SEQUENTIAL ENERGY ALLOCATION STRATEGIES FOR CHANNEL ESTIMATION

Raghuram Rangarajan, Raviv Raich, and Alfred O. Hero III

Department of EECS, University of Michigan, Ann Arbor, MI 48109-2122, USA
{rangaraj, ravivr, hero}@eecs.umich.edu

ABSTRACT

The context of this paper is adaptive waveform design for estimating parameters of an unknown channel under average energy constraints. This paper focuses on the simpler problem of adaptive waveform-amplitude design for which we obtain interesting analytical results. We treat an N-step design problem where a fixed waveform can be transmitted into the channel N times with amplitudes that can be chosen as a function of past channel outputs. For $N = 2$ and a linear Gaussian channel model, we derive the optimal amplitude to transmit at the second step as a function of the first measurement. This adaptive 2-step energy allocation strategy yields a mean-squared error (MSE) improvement of at least 1.7dB relative to the optimal non-adaptive strategy. Motivated by the optimal two-step strategy we propose a suboptimal adaptive N-step strategy that can achieve an MSE improvement of more than 5dB for $N = 50$. Applications of our results to MIMO and inverse scattering channel models are discussed.

***Index Terms*—** Parameter estimation, adaptive control, energy allocation, maximum likelihood, MMSE.

1. INTRODUCTION

One of the important components in adaptive sensing is the need for energy management. Most applications are limited by peak power or average power. Hence it is important to consider energy limitations in waveform design problems. Most previous research has focussed on waveform design under peak power constraints [1, 2]. There has been little effort in adaptive energy management strategies that allocate different amounts of energy to the waveforms over time. In this paper, we find optimal sequential energy allocation strategies for a general class of estimation problems under an average power constraint and show performance gains over non-adaptive design strategies.

Measurement-adaptive estimation has countless number of important applications in a wide variety of areas such as communications and control, medical imaging, radar systems, system identification, and inverse scattering. By measurement-adaptive estimation we mean that one has control over the way measurements are made, e.g., through the selection of waveforms, projections, or transmitted energy. The standard solution for estimating parameters from adaptive measurements is the maximum likelihood (ML) estimator. For the case of classic linear Gaussian model, i.e., a Gaussian observation with unknown mean and known variance, it is well-known [3] that the ML estimator is unbiased and achieves the unbiased Cramér Rao lower bound (CRB). Many researchers have looked at improving parameter estimation performance by adding a small estimator bias to reduce the MSE. Stein showed that this leads to better estimators that achieve lower MSE than the linear least squares (LS) estimator for estimating the mean in a multivariate Gaussian distribution with dimension greater than two [4]. Other alternatives such as shrinkage estimator [5], Tikhonov regularization [6], and covariance shaping least squares (CSLS) estimator [7] have also been proposed in the literature. None of these approaches to improve performance incorporate the notion of sequential energy allocation in their work.

In this paper, we formulate a problem of adaptively selecting waveform amplitudes for estimating parameters of a linear Gaussian channel model under an average energy constraint over the waveforms and over the number of transmissions. Waveform amplitude design can be cast as sequential parameter estimation where a transmitted waveform is measured at a receiver after passing through a channel having unknown parameters. We first obtain closed-form expressions for the MSE of the optimal two-step sequential energy allocation strategy for a scalar parameter in a multivariate linear Gaussian model. We then extend these results to the case of vector parameters. Furthermore we provide an N-step sequential strategy which yields more than 5dB gain over non-adaptive methods. We conclude by providing applications to channel estimation and imaging. The results in this paper summarize the results of [8] and represent a significant extension of our previous paper [9].

2. PROBLEM SETTING FOR ESTIMATION

We denote vectors in $\mathbb{C}^M$ by boldface lower case letters and matrices in $\mathbb{C}^{M\times N}$ by boldface uppercase letters. The symbol $\|\cdot\|$ refers to the l_2-norm of a vector, i.e., $\|\mathbf{x}\| = \sqrt{\mathbf{x}^H\mathbf{x}}$, where $(\cdot)^H$ denotes the conjugate transpose. Let $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_M]^T$ be the M-element vector of unknown parameters, where $(\cdot)^T$ denotes the transpose. The problem of waveform design is to select the sequence of waveforms $\{\mathbf{x}_i\}_{i=1}^N$ in order to best estimate the parameters $\boldsymbol{\theta}$ in the model

$$\mathbf{y}_i = \mathbf{H}(\mathbf{x}_i)\boldsymbol{\theta} + \mathbf{n}_i, \quad i = 1, 2, \ldots, N, \tag{1}$$

where $\mathbf{H}(\mathbf{x}_i) = [\mathbf{h}_1(\mathbf{x}_i), \mathbf{h}_2(\mathbf{x}_i), \ldots, \mathbf{h}_M(\mathbf{x}_i)]$ is a known $K \times M$ matrix and N indicates the number of time steps. The T-element design vectors, $\{\mathbf{x}_i\}_{i=1}^N$ can depend on the past measurements: $\mathbf{x}_i = \mathbf{x}_i(\mathbf{y}_1, \ldots, \mathbf{y}_{i-1})$, where $\mathbf{y}_i$ is the i^{th} K-element received signal vector. The K-element noise vectors $\{\mathbf{n}_i\}_{i=1}^N$ are independent identically distributed (i.i.d) circularly symmetric complex Gaussian random variables with zero mean and variance σ^2 denoted by $\mathbf{n}_i \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$. When $\mathbf{H}(\mathbf{x})$ is linear in $\mathbf{x}$, we can write $\mathbf{h}_j(\mathbf{x}) = \boldsymbol{H}_j\mathbf{x}$, $j = 1, 2, \ldots, M$. In this case $\mathbf{H}(\cdot)$ is uniquely determined by the $K \times T$ matrices $\{\boldsymbol{H}_1, \boldsymbol{H}_2, \ldots, \boldsymbol{H}_M\}$. For the case of a scalar parameter θ_1, the measurements are

$$\mathbf{y}_i = \mathbf{h}_1(\mathbf{x}_i)\theta_1 + \mathbf{n}_i, \quad i = 1, 2, \ldots, N. \tag{2}$$

We evaluate the performance of the measurement scheme in terms of the MSE of the ML estimator of θ_1 given $\{\mathbf{y}_i\}_{i=1}^N$ subject to the

This research was partially supported by ARO-DARPA MURI Grant #DAAD19-02-1-0262

energy constraint, $\mathrm{E}\left[\sum_{i=1}^{N}\|\mathbf{x}_i\|^2\right] \leq E_0$, where E_0 is the total available energy and $\mathrm{E}[\cdot]$ denotes the statistical expectation. The ML estimator of θ_1 for the N-step procedure is given by

$$\hat{\theta}_1^{(N)} = \frac{\sum_{i=1}^{N}\mathbf{h}_1(\mathbf{x}_i)^H\mathbf{y}_i}{\sum_{i=1}^{N}\|\mathbf{h}_1(\mathbf{x}_i)\|^2} \quad (3)$$

and the corresponding $\mathrm{MSE} = \mathrm{E}\left[|\hat{\theta}_1^{(N)} - \theta_1|^2\right]$ is

$$\mathrm{MSE}^{(N)}\left(\{\mathbf{x}_i\}_{i=1}^{N}\right) = \mathrm{E}\left[\left|\frac{\sum_{i=1}^{N}\mathbf{h}_1(\mathbf{x}_i)^H\mathbf{n}_i}{\sum_{i=1}^{N}\|\mathbf{h}_1(\mathbf{x}_i)\|^2}\right|^2\right]. \quad (4)$$

Denote $E_i(\mathbf{y}_1,\ldots,\mathbf{y}_{i-1}) = \|\mathbf{x}_i(\mathbf{y}_1,\ldots,\mathbf{y}_{i-1})\|^2$, where E_i represents the energy allocated at each time step i. The total energy in the measurements is given by

$$\mathcal{E}\left[\{\mathbf{x}_i(\mathbf{y}_1,\ldots,\mathbf{y}_{i-1})\}_{i=1}^{N}\right] = \mathrm{E}\left[\sum_{i=1}^{N}E_i(\mathbf{y}_1,\ldots,\mathbf{y}_{i-1})\right]. \quad (5)$$

Our goal is to find the best sequence of the transmitted signals $\{\mathbf{x}_i\}_{i=1}^{N}$ to minimize the $\mathrm{MSE}^{(N)}$ in (4) under the average energy constraint $\mathrm{E}\left[\sum_{i=1}^{N}\|\mathbf{x}_i\|^2\right] \leq E_0$. Define $\mathrm{SNR}\left(\{\mathbf{x}_i\}_{i=1}^{N}\right)$ as

$$\mathrm{SNR}^{(N)}\left(\{\mathbf{x}_i\}_{i=1}^{N}\right) = \frac{\mathcal{E}\left[\{\mathbf{x}_i(\mathbf{y}_1,\ldots,\mathbf{y}_{i-1})\}_{i=1}^{N}\right]}{\sigma^2}. \quad (6)$$

The average energy constraint can be rewritten as $\mathrm{SNR}^{(N)} \leq \mathrm{SNR}_0$, where $\mathrm{SNR}_0 = E_0/\sigma^2$. Minimizing $\mathrm{MSE}^{(N)}$ subject to an SNR constraint $\mathrm{SNR}^{(N)} \leq \mathrm{SNR}_0$ is equivalent to minimizing $\mathrm{MSE}^{(N)} \times \mathrm{SNR}^{(N)}$ [8]. The product of $\mathrm{MSE}^{(N)}$ and $\mathrm{SNR}^{(N)}$ is given by

$$\mathrm{MSE}^{(N)}\times\mathrm{SNR}^{(N)} = \mathrm{E}\left[\left|\frac{\sum_{i=1}^{N}\mathbf{h}_1(\mathbf{x}_i)^H\mathbf{n}_i}{\sum_{i=1}^{N}\|\mathbf{h}_1(\mathbf{x}_i)\|^2}\right|^2\right]\frac{\mathrm{E}\left[\sum_{i=1}^{N}\|\mathbf{x}_i\|^2\right]}{\sigma^2}.$$

As a benchmark for comparison, we consider the non-adaptive case where $\mathbf{x}_i(\mathbf{y}_1,\ldots,\mathbf{y}_{i-1}) = \sqrt{E_i}\,\bar{\mathbf{x}}_i$. Here $\bar{\mathbf{x}}_i$, E_i are deterministic quantities, independent of $\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_{i-1}$, $\|\bar{\mathbf{x}}_i\| = 1$, and $\sum_{i=1}^{N}E_i \leq E_0$. For the model (2), $\mathrm{MSE}^{(N)}$ is given by

$$\mathrm{MSE}^{(N)} = \frac{\sigma^2}{\sum_{i=1}^{N}\|\mathbf{h}_1(\mathbf{x}_i)\|^2} = \frac{\sigma^2}{\sum_{i=1}^{N}E_i\frac{\|\mathbf{h}_1(\bar{\mathbf{x}}_i)\|^2}{\|\bar{\mathbf{x}}_i\|^2}} \geq \frac{\sigma^2}{E_0\lambda_\mathrm{m}}, \quad (7)$$

where equality is achieved iff $\forall i\ \bar{\mathbf{x}}_i = \mathbf{v}_\mathrm{m}$, the normalized eigenvector corresponding to λ_m, the maximum eigenvalue of the channel matrix $\boldsymbol{H}_1^H\boldsymbol{H}_1$. Note $\lambda_\mathrm{m} = \max_{\mathbf{x}}(\mathbf{x}^H\boldsymbol{H}_1^H\boldsymbol{H}_1\mathbf{x})/(\mathbf{x}^H\mathbf{x}) = \max_{\mathbf{x}}\|\mathbf{h}_1(\mathbf{x})\|^2/\|\mathbf{x}\|^2$. Furthermore, the performance of the ML estimator does not depend on the energy allocation. Hence, without loss of generality we can assume that all energy is allocated to the first transmission. The minimum MSE of the one-step (or non-adaptive N-step) strategy for a scalar parameter is then given by $\mathrm{MSE}_{\min}^{(1)} = 1/\widetilde{\mathrm{SNR}}_0$, where $\widetilde{\mathrm{SNR}}_0 = \lambda_\mathrm{m}\mathrm{SNR}_0$. We first look at a two-step sequential design procedure.

3. OMNISCIENT TWO-STEP SEQUENTIAL STRATEGY

In the two-step sequential procedure, we have $N = 2$ time steps where in each time step $i = 1, 2$, we can control input waveform $\mathbf{x}_i$ to obtain signal $\mathbf{y}_i$. The two-step ML estimator of θ_1 from (3) is

$$\hat{\theta}_1^{(2)} = \frac{\mathbf{h}_1(\mathbf{x}_1)^H\mathbf{y}_1 + \mathbf{h}_1(\mathbf{x}_2)^H\mathbf{y}_2}{\|\mathbf{h}_1(\mathbf{x}_1)\|^2 + \|\mathbf{h}_1(\mathbf{x}_2)\|^2} \quad (8)$$

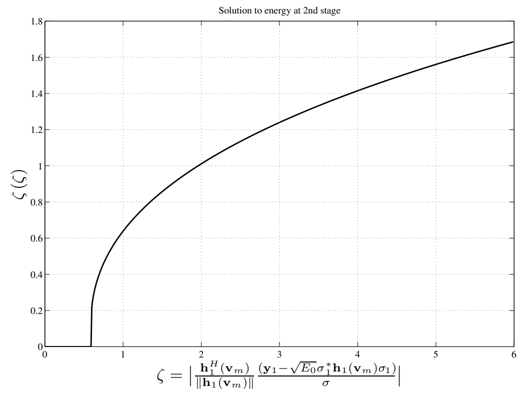

Fig. 1. Plot of the optimal solution to the normalized energy transmitted at the second stage as a function of received signal at first stage.

and the corresponding $\mathrm{MSE}^{(2)}$ to be minimized from (4) is

$$\mathrm{MSE}^{(2)} = \mathrm{E}\left[\frac{|\mathbf{h}_1(\mathbf{x}_1)^H\mathbf{n}_1 + \mathbf{h}_1(\mathbf{x}_2)^H\mathbf{n}_2|^2}{(\|\mathbf{h}_1(\mathbf{x}_1)\|^2 + \|\mathbf{h}_1(\mathbf{x}_2)\|^2)^2}\right]. \quad (9)$$

We assume that the shape of the optimal designs, i.e., $\{\mathbf{x}_i/\|\mathbf{x}_i\|\}$ is the one-step optimum given by $\mathbf{v}_\mathrm{m}$ defined below (7) and minimize the MSE over the energy of the waveforms. Denote $\|\mathbf{x}_1\| = \sqrt{E_0}\alpha_1$ and $\|\mathbf{x}_2(\mathbf{y}_1)\| = \sqrt{E_0}\alpha_2(\mathbf{y}_1)$. The average energy constraint, $\mathcal{E}[\mathbf{x}_1,\mathbf{x}_2(\mathbf{y}_1)] = \mathrm{E}\left[\|\mathbf{x}_1\|^2 + \|\mathbf{x}_2\|^2\right] \leq E_0$ can be rewritten as $\alpha_1^2 + \mathrm{E}\left[\alpha_2^2(\mathbf{y}_1)\right] \leq 1$. We use Lagrangian multipliers to minimize the $\mathrm{MSE}^{(2)}$ in (9) with respect to α_1 and $\alpha_2(\cdot)$ under this energy constraint. The optimal design for the two-step procedure is $\mathbf{x}_2^*(\mathbf{y}_1) = \sqrt{E_0}\alpha_2^*(\mathbf{y}_1)\mathbf{v}_\mathrm{m}$, where

$$\alpha_2^*(\mathbf{y}_1) = \beta\left(\left|\frac{\mathbf{h}_1(\mathbf{v}_\mathrm{m})^H}{\|\mathbf{h}_1(\mathbf{v}_\mathrm{m})\|}\frac{(\mathbf{y}_1 - \sqrt{E_0}\alpha_1^*\mathbf{h}_1(\mathbf{v}_\mathrm{m})\theta_1)}{\sigma}\right|\right) \quad (10)$$

has MSE satisfying $\mathrm{MSE}_{\min}^{(2)} \times \widetilde{\mathrm{SNR}}_0 = \eta_2^* \approx 0.68$, and $\alpha_1^* \approx 0.7421$. The optimal solution in terms of $\beta(\cdot)$ is shown in Fig. 1. This solution depends on the unknown parameter θ_1 and thus we will call this minimizer an "omniscient" energy allocation strategy. The two-step strategy yields a 32% improvement in performance or a 1.7dB gain in terms of SNR. The product $\mathrm{MSE}^{(2)} \times \widetilde{\mathrm{SNR}}_0$ is plotted for various values of α_1 using both simulations (dotted) and theory (solid) in Fig. 2. The details of the derivation can be found in [8].

The "omniscient" solution (10) depends on the parameter to be estimated. Here, we prove that we can approach the optimal two-step solution by implementing a θ_1-independent energy allocation strategy when θ_1 is bounded, i.e., $\theta_1 \in [\theta_a,\theta_b]$, $\theta_a,\theta_b \in \mathbb{R}$. Since we do not know the value of the actual parameter, we replace θ_1 by a 'guess' of θ_1 say θ_g in the optimal solution to the energy at the second stage given in (10). The resulting suboptimal design is

$$\mathbf{x}_2 = \sqrt{E_0}\,\beta\left(\left|\tilde{n}_1 + \frac{\alpha_1^*\|\mathbf{h}_1(\mathbf{v}_\mathrm{m})\|\sqrt{E_0}}{\sigma}(\theta_1 - \theta_g)\right|\right)\mathbf{v}_\mathrm{m}, \quad (11)$$

where $\tilde{n}_1 = \mathbf{h}_1(\mathbf{v}_\mathrm{m})^H\left(\mathbf{y}_1 - \sqrt{E_0}\alpha_1^*\mathbf{h}_1(\mathbf{v}_\mathrm{m})\theta_1\right)/\lambda_m\sigma \sim \mathcal{CN}(0,1)$. When the optimal two-step design is used with θ_g in place of θ_1, $\eta(z) = \mathrm{MSE}^{(2)} \times \widetilde{\mathrm{SNR}}^{(2)}$ is

$$\eta(z) = \mathrm{E}\left[\frac{\alpha_1^{*2}|\tilde{n}_1|^2 + \beta^2(|\tilde{n}_1 + z|)}{(\alpha_1^{*2} + \beta^2(|\tilde{n}_1 + z|))^2}\right]\mathrm{E}\left[\alpha_1^{*2} + \beta^2(|\tilde{n}_1 + z|)\right],$$

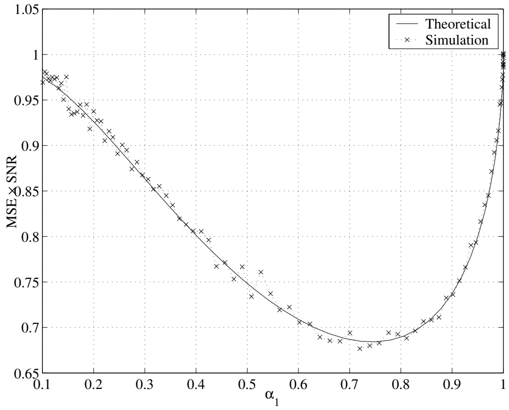

Fig. 2. Reduction in MSE for varying values of energy transmitted at the first stage, α_1.

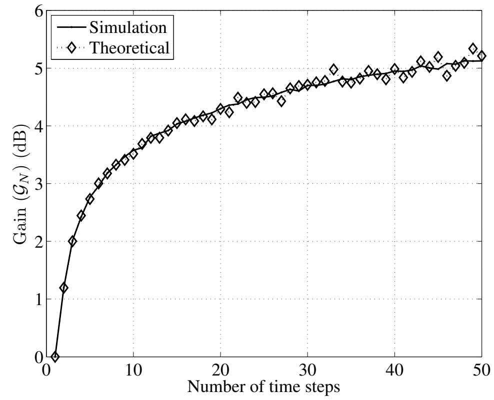

Fig. 3. Plot of gain over non-adaptive energy allocation strategy obtained by implementing the adaptive N-step procedure as a function of N through theory [8] and simulations.

where $z = \alpha_1^* \sqrt{E_0} \|\mathbf{h}_1(\mathbf{v}_\mathrm{m})\| (\theta_1 - \theta_g)/\sigma = \alpha_1^* \sqrt{\widetilde{\mathrm{SNR}}_0} (\theta_1 - \theta_g)$ and $\widetilde{\mathrm{SNR}}^{(2)} = \lambda_\mathrm{m} \mathrm{SNR}^{(2)}$. Therefore, $\mathrm{MSE}^{(2)} \times \widetilde{\mathrm{SNR}}^{(2)}$ is a function of z only and is optimum at $z = 0$. Note that when SNR becomes sufficiently small $\mathrm{MSE}^{(2)} \times \widetilde{\mathrm{SNR}}^{(2)}$ approaches its minimal value. Hence instead of performing one two-step procedure, we perform a set of N independent two-step procedures with equal energy E_0/N and average the estimates from each step to obtain the new estimate. In such a way, we reduce the SNR at each stage, thereby eliminating the effect of the unknown parameter θ_1. As $N \to \infty$, $z \to 0$ and the optimum two-step $\mathrm{MSE}^{(2)} \times \widetilde{\mathrm{SNR}}^{(2)} = \eta_2^*$ is achieved. The complete proof can be found in [8].

4. DESIGN OF N-STEP PROCEDURE

In Section 3, we looked at the optimal two-stage sequential design procedure for energy allocation and proved that we can achieve the optimal performance using an $N \times 2$-step strategy. In this section, we generalize the solution from the two-step case to an N-step strategy. We assume that the shape of the transmitted waveform is fixed and look at the energy allocation among the various steps. Let the energy at step k be denoted as $\alpha_k^2(\mathbf{y}_1, \dots, \mathbf{y}_{k-1})$, i.e., $\mathbf{x}_k = \mathbf{v}_\mathrm{m} \alpha_k(\mathbf{y}_1, \dots, \mathbf{y}_{k-1})$, $1 \le k \le N$. Then

$$\alpha_1 = A_1,\ \alpha_k = A_k \mathbf{I}\left(\frac{|\sum_{i=1}^{k-1} \mathbf{h}_1(\mathbf{x}_i)^H \mathbf{n}_i|^2}{\sum_{i=1}^{k-1} \|\mathbf{h}_1(\mathbf{x}_i)\|^2 \sigma^2} \ge \rho_k \right), \quad k \ge 2.$$

Note that the definition of the energy at each stage is recursive. This suboptimal energy allocation for the N-step case is an approximation to the optimal threshold like solution for the two-step case. We choose $\mathbf{A} = [A_1, \dots, A_N]$ and $\boldsymbol{\rho} = [\rho_1, \dots, \rho_N]$ appropriately to satisfy the average energy constraint. The intuition behind the choice of $\mathbf{A}, \boldsymbol{\rho}$ is motivated by an asymptotic result in [8]. We evaluate the performance of this suboptimal approach using simulations. Performance gain $\mathcal{G}_N$ (in dB) is presented in Fig. 3. We see that in 50 steps, we are able to achieve a gain of more than 5dB! Moreover, by the same argument presented in Section 3, SNR decreases at each step which implies that as the number of steps increases, the lack of knowledge on θ_1 has a limited effect on the overall performance.

5. VECTOR PARAMETER CASE

A general N-step procedure for the case of M unknown parameters is defined in (1). For the multiple parameter case, we consider the trace of the MSE matrix as a measure of performance. The problem of multiple parameter estimation is more complicated than estimation of a single parameter for the following reason. We showed in Section 2 that independent of the shape of $\mathbf{x}_i$, any non-adaptive energy allocation strategy is to assign all energy to the first step, i.e., a one step strategy with total energy E_0. But this is not true for a multiple parameter setting. Let us consider a simple example of estimating two parameters $\boldsymbol{\theta} = [\theta_1\ \theta_2]^T$ in the model $\mathbf{y} = \mathbf{H}(\mathbf{x})\boldsymbol{\theta} + \mathbf{n}$, where

$$\mathbf{H}(\mathbf{x}) = \begin{bmatrix} x_1 & x_2 \\ 0 & x_2 \end{bmatrix}, \quad (12)$$

$\mathbf{x} = [x_1\ x_2]^T$, $\mathbf{y} = [y_1\ y_2]^T$, and $\mathbf{n} = [n_1\ n_2]^T \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I})$. Then for a one-step process, we have $\mathrm{MSE}^{(1)}(\theta_1) = 2\sigma^2/x_1^2$ and $\mathrm{MSE}^{(1)}(\theta_2) = \sigma^2/x_2^2$. Minimizing the $\mathrm{tr}(\mathrm{MSE}^{(1)}) = \mathrm{MSE}^{(1)}(\theta_1) + \mathrm{MSE}^{(1)}(\theta_2)$ (tr denotes the trace) over the energy constraint $\|\mathbf{x}\|^2 \le E_0 = 1$, we obtain $x_1 = x_2 = 1/\sqrt{2}$ and $\mathrm{tr}(\mathrm{MSE}^{(1)}_{\mathrm{min}}) = 6\sigma^2$. Now consider the following two-step non-adaptive energy design.

Step 1. $\quad \mathrm{Tx}: \mathbf{x} = [x_1\ 0]^T,\ \mathrm{Rx}:\ y_1 = x_1\theta_1 + n_1,$

Step 2. $\quad \mathrm{Tx}: \mathbf{x} = [0\ x_2]^T,\ \mathrm{Rx}:\ [1\ 1]\mathbf{y}_2 = 2x_2\theta_2 + [1\ 1]\mathbf{n}_2.$

Minimizing $\mathrm{tr}(\mathrm{MSE}^{(2)}) = \mathrm{MSE}^{(2)}(\theta_1) + \mathrm{MSE}^{(2)}(\theta_2) = \sigma^2/x_1^2 + \sigma^2/2x_2^2$ over the energy constraint, we obtain $x_1 = x_2 = 1/\sqrt{2}$ and $\mathrm{tr}(\mathrm{MSE}^{(2)}_{\mathrm{min}}) = 3\sigma^2$. This translates to a 3dB gain in SNR for the two-step non-adaptive strategy over the one step approach. We control the input $\mathbf{x} = [x_1\ x_2]^T$ such that we have different energy allocation for each column of the matrix $\mathbf{H}$. By specifically designing the two-step non-adaptive strategy given in step 1 and step 2, we have reduced the estimation of the vector parameter $\boldsymbol{\theta} = [\theta_1, \theta_2]$ to two independent problems of estimating scalar parameters θ_1 and θ_2 respectively. For each of these scalar estimators, we design two N-step sequential procedures as in Section 4 for scalar controls x_1 and x_2 to obtain an improvement in performance of estimating $\boldsymbol{\theta}$. Applying the N-step design to both x_1 and x_2, we have $\mathrm{MSE}^{(N)}(\theta_i) = \mathcal{G}_N \mathrm{MSE}^{(2)}_{\mathrm{min}}(\theta_i)$ and hence $\mathrm{tr}(\mathrm{MSE}^{(N)}) = \mathcal{G}_N \mathrm{tr}(\mathrm{MSE}^{(2)}_{\mathrm{min}})$. In other words, we would obtain the gains of the N-step procedure over non-adaptive strategies for the vector parameter case as well.

6. APPLICATIONS OF SEQUENTIAL ESTIMATION

6.1. MIMO Channel Estimation

One important component in a MIMO system is the need to accurately estimate the channel state information (CSI) at the transmitter and receiver. Recently, [2] proposed four different training based methods for the flat block-fading MIMO system including the least squares and best linear unbiased estimator (BLUE) approach for the case of multiple LS channel estimates. In order to estimate the $r \times t$ channel matrix $\mathbf{\Theta}$ with t transmit and r receive antennas, $N \geq t$ training vectors $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]$ are transmitted. The corresponding received signal is $\mathbf{R} = \mathbf{\Theta X} + \mathbf{M}$, where $\mathbf{R} = [\mathbf{r}_1, \ldots, \mathbf{r}_N]$ is a $r \times N$ matrix, $\mathbf{M} = [\mathbf{m}_1, \ldots, \mathbf{m}_N]$ is the $r \times N$ matrix of sensor noise, $\mathbf{x}_i$ is the $t \times 1$ complex vector of transmitted signals and $\mathbf{m}_i$ is the $r \times 1$ complex zero-mean white noise vector. Let P_0 be the transmitted training power constraint, i.e., $\|\mathbf{X}\|_{\text{F}}^2 = P_0$, $\|\cdot\|_F$ indicates Frobenius norm ($\|\mathbf{X}\|_{\text{F}} = \sqrt{\text{tr}(\mathbf{X}^H\mathbf{X})}$) and σ^2 denote the variance of receiver noise. Assuming co-located transmitter and receiver arrays and multiple training periods available within the same coherency time (quasi-static) to estimate the channel, the set of received signals can be rewritten in the following form:

$$\mathbf{y}_i = \mathbf{H}(\mathbf{X}_i)\boldsymbol{\theta} + \mathbf{n}_i \quad i = 1, 2, \ldots, K, \tag{13}$$

where $\mathbf{y}_i = \text{vec}(\mathbf{R}_i), \boldsymbol{\theta} = \text{vec}(\mathbf{\Theta}), \mathbf{n}_i = \text{vec}(\mathbf{M}_i)$, $\text{vec}(\cdot)$ denotes the column-wise concatenation of the matrix, $\mathbf{H}(\mathbf{X}_i) = (\mathbf{X}_i \otimes \mathbf{I})^T$ is a linear function of the input $\mathbf{X}_i$, and $\otimes$ denotes the Kronecker product, which is the same model described in (1). In [2], a method of linearly combining the estimates from the K transmission stages was proposed and the MSE of the K-step estimator was shown to be $\text{MSE}_{\text{LS}}^{(K)} = \sigma^2 t^2 r / P_0$, where P_0 is the total power used in the K steps, i.e., $\sum_{i=1}^{K} \|\mathbf{X}_i\|_{\text{F}}^2 \leq P_0$. If we have enough training samples, we could completely control the matrix $\mathbf{H}(\mathbf{X}_i)$ through our input $\mathbf{X}_i$ and we can make $\mathbf{H}(\mathbf{X}_i)$ orthogonal. In this case (13) along with the average power constraint $\text{E}\left[\sum_i \|\mathbf{X}_i\|_{\text{F}}^2\right] \leq P_0$ falls in the framework of the problem of adaptive energy allocation in Sections 4 and 5 where the problem is then separable into N independent estimation problems of scalar parameters. Having K steps in the training sequence also directly enables us to implement our K-step strategy to achieve optimal performance. Hence it directly follows that using our strategy we are guaranteed to achieve the optimal error given by $\text{MSE}_{\text{LS}}^{(K)} \approx \mathcal{G}_K \sigma^2 t^2 r / P_0$, which we have shown to be at least 5dB (in 50 steps) better than any non-adaptive strategy.

6.2. Inverse Scattering Problem

The problem of imaging a medium using an array of transducers has been widely studied in many research areas such as mine detection, ultrasonic medical imaging, and non-destructive testing. The goal in imaging is to detect and image small scatterers in a known background medium. We apply our concept of designing a sequence of measurements to image a medium of multiple scatterers using an array of transducers. The imaging area (or volume) is divided into V voxels and the channel, denoted $\mathbf{a}_i$, between a candidate voxel i and the N transducers is given by the homogeneous Green's function and ignores the effect of multiple scattering. Each voxel can be characterized by its scatter coefficient, e.g., radar cross-section, $\{\theta_i\}_{i=1}^{V}$, which indicates the proportion of the received field that is re-radiated. Thus the channel between the transmitted field and the measured backscattered field at the transducer array is $\mathbf{A}\text{diag}(\boldsymbol{\theta})\mathbf{A}^T$, where $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_V]$, $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_V]^T$, and $\text{diag}(\boldsymbol{\theta})$ is a $V \times V$ diagonal matrix with θ_i as its i^{th} diagonal element. The probing mechanism for imaging of the scatter cross section generates the following sequence of noise contaminated signals

$$\mathbf{y}_i = \mathbf{A}\text{diag}(\boldsymbol{\theta})\mathbf{A}^T\mathbf{x}_i + \mathbf{n}_i = \mathbf{H}(\mathbf{x}_i)\boldsymbol{\theta} + \mathbf{n}_i, \tag{14}$$

where $\mathbf{H}(\mathbf{x}_i) = \mathbf{A}\text{diag}(\mathbf{A}^T\mathbf{x}_i)$. The goal is to find estimates for the scattering coefficients $\boldsymbol{\theta}$ under the average energy constraint to minimize the MSE. If $\mathbf{A}$ is an invertible square matrix, then we can condition $\text{diag}(\mathbf{A}^T\mathbf{x}_i)$ to have a single non zero component on any one of the diagonal elements, which translates to isolating the i^{th} column for any i. As in Section 5, we can perform V independent N-step experiments to guarantee the N-step gains of at least 5dB over the standard single step ML estimation for imaging [10].

7. CONCLUSIONS

In this paper we considered the N-step adaptive waveform-amplitude design problem for estimating parameters of an unknown channel under average energy constraints. For $N = 2$ and a linear Gaussian channel model, we found the optimal amplitude to transmit at the second step as a function of the first measurement for a scalar parameter case. We showed that this two-step adaptive strategy obtained an improvement of at least 1.7dB over any non-adaptive strategy. We then designed a suboptimal N-stage procedure based on the two-step approach and proved gains of more than 5dB in $N = 50$ steps. Furthermore, we extended our results to the case of vector parameters. To conclude, we provided applications of our design to MIMO and inverse scattering channel models.

8. REFERENCES

[1] B. L. Scala, M. Rezaeian, and B. Moran, "Optimal adaptive waveform selection for target tracking," *Proc. Intl. Conf. Information Fusion*, vol. 1, pp. 552–557, 2005.

[2] M. Biguesh and A. B. Gershman, "Training based mimo channel estimation: A study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Processing*, vol. 54, no. 3, pp. 884–892, 2006.

[3] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York, N.Y: Springer-Verlag, 1988.

[4] C. M. Stein, "Inadmissibility of the usual estimator for the mean of a multivariate normal distribution," *Proc. 3rd Berkeley Symp. Math. Stat. Prob.*, vol. 1, pp. 197–206, 1956.

[5] L. S. Mayer and T. A. Willke, "On biased estimation in linear models," *Technometr.*, vol. 15, pp. 497–508, 1973.

[6] A. N. Tikhonov and V. Y. Arsenin, *Solution to Ill-Posed Problems*. Washington D.C: V. H. Winston, 1977.

[7] Y. C. Eldar and A. V. Oppenheim, "Covariance shaping least-squares estimation," *IEEE Trans. Signal Processing*, vol. 51, no. 3, pp. 686–697, 2003.

[8] R. Rangarajan, R. Raich, and A. O. Hero, "Optimal sequential energy allocation for inverse problems," accepted with minor revisions to *IEEE J. Select. Topics in Signal Processing*, May, 2007.

[9] ——, "Optimal experimental design for an inverse scattering problem," *Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, vol. 4, pp. 1117–1120, 2005.

[10] G. Shi and A. Nehorai, "Maximum likelihood estimation of point scatterers for computational time-reversal imaging," *Communications in Information and Systems*, vol. 5, no. 2, pp. 227–256, 2005.