

SINGLE-STAGE WAVEFORM SELECTION FOR ADAPTIVE RESOURCE CONSTRAINED STATE 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

We consider the problem of optimal waveform selection. We would like to choose a small subset from a given set of waveforms that minimizes state prediction mean squared error (MSE) given the past observations. This differs from previous approaches to this problem since the optimal waveforms cannot be computed offline; it requires the previous observations. Since the optimal solution to this subset selection problem is combinatorially complex, we propose a convex relaxation of the problem and provide a low complexity suboptimal solution. We present a specific model and show that the performance of this suboptimal procedure approaches that of the optimal waveforms.

1. INTRODUCTION

Over the past decade, the problem of optimal waveform design has found important applications in synthetic aperture radar (SAR), automatic target recognition and radar astronomy [1]. Based on the application, waveform design may depend on various optimality criteria, e.g., target classification [2], accurate reconstruction of a high resolution radar image [3], or estimating a set of target parameters. One implication of choosing the set of transmitted waveforms optimally is that the backscattered signals will contain maximum target information.

Most of the work in the area of waveform design involves finding the best functional form of the waveforms suited to a particular task, e.g., design of waveforms from the radar ambiguity function for narrowband signals [4] or design of wideband waveforms to resolve targets in dense target environments [5]. In this paper, we focus on the optimal waveform selection problem rather than the design of actual waveforms. We would like to choose only a small subset from a given set of waveforms. This restriction is typical in radar systems where there is a constraint on resources such as energy. To assess the performance of a particular subset of waveforms, we need to define an optimization criterion such as expected reward or risk.

The problem of choosing p out of M possible waveforms becomes a high complexity combinatorial optimization problem. E.g., if there are $M = 128$ waveforms and we need to select $p = 32$ element subset, there are more than 10^{30} combinations of indices that need to be checked. As a result, significant work has been focussed on approximation methods based on convex relaxation which lead to sparse solutions. Complexity penalties have also been used to find sparse solutions to such problems [6]. One type of convex penalty is the lasso, a shrinkage method which imposes an l_1 -norm constraint

on the optimization problem [7]. By nature of the constraint, making the weighting of the constraint larger causes some of the coefficients to be zero thus giving rise to a suboptimal sparse solution to the subset selection problem. Recent work advocates the use of l_1 -norm constrained convex optimization problems to obtain sparse representations [6]. Most of these problems deal with sparse regression and are offline strategies where the solution to the problem is found based on accumulated data.

In this paper, we consider the expected state prediction MSE as a measure of performance and impose the problem of finding the optimal subset that minimizes this expected reward given the past measurements (online strategy). We relax this combinatorially complex problem into an optimization problem under l_1 -norm constraint and propose a low complexity suboptimal solution whose performance approaches that of the optimal subset selection. We then consider a numerical example of this approach and provide simulation results to compare the various solutions.

The organization of the paper is as follows: In Section 2, we present the waveform selection problem. Section 3 proposes a suboptimal solution. In Section 4, we solve the problem for a specific model. Section 5 addresses the computational complexity of the proposed solution and Section 6 provides simulation results. We conclude this paper in Section 7.

2. PROBLEM FORMULATION

We consider the waveform selection problem for a hyper-spectral radar system, where the radar can transmit and receive energy over multiple channels simultaneously. We restrict the number of waveforms transmitted at any time to be a small subset of p out of M available waveforms. Denote the state at time t as $\mathbf{s}_t$ and let the received signals corresponding to a single transmit waveform ϕ_i be denoted as $\mathbf{y}_t^i$, $i = 1, \dots, M$. We restrict our attention to single stage policies, i.e., myopic policies that seek to maximize an expected reward conditioned on the immediate past.

Let $\{i_1, \dots, i_p\} \in \{1, \dots, M\}$ denote the indices of the p different waveforms taken from a set of M ($M \geq p$) waveforms. We solve the optimal subset selection problem by maximizing the expected reduction in the variance of the optimal state estimator after an action (choosing p out of M waveforms) is taken:

$$\max_{i_1, \dots, i_p} \left\{ \mathbb{E} \left[\|\mathbf{s}_t - \mathbb{E}[\mathbf{s}_t | \mathbf{y}_{t-1}]\|^2 | \mathbf{y}_{t-1} \right] - \mathbb{E} \left[\left\| \mathbf{s}_t - \mathbb{E}[\mathbf{s}_t | \mathbf{y}_t^{i_1}, \dots, \mathbf{y}_t^{i_p}, \mathbf{y}_{t-1}] \right\|^2 | \mathbf{y}_{t-1} \right] \right\}. \quad (1)$$

Since the first term is independent of $\{i_1, \dots, i_p\}$, the maximization

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

in (1) can be equivalently expressed as

$$\min_{i_1, \dots, i_p} \mathbb{E} [\| \mathbf{s}_t - \hat{\mathbf{s}}_t(i_1, \dots, i_p) \|^2 | \mathbf{y}_{t-1}], \quad (2)$$

where

$$\hat{\mathbf{s}}_t(i_1, \dots, i_p) = \mathbb{E} [\mathbf{s}_t | \mathbf{y}_t^{i_1}, \dots, \mathbf{y}_t^{i_p}, \mathbf{y}_{t-1}]. \quad (3)$$

The minimization in (2) requires one to evaluate (3) for all $\binom{M}{p}$ possibilities of $i_1, \dots, i_p$. Two fundamental difficulties are encountered in solving (2): computation of the conditional expectation (3); and combinatorial minimization of (2). In the tracking examples considered here the computation of (3) is not difficult. Since the complexity of problem is exponential in M (for fixed p/M), we propose a low complexity suboptimal solution for (2) whose performance approaches that of the optimal one.

3. PROPOSED SOLUTION

As an alternative to exhaustively searching over $\binom{M}{p}$ possible subsets we pose the following sparsity constrained prediction surrogate:

$$\min_{\gamma} \mathbb{E} \left[\left\| \mathbf{s}_t - \sum_i \gamma_i g_i(\mathbf{y}_t^1, \dots, \mathbf{y}_t^M, \mathbf{y}_{t-1}) \right\|^2 | \mathbf{y}_{t-1} \right] + \beta \|\gamma\|_l, \quad (4)$$

where $\beta \geq 0$, $\|\gamma\|_l$, $0 \leq l \leq 1$ is a sparseness inducing penalty and $\{g_i\}$ is a set of base predictors of $\mathbf{s}_t$ and the linear combination of these predictors approximates the exact solution in (3). When $\|\gamma\|_l = \|\gamma\|_0$ is the l_0 -norm, $g_i(\mathbf{y}_t^1, \dots, \mathbf{y}_t^M, \mathbf{y}_{t-1}) = g_i(\mathbf{y}_t^{i_1}, \dots, \mathbf{y}_t^{i_p}, \mathbf{y}_{t-1}) = \hat{\mathbf{s}}_t$ in (3), i indexes over the $\binom{M}{p}$ combinations of indices $i_1, \dots, i_p$, the solution of (4) yields the optimal solution (2) for sufficiently large β . A surrogate investigated by many [7, 8] for the l_0 -norm penalty is the l_1 -norm penalty $\|\gamma\|_1$ which will be adopted here. In the special case that g_i depends only on a single variable $\mathbf{y}_t^i$ the regression in (4) is equivalent to using a simple generalized additive model (GAM) [9]. We further assume that $g_i(\mathbf{y}_t^i) = \mathbb{E} [\mathbf{s}_t | \mathbf{y}_t^i, \mathbf{y}_{t-1}]$. Thus the constrained prediction problem can be formulated as

$$\min_{\gamma} \mathbb{E} \left[\left\| \mathbf{s}_t - \sum_{i=1}^M \gamma_i \mathbb{E} [\mathbf{s}_t | \mathbf{y}_t^i, \mathbf{y}_{t-1}] \right\|^2 | \mathbf{y}_{t-1} \right] + \beta \|\gamma\|_1, \quad (5)$$

and β is chosen such that exactly p out of the M γ_i 's are nonzero. This quadratic optimization in γ under l_1 -norm constraint is a convex problem and can be evaluated in a straightforward fashion using standard techniques, e.g., [7, 8, 10, 11]. We first find the range of β that gives rise to a sparse solution with exactly p nonzero elements and fix it to that value in the range which gives the minimum unconstrained error. We take the indices of the p nonzero components of γ corresponding to this β as the solution to the waveform subset selection problem in (2).

4. NUMERICAL STUDY

To illustrate this approach, we consider the following problem: At time $t = 1$, we assume without loss of generality that an arbitrary waveform index η from $\{1, \dots, M\}$ is chosen and waveform ϕ_η is

transmitted into the medium. The received signal at the first stage can then be written as

$$\mathbf{y}_1 = \mathbf{L}(\phi_\eta) \mathbf{s}_1 + \mathbf{n}_1 = \mathbf{L}_\eta \mathbf{s}_1 + \mathbf{n}_1, \quad (6)$$

where $\mathbf{L}(\cdot)$ is based on the channel model, $\mathbf{n}_1$ is receiver noise, and $\mathbf{s}_1$ is the initial state. We consider the state update equation as a hidden Markov model (HMM), equivalent to a Gaussian mixture model, defined as

$$\mathbf{s}_t = \mathbf{A} \mathbf{s}_{t-1} + \mathbf{I}_t \mathbf{w}_{1,t} + (1 - \mathbf{I}_t) \mathbf{w}_{0,t}, \quad t = 2, 3, \dots, \quad (7)$$

where $\{\mathbf{w}_{i,t}, i = 0, 1\}_t$ are independent Normal random vectors with mean $\boldsymbol{\mu}_i$ and covariance matrix $\mathbf{R}_{\mathbf{w}_i}$, $\mathbf{A}$ is a fixed matrix and $\mathbf{I}_t$ are i.i.d Bernoulli random variables with success probability q .

Assume the initial state $\mathbf{s}_1$ is a Normal random vector with zero mean and covariance matrix $\mathbf{R}_s$. Receiver noises $\{\mathbf{n}_t\}$ are i.i.d Normal with zero mean and covariance matrix $\mathbf{R}_n$ and $\{\mathbf{n}_t, \{\mathbf{w}_{i,t}, i = 0, 1\}, \mathbf{I}_t, \mathbf{s}_1\}$ are all independent. The model (7) captures the non-Gaussian nature of the tracking problem where the state dynamics switch at random between the hidden states $\mathbf{I}_t = 1$ and $\mathbf{I}_t = 0$. The received signal at time $t = 2$ corresponding to transmission of waveform ϕ_i can be written as

$$\mathbf{y}_2^i = \mathbf{L}_i \mathbf{s}_2 + \mathbf{n}_2^i, \quad i = 1, \dots, M. \quad (8)$$

Our goal is to maximize expected reduction in the variance of the state estimator after sending the waveforms $\{\phi_{i_k}\}_{k=1}^p$ and receiving the backscatter $\mathbf{y}_2^{i_1}, \dots, \mathbf{y}_2^{i_p}$, i.e.,

$$\min_{i_1, \dots, i_p} \mathbb{E} \left[\left\| \mathbf{s}_2 - \mathbb{E} [\mathbf{s}_2 | \mathbf{y}_2^{i_1}, \dots, \mathbf{y}_2^{i_p}, \mathbf{y}_1] \right\|^2 | \mathbf{y}_1 \right]. \quad (9)$$

For the proposed GAM prediction problem under l_1 -norm constraint, we need to minimize

$$\mathbb{E} \left[\left\| \mathbf{s}_2 - \sum_{i=1}^M \gamma_i \mathbb{E} [\mathbf{s}_2 | \mathbf{y}_2^i, \mathbf{y}_1] \right\|^2 | \mathbf{y}_1 \right] + \beta \|\gamma\|_1 \quad (10)$$

with respect to γ and use the nonzero indices obtained through this method as our solution to the subset selection problem.

Given $\mathbf{I}_2 = k \in \{0, 1\}$, the random vectors $\mathbf{x}_2$, $\mathbf{y}_2^i$ and $\mathbf{y}_1$ are jointly Gaussian. Let $\mathbf{y} = [\mathbf{y}_2^{i_1 T}, \dots, \mathbf{y}_2^{i_p T}, \mathbf{y}_1^T]^T$. Then the joint distribution can be written as

$$\begin{pmatrix} \mathbf{s}_2 \\ \mathbf{y} \end{pmatrix}_{\mathbf{I}_2=k} = \mathcal{N} \left[\begin{pmatrix} \boldsymbol{\mu}_k \\ \boldsymbol{\mu}_{\mathbf{y},k} \end{pmatrix}, \begin{pmatrix} \mathbf{R}_{\mathbf{s}_2,k} & \mathbf{R}_{\mathbf{s}_2,k,\mathbf{y}} \\ \mathbf{R}_{\mathbf{s}_2,k,\mathbf{y}}^H & \mathbf{R}_{\mathbf{y},k} \end{pmatrix} \right], \quad (11)$$

where

$$\boldsymbol{\mu}_{\mathbf{y},k} = \begin{bmatrix} [\mathbf{L}_{i_1}^H, \dots, \mathbf{L}_{i_p}^H]^H \boldsymbol{\mu}_k \\ 0 \end{bmatrix}, \quad (12)$$

$$\mathbf{R}_{\mathbf{s}_2,k,\mathbf{y}} = [(\mathbf{R}_{\mathbf{s}_2,k}) [\mathbf{L}_{i_1}^H, \dots, \mathbf{L}_{i_p}^H], \mathbf{A} \mathbf{R}_s \mathbf{L}_\eta^H], \quad (13)$$

$$\mathbf{R}_{\mathbf{s}_2,k} = \mathbf{R}_{\mathbf{w}_k} + \mathbf{A} \mathbf{R}_s \mathbf{A}^H. \quad (14)$$

If $\mathbf{y}_1$ is a $N \times 1$ vector, then $\mathbf{R}_{\mathbf{y}}$ is a $N(p+1) \times N(p+1)$ matrix whose mn -th block is given by

$$\mathbf{R}_{\mathbf{y},k,m,n} = \mathbf{L}_{i_m} \mathbf{R}_{\mathbf{s}_2,k} \mathbf{L}_{i_n}^H + \mathbf{R}_n \delta(m-n), \quad 1 \leq m, n \leq p$$

$$\mathbf{R}_{\mathbf{y},k,m,p+1} = \mathbf{R}_{\mathbf{y},k,p+1,m}^H = \mathbf{L}_{i_m} \mathbf{A} \mathbf{R}_s \mathbf{L}_\eta^H, \quad 1 \leq m \leq p.$$

$$\mathbf{R}_{\mathbf{y},k,p+1,p+1} = \mathbf{L}_\eta \mathbf{R}_s \mathbf{L}_\eta^H + \mathbf{R}_n.$$

Since the random vectors $\mathbf{s}_2, \mathbf{y}_2^{i_1}, \dots, \mathbf{y}_2^{i_p}, \mathbf{y}_1$ are jointly Gaussian, the conditional mean of $\mathbf{s}_2$ given $\mathbf{y}$ and $I_2 = k$ can be evaluated as

$$\begin{aligned} \mathbb{E} [\mathbf{s}_2 | \mathbf{y}_2^{i_1}, \dots, \mathbf{y}_2^{i_p}, \mathbf{y}_1, I_2 = k] \\ = \boldsymbol{\mu}_k + \mathbf{R}_{\mathbf{s}_2, k, \mathbf{y}} \mathbf{R}_{\mathbf{y}, k}^{-1} (\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}, k}), \end{aligned}$$

and the conditional mean estimator is

$$\begin{aligned} \mathbb{E} [\mathbf{s}_2 | \mathbf{y}_2^{i_1}, \dots, \mathbf{y}_2^{i_p}, \mathbf{y}_1] = \sum_{k=0}^1 \mathbb{E} [\mathbf{s}_2 | \mathbf{y}_2^{i_1}, \dots, \mathbf{y}_2^{i_p}, \mathbf{y}_1, I_2 = k] \\ \mathbb{P}(I_2 = k | \mathbf{y}_2^{i_1}, \dots, \mathbf{y}_2^{i_p}, \mathbf{y}_1), \quad (15) \end{aligned}$$

where the conditional probability of I_2 can be found using Bayes formula:

$$\begin{aligned} \Pi_k(\mathbf{y}) &= \mathbb{P}(I_2 = k | \mathbf{y}) \\ &= \mathbb{P}(I_2 = k | \mathbf{y}_2^{i_1}, \dots, \mathbf{y}_2^{i_p}, \mathbf{y}_1) \\ &= \frac{f(\mathbf{y} | I_2 = k) \mathbb{P}(I_2 = k)}{\sum_i f(\mathbf{y} | I_2 = i) \mathbb{P}(I_2 = i)}, \quad (16) \end{aligned}$$

where $f(\mathbf{y} | I_2 = k)$

$$= \frac{|\mathbf{R}_{\mathbf{y}, k}|^{-0.5}}{(\sqrt{2\pi})^{N/2}} \exp \left(-0.5(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}, k})^H \mathbf{R}_{\mathbf{y}, k}^{-1} (\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}, k}) \right)$$

and $\mathbb{P}(I_2 = 1) = q$. Thus equation (15) can be rewritten as

$$\mathbb{E} [\mathbf{s}_2 | \mathbf{y}] = \sum_{k=0}^1 \Pi_k(\mathbf{y}) (\boldsymbol{\mu}_k + \mathbf{R}_{\mathbf{s}_2, k, \mathbf{y}} \mathbf{R}_{\mathbf{y}, k}^{-1} (\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}, k})). \quad (17)$$

The MSE criterion in (9) can now be evaluated by substituting for the conditional expectation from (17). For the suboptimal criterion in (10), we need to find $\mathbb{E} [\mathbf{s}_2 | \mathbf{y}_2^i, \mathbf{y}_1]$ which is a specific case of (17) with $p = 1$, i.e., $\mathbf{y} = [\mathbf{y}_2^i, \mathbf{y}_1]$. It is worthwhile to note that even in the case of $q = 0$ or 1 for which the target dynamics are linear Gaussian, the solution to (10) is suboptimal, i.e., it is not equivalent to the conditional expectation (3). This is because the predictor does not take into account the spatial correlation between the received signals $\mathbf{y}_2^1, \dots, \mathbf{y}_2^M$. However, if the received signals are scalars, then the l_1 -norm constrained solution to (10) can be shown to be the optimal solution for the Gaussian case.

5. COMPUTATIONAL COMPLEXITY

The estimator given in (17) is in closed-form and hence the major complexity in finding the optimal solution is in its evaluation for all $\binom{M}{p}$ possible combinations of waveforms. Instead we use the suboptimal solution given by (10) to find the best p waveforms to be transmitted at the second stage. We use the recently proposed LARS algorithm (Least Angle Regression) [10] to solve for (10) which requires only the same order of magnitude of computational effort as the ordinary least squares solution. The algorithm uses the fact that the solution to (10) is piecewise linear in β and hence one can obtain the exact solution in $\min(p, M - p)$ steps either by doing a forward selection or backward elimination procedure.

6. SIMULATION RESULTS

Based on the formulation in Section 4, we perform a simulation for the simple case of $M = 5$ different waveforms. This will allow us to quantify the gap between the optimal solution (3) and the solution to the approximation (10). We assume a radar receiver array with $N = 25$ antenna elements so that the received signals $\mathbf{y}_1, \mathbf{y}_2$ are 25×1 vectors. The state vector is assumed to be a $N_s \times 1$ vector with $N_s = 10$. The correlation matrices $\mathbf{R}_n, \mathbf{R}_{\mathbf{w}_0}, \mathbf{R}_{\mathbf{w}_1}, \mathbf{R}_s$ are identity matrices. The mean vectors $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$ are 10×1 vectors consisting of all zeros and all 0.1 respectively. The Bernoulli random variables I_t takes the value 1 with probability $q = 0.4$. We assumed the channel model to be linear and selected the waveforms $\{\phi_i\}_{i=1}^M$ at random over 25 dimensional unit sphere. These waveforms are unit norm and have cross correlation less than 0.1. We simulated the performance of the optimal subset selector along with the l_1 -norm constrained convex problem under this setting. The performance criteria considered in the simulations is shown in Table 1.

We first present the MSE of the l_1 -norm penalized solution found from (5) (solid line, GAM with l_1) as a function of the sparseness regularization parameter β in Fig. 1. For each value of β , we also show the corresponding l_0 -norm of optimal γ (on top of the solid line) in the figure. The MSE is an increasing function of β and as explained earlier, we notice that increasing β induces more sparseness in the solution. When β is large, the MSE converges to the variance of the state parameter. We also plot the MSE of the optimal subset selection solution (dashed line) corresponding to the l_0 -norm obtained through the l_1 -norm constrained solution. We see a clear difference in performance between the two techniques. This is because of two main reasons: The primary reason is the fact that we find a suboptimal solution by assuming the GAM estimator of the form in (5) rather than the optimal estimator given in (3). The other reason is due to the fact that we solve the minimization problem subject to an l_1 -norm constraint rather than an l_0 -norm constraint.

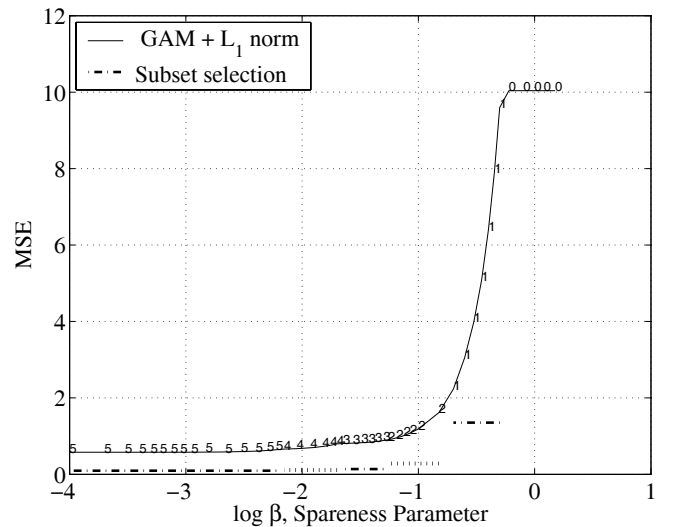


Fig. 1. Minimum MSE for optimal subset selection (dotted and dashed-dotted line) and l_1 -norm constrained solution (solid line) with respect to β . $\|\gamma\|_0$, corresponding to the number of nonzero components in the optimal solution of γ for constrained optimization is shown adjacent to the solid line as a function of β .

Approach	Form of predictor	Constr.
Subset Selection	$E[s_2 y_2^{i_1}, \dots, y_2^{i_p}, y_1]$	-
GAM + l_0	$\sum_i \gamma_i E[s_2 y_2^i, y_1]$	$\ \gamma\ _0$
GAM + l_1	$\sum_i \gamma_i E[s_2 y_2^i, y_1]$	$\ \gamma\ _1$
Optimal predictor	$E[s_2 y_2^1, \dots, y_2^M, y_1]$	-
Proposed Solution	Use optimal from GAM + l_1 in subset selection	$\ \gamma\ _1$

Table 1. Form of predictors

In Fig. 2, we plot the performance of state estimators mentioned in Table 1. We observe that the performance of GAM under l_0 -norm constraint is indeed found to be optimal for $\|\gamma\|_0 = 1$ case and clearly suboptimal for other cases due to the restrictive additive model. Finally we see that our proposed solution has a significant performance gain as compared to the simple l_1 -norm constrained minimization and approaches the optimal subset selection performance. This suggests that we can considerably reduce the computational complexity of the problem and at the same time achieve nearly optimal performance using such a design approach.

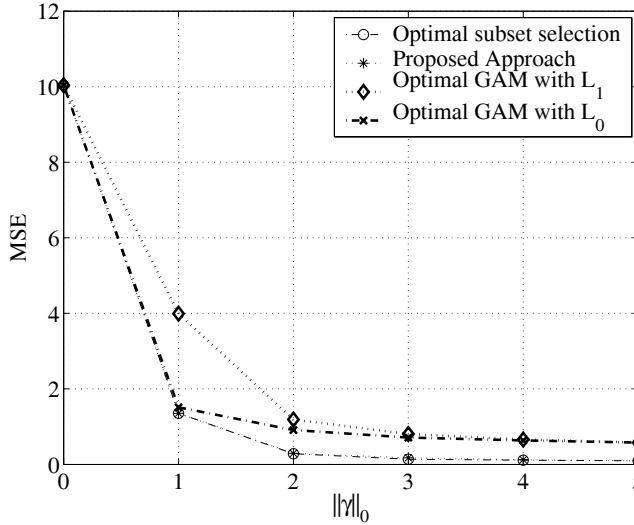


Fig. 2. Minimum MSE for the optimal subset selection problem (circle), optimal GAM with l_1 constraint (diamond), optimal GAM with l_0 constraint (cross) and the proposed approach (star) as a function of $\|\gamma\|_0$.

7. CONCLUSIONS

We considered the problem of optimal waveform selection. We optimally choose a small subset of waveforms that minimizes the state prediction MSE given the past observations. We observe that the

optimal subset selection is a combinatorially complex optimization problem and hence infeasible. We proposed a suboptimal solution through convex relaxation which achieves near optimal performance. We considered a particular model and compared the performance of the various strategies through simulation. This problem is a natural extension to the problem of optimal energy allocation between two stages of transmission under energy constraints using sequential design strategies [12, 13]. One extension is to solve this problem simultaneously for both optimal waveform selection and optimal energy allocation.

8. REFERENCES

- [1] M. R. Bell, "Information theory and radar waveform design," *IEEE Trans. on Inform. Theory*, vol. 39, no. 5, pp. 1578–1597, Sep. 1993.
- [2] S. M. Sowelam and T. H. Tewfik, "Waveform selection in radar target classification," *IEEE Trans. on Inform. Theory*, vol. 46, no. 3, pp. 1014–1029, May 2000.
- [3] S. M. Sowelam and T. H. Tewfik, "Optimal waveforms for wideband radar imaging," *J. Franklin Institute - Engg. App. Math.*, vol. 335B, no. 8, pp. 1341–1366, Nov. 1998.
- [4] R. E. Blahut, W. M. Miller, and C. H. Wilcox, *Radar and Sonar, Part I*, Springer-Verlag, New York, 1991.
- [5] H. Naparst, "Dense target signal processing," *IEEE Trans. on Inform. Theory*, vol. IT-37, no. 2, pp. 317–327, Mar. 1991.
- [6] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.
- [7] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. R. Statist. Soc.*, vol. 58, pp. 267–288, Nov. 1996.
- [8] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Comm. Pure App. Math.*, vol. 57, no. 11, pp. 1413–1457, Nov. 2004.
- [9] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, Springer Series in Statistics. Springer Verlag, New York, 2000.
- [10] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Ann. Statist.*, vol. 32, no. 2, pp. 407–499, 2004.
- [11] M. Osborne, B. Presnell, and B. Turlach, "A new approach to variable selection in least squares problems," *IMA J. Numeric. Anal.*, vol. 20, no. 3, pp. 389–403, 2000.
- [12] R. Rangarajan, R. Raich, and A. O. Hero III, "Sequential design for a rayleigh inverse scattering problem," *Proc. IEEE Workshop on Stat. Signal Processing*, July 2005.
- [13] R. Rangarajan, R. Raich, and A. O. Hero III, "Optimal experimental design for an inverse scattering problem," *Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, vol. 4, pp. 1117–1120, 2005.