Improvement of Doppler Estimation through Repeat-Sequence Coding

J. A. Smith and R. Pinkel
Scripps Institution of Oceanography, La Jolla, CA 92093-0213

Abstract

Repeat sequence coding offers a robust method for improving the precision of velocity estimates from incoherent Doppler sounders, while retaining the simplicity of the complex covariance estimation technique. This method involves transmitting a number of repeats of a broad band "subcode." The Doppler shift is estimated from the complex autocovariance value at a time lag equal to the subcode duration. The repeat sequence code is an extension of the simple pulse trains developed in the early days of radar. By transmitting codes, rather than discrete pulses, greater average transmitted power is achieved. The performance is enhanced roughly in proportion to the usable bandwidth. Recent field work demonstrates performance improvements in both open ocean and shallow water applications.

1. Introduction

Pulse to pulse incoherent sonar was first considered for oceanographic research by Emmanuel and Mandics.[1] Development of prototype systems followed.[2,3] Performance bounds on simple Doppler systems can be calculated using methods derived for atmospheric acoustic and radar research.[4,5] Theriault[6] gives a simple approximate relation for the incoherent sonar

$$\Delta V \, \Delta R = K_1 \, c^2 \, f^{-1} \, P^{-1/2} \qquad (1)$$

where

ΔV is the rms velocity imprecision
ΔR is the effective range resolution $\approx cT/2$
T is the duration of the transmitted pulse
c is the speed of sound
f is the acoustic frequency
P is the number of independent incoherent averages used in forming the velocity estimate (e.g., the number of transmissions).
K_1 is a constant. With f in Hz, K_1 is ideally $(8\pi)^{-1}$.

Velocity precision increases with acoustic operating frequency. However, both acoustic attenuation and background ambient noise increase with increasing frequency in the sea, at frequencies above 90 kHz. These factors give rise to a tradeoff between the maximum range achievable with a given system and the expected range-velocity precision. The only way to improve on this is to increase the information content of the returning echo.

Here a simple coding scheme is investigated: repeat sequence coding. While the method is sub-optimal from a signal processing viewpoint, it is easy to implement and the performance improvement survives signal distortion by the real ocean. Repeat sequence codes are produced by taking broad-band "subcodes" and repeating them sequentially M times. A variety of codes have been developed with the objective of having minimal self-correlation, except at zero lag. We consider here codes generated by reversing the sign of the carrier frequency of the sonar at controlled intervals. The sonar output signal can be thought of as a sinusoid which is multiplied by either 1 or -1, depending on the dictates of the code. The seven-bit Barker code is shown in Fig. 1 as an example.[7] For a digitally synthesized subcode of bandwidth τ^{-1} and duration $L\tau$, the time bandwidth product is the number of "bits" in the subcode, L. The time bandwidth product of the overall transmission is ML. If the subcode is properly chosen, the autocorrelation of the overall transmission is nearly zero, except at lags $t = nL\tau$, $n = 0$ to M-1 (Fig. 1b). The number of peaks in the code spectrum is roughly proportional to the number of bits, L, in the subcode (Fig. 1c).

Repeat sequence codes are contemporary analogs of the simple "pulse train" sequences developed in the early days of radar.[5,7] A pulse train is a sequence of M sinusoidal pulses, each of length τ, transmitted at intervals of $L\tau$. The pulse train has an autocorrelation which is exactly zero except at lags $NL\tau$, N=1 to M-1. These pulse trains represent an ideal limit of the repeat sequence codes discussed here. However, the average power transmitted in a pulse train is low relative to the peak power. Repeat sequence codes approximate the pulse train ideal, but the average power transmitted is a factor of L greater than the pulse train, with little loss in velocity precision. This can extend the usable range of a practical sounder.

Repeat sequence codes also generalize the pulse-pair approach investigated by Edwards and others.[5,8] Here we consider the option of repeating the pulse or sequence more than just once. This enables the overall transmission length to be specified independent of the maximum unambiguous velocity, which is determined by the interpulse spacing, $L\tau$. This generalization is useful in adapting systems to different environmental conditions.

2. Performance Assessment

Methods for estimating a Cramer-Rao lower bound for precision are well known.[4,5,6,9] Here we take a different approach, based on the sample error inherent to covariance estimation from real data. This approach is in some ways more appealing, since it applies directly to the technique employed. The results explain part of the discrepency between observed error levels and the Cramer-Rao lower bound estimate.

a) Sample Error

Consider first a filtered, discretely sampled echo, from an uncoded transmission of duration $M\tau$, where τ is the sample interval. We conceptually divide the echo into "range cells" of length $c\tau/2$. The echo from each cell typically arises from sums over many scatterers. Regardless of the actual distribution of scatterer cross-sections, the central limit theorem implies that the complex returns from these cells are nearly normally distributed. For spatially homogeneous cells with average echo

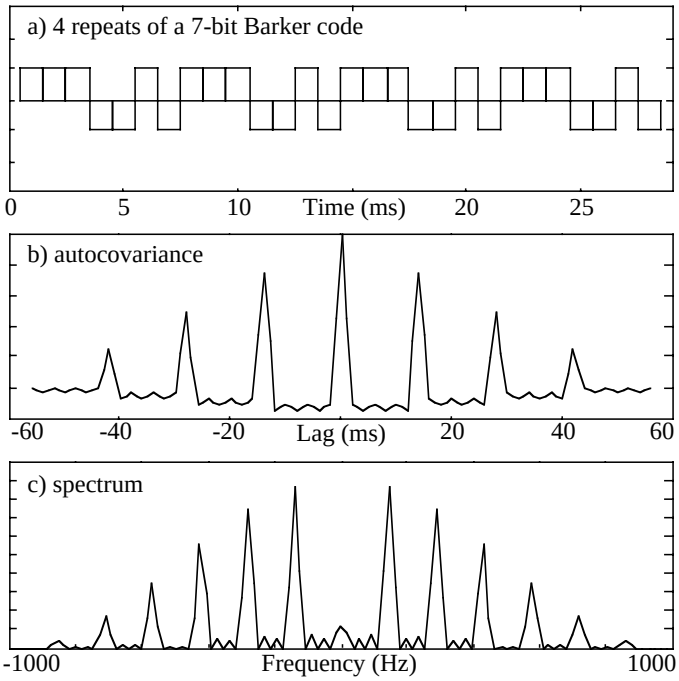

Figure 1. a) Schematic Representation of a repeat sequence code. The transmitted waveform consists of a seven bit Barker code repeated four times. b) The corresponding autocorrelation function. Values are small except at even multiples of the subcode length. c) The frequency spectrum of the code. The spectrum of an uncoded sinusoidal pulse of the same duration would consist of a single peak of width equal to the width of any individual peak in the above spectrum.

intensity "B," a lag-Q covariance sample has the expected value

$$<C_Q> = e^{iQ\omega\tau} B(M-Q), \quad (2)$$

where ω is the frequency of the signal. Samples of C_Q will most likely differ from its expected value. This difference is termed sampling error. An estimate of the sample error variance is obtained by examining

$$<|C_Q|^2> = |<C_Q>|^2 + |<C_0>|^2 \approx B^2\left((M-Q)^2 + M^2\right) \quad (3)$$

(this is a fourth-order statistic of the complex returns from each cell). The first term is just the magnitude-squared of the expected covariance (2). The second term is the sample error (here C_0 is the lag-0 covariance, so the error variance equals the square of the intensity).

Not all of this error contributes to velocity uncertainty; only "phase error" does. Phase error is primarily related to the variance perpendicular to $<C_Q>$ on the complex plane. For simplicity, consider zero signal frequency. Then the perpendicular variance is that of the imaginary part of C_Q:

$$E_Q \equiv \tfrac{1}{4}<|C_Q - C_Q^*|^2> = \tfrac{1}{2}\left(<|C_Q|^2> - \mathrm{Re}\{<C_Q C_Q>\}\right). \quad (4)$$

The first term in (4) is given by (3); the second is

$$\mathrm{Re}\{<C_Q C_Q>\} = 2|<C_Q>|^2 \approx 2B^2(M-Q)^2. \quad (5)$$

We obtain

$$E_Q = \tfrac{1}{2}\left(|<C_0>|^2 - |<C_Q>|^2\right) \approx B^2 Q(M - \tfrac{1}{2}Q). \quad (6)$$

From a detailed treatment,[10] it is found that the perpendicular error variance arises entirely from products involving "end-cells," cells 1 to Q and (M+1) to (M+Q) (say) of the total of (M+Q) cells involved in a given sample of C_Q. The signal arises only from the (M-Q) overlapped cells. This point bears on the effective range resolution obtained (discussed below). The last approximation in (6) depends strongly on the end-cells being of average magnitude.

Now consider a coding scheme in which the transmission consists of a sinusoid multiplied by a series of "bits," z_i. This sequence of bits is made up of a "subcode" of L bits, repeated M times. Thus,

$$z_{i+L} = z_i. \quad (7)$$

These might (in general) be complex, although we consider here only real codes (z_i=±1; 1 or 0 for the pulse train).

The above analysis remains essentially intact, except that a "subcode" of L samples replaces each of the uncoded samples. With coding, the covariance will usually be estimated at lag Q=L (but here we shall also consider lags Q=nL). The expected value of the covariance at lag Q becomes

$$<C_Q> = B e^{iQ\omega\tau} \sum_{m=1}^{ML-Q} z^*_{m+Q} z_m \approx \begin{cases} (ML-Q)\, B e^{iQ\omega\tau} \\ 0 \end{cases} \begin{array}{l} \text{if } Q=nL \text{ or} \\ \text{otherwise.} \end{array} \quad (8)$$

The last approximation holds for an ideal code with bits of unit magnitude; otherwise, the sum can be evaluated numerically.

b) Range averaging

Now consider the effect of averaging covariance samples in range. With L=1, the perpendicular variance of any given sample arises from the (2M-1) products of unlike cells involving the 2 end-cells. In each of the next (M-1) samples of the covariance, exactly one of these pairs again contributes to this noise. This gives rise to correlation between consecutive samples of noise, reducing the effect of averaging. Again, for zero signal frequency, the perpendicular variance involves just the imaginary part of (say) C_L. The noise covariance between covariance samples $C_L(t)$ and $C_L(t+k)$ is then

$$e_{kL} \equiv <\mathrm{Im}\{C_L(t)\}\,\mathrm{Im}\{C_L(t+k)\}>$$
$$= \tfrac{1}{2}\left(<C_k>^2 - <C_{k+L}><C_{k-L}>\right). \quad (9)$$

Note that $e_{0L} \equiv E_L \approx B^2L^2(M - 1/2)$. For k equal to any multiple of L up to (M-1)L, the correlated part of the noise is $B^2L^2/2$. Only at these multiples do the codes overlay, so the correlation is larger there than in between. For an ideal code, the correlations are zero at all other lags.

After averaging $C_L(t)$ over NL samples in range, the variance of the imaginary part of the mean becomes

$$\overline{E_L} = NL^{-1}\left[e_{0L} + \frac{2}{NL}\sum_{k=1}^{NL-1}(NL - k)\, e_{kL}\right]. \quad (10)$$

This can be applied to any sequence, using numerically generated values of e_{kL} from (8) and (9).

A simple case is an ideal code. Then $e_{kL} = B^2L^2/2$ for k=nL, 1≤n≤M-1, and zero otherwise. For N≤M, the above becomes

$$\overline{E_L}(\text{ideal}) \to \frac{e_{0L}}{NL}\left(1 + \frac{N-1}{2M-1}\right) = \frac{B^2L}{N}\left(M - 1 + \tfrac{1}{2}N\right). \quad (11)$$

The e_{kL} are zero at higher multiples of L. This makes the equivalent of (11) more complicated for longer averaging times, but the result is straightforward to calculate via (10).

A code to consider is "1" followed by L-1 zeros (a "pulse train"). This behaves like an ideal code, except that both the signal and Doppler noise are reduced by L in magnitude, relative to ideal codes with $|z|^2 \equiv 1$. The simplicity of these codes is appealing, and in some circumstances they are appropriate (e.g., some atmospheric sounding systems). However, in oceanic applications, the maximum range is limited by attenuation of the signal below the ambient noise level.[9] Thus, the slightly degraded performance of Barker codes is compensated for by the factor L increase in the signal strength relative to the ambient noise (not to be confused with the Doppler noise). This enhances the maximum range attained.

For uncoded transmissions ($z_i \equiv 1$), (10) is also straightforward. For the case M=2 and N=1, (10) simplifies to

$$\overline{E_L} = e_{0L}\left(\frac{2+L^{-2}}{3}\right). \quad (12)$$

Holding the lag, pulse duration, and averaging time constant (so e_{kL} remains the same), but increasing the sample rate (and hence L), the result approaches 2/3. This limiting result essentially matches the Cramer-Rao lower bound (see below). With more repeats, the improvement rendered by oversampling decreases. With M=N+1 and increasing N, some numerical experimentation indicates that the fraction of the error which is removed, in limit of continuous sampling, is roughly given by $(4N-1)/(3N)^2$. Some more numerical experimentation indicates that this trend applies to coded pulses as well, where M and N refer to the number of code repeats (i.e., using "fat bits" of more than one sample each). However, it should be noted that this is a poor improvement compared to using a correspondingly longer code. The preferred approach is to use the additional bandwidth to increase the information content of the echo.

An estimate of the radian angular error variance is obtained from the ratio of $\overline{E_L}$ to the squared magnitude of the expected covariance:

$$\sigma_\phi^2 \approx \frac{\overline{E_L}}{|<C_L>|^2} \approx \frac{M-1+\frac{1}{2}N}{LN(M-1)^2} \quad (13)$$

(the last is for an ideal code, using 11). This is not valid for the phase error in a single covariance sample, but becomes accurate after some amount of averaging.

The signal frequency is estimated by dividing the covariance by the lag Lτ, so (13) implies a net frequency error (for ideal codes) of

$$\sigma_\omega^2 \approx \frac{1 + T_{ave}/2T_{ovl}}{LT_{ave}T_{ovl}} \quad (14)$$

where $T_{ave} \equiv NL\tau$ is the range-averaging time, and $T_{ovl} \equiv (M-1)L\tau$ is the covariance overlap time. This becomes $3/(2LT_{ovl}^2)$ for the case $T_{ave} = T_{ovl}$, independent of the number of code repeats. Ideal coding reduces the

variance of the frequency estimate by 1/L compared to the uncoded case, holding $T_{lag}=L\tau$ constant.

For comparison, a lower bound estimate in the uncoded case (valid for $N \le M$) is given by[9]

$$\frac{1}{\sigma_\omega^2} = T_{ave}T_{ovl} - T_{ave}^2\left(\frac{T_{pulse}- T_{ave}}{2T_{pulse}- T_{ave}}\right), \qquad (15)$$

where $T_{pulse}=ML\tau$ is the duration of the whole transmitted sequence. For N=M, this is about 2/3 the value given by (14).

Equation (14) applies only to an ideal code, lag-L estimate. With Barker codes, or a longer lag, (10) must be evaluated using (8) and (9). Figure 2 shows the relative performance of "ideal codes" (pulse trains) and Barker codes, in the form $(LT_{ave}T_{ovl})\sigma_\omega^2$ versus averaging time T_{ave}. The Barker codes perform nearly as well as ideal codes. The 7-bit code falls farthest off the line, yet is only about 11% worse than ideal (i.e., it performs more like a 6-bit "ideal" code).

The RMS velocity error corresponding to (13) is

$$\Delta V = \left(\frac{c}{4\pi f_0 \tau L}\right)\sigma_\phi = \left(\frac{c}{4\pi f_0}\right)\left(\frac{1 + T_{ave}/2T_{ovl}}{LT_{ave}T_{ovl}}\right)^{1/2} \qquad (16)$$

where f_0 is the acoustic carrier frequency in Hz and c is the speed of sound.

c) Range Resolution

To cast the results into the form of (1) while allowing for different averaging times, the effective range resolution must be defined. A conservative definition is the maximum of the averaging time $T_{ave}=\tau LN$ and the covariance "overlap time" $T_{ovl}=\tau L(M-1)$. This corresponds to half amplitude points, or 6 db down in power. Let

$$\Delta R \equiv \tfrac{1}{2} c\tau L \max[\, N, (M-1)\,] \equiv \tfrac{1}{2} cT_{max}, \qquad (17)$$

where T_{max} is the larger of T_{ave} and T_{ovl}. Then the range-velocity error product estimate becomes (for an ideal code with $T_{ave} \le T_{pulse} \equiv ML\tau$):

$$\Delta R\, \Delta V = K_1 c^2 f^{-1}(PL)^{-1/2}\left(\frac{T_{max}}{T_{ovl}}\right)\left(\frac{T_{ovl}}{T_{ave}} + \frac{1}{2}\right)^{1/2}. \qquad (18)$$

Here, P is the number of echoes averaged, L is the number of bits in the subcode, K_1 is a constant (ideally $(8\pi)^{-1}$), c is the speed of sound, f is the acoustic frequency, M is the number of subcode repeats in the total transmission, T_{ave} is the range-averaging time within an individual echo, and T_{ovl} is the overlap time at one subcode lag.

The above definition of ΔR postulates that $T_{ovl} \equiv (M-1)L\tau$ affects the range resolution, rather than $T_{pulse} \equiv ML\tau$. Physically, the signal in the individual covariance samples comes only from the overlapped range cells (M-1 subcodes long); the rest contributes to noise. This being so, the most averaging which can be done without degrading the range resolution corresponds to setting $T_{ave} = T_{ovl}$. Thus, this setting yields the optimal range-velocity error product, as shown by (18). The factors in braces together then become 3/2. One is led to use the longest possible code, with the fewest possible repeats. This results in an expected triangular weighting of the signal in range.

Often, T_{ave} has been set equal to T_{pulse} (M=N). This may be because little distinction has been made between T_{ovl} and T_{pulse} in the past. In this case, (18) permits minimal error at some number of code repeats. This arises because, with T_{pulse} fixed, longer codes imply a smaller overlap time T_{ovl}. Thus, as the code length is increased, the potential gain in the error product is offset by a mismatch in range resolution between T_{ovl} and T_{ave}. With the introduction of long subcodes, the distinction between T_{ovl} and T_{pulse} can no longer be ignored.

These two settings are illustrated in Table 1, using a sample interval τ of 1 ms (the covariance lag is L milliseconds). The error values are for single transmissions with various subcode lengths L and number of repeats M. The averaging time NL is held roughly constant. Modeled Barker code performance is used. With N=M, a minimum is indeed seen, at 4 repeats of a 7 bit code. Again, this arises primarily because of the decrease in T_{ovl} with code-length in this case. In the other case, where the transmission is lengthened to maintain a constant range resolution (i.e. M=N+1), the variances decrease almost like L^{-1}.

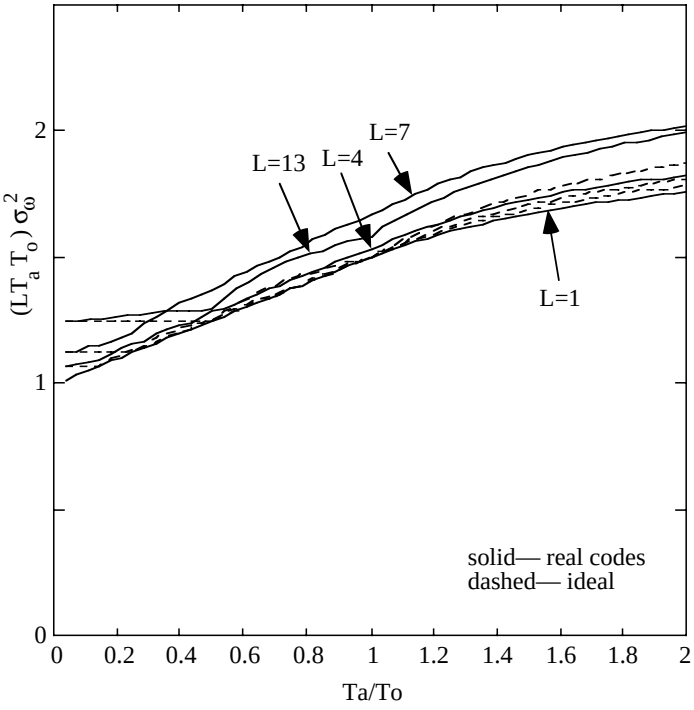

Figure 2. Performance of Barker Codes vs. equivalent ideal (pulse train) codes: σ_ω^2 normalized by $LT_{ave}T_{ovl}$. Sequences are 2 13-bit, 4 7-bit, and 7 4-bit subcodes. In practice T_{ave} should be set equal to T_{ovl}.

d) Optimal Number of Subcode Repeats

One wishes to know the settings that minimize the net error. Equation 18 indicates that optimal performance is obtained with $T_{ave} = T_{ovl}$ (i.e., M=N+1), and with both corresponding to the desired range resolution. With this constraint, (14) implies the use of the longest code possible; e.g., a subcode equal in length to the desired range resolution, with two subcodes transmitted (N=1 and M=2). The trend in Table 1 supports this conclusion. For a given system bandwidth, the improvement with code length comes largely from using a longer lag, since the squared lag enters in the denominator of the frequency error estimate (14). Thus, one could use a pair of pulses separated by some long lag to achieve similar or better results (this approach is often used in atmospheric radar systems). However, N=1 is not necessarily the best choice.

Two factors work against the use of very long lags: (1) The many scatterers in each cell move relative to each other, introducing a "natural decorrelation time" of the backscatter cells b_n. For a given level of RMS velocity fluctuation, the decorrelation time is inversely proportional to the acoustic frequency used. This time must be large compared to the lag to assure proper performance of the code. In the oceans, this natural

Table 1: Noise variance with τ = 1 ms.

code L	averaging N	M=N σ_V^2 (cm/s)^2	M=N+1 σ_V^2 (cm/s)^2
1	28	49.53	47.18
4	7	14.89	12.06
7	4	11.09	7.49
13	2	11.86	4.45

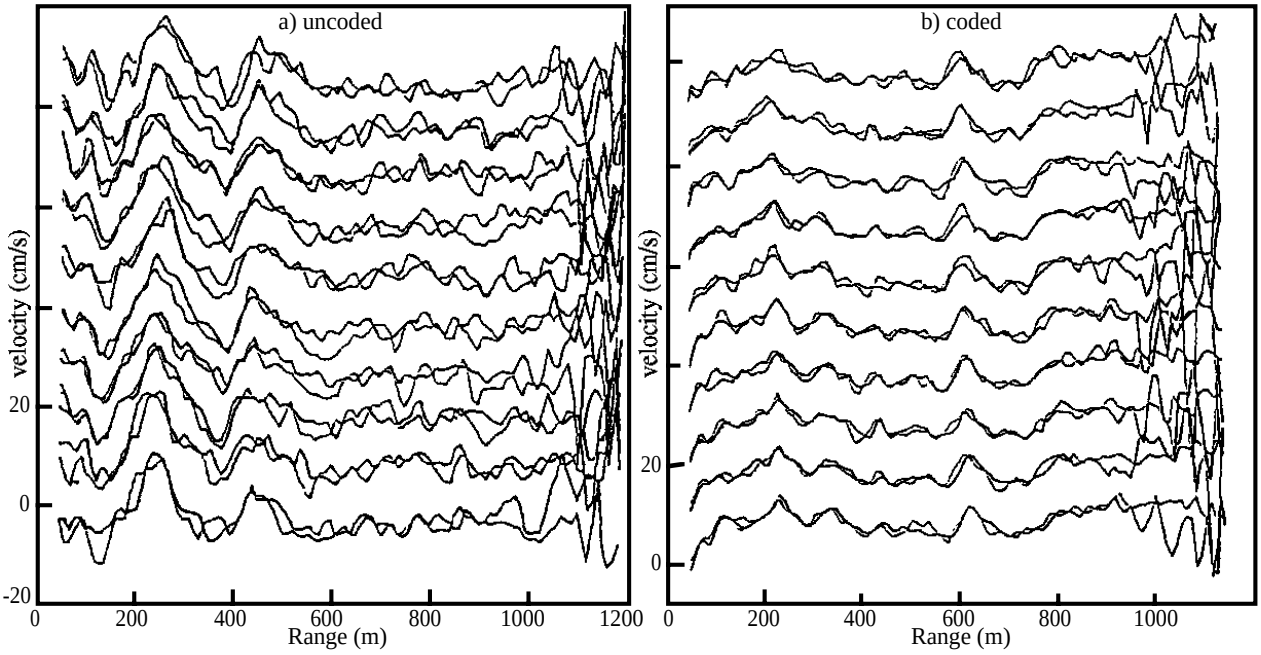

Figure 3. A comparison of (a) uncoded and (b) coded pulse performance. Velocity profiles estimated from the 71.5 kHz and 75 kHz systems are overplotted. The disparity between profiles is greatly reduced through coding.

decorrelation time is often a limiting factor. (2) The lag used should be small enough that the phase of the covariance is "comfortably" less than $\pm\pi$, so that the direction of the frequency shift can be unambiguously determined. This depends on the range of velocities expected and the acoustic carrier frequency employed. For example, if the velocities are always smaller than ±10 cm/sec, and the sonar is operated at 75kHz, a maximum of 1/4 cycle signal phase shift would occur at a lag of about 25 ms. For surface wave measurements, on the other hand, velocities of 2 m/s are possible, reducing the allowable lag to about 1 ms. For oceanic applications, the maximum usable lag, set by one of the above constraints, is usually smaller than the time corresponding to a range bin (the latter is set by balancing $\Delta R \Delta V$ against the expected spectrum of the geophysical signal of interest[9]). In this case, the "overlap region" must be spanned by several repeats of the code; i.e., values of N somewhat larger than 1 will normally be used (as in Table 1).

There is another caveat concerning "small N" schemes: a strong scatterer can alter the expected balance between error and signal. For example, with N=1 (only one subcode overlap), a bright "end cell" can make noise comparable to the signal level in both neighboring range bins, in addition to dominating the signal in its own bin (bright spots, for example fish, frequently move with a velocity not representative of the overall medium). The former problem is reduced for larger values of N: the scatterer must be more than about N times brighter than the N "signal cells" to dominate.

3. Field Trials

We have used coded transmissions in a variety of field trials and scientific experiments. Surface waves, Langmuir circulation, internal waves, and other opper-ocean phenomena have been measured. The primary scatterers have included sub-surface bubbles, plankton, and oceanic turbulence.

To document code performance, we discuss here two sets of field trials: one in the open ocean in May 1988, and the other in shallow water in August 1990. The open ocean (volume reverberation) tests are presented first, then the shallow water (surface reverberation) results. Data from both trials are compared with theory.

In May 1988, tests of repeat sequence codes were conducted aboard the Research Platform FLIP off the coast of San Diego, California. Two sonars with parallel beams were used for the tests. The same volume of the ocean was ensonified simultaneously at two different carrier frequencies. Independent estimates of the velocity profile were calculated. Differences between the two estimates result from both hardware differences and the fundamental range-velocity limit of interest here. The mean-square velocity difference between the two systems is the sum of the velocity error variances of each sonar (fig. 3).

The sonars were operated at carrier frequencies of 71.5 and 75 kHz. They transmitted at two second intervals. Useful data were obtained to ranges in excess of 1100m. The data were processed using the complex covariance technique, modified for pulse to pulse incoherent operation.[11,12] Received echo signals were complex demodulated and digitized at 2kHz per channel. At the time of the tests, the useful bandwidth of the sonar system was of order 1.2kHz.

A number of data runs of about 10 min. (300 ping) duration were conducted. Velocity profiles were calculated every minute (30 pings), and also averaged over each ten minute run, for each sonar. Differences in the 10 minute average profiles provide, primarily, a range dependent index of system problems (Fig. 4b). Sonar to sonar differences in the one minute profiles provide a range dependent (signal to noise dependent) picture of velocity estimator precision. The mean square velocity difference is presented in Fig. 4c, using a representative pulse. It is seen that estimator

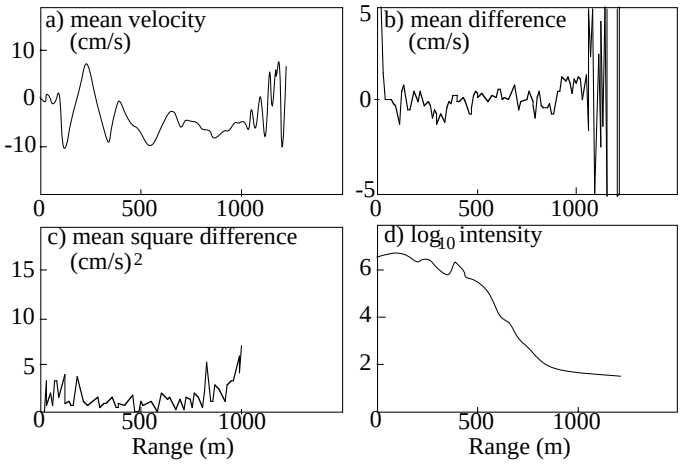

Figure 4. (a) The ten minute mean velocity profile, as estimated by averaging results from both the 71.5 and 75kHz sonars. (b) The difference in the ten minute means of individual sonars. (c) The variance of the difference velocity as estimated from 10 one minute profile pairs. (d) The ten minute average acoustic intensity. Variations in the first 200m are obscured by clipping in the A/D converters of the data recording computer.

error decreases from the nearest ranges out to approximately 200m. In this region, the return echo was clipped, limited by the maximum range of the A/D converters (fig. 4d). Both coded and uncoded pulses show this range dependence. In both cases, minimum error (maximum precision) is achieved over roughly 200-800m. The performance of the coded pulses begins to degrade at a shorter range than an uncoded pulse. At greater ranges, error gradually increases as the signal decays into the background noise. The receive noise level is higher in the coded system due to the broader system frequency bandwidth necessary to receive the coded transmission (the uncoded signal was filtered to 200Hz bandwidth, while the coded signals used the full 1.2kHz bandwidth available).

Tests were run with uncoded sinusoidal pulses, and with subcodes consisting of four, seven and thirteen bits (fig. 5). Bit lengths of half, one, and two ms were tried. The lengths of the transmitted sequences were set near 28 ms (T_{pulse}). The autocovariance was averaged over $T_{ave}=T_{pulse}$ (rather than T_{ovl}) prior to calculating velocities. All of the codes performed significantly better than the uncoded pulse (fig. 5). Performance differences between codes were minor, as predicted by Table 1.

As a test of the robustness of repeat sequence coding, a 7-bit code was tried with bit lengths of 0.5ms. This has an inherent bandwidth requirement of 2kHz, compared to the actual system bandwidth of about 1.2kHz (as configured, the system had roughly an "RC-filter" response, 6 dB down at 1.2kHz). Code performance was degraded relative to theory, but not catastrophically. The 7-bit 1/2ms code performed roughly as well as an ideal 5-bit code. Apparently, when the usable bandwidth is exceeded, the performance is still improved in proportion to that bandwidth available.

To test code performance with a different system, and in the harsh environment inherent to surface reverberations, a second series of tests were conducted in August 1990. In this experiment, a 195kHz fan-beam sonar was deployed in about 12m water depth, near the end of the Scripps Pier. The beam was oriented offshore, almost directly into both the incoming swell and locally generated wind waves. The operating bandwidth was 5kHz (bit lengths of 0.2 ms). Uncoded (1-bit), 4-bit, and 7-bit coded transmissions were tried. The averaging and overlap times were held nearly constant, at $T_{ave}=T_{ovl}\approx 4$ ms (20 bits for the uncoded and 4-bit cases, 21 bits for the 7-bit case). With $T_{ave}=T_{ovl}$ rather than T_{pulse}, systematic variation with subcode length is more evident than in the May 1988 tests (as is indicated by Table 1; see fig. 6).

As with the May 1988 data, a range interval was selected over which the sonar performance appears nearly optimal (roughly 40 to 200m). The Doppler noise estimate was extracted by assuming the "signal" part of the return velocities to be due entirely to waves progressing toward the shore: the magnitudes of frequency cross-spectra between estimates at different ranges were subtracted from the geometric means of the frequency spectra at those pairs of ranges (see figure 6). Results using a variety of cross-spectral "separations" were then extrapolated to zero offset in range (the extrapolation had minor effect). To minimize the effects of a non-stationary environment, two codes were run at a time, with interlaced "pings" of each type (this also proved unnecessary). Some physical variance passes through this "filter," since not all the variance comes from unidirectional waves. Thus, the noise estimates are slightly higher than for the "shared volume" tests of May 1988. With M=N+1, the error decreases monotonically with code length (figure 6).

A summary of both sets of tests is given in figure 7, in the form of normalized error vs. the inverse number of bits, 1/L. The error is normalized by converting the appropriate mean-square velocity-error back to a radian frequency error, and multiplying by $PT_{ave}T_{ovl}/(1+T_{ave}/2T_{ovl})$ (in the May 1988 tests, P=30; in the August 1990 experiment, P=1). This normalization should remove all but the dependence on L, the code length. The corresponding theoretical predictions are also shown for real (Barker) and ideal (pulse train) codes. Note that the 4-bit code yields estimates closest to theory in both sets of data, while the 7-bit codes seem to fare worse. Although this pattern is reflected in the "real vs. ideal" code comparison (fig. 2), it appears to be exaggerated in the actual data. With this normalization, which compensates for the variations of T_{ovl} in the May 1988 data, we nevertheless see a monotonic increase in performance with code length.

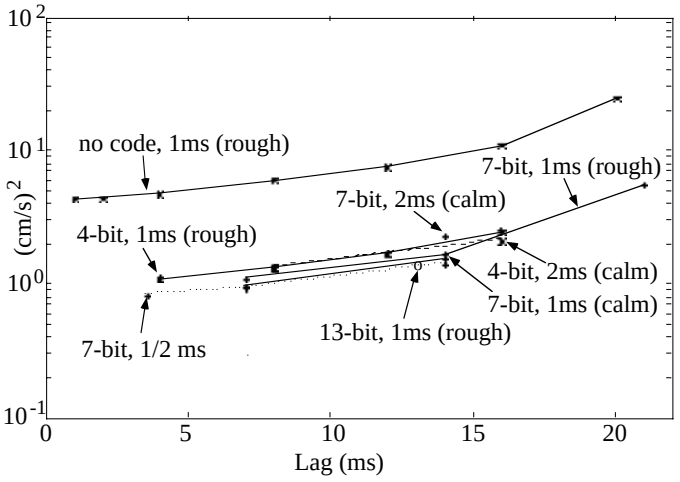

Figure 5. Mean-square velocity difference from the May 1988 data, for various codes and lags. Run parameters are noted in the figure; for example, "7-bit, 1ms" refers to a run with 1ms bit-length, repeating a 7-bit Barker code enough times to approximate 28ms total (4 repeats).

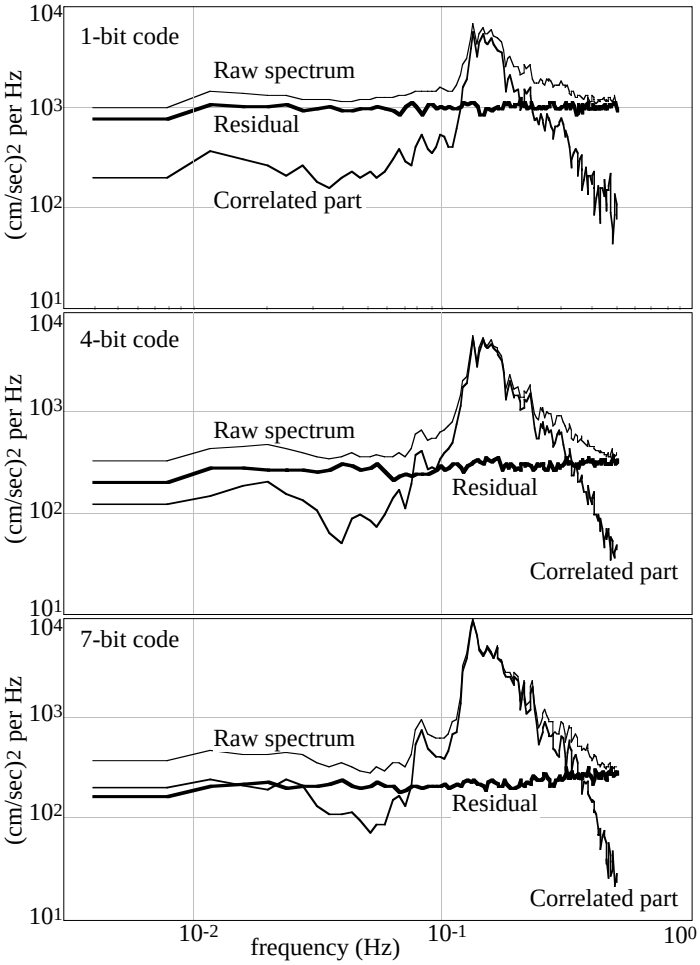

Figure 6. Illustration of the "Cross-spectral" method for obtaining the Doppler noise from the August 1990 data (from the SIO pier, 12m water depth). The magnitude of the averaged velocity cross spectra at a fixed separation in range is subtracted from the mean spectrum (shown are results using a 3m range interval). The surface wave signal is reduced, leaving a spectrum which is nearly flat in frequency. (The average spectral densities, neglecting the two lowest frequency estimates, are: 1-bit 1023; 4-bit 289; and 7-bit 249 $(cm/sec)^2$ per Hz. The corresponding RMS velocity errors are 22.6, 12.0, and 11.1 cm/sec, respectively).

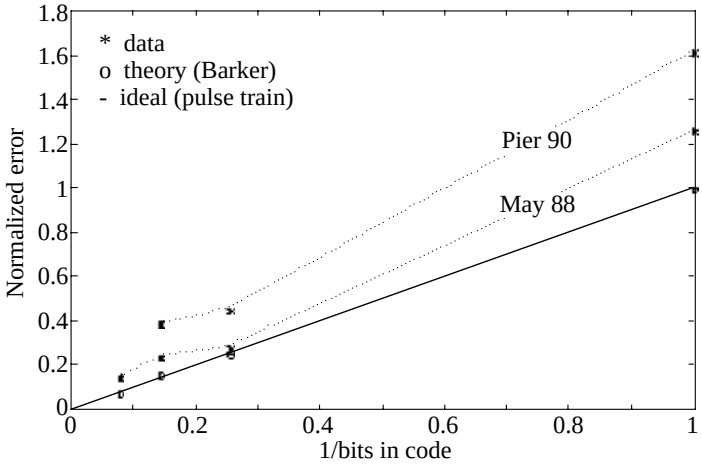

Figure 7. Normalized error vs. 1/code length. The radian frequency errors are normalized by $PT_{ave}T_{ovl}/(1+T_{ave}/2T_{ovl})$ (see text). The data are effectively "collapsed," although taken under widely different conditions, with different systems, and using different acoustic frequencies.

Do the differences between theory and data arise from an unmodelled but systematic variation with T_{lag}, as might result from the decorrelation time of the medium, or are they linked to the codes themselves? A plot of the ratio of measured to theoretical performance vs. lag is presented in Figure 8, using just May 1988 data. The corresponding predictions of the theory for the actual codes were used. Figure 8 shows a slight trend in performance vs. lag for each code. However, this is small compared to the variation between codes: the systematic trend with lag accounts for only about 1/3 of the total variation. For example, the actual performance of the 4-bit code at a lag of two subcodes (lag 8) is nearly as good as the 7-bit code at one subcode (lag 7), even though the theory predicts worse performance. It appears that the discrepancies are linked to the code used, and not just to the lag.

Is there something inherent to the 4-bit code which makes it perform close to theory, while the 7-bit code is always farther from theory? Although the "4-bit, 1ms" data are consistently closest to the theory, the "4-bit, 2ms" data are quite far off. The 2ms data were gathered during a "calm" period, and figure 8 shows that the "4-bit, 2ms" case is even farther from the theory than the corresponding "7-bit, 1ms" case (the "calm-weather" case, or lower of the two "7-bit 12ms" curves). This shows that there is nothing special about the 4-bit code.

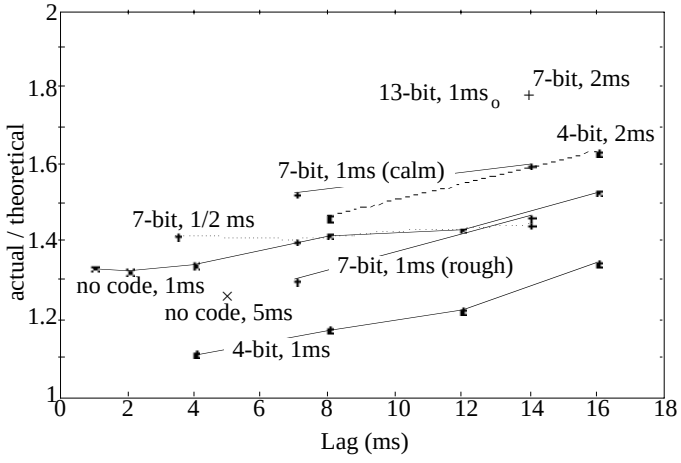

Figure 8. Ratios of measured performances to predictions from the theory. Only May 1988 data are plotted, showing a variety of bandwidths and code lengths. The "x" plotted at lag 5, 1.25 ratio corresponds to the uncoded results, interpolated to lag-5 and compared to the theory for data sampled every 5 ms (see text). An overall trend is seen with lag, but a larger variation occurs from code to code. The considerable scatter independent of lag indicates that another parameter is important.

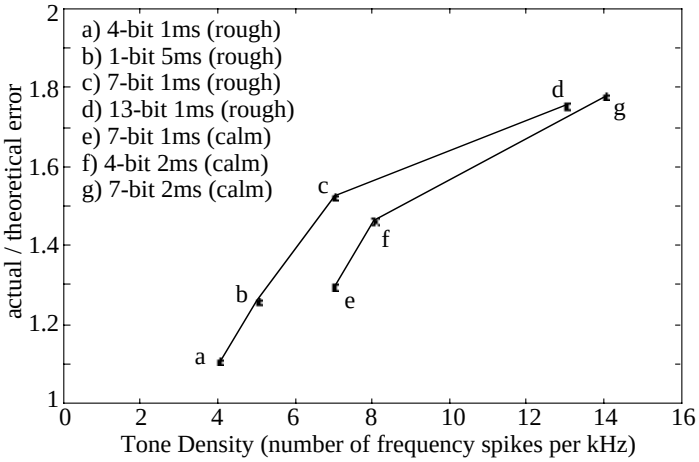

Figure 9. Performance ratios vs. "Tone density," or the number of peaks in the code's spectrum per kHz of available bandwidth. The two curves are for the "rough" conditions (upper curve) and "calm" (lower). Note that the uncoded case is plotted according to its (filtered) bandwidth of 200Hz, using the interpolated "lag 5" value. A simple pattern emerges which is apparently dependent on conditions. A simple explanation has not yet emerged.

Also shown in figure 8 is an "artificial point" constructed for the uncoded case. Before proceeding, this point must be explained. The theory developed above assumes independent samples. The uncoded data were filtered with 200Hz total bandwidth; thus, independent samples occur only every 5 ms. For comparison with theory, a value for lag-5 was interpolated from the lag-4 and lag-8 data points. Even though the covariances were averaged at 1 ms increments, there is no more information than would have (in theory) been obtained with 5 ms samples. The interpolated "lag 5" value is compared to the theory for 5ms sample rate. The ratio (about 1.25) is plotted at "lag 5" in figure 8. This point is helpful in the following.

A pattern of performance is found by separating the data into "calm" and "rough" conditions, and plotting the smallest lag from each group against the "tone density" (figure 9). "Tone density" is defined as the number of peaks in the code frequency-spectrum divided by the available bandwidth. For the uncoded case, the total bandwidth is 200Hz (yielding a tone-density of 5 "tones per kHz"). For the 2ms bit-length data, the tones are effectively compressed into the middle half of the total bandwidth; i.e., in this case the tone-density corresponds to one over the spacing between the peaks in the code spectrum. Figure 9 shows that (a) the performance is linked to the separation of peaks in frequency, and (b) there is a "cutoff" spacing which varies with conditions. Above this cutoff, the performance is degraded relative to theory; however, there is no clear indication that the performance continues to degrade without limit as the tone density increases. (Note that the distinction between lag and tone density could only be made by examining results from lags of more than one subcode.)

While the tone density nicely divides the data, the reason for its importance remains unexplained. It should be remarked that the longer codes always perform better than the shorter ones, taking range resolution into account; it is only that the improvement is not as great as the "pure" theory implies (figs 6, 7).

4. Discussion

An objective here is to establish guidelines for the design and operation of a Doppler acoustic sounder that will help minimize range-velocity error, within limits imposed by hardware and the environment. We envision guidelines such as the following:

(1) An operating frequency is selected, based on maximum range requirements.

(2) The maximum usable lag is estimated or measured for the environment of interest, as determined by either the decorrelation time of the scatterers or the "ambiguity time" associated with the maximum velocity expected (both for the given frequency).

(3) The usable bandwidth of the sonar system is specified. This may be limited either by available hardware, or by constraints such as the frequency-dependence of acoustic attenuation with range, or the difficulties of beam-forming over a broad range of frequencies. This sets the time interval τ between independent samples.

(4) From 2 and 3, an initial subcode length L is specified, and from equation 18, and initial estimate of the range-velocity error product is obtained.

(5) Comparing this with the expected wavenumber spectrum of velocity for the geophysical signal to be measured, the desired range resolution can be specified.[9] If this resolution is smaller than the usable lag specified, then the subcode is shortened and steps 4 and 5 are repeated. More commonly, the desired resolution is several times longer than the lag, and so some number of overlapped code repeats (N) is specified.

(6) The transmitted pulse is made one code-length longer (M=N+1), and the one-code-lag covariance is averaged over a time equal to the "overlap time" (NL samples). A key point is that the range resolution ($T_{ovl} = T_{ave}$) is fixed, and the transmitted sequence is adjusted to minimize the error (so M=N+1).

The above guideline is simple and appealing. However, the codes perform slightly worse than the theoretical predictions, in initial tests. More significant, perhaps, is the suggestion arising from the data that the performance relative to theory might degrade systematically with code length. Before relying on the modeled performance predictions it is wise to assess the causes of the observed degradation.

a) Uncoded Pulse Performance

After one minute of pulse to pulse averaging, the uncoded velocity precision measured was about twice the theoretical (Cramer-Rao) lower bound (15). Discrepancies of this sort have been previously reported.[13] It has been surmised that the difference might be due to the form of processing employed to extract the Doppler information. The present error analysis is not a lower bound, but is appropriate to complex covariance processing. Most of the discrepancy appears to be explained (a factor of 1.5), but some remains.

One possible source of degradation is the addition of broadband "noise" to the signal. Noise is usually considered to be a problem only at great range. Sonar signal levels typically vary over 60-100 dB before falling to the noise level. Hence, at mid ranges, signal levels should be 20-50 dB above the noise. The effect of broadband noise on the return should be negligible over most of the usable range.

The above assumes that the relevant noise is that which exists in the absence of active transmitting (e.g., electronic noise or ambient acoustic noise). Another possible source is the relative motion of the scatters within the overlap volume. This broadens the frequency bandwidth of the return echo at all ranges. Unless the peaks get so broad as to be truncated by the frequency pass-band filtering, however, this should have little net effect on the Doppler estimate for the uncoded case.

There may also be "distortion noise," associated with the imperfect transmission or reception of the acoustic pulse. This noise might be "broad band" yet have an intensity which is some percentage of the signal power. It would degrade the precision of the velocity estimate over the useful range of the sonar, yet vanish along with the signal at great range. Rough calculations indicate that a distortion noise to signal ratio of -25 dB would be sufficient to cause the observed 1 $(cm/s)^2$ of additional error variance in the velocity estimates.[10]

b) Coded Pulse Performance

The observations indicate that subcodes of longer duration are degraded slightly relative to the theory, while subcodes of shorter duration are degraded less, even when evaluated at lags comparable to that needed for the longer duration subcode. The relevant parameter appears to be "tone density," the spacing of the peaks in the frequency-spectrum of the code used. A transition appears in the data, at a tone density which is dependent on the environmental conditions. Although this remains to be understood, we surmise that it may have to do with platform motion and/or the motion of the individual scatterers in the turbulent velocity field. In spite of this degradation, the measurements indicate a steady improvement of performance with code length, provided that T_{ave} is set to equal T_{ovl}.

As increased bandwidth is contemplated for new systems, one must contend with non-negligible variations in acoustic properties of the medium with frequency. These include, for example, differential attenuation of the frequency components of the code, and additional difficulties in beamforming. Would the "real world" performance keep improving with bandwidth, or is there another aspect of the sonar system or environment which will eventually limit the accuracy regardless of bandwidth? We feel this question can be answered plausibly only through actual trials. However, the results of the half-millisecond tests (which exceeded the available electronic bandwidth) are encouraging: performance was degraded relative to the full code, but it was nontheless enhanced, roughly in proportion to the actual available bandwidth. Thus, even if the ocean imposes an upper bound on useful bandwidth, smaller than that of the sonar system, there is still reason to expect performance improvement in proportion to the actual usable bandwidth. The only real penalty incurred by trying too wide a bandwidth is reduced maximum range (plus, perhaps, other compromises in system design).

c) Operational Concerns

In the present study, range resolution ($NL\tau$) and acquisition time (P) were held fixed, and the enhanced performance of the codes reduced velocity variance. Alternatively, one could improve range resolution or decrease acquisition time, keeping ΔV constant.

Not explicit in the above is another option. Coded pulses at a lower frequency can provide error-product levels comparable to a higher frequency uncoded device. Given that both acoustic attenuation and ambient noise (at frequencies above 100kHz) decrease with decreasing frequency, significantly greater maximum ranges can be attained by decreasing the operating frequency of the sonar. For example, the 75kHz sonars reported here achieved performance comparable to a 150 kHz system operating at the same pulse repetition rate. Yet the maximum useful range is more than twice that typically reported for 150 kHz systems. (Of course, one pays a price for longer range. To achieve longer ranges, the pulse repetition frequency must be reduced. Fewer independent echoes, P, are available to average into a fixed duration velocity estimate. This can partially offset the gain from the coding.) In many situations, external limits are imposed on the range and velocity resolutions attainable. For example, variations in ship motion limit both for ship mounted sonars. Given such limits, it's attractive to use the improved capability provided by codes to achieve greater maximum range, by shifting to lower frequency sonars than are presently used. With efforts toward adapting larger transducers to shipboard mounts, and in adaptively cancelling ship machinery noise, significant gains can be anticipated.

Measurements sometimes suffer from platform motions. These introduce range-independent velocities which, at surface wave frequencies, may be much larger than (say) the deep ocean velocities of interest. If the code length is selected on the basis of the deep ocean environment, the instantaneous velocities may be far beyond the unambiguous limit. Although time averages should approach the correct (small) phases, these motions reduce the magnitude of the averaged covariances, and hence increase the net error. This is a non-linear effect, which is hard to evaluate. There is a way to reduce this effect, however. First, a mean velocity is calculated over a selected (large but reliable) range interval for each ping. Then the complex covariances for the ping are rotated on the complex plane, so that the phase corresponding to the range mean velocity becomes zero. The range mean is saved and time averaged along with the rest of the profile. Each of the estimates along the profile is then subject only to the "local" velocity magnitude, in line with the selected code length, and the time averages suffer much less from such "velocity aliasing." Since the range mean is an average over many more samples, the effect of the large velocities on the final range-time mean velocity estimate is also minimized. (It should be noted that, in this scenario, the overall mean estimate can retain more error than the averaged deviations from the mean; this is somewhat unusual). We have implemented this technique (but not in the

tests discussed here), and found that is does seem to reduce noise levels of one-minute averages.

In some circumstances (e.g., measurement of surface waves), ambiguity of the velocity estimate is the limiting factor. The covariance lag must be short enough to unambiguously determine the sign of the velocity over the full range of velocities encountered. Orbital velocities of surface waves can be as large as 2 m/s. At 75kHz, for example, this corresponds to a Doppler shift of 200Hz, or one-quarter cycle in only 1.25 ms. Even risking more than a quarter cycle phase shift, the subcode duration (T_{lag}) is restricted to less than about 1.5 ms. A "two-code" system might help overcome this difficulty. This would consist of a combination of a "short code" (set by the "ambiguity time") and a "long code," where the latter is set by the natural decorrelation time, rather than the arctangent ambiguity time. The shorter lag provides an unambiguous but noisier velocity estimate, and the longer lag can then be used to refine the estimate. We note that (a) the noise of the noisier estimate must be smaller than the ambiguity time of the longer code, and (b) if both codes are transmitted in a single (double-length) sequence, the presence of one code degrades the performance of the other, making it difficult to benefit from such a scheme. If the time between pings is short compared to the timescale of the field being measured, alternate pings could have alternate code lengths. Unfortunately, this approach appears to be of limited value for surface wave measurements, which are the ones most in need of it.

5. Conclusions

Repeat sequence coding offers substantial resolution improvements over uncoded transmissions. It retains both simplicity of implementation and robustness with respect to environmental distortion. A variation of the standard complex covariance technique is used for estimating frequency: covariance is estimated at a lag corresponding to one subcode length $L\tau$, rather from adjacent data samples. In practice, the improvement appears to be a robust feature of the technique, in an environment which has so far thwarted more ambitious (and complex) coding schemes. This technique may also serve in atmospheric acoustic sounding systems (SODAR).

Several conclusions arise from this study:

(1) The theoretical error level for complex-covariance processing is about 1.5 times the Cramer-Rao lower bound. Practically achieved errors varied between about 1.6 to 2.6 times this lower bound. Thus, part of the discrepancy between observations and the lower bound estimates has been explained. The remaining portion is quite variable.

(2) It is the "overlap time" $T_{ovl} \equiv (M-1)L\tau$ which enters in determining range resolution, not T_{pulse}. Thus, it is appropriate to set the overlap and averaging time equal, $T_{ave} = T_{ovl}$ or $M = N+1$, in order to minimize velocity error for a given range resolution.

(3) Repeat-sequence coding improves the range-velocity error product by a factor roughly proportional to $L^{-1/2}$, where L is the number of bits in the subcode used. This means that even short codes help significantly. Many existing systems can be easily modified to take advantage of 4 or 7 bit codes. In general, both the bandwidth and time-lag used should be maximized (within constraints applied by the ocean, electronics, or resources), to obtain the largest value of L possible.

(4) Sonars designed with very wide inherent bandwidth may encounter bandwidth limitations imposed by the ocean (for example, by the differential attenuation of sound at different frequencies). A test run in which the system bandwidth was deliberately exceeded, however, indicates that the coding still enhances performance, in proportion to the actual usable bandwidth.

(5) Low-pass filtering the signal before forming covariances (as in the uncoded May 1988 test) increases the error product of the resulting estimates slightly, relative to the "oversampled" case. This is offset by reducing the sample rate required and increasing the maximum range attainable. (Of course, such available bandwidth is put to better use with repeat-sequence coding.)

The proven performance enhancement has significant implications for the advancement of knowledge of the sea. For example, it may permit probing of the internal wave field throughout the 4 km water column, with the same resolution as presently used over the top quarter (by moving to a lower acoustic frequency). Alternatively, one might observe motions throughout the upper thermocline, resolving the scales at which the transition to turbulence occurs.

References

[1] Emmanuel, C.B., and P.A. Mandics. A feasibility study for the remote measurement of underwater currents using acoustic Doppler techniques. NOAA Tech. rep. ERL 278-WPL25, august 1973.

[2] Pinkel, R. and F. N. Spiess. "Space-time Measurement of Oceanic Motions from a Range-Gated Doppler Sonar." J. Acoustical Soc. Am., vol. 59, Supl. 1, 1976

[3] Scherer, W.D., and K.A. Sage, and D.E. Pryor. "An intercomparison of an acoustic remote current sensor and Aanderaa current meters in an estuary." Presented at 98th meeting of the Acoustical Society of America, Salt Lake City, Utah, November 1979.

[4] Miller, K. S. and M. M. Rochwarger. "A Covariance Approach to Spectral Moment Estimation." IEEE Transactions on Information Theory, vol. IT-18, no. 5, pp. 588-596, 1972.

[5] Doviac, R.J. and D.S. Zrnic. Doppler Radar and Weather Observations. Academic Press, Orlando Florida, 1984, pp.103-107.

[6] Theriault, K. B. "Incoherent Multibeam Doppler current Profiler Performance. Part I: Estimate Variance." J. Ocean Eng., vol. 0E-11(1), pp. 7-15, 1986.

[7] Rihaczek, A. W., Principles of High Resolution Radar. McGraw-Hill, New York, 1969, p. 328.

[8] Edwards, J.A. "Remote measurement of water currents using correlation sonar." Presented at 98th meeting of the Acoustical Society of America, Salt Lake City, Utah, November 1979.

[9] Smith, J. A. "Doppler Sonar and Surface Waves: Range and Resolution." J. Atm and Oceanic Tech., vol. 6(4), pp. 680-696, 1989.

[10] Pinkel, R., and J. A. Smith. "Repeat Sequence Codes for Improved Performance of Doppler Sounders." J. Atm. and Oceanic Tech., (submitted, 1991).

[11] Rummler, W. D. Introduction of a New Estimator for Velocity Spectral Parameters. Tec. Memo MM-68-4141-5, Bell Telephone Labs, no. 24, 1968.

[12] Pinkel, R. "On the use of Doppler Sonar for Internal Wave Measurements." Deep Sea Research, vol. 28A, pp. 269-289, 1981.

[13] Hanson, D. S. "Oceanic Incoherent Doppler Sonar Spectral Analysis by Conventional and Finite Parameter Modeling Methods." IEEE Journal of Oceanic Engineering, vol. 0E-11, No. 1, pp. 26-40, 1986.