1-1-2022

Performance of XOR Rule for Decentralized Detection of Deterministic Signals in Bivariate Gaussian Noise

Xingjian Sun
University of Mississippi

Shailee Yagnik
University of Mississippi

Ramanarayanan Viswanathan
University of Mississippi

Lei Cao
University of Mississippi

Recommended Citation

Sun, X., Yagnik, S., Viswanathan, R., & Cao, L. (2022). Performance of xor rule for decentralized detection of deterministic signals in bivariate gaussian noise. IEEE Access, 10, 8092–8102. https://doi.org/10.1109/ACCESS.2022.3143105

Performance of XOR Rule for Decentralized Detection of Deterministic Signals in Bivariate Gaussian Noise

XINGJIAN SUN[ID], (Student Member, IEEE), **SHAILEE YAGNIK**, (Student Member, IEEE),
RAMANARAYANAN VISWANATHAN[ID], (Life Fellow, IEEE),
AND LEI CAO[ID], (Senior Member, IEEE)

Department of Electrical and Computer Engineering, University of Mississippi, Oxford, MS 38655, USA

Corresponding author: Xingjian Sun (xsun2@go.olemiss.edu)

ABSTRACT In this paper, we consider the performance of exclusive-OR (XOR) rule in detecting the presence or absence of a deterministic signal in bivariate Gaussian noise. Signals, when present at the two sensors, are assumed unequal, whereas the noise components have identical marginal distribution but are correlated. The sensors send their one-bit quantized data to a fusion center, which then employs the XOR rule to arrive at the final decision. Here we prove that, in the limit as the correlation coefficient r approaches 1, the optimum fusion rule for both parallel and tandem topologies is XOR with identical, alternating partitions (XORAP) of the observations at the sensors. We further quantify the asymptotic decrease of the Bayes error of XORAP towards zero as a constant multiplied by $\sqrt{1-r}$, as r approaches 1. When compared to the asymptotic Bayes error of CLRT, which decreases to zero exponentially fast, as a function of $1/(1-r)$, the Bayes error of XORAP decreases to zero much slower.

INDEX TERMS Gaussian noise, multiple sensors, noise correlation, signal detection.

I. INTRODUCTION

A collection of sensors are employed in a variety of situations in order to enhance information gathering and processing operations. Depending upon the capability of sensor networks, either condensed information or the raw data from the sensor sites may be transmitted to a fusion center (FC), where a final decision with regard to the presence (hypothesis $\mathcal{H}_1$) or absence (hypothesis $\mathcal{H}_0$) of a signal (or target) is made. The former case is usually referred to as decentralized detection, whereas the latter case is termed as the centralized detection [1]. Due to the bandwidth constraint in the reporting channels, decentralized detection is often expected in practice. It is well known that the complexity of finding optimal solutions to decentralized detection problems is, in general, formidable [1]–[3]. The usual performance criterion employed is either Bayes error (or cost) or Neyman-Pearson (NP) test.

In this paper, we employ Bayes error criterion and restrict consideration to a two-sensor network, with one-bit quantiza-
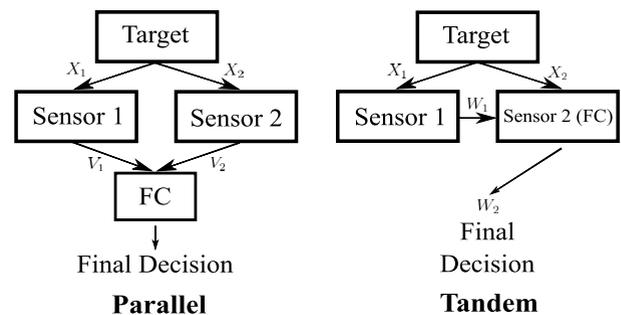
FIGURE 1. Diagram of a two-sensor decentralized detection system with the parallel topology (left) and with the tandem topology (right). FC stands for the fusion center. In tandem topology, the second sensor also acts as the FC.

tion of sensor observations. Fig. 1 shows configurations for parallel and tandem topologies in a decentralized detection system. Even in our case of two sensors, with each sensor sending one-bit quantized information to a FC, which then makes a final decision on the true hypothesis, finding the optimal decentralized decision scheme, i.e., quantization schemes at the two sensors and the decision rule at the FC,

is still a difficult problem. In the parallel topology, the optimal quantizers at the two sensors need not be identical, even if the marginal distributions are identical [1], [2], [4]. This is applicable for both correlated observations and statistically independent observations [1]. In the case of one-bit quantized data from two sensors, the number of possible FC Boolean rules are limited to 2^{2^2}, out of which the non-trivial ones are the Boolean AND, OR, XOR, and "ignoring one of the sensors," excluding the other equivalent rules. A method to find the optimal solution to the decentralized detection problem is to assume a particular fusion rule, optimize the quantizers at the sensors for that fusion rule, and then pick the best fusion rule along with the corresponding quantizers at the sensors as the optimal solution. In the simplest case of each sensor sending one-bit information, the optimal quantizer turns out as the partitioning of the real line (i.e., the observation) into quantization intervals, where the intervals are represented as a bit '1' (possibly also for the intervals represented by the bit '0') could be non-contiguous.

At sensor 1 (2), assume that a known signal s_1 (s_2) is present in Gaussian noise, when the hypothesis $\mathcal{H}_1$ is true. When $\mathcal{H}_0$ is true, there is no signal at each sensor, and the sensor observation is simply noise. The noise components at the sensors are assumed jointly Gaussian with a correlation coefficient r and identical variance (non-identical variance case can simply be handled by dividing each observation at a sensor with the corresponding standard deviation of the noise component). In the seminal paper by Willett *et al.* [4], for the parallel topology, the authors identified regions for the optimality of the Boolean AND (or OR) rule in the (s_1, s_2) signal plane, which includes the region where the quantization intervals for both bits '1' and '0' are contiguous (called simply the 'Good' region), the region where it may be optimal to ignore one sensor data ('Bad' region), and the region where the exclusive-OR (XOR) rule could outperform both AND and OR rule. However, [4] did not examine the performances of those rules when the correlation coefficient r is closer to 1. As will be shown in this paper, the performance of XOR rule is of significance, as r approaches 1.

Reference [5] extended the results in [4] to two-sensors tandem case and obtained 'Good' region solution. Reference [6] discussed the detection of Gaussian signals in Gaussian noise for the same tandem case. It was pointed out in [6] that [7] made an incorrect assumption that the first sensor in a tandem network could employ a simple likelihood ratio test without compromising the global optimality. A recent contribution discusses the joint estimation of unknown parameters and detection in a joint copula probability distribution model [8]. All the above results and those in [9]–[12] show the difficulty in finding an optimal solution when the sensors' observations are correlated. In this paper, we consider the detection problem with completely known parameters and the XORAP rule. It must be mentioned that the term XOR refers to the case where the Boolean fusion rule at the FC is exclusive-OR, without any specific reference to the one-bit quantization schemes that the sensors would employ. XORAP
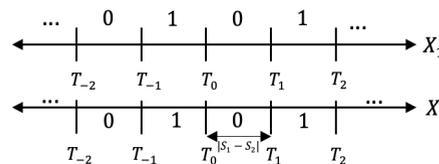
FIGURE 2. Local one-bit quantizers at the two sensors for XOR fusion in XORAP rule. Partition of $X_i (i = 1, 2)$: $V_i = 1, X_i \in R_{V_i}$; $V_i = 0, X_i \in \overline{R_{V_i}}$.

specifies identical alternating quantization intervals of length, $D \triangleq |s_1 - s_2|$, representing bits '1' and '0' at both the sensors, in addition to the specification of XOR fusion rule (see Fig. 2).

A. MOTIVATION FOR INVESTIGATING THE XOR PERFORMANCE

The first reason for the investigation is the optimality of the XORAP rule. If each sensor makes its decision based on the comparison of its observation against a single threshold, then XOR as a fusion rule is not meaningful in the case of independent observations. This is because the XOR rule is not a monotone fusion rule, a condition needed for optimality in the case of independent observations [1], [13]. Consider the situation when sensors utilize multiple quantization intervals, each interval designated as either a bit '1' or a bit '0'. In that case, for example, a bit '1' received at the FC from a sensor might be the result of its observation falling in any one of the multiple intervals marked as bit '1' by the sensor, and therefore this bit is not a local decision in favor of a hypothesis. The alternating partitioning intervals of span D for XOR decision was first mentioned in [14]. It showed the optimality of XORAP rule for the case of two-sensor parallel topology, in the sense of achieving perfect decision, when $r = 1$. The partitioning was arrived at heuristically. A theoretical proof that it is optimal for both parallel and tandem topologies is needed.

The second reason for the investigation is the behavior of different fusion rules as the correlation coefficient increases from 0 to 1. As shown in [15], for the case of centralized likelihood ratio test (CLRT), as r increases from 0, the Bayes error keeps increasing up to the point $r = \min(s_1, s_2)/\max(s_1, s_2)$, after which it monotonically decreases towards 0. That is, the increasing correlation is detrimental to the probability of error performance in detection, but that is only up to the point determined by the signal levels at the two sensors. This behavior is seen for the number of sensors exceeding 2 as well, but the point of inflection is not the above simple formula given for two sensors. Any increase in correlation coefficient beyond that point aids in the detection of signals, ultimately leading to zero Bayes error as $r \to 1$. Similar behavior is also exhibited when sensors employ multi-bit quantization, see [14], [16]. Hence, given that XORAP is the best one-bit fusion rule as $r \to 1$, and that the CLRT is the best fusion rule without

any quantization of sensor data, it will be of interest to compare the performances of the two for large values of r, knowing well that the performance of any reasonable scheme improves, once the correlation becomes increasingly larger.

Finally, it should be noted that although the correlation coefficient close to 1 might be rarely seen in practice, investigation of this limiting case has its own theoretical importance that contributes to a better understanding of how the local decisions and the fusion rules would change with respect to the change of correlation coefficient, as well as the corresponding performance.

B. CONTRIBUTION

Contribution in this paper is two-fold. First, in *Theorem 1* to be presented in the next section, we will present a proof asserting that XORAP is the optimal rule for the two-sensor tandem topology, as $r \rightarrow 1$. For a two-sensor network, since the optimal rule in the parallel case can never outperform the optimal rule in the tandem case (see [1], [13]) and the XORAP is also a parallel decision rule, other parallel fusion rules with one-bit quantization cannot outperform the XORAP rule in the parallel case, in terms of the rate of decrease of the Bayes error as r approaches 1. This proves that the XORAP is the optimal rule for both tandem and parallel cases as $r \rightarrow 1$. Second, in *Theorem 2* we will prove that the Bayes error of the XORAP rule decreases to zero as a constant multiplied by $\sqrt{1-r}$, as $r \rightarrow 1$ and then compare it with the case of centralized detection.

The rest of the paper is organized as follows. Section II describes the problem formulation, proof of the optimality of the XORAP rule for the two-sensor tandem topology, and the derivation of the asymptotic (as $r \rightarrow 1$) Bayes error for both CLRT and XORAP. Additionally, random number generation to corroborate the theoretical results, graphical explanation for the asymptotic error behavior of XORAP, and results of suboptimal solution obtained from a genetic algorithm based simulation study are also presented in this section. Section III concludes this paper.

II. PROBLEM FORMULATION AND SOLUTION

Consider two sensors monitoring a region of interest to ascertain the presence of a signal (hypothesis $\mathcal{H}_1$) or its absence ($\mathcal{H}_0$). When a signal is present in the region, the signal components at the two sensors, which are assumed to be deterministic and known, are received in additive Gaussian noise. Otherwise, the observations at the sensors have only noise components. The two hypotheses can be stated as the following equation:

$$\mathcal{H}_1 : X = S + N$$
$$\mathcal{H}_0 : X = N, \qquad (1)$$

where $X^T = [X_1, X_2]$, $N^T = [N_1, N_2]$, $S^T = [s_1, s_2]$. As in [4], N is assumed to be distributed as bivariate Gaussian with zero means, unit variances and correlation coefficient r. Since s_1 and s_2 could assume negative or

TABLE 1. Definition of symbols.

Symbols	Definition
$X_i, i = 1, 2$	Observation at sensor i
s_i	Signal at sensor i, when present
$D = |s_1 - s_2|$	Absolute difference of signals
N_i	Noise at sensor i
r	Correlation coefficient between N_1 and N_2
$P(\mathcal{H}_j) = \pi_j$	Prior probability of hypothesis $\mathcal{H}_j, j = 0, 1$
η	Ratio of prior probabilities, $\eta = \pi_0/\pi_1$
P_e	Bayes error,
	P_{eC} (CLRT), P_{eX} (XORAP)
Parallel Topolgy	
$V_i, i = 1, 2$	Report of sensor i to the FC
Tandem Topology	
W_1	Report of sensor 1 to the FC
$R_{W_1}, \overline{R_{W_1}}$	The region in which sensor 1 makes a decision favoring $\mathcal{H}_1$ and $\mathcal{H}_0$, respectively
$R_{W_2|W_1}, \overline{R_{W_2|W_1}}$	Given W_1, the region in which sensor 2 makes a decision favoring $\mathcal{H}_1$ and $\mathcal{H}_0$, respectively

positive values, without any loss of generality we can restrict r to be in the interval $[0, 1]$. Assuming prior probabilities for the hypotheses, namely $P(\mathcal{H}_0) = \pi_0, P(\mathcal{H}_1) = \pi_1 = 1 - \pi_0$, we consider the Bayes error $P_e = \pi_0 P(D_1|\mathcal{H}_0) + \pi_1 P(D_0|\mathcal{H}_1)$ as the performance criterion for a decision rule, where $D_j, (j = 0, 1)$, is the decision favoring hypothesis $\mathcal{H}_j$. Here, the first conditional probability is the probability of false alarm P_f, that is, the probability that the decision rule decides hypothesis $\mathcal{H}_1$, when the true hypothesis is $\mathcal{H}_0$. Similarly, the second conditional error probability is the probability of miss, P_m. As shown below, the optimality of XORAP and its asymptotic error are valid for any prior probabilities. For simplicity, the asymptotic error rate of CLRT is derived for the case of equal prior. Similarly, in the simulation studies of sections II. *D* and II. *F* we assume equal prior. If needed, similar analyses can easily be done for unequal prior probabilities. The notations used in this paper are listed in Table 1.

A. XORAP RULE AND ITS OPTIMALITY

In a tandem decentralized detection system with two sensors, see Fig. 1, assume the decision rule of sensor 2 (FC) is denoted by W_2 and the decision rule of sensor 1 is given by W_1, with R_{W_1} as the region in which sensor 1 makes a decision favoring $\mathcal{H}_1$. Since we consider one-bit sensor reports, both W_1 and W_2 take values from $\{0, 1\}$. Referring to Fig. 2, we have the following Theorem.

Theorem 1: For a two-sensors decentralized detection system with tandem topology, as $r \rightarrow 1$, the optimum decision rule W_2 can be expressed as the Exclusive-OR (XOR) rule between W_1 and $I(x_2 \in R_{W_1})$, where $I(\cdot)$ is the indicator function which takes value 1, if $x_2 \in R_{W_1}$, and value 0, otherwise. The optimum partitions of both sensors' observations, as bits '1' and '0', are the same and are given by alternating segments of fixed length, $D = |s_2 - s_1|$, assuming that the length is not zero.

Proof: Appendix A proves in detail that the optimal detection rule for the two-sensor tandem topology, with

one-bit quantization, is exactly the XOR rule with alternating partitions as shown in Fig. 2, as $r \to 1$. This result is valid for any prior probability. $\qquad \square$

B. ASYMPTOTIC ERROR ANALYSIS OF CLRT

Let the covariance matrix of noise N be denoted as Σ, with $\Sigma_{11} = \Sigma_{22} = 1$ and $\Sigma_{12} = \Sigma_{21} = r$. Since CLRT is a LRT based on the observation X, it can be easily seen from [15] that the Bayes error P_{eC} for CLRT, with $\pi_0 = \pi_1 = 1/2$, is given by $Q(\sqrt{K}/2)$, where K is the SNR defined as $K = S^T \Sigma^{-1} S$, $Q(z) = 1 - \Phi(z)$, and $\Phi(z)$ is the standard normal cumulative distribution function (CDF). By simplifying K, we get $K = (s_1^2 - 2rs_1s_2 + s_2^2)/(1 - r^2)$. As $r \to 1$, $P_{eC} \to Q(|s_1 - s_2|/(2\sqrt{1-r^2}))$. Using asymptotic expansion for the $Q(\cdot)$ function as the argument becomes infinitely large, we get $\lim_{r \to 1} P_{eC} = e^{-\frac{D^2}{16(1-r)}}$. Hence, the Bayes error decreases to zero exponentially fast, with increasing $1/(1-r)$.

C. ASYMPTOTIC ERROR ANALYSIS OF XORAP

We begin by evaluating the probability of correct decision of the XORAP rule when $\mathcal{H}_0$ is true, that is the probability, $P_{c0} = P(\text{XORAP decides } \mathcal{H}_0 | \mathcal{H}_0)$. The contributions to this probability come from two distinct probabilities P_1 and P_2, which are defined below in the following two equations.

$$P_1 = \sum_{i=-\infty}^{\infty} \int_{iD}^{(i+1)D} \int_{iD}^{(i+1)D} f(x, y; r) dx dy \qquad (2)$$

$$P_2 = \sum_{i=-\infty}^{\infty} \sum_{k=1}^{\infty} \int_{iD}^{(i+1)D} \int_{iD \pm 2kD}^{(i+1)D \pm 2kD} f(x, y; r) dx dy, \quad (3)$$

where $f(x, y; r)$ denotes bivariate Gaussian density with zero means, unit variances and correlation coefficient r. P_1 is nothing but the probability that both the sensor observations fall in the same interval $(iD, (i+1)D)$ in Fig. 2. Therefore, when this happens, XORAP will correctly decide $\mathcal{H}_0$. By looking at Fig. 2, it is plain that, under $\mathcal{H}_0$, XORAP will also decide correctly when $X_1 \in (iD, (i+1)D)$ and $X_2 \in (iD \pm 2kD, (i+1)D \pm 2kD)$, k is an integer, thereby contributing to the probability P_2.

Theorem 2: By using the result by G Pólya in [17], we show below that

$$\lim_{r \to 1} 1 - P_1 = \left[\frac{\sqrt{2}}{\pi} (1-r)^{\frac{1}{2}} \left(1 + 2 \sum_{i=1}^{\infty} (e^{-0.5D^2})^{i^2} \right) \right]. \quad (4)$$

Proof: Based on the partition and the XOR rule shown in Fig. 2, we have $P_1 = \sum_{i=-\infty}^{\infty} J_i$, where

$$J_i \triangleq \int_{iD}^{(i+1)D} \int_{iD}^{(i+1)D} f(x, y; r) dx dy. \quad (5)$$

We will be utilizing a result from [17] to evaluate J_i, as $r \to 1$. The notation $L(\cdot)$ below and its arguments from this reference are restated here for our use.

$$L(a, a'; b, b'; r) \triangleq \int_a^b \int_{a'}^{b'} f(x, y; r) dx dy. \quad (6)$$

In our case, $a = a' = iD$, $b = b' = (i+1)D$. Therefore,

$$J_i = L(a, a; b, b; r). \quad (7)$$

Define the following three variables [17],

$$\alpha \triangleq \frac{a + a'}{2^{1/2}(1+r)^{1/2}}, \beta \triangleq \frac{b + b'}{2^{1/2}(1+r)^{1/2}}, \delta' \triangleq \frac{-a + b'}{\sqrt{2(1-r)}}.$$

We have

$$\lim_{r \to 1} \alpha = iD, \ \lim_{r \to 1} \beta = (i+1)D.$$

Direct application of (6.8) in [17] gives

$$\lim_{r \to 1} J_i \sim 2J_{i1} + 2J_{i2}, \quad (8)$$

where

$$J_{i1} = [G(\beta) - G(\alpha)]G(\delta'),$$
$$J_{i2} = -\rho(g(\alpha) + g(\beta))(g(0) - g(\delta')).$$

$G(x)$ is defined as the area under the standard Gaussian density, $g(y)$, over $y \in (0, x)$, and $\rho = \left(\frac{1-r}{1+r} \right)^{\frac{1}{2}}$. Based on (2), (5)-(8) and considering $r \to 1$,

$$P_1 = I_1 - \frac{2\rho}{\sqrt{2\pi}} I_2, \quad (9)$$

where

$$I_1 = \sum_{i=-\infty}^{\infty} [G((i+1)D) - G(iD)]$$
$$= G(\infty) - G(-\infty) = 1,$$
$$I_2 = \sum_{i=-\infty}^{\infty} g(iD) + g((i+1)D).$$

We will show that $I_2 < \infty$. Since $g(-iD) = g(iD)$,

$$I_2 = 4 \sum_{i=1}^{\infty} g(iD) + 2g(0), \quad (10)$$

or, $I_2 = \frac{4}{\sqrt{2\pi}} \sum_{i=1}^{\infty} (e^{-\frac{D^2}{2}})^{i^2} + \frac{2}{\sqrt{2\pi}} < \infty$. Now, using (9), $I_1 = 1$, ρ given above and this asymptotic expansion of I_2, we arrive at (4). $\qquad \square$

We will show in Appendix B that P_2 decreases to zero exponentially fast and can be neglected when compared to $1 - P_1$. Hence, the probability of false alarm, $P_f = 1 - P_{c0} = 1 - P_1 - P_2$ goes to zero as a constant multiplied by $\sqrt{1-r}$ (4). The other conditional probability of error, the probability of miss, is exactly the same as P_f. This can be seen from the following argument. Since the error in XORAP is caused by the pair of noise samples, whenever XORAP makes an error under $\mathcal{H}_0$, the same noise pair would also cause an error if $\mathcal{H}_1$ were true. Similar comments apply to non-error-causing events. Hence, the Bayes error of XORAP is $P_{eX} = P_f$, for any prior probability.

TABLE 2. Bayes error performance of CLRT and XORAP for $s_1 = 1$, $s_2 = 0.5$, and different r values.

r	P_{ec}	P_{eX}	$1 - P_1$
0.3	0.3047	0.5	-
0.5	0.3085	0.5	-
0.9	0.2487	0.49	-
0.99	0.0306	0.2272	0.2257
0.999	1.05 E-08	0.071	0.0714
0.9999	1E-70	0.023	0.0226
0.99999	-	0.007	0.0071

D. SIMULATION OF XORAP ERROR

In this subsection, we study the probability of error behavior of XORAP, by generating bivariate Gaussian samples with different values of correlation coefficient r, which can be simulated by using the following equation.

$$N_2 = N_1 + \epsilon Z, \tag{11}$$

where N_1 and Z are i.i.d as Gaussian with zero mean and unit variance, i.e. $\sim$ Gaussian $(0, 1)$, and $\epsilon = \sqrt{1 - r^2}/r$. Table 2 shows how the Bayes error P_{eC} and P_{eX}, respectively for CLRT and XORAP, decrease towards zero as r increases towards 1. The Bayes error probabilities for CLRT are based on exact theoretical expression, assuming equal prior, $\pi_0 = \pi_1 = \frac{1}{2}$. The P_{eX} values shown in the Table 2 for $r = (0.99, 0.999, 0.9999, 0.99999)$ are remarkably close to what one gets from the asymptotic error expression for $P_{eX} \approx 1 - P_1$, shown as the last column in the Table 2. Since $P_f = P_m$ for XORAP, P_{eX} is independent of prior probability.

In this simulation study and the asymptotic error analysis, we assumed that the origin coincided with T_0 in Fig. 2. As can be seen from the proof of Theorem 2, if the origin were offset from T_0, then the asymptotic error still decays as $\sqrt{1-r}$, but the proportionality constant in (4) will slightly be altered. Similarly, P_{eX} values in Table 2 will change slightly.

The following subsection provides a graphical illustration of why Bayes error of XORAP decreases at a much slower rate when compared to the CLRT rate of decrease.

E. GRAPHICAL ILLUSTRATION OF DIFFERENCE IN PERFORMANCES OF CLRT AND XORAP

Consider an example with $s_1 = 1$, $s_2 = 0.5$, $\eta = \pi_0/\pi_1 = 1$. The illustration in Fig. 3 shows the decision regions of CLRT and XORAP in the (X_1, X_2) plane with $D = |s_1 - s_2| = 0.5$.

The CLRT separation line is given by [4]

$$(s_1 - rs_2)X_1 + (s_2 - rs_1)X_2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}}$$
$$(1 - r^2)\ln\eta + \frac{s_1^2 + s_2^2 - 2rs_1s_2}{2}. \tag{12}$$

Assume $\epsilon = 1 - r$, as $r \to 1$, $\epsilon \to 0$. Then (12) is reduced to

$$X_1 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \frac{1 - 2\epsilon}{1 + \epsilon} X_2 + \frac{0.25 + \epsilon}{1 + \epsilon}. \tag{13}$$
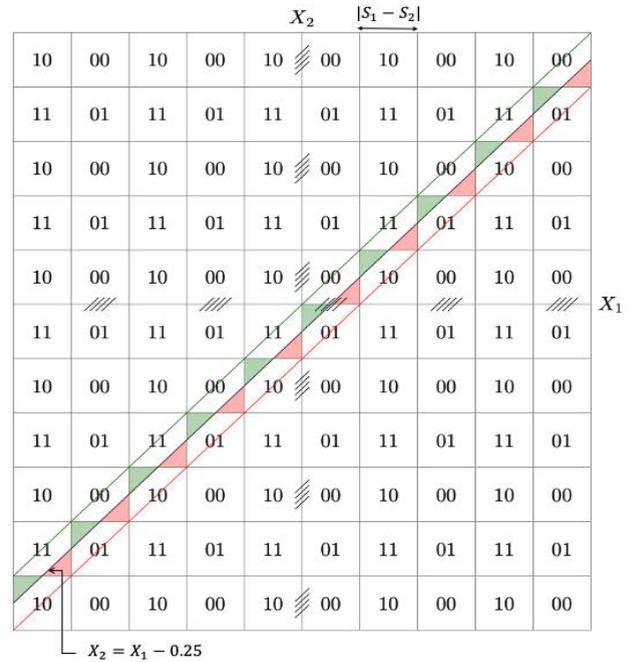
FIGURE 3. Graphical Illustration of XORAP and CLRT: Hatched green area XORAP decides $\mathcal{H}_1$, CLRT decides $\mathcal{H}_0$; Hatched red area XORAP decides $\mathcal{H}_0$, CLRT decides $\mathcal{H}_1$.

According to Taylor series expansion, by ignoring the second and higher orders of ϵ, the CLRT separation line can be obtained as

$$X_1 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} (1 - 3\epsilon)X_2 + \frac{1 + 3\epsilon}{4}. \tag{14}$$

The CLRT divides the plane into two parts, with those points (X_1, X_2) falling below the straight line $(X_2 = X_1 - 0.25, \epsilon \to 0)$ classified as decision $\mathcal{H}_1$ and those falling above the line classified as decision $\mathcal{H}_0$.

In Fig. 3, the line above the CLRT separation line is the line through the origin, with slope 1 and the line below the CLRT separation line is represented as $X_2 = X_1 - 0.5$. At $r = 1$, the probability distribution of (X_1, X_2) is degenerate along the line $X_1 = X_2$, when $\mathcal{H}_0$ is true, and the distribution is degenerate along the line $X_2 = X_1 - 0.5$, when $\mathcal{H}_1$ is true. The shaded green area corresponds to the region where XORAP decides $\mathcal{H}_1$, whereas CLRT decides $\mathcal{H}_0$. The opposite is true for the shaded red area. Under $\mathcal{H}_0$, as r gets closer and closer to 1, the probability distribution keeps increasingly getting concentrated along the $X_1 = X_2$ line. Hence, the CLRT will make less and less error (error happens when samples fall below the CLRT separation line) as r gets closer and closer to 1. In the area above the $X_1 = X_2$ line, there are regions labeled as '01' or '10', where the XORAP decides $\mathcal{H}_1$, resulting in decision errors for XORAP, when $\mathcal{H}_0$ is true. However, the probability of this happening will be negligible and will decrease exponentially to zero as $r \to 1$.

Now, consider hypothesis $\mathcal{H}_0$ being true. What differentiates the rates of decrease of the probabilities of false alarm

of CLRT and XORAP is the difference in the probability mass over the areas within the band, flanked by the lines, $X_2 = X_1$ and $X_2 = X_1 - 0.5$, where the two decision schemes make decision $\mathcal{H}_1$. The probability mass over the green shaded area, where XORAP decision differs from CLRT, contributes to XORAP making an incorrect decision, $\mathcal{H}_1$. In the shaded red area, where XORAP decision differs from CLRT, the XORAP makes the correct decision, $\mathcal{H}_0$. Since the probability mass is increasingly getting concentrated along the $X_2 = X_1$ line, as r tends to 1, the probability mass over the red area, which is below, and touching at points along the $X_2 = X_1 - 0.25$ line, decreases to zero very fast. Hence, this addition to the probability of correct decision for the XORAP has no impact on the rate at which XORAP probability of error (P_f for XORAP) approaches zero. In contrast, as r tends to 1, the probability mass over the green area, which is closer to, and touching at points along the $X_2 = X_1$ line, goes to zero at a slower rate. Although one cannot predict from the graph the rate at which the error would decrease to zero, the analytical expression in *Theorem 2* shows that this is of the order $\sqrt{1-r}$. A similar conclusion can be drawn when $\mathcal{H}_1$ is true.

F. SUBOPTIMAL SOLUTIONS FROM THE GENETIC ALGORITHM (GA)

Finding the sensor decision rules and the fusion rule that minimize the Bayes error is an non-linear integer programming problem [14]. In this subsection, the GA is used to find suboptimal solutions for both parallel and tandem topologies. The flowchart of this GA-based method is shown in Fig. 4.

1) GA OPERATIONS

Key operations carried out in the GA are as follows.

Initialization: In the beginning, n_p numbers of solutions are randomly generated, each of which is a binary sequence that is converted from the sensors' decision rules as follows.

1) Divide the observation range of the k-th ($k = 1, 2$) sensor into a number of sub-intervals, such as S_k, with equal space division.
2) Assume that sensor k has B_k reporting bits. Sensor observation X_k is quantized by assigning each sub-interval with one value from the set $\{1, 2, \ldots, Q_k\}$ where $Q_k = 2^{B_k}$. Quantized values of all sub-intervals are concatenated into a sequence.
3) Represent each value of the above sequence with its binary format to obtain a binary sequence b_k with a length of $B_k S_k$ bits, $k = 1, 2$.
4) In the parallel topology, the solution is represented as $b^P = [b_1 b_2]$, which is obtained by concatenating binary sequences from both sensors and is of length $B_1 S_1 + B_2 S_2$. In the tandem topology, $b^T = b_1$ with length of $B_1 S_1$ only, because given the decision rule of sensor 1, the rule of sensor 2 can be obtained by utilizing a person-by-person optimal decision, according to the tests in (15) and (16), see Appendix A.
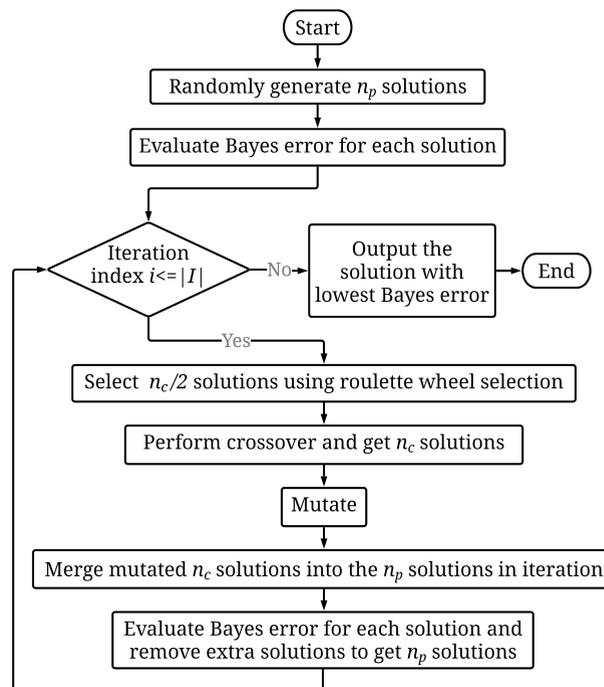
FIGURE 4. The flowchart of the GA implemented for decentralized detection problem.

Bayes error calculation: As the fitness criterion, the Bayes error for each of the n_p solutions is evaluated using the log-likelihood test at the FC.

Selection: A number of $\frac{n_c}{2}$ pairs of solutions (i.e., parents), as determined by the crossover probability p_c, are selected by the roulette wheel selection, where solutions with a lower Bayes error are more likely to be selected.

Crossover and mutation: For each pair of the selected solutions, uniform crossover [18] with a crossover probability p_c is performed to generate a new pair of solutions (i.e., offsprings). In the mutation step, each bit in one solution will be flipped with a probability μ_c from '0' to '1' or vice versa [18].

As a final step in each iteration, the total of n_c new offsprings are merged with the n_p solutions at the beginning of the iteration, among which n_p solutions with lower Bayes error are chosen as input for the next iteration. At the end of the final iteration, the solution with the lowest Bayes error will be output as the final sub-optimal solution. Furthermore, for one-bit quantization, by checking the result of the log-likelihood test used for this final solution, the corresponding fusion rule, either AND rule, OR rule, XOR rule, or "ignoring one of the sensors" rule, can be decided. A similar algorithm and discussion of complexity can be found in [14].

2) SIMULATION RESULTS

The parameters in this simulation are $n_p = 500$, $p_c = 0.8$, $n_c = 400$, $\mu_c = 0.02$, and $|I| = 1000$.
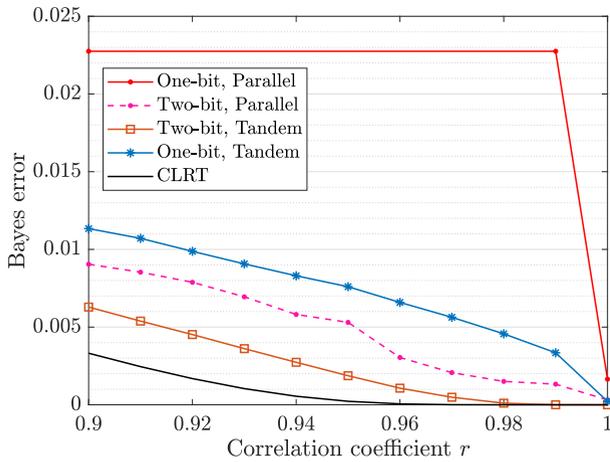
FIGURE 5. Rates of Bayes error decreasing to zero as correlation coefficient approaches 1 from 0.9: decentralized detection cases and CLRT.

First, we consider one-bit quantization for sensor observations, i.e., $B_1 = B_2 = 1$. For several testing points of correlation coefficient $r \in (0.93, 0.9999)$, $s_1 = 0.5, s_2 = 1, \sigma^2 = 1$, and $\eta = 1$ (same parameters used in Table 2), we studied how one-bit quantization pattern might change if LRT fusion with one-bit quantized data is optimized using the GA in the two-sensor parallel topology, like the one used in [16]. It is interesting to notice that the best rules identified by GA are as follow. It is "ignoring sensor 1" for $r = 0.93$, AND rule for $r = 0.95$, and XORAP for $r \geq 0.999$. For $r = 0.99$, GA yields the XOR rule which is slightly different from XORAP, where the quantization intervals of the GA alternate over a certain segment around the origin, but at the two ends, only one interval happens for the remaining parts of the real line. Although the GA can only claim sub-optimal solutions, nevertheless, it also shows that XORAP could be the optimal one-bit fusion rule for r close to 1.

Second, we study the variation of the Bayes error when using both one-bit and two-bit quantizations as r approaches 1 from 0.9, and $s_1 = 2, s_2 = 4, \sigma^2 = 1, \eta = 1$. From Fig. 5, we can observe that the rates of the Bayes errors decreasing to zero, for the two-sensor decentralized detection, with the quantized case are very slow when compared to the exponential decrease exhibited by the CLRT. The coarseness of one-bit parallel scheme is evident from Fig. 5. The average Bayes error remains high for values of r up to about 0.99, before it starts to drop. In contrast, tandem schemes with both one-bit and two-bit sensor reports, as well as the parallel scheme with two-bit sensor reports, show decreasing Bayes error trends over the span of r. Hence, the benefits of using two-bit quantization over one-bit quantization can be observed. Moreover, as to be expected, the tandem scheme exhibits superior performance when compared to the parallel scheme, when both schemes utilize the same number of bits for quantization.

G. DISCUSSION
A special case is what happens when $s_1 = s_2$ and $r \to 1$. Partition interval in the XORAP rule cannot be of zero length.

Since observations at both sensors now will be identical under hypothesis $\mathcal{H}_1$ and identical under $\mathcal{H}_0$ (identical but with only noise) also, the FC essentially has information from any one sensor. The CLRT in such a case will have the performance of a single sensor. Hence, the Bayes error reaches a threshold that is bounded away from zero, determined only by the signal-to-noise ratio (SNR) of a single sensor, as $r \to 1$.

An extension of analysis to the general case of $n > 2$ sensors would have to address several complex issues. The first issue one has to deal with is the general nature of $n \times n$ covariance matrix of the noise components at the n sensors. For analyzing the effect of strong noise correlations, in general, one has to examine the effect of $n(n-1)/2$ pairwise correlation parameters on the Bayes error performance. Basically, the covariance matrix is not dependent upon one correlation, but several correlation parameters. One can restrict attention to some structured covariance matrices, such as the one based on an autoregressive model (AR-1) or an equicorrelated model [15]. In both cases, the covariance depends upon a single correlation parameter r. In [14] GA based solution for three sensors has been presented. For the CLRT, the SNR parameter K can be calculated, as was done for the two-sensor case, and its variation with respect to r examined. In general, K grows unbounded, as $r \to 1$, except for the case of all identical signals at the sensors. However, for any decentralized detection rule, finding the rate of convergence of P_e, as r approaches 1, is a difficult problem. For one reason, the optimal partitions for quantizers at the $n > 2$ sensors and the corresponding fusion center rule are difficult to find, even for one-bit quantization. XOR rule can be defined only for $n = 2$.

III. CONCLUSION
In this paper, we considered the performance of an exclusive-OR fusion rule for detecting the presence or absence of a known deterministic signal in correlated Gaussian noise. The XORAP's performance is perfect when the correlation coefficient r equals 1, but far worse than the centralized likelihood ratio test when $0.99 < r < 1$. Our contributions are 1) to prove the optimality of the XORAP rule for two-sensor tandem and parallel topologies; and 2) to analytically quantify the asymptotic Bayes error of the XORAP, as $r \to 1$, and observe the pronounced difference in the asymptotic performances between XORAP and the centralized likelihood ratio test. Since XORAP is optimal for both two-sensor parallel and tandem topologies, as $r \to 1$, this error convergence result applies to both network topologies.

APPENDIX A
PROOF OF THEOREM 1
Before we prove Theorem 1, we prove the necessary conditions that the decision rules at the two sensors in the tandem configuration need to satisfy for the final decision to be globally optimal. Although these conditions have been stated in [5], [16], we state this as a theorem and provide the

proof here. Also, these conditions are for any bivariate distributions, not necessarily restricted to the bivariate Gaussian family.

Theorem 3: For the global optimality of two-sensor tandem scheme, the decision rules at the two sensors need to satisfy the following coupled equations

$$R_{W_2|W_1=1} = \left\{ x_2 : L_1(x_2) \equiv \frac{\int_{R_{W_1}} p(x_1, x_2|\mathcal{H}_1)dx_1}{\int_{R_{W_1}} p(x_1, x_2|\mathcal{H}_0)dx_1} \geq \eta \right\}, \tag{15}$$

$$R_{W_2|W_1=0} = \left\{ x_2 : L_0(x_2) \equiv \frac{\int_{\overline{R_{W_1}}} p(x_1, x_2|\mathcal{H}_1)dx_1}{\int_{\overline{R_{W_1}}} p(x_1, x_2|\mathcal{H}_0)dx_1} \geq \eta \right\}, \tag{16}$$

where $\eta = \frac{\pi_0}{\pi_1}$ is the decision threshold. L_1 and L_0 are corresponding functions in (15) and (16),

$$\begin{aligned} R_{W_1} = \Big\{ x_1 : &\pi_1 \Big[\int_{R_{W_2|W_1=1}} p(x_1, x_2|\mathcal{H}_1)dx_2 \\ &- \int_{R_{W_2|W_1=0}} p(x_1, x_2|\mathcal{H}_1)dx_2 \Big] \\ > &\pi_0 \Big[\int_{R_{W_2|W_1=1}} p(x_1, x_2|\mathcal{H}_0)dx_2 \\ &- \int_{R_{W_2|W_1=0}} p(x_1, x_2|\mathcal{H}_0)dx_2 \Big] \Big\}. \end{aligned} \tag{17}$$

Proof: The decision variables of sensors, W_i, $i = 1, 2$, are defined in Table 1. Since the final decision is W_2, the Bayes error can be written as

$$\begin{aligned} C(W_2) &= \pi_0 P(W_2 = 1|\mathcal{H}_0) + \pi_1 P(W_2 = 0|\mathcal{H}_1) \\ &= \pi_1 + \pi_0 P(W_2 = 1|\mathcal{H}_0) - \pi_1 P(W_2 = 1|\mathcal{H}_1). \end{aligned} \tag{18}$$

The person-by-person optimal solution is obtained by first treating one of the sensor rules, say sensor 1, as fixed and then sensor 2 rule is optimized for minimum Bayes error, followed by the reverse operation of treating the sensor 2 rule being fixed and optimizing the sensor 1 rule. In the tandem structure, since the sensor 2 decision W_2 depends upon the first sensor decision, which could be either a bit '1' or a bit '0,' decision regions where $W_2 = 1$, that is, where sensor 2 decides $\mathcal{H}_1$, would generally be different for $W_1 = 0$ and $W_1 = 1$. Hence, the decision regions where W_2 decides $\mathcal{H}_1$ are denoted as $R_{W_2|W_1=i}$, $i = 0, 1$. For sensor 1, the decision region for $\mathcal{H}_1$ decision is denoted as R_{W_1}. Obviously, the regions where sensors decide $\mathcal{H}_0$ are denoted as the complements of these regions. First, assuming that the sensor 2 decision rules are fixed, we will minimize $C(W_2)$ by optimizing sensor 1 decision rule. Expanding equation (18), we get

$$\begin{aligned} P(W_2 = 1|\mathcal{H}_0) = \sum_{W_1} \int_{x_1} \int_{x_2} P(W_2 = 1|W_1, x_2, x_1) \\ \cdot P(W_1|x_2, x_1) \cdot p(x_1, x_2|\mathcal{H}_1)dx_2 dx_1. \end{aligned} \tag{19}$$

Since W_2 depends upon W_1 and x_2 only, and W_1 depends upon x_1 only, equation (19) can be reduced to

$$\begin{aligned} &P(W_2 = 1|\mathcal{H}_0) \\ &= \sum_{W_1} \int_{x_1} \int_{x_2} P(W_2 = 1|W_1, x_2) \\ &\quad \cdot P(W_1|x_1) \cdot p(x_1, x_2|\mathcal{H}_0)dx_2 dx_1 \\ &= \int_{x_1} \int_{x_2} P(W_2 = 1|W_1 = 0, x_2) p(x_1, x_2|\mathcal{H}_0)dx_2 dx_1 \\ &\quad + \int_{x_1} P(W_1 = 1|x_1)dx_1 \cdot \int_{x_2} \Big[P(W_2 = 1|W_1 = 1, x_2) \\ &\quad - P(W_2 = 1|W_1 = 0, x_2) \Big] \cdot p(x_1, x_2|\mathcal{H}_0)dx_2. \end{aligned} \tag{20}$$

Similarly,

$$\begin{aligned} &P(W_2 = 1|\mathcal{H}_1) \\ &= \int_{x_1} \int_{x_2} P(W_2 = 1|W_1 = 0, x_2) \\ &\quad \cdot p(x_1, x_2|\mathcal{H}_1)dx_2 dx_1 + \int_{x_1} P(W_1 = 1|x_1)dx_1 \\ &\quad \cdot \int_{x_2} \Big[P(W_2 = 1|W_1 = 1, x_2) - P(W_2 = 1|W_1 = 0, x_2) \Big] \\ &\quad \cdot p(x_1, x_2|\mathcal{H}_1)dx_2. \end{aligned} \tag{21}$$

Therefore,

$$\begin{aligned} &C(W_2) \\ &= \pi_1 + \pi_0 \int_{x_1} \int_{R_{W_2|W_1=0}} p(x_1, x_2|\mathcal{H}_0)dx_2 dx_1 \\ &\quad - \pi_1 \int_{x_1} \int_{R_{W_2|W_1=0}} p(x_1, x_2|\mathcal{H}_1)dx_2 dx_1 \\ &\quad + \int_{x_1} P(W_1 = 1|x_2) \cdot \Big\{ \pi_0 \Big[\int_{R_{W_2|W_1=1}} p(x_1, x_2|\mathcal{H}_0)dx_2 \\ &\quad - \int_{R_{W_2|W_1=0}} p(x_1, x_2|\mathcal{H}_0)dx_2 \Big] \\ &\quad - \pi_1 \Big[\int_{R_{W_2|W_1=1}} p(x_1, x_2|\mathcal{H}_1)dx_2 \\ &\quad - \int_{R_{W_2|W_1=1}} p(x_1, x_2|\mathcal{H}_0)dx_2 \Big] \Big\} dx_1. \end{aligned} \tag{22}$$

To minimize $C(W_2)$, if the item inside the curly bracket of equation (22) is negative, $P(W_1 = 1|x_2) = 1$; otherwise, $P(W_1 = 1|x_2) = 0$. The condition in (17) is reached. Next, we assume that sensor 1 rule is fixed so that we can find the sensor 2 decision regions in order to minimize $C(W_2)$. Rewrite the Bayes error as

$$\begin{aligned} C(W_2) &= \pi_1 + \Big(\int_{R_{W_2|W_1=1}} [\pi_0 \int_{R_{W_1}} p(x_1, x_2|\mathcal{H}_0)dx_1 \\ &\quad - \pi_1 \int_{R_{W_1}} p(x_1, x_2|\mathcal{H}_1)dx_1]dx_2 \Big) \\ &\quad + \Big(\int_{R_{W_2|W_1=0}} [\pi_0 \int_{\overline{R_{W_1}}} p(x_1, x_2|\mathcal{H}_0)dx_1 \\ &\quad - \pi_1 \int_{\overline{R_{W_1}}} p(x_1, x_2|\mathcal{H}_1)dx_1]dx_2 \Big). \end{aligned} \tag{23}$$

Each of the two integrals within the parentheses in the above equation, which contribute to the Bayes error, can be individually minimized as they correspond to the cases of $W_1 = 1$ and $W_1 = 0$, respectively. The first integral is minimized by assigning all points x_2 to $R_{W_2|W_1=1}$ that make the integrand within the square brackets negative. This leads to equation (15) above. Similarly, the second integral is minimized by assigning all points x_2 to $R_{W_2|W_1=0}$ that make the integrand within the square brackets negative. This leads to the equation (16) above. □

Degenerate distributions: As shown below, bivariate Gaussian distribution becomes degenerate distributions, as $r \to 1$, under both the hypotheses. Let $\epsilon = 1 - r$, then $\epsilon \to 0$ as $r \to 1$. The joint probability density function (PDF) of bivariate Gaussian is given by

$$\lim_{r \to 1} p(x_1, x_2 | \mathcal{H}_1) = \lim_{\epsilon \to 0} p(x_1, x_2 | \mathcal{H}_1)$$

$$= \lim_{\epsilon \to 0} \frac{1}{2\pi \sqrt{1 - (1-\epsilon)^2}}$$

$$\cdot e^{-\frac{(x_1-s_1)^2 - 2(1-\epsilon)(x_1-s_1)(x_2-s_2) + (x_2-s_2)^2}{2[1-(1-\epsilon)^2]}}$$

$$\sim \frac{1}{\sqrt{2\pi}\sqrt{2\epsilon}} e^{-\frac{[(x_1-s_1)-(x_2-s_2)]^2}{4\epsilon}}$$

$$\cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{2\epsilon(x_1-s_1)(x_2-s_2)}{4\epsilon}}$$

$$\sim \delta(x_1 - (x_2 + s_1 - s_2)) \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_2-s_2)^2}{2}}$$

$$\sim \delta(x_1 - (x_2 + s_1 - s_2)) \cdot p(x_2 | \mathcal{H}_1), \quad (24)$$

where $p(x_2|\mathcal{H}_1)$ is the marginal PDF of sensor 2 under $\mathcal{H}_1$. Similarly, under $\mathcal{H}_0$,

$$\lim_{r \to 1} p(x_1, x_2 | \mathcal{H}_0) \sim \delta(x_1 - x_2) \cdot p(x_2 | \mathcal{H}_0) \quad (25)$$

Therefore, the joint pdf $p(x_1, x_2)$ is degenerate with support $x_1 = x_2 + s_1 - s_2$ under $\mathcal{H}_1$ and $x_1 = x_2$ under $\mathcal{H}_0$.

We now proceed to prove Theorem 1 by showing that the XORAP rule is the unique solution satisfying both (15) and (16) when $r \to 1$, and then showing that this solution also satisfies (17) when $r \to 1$, which hence establishes the optimality of the XORAP rule in the two-sensor tandem topology.

First, by applying (24) and (25) to (15) and (16), we have

$$L_1(x_2) = \frac{p(x_2|\mathcal{H}_1)}{p(x_2|\mathcal{H}_0)} \cdot$$

$$\frac{\int_{R_{W_1}} \delta(x_1 - (x_2 + s_1 - s_2)) dx_1}{\int_{R_{W_1}} \delta(x_1 - x_2) dx_1}, \quad (26)$$

and

$$L_0(x_2) = \frac{p(x_2|\mathcal{H}_1)}{p(x_2|\mathcal{H}_0)} \cdot \frac{\int_{\overline{R_{W_1}}} \delta(x_1 - (x_2 + s_1 - s_2)) dx_1}{\int_{\overline{R_{W_1}}} \delta(x_1 - x_2) dx_1}.$$

$$(27)$$

Decision W_2 is arrived by examining $L_1(x_2)$ in (26) and $L_0(x_2)$ in (27), as long as they are not indeterminate.

We consider $W_1 = 1$ and $W_1 = 0$ within each of the following four cases.

1) $x_2 + s_1 - s_2 \in R_{W_1}$ AND $x_2 \in R_{W_1}$
When $W_1 = 0$, due to $\int_{\overline{R_{W_1}}} \delta(x_1 - (x_2 + s_1 - s_2)) dx_1 = \int_{\overline{R_{W_1}}} \delta(x_1 - x_2) dx_1 = 0$, $L_0(x_2)$ is indeterminate and after multiplying both sides with the denominator of the left-hand side of (16), it shows that W_2 can be either a '1' or a '0', without affecting the Bayes error.
When $W_1 = 1$,

$$L_1(x_2) = e^{-\frac{s_2^2}{2}} \cdot e^{s_2 x_2} \cdot 1 \underset{W_2=0}{\overset{W_2=1}{\gtrless}} \eta. \quad (28)$$

2) $x_2 + s_1 - s_2 \in R_{W_1}$ AND $x_2 \notin R_{W_1}$
When $W_1 = 1$, $L_1(x_2) = \infty \Rightarrow W_2 = 1$, and when $W_1 = 0$, $L_0(x_2) = 0 \Rightarrow W_2 = 0$.

3) $x_2 + s_1 - s_2 \notin R_{W_1}$ AND $x_2 \in R_{W_1}$
When $W_1 = 1$, $L_1(x_2) = 0 \Rightarrow W_2 = 0$, and when $W_1 = 0$, $L_0(x_2) = \infty \Rightarrow W_2 = 1$.

4) $x_2 + s_1 - s_2 \notin R_{W_1}$ AND $x_2 \notin R_{W_1}$
When $W_1 = 1$, since $\int_{R_{W_1}} \delta(x_1 - (x_2 + s_1 - s_2)) dx_1 = \int_{R_{W_1}} \delta(x_1 - x_2) dx_1 = 0$, $L_1(x_2)$ is indeterminate and after multiplying both sides with the denominator of the left-hand side of (15), it shows that W_2 can be either a '1' or a '0', without affecting the Bayes error. When $W_1 = 0$,

$$L_0(x_2) = e^{-\frac{s_2^2}{2}} \cdot e^{s_2 x_2} \cdot 1 \underset{W_2=0}{\overset{W_2=1}{\gtrless}} \eta. \quad (29)$$

In case 1), $W_2 = 1$ only when $L_1(x_2) \geq \eta$, with x_2 satisfying condition 1). Hence, such a decision in this case will not lead to the zero probability of error, as $r \to 1$. Contrast this with the XORAP rule, which will have zero probability of error as $r \to 1$ (see section II and [14] when $r = 1$). Hence, the case 1) $x_2 + s_1 - s_2 \in R_{W_1}$ and $x_2 \in R_{W_1}$ cannot be in the optimal rule as $r \to 1$.

Similarly in case 4), $W_2 = 1$ only when $L_0(x_2) \geq \eta$, with x_2 satisfying condition 4). Hence, such a decision in this case will not lead to the zero probability of error, as $r \to 1$. Contrast this with the XORAP rule, which will have zero probability of error as $r \to 1$. Hence, the case 4) $x_2 + s_1 - s_2 \notin R_{W_1}$ and $x_2 \notin R_{W_1}$ cannot be in the optimal rule as $r \to 1$.

Therefore, optimal decision rule of sensor 1 will not encounter the cases of $(x_2 + s_2 - s_1 \in R_{W_1}, x_2 \in R_{W_1})$ and $(x_2 + s_2 - s_1 \in \overline{R_{W_1}}, x_2 \in \overline{R_{W_1}})$. Since x_2 is an arbitrary real number, both $x_2 + s_2 - s_1$ and x_2 cannot simultaneously fall in R_{W_1} or simultaneously fall in $\overline{R_{W_1}}$. This tells us that the optimum partition of X_1 to obtain R_{W_1} and $\overline{R_{W_1}}$ has to be alternating segments of length $|s_1 - s_2|$ as shown in Fig. 2.

Since only either $x_2 + s_1 - s_2$ or x_2 can be in any particular interval, let us focus on where x_2 falls in the partition in Fig. 2. Consider the case $W_1 = 1$, if $x_2 \in R_{W_1}$ (i.e., x_2 falls in decision region '1'), from case 3), $W_2 = 0$ due to $L_1(x_2) = 0$;

if $x_2 \in \overline{R_{W_1}}$, from case 2), $W_2 = 1$ due to $L_1(x_2) = \infty$. As a result,

$$W_2 = (W_1 = 1) \oplus I(x_2 \in R_{W_1}). \qquad (30)$$

Now consider the case $W_1 = 0$, if $x_2 \in \overline{R_{W_1}}$ from case 2), $W_2 = 0$ due to $L_0(x_2) = 0$; if $x_2 \in R_{W_1}$, from case 3), $W_2 = 1$ due to $L_0(x_2) = \infty$. As a result,

$$W_2 = (W_1 = 0) \oplus I(x_2 \in R_{W_1}). \qquad (31)$$

Therefore, according to (30) and (31), decision W_2 can be expressed as the Exclusive-OR (XOR) rule between W_1 and the condition of $x_2 \in R_{W_1}$, i.e. $W_2 = W_1 \oplus I(x_2 \in R_{W_1})$.

Next, we show that this XOR rule with alternating partition also satisfies the other coupling condition given by (17). Applying the degenerate distributions to (17), we obtain

$$
\begin{aligned}
R_{W_1} = \Big\{ x_1 : \Big[&\int_{R_{W_2|W_1=1}} \delta(x_2 - (x_1 + s_2 - s_1)) dx_2 \\
&- \int_{R_{W_2|W_1=0}} \delta(x_2 - (x_1 + s_2 - s_1)) dx_2 \Big] \\
> &\frac{\pi_0}{\pi_1} \frac{p(x_1|H_0)}{p(x_1|H_1)} \Big[\int_{R_{W_2|W_1=1}} \delta(x_2 - x_1) dx_2 \\
&- \int_{R_{W_2|W_1=0}} \delta(x_2 - x_1) dx_2 \Big] \Big\}.
\end{aligned}
\qquad (32)
$$

Now, starting from the XOR rule with alternating partition obtained in the above equations (30) and (31), and noticing that $R_{W_2|W_1=1} = \overline{R_{W_1}}$ and $R_{W_2|W_1=0} = R_{W_1}$, it is easy to see that the XORAP rule satisfies the inequality in (32). In other words, the W_1 partition of X_1 by applying (17) to the original partition and XOR rule that are obtained from (15), (16), results in no change to the original XORAP partition. Hence, identical XORAP partition for W_1 and W_2 satisfy the coupled necessary conditions for the optimality of tandem system, which closes the loop and asserts the global optimality of the rule when $r \to 1$.

APPENDIX B
RATE OF DECREASE OF P_2 AS r APPROACHES 1
Consider the term $P_2 = \sum_{i=-\infty}^{\infty} \sum_{k=1}^{\infty} S_{ik}$, where $S_{ik} = \int_{iD}^{(i+1)D} \int_{iD \pm k(2D)}^{(i+1)D \pm k(2D)} f(x, y; r) dx dy$.

Since $f(x, y; r) = f(x) f(y|x; r)$, where $f(x) \sim$ Gaussian$(0, 1)$, $f(y \mid x; r) \sim$ Gaussian$(rx, (1 - r^2))$,

$$
S_{ik} = \int_{iD}^{(i+1)D} g(x) \Big[\Phi \Big(\frac{(i+1) D \pm k (2D) - rx}{\sqrt{1 - r^2}} \Big) \\
- \Phi \Big(\frac{iD \pm k (2D) - rx}{\sqrt{1 - r^2}} \Big) \Big] dx, \qquad (33)
$$

where $\Phi(\cdot)$ is the CDF of standard Gaussian random variable and $g(\cdot)$ is the standard Gaussian density. Without loss of generality, consider $+k(2D)$ in the above integral. Similar analysis for $-k(2D)$ leads to the same conclusion.

As $r \to 1$,

$$
S_{ik} = \int_{iD}^{(i+1)D} g(x) \Big[Q \Big(\frac{(i + 2k)D - x}{\sqrt{1 - r^2}} \Big) \\
- Q \Big(\frac{(i + 2k)D + D - x}{\sqrt{1 - r^2}} \Big) \Big] dx. \qquad (34)
$$

Let $z = x - iD$, $t = \frac{2kD - z}{\sqrt{1 - r^2}}$. Observe that $k = 1, 2, \dots$,

$$
\lim_{r \to 1} S_{ik} = \int_0^D g(z + iD) \\
\cdot \Big[Q(t) - Q \Big(t + \frac{D}{\sqrt{1 - r^2}} \Big) \Big] dz. \qquad (35)
$$

As $r \to 1$, $t \to \infty$, both $Q(\cdot)$ terms above go towards 0. Using the asymptotic expansion of $Q(\cdot)$ function, we get the difference of two $Q(\cdot)$ terms as

$$
\lim_{t \to \infty} \frac{1}{t\sqrt{2\pi}} e^{\frac{-t^2}{2}} - \frac{1}{\sqrt{2\pi} \Big(t + \frac{D}{\sqrt{1-r^2}} \Big)} e^{-\frac{1}{2} \Big(t + \frac{D}{\sqrt{1-r^2}} \Big)^2}
$$
$$
= \lim_{t \to \infty} \frac{1}{t\sqrt{2\pi}} e^{-\frac{t^2}{2}}.
$$

Also, $g(z + iD) \leq g(iD)$, $\frac{1}{2kD - z} \leq \frac{1}{2kD - D}$. Hence,

$$
\lim_{r \to 1} S_{ik} \leq \frac{g(iD)}{\sqrt{2\pi} ((2k - 1) D)} \sqrt{1 - r^2} \\
\cdot \int_0^D e^{-\frac{(2kD - z)^2}{2(1 - r^2)}} dz. \qquad (36)
$$

Since $\int_0^D e^{-\frac{(2kD - z)^2}{2(1-r^2)}} dz < D e^{-\frac{[(2k-1)D]^2}{2(1-r^2)}}$,

$$
\lim_{r \to 1} S_{ik} \leq g(iD) \frac{\sqrt{1 - r^2}}{\sqrt{2\pi}(2k - 1)} e^{-\frac{((2k-1)D)^2}{2(1-r^2)}}. \qquad (37)
$$

Consider $T^* \triangleq \sum_{k=1}^{\infty} \frac{\sqrt{1-r^2}}{\sqrt{2\pi}(2k-1)} e^{-\frac{((2k-1)D)^2}{2(1-r^2)}}$. T^* is independent of i and the exponential term goes to zero as $r \to 1$. Hence, T^* goes to zero exponentially. Therefore,

$$
\lim_{r \to 1} P_2 \leq \sum_{i=-\infty}^{\infty} g(iD) T^* \\
= \Big(1 + \frac{2}{\sqrt{2\pi}} (e^{-\frac{D^2}{2}} + e^{-\frac{4D^2}{2}} + ..) \Big) T^*. \qquad (38)
$$

goes to zero exponentially fast.

REFERENCES
[1] P. Varshney, *Distributed Detection and Data Fusion*. New York, NY, USA: Springer, 1997.
[2] J. N. Tsitsiklis and M. Athans, "On the complexity of decentralized decision making and detection problems," *IEEE Trans. Autom. Control*, vol. 30, no. 5, pp. 440–446, May 1985.
[3] L. Cao and R. Viswanathan, "Divergence-based soft decision for error resilient decentralized signal detection," *IEEE Trans. Signal Process.*, vol. 62, no. 19, pp. 5095–5106, Oct. 2014.
[4] P. Willett, P. F. Swaszek, and R. S. Blum, "The good, bad and ugly: Distributed detection of a known signal in dependent Gaussian noise," *IEEE Trans. Signal Process.*, vol. 48, no. 12, pp. 3266–3279, Dec. 2000.

[5] S. Yagnik, R. Viswanathan, and L. Cao, "Performance of two-sensors tandem network for detecting deterministic signals in correlated Gaussian noise," in *Proc. 54th Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2020, pp. 1–6.

[6] E. Akofor and B. Chen, "On optimal fusion architecture for a two-sensor tandem distributed detection system," in *Proc. IEEE Global Conf. Signal Inf. Process.*, Dec. 2013, pp. 129–132.

[7] E. Song, X. Shen, J. Zhou, Y. Zhu, and Z. You, "Performance analysis of communication direction for two-sensor tandem binary decision system," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4777–4785, Oct. 2009.

[8] S. Sobhiyeh and M. Naraghi-Pour, "Online detection and parameter estimation with correlated observations," *IEEE Syst. J.*, vol. 15, no. 1, pp. 180–191, Mar. 2021.

[9] E. Akofor and B. Chen, "Interactive fusion in distributed detection: Architecture and performance analysis," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 4261–4265.

[10] S. Venkateswaran and U. Madhow, "Distributed detection with a minimalistic signal model: A framework for exploiting correlated sensing," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 1597–1601.

[11] P.-N. Chen and A. Papamarcou, "Likelihood ratio partitions for distributed signal detection in correlated Gaussian noise," in *Proc. IEEE Int. Symp. Inf. Theory*, Sep. 1995, p. 118.

[12] F. Peng and B. Chen, "Decentralized estimation with dependent Gaussian observations," *IEEE Trans. Signal Process.*, vol. 65, no. 5, pp. 1172–1182, Mar. 2017.

[13] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors Part I. Fundamentals," *Proc. IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.

[14] H. Kasasbeh, L. Cao, and R. Viswanathan, "Soft-decision-based distributed detection with correlated sensing channels," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 3, pp. 1435–1449, Jun. 2019.

[15] H. Kasasbeh, R. Viswanathan, and L. Cao, "Noise correlation effect on detection: Signals in equicorrelated or autoregressive(1) Gaussian," *IEEE Signal Process. Lett.*, vol. 24, no. 7, pp. 1078–1082, Jul. 2017.

[16] X. Sun, S. Yagnik, L. Cao, and R. Viswanathan, "Coupling effect of single-threshold sensor rules in two-sensor tandem network," in *Proc. 55th Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2021, pp. 1–6.

[17] G. Pólya, "Remarks on computing the probability integral," in *Proc. Berkely Symp. Math. Staistics Probab.* Berkeley, CA, USA: Univ. California Press, 1949, pp. 63–78.

[18] T. Point. *Genetic Algorithms Tutorial*. Accessed: Dec. 7, 2021. [Online]. Available: https://www.tutorialspoint.com/genetic_algorithms/

SHAILEE YAGNIK (Student Member, IEEE) received the M.S. degree in electrical engineering from the University of Mississippi, Oxford, MS, USA, in 2019, where she is currently pursuing the Ph.D. degree in electrical engineering. Her research interests include distributed signal detection, wireless sensor networks, and wireless communications.

RAMANARAYANAN VISWANATHAN (Life Fellow, IEEE) received the Ph.D. degree from Southern Methodist University, Dallas, TX, USA, in 1983. He is currently the Chair and a Professor of electrical and computer engineering at the University of Mississippi. Prior to his current assignment, he was a Professor of electrical and computer engineering at Southern Illinois University, Carbondale, IL, USA. His research interests include signal detection and estimation, wireless sensor networks, and wireless communications.

XINGJIAN SUN (Student Member, IEEE) received the bachelor's degree in applied physics from Xidian University, Xi'an, China, in 2012, the master's degree in information and communications engineering from the University of Leicester, U.K., in 2013, and the Ph.D. degree in engineering science with an emphasis in electrical engineering from the University of Mississippi, Oxford, MS, USA. His research interests include wideband and narrowband spectrum sensing, and distributed detection.

LEI CAO (Senior Member, IEEE) received the B.E. degree in electrical engineering from the Hefei University of Technology, Hefei, China, in 1990, the M.S. degree in computer science from the University of Science and Technology of China, Hefei, in 1993, and the Ph.D. degree in electrical engineering from the University of Missouri–Columbia, Columbia, MO, USA, in 2002.

From 1993 to 1998, he was a Lecturer in electronic engineering with the University of Science and Technology of China. Since 2002, he has been working with the University of Mississippi, Oxford, MS, USA, where he is currently a Professor of electrical engineering. His research interests include source and channel coding, wireless communications, and applied probability modeling.

・・・