DISTRIBUTED DETECTION WITH COMMON OBSERVATIONS

Hao Chen

Boise State University
Depart. of Electrical and Computer Engineering
Boise, ID 83712
haochen@boisestate.edu

Tsang-Yi Wang

National Sun Yat-sen University
Inst. of Communications Engineering
Taiwan, ROC
tcwang@faculty.nsysu.edu.tw

ABSTRACT

Distributed detection with dependent observations is always a challenging problem. In this paper, we consider a special dependent case where sensors share some common information. Specifically, we investigate a tandem network with sensor 1 sending a one-bit decision to sensor 2 where the final decision is made. Along with their common observation X_2, sensors 1 and 2 possess their conditionally independent measurements X_1 and X_3, respectively. After obtaining the relationship between the optimal sensor 1 rule and sensor 2 fusion rule, we derive the necessary condition for the optimal sensor decision rules for both sensors. We compare the optimal rules with two suboptimal rules for distributed detection of a constant DC signal in Gaussian noise with various of signal-to-noise ratios.

Index Terms— Distributed Detection, Common Information, Tandem Network, Conditionally Dependent Observations

1. INTRODUCTION

Distributed detection has been an active research area with a focus on the analysis and optimization of the detection performance via the design of local decision rules as well as the fusion rule [1]. While the optimal fusion rule is known to be the likelihood-ratio test (LRT) at the fusion center [2–4], designing local sensor decision rules is much more complicated because of the distributed nature. Most of the results are obtained under the assumption that local sensor observations are assumed to be conditionally independent (CI) given the underlying hypothesis (see [1, 5, 6] and the references therein). For the binary hypotheses testing problem, the optimality of local LR quantization has been established under various conditions [5–8].

The problem of designing the optimal local decision rules becomes significantly more difficult without the conditional independence assumption. The optimal design problem is shown to be an NP-hard problem in general [9]. In addition, *the form* of the optimal sensor decision rules is often unknown and is coupled with other sensor rules and the fusion rule. Even for the binary hypotheses testing problem with binary sensor output, local LRTs are often no longer optimal [10, 11]. Suboptimal approaches are often employed to solve the optimization problem. In [12], some numerical approaches are proposed for distributed detection of weak signals. Another popular approach is to fix the local sensor rules to be local LRTs and to try to tune their thresholds [13, 14]. The remarkable complexity of the optimization problem was demonstrated via a distributed binary hypotheses testing problem where two sensors observe a shift in mean of dependent Gaussian noises [15, 16]. Recently, it was discovered that the local decision rules may admit similar simplified structures as in the CI case for a wide range of scenarios, by modeling the detection system using a novel hierarchical conditional independence (HCI) framework with the addition of a hidden variable [17].

In this paper, we consider another class of distributed detection problems where the conditional dependency is introduced due to the common information shared between sensors. This type of shared information arises, for instance, when sensors have overlapped measurements, e.g., surveillance cameras monitoring from different angles. In such cases, each sensor possesses both the common shared information and their individual "private" information. The design problem is how to best utilize the common information at both the sensors and the fusion center to achieve best possible performance. As a first attempt, we investigate a 2-sensor tandem network with sensor 1 sending a one-bit decision to sensor 2 where the final decision is made. Along with their common observation X_2, sensors 1 and 2 possess their conditionally independent measurements X_1 and X_3, respectively. After obtaining the relationship between the optimal sensor 1 rule and sensor 2 fusion rule, we derive the necessary condition for the optimal sensor decision rules for both sensors. We compare the optimal rules with two suboptimal rules for distributed detection of a constant signal in Gaussian noise.

This work was supported in part by the National Science Foundation under grant ECCS 0925854

2. PROBLEM SETUPS

We consider a binary distributed detection problem consisting of two sensors connected in a serial structure. Sensor 1 observes $\{X_1, X_2\}$ and sensor 2 observes $\{X_2, X_3\}$ where X_2 is the shared common information and X_1, X_3 are "private" information, X_1, X_2, X_3 are jointly conditionally independent given the underlying hypothesis $H \in \{0, 1\}$ such that

$$p(X_1, X_2, X_3|H) = \prod_{i=1}^{3} p(X_i|H). \tag{1}$$

Sensor 1 sends a one-bit decision U_1 based on its own observation to sensor 2 where the final decision $U_0 \in \{0, 1\}$ is made based on the received U_1 and its own observation X_2, X_3. The system design goal is to *jointly design* the sensor 1's decision rule $P(U_1|X_1, X_2) = \gamma_1(X_1, X_2)$ and the corresponding fusion rule $P(U_0|U_1, X_2, X_3) = \gamma_0(U_1, X_2, X_3)$ at sensor 2. Given γ_1, the optimal γ_0 is just a likelihood ratio test (LRT) at sensor 2.

2.1. Intuitive Suboptimal Sensor Rule Designs

Obviously, due to the common information X_2, the observations at sensor 1 and sensor 2 are no longer conditionally independent even if X_1, X_2, X_3 are jointly conditionally independent. As X_2 is completely available at sensor 2, one may tempt to draw the conclusion such that the optimal design γ_1^1 should ignore X_2 at sensor 1 by sending U_1 as a function of X_1 only. It is straightforward to show that the optimal γ_1^1 under this approach is obtained by choosing γ_1 as a LRT based on X_1 with a suitable threshold τ_1. It can be further shown that γ_1^1 can be a locally optimal design with an optimized τ_1. Such approach, however, is often not optimal.

Another somewhat intuitive approach is to ignore the fact that X_2 is also available at sensor 2. As a result, this approach utilizes both the X_1 and X_2 information at sensor 1 by performing a LRT $\gamma_1^2(X_1, X_2)$ based on X_1, X_2. Surprisingly, this approach often yields a better performance than γ_1^1. However, as we will show later, neither is this approach is optimal in general.

3. THE OPTIMAL DECISION RULES

3.1. Evaluating the Distributed Detection Performance

The key to solve the optimal decision rules design problem is to derive the performance measure $P(U_0|H)$. For any given γ_1 and γ_0, we have

$$\begin{aligned}
P(U_0|H) &= \sum_{U_1=0}^{1} \int P(U_0, U_1, x_1, x_2, x_3|H) dx_1 x_2 x_3 \\
&= \sum_{U_1=0}^{1} \int P(U_0|U_1, x_1, x_2, x_3, H) \\
&\qquad \times p(U_1, x_1, x_2, x_3|H) dx_1 x_2 x_3 \\
&\overset{(a)}{=} \sum_{U_1=0}^{1} \int P(U_0|U_1, x_2, x_3) p(U_1|x_1, x_2, x_3, H) \\
&\qquad \times p(x_1, x_2, x_3|H) dx_1 x_2 x_3
\end{aligned} \tag{2}$$

$$\begin{aligned}
&\overset{(b)}{=} \sum_{U_1=0}^{1} \int P(U_0|U_1, x_2, H) p(U_1|x_1, x_2) \\
&\qquad \times p(x_1|H) p(x_2|H) dx_1 x_2,
\end{aligned}$$

where (a) and (b) are obtained by Markov Chain property of the problem.

3.2. Optimal Sensor 1 Rule for a Given Fusion Rule

We examine the problem under the Bayesian framework where the prior probability $P(H = 0) = \pi_0 > 0$ and $P(H = 1) = \pi_1 = 1 - \pi_0 > 0$ are given. The probability of error P_e can be written as

$$\begin{aligned}
P_e &= P(H = 0) P(U_0 = 1|H = 0) \\
&\qquad + P(H = 1) P(U_0 = 0|H = 1) \\
&= \pi_1 - \pi_0 \Big(\frac{\pi_1}{\pi_0} P(U_0 = 1|H = 1) - P(U_0 = 1|H = 0)\Big).
\end{aligned} \tag{3}$$

From (3), minimizing P_e is equivalent to maximizing $C = \frac{\pi_1}{\pi_0} P(U_0 = 1|H = 1) - P(U_0 = 1|H = 0)$. With the expanded expression (2), we have

$$\begin{aligned}
C &= \frac{\pi_1}{\pi_0} P(U_0 = 1|H = 1) - P(U_0 = 1|H = 0) \\
&= \sum_{U_1=0}^{1} \int \Big\{ \frac{\pi_1}{\pi_0} P(U_0 = 1|U_1, x_2, H = 1) \times \\
&\quad p(x_1|H = 1) p(x_2|H = 1) - P(U_0 = 1|U_1, x_2, H = 0) \times \\
&\quad p(x_1|H = 0) p(x_2|H = 0) \Big\} p(U_1|x_1, x_2) dx_1 x_2.
\end{aligned} \tag{4}$$

Given the fusion rule $P(U_0 = 1|U_1, x_2, H)$, to minimize the probability of error, we get the optimal sensor 1 decision rule as given in (5). While $\frac{p(x_1|H=1)p(x_2|H=1)}{p(x_1|H=0)p(x_2|H=0)}$ indeed is the LRT based on X_1, X_2, the optimal local sensor rule is *not* because the other factor $\frac{P(U_0=1|U_1=1,x_2,H=1)-P(U_0=1|U_1=0,x_2,H=1)}{P(U_0=1|U_1=1,x_2,H=0)-P(U_0=1|U_1=0,x_2,H=0)}$ is in general a function of x_2 as well.

$$\gamma_1^o = P(U_1=1|x_1,x_2) = \begin{cases} 1, & \frac{P(U_0=1|U_1=1,x_2,H=1)-P(U_0=1|U_1=0,x_2,H=1)}{P(U_0=1|U_1=1,x_2,H=0)-P(U_0=1|U_1=0,x_2,H=0)} \frac{p(x_1|H=1)p(x_2|H=1)}{p(x_1|H=0)p(x_2|H=0)} > \frac{\pi_0}{\pi_1}; \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

3.3. Optimal Decision Rules for the Gaussian case

Eqn. (5) gives the form of the local sensor rule if the fusion rule is fixed. Next, we consider the scenario where both sensor rules can be jointly optimized. In this case, the optimal fusion rule is the LRT of U_1, X_2, X_3 which in turn is a function of $p(U_1, X_2, X_3|H)$ and eventually, γ_1^o. As a result, γ_1^o is therefore more complicated than a simple form of X_1 or X_1, X_2. Furthermore, there may exist a set of suboptimal rules which are locally optimal but not global optimal.

To shed some light into this somewhat complicated optimization problem, we consider the classic problem of detecting a fixed and known signal in an additive Gaussian noise such that X_1, X_2, X_3 are conditionally independent and,

$$X_i \sim \mathcal{N}(s_i H, 1), \tag{6}$$

where $s_i > 0$ is the signal strength, $i = 1, 2, 3$.

In general, the sensor 2 decision rule is a monotonic increasing function of U_1 such that $P(U_0 = 1|U_1 = 1, x_2, H) - P(U_0 = 1|U_1 = 0, x_2, H) > 0$. In other words, sending $U_1 = 1$ results a higher chance of deciding on $H = 1$ by sensor 2 than sending $U_1 = 0$. Define the "extra" term in (5) as $\exp(g(x_2))$, and $\varphi = \ln \frac{\pi_0}{\pi_1}$ as the log ratio of the two prior probabilities.

Taking the logarithm of (5) and simplifying, we have the optimal sensor 1's rule for a fixed sensor 2 rule is given by

$$\begin{aligned} \gamma_1^o &= P(U_1=1|x_1,x_2) \\ &= \begin{cases} 1, & x_1 > \frac{s_2^2+s_1^2}{2s_1} + \frac{\varphi}{s_1} - \frac{s_2}{s_1}x_2 - \frac{g(x_2)}{s_1}; \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{8}$$

As a result, we have

$$P(U_1=1|x_2,H) = \begin{cases} Q(\frac{s_2^2-s_1^2}{2s_1} + \frac{\varphi}{s_1} - \frac{s_2}{s_1}x_2 - \frac{g(x_2)}{s_1}), & H=1; \\ Q(\frac{s_2^2+s_1^2}{2s_1} + \frac{\varphi}{s_1} - \frac{s_2}{s_1}x_2 - \frac{g(x_2)}{s_1}), & \text{otherwise.} \end{cases} \tag{9}$$

where $Q(\cdot)$ is the standard Gaussian compliment distribution function.

At sensor 2, the likelihood ratio is

$$\frac{p(U_1,X_2,X_3|H=1)}{p(U_1,X_2,X_3|H=0)} = \frac{p(U_1|X_2,H=1)}{p(U_1|X_2,H=0)} \frac{p(X_2,X_3|H=1)}{p(X_2,X_3|H=1)}. \tag{10}$$

The optimal test is the well known LRT such that

$$U_0 = \begin{cases} 1, & x_3 > \frac{s_2^2+s_3^2}{2s_3} + \frac{\varphi}{s_3} - \frac{\ln \frac{p(U_1|X_2,H=1)}{p(U_1|X_2,H=0)}}{s_3} - \frac{s_2}{s_3}x_2; \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

Therefore, given U_1, X_2, H at sensor 2, the final detection performance is

$$P(U_0=1|U_1,x_2,H) = \begin{cases} Q(\frac{s_2^2-s_3^2}{2s_3} + \frac{\varphi}{s_3} - \frac{\ln \frac{p(U_1|X_2,H=1)}{p(U_1|X_2,H=0)}}{s_3} - \frac{s_2}{s_3}x_2), & H=1; \\ Q(\frac{s_2^2+s_3^2}{2s_3} + \frac{\varphi}{s_3} - \frac{\ln \frac{p(U_1|X_2,H=1)}{p(U_1|X_2,H=0)}}{s_3} - \frac{s_2}{s_3}x_2), & H=0. \end{cases} \tag{12}$$

Plugging equations (9) and (12) to (7), we obtain a function of $g(x_2)$ to be solved. Although an analytical solution is not easy to obtain, this function is still relatively simple to be solved with numerical tools.

4. PERFORMANCE EVALUATION

In this section, the performance of the scheme employing the optimal sensor 1 rule (i.e., γ_1^o given in (8)) and the fusion rule given in (12) is compared with the other two simplified suboptimal schemes (mentioned in Section 2), which are as defined below.

Scheme A: In this scheme, sensor 1 completely ignores X_2 as it is available at sensor 2 already. Therefore, sensor 1 makes decision U_1 by performing LRT on X_1 (equivalent to quantize X_1). Sensor 2 takes the optimal LRT on U_1, X_2, X_3. As a result, the likelihood function of Sensor 2 can be written as $P(U_1|H)p(X_2|H)p(X_3|H)$. Notably, in Scheme A, we have sensor 1's decision rule given by

$$\gamma_1^A = P(U_1=1|x_1,x_2) = \begin{cases} 1, & x_1 > \frac{\varphi}{s_1} + \frac{s_1}{2}; \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

Accordingly, when comparing with (8), the corresponding $g(x_2)$ for Scheme A is $g(x_2) = \frac{s_2^2}{2}s^2 - s_2x_2$.

Scheme B: In this scheme, sensor 1 ignores the fact that Sensor 2 (the fusion center) also has the common observation X_2 and makes decision U_1 based on the optimal LRT of (X_1, X_2). Same as the other cases, at the fusion center, sensor 2 performs optimal LRT based on U_1, X_1, X_2. The likelihood function can be written as $P(u_1|X_2, H)P(X_2|H)p(X_3|H)$. Notably, in Scheme B, we have sensor 1's decision rule given by

$$\gamma_1^B = P(U_1=1|x_1,x_2) = \begin{cases} 1, & x_1 > \frac{\varphi}{s_1} + \frac{s_2^2+s_1^2}{2s_1} - \frac{s_2}{s_1}x_2; \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

Accordingly, when comparing with (8), the corresponding $g(x_2)$ is always 0 for Scheme B.

For comparison purposes, the performance loss of each scheme with respective to a benchmark scheme is illustrated.

$$g(x_2) \equiv \ln \frac{P(U_0 = 1|U_1 = 1, x_2, H = 1) - P(U_0 = 1|U_1 = 0, x_2, H = 1)}{P(U_0 = 1|U_1 = 1, x_2, H = 0) - P(U_0 = 1|U_1 = 0, x_2, H = 0)}. \tag{7}$$

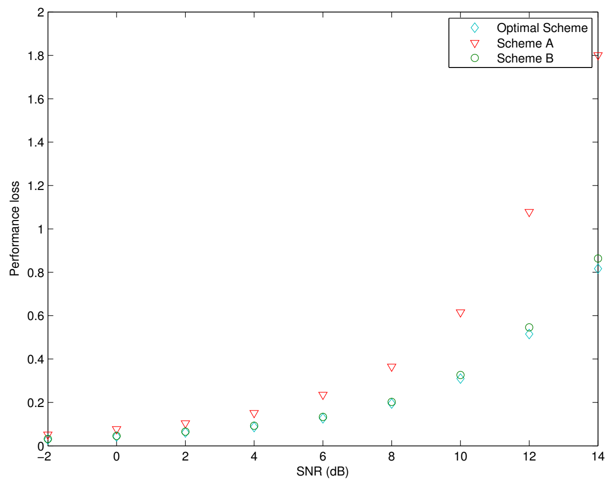

Fig. 1. Performance loss ratios of optimal scheme, Scheme A, and Scheme B for case in which $c_1 = c_2 = 1$ and $c_3 = 0.5$.

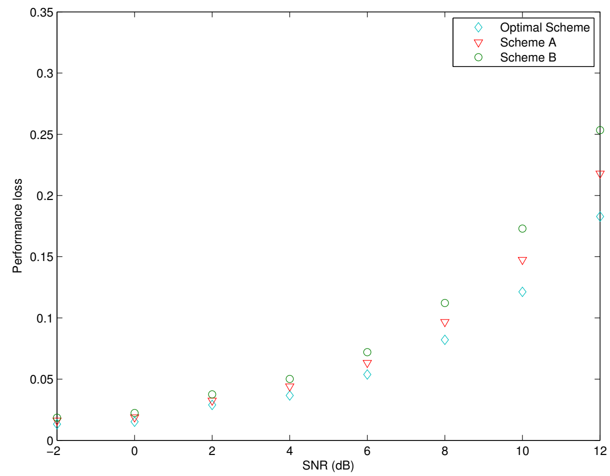

Fig. 2. Performance loss ratios of optimal scheme, Scheme A, and Scheme B for case in which $c_1 = c_3 = 0.5$ and $c_2 = 1$.

The *benchmark scheme* is the *centralized test* in which all the observation X_1, X_2, X_3 are assumed to be available to sensor 2, or equivalently, the bandwidth between sensors 1 and 2 are unlimited. The sensor makes a decision using LRT based on X_1, X_2, X_3, which is given by

$$U_0 = \begin{cases} 1, & s_1x_1 + s_2x_2 + s_3x_3 > \varphi + \frac{s_1^2+s_2^2+s_3^2}{2}; \\ 0, & \text{otherwise.} \end{cases} \tag{15}$$

As a result, the probability of error of the centralized test ($P_{e,\text{cen}}$) is given by

$$P_{e,\text{cen}} = \pi_1 \left(1 - Q\left(\frac{\varphi}{\sqrt{s_1^2+s_2^2+s_3^2}} - \frac{\sqrt{s_1^2+s_2^2+s_3^2}}{2}\right)\right) + \pi_0 Q\left(\frac{\varphi}{\sqrt{s_1^2+s_2^2+s_3^2}} + \frac{\sqrt{s_1^2+s_2^2+s_3^2}}{2}\right).$$

Given the probability of error of the centralized test ($P_{e,\text{cen}}$), we define the performance loss ratio of each scheme as $(P_e - P_{e,\text{cen}})/P_{e,\text{cen}}$, where P_e is the error probability of Scheme A, Scheme B, or the optimal Scheme.

In evaluating the performance, we set s_i given in (6) as $s_i = c_i s$ and define signal-to-noise ratio (SNR) as s^2 (because of unit variance given in (6)). Figure 4 plots the performance loss ratios of the optimal scheme, Scheme A, and Scheme B for the case in which $c_1 = c_2 = 1$ and $c_3 = 0.5$. As can be seen, the optimal scheme outperforms the other two schemes. In addition, it can be observed that Scheme B yields a near-optimal performance when the SNR of X_3 is low compared to others. Figure 4 plots the performance loss ratios of the optimal scheme, Scheme A, and Scheme B for the case in which $c_1 = c_3 = 0.5$ and $c_2 = 1$. The optimal scheme again outperforms the other two schemes. On the other hand, it can be observed that Scheme A, beside of its simplicity, appears to be near-optimal when the SNR of X_2 is low.

5. CONCLUSION

We have studied a distributed detection problem in which the participating sensors share some common information. Specifically, we have investigated a 2-sensor tandem network with sensor 1 sending a one-bit decision to sensor 2 where the final decision is made. Along with their common observation X_2, sensor 1 and 2 possess their conditionally independent measurements X_1 and X_3, respectively. We have derived the necessary condition for the optimal sensor decision rules for both sensors in the considered tandem network. Finally, We have compared the derived optimal rules with two suboptimal rules for the case of detecting a constant DC signal in Gaussian noises with potentially different SNRs, and the obtained results confirm the optimality of the derived optimal rules.

6. REFERENCES

[1] P. K. Varshney, *Distributed Detection and Data Fusion*. Springer, 1997.

[2] Z. Chair and P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, pp. 98–101, 1986.

[3] E.Drakopoulos and C. C. Lee, "Optimum multisensor fusion of correlated local decisions," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 27, no. 4, pp. 593 – 606, July 1991.

[4] M. Kam, Q. Zhu, and W. S. Gray, "Optimal data fusion of correlated local decisions in multiple sensor detection systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 3, pp. 916–920, July 1992.

[5] J. N. Tsitsiklis, "Decentralized detection," in *Advances in Signal Processing*, H. V. Poor and J. B. Thomas, Eds. JAI Press, 1993, vol. 2, pp. 297–344.

[6] D. Warren and P. Willett, "Optimum quantization for detector fusion: some proofs, examples and pathology," *Journal of the Franklin Institute*, vol. 336, pp. 323–359, 1999.

[7] B. Chen, L. Tong, and P. K. Varshney, "Channel-aware distributed detection in wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 16–26, July 2006.

[8] H. Chen, B. Chen, and P. K. Varshney, "Further results on the optimality of the likelihood-ratio test for local sensor decision rules in the presence of nonideal channels," *IEEE Trans. Info. Theory*, vol. 55, Feb. 2009.

[9] J. N. Tsitsiklis and M. Athans, "On the complexity of decentralized decision making and detection problems," *IEEE Transactions on Automatic Control*, vol. AC-30, no. 5, pp. 440–446, May 1985.

[10] Z.-B. Tang, K. R. Pattipati, and D. L. Kleinman, "A distributed m-ary hypotheses testing problem with dependent observations," *IEEE Transactions on Automatic Control*, vol. 37, no. 7, pp. 1042–1046, July 1992.

[11] Q. Yan and R. S. Blum, "Distributed signal detection under the neyman-pearson criterion," *IEEE Trans. Info. Theory*, vol. 47, no. 4, pp. 1368–1377, May 2001.

[12] R. S. Blum and S. Kassam, "Optimum distributed detection of weak signals in dependent sensors," *IEEE Trans. Info. Theory*, vol. 38, pp. 1066–1079, May 1992.

[13] V. Aalo and R. Viswanathan, "On distributed detection with correlated sensors: two examples," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 25, no. 3, pp. 414 – 421, May 1989.

[14] X. Lin and R. S. Blum, "Numerical solutions for optimal distributed detection of known signals in dependent t-distributed noise-the two sensor problem," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Nov. 1998, pp. 613–617.

[15] P.-N. Chen and A. Papamarcou, "Likelihood ratio partitions for distributed signal detection in correlated gaussian noise," in *Proc. IEEE Int. Symp. Inform. Theory*, Oct. 1996, p. 118.

[16] 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. Sign. Proc.*, vol. 48, no. 12, pp. 3266–3279, Dec. 2000.

[17] H. Chen, B. Chen, and P. Varshney, "A new framework for distributed detection with conditionally dependent observations," *IEEE Transactions on Signal Processing*, vol. 60, pp. 1409–1419, 2012.