Consensus+Innovations Detection: Phase Transition Under Communication Noise

Dušan Jakovetić*,†, José M. F. Moura*, João Xavier†

Abstract—We consider the tradeoffs between sensing and communication in a consensus+innovations distributed detection problem when the local communications among agents are noisy. Intuitively, we can expect that the error performance of the distributed detector is affected by both the sensing noise and the noise corrupting the communication among agents in the network. Too little communication (cooperation) and the distributed detector error performance will be dominated by the sensing noise. Too much communication and the detector error performance is dominated by the communication noise. We make this tradeoff precise through a large deviations analysis, i.e., by studying the exponential decay rate of the probability of error of the consensus+innovations distributed detector at each agent. Under a mild assumption of network connectedness, we show: 1) the weight sequences affecting the consensus and innovations potentials in the distributed detector need to be carefully designed for the error probability at every agent detector to decay exponentially fast; 2) the network exhibits a phase transition with respect to the communication noise power. Below a threshold on the communication noise power, cooperation (communication) among agents improves the error detection performance; above threshold, inter-agent communication does not enhance the error detection performance.

I. INTRODUCTION

We study the large deviations performance of consensus+innovations distributed detection under communication noise. We consider N agents, connected in a generic network; agents sense the environment and communicate over *noisy* links to detect the event of interest. The agents employ a consensus+innovations distributed detector, which operates as follows. At a time k, agent i first communicates its decision variable to its neighbors, and receives from neighbors their decision

The work of the first and third authors is partially supported by: the Carnegie Mellon—Portugal Program under a grant from the Fundação para a Ciência e Tecnologia (FCT) from Portugal; by FCT grants CMU-PT/SIA/0026/2009; and by ISR/IST plurianual funding (POSC program, FEDER). The work of the first and second authors is partially supported by NSF under grants CCF-1011903 and CCF-1018509, and by AFOSR grant FA95501010291. D. Jakovetić holds a fellowship from FCT.

*Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213-3890, USA djakovet@andrew.cmu.edu, moura@ece.cmu.edu.

†Institute for Systems and Robotics (ISR), Instituto Superior Técnico (IST), Technical University of Lisbon, Lisbon, Portugal jxavier@isr.ist.utl.pt.

variables. This inter-agent communication is corrupted by an additive communication noise with certain variance. Upon reception of the neighbors' noisy copies of their decision variables, agent *i* makes a *weighted average* of its own and the neighbors' decision variables, and incorporates the log-likelihood ratio from its newly acquired measurement.

In this paper, we ask the following two questions. First, can we design the time varying consensus weights so that, at each agent i, detection error probability decays to zero exponentially fast $\sim e^{-kC_{i,\text{coop}}}$ (with $C_{i,\text{coop}}$ strictly positive), even if all except at least one agent cannot detect the event of interest in isolation? Second, under what conditions do we have a communication payoff, so that the worst agent under communication performs better than the best agent without communication? More precisely, suppose that agent i has the detection error's exponential decay rate $C_{i,isol}$ when in isolation, and $C_{i,\text{coop}}$ with the consensus+innovations detector. Then, we ask when is $\min_{i=1,...,N} C_{i,\text{coop}} >$ $\max_{i=1,\ldots,N} C_{i,\text{isol}}$? Regarding the first question, we answer it affirmatively. We show that, with the weights of type $\alpha_k = b_0/(a+k)$, where a > 0, and $b_0 > 0$ is greater than a certain threshold, all agents i achieve a positive decay rate $C_{i,\text{coop}}$, irrespective of the communication noise power. With respect to the second question, we consider equal agents (with equal positive $C_{i,isol}$'s), and we show that a phase transition occurs: below a threshold on the communication noise power, agents achieve a communication payoff. Further, we quantify this threshold in terms of the system parameters sensing signal and noise, and the network's connectivity.

We further show by a numerical example that the agents' exponential decay rates $C_{i,\text{coop}}$'s significantly depend on the choice of the parameter b_0 . We also analytically demonstrate how the optimal choice of b_0 balances two opposing effects: the increase of b_0 increases the "information flow" from the neighbors (positive effect), but it also injects more communication noise in the agent i's decision variable (negative effect).

We end the introduction with a brief literature review on the consensus+innovations distributed detection. References [1], [2], [3] construct distributed detectors based on the LMS and RLS based diffusion

distributed estimators, e.g., [4]. Reference [5] proposes a distributed algorithm for the change detection problem. Reference [6] proposes and analyzes the running consensus distributed detector; running consensus is further analyzed in, e.g., [7], [8]. All the references above assume no additive communication noise, and either static or failing underlying network's links. In contrast, we study here, and also in a companion journal paper [9], consensus+innovations distributed detection under additive communication noise (and no link failures). Reference [10] also allows for additive communication noise (and no link failures). Unlike [10], where the detector proposed therein has a sub-exponential decay of the error probability under unequal agents (unequal $C_{i,isol}$'s), the distributed detector here achieves an exponential decay of the error probability, at each agent i.

The remainder of the paper is organized as follows. The next paragraph introduces notation. Section II describes the problem model. Section III presents the consensus+innovations distributed detector. Section IV presents our main results, by addressing our first question – exponential decay of the error probability at each agent, and our second question – communication payoff. Section IV also illustrates by a numerical example the performance of the distributed detector. Finally, Section V concludes the paper.

Throughout, we adopt the following notation. We use lower and upper boldface letters to represent vectors and matrices; A_{ij} or $[\mathbf{A}]_{ij}$ are the (i,j)-th entry of a matrix A; a_i or $[a]_i$ are the *i*-th entry of a vector $\mathbf{a}; \mathbf{A}^{\top}$ and \mathbf{A}^{-1} are the transpose and inverse of $\mathbf{A};$ $A \succ 0$ means that the matrix A is positive definite; **I**, **1**, and e_i are the identity matrix, the column vector with unit entries, and the *i*-th column of \mathbf{I} ; $\mathbf{J} := (1/N)\mathbf{1}\mathbf{1}^{\top}$ is the $N \times N$ ideal averaging matrix; $\|\cdot\| = \|\cdot\|_2$ is the Euclidean (respectively, spectral) norm; $\lambda_i(\cdot)$ the i-th smallest eigenvalue; \otimes the Kronecker product of matrices; Diag (a) the diagonal matrix with the diagonal equal to the vector \mathbf{a} ; $\mathbf{a} = \mathbf{Vec}(\mathbf{A})$ is the vector that stacks columns of A, and the "inverse" operation is $\mathbf{A} = \mathbf{Vec}^{-1}(\mathbf{a})$; $|\mathcal{A}|$ is the cardinality of \mathcal{A} ; $\mathbb{E}[\cdot]$, $Var(\cdot)$, $Cov(\cdot)$, and $\mathbb{P}(\cdot)$ are the expected value, the variance, the covariance, and probability operators.

II. PROBLEM MODEL

Subsection II-A introduces the sensing and communication models that we assume, and Subsection II-B presents the consensus+innovations distributed detector.

Sensing model. We consider the binary hypothesis test H_1 versus H_0 . Each agent i, at each time k, obtains a scalar measurement $y_i(k)$, modeled as follows:

$$H_1: y_i(k) = m_i + \zeta_i(k)$$
 (1)

$$H_0: y_i(k) = \zeta_i(k). \tag{2}$$

Here, m_i is a constant signal, and $\zeta_i(k)$ is a zero mean additive *sensing noise*. The prior probabilities are $0 < P(H_1), P(H_0) < 1$. Introduce the following compact notation:

$$\mathbf{y}(k) = (y_1(k), \dots, y_N(k))^{\top}$$

$$\mathbf{m} = (m_1, \dots, m_N)^{\top}$$

$$\boldsymbol{\zeta}(k) = (\zeta_1(k), \dots, \zeta_N(k))^{\top}.$$
(3)

We assume the following sensing model.

Assumption 1 (Sensing model) The sensing noise $\{\zeta(k)\}$ is a zero mean i.i.d. Gaussian sequence (possibly spatially correlated) with $\mathbf{Cov}(\zeta(k)) = \mathbf{S}_{\zeta} \succ 0$. Furthermore, there exists an agent i such that $m_i \neq 0$.

Note that certain entries m_i may be equal zero.

Communication graph. Agents $i \in \mathcal{V} := \{1, ..., N\}$ are situated in a graph $G = (\mathcal{V}, \mathcal{E})$, where \mathcal{E} is the set of undirected links. We assume the following.

Assumption 2 (Communication graph) Graph G is connected, undirected, and simple (no self/multiple links).

Introduce also the $N \times N$ symmetric graph Laplacian matrix \mathcal{L} : $[\mathcal{L}]_{ij} = -1$, if $\{i,j\} \in \mathcal{E}$, $i \neq j$; $[\mathcal{L}]_{ij} = 0$ if $\{i,j\} \notin \mathcal{E}$, $i \neq j$; and $[\mathcal{L}]_{ii} = |O_i|$, where O_i is the agent i's neighborhood set $O_i := \{j : \{i,j\} \in \mathcal{E}, j \neq i\}$.

Communication noise. With our consensus+innovations distributed detector, each agent i, at each time k, receives its neighbor j's decision variable $x_j(k)$ $(j \in O_i)$, corrupted by additive communication noise:

$$z_{ij}(k) = x_j(k) + \nu_{ij}(k).$$
 (4)

Here $\nu_{ij}(k)$ is the communication noise. Similarly, agent $j \in O_i$ at time k receives from agent i $z_{ji}(k) = x_i(k) + \nu_{ji}(k)$, where $\nu_{ij}(k)$ and $\nu_{ji}(k)$ are different random variables. Note that, for each k, there is one distinct variable $\nu_{ij}(k)$ per each directed pair (i,j) such that $\{i,j\} \in \mathcal{E}$. Also, introduce:

$$v_i(k) = \sum_{j \in O_i} \nu_{ij}(k), \quad i = 1, \dots, N$$

$$\mathbf{v}(k) = (v_1(k), \dots, v_N(k))^{\top}.$$
(5)

We make the following assumption on the communication noise.

Assumption 3 (Communication noise) The communication noise $\{\mathbf{v}(k)\}$ is a Gaussian, temporally i.i.d. sequence, with $\mathbf{Cov}(\mathbf{v}(k)) =: \mathbf{S_v} \succ 0$. Furthermore, $v_i(k)$ and $\zeta_j(s)$ are mutually independent over all i, j, k, s.

III. CONSENSUS+INNOVATIONS DISTRIBUTED DETECTOR

We now present our consensus+innovations distributed detector. The detector has a form similar to that of the running consensus detector [6]. The difference is in the time varying consensus weights; the choice of the weights will be specified in Subsection III-A. Specifically, at each time k, each agent i thresholds its decision variable against the zero threshold to decide between the two hypothesis:

$$x_i(k) \underset{H_0}{\overset{H_1}{\gtrless}} 0. \tag{6}$$

Agent i updates recursively the decision variable $x_i(k)$ through the following rule:

$$x_i(k+1) = \frac{k}{k+1} (1 - d_i \alpha_k) x_i(k)$$
 (7)

$$+ \frac{k}{k+1} \alpha_k \sum_{j \in O_i} (x_j(k) + \nu_{ij}(k)) + \frac{1}{k+1} \eta_i(k+1),$$

for k = 1, 2, ..., with the initialization $x_i(1) = \eta_i(1)$. Here, α_k is the time varying consensus weight, specified in Subsection III-A, and $\eta_i(t)$ is the local innovation:

$$\eta_i(t) = \left[\mathbf{S}_{\zeta}^{-1} \mathbf{m} \right]_i \left(y_i(t) - \frac{m_i}{2} \right).$$
(8)

Note that $\sum_{i=1}^{N} \eta_i(t) = \mathbf{m}^{\top} \mathbf{S}_{\zeta}^{-1} \left(\mathbf{y}(t) - \frac{\mathbf{m}}{2} \right)$ – the (hypothetical) log likelihood ratio based on the sample from all agents $\mathbf{y}(t)$. Quantity $\eta_i(t)$ is an affine function of agent i's locally available measurement $y_i(t)$. To calculate $\eta_i(t)$, agent i needs the statistics $\left[\mathbf{S}_{\zeta}^{-1} \mathbf{m} \right]_i$ and m_i . We assume that these are acquired in the network training period.

Matrix format. We re-write (7) in matrix format. Define the decision vector $\mathbf{x}(k)$, the innovations vector $\boldsymbol{\eta}(k)$, and the weight matrix $\mathbf{W}(k)$:

$$\mathbf{x}(k) = (x_1(k), \cdots, x_N(k))^{\top}$$

$$\boldsymbol{\eta}(k) = (\eta_1(k), ..., \eta_N(k))^{\top}$$

$$\mathbf{W}(k) = \mathbf{I} - \alpha_k \boldsymbol{\mathcal{L}}.$$

$$(9)$$

Innovations' statistics. For subsequent analysis, we need the first and second moments of $\eta(k)$; it can be shown that these equal:

$$\mathbb{E}\left[\boldsymbol{\eta}(k)|H_{1}\right] = -\mathbb{E}\left[\boldsymbol{\eta}(k)|H_{0}\right] := \mathbf{m}_{\boldsymbol{\eta}}$$
(10)
$$= \frac{1}{2}\operatorname{Diag}\left(\mathbf{S}_{\boldsymbol{\zeta}}^{-1}\mathbf{m}\right)\mathbf{m}$$

$$\operatorname{Cov}(\boldsymbol{\eta}(k)) := \mathbf{S}_{\boldsymbol{\eta}}$$
(11)
$$= \operatorname{Diag}\left(\mathbf{S}_{\boldsymbol{\zeta}}^{-1}\mathbf{m}\right)\mathbf{S}_{\boldsymbol{\zeta}}\operatorname{Diag}\left(\mathbf{S}_{\boldsymbol{\zeta}}^{-1}\mathbf{m}\right).$$

$$\text{IV. Main results}$$

Subsection III-A addresses our first question: Can we design the weights α_k in (7) such that each agent

achieves an exponential decay of the error probability. Subsection III-B considers equal agents and addresses the second question: When does the communication pay off?

A. Exponential decay of the error probability

In this subsection, we design the weights α_k in (7), so that each agent achieves an exponential decay of the error probability. We first define formally the exponential decay rate. Consider $P^e_{i,\text{coop}}(k)$ – the (Bayes) error probability with detector (7) at time k and agent i. We define the agent i's exponential decay of the error probability by:

$$0 \le C_{i,\text{coop}} := \lim_{k \to \infty} -\frac{1}{k} \log P_{i,\text{coop}}^e(k).$$

It can be shown that the above limit exists. Further, if $C_{i,\text{coop}} = 0$, then detection error probability either stays non zero, or decays to zero slower than exponentially. The next Theorem finds exactly $C_{i,\text{coop}}$, for all agents i. (For a proof of the Theorem, see [9].)

Theorem 1 (Exponential decay rate) Consider the consensus+innovations distributed detector in (7) under Assumptions 1–3. Set the weights α_k to:

$$\alpha_k = \frac{b_0}{a+k}, \ k = 1, 2, ...,$$
 (12)

where $b_0, a > 0$. Then:

(1) For every agent i,

$$C_{i,\text{coop}} = \frac{1}{2} \frac{[\boldsymbol{\mu}_{\infty}]_i^{2,+}}{[\boldsymbol{\Sigma}_{\infty}]_{ii}},$$

where $z^{2,+} = (\max(z,0))^2$ and:

$$\mu_{\infty} = \lim_{k \to \infty} \mathbb{E} \left[\mathbf{x}(k) | H_1 \right] = \left(\mathbf{I} + b_0 \, \mathcal{L} \right)^{-1} \mathbf{m}_{\eta}$$

$$\Sigma_{\infty} = \lim_{k \to \infty} k \, \mathbf{Cov} \left(\mathbf{x}(k) \right)$$

$$= \mathbf{Vec}^{-1} \left\{ \left(\mathbf{I} + b_0 \, (\mathcal{L} \otimes \mathbf{I} + \mathbf{I} \otimes \mathcal{L}) \right)^{-1} \right.$$

$$\times \left. \left(\mathbf{Vec} \left(\mathbf{S}_{\eta} \right) + b_0^2 \mathbf{Vec} \left(\mathbf{S}_{\mathbf{v}} \right) \right) \right\}.$$

(2) If the parameters b_0 , a satisfy:

$$b_0 > \max \left\{ 0, \frac{\frac{2N \|\mathbf{m}_{\boldsymbol{\eta}}\|}{\mathbf{m}^{\mathsf{T}} \mathbf{S}_{\boldsymbol{\zeta}}^{-1} \mathbf{m}} - 1}{\lambda_2(\mathcal{L})} \right\}$$

$$a > b_0 \lambda_N(\mathcal{L}),$$

then $C_{i,\text{coop}} > 0$, $\forall i$, i.e., each agent achieves the exponential decay rate.

Theorem 1 finds the exponential decay rate $C_{i,\text{coop}} \geq 0$ in terms of the first and second moments of the vector decision variable $\mathbf{x}(k)$. Furthermore, the Theorem says that, if b_0 and a are above certain thresholds, then

 $C_{i,\text{coop}}$ is strictly positive, for all i, and the exponential decay is achieved.

Numerical demonstration. We now illustrate by a numerical example that the weight choice α_k in (12) leads to an exponential decay at every agent under communication noise. We consider a geometric graph with N=10 agents deployed uniformly randomly over a 2D unit square; the agent whose distance is less than a radius are connected by an edge. The resulting graph is connected and has 20 undirected links. The sensing m_i signal is generated randomly. Each entry is generated independently; for each entry i, we toss an unfair coin with Heads probability 0.2; if Heads come out, m_i from the uniform distribution over [0,1]; else, m_i is set to zero. The resulting m_i has 3 nonzero and 7 zero entries, and $\|\mathbf{m}\| \approx 0.84$. The sensing noise is zero mean Gaussian, spatio-temporally independent; the covariance \mathbf{S}_{ζ} is diagonal; each entry $[\mathbf{S}_{\zeta}]_{ii}$ is generated randomly form the uniform distribution on [0,6]. The resulting \mathbf{S}_{ζ} has the norm $\|\mathbf{S}_{\zeta}\| \approx 5.50$. With each directed pair (i,j) ($\{i,j\} \in \mathcal{E}$), we generate the communication noise to be spatio-temporally independent; the noise $\nu_{ii}(k)$ is Gaussian, with zero mean and variance 0.04; $\|\mathbf{S_v}\| = 0.24$. We set $b_0 = 5$, and a = 36.73. We estimate the detection error probability at each agent via 8,000 Monte Carlo runs per each hypothesis. Figure 1 plots the estimated detection error probability versus time steps k for each agent i in a semi-log scale. With each agent i, we can observe a straight line decay, which indicates an exponential decay of the error probability.

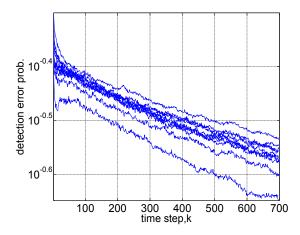


Fig. 1. Monte Carlo estimate of the detection error probability versus time step k for each agent i in the network.

B. Communication payoff

We now address our second question – whether communication (cooperation) through the consen-

sus+innovations detector pays off. We consider the identical agents case, with $m_i = m_j =: \overline{m} \neq 0$, for all $i \neq j$, and $\mathbf{S}_{\zeta} = \sigma_{\zeta}^2 \mathbf{I}$. The case of nonidentical agents is treated numerically in [9].

We first define formally the communication payoff. Consider the scenario where agent i does not communicate with other agents; rather, agent i performs detection based only on its own measurements $y_i(1), y_i(2), ...$ It is well known (see, e.g., [9]) that, under identical agents with $m_i \neq 0$, Gaussian sensing noise, and the local log-likelihood ratio detector, detection error probability $P_{i,\mathrm{isol}}^e(k)$ has the following exponential decay rate:

$$C_{i,\text{isol}} := \lim_{k \to \infty} -\frac{1}{k} \log P_{i,\text{isol}}^e(k) = \frac{\overline{m}^2}{8\sigma_{\ell}^2} > 0.$$

Denote, as before, by $C_{i,\text{coop}}$ the exponential decay rate at agent i with the distributed detector (7). We say that agent i achieves a communication payoff if:

$$C_{i,\text{coop}} > C_{i,\text{isol}}$$
.

We have the following Theorem. The Theorem can be derived from Corollary 4 in [9].

Theorem 2 (Communication payoff) Consider the consensus+innovations distributed detector in (7) under Assumptions 1–3, with the weights α_k as in (12). Further, suppose that the agents are identical, with the sensing signal $\overline{m} \neq 0$, and the sensing variance $\sigma_{\zeta}^2 \neq 0$. Suppose that:

$$\|\mathbf{S_v}\| \leq \frac{1}{8} \left(\frac{N-1}{2N}\right)^3 \frac{\overline{m}^2}{\sigma_{\zeta}^2} \left(\lambda_2(\mathcal{L})\right)^2.$$

Then, for b_0 set to

$$b_0^{\bullet} = \frac{1}{(\lambda_2(\mathcal{L}))^{1/3} 2^{1/3}} \left(\frac{\overline{m}^2}{\sigma_{\zeta}^2 \|\mathbf{S}_{\mathbf{v}}\|} \right)^{1/3},$$

each agent i achieves a communication payoff.

Theorem 2 says that there is a threshold on the communication noise power below which each agent is guaranteed to achieve a communication payoff. The Theorem is intuitive. For example, when the sensing quality is better $(\frac{\overline{m}^2}{\sigma_{\zeta}^2})$ is higher, the communication noise payoff threshold increases; the latter means that better sensing quality can sustain more communication noise.

V. CONCLUSION

We studied consensus+innovations distributed detection under noisy communication links. We addressed two fundamental questions. First, can we design the time-varying consensus weights, so that each agent i achieves an exponential decay of the detection error probability, even when almost all (except at least one)

agent(s) cannot detect the event of interest in isolation? Second, we asked when do we have a communication payoff, i.e., when is the worst agent under communication better than the best agent without communication? For the first question, we showed that each agent achieves a strictly positive exponential decay when the consensus weights are set to $b_0/(a+k)$, with b_0 larger than a certain threshold. For the second question, we quantified a communication noise power threshold below which agents achieve a communication payoff.

REFERENCES

- [1] F. S. Cattivelli and A. H. Sayed, "Distributed detection over adaptive networks based on diffusion estimation schemes," in Proc. IEEE SPAWC '09, 10th IEEE International Workshop on Signal Processing Advances in Wireless Communications, Perugia, Italy, June 2009, pp. 61–65.
- [2] —, "Diffusion LMS-based detection over adaptive networks," in *Proc. 39th Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, October 2009, pp. 171–175.
- [3] F. S. Cattivelli and A. H. Sayed, "Distributed detection over adaptive networks using diffusion adaptation," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 1917–1932, May 2011.
- [4] C. G. Lopes and A. H. Sayed, "Diffusion least-mean squares over adaptive networks: formulation and performance analysis," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3122–3136, July 2008.
- [5] S. Stankovic, N. Ilic, M. S. Stankovic, and K. H. Johansson, "Distributed change detection based on a consensus algorithm," *IEEE Transactions o Signal Processing*, vol. 59, no. 12, pp. 5686–5697, December 2011.
- [6] P. Braca, S. Marano, V. Matta, and P. Willet, "Asymptotic optimality of running consensus in testing binary hypothesis," *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 814–825, February 2010.
- [7] D. Bajovic, D. Jakovetic, J. Xavier, B. Sinopoli, and J. M. F. Moura, "Distributed detection via Gaussian running consensus: large deviations asymptotic analysis," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4381–4936, September 2011.
- [8] D. Bajovic, D. Jakovetic, J. M. F. Moura, J. Xavier, and B. Sinopoli, "Large deviations performance of consensus+innovations distributed detection with non-Gaussian observations," to appear in IEEE Transactions on Signal Processing, 2012, DOI: 10.1109/TSP.2012.2210885.
- [9] D. Jakovetic, J. M. F. Moura, and J. Xavier, "Distributed detection over noisy networks: Large deviations analysis," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4306–4320, August 2012.
- [10] S. Kar, R. Tandon, H. V. Poor, and S. Cui, "Distributed detection in noisy sensor networks," in *Proc. IEEE ISIT 2011, Interna*tional Symposium on Information Theory, Saint Petersburgh, Russia, August 2011.