A Computational Framework for Quality of Information Analysis for Detection-oriented Sensor Networks

Sadaf Zahedi *and* Mani B. Srivastava
Univ. of California, Los Angeles
Electrical Engineering Department
Los Angeles, CA 90095-1594, USA
{*szahedi,mbs*}*@ee.ucla.edu*

Chatschik Bisdikian
IBM Corp.
T. J. Watson Research Center
Hawthorne, NY 10514, USA
bisdik@us.ibm.com

***Abstract*—An application planner's point of view of sensor-enabled detection systems is considered and a hypothesis-testing-based computational framework for evaluating the quality of information (QoI) supported by a sensor network deployment is explored. Through a common, modular analysis framework, that decomposes the computational burden of QoI analysis, the QoI properties of various decision architectures are investigated and trade-offs explored at the sensor, cluster, and system-level. Both finite and infinite-sized networks are considered and extensions of the analysis framework to faulty sensor and the impact of calibration are also investigated.**[1]

I. Introduction

Low cost, intelligent sensor systems find their way in a multitude of application domains such as habitant monitoring, public land monitoring, utility grid monitoring, environmental control, machinery control, intelligence gathering and enemy activity surveillance. As a result, there is a broad spectrum of research efforts for these systems covering such topics as ad hoc deployment and autonomic operation, energy-aware designs, coverage and localization, efficient query dissemination, sensor calibration and data cleaning.

The current research efforts notwithstanding, there is a need to develop general computational tools (or toolkits) to aid sensor-application planners and designers in evaluating the impact on applications of their sensor network designs in advance of deploying their sensor networks in the field. Consider the following use scenario: Jane, a *planner* for intelligence, surveillance, and reconnaissance operations, is contemplating the deployment of sensor-enabled surveillance mission. Jane employs Joe, a sensor system *designer*, to design a sensor network that will satisfy the information needs required by Jane's mission. Concerned with the *effectiveness* of her mission, Jane expresses the requirements for her mission's information needs in the form of the *quality of* the *information* (QoI) provided by the sensor network to be designed. QoI broadly captures the capability of sensor-derived information to describe the important features of situations of interest (e.g., the detection of an explosion) sufficient well to allow the related missions (and more general sensor-enabled applications) to perform their tasks at a desired level of effectiveness.

In considering Joe's toolkit, we introduced in [1] a first instance of a QoI-inspired, layered analysis framework comprising a decomposable series of computational steps that serve as the foundation for the computational toolkit. The framework considers detection of events possessing a persistent or time-limited (transient) energy signature by exploiting traditional hypothesis testing techniques [2], [3]. For such systems, we have selected the *detection probability* P_d and *false alarm rate* P_f as quantifiable attributes that describe their QoI.

In this paper, we extend our earlier work with the following contributions: derivation of analytical results for QoI for networks with finite number of sensors and arbitrary topology, along with a computationally attractive iterative procedure; study of their limiting behavior and derivation of fundamental dominance relationships for establishing bounding approximations to the system performance; study of the relative performance of various fusion architectures and the use of equivalent sensors to simplify their analysis; and finally, a sensor model generalization to accommodates potential faulty sensor operations, e.g., uncalibrated sensors.

The area of QoI for sensor networks is a novel and challenging one and to the best of our knowledge, no prior work on a computational toolkit for it, other than [1] is available. Nevertheless, considerable work on sensor-based detection system has been amassed considering various parameters and models for channel fading, spatial correlations, sensor detection models, signal-to-noise ratio, etc., [4], [5], [6], [7], [8], [9], [6], [10]. Our work does not seek to

[1]Research was sponsored in part by the U.S. Army Research Laboratory and the U.K. Ministry of Defence (under Agreement Number W911NF-06-3-0001), by the NSF (under award CNS-0520006), and by the Center for Embedded Networked Sensing at UCLA. The views and conclusions contained in this document are those of the author(s) and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

replace any previous analyses but rather provide a common computational approach that can leverage these analysis techniques whenever appropriate or use new techniques for building a modular mix-and-match computational toolkit to support Joe's line of work.

The organization of the paper is as follows: In section II, we present the detection system model, the general toolkit framework, and the core analysis approach based on hypothesis testing. In section III, we discuss the application of the core QoI analysis in characterizing and comparing different fusion architectures. In section IV, we study the case of faulty sensors. Finally we conclude in section V with some concluding remarks.

II. The detection system model and QoI analysis methodology

Our analysis framework is based on a fairly general detection system model, which we introduce next.

A. The high-level detection system model

Figure 1 shows the high-level architecture of the sensor-based detector system under consideration. It partitions the sensing operation into three functional subsystems: (a) the *sensor subsystem* or *sampler*, comprising $M \geq 1$ sensor nodes; (b) the *fusion subsystem*, comprising $L \geq 1$ fusion centers; and (c) the *detection subsystem*. The sensor subsystem samples the physical world (in search for an event signature) and passes these samples to the fusion subsystem. The fusion subsystem operates on the sensor readings (samples), which could be corrupted by noise, to produce a "summary" description of the samples. Finally, this summary is used by the detection subsystem to decide whether an event of interest has occurred or not. According to the detection system model, when an event with signature $s^*(t)$ occurs, the k-th sensor, $k \in \{1, \ldots, M\}$, senses its *projection* $s^k(t)$, which encompasses the impact of the system environment and geography on the original event signature. As discussed in [1], the projection metaphor is key in developing our layered computational framework, to be highlighted discussed shortly.

B. The QoI analysis framework

Prior research in the area has studied very specific system models, e.g., for constant amplitude signal signature or a specific detection architecture. However, the specific analysis assumptions may bear limited (if any) resemblance to the realities of a particular deployment instance. Therefore, we have opted to consider a "higher-level" analysis methodology, or *analysis framework*, that will allow us to compose specific solutions from any number of existing or new analysis techniques. In [1] we highlighted such an analysis framework comprised three major processing layers: (a) *input pre-processing*; (b) *detection QoI analysis*

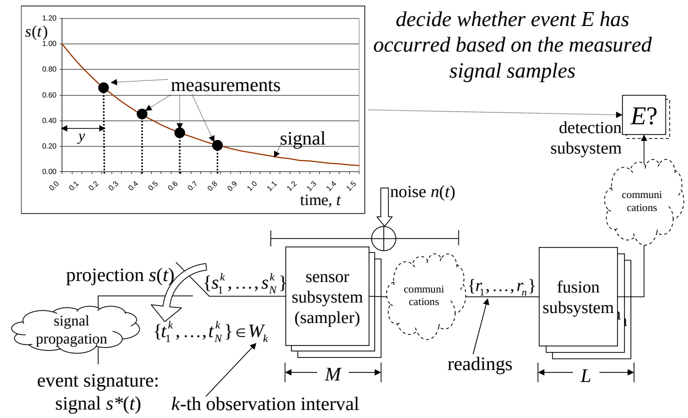

Fig. 1. High level functional architecture of the detection system.

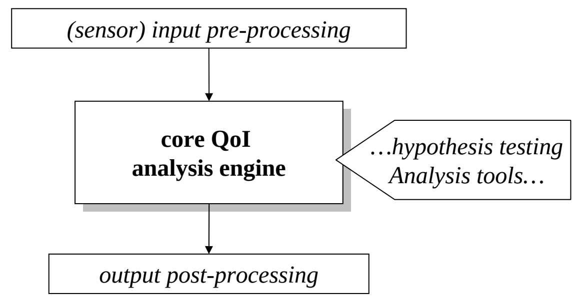

Fig. 2. The QoI analysis framework and toolkit.

(the core analysis engine); and (c) *output post-processing*; see figure 2.

The input pre-processing computational layer pertains to anything that may influence the original event signature until it is recorded by the sensor including the deployment and observation topologies, the signal propagation and attenuation, the noise, the sensor faults, etc. The QoI core analysis engine layers pertains to the computational efforts for calculating the QoI attributes from the recorded event signature projections. Finally, the post-processing layer pertains to any additional QoI-related computations that may be necessary, for example, calculating averages over an observation region, or calculating optimal position of sensors, performing sensitivity analysis, etc.

Each of the processing layers can be studied on its own right separately to the extend necessary, e.g., utilize portions of existing related analysis studies. Under the framework umbrella, these separate studies can be combined to provide answers to a variety of deployment situations in a reusable fashion. Note that prior studies that focus on a specific set of system assumptions represent monolithic (non-layered), top-down interpretations of the framework which increases analysis complexity (the entire problem needs to be studied in one-shot) and reduces application adaptability.

We are closing this section, by highlighting those elements of our current instantiation of the QoI analysis engine that are needed in this paper. The engine is based on binary hypothesis testing analysis techniques [2], [3]. However, alternative techniques may also be employed including multi-hypothesis testing (e.g., for multiple event detection cases) and composite hypothesis testing (e.g., for models with unknown parameters); we view such alternatives as augmentations of the core engine.

C. The core QoI analysis engine: Likelihood ratio tests & hypothesis testing

In order to test the hypothesis that an event of interest occurred (hypothesis H_1), or not (the *null* hypothesis H_0) the detector accumulates observations $r_1, \ldots, r_N$:

$$\text{under hypothesis } H_1 : r_i = s_i + n_i, \text{ or,} \tag{1a}$$

$$\text{under hypothesis } H_0 : r_i = \quad n_i, \tag{1b}$$

where s_i is the i-the sample of the projection of the event signature (under H_1), and n_i is an additive noise component with covariance matrix $\mathbf{C} = E((\mathbf{n}_1^N)^T(\mathbf{n}_1^N))^2$. Considering a Bayesian-based hypothesis test, which minimizes the average cost/risk in making an event occurrence decision (other tests are also possible), and assuming a zero mean, white, additive Gaussian (AWG) noise process, the test reduces to the following *sufficient statistics* expression [2]:

$$l \triangleq \mathbf{r}^T\mathbf{C}^{-1}\mathbf{s} \underset{H_0}{\overset{H_1}{\gtrless}} \eta^* \triangleq \ln(\eta) + 0.5\,\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}. \tag{2}$$

where (for the special case of penalizing only the erroneous decisions) $\eta = P_0/P_1$, where the a priori probabilities for the two hypotheses, P_0 and P_1 are assumed known. The expression $\psi^2 = \mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}$ represents the *signal-to-noise ratio* (SNR) for this system.

For the case of a single-sensor detection systems, whose QoI performance [11] serves as the starting block for the analysis of the multi-sensor systems, (2) results to the following expressions for the probability of detection P_d and false alarm rate P_f [2], [3]:

$$P_d = \Pr(l \geq \eta^\star | H_1) = 1 - \Phi\left(\frac{\ln(\eta)}{\psi} - \frac{\psi}{2}\right), \text{ and}$$

$$P_f = \Pr(l \geq \eta^\star | H_0) = 1 - \Phi\left(\frac{\ln(\eta)}{\psi} + \frac{\psi}{2}\right), \tag{3}$$

respectively, where $\Phi(\cdot)$ is the cumulative distribution function of the normalized Gaussian random variable $\mathcal{N}(0,1)$. For uncorrelated noise samples, the SNR ψ^2 is given by $\psi^2 = (\sum_{i=1}^N s_i^2)/\sigma^2$.

[2]The notation $\mathbf{x}_i^j$ stands for $[x_i, x_{i+1}, \cdots, x_j]^T$. The indexes i and j will be omitted when apparent from the context, e.g., when considering the N-dimensional vectors of signal, observations, noise samples.

III. Fusion architectures for collaborative detection sensor-systems

Next we show how the computational procedures implied by the core analysis engine can be adjusted to reflect various sensor fusion architectures. Beyond its relation to the framework, this section contains performance results of their own, independent merit.

A. L=1: Centralized multi-sensor detection system

The centralized multi-sensor system with only one fusion subsystem ($L = 1$) and one detection subsystem serves as a performance benchmark for any detection architecture and we highlight here a key result from [1]. Specifically, the expressions in (3) still holds true with $\psi^2 = \mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}$ and for a diagonal covariance matrix of the form $\mathbf{C} = \text{diag}(\sigma_1^2\mathbf{I}_{N_1}, \sigma_2^2\mathbf{I}_{N_2}, \ldots, \sigma_M^2\mathbf{I}_{N_M})$, the system-wide SNR is given by

$$\psi_{sys}^2 = \mathbf{s}^T\mathbf{C}^{-1}\mathbf{s} = \sum_{k=1}^{M}\left\{\frac{1}{\sigma_k^2}\sum_{i=1}^{N_k}(s_i^k)^2\right\} = \sum_{k=1}^{M}\psi_k^2, \tag{4}$$

where ψ_k^2 is the SNR attributed to samples from sensor k. In other words, the system-wide SNR is decomposable to the SNRs at the individual sensor level.

In [1], we introduced the concept of an *equivalent sensor* for a collection of networked sensors to describe a sensor whose QoI performance coincides to that of the collection. It follows from (4) and (3) that:

Corollary 1. *A centralized multi-sensor detection system with a Gaussian noise process possesses an equivalent sensor with SNR* $\psi^2 = \mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}$*. If the noise process is an independent process, the SNR is decomposable as in* (4).

B. L=M: Fully distributed multi-sensor detection system

At the other end of the spectrum of detection architectures lies the fully distributed (or, simply, distributed) $(L = M)$ detector. In the distributed detector (and more generally in any non-centralized detector), decision is made in (at least) two steps. In the first step, sensors make *sensor-* or *local-level decision* and then the local decisions are fused to obtain the *system-level* ones. Clearly the QoI performance at the sensor level follows from the study of the single-sensor detector in section II-C, where the SNR ψ_k (instead of ψ) in (3) is calculated separately for each sensor k.

Fusing the local decisions to derive the system-level decision requires the use of a *detection policy*. A variety of policies can and have been considered which typically will involve some form of a counting strategy, e.g., decide that the event has occurred if at least, say, Q out of the M sensors indicate that it has. We refer to the latter as an $\{Q, M\}$ detection policy or system. A more general weighted sum approach has been studied in [9] where each sensor's decision (+1 in favor of H_1 and -1 in favor of H_0) is weighted by the probability of miss and false alarms.

The rather elaborate expressions from the weighted policy in [9] are not particularly amenable to analytical evaluation with regard to the QoI analysis. Instead, next we will study the QoI performance of a counting-based $\{Q, M\}$ detection policy. Let P_k^z represent the probability of detection $(z = d)$ and false alarm $(z = f)$ for sensor k. The system-wide probability $P_z(Q; M)$ is equal to the probability that there exists a collection of sensor containing at least Q sensors all of which declare in favor of the event occurrence, i.e., (where $y = 1$ when $z = d$, and 0 when $z = f$):

$$\begin{aligned} P_z(Q; M) &= Pr(q \geq Q | \text{hypothesis } H_y) \\ &= \sum_{q=Q}^{M} \Big\{ \sum_{\mathbf{x}_q \in \mathcal{S}_q^M} \Big[\big(\prod_{\substack{\text{sensor} \\ m \in \mathbf{x}_q}} P_m^z \big) \big(\prod_{\substack{\text{sensor} \\ n \notin \mathbf{x}_q}} (1 - P_n^z) \big) \Big] \Big\}, \end{aligned} \quad (5)$$

where $\mathbf{x}_q$ is a set of sensors of size q and $\mathcal{S}_q^M$ is the set of all such sets. As the number of sensors M increases, the computation burden of (5) becomes daunting. However, assuming knowledge of the QoI performance of the $\{Q - 1, M - 1\}$ and $\{Q, M - 1\}$ detection systems and tallying the cases that will bring them up to the $\{Q, M\}$ system, we can derive the following simple recursion:

$$P_z(Q; M) = (1 - P_M^z) P_z(Q; M - 1) + P_M^z P_z(Q - 1; M - 1). \quad (6)$$

At the boundaries $Q = 1$ or M, (6) can be be evaluated directly from (5), which reduces to simple binomial expressions.

In the special, albeit unlikely case, where the probabilities P_z^k are independent of the sensor k, i.e., $P_k^z = p_z$ for all $k \in \{1, \ldots, M\}$, (5) reduces to the well known tail of a binomially distributed random variable:

$$P_z^{eq}(Q; M) = \sum_{q=Q}^{M} \Big[\binom{M}{q} (p_z)^q (1 - p_z)^{M-q} \Big]. \quad (7)$$

Such *equiprobable* detection systems have been considered in [6], [7], [8]. For such a system, it follows from the *central limit theorem* (CLT) that as M increases:

$$P_z^{eq}(Q; M) \xrightarrow[M \to \infty]{} 1 - \Phi\Big(\frac{Q - M p_z}{\sqrt{M p_z (1 - p_z)}} \Big). \quad (8)$$

However, the above behavior is not unique to the equiprobable system only and from from the generalized CLT [12],

$$P_z(Q; M) \xrightarrow[M \to \infty]{} 1 - \Phi\Big(\frac{Q - \sum_{i=1}^M P_i^z}{\sqrt{\sum_{i=1}^M P_i^z (1 - P_i^z)}} \Big). \quad (9)$$

For a given original distributed detection system, let a corresponding equiprobable system be constructed that satisfies $P_z^{eq} = (\sum_{k=1}^M P_z^k)/M$, then (see [13] for proof):

Corollary 2. *As M increases, the QoI performance for the $\{Q^{eq}, M\}$ equiprobable system is a close approximation to that of the $\{Q, M\}$ original system, when*

$$\begin{aligned} Q &= M P_z^{eq} + \frac{\sigma}{\sigma^{eq}} (Q^{eq} - M P_z^{eq}) \\ &= \sum_{i=1}^{M} P_i^z + \frac{\sigma}{\sigma^{eq}} (Q^{eq} - \sum_{i=1}^{M} P_i^z), \end{aligned} \quad (10)$$

where $\sigma = \sqrt{\sum_{i=1}^M P_i^z (1 - P_i^z)}$ *and* $\sigma^{eq} = \sqrt{M P_z^{eq} (1 - P_z^{eq})}$.

C. $1 < L < M$: Hybrid multi-sensor detection systems

For the hybrid detection system, the number of fusion subsystems are between 1 and M. Assuming a two-tier system, this means that *clusters* of sensors contribute their samples to lower-tier fusion subsystems, i.e., operate in a centralized detection manner analogous to section III-A. Also, there may be individual sensors that make decisions locally analogous to section II-C. Decisions from the clusters then fuse with the decisions from the individual sensors at a higher-tier fusion subsystem in a manner analogous to the second decision step discussed in section III-B. Thus, hybrid detection systems combine elements from all previous architectures presented and their QoI performance analysis follows from a combination of the previous analysis. By substituting entire clusters of sensors with their equivalent sensors, this case reduces to the study of a distributed sensor system with L sensors.

The assignment of sensors to clusters is an open research activity. We expect that both QoI performance objectives and geography constraints will influence the assignment. Additionally, there will be a trade-off between communication cost and achievable QoI, as it "costs" more to transmit data to a central location to achieve the higher QoI performance of the centralized system.

Next we study the QoI performance of the various architectures through some numerical examples.

D. Comparison of sensor fusion architectures

We will compare the QoI performance of two hybrid systems with the corresponding fully and centralized systems. For these comparisons, we assume an event occurs at a specific location and the sensors are deployed at specific locations. This is indeed a simple case, but since we are interested in comparing different systems against this case, it provides an unambiguous basis for the comparison. More elaborate deployment and observation field are under study.

Let the physical topology of the system be represented by the distance vector $\mathbf{d} = [d_1, \ldots, d_M]^T$, where d_k is the distance of the path that a signal takes from the event location to sensor k. Over this path, let $a_k(t)$ be the *attenuation* that the signal experiences, and let v be the propagation speed for the signal, hence the *propagation delay* is $\tau_k = d_k / v$. Assuming that the event occurs at time $t = 0$ and possesses the (transient) signature $s^*(t)$, the

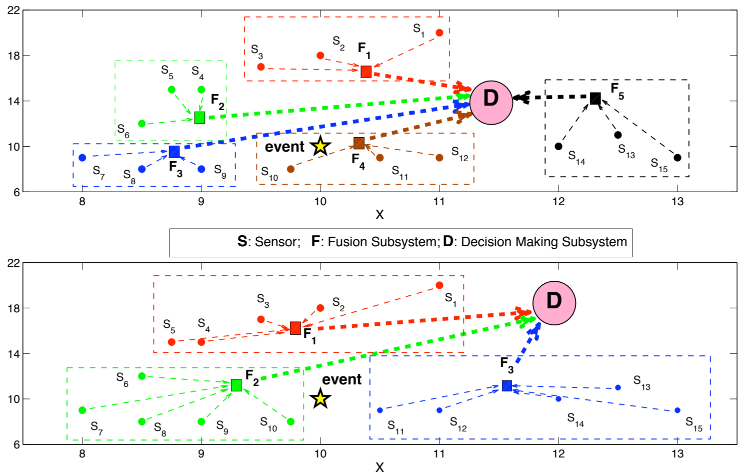

Fig. 3. Topology of the network and two different detection architectures

signal signature seen by sensor k (the k-th event signature projection) would be (excluding any noise components):

$$s_k(t) = a_k(t)s^*(t - \tau_k)u(t - \tau_k), \quad (11)$$

where $u(t)$ is the unit step function. The expression in (11) is an example of a pre-processing operation in our framework that projects the original event signature from the location of the event occurrence to the location of the sensor. For our comparisons, we assume the attenuation function to be $a_k(t) = a_k = 1/(1 + d_k^2)$.

We start with the multi-sensor system ($M = 15$) in figure 3. We assume that the event signature is the transient $s^*(t) = 1 - t^2$ with a lifespan of one time unit) and that each sensor contributes $N = 20$ samples. The number of fusion subsystem varies $L = [1, 3, 5, 15]$; the two end-cases represent the centralized and distributed architectures while the rest the hybrid ones. When $L > 1$, the decision subsystems uses a Q-count policy to make its system-wide decision. The top part of figure 3 shows a case with 5 clusters, with each cluster having 3 sensors contributing samples to their fusion subsystem. The fusion subsystems, then, send their local decisions to a single system-wide decision subsystem. The bottom part of the figure shows a case where we have 3 clusters of sensors. The clusters in both cases were formed based on the relative distance between the neighboring sensors and the event location.

Figure 4 summarizes the QoI performance in this case, where we assume AWG noise with the same variance level σ^2 for each sensor. We consider $\eta = P_0/P_1 = 1$, which serves as our reference point, $\eta = 1.3$ and $\eta = 0.7$. We describe QoI through a single metric, the *probability of error* P_e, which aggregates the QoI attributes:

$$P_e = P_0P_f + P_1(1 - P_d) = P_1(\eta P_f + 1 - P_d); \quad (12)$$

for brevity, we have dropped the Q and M from the notation. As expected, centralized architecture achieves the best (i.e., smallest) P_e of all architectures as it makes best use of the *all* available information.

For $\eta = 1$, the Q_{opt}'s that achieve the highest QoI (lowest P_e) for $L = 5$, 3 and 15 are $Q = 3$, 2 and 8, respectively Note that P_f and P_d (not shown in the figure) decrease with increasing Q as the detection policy decides in favor of H_0 more liberally; this trend holds true for all cases of η studied.

When $\eta > 1$, i.e., $P_0 > P_1$, as in the third column of figure 4, the threshold for deciding in favor of H_1 becomes harder to reach when compared with the $\eta = 1$ case, see (2). Therefore, P_d and P_f decrease in magnitude when compared with these probabilities when $\eta = 1$ (for the same σ) for both architectures. The best detection policy for all three cases is attained at a Q value that is smaller than the best Q of the similar case but with $\eta = 1$. This reduction of the optimal Q value follows from the fact that, with decreasing P_1, it becomes harder for any sensor to declare in favor of H_1. Therefore, having even a small number of sensors declaring in favor of H_1 is reason enough to decide in favor of H_1 system-wide. As expected, the centralized one achieves the best performance.

Finally, the case where $\eta = 0.7$ is shown in figure 4. Arguing as before, the behavior of the QoI metrics in this case relative to those of the $\eta = 1$ case is in reverse order to the behavior experienced when $\eta = 1.3$. With respect to P_e, while the centralized architecture still performs the best.

Next we consider one very important pre-precessing procedure of our framework, specifically, the case of faulty sensor operation causing erroneous measurements. Again, in addition to its relation to the framework, this case represents a contribution with merit of its own in the related art.

IV. QoI Analysis Methodology with Faulty Sensor Measurements

In this section, we assume that in addition to the additive ambient noise, the sensor system itself exhibits a faulty behavior. Enriching our framework with sensor fault models will allow us to analyze the affects of faults, such as due to measurement calibration (or, lack of), on the QoI, and moreover evaluate techniques for reducing their effects.

We start our study with a single sensor system; extensions to multi-sensor systems are under consideration. Furthermore, we start with the linearly biased measurement model [14] according to which the sensor response exhibits an unknown gain a and offset b. Incorporating this fault model as a preprocessing step in our framework, the fundamental hypothesis testing formulation in (1) becomes:

$$\begin{aligned} &\text{under hypothesis } H_1 : r_i = as_i + b + n_i, \text{ or,} \\ &\text{under hypothesis } H_0 : r_i = \qquad\; b + n_i. \end{aligned} \quad (13)$$

Since the gain and offset parameters are unknown, the hypothesis test in (13) will be a composite one, where we not only decide in favor of one of the hypotheses but estimate the unknown parameters a and b as well. Let $\mathbf{r}_{1;H_i}^N$ represent

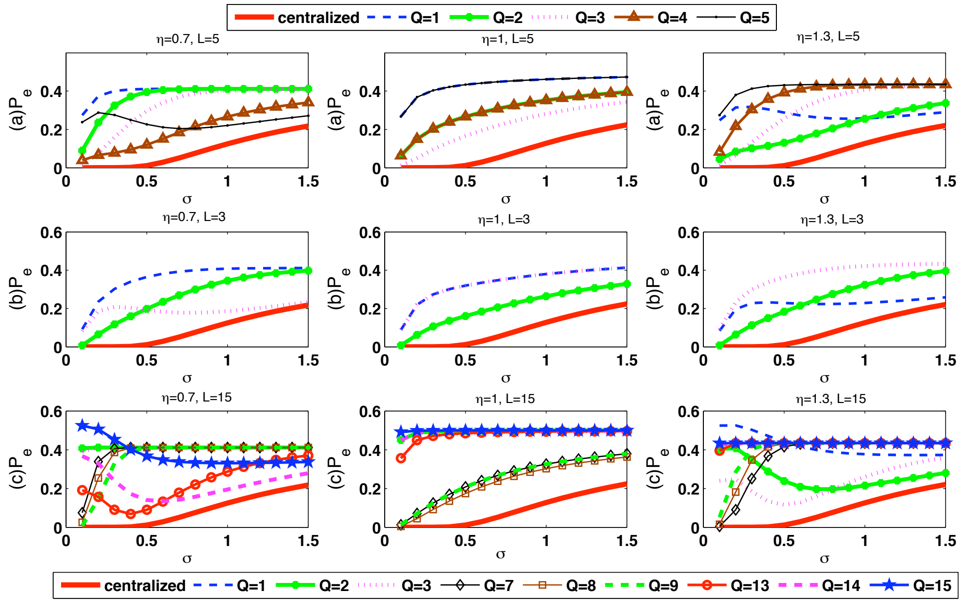

Fig. 4. Performance of hybrid architectures vs. σ for $M = 15$, $\eta = [.7, 1, 1.3]$, and: (a) $L = 5$; (b) $L = 3$; and (c) $L = 15$.

the vector of N observations under hypothesis H_i (we will subsequently drop the indexes 1 and N), then (13) can be written as:

$$\begin{aligned} \mathbf{r}_{1;H_i}^N &= \mathbf{G}_i \boldsymbol{\theta}_i + \mathbf{n}_1^N, \quad i \in \{0,1\}, \\ \mathbf{G}_1 &= [\mathbf{s}_1^N \,|\, \mathbf{1}_1^N] \quad \mathbf{G}_0 = [\mathbf{0}_1^N \,|\, \mathbf{1}_1^N], \text{ and} \\ \boldsymbol{\theta}_1 &= [a,b]^T, \quad \boldsymbol{\theta}_0 = [0,b]^T. \end{aligned} \tag{14}$$

The noise vector $\mathbf{n}$ has a $\mathcal{N}(\mathbf{0}, \mathbf{C})$ distribution and is independent of $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta}_1$. If the calibration parameters a and b were deterministic, the linear model would become a classical linear model. However there exists no *uniformly most powerful* (UMP) test for a classical linear model [3], and hence, in such a case, the accuracy of their estimate cannot be guaranteed with a relatively small number of samples. However, early lab experiments and available training data indicate that these parameters may be treated as random variables that exhibit a Gaussian behavior, i.e., $\boldsymbol{\theta}_i \sim \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{\theta}_i}, \mathbf{C}_{\boldsymbol{\theta}_i})$. In such a case, the model (14) can be treated as a Bayesian linear model, and therefore the posterior distribution $p(\boldsymbol{\theta}_i|\mathbf{r})$ will also be Gaussian with mean $E(\boldsymbol{\theta}_i|\mathbf{r})$ and covariance matrix $C(\boldsymbol{\theta}_i|\mathbf{r})$ given by (see [15, Th. 10.3]):

$$\begin{aligned} E(\boldsymbol{\theta}_i|\mathbf{r}) &= \boldsymbol{\mu}_{\boldsymbol{\theta}_i} + \mathbf{C}_{\boldsymbol{\theta}_i}\mathbf{G}_i^T(\mathbf{G}_i\mathbf{C}_{\boldsymbol{\theta}_i}\mathbf{G}_i^T + \mathbf{C})^{-1}(\mathbf{r} - \mathbf{G}_i\boldsymbol{\mu}_{\boldsymbol{\theta}_i}), \\ C(\boldsymbol{\theta}_i|\mathbf{r}) &= \mathbf{C}_{\boldsymbol{\theta}_i} - \mathbf{C}_{\boldsymbol{\theta}_i}\mathbf{G}_i^T(\mathbf{G}_i\mathbf{C}_{\boldsymbol{\theta}_i}\mathbf{G}_i^T + \mathbf{C})^{-1}(\mathbf{G}_i\mathbf{C}_{\boldsymbol{\theta}_i}). \end{aligned} \tag{15}$$

The minimum mean square error (MMSE) estimate of $\boldsymbol{\theta}_i$ will be $\hat{\boldsymbol{\theta}}_i = E(\boldsymbol{\theta}_i|\mathbf{r})$, where $i \in \{0,1\}$. We can now substitute these estimates in (13) and the decision test becomes:

$$\hat{l} \triangleq \mathbf{r}^T\mathbf{C}^{-1}(\mathbf{s}_1(\hat{\boldsymbol{\theta}}_1) - \mathbf{s}_0(\hat{\boldsymbol{\theta}}_0)) \underset{H_0}{\overset{H_1}{\gtrless}} \eta^* \triangleq \ln(\eta) + \\ + \frac{1}{2}[\mathbf{s}_1(\hat{\boldsymbol{\theta}}_1)^T\mathbf{C}^{-1}\mathbf{s}_1(\hat{\boldsymbol{\theta}}_1) - \mathbf{s}_0(\hat{\boldsymbol{\theta}}_0)^T\mathbf{C}^{-1}\mathbf{s}_0(\hat{\boldsymbol{\theta}}_0)], \tag{16}$$

where $\mathbf{s}_i(\hat{\boldsymbol{\theta}}_i) = \mathbf{G}_i\hat{\boldsymbol{\theta}}_i$. Finally, the QoI performance becomes:

$$P_z = \Pr(\hat{l} \geq \eta^\star | H_i) = 1 - \Phi\left(\frac{\eta^* - \mu_{\hat{l}|H_i}}{\sigma_{\hat{l}|H_i}}\right), \tag{17}$$

where $\mu_{\hat{l}|H_i} = \mathbf{s}_i(\hat{\boldsymbol{\theta}}_i)^T\mathbf{C}^{-1}\boldsymbol{\Delta}$, $\sigma^2_{\hat{l}|H_i} = \boldsymbol{\Delta}^T\mathbf{C}^{-1}\boldsymbol{\Delta}$, and $\boldsymbol{\Delta} = \mathbf{s}_1(\hat{\boldsymbol{\theta}}_1) - \mathbf{s}_0(\hat{\boldsymbol{\theta}}_0)$, and $i = 1$ when $z = d$, and 0 when $z = f$.

If we were not aware of calibration errors, then $a = 1$ and $b = 0$ could be considered, which results in $\mathbf{s}_1(\breve{\boldsymbol{\theta}}_1) = \mathbf{s}$, and $\mathbf{s}_0(\breve{\boldsymbol{\theta}}_0) = \mathbf{0}$. But, since the observations come from a sensor with gain a and offset b, $\mu_{\hat{l}|H_1} = (a\mathbf{s} + b\mathbf{1})^T\mathbf{C}^{-1}\mathbf{s}$ and $\mu_{\hat{l}|H_0} = b\mathbf{1}^T\mathbf{C}^{-1}\mathbf{s}$ and $\sigma^2_{\hat{l}|H_1} = \sigma^2_{\hat{l}|H_0} = \sigma^2 = \mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}$. Finally, using an approach similar to (17), we can compute the probability of detection and false alarm for this case.

We apply the prior analysis to a single-sensor system for detecting events with transient signature $s^*(t) = 1 - t^2$, located at distance $d = 3$ from the event location that contributes $N = 20$ samples to the detection process. The sensor exhibits a faulty measurement operation with a gain and offset of $a = 1.4$ and $b = 0.4$, respectively, and we have prior knowledge that the parameters behave as in $\boldsymbol{\theta}_1 \sim \mathcal{N}([1, 0.2], \mathrm{diag}(1,1))$. Figure 5 shows the probability of error P_e as a function of the noise level σ. Three cases are considered: (a) there is complete knowledge of the sensor's true gain a and offset b calibration parameters – to serve as the benchmark case; (b) only the distributions of the calibration parameters are known and their estimated

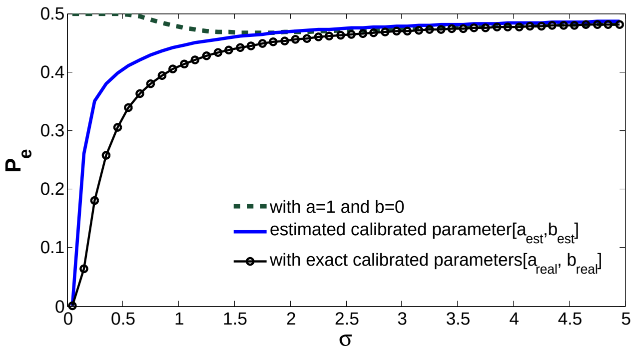

Fig. 5. P_e analysis of the un-calibrated sensors

parameter values are used instead; and (c) there is no knowledge of any sensor fault and the ideal values $a = 1$ and $b = 0$ are used instead.

As figure 5 shows, if we are unaware of have any sensor calibration faults, we may fail to provide a reliable QoI analysis for the system. However, using the estimates of the calibration parameters contributes significantly in approaching the real performance curve. As expected, as the noise level increases, the QoI performance becomes less reliable as the samples collected does not affect the posterior probability of the hypotheses. performance it affects the estimator performance. Also note that using larger number of samples from observations can improve the estimator performance. Finally, this analysis points to the importance of considering sensor faults and the use of proper fault models and measurement errors in the system for more reliable QoI performance analysis.

V. Concluding Remarks

Taking an application-centric view of sensor-derived information, we have presented a computational framework for evaluating the QoI that a detection-based sensor system is expected to achieve. Broadly speaking QoI represents the level of confidence that a sensor-data-dependent application may place on information derived from the sensor network(s) that (may) support the information needs of the application. Knowledge of the QoI allows mission/application planners to enhance their system with contingency plans in anticipation of "errors" that the network of sensors and their fusion modules accurately capturing and interpreting the real world.

Considering the rather broad application space of detection of signal-producing events, we have proposed and studied a layered analysis framework that allows the composition of performance analysis solutions from a number of constituent analysis techniques. The layered approach *decomposes* the computational burden of the entire solution approach into three steps: a preprocessing step that focuses on what affects signals from their source to their (sensor) destination; a core analysis of QoI; and, finally, a post processing layer that operates on the results of the QoI analysis to build the final desired solution.

We have based our QoI analysis on hypothesis testing, and derived the QoI performance, in the form of the probability of detection, false alarm, and error, for a number of fusion architectures. Within this context and in addition to their relation to the analysis framework, we presented new results studying both finite and "infinite" sensor systems, deriving a continuum of computationally attractive solutions spanning both small and large systems. Furthermore, we derived dominance relationships that allow to calculate QoI metrics by constructing simpler to analyze equiprobable detection systems. Finally, we investigated the impact of calibration faults on the QoI performance of a single-sensor system, which also represents a new result in the study of sensor-based detection systems.

References

[1] S. Zahedi and C. Bisdikian, "A framework for QoI-inspired analysis for sensor network deployment planning," in *2nd Int'l Workshop on Performance Control in Wireless Sensor Networks (PWSN)*, Austin, TX, USA, Oct. 23, 2007.

[2] H. L. van Trees, *Detection, Estimation, and Modulation Theory, part I*. John Wiley & Sons, 1968.

[3] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection theory*. Prentice Hall, 1998, vol. 2.

[4] R. Viswanathan and P. Varshney, "Distributed detection with multiple sensors:part i-fundamentals," in *Proceedings of the IEEE*, vol. 85, no. 1, Jan. 1997, pp. Page(s):54 – 63.

[5] R. Blum, S. Kassamand, and H. Poor, "Distributed detection with multiple sensors:part ii-advanced topics," in *Proceedings of the IEEE*, vol. 85, no. 1, Jan. 1997, pp. Page(s):64–79.

[6] R. Niu and P. Varshney, "Distributed detection and fusion in a large wireless sensor network of random size," in *EURASIP Journal on Wireless Communications and Networking*, vol. 5, no. 4, 2005, pp. Page(s):462 –472.

[7] S. K. Jayaweera and K. Al-Tarazi, "Large system decision fusion performance in inhomogeneous sensor networks," in *Proc. of the IEEE Int'l Conf. on Video and Signal Based Surveillance*, 2006.

[8] N. Katenka, E. Levina, and G. Michailidis, "Local vote decision fusion for target detection in wireless sensor networks," in *Joint Research Conference on Statistics in Quality Industry and Technology*, Knoxville, TN, USA, June 7-9, 2006.

[9] Z. Chair and P. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Transactions on Aerospace and Electronic Systems*, pp. 98–101, 1986.

[10] L.Yu, L.Yuang, G. Qu, and A. Ephremides, "Energy-deriven detection scheme with with gauranteed accuracy," in *in Proc. IPSN'06*, Nashville,TN, USA, April 9-21, 2006.

[11] C. Bisdikian, "Quality of information trade-offs in the detection of transient phenomena," in *SPIE Unattended Ground, Sea, and Air Sensor Technologies And Applications IX Conference*, Orlando, FL, USA, April 9-13, 2007.

[12] X.Guyon, *Random Fields on a network: Modeling, Statistictics, and Applications*. Springer-Verlag, 1995.

[13] S. Zahedi, M. B. Srivastava, and C. Bisdikian, "A framework for quality of information analysis for detection-oriented sensor network deployments," IBM Research, Tech. Rep. RC24404, Nov. 2007.

[14] L. Balzano and R. Nowak, "Blind calibration of sensor networks," *Proc. Information Processing in Sensor Networks (IPSN)*, 2007.

[15] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993, vol. 1.