LIKELIHOOD UPDATING FOR GAUSS-GAUSS DETECTION

Nick Klausner and Mahmood Azimi-Sadjadi

Department of Electrical and Computer Engineering
Colorado State University
Electrical and Computer Engineering
Fort Collins, CO 80523-1373
azimi@engr.colostate.edu

ABSTRACT

This paper investigates the effects of incrementally adding new data to the classical Gauss-Gauss detector for testing between the known covariance matrices in competing multivariate models. We show that updating the likelihood ratio and J-divergence as a result of general data augmentation inherently involves linearly estimating the new data from the old. Using the change in divergence and the eigenstructure of a whitened error covariance matrix, a reduced-rank version of the update is built. A simulation example of a single narrow-band source in the sensing environment of multiple uniform linear arrays (ULA's) is given showing the practicality of adding data in multi-static sonar applications.

Index Terms— binary hypothesis testing, detection, likelihood updating, multi-static sonar

1. INTRODUCTION

In multivariate detection [1], updating the likelihood function and J-divergence, a measure of discrimination, when observations are augmented by new data is of great interest in order to decide in-situ if and when the detector reaches a point of diminishing return in adding sensory channels or which channels bring the largest increases in discriminatory information.

The evaluation of likelihood functions for the detection and parameter estimation of Gaussian signals observed in the presence of additive Gaussian noise is considered in [2]. The author developed a detector for a signal-plus-noise model both for discrete-time and continuous-time cases and derived the updating equations for a signal modeled as a Markov process. Similarly, in [3] a causal estimator-correlator version of the discrete-time Gauss-Markov likelihood ratio is derived that replaces the least squares estimator with one which is locally stable. This locally stable estimate is shown to be a compromise between the *a priori* and *a posteriori* state estimators. When the observation noise covariance is unknown, an integral equation is derived by assigning a Wishart prior to the innovations covariance which results in an appealing parallel approximation of the likelihood ratio.

In this paper, we address the problem of updating the likelihood function and J-divergence by showing that, for a general augmentation in our observation, we can always update the likelihood ratio by adding quadratic error terms produced from a discrete Wiener filter. The change in J-divergence is then shown to involve error covariance matrices that are matched/mismatched to the given hypothesis. This change gives us some insight as to how the performance of the detector incrementally changes as we add additional measurements to our observation. Additionally, when adding measurements from disparate sources of information, such as multiple sensor platforms, the change in J-divergence can be used to determine which platform's observation should be added or when adding observations from platforms reaches a point of diminishing return. Using a constructive procedure similar to that presented in [1] for building low-rank detectors, we extend the idea to build reduced-rank approximations of the likelihood update. Simulation of a single narrow-band source in the sensing environment of multiple ULAs is then considered showing the practicality of analyzing the incremental addition of multiple observations for in-situ detection in multi-static sonar.

2. FULL-RANK LIKELIHOOD UPDATING

Consider the vector $\mathbf{z}_k = \left[\mathbf{x}_1^H \cdots \mathbf{x}_k^H\right]^H \in \mathbb{C}^m$ where $\mathbf{x}_i$ represents any arbitrary set of measurements, e.g. the i^{th} realization of a multi-variate time series, the observation from the i^{th} sensory channel, etc. Here, we assume that $\mathbf{z}_k \sim \mathcal{CN}\left(\mathbf{0}, R_{\mathbf{z}_k\mathbf{z}_{k_i}}\right)$, i.e. a zero-mean complex normal random vector with covariance matrix $R_{\mathbf{z}_k\mathbf{z}_{k_1}} = E_{H_1}\mathbf{z}_k\mathbf{z}_k^H$ under H_1 (e.g., signal plus noise) and $R_{\mathbf{z}_k\mathbf{z}_{k_0}} = E_{H_0}\mathbf{z}_k\mathbf{z}_k^H$ under H_0 (e.g., noise alone). Note that E_{H_1} and E_{H_0} denote expectation conditioned upon the H_1 and H_0 hypotheses, respectively. We then consider adding the vector $\mathbf{x}_{k+1} \in \mathbb{C}^n$ to form

This work was supported by the Office of Naval Research, Code 321OE under contract N00014-09-1-0087.

the new augmented observation $\mathbf{z}_{k+1} = \left[\mathbf{z}_k^H \ \mathbf{x}_{k+1}^H\right]^H$ which we assume to remain complex normal with covariance structure $R_{\mathbf{z}_{k+1}\mathbf{z}_{k+1_1}}$ under H_1 and $R_{\mathbf{z}_{k+1}\mathbf{z}_{k+1_0}}$ under H_0. No specific structure is assumed for any covariance matrix with the only restriction that they be positive definite. To decide which models most likely generated the measurements $\mathbf{z}_k$ and $\mathbf{z}_{k+1}$, we construct the log-likelihood ratio functions representing simple quadratic detectors [1]

$$l(\mathbf{z}_k) = \mathbf{z}_k^H \left(R^{-1}_{\mathbf{z}_k\mathbf{z}_{k_0}} - R^{-1}_{\mathbf{z}_k\mathbf{z}_{k_1}}\right)\mathbf{z}_k$$

and

$$l(\mathbf{z}_{k+1}) = \mathbf{z}_{k+1}^H \left(R^{-1}_{\mathbf{z}_{k+1}\mathbf{z}_{k+1_0}} - R^{-1}_{\mathbf{z}_{k+1}\mathbf{z}_{k+1_1}}\right)\mathbf{z}_{k+1},$$

respectively. Taking advantage of the inherent block structure in these covariance matrices, we apply the matrix inversion identity [4]

$$R^{-1}_{\mathbf{z}_{k+1}\mathbf{z}_{k+1}} = \begin{bmatrix} R^{-1}_{\mathbf{z}_k\mathbf{z}_k} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} + \begin{bmatrix} -W^H \\ I \end{bmatrix} Q^{-1}\left[-W \ I\right]$$

where $W = R_{\mathbf{x}_{k+1}\mathbf{z}_k} R^{-1}_{\mathbf{z}_k\mathbf{z}_k}$ is a discrete Wiener filtering matrix that estimates the new data from the old and $Q = R_{\mathbf{x}_{k+1}\mathbf{x}_{k+1}} - R_{\mathbf{x}_{k+1}\mathbf{z}_k} R^{-1}_{\mathbf{z}_k\mathbf{z}_k} R^H_{\mathbf{x}_{k+1}\mathbf{z}_k}$ is its associated error covariance matrix. From here, it is trivial to show that the change in the log-likelihood ratio as a result of the augmentation becomes [2]

$$\begin{aligned} \Delta l\left(\mathbf{z}_{k+1}, \mathbf{z}_k\right) &= l\left(\mathbf{z}_{k+1}\right) - l\left(\mathbf{z}_k\right) \\ &= \mathbf{e}_0^H Q_0^{-1}\mathbf{e}_0 - \mathbf{e}_1^H Q_1^{-1}\mathbf{e}_1 \end{aligned} \quad (1)$$

where $W_0 = R_{\mathbf{x}_{k+1}\mathbf{z}_{k_0}} R^{-1}_{\mathbf{z}_k\mathbf{z}_{k_0}}$ and $W_1 = R_{\mathbf{x}_{k+1}\mathbf{z}_{k_1}} R^{-1}_{\mathbf{z}_k\mathbf{z}_{k_1}}$ are Wiener filters conditioned upon H_0 and H_1, respectively, and $\mathbf{e}_0 = \mathbf{x}_{k+1} - W_0\mathbf{z}_k$ and $\mathbf{e}_1 = \mathbf{x}_{k+1} - W_1\mathbf{z}_k$ are the error vectors produced by these smoothing matrices with covariance matrices

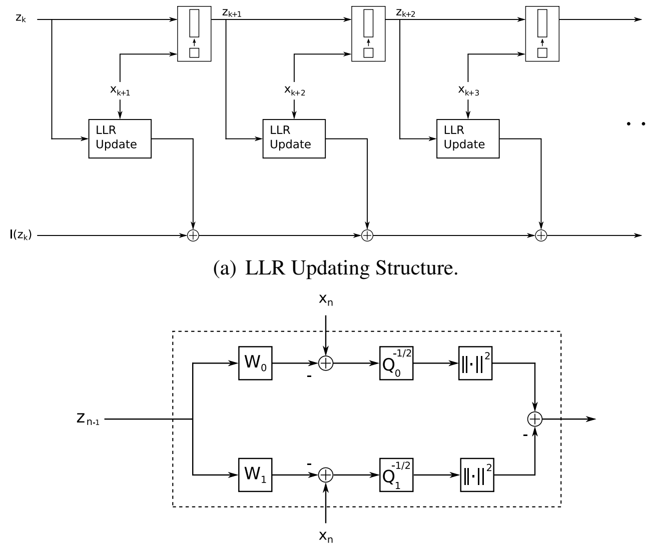
$$Q_0 = E_{H_0}\left[\mathbf{e}_0\mathbf{e}_0^H\right] = R_{\mathbf{x}_{k+1}\mathbf{x}_{k+1_0}} - R_{\mathbf{x}_{k+1}\mathbf{z}_{k_0}} R^{-1}_{\mathbf{z}_k\mathbf{z}_{k_0}} R^H_{\mathbf{x}_{k+1}\mathbf{z}_{k_0}}$$

and

$$Q_1 = E_{H_1}\left[\mathbf{e}_1\mathbf{e}_1^H\right] = R_{\mathbf{x}_{k+1}\mathbf{x}_{k+1_1}} - R_{\mathbf{x}_{k+1}\mathbf{z}_{k_1}} R^{-1}_{\mathbf{z}_k\mathbf{z}_{k_1}} R^H_{\mathbf{x}_{k+1}\mathbf{z}_{k_1}},$$

respectively. Figures 1(a) and 1(b) graphically depict this updating process wherein one forms a residual under each hypothesis by linearly estimating the additional channel from the previous channels, whitens with the appropriate error covariance, and finally updates the likelihood ratio by computing the energy in each whitened residual and forming their difference.

The J-divergence [1] provides a tractable measure of the amount of discriminatory information among the two hypotheses by simply measuring the difference in means of the log-likelihood ratio under both hypotheses, i.e. $J(\mathbf{z}) = E_{H_1} l(\mathbf{z}) - E_{H_0} l(\mathbf{z})$. Using the results given above, it is



(a) LLR Updating Structure.

(b) LLR Update Block.

Fig. 1. Log-Likelihood Ratio Updating.

easy to show that the change in divergence when adding the observation $\mathbf{x}_{k+1}$ becomes

$$\begin{aligned} \Delta J\left(\mathbf{z}_{k+1}, \mathbf{z}_k\right) &= J\left(\mathbf{z}_{k+1}\right) - J\left(\mathbf{z}_k\right) \\ &= \mathrm{tr}\left(-2I + Q_0^{-1}Q_{10} + Q_1^{-1}Q_{01}\right) \end{aligned} \quad (2)$$

where

$$\begin{aligned} Q_{10} &= E_{H_1}\left[\mathbf{e}_0\mathbf{e}_0^H\right] \\ &= R_{\mathbf{x}_{k+1}\mathbf{x}_{k+1_1}} - W_0 R^H_{\mathbf{x}_{k+1}\mathbf{z}_{k_1}} - R_{\mathbf{x}_{k+1}\mathbf{z}_{k_1}} W_0^H \\ &\quad + W_0 R_{\mathbf{z}_k\mathbf{z}_{k_1}} W_0^H \\ Q_{01} &= E_{H_0}\left[\mathbf{e}_1\mathbf{e}_1^H\right] \\ &= R_{\mathbf{x}_{k+1}\mathbf{x}_{k+1_0}} - W_1 R^H_{\mathbf{x}_{k+1}\mathbf{z}_{k_0}} - R_{\mathbf{x}_{k+1}\mathbf{z}_{k_0}} W_1^H \\ &\quad + W_1 R_{\mathbf{z}_k\mathbf{z}_{k_0}} W_1^H \end{aligned}$$

are cross terms representing the error covariance when using the wrong smoothing filter. That is, Q_{10} is the error covariance matrix incurred when filtering with W_0 given that it is actually the H_1 model that produced the data and vice versa for Q_{01}. Since we are filtering with a sub-optimal smoother in such situations, we have the following inequalities

$$\begin{aligned} \mathbf{x}^H Q_{10}\mathbf{x} &\geq \mathbf{x}^H Q_1\mathbf{x} \\ \mathbf{x}^H Q_{01}\mathbf{x} &\geq \mathbf{x}^H Q_0\mathbf{x} \end{aligned} \quad (3)$$

for any non-zero $\mathbf{x} \in \mathbb{C}^n$.

3. REDUCED-RANK LIKELIHOOD UPDATING

In situations where n is large, it is advantageous to employ reduced-rank methods to approximate the update. To achieve this, we use the same constructive procedure presented in [1] to define a change in coordinate system built from the eigenstructure of a normalized covariance matrix and use divergence as a performance criterion for rank reduction. We begin by removing the contribution from the H_0 hypothesis via a whitening procedure

$$\begin{aligned}\mathbf{w}_0 &= Q_0^{-1/2}\mathbf{e}_0 \;:\; E_{H_0}\mathbf{w}_0\mathbf{w}_0^H = I\\ \mathbf{w}_1 &= Q_0^{-1/2}\mathbf{e}_1 \;:\; E_{H_1}\mathbf{w}_1\mathbf{w}_1^H = \Gamma\end{aligned}$$

where $\Gamma = Q_0^{-1/2}Q_1Q_0^{-H/2}$ and matrix $Q_0^{1/2}$ satisfies the equality $Q_0 = Q_0^{1/2}Q_0^{H/2}$. The change in log-likelihood can then be written as

$$\Delta l\left(\mathbf{z}_{k+1}, \mathbf{z}_k\right) = \mathbf{w}_0^H\mathbf{w}_0 - \mathbf{w}_1^H\Gamma^{-1}\mathbf{w}_1$$

We then take the eigenvalue decomposition of matrix Γ such that $\Gamma = U\Sigma U^H$ for some orthonormal matrix $U = [\mathbf{u}_1 \;\cdots\; \mathbf{u}_n]$ and diagonal matrix $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$. The whitened versions of both error vectors are then represented in this new basis

$$\begin{aligned}\mathbf{y}_0 &= U^H\mathbf{w}_0 \;:\; E_{H_0}\mathbf{y}_0\mathbf{y}_0^H = I\\ \mathbf{y}_1 &= U^H\mathbf{w}_1 \;:\; E_{H_1}\mathbf{y}_1\mathbf{y}_1^H = \Sigma\end{aligned}$$

resulting in the change in log-likelihood ratio

$$\Delta l\left(\mathbf{z}_{k+1}, \mathbf{z}_k\right) = \mathbf{y}_0^H\mathbf{y}_0 - \mathbf{y}_1^H\Sigma^{-1}\mathbf{y}_1. \qquad (4)$$

To find the corresponding expressions for the change in J-divergence, we define the following two matrices

$$\begin{aligned}\Gamma_{10} &= E_{H_1}\left[Q_0^{-\frac{1}{2}}\mathbf{e}_0\mathbf{e}_0^HQ_0^{-\frac{H}{2}}\right] = Q_0^{-\frac{1}{2}}Q_{10}Q_0^{-\frac{H}{2}}\\ \Gamma_{01} &= E_{H_0}\left[Q_0^{-\frac{1}{2}}\mathbf{e}_1\mathbf{e}_1^HQ_0^{-\frac{H}{2}}\right] = Q_0^{-\frac{1}{2}}Q_{01}Q_0^{-\frac{H}{2}}\end{aligned}$$

Using the expression given in (2), the change in J-divergence can be written as

$$\Delta J\left(\mathbf{z}_{k+1}, \mathbf{z}_k\right) = \sum_{i=1}^{n} -2 + \mathbf{u}_i^H\left(\Gamma_{10} + \sigma_i^{-1}\Gamma_{01}\right)\mathbf{u}_i \qquad (5)$$

Therefore, to find the best low-rank approximation of Γ that maximizes the change in divergence, we find that it is not the value of σ_i that determines a dominant mode but rather the quadratic form $\mathbf{u}_i^H\left(\Gamma_{10} + \sigma_i^{-1}\Gamma_{01}\right)\mathbf{u}_i$. Using this decomposition of the divergence and the two inequalities given in (3), we can lower bound the change in divergence in terms of a sum only involving the eigenvalues of Γ

$$\Delta J\left(\mathbf{z}_{k+1}, \mathbf{z}_k\right) \geq \sum_{i=1}^{n} -2 + \sigma_i + \sigma_i^{-1},$$

and as the function $-2 + \sigma_i + \sigma_i^{-1}$ is non-negative definite for any $\sigma_i \geq 0$, it then follows as a corollary that $\Delta J\left(\mathbf{z}_{k+1}, \mathbf{z}_k\right) \geq 0$. Thus, adding data to this quadratic detector can never have a negative impact on divergence.

To perform the rank-p update, we decompose the coordinate system as follows

$$U = \left[\begin{array}{cc} U_p & U_{p+1}\end{array}\right]$$

where $U_p = [\mathbf{u}_1 \;\cdots\; \mathbf{u}_p]$ and $U_{p+1} = [\mathbf{u}_{p+1} \;\cdots\; \mathbf{u}_n]$. Also,

$$\Sigma = \left[\begin{array}{cc}\Sigma_p & \mathbf{O}\\ \mathbf{O} & \Sigma_{p+1}\end{array}\right]$$

where $\Sigma_p = \operatorname{diag}[\sigma_1 \;\cdots\; \sigma_p]$ and $\Sigma_{p+1} = \operatorname{diag}[\sigma_{p+1} \;\cdots\; \sigma_n]$. Without loss in generality, we assume that the coordinates are sorted in a descending fashion such that

$$\mathbf{u}_1^H\left(\Gamma_{10} + \sigma_1^{-1}\Gamma_{01}\right)\mathbf{u}_1 > \cdots > \mathbf{u}_n^H\left(\Gamma_{10} + \sigma_n^{-1}\Gamma_{01}\right)\mathbf{u}_n.$$

The filter $U_p^HQ_0^{-1/2} : \mathbb{C}^n \rightarrow \mathbb{C}^p$ yields the low-rank approximations of the error vectors as

$$\tilde{\mathbf{y}}_0 = U_p^HQ_0^{-1/2}\mathbf{e}_0 \;\text{ and }\; \tilde{\mathbf{y}}_1 = U_p^HQ_0^{-1/2}\mathbf{e}_1$$

Using (1), the change in log-likelihood in this reduced-rank subspace becomes

$$\Delta l_p\left(\mathbf{z}_{k+1}, \mathbf{z}_k\right) = \tilde{\mathbf{y}}_0^H\tilde{\mathbf{y}}_0 - \tilde{\mathbf{y}}_1^H\Sigma_p^{-1}\tilde{\mathbf{y}}_1 \qquad (6)$$

with an associated change in J-divergence

$$\Delta J_p((\mathbf{z}_{k+1}, \mathbf{z}_k) = \sum_{i=1}^{p} -2 + \mathbf{u}_i^H\left(\Gamma_{10} + \sigma_i^{-1}\Gamma_{01}\right)\mathbf{u}_i \qquad (7)$$

4. SIMULATION RESULTS

To demonstrate a situation where log-likelihood updating may be useful, we consider the problem of detecting the presence of a $10kHz$ source using multiple uniform linear arrays (ULAs). We assume that each platform individually pings the environment and collects far-field measurements via a 16-element ULA at a half-wavelength spacing of $7.5cm$. For this multi-static sonar simulation, we consider a situation where 20 ULAs are oriented in the same direction and are all above the seafloor at an elevation of $5m$. The platforms are uniformly spaced across a $40m$ distance in the cross-track direction and each platform is located $1m$ behind the platform to the left. We assume that the source is located at the origin and all platforms are moving in the same along-track direction with a speed of $1.5m/s$. A three-dimensional perspective of the problem setup is given in Figure 2.

Assuming that each platform synchronously pings the environment with the same transmit signal via a global clock and each array knows the location of the others, then it may be possible in such a situation to account and equalize the

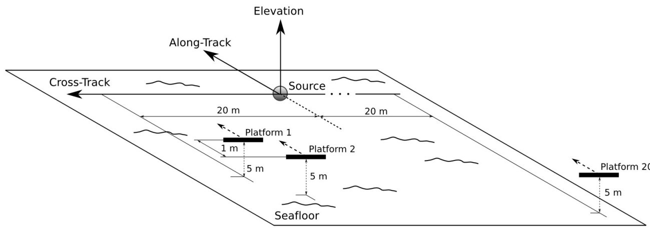


Fig. 2. Multi-Static Sonar Setup.

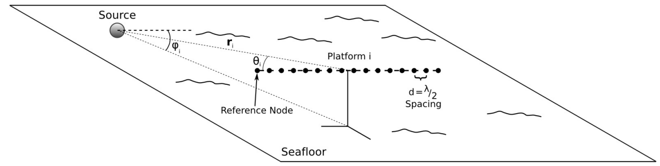


Fig. 3. Geometry of Multi-Platform Model.

effects of direct path propagation. We assume that such a situation exists and, for the k^{th} observation from the l^{th} array, consider the detection problem

$$\begin{aligned} H_1 &: \quad \mathbf{y}_l[k] = \mathbf{h}\left(\theta_l\right) \boldsymbol{\alpha}_l^T \mathbf{s}[k] + \mathbf{n}_l[k] \\ H_0 &: \quad \mathbf{y}_l[k] = \mathbf{n}_l[k] \end{aligned}$$

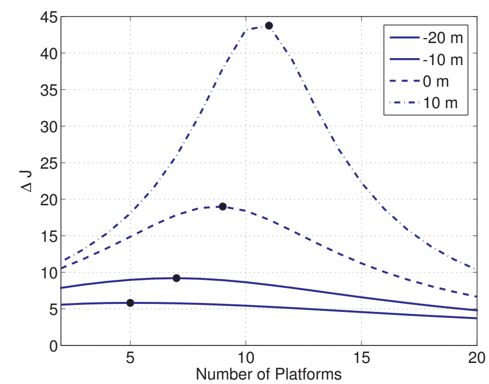
where $\mathbf{n}_l[k] \in \mathbb{C}^L$ is a zero-mean complex Gaussian random vector with covariance structure $E\left[\mathbf{n}_l[k]\mathbf{n}_i[k]^H\right] = \delta_{l-i}\sigma_n^2 I$ and $\mathbf{h}\left(\theta_l\right) = \left[1 \; e^{j\pi\cos(\theta_l)} \; \cdots \; e^{j(L-1)\pi\cos(\theta_l)}\right]^H \in \mathbb{C}^L$ is the steering vector of the ULA at direction of arrival (DOA) θ_l and at half-wavelength spacing ($d = \lambda/2$). Assuming that there are N platforms in the sensing environment, the vector $\boldsymbol{\alpha}_l \in \mathbb{R}^N$ is given to be $\boldsymbol{\alpha}_l = \left[(||\mathbf{r}_l|| + ||\mathbf{r}_1||)^{-1} \; \cdots \; (||\mathbf{r}_l|| + ||\mathbf{r}_N||)^{-1}\right]^T$ where each element is an attenuation (or fading) weighting assumed to be the inverse of the transmitter-receiver path-length and $\mathbf{r}_i$, $i = 1, \ldots, N$, is a vector describing the platform's location with respect to the source. Finally, the vector $\mathbf{s}[k] \in \mathbb{R}^N$ is a zero-mean Gaussian random vector containing the source-signals from each platform at the k^{th} observation and has covariance matrix $R_{\mathbf{s}} = E\left[\mathbf{s}[k]\mathbf{s}[k]^T\right]$ given by

$$R_{\mathbf{s}} = \sigma_s^2 \begin{bmatrix} 1 & \rho(\phi_1 - \phi_2) & \cdots & \rho(\phi_1 - \phi_N) \\ \rho(\phi_2 - \phi_1) & 1 & \cdots & \rho(\phi_2 - \phi_N) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(\phi_N - \phi_1) & \rho(\phi_N - \phi_2) & \cdots & 1 \end{bmatrix}$$

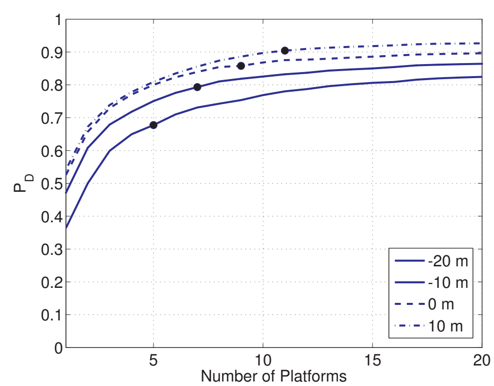
where ϕ_l is the aspect angle among the source and the l^{th} platform and $\rho(\cdot)$ is a correlation coefficient assumed to be a Gaussian function of the form $\rho(\phi_l - \phi_i) = e^{-\frac{(\phi_l - \phi_i)^2}{\Omega}}$ where Ω is a parameter loosely describing the target structure. The geometry of the problem being considered under H_1 is shown in Figure 3 for one array. For the sake of simplicity, we assume all the parameters of the model are known *a priori*.

4.1. Systematic Channel Updating

For simulation, we build a detector to handle data from $Platform$ 1 and subsequently perform updating to account for the addition of the data from $Platform$ 2 and so on until the data from all 20 arrays have been taken into account. Figures 4(a) and (b) display ΔJ and P_D corresponding to a false alarm rate of 5% at an SNR value of 5 dB when $Platform$ 1 is located at -20, -10, 0, and 10 m in the along-track direction.



(a) ΔJ vs. # of Platforms.



(b) P_D vs. # of Platforms.

Fig. 4. Detection Performance as a Function of the Number of Platforms.

As can be observed, the performance of the detector always improves the closer the platforms approach the target and always increases with a larger number of platforms. The change in J-divergence also always increases the closer we approach the target but does not always increase with the number of platforms. The point at which the change in J-divergence reaches its maximum value (denoted by a black dot in Figures 4(a) and (b)) signifies the point of diminishing return as it is after this point that the J-divergence increases but at a *decreasing* rate. Looking at Figure 4(b), one can observe that the addition of new platforms beyond this point brings little improvement in detection performance. Therefore, when adding platforms, we can see that the change in J-divergence provides an effective measure for recognizing when adding an additional platform would not bring tangible improvement in detection performance.

4.2. Selective Channel Updating

For the next simulation, we assume that the system can only use a subset of the total number of platforms available but we are given the opportunity to choose which platforms to use when performing detection. Rather than searching through all possible combinations of channels and choosing the one that yields maximal detection performance, we instead recursively search through all the platforms and sequentially add the observation that yields the largest increase in J-divergence. As before, the system is initialized by building a detector to handle data from $Platform\ 1$. We then search through all the other 19 platforms, measuring the increase in J-divergence that would be seen if we were to add the observation from that platform. We then choose the one that gives the largest increase in divergence. Then, the likelihood ratio is updated accordingly by augmenting that observation to the observation of $Platform\ 1$. We then search through all the remaining 18 platforms and measure the increase in divergence that would be seen if we were to add the observation from that platform given the new augmented measurement. Again, we choose the one that gives the largest increase in J-divergence, incrementally update the likelihood ratio, and stack the observation from that platform with the augmented measurement from the previous iteration and so on. This selective platform-allocation scheme is compared to a scheme where platforms are chosen at random, i.e. integers ranging from 2 to 20 are selected at random without replacement.

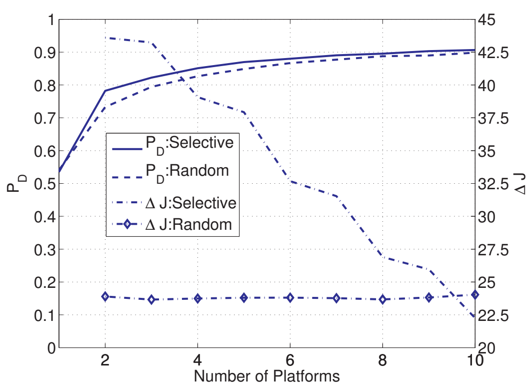


Fig. 5. P_D and ΔJ versus Number of Platforms.

Figure 5 displays the results of this simulation in terms of P_D and ΔJ versus the number of platforms when $Platform\ 1$ is located at 10 m in the along-track direction. We can see that, when selectively adding platforms, the change in J-divergence starts out large and subsequently decreases at a faster rate compared to that when we choose at random which begins at a lower change in J-divergence and remains fairly constant throughout. We can observe in Figure 5 that the performance of the detector associated with selective allocation always exceeds that of random platform selection. We can also see that the difference in P_D among selective and random platform allocation diminishes as the number of platforms grows large. Thus, if we can only take advantage of a small number of platforms, the selective allocation scheme can give a significant increase in detection performance compared to a situation where platforms are arbitrarily chosen, i.e. at random. Stated slightly differently, for a given P_D we can also see that selectively choosing platforms generally requires a smaller number of platforms than when we choose at random. Therefore, we again see that the recursive framework of log-likelihood updating and the corresponding change in J-divergence can be effective tools for deciding which platforms we wish to use when performing detection in such a framework.

5. CONCLUSIONS

In this work, we considered the problem of likelihood updating in Gauss-Gauss detection when new data channels become available. We showed that this process inherently involved linearly estimating the new channel data based upon the old ones. The change in divergence is also derived in terms of error covariance matrices that are matched/mismatched to the given hypothesis and can be used to build low-rank approximations of the update. Simulation examples involving multi-static sonar were considered to illustrate the practicality of the proposed updating methods for not only deciding if and when one reaches a point of diminishing return in adding new sonar platform data but also for deciding which platforms to add.

6. REFERENCES

[1] L. L. Scharf and B. D. Van Veen, "Low rank detectors for Gaussian random vectors," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 11, pp. 1579–1582, Nov 1987.

[2] F. Schweppe, "Evaluation of likelihood functions for gaussian signals," *IEEE Transactions on Information Theory*, pp. 61–70, 1965.

[3] L. Scharf and L. Nolte, "Likelihood ratios for sequential hypothesis testing on markov sequences," *IEEE Transactions on Information Theory*, vol. 23, no. 1, pp. 101–109, 1977.

[4] T. Kailath, A. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall Information and System Sciences Series, 2000.