{1}{|\omega|^{2H+1}}.$$

When we restrict our attention to the case in which $B(\omega)$ is the ideal bandpass filter (3.24), we see that $y(t)$ is not only stationary, but has finite variance. This establishes that any fractional Brownian motion $x(t)$ satisfies the definition of a $1/f$ process.

That the generalized derivative, fractional Gaussian noise $x'(t)$, is also a $1/f$ process follows almost immediately. Indeed, when $x'(t)$ is processed by the LTI filter with impulse response $b(t)$ described above, the output is $y'(t)$, the derivative of (B.10). Since $y(t)$ is stationary, so is $y'(t)$. Moreover, $y'(t)$ has spectrum

$$S_{y'}(\omega) = |B(\omega)|^2 \cdot \frac{1}{|\omega|^{2H'+1}}.$$

where H' is as given by (3.20). Again, when $B(\omega)$ is given by (3.24), $y'(t)$ is not only stationary, but has finite variance, which is our desired result.

■

B.3 PROOF OF THEOREM 3.4

Without loss of generality, let us assume $\sigma^2 = 1$. Next, we define

$$x_M(t) = \sum_{m=-M}^{M} \sum_{n} x_n^m \psi_n^m(t) \tag{B.11}$$

as a resolution-limited approximation to $x(t)$ in which information at resolutions coarser than 2^{-M} and finer than 2^M is discarded, so

$$x(t) = \lim_{M\to\infty} x_M(t) = \sum_m \sum_n x_n^m \psi_n^m(t).$$

Since for each m the wavelet coefficient sequence x_n^m is wide-sense stationary with spectrum $2^{-\gamma m}$, the approximation $x_M(t)$ is cyclostationary [40] with period 2^M, has finite variance, and has the associated time-averaged spectrum

$$S_M(\omega) = \sum_{m=-M}^{M} 2^{-\gamma m} |\Psi(2^{-m}\omega)|^2. \tag{B.12}$$

The limiting time-averaged spectrum

$$S_x(\omega) = \lim_{M\to\infty} S_M(\omega)$$

gives the desired spectrum expression (3.36), and corresponds to the time-averaged spectrum of $x(t)$ as measured at the output of a bandpass filter for each frequency ω in the passband. The desired octave-spaced ripple relation (3.38) for arbitrary integer k follows immediately from (3.36).

To establish (3.37), we begin by noting that, given ω, we can choose m_0 and ω_0 such that $\omega = 2^{m_0}\omega_0$ and $1 \le |\omega_0| < 2$. Hence, using (3.38) we see

$$S_x(\omega) = 2^{-m_0\gamma} S_x(\omega_0)$$

from which it follows that

$$\left[\inf_{1\le|\omega_0|<2} S_x(\omega_0)\right] \frac{1}{|\omega|^\gamma} \le S_x(\omega) \le \left[\sup_{1\le|\omega_0|<2} S_x(\omega_0)\right] \frac{2^\gamma}{|\omega|^\gamma}.$$

It suffices, therefore, to find upper and lower bounds for $S_x(\omega_0)$ on $1 \le |\omega_0| < 2$.

Since $\psi(t)$ is Rth-order regular, $\Psi(\omega)$ decays at least as fast as $1/\omega^R$ as $\omega \to \infty$. This, together with the fact that $\Psi(\omega)$ is bounded according to (2.8a), implies that

$$|\Psi(\omega)| \le \frac{C}{1+|\omega|^R}.$$

for some $C \ge 1$. Using this with (2.14a) in (3.36) leads to the upper bound

$$S_x(\omega_0) \le \sum_{m=0}^{\infty} 2^{-\gamma m} + \sum_{m=1}^{\infty} 2^{\gamma m} C^2 2^{-2Rm} < \infty.$$

To establish the lower bound it suffices to show $S_x(\omega) > 0$ for every $1 \le \omega \le 2$, which we establish by contradiction.

Suppose for some $1 \le \omega_0 \le 2$,

$$S_x(\omega_0) = \sum_m 2^{-\gamma m} |\Psi(2^{-m}\omega_0)|^2 = 0$$

Then since all the terms in the sum are non-negative, this would imply that each term is zero, from which we could conclude

$$\sum_m |\Psi(2^{-m}\omega_0)|^2 = 0.$$

However, this contradicts the wavelet basis identity (2.9). Hence, we must have that $S(\omega) > 0$ for every $\pi \leq \omega_0 \leq 2\pi$. The complete theorem follows.
■

B.4 PROOF OF THEOREM 3.5

We begin by defining the process $x_K(t)$ as the result of filtering $x(t)$ with the ideal bandpass filter whose frequency response is given by

$$B_K(\omega) = \begin{cases} 1 & 2^{-K} < |\omega| \leq 2^K \\ 0 & \text{otherwise} \end{cases}$$

so that

$$\lim_{K\to\infty} x_K(t) = x(t).$$

Then by Theorem 3.2, $x_K(t)$ is wide-sense stationary and has power spectrum

$$S_K(\omega) = \begin{cases} \sigma_x^2/|\omega|^\gamma & 2^{-K} < |\omega| \leq 2^K \\ 0 & \text{otherwise} \end{cases}. \tag{B.13}$$

If we denote its corresponding autocorrelation by

$$R_K(\tau) = E\left[x_K(t)x_K(t-\tau)\right]$$

and its wavelet coefficients by

$$x_n^m(K) = \int_{-\infty}^{\infty} x_K(t)\,\psi_n^m(t)\,dt,$$

the correlation between wavelet coefficients may be expressed as

$$\begin{aligned} E\left[x_n^m(K)x_{n'}^{m'}(K)\right] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi_n^m(t)\,R_K(t-\tau)\,\psi_{n'}^{m'}(\tau)\,dt\,d\tau \\ &= \int_{-\infty}^{\infty} \psi_n^m(t)\cdot\left[R_K(t) * \psi_{n'}^{m'}(t)\right]\,dt. \end{aligned} \tag{B.14}$$

Applying Parseval's theorem and exploiting (B.13), we may rewrite (B.14) in the frequency domain as

$$\begin{aligned} E\left[x_n^m(K)x_{n'}^{m'}(K)\right] &= \frac{2^{-(m+m')/2}}{2\pi}\left\{\int_{-2^K}^{-2^{-K}} \frac{\sigma_x^2}{|\omega|^\gamma}\Psi(2^{-m}\omega)\,\Psi^*(2^{-m'}\omega)\,d\omega\right. \\ &\quad \left. + \int_{2^{-K}}^{2^K} \frac{\sigma_x^2}{|\omega|^\gamma}\Psi(2^{-m}\omega)\,\Psi^*(2^{-m'}\omega)\,d\omega\right\}. \end{aligned} \tag{B.15}$$

Interchanging limits, we get

$$x_n^m = \lim_{K\to\infty} x_n^m(K)$$

and, in turn,

$$E\left[x_n^m x_{n'}^{m'}\right] = \lim_{K\to\infty} E\left[x_n^m(K) x_{n'}^{m'}(K)\right]. \tag{B.16}$$

Substituting (B.15) into (B.16) yields (3.40). Since

$$|E\left[x_n^m x_{n'}^{m'}\right]|^2 \le \operatorname{var} x_n^m \cdot \operatorname{var} x_{n'}^{m'},$$

and since

$$\operatorname{var} x_n^m = \frac{\sigma^2\sigma_x^2 2^{-\gamma m}}{\pi} J$$

where

$$J = \int_0^\infty \omega^{-\gamma}\,|\Psi(\omega)|^2\,d\omega, \tag{B.17}$$

it suffices to show that (B.17) converges. Because $\psi(t)$ has R vanishing moments, there exist constants C_0 and C_1 such that

$$|\Psi(\omega)| < C_0|\omega|^R \tag{B.18a}$$
$$|\Psi(\omega)| < C_1|\omega|^{-R}. \tag{B.18b}$$

Using (B.18) in (B.17), we obtain, for $0 < \gamma < 2R$ and $R \ge 1$,

$$J = \int_0^1 C_0^2\,\omega^{2R-\gamma}\,d\omega + \int_1^\infty C_1^2\,\omega^{-2R-\gamma}\,d\omega < \infty$$

as required.

■

B.5 PROOF OF THEOREM 3.6

Let us define

$$\Delta = 2^{-m}n - 2^{-m'}n'$$

and

$$\Xi(\omega) = \omega^{-\gamma}\Psi(2^{-m}\omega)\Psi^*(2^{-m'}\omega)$$

for $\omega > 0$, so that (3.41) may be expressed, via (3.40), as

$$\rho_{n,n'}^{m,m'} = \frac{\sigma_x^2}{\pi\sigma^2}\operatorname{Re} I(\Delta) \tag{B.19}$$

where

$$I(\Delta) = \int_0^\infty \Xi(\omega)\,e^{-j\Delta\omega}\,d\omega. \tag{B.20}$$

Thus, to establish the desired result, it suffices to show that $I(\Delta)$ has the appropriate decay.

We first note that if $\gamma \geq 2R+1$, then we cannot even guarantee that $I(\Delta)$ converges for any Δ. Indeed, since

$$\Xi(\omega) \sim \mathcal{O}\left(\omega^{2R-\gamma}\right), \qquad \omega \to 0$$

we see that $I(\Delta)$ is not absolutely integrable. However, provided $\gamma \leq 2R$, $I(\Delta)$ is absolutely integrable, i.e.,

$$\int_0^\infty |\Xi^{(Q)}(\omega)|\, d\omega < \infty.$$

In this case, we have, by the Riemann-Lebesgue lemma [50], that

$$I(\Delta) \to 0, \qquad \Delta \to \infty.$$

When $0 < \gamma < 2R$, we may integrate (B.20) by parts Q times, for some positive integer Q, to obtain

$$\begin{aligned} I(\Delta) &= \frac{1}{(j\Delta)^Q} \int_0^\infty \Xi^{(Q)}(\omega)\, e^{-j\Delta\omega}\, d\omega \\ &\quad + \sum_{q=0}^{Q-1} \frac{1}{(j\Delta)^q} \left\{ \lim_{\omega\to 0} \left[\Xi^{(q)}(\omega)\, e^{-j\Delta\omega}\right] - \lim_{\omega\to\infty} \left[\Xi^{(q)}(\omega)\, e^{-j\Delta\omega}\right] \right\}. \end{aligned} \tag{B.21}$$

Due to the vanishing moments of the wavelet we have

$$\Xi^{(q)}(\omega) \sim \mathcal{O}\left(\omega^{2R-\gamma-q}\right), \qquad \omega \to 0 \tag{B.22}$$

while due to the regularity of the wavelet, $\Psi(\omega)$ decays at least as fast as $1/\omega^R$ as $\omega \to \infty$, whence

$$\Xi^{(q)}(\omega) \sim \mathcal{O}\left(\omega^{-2R-\gamma-q}\right), \qquad \omega \to \infty. \tag{B.23}$$

Hence, the limit terms in (B.21) for which $-2R-\gamma < q < 2R-\gamma$ all vanish.

Moreover, when we substitute $q = Q$, (B.22) and (B.23) imply that $\Xi^{(Q)}(\omega)$ is absolutely integrable, i.e.,

$$\int_0^\infty |\Xi^{(Q)}(\omega)|\, d\omega < \infty, \tag{B.24}$$

whenever $-2R-\gamma+1 < Q < 2R-\gamma+1$, which implies, again via the Riemann-Lebesgue lemma, that the integral in (B.21) vanishes asymptotically, i.e.,

$$\int_0^\infty \Xi^{(Q)}(\omega)\, e^{-j\Delta\omega}\, d\omega \to 0, \qquad \Delta \to \infty. \tag{B.25}$$

Hence, choosing $Q = \lceil 2R-\gamma \rceil$ in (B.21) (so $2R-\gamma \leq Q < 2R-\gamma+1$) allows us to conclude

$$I \sim \mathcal{O}\left(\Delta^{-\lceil 2R-\gamma \rceil}\right), \qquad \Delta \to \infty. \tag{B.26}$$

Substituting (B.26) into (B.19) then yields the desired result.

■

Appendix C

The EM Parameter Estimation Algorithm

In this appendix, we derive the EM algorithm for the estimation of the signal and noise parameters $\Theta = \{\beta, \sigma^2, \sigma_w^2\}$ for the scenario described in Section 4.3.

We begin by defining our observed (incomplete) data to be

$$\mathbf{r} = \{r_n^m,\ m, n \in \mathcal{R}\},$$

and our complete data to be $(\mathbf{x}, \mathbf{r})$ where

$$\mathbf{x} = \{x_n^m,\ m, n \in \mathcal{R}\}.$$

Consequently, the EM algorithm for the problem is defined as [80]

E step: Compute

$$U(\Theta, \hat{\Theta}^{[l]}).$$

M step:

$$\max_{\Theta} U(\Theta, \hat{\Theta}^{[l]}) \rightarrow \hat{\Theta}^{[l+1]}$$

where

$$U(\Theta, \tilde{\Theta}) \triangleq E\left[\ln p_{\mathbf{r},\mathbf{x}}(\mathbf{r}, \mathbf{x}; \Theta) | \mathbf{r}; \tilde{\Theta}\right].$$

For our case, U is obtained conveniently via

$$U(\Theta, \tilde{\Theta}) = E\left[\ln p_{\mathbf{r}|\mathbf{x}}(\mathbf{r}|\mathbf{x}; \Theta) + \ln p_{\mathbf{x}}(\mathbf{x}; \Theta) | \mathbf{r}; \tilde{\Theta}\right]$$

with

$$p_{\mathbf{r}|\mathbf{x}}(\mathbf{r}|\mathbf{x}; \Theta) = \prod_{m,n \in \mathcal{R}} \frac{1}{\sqrt{2\pi\sigma_w^2}} \exp\left\{-\frac{(r_n^m - x_n^m)^2}{2\sigma_w^2}\right\}$$

and

$$p_{\mathbf{x}}(\mathbf{x};\Theta) = \prod_{m,n\in\mathcal{R}} \frac{1}{\sqrt{2\pi\sigma^2\beta^{-m}}} \exp\left\{-\frac{(x_n^m)^2}{2\sigma^2\beta^{-m}}\right\}.$$

Then

$$U(\Theta,\tilde{\Theta}) = -\frac{1}{2}\sum_{m\in\mathcal{M}} N(m)\left\{\frac{1}{\sigma_w^2}S_m^w(\tilde{\Theta}) + \ln 2\pi\sigma_w^2 + \frac{1}{\sigma^2\beta^{-m}}S_m^x(\tilde{\Theta}) + \ln 2\pi\sigma^2\beta^{-m}\right\} \tag{C.1}$$

where

$$\begin{aligned} S_m^w(\Theta) &= \frac{1}{N(m)}\sum_{n\in\mathcal{N}(m)} E\left[(w_n^m)^2|r_n^m;\Theta\right] \\ S_m^x(\Theta) &= \frac{1}{N(m)}\sum_{n\in\mathcal{N}(m)} E\left[(x_n^m)^2|r_n^m;\Theta\right] \end{aligned}$$

are (quasi) conditional sample-variance estimates from the data based upon the model parameters Θ. Evaluating the expectations we get

$$\begin{aligned} S_m^w(\Theta) &= A_m(\Theta) + B_m^w(\Theta)\hat{\sigma}_m^2 \\ S_m^x(\Theta) &= A_m(\Theta) + B_m^x(\Theta)\hat{\sigma}_m^2 \end{aligned}$$

where

$$\begin{aligned} A_m(\Theta) &= \frac{\sigma_w^2\cdot\sigma^2\beta^{-m}}{\sigma_w^2+\sigma^2\beta^{-m}} \\ B_m^w(\Theta) &= \left(\frac{\sigma_w^2}{\sigma_w^2+\sigma^2\beta^{-m}}\right)^2 \\ B_m^x(\Theta) &= \left(\frac{\sigma^2\beta^{-m}}{\sigma_w^2+\sigma^2\beta^{-m}}\right)^2, \end{aligned}$$

which completes our derivation of the E step.

To derive the structure of the M step, we maximize $U(\Theta,\tilde{\Theta})$ as given by (C.1). This maximization is always well defined as $U(\Theta,\tilde{\Theta}) \le L(\Theta)$ for any $\Theta,\tilde{\Theta}$.

The local extrema are obtained by differentiating $U(\Theta,\tilde{\Theta})$ with respect to each of the parameters of Θ. Since (C.1) expresses $U(\Theta,\tilde{\Theta})$ as the sum of two terms, one of which depends only on σ_w^2 and the other of which depends only on β and σ^2, the maximization can be broken down into two independent parts.

Considering first our maximization over σ_w^2, we readily obtain the maximizing $\hat{\sigma}_w^2$ as the sample-average

$$\hat{\sigma}_w^2 = \frac{\sum_{m \in \mathcal{M}} N(m) S_m^w(\tilde{\Theta})}{\sum_{m \in \mathcal{M}} N(m)}.$$

Turning next to β and σ^2, we find that the maximizing parameters $\hat{\beta}$ and $\hat{\sigma}^2$ satisfy

$$\sum_{m \in \mathcal{M}} N(m) S_m^x(\tilde{\Theta}) \beta^m = \sigma^2 \sum_{m \in \mathcal{M}} N(m) \tag{C.2a}$$

$$\sum_{m \in \mathcal{M}} m N(m) S_m^x(\tilde{\Theta}) \beta^m = \sigma^2 \sum_{m \in \mathcal{M}} m N(m). \tag{C.2b}$$

Eliminating σ^2 we obtain that $\hat{\beta}$ is the solution of the polynomial equation

$$\sum_{m \in \mathcal{M}} C_m N(m) S_m^x(\tilde{\Theta}) \beta^m = 0, \tag{C.3}$$

where C_m is as defined in (4.16). The eliminated variable $\hat{\sigma}^2$ is trivially obtained by back-substitution:

$$\hat{\sigma}^2 = \frac{\sum_{m \in \mathcal{M}} N(m) S_m^x(\tilde{\Theta}) \hat{\beta}^m}{\sum_{m \in \mathcal{M}} N(m)}.$$

Finally, to show that the maximizing parameters are the only solution to (C.2) it suffices to show that the solution to (C.3) is unique, which we establish via the following lemma.

Lemma C.1 *Any polynomial equation of the form*

$$\sum_{m \in \mathcal{M}} C_m K_m \beta^m = 0 \tag{C.4}$$

where C_m is given by (4.16) and $K_m \geq 0$ has a unique positive real solution provided $M \geq 2$ and not all K_m are zero.

Proof: Let

$$m_* = \frac{\sum_{m \in \mathcal{M}} m N(m)}{\sum_{m \in \mathcal{M}} N(m)}$$

be a weighted average of the $m \in \mathcal{M}$, so $m_1 < m_* < m_M$. Then, from (4.16), for $m > m_*$, $C_m > 0$, while for $m < m_*$, $C_m < 0$. Hence, $C_m(m - m_*) \geq 0$ with

strict inequality for at least two values of $m \in \mathcal{M}$ from our hypothesis. Now let $f(\beta)$ be the left-hand side of (C.4), and observe that

$$\tilde{f}(\beta) \triangleq f(\beta)\beta^{-m_*}$$

is increasing for $\beta > 0$, i.e.,

$$\tilde{f}'(\beta) = \sum_{m \in \mathcal{M}} C_m(m - m_*)N(m)\hat{\sigma}_m^2 \beta^{m-m_*-1} > 0.$$

Then, since $\tilde{f}(0) = -\infty$ and $\tilde{f}(\infty) = \infty$, we see $\tilde{f}(\beta)$ has a single real root on $\beta > 0$. Since $f(\beta)$ shares the same roots on $\beta > 0$, we have the desired result. ■

This completes our derivation for the M step. The complete algorithm follows directly.

Appendix D

Proofs for Chapter 5

D.1 PROOF OF THEOREM 5.2

To show that $y(t)$ has finite energy, we exploit an equivalent synthesis for $y(t)$ as the output of a cascade of filters driven by $x(t)$, the first of which is an ideal bandpass filter whose passband includes $\omega_L < |\omega| < \omega_U$, and the second of which is the filter given by (5.4).

Let $b_m(t)$ be the impulse response of a filter whose frequency response is given by

$$B_m(\omega) = \begin{cases} 1 & 2^m\pi < |\omega| \leq 2^{m+1}\pi \\ 0 & \text{otherwise} \end{cases}, \tag{D.1}$$

and let $b(t)$ be the impulse response corresponding to (5.4). Furthermore, choose finite integers M_L and M_U such that $2^{M_L}\pi < \omega_L$ and $\omega_U < 2^{M_U+1}\pi$. Then, using $*$ to denote convolution,

$$\begin{aligned} y(t) &= b(t) * \left[\sum_{m=M_L}^{M_U} b_m(t) \right] * x(t) \\ &= b(t) * \sum_{m=M_L}^{M_U} \tilde{x}_m(t) \end{aligned} \tag{D.2}$$

where

$$\tilde{x}_m(t) = x(t) * b_m(t) = 2^{-mH}\tilde{x}_0(2^m t), \tag{D.3}$$

and where the last equality in (D.3) results from an application of the self-similarity relation (5.2) and the identity

$$b_m(t) = 2^m b_0(2^m t).$$

Because $x(t)$ is energy-dominated, $\tilde{x}_0(t)$ has finite energy. Hence, (D.3) implies that every $\tilde{x}_m(t)$ has finite energy. Exploiting this fact in (D.2) allows us to conclude that $y(t)$ must have finite energy as well.

To verify the spectrum relation (5.5), we express (D.2) in the Fourier domain. Exploiting the fact that we may arbitrarily extend the limits in the summation in (D.2), we get

$$Y(\omega) = B(\omega) \sum_{m=-\infty}^{\infty} \tilde{X}_m(\omega) = \begin{cases} X(\omega) & \omega_L < |\omega| < \omega_U \\ 0 & \text{otherwise} \end{cases}$$

where $\tilde{X}_m(\omega)$ denotes the Fourier transform of $\tilde{x}_m(t)$, and where

$$X(\omega) \triangleq \sum_{m=-\infty}^{\infty} \tilde{X}_m(\omega). \tag{D.4}$$

The right-hand side of (D.4) is, of course, pointwise convergent because for each ω at most one term in the sum is non-zero. Finally, exploiting (D.3) in (D.4) gives

$$X(\omega) = \sum_m 2^{-m(H+1)} \tilde{X}_0(2^{-m}\omega),$$

which, as one can readily verify, satisfies (5.6).

■

D.2 PROOF OF THEOREM 5.3

To prove the "only if" statement, we suppose $x(t) \in \mathbf{E}^H$ and begin by expressing $x(t)$ in terms of the ideal bandpass wavelet basis. In particular, we let

$$x(t) = \sum_m \tilde{x}_m(t)$$

where

$$\tilde{x}_m(t) = \beta^{-m/2} \sum_n \tilde{q}[n] \tilde{\psi}_n^m(t)$$

and where $\tilde{q}[n]$, the generating sequence in this basis, has energy $\tilde{E} < \infty$. The new generating sequence $q[n]$ can then be expressed as

$$q[n] = \sum_m q_m[n] \tag{D.5}$$

where

$$q_m[n] = y_m(t)|_{t=n}$$

and

$$y_m(t) = \tilde{x}_m(t) * \psi(-t).$$

For each m, since $\tilde{x}_m(t)$ is bandlimited, $y_m(t)$ and $q_m[n]$ each have finite energy and Fourier transforms $Y_m(\omega)$ and $Q_m(\omega)$ respectively. Hence,

$$Q_m(\omega) = \sum_k Y_m(\omega - 2\pi k) \tag{D.6}$$

where

$$Y_m(\omega) = \begin{cases} (2\beta)^{-m/2}\Psi^*(\omega)\tilde{Q}(2^{-m}\omega) & 2^m\pi < |\omega| \leq 2^{m+1}\pi \\ 0 & \text{otherwise} \end{cases}$$

with $\tilde{Q}(\omega)$ denoting the Fourier transform of $\tilde{q}[n]$, and $\Psi^*(\omega)$ the complex conjugate of $\Psi(\omega)$.

In deriving bounds on the energy E_m in each sequence $q_m[n]$ for a fixed m, it is convenient to consider the cases $m \leq -1$ and $m \geq 0$ separately. When $m \leq -1$, the sampling by which $q_m[n]$ is obtained involves no aliasing. Since on $|\omega| \leq \pi$ we then have

$$Q_m(\omega) = Y_m(\omega),$$

we may deduce that $q_m[n]$ has energy

$$E_m = \sum_n |q_m[n]|^2 = \frac{(2\beta)^{-m}}{\pi} \int_{2^m\pi}^{2^{m+1}\pi} |\Psi(\omega)|^2 \, |\tilde{Q}(2^{-m}\omega)|^2 \, d\omega. \tag{D.7}$$

Because $\psi(t)$ has R vanishing moments, there exists a $0 < \epsilon_0 < \infty$ such that

$$|\Psi(\omega)| \leq \epsilon_0 |\omega|^R \tag{D.8}$$

for all ω. Exploiting this in (D.7) we obtain

$$E_m \leq C_0 2^{(2R-\gamma)m} \tilde{E} \tag{D.9}$$

for some $0 \leq C_0 < \infty$.

Consider, next, the case corresponding to $m \geq 0$. Since $\psi(t)$ has R vanishing moments, there also exists a $0 < \epsilon_1 < \infty$ such that

$$|\Psi(\omega)| \leq \epsilon_1 |\omega|^{-R} \tag{D.10}$$

for all ω. Hence, on $2^m\pi < |\omega| \leq 2^{m+1}\pi$,

$$|Y_m(\omega)| \leq \epsilon_1 \, \pi^{-R} \, 2^{-(\gamma+1+2R)m/2} \, |\tilde{Q}(2^{-m}\omega)|. \tag{D.11}$$

From (D.6), we obtain

$$|Q_m(\omega)| \leq \epsilon_1 \, \pi^{-R} \, 2^{-(\gamma+1+2R)m/2} \sum_{k=0}^{2^m-1} |\tilde{Q}(2^{-m}\omega + 2\pi k 2^{-m})| \tag{D.12}$$

by exploiting, in order, the triangle inequality, the bound (D.11), the fact that only 2^m terms in the summation in (D.6) are non-zero since $y_m(t)$ is

bandlimited, and the fact that $\tilde{Q}(\omega)$ is 2π-periodic. In turn, we may use, in order, (D.12), the Schwarz inequality, and again the periodicity of $\tilde{Q}(\omega)$ to conclude that

$$\begin{aligned} E_m &\leq \epsilon_1^2 \pi^{-2R} 2^{-(\gamma+1+2R)m} \left[\sum_{k=0}^{2^m-1} \sqrt{\frac{1}{2\pi} \int_{-\pi}^{\pi} |\tilde{Q}(2^{-m}\omega + 2\pi k 2^{-m})|^2 \, d\omega} \right]^2 \\ &\leq C_1 2^{-(\gamma-2+wR)m} \tilde{E} \end{aligned} \tag{D.13}$$

for some $0 \leq C_1 < \infty$.

Using (D.5), the triangle inequality, and the Schwarz inequality, we obtain the following bound on the energy in $q[n]$

$$E = \sum_n |q[n]|^2 \leq \left[\sum_m \sqrt{E_m} \right]^2$$

which from (D.13) and (D.9) is finite provided $0 < \gamma < 2R$ and $R \geq 1$.

Let us now show the converse. Suppose $q[n]$ has energy $E < \infty$, and express $x(t)$ as

$$x(t) = \sum_m x_m(t)$$

where

$$x_m(t) = \beta^{-m/2} \sum_n q[n] \, \psi_n^m(t).$$

If we let

$$\tilde{y}_m(t) = b_0(t) * x_m(t)$$

where $b_0(t)$ is the impulse response of the ideal bandpass filter in Definition 5.1, it suffices to show that

$$\tilde{y}(t) = \sum_m \tilde{y}_m(t) \tag{D.14}$$

has finite energy.

For each m, we begin by bounding the energy in $\tilde{y}_m(t)$, which is finite because $x_m(t)$ has finite energy. Since $\tilde{y}_m(t)$ has Fourier transform

$$\tilde{Y}_m(\omega) = \begin{cases} (2\beta)^{-m/2} Q(2^{-m}\omega) \, \Psi(2^{-m}\omega) & \pi < |\omega| \leq 2\pi \\ 0 & \text{otherwise} \end{cases}$$

where $Q(\omega)$ is the discrete-time Fourier transform of $q[n]$, we get

$$\tilde{E}_m = \frac{2^{-\gamma m}}{\pi} \int_{2-m\pi}^{2^{-m+1}\pi} |Q(\omega)|^2 \, |\Psi(2^{-m}\omega)|^2 \, d\omega.$$

Again, it is convenient to consider the cases corresponding to $m \leq -1$ and $m \geq 0$ separately. For $m \leq -1$, most of the energy in $x_m(t)$ is at frequencies

below the passband of the bandpass filter. Hence, using the bound (D.10) and exploiting the periodicity of $Q(\omega)$ we obtain

$$\tilde{E}_m \leq \tilde{C}_0 2^{(2R-1-\gamma)m} E. \tag{D.15}$$

for some $0 \leq \tilde{C}_0 < \infty$. For $m \geq 0$, most of the energy in $x_m(t)$ is at frequencies higher than the passband of the bandpass filter. Hence, using the bound (D.8) we obtain

$$\tilde{E}_m \leq \tilde{C}_1 2^{-(\gamma+2R+1)m} E. \tag{D.16}$$

for some $0 \leq \tilde{C}_1 < \infty$.

Finally, using (D.14), the triangle inequality, and the Schwarz inequality, we obtain the following bound on the energy in $\tilde{y}(t)$

$$\tilde{E} = \int_{-\infty}^{\infty} |\tilde{y}(t)|^2 \, dt \leq \left[\sum_m \sqrt{\tilde{E}_m} \right]^2$$

which, from (D.16) and (D.15) is finite provided $0 < \gamma < 2R - 1$ since $R \geq 1$.

■

D.3 PROOF OF THEOREM 5.5

Following an approach analogous to the proof of Theorem 5.2, let $b_m(t)$ be the impulse response of a filter whose frequency response is given by (D.1), and let $b(t)$ be the impulse response corresponding to (5.4). By choosing finite integers M_L and M_U such that $2^{M_L}\pi < \omega_L$ and $\omega_U < 2^{M_U+1}\pi$, we can again express $y(t)$ in the form of eq. (D.2). Because $x(t)$ is power-dominated, $\tilde{x}_0(t)$ has finite power. Hence, (D.3) implies that every $\tilde{x}_m(t)$ has finite power. Exploiting this fact in (D.2) allows us to conclude that $y(t)$ must have finite power as well.

To verify the spectrum relation (5.23), we use (D.2) together with the fact that the $\tilde{x}_m(t)$ are uncorrelated for different m to obtain

$$S_y(\omega) = |B(\omega)|^2 \sum_{m=-\infty}^{\infty} S_{\tilde{x}_m}(\omega) = \begin{cases} S_x(\omega) & \omega_L < |\omega| < \omega_U \\ 0 & \text{otherwise} \end{cases}$$

where $S_{\tilde{x}_m}(\omega)$ denotes the power spectrum of $\tilde{x}_m(t)$, and where

$$S_x(\omega) \triangleq \sum_{m=-\infty}^{\infty} S_{\tilde{x}_m}(\omega). \tag{D.17}$$

Again we have exploited the fact that the upper and lower limits on the summation in (D.2) may be extended to ∞ and $-\infty$, respectively. The right-hand side of (D.17) is, again, pointwise convergent because for each ω at most one term in the sum is non-zero. Finally, exploiting (D.3) in (D.17) gives

$$S_x(\omega) = \sum_m 2^{-\gamma m} S_{\tilde{x}_0}(2^{-m}\omega)$$

which, as one can readily verify, satisfies (5.24).

■

D.4 PROOF OF THEOREM 5.6

We first establish some notation. Let us denote the cross-correlation between two finite-power signals $f(t)$ and $g(t)$ by

$$R_{f,g}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(t)\, g(t-\tau)\, dt.$$

Its Fourier transform is the corresponding cross-spectrum $S_{f,g}(\omega)$. Similarly

$$R_{a,b}[k] = \lim_{L\to\infty} \frac{1}{2L+1} \sum_{n=-L}^{L} a[n]\, b[n-k]$$

will denote the cross-correlation between two finite-power sequences $a[n]$ and $b[n]$.

We begin by expressing $x(t)$ as

$$x(t) = \sum_m x_m(t)$$

where

$$x_m(t) = \beta^{-m/2} \sum_n q[n] \psi_n^m(t).$$

Then the deterministic power spectrum of $x(t)$ is given by

$$S_x(\omega) = \sum_m \sum_{m'} S_{x_m, x_{m'}}(\omega). \tag{D.18}$$

We will proceed to evaluate these various terms. Because of the dilational relationships among the $x_m(t)$, viz.,

$$x_m(t) = 2^{m/2} \beta^{-m/2} x_0(2^m t),$$

it will suffice to consider a single term of the form $S_{x_0, x_m}(t)$, for some $m \geq 0$.

Hence, let

$$v_m(t) = \beta^{-m/2} \sum_n q[n] \delta(t - 2^{-m} n)$$

and note that

$$v_0(t) = \sum_n \tilde{q}[n] \delta(t - 2^{-m} n)$$

where $\tilde{q}[n]$ is an upsampled version of $q[n]$, i.e.,

$$\tilde{q}[n] = \begin{cases} q[2^{-m} n] & n = 2^m l, \quad l = \ldots, -1, 0, 1, 2, \ldots \\ 0 & \text{otherwise} \end{cases}.$$

Hence,

$$R_{v_0, v_m}(\tau) = \sum_k R_{\tilde{q}, q}[k] \delta(t - 2^{-m} k)$$

where

$$\begin{aligned} R_{\tilde{q},q}[k] &= \lim_{L\to\infty} \frac{1}{2L+1} \sum_{|n|\le L, n=2^m l} q[2^{-m}n]\, q[n-k] \\ &= \lim_{L\to\infty} \frac{1}{2L+1} \sum_{|l|\le 2^{-m}L} q[l]\, q[2^m l - k]. \end{aligned}$$

Since $q[n]$ is correlation-ergodic, we may replace this correlation with its expected value:

$$R_{\tilde{q},q}[k] = \lim_{L\to\infty} \frac{1}{2L+1} \sum_{|l|\le 2^{-m}L} \delta[(2^m-1)l - k] = \begin{cases} \delta[k] & m=0 \\ 0 & \text{otherwise} \end{cases}.$$

Hence,

$$S_{v_0,v_m}(\omega) = \begin{cases} 1 & m=0 \\ 0 & \text{otherwise} \end{cases}$$

where, without loss of generality, we have set $\sigma^2 = 1$. Then, using

$$S_{x_0,x_m}(\omega) = (2\beta)^{-m/2}\, \Psi(\omega)\, \Psi^*(2^{-m}\omega)\, S_{v_0,v_m}(\omega)$$

we get that

$$S_{x_0,x_m}(\omega) = \begin{cases} |\Psi(\omega)|^2 & m=0 \\ 0 & \text{otherwise} \end{cases}. \tag{D.19}$$

Finally, we note that

$$S_{x_m,x_{m'}}(\omega) = S^*_{x_{m'},x_m}(\omega)$$

and that

$$S_{x_m,x_{m'}}(\omega) = \beta^{-m'} S_{x_0,x_{m-m'}}(2^{-m'}\omega).$$

Using these identities together with (D.19) in (D.18) yields

$$S_x(\omega) = \sum_m \beta^{-m} |\Psi(2^{-m}\omega)|^2$$

as desired.

■