2pUWa4. Synthetic-array beamforming for bottom-loss estimation using marine ambient noise

Lanfranco Muzi* and Martin Siderius

*Corresponding author’s address: Electrical and Computer Engineering, Portland State University, 1900 SW 4th Avenue, Portland, Oregon 97201, muzi@pdx.edu

Previous studies have shown that, using a vertical line array, the bottom loss can be estimated in an ambient-noise field from the output power of beams steered towards the sea surface and beams reflected off the seabed. With short arrays, the low angular resolution of the bottom-loss estimate is one of the main limitations of the approach. Synthetic-array processing is proposed as a technique that can improve the angular resolution of the bottom-loss estimate to a level comparable to that of an array with twice as many physical sensors (at equal inter-sensor spacing). The proposed technique follows naturally from a new derivation in frequency-wavenumber domain of the bottom-loss estimation procedure. The conditions under which the approach can be successfully applied are analyzed, with particular regard to the need for the array cross spectral density matrix to be close to Toeplitz. The technique is illustrated and then applied to data from experimental campaigns. Results show that a 16-element synthetic array can achieve an angular resolution comparable to that of a 32-element array.
INTRODUCTION

The seabed bottom loss is an important quantity for predicting transmission loss in the ocean. A simple passive technique for estimating the bottom loss has been developed by Harrison and Simons (in the sonar field the word passive designates techniques that do not make use of artificial acoustic sources such as sound projectors or explosive charges.) In this technique, the marine ambient-noise field, mainly originating from breaking waves, wind and rain at the surface, is sampled at discrete locations in space by a vertical line array of hydrophones. The data are then beamformed to obtain estimates of the power impinging on the array from different grazing angles. The ratio between averaged noise coming directly from the seabed with that coming from the surface (at opposite grazing angles) reveals the loss due to interaction with the seabed, which — by definition — is the bottom loss.

With increasing interest for short arrays, which can be more easily deployed (even on autonomous underwater vehicles) and potentially eliminate array-mismatch errors due to geometric deformation of the array, poor angular resolution becomes a matter of concern. The consequences on the estimated bottom loss can include a shift in the location of the critical angle and, if the seabed is layered, a reduction of the level of interference features in the computed bottom loss. These effects can introduce significant errors when the estimated bottom loss is used directly in propagation models, or in an inversion scheme to estimate geo-acoustic properties of the seabed.

This paper proposes synthetic-array processing as a technique that can improve the angular resolution of an array to a level comparable to that of an array with twice as many sensors (at equal inter-sensor spacing). This is done by properly “augmenting” the cross-spectral-density (CSD) matrix measured by the array. For the technique to work, the prerequisite is that the CSD matrix be (approximately) Toeplitz, which implies that the noise spatial coherence function between two hydrophones depends only on the distance between the hydrophones, and not their absolute position in the water column. This property of the CSD matrix was asserted by Buckingham for marine ambient noise in deep water. Harrison estimated that the spatial coherence function becomes weakly dependent on sensor depth at a distance from the waveguide boundaries of the order of a few wavelengths. When this condition is met, this study found that the CSD matrix can be sufficiently close to Toeplitz to allow synthetic-array processing also in shallow water.

DERIVATION OF THE POWER REFLECTION COEFFICIENT FROM THE NOISE SPATIAL COHERENCE FUNCTION

This section presents the derivation of a formula for computing the power reflection coefficient from the spatial coherence function (or cross spectral density) of the surface-generated marine noise field in a lossless water column of constant sound speed. A more complete derivation of this result, as well as the extension to lossy media in the presence of a sound-speed profile, can be found in a separate publication.

The spatial coherence function of the pressure field \( p(\mathbf{r},t) \) between two points in space \( \mathbf{r}_1 \) and \( \mathbf{r}_2 \) is defined as the ensemble average of the product \( p_\omega(\mathbf{r})p_\omega^*(\mathbf{r}) \):

\[
C_\omega(\mathbf{r}_1,\mathbf{r}_2) = \langle p_\omega(\mathbf{r}_1)p_\omega^*(\mathbf{r}_2) \rangle,
\]

where \* indicates complex conjugate and \( p_\omega(\mathbf{r}) \) is the coefficient of the Fourier expansion of \( p(\mathbf{r},t) \) at angular frequency \( \omega \).

Harrison derived a formula for the spatial coherence function of surface generated noise in the ocean, which for the case of two hydrophones joined by a perfectly vertical line and separated by a distance \( z \) is written:

\[
C_\omega (z) = \int_{0}^{\pi/2} 2\pi \sin \theta \cos \theta \frac{2}{1 - R(\theta)R(\theta)e^{-\omega z}} \left[ e^{i(\omega \theta) - \omega z} + R(\theta)e^{-\omega z} e^{i(\omega \theta) - \omega z} \right] d\theta.
\]

In Eq.(2), \( C_\omega (z) \) introduces a more compact notation, defined as follows:

\[
C_\omega (z) = C_\omega (\mathbf{r}_1, \mathbf{r}_2);
\]

\[
\mathbf{r}_1 = (0,0), \mathbf{r}_2 = (0, z).
\]
which means that, when computing $C_n(z)$, the hydrophone pair is assumed to be aligned with the $z$ axis, with the first hydrophone at $z = 0$. (Given the reference defined in FIGURE 1, the position of the second hydrophone is below the first one when $z$ is positive, and above when $z$ is negative.) In Eq.(2) $\theta_s$, $\theta_i$, and $\theta_b$ are the ray angles at the receiver, the surface, and the bottom; $s_\theta$ and $s_r$ are the complete and partial ray-path lengths; $\omega$ is the angular frequency; $c$ is the sound speed at the receiver in the medium, and $R$ and $R_s$ are the bottom and surface power reflection coefficients. In general, besides the ray angle, the reflection coefficients are also a function of frequency, but for the sake of simplicity this dependence will not be indicated explicitly. Note that $a$ is the power attenuation coefficient per unit length.

The model assumes that the hydrophones are “close”, so the ray paths and the sound speed are unambiguously defined (see FIGURE 1 for the definition of the coordinate system and all geometric quantities).

As a first approximation, the volume attenuation can be neglected in Eq.(2), and let:

$$G(\theta_s, \theta_i) = \frac{2\pi \sin \theta_i}{1 - R_s(\theta_i) R(\theta_i)},$$

and:

$$k = \frac{\omega}{2\pi \lambda} \sin \theta_i = \frac{\sin \theta_i}{\lambda},$$

(where $\lambda$ is the signal wavelength) so that:

$$\theta_s = a \sin (\lambda k),$$

$$d\theta_s = -\frac{\lambda}{\cos \theta_s} dk.$$  

Assuming constant sound speed in the water column yields:

$$\theta_b = \theta_i = |\theta|.$$
explicitly showing that, by definition, $\theta_0$ and $\theta_1$ only span values between zero and $\pi/2$, and are even functions of $\theta$ (see FIGURE 1). Letting $\theta = |\theta|$ and switching the independent variable to $k$ by use of Eq.(6) yields:

$$\tilde{G}(k) = \frac{2\pi|k|}{\lambda} 1 - R(k) R(k),$$

(8)

where, to avoid an excessive number of symbols, with a slight abuse of notation the reflection coefficients are designated with the same symbols as in Eq.(4). Note that, given Eq.(5) and Eq.(7), $R(k)$ and $R'(k)$ are even functions of $k$, therefore $\tilde{G}(k)$ is even in $k$ too. Introducing the generalized rectangle function $r(x)$ with height $h=1$, full width $b=1/\lambda$, center $c=1/(2\lambda)$:

$$r(x) = h\Pi\left[\frac{(x-c)}{b}\right],$$

(9)

$$\Pi(x) = \begin{cases} 0 & \text{for } |x| > 1/2, \\ 1 & \text{for } |x| \leq 1/2, \end{cases}$$

and using the Fourier-Transform pair:

$$F(y) = \mathcal{F}\{f(x)\} = \int_{-\infty}^{\infty} f(x)e^{-2\pi iy} dx$$

$$f(x) = \mathcal{F}^{-1}\{F(y)\} = \int_{-\infty}^{\infty} F(y)e^{2\pi iy} dy,$$

(10)

the Fourier transform of Eq.(2) can be written as:

$$C_o(k) = \mathcal{F}\{C_o(z)\} = \mathcal{F}\{\tilde{G}(k)\Pi(\lambda k - 1/2)\} + \mathcal{F}\{\tilde{G}(k)R(k)\Pi(\lambda k - 1/2)\}$$

$$= \mathcal{F}^-(k) + \mathcal{F}^+(k),$$

(11)

where:

$$\mathcal{F}^-(k) = \mathcal{F}\{\tilde{G}(k)\Pi(\lambda k - 1/2)\} = \tilde{G}(k)\Pi(\lambda k - 1/2),$$

$$\mathcal{F}^+(k) = \mathcal{F}\{\tilde{G}(k)R(k)\Pi(\lambda k - 1/2)\} = \tilde{G}(k)R(k)\Pi(\lambda k + 1/2).$$

(12)

Equations (11) and (12) show that $C_o(k)$, the $k$-spectrum of the coherence function, is split into a portion $\mathcal{F}^+(k)$, which is nonzero only for positive $k$ values, and a portion $\mathcal{F}^-(k)$, which is nonzero only for negative $k$ values. $R(k)$ can now be computed as the ratio:

$$R(k) = \mathcal{F}^-(k)/\mathcal{F}^+(k), \quad k \in \left[0, 1/\lambda\right],$$

(13)

where use has been made of the fact that $\tilde{G}(k)$ and $R(k)$ are even functions of $k$.

Since negative values of $k$ correspond to waves reaching the hydrophones after reflection from the bottom — i.e., $\theta_0 < 0$ in Eq.(5) — and positive values of $k$ correspond to reflection from the surface — i.e., $\theta_0 > 0$ in Eq.(5) — the result in Eq.(13) is equivalent to the method for estimating $R$ described by Harrison and Simons. They derived it through an energy-flux argument, whereas here a frequency-wavenumber domain derivation is presented, from which the developments presented in this study follow more naturally. Note that, because of the rectangle functions in Eq.(12), the power reflection coefficient $R(k)$ in Eq. (13) is defined only for $k \in \left[0, 1/\lambda\right]$, i.e. $\theta \in \left[0, \pi/2\right]$, which were the integration limits in Eq.(2).
BOTTOM-LOSS ESTIMATION AND SYNTHETIC-ARRAY PROCESSING

Beamforming and Power Reflection Coefficient

Bottom loss estimation is the application of interest to this study. For a wave front of frequency $\omega$ incident upon the bottom at grazing angle $\theta_b > 0$ (see FIGURE 1), the bottom loss is defined as:

$$BL(\theta_b,\omega) = -10 \log R(\theta_b,\omega),$$  \hspace{1cm} (14)

where $R(\theta_b,\omega)$ is the power reflection coefficient of the bottom. Harrison and Simons show that the bottom loss can successfully be computed from an estimate $\hat{R}(\theta_b,\omega)$ of the power reflection coefficient obtained from array data as the ratio of the downward and upward beam powers$^1$:

$$\hat{R}(\theta_b,\omega) = \frac{B(-\theta_b,\omega)}{B(\theta_b,\omega)}.$$  \hspace{1cm} (15)

The beam power $B(\phi,\omega)$ is defined as (for the sake of simplicity, in the following the dependence on frequency and grazing angle will often be dropped in the right hand side of equations):

$$B(\phi,\omega) = w^T p(w^T p)^T = w^T (pp^T) w.$$  \hspace{1cm} (16)

In Eq.(16) $H$ denotes the conjugate transpose operation and $w(\phi,\omega) = [w_1, w_2, \ldots, w_M]^T$ is the weight vector for the steering angle $\phi$ ($T$ denotes the transpose operation)$^1$. The angle $\phi = 0$ corresponds to the array being steered towards broadside, $\phi > 0$ towards the surface, and $\phi < 0$ towards the bottom. The vector $p(\omega) = [p_1(\omega), p_2(\omega), \ldots, p_M(\omega)]^T$ represents the data from the $M$ hydrophones in the array.

For the classical delay-and-sum beamformer, (here also referred to as “conventional beamformer”, or “CBF”), the $m$-th component of the weight vector (i.e., the weight for the $m$-th element in the array) is computed as:

$$w_m(\phi,\omega) = \frac{1}{\sqrt{M}} e^{-i(m-1)\omega c \sin \phi},$$  \hspace{1cm} (17)

where $c$ is the sound speed and $d$ is the array inter-element spacing.

The spatial coherence matrix (or cross-spectral-density matrix, hereafter also referred to as “CSD matrix”) $C_\omega$ is defined as the expected value of the outer product $p(\omega)p^H(\omega)$:

$$C_\omega = E[pp^H]$$  \hspace{1cm} (18)

Typically, in real-world applications several realizations $p_i(\omega)$ of $p(\omega)$ (each called a snapshot) are collected, and an estimate $\hat{C}_\omega$ of $C_\omega$ is obtained by averaging $p_i(\omega)p_i^H(\omega)$ over the entire collection time; this estimate is then used to replace $[pp^H]$ in Eq.(16), yielding:

$$B(\phi,\omega) = w^T \hat{C}_\omega w.$$  \hspace{1cm} (19)

$^1$ $\phi$ and $\theta_b$ are conceptually different, but, given the way the two angles are defined, if the array is steered in the direction of propagation of the plane wave, then $\phi = \theta_b$. 

---

1. L. Muzi and M. Siderius

Equation (15) shows that, in bottom-loss estimation, the ratio of the beamformer output power is used to estimate the power ratio of (plane) wave fronts incident upon the array from symmetric grazing angles. More consistently with our treatment, the angle of incidence can be expressed in terms of the vertical wavenumber $k$, (note that Eq.(5) defines $k$ as a scaled vertical wavenumber at the receiver: $k = k_L / 2\pi$). The ability of the beamformer to discriminate between wave fronts incident from closely spaced angles is the beamformer resolution. Several ways of defining the resolution exist; adopting the definition based on the Rayleigh criterion, the resolution in wavenumber domain for a linear array is:

$$\Delta k = 2\pi / L,$$

where $\Delta k$ is the difference between the two closest values of $k$ that can be resolved and $L = d(M - 1)$ is the total length of the array.

**Synthetic-Array Processing**

It has been shown above that it is possible, for a noise-only field and under certain conditions, to recover $R(k)$ by computing the Fourier transform of the coherence function $C_o(z)$. This theoretical link between $R(k)$ and $C_o(z)$ finds experimental confirmation in Harrison and Simons’ method for bottom-loss estimation, where the CSD matrix $C_o$ is included in the beamforming process.

When working with array data, measurements are only available at the locations of the sensors, so the coherence function $C_o(z)$ is sampled at discrete intervals (integer multiples of $d$) in the $z$ variable, and its Fourier transform in Eq.(11) becomes a discrete Fourier transform (DFT). The resolution in $k$ of the DFT increases with the number of samples, which is equivalent to increasing the number of elements in the array. This is consistent with Eq.(20), which shows that, given the inter-element spacing $d$, the beamformer resolution increases with the number of elements.

Increasing the resolution, at this point, may seem as easy as zero-padding $C_o(z)$ in Eq.(11), but in this application this approach does not necessarily yield acceptable results. The explanation is found in the more intuitive beamforming approach, where increasing the number of samples of $C_o(z)$ would correspond to “augmenting” the CSD matrix obtained from array data by adding new elements (i.e., zeros) to its rows and columns. Since the matrix must remain square, this poses the problem of how all the additional elements should be filled, especially below the $M$–th row (an illustration of the problem, for the case of a simple, 3-element array, is given in FIGURE 2).

$$C_o = \begin{bmatrix} c_{1,1} & c_{1,2} & c_{1,3} \\ c_{2,1} & c_{2,2} & c_{2,3} \\ c_{3,1} & c_{3,2} & c_{3,3} \end{bmatrix} \Rightarrow \begin{bmatrix} c_{1,1} & c_{1,2} & c_{1,3} \\ c_{1,2} & c_{2,2} & c_{2,3} \\ c_{1,3} & c_{2,3} & c_{3,3} \end{bmatrix} \Rightarrow \begin{bmatrix} c_{1,1} & c_{1,2} & c_{1,3} & 0 & 0 & 0 \\ c_{1,2}^* & c_{2,2} & c_{2,3} & ? & ? & ? \\ c_{1,3}^* & c_{2,3} & c_{3,3} & ? & ? & ? \\ ? & ? & ? & ? & ? & ? \\ ? & ? & ? & ? & ? & ? \end{bmatrix}$$

**FIGURE 2.** CSD matrix “augmentation” for a 3-element array ($M = 3$). The “$H$” equality can be stated due to the Hermitian character of $C_o$, but augmenting the matrix to dimension $6\times6$ is not possible without further assumptions.

One useful consideration is that, being the result of the outer product $p(\omega)p^*(\omega)$, the CSD matrix $C_o$ is Hermitian by construction. If negligible volume attenuation is assumed, the dependence of Eq.(2) on hydrophone-pair depth is quite weak (Buckingham had reached this conclusion for deep water). In these conditions, the coherence function between two hydrophones is almost exclusively a function of the distance between the two, and does not vary appreciably with their depth over the length of the array, i.e.: $C_o(r_i, r_j) = C_o(r, r_n)$ when $(i - j) = (l - m)$. This has significant consequences for the structure of $C_o$, implying that, besides being Hermitian, the matrix is also Toeplitz.
If $\omega C$ is Hermitian Toeplitz, then the augmentation step in FIGURE 3 is possible, reducing significantly the number of “arbitrary” zeros in the augmented matrix.

$$
\begin{bmatrix}
c_{1,1} & c_{1,2} & c_{1,3} \\
c_{1,2}^* & c_{2,2} & c_{2,3} \\
c_{1,3}^* & c_{2,3}^* & c_{3,3}
\end{bmatrix} 
\rightarrow 
\begin{bmatrix}
c_{1,1} & c_{1,2} & 0 & 0 & 0 \\
c_{1,2}^* & c_{1,2} & c_{1,3} & 0 & 0 \\
c_{1,3}^* & c_{1,3}^* & c_{1,1} & c_{1,2} & c_{1,3} \\
0 & 0 & c_{1,3} & c_{1,2}^* & c_{1,1} \\
0 & 0 & 0 & c_{1,3}^* & c_{1,2}^* \\
0 & 0 & 0 & 0 & c_{1,1}^*
\end{bmatrix}
$$

FIGURE 3. CSD matrix “augmentation” for a 3-element array ($M = 3$). The augmentation step indicated by “$T$” is possible under the Toeplitz assumption and reduces significantly the number of arbitrary values in the augmented matrix.

The size of the new CSD matrix now corresponds to a “synthetic” $2M − 1$-element array which, according to Eq.(20) should afford a higher angular resolution. Therefore, successful application of the synthetic-array technique requires a CSD matrix that is (approximately) Toeplitz — i.e., weak dependence of the coherence function on the depth of sensors.

Of course, the CSD matrix for the synthetic array is not actually complete, because the coherence function is only computable from data for a maximum hydrophone spacing of $|r_i − r_j|_{\text{max}} = d(M − 1)$, i.e., the total length of the original array. Nevertheless, the magnitude of $C_{\omega}(r_i, r_j)$ decreases with increasing $|r_i − r_j|$ (see, e.g., Buckingham and Harrison — this effect is more apparent at higher frequencies), so that the terms replaced by zeros do not contribute as much as those clustered around the main diagonal. Another caveat of the methodology is the discontinuity in each row between the first or last nonzero element and the neighboring zero. This discontinuity can be quite severe especially at low frequencies, and it is not observed in real CSD matrices. In such cases, it proved beneficial in this study to smooth the transition by applying an appropriate window that tapers the nonzero portion of each row.

APPLICATION TO MEASURED DATA

The application of synthetic-array processing to synthetic data from OASN simulations has been shown to produce bottom-loss profiles that reproduce more accurately the theoretically-predicted results. This section shows the result of the application of synthetic array processing to array data acquired during the the NATO Undersea Research Centre’s Boundary 2003 experiment. The data were collected at a sampling frequency of 12kHz by a drifting vertical line array with 32 elements and 0.18m spacing.

FIGURE 4 shows the bottom-loss profiles computed by beamforming the CSD matrices obtained from two different 5-minute snapshots from the Boundary 2003 data. In both cases, analysis of the beam-power plots showed strong surface noise, noise notch and no discernible interferers. The two “CBF” curves correspond to profiles obtained using the full array (32 elements), and a shorter array composed of only the first 16 elements. No ground truth is available for these data, so the estimate from the longer array is assumed to be the better one. In the case of the shorter array, the curves show evidence of a marked degradation in angular resolution, in the form of less pronounced, wider peaks and valleys, and a generally lower loss estimated towards endfire, where the beams become wider. The “CBF-SA” curves correspond to profiles obtained using only the first 16 elements of the array, but “augmenting” the CSD matrix to $32 \times 32$ elements by synthetic-array processing. The curves appear largely immune from the degradation experienced by the physical 16-element array, very closely resembling the performance of the 32-element physical array.
FIGURE 4. Bottom loss profiles computed from two 5-minute averages (data from the Boundary 2003 experiment) at 2156Hz (left) and 2250Hz (right): Conventional beamforming for 32-element (solid) and 16-element (dotted) physical array vs. conventional beamformer 16-element (dashed) synthetic array. The results from the shorter physical array show less pronounced peaks and valleys and an underestimation of the bottom loss at angles close to endfire. The synthetic 16-element array appears to correct these errors quite effectively.

CONCLUSIONS

A new derivation in frequency-wavenumber domain of a known technique for bottom-loss estimation has been developed, and synthetic-array processing presented as a technique for improving the vertical-wavenumber (or grazing-angle) resolution of the estimated bottom loss. When the estimated cross-spectral-density matrix obtained from array data is sufficiently close to Toeplitz, experimental results show that a 16-element synthetic array can improve the estimated bottom loss, achieving an angular resolution comparable to that of a 32-element array.

ACKNOWLEDGMENTS

This work has been partially supported by the Office of Naval Research.

REFERENCES