Sonar pulse wave form optimization in cluttered environments

Peter B. Weichman

BAE Systems, Advanced Information Technologies, 6 New England Executive Place, Burlington, Massachusetts 01803, USA

(Received 18 February 2006; revised manuscript received 15 May 2006; published 26 September 2006)

A theory of active sonar (or radar) pulse wave form design, for optimal target detection in cluttered environments, is presented. The received target signal is maximized via a cost function $\mathcal{L}$ that incorporates both the signal-to-noise ratio and a generalization of the Heisenberg uncertainty principle, which is used to balance bandwidth (or range resolution) against signal gain. The optimal pulse wave form is the ground state solution to a one-dimensional Schrödinger-type equation in frequency space, with an effective potential energy that tends to concentrate pulse energy in frequency bands where the target reflectivity dominates the clutter reflectivity.

DOI: 10.1103/PhysRevE.74.036619 PACS number(s): 43.38.+n, 43.50.+y, 43.60.+d, 84.40.Xb

I. INTRODUCTION

The issue of optimal configuration and application of sonar, radar, and other wave-based remote sensing tools, is receiving renewed attention, with significant research programs designed to address, for obvious tactical reasons, the problem of target detection and identification in cluttered (shallow water, urban, forest, etc.) environments. Hardware advances must go hand in hand with modeling and signal processing advances for optimal use of these technologies. Available degrees of freedom include number of transmitters and receivers; their placement and/or flight path; transmitter aperture and focus; and transmitted pulse repetition rate and wave form design. Each plays an important role in defining the ultimate accuracy of the measurement. Much work, for example, has gone into improving spatial resolution of synthetic aperture radar images [1]; detection, via Doppler shift, of moving targets in a stationary background [2]; and imaging using large sonar arrays [3]. Here I consider a problem that has received far less attention, namely, optimal tailoring of the transmitted pulse wave form to the scattering characteristics of the scene. In combination with the other degrees of freedom, the derived solution may be used to further enhance the quality of measurement.

The focus here is on sonar problems, but the methodology is equally applicable to radar and other remote sensing tools [4]. Target detection and location via active sonar in littoral and other high-noise acoustic environments is complicated by high-clutter returns that may swamp the target signal of interest. To mitigate this, pulses with large bandwidth $\Delta\omega$ are used, thereby increasing range resolution $\Delta r \sim 2\pi c/\Delta\omega$ (c is the sound speed) and target visibility, by reducing the number of clutter returns in a given range-bearing cell.

High bandwidth also provides the opportunity for tailoring the frequency spectrum of the pulse to the target of interest [5–9]: for example, target and clutter returns generally have different frequency dependence, and the signal-to-noise ratio (SNR) can be increased by focusing greater pulse energy in bands with enhanced target reflectivity. Here a more general continuum formulation of this optimization problem is presented that exhibits clearly the interplay between the opposing requirements of high SNR and high range resolution. An objective functional $\mathcal{L}[\hat{P}]$ is constructed containing terms that control the relative importance of these effects, and allows one to efficiently determine optimal forms for the pulse spectrum $\hat{P}$ under a variety of conditions.

For moving targets one may be interested in both its range and its relative velocity. The radial component of the latter may be inferred from the Doppler shift ω_D of the reflected pulse. Since it is essentially a wave front counting exercise, accurate measurement of ω_D is limited by the total measurement time, not by the bandwidth, and is equally well accomplished with a monotone, continuous wave source [10]. In particular, the optimal wave forms derived here pose no limitation on Doppler measurements, and may in fact improve the measurement by their enhancement of the target SNR.

The outline of the remainder of this paper is as follows. In Sec. II general properties of the commonly used frequency-modulated pulse wave forms, and the corresponding pulse compression operations, are reviewed. In Sec. III the wave form optimization problem is formulated mathematically in terms of an objective function whose maximization determines the desired pulse. The objective function incorporates both deterministic and statistical knowledge of the scatterers. It also contains weighting parameters that allow one to adjust the desired balance between the generally competing requirements of high spatial resolution and high SNR. Considerations entering a Doppler shift measurement are also discussed. Lacking, at this stage, measured data appropriate for a realistic illustration of the theory, applications using a simulated example are discussed in Sec. IV. The paper is concluded in Sec. V.

II. SIGNAL WAVE FORMS AND PULSE COMPRESSION

The complex time-domain signal

$$p(t) = A(t)e^{-i\phi(t)} \tag{2.1}$$

emitted by a sonar system would ideally take the form of a δ function. The measured target echo $S(t)$ would then also be a δ function, and an accurate target range estimate could be obtained from simple time of flight. However, such pulses have very high energy density, and may induce nonlinear behavior in the medium (in extreme cases, cavitation). To avoid this, extended time-domain pulses are desirable, with

 ©2006 The American Physical Society

range resolution to be restored in postprocessing. Suppose that the amplitude $A(t)$ (which, for generality, need not be assumed real) and time-dependent (angular) frequency

$$\omega_\phi(t) = \partial_t \phi(t) \tag{2.2}$$

are both slowly varying, in the sense that

$$\frac{\partial_t \omega_\phi(t)}{\omega_\phi(t)^2} = O(\epsilon), \quad \frac{(t)\partial_t A(t)}{\omega_\phi A(t)} = O(\epsilon), \tag{2.3}$$

where $\epsilon \ll 1$ is a small parameter: the phase undergoes a large number $N = O(1/2\pi\epsilon)$ cycles without significant change in amplitude or frequency, and locally in time the signal looks like a simple sinusoid with well-defined frequency $\omega_\phi(t)$. As will be seen below, under this condition there is a one-to-one relation between $\omega_\phi(t)$ and the spectrum $\hat{p}(\omega)$.

Let the origin of time be chosen so that $A(t)$ is nonzero over a range $-\tau_p/2 < t < \tau_p/2$, where τ_p is the pulse length, during which the frequency ranges over a band $\omega_{\min} < \omega_\phi(t) < \omega_{\max}$. The center frequency and bandwidth are defined by

$$\omega_0 = (\omega_{\min} + \omega_{\max})/2,$$

$$\Delta\omega = \omega_{\max} - \omega_{\min}. \tag{2.4}$$

It is assumed that the mapping $t \to \omega_\phi(t)$ is invertible, associating a unique time $t_\phi(\omega)$ with each frequency in the band. Thus, $\omega_\phi(t)$ should be monotonic, typically increasing, with $\omega_{\min} = \omega_\phi(-\tau_p/2)$ and $\omega_{\max} = \omega_\phi(\tau_p/2)$.

A. Pulse spectrum

The slow variation conditions allow one to infer a one-to-one correspondence between $\omega_\phi(t)$ and the pulse spectrum,

$$\hat{p}(\omega) = \int dt\, e^{i\omega t} p(t) = \frac{1}{\epsilon\omega_0} \int ds\, \mathcal{A}(s) e^{-(i/\epsilon)[f(s) - \zeta s]}, \tag{2.5}$$

where $s = \epsilon\omega_0 t$, $\zeta = \omega/\omega_0$, $\mathcal{A}(s) = A(s/\epsilon\omega_0)$, and $f(s) = \epsilon\phi(s/\epsilon\omega_0)$. The scaling ensures that $\mathcal{A}(s)$ and $f(s)$ vary on a scale of order unity. For small ϵ the integral is dominated by the stationary point of the exponent, $\partial_s f(s) = \zeta$. In terms of the original variables, this condition specifies the time $t_\phi(\omega)$, which is the solution to

$$\omega_\phi(t) = \omega, \tag{2.6}$$

where a unique solution is guaranteed by monotonicity. Let $\phi_2(\omega) = \partial_t^2 \phi[t_\phi(\omega)] = \partial_t \omega_\phi[t_\phi(\omega)]$ be the second derivative of the phase at the stationary point. Then the stationary phase approximation to (2.3) is [11]

$$\hat{p}(\omega) \approx \sqrt{\frac{2\pi}{i\phi_2(\omega)}} A[t_\phi(\omega)] e^{-i\{\phi[t_\phi(\omega)] - \omega t_\phi(\omega)\}}, \tag{2.7}$$

demonstrating a one-to-one correspondence between the frequency-domain and time-domain pulse amplitudes.

B. Examples

For a linear frequency-modulated (LFM) pulse, or chirp, one has

$$\phi(t) = \omega_0 t + \pi K_c t^2,$$

$$\omega_\phi(t) = \omega_0 + 2\pi K_c t,$$

$$t_\phi(\omega) = (\omega - \omega_0)/2\pi K_c,$$

$$\phi_2(\omega) = 2\pi K_c, \tag{2.8}$$

and (2.7) yields

$$\hat{p}_{\mathrm{LFM}}(\omega) \approx \frac{1}{\sqrt{iK_c}} A\left(\frac{\omega - \omega_0}{2\pi K_c}\right) e^{i(\omega - \omega_0)^2/4\pi K_c}, \tag{2.9}$$

valid for $\omega_0^2/2\pi K_c \gg 1$. The bandwidth is $\Delta\omega = 2\pi K_c \tau_p$, centered on ω_0. The amplitude $A(t)$ is often taken to be constant, but should be rounded at the edges to ensure validity of the second condition in (2.3). Otherwise the spectrum will contain high-frequency side lobes not accounted for in (2.9).

For a hyperbolic frequency-modulated (HFM) pulse one has

$$\phi(t) = -\frac{\omega_0}{K_\infty} \ln(1 - K_\infty t),$$

$$\omega_\phi(t) = \frac{\omega_0}{1 - K_\infty t},$$

$$t_\phi(\omega) = \frac{\omega - \omega_0}{K_\infty \omega},$$

$$\phi_2(\omega) = -\frac{K_\infty \omega^2}{\omega_0}, \tag{2.10}$$

whence

$$\hat{p}_{\mathrm{HFM}}(\omega) = \sqrt{\frac{2\pi i\omega_0}{K_\infty \omega^2}} A\left(\frac{\omega - \omega_0}{K_\infty \omega}\right) e^{i(\omega_0/K_\infty)[\omega/\omega_0 - 1 - \ln(\omega/\omega_0)]}. \tag{2.11}$$

Validity of the stationary phase approximation requires that $\omega_0/K_\infty \gg 1$. A useful property of HFM pulses is the following scale invariance property: for any $\eta > 0$

$$\phi_{\mathrm{HFM}}(\eta t) = \phi_{\mathrm{HFM}}(t - t_\eta) + \phi_\eta,$$

$$\phi_\eta = -\frac{\omega_0}{K_\infty} \ln(\eta), \quad t_\eta = \frac{1 - \eta}{K_\infty \eta}. \tag{2.12}$$

A moving target or platform leads to a Doppler shift, which corresponds to a nonunit value of η. The phase of the HFM pulse is therefore Doppler invariant, up to an additive constant and a shift in the origin of time. If the amplitude A is constant, the pulse is invariant, except for corrections near its edges. The latter will be small if $|1 - \eta| \propto v_r/c$ is small, where v_r is the relative target-platform radial velocity.

C. Pulse compression

Frequency-modulated pulses are examples of functions that are broad in both frequency and time domains. The Heisenberg uncertainly principle [12] requires that $\tau_p \Delta\omega \geqslant \pi$, but places no upper bound on this product. As alluded to above, practical considerations forbid short, high-energy pulses. However, long FM pulses still achieve the necessary range resolution while also allowing the pulse energy to be distributed over a longer time interval. To see this, consider the overlap integral

$$P(t) = \int ds\, p(s) p^*(s-t), \tag{2.13}$$

which has Fourier transform

$$\hat{P}(\omega) = |\hat{p}(\omega)|^2 \approx \frac{2\pi |A[t_\phi(\omega)]|^2}{|\phi_2(\omega)|}. \tag{2.14}$$

The latter is a positive, smoothly varying function distributed over the interval $(\omega_{\min}, \omega_{\max})$. The inverse Fourier transform $P(t)$ will therefore have a time-domain width close to the uncertainty limit $\pi/\Delta\omega$. For a rectangular LFM pulse one obtains

$$P_{\mathrm{LFM}}(t) = |A_0|^2 \tau_p e^{-i\omega_0 t} \frac{\sin(\Delta\omega t/2)}{\Delta\omega t/2}, \tag{2.15}$$

which approaches a δ function for large bandwidth, and achieves essentially maximal range resolution $\pi c/\Delta\omega$. Smoothing of the rectangular pulse can be used to reduce the side lobes of the sinc function without loss of resolution.

Pulse compression refers to the property (2.13) that the arbitrarily broad function $p(t)$ is compressed into the narrow function $P(t)$ by convolving it with its complex conjugate. In order to take advantage of the pulse compression property, let the received signal and its Fourier transform be modeled in the form

$$S(t) = \int \mathcal{M}(t-s) p(s) ds, \quad \hat{S}(\omega) = \hat{\mathcal{M}}(\omega)\hat{p}(\omega), \tag{2.16}$$

where $\mathcal{M}(t)$ is the scene response function at delay time t, vanishing for $t<0$. For example, for single scattering from a scene with frequency-dependent reflectivity $\hat{f}(\mathbf{x},\omega)$, one obtains

$$\hat{\mathcal{M}}(\omega) = \int d\mathbf{x}\, \hat{f}(\mathbf{x},\omega) \mathcal{B}(\mathbf{x}) e^{2i\omega\tau(\mathbf{x})},$$

$$\mathcal{M}(t) = \int d\mathbf{x}\, f[\mathbf{x}, t - 2\tau(\mathbf{x})] \mathcal{B}(\mathbf{x}), \tag{2.17}$$

where $f(\mathbf{x},t)$ is the Fourier transform of $\hat{f}(\mathbf{x},\omega)$ (again vanishing for $t<0$) and represents the "microscopic" scene response at point $\mathbf{x}$. If the reflectivity is frequency independent, $\hat{f}(\mathbf{x},\omega)=\hat{f}(\mathbf{x})$, then $f(\mathbf{x},t)=\hat{f}(\mathbf{x})\delta(t)$ has no memory. The transmitter and receiver are both assumed here to be isotropic and collocated at the origin (limitations that are easily relaxed). The time delay $\tau(\mathbf{x})$ is the ray propagation time from the origin to $\mathbf{x}$, and $\mathcal{B}(\mathbf{x})$ is the appropriate two-way geometric ray spreading amplitude [13]. In a homogeneous medium one has $\tau(\mathbf{x})=|\mathbf{x}|/c$ and $\mathcal{B}(\mathbf{x})=1/(4\pi|\mathbf{x}|)^2$, and these are reasonable approximations more generally for not too distant targets. For more distant targets, heterogeneous propagation environments and accompanying ray bending effects may significantly complicate the relation between the scene reflectivity $f(\mathbf{x})$ and the response function $\mathcal{M}(t)$, but the signal structure (2.16) remains valid, and the wave form optimization procedure described below is relatively insensitive to such effects.

The signal at time t contains contributions from a broad swathe of width $c\tau_p/2$. However, the compressed signal

$$Y(t) = \int ds\, S(s) p^*(s-t) = \int ds\, \mathcal{M}(t-s) P(s),$$

$$\hat{Y}(\omega) = \hat{\mathcal{M}}(\omega)\hat{P}(\omega), \tag{2.18}$$

obtained by convolving $S(t)$ with $p^*(t)$, is sensitive only to the much narrower range swath of width $\pi c/2\Delta\omega$, limited only by the signal bandwidth.

There are other convolution operations that can be used to achieve the same purpose. The key is to compensate for the rapidly varying phase. Thus, in place of $p^*(t-s)$ one could use $\tilde{p}^*(t-s)$, where $\tilde{p}(t)=\tilde{A}(t)e^{-i\phi(t)}$ has the same phase, but different smooth amplitude $\tilde{A}(t)$. This leads to an overlap integral with Fourier transform

$$\tilde{P}(\omega) = \tilde{p}^*(\omega)\hat{p}(\omega) \approx \frac{2\pi \tilde{A}^*[t_\phi(\omega)] A[t_\phi(\omega)]}{|\phi_2(\omega)|}, \tag{2.19}$$

which is also smoothly varying. Common choices for $\tilde{A}(t)$ include rectangular pulse shapes, or a shape that yields a desired window form for $\tilde{P}(\omega)$.

More importantly, Eq. (2.19) also provides a postprocessing software implementation of the optimal wave form design derived below. In this paper, it is implicitly assumed that the pulse wave form design is performed in hardware, since this would appear to make optimal use of the pulse energy by distributing it optimally over frequency at the outset. However, simply replacing $\hat{P}(\omega)$ by $\tilde{P}(\omega)$ in all that follows results in the identical optimization condition—see Eqs. (3.10) and (3.11) below. However, rather than a hardware adjustment of $\hat{p}(\omega)$, one may now use a software adjustment (or some combination of the two) of the compression envelope $|\tilde{p}(\omega)|$ to obtain the optimal wave form.

III. WAVE FORM OPTIMIZATION

Consider now the weak target detection problem. The scene is considered to be made up of target and clutter contributions:

$$\hat{\mathcal{M}}(\omega) = \hat{\mathcal{M}}^C(\omega) + \hat{\mathcal{M}}^T(\omega),$$

$$\hat{f}(\mathbf{x},\omega) = \hat{f}^C(\mathbf{x},\omega) + \hat{f}^T(\mathbf{x},\omega), \tag{3.1}$$

which generate corresponding returns $\hat{Y}(\omega) = \hat{Y}^C(\omega) + \hat{Y}^T(\omega)$. An example to keep in mind is the discrete isotropic point scatterer model

$$f^{C,T}(\mathbf{x},\omega) = \sum_{i=1}^{N_{C,T}} \hat{f}_i^{C,T}(\omega)\,\delta(\mathbf{x} - \mathbf{x}_i^{C,T}) \tag{3.2}$$

in which the ensemble is specified by a joint probability distribution for the numbers N_C, N_T, positions $\mathbf{x}_i^C$, $\mathbf{x}_j^T$, and individual reflectivities $\hat{f}_i^C$, $\hat{f}_j^T$.

The following question is posed: If $\hat{f}^C$ and $\hat{f}^T$ have different frequency dependence, how can one optimize the pulse to best detect the target? The optimization is in general performed in a statistical sense, using an ensemble of scene realizations. We consider the frequency- and ensemble-averaged SNR, defined by

$$\mathcal{R} = \frac{\displaystyle\int \frac{d\omega}{2\pi} \langle |\hat{Y}^C(\omega) + \hat{Y}^T(\omega)|^2 \rangle}{\displaystyle\int \frac{d\omega}{2\pi} \langle |\hat{Y}^C(\omega)|^2 \rangle} - 1 = \frac{\mathcal{F}^T}{\mathcal{F}^C},$$

$$\mathcal{F}^{C,T} \equiv \int \frac{d\omega}{2\pi} |\hat{P}(\omega)|^2 \langle |\hat{\mathcal{M}}^{C,T}(\omega)|^2 \rangle, \tag{3.3}$$

in which $\langle\cdot\rangle$ represents the ensemble average, and in the second equality on the first line the reasonable assumption that the target and clutter distributions are uncorrelated has been made. If one adopts the model (3.2), and further assumes that the individual scatterers are uncorrelated, one obtains

$$\langle |\hat{\mathcal{M}}^{C,T}(\omega)|^2 \rangle = \bar{N}_{C,T} \langle |\hat{f}^{C,T}(\omega)|^2 \mathcal{B}(\mathbf{x}^{C,T})^2 \rangle \tag{3.4}$$

in which $\bar{N}_C$, $\bar{N}_T$ are the mean numbers of scatterers. The mean square signal reflects both the spatial and reflectivity distributions.

If the only object were to choose the form of $\hat{P}(\omega)$ that maximizes $\mathcal{R}$, the solution would be very simple: simply use a monochromatic signal at that value of the frequency which maximizes the ratio $\langle |\hat{\mathcal{M}}^T(\omega)|^2 \rangle / \langle |\hat{\mathcal{M}}^C(\omega)|^2 \rangle$. However, this choice leads to an infinitely broad compressed pulse, and spatial resolution is completely lost.

Restoring range resolution requires broadening the pulse bandwidth, which necessarily incorporates frequencies for which $\mathcal{R}$ is suboptimal, and hence necessarily leads to a lower SNR. However, there are many different ways to distribute the pulse energy over its bandwidth to achieve a desired resolution, and the object now is to choose the one that leads to the least reduction in SNR.

A. Objective functional

To accomplish this task, an objective functional $\mathcal{L}$ will be defined whose maximum is attained by the desired wave form. This functional will contain parameters that allow one to adjust the balance between SNR and resolution, yielding a quantitative tradeoff curve (see the center panel of Fig. 1). One may then select a particular operational point on this curve best suited to the application at hand.

To design an objective functional, one first needs to constrain the maximization by maintaining the pulse compressed property of $P(t)$, i.e., that it have width in time of order $\pi/\Delta\omega$. This is enforced through the "Heisenberg functional" [12]

$$\mathcal{H} = \Omega_2 \mathcal{T}_2, \tag{3.5}$$

in which

$$\Omega_2 \equiv \int \frac{d\omega}{2\pi} (\omega - \bar{\omega})^2 |\hat{P}(\omega)|^2,$$

$$\mathcal{T}_2 \equiv \int dt\, t^2 |P(t)|^2 = \int \frac{d\omega}{2\pi} \left| \frac{d\hat{P}(\omega)}{d\omega} \right|^2 \tag{3.6}$$

are, respectively, the spectral and temporal variances of the compressed pulse. By varying $\bar{\omega}$ one can control the first moment

$$\Omega_1 = \int \frac{d\omega}{2\pi} \omega |\hat{P}(\omega)|^2, \tag{3.7}$$

and hence the center frequency of the pulse. A well-defined minimum of $\mathcal{H}$ occurs only if the normalization

$$\mathcal{N} = \int \frac{d\omega}{2\pi} |\hat{P}(\omega)|^2 = \int dt |P(t)|^2 \tag{3.8}$$

is fixed. Notice that $\mathcal{N}$ cancels in $\mathcal{R}$ but not in $\mathcal{H}$. Minimization of $\mathcal{H}$ with fixed $\mathcal{N}$ produces a family of Gaussian pulses with minimal time-bandwidth product (TBP) (see below). The bandwidth $\Delta\omega$ itself remains arbitrary, but may be specified by fixing $\Omega_2/\mathcal{N}$. The range resolution is defined in terms of the time-width via

$$\Delta r = c\sqrt{\mathcal{T}_2/\mathcal{N}}, \tag{3.9}$$

and is therefore directly accounted for in $\mathcal{H}$.

Therefore, in order to balance the requirements of large $\mathcal{R}$ and small $\mathcal{H}$, while at the same time controlling the bandwidth and center frequency, one maximizes the combined functional

$$\mathcal{L} = \mu_R \mathcal{R} - \mathcal{H} + \lambda_\Omega \Omega_2 + \lambda_N \mathcal{N}, \tag{3.10}$$

in which μ_R determines the relative impact of the reflectivity terms on the wave form, while λ_Ω and λ_N are Lagrange multipliers used to enforce desired values of $\mathcal{N}$ and Ω_2. The parameter $\bar{\omega}$ inside Ω_2 may be viewed as another Lagrange multiplier controlling the center frequency. One expects, as will be confirmed below, that for small μ_R the SNR plays little role, and $\mathcal{L}$ will produce standard minimum TBP pulses, while for large μ_R the SNR will dominate, and a near-monchromatic pulse will result.

B. Euler-Lagrange equation

Applying the Euler-Lagrange equation [14] to (3.10), one obtains the one-dimensional Schrödinger-type equation [12]

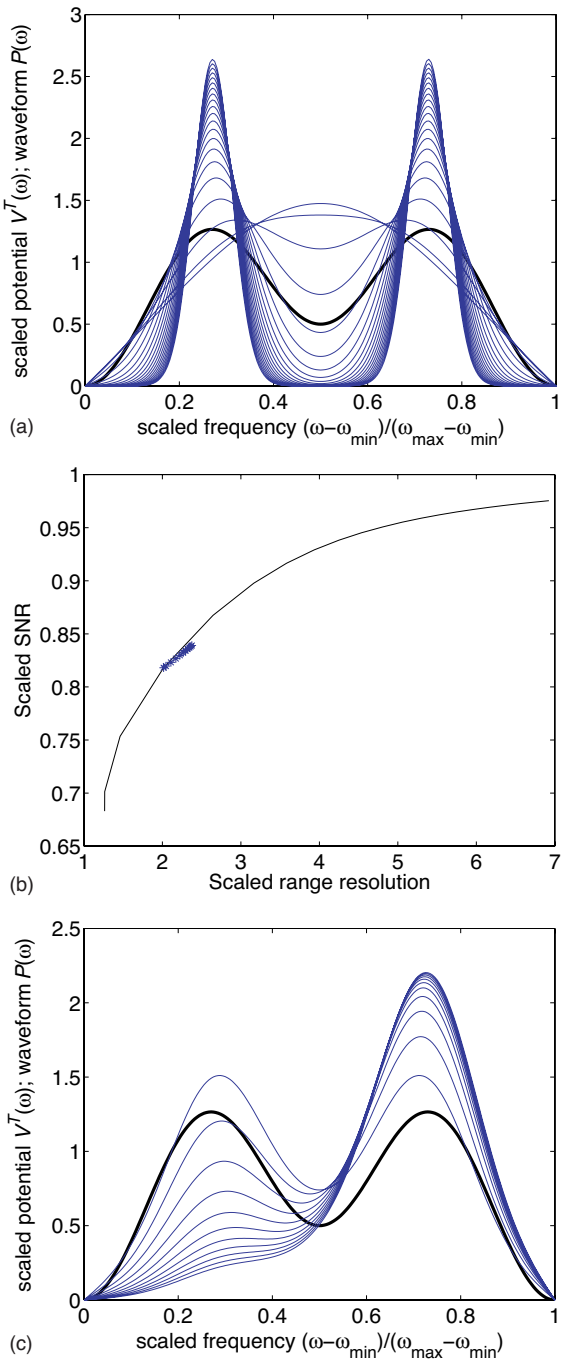

FIG. 1. (Color online) Examples of solutions to the scaled equation (4.2) using a uniform clutter potential $V^C(\omega)\equiv\mathcal{F}^C$, and a symmetric double-peaked target potential $V_T(\omega)$ (thick black line in upper and lower panels). Upper panel: Optimized wave forms $\hat{P}(\omega)$ for fixed $K_0=100$, $\bar{x}=0.5$, and a sequence of $V_0=25j^2$, $j=0,1,2,\ldots,20$. For $V_0=0$, $\hat{P}(\omega)$ is a simple half-sine function. As V_0 increases, the solution concentrates more and more under the peaks of V^T. Lower panel: Optimized wave forms for fixed $K_0=100$, $V_0=225$ (curve $j=3$ in the upper panel), and increasing $\bar{x}=0.5(1+k/20)$, $k=0,1,2,\ldots,10$. Center panel: Illustration of tradeoff between (scaled) SNR (4.5) and range resolution (4.6). Solid line shows increasing SNR, but decreasing resolution, for the solutions in the upper panel as V_0 increases. The cluster of symbols shows the much narrower variation corresponding to the curves in the lower panel.

$$-\Omega_2\frac{d^2\hat{P}}{d\omega^2}+V(\omega)\hat{P}=\lambda_N\hat{P}, \tag{3.11}$$

in which the "potential energy" function is given by

$$V(\omega)=\frac{\mu_R}{(\mathcal{F}^C)^2}[\mathcal{F}^TV^C(\omega)-\mathcal{F}^CV^T(\omega)]+(\mathcal{T}_2-\lambda_\Omega)(\omega-\bar{\omega})^2, \tag{3.12}$$

where $V^{C,T}(\omega)=\langle|\hat{\mathcal{M}}^{C,T}(\omega)|^2\rangle$ are the clutter and target effective potentials, and λ_N plays the role of an eigenvalue. The maximal solution corresponds to the ground state, i.e., the smallest eigenvalue, which is nodeless and may always be chosen real.

We note in passing that one is free to choose alternative, nonquadratic forms [e.g., $(\omega-\bar{\omega})^4$, or perhaps a square well form vanishing in a fixed interval and diverging outside of it] to replace $(\omega-\bar{\omega})^2$ in Ω_2. This same form would then appear as the second term in (3.12) and would change the detailed shape of $\hat{P}$ away from the band center, but not the essence of the SNR-range resolution balance.

By multiplying both sides of (3.11) by $\hat{P}(\omega)^*$, integrating over ω, and using (3.3), one obtains the constraint

$$(2\mathcal{T}_2-\lambda_\Omega)\Omega_2=\lambda_N\mathcal{N}, \tag{3.13}$$

which is independent of the reflectivity terms. This is the generalization of the time-bandwidth product constraint that leads to the uncertainty principle [12]. Furthermore, although solutions to (3.11) may generally be found for arbitrary values of $\mathcal{F}^C$, $\mathcal{F}^T$, Ω_2, $\mathcal{T}_2$, only for particular values are (3.3) and (3.6) satisfied. After enforcing all constraints, the optimal wave form can depend only on the single parameter μ_R.

Note that the optimization procedure, as defined, constrains only the magnitude $|\hat{p}(\omega)|=\sqrt{\hat{P}(\omega)}$ of the pulse spectrum, leaving one free to assign its phase in whatever way is convenient. Thus there are many ways to encode the optimal $\hat{P}(\omega)$ in the original wave form $p(t)$, via different choices of $A(t),\phi(t)$ for input into (2.7). One could, for example, choose a LFM pulse (hence constant ϕ_2), and then use $A(t)$ to provide the correct envelope. Alternatively, one could use a rectangular pulse and adjust $\phi(t)$ in such a way that ϕ_2 provides the correct envelope. The latter may correspond more closely with the operating characteristics of existing sonar systems. It is also closest to the approach used in Ref. [5], where $\omega_\phi(t)$ was taken as a sequence of N linear ramps (i.e., rectangular LFM subpulses) of fixed bandwidth $B=\Delta\omega/2\pi N$, on time intervals (t_{i-1},t_i), and hence with ramp rate $K_i=B/(t_i-t_{i-1})$. The optimal pulse was obtained by maximizing $\mathcal{R}$ over the K_i, with the constraint of fixed pulse length $\tau_p=t_N-t_0$.

C. Doppler resolution and ambiguity function

Sensitivity to Doppler shifted signals is incorporated by generalizing (2.13) to

$$\mathcal{P}(t,\eta) = \int ds\, p(s)p^*[\eta(s-t)], \tag{3.14}$$

with Fourier transform

$$\hat{\mathcal{P}}(\omega,\eta) = \hat{p}(\omega)\hat{p}^*(\omega/\eta). \tag{3.15}$$

For a narrowband pulse, $\Delta\omega \ll \omega_0$, one obtains $p(\eta s) \approx e^{-i\omega_D t}p(s)$, with Doppler shift $\omega_D = \omega_0(\eta - 1)$. The quantity $|P(t,\eta)|^2$ is known as the ambiguity function [10].

The reflected pulse from a moving target with radial velocity v_r yields the time stretching or compression factor $\eta_T = 1 - 2v_r/c$, and the received signal can be modeled in the form [compare (2.16)]

$$S(t) = \int ds\, \mathcal{M}(t-s)p(\eta_T s). \tag{3.16}$$

The corresponding two-parameter detection statistic is [compare (2.18)]

$$Y(t,\eta) = \int ds\, S(s)p^*[\eta(s-t)]$$

$$= \frac{1}{\eta_T}\int ds \mathcal{M}(t-s)\mathcal{P}(\eta_T s, \eta/\eta_T). \tag{3.17}$$

We have so far considered the issue of range resolution, defined via (3.9) in terms of the temporal width of $P(t) = \mathcal{P}(t,1)$. Similarly, Doppler sensitivity may be defined in terms of the width of $\mathcal{P}(0,\eta)$ about $\eta = 1$. However, unlike the former, which is fundamentally limited by the bandwidth $\Delta r \geqslant \pi c/\Delta\omega$, the latter is governed by a different uncertainty relation $\Delta v_r \geqslant \pi c/2\omega_0\tau_p$, which limits the velocity resolution by the total number of wave fronts $[\phi(\tau_p/2) - \phi(-\tau_p/2)]/2\pi \approx \omega_0\tau_p/2\pi$ in the pulse. This can be seen formally by noting that the width of $\mathcal{P}(0,\eta)$ can be estimated using the approximation $\phi(\eta s) - \phi(s) \approx (\eta - 1)s\omega_\phi(s) \approx (\eta - 1)\omega_0 s$ so that

$$\mathcal{P}(0,\eta) \approx \int ds |A(s)|^2 e^{i\omega_D s}. \tag{3.18}$$

Since the amplitude $A(s)$ has width τ_p, one obtains $\tau_p\Delta\omega_D \geqslant \pi$, from which the quoted uncertainty relation then follows. Since $|A(s)|^2$ is a positive smooth function, one expects rough achievement of the bound, $\Delta\omega_D \approx \pi/\tau_p$.

It follows that the Doppler resolution may be increased arbitrarily by extending the temporal pulse width τ_p, while keeping the spectrum (and hence the range resolution) fixed. This corresponds to reducing the value of the small parameter ϵ in (2.3) and (2.5), while keeping the functions $f(s)$ and $\mathcal{A}(s)$ fixed. The spectrum changes only by an overall multiplicative factor [via the $1/\epsilon$ prefactor in (2.5)].

In summary, given an optimal wave form $p_{\text{opt}}(t)$, constructed from the solution $\hat{P}_{\text{opt}}(\omega)$ to (3.11), the corresponding ambiguity function $|\mathcal{P}_{\text{opt}}(t,\eta)|^2$ is specified by (3.14). The detection statistic (3.17) for a compact target will be peaked in range, as prescribed by $P_0(t)$, and in Doppler shift near $\eta = \eta_T$, with resolution $\approx \pi/\omega_0\tau_p$ that may be increased arbitrarily by increasing τ_p. Since Doppler resolution of a moving target in clutter must also rely on maximizing its signal relative to that of the background, the optimal wave forms derived in this work, which tend to brighten images of targets relative to clutter, may be expected to similarly aid such a measurement.

IV. APPLICATIONS

If $\mu_R = 0$ in (3.12) one obtains the usual formulation of the uncertainty principle. Equation (3.9) becomes the standard harmonic oscillator equation. Setting the normalization $\mathcal{N} = 1$, one finds $\Omega_1 = \bar{\omega}$ and $\lambda_\Omega = 0$. The constraint (3.13) takes the form $\Omega_2\mathcal{T}_2 = \lambda_N = 1/4$ (equivalent to the minimal TBP of π), and $\hat{P}(\omega)$, $P(t)$ take the Gaussian forms

$$\hat{P}(\omega) = \frac{e^{-(\omega-\Omega_1)^2/4\Omega_2}}{(\Omega_2/2\pi)^{1/4}}, \quad P(t) = e^{-i\Omega_1 t}\frac{e^{-t^2/4\mathcal{T}_2}}{(2\pi\mathcal{T}_2)^{1/4}}. \tag{4.1}$$

For nonzero μ_R, the reflectivity terms will in general distort $\hat{P}(\omega)$ away from the pure Gaussian form (4.1). In particular, $|\hat{P}(\omega)|^2$ will be enhanced in regions where $V(\omega)$ is minimal, i.e., where, as might be expected, $V^T(\omega)$ dominates $V^C(\omega)$. The total potential (3.12), of course, includes the quadratic term whose coefficient determines the spread of $\hat{P}(\omega)$ around its minimum.

Equation (3.11) is easily implemented using standard symmetric matrix diagonalization routines. To illustrate quantitatively the evolution of the wave form with μ_R, simulated examples are shown in Fig. 1. For simplicity, the clutter potential has been chosen frequency independent, $V^C(\omega) \equiv \mathcal{F}^C$, while the target potential has been given a symmetric double-peaked structure. Boundary conditions have been applied so that $\hat{P}(\omega)$ vanishes on the boundaries of the frequency band. For clarity, the plots are based on the rescaled version of (3.11):

$$-\frac{\partial^2\hat{P}}{dx^2} + [K_0(x-\bar{x})^2 - V_0V^T(x)]\hat{P} = \lambda\hat{P} \tag{4.2}$$

in which the scaled frequency

$$x = \frac{\omega - \omega_{\min}}{\Delta\omega} \tag{4.3}$$

lies in the interval [0,1], and

$$V_0 = \mu_R\Delta\omega^2/\Omega_2\mathcal{F}^C,$$

$$K_0 = \frac{(\mathcal{T}_2 - \lambda_\Omega)\Delta\omega^4}{\Omega_2},$$

$$\lambda = \frac{\lambda_N\Delta\omega^2 - \mu_R\mathcal{F}^T/\mathcal{F}^C}{\Omega_2}. \tag{4.4}$$

The amplitude V_0 now controls the relative influence of the SNR and reflectivity terms, K_0 controls the scaled bandwidth, and $\bar{x}$ controls the scaled center frequency. Figure 1

illustrates the variation of the pulse spectrum with these parameters. The upper panel of the figure illustrates effects of biasing the wave form toward higher SNR. The thick line shows the chosen reflectivity potential $V^T(x)$, while the remaining curves show numerical (ground state) solutions $\hat{P}$ to (4.2) for fixed $K_0=100$, $\bar{x}=0.5$, and a sequence of values of V_0 in the range $0\leqslant V_0\leqslant 10\,000$. For $V_0=0$ the wave form is completely unbiased by the scene, and is comparable to standard wave forms, such as those discussed in Sec. II B. As V_0 increases, the wave form spectrum narrows, becoming more concentrated near the maxima of V^T. The lower panel of the figure shows the evolution of the spectrum with center frequency at fixed SNR bias. The same $V^T(x)$ is plotted, along with a sequence of solutions for fixed $K_0=100$, $V_0=225$, and a sequence of values of $\bar{x}$ in the range $0.5\leqslant\bar{x}\leqslant 0.75$. As $\bar{x}$ increases the solution is pushed to the right and concentrates under the right-hand peak. For all curves, the normalization is chosen so that $\int_0^1 (dx/2\pi)\hat{P}^2=1$.

The center plot in the figure shows the corresponding tradeoff curve in which the variation of the SNR,

$$\frac{\mathcal{R}}{\mathcal{R}_{\max}}=\frac{1}{V^T_{\max}}\int_0^1 \frac{dx}{2\pi}\hat{P}(x)^2V^T(x), \tag{4.5}$$

scaled by its maximum value $V^T_{\max}/\mathcal{F}^C$, is plotted against the range resolution of the compressed pulse,

$$\frac{\Delta r}{(c/\Delta\omega)}\equiv\sqrt{\int_0^1 \frac{dx}{2\pi}\left(\frac{d\hat{P}}{dx}\right)^2}, \tag{4.6}$$

scaled by $c/\Delta\omega$, the rough minimum value achieved for $V_0=0$. The solid line corresponds to the curves in the upper plot, with larger values of V_0 leading to higher SNR, but lower range resolution. The cluster of symbols corresponds to the curves in the lower plot: since V_0 is fixed, there is little variation in this case. More general unbiased wave forms, such as those described in Sec. II B, would cluster further down the high-resolution, low-SNR, end of the curve.

V. DISCUSSION AND CONCLUSIONS

The results shown in the figure illustrate the substantial freedom available to tailor pulses according to desired bandwidth, band center, and influence of target scattering properties. In the example, there is a 44% increase in SNR, achieving 98% of the maximum possible SNR, with a factor of 5.5 reduction in range resolution. More than 90% of the maximal SNR is achieved with only a factor of 3 reduction in resolution.

Clearly the level of enhancement in any given application will be strongly dependent on the frequency variability of the contrast $V(\omega)$, and its significance depends on the concept of operation (e.g., target size and clutter density). The qualitative properties of the method, and the degree of potential resolution tradeoff required to improve the SNR, should, however, be clear from the examples.

In high-clutter environments, where compromising range resolution is undesirable, one may nevertheless be able to make productive use of SNR-optimized pulses via a follow-on interrogation procedure. Thus, once a subset of possible targets is resolved by an "ordinary" high-resolution pulse, the scene could be interrogated (or simply postprocessed, as described at the end of Sec. II C) with optimized pulses to see if returns from some regions "brighten" relative to others. If the elements of this subset are sufficiently well separated compared to the total number of clutter returns, this may aid in isolating the true target.

Exploring such issues using more realistic (preferably data-based) examples will be a topic of future work. Typically $V^C(\omega)$ would be obtained from a series of prior measurements in a region of interest, while $V^T(\omega)$ would be estimated, e.g., from (3.4) using prior knowledge of expected target reflectivity characteristics, together with assumptions regarding its expected spatial distribution. With this knowledge, $V(\omega)$ can be constructed, and an optimal wave form rapidly found, constrained by the desired frequency band characteristics.

ACKNOWLEDGMENT

Support from the U.S. Navy, under Contract No. N00024-02-C-4103, is gratefully acknowledged.

[1] See, e.g., M. Soumekh, *Synthetic Aperture Radar Signal Processing* (John Wiley & Sons, New York, 1999).
[2] See, e.g., I. S. Reed, J. D. Mallett, and L. E. Brennan, IEEE Trans. Aerosp. Electron. Syst. **10**, 853 (1974).
[3] See, e.g., W. Munk, P. Worcester, and C. Wunsch, *Ocean Acoustic Tomography* (Cambridge University Press, Cambridge, U.K., 1995).
[4] Additional polarization degrees of freedom are available to electromagnetic sensors that could also be optimized.
[5] P. M. Baggenstoss, IEEE J. Ocean. Eng. **23**, 1 (1998).
[6] T. Collins and P. Atkins, IEE Proc., Radar Sonar Navig. **145**, 347 (1998).
[7] R. J. Bonneau, in IEEE Antennas and Propagation Society International Symposium, 2001, Vol. 4, pp. 814–816. Available online at http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=959589
[8] R. J. Bonneau and M. C. Wicks, in Proceedings of the IEEE Radar Conference, 2001, pp. 448–450. Available online at http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=923021
[9] B. Berksoy and L. Wei, in Proceedings of the 58th Vehicular Technology Conference, 2003, Vol. 5, p. 3174. Available online at http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1286218.
[10] For background and further references on Doppler effects see, e.g., X. Lurton, *Introduction to Underwater Acoustics*

(Springer, Praxis, 2002), especially Chap. 6. In the radar context, see, e.g., M. I. Skolnik, *Introduction to Radar Systems*, 3rd ed. (McGraw-Hill, New York, 2001), especially Chaps. 3 and 6. The basic concepts are essentially identical for both wave types.

[11] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, New York, 1978).

[12] E. Merzbacher, *Quantum Mechanics*, 2nd ed. (John Wiley and Sons, New York, 1970).

[13] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (Springer-Verlag, New York, 1993), especially Chap. 3.

[14] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), especially Part I, Chap. 3.