Performance of a Linear-Detector Joint Radar-Communication System in Doubly Selective Channels

ANDREW D. HARPER, Student Member, IEEE
Georgia Institute of Technology, Atlanta, GA, USA

JEREMY T. REED, Member, IEEE
JONATHAN L. ODOM, Member, IEEE
Georgia Tech Research Institute, Atlanta, GA, USA

AARON D. LANTERMAN, Member, IEEE
XIAOLI MA, Fellow, IEEE
Georgia Institute of Technology, Atlanta, GA, USA

When radar and communication systems are colocated and operating simultaneously in the same frequency band, interference can be addressed by designing the systems to share resources. This paper develops a framework for cooperative operation of single-input single-output bistatic radar and wireless communication systems. We adopt a low-complexity linear minimum mean square error (LMMSE) optimal pilot symbol aided modulation scheme, and show the achievability region of the joint radar-communication system designed for simultaneous operation in a wireless channel that is both frequency and time-selective. Since the LMMSE estimate requires accurate knowledge of channel statistics, we give a lower bound on performance using least-squares estimation that operates purely on the received signal. Finally, to better match power levels between systems, we compare the performance using optimal training sequences with a suboptimal scheme using Barker sequences.

Manuscript received August 26, 2015; revised February 7, 2016 and August 3, 2016; released for publication August 19, 2016. Date of publication February 6, 2017; date of current version April 27, 2017.

DOI. No. 10.1109/TAES.2017.2664481

Refereeing of this contribution was handled by A. Charlish.

This work was supported by the Defense Advanced Research Projects Agency under the Shared Spectrum Access for Radar and Communications program, Contract HR0011-14-C-0026.

Authors' addresses: A. D. Harper, A. D. Lanterman, and X. Ma are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA, E-mail: (andrewharper@gatech.edu; lanterma@ece.gatech.edu; xiaoli@gatech.edu); J. T. Reed and J. L. Odom are with Georgia Tech Research Institute, Atlanta, GA 30318 USA, E-mail: (jeremy.reed@gtri.gatech.edu; jonathan.odom@gtri.gatech.edu).

I. INTRODUCTION

As the prevalence of wireless technologies continues to grow, the spectrum becomes increasingly crowded. Hence, interference between systems operating in the same frequency band and at the same time is also an increasingly important issue. This paper investigates the performance of a cooperative, jointly designed communication-radar system. At the core of this investigation is the notion that, despite major differences in overall mission, the processes involved in radar detection mirror those of channel estimation in a communication system in many ways. By cooperating, the redundant processing can be shared and interference reduced.

If both systems operate independently, in overlapping frequency bands, and without regard to the other, interference is unavoidable. The most straightforward solution to eliminate interference is to divide the resources available into separate regions of time or frequency through policy, such as the federal communications commission spectral allocation or dynamic spectrum access [1], [2]. In an environment where it is desired to operate both radar and communication systems, this amounts to either the systems taking turns using the whole frequency band or allocating two nonoverlapping portions of the band for simultaneous operation. Here, we explore the possibility of dividing neither time nor frequency; our system allows both systems to share the entire available spectrum at all times.

In a frequency-selective channel, the transmitted signal is spread in time, causing inter-symbol interference. In a time-selective channel, the transmitted signal is spread in frequency (i.e., Doppler shift). Signals propagating in doubly selective channels are spread in both time and frequency, both of which must be corrected for at the receiver. If the receiver does not know the channel *a priori*, it is common to devote a portion of the transmitted signal to training (i.e., pilot) symbols. Pilot symbol aided modulation (PSAM) is a popular choice for pilot symbols, and is able to capture time-domain channel variations by spreading the pilot symbols throughout a transmission block [3], [4]. Similarly, radar uses repeated pulses like those in PSAM to identify Doppler shifts induced by moving targets [5]. While the goals of radar and communication are inherently different, both require characterization of the multipath channel; that is, both systems require a channel estimate. Radar systems use the channel estimate to detect and locate targets, whereas communications systems use the estimate to account for the effects of fading on the transmitted data. This forms the basis for our joint radar-communication system, where a single channel estimation process is shared between the two systems.

We assume the multipath environment is doubly selective, i.e., the channel varies over both frequency and time. Further, we design our transmit signal so that the received signal can be easily decomposed into data-only and training-only signals. This decouples the channel and data estimation processes. In the decoupled system, we allocate a portion of power to estimate the channel and use the

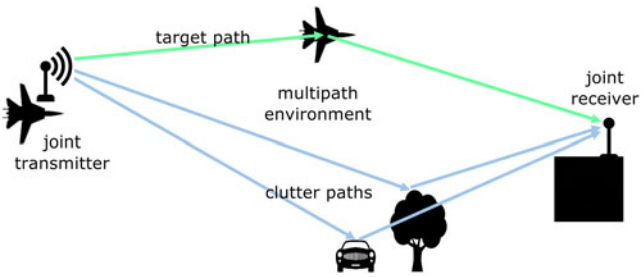


Fig. 1. Problem setup. The joint system uses a single antenna at both the mobile transmitter and stationary receiver, and designs a single waveform to be used by both the wireless communication and bistatic radar systems.

remainder to transmit data. For a block transmission scheme with decoupled linear estimation, equipowered and equispaced training symbols have been shown to be both the minimum mean-square error (MMSE) and maximum-capacity (lower bound) strategy [6]. We begin with this equipowered and equispaced single-pulse strategy, which has been shown to be optimal for communication. Because radar detection requires more energy than channel estimation for communication, this optimal scheme may not be feasible to implement using actual radar amplifiers. To explore more practical solutions, we consider suboptimal training schemes with energy spread over multiple consecutive pulses. In particular, we compare the capacity lower bound from optimal training with those achieved using Barker sequences.

The combination and/or cooperation of radar and communication is not a new topic. Previous studies have primarily addressed either waveform design or radar-communication coexistence issues (i.e., interference mitigation). The joint operation we study in this paper requires full cooperation between the radar and communication in both waveform design and hardware. The problem setup is illustrated in Fig. 1. To highlight the contributions of our work, we enumerate the following features of our chosen configuration:

1) One single-antenna transmitter shared by radar and communication systems;
2) One transmitted signal consisting of both data and pilot symbols;
3) Narrowband transmission;
4) One shared signal-antenna receiver;
5) One received signal, which is distributed to each system after reception for independent processing;
6) Elimination of interference in the time domain;
7) One (at least) mobile terminal.

Since we design our one transmitted signal for both systems, both systems share the full spectrum at all times. In contrast to previous works that propose dividing the available spectrum between systems (e.g., [7]–[9]), our joint design *capitalizes* on redundancy between channel estimation and radar detection by allowing both systems to use the same training signal. Our single antenna transmission design is desirable in that it is low complexity, helps keep hardware costs low, and is more applicable to the predominant existing hardware; however, it preempts use of MIMO techniques such as the orthogonal waveform interference mitigation in [10], or multiple-access channel models as in [11]. We also show that a narrowband approach is effective, in contrast to wideband or ultrawideband techniques in [8], [10], [12].

With the time-selective channel model we use, OFDM techniques such as those used in [13]–[16] are likely to experience interference from spectral leakage, and system performance may suffer. Therefore, we employ single carrier transmission. To capture the variations of the channel in both the time and frequency domains, we employ a basis-expansion [6]. We use the channel estimate to both improve data decoding through channel equalization and to detect moving targets. However, the transmit scheme that both minimizes the channel estimation error and maximizes a lower bound on capacity may result in a waveform that stresses current amplifier designs because of a high peak-to-average power ratio (PAPR). Therefore, similar to [13] and [17], we consider suboptimal training sequences. In particular, we compare the effects of Barker codes of varying lengths.

Finally, the majority of previous studies assume a monostatic radar that uses a portion of its transmit signal (or bandwidth) to communicate to a separate receiver that is uninterested in the radar signal. In contrast, we assume the bistatic case where the receiver is both interested in the data sent by the transmitter and detecting targets in the multipath environment. Note that the receiver knows the training portion of the waveform, but does not know the data. Rather than dividing time or spectrum between the two systems, we keep the bandwidth and structure of the waveform constant while dividing the total power between data and pilot symbols. The bistatic radar model we use significantly complicates Doppler calculations. For a comprehensive treatment of bistatic clutter and bistatic Doppler derivations, respectively, see [18, Ch. 11] and [19, Ch. 5].

Notation: Bold-face capital (lower-case) letters denote matrix (vector) values, and plain text denotes scalar values. Superscript $^{\mathrm{H}}(^{\mathrm{T}})$ denotes the Hermitian (regular) transpose, and superscript * is used for conjugation. $\lfloor\cdot\rfloor$ is the integer floor. E[$\cdot$] denotes expectation. The trace of a matrix is denoted tr[$\cdot$]. We use $\|\cdot\|$ to denote the Euclidean norm of a vector, and $\otimes$ to denote the Kronecker product. We use diag[$\mathbf{x}$] to denote a square matrix with vector $\mathbf{x}$ as its diagonal elements and zeros elsewhere.

II. SYSTEM MODEL

A. Doubly Selective Channel Model

The impulse response of a doubly selective channel varies with time parameter t and delay parameter τ. Let the time- and delay-varying channel impulse response be $h(t;\tau)$. The scattering function

$$S_h(f;\tau) \triangleq \int_{-\infty}^{\infty} A_h(\Delta t;\tau)\mathrm{e}^{-j2\pi f\Delta t}d\Delta t \qquad (1)$$

describes the average output power of the channel for Doppler frequency f and delay τ, where

$$A_h(\Delta t;\tau) \triangleq E\left[h^*(t;\tau)h(t+\Delta t;\tau)\right] \tag{2}$$

is the two-dimensional (2-D) autocorrelation function with respect to time shift Δt. The average and root-mean-square delay spreads are found using the standard definitions in [20, Ch. 3]. We assume $|S_h(f;\tau)| \approx 0$ for $|\tau| > \tau_{\max}$ and $\forall\, f$, and $|S_h(f;\tau)| \approx 0$ for $|f| > f_{\max}$ and $\forall\, \tau$. For a given channel, the values of $\tau_{\max}$ and $f_{\max}$ are easily measured empirically, and thus, assumed to be known. We assume the channel remains constant for some duration NT_s, where T_s is the sampling period at the receiver and N is a positive integer. This model invites block-based processing, where a contiguous block of N transmitted symbols is processed concurrently.

We use a Fourier basis expansion model (BEM) to capture the multipath delay and Doppler characteristics of the channel. With knowledge of $f_{\max}$ and $\tau_{\max}$, the channel can be represented with $L+1$ delay samples and $Q+1$ Doppler samples where $L = \lfloor \tau_{\max}/T_s \rfloor$ and $Q = \lceil f_{\max}/\Delta f \rceil$, $1/\Delta f = NT_s$, and where Δf is the Doppler sample spacing. Therefore, all multipath reflections arrive within L delay samples after the pulse, with a Doppler-sample between $-Q/2$ and $Q/2$. Let i denote the received sample index, and define the block index $k = \lfloor i/N \rfloor$. Using the BEM, each $h(i;l)$ is defined by its Fourier coefficients $h_q(k;l)$ to capture the time-selectivity of the channel:

$$h(i;l) = \sum_{q=0}^{Q} h_q(k;l)\mathrm{e}^{j\omega_q i} \quad \forall\, l \in \{0, L\} \tag{3}$$

where

$$\omega_q \triangleq \frac{2\pi}{N}(q - Q/2) \tag{4}$$

is the qth frequency component. Thus, (3) models the channel with a set of complex exponentials located on a 2-D grid of delay and Doppler locations. In pulsed radar systems, the delay is measured by match filtering the received data. Further, for a given delay, the temporal evolution of the match filtered response is evaluated over a series of pulses because it is generally assumed that the Doppler frequency of scatterers is very small compared with the bandwidth of the signal [5]. Therefore, the received data from a single channel is viewed as 2-D grid of samples, where one of the dimensions is called "fast-time," which is used to measure the l-dimension, and the other dimension is called "slow-time," which is used to measure the q-dimension.

We also assume that $h(i;l) = c(i;l) + s(i;l)$ is the summation of clutter, $c(i;l)$, and target $s(i;l)$. Similar to $h(i;l)$, the BEM expands $c(i;l)$ and $s(i;l)$ as

$$c(i;l) = \sum_{q=0}^{Q} c_q(k;l)\mathrm{e}^{j\omega_q i} \tag{5}$$

$$s(i;l) = \sum_{q=0}^{Q} s_q(k;l)\mathrm{e}^{j\omega_q i} \tag{6}$$

where $c_q(k;l)$ and $s_q(k;l)$ are the Fourier coefficients of the kth block for the clutter and target, respectively. Each $c_q(k;l)$ is an independent, but not identically distributed, zero-mean circularly symmetric complex Gaussian random variable with variance $\sigma^2_{c_q(k,l)}$. We assume the target is a Swerling I point target [5] independent of all clutter. Let l_0 and q_0, respectively, denote the delay bin and Doppler basis of the target. Then the sum in (6) simplifies to

$$s(i;l) = \alpha(k)\delta(l - l_0)\mathrm{e}^{j\omega_{q_0} i} \tag{7}$$

where $\alpha(k) \sim \mathcal{CN}(0, \sigma^2_\alpha)$ is the effective complex target gain that incorporates the target radar cross section.

B. Block Transmission Model

As stated in Section II-A, the channel model invites block-based processing. Define $\mathbf{u}_k$ as the kth transmitted block vector. Our transmission scheme is the same as [6], which breaks $\mathbf{u}_k$ into M subblocks. The first and last L symbols in the training portion of each subblock are set to zero (i.e., zero-padded), which prevents the returns from the information and training portion of the subblocks from interfering with each other. Furthermore, this ensures there is no interference between blocks or between subblocks within a block. Although it is also possible to decouple equally effectively by using a cyclic prefix, zero-padding is more power efficient since all transmitted cyclic prefix symbols are discarded at the receiver. This training strategy, therefore, effectively decouples channel estimation from symbol decoding.

The motivation for designing decoupled signaling is to reduce decoding complexity at the receiver. In general, estimating the channel and transmitted data is a nonlinear estimation problem (see, e.g., [21], [22]). By padding the space between information symbols and the training symbols with L zeros, the problem splits into two separate linear estimation problems by simply selecting the appropriate samples at the receiver.

Since our assumed design eliminates interblock interference, and since we will process each received block in an identical manner, we, hereafter, focus on one block and omit the block index k. Therefore, we may write the transmit sequence as

$$\mathbf{u} = [\mathbf{d}_1^{\mathrm{T}}\ \mathbf{b}_1^{\mathrm{T}}\ \mathbf{d}_2^{\mathrm{T}}\ \mathbf{b}_2^{\mathrm{T}}\ \ \ldots\ \mathbf{d}_M^{\mathrm{T}}\ \mathbf{b}_M^{\mathrm{T}}]^{\mathrm{T}} \tag{8}$$

where $\mathbf{d}_m$ is the mth data subblock and $\mathbf{b}_m$ is the mth training subblock. With this transmission scheme, the ith received sample is

$$y(i) = \sum_{l=0}^{L} h(i;l)u(i-l) + w(i) \tag{9}$$

where $w(i)$ is zero-mean circularly symmetric complex Gaussian with variance σ_w^2. In matrix-vector form, the received block is

$$\mathbf{y} = \mathbf{H}\mathbf{u} + \mathbf{w} \tag{10}$$

where $\mathbf{y} = [y(1)\ldots y(N)]^{\mathrm{T}}$ and $\mathbf{w} = [w(1)\ldots w(N)]^{\mathrm{T}}$. Defining $\mathbf{D}_q = \mathrm{diag}[1\ \mathrm{e}^{j\omega_q}\ \ldots\ \mathrm{e}^{j\omega_q(N-1)}]$, the Fourier-basis

channel expansion is

$$\mathbf{H} = \sum_{q=0}^{Q} \mathbf{D}_q \mathbf{H}_q. \tag{11}$$

Each $\mathbf{H}_q$ is a lower-triangular $N \times N$ Toeplitz matrix with first column $[h_q(0) \ \ldots \ h_q(L) \ \mathbf{0}_{1\times(N-L)}]^{\mathrm{T}}$. Note that for any q, the Fourier coefficients are functions only of delay τ, and the time-selectivity of the channel is described in the frequency domain.

Define the block data symbol vector $\mathbf{d} \triangleq [\mathbf{d}_1^{\mathrm{T}} \ \mathbf{d}_2^{\mathrm{T}} \ \ldots \ \mathbf{d}_M^{\mathrm{T}}]^{\mathrm{T}}$ and the block training vector $\mathbf{b} \triangleq [\mathbf{b}_1^{\mathrm{T}} \ \mathbf{b}_2^{\mathrm{T}} \ \ldots \ \mathbf{b}_M^{\mathrm{T}}]^{\mathrm{T}}$, and let $\bar{N}_m^d$ and $\bar{N}_m^b$ denote the length of the mth data and training subblocks, respectively. The total number of data and training symbols are then $N_d = \sum_{m=1}^{M} \bar{N}_m^d$ and $N_b = \sum_{m=1}^{M} \bar{N}_m^b$, respectively. The symbols arriving at the receiver only as a result of training symbols (i.e., the portion of the received signal containing no multipath delay from data symbol transmission) are $\mathbf{y_b} = \mathbf{H_b b} + \mathbf{w_b}$, which are easily extracted from $\mathbf{y}$ by simply selecting the rows with indices corresponding to integers on the interval $[r_1^b, r_2^b]$, where

$$r_1^b = (m-1)(\bar{N}_m^d + \bar{N}_m^b) + \bar{N}_m^d + L + 1 \tag{12}$$

$$r_2^b = m(\bar{N}_m^d + \bar{N}_m^b) \tag{13}$$

for each $m \in [1, M]$. We may write $\mathbf{y_b} = [(\mathbf{y}_0^b)^{\mathrm{T}} \ (\mathbf{y}_1^b)^{\mathrm{T}} \ \ldots \ (\mathbf{y}_{M-1}^b)^{\mathrm{T}}]^{\mathrm{T}}$, where

$$\mathbf{y}_m^b = \mathbf{H}_m^b \mathbf{b}_m + \mathbf{w}_m^b \tag{14}$$

the portion of the channel matrix that only affects the mth subblock is $\mathbf{H}_m^b = \sum_{q=0}^{Q} \mathbf{D}_{q,m}^b \mathbf{H}_{q,m}^b$, $\mathbf{w}_m^b$ contains the noise samples for the training portion of the mth subblock, and $\mathbf{H}_{q,m}^b$ and $\mathbf{D}_{q,m}^b$ are the submatrices of $\mathbf{H}_q$ and $\mathbf{D}_q$ due to the training portion of the mth subblock, respectively. Fig. 2 provides an illustration of how zero-padding allows the total received signal $\mathbf{y}$ to be easily split into data vector $\mathbf{y}_d$ and training vector $\mathbf{y}_b$; the channel length L determines the number of zeros added so that, for each subblock, the multipath reflections from one transmitted signal type (i.e., training or data) die out before the other type begins. Expanding the pilot-only signal across both Fourier bases and subblock index, we have

$$\begin{aligned}
\mathbf{y_b} &= \sum_{q=0}^{Q} \mathbf{D}_q^b \mathbf{H}_q^b \mathbf{b} \\
&= \sum_{q=0}^{Q} \begin{bmatrix} \mathbf{D}_{q,1}^b \mathbf{H}_{q,1}^b \mathbf{b}_1 \\ \vdots \\ \mathbf{D}_{q,M}^b \mathbf{H}_{q,M}^b \mathbf{b}_M \end{bmatrix} \\
&= \sum_{q=0}^{Q} \begin{bmatrix} \mathbf{D}_{q,1}^b \mathbf{B}_1 \\ \vdots \\ \mathbf{D}_{q,M}^b \mathbf{B}_M \end{bmatrix} \mathbf{h}_q
\end{aligned} \tag{15}$$

where the last equality of (15) follows from the commutativity of convolution ($\mathbf{H}_{q,m}^b \mathbf{b}_m = \mathbf{B}_m \mathbf{h}_q$). By expanding the sum over Q, we can write the received training signal in vector matrix form as

$$\mathbf{y_b} \triangleq \Phi_\mathbf{b} \mathbf{h} + \mathbf{w_b} \tag{16}$$

where $\Phi_\mathbf{b}$ is a *known* block matrix with $[\Phi_\mathbf{b}]_{m,q} = \mathbf{D}_{q,m}\mathbf{B}_m$, $\mathbf{h}_q \triangleq [h_q(0) \ \ldots \ h_q(L)]^{\mathrm{T}}$, and $\mathbf{h} = [\mathbf{h}_0^{\mathrm{T}} \ \mathbf{h}_1^{\mathrm{T}} \ \ldots \ \mathbf{h}_Q^{\mathrm{T}}]^{\mathrm{T}}$.

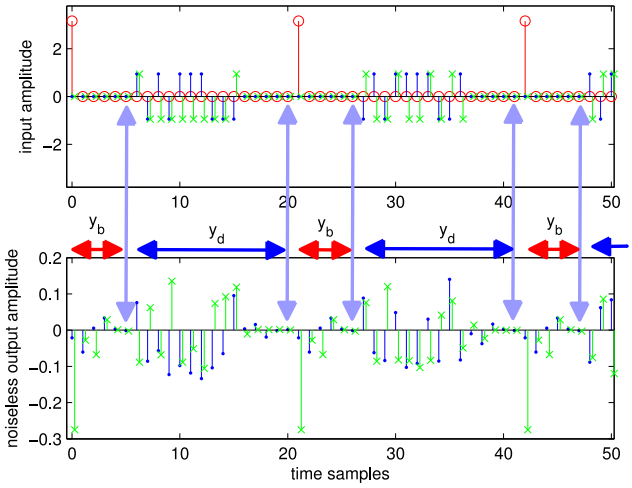


Fig. 2. Subblock structure of an example transmit waveform using optimal training (top) and the resulting noise-free channel output (bottom), for channel length 5 and data subblock length 10. Red circles show the single training symbol per subblock. Blue dots and green x's show the real and imaginary parts of the data symbols using QPSK modulation. Vertical arrows match the end of the zero padding regions (top) with corresponding spots where received channel energy returns to zero. Horizontal arrows indicate the partitioning of the received (noise-free) signal into data and training.

III. SYSTEM PERFORMANCE

Define the total power transmitted as

$$P \triangleq \mathrm{E}\left[\|\mathbf{u}\|^2\right] \tag{17}$$

which is split between power allocated to data symbols ($P_d \triangleq \mathrm{E}[\|\mathbf{d}\|^2]$) and training symbols ($P_b \triangleq \|\mathbf{b}\|^2$). To quantify the effect of sharing power between systems, we define the power allocation parameter ρ such that $P_d \triangleq \rho P$ and $P_b \triangleq (1-\rho)P$. Denote the average power per data and training symbol, respectively, as $\bar{P}_d = P_d/N_d$ and $\bar{P}_b = P_b/N_b$. Further, define $\mathbf{H_d}$ as the portion of the channel matrix affecting only data symbols and $\mathbf{y_d} = \mathbf{H_d d} + \mathbf{w_d}$ as the resulting received signal; these can be formed by selecting from $\mathbf{H}$, $\mathbf{y}$, and $\mathbf{w}$ the rows with indices corresponding to integers on the interval $[r_1^d, r_2^d]$, where

$$r_1^d = (m-1)(\bar{N}_m^d + \bar{N}_m^b) + 1 \tag{18}$$

$$r_2^d = (m-1)(\bar{N}_m^d + \bar{N}_m^b) + \bar{N}_m^d + L \tag{19}$$

for each $m \in [1, M]$ and column indices on the intervals $[c_1^d, c_2^d]$, where

$$c_1^d = (m-1)(\bar{N}_m^d + \bar{N}_m^b) + 1 \tag{20}$$

$$c_2^d = (m-1)(\bar{N}_m^d + \bar{N}_m^b) + \bar{N}_m^d \tag{21}$$

for each $m \in [1, M]$.

In order to analyze performance, signal-to-noise ratio (SNR) is used by both the radar and communication communities. However, each community has a different signal of interest; therefore, the term is vague when analyzing the performance of both systems. We, therefore, choose SNR to refer to the ratio of the total received power (including multipath reflections) to the noise power. We further refer to the ratio of the received power to due reflections from the target to the noise power as the target-to-noise ratio.

A. Radio Optimization

1) *Linear Minimum Mean Square Error (LMMSE) Channel Estimation:* For a given channel realization, the channel capacity is a function of the transmitted power and random channel gains. Averaging over the channel realization yields the ergodic capacity

$$\overline{C} \triangleq \frac{1}{N}\mathrm{E}\left[\log\det\left(\mathbf{I}+\frac{\bar{P}_d}{\sigma_w^2}\mathbf{H_d}\mathbf{H_d}^{\mathrm{H}}\right)\right] \text{ bits/s/Hz} \quad (22)$$

where expectation is over random channel gains $\mathbf{H_d}$ (i.e., the parts of $\mathbf{H}$ contributing to $\mathbf{H_d}$). The overbar notation signifies that $\overline{C}$ represents an *upper bound* on capacity since $\mathbf{H_d}$ is unknown and must be estimated. Since no channel information is assumed known at the transmitter, (22) implicitly assumes uniform power allocation (i.e., $\mathrm{E}\left[\mathbf{d}\mathbf{d}^{\mathrm{H}}\right]=\bar{P}_d\mathbf{I}$). This upper bound provides a benchmark for performance with channel estimation. The LMMSE channel estimate is [6], [23]

$$\hat{\mathbf{h}} = \frac{1}{\sigma_w^2}\left(\mathbf{R}_\mathbf{h}^{-1}+\frac{1}{\sigma_w^2}\Phi_\mathbf{b}^{\mathrm{H}}\Phi_\mathbf{b}\right)^{-1}\Phi_\mathbf{b}^{\mathrm{H}}\mathbf{y} \quad (23)$$

where $\mathbf{R_h}=\mathrm{E}\left[\mathbf{h}\mathbf{h}^{\mathrm{H}}\right]$. The corresponding mean square error (MSE) is

$$\epsilon_{\mathrm{MMSE}} = \mathrm{tr}\left[\left(\mathbf{R}_\mathbf{h}^{-1}+\frac{1}{\sigma_w^2}\Phi_\mathbf{b}^{\mathrm{H}}\Phi_\mathbf{b}\right)^{-1}\right]. \quad (24)$$

If $\mathbf{R_h}$ is diagonal, the MSE in (24) is minimized by designing the product $\Phi_\mathbf{b}^H\Phi_\mathbf{b}$ to be diagonal [24, App. I]. This can be seen intuitively by noting that $\Psi \triangleq \Phi_\mathbf{b}^H\Phi_\mathbf{b}$ is positive definite, and by the Hadamard inequality the diagonal terms $[\Psi]_{i,i}$ are at a maximum when Ψ is diagonal; thus, the trace of the inverse is minimized under the same condition.

The formula for the capacity of a channel is a function of its random channel gains, and does not traditionally account for effects of channel estimation. However, poor channel estimation will surely result in lower achievable rates. Note that $\mathbf{y}_b$ is not a function of the estimation error. To incorporate the effect channel estimation error has on communication rates, we use the lower bound capacity from [6]

$$\underline{C} \triangleq \frac{1}{N}\mathrm{E}\left[\log\det\left(\sigma_w^2\mathbf{I}+\bar{P}_d\mathbf{R}_\mathbf{v}^{-1}\hat{\mathbf{H}}_\mathbf{d}\hat{\mathbf{H}}_\mathbf{d}^{\mathrm{H}}\right)\right] \text{ bits/s/Hz} \quad (25)$$

where $\hat{\mathbf{H}}_\mathbf{d}$ is the estimate of $\mathbf{H_d}$, $\mathbf{R_v} \triangleq \bar{P}_d\mathrm{E}[\tilde{\mathbf{H}}_\mathbf{d}\tilde{\mathbf{H}}_\mathbf{d}^{\mathrm{H}}]+\sigma_w^2\mathbf{I}$, and $\tilde{\mathbf{H}}_\mathbf{d} \triangleq \hat{\mathbf{H}}_\mathbf{d}-\mathbf{H_d}$. The expectation in (25) is again over random channel realizations $\mathbf{H_d}$, but also the noise $\mathbf{w_d}$; the estimates $\hat{\mathbf{H}}_\mathbf{d}$, and hence error matrices $\tilde{\mathbf{H}}_\mathbf{d}$, are functions of $\mathbf{H_d}$ and noise $\mathbf{w_d}$. At high SNR, the LMMSE estimate that results in maximal lower bound on ergodic channel capacity is attained [6] if:

C1: all data subblocks have the same number of symbols (i.e., $\bar{N}_1^d=\bar{N}_2^d=\ldots=\bar{N}_M^d \triangleq \bar{N}_d$);
C2: all training subblocks have only one nonzero symbol (i.e., $\bar{N}_1^b=\bar{N}_2^b=\ldots=\bar{N}_M^b \triangleq \bar{N}_b=2L+1$) and are allocated equal power (i.e., $\|\mathbf{b}_m\|^2=P_b/M$ for each $m\in[1,M]$), and
C3: the number of subblocks equals the number of Doppler bases (i.e., $M=Q+1$).

Thus, for *optimal* training, the transmit vector is of the form

$$\mathbf{u}=[\mathbf{d}_1^{\mathrm{T}}\ \mathbf{0}_{1\times L}^{\mathrm{T}}\ b\ \mathbf{0}_{1\times L}^{\mathrm{T}}\ \ldots\ \mathbf{d}_M^{\mathrm{T}}\ \mathbf{0}_{1\times L}^{\mathrm{T}}\ b\ \mathbf{0}_{1\times L}^{\mathrm{T}}]^{\mathrm{T}} \quad (26)$$

where $b=P_b/M$.

2) *Least-Squares (LS) Channel Estimation:* The LMMSE channel estimate assumes that the channel covariance matrix is known. There are instances, such as a sudden shift in environment, where this assumption is unwarranted. We can lower bound the achievable communication rates by removing the known channel statistics and using data-driven LS channel estimation. The LS channel estimate is [25]

$$\hat{\mathbf{h}}_{\mathrm{LS}} = \left(\Phi_\mathbf{b}^{\mathrm{H}}\Phi_\mathbf{b}\right)^{-1}\Phi_\mathbf{b}\mathbf{y}_\mathbf{b}. \quad (27)$$

Note that $\hat{\mathbf{h}}_{\mathrm{LS}}\sim\mathcal{CN}(\mathbf{h},\sigma_w^2\left(\Phi_\mathbf{b}^{\mathrm{H}}\Phi_\mathbf{b}\right)^{-1})$, so the LS estimation error is

$$\epsilon_{\mathrm{LS}} \triangleq \mathrm{tr}\left[\sigma_w^2\left(\Phi_\mathbf{b}^{\mathrm{H}}\Phi_\mathbf{b}\right)^{-1}\right] \quad (28)$$

$$= \sum_{l=0}^{(Q+1)(L+1)}\frac{1}{\left[\frac{1}{\sigma_w^2}\Phi_\mathbf{b}^{\mathrm{H}}\Phi_\mathbf{b}\right]_{l,l}} \quad (29)$$

$$\geq \sum_{l=0}^{(Q+1)(L+1)}\frac{1}{\left[\mathbf{R}_\mathbf{h}^{-1}\right]_{l,l}+\left[\frac{1}{\sigma_w^2}\Phi_\mathbf{b}^{\mathrm{H}}\Phi_\mathbf{b}\right]_{l,l}} = \epsilon_{\mathrm{MMSE}} \quad (30)$$

where (29) follows from conditions C1-3, and the inequality in (30) follows from the fact that $\mathbf{R_h}$ is a covariance matrix with all diagonal elements greater than or equal to zero. There is no direct link between LS estimation error in (28) and LS estimate (27) and the capacity bound in (25). However, since the lower bound proof in [6, App. C] relies on an assumption of high SNR, and since at high SNR the LMMSE and LS estimates converge, the effect of the channel estimates on the lower bound capacity is negligible. Thus, minimizing $\mathbf{R_v}$ (by plugging in the errors of LS estimates) can maximize the lower bound capacity with LS approximation.

B. Radar Target Detection

We can decompose the pilot portion of the received signal in (16) into its clutter and target components

$$\mathbf{y_b} = \Phi_\mathbf{b}\mathbf{s}+\Phi_\mathbf{b}\mathbf{c}+\mathbf{w_b} \quad (31)$$

$$= \Phi_\mathbf{b}\mathbf{s}+\tilde{\mathbf{w}} \quad (32)$$

where $\mathbf{c} = [c(0) \dots c(L)]^{\mathrm{T}}$, $\mathbf{s} = [s(0) \dots s(L)]^{\mathrm{T}}$, and $\tilde{\mathbf{w}} \triangleq \Phi_{\mathbf{b}}\mathbf{c} + \mathbf{w}_{\mathbf{b}}$ with covariance

$$\mathbf{R}_{\tilde{\mathbf{w}}} = \mathrm{E}\left[(\Phi_{\mathbf{b}}\mathbf{c} + \mathbf{w}_{\mathbf{b}})(\Phi_{\mathbf{b}}\mathbf{c} + \mathbf{w}_{\mathbf{b}})^{\mathrm{H}}\right]. \tag{33}$$

The detection problem then becomes

$$\mathcal{H}_0 : \mathbf{y}_{\mathbf{b}} = \tilde{\mathbf{w}} \tag{34}$$

$$\mathcal{H}_1 : \mathbf{y}_{\mathbf{b}} = \Phi_{\mathbf{b}}\mathbf{s} + \tilde{\mathbf{w}} \tag{35}$$

where the null hypothesis $\mathcal{H}_0$ represents the case where the received signal is clutter and noise only, and the alternative hypothesis $\mathcal{H}_1$ represents the presence of a target in clutter and noise. The Neyman–Pearson [26] detector decides $\mathcal{H}_1$ if the likelihood ratio test statistic exceeds a threshold

$$\gamma < \frac{p(\mathbf{y}|\mathcal{H}_1)}{p(\mathbf{y}|\mathcal{H}_0)}. \tag{36}$$

The detection performance when $\mathbf{s}$ is a known vector is easily found. However, from (7) it is clear that both the magnitude and the phase of the target return are unknown quantities. Whether or not the target magnitude $|\alpha(k)|$ is known has no effect on the structure of the detector; thus $|\alpha(k)|$ can be assumed known when deriving the detector, and accounted for using noncoherent integration [5, Ch. 6.2]. Knowing the phase $\angle\alpha(k)$ requires knowing the range to the target to incredible precision, since only a change of a quarter wavelength causes a complete reversal of phase in the returned signal. To account for this unknown phase, we use envelope detection assuming the magnitude is a known vector $\mathbf{m}$, and examine detection performance when $\mathbf{s} = \mathrm{e}^{j\theta}\mathbf{m}$, where $\theta \sim \mathcal{U}(0, 2\pi)$. The probability of detection for a Swerling I target is then [5, Ch. 6.3]

$$P_D = \exp\left[\frac{-\mathcal{T}'}{1+\chi}\right] \tag{37}$$

where $\mathcal{T}'$ is a threshold selected to give a desired probability of false alarm P_{FA} (i.e., $\mathcal{T}' = \mathcal{T}'(P_{\mathrm{FA}})$), and χ is the target signal to interference (clutter) plus noise ratio (TINR). Note that (37) is monotonic in χ. Using the Cauchy–Schwarz inequality, it is straightforward to show that the maximum TINR achievable is

$$\chi = \mathbf{s}^{\mathrm{H}}\Phi_{\mathbf{b}}^{\mathrm{H}}\mathbf{R}_{\tilde{\mathbf{w}}}^{-1}\Phi_{\mathbf{b}}\mathbf{s}. \tag{38}$$

If we assume $\mathbf{R}_{\tilde{\mathbf{w}}}$ is diagonal (i.e., each $c_q(k, l)$ is independent), optimal training signals are used, and the target is a point target, then, using the definitions in (5)–(7), the TINR in (38) simplifies to

$$\chi = \frac{\sigma_\alpha^2 P_b}{\sigma_{c_{q_0}(k,l_0)}^2 P_b + \sigma_w^2}. \tag{39}$$

Note that (39) is monotonic in P_b. Thus, both the radar performance and the communications performance are tied to the proportion of transmit power allocated to channel estimation. The performance of both systems is considered jointly over P_b in the next section.

Since the two systems fully cooperate, it is theoretically possible to use the entire transmit waveform for detection after the communication system has finished decoding the data. However, we omit the use of data symbols for detection for the following reasons. First, this paper focuses on low-complexity linear approach. Second, the latency resulting from decoding the data before sending to the radar system would likely be unsatisfactory in many radar applications. Finally, feedback paths in hardware design are generally frowned upon due to timing and synchronization issues. We, therefore, leave detection using the entire waveform for future work.

C. Optimality Considerations

When finding an optimal solution, it is important to make explicit *in what sense* a solution is optimal. At the outset we adopted a training scheme shown to be optimal for communication systems; however, other training schemes may perform better for radar purposes because of the higher energy required. In what follows, we will continue to refer to the training scheme that achieves the LMMSE channel estimate as "optimal," while acknowledging that other training schemes may perform better on other metrics.

1) *Suboptimal Training:* For radar, probability of detection is a function of the total amount of reflected energy from the target, rather than the peak transmission power. Using the repeated single pulse optimal for communication, the peak power required for detection may greatly exceed the capability of amplifiers standard in radar hardware. Instead, the optimal signaling structure can be relaxed by allowing the total power needed to detect targets to be distributed over multiple consecutive training symbols.

As an example, we use Barker sequences [27] as a suboptimal waveform in our simulations. Barker sequences have two properties that make them a good choice for spreading power allocated to training over multiple symbols: 1) they are constant modulus, facilitating amplification, and 2) they exhibit autocorrelation sidelobes with maximum power inversely proportional to the code length, and therefore, provide a good approximation to a single-pulse train. However, Barker sequences are not as Doppler-tolerant and can have more spectral splatter as other coded waveforms for pulse compression [28]. The Barker sequences we employ here differ from Barker coding traditionally used in radar for pulse compression in that the chip duration is equivalent to the symbol duration, so the bandwidth is equivalent for any Barker (as well as optimal) sequence length. Rather than increasing bandwidth, sequences of increasing length instead produce longer overall block lengths. This allows fair comparison of differing sequence lengths, since the capacity lower bound is normalized by block length.

Note that the data symbol subblocks are unaffected by the use of Barker sequences, since we apply them only to training. However, using Barker sequences the lower bound on capacity in (25) is likely to suffer, since a larger percentage of *time* is spent on training. Fig. 3 shows the effect of using a Barker-7 sequence for training. Compared with the optimal training in Fig. 2, the instantaneous power of each training symbol is reduced to a level more comparable

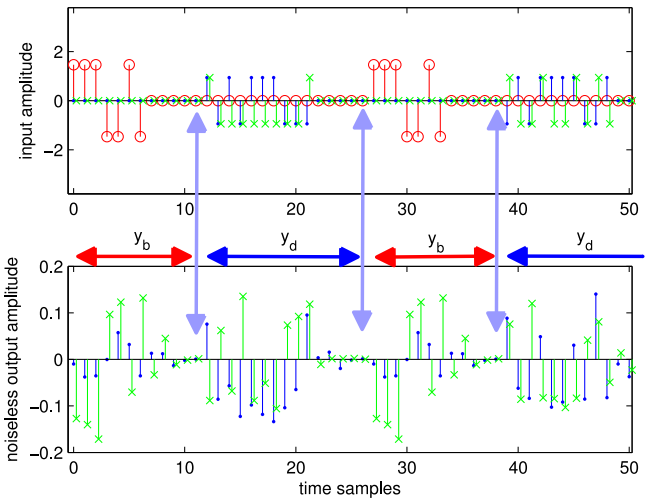


Fig. 3. Barker-sequence counterpart to Fig. 2. Here, a Barker sequence of length 7 is shown, and all other parameters are held constant. Note the increased length of the received training signal compared with the single-pulse optimal training used in Fig. 2.

to that of the data symbols, but the training sequence is six symbols longer for each subblock. In addition to total time spent transmitting training symbols, suboptimal training also affects channel estimation error, since Barker sequences have good (but not ideal, like a single optimal training pulse) autocorrelation properties. Using Barker sequences also affects the probability of detection. From (39), it is apparent that sequence length has no effect on probability of detection in the radar system if the clutter plus noise covariance is uncorrelated, and so long as $\Phi^{\mathrm{H}}\Phi$ is diagonal. Despite their desirable autocorrelation properties, however, Barker sequences do not produce a strictly diagonal $\Phi^{\mathrm{H}}\Phi$, so we must revert to (38) when calculating TINR. We expect that the use of Barker sequences is likely to have a more detrimental effect on the lower bound on capacity than on probability of detection, since capacity is affected both by time and quality of estimation, whereas the $\Phi^{\mathrm{H}}\Phi$ will still be *nearly* diagonal.

2) *Sampling Doppler Outside the Clutter Region:* For both communication and radar systems, keeping Q to a minimum avoids unnecessary overhead. In communication-only systems, the value is often set to

$$Q = \lceil f_{\max} N T_s \rceil \tag{40}$$

so that the highest Dopper frequency sampled coincides with the maximum clutter Doppler component, and the Doppler samples are evenly spaced on the interval $[-f_{\max}, f_{\max}]$. Thus, in many communication systems Q is determined by the speed of the transmitter and/or the receiver. Here, we use the block length N to be the fast Fourier transform (FFT) size used in the BEM. It has been noted that using a denser FFT and higher Q values can increase the accuracy of the BEM [29]. However, to estimate $Q+1$ Doppler bases we need $Q+1$ subblocks, so increasing Q results in progressively longer block lengths and processing complexity.

From a radar perspective, choosing a larger Q may also be required to identify a target producing a Doppler shift that exceeds the maximum Doppler shift produced by clutter. Say, for example, that we are interested in a target with Doppler shift $(1+\delta) f_{\max}$, for any $\delta > 0$. Then if we define Q as in (40), our target Doppler will alias. We might then be tempted to accept longer block lengths and increase Q to avoid aliasing, but with the exception of the target Doppler, all Doppler values outside the clutter region are modeled (in Jakes' model) as exactly zero. Thus, $\mathbf{h}$ would become a *degenerate* Gaussian random vector. While these are interesting issues to be solved, they lie beyond the scope of this paper. However, in practice, for joint radar and communication operations, one would choose Q to describe the fastest scatterer, whether it is clutter or a hypothesized target. Further, in practice, one might perform two different estimation problems on the data: one assuming there are no targets in the data and one assuming a target at a hypothesized location. In this case, there are two potential lower bounds on capacity that would result: one assuming a target is in the channel and one assuming one is not in the channel. However, in practice, the two lower bounds are likely to be very similar since there is rarely significant energy from a target in excess to the total combined energy from clutter. In Section IV, we consider cases where the target is located within the Doppler spread of the clutter; the study of targets outside of this spread is left for future work.

IV. NUMERICAL EXAMPLES

We simulated the joint system using 60 Monte Carlo trials. We varied the total block-duration power P in (17) from 20 W to 120 kW, and percentage of power allocated to data over $0 < \rho < 1$. The delay power in our channel was modeled as decaying exponentially as $A_h(0;\tau) = \beta_1 \exp(-\tau/\beta_2)$, with β_1 and β_2 representing the initial power and power decay parameters, respectively. For an exponential decay model with average delay spread μ_τ and rms delay spread σ_τ, we have $\mu_\tau = \sigma_\tau = \beta_2$. For our simulations we chose β_1 to be $1/10$ of the line of sight power at the receiver, and $\beta_2 = 2\,\mu$s. Fig. 4 shows the Doppler power spectrum, normalized such that the total Doppler power sums to 1, and the power delay profile (PDP), normalized by the line of sight power.

We used an initial block size of $N = 2048$, with channel length $L = 30$ and Doppler resolution $Q = 16$. For our system parameters, we used a carrier frequency of 2 GHz with a bandwidth of 1 MHz; transmit and receive gains of 5 dB; system losses of 2 dB; and a noise figure of 3 dB. Suboptimal training was implemented by simply replacing each scalar b in (26) with an appropriately power-scaled Barker sequence of desired length. For example, if a Barker-4 sequence was used, each b would be replaced with $\frac{\tilde{P}_b}{4}[+1,+1,-1,+1]$. To keep the block length N relatively constant, we first set the number of subblocks $M = Q+1$ and then selected N_d to yield the smallest $N_d + M(2L + L_B) \geq N$, where L_B is the Barker sequence length.

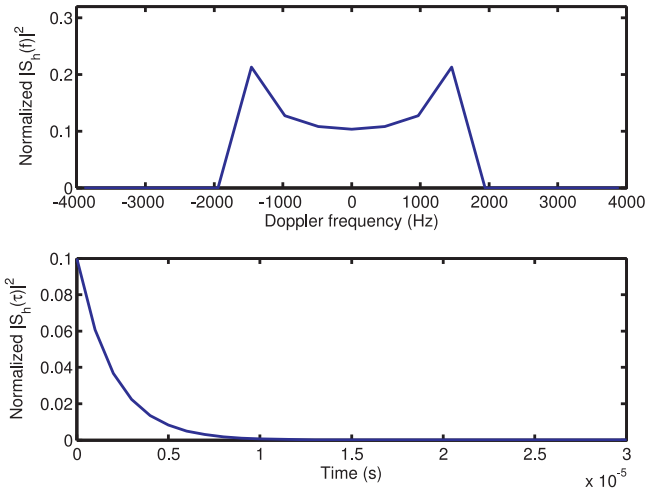


Fig. 4. Clutter Doppler power (top) and PDP (bottom).

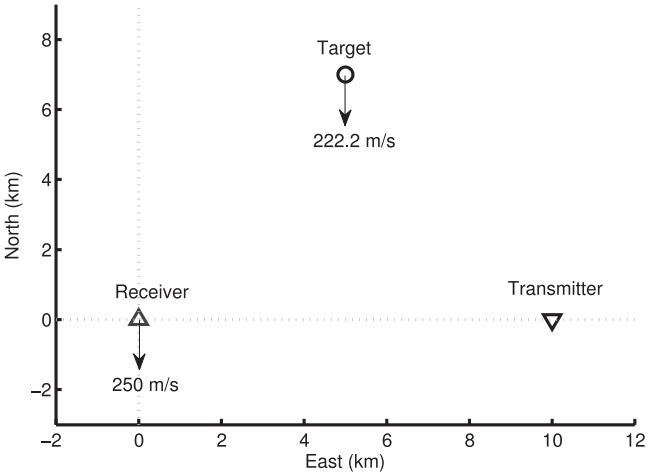


Fig. 5. Simulation scenario.

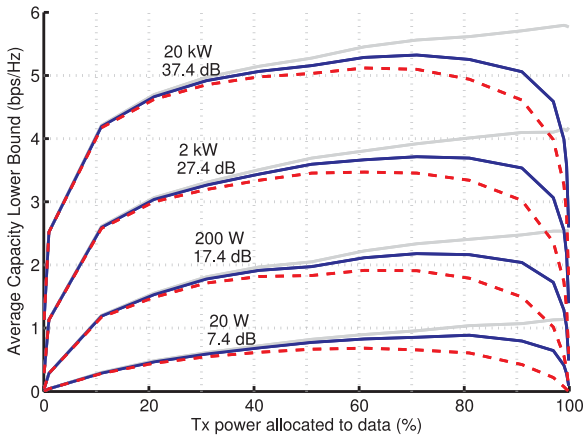


Fig. 6. Capacity lower bound as a function of power allocation for various total transmission powers and corresponding receiver SNR per symbol. MMSE: solid blue; LS: dotted red; upper bound: solid light-gray.

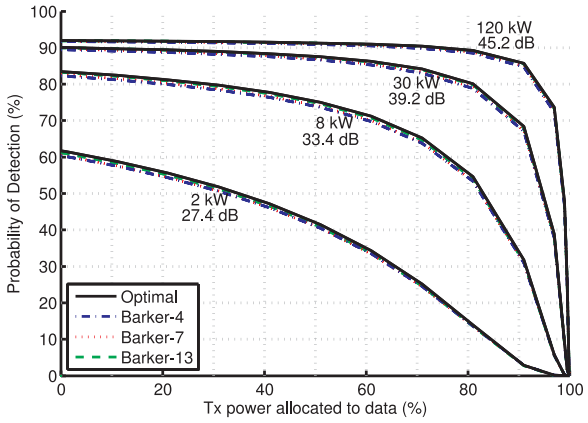


Fig. 7. Probability of detection as a function of power allocation for various total transmission powers and corresponding receiver SNR per symbol.

The transmitter and receiver were positioned at (10, 0) and (0, 0) km, respectively, and the target was positioned at (5, 7) km, as shown in Fig. 5. We assumed a stationary transmitter, and mobile receiver and target with instantaneous velocity vectors (0, −900) and (0, −800) km/hr, respectively. With these velocities and positions, the bistatic Doppler shift of the target lies within the total Doppler spread arising from clutter. This allows us to adopt Jakes' model [30] for the channel Doppler spread and find the MMSE channel estimate.

The effect of the power allocation ρ is shown in Fig. 6 for the capacity lower bound. The communication system functions well in the lower total power range. Note that the power values given represent fundamentally different quantaties: The value given in W or kW is the total power transmitted over the entire block duration; the value in dB is the SNR per symbol at the receiver. The figure shows three curves: the light gray curve is the upper bound from (22) assuming the channel is known perfectly at the receiver and does not require estimation. The solid blue and dotted red curves show the effect of imperfect MMSE and LS estimation, respectively. The left portions of the curves are *power* limited in that there is negligible difference between performance with perfect channel knowledge compared with the estimated channel performance, and rate increases are achieved only by allocating more power to the data signal. The right portions of the curves are *estimation* limited in that the effect of channel estimation becomes salient. These trends appear largely insensitive to overall transmit power across the moderate to high receiver SNR cases tested.

The radar detection probability shown in Fig. 7 is monotonic in total power allocated to training. For low total power, allocating more power to training greatly increases detection probability, whereas for high total power the benefit of allocating more power to training saturates quickly. Note that the overall power expenditures necessary for the radar system are much higher than those needed for reliable communication. From (37) and (39), it is also clear that detection probability is a function only of the channel realization and total power P_b allocated to training, and thus, is unaffected by method of channel estimation so long as optimal training sequences are transmitted. The detection probabilities for Barker sequences show slight performance

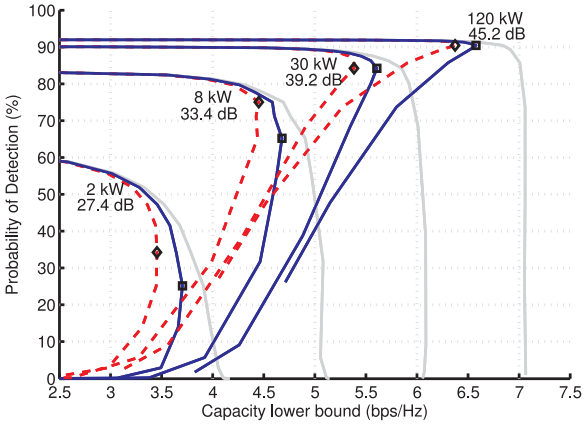


Fig. 8. Joint radar-communication system performance for various total transmission powers and corresponding receiver SNR per symbol. MMSE: solid blue; LS: dotted red; upper bound: solid light-gray.

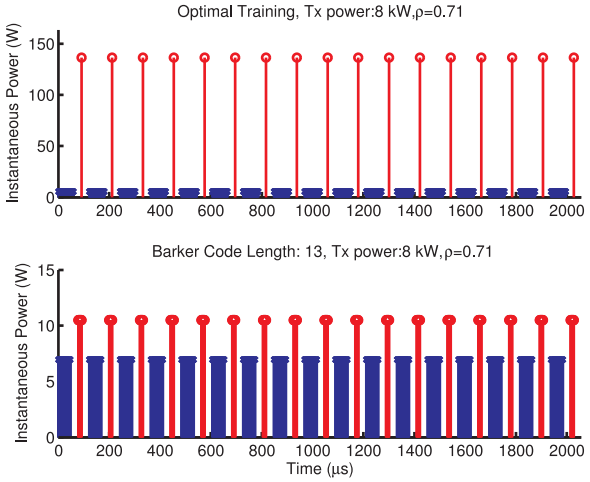


Fig. 9. Comparison of training symbol instantaneous power levels using a single pulse per subblock optimal for communications systems (top) and after spreading power among multiple training symbols using a Barker sequence (bottom).

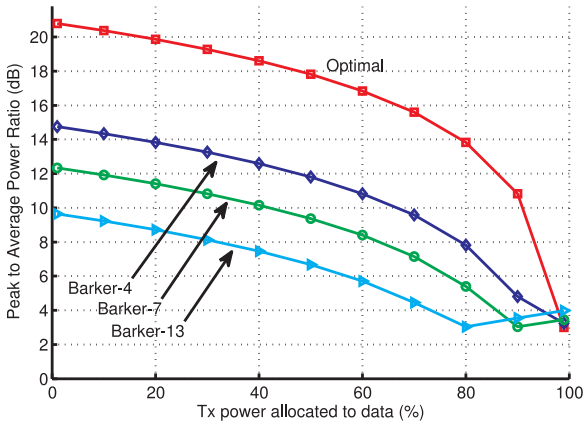


Fig. 10. PAR for optimal training and three Barker code training signals.

degradation, but do not significantly change the shapes of the curves. Note that longer Barker sequences actually perform closer to optimal training than shorter ones, since the autocorrelation sidelobes decrease with sequence length.

The joint communication-radar performance is shown in Fig. 8. The curves shown are parameterized over $0 < \rho < 1$. Values marked by squares (diamonds) represent the capacity-maximizing values for the MMSE (LS) case. For the MMSE case, the maximizing values for the two lower-power curves are obtained by setting $\rho = 0.71$; for the two higher power curves, the maximizing $\rho = 0.61$. For LS, all maximizing values arise from $\rho = 0.61$. These maximizing values show the diminishing returns of incorporating prior knowledge of channel statistics. We note that the actual achievability region is obtained by dropping a vertical line from the point of maximized capacity lower bound, since anywhere below that point is (unadvisably) achievable simply by selecting an inferior detection method. When total power grows large, inflection points begin to appear in the performance curves. These points indicate the allocation values at which further allocating power away from training becomes more detrimental to detection probability than to the capacity lower bound.

Fig. 10 shows the PAPR for blocks using optimal training versus Barker coded training signals. For optimal training, PAPR is monotonically decreasing with power allocated to data. Barker sequences reduce PAPR by up to 10 dB. Codes with length 7 or 13 also achieve the lowest PAPR at lower data allocation values.

Fig. 9 shows the time-domain discrete baseband transmitted signal for optimal training versus a Barker code of length 13 for 8 kW total transmit power and $\rho = 0.71$. Even though the power and chosen allocation level are the same for each case, the dynamic range of the optimal training signal is far higher than that of the Barker coded signal, since the optimal training signal concentrates all its power into a single symbol while the Barker code spreads the power over many training symbols. Spreading training energy over multiple symbols could greatly facilitate the amplification process in practical systems.

The effect of using Barker sequences as a suboptimal training scheme is shown in Figs. 11 and 12 for the MMSE and LS cases, respectively. In both cases, there is a penalty paid for increasing the length of the nonzero training signal. The penalty at low SNR is negligible and becomes more substantial as overall transmitted power grows. However, the gaps from the Barker curves to the optimal training curve appear relatively insensitive to allocation level. This signifies that lower bounds on capacity for higher code lengths become *time* limited in that the penalty appears to have little relation to channel estimation quality.

To investigate the potential impact on performance due to clutter, we define the target-to-clutter ratio (TCR) as ratio of the variance of $c_{q_0}(k, l_0)$ to σ_α^2. Thus, the TCR is the total energy received due to clutter that is located in the delay bin of the target divided by the noise power. We repeat the experiment using the optimal waveform and $P = 8$ kW, but we vary the decay rate, β_2, as 2, 2.25, and 2.5 μs, which results in an TCR of 22.5, 17.0, and 12.5 dB, respectively. The results in Fig. 13 demonstrate that target detection performance decreases significantly as the TCR is

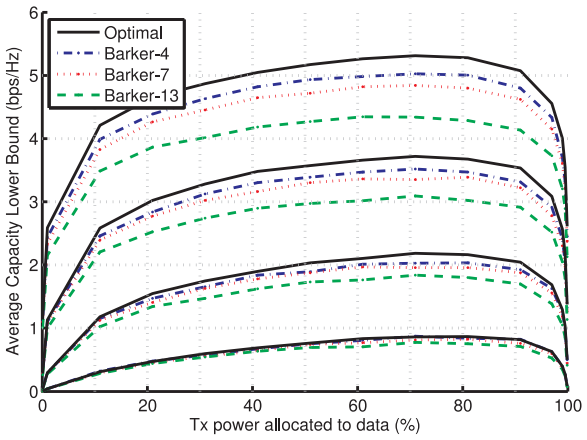


Fig. 11. Capacity lower bound using Barker training sequences compared with the optimal training scheme for MMSE channel estimation.

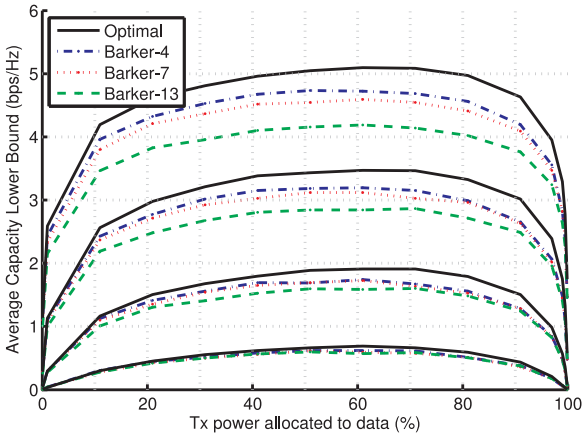


Fig. 12. Capacity lower bound using Barker training sequences compared with the optimal training scheme for LS channel estimation.

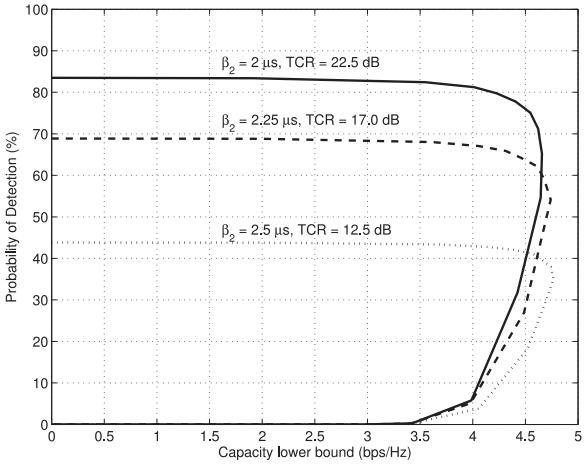


Fig. 13. Capacity lower bound using Barker training sequences with varying TCR.

increased. It should also be noted that increasing β_2 results in a higher per-symbol SNR of the communications signal, which increases the capacity lower bound; however, the gains are relatively modest.

V. CONCLUSION

We have presented an analysis framework for a cooperative radar-communication system operating simultaneously in the same frequency band. Our design leverages decoupled training and data to yield low-complexity linear processing for the communication system, and assumes a point-target model, which yields a simple detector structure for the radar system. The specific performance of each system depends on the power allocation, but mutually successful coexistence is possible. Optimal training signals are shown to produce the greatest lower bound on communication rate as well as the maximum probability of detection. Given that amplification is likely to impose real-world constraints on instantaneous transmitted power, we have shown that Barker training sequences are a viable option. The penalty for using Barker sequences is most apparent in the lower bound on achievable communication rate, and less significant for radar detection. Despite being suboptimal, the longer Barker sequences have the advantage of greatly reducing PAPR and retaining near-optimal detection performance. Future work can extend the framework to more more complex scenarios, specifically addressing coupled data and symbol detection with nonlinear processing to compare performance gains with increased complexity, expanding beyond the simplistic point target model, and addressing the possibility of multiple targets.

ACKNOWLEDGMENT

The authors would like to thank Dr. J. Barry and Dr. R. Causey for their advice. In particular, we would like to thank Dr. R. Baxley for his assistance in framing the problems tackled in this paper. The views, opinions, and/or findings contained in this article are those of the author(s) and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. government. Distribution Statement A: Approved for Public Release, Distribution Unlimited.

REFERENCES

[1] FCC Spectrum Policy Task Force Report of the spectrum efficiency working group Nov. 2002. [Online]. Available: https://transition.fcc.gov/sptf/files/SEWGFinalReport_1.pdf

[2] Q. Zhao and A. Swami
A survey of dynamic spectrum access: Signal processing and networking perspectives
In *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Apr. 2007, vol. 4, pp. IV–1349–IV–1352.

[3] J. Cavers
An analysis of pilot symbol assisted modulation for Rayleigh fading channels [mobile radio]
IEEE Trans. Veh. Technol., vol. 40, no. 4, pp. 686–693, Nov. 1991.

[4] Y. Chen and N. Beaulieu
Optimum pilot symbol assisted modulation
IEEE Trans. Commun., vol. 55, no. 8, pp. 1536–1546, Aug. 2007.

[5] M. A. Richards
Fundamentals of Radar Signal Processing. New York, NY, USA: McGraw-Hill, 2005.

[6] X. Ma, G. Giannakis, and S. Ohno
Optimal training for block transmissions over doubly selective wireless fading channels
IEEE Trans. Signal Process., vol. 51, no. 5, pp. 1351–1366, May 2003.

[7] J. Johnson, C. Baker, H. Wang, L. Ye, and C. Zhang
Assessing the potential for spectrum sharing between communications and radar systems in the l-band portion of the RF spectrum allocated to radar
In *Proc. Int. Conf. Electromagn. Adv. Appl.*, Aug. 2014, pp. 331–334.

[8] S. Surender, R. Narayanan, and C. Das
Performance analysis of communications & radar coexistence in a covert UWB OSA system
In *Proc. IEEE Glob. Telecommun. Conf.*, Dec. 2010, pp. 1–5.

[9] L. Wang, J. McGeehan, C. Williams, and A. Doufexi
Application of cooperative sensing in radar-communications coexistence
IET Commun., vol. 2, no. 6, pp. 856–868, Jul. 2008.

[10] H. Deng and B. Himed
Interference mitigation processing for spectrum-sharing between radar and wireless communications systems
IEEE Trans. Aerosp. Electron. Syst., vol. 49, no. 3, pp. 1911–1919, Jul. 2013.

[11] D. Bliss
Cooperative radar and communications signaling: The estimation and information theory odd couple
In *Proc. IEEE Radar Conf.*, May 2014, pp. 0050–0055.

[12] C. Rossler, E. Ertin, and R. Moses
A software defined radar system for joint communication and sensing
In *Proc. IEEE Radar Conf.*, May 2011, pp. 1050–1055.

[13] B. Donnet and I. Longstaff
Combining MIMO radar with OFDM communications
In *Proc. 3rd Eur. Radar Conf.*, Sep. 2006, pp. 37–40.

[14] C. Sturm, T. Zwick, and W. Wiesbeck
An OFDM system concept for joint radar and communications operations
In *Proc. IEEE 69th Veh. Technol. Conf.*, Apr. 2009, pp. 1–5.

[15] Y. Sit, C. Sturm, and T. Zwick
One-stage selective interference cancellation for the OFDM joint radar-communication system
In *Proc. 7th German Microw. Conf.*, Mar. 2012, pp. 1–4.

[16] S. Liu, Z. Huang, and W. Zhang
A power-efficient radar waveform compatible with communication
In *Proc. Int. Conf. Commun. Circuits Syst.*, vol. 2, Nov. 2013, pp. 1–4.

[17] C. Sturm and W. Wiesbeck
Waveform design and signal processing aspects for fusion of wireless communications and radar sensing
In *Proc. IEEE*, vol. 99, no. 7, pp. 1236–1259, Jul. 2011.

[18] N. J. Willis and H. D. Griffiths Eds.
Advances in Bistatic Radar. Raleigh, NC, USA: SciTech Publishing, 2007.

[19] N. J. Willis
Bistatic Radar. Raleigh, NC, USA: SciTech Publishing, 2005.

[20] A. Goldsmith
Wireless Communications. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[21] M. Valenti and B. Woerner
Iterative channel estimation and decoding of pilot symbol assisted turbo codes over flat-fading channels
IEEE J. Sel. Areas Commun., vol. 19, no. 9, pp. 1697–1705, Sep. 2001.

[22] H.-J. Su and E. Geraniotis
Low-complexity joint channel estimation and decoding for pilot symbol-assisted modulation and multiple differential detection systems with correlated Rayleigh fading
IEEE Trans. Commun., vol. 50, no. 2, pp. 249–261, Feb. 2002.

[23] S. M. Kay
Fundamentals Statistical Signal Processing, Volume I: Estimation Theory. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

[24] S. Ohno and G. Giannakis
Capacity maximizing MMSE-optimal pilots for wireless OFDM over frequency-selective block Rayleigh-fading channels
IEEE Trans. Inf. Theory, vol. 50, no. 9, pp. 2138–2145, Sep. 2004.

[25] L. L. Scharf
Statistical Signal Processing: Detection, Estimation, and Time Series Analysis. Reading, MA, USA: Addison-Wesley, 1991.

[26] S. M. Kay
Fundamentals Statistical Signal Processing, Volume II: Detection Theory. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

[27] S. Golomb and R. Scholtz
Generalized barker sequences
IEEE Trans. Inf. Theory, vol. 11, no. 4, pp. 533–537, Oct. 1965.

[28] J. Taylor and H. Blinchikoff
Quadriphase code—A radar pulse compression signal with unique characteristics
IEEE Trans. Aerosp. Electron. Syst., vol. 24, no. 2, pp. 156–170, Mar. 1988.

[29] O. Rousseaux, G. Leus, and M. Moonen
Estimation and equalization of doubly selective channels using known symbol padding
IEEE Trans. Signal Process., vol. 54, no. 3, pp. 979–990, Mar. 2006.

[30] W. C. Jakes Ed.
Microwave Mobile Communications. Hoboken, NJ, USA: Wiley, 1975.

Andrew D. Harper (S'07) received the B.S. degree in engineering with electrical specialty from Colorado School of Mines, Golden, CO, USA, and the M.S. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2008 and 2010, respectively. He is currently working toward the Ph.D. degree in electrical engineering at the Georgia Institute of Technology.

He spent summer 2007 as a Technical Intern in signal and image processing at Ball Aerospace in Dayton, OH, USA, and summer 2012 as a Graduate Intern at the Naval Research Laboratory in Washington, D.C., USA. His research interests include statistical signal processing for wireless communications and radar, information and coding theory, and physical-layer security in wireless communications.

Jeremy T. Reed (M'09) received the B.E., M.S., and Ph.D. degrees in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2002, 2005, and 2010, respectively.

From 2002 to 2003, he worked as a Design Engineer with Moog, Inc. in the Aircraft Group. He has also worked for Johnson & Johnson, Lockheed-Martin, and VT Silicon. He currently works in the Sensors and Electromagnetic Applications Laboratory, Georgia Tech Research Institute in Smyrna, GA, USA. His major research interests include adaptive radar signal processing, shared spectrum technologies for radar and communications, automatic target recognition, modeling and simulation, and decision support for sensor applications.

Dr. Reed served as a Chair for the Atlanta Chapter of the IEEE Signal Processing Society from 2011 to 2013.

Jonathan L. Odom (S'11–M'14) received the B.S.E., M.S., and Ph.D. degrees in electrical and computer engineering from Duke University, Durham, NC, USA, in 2009, 2011, and 2013, respectively.

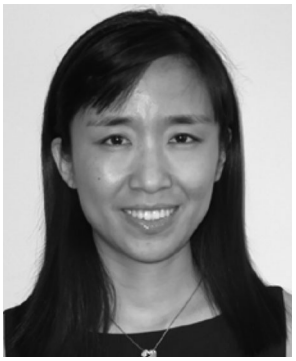
From 2009 to 2013, he was a Research Assistant in the Sensor Array and Multipath Lab at Duke University. He spent summer 2012 at The Johns Hopkins Applied Physics Laboratory, Baltimore, MD, USA. In 2014, he was an Adjunct Assistant Professor with Duke University. He is currently with Georgia Tech Research Institute, Atlanta, GA, USA. His research interests include array and statistical signal processing with applications to sonar and radar.

Dr. Odom is a member of the Acoustical Society of America, Eta Kappa Nu, and Tau Beta Pi. He received the Best Young Presenter Paper in Signal Processing award at the 166th Meeting of the Acoustical Society of America in 2013.

Aaron D. Lanterman (M'91) received the B.A. degree in music, the B.S. degree in computer science, the B.S. degree in electrical engineering, in 1993, the M.S. degree, and the D.Sc. degree in electrical engineering from Washington University in St. Louis, St. Louis, MO, USA, 1995 and 1998, respectively.

He is an Associate Professor of electrical and computer engineering at the Georgia Institute of Technology, Atlanta, GA, USA, which he joined in the fall of 2001. In 2004, he was chosen to hold the Demetrius T. Paris Professorship, a special chaired position for the development of young faculty. His thesis focused on target recognition for infrared imagery as part of the multi-university U.S. Army Center for Imaging Science. After graduation, he joined the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign as a Postdoctoral Research Associate and then as a Visiting Assistant Professor, where he managed a project on radar systems exploiting "illuminators of opportunity" such as television and FM radio signals. His other research interests include target tracking, image reconstruction, music synthesis, and video game hardware.

Dr. Lanterman received the School of Electrical and Computer Engineering Richard M. Bass/Eta Kappa Nu Outstanding Junior Teacher Award in 2006, and was named a 2009 Hesburgh Award Teaching Fellow. He has served as an Associate Editor for radar for the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS.



Xiaoli Ma (F'15) received the B.S. degree in automatic control from Tsinghua University, Beijing, China, in 1998, the M.S. degree in electrical engineering from the University of Virginia, Charlottesville, VA, USA, in 2000, and the Ph.D. degree in electrical engineering from the University of Minnesota, Minneapolis, MN, USA, in 2003.

After receiving the Ph.D. degree, she joined the Department of Electrical and Computer Engineering, Auburn University, where she served as an Assistant Professor until December 2005. Since January 2006, she has been with the School of Electrical and Computer Engineering at Georgia Tech, Atlanta, GA , USA, where she is currently a Professor. Her research interests include networking and communications, including network performance analysis, transceiver designs for wireless time- and frequency-selective channels, channel estimation and equalization algorithms, carrier frequency synchronization for OFDM systems, performance analysis, and cooperative designs for wireless networks.

Dr. Ma received the Lockheed Martin Aeronautics Company Dean's Award for Teaching Excellence by the College of Engineering in 2009, and Outstanding Junior Faculty Award by the School of Electrical and Computer Engineering in 2010, at Georgia Tech. She is a Senior Area Editor for the IEEE SIGNAL PROCESSING LETTERS since 2014 and Elsevier Digital Signal Processing since June 2012, and has been an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS (2007–2009) and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2008–2013).