An Exact Solution for Sparse Sampling for Optimal Detection of Known Signals in Gaussian Noise

Kaushallya Adhikari, *Member, IEEE*, and Steven Kay, *Life Fellow, IEEE*

***Abstract*—Detection of known signals of interest that are embedded in colored noise involves whitening the received samples and matched-filtering. In many applications, due to computational constraints, it is critical to select only a subset of the received samples for detection. This letter addresses the problem of selecting only a given number of temporal or spatial samples while maximizing detection performance for deterministic signals in first-order autoregressive Gaussian noise. The direct solution of this entails a combinatorial search, where the deflection coefficient is evaluated for each possible combination of sparse samples. This approach is infeasible when the number of samples is large since the number of possible combinations increases factorially with the number of samples. We present an efficient method to whiten Gaussian noise samples and express deflection coefficient in a form that is amenable to dynamic programming. Exploiting dynamic programming, we specify a feasible and efficient procedure to find optimal sparse samples where the number of computational steps increases linearly with the number of samples. Also, conditions under which uniform sampling is optimal is given.**

***Index Terms*—Autoregressive process, deflection coefficient, dynamic programming, signal detection, sparse samples, whitening.**

I. Introduction

THE problem of selecting a subset of temporal or spatial samples from available data sets to optimize a parameter of interest has a rich history [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22]. In this research, we address the problem of selecting M samples out of N samples while maximizing detection in a binary hypothesis problem, where the signal of interest is deterministic and the noise is Gaussian. For this detection problem, maximizing detection is equivalent to maximizing the deflection coefficient [23].

A direct method to choose M samples out of $N > M$ samples while maximizing detection performance is the exhaustive combinatorial search over all $N!/(M!(N-M)!)$ possibilities. To see why the direct method could get infeasible, consider the deflection coefficient $d^2 = \mathbf{s}^T\mathbf{R}^{-1}\mathbf{s}$, where the $N \times 1$ vector $\mathbf{s}$ represents the Nyquist samples of the signal of interest, and the $N \times N$ matrix $\mathbf{R}$ is the noise covariance matrix. The matrix $\mathbf{R}^{-1}$ is needed when the noise is not white; this matrix effectively whitens the noise. For small values of N, this direct approach yields the solution in a reasonable number of steps. However, if we consider more practical values for N and M, the direct approach becomes infeasible. For instance, if we desire to select $M = 25$ samples out of $N = 50$ samples, we have to consider a formidable number of combinations: 1.2641×10^{14}. Furthermore, for each possible combination, a matrix inversion is required to whiten the received data, which is computationally expensive. We present an alternative method to solve this problem efficiently, where the number of possible combinations that needs to be analyzed increases linearly with N and M. The crux of our method is an alternative procedure to whiten colored noise and a novel way to express the deflection coefficient. The unique representation we provide for the deflection coefficient is amenable to *Dynamic Programming* (DP), which enables solving the sample selection problem in a time that increases linearly with M [24, pp. 37].

Some of the approaches that attempt to optimally choose M samples out of N samples or design optimal sampling patterns can be found in [15], [16], [17], [18], [19], [20], [21], [22]. They have met with varying degrees of success and cannot be said to be optimal in general. The rate of the convergence of the detection metric of the discrete-time samples to that of the continuous-time observations is derived in [15]. Their results are of an asymptotic nature and they do not provide any method to find the sampling pattern that maximizes the detection metric. We provide a procedure for selecting optimal samples from finite data samples. Also, we have implicitly assumed the availability of only Nyquist samples and hence a discrete-time noise model with no assumptions about the underlying continuous-time origin of the data. The sampling schemes in [16], [17] are applicable only when $N \rightarrow \infty$. [18] designs sparse sampling patterns that maximize I-divergence, J-divergence, Bhattacharya distance, and Chernoff distance. [19] considers Kullback-Leibler divergence (KLD) and Chernoff distance as the optimality criteria and proves that the problem is NP hard. They provide an algorithm to provide the suboptimal solution to the problem of sample selection while optimizing KLD and Chernoff distance. [20], [21], [22] use KLD and Bhattacharya distance as optimilaity criteria for Gaussian observations using submodular optimization. As a result, the approaches in [18], [19], [20], [21], [22] are suboptimal. Furthermore, the computation necessary to obtain the asymptotic solutions is considerable and not guaranteed to

Manuscript received 16 February 2023; revised 23 March 2023; accepted 29 March 2023. Date of publication 3 April 2023; date of current version 10 April 2023. This work was supported by the U.S. Office of Naval Research under Grant N00014-20-1-2820. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Xiangui Kang. *(Corresponding author: Kaushallya Adhikari.)*

The authors are with the Department of Electrical, Computer, and Biomedical Engineering, University of Rhode Island, Kingston, RI 02881 USA (e-mail: kadhikari@uri.edu; kay@ele.uri.edu).

Digital Object Identifier 10.1109/LSP.2023.3264106

lead to the true asymptotic solution. *Our approach applies to finite N, guarantees optimality, and is feasible for any value of N.* Also, our method uses an autoregressive (AR) noise model of order $p = 1$. With high enough p, an AR model approximates any reasonable noise process [25]. An AR(1) process is commonly used to model noise processes with a single peak in the power spectrum, which arise in applications such as sonar, radar, and speech processing [26], [27]. Details on estimation of the AR model for a given data set are given in [25]. Future work will attempt to extend our results to higher order AR processes. The overall contributions of this letter are

1) An alternative method to whiten the data for a signal embedded in AR(1) noise (Section II).
2) A novel representation of deflection coefficient that lends itself to DP (Section II).
3) A DP method to the find the subset of M samples that maximizes detection of deterministic signals (Section III).
4) An analytical method to find the subset of M samples that maximizes detection of known constant signals (Section IV-A).

Our method applies to both temporal and spatial sampling. We use the notation and terminology for temporal sampling.

II. Signal Model and an Alternative Expression for Deflection Coefficient

Consider the following discrete-time binary hypothesis problem for a known deterministic signal in noise given as

$$\mathcal{H}_0 : x[n] = w[n]$$
$$\mathcal{H}_1 : x[n] = s[n] + w[n], \text{ for } n = 1, 2, \ldots, N, \tag{1}$$

where $\mathcal{H}_0$ and $\mathcal{H}_1$ represent the null and alternative hypotheses, respectively, $w[n]$ is a zero-mean discrete-time AR(1) process which satisfies the first order difference equation

$$w[n] = -a[1]w[n-1] + u[n] \tag{2}$$

and $u[n]$ is white Gaussian noise (WGN) with variance σ_u^2 and mean 0. We desire to choose samples $\{x[n_1], x[n_2], \ldots, x[n_M]\}$ so that the detection performance is maximized. Without loss of generality, we assume that the samples are arranged in ascending order, i.e.,

$$1 \le n_1 < n_2 < \ldots < n_M \le N. \tag{3}$$

We have also assumed in (3) that the samples are distinct. The autocovariance sequence of $w[n]$ is [25]

$$r_w[k] = \frac{\sigma_u^2}{1 - a^2[1]}(-a[1])^{|k|} = r_w[0](-a[1])^{|k|} \tag{4}$$

and the autcorrelation coefficient sequence is

$$\rho[k] = \frac{r_w[k]}{r_w[0]} = (-a[1])^{|k|}. \tag{5}$$

The optimal detector consists of a whitening filter followed by a replica-correlator [23]. The detection metric is the deflection coefficient. To find the M samples that are optimal for detection, we need to evaluate the deflection coefficient, which entails computing the inverse of a covariance matrix. To circumvent the covariance matrix inversion and facilitate a closed form representation of the deflection coefficient, we employ an alternative approach to whiten the data samples. This approach to whiten noise has been exploited by Grenander in [28, pg. 118] to find the Radon-Nikodym derivative for a continuous-time problem. Using this approach, as derived in Appendix, al alternative expression for the deflection coefficient is

$$d_d^2 = \frac{s^2[n_1]}{r_w[0]} + \sum_{i=1}^{M-1} \frac{(s[n_{i+1}] - \rho[n_{i+1} - n_i]s[n_i])^2}{r_w[0]\left(1 - \rho^2[n_{i+1} - n_i]\right)}. \tag{6}$$

III. DP to Select Samples That Maximize d_d^2

Expressing deflection coefficient in the form given in (6) reveals its special structure. The computation of each term in the summation of (6) involves two successive samples $s[n_{i+1}]$ and $s[n_i]$. Thus, the terms in d_d^2 are linked in a particular way. This feature of d_d^2 is the result of the Markovian property of the AR(1) noise process. The special structure in d_d^2 makes the evaluation of d_d^2 amenable to DP, where the computation time increases linearly, instead of exponentially, with the number of terms.

To maximize d_d^2 over the integer variables $n_1, n_2, \ldots, n_M$, we frame the maximization problem as

$$I_{M-1}(n_M) = \max_{\substack{n_1, n_2, \ldots, n_{M-1} \\ 1 \le n_1 < n_2 < \ldots < n_M \le N}} \Bigg[g(0, n_1) + \sum_{i=1}^{M-2} g(n_i, n_{i+1}) + g(n_{M-1}, n_M) \Bigg], \tag{7}$$

where $g(n_i, n_{i+1}) = \dfrac{(s[n_{i+1}] - \rho[n_{i+1} - n_i]s[n_i])^2}{r_w[0](1 - \rho^2[n_{i+1} - n_i])}$ and $g(0, n_1) = \dfrac{s^2[n_1]}{r_w[0]}$. Then, the maximum value of d_d^2 is

$$\max_{n_M} I_{M-1}(n_M). \tag{8}$$

To express (7) in a recursive form, we rewrite it as

$$I_{M-1}(n_M) = \max_{n_{M-1}} \Bigg(\Bigg\{ \max_{n_1, n_2, \ldots, n_{M-2}} \Bigg[g(0, n_1) + \sum_{i=1}^{M-3} g(n_i, n_{i+1}) + g(n_{M-2}, n_{M-1}) \Bigg\} + g(n_{M-1}, n_M) \Bigg). \tag{9}$$

Since the term inside the braces can be expressed as $I_{M-2}(n_{M-1})$, we have

$$I_{M-1}(n_M) = \max_{n_{M-1}} \left(I_{M-2}(n_{M-1}) + g(n_{M-1}, n_M) \right). \tag{10}$$

The successive backward recursion provides the expressions

$$I_{M-2}(n_{M-1}) = \max_{n_{M-2}} \left(I_{M-3}(n_{M-2}) + g(n_{M-2}, n_{M-1}) \right)$$
$$\vdots$$
$$I_2(n_3) = \max_{n_2} \left(I_1(n_2) + g(n_2, n_3) \right)$$
$$I_1(n_2) = \max_{n_1} \left(I_0(n_1) + g(n_1, n_2) \right). \tag{11}$$

The initialization expression is $I_0(n_1) = g(0, n_1)$. To find the optimal indices n_1^*, n_2^*, …, n_M^*, we first evaluate $I_0(n_1)$ for each possible value of n_1. The possible values of n_1 are $1 \le$

$n_1 \leq (N-M)+1$. The next step is to evaluate $I_1(n_2)$ for each possible value of n_2. The possible values of n_2 are $2 \leq n_2 \leq (N-M)+2$. Similarly, $I_2(n_3)$, $I_3(n_4)$, …, $I_{M-1}(n_M)$ are computed. The possible values for $n_3, n_4, \ldots, n_M$ are $3 \leq n_3 \leq (N-M)+3$, $4 \leq n_4 \leq (N-M)+4$, …, $M \leq n_M \leq N$, respectively.

Finally, the values for n_M^*, n_{M-1}^*, …, n_1^*, respectively, are found through *backtracking*. The first step in backtracking is finding the solution of $\max_{n_M} I_{M-1}(n_M)$, which yields n_M^*. Then, n_{M-1}^* is found by solving

$$\max_{n_{M-1}} I_{M-2}(n_{M-1}). \tag{12}$$

Similarly, n_{M-2}^*, n_{M-3}^*, …, n_2^*, n_1^*, are sequentially computed by maximizing $I_{M-3}(n_{M-2})$, $I_{M-4}(n_{M-3})$, …, $I_1(n_2)$, $I_0(n_1)$, respectively.

A. Detection of Constant Signals

Consider the maximization of (6) for $s[n]=A$ over the M sparse samples. Then,

$$\begin{aligned} d_d^2 &= \frac{A^2}{r_w[0]} + \frac{A^2}{r_w[0]} \sum_{i=1}^{M-1} \frac{1-\rho[n_{i+1}-n_i]}{1+\rho[n_{i+1}-n_i]} \\ &= \frac{A^2}{r_w[0]} + \frac{A^2}{r_w[0]} \sum_{i=1}^{M-1} \frac{1-(-a[1])^{n_{i+1}-n_i}}{1+(-a[1])^{n_{i+1}-n_i}} \end{aligned}$$

where it is assumed that $0 < -a[1] < 1$ (the noise power spectral density is lowpass). Now let (with a slight abuse of notation)

$$g(n_{i+1}-n_i) = \frac{1-(-a[1])^{n_{i+1}-n_i}}{1+(-a[1])^{n_{i+1}-n_i}}$$

and therefore we wish to maximize

$$J(n_1, n_2, \ldots, n_M) = \sum_{i=1}^{M-1} g(n_{i+1}-n_i).$$

It is easily shown that for

$$g(t) = \frac{1-(-a[1])^t}{1+(-a[1])^t}$$

defined over $1 \leq t \leq T$, where t is a continuous variable that $g(t)$ is a *strictly increasing function* as well as a *strictly concave function* of t. With these observations we have the following theorem.

Theorem 1 (Optimality of uniform samples for DC Level Signal): Assume that $g(t)$ is a strictly increasing and strictly concave function over the continuous interval $[1, T]$. To maximize

$$J(t_1, t_2, \ldots, t_M) = \sum_{i=1}^{M-1} g(t_{i+1}-t_i)$$

for $1 \leq t_1 \leq t_2 \leq \cdots \leq t_M \leq T$, one should choose $t_i^* = 1 + (i-1)\Delta^*$, where $\Delta^* = (T-1)/(M-1)$. Thus, the optimal sparse samples will be uniformly spaced. Furthermore, assuming that T is an integer, denoted by N and that Δ^* is an integer, the optimal samples will be at the integer samples, $n_i^* = 1 + (i-1)(N-1)/(M-1)$. Note that the first samples is at $n_1^* = 1$ and the last sample is at $n_M^* = N$.

Proof: Using the one-to-one transformation, $\Delta_1 = t_2 - t_1 \geq 0, \ldots, \Delta_{M-1} = t_M - t_{M-1} \geq 0$ and keeping t_1 unchanged yields

$$J'(t_1, \Delta_1, \ldots, \Delta_{M-1}) = \sum_{i=1}^{M-1} g(\Delta_i),$$

where the constraints are now $t_1 \geq 1$ and since $t_M \leq T$, we have equivalently $t_M = t_1 + \sum_{i=1}^{M-1} \Delta_i \leq T$. But since $\sum_{i=1}^{M-1} \Delta_i \leq T - t_1$, and $g(\cdot)$ is strictly monotonically increasing, the maximizing value for t_1 must be its minimum value or $t_1^* = 1$. Thus, we need to maximize $\sum_{i=1}^{M-1} g(\Delta_i)$ over only the Δ_i. Since $g(\cdot)$ is strictly concave, we have that for all $\lambda_i \geq 0$ and $\sum_{i=1}^{M-1} \lambda_i = 1$ that

$$\sum_{i=1}^{M-1} \lambda_i g(\Delta_i) \leq g\left(\sum_{i=1}^{M-1} \lambda_i \Delta_i\right)$$

and letting $\lambda_i = 1/(M-1)$ yields

$$\sum_{i=1}^{M-1} g(\Delta_i) \leq (M-1) g\left(\frac{1}{M-1} \sum_{i=1}^{M-1} \Delta_i\right)$$

with equality if and only if $\Delta_1 = \Delta_2 = \cdots = \Delta_{M-1} = \Delta^*$. And finally, if the sample times are integers, which will be the case if $\Delta^* = (T-1)/(M-1) = (N-1)/(M-1)$ is an integer, then the optimal sample times are as given in the theorem.

IV. Simulations

A. Detection of Sinusoidal Signals

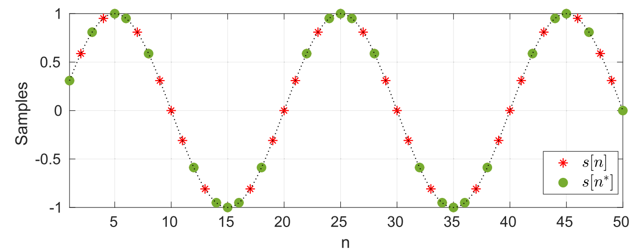

Fig. 1. Optimal $M = 25$ samples out of $N = 50$ samples for detection of a sinusoidal signal in colored Gaussian noise.

Consider a sinusoidal signal $s[n] = A\sin(2\pi f_0 n)$, for $n = 1, 2, \ldots, N$. Let $A = 1$, $f_0 = 0.05$ Hz, $\rho = 0.88$, $N = 50$, and $M = 25$. The signal samples are depicted in Fig. 1. The number of ways in which we can select 25 samples out of 50 samples is staggeringly large—1.2614×10^{14}. Thus, brute force solution of this problem requires tremendous computing resources. However, with DP, we obtain the precise solution for this problem in 2.9 ms using MATLAB on an Intel(R) Core(TM) i7-8750H CPU @ 2.20 GHz machine. The optimal 25 samples are depicted in Fig. 1.

Fig. 2 plots the receiver operation characteristic (ROC) curve for $M = 20$, obtained using 100,000 realizations of the test statistic for each hypothesis. The ROC curve obtained using the theoretical value of d_d^2 aligns with the empirical ROC curve. The figure also plots the ROC curve corresponding to the naive sampling approach, where M samples with the largest signal powers are selected for detection. The optimal approach outperforms the naive approach as evidenced by the figure.

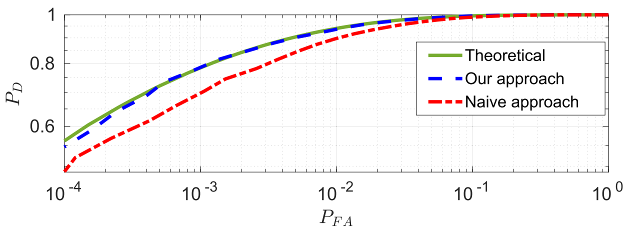

Fig. 2. Comparison of ROC curves for $M = 20$.

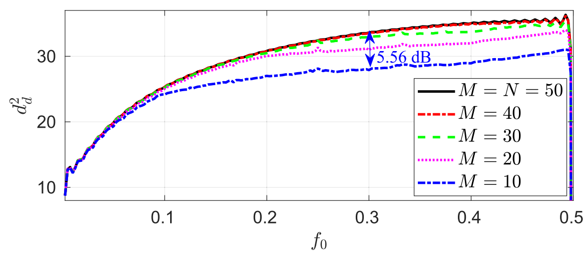

Fig. 3. Variation of the detection metric with f_0 for 10, 20, 30, and 40 samples out of $N = 50$ samples.

In the next example for $s[n] = A\sin(2\pi f_0 n)$, we observe the variation of d_d^2 with frequency for fixed $N = 50$. We consider the frequency in the range $0.0025 \leq f_0 \leq 0.5$. d_d^2 increases with f_0 and M as evidenced by Fig. 3. When f_0 is an integer multiple of 0.5, $s[n] = 0$ and $d_d^2 = 0$, as expected. Our approach allows for the computation of the detection loss when using $M < N$ optimal samples. As depicted in the figure, detection loss in using 10 optimal samples out of 50 at $f_0 = 0.3$ is 5.56 dB.

V. Conclusion

We presented an alternative method to whiten sparse data samples for additive AR(1) colored Gaussian noise. The presented method is computationally efficient since it does not require the matrix inversion that is involved in whitening using the standard technique. We then used DP to reduce the number of computational steps in selecting M samples out of N samples to maximize detection. The computing time is proportional to M and thereby replaces infeasible, expensive exhaustive searches by an efficient algorithm. This algorithm requires observations $x[n]$ corresponding to optimal sampling points only. We demonstrated examples with sinusoidal signals and analyzed the variation of deflection coefficient with frequency. Finally, we provided an analytical result to choose optimal M samples out of N samples for detection of constant signals where $(N-1)/(M-1)$ is an integer.

Appendix
Derivation of the Alternative Expression for d_d^2

Consider the following stochastic variables

$$x[n_1], y[n_1], y[n_2], \ldots, y[n_M], \tag{13}$$

where $y[n_m] = x[n_m] - \rho[n_m - n_{m-1}]x[n_{m-1}]$. The variables in (13) are uncorrelated. To see this, we compute $\text{cov}(y[n_{m+h}], y[n_m])$, which can be expressed as

$$\begin{aligned}\text{cov}\left(y[n_{m+h}], y[n_m]\right) &= r_w[0]\left[(-a[1])^{n_{m+h}-n_m}\right.\\ &\quad + (-a[1])^{n_{m+h}+n_m-2n_{m-1}} - (-a[1])^{n_{m+h}-n_m}\\ &\quad \left.-(-a[1])^{n_m-n_{m-1}+n_{m+h}-n_{m-1}}\right].\end{aligned} \tag{14}$$

Substituting $h = 0$ in (14), we obtain the variance of $y[n_m]$ for $m = 1, 2, \ldots, M$, which is

$$\text{var}(y[n_m]) = r_w[0]\left(1 - \rho^2[n_m - n_{m-1}]\right). \tag{15}$$

Substituting $h > 0$ in (14), we get $\text{cov}(y[n_{m+1}], y[n_m]) = 0$. Thus, the variables $y[n_1], y[n_2], \ldots, y[n_M]$ are uncorrelated. To see that they are uncorrelated from $x[n_1]$, we compute the covariance between $x[n_1]$ and $y[n_m]$ as

$$\begin{aligned}&\text{cov}(x[n_1], y[n_m])\\ &= \text{cov}(x[n_1], x[n_m]) - \rho[n_m - n_{m-1}]\,\text{cov}(x[n_1], x[n_{m-1}])\\ &= r_w[0]\left[(-a[1])^{n_m-n_1} - (-a[1])^{n_m-n_{m-1}+n_{m-1}-n_1}\right] = 0.\end{aligned} \tag{16}$$

Therefore, the variables in (13) are uncorrelated. We divide the variables in (13) by their standard deviations to obtain the following whitened samples

$$\begin{aligned}\epsilon[n_1] &= \frac{x[n_1]}{\sqrt{r_w[0]}}\\ \epsilon[n_2] &= \frac{x[n_2] - \rho[n_2 - n_1]x[n_1]}{\sqrt{r_w[0](1-\rho^2[n_2 - n_1])}}\\ \epsilon[n_3] &= \frac{x[n_3] - \rho[n_3 - n_2]x[n_2]}{\sqrt{r_w[0](1-\rho^2[n_3 - n_2])}}\\ \vdots &= \vdots\\ \epsilon[n_M] &= \frac{x[n_M] - \rho[n_M - n_{M-1}]x[n_{M-1}]}{\sqrt{r_x[0](1-\rho^2[n_M - n_{M-1}])}}.\end{aligned} \tag{17}$$

In (17) converts the sequence of M non-white samples $\{x[n_1], x[n_2], \ldots, x[n_M]\}$ to a sequence of M white samples $\{\epsilon[n_1]\,\epsilon[n_2]\,\ldots, \epsilon[n_M]\}$. Since the transformation in (17) is linear and $\mathbf{x}_s = [x[n_1]\,x[n_2]\ldots x[n_M]]^T$ is a Gaussian vector, the transformed vector $\boldsymbol{\epsilon} = [\epsilon[n_1]\,\epsilon[n_2]\,\ldots, \epsilon[n_M]]^T$ is also Gaussian with covariance $\mathbf{I}_M$. The mean of $\boldsymbol{\epsilon}$ is $\mathbf{0}$ under $\mathcal{H}_0$ and

$$\boldsymbol{\mu}_\epsilon = E\{\boldsymbol{\epsilon}\} = \begin{bmatrix}\dfrac{s[n_1]}{\sqrt{r_w[0]}}\\ \dfrac{s[n_2] - \rho[n_2 - n_1]s[n_1]}{\sqrt{r_w[0](1-\rho^2[n_2 - n_1])}}\\ \dfrac{s[n_3] - \rho[n_3 - n_2]s[n_2]}{\sqrt{r_w[0](1-\rho^2[n_3 - n_2])}}\\ \vdots\\ \dfrac{s[n_M] - \rho[n_M - n_{M-1}]s[n_{M-1}]}{\sqrt{r_w[0](1-\rho^2[n_M - n_{M-1}])}}\end{bmatrix} \tag{18}$$

under $\mathcal{H}_1$. Thus the distributions of the random vector $\boldsymbol{\epsilon}$ under $\mathcal{H}_0$ and $\mathcal{H}_1$ are $\mathcal{N}(\mathbf{0}, \mathbf{I}_M)$ and $\mathcal{N}(\boldsymbol{\mu}_\epsilon, \mathbf{I}_M)$, respectively. Therefore, the optimal detector is $T(\boldsymbol{\epsilon}) = \boldsymbol{\epsilon}^T\boldsymbol{\mu}_\epsilon > \gamma$ and the deflection coefficient is $d_d^2 = \boldsymbol{\mu}_\epsilon^T\boldsymbol{\mu}_\epsilon$ [23]. This is the detection metric also for the case when the amplitude of $s[n]$ is not known [23]. Using (18), an efficient closed form expression for deflection coefficient is given by (6).

REFERENCES

[1] A. Berman and C. S. Clay, "Theory of time averaged product arrays," *J. Acoust. Soc. Amer.*, vol. 29, no. 7, pp. 805–812, 1957.

[2] D. Davies and C. Ward, "Low sidelobe patterns from thinned arrays using multiplicative processing," *IEE Proc. F Commun., Radar Signal Process.*, vol. 127, no. 1, pp. 9–23, Feb. 1980.

[3] A. Maffett, "Array factors with nonuniform spacing parameter," *IRE Trans. Antennas Propag.*, vol. 10, no. 2, pp. 131–136, 1962.

[4] A. Moffet, "Minimum-redundancy linear arrays," *IEEE Trans. Antennas Propag.*, vol. 16, no. 2, pp. 172–175, Mar. 1968.

[5] S. Pillai and F. Haber, "Statistical analysis of a high resolution spatial spectrum estimator utilizing an augmented covariance matrix," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 35, no. 11, pp. 1517–1523, Nov. 1987.

[6] S. K. Mitra, K. Mondal, M. K. Tchobanou, and G. J. Dolecek, "General polynomial factorization-based design of sparse periodic linear arrays," *IEEE Trans. Ultrasonics, Ferroelect., Freq. Control*, vol. 57, no. 9, pp. 1952–1966, Sep. 2010.

[7] D. King, R. Packard, and R. Thomas, "Unequally-spaced, broad-band antenna arrays," *IRE Trans. Antennas Propag.*, vol. 8, no. 4, pp. 380–384, 1960.

[8] B. P. Kumar and G. R. Branner, "Design of unequally spaced arrays for performance improvement," *IEEE Trans. Antennas Propag.*, vol. 47, no. 3, pp. 511–523, Mar. 1999.

[9] S. Kay and S. Saha, "Design of sparse linear arrays by monte carlo importance sampling," *IEEE J. Ocean. Eng.*, vol. 27, no. 4, pp. 790–799, Oct. 2002.

[10] P. Vaidyanathan and P Pal, "Sparse sensing with co-prime samplers and arrays," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 573–586, Feb. 2011.

[11] K. Adhikari and B. Drozdenko, "Symmetry-imposed rectangular coprime and nested arrays for direction of arrival estimation with multiple signal classification," *IEEE Access*, vol. 7, pp. 153217–153229, 2019.

[12] P. Pal and P. Vaidyanathan, "Nested arrays: A novel approach to array processing with enhanced degrees of freedom," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4167–4181, Aug. 2010.

[13] K. Adhikari, "Beamforming with semi-coprime arrays," *J. Acoust. Soc. Amer.*, vol. 145, no. 5, pp. 2841–2850, 2019, doi. org/10.1121/1.5100281.

[14] K. Adhikari and S. Kay, "Optimal sparse sampling for detection of a known signal in nonwhite Gaussian noise," *IEEE Signal Process. Lett.*, vol. 28, pp. 1908–1912, 2021.

[15] S. Cambanis and E. Masry, "Sampling designs for the detection of signals in noise," *IEEE Trans. Inf. Theory*, vol. 29, no. 1, pp. 83–104, Jan. 1983.

[16] R. K. Bahr and J. A. Bucklew, "Optimal sampling schemes for the Gaussian hypothesis testing problem," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1677–1686, Oct. 1990.

[17] Y. Sung, X. Zhang, L. Tong, and H. V. Poor, "Sensor configuration and activation for field detection in large sensor arrays," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 447–463, Feb. 2008.

[18] ang Yu and P. K. Varshney, "Sampling design for Gaussian detection problems," *IEEE Trans. Signal Process.*, vol. 45, no. 9, pp. 2328–2337, Sep. 1997.

[19] D. Bajovic, B. Sinopoli, and J. Xavier, "Sensor selection for event detection in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4938–4953, Oct. 2011.

[20] S. P. Chepuri and G. Leus, "Sparse sensing for distributed Gaussian detection," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2015, pp. 2394–2398.

[21] S. P. Chepuri and G. Leus, "Sparse sensing for distributed detection," *IEEE Trans. Signal Process.*, vol. 64, no. 6, pp. 1446–1460, Mar. 2016.

[22] M. Coutino, S. P. Chepuri, and G. Leus, "Submodular sparse sensing for Gaussian detection with correlated observations," *IEEE Trans. Signal Process.*, vol. 66, no. 15, pp. 4025–4039, Aug. 2018.

[23] S. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, vol. II. Englewood Cliffs, NJ, USA: Prentice Hall, 1998.

[24] R. E. Bellman and S. E. Dreyfus, *Applied Dynamic Programming*. Princeton, NJ, USA: Princeton Univ. Press, 1962.

[25] S. Kay, *Modern Spectral Estimation Theory and Application*. Englewood Cliffs, NJ, USA: Prentice Hall, 1988.

[26] R. Urick, *Principles of Underwater Sound*. New York, NY, USA: McGraw-Hill, 1983.

[27] S. Kay and J. Salisbury, "Improved active sonar detection using autoregressive prewhiteners," *J. Acoust. Soc. Amer.*, vol. 87, no. 4, pp. 1603–1611, 1990, doi. org/10.1121/1.399408.

[28] U. Grenander, *Abstract Inference*. Hoboken, NJ, USA: Wiley, 1981.