Distributed Model-Invariant Detection of Unknown Inputs in Networked Systems

James Weimer, Damiano Varagnolo, Karl Henrik Johansson *†

ABSTRACT

This work considers hypothesis testing in networked systems under severe lack of prior knowledge. In previous work we derived a centralized Uniformly Most Powerful Invariant (UMPI) approach to testing unknown inputs in unknown Linear Time Invariant (LTI) networked dynamics subject to unknown Gaussian noise. The detector was also shown to have Constant False Alarm Rate (CFAR) properties. Nonetheless, in large-scale systems, centralized testing may be infeasible or undesireable. Thus, we develop a distributed testing version of our previous work that utilizes a statistic that is maximally invariant to the unknown parameters and the non-local/neighboring measurements. Similar to the centralized approach, the distributed test is shown to have CFAR properties and to have performance that asymptotically approaches that of the centralized test. Simulation results illustrate that the performance of the distributed approach suffers marginal performance degradation in comparison to the centralized approach. Insight to this phenomena is provided through a discussion.

Keywords

distributed hypothesis testing, invariant tests, linear systems, time invariant systems, networked systems

*All the authors are with the ACCESS Linnaeus Centre, School of Electrical Engineering, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden. Emails: { weimerj, damiano, kallej } @kth.se.

†The research leading to these results has received funding from the European Union Seventh Framework Programme [FP7/2007-2013] under grant agreement n°257462 HYCON2 Network of Excellence. The research was also supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation.

1. INTRODUCTION

Driven by the possibility of augmenting the flexibility and the reconfiguration capabilities of very complex systems, in many applications the current trend is to exploit multitudes of sensors and actuators, as in environmental monitoring [1], building energy management [2, 3], wireless communications [4] and power grids [5, 6]. The trend, however, comes with drawbacks: the high number of devices induces an increased possibility of faults with potentially disruptive ripple effects, like extended blackouts in power systems. There is thus a factual need for distributed fault detection algorithms.

We then consider that in every system, including dynamically networked ones such as the smart grid and building thermal dynamics, fault detection algorithms undoubtedly benefit from the knowledge of accurate models [6, 1, 3]. However, obtaining accurate models is often difficult or unrealistic due to the complexity of the system itself or the effects of environmental disturbances. For instance, in the smart grid security domain, it is common to assume the admittance of a transmission line is known [6]; however, the power line admittance is known to change with the temperature, humidity, and power flow, which leads to inaccurate models. Similarly, in building thermal dynamic modeling, even the simplest first-order heat equation model requires the knowledge of inter air-mass interactions, which change with the state of windows and doors (open or closed), the prevailing winds, the temperature, and the humidity. Thus, it is necessary to design fault detection schemes robust to these complex interactions.

If one were to consider large-scale networked systems, centralized approaches which apply model identification techniques in cascade with hypothesis testing may not be feasible. Similarly, when there are limited measurements, these identification and testing approaches tend to yield unexpected results, primarily due to the lack of information suitable for accurate parameter identification, see, e.g., [7, Example 1, page 46]. In this situation, distributed testing approaches that are designed to be invariant to the actual model parameters can result in better performance. In this paper we thus analyze if it is possible to derive distributed decision rules that do not depend on the model parameters and that are, in some sense to be defined, optimal with respect to the available information.

Literature review. Centralized classical hypothesis testing approaches usually exploit Generalized Likelihood Ratio (GLR) strategies, relying on obtaining Maximum Likelihood (ML) estimates of the unknown parameters under the various hypotheses and then testing their likelihood ratios. Maximally Invariant (MI) tests [8, Sec. 4.8] instead perform some additional preliminary operations so that the test is not influenced by the nuisance parameters. If MI tests are Uniformly Most Powerful Invariant (UMPI), then when the Signal to Noise Ratio (SNR) tends to infinity (e.g., when the number of measurements approaches infinity, see [9]), GLR and UMPI strategies are asymptotically equivalent. When small datasets are available, nonetheless, MI tests can outperform GLR approaches [10].

Invariant strategies have been used in several applications, like detection of structural changes in linear regression models [11] or in spectral properties of disturbances [12]. The literature focuses mainly on finding invariant methods in linear models with unknown or partially known covariance matrices [13, 14, 15, 16, 17], with efforts specially in finding tests that exploit maximally invariant statistics and that have Constant False Alarm Rate (CFAR) properties.

Recently, there has been substantial research in distributed GLR tests for networked systems, e.g., in environmental monitoring, smart grid fault detection, and building HVAC failure detection and diagnostics applications. While all these approaches yield asyptotically accurate results as the number of measurements increases, their performance under limited measurements is sporadic and unpredictable. This motivates the need for distributed testing techniques which have predictable performance regardless of the number of measurements.

In our previous work [18], we considered the centralized detection of unknown inputs in unknown dynamically networked Linear Time Invariant (LTI) Gaussian systems and developed a UMPI test with CFAR properties. This work not only showed the existence on a UMPI test, but also established an upper bound on the performance of any distributed detection scheme.

Statement of contributions. here we again focus on LTI-Gaussian models, but reduce the prior information to be the smallest possible. More precisely, we assume the knowledge of *just* the fact that the system dynamics is networked, LTI with Gaussian driving noises and, furthermore, a weak knowledge on the structure of the input fault. We thus develop a distributed CFAR test that is invariant to the unknown parameters and the non-local/neighboring measurements describing the system. The distributed test is then numerically evaluated against the centralized test developed in [18] as well as the best case (assuming a known model) and the worst case (assuming no model) scenarios, where it is shown empirically that the distributed test approaches the performance of the centralized UMPI test.

Structure of the paper. Section 2 reports the needed basic results and definitions on invariant hypothesis testing. Section 3 formulates precisely the problem considered. We propose our testing technique along with its statistical characterization in Section 4. Section 5 numerically compares the performance of the distributed detector against the performance of the centralized UMPI detector in [18] and strategies endowed with more prior information and no prior information for different operating points and systems. Finally, Section 6 reports some concluding remarks and proposes future extensions.

Notation. we use plain lower case italic fonts to indicate scalars or functions with scalar range, bold lower case italic fonts to indicate vectors or functions with vectorial range, and plain upper case italic fonts to indicate matrices. We also use $\otimes$ to denote Kronecker products, and $\boldsymbol{e}_{i,j}$ to denote the elementary vector of dimension i consisting of all zeros with a single unit entry in the j-th position.

2. HYPOTHESIS TESTING PRELIMINARIES

Commiserate with [8], we recall the definitions and methodology employed in designing UMPI tests. Let $\boldsymbol{y}$ be a r.v. with probability density $f(\boldsymbol{y}\,;\,\boldsymbol{d},\boldsymbol{\delta})$ parametrized in $\boldsymbol{d}$, $\boldsymbol{\delta}$. We define $\boldsymbol{d}$ to be the set of parameters of interest, and thus $\boldsymbol{\delta}$ to be the set of nuisance parameters, which induce a *transformation group* G, i.e., a set of endomorphisms g on the space of the realizations $\boldsymbol{y}$ [8, Sec. 4.8]. This group of transformations partitions the measurement space into equivalence classes (or orbits) where points are considered equal if there exist $g, g' \in G$ mapping the first into the second and vice versa.

Definition 1 (Maximally Invariant Statistic [8, Sec. 4.8]) A statistic $T[\boldsymbol{y}]$ is said to be maximally invariant w.r.t. a transformation group G if it is:

invariant: $T[g(\boldsymbol{y})] = T[\boldsymbol{y}], \quad \forall g \in G$

maximal: $T[\boldsymbol{y}'] = T[\boldsymbol{y}''] \Rightarrow \exists g \in G$ s.t. $\boldsymbol{y}'' = g(\boldsymbol{y}')$.

A statistical test, ϕ, based on an invariant statistic is said to be an invariant test:

Definition 2 (Invariant Test [8, Sec. 4.8]) Let G be a transformation group, $T[\boldsymbol{y}]$ a statistic and $\phi(\cdot)$ a hypothesis test. ϕ is said to be invariant w.r.t. G if

$$\phi\big(T[g(\boldsymbol{y})]\big) = \phi\big(T[\boldsymbol{y}]\big) \tag{1}$$

for every $g \in G$.

The statistical performance of an invariant test ϕ is measured in terms of its *size* and *power*, where an invariant test is desired to be Uniformly Most Powerful Invariant (UMPI):

Definition 3 (Uniformly Most Powerful Invariant (UMPI) Test [8, Sec. 4.8]) Let G be a transformation group, $T[\boldsymbol{y}]$ a statistic and $\phi(\cdot)$ a test for deciding between H_0 and H_1 that is invariant w.r.t. G. Then $\phi\big(T[\boldsymbol{y}]\big)$ is said to be an *uniformly most powerful invariant* (UMPI) test of size α if for every competing invariant test $\phi'\big(T[\boldsymbol{y}]\big)$ it holds that

$$\begin{aligned}
&\text{(size)} \sup_{\boldsymbol{d},\boldsymbol{\delta} \text{ under } H_0} \Pr\Big[\phi\big(T[\boldsymbol{y}]\big) = H_1 \;\Big|\; \boldsymbol{d},\boldsymbol{\delta}\Big] = \alpha; \\
&\sup_{\boldsymbol{d},\boldsymbol{\delta} \text{ under } H_0} \Pr\Big[\phi'\big(T[\boldsymbol{y}]\big) = H_1 \;\Big|\; \boldsymbol{d},\boldsymbol{\delta}\Big] \leq \alpha;
\end{aligned} \tag{2}$$

$$\begin{aligned}
&\text{(power)} \Pr\Big[\phi\big(T[\boldsymbol{y}]\big) = H_1 \;\Big|\; \boldsymbol{d},\boldsymbol{\delta} \text{ under } H_1\Big] \geq \\
&\Pr\Big[\phi'\big(T[\boldsymbol{y}]\big) = H_1 \;\Big|\; \boldsymbol{d},\boldsymbol{\delta} \text{ under } H_1\Big].
\end{aligned} \tag{3}$$

As a remark, thanks to the Karlin-Rubin theorem [8, Sec. 4.7, page 124], a scalar maximally invariant statistic whose likelihood ratio is monotone can be used to construct an UMPI test.

3. PROBLEM FORMULATION AND NOTATION

This section introduces a distributed hypothesis testing problem for deciding whether a signal, driven by unknown LTI networked Gaussian dynamics, lies also in a given subspace. Specifically, we consider a system of M interconnected nodes for which there exists an underlying interconnection graph, $\mathcal{G}(\mathcal{V},\mathcal{E})$, between the M nodes, where $\mathcal{V} := \{1,\ldots,M\}$ is the vertex set, with $i \in \mathcal{V}$ corresponding to node i, and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the edge set of the graph. The undirected edge $\{i,j\}$ is incident on vertices i and j if nodes i and j share an interconnection, such that the neighborhood of node i, $\mathcal{N}_i$, is defined as

$$\mathcal{N}_i := \big\{ j \in \mathcal{V} \;\big|\; \{i,j\} \in \mathcal{E} \big\} \tag{4}$$

The inter-node dynamics are governed by discrete-time LTI-Gaussian dynamics

$$\begin{aligned}
x_j(k+1) &= x_j(k) + m_j \sum_{i \in \mathcal{N}_j} a_{ji}\Big(x_i(k) - x_j(k)\Big) \\
&\quad + b_j d_j(k) + w_j(k) \\
y_j(k) &= x_j(k) + v_j(k)
\end{aligned} \tag{5}$$

where:

- $k = 0,\ldots,T$ is the time index (T even for notational simplicity[1]);
- $j = 1,\ldots,M$ is the agent index;
- the states $x_j(k)$'s, measurements $y_j(k)$'s and inputs $d_j(k)$'s are scalar;
- $m_j a_{ji} = m_j a_{ij} \in \mathbb{R}$ and $b_j \in \mathbb{R}$ denote respectively the gains between $x_i(k)$ and $x_j(k{+}1)$, and between $d_j(k)$ and $x_j(k+1)$;
- $w_j(k), v_j(k) \in \mathbb{R}$ are uncorrelated i.i.d. Gaussian process noise and measurement noise with moments

$$\mathbb{E}\left[w_j(k)\right] = \chi_{j,w} \quad \mathbb{E}\left[v_j(k)\right] = \chi_{j,v},$$

$$\mathbb{E}\left[\Big(w_j(k) - \overline{w}_j\Big)^2\right] = \sigma^2_{j,w} \quad \mathbb{E}\left[\Big(v_j(k) - \overline{v}_j\Big)^2\right] = \sigma^2_{j,v}.$$

To compact the notation we let, for $j = 1,\ldots,M$,

$$\begin{aligned}
A &:= \begin{bmatrix} \alpha_{ij} \end{bmatrix} \\
\alpha_{ij} &:= \begin{cases} 1 - m_j \displaystyle\sum_{n \in \mathcal{N}_j} a_{nj} & \text{if } i = j \\ m_j a_{ij} & \text{if } i \in \mathcal{N}_j, \quad i \neq j \\ 0 & \text{otherwise} \end{cases} \\
B &:= \operatorname{diag}\left[b_1,\ldots,b_M\right] \\
\boldsymbol{y}_j &:= \left[y_j(0),\ldots,y_j(T)\right]^\top \\
\boldsymbol{d}_j &:= \left[d_j(0),\ldots,d_j(T)\right]^\top.
\end{aligned}$$

Additionally, we consider the following quantities: let $\mathcal{N}_j = \{i_1,\ldots,i_J\}$ be the sorted list of neighbors of agent j. Then

$$\begin{aligned}
\vec{\alpha}_j &:= \left[\alpha_{i_1 j},\ldots,\alpha_{i_J j}\right]^\top \\
\vec{y}_j(k) &:= \left[y_{i_1}(k),\ldots,y_{i_J}(k)\right]^\top \\
\vec{\boldsymbol{y}}_j &:= \left[\boldsymbol{y}^T_{i_1},\ldots,\boldsymbol{y}^T_{i_J}\right]^\top,
\end{aligned}$$

[1] For ease of notation and without loss of generality we assume that the available measurements are over a given period whose length is fixed *ex ante*.

i.e., $\vec{y}_j(k)$ is the set of the measurements of agent j *and* its neighbors (sorted lexicographically) at time k, while $\vec{\boldsymbol{y}}_j$ is the set of *all* the measurements of agent j *and* its neighbors (again sorted lexicographically).

Consider then a *specific* agent $\ell \in \{1, \ldots, M\}$. The structure of the input $\boldsymbol{d}_\ell$ is assumed to be as follows:

- $\boldsymbol{u}_\ell := \left[u_\ell(0), \ldots, u_\ell(T)\right]^\top$ is a *desired* and *known* input signal;
- $\boldsymbol{s}_\ell^f := \left[s_\ell^f(0), \ldots, s_\ell^f(T)\right]^\top$, $f = 1, \ldots, N_\ell$ are some *known* signals defining the space of signals
$$\operatorname{span}\left\langle \boldsymbol{s}_\ell^1, \ldots, \boldsymbol{s}_\ell^{N_\ell} \right\rangle$$
(with $S_\ell := \left[\boldsymbol{s}_\ell^1, \ldots, \boldsymbol{s}_\ell^{N_\ell}\right]$ being a shorthand for the $\boldsymbol{s}_\ell^f$'s);
- $\boldsymbol{\theta}_\ell \in \mathbb{R}^{N_\ell}$ is an unknown (but constant) signal selection parameter.

Then

$$\boldsymbol{d}_\ell = S_\ell \boldsymbol{\theta}_\ell + \mu_\ell \boldsymbol{u}_\ell \tag{6}$$

where the scalar μ_ℓ is an unknown parameter.

Summarizing, the information owned by agent ℓ is either *available* or *unavailable* as follows:

Assumption 4 Available information:

- the time-series measurements $\vec{\boldsymbol{y}}_\ell$
- the local desired input signal $\boldsymbol{u}_\ell$;
- the local nuisance subspace S_ℓ;
- the local weight m_ℓ;
- the fact that the state dynamics are LTI-Gaussian, constant in time, and with $b_\ell \neq 0$.

Assumption 5 Unavailable information:

- all the time-series measurements but $\vec{\boldsymbol{y}}_j$;
- all the local desired input signals but $\boldsymbol{u}_\ell$;
- all the local nuisance subspaces but S_ℓ;
- all the local weights but m_ℓ;
- the weights A and B;
- the moments of the process and measurement noises $\chi_{j,w}$, $\chi_{j,v}$, $\sigma_{j,w}^2$, $\sigma_{j,v}^2$, $j = 1, \ldots, M$;
- the parameters $\boldsymbol{\theta}_j$ and μ_j;
- the initial conditions $x_1(0), \ldots, x_M(0)$;
- the input signals $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_M$.

We then assume the unknown μ_ℓ to be either 0 or 1 and pose the following binary hypothesis testing problem:

Assumption 6 Structure of the fault μ_ℓ satisfies either one of the two following hypotheses:

H_0 **(null hypothesis):** $\mu_\ell = 0$

H_1 **(alternative hypothesis):** $\mu_\ell = 1$

In words, both hypotheses assume the actual $\boldsymbol{d}_\ell$ to be unknown, since $\boldsymbol{\theta}_\ell$ is unknown, but with a fixed and known functional structure. H_1 additionally assumes the presence of a known input $\boldsymbol{u}_\ell$.

Our aim is thus: develop a distributed test that considers a ***specific*** *agent $\ell \in \{1, \ldots, M\}$, and decides among the hypotheses H_0 vs. H_1 in Assumption 6 using only the information in Assumption 4 and, at the same time, being invariant to the unavailable information in Assumption 5.*

We note that the problem formulated in this section is fundamentally different from the problem formulated in [18]. Indeed, the novel test should be computable distributedly *and* should be invariant also to the non-local measurements (in addition to all the unavailable information in [18]).

We thus aim to find a test that detects whether node ℓ has a fault independently of whether a fault exists at any other node $j \neq \ell$ (fault isolation) *and* maximizes the probability of detection (power) for any probability of false alarm (size), i.e., we require the detector to be UMPI. Formally, thus, we aim to solve the following:

Problem 7

1. find a statistic $T[\vec{\boldsymbol{y}}_\ell]$ that satisfies Definition 1 (maximal invariance) w.r.t. the transformation group induced by nuisance parameters in Assumption 5;

2. find a test $\phi\Big(T[\vec{\boldsymbol{y}}_\ell]\Big)$ that satisfies Definition 3 (UMPI test) w.r.t. to the class of tests based on the previously introduced maximal invariant statistic $T[\vec{\boldsymbol{y}}_\ell]$.

4. DISTRIBUTED INVARIANT TESTING

In this section we solve the previously posed problem and develop a distributed UMPI test that uses only local and neighboring measurements. The algorithm is based on the following novel result, solving the first part of Problem 7:

Theorem 8 A maximally invariant statistic that solves Problem 7-1 is

$$T[\boldsymbol{z}_\ell] = \frac{\boldsymbol{z}_\ell^\top P_\ell \boldsymbol{z}_\ell}{\frac{1}{N_\ell - 1}\boldsymbol{z}_\ell^\top \big(I_{N_\ell} - P_\ell\big)\boldsymbol{z}_\ell} \tag{7}$$

with

$$\begin{aligned} \boldsymbol{z}_\ell &:= F_\ell Q \boldsymbol{y}_\ell \\ P_\ell &:= \frac{F_\ell Q \boldsymbol{u}_\ell \boldsymbol{u}_\ell^\top Q^\top F_\ell^\top}{\boldsymbol{u}_\ell^\top Q^\top F_\ell^\top F_\ell Q \boldsymbol{u}_\ell} \\ N_\ell &:= \frac{k}{2} - \|\mathcal{N}_\ell\|_0 \end{aligned} \tag{8}$$

and where the exploited quantities satisfy

$$\begin{aligned} F_\ell^\top F_\ell &= I_{\frac{k}{2}} - \vec{Y}_\ell (\vec{Y}_\ell^\top \vec{Y}_\ell)^{-1} \vec{Y}_\ell^\top \\ Q &= I_{\frac{k}{2}} \otimes [\,0 \quad 1\,] \\ \vec{Y}_\ell &= \begin{bmatrix} \vec{y}_\ell^\top(0) & (s_\ell^f(0))^\top & 1 \\ \vec{y}_\ell^\top(2) & (s_\ell^f(2))^\top & 1 \\ \vec{y}_\ell^\top(4) & (s_\ell^f(4))^\top & 1 \\ \vdots & \vdots & \vdots \\ \vec{y}_\ell^\top(T) & (s_\ell^f(T))^\top & 1 \end{bmatrix} \end{aligned} \tag{9}$$

Proof. The proof follows a similar flow to the centralized test proof in [18]. The proof is omitted due to space constraints in this extended abstract. If accepted, the final version of this work will include a proof for Theorem 8.

We observe that the maximally invariant statistic in (7) can be equivalently written as a ratio of independent chi-square random variables. This particular ratio is known to follow an F-distribution, which has a monotone likelihood ratio [8]. Thus we solve the second part of Problem 7 by applying the Karlin-Rubin theorem, obtaining directly the following:

Corollary 9 A distributed UMPI test of size α for Problem 7-2 is

$$\phi_\ell(\boldsymbol{z}_\ell) = \begin{cases} H_0 & \text{if } \; T_\ell[\boldsymbol{z}_\ell] < \mathcal{F}^{-1}_{1,N_\ell-1}(\alpha) \\ H_1 & \text{otherwise.} \end{cases} \tag{10}$$

where $\mathcal{F}^{-1}_{n,m}(\alpha)$ is the inverse central cumulative F-distribution of dimensions n and m.

We remark that, w.r.t. the algorithm proposed in [18], test (10) can be performed in parallel and it is invariant to the non-local measurements. This comes with a price: the test exploits only about half of the available measurements (either local or from neighbors). The remaining local and neighbors' measurements are in fact lost in the attempt of obtaining invariance. Since the dataset is smaller than the one exploited in [18], it is expected that the novel test will perform worse. In the following section we then numerically evaluate this loss.

5. NUMERICAL EXAMPLES

We perform three Monte-Carlo characterizations as follows:

1. we fix a desired probability of false alarms α (0.01, 0.1 and 0.25);

2. we randomly generate 500 stable networked systems of 10 agents like (5) as described in Table 1 (i.e., we discarded the unstable realizations);

3. for each of the 500 systems (5) we generated exactly one realization $y_j(1), \ldots, y_j(500)$, $j = 1, \ldots, 10$;

4. for each $T = 1, \ldots, 500$ and each of the 500 systems (5) we executed the following four tests, all with the same desired probability of false alarms α:

   (a) full information test: assume the perfect knowledge of the weights A and B; the moments of the process and measurement noises $\chi_{j,w}$, $\chi_{j,v}$, $\sigma^2_{j,w}$, $\sigma^2_{j,v}$; the parameters $\boldsymbol{\theta}_j$; the initial conditions $x_1(j)$ ($j = 1, \ldots, 10$). Then design the Uniformly Most Powerful (UMP) test for testing H_0 vs. H_1 given all this information;

$a_j, b_j \sim \mathcal{U}[-0.5, 0.5]$	$m_j \sim \mathcal{U}[1, 2]$
$\chi_{j,w}, \chi_{j,v} \sim \mathcal{N}(0, 1)$	$\sigma^2_{j,w}, \sigma^2_{j,v} \sim \mathcal{U}[0.1, 1]$

Table 1: Random extraction mechanisms for the generation of the systems (5). $\mathcal{N}$ indicates Gaussian distributions, $\mathcal{U}$ uniform distributions. All the quantities are extracted independently.

(b) centralized UMPI test: the UMPI test developed in [18], which is provided in the appendix using the notation introduced within this work;

(c) distributed UMPI test (DUMPI): our test (10);

(d) no information test: perform a weighted coin flip s.t. the desired probability of false alarms α is met.

The outcomes are then summarized in the following Figures 1, 2 and 3, that plot for each test and each T the average correct detection rate reached over the 500 considered realizations of system 5.

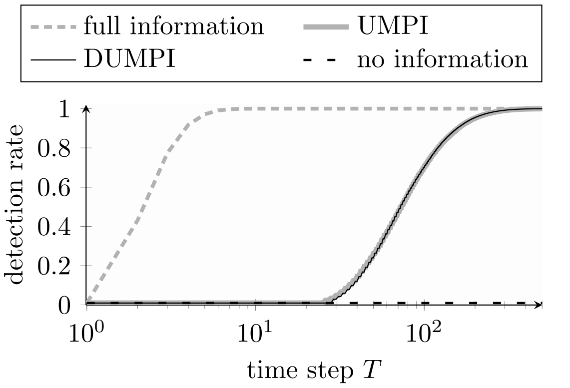

Figure 1: Monte-Carlo characterization of the detection tests given $\alpha = 0.01$.

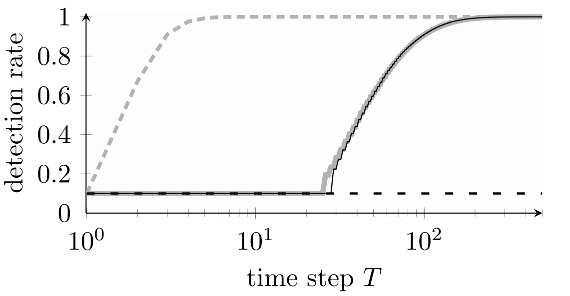

Figure 2: Monte-Carlo characterization of the detection tests given $\alpha = 0.1$. Legend as in Figure 1.

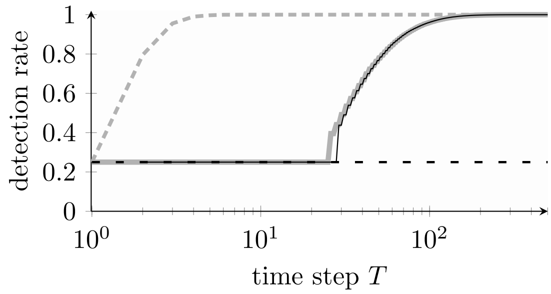

Figure 3: Monte-Carlo characterization of the detection tests given $\alpha = 0.25$. Legend as in Figure 1.

From the previous graphics we draw the following conclusions. Before the number of measurements (proportional to T) passes the threshold $\frac{T}{2} - N_\ell - M + 1$ (independent of the chosen α), both the centralized and distributed UMPI tests are equivalent to a coin flipping (since the amount of information is insufficient to take meaningful decisions). After that threshold, instead, the two test start increasing their correct detection rates (with different speeds, depending on the selected probability of false alarms), discerning better and better. Eventually they reach the same performance of the full information-based test, i.e., the best one might desire. We then notice that the difference in the correct detection rates between the centralized and distributed approaches starts small and vanishes quickly. This indicates that, from practical purposes, the distributed strategy performs well. The reason for such a similar performance between the centralized and distributed approaches lies in that the centralized appraoch from [18] (also provided in the appendix of this extended abstract), effectively disregards half of the measurements to achieve maximal invariance. In the distributed approach, the same measurements that are discarded by the centralized approach are employed to provide invariance to the local inter-node dynamics.

6. DISCUSSION AND FUTURE WORKS

We considered fault detection in networked Linear Time Invariant-Gaussian systems. More precisely, we defined a hypothesis testing problem over the structure of the inputs of the agents, and then derived a distributed Uniformly Most Powerful Invariant detector with Constant False Alarm Rate properties that is invariant to most of the parameters of the systems. We address the situation where there is little prior information available, and develop a distributed test starting from our previous centralized results described in [18]. 'Remarkably we obtain a distributed algorithm that has some capability of detecting faults even if knowledge of the overall system is really uncertain and the number of measurements is limited.

As in the centralized case, tests that exploit information of the system have better performance in terms of false positives / negatives rates. Nonetheless, the more measurements that are taken the more the distributed detector is shown to be perform better, achieving performance of its centralized counterpart quickly.

The value of the proposed strategy relies in its optimality properties, being in fact based on a maximally invariant statistic and being uniformly most powerful. This implies that in a certain sense it characterizes the performance that can be achieved when testing the posed hypotheses under the severe lack of knowledge assumed here.

The main future direction is thus to compare the developed strategy, both from practical and theoretical aspects, with the distributed fault detection algorithm that are based on dynamically identified systems. It is in fact necessary to understand if there are conditions s.t. the invariant test developed here is guaranteed to perform better than algorithms that start identifying the test and then perform tests on the identified model.

7. REFERENCES

[1] J. Weimer, B. Sinopoli, and B. Krogh, "An approach to leak detection using wireless sensor networks at carbon sequestration sites," *International Journal of Greenhouse Gas Control*, vol. 9, pp. 243–253, 2012.

[2] F. Oldewurtel, A. Parisio, C. N. Jones, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and M. Morari, "Use of model predictive control and weather forecasts for energy efficient building climate control," *Energy and Buildings*, vol. 45, no. 0, pp. 15 – 27, 2012.

[3] J. Weimer, S.Ahmadi, J. Araujo, F. Mele, D. Papale, I. Shames, H. Sandberg, and K. Johansson, "Active actuator fault detection and diagnostics in hvac systems," in *ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings (BuildSys)*, 2012.

[4] T. Arampatzis, J. Lygeros, and S. Manesis, "A survey of applications of wireless sensors and wireless sensor networks," in *IEEE International Symposium on Intelligent Control, MCCA*, 2005, pp. 719–724.

[5] S. Bolognani, A. Carron, A. Di Vittorio, D. Romeres, L. Schenato, and S. Zampieri, "Distributed multi-hop reactive power compensation in smart micro-grids subject to saturation constraints," in *51st IEEE Conference on Decision and Control*, 2012, pp. 785–790.

[6] I. Shames, A. Teixeira, H. Sandberg, and K. Johansson, "Distributed fault detection for interconnected systems," *Automatica*, vol. 47, no. 12, pp. 2757–2764, 2011.

[7] V. N. Vapnik, *Statistical learning theory.* New York: Wiley, 1998.

[8] L. L. Scharf, *Statistical Signal Processing, Detection, Estimation, and Time Series Analysis.* Addison-Welsley Publishing Company Inc., Reading, Massachusetts, 1991.

[9] J. Gabriel and S. Kay, "On the relationship between the glrt and umpi tests for the detection of signals with unknown parameters," *Signal Processing, IEEE Transactions on*, vol. 53, no. 11, pp. 4194–4203, 2005.

[10] H. Kim and A. Hero III, "Comparison of glr and invariant detectors under structured clutter covariance," *Image Processing, IEEE Transactions on*, vol. 10, no. 10, pp. 1509–1520, 2001.

[11] M. Hušková, "Some invariant test procedures for detection of structural changes," *Kybernetika*, vol. 36, no. 4, pp. 401–414, 2000.

[12] N. Begum and M. L. King, "Most mean powerful invariant test for testing two-dimensional parameter spaces," *Journal of Statistical Planning and Inference*, vol. 134, no. 2, pp. 536 – 548, 2005.

[13] S. Bose and A. Steinhardt, "A maximal invariant framework for adaptive detection with structured and unstructured covariance matrices," *Signal Processing, IEEE Transactions on*, vol. 43, no. 9, pp. 2164–2175, 1995.

[14] K. Noda and H. Ono, "On ump invariant f-test procedures in a general linear model," *Communications in Statistics-Theory and Methods*, vol. 30, no. 10, pp. 2099–2115, 2001.

[15] E. Conte, A. De Maio, and C. Galdi, "Cfar detection of multidimensional signals: an invariant approach," *Signal Processing, IEEE Transactions on*, vol. 51, no. 1, pp. 142–151, 2003.

[16] A. De Maio, "Rao test for adaptive detection in gaussian interference with unknown covariance matrix," *Signal Processing, IEEE Transactions on*, vol. 55, no. 7, pp. 3577–3584, 2007.

[17] A. De Maio and E. Conte, "Adaptive detection in gaussian interference with unknown covariance after reduction by invariance," *Signal Processing, IEEE Transactions on*, vol. 58, no. 6, pp. 2925–2934, 2010.

[18] J. Weimer, D. Varagnolo, M. Stankovic, and K. Johansson, "Model-invariant detection of unknown inputs in networked systems," in *European Control Conference (under review)*, 2013.

Appendix

This appendix provides a proof for 8. We begin by writing the measurement dynamics in 5 as

$$y_j(k+1) = x_j(k) + m_j \sum_{i\in\mathcal{N}_j} a_{ji}\Big(y_i(k) - y_j(k)\Big) + b_j d_j(k) + n_j(k) \tag{11}$$

where

$$n_j(k) = w_j(k) + v_j(k+1) - m_j \sum_{i\in\mathcal{N}_j} a_{ji}\left(v_i(k) - v_j(k)\right) \tag{12}$$

$$n_j(k) = w_j(k) + v_j(k+1) - \left(1 - m_j \sum_{i\in\mathcal{N}_j} a_{ji}\right) v_j(k) - m_j \sum_{i\in\mathcal{N}_j} a_{ji} v_i(k). \tag{13}$$

Since the noise correlation is unknown, we whiten the measurements by using only every other measurement and write the resulting time-series measurements as

$$\boldsymbol{y}_\ell = \vec{Z}_\ell \boldsymbol{\theta} + b_j \mu_j \boldsymbol{u}_\ell + \boldsymbol{n}_\ell \tag{14}$$

where

$$\mathrm{Cov}\left[\boldsymbol{n}_\ell\right] = \sigma_0^2 I + \sigma_1^2 \sum_{i=0}^{\frac{T}{2}} \left(\boldsymbol{e}_{2i}\boldsymbol{e}_{2i+1}^\top + \boldsymbol{e}_{2i+1}\boldsymbol{e}_{2i}^\top\right) \tag{15}$$

The unknown parameters induce a group of transformation on the measurements,

Since at the time of submission of this extended abstract the previous work in [18] is under review, this appendix provides a centralized maximally invariant statistic for detection of unknown inputs in LTI-Gaussian networked systems.

Specifically, the maximally invariant statistic is

$$T_c[\boldsymbol{r}_\ell] = \frac{\boldsymbol{r}_\ell^\top R_\ell \boldsymbol{r}_\ell}{\frac{1}{N_\ell^c - 1}\boldsymbol{r}_\ell^\top \left(I - R_\ell\right)\boldsymbol{r}_\ell} \tag{16}$$

with

$$\begin{aligned} \boldsymbol{z}_\ell &:= H_\ell G E D \boldsymbol{y} \\ R_\ell &:= \frac{H_\ell G E \boldsymbol{u}_\ell \boldsymbol{u}_\ell^\top E^\top G^\top H_\ell^\top}{\boldsymbol{u}_\ell^\top E^\top G^\top H_\ell^\top H_\ell G E \boldsymbol{u}_\ell} \\ N_\ell^c &:= \frac{T}{2} - N_\ell - M + 1 \end{aligned} \tag{17}$$

and where the exploited quantities satisfy

$$\begin{aligned} \boldsymbol{y} &= [y_1(0), \ldots, y_M(0), y_1(1), \ldots, y_M(1), \ldots, y_1(T), \ldots, y_M(\\ H_\ell^\top H_\ell &= I - \left[\, \boldsymbol{u}_1 \ \ldots \ \boldsymbol{u}_{\ell-1} \ \boldsymbol{u}_{\ell+1} \ \ldots \ \boldsymbol{u}_M \ S_1 \ \ldots \ S_M \ \mathbf{1} \right] \\ G &= I \otimes [\, 0 \quad 1 \,] \\ \boldsymbol{p} &= \left[\frac{1}{m_1}, \ldots, \frac{1}{m_M}\right]^\top \cdot \left(\sum_{j=1}^{M} m_j^{-2}\right)^{-\frac{1}{2}} \\ E &= \begin{bmatrix} \boldsymbol{p}^\top & & \\ & \ddots & \\ & & \boldsymbol{p}^\top \end{bmatrix} \\ D &= \begin{bmatrix} I & & & \\ -I & I & & \\ & \ddots & \ddots & \\ & & -I & I \end{bmatrix} \end{aligned} \tag{18}$$

A UMPI test of size α for the centralized detector is

$$\phi_\ell(\boldsymbol{z}_\ell) = \begin{cases} H_{\ell,0} & \text{if } \ T_\ell[\boldsymbol{z}_\ell] < \mathcal{F}^{-1}_{1,N_\ell^c-1}(\alpha) \\ H_{\ell,1} & \text{otherwise.} \end{cases} \tag{19}$$

7.1 Supporting Lemmas

This subsection sequentially introduces lemmas to:

1. obtain composed maximally invariant statistics;
2. obtain maximal invariance w.r.t. an unknown subspace bias;
3. obtain maximal invariance w.r.t. an unknown correlated noise;
4. obtain maximal invariance w.r.t. an unknown subspace gain;
5. obtain maximal invariance w.r.t. an unknown measurement scaling.

In this subsection we use the notation $\boldsymbol{r}$ to denote a generic measurement vector or a linear combination of measurements. Additionally, in each of the following Lemmas we re-use the same variables names to denote different objects, in order to lessen the notational overhead. Each lemma, thus, is an independent statement.

7.1.1 Composed maximally invariant statistic

If $\boldsymbol{\delta}$ in Section 2 is composed by several nuisance parameters it is then convenient to obtain a statistic, $T[\cdot]$, invariant to $\boldsymbol{\delta}$ from the composition of other invariant statistics, say $T_1[\cdot], T_2[\cdot], \ldots$, where each statistic is invariant to some of the nuisance parameters in $\boldsymbol{\delta}$. The following lemma states some sufficient conditions for a composition of statistics to be maximally invariant to $\boldsymbol{\delta}$:

Lemma 10 (Composed Maximally Invariant Statistics) Let

- $\boldsymbol{\delta}_0$, $\boldsymbol{\delta}_1$ be two set of nuisance parameters;
- G_0, G_1 be two group of transformations, respectively induced by the nuisance parameters $\boldsymbol{\delta}_0$, $\boldsymbol{\delta}_1$;
- $T_0[\boldsymbol{r}] = Q_0\boldsymbol{r}$ be a statistic that is maximally invariant w.r.t. the transformation group G_0;

The statistic $T[\boldsymbol{r}] = T_1\left[T_0[\boldsymbol{r}]\right]$ is maximally invariant w.r.t. the transformation group G_1G_0 if

- $Q_0Q_0^\top = I$ (i.e., Q_0 is unitary);
- $T[\boldsymbol{r}]$ is maximally invariant w.r.t. the group
$$\widehat{G} := \left\{ g \mid g(\boldsymbol{r}) = Q_0^\top Q_0 g_1(\boldsymbol{r}), \quad g_1 \in G_1 \right\}.$$

Proof. Invariance:

$$\begin{aligned} g_0 \in G_0 :\ & T\left[g_0(\boldsymbol{r})\right] = T_1\Big[T_0\left[g_0(\boldsymbol{r})\right]\Big] = T_1\left[T_0[\boldsymbol{r}]\right] = T[\boldsymbol{r}] \\ g_1 \in G_1 :\ & T\left[g_1(\boldsymbol{r})\right] = T_1\Big[T_0\left[g_1(\boldsymbol{r})\right]\Big] = T_1\left[Q_0 g_1(\boldsymbol{r})\right] \\ & = T_1\Big[T_0\left[\widehat{g}(\boldsymbol{r})\right]\Big], \quad \exists \widehat{g} \in \widehat{G} \\ & = T[\boldsymbol{r}]. \end{aligned} \tag{20}$$

Maximality:

$$\begin{aligned} & T\left[\widehat{\boldsymbol{r}}\right] = T\left[\boldsymbol{r}\right] \longrightarrow T\left[\widehat{\boldsymbol{r}}\right] = T_1\left[Q_0\boldsymbol{r}\right] \\ & \longrightarrow T\left[\widehat{\boldsymbol{r}}\right] = T_1\left[Q_0Q_0^\top Q_0\boldsymbol{r}\right] \\ & \longrightarrow T\left[\widehat{\boldsymbol{r}}\right] = T\left[Q_0^\top Q_0\boldsymbol{r}\right] \\ & \longrightarrow \widehat{\boldsymbol{r}} = g_1(\boldsymbol{r}), \quad \exists g_1 \in G_1. \end{aligned} \tag{21}$$

With a similar logic we can derive the following corollary, that can be applied to statistics resulting from invertible transformations:

Corollary 11 (Maximality of Invertible Statistics) With the same premises as in Lemma 10, if $T_0[\boldsymbol{r}]$ is maximally invariant w.r.t. the group G_0 and Q_1 is invertible, then the composed statistic $T[\boldsymbol{r}] = T_1\left[T_0[\boldsymbol{r}]\right] = Q_1T_0[\boldsymbol{r}]$ is maximally invariant w.r.t. G_0.

7.1.2 Maximal invariance w.r.t. a subspace bias

Consider measurements generated according to

$$\boldsymbol{r} = \widehat{\boldsymbol{s}} + H\boldsymbol{\theta} \tag{22}$$

where $\boldsymbol{\theta}$ is a vector of nuisance parameters, H is a known subspace of appropriate dimension, and $\widehat{\boldsymbol{s}}$ is an arbitrary signal. The nuisance parameter $\boldsymbol{\theta}$ induces the group of transformations

$$G = \left\{ g \;\middle|\; g(\boldsymbol{r}) = \boldsymbol{r} + H\boldsymbol{\theta} \right\}. \tag{23}$$

It then follows that:

Lemma 12 (Maximal invariance w.r.t. a subspace bias) Let Q be s.t.

$$Q^\top Q = I - H\left(H^\top H\right)^{-1}H^\top, \qquad QQ^\top = I. \tag{24}$$

Then the statistic $T[\boldsymbol{r}] = Q\boldsymbol{r}$ is maximally invariant w.r.t. G.

Proof. Invariance:

$$T\left[g(\boldsymbol{r})\right] = Q\left(\boldsymbol{r} + H\boldsymbol{\theta}\right) = Q\boldsymbol{r} = T[\boldsymbol{r}]. \tag{25}$$

Maximality:

$$\begin{aligned} & T\left[\widehat{\boldsymbol{r}}\right] = T[\boldsymbol{r}] \longrightarrow Q\widehat{\boldsymbol{r}} = Q\boldsymbol{r} \\ & \longrightarrow \widehat{\boldsymbol{r}} = \boldsymbol{r} + \left(Q^\top Q - I\right)\left(\boldsymbol{r} - \widehat{\boldsymbol{r}}\right) \\ & \longrightarrow \widehat{\boldsymbol{r}} = g(\boldsymbol{r}), \quad \exists g \in G. \end{aligned} \tag{26}$$

7.1.3 Maximal invariance w.r.t. a correlated noise

Consider measurements generated according to

$$\boldsymbol{r} = \widehat{\boldsymbol{s}} + \boldsymbol{n} \tag{27}$$

where $\boldsymbol{n}$ is a vector of Gaussian random variables with covariance

$$\mathrm{Cov}\left[\boldsymbol{n}\right] = \sigma_0^2 I + \sigma_1^2\left(\boldsymbol{e}_j\boldsymbol{t}_j^\top + \boldsymbol{t}_j\boldsymbol{e}_j^\top\right) \tag{28}$$

where $\sigma_0, \sigma_1 \in \mathbb{R}_{++}$ are unknown, $\boldsymbol{t}_j$ is an arbitrary vector of appropriate dimension, and $\boldsymbol{e}_j$ is the elementary vector with a single unit entry in the j-th element. The correlation induces the group of transformations

$$G = \left\{ g \;\middle|\; g(\boldsymbol{r}) = \left(I + \boldsymbol{e}_j\boldsymbol{t}_j^\top\right)\boldsymbol{r} \right\}. \tag{29}$$

It then follows that:

Lemma 13 (Maximal invariance w.r.t. a correlated noise) Let Q be s.t.

$$Q^\top Q = I - \boldsymbol{e}_j\boldsymbol{e}_j^\top, \qquad QQ^\top = I. \tag{30}$$

Then the statistic $T[\boldsymbol{r}] = Q\boldsymbol{r}$ is maximally invariant w.r.t. G.

Proof. Invariance:

$$T\left[g(\boldsymbol{r})\right] = Q\left(I + \boldsymbol{e}_j\boldsymbol{t}_j^\top\right)\boldsymbol{r} = Q\boldsymbol{r} = T[\boldsymbol{r}]. \tag{31}$$

Maximality:

$$\begin{aligned} T\big[\widehat{\boldsymbol r}\big] = T\big[\boldsymbol r\big] &\longrightarrow Q\widehat{\boldsymbol r} = Q\boldsymbol r \\ &\longrightarrow \widehat{\boldsymbol r} = \boldsymbol r + \Big(Q^\top Q - I\Big)\Big(\boldsymbol r - \widehat{\boldsymbol r}\Big) \\ &\longrightarrow \widehat{\boldsymbol r} = g(\boldsymbol r), \quad \exists g \in G. \end{aligned} \tag{32}$$

7.1.4 Maximal invariance w.r.t. an unknown subspace gain

Consider measurements generated according to

$$\boldsymbol r = \big(I + H\big)\widehat{\boldsymbol s} \tag{33}$$

where $\widehat{\boldsymbol s}$ is an arbitrary signal, H is an unknown subspace of appropriate dimension with a known left eigenvector, $\boldsymbol v^\top$, corresponding to the unique zero eigenvalue of H. The nuisance parameter H induces the group of transformations

$$G = \Big\{ g \;\Big|\; g(\boldsymbol r) = \big(I + H\big)\boldsymbol r,\ \boldsymbol v^\top H = 0,\ \boldsymbol v^\top \boldsymbol v = 1 \Big\}. \tag{34}$$

It then follows that:

Lemma 14 (Maximal invariance w.r.t. an unknown subspace gain) The statistic

$$T\left[\boldsymbol r\right] = \boldsymbol v^\top \boldsymbol r \tag{35}$$

is maximally invariant to G.

PROOF. Invariance:

$$T\big[g(\boldsymbol r)\big] = \boldsymbol v^\top\big(\boldsymbol r + H\boldsymbol r\big) = \boldsymbol v^\top \boldsymbol r = T\left[\boldsymbol r\right]. \tag{36}$$

Maximality:

$$\begin{aligned} T\big[\widehat{\boldsymbol r}\big] = T\big[\boldsymbol r\big] &\longrightarrow \boldsymbol v^\top\widehat{\boldsymbol r} = \boldsymbol v^\top\boldsymbol r \\ &\longrightarrow \widehat{\boldsymbol r} = \boldsymbol r + \Big(I - \boldsymbol v\boldsymbol v^\top\Big)\Big(\boldsymbol r - \widehat{\boldsymbol r}\Big) \\ &\longrightarrow \widehat{\boldsymbol r} = g(\boldsymbol r), \quad \exists g \in G. \end{aligned} \tag{37}$$

7.1.5 Maximal invariance w.r.t. a measurement scaling

Consider measurements generated according to

$$\boldsymbol r = \sigma\big(\mu\boldsymbol s + \boldsymbol n\big) \tag{38}$$

where $\boldsymbol n$ is a vector of zero-mean white Gaussian r.v.s., $\boldsymbol s$ is a known signal, and $\mu, \sigma \in \mathbb{R}$ are unknown constants. The nuisance parameter σ induces the group of transformations

$$G = \Big\{ g \;\Big|\; g(\boldsymbol r) = c\boldsymbol r, \quad c \in \mathbb{R} \Big\}. \tag{39}$$

It then follows that:

Lemma 15 (Maximal invariance w.r.t. a measurement scaling) Let P be s.t.

$$P = I - \boldsymbol s\big(\boldsymbol s^\top \boldsymbol s\big)\boldsymbol s^\top. \tag{40}$$

Then the statistic

$$T\left[\boldsymbol r\right] = \frac{\boldsymbol r^\top P \boldsymbol r}{\boldsymbol r^\top\big(I - P\big)\boldsymbol r} \tag{41}$$

is maximally invariant w.r.t. G.

PROOF. Invariance:

$$T\big[g(\boldsymbol r)\big] = \frac{c^2\boldsymbol r^\top P \boldsymbol r}{c^2\boldsymbol r^\top\big(I - P\big)\boldsymbol r} = \frac{\boldsymbol r^\top P \boldsymbol r}{\boldsymbol r^\top\big(I - P\big)\boldsymbol r} = T\left[\boldsymbol r\right]. \tag{42}$$

Maximality:

$$\begin{aligned} & T\big[\widehat{\boldsymbol r}\big] = T\left[\boldsymbol r\right] \\ &\longrightarrow \frac{\boldsymbol r^\top P \boldsymbol r}{\boldsymbol r^\top\big(I - P\big)\boldsymbol r} = \frac{\widehat{\boldsymbol r}^\top P \widehat{\boldsymbol r}}{\widehat{\boldsymbol r}^\top\big(I - P\big)\widehat{\boldsymbol r}} \\ &\longrightarrow \widehat{\boldsymbol r}^\top\left(P - I\frac{\boldsymbol r^\top P \boldsymbol r}{\boldsymbol r^\top \boldsymbol r}\right)\widehat{\boldsymbol r} = 0 \\ &\longrightarrow \widehat{\boldsymbol r} = g(\boldsymbol r), \quad \exists g \in G. \end{aligned} \tag{43}$$

7.2 Proof of Theorem 8

We employ the following notation to rewrite the measurements as time-series:

$$\begin{aligned} \boldsymbol x(k) &:= \left[x_1(k), \ldots, x_M(k)\right]^\top \\ \boldsymbol y(k) &:= \left[y_1(k), \ldots, y_M(k)\right]^\top \\ \boldsymbol d(k) &:= \left[d_1(k), \ldots, d_M(k)\right]^\top \\ \boldsymbol w(k) &:= \left[w_1(k), \ldots, w_M(k)\right]^\top \\ \boldsymbol v(k) &:= \left[v_1(k), \ldots, v_M(k)\right]^\top \end{aligned}$$

These quantities shall not be confused with $\boldsymbol y_j$, $\boldsymbol d_j$, $\boldsymbol w_j$, $\boldsymbol v_j$. The latter in fact correspond to, e.g., the set of measurements of the specific agent j and all the times $k = 1, \ldots, T$, while $\boldsymbol y(k)$ corresponds to the set of measurements relative to the specific time k and all the agents $j = 1, \ldots, M$.

We begin with all the information contained in the time-series measurements and apply an invertible transformation such that

$$D\begin{bmatrix} \boldsymbol y(0) \\ \vdots \\ \boldsymbol y(T) \end{bmatrix} = (I + H)\begin{bmatrix} \boldsymbol x(0) \\ \boldsymbol d(0) + \boldsymbol w(0) \\ \vdots \\ \boldsymbol d(T-1) + \boldsymbol w(T-1) \end{bmatrix} + D\begin{bmatrix} \boldsymbol v(0) \\ \vdots \\ \boldsymbol v(T) \end{bmatrix} \tag{44}$$

where

$$H := \begin{bmatrix} 0 & & & \\ A - I & 0 & & \\ \vdots & \ddots & \ddots & \\ (A-I)A^{(T-2)} & \ldots & A - I & 0 \end{bmatrix} \tag{45}$$

and D was defined in Theorem 8. We begin by asking for invariance to the unknown subspace gain induced by H. Observing that $A - I$ is the network Laplacian, it follows that it has a single zero eigenvalue corresponding to the left eigenvector $\boldsymbol{p}^\top$. Thus we can directly apply Lemma 14 and write

$$T_0[\boldsymbol{y}] = RD \begin{bmatrix} \boldsymbol{y}(0) \\ \vdots \\ \boldsymbol{y}(T) \end{bmatrix} = R \begin{bmatrix} \boldsymbol{x}(0) \\ \boldsymbol{d}(0) + \boldsymbol{w}(0) \\ \vdots \\ \boldsymbol{d}(T-1) + \boldsymbol{w}(T-1) \end{bmatrix} + RD \begin{bmatrix} \boldsymbol{v}(0) \\ \vdots \\ \boldsymbol{v}(T) \end{bmatrix} \tag{46}$$

We then observe that the statistic has unknown correlated noise, written as

$$\mathrm{Cov}\left[T_0[\boldsymbol{y}]\right] = \begin{bmatrix} \sigma_v + \sigma_w & -\sigma_v & & & \\ -\sigma_v & 2\sigma_v + \sigma_w & -\sigma_v & & \\ & \ddots & \ddots & \ddots & \\ & & -\sigma_v & 2\sigma_v + \sigma_w & -\sigma_v \\ & & & -\sigma_v & 2\sigma_v + \sigma_w \end{bmatrix} \tag{47}$$

Applying thus Lemma 13 we obtain a statistic which is invariant to the correlated noise induced by $T_0[\boldsymbol{y}]$ by writing

$$T_1\left[T_0[\boldsymbol{y}]\right] = QRD \begin{bmatrix} \boldsymbol{y}(0) \\ \vdots \\ \boldsymbol{y}(T) \end{bmatrix} = QR \begin{bmatrix} \boldsymbol{x}(0) \\ \boldsymbol{d}(0) + \boldsymbol{w}(0) \\ \vdots \\ \boldsymbol{d}(T-1) + \boldsymbol{w}(T-1) \end{bmatrix} + 2(\sigma_v + \sigma_w) Q \begin{bmatrix} \boldsymbol{n}(0) \\ \vdots \\ \boldsymbol{n}(T) \end{bmatrix}. \tag{48}$$

Next we observe that the noise mean and non-local inputs induce a bias in the subspace

$$I - F_\ell^\top F_\ell = \begin{bmatrix} \boldsymbol{u}_1 & \dots & \boldsymbol{u}_{\ell-1} & \boldsymbol{u}_{\ell+1} & \dots & \boldsymbol{u}_M & S_1 & \dots & S_M & \mathbf{1} \end{bmatrix}. \tag{49}$$

We thus apply Lemma 12 and obtain the composed statistic

$$T_2\Big[T_1\left[T_0[\boldsymbol{y}]\right]\Big] = F_\ell QRD \begin{bmatrix} \boldsymbol{y}(1) \\ \vdots \\ \boldsymbol{y}(T) \end{bmatrix} = \mu_\ell F_\ell QR \begin{bmatrix} \boldsymbol{x}(0) \\ \boldsymbol{u}_\ell(0) \\ \vdots \\ \boldsymbol{u}_\ell(T-1) \end{bmatrix} + 2(\sigma_v + \sigma_w) Q \begin{bmatrix} \boldsymbol{n}(0) \\ \vdots \\ \boldsymbol{n}(T) \end{bmatrix}. \tag{50}$$

Lastly, we observe that $(\sigma_v + \sigma_w) \in \mathbb{R}$ induces a measurement scaling. We thus can apply Lemma 15 and obtain the $T[\boldsymbol{z}_\ell]$ in (7).

To prove then that the resulting test is maximally invariant, we notice that $T_2[\cdot], T_1[\cdot], T_0[\cdot]$ are all unitary. Additionally, we observe that $T_1[\cdot]$ was designed after applying $T_0[\cdot]$ and, similarly, $T_2[\cdot]$ was designed after applying $T_1[\dots]$, etc., and thus the second requirement of Lemma 10 is by construction satisfied. Thus the composed statistic is guaranteed to be maximally invariant.