Innovative Attack Detection Solutions for Wireless Networks With Application to Location Security

Danilo Orlando, *Senior Member, IEEE*, Stefania Bartoletti, *Member, IEEE*, Ivan Palamà, Giuseppe Bianchi, and Nicola Blefari Melazzi

***Abstract*— Modern wireless communication networks are threatened by new generations of radio hackers. These are skilled attackers equipped with low-cost software radios, suitably instrumented so as to monitor, degrade, or even alter the radio signals. The aim of this paper is to devise innovative detection architectures against the most common classes of threats: broadband noise jammers, whose goal is to reduce the signal-to-noise ratio, and spoofing/meaconing attacks, which aim to inject false or incorrect information into the receiver. To this end, we resort to the hypothesis testing theory and solve the associated problems by means of the GLRT possibly accounting for penalty terms. The resulting decision schemes represent the main technical novelty of this work. The analysis of their performance focuses on a location security case study for 4G/5G cellular networks. To this end, we leverage measurement models from the cellular localization literature and generate data according to these models. The numerical results show the effectiveness of the proposed approaches in comparison with suitable counterparts.**

***Index Terms*— Attack detection, generalized likelihood ratio test, location security, meaconing, model order selection, noise jamming, spoofing, wireless networks, 4G/5G.**

I. Introduction

THE availability of ultra-cheap software-defined radio boards along with the simultaneous development of open-source implementations of the standard protocol stacks, such as for the 5G cellular standard, come with a bleak side: these instruments may also be used to attack deployed systems. As a matter of fact, there exists a plethora of works in the open literature [1]–[13] showing how easy is for a tech-savvy opponent to build ultra-low cost jammers or even LTE/5G "rogue" base stations [14] capable of generating fake signals (with counterfeit information) or interfering with the legitimate ones (with a consequent quality of service degradation). It readily follows that next generation networks should be empowered with suitable countermeasures acting at different stack layers and that can preserve data integrity. A tangible example of the aforementioned situations is provided by the 5G location services where the user equipment (UE) position is estimated by combining timing, angle, and power measurements of the signals received by other UEs and/or gNBs (either in uplink or downlink). Specifically, the location management function is responsible for such a combination [15]. Now, in the presence of an attack, the quality of the UE position estimate might be heavily impaired due to induced low signal-to-noise ratio (SNR) values or false measurements. It is also important to observe that other kinds of information can be targeted by hostile platforms.

Generally speaking, an electronic attack consists of two main activities:[1] *Jamming* and *Deception* [16]. The attacks belonging to the first category intentionally inject noise into the channel with the consequent reduction of the SNR and, hence, disruption of the receiver functionalities [17]–[19]. They comprise the transmission of wideband noise, also known as broadband noise jammer (BBNJ), partial-band noise, and narrowband noise. The first two cases apply to a nonagile jammer where the interfering signals occupy a portion of or the entire spectrum in use by the communication system and stay in one place of the spectrum. The third case is associated with a jammer attempting to follow a frequency-hopping target [20]. Most of the existing detection and mitigation strategies against noise-like jamming take place at physical and/or protocol layer in (see [21, and references therein]). For instance, in [21], the authors investigate the effects of noise-like interference on the Physical Uplink Control Channel in LTE and propose a mitigation strategy based upon the Radio Link Control protocol. Anti-jamming techniques in the context of cognitive radio networks are addressed in [22], where several mitigation techniques are reviewed and a new anti-jamming protocol relying on probabilistic pairing and frequency tuning is proposed. Other examples of electronic counter-countermeasures are given in [23], [24] where an approach based on game theory is applied at the design stage. Mitigation solutions conceived at the physical layer can be

Manuscript received 24 December 2021; revised 13 May 2022; accepted 15 July 2022. Date of publication 26 July 2022; date of current version 9 January 2023. This work was supported by the European Union's Horizon 2020 Research and Innovation Program under Grant 871249. The associate editor coordinating the review of this article and approving it for publication was K. Xue. *(Corresponding author: Danilo Orlando.)*

Danilo Orlando is with the Università degli Studi "Niccolò Cusano," 00166 Roma, Italy (e-mail: danilo.orlando@unicusano.it).

Stefania Bartoletti is with the National Research Council of Italy, CNR-IEIIT, CNIT, 40136 Bologna, Italy (e-mail: stefania.bartoletti@cnr.it).

Ivan Palamà, Giuseppe Bianchi, and Nicola Blefari Melazzi are with the Department of Electronic Engineering, CNIT—University of Roma Tor Vergata, 00133 Rome, Italy (e-mail: ivan.palama@uniroma2.it; giuseppe.bianchi@uniroma2.it; blefari@uniroma2.it).

This article has supplementary material provided by the authors and color versions of one or more figures available at https://doi.org/10.1109/TWC.2022.3192225.

Digital Object Identifier 10.1109/TWC.2022.3192225

[1] Actually, this classification is not exhaustive since other kinds of electronic attacks are possible such as *directed energy* whose application leads to the permanent destruction of the communication equipment.

found in [25], where the received signal is classified by means of a deep convolutional neural network fed by the signal features in the wavelet domain. In [26], the authors leverage random matrix theory tools to conceive a multiple hypothesis test for jamming detection. To this end, they estimate the jammer subspace through the sample covariance matrix [27] and project the received data onto the user subspace in order to mitigate the jamming components. Finally, the authors of [20] apply change-point detection algorithms to power-related (high-level) measurements provided by the UE in order to declare that a jamming attack is ongoing. More importantly, such algorithms are rather general and (almost completely) disregard the underlying wireless technology.

The intent of the deception attacks is to mislead an opponent by creating a ruse (*spoofing/meaconing*) [19], [28], [29]. To this end, fake information is injected into the victim system. False communication signals are an important part of any tactical deception activity. For instance, in respect of location services, a spoofer can intercept the positioning messages exchanged by two legitimate actors and suitably delay or modify them to generate false positions. The received signal strength (RSS) information is widely used for spoofing detection in wireless networks [30]–[32, and references therein]. In [30], the authors apply K-means cluster analysis to detect the spoofer and exploit this detector also for localization purposes. The detection of spoofing attacks in mobile wireless networks is addressed in [31], where the proposed algorithm cluster the RSS readings using the Otsu method. Detection algorithms relying on the spatial correlation properties of RSS are conceived in [32]. More recently, in [33], physical-layer authentication exploiting radio channel information to detect spoofing attacks in wireless networks is investigated. In addition, reinforcement learning techniques are used to find the optimal strategy in a dynamic environment with incomplete information. Finally, in [34], [35], algorithms against pilot spoofing attacks based upon the likelihood ratio test are proposed. These algorithms are fed by raw data collected by (possibly massive) Multiple-Input-Multiple-Output systems.

Summarizing, these attacks (jamming and spoofing) can be thwarted at data and/or signal levels, which differ in the amount of both the available information and data samples to be processed. In fact, if on one side, at signal level, data contains all the available information, on the other side, such volume of samples might require time-demanding algorithms [21], [25], [26]. On the contrary, at data level, an estimation has been already performed with a consequent loss of information (with respect to signal samples) and a lightening of the computational load. More importantly, in many situations of practical interest raw data are not available due to several factors that limit the access to them.

With the above remarks in mind and given the in-progress development of the new generations communication networks, we further extend the ideas behind [20] and devise new decision schemes operating on high-level data and tailored to specific attacks. The main difference with respect to [20] resides in the models assumed at the design stage that, in this case, more accurately adhere to the real behavior of the considered threats. As a result, the proposed decision rules represent an improvement of the baseline in [20]. Thus, to define such models, we need to clarify the possible effects of the two main disturbance activities on collected data. Starting from jamming attacks, we notice that the impairment of the SNR generated by the jammer might lead to an increase of the measurement/information uncertainty and, hence, of the related variance. On the other hand, a spoofer would leave unaltered measurement/information uncertainty while replacing part or decrease the entire set of the expected values with counterfeit measurement/information. Now, assuming an adaptive monitoring of the incoming data within a suitable temporal sliding window, when an attack takes place at a given time instant, the window under test will contain a discontinuity point in data distribution. Otherwise stated, data window can be partitioned into two subsets whose distribution parameters are affected by an abrupt change depending on the situation before the attack and on the attack itself. As a consequence, we are faced with a change detection problem containing only one change time instant that can be estimated through a linear search grid [36]–[40, and references therein]. This problem is solved by means of parametric methods and, in the specific case, we resort to a suboptimal design criterion since a uniformly most powerful test (if exists) is difficult to find. To be more definite, we apply the generalized likelihood ratio test (GLRT) that is widely adopted in signal processing and, in most cases, returns excellent detection performance [36]. Although the GLRT is a well-established design technique, its application to the considered change point detection problems leads to new detection architectures and derivations that represent the main technical novelty of this paper. In fact, the GLRT for BBNJ attack detection does not admit closed-form expression (at least to the best of authors' knowledge) and, hence, we exploit suitable approximations that are a good compromise between computational load and detection performance. As for the spoofer detection, the considered problem still contains a single change point but it comprises one null hypothesis and several alternative hypotheses arising from the fact that not all the components of the mean vector are counterfeit. This new problem is solved by applying the elegant design framework developed in [41] where the so-called penalized GLRT is introduced. Such an approach allows us to overcome the limitation of the maximum likelihood framework in the presence of nested hypotheses and, as side information, returns which components are spoofed. Remarkably, such components can be discarded by the system. Finally, the effectiveness of the proposed techniques also in comparison with conventional competitors is proved exploiting a case study related to location security in 5G networks, where data are generated according to the state-of-the-art experimental models from the literature.

The remainder of the paper is organized as follows. Section II is devoted to sensor model description and to the formal statement of the detection problems at hand, while the detection rules for jamming and spoofing attacks are provided in Section III and IV, respectively. In Section V, some numerical results related to a location security example are given to show the effectiveness of the proposed strategies.

Section VI contains concluding remarks and charts a course for future works. Analytical derivations are confined to the appendices (the supplemental material contains the derivations of the considered competitors).

A. Notation

In the sequel, vectors and matrices are denoted by boldface lower-case and upper-case letters, respectively. Symbols $\det(\cdot)$, $\mathrm{Tr}\,(\cdot)$, and $(\cdot)^T$ denote the determinant, trace, and transpose, respectively. $\mathbb{R}$ is the set of real numbers and $\mathbb{R}^{N\times M}$ is the Euclidean space of $(N\times M)$-dimensional real matrices (or vectors if $M=1$). The Euclidean norm of a generic vector $\boldsymbol{x}$ is denoted by $\|\boldsymbol{x}\|$ whereas the modulus of a real number x is denoted by $|x|$. If $\mathcal{A}$ is a set, then its cardinality is $|\mathcal{A}|$, while $\cdot\backslash\cdot$ denotes the difference between sets. For any N-dimensional vector $\boldsymbol{x}$, $\boldsymbol{X}=\mathbf{diag}\,\{\boldsymbol{x}\}$ is a $(N\times N)$-dimensional diagonal matrix whose principal diagonal contains the elements of $\boldsymbol{x}$. The gradient of a real-valued function with vector argument, $g(\boldsymbol{x})$, is denoted by $\partial g(\boldsymbol{x})/\partial \boldsymbol{x}$. Symbols $\boldsymbol{I}$ and $\mathbf{0}$ indicate the identity matrix and the null matrix or vector, respectively, whose size depends on the context; $\mathbf{1}$ is a vector whose entries are equal to 1. For any Hermitian matrix $\boldsymbol{A}$, $\boldsymbol{A}\succ\mathbf{0}$ means that $\boldsymbol{A}$ is a positive definite matrix. Finally, we write $\boldsymbol{x}\sim\mathcal{N}_N(\boldsymbol{m},\boldsymbol{\Sigma})$ if $\boldsymbol{x}$ is a N-dimensional Gaussian vector with mean $\boldsymbol{m}$ and covariance matrix $\boldsymbol{\Sigma}\succ\mathbf{0}$, whereas, if $\boldsymbol{X}\in\mathbb{R}^{N\times K}$, $\boldsymbol{X}\sim\mathcal{N}_{N,K}(\boldsymbol{m},\boldsymbol{\Sigma},\boldsymbol{I})$ means that the columns of $\boldsymbol{X}$ are Independent and Identically Distributed (IID) random vectors following the Gaussian distribution with mean $\boldsymbol{m}$ and covariance matrix $\boldsymbol{\Sigma}\succ\mathbf{0}$.

II. Sensor Model and Problem Formulation

Let us assume that an UE is gathering information from the network infrastructure and denote by $\boldsymbol{Z}=[\boldsymbol{z}_1,\boldsymbol{z}_2,\ldots,\boldsymbol{z}_K]\in\mathbb{R}^{N\times K}$ the data matrix whose kth column, $k\in\Omega=\{1,\ldots,K\}$, contains the values of the (high-level) measurements of interest acquired at the kth discrete time instant. For instance, in a localization system, such measurements can include the direction of arrival (DOA), the time of arrival (TOA), the observed time difference of arrival (OTDOA), the reference signal received power (RSRP), or other parameters useful for localization purposes (see [42]–[44, and references therein]). Moreover, we assume that the measurement errors are statistically independent over the time and obey the Gaussian distribution with zero mean and positive definite covariance matrix.[2]

Therefore, when data are collected in the absence of intentional interfering signals (namely, under the null hypothesis), all the measurements are IID within the observation time window, namely[3] $\boldsymbol{Z}\sim\mathcal{N}_{N,K}(\boldsymbol{m}_0,\boldsymbol{\Sigma}_0,\boldsymbol{I})$, where $\boldsymbol{m}_0\in\mathbb{R}^{N\times 1}$ contains the actual values of the parameters of interest and $\boldsymbol{\Sigma}_0\in\mathbb{R}^{N\times N}$ is the positive definite error covariance matrix, which can exhibit either a generic symmetric or diagonal structure. The first structure accounts for a possible correlation among the measurements, whereas the second one is used to model independent parameters. Under the alternative hypothesis, namely in the presence of a BBNJ or spoofing attack, the distribution of $\boldsymbol{Z}$ modifies as described in the next subsections.

[2]This assumption is useful for the analytical tractability of the problems. Nevertheless, the performance of the proposed methods will be assessed also under non-Gaussian models inferred from real data that can be found in the open literature.

[3]Notice also that we are assuming that the unintentional interference affecting the received signals is stationary at least within the observation time interval.

A. BBNJ Attack Model

When at a certain time index $K_0+1\in\Omega$, a BBNJ attack aimed at disrupting the receiver functionalities is performed, the quality of the measurements provided by the sensors would impair due to an increase of the disturbance power or, equivalently, a decrease of the SNR. Otherwise stated, an abrupt change in the measurement covariance matrix occurs. As a consequence, data matrix can be partitioned as $\boldsymbol{Z}=[\boldsymbol{Z}_{1:K_0},\boldsymbol{Z}_{K_0+1:K}]$, where $\boldsymbol{Z}_{1:K_0}=[\boldsymbol{z}_1,\ldots,\boldsymbol{z}_{K_0}]\sim\mathcal{N}_{N,K_0}(\boldsymbol{m}_1,\boldsymbol{\Sigma}_1,\boldsymbol{I})$ and $\boldsymbol{Z}_{K_0+1:K}=[\boldsymbol{z}_{K_0+1},\ldots,\boldsymbol{z}_K]\sim\mathcal{N}_{N,K_1}(\boldsymbol{m}_1,\boldsymbol{\Sigma}_2,\boldsymbol{I})$ with $K_1=K-K_0$, $\boldsymbol{\Sigma}_2-\boldsymbol{\Sigma}_1\succ\mathbf{0}$, and $K_0\in\Omega$ being unknown. Finally, $\boldsymbol{m}_1$ and $\boldsymbol{\Sigma}_i$, $i=1,2$, are the unknown mean vector (containing the information of interest) and the unknown covariance matrices under H_1, respectively. As a consequence, we can formulate the following binary hypothesis test

$$\begin{cases} H_0:\boldsymbol{Z}\sim\mathcal{N}_{N,K}(\boldsymbol{m}_0,\boldsymbol{\Sigma}_0,\boldsymbol{I}), \\ H_1:\begin{cases}\boldsymbol{Z}_{1:K_0}\sim\mathcal{N}_{N,K_0}(\boldsymbol{m}_1,\boldsymbol{\Sigma}_1,\boldsymbol{I}),\\ \boldsymbol{Z}_{K_0+1:K}\sim\mathcal{N}_{N,K_1}(\boldsymbol{m}_1,\boldsymbol{\Sigma}_2,\boldsymbol{I}).\end{cases}\end{cases} \tag{1}$$

Finally, we stress that $\boldsymbol{\Sigma}_2-\boldsymbol{\Sigma}_1\succ\mathbf{0}$, namely, data affected by BBNJ might exhibit an increased uncertainty.

B. Spoofing/Meaconing Attack Model

The second situation under consideration encompasses the presence of a spoofing/meaconing attack that consists of injecting false information into the network receivers. This injection is modeled in terms of a variation of the mean vectors starting from a time instant $K_0+1\in\Omega$. However, it is important to highlight that even a subset of components of the mean vector can be falsified by the spoofer. For instance, focusing on the location data, the spoofer might counterfeit either DOA or TOA or both measurements. Thus, let us define $\Gamma_N=\{1,\ldots,N\}$ and denote by $\mathcal{P}(\Gamma_N)$ the set obtained by excluding the empty set from the powerset of Γ_N; it follows that $|\mathcal{P}(\Gamma_N)|=N_\Gamma=2^N-1$. The generic element of $\mathcal{P}(\Gamma_N)$ is indicated with Γ_i, $i=1,\ldots,N_\Gamma$. Now, given a value of i, we assume that a spoofing attack can affect only the components of the mean vector indexed by Γ_i. Therefore, denoting by $\boldsymbol{m}_{\Gamma_i}\in\mathbb{R}^{N\times 1}$ the mean vector of $\boldsymbol{z}_k$, $k>K_0$ (i.e., after the attack), it has the same values as the components of $\boldsymbol{m}_1\in\mathbb{R}^{N\times 1}$ (that is the mean value of $\boldsymbol{z}_k$ for $k\le K_0$) except for those indexed by Γ_i, that are spoofed. It follows that we can model the spoofer detection problem as a multiple hypothesis testing problem with only one null hypothesis and

several (possibly nested) alternative hypotheses, namely

$$\begin{cases} H_0 : \boldsymbol{Z} \sim \mathcal{N}_{N,K}(\boldsymbol{m}_0, \boldsymbol{\Sigma}_0, \boldsymbol{I}), \\ H_{1,i} : \begin{cases} \boldsymbol{Z}_{1:K_0} \sim \mathcal{N}_{N,K_0}(\boldsymbol{m}_1, \boldsymbol{\Sigma}_1, \boldsymbol{I}), \\ \boldsymbol{Z}_{K_0+1:K} \sim \mathcal{N}_{N,K_1}(\boldsymbol{m}_{\Gamma_i}, \boldsymbol{\Sigma}_1, \boldsymbol{I}), \end{cases} i = 1, \ldots, N_\Gamma, \end{cases} \tag{2}$$

where $\boldsymbol{m}_0$, $\boldsymbol{m}_1$, and $\boldsymbol{\Sigma}_0$ have been already defined and $\boldsymbol{\Sigma}_1 \in \mathbb{R}^{N\times N}$ is the unknown covariance matrix of $\boldsymbol{Z}$ under $H_{1,i}$.

Before concluding this section, we provide some definitions which are used in the ensuing developments. More precisely, the probability density function (PDF) of $\boldsymbol{Z}$ under H_0 for all the considered problems has the following expression $f_0(\boldsymbol{Z}; \boldsymbol{m}_0, \boldsymbol{\Sigma}_0) = \prod_{k=1}^{K} f(\boldsymbol{z}_k; \boldsymbol{m}_0, \boldsymbol{\Sigma}_0)$, while the PDF of $\boldsymbol{Z}$ under H_1 for problem (1) and that under the generic $H_{1,i}$ for problem (2) are given by $f_1(\boldsymbol{Z}; \boldsymbol{m}_1, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, K_0) = \prod_{k=1}^{K_0} f(\boldsymbol{z}_k; \boldsymbol{m}_1, \boldsymbol{\Sigma}_1) \times \prod_{k=K_0+1}^{K} f(\boldsymbol{z}_k; \boldsymbol{m}_1, \boldsymbol{\Sigma}_2)$ and $f_{1,i}(\boldsymbol{Z}; \boldsymbol{m}_1, \boldsymbol{m}_{\Gamma_i}, \boldsymbol{\Sigma}_1, K_0) = \prod_{k=1}^{K_0} f(\boldsymbol{z}_k; \boldsymbol{m}_1, \boldsymbol{\Sigma}_1) \prod_{k=K_0+1}^{K} f(\boldsymbol{z}_k; \boldsymbol{m}_{\Gamma_i}, \boldsymbol{\Sigma}_1)$, respectively, where

$$f(\boldsymbol{z}; \boldsymbol{m}, \boldsymbol{\Sigma}) = \frac{\exp\left\{-\frac{1}{2}\mathrm{Tr}\left[\boldsymbol{\Sigma}^{-1}(\boldsymbol{z}-\boldsymbol{m})(\boldsymbol{z}-\boldsymbol{m})^T\right]\right\}}{(2\pi)^{N/2}[\det(\boldsymbol{\Sigma})]^{1/2}}. \tag{3}$$

III. BBNJ Detector Designs

In this section, we focus on problem (1) and solve it by applying design procedures grounded on the GLRT, whose general structure is

$$\frac{\max\limits_{K_0\in\Omega_0} \max\limits_{\boldsymbol{m}_1} \max\limits_{\boldsymbol{\Sigma}_1} \max\limits_{\boldsymbol{\Sigma}_2} f_1(\boldsymbol{Z}; \boldsymbol{m}_1, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, K_0)}{\max\limits_{\boldsymbol{m}_0} \max\limits_{\boldsymbol{\Sigma}_0} f_0(\boldsymbol{Z}; \boldsymbol{m}_0, \boldsymbol{\Sigma}_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{4}$$

where η is the detection threshold[4] to be set in order to ensure a preassigned probability of false alarm (P_{fa}). Moreover, as stated before, two classes for the covariance structure are considered: (i) the available measurements are uncorrelated leading to diagonal covariance matrices; (ii) there exists a correlation among the measurements, i.e., the covariance matrices are generally symmetric and, in this case, we also assume that $\min\{K_0, K_1\} > N$ (see Appendix B).

A. BBNJ Detector: Uncorrelated Measurements

Let us assume that measurements are uncorrelated, then, we can write $\boldsymbol{\Sigma}_0 = \mathbf{diag}\{\sigma_{0,1}^2, \ldots, \sigma_{0,N}^2\}$, $\boldsymbol{\Sigma}_1 = \mathbf{diag}\{\sigma_{1,1}^2, \ldots, \sigma_{1,N}^2\}$, and $\boldsymbol{\Sigma}_2 = \mathbf{diag}\{\sigma_{1,1}^2 + \Delta\sigma_1^2, \ldots, \sigma_{1,N}^2 + \Delta\sigma_N^2\}$ with $\Delta\sigma_n^2 > 0$, $n = 1, \ldots, N$.

As shown in Appendix A, the application of the plain GLRT to this case leads to time demanding estimation procedures for the unknown parameters. For this reason, we resort to a suboptimal approximation of the compressed log-likelihood under H_1 allowing for a reasonable compromise between detection performance and computational requirements. Specifically, such an approximation has the following expression

$$\begin{aligned} \max_{K_0} \Bigg\{ & -\frac{K_0}{2} \sum_{n\in\tilde{\Gamma}(\widehat{\mathcal{B}})} \log\left[\frac{1}{K_0}\sum_{k=1}^{K_0}(z_{k,n} - \widehat{m}_{1,n})^2\right] \\ & -\frac{K_1}{2} \sum_{n\in\tilde{\Gamma}(\widehat{\mathcal{B}})} \log\left[\frac{1}{K_1}\sum_{k=K_0+1}^{K}(z_{k,n} - \widehat{m}_{1,n})^2\right] \\ & +\frac{K}{2} \sum_{n\in\tilde{\Gamma}(\widehat{\mathcal{B}})} \log\left[\frac{1}{K}\sum_{k=1}^{K}(z_{k,n} - \widehat{m}_{0,n})^2\right] \Bigg\} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \end{aligned} \tag{5}$$

where $z_{k,n}$ is the nth component of vector $\boldsymbol{z}_k$, $\widehat{m}_{i,n}$, $i = 0, 1$, is the estimate of the nth component of $\boldsymbol{m}_i$ under H_i (Appendix A), and $\tilde{\Gamma}(\widehat{\mathcal{B}})$ is a suitable estimate of $\Gamma(\mathcal{B})$ defined by (16). Note that if $\tilde{\Gamma}(\widehat{\mathcal{B}}) = \emptyset$, then the decision statistic is equal to zero.

In the following, we refer to this decision rule as BBNJ detector for uncorrelated measurements (BBNJ-D-UM).

B. BBNJ Detector: Correlated Measurements

In this subsection, $\boldsymbol{\Sigma}_i$, $i = 0, 1, 2$, in (1) are no longer diagonal but positive definite symmetric matrices. In addition, we assume that $\min\{K_0, K_1\} > N$. This constraint ensures that the sample covariance matrices based upon $\boldsymbol{Z}_{1:K_0}$ and $\boldsymbol{Z}_{K_0+1:K}$ are nonsingular with probability 1 [45] as required in Appendix B. It follows that $K_0 \in \Omega_0 = \{N+1, \ldots, K-N-1\}$.[5]

In this case, the likelihood maximization under H_1 cannot be conducted in closed form (at least to the best of authors' knowledge) due to the intractable mathematics. For this reason, we again resort to an approximate GLRT that allows us to come up with a simplified expression for the decision statistic. In Appendix B, we devise such a decision rule whose expression is

$$\max_{K_0\in\Omega_0} \Lambda(\boldsymbol{Z}; K_0) \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{6}$$

where

$$\begin{aligned} & \Lambda(\boldsymbol{Z}; K_0) \\ & = -\frac{K_0}{2}\log\det\left[\frac{1}{K_0}\sum_{k=1}^{K_0}(\boldsymbol{z}_k - \bar{\boldsymbol{m}}_1)(\boldsymbol{z}_k - \bar{\boldsymbol{m}}_1)^T\right] \\ & \quad -\frac{K_1}{2}\log\det\left[\frac{1}{K_1}\sum_{k=K_0+1}^{K}(\boldsymbol{z}_k - \bar{\boldsymbol{m}}_1)(\boldsymbol{z}_k - \bar{\boldsymbol{m}}_1)^T\right] \\ & \quad +\frac{K_0 N}{2}\log K_0 \\ & \quad +\frac{K}{2}\log\det\left[\sum_{k=1}^{K}(\boldsymbol{z}_k - \bar{\boldsymbol{m}}_0)(\boldsymbol{z}_k - \bar{\boldsymbol{m}}_0)^T\right] \\ & \quad +\frac{K_1 N}{2}\log K_1, \end{aligned} \tag{7}$$

[4]Hereafter, we denote by η the generic detection threshold.

[5]From an operating point of view and for sufficiently wide data windows, the above requirement can be fulfilled.

if $\frac{1}{K_1}\sum_{k=K_0+1}^{K}(\boldsymbol{z}_k-\bar{\boldsymbol{m}}_1)(\boldsymbol{z}_k-\bar{\boldsymbol{m}}_1)^T-\frac{1}{K_0}\sum_{k=1}^{K_0}(\boldsymbol{z}_k-\bar{\boldsymbol{m}}_1)(\boldsymbol{z}_k-\bar{\boldsymbol{m}}_1)^T \succ \mathbf{0}$, and $\Lambda(\boldsymbol{Z};K_0)=0$, otherwise. In (7), $\bar{\boldsymbol{m}}_0$ is the maximum likelihood estimate (MLE) of $\boldsymbol{m}_0$, whereas $\bar{\boldsymbol{m}}_1$ is the estimate of $\boldsymbol{m}_1$ given by (34).

In what follows, we refer to this decision scheme as BBNJ detector for correlated measurements (BBNJ-D-CM).

IV. Spoofing Detector Designs

Before providing the expressions of the spoofer detectors, it is worth noticing that problem (2) is a multiple hypothesis test with only one null hypothesis and several (possibly nested) alternative hypotheses. In this case, the GLRT, which is based upon the maximum likelihood (ML) approach, might prevent a reliable estimation of the actual alternative hypothesis. Therefore, in this case, we resort to the elegant and systematic framework devised in [41] that provides a theoretical justification for the so-called penalized likelihood ratio tests. The general structure of such decision statistics is the difference between a compressed log-likelihood ratio and a penalty term borrowed from the model order selection (MOS) rules [46] as the Akaike information criterion (AIC), Bayesian information criterion (BIC), and generalized information criterion (GIC), namely

$$\max_{i=1,\ldots,N_\Gamma}\{\log\Lambda_i(\boldsymbol{Z})-\gamma\cdot N_{p,i}\}\underset{H_0}{\overset{H_{1,\hat{i}}}{\gtrless}}\eta, \tag{8}$$

where $\hat{i}$ is the maximizer of the left-hand side of (8),

$$\Lambda_i(\boldsymbol{Z})=\frac{\max\limits_{K_0\in\Omega_0}\max\limits_{\boldsymbol{m}_1}\max\limits_{\boldsymbol{m}_2}\max\limits_{\boldsymbol{\Sigma}_1} f_{1,i}(\boldsymbol{Z};\boldsymbol{m}_1,\boldsymbol{m}_{\Gamma_i},\boldsymbol{\Sigma}_1,K_0)}{\max\limits_{\boldsymbol{m}_0}\max\limits_{\boldsymbol{\Sigma}_0} f_0(\boldsymbol{Z};\boldsymbol{m}_0,\boldsymbol{\Sigma}_0)}, \tag{9}$$

$N_{p,i}$ is the number of unknown parameters under $H_{1,i}$, and γ is a factor depending on which MOS rule is used to obtain (8) [41], namely $\gamma=1$ for AIC-based detector, $\gamma=\log(NK)/2$ for BIC-based detector, and $\gamma=(1+\rho)/2$, $\rho>1$, for GIC-based detector. Similarly to the BBNJ detection problem, we consider uncorrelated and correlated measurements and, in the last case, we also assume that $K>N+1$. Finally, notice that in some cases we approximate $\Lambda_i(\boldsymbol{Z})$ (and, hence, (8)) since solving the related maximization problems involves intractable mathematics.

A. Spoofing Detector: Uncorrelated Measurements

Uncorrelated measurements yield the following covariance matrices: $\boldsymbol{\Sigma}_0=\mathbf{diag}\{\sigma_{0,1}^2,\sigma_{0,2}^2,\ldots,\sigma_{0,N}^2\}$ under H_0 and $\boldsymbol{\Sigma}_1=\mathbf{diag}\{\sigma_{1,1}^2,\sigma_{1,2}^2,\ldots,\sigma_{1,N}^2\}$ under $H_{1,i}$, $i=1,\ldots,N_\Gamma$. In Appendix C, we prove that the logarithm of (9) can be written as

$$\begin{aligned}\log\Lambda_i(\boldsymbol{Z})&=\max_{K_0}\left\{\frac{K}{2}\sum_{n=1}^{N}\log\left[\frac{1}{K}\sum_{k=1}^{K}(z_{k,n}-\widehat{m}_{0,n})^2\right]\right.\\&\quad-\frac{K}{2}\left\{\sum_{n\in\Gamma_N\setminus\Gamma_i}\log\left[\frac{1}{K}\sum_{k=1}^{K}(z_{k,n}-\widehat{m}_{1,n})^2\right]\right.\\&\quad+\sum_{n\in\Gamma_i}\log\left[\frac{1}{K}\left(\sum_{k=1}^{K_0}(z_{k,n}-\widehat{m}_{1,n})^2\right.\right.\\&\quad\left.\left.\left.\left.+\sum_{k=K_0+1}^{K}(z_{k,n}-\widehat{m}_{\Gamma_i,n})^2\right)\right]\right\}\right\},\end{aligned} \tag{10}$$

where $\widehat{m}_{0,n}$, $\widehat{m}_{1,n}$, and $\widehat{m}_{\Gamma_i,n}$ are given by (13), (38), and (39), respectively. Finally, in order to write (8), we need the number of unknown parameters under $H_{1,i}$ that is given by $N_{p,i}=2N+|\Gamma_i|$.

In the next sections, we will refer to this decision scheme as AIC/BIC/GIC-based spoofing detector for uncorrelated measurements (SP-D-UM).

B. Spoofing Detector: Correlated Measurements

In this last subsection, we address the case that data covariance matrices exhibit a general symmetric structure. Similarly to what observed in Subsection III-B, obtaining closed-form expression for $\Lambda_i(\boldsymbol{Z})$ is not an easy task due to the fact that the maximization of its numerator leads to intractable mathematics (at least to the best of authors' knowledge). Therefore, we resort to an approximation of $\Lambda_i(\boldsymbol{Z})$ where the unknown parameters are estimated by means of an alternating procedure that returns a nondecreasing sequence of log-likelihood values as described in Appendix D. Moreover, we assume that $K>N+1$. The final result is

$$\begin{aligned}\log\Lambda_i(\boldsymbol{Z})\approx&-\frac{K}{2}\log\det\left[\sum_{k=1}^{K_0}(\boldsymbol{z}_k-\widehat{\boldsymbol{m}}_1)(\boldsymbol{z}_k-\widehat{\boldsymbol{m}}_1)^T\right.\\&\left.+\sum_{k=K_0+1}^{K}(\boldsymbol{z}_k-\widehat{\boldsymbol{m}}_{\Gamma_i})(\boldsymbol{z}_k-\widehat{\boldsymbol{m}}_{\Gamma_i})^T\right]\\&+\frac{K}{2}\log\det\left[\sum_{k=1}^{K}(\boldsymbol{z}_k-\bar{\boldsymbol{m}}_0)(\boldsymbol{z}_k-\bar{\boldsymbol{m}}_0)^T\right]\end{aligned} \tag{11}$$

where $\bar{\boldsymbol{m}}_0$ is computed in Appendix B, while $\widehat{\boldsymbol{m}}_{\Gamma_i}$ and $\widehat{\boldsymbol{m}}_1$ are the estimates obtained through the cyclic procedure of Appendix D. Finally, notice that in this case $N_{p,i}=N(N+1)/2+N+|\Gamma_i|$.

The above decision scheme is referred to in the following as AIC/BIC/GIC-based spoofing detector for correlated measurements (SP-D-CM).

V. An Application: Location Security

In this section, we present a performance evaluation of the proposed detectors by focusing on two different operating scenarios that differ in data distribution models: (i) a general Gaussian scenario (i.e., data are generated according to the design assumptions) and (ii) a 5G localization scenario (i.e., data are generated by exploiting experimental models from the literature of cellular localization). The Gaussian scenario allows us to gain an assessment of the nominal detectors' performance, whereas the analysis based on experimental models allows us to appreciate the performance sensitivity with respect to realistic operating conditions. The performance metrics are the probability of detection (P_d) given a pre-assigned value for

the P_{fa}, the root-mean-square error (RMSE) for the estimates of the parameter of interest, and the probability of correct classification of the spoofed components (see the next subsections for the precise definition). For comparison purposes, we consider other parametric change point detection schemes that can be obtained through straightforward generalizations of existing/conventional approaches (and, hence, can be considered a baseline); in addition, more advanced competitors based upon a clustering approach are considered for both kinds of attacks. Due to the lack of space, we confine the derivations and final expressions of these competitors to the supplemental material. In what follows, they are referred to as BBNJ naive change detector (BBNJ-NCD), BBNJ latent variable model (LVM) change detector (BBNJ-LVM), spoofing naive detector for uncorrelated measurements (SP-NCD-UM), spoofing naive detector for correlated measurements (SP-NCD-CM), and the spoofing LVM change detector (SP-LVM).

A. Simulation Setting and Operating Scenarios

We assume a slow-moving UE, which is being tracked by the network infrastructure. UE derives range and DOA positioning measurements from signals sent by access node (AN) nodes. As described in Section II, we assume that at a certain time instant a malicious platform performs a jamming or a spoofing attack. According to 3GPP standard, we consider a scenario with the UE localized exploiting range (RSRP measurements) and DOA estimates (azimuth and elevation angles). As a consequence, the size of the generic vector $\boldsymbol{z}_k$ is $N = 3$ and the procedure to generate it is described below.

Since closed-form expressions for the performance metrics are not available, we resort to standard Monte Carlo counting techniques. More precisely, we exploit 5000 independent trials to estimate the P_{d} and the probability of correct classification, 1000 independent trials for the RMSE values, and $100/P_{\text{fa}}$ independent trials for the detection thresholds with $P_{\text{fa}} = 10^{-2}$. The proposed decision schemes are assessed accounting for two different lengths of the sliding window (namely, $K = 24, 32$) and several values for K_0. The BBNJ attack is simulated by varying the variance of the noise affecting the measurements. Specifically, starting from a diagonal covariance matrix, $\boldsymbol{\Sigma}_0$ say, set using the results of [47], [48], we modify the latter through a scaling factor γ such that[6] $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_2 = \gamma\boldsymbol{\Sigma}_0$ in (1). Further details on the construction of the covariance matrices are given with the *5G-based Scenario* below.

On the other hand, in the case of spoofing attack, the numerical examples are obtained by varying the mean value of the original signal. More precisely, after the change point, the entries of the mean measurement vector under H_0 are scaled by a factor ν, which represents "the amount of fake information" injected by the spoofer. As for the operating scenarios, we consider two cases.

1) Gaussian Scenario: time difference of arrival (TDOA) and DOA measurements are modeled as statistically independent Gaussian random variables with $\boldsymbol{m}_0$ and $\boldsymbol{\Sigma}_0$ obtained from the exemplary models of the 5G-based scenario as described below.

2) 5G-Based Scenario: we borrow distribution models from existing literature [47] where ranging is performed through downlink (DL)-TDOA measurements of 5G positioning reference signal in a urban macro environment with line-of-sight conditions (the SNR is 20 dB); and [48] where the angle estimates are obtained through a beam-RSRP of DL with 16 UE beams (the SNR is 20 dB). In particular, the ranging and angle errors used in the simulations are generated using the empirical cumulative distribution function (CDF) of elevation/azimuth angle error and 2D ranging error from papers [47], [48], respectively. As for the realization of the range and angle measurements, we consider the true range and true DOA (elevation and azimuth) when the distance between the AN and UE is equal to $d_0 = 200$ m and the DOA is 0 degrees. Then, the measurement error is added as two mutually independent random variables modeled according to the exemplary error PDFs proposed in [47] for the TDOA and [48] for the DOA. Finally, a dataset composed of 10^4 TDOA and DOA realizations is used to compute the empirical mean $\boldsymbol{m}_0 = \boldsymbol{m}_1$ and the empirical covariance matrix $\boldsymbol{\Sigma}_0 = \boldsymbol{\Sigma}_1$.

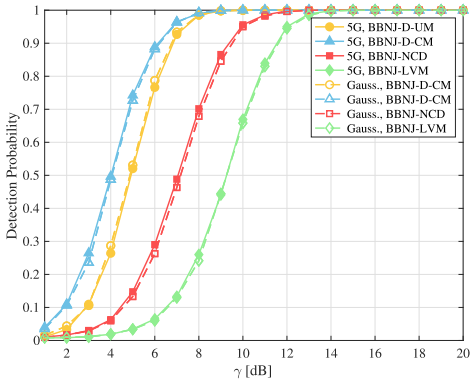


Fig. 1. P_{d} versus γ assuming $K = 32$ and $K_0 = 16$ (BBNJ detectors).

B. Performance of the BBNJ Detection Architectures

The performance of the BBNJ-D-UM, BBNJ-D-CM, BBNJ-NCD, and BBNJ-LVM is shown in Figs. 1 and 2 under both the Gaussian and 5G models. Fig. 1 shows P_{d} versus the parameter γ for $K = 32$ and $K_0 = 16$. The figure indicates that the nominal behavior of the considered decision rules is quite similar to that obtained in the 5G scenario. The BBNJ-D-CM exhibits the best performance and both BBNJ-D-CM and BBNJ-D-UM outperform the (baseline) BBNJ-NCD and the BBNJ-LVM. Fig. 2 shows the RMSE for the estimate of K_0 as a function of γ. Notice that the estimation performance of the LVM-based detector is not reported due to the fact that it does not return an estimate of K_0 but it clusters data without any constraint. Thus, a further processing stage would be required to estimate the change point position by means of the assigned labels. The figure shows that the estimation performance under the nominal conditions is comparable to that obtained using 5G measurement distributions. The BBNJ-D-CM exhibits higher RMSE values than the BBNJ-NCD, whereas the BBNJ-D-UM

[6]Actually, γ is related to the BBNJ transmitted power and is a function of the SNR through the underlying physical model.

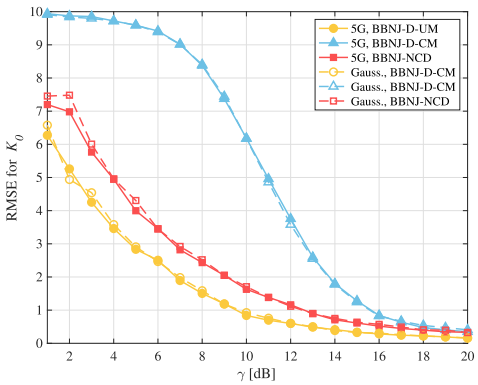


Fig. 2. RMSE versus γ for $K = 32$ and true value $K_0 = 16$ (BBNJ detectors).

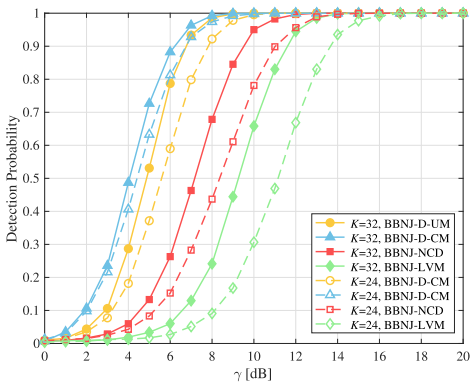


Fig. 5. $P_{\rm d}$ versus γ assuming $K_0 = 0.5K$, $K = 24, 32$, and the 5G scenario (BBNJ detectors).

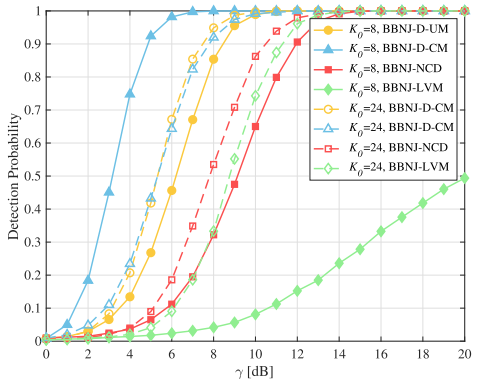


Fig. 3. $P_{\rm d}$ versus γ assuming $K = 32$, $K_0 = 8, 24$, and the 5G scenario (BBNJ detectors).

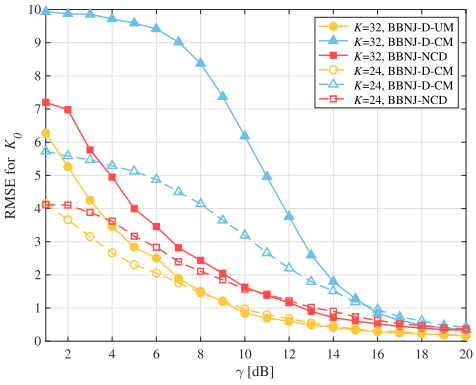


Fig. 6. RMSE versus γ assuming $K_0 = 0.5K$, for $K = 24, 32$, and the 5G scenario (BBNJ detectors).

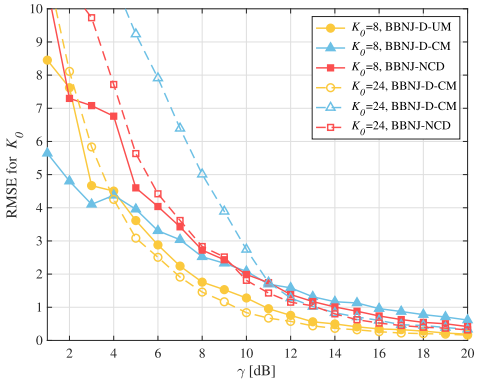


Fig. 4. RMSE versus γ assuming $K = 32$, $K_0 = 8, 24$, and the 5G scenario (BBNJ detectors).

provides the lowest RMSE values. The behavior of the BBNJ-D-CM observed in this figure can be explained by the fact that it is an approximation of the GLRT that breaks the maximum likelihood principle (MLP). Thus, the related estimates can be considered suboptimum with respect to the MLP.

Since the performance obtained under the general Gaussian assumption is comparable to that obtained using 5G data distributions, in the next figures we will report only the results obtained under the 5G scenario. Fig. 3 is analogous to Fig. 1 but different values for K_0 are considered. It turns out that the $P_{\rm d}$ increases with K_0 for BBNJ-D-UM and BBNJ-NCD, whereas it decreases with K_0 for the BBNJ-D-CM. For $K_0 = 24$, the curves of BBNJ-D-CM and BBNJ-D-UM are almost overlapped. In all the other cases, the BBNJ-D-CM outperforms both BBNJ-D-UM, BBNJ-NCD, and BBNJ-LVM. Fig. 4 shows the RMSE curves as a function of γ for $K_0 = 8$ and $K_0 = 24$ (again, we do not consider the BBNJ-LVM). It is possible to observe that the effect of the true value of K_0 on the related estimate is more evident for lower values of γ. The BBNJ-D-UM outperforms both the BBNJ-D-CM and BBNJ-NCD, at least when $\gamma > 6$ dB. For $K_0 = 8$ and $\gamma < 6$ dB, the BBNJ-D-CM experiences better performance than the BBNJ-D-UM. In any configurations, the proposed schemes are superior to the naive detector. Finally, in Figs. 5 and 6, we analyze the effect of K on the detection and estimation performance, respectively. These figures corroborate the hierarchy arisen from the previous examples and show that, as expected, increasing K is a blessing and a curse. In fact, high values of K improve the estimation quality of parameters such as the mean and the covariance matrix (and, hence, the detection performance), but, at the same time, extend the range of possible values for K_0 yielding more uncertainty and computational complexity.

As final remark, it is important to underline that the BBNJ-D-CM is designed assuming the most general structure

for the covariance matrix and its detection performance is better than those provided by the BBNJ-D-UM under the 5G data, even though the data was uncorrelated. Therefore, we would single out the BBNJ-D-CM as the best architecture in terms of detection probability. Nevertheless, the BBNJ-D-UM shows better estimation performance due to the fact that the BBNJ-D-CM is an approximation of the GLRT.

C. Performance of the Spoofing Detection Architectures

This last subsection deals with the performance assessment of the AIC/BIC/GIC-based SP-D-UM and SP-D-CM also in comparison with the counterparts mentioned at the beginning of this section and derived in the supplemental material. In the following, the scenario is generated according to the 5G model only because the performances under the general Gaussian model are rather similar (at least for the considered parameters) and, hence, are not reported here for brevity; in addition, we set $K = 32$, $K_0 = 16$, and evaluate the performance metrics by varying the parameter ν, which represents the level of fake information (see Subsection V-A).

In Fig. 7, we show the detection probability versus ν. It turns out that all the considered detectors share the same detection performance except for the SP-LVM that returns the lowest $P_{\rm d}$ values confirming what observed for the BBNJ attacks. Notice also that there exists a first floor of 0.65 when 0.2 dB $< \nu <$ 2 dB, which depends on the fact that at each Monte Carlo trial we counterfeit a random number of components. Specifically, we have verified that when all the three components are spoofed, a spoofer attack is detected even with small value of ν, i.e., $\nu \leq 1$ dB, while when one or two components are spoofed, the attack is detected only for higher values of ν, e.g., $\nu \geq 2$ dB.

The RMSE values for the estimate of K_0 are provided in Fig. 8, where it can be observed that the RMSE curves are bounded from above by 1.1 and for $\nu > 5$ the errors are lower than 1 and experience a limited variability. Moreover, a local maximum for the the RMSE around $\nu = 3$ dB is present. Indeed, when $\nu \leq 2$ dB, the error only accounts for the cases where all the components are spoofed (i.e., they are the only detected ones), while for $\nu \geq 2$ dB also the cases where less than three spoofed components are detected. Therefore, the error, when all the components are spoofed, is lower than the error obtained when fewer than three components are spoofed for low values of ν. Then, as ν takes on high value, as expected, the estimation error decreases.

It is important to underline here that even though all the considered detectors (excluding the SP-LVM) share almost the same detection and estimation performance, the competitors are not capable of identifying the spoofed components unlike the proposed decision schemes. Remarkably, such an information can be suitably used to mitigate the effects of the attack on the position estimation. For this reason, in Fig. 9, we investigate the probability of correct classification of the spoofed components defined as the percentage of trials where the spoofed components among the N data components are correctly detected. It turns out that AIC/BIC/GIC-based SP-D-UM and SP-D-CM can provide a probability of correct

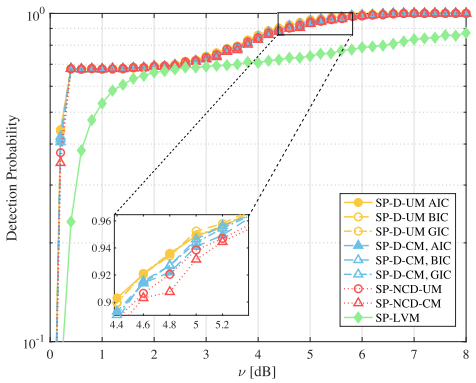


Fig. 7. $P_{\rm d}$ versus ν assuming $K = 32$, $K_0 = 16$, and the 5G scenario (spoofing detectors).

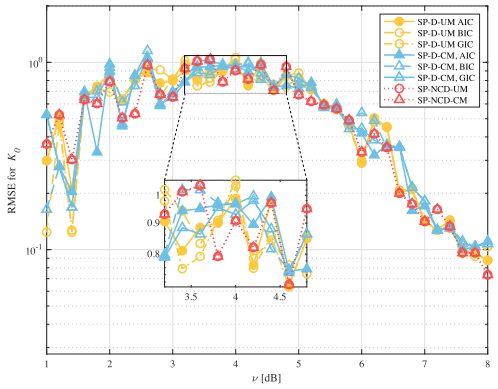


Fig. 8. RMSE versus ν assuming $K = 32$, $K_0 = 16$, and the 5G scenario (spoofing detectors).

classification greater than 0.8 for $\nu \geq 1$ dB. Additional results for different values of K_0 and K, not reported here for brevity, confirm the observed behavior.

Summarizing, in the presence of a spoofing attack, the proposed detectors share almost the same performance as the naive detectors, while the SP-LVM experiences the worst detection performance. However, it is important to stress again that, unlike the considered competitors, the SP-D-CM and SP-D-UM are capable of establishing which data entry is corrupted. Such an information becomes of primary importance to counteract the spoofing attacks.

Two concluding remarks are now in order. It is worth recalling that the attacks considered here lead to an abrupt change in the measurements of interest (and, hence, are easy to be implemented). Thus, in the presence of smart attacks that smoothly modify the information of interest, the proposed algorithms might fail. Facing this kind of attack requires to suitably re-formulate the detection problems and to derive new detection architectures. Finally, it is important to observe that the proposed methods are rather general from the application point of view. In fact, another important application can be the network security, where jammers force the user to disconnect from the legitimate base station and to connect to a rogue base station.

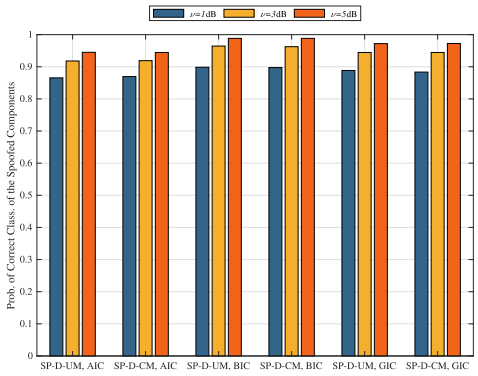


Fig. 9. Probability of correct classification of the spoofed components assuming $K = 32$, $K_0 = 16$, and the 5G scenario (spoofing detectors).

VI. Conclusion

In this paper, we have devised innovative BBNJ and spoofing detection architectures fed by high-level data in wireless networks. At the design stage, we framed this problem in the context of detection theory and formulate it in terms of a hypothesis test where under the null hypothesis, data are homogeneous, while under the alternative (possibly multiple) hypothesis, only one abrupt variation occurs in data under test. The newly proposed algorithms, which represent the main technical novelty of this work, have been obtained by resorting to GLRT-based design procedures that include penalized tests. Due to mathematical and/or computational issues, we used suitable approximations to come up with closed-form expressions of practical value. The performance of the proposed detectors has been assessed in a location security case study and in comparison with natural counterparts. Besides the analysis under nominal conditions, we simulated a more realistic scenario adhering to the 5G environment to measure the deviation from the nominal behavior. In both cases, the analysis highlighted the superiority of the proposed detectors over the competitors at least for the considered parameters. Remarkably, the proposed spoofing detectors are capable of identifying with high probability the counterfeit components and such an information can be exploited to counter this kind of attack.

Future research tracks might include the design of architectures that can deal with smart attacks that slowly modify the information of interest. In this context, the attack classification could be useful to draw a complete picture of the operating scenario.

Appendix A
Derivation of BBNJ-D-UM (5)

Let us focus on the left-hand side of (4) and, for computational convenience, consider the logarithm of it. Thus, neglecting the irrelevant constants, the log-likelihood function under H_0 can be recast as

$$\mathcal{L}_0(\boldsymbol{m}_0, \boldsymbol{\Sigma}_0; \boldsymbol{Z}) \approx -\frac{K}{2}\sum_{n=1}^{N} \log \sigma_{0,n}^2 - \frac{1}{2}\sum_{k=1}^{K}\sum_{n=1}^{N}\frac{(z_{k,n}-m_{0,n})^2}{\sigma_{0,n}^2}, \quad (12)$$

where recall that $m_{0,n}$ and $z_{k,n}$, $n = 1, \ldots, N$, are the entries of $\boldsymbol{m}_0$ and $\boldsymbol{z}_k$, respectively. It is not difficult to show that the MLEs of $m_{0,n}$ and $\sigma_{0,n}^2$, $n = 1, \ldots, N$, are given by

$$\widehat{m}_{0,n} = \frac{1}{K}\sum_{k=1}^{K} z_{k,n} \text{ and } \widehat{\sigma}_{0,n}^2 = \frac{1}{K}\sum_{k=1}^{K}(z_{k,n} - \widehat{m}_{0,n})^2, \quad (13)$$

respectively. As for the maximization of the log-likelihood under H_1, notice that, neglecting the irrelevant constants, it can be written as

$$\begin{aligned}
&\mathcal{L}_1(\boldsymbol{m}_1, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, K_0; \boldsymbol{Z}) \\
&\approx -\frac{K_0}{2}\sum_{n=1}^{N}\log\sigma_{1,n}^2 \\
&\quad -\frac{K_1}{2}\sum_{n=1}^{N}\log(\sigma_{1,n}^2 + \Delta\sigma_n^2) - \frac{1}{2}\sum_{k=1}^{K_0}\sum_{n=1}^{N}\frac{(z_{k,n}-m_{1,n})^2}{\sigma_{1,n}^2} \\
&\quad -\frac{1}{2}\sum_{k=K_0+1}^{K}\sum_{n=1}^{N}\frac{(z_{k,n}-m_{1,n})^2}{\sigma_{1,n}^2 + \Delta\sigma_n^2}, \quad (14)
\end{aligned}$$

where $m_{1,n}$, $n = 1, \ldots, N$, are the components of $\boldsymbol{m}_1$. Before proceeding with the search of the stationary points of the above function, let us notice that $\lim_{\|\boldsymbol{m}\|\to+\infty}\mathcal{L}_1(\boldsymbol{m}_1, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, K_0; \boldsymbol{Z}) = -\infty$, $\lim_{\sigma_{1,n}^2\to+\infty}\mathcal{L}_1(\boldsymbol{m}_1, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, K_0; \boldsymbol{Z}) = -\infty$, $n = 1, \ldots, N$, $\lim_{\sigma_{1,n}^2\to 0}\mathcal{L}_1(\boldsymbol{m}_1, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, K_0; \boldsymbol{Z}) = -\infty$, $n = 1, \ldots, N$, $\lim_{\Delta\sigma_n^2\to+\infty}\mathcal{L}_1(\boldsymbol{m}_1, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, K_0; \boldsymbol{Z}) = -\infty$, $n = 1, \ldots, N$, and $\lim_{\Delta\sigma_n^2\to 0}\mathcal{L}_1(\boldsymbol{m}_1, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, K_0; \boldsymbol{Z}) = C \in \mathbb{R}$, $n = 1, \ldots, N$. Thus, we can search the stationary points in the interior of the log-likelihood domain. Let us start from the maximization with respect to $\Delta\sigma_n^2$, $n = 1, \ldots, N$. The maximizers are the solutions of $\partial/\partial\Delta\sigma_n^2[\mathcal{L}_1(\boldsymbol{m}_1, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, K_0; \boldsymbol{Z})] = 0$, $n = 1, \ldots, N$, and are given by[7]

$$\widehat{\Delta\sigma_n^2}(\sigma_{1,n}^2, m_{1,n}, K_0) = \begin{cases} \dfrac{1}{K_1}\displaystyle\sum_{k=K_0+1}^{K}(z_{k,n} - m_{1,n})^2 - \sigma_{1,n}^2, \\ \text{if } \widehat{\Delta\sigma^2}_n > 0, \\ 0, \text{ otherwise}, \end{cases} \quad (15)$$

$n = 1, \ldots, N$. Given the set $\mathcal{B} = \{K_0, \boldsymbol{\sigma}_1^2, \boldsymbol{m}_1\}$ with $\boldsymbol{\sigma}_1^2 = [\sigma_{1,1}^2, \ldots, \sigma_{1,N}^2]^T$, we define

$$\Gamma(\mathcal{B}) = \{n \in \Gamma_N : \widehat{\Delta\sigma_n^2}(\sigma_{1,n}^2, m_{1,n}, K_0) > 0\}, \quad (16)$$

$$\bar{\Gamma}(\mathcal{B}) = \Gamma_N \setminus \Gamma(\mathcal{B}), \quad (17)$$

and recast the right-hand side of (14) as (neglecting the constants)

$$(14) \approx -\frac{K_0}{2}\sum_{n=1}^{N}\log\sigma_{1,n}^2$$

[7]From the sign of the derivative, it is possible to show that (15) is a maximum point.

$$-\frac{K_1}{2}\sum_{n\in\Gamma(\mathcal{B})}\log\left[\frac{1}{K_1}\sum_{k=K_0+1}^{K}(z_{k,n}-m_{1,n})^2\right]$$
$$-\frac{K_1}{2}\sum_{n\in\bar{\Gamma}(\mathcal{B})}\log\sigma_{1,n}^2-\frac{1}{2}\sum_{n=1}^{N}\sum_{k=1}^{K_0}\frac{(z_{k,n}-m_{1,n})^2}{\sigma_{1,n}^2}$$
$$-\frac{1}{2}K_1|\Gamma(\mathcal{B})|-\frac{1}{2}\sum_{n\in\bar{\Gamma}(\mathcal{B})}\sum_{k=K_0+1}^{K}\frac{(z_{k,n}-m_{1,n})^2}{\sigma_{1,n}^2}. \tag{18}$$

Let us proceed by setting to zero the first derivative of (18) with respect to $\sigma_{1,n}^2$, $n\in\Gamma_N$, to obtain

$$\widehat{\sigma}_{1,n}^2(m_{1,n},K_0)=\frac{1}{K_0}\sum_{k=1}^{K_0}(z_{k,n}-m_{1,n})^2,\ n\in\Gamma(\mathcal{B}),$$
$$\widehat{\sigma}_{1,n}^2(m_{1,n},K_0)=\frac{1}{K}\sum_{k=1}^{K}(z_{k,n}-m_{1,n})^2,\ n\in\bar{\Gamma}(\mathcal{B}). \tag{19}$$

It follows that the compressed log-likelihood function can be approximated as

$$\mathcal{L}_1(\boldsymbol{m}_1,\widehat{\boldsymbol{\Sigma}}_1,\widehat{\boldsymbol{\Sigma}}_2,K_0;\boldsymbol{Z})$$
$$\approx-\frac{K_0}{2}\sum_{n\in\Gamma(\mathcal{B})}\log\left[\frac{1}{K_0}\sum_{k=1}^{K_0}(z_{k,n}-m_{1,n})^2\right]$$
$$-\frac{K_1}{2}\sum_{n\in\Gamma(\mathcal{B})}\log\left[\frac{1}{K_1}\sum_{k=K_0+1}^{K}(z_{k,n}-m_{1,n})^2\right]$$
$$-\frac{K}{2}\sum_{n\in\bar{\Gamma}(\mathcal{B})}\log\left[\frac{1}{K}\sum_{k=1}^{K}(z_{k,n}-m_{1,n})^2\right]. \tag{20}$$

It still remains to maximize the log-likelihood function with respect to $\boldsymbol{m}_1$ and K_0. The latter problem can be carried out by means of a 1-dimensional grid search, whereas the former problem can be solved by finding the solutions of

$$\frac{\partial}{\partial m_{1,n}}[\mathcal{L}_1(\boldsymbol{m}_1,\widehat{\boldsymbol{\Sigma}}_1,\widehat{\boldsymbol{\Sigma}}_2,K_0;\boldsymbol{Z})]=0,\ n\in\Gamma_N. \tag{21}$$

Observe that when $n\in\bar{\Gamma}(\mathcal{B})$, the final result is $\widehat{m}_{0,n}$ given by (13). On the other hand, in the case that $n\in\Gamma(\mathcal{B})$, we obtain

$$\frac{1}{K_1}\sum_{k=K_0+1}^{K}(z_{k,n}-m_{1,n})^2\left[\sum_{k=1}^{K_0}z_{k,n}-K_0m_{1,n}\right]$$
$$+\frac{1}{K_0}\sum_{k=1}^{K_0}(z_{k,n}-m_{1,n})^2\left[\sum_{k=K_0+1}^{K}z_{k,n}-K_1m_{1,n}\right]=0. \tag{22}$$

Now, let us define the following quantities $A_{K_0,n}=\sum_{k=1}^{K_0}z_{k,n}$, $B_{K_0,n}=\sum_{k=1}^{K_0}z_{k,n}^2$, $A_{K_1,n}=\sum_{k=K_0+1}^{K}z_{k,n}$, and $B_{K_1,n}=\sum_{k=K_0+1}^{K}z_{k,n}^2$, then, (22) can be recast as

$$C_3m_{1,n}^3+C_2m_{1,n}^2+C_1m_{1,n}+C_0=0, \tag{23}$$

where $C_0=\frac{B_{K_1,n}A_{K_0,n}}{K_1}+\frac{B_{K_0,n}A_{K_1,n}}{K_0}$, $C_1=-\frac{K_0}{K_1}B_{K_1,n}-\frac{K_1}{K_0}B_{K_0,n}-\left(\frac{2}{K_1}+\frac{2}{K_0}\right)A_{K_1,n}A_{K_0,n}$, $C_2=A_{K_0,n}+A_{K_1,n}+2\frac{K_0}{K_1}A_{K_1,n}+2\frac{K_1}{K_0}A_{K_0,n}$, and $C_3=-(K_0+K_1)$. The solutions of the above equations can be explicitly obtained resorting to Cardano's method [49] and, then, we choose that one, $\widehat{m}_1$ say, leading to the maximum of $\mathcal{L}_1(\boldsymbol{m}_1,\widehat{\boldsymbol{\Sigma}}_1,\widehat{\boldsymbol{\Sigma}}_2,K_0;\boldsymbol{Z})$, for all admissible values of $K_0\in\Omega$.[8]

Finally, notice that $\Gamma(\mathcal{B})$ is not known and, hence, in principle, the above procedure should be repeated for each K_0 and each partition of Γ_N. Then, we can select the combination of estimates returning the maximum value for the log-likelihood. However, this approach is prohibitive from a computational point of view. For this reason, we pursue an alternative strategy that is suboptimum. To be more precise, given a generic index n and K_0, we proceed by assuming that $n\in\Gamma(\mathcal{B})$ and by estimating the unknown parameters indexed by n accordingly. Then, we evaluate the following inequality (see (15))

$$\frac{1}{K_1}\sum_{k=K_0+1}^{K}(z_{k,n}-\widehat{m}_{1,n})^2-\frac{1}{K_0}\sum_{k=1}^{K_0}(z_{k,n}-\widehat{m}_{1,n})^2>0. \tag{24}$$

If the above inequality is valid, we repeat the above reasoning for the next index $n+1$, otherwise, n is classified as belonging to $\bar{\Gamma}(\mathcal{B})$ and the estimation is performed under this assumption. These steps continue until $n\leq N$ and we denote the obtained sets by $\tilde{\Gamma}(\widehat{\mathcal{B}})$ and $\tilde{\bar{\Gamma}}(\widehat{\mathcal{B}})$ with $\widehat{\mathcal{B}}$ the corresponding estimate of $\mathcal{B}$.

The final expression of the compressed log-likelihood under H_1 (neglecting irrelevant terms) is given by the left-hand side of (20) with the remaining unknown parameters replaced by the respective estimates. Decision statistic of (5) naturally follows.

APPENDIX B
DERIVATION OF BBNJ-D-CM (6)

The optimization problem at the denominator of the left-hand side of (4) is rather conventional and the MLEs of $\boldsymbol{m}_0$ and $\boldsymbol{\Sigma}_0$ are given by $\bar{\boldsymbol{m}}_0=\frac{1}{K}\sum_{k=1}^{K}\boldsymbol{z}_k$ and $\bar{\boldsymbol{\Sigma}}_0=\frac{1}{K}\sum_{k=1}^{K}(\boldsymbol{z}_k-\bar{\boldsymbol{m}}_0)(\boldsymbol{z}_k-\bar{\boldsymbol{m}}_0)^T$ (see, e.g. [45]), respectively.

The situation under H_1 is different. As a matter of fact, the exact expressions for the MLEs of the unknown parameters are not available (at least to the best of authors' knowledge). For this reason, as shown below, we resort to suitable approximations of the partially-compressed log-likelihood function. Therefore, under H_1, the log-likelihood function can be recast (up to constants) as

$$\mathcal{L}_1(\boldsymbol{m}_1,\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2,K_0;\boldsymbol{Z})$$
$$\approx-\frac{K_0}{2}\log\det(\boldsymbol{\Sigma}_1)$$
$$-\frac{K_1}{2}\log\det(\boldsymbol{\Sigma}_1+\boldsymbol{R})$$
$$-\frac{1}{2}\sum_{k=1}^{K_0}(\boldsymbol{z}_k-\boldsymbol{m}_1)^T\boldsymbol{\Sigma}_1^{-1}(\boldsymbol{z}_k-\boldsymbol{m}_1)$$

[8]In the case of uncorrelated measurements, K_0 is not subject to any constraint and, hence, it takes on value in Ω.

$$-\frac{1}{2}\sum_{k=K_0+1}^{K}(\boldsymbol{z}_k-\boldsymbol{m}_1)^T(\boldsymbol{\Sigma}_1+\boldsymbol{R})^{-1}(\boldsymbol{z}_k-\boldsymbol{m}_1), \tag{25}$$

where $\boldsymbol{R}=\boldsymbol{\Sigma}_2-\boldsymbol{\Sigma}_1\succ\boldsymbol{0}$, and let us proceed as follows. Assume that $\boldsymbol{\Sigma}_1$ is known while $\boldsymbol{R}$ is positive definite and unknown, then $\boldsymbol{\Sigma}_2$ is a positive definite matrix such that $\boldsymbol{\Sigma}_2-\boldsymbol{\Sigma}_1$ is an arbitrary positive definite matrix. Thus, we can maximize $\mathcal{L}_1(\boldsymbol{m}_1,\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2;\boldsymbol{Z})$ with respect to $\boldsymbol{\Sigma}_2$ to obtain [45]

$$\bar{\boldsymbol{\Sigma}}_2(\boldsymbol{m}_1,K_0)=\frac{1}{K_1}\sum_{k=K_0+1}^{K}(\boldsymbol{z}_k-\boldsymbol{m}_1)(\boldsymbol{z}_k-\boldsymbol{m}_1)^T. \tag{26}$$

It readily follows that

$$\bar{\boldsymbol{R}}(\boldsymbol{\Sigma}_1,\boldsymbol{m}_1,K_0)=\begin{cases}\bar{\boldsymbol{\Sigma}}_2(\boldsymbol{m}_1,K_0)-\boldsymbol{\Sigma}_1, & \text{if } \bar{\boldsymbol{R}}\succ\boldsymbol{0},\\ \boldsymbol{0}, & \text{otherwise.}\end{cases} \tag{27}$$

In the case that $\bar{\boldsymbol{R}}=\boldsymbol{0}$ (namely, $\boldsymbol{\Sigma}_1=\boldsymbol{\Sigma}_2$), the problem is tantamount to that under H_0. As a consequence, the final statistic is a constant equal to zero.

In the opposite case ($\bar{\boldsymbol{R}}\succ\boldsymbol{0}$), we proceed by maximizing

$$\begin{aligned}&\mathcal{L}_1(\boldsymbol{m}_1,\boldsymbol{\Sigma}_1,\bar{\boldsymbol{\Sigma}}_2,K_0;\boldsymbol{Z})\\&\quad\approx-\frac{K_0}{2}\log\det(\boldsymbol{\Sigma}_1)-\frac{1}{2}K_1N\\&\qquad-\frac{K_1}{2}\log\det(\bar{\boldsymbol{\Sigma}}_2(\boldsymbol{m}_1,K_0))\\&\qquad-\frac{1}{2}\sum_{k=1}^{K_0}\|\boldsymbol{\Sigma}_1^{-1/2}(\boldsymbol{z}_k-\boldsymbol{m}_1)\|^2,\end{aligned} \tag{28}$$

with respect to $\boldsymbol{\Sigma}_1$ to obtain (see [45])

$$\bar{\boldsymbol{\Sigma}}_1(\boldsymbol{m}_1,K_0)=\frac{1}{K_0}\sum_{k=1}^{K_0}(\boldsymbol{z}_k-\boldsymbol{m}_1)(\boldsymbol{z}_k-\boldsymbol{m}_1)^T. \tag{29}$$

Replacing the above estimate into (28) yields (up to irrelevant constants)

$$\begin{aligned}&\mathcal{L}_1(\boldsymbol{m}_1,\bar{\boldsymbol{\Sigma}}_1,\bar{\boldsymbol{\Sigma}}_2,K_0;\boldsymbol{Z})\\&\quad\approx-\frac{K_0}{2}\log\det\left[\frac{1}{K_0}\sum_{k=1}^{K_0}(\boldsymbol{z}_k-\boldsymbol{m}_1)(\boldsymbol{z}_k-\boldsymbol{m}_1)^T\right]\\&\qquad-\frac{K_1}{2}\log\det\left[\frac{1}{K_1}\sum_{k=K_0+1}^{K}(\boldsymbol{z}_k-\boldsymbol{m}_1)(\boldsymbol{z}_k-\boldsymbol{m}_1)^T\right].\end{aligned} \tag{30}$$

Now, let us focus on the terms depending on $\boldsymbol{m}_1$ and define

$$\begin{aligned}&g(\boldsymbol{m}_1,K_0)\\&\quad=-\frac{\log\det\left(\boldsymbol{S}_0-\boldsymbol{s}_0\boldsymbol{m}_1^T-\boldsymbol{m}_1\boldsymbol{s}_0^T+K_0\boldsymbol{m}_1\boldsymbol{m}_1^T\right)}{2/K_0}\\&\qquad-\frac{\log\det\left(\boldsymbol{S}_1-\boldsymbol{s}_1\boldsymbol{m}_1^T-\boldsymbol{m}_1\boldsymbol{s}_1^T+K_1\boldsymbol{m}_1\boldsymbol{m}_1^T\right)}{2/K_1}\end{aligned} \tag{31}$$

where $\boldsymbol{S}_0=\sum_{k=1}^{K_0}\boldsymbol{z}_k\boldsymbol{z}_k^T$, $\boldsymbol{S}_1=\sum_{k=K_0+1}^{K}\boldsymbol{z}_k\boldsymbol{z}_k^T$, $\boldsymbol{s}_0=\sum_{k=1}^{K_0}\boldsymbol{z}_k$, and $\boldsymbol{s}_1=\sum_{k=K_0+1}^{K}\boldsymbol{z}_k$. Completing the quadratic forms, we can write

$$(31)=-\frac{\log\det\left(\boldsymbol{M}_0+\boldsymbol{u}_0\boldsymbol{u}_0^T\right)}{2/K_0}-\frac{\log\det\left(\boldsymbol{M}_1+\boldsymbol{u}_1\boldsymbol{u}_1^T\right)}{2/K_1}, \tag{32}$$

where $\boldsymbol{M}_0=\boldsymbol{S}_0-(1/K_0)\boldsymbol{s}_0\boldsymbol{s}_0^T$, $\boldsymbol{M}_1=\boldsymbol{S}_1-(1/K_1)\boldsymbol{s}_1\boldsymbol{s}_1^T$, $\boldsymbol{u}_0=(1/\sqrt{K_0})\boldsymbol{s}_0-\sqrt{K_0}\boldsymbol{m}_1$, and $\boldsymbol{u}_1=(1/\sqrt{K_1})\boldsymbol{s}_1-\sqrt{K_1}\boldsymbol{m}_1$. Now, since $\boldsymbol{M}_0=(\boldsymbol{Z}_{1:K_0}-(1/K_0)\boldsymbol{s}_0\boldsymbol{1}^T)(\boldsymbol{Z}_{1:K_0}-(1/K_0)\boldsymbol{s}_0\boldsymbol{1}^T)^T$, $\boldsymbol{M}_1=(\boldsymbol{Z}_{K_0+1:K}-(1/K_1)\boldsymbol{s}_1\boldsymbol{1}^T)(\boldsymbol{Z}_{K_0+1:K}-(1/K_1)\boldsymbol{s}_1\boldsymbol{1}^T)^T$, and $\min\{K_0,K_1\}>N$, they are also positive definite with probability 1 [45] and, hence, $g(\boldsymbol{m}_1,K_0)$ becomes

$$\begin{aligned}&g(\boldsymbol{m}_1,K_0)\\&\quad=-\frac{K_0}{2}\log\det(\boldsymbol{M}_0)-\frac{K_1}{2}\log\det(\boldsymbol{M}_1)\\&\qquad-\frac{\log\left(1+\boldsymbol{u}_0^T\boldsymbol{M}_0^{-1}\boldsymbol{u}_0\right)}{2/K_0}-\frac{\log\left(1+\boldsymbol{u}_1^T\boldsymbol{M}_1^{-1}\boldsymbol{u}_1\right)}{2/K_1},\end{aligned} \tag{33}$$

where the last equality comes from the fact [50] that $\forall\boldsymbol{A}\in\mathbb{R}^{N\times M}$, $\boldsymbol{B}\in\mathbb{R}^{M\times N}$: $\det(\boldsymbol{I}+\boldsymbol{AB})=\det(\boldsymbol{I}+\boldsymbol{BA})$. The most right-hand side of (33) clearly unveils the radially unbounded nature of $g(\boldsymbol{m}_1,K_0)$ with respect to $\boldsymbol{m}_1$, thus we can search the maximum points in the interior of its domain. However, solving the following system of equations $\partial/\partial\boldsymbol{m}_1[g(\boldsymbol{m}_1,K_0)]=\boldsymbol{0}$ is not straightforward and can lead to additional complexity. For this reason, we resort to the approximation $\log(1+x)\approx x$ to obtain $\tilde{g}(\boldsymbol{m}_1,K_0)=-\frac{K_0}{2}\log\det(\boldsymbol{M}_0)-\frac{K_1}{2}\log\det(\boldsymbol{M}_1)-\frac{K_0}{2}\boldsymbol{u}_0^T\boldsymbol{M}_0^{-1}\boldsymbol{u}_0-\frac{K_1}{2}\boldsymbol{u}_1^T\boldsymbol{M}_1^{-1}\boldsymbol{u}_1$, and solve $\partial/\partial\boldsymbol{m}_1[\tilde{g}(\boldsymbol{m}_1,K_0)]=\boldsymbol{0}$. The above equation can be recast as $K_0\boldsymbol{M}_0^{-1}(\boldsymbol{s}_0-K_0\boldsymbol{m}_1)+K_1\boldsymbol{M}_1^{-1}(\boldsymbol{s}_1-K_1\boldsymbol{m}_1)=\boldsymbol{0}$ and, hence,

$$\bar{\boldsymbol{m}}_1=\left(K_0^2\boldsymbol{M}_0^{-1}+K_1^2\boldsymbol{M}_1^{-1}\right)^{-1}\left(K_0\boldsymbol{M}_0^{-1}\boldsymbol{s}_0+K_1\boldsymbol{M}_1^{-1}\boldsymbol{s}_1\right). \tag{34}$$

It is also straightforward to show that the Hessian of $\tilde{g}(\boldsymbol{m}_1,K_0)$ with respect to $\boldsymbol{m}_1$ is negative definite, in fact, recalling that $\boldsymbol{M}_i^{-1}\succ\boldsymbol{0}$, $i=0,1$, it turns out that

$$\frac{\partial^2}{\partial\boldsymbol{m}_1\partial\boldsymbol{m}_1^T}[\tilde{g}(\boldsymbol{m}_1,K_0)]=-K_0^2\boldsymbol{M}_0^{-1}-K_1^2\boldsymbol{M}_1^{-1}\prec\boldsymbol{0}. \tag{35}$$

The final expression for the approximate partially-compressed log-likelihood is obtained by replacing $\boldsymbol{m}_1$ in (30) with $\bar{\boldsymbol{m}}_1$. Recalling that the compressed log-likelihood under H_0 is (up to irrelevant constants)

$$-\frac{K}{2}\log\det\left[\sum_{k=1}^{K}(\boldsymbol{z}_k-\bar{\boldsymbol{m}}_0)(\boldsymbol{z}_k-\bar{\boldsymbol{m}}_0)^T\right], \tag{36}$$

(6) naturally follows.

APPENDIX C
DERIVATION OF SP-D-UM (10)

The maximization at the denominator of (9) is the same as that of Subsection III-A and the estimates of $\boldsymbol{m}_0$ and $\boldsymbol{\Sigma}_0$ are given by (13). On the other hand, under the generic

$H_{1,i}$, the log-likelihood function can be written (neglecting the irrelevant constants) as

$$\begin{aligned}&\mathcal{L}_{1,i}(\boldsymbol{m}_1, \boldsymbol{m}_{\Gamma_i}, \boldsymbol{\Sigma}_1, K_0; \boldsymbol{Z})\\ &\approx -\frac{K}{2}\sum_{n=1}^{N}\log\sigma_{1,n}^2 - \frac{1}{2}\\ &\quad\times\sum_{k=1}^{K_0}\sum_{n\in\Gamma_i}\frac{(z_{k,n}-m_{1,n})^2}{\sigma_{1,n}^2} - \frac{1}{2}\sum_{k=1}^{K}\sum_{n\in\Gamma_N\setminus\Gamma_i}\frac{(z_{k,n}-m_{1,n})^2}{\sigma_{1,n}^2}\\ &\quad-\frac{1}{2}\sum_{k=K_0+1}^{K}\sum_{n\in\Gamma_i}\frac{(z_{k,n}-m_{\Gamma_i,n})^2}{\sigma_{1,n}^2},\end{aligned} \tag{37}$$

where $m_{1,n}$, $n=1,\ldots,N$, is the nth entry of $\boldsymbol{m}_1$ and $m_{\Gamma_i,n}$, $n\in\Gamma_i$, is the nth entry of $\boldsymbol{m}_{\Gamma_i}$ that is different from $m_{1,n}$. It is not difficult to show that the maximization of $\mathcal{L}_{1,i}(\boldsymbol{m}_1, \boldsymbol{m}_{\Gamma_i}, \boldsymbol{\Sigma}_1, K_0; \boldsymbol{Z})$ with respect to $\boldsymbol{m}_1$, $\boldsymbol{m}_{\Gamma_i}$, and $\boldsymbol{\Sigma}_1$, can be conducted by setting to zero the first derivative with respect to the unknown parameters. Therefore, starting from $\boldsymbol{m}_1$ and $\boldsymbol{m}_{\Gamma_i}$, we obtain that

$$\widehat{m}_{1,n} = \begin{cases}\dfrac{1}{K_0}\displaystyle\sum_{k=1}^{K_0} z_{k,n}, & \text{if } n\in\Gamma_i,\\ \dfrac{1}{K}\displaystyle\sum_{k=1}^{K} z_{k,n}, & \text{if } n\in\Gamma_N\setminus\Gamma_i,\end{cases} \tag{38}$$

$$\widehat{m}_{\Gamma_i,n} = \frac{1}{K_1}\sum_{k=K_0+1}^{K} z_{k,n}, \quad n\in\Gamma_i. \tag{39}$$

It follows that the partially-compressed log-likelihood can be written as

$$\begin{aligned}&\mathcal{L}_{1,i}(\widehat{\boldsymbol{m}}_1, \widehat{\boldsymbol{m}}_{\Gamma_i}, \boldsymbol{\Sigma}_1, K_0; \boldsymbol{Z})\\ &\quad\approx -\frac{K}{2}\sum_{n=1}^{N}\log\sigma_{1,n}^2\\ &\qquad-\frac{1}{2}\sum_{k=1}^{K_0}\sum_{n\in\Gamma_i}\frac{(z_{k,n}-\widehat{m}_{1,n})^2}{\sigma_{1,n}^2}\\ &\qquad-\frac{1}{2}\sum_{k=1}^{K}\sum_{n\in\Gamma_N\setminus\Gamma_i}\frac{(z_{k,n}-\widehat{m}_{1,n})^2}{\sigma_{1,n}^2}\\ &\qquad-\frac{1}{2}\sum_{k=K_0+1}^{K}\sum_{n\in\Gamma_i}\frac{(z_{k,n}-\widehat{m}_{\Gamma_i,n})^2}{\sigma_{1,n}^2}.\end{aligned} \tag{40}$$

Setting to zero the first derivative of the above function with respect to $\sigma_{1,n}^2$ leads to

$$\widehat{\sigma^2}_{1,n} = \begin{cases}\dfrac{1}{K}\left[\displaystyle\sum_{k=1}^{K_0}(z_{k,n}-\widehat{m}_{1,n})^2 + \sum_{k=K_0+1}^{K}(z_{k,n}-\widehat{m}_{\Gamma_i,n})^2\right], & \text{if } n\in\Gamma_i,\\ \dfrac{1}{K}\displaystyle\sum_{k=1}^{K}(z_{k,n}-\widehat{m}_{1,n})^2, & \text{if } n\in\Gamma_N\setminus\Gamma_i.\end{cases} \tag{41}$$

The final expression of the log-likelihood function compressed with respect to $\boldsymbol{m}_1$, $\boldsymbol{m}_{\Gamma_i}$, and $\boldsymbol{\Sigma}_1$, is (up to constants independent of K_0 and Γ_i)

$$\begin{aligned}&-\frac{K}{2}\left\{\sum_{n\in\Gamma_N\setminus\Gamma_i}\log\left[\frac{1}{K}\sum_{k=1}^{K}(z_{k,n}-\widehat{m}_{1,n})^2\right]\right.\\ &\left.+\sum_{n\in\Gamma_i}\log\left[\frac{1}{K}\left(\sum_{k=1}^{K_0}(z_{k,n}-\widehat{m}_{1,n})^2+\sum_{k=K_0+1}^{K}(z_{k,n}-\widehat{m}_{\Gamma_i,n})^2\right)\right]\right\}.\end{aligned}$$

Recalling that the compressed log-likelihood function under H_0 is given by (12) with the unknown parameters replaced by (13), it is straightforward to obtain (10).

Appendix D
Derivation of SP-D-CM (11)

First, notice that the maximization under H_0 is the same as that in Appendix B. Thus, we focus on the numerator of $\Lambda_i(\boldsymbol{Z})$ and write the log-likelihood under $H_{1,i}$ (up to irrelevant constants) as

$$\begin{aligned}&\mathcal{L}_{1,i}(\boldsymbol{m}_1, \boldsymbol{m}_{\Gamma_i}, \boldsymbol{\Sigma}_1, K_0; \boldsymbol{Z})\\ &\quad\approx -\frac{K}{2}\log\det(\boldsymbol{\Sigma}_1)\\ &\qquad-\frac{1}{2}\mathrm{Tr}\left\{\boldsymbol{\Sigma}_1^{-1}\left[\sum_{k=1}^{K_0}(\boldsymbol{z}_k-\boldsymbol{m}_1)(\boldsymbol{z}_k-\boldsymbol{m}_1)^T\right.\right.\\ &\qquad\left.\left.+\sum_{k=K_0+1}^{K}(\boldsymbol{z}_k-\boldsymbol{m}_{\Gamma_i})(\boldsymbol{z}_k-\boldsymbol{m}_{\Gamma_i})^T\right]\right\}.\end{aligned} \tag{42}$$

The maximization of the above function with respect to $\boldsymbol{\Sigma}_1$ yields [45]

$$\begin{aligned}&\mathcal{L}_{1,i}(\boldsymbol{m}_1, \boldsymbol{m}_{\Gamma_i}, \widehat{\boldsymbol{\Sigma}}_1, K_0; \boldsymbol{Z})\\ &\quad\approx -\frac{K}{2}\log\det\left[\sum_{k=1}^{K_0}(\boldsymbol{z}_k-\boldsymbol{m}_1)\right.\\ &\qquad\left.\times(\boldsymbol{z}_k-\boldsymbol{m}_1)^T+\sum_{k=K_0+1}^{K}(\boldsymbol{z}_k-\boldsymbol{m}_{\Gamma_i})(\boldsymbol{z}_k-\boldsymbol{m}_{\Gamma_i})^T\right]\\ &\quad= g(\boldsymbol{m}_1, \boldsymbol{m}_{\Gamma_i}, K_0; \boldsymbol{Z}).\end{aligned} \tag{43}$$

Now, for each Γ_i, let us define a permutation matrix, $\boldsymbol{P}_{\Gamma_i}\in\mathbb{R}^{N\times N}$ say, such that the first $|\Gamma_i|$ components of the vector

$$\tilde{\boldsymbol{m}}_{\Gamma_i} = \boldsymbol{P}_{\Gamma_i}\boldsymbol{m}_{\Gamma_i} = \begin{bmatrix}\tilde{\boldsymbol{\mu}}_{\Gamma_i,1}\\ \boldsymbol{0}\end{bmatrix} + \begin{bmatrix}\boldsymbol{0}\\ \tilde{\boldsymbol{\mu}}_{\Gamma_i,2}\end{bmatrix} = \tilde{\boldsymbol{m}}_{\Gamma_i,1} + \tilde{\boldsymbol{m}}_{\Gamma_i,2}$$

are those of $\boldsymbol{m}_{\Gamma_i}$ indexed by Γ_i and form $\tilde{\boldsymbol{\mu}}_{\Gamma_i,1}\in\mathbb{R}^{|\Gamma_i|\times 1}$; in addition, $\tilde{\boldsymbol{\mu}}_{\Gamma_i,2}$ contains the components of $\boldsymbol{m}_{\Gamma_i}$ indexed by $\Gamma_N\setminus\Gamma_i$. Similarly, let $\tilde{\boldsymbol{m}}_1$ be

$$\tilde{\boldsymbol{m}}_1 = \boldsymbol{P}_{\Gamma_i}\boldsymbol{m}_1 = \begin{bmatrix}\tilde{\boldsymbol{\nu}}_{\Gamma_i,1}\\ \boldsymbol{0}\end{bmatrix} + \begin{bmatrix}\boldsymbol{0}\\ \tilde{\boldsymbol{\mu}}_{\Gamma_i,2}\end{bmatrix} = \tilde{\boldsymbol{n}}_{\Gamma_i,1} + \tilde{\boldsymbol{m}}_{\Gamma_i,2},$$

where $\tilde{\boldsymbol{\nu}}_{\Gamma_i,1}\in\mathbb{R}^{|\Gamma_i|\times 1}$ contains the components of $\boldsymbol{m}_1$ indexed by Γ_i. Since $\boldsymbol{P}_{\Gamma_i}$ is an orthogonal matrix and letting $\tilde{\boldsymbol{z}}_k = \boldsymbol{P}_{\Gamma_i}\boldsymbol{z}_k$, we can recast (43) as follows

$$\begin{aligned}&g(\boldsymbol{m}_1, \boldsymbol{m}_{\Gamma_i}, K_0; \boldsymbol{Z})\\ &= -\frac{K}{2}\log\det\left\{\boldsymbol{P}_{\Gamma_i}\left[\sum_{k=1}^{K_0}(\boldsymbol{z}_k-\boldsymbol{m}_1)\right.\right.\end{aligned}$$

$$\times(\boldsymbol{z}_k-\boldsymbol{m}_1)^T+\sum_{k=K_0+1}^{K}(\boldsymbol{z}_k-\boldsymbol{m}_{\Gamma_i})(\boldsymbol{z}_k-\boldsymbol{m}_{\Gamma_i})^T\Bigg]\boldsymbol{P}_{\Gamma_i}^T\Bigg\}$$
$$=-\frac{K}{2}\log\det\Bigg[\sum_{k=1}^{K_0}(\tilde{\boldsymbol{z}}_k-\tilde{\boldsymbol{n}}_{\Gamma_i,1}-\tilde{\boldsymbol{m}}_{\Gamma_i,2})$$
$$\times(\tilde{\boldsymbol{z}}_k-\tilde{\boldsymbol{n}}_{\Gamma_i,1}-\tilde{\boldsymbol{m}}_{\Gamma_i,2})^T+\sum_{k=K_0+1}^{K}(\tilde{\boldsymbol{z}}_k-\tilde{\boldsymbol{m}}_{\Gamma_i,1}-\tilde{\boldsymbol{m}}_{\Gamma_i,2})$$
$$\times(\tilde{\boldsymbol{z}}_k-\tilde{\boldsymbol{m}}_{\Gamma_i,1}-\tilde{\boldsymbol{m}}_{\Gamma_i,2})^T\Bigg]=g(\tilde{\boldsymbol{n}}_{\Gamma_i,1},\tilde{\boldsymbol{m}}_{\Gamma_i,1},\tilde{\boldsymbol{m}}_{\Gamma_i,2},K_0;\boldsymbol{Z}).$$

The joint maximization of $g(\cdot)$ with respect to $\tilde{\boldsymbol{n}}_{\Gamma_i,1}$, $\tilde{\boldsymbol{m}}_{\Gamma_i,1}$, and $\tilde{\boldsymbol{m}}_{\Gamma_i,2}$ represents a difficult task from a mathematical point of view (at least to the best of authors' knowledge). For this reason, we resort to a cyclic optimization procedure giving rise to a nondecreasing sequence of values for $g(\cdot)$ [51]. This procedure consists in repeating the following two steps until a stopping criterion is not satisfied

- assume that $\tilde{\boldsymbol{n}}_{\Gamma_i,1}$ and $\tilde{\boldsymbol{m}}_{\Gamma_i,1}$ are known (and equal to the estimates obtained at the previous cycle) and maximize $g(\cdot)$ with respect to $\tilde{\boldsymbol{m}}_{\Gamma_i,2}$;
- assume that $\tilde{\boldsymbol{m}}_{\Gamma_i,2}$ is known (and equal to the estimate at the previous step) and maximize $g(\cdot)$ with respect to $\tilde{\boldsymbol{n}}_{\Gamma_i,1}$ and $\tilde{\boldsymbol{m}}_{\Gamma_i,1}$.

The entire procedure may terminate when $\Delta g(h)=|g(\widehat{\tilde{\boldsymbol{n}}}_{\Gamma_i,1}^{(h)},\widehat{\tilde{\boldsymbol{m}}}_{\Gamma_i,1}^{(h)},\widehat{\tilde{\boldsymbol{m}}}_{\Gamma_i,2}^{(h)},K_0;\boldsymbol{Z})-g(\widehat{\tilde{\boldsymbol{n}}}_{\Gamma_i,1}^{(h-1)},\widehat{\tilde{\boldsymbol{m}}}_{\Gamma_i,1}^{(h-1)},\widehat{\tilde{\boldsymbol{m}}}_{\Gamma_i,2}^{(h-1)},K_0;\boldsymbol{Z})|/|g(\widehat{\tilde{\boldsymbol{n}}}_{\Gamma_i,1}^{(h)},\widehat{\tilde{\boldsymbol{m}}}_{\Gamma_i,1}^{(h)},\widehat{\tilde{\boldsymbol{m}}}_{\Gamma_i,2}^{(h)},K_0;\boldsymbol{Z})|<\epsilon$, where $\epsilon>0$, $\widehat{\tilde{\boldsymbol{n}}}_{\Gamma_i,1}^{(h)}$, $\widehat{\tilde{\boldsymbol{m}}}_{\Gamma_i,1}^{(h)}$, and $\widehat{\tilde{\boldsymbol{m}}}_{\Gamma_i,2}^{(h)}$ are the estimates of $\tilde{\boldsymbol{n}}_{\Gamma_i,1}$, $\tilde{\boldsymbol{m}}_{\Gamma_i,1}$, and $\tilde{\boldsymbol{m}}_{\Gamma_i,2}$ at the hth step, respectively; or when $h\geq h_{\max}$ where $h_{\max}$ is the maximum allowed number of iterations that could be set to guarantee a good compromise between estimation fidelity and computational load.

Therefore, let us start from the first step and maximize $g(\cdot)$ with respect to $\tilde{\boldsymbol{m}}_{\Gamma_i,2}$ (or, equivalently, to $\tilde{\boldsymbol{\mu}}_{\Gamma_i,2}$) assuming that $\tilde{\boldsymbol{n}}_{\Gamma_i,1}=\widehat{\tilde{\boldsymbol{n}}}_{\Gamma_i,1}^{(h-1)}$ and $\tilde{\boldsymbol{m}}_{\Gamma_i,1}=\widehat{\tilde{\boldsymbol{m}}}_{\Gamma_i,1}^{(h-1)}$. To this end, let $\boldsymbol{x}_k=\tilde{\boldsymbol{z}}_k-\widehat{\tilde{\boldsymbol{n}}}_{\Gamma_i,1}^{(h-1)}$ for $k=1,\ldots,K_0$ and $\boldsymbol{x}_k=\tilde{\boldsymbol{z}}_k-\widehat{\tilde{\boldsymbol{m}}}_{\Gamma_i,1}^{(h-1)}$ for $k=K_0+1,\ldots,K$. Thus, we obtain that[9]

$$g(\widehat{\tilde{\boldsymbol{n}}}_{\Gamma_i,1}^{(h-1)},\widehat{\tilde{\boldsymbol{m}}}_{\Gamma_i,1}^{(h-1)},\tilde{\boldsymbol{m}}_{\Gamma_i,2},K_0;\boldsymbol{Z})=-\frac{K}{2}$$
$$\times\log\det\left[\boldsymbol{X}+\left(\bar{\boldsymbol{x}}-\sqrt{K}\tilde{\boldsymbol{m}}_{\Gamma_i,2}\right)\left(\bar{\boldsymbol{x}}-\sqrt{K}\tilde{\boldsymbol{m}}_{\Gamma_i,2}\right)^T\right],\tag{44}$$

where $\boldsymbol{X}=\sum_{k=1}^{K}\boldsymbol{x}_k\boldsymbol{x}_k^T-\bar{\boldsymbol{x}}\bar{\boldsymbol{x}}^T$ and $\bar{\boldsymbol{x}}=(1/\sqrt{K})\sum_{k=1}^{K}\boldsymbol{x}_k$. Now, observe that

$$\sqrt{K}\tilde{\boldsymbol{m}}_{\Gamma_i,2}=\begin{bmatrix}\boldsymbol{0}\\ \sqrt{K}\boldsymbol{I}\end{bmatrix}\tilde{\boldsymbol{\mu}}_{\Gamma_i,2}=\boldsymbol{H}_2\tilde{\boldsymbol{\mu}}_{\Gamma_i,2},\tag{45}$$

[9]Actually, $\boldsymbol{x}_k$ depends on h, but we omit this dependence in order not to burden the notation.

where $\boldsymbol{H}_2=[\boldsymbol{0}\ \sqrt{K}\boldsymbol{I}]^T\in\mathbb{R}^{N\times(N-|\Gamma_i|)}$. It follows that

$$\underset{\tilde{\boldsymbol{\mu}}_{\Gamma_i,2}}{\operatorname{argmax}}-\frac{K}{2}\log\det\left[\boldsymbol{X}+\left(\bar{\boldsymbol{x}}-\boldsymbol{H}_2\tilde{\boldsymbol{\mu}}_{\Gamma_i,2}\right)\left(\bar{\boldsymbol{x}}-\boldsymbol{H}_2\tilde{\boldsymbol{\mu}}_{\Gamma_i,2}\right)^T\right]$$
$$=\underset{\tilde{\boldsymbol{\mu}}_{\Gamma_i,2}}{\operatorname{argmin}}\frac{K}{2}\log\Big[1+\left(\bar{\boldsymbol{x}}-\boldsymbol{H}_2\tilde{\boldsymbol{\mu}}_{\Gamma_i,2}\right)^T$$
$$\times\boldsymbol{X}^{-1}\left(\bar{\boldsymbol{x}}-\boldsymbol{H}_2\tilde{\boldsymbol{\mu}}_{\Gamma_i,2}\right)\Big]$$
$$=(\boldsymbol{H}_2^T\boldsymbol{X}^{-1}\boldsymbol{H}_2)^{-1}\boldsymbol{H}_2^T\boldsymbol{X}^{-1}\bar{\boldsymbol{x}}=\widehat{\tilde{\boldsymbol{\mu}}}_{\Gamma_i,2}^{(h)},\tag{46}$$

where we have used the fact that $\boldsymbol{X}$ is positive definite with probability 1 since $K>N+1$ and the distribution of the $\boldsymbol{z}_k$s is continuous. The update rule for $\tilde{\boldsymbol{m}}_{\Gamma_i,2}$ is

$$\widehat{\tilde{\boldsymbol{m}}}_{\Gamma_i,2}^{(h)}=\begin{bmatrix}\boldsymbol{0}\\ \widehat{\tilde{\boldsymbol{\mu}}}_{\Gamma_i,2}^{(h)}\end{bmatrix}\tag{47}$$

and notice that it depends on $\widehat{\tilde{\boldsymbol{n}}}_{\Gamma_i,1}^{(h-1)}$ and $\widehat{\tilde{\boldsymbol{m}}}_{\Gamma_i,1}^{(h-1)}$.

Now, we are ready to perform the second step of the procedure that pertains the maximization over $\tilde{\boldsymbol{n}}_{\Gamma_i,1}$ and $\tilde{\boldsymbol{m}}_{\Gamma_i,1}$ assuming that $\tilde{\boldsymbol{m}}_{\Gamma_i,2}=\widehat{\tilde{\boldsymbol{m}}}_{\Gamma_i,2}^{(h)}$. Defining $\boldsymbol{y}_k=\tilde{\boldsymbol{z}}_k-\widehat{\tilde{\boldsymbol{m}}}_{\Gamma_i,2}^{(h)}$, $k=1,\ldots,K$, we can recast $g(\tilde{\boldsymbol{n}}_{\Gamma_i,1},\tilde{\boldsymbol{m}}_{\Gamma_i,1},\widehat{\tilde{\boldsymbol{m}}}_{\Gamma_i,2}^{(h)},K_0;\boldsymbol{Z})$ as

$$\frac{K}{2}\log\det\Bigg[\boldsymbol{Y}+\left(\bar{\boldsymbol{y}}_1-\sqrt{K_0}\tilde{\boldsymbol{n}}_{\Gamma_i,1}\right)\left(\bar{\boldsymbol{y}}_1-\sqrt{K_0}\tilde{\boldsymbol{n}}_{\Gamma_i,1}\right)^T$$
$$+\left(\bar{\boldsymbol{y}}_2-\sqrt{K_1}\tilde{\boldsymbol{m}}_{\Gamma_i,1}\right)\left(\bar{\boldsymbol{y}}_2-\sqrt{K_1}\tilde{\boldsymbol{m}}_{\Gamma_i,1}\right)^T\Bigg],\tag{48}$$

where $\boldsymbol{Y}=\sum_{k=1}^{K}\boldsymbol{y}_k\boldsymbol{y}_k^T-\bar{\boldsymbol{y}}_1\bar{\boldsymbol{y}}_1^T-\bar{\boldsymbol{y}}_2\bar{\boldsymbol{y}}_2^T$, $\bar{\boldsymbol{y}}_1=(1/\sqrt{K_0})\sum_{k=1}^{K_0}\boldsymbol{y}_k$, and $\bar{\boldsymbol{y}}_2=(1/\sqrt{K_1})\sum_{k=K_0+1}^{K}\boldsymbol{y}_k$. Moreover, since

$$\sqrt{K_0}\tilde{\boldsymbol{n}}_{\Gamma_i,1}=\begin{bmatrix}\boldsymbol{I}\\ \boldsymbol{0}\end{bmatrix}\sqrt{K_0}\tilde{\boldsymbol{\nu}}_{\Gamma_i,1}=\boldsymbol{H}_1\tilde{\boldsymbol{\nu}}'_{\Gamma_i,1},\tag{49}$$

$$\sqrt{K_1}\tilde{\boldsymbol{m}}_{\Gamma_i,1}=\begin{bmatrix}\boldsymbol{I}\\ \boldsymbol{0}\end{bmatrix}\sqrt{K_1}\tilde{\boldsymbol{\mu}}_{\Gamma_i,1}=\boldsymbol{H}_1\tilde{\boldsymbol{\mu}}'_{\Gamma_i,1},\tag{50}$$

where $\boldsymbol{H}_1=[\boldsymbol{I}\ \boldsymbol{0}]^T\in\mathbb{R}^{N\times|\Gamma_i|}$, (48) becomes

$$g\left(\tilde{\boldsymbol{n}}_{\Gamma_i,1},\tilde{\boldsymbol{m}}_{\Gamma_i,1},\widehat{\tilde{\boldsymbol{m}}}_{\Gamma_i,2}^{(h)},K_0;\boldsymbol{Z}\right)$$
$$=-\frac{K}{2}\log\det\left[\boldsymbol{Y}+\left(\bar{\boldsymbol{Y}}-\boldsymbol{H}_1\boldsymbol{V}\right)\left(\bar{\boldsymbol{Y}}-\boldsymbol{H}_1\boldsymbol{V}\right)^T\right],\tag{51}$$

where $\bar{\boldsymbol{Y}}=[\bar{\boldsymbol{y}}_1\ \bar{\boldsymbol{y}}_2]\in\mathbb{R}^{N\times2}$ and $\boldsymbol{V}=[\tilde{\boldsymbol{\nu}}'_{\Gamma_i,1}\ \tilde{\boldsymbol{\mu}}'_{\Gamma_i,1}]\in\mathbb{R}^{|\Gamma_i|\times2}$. The constraint $K>N+1$ ensures that $\boldsymbol{Y}$ is positive definite [45] and, hence, maximizing the above function with respect to $\boldsymbol{V}$ is tantamount to $\min_{\boldsymbol{V}}\det\left[\boldsymbol{I}+\left(\bar{\boldsymbol{Y}}-\boldsymbol{H}_1\boldsymbol{V}\right)^T\boldsymbol{Y}^{-1}\left(\bar{\boldsymbol{Y}}-\boldsymbol{H}_1\boldsymbol{V}\right)\right]$. Now, since

$$\det\left[\boldsymbol{I}+\left(\bar{\boldsymbol{Y}}-\boldsymbol{H}_1\boldsymbol{V}\right)^T\boldsymbol{Y}^{-1}\left(\bar{\boldsymbol{Y}}-\boldsymbol{H}_1\boldsymbol{V}\right)\right]$$
$$=\det\Big[\boldsymbol{I}+\bar{\boldsymbol{Y}}^T\boldsymbol{Y}^{-1}\bar{\boldsymbol{Y}}+\left(\boldsymbol{H}_1^T\boldsymbol{Y}^{-1}\bar{\boldsymbol{Y}}-\left(\boldsymbol{H}_1^T\boldsymbol{Y}^{-1}\boldsymbol{H}_1\right)\boldsymbol{V}\right)^T$$
$$\times\left(\boldsymbol{H}_1^T\boldsymbol{Y}^{-1}\boldsymbol{H}_1\right)^{-1}\left(\boldsymbol{H}_1^T\boldsymbol{Y}^{-1}\bar{\boldsymbol{Y}}-\left(\boldsymbol{H}_1^T\boldsymbol{Y}^{-1}\boldsymbol{H}_1\right)\boldsymbol{V}\right)$$

$$-\bar{\boldsymbol{Y}}^T\boldsymbol{Y}^{-1}\boldsymbol{H}_1\left(\boldsymbol{H}_1^T\boldsymbol{Y}^{-1}\boldsymbol{H}_1\right)^{-1}\boldsymbol{H}_1^T\boldsymbol{Y}^{-1}\bar{\boldsymbol{Y}}\Big], \tag{52}$$

it follows that

$$\begin{aligned}\widehat{\boldsymbol{V}} &= \underset{\boldsymbol{V}}{\operatorname{argmin}} \det\left[\boldsymbol{I}+\left(\bar{\boldsymbol{Y}}-\boldsymbol{H}_1\boldsymbol{V}\right)^T\boldsymbol{Y}^{-1}\left(\bar{\boldsymbol{Y}}-\boldsymbol{H}_1\boldsymbol{V}\right)\right] \\ &= \left(\boldsymbol{H}_1^T\boldsymbol{Y}^{-1}\boldsymbol{H}_1\right)^{-1}\boldsymbol{H}_1^T\boldsymbol{Y}^{-1}\bar{\boldsymbol{Y}} = \left[\widehat{\tilde{\boldsymbol{\nu}}}'_{\Gamma_i,1}\ \widehat{\tilde{\boldsymbol{\mu}}}'_{\Gamma_i,1}\right]. \end{aligned} \tag{53}$$

At the end of the procedure, the final estimates of $\boldsymbol{m}_1$ and $\boldsymbol{m}_{\Gamma_i}$ are denoted by $\widehat{\boldsymbol{m}}_1$ and $\widehat{\boldsymbol{m}}_{\Gamma_i}$, respectively, and are used in the compressed log-likelihood under $H_{1,i}$ (up to irrelevant constants), i.e.,

$$\begin{aligned}-\frac{K}{2}\log\det\Bigg[&\sum_{k=1}^{K_0}(\boldsymbol{z}_k-\widehat{\boldsymbol{m}}_1)(\boldsymbol{z}_k-\widehat{\boldsymbol{m}}_1)^T \\ &+\sum_{k=K_0+1}^{K}(\boldsymbol{z}_k-\widehat{\boldsymbol{m}}_{\Gamma_i})(\boldsymbol{z}_k-\widehat{\boldsymbol{m}}_{\Gamma_i})^T\Bigg].\end{aligned} \tag{54}$$

Exploiting the above result in conjunction with the compressed log-likelihood under H_0 leads to the approximation (11).

REFERENCES

[1] M. Lichtman, R. P. Jover, M. Labib, R. Rao, V. Marojevic, and J. H. Reed, "LTE/LTE-A jamming, spoofing, and sniffing: Threat assessment and mitigation," *IEEE Commun. Mag.*, vol. 54, no. 4, pp. 54–61, Apr. 2016.

[2] F. M. Aziz, J. S. Shamma, and G. L. Stüber, "Resilience of LTE networks against smart jamming attacks: Wideband model," in *Proc. IEEE 26th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Aug. 2015, pp. 1344–1348.

[3] R. P. Jover, "Security attacks against the availability of LTE mobility networks: Overview and research directions," in *Proc. 16th Int. Symp. Wireless Pers. Multimedia Commun. (WPMC)*, Jun. 2013, pp. 1–9.

[4] D. Rupprecht, K. Kohls, T. Holz, and C. Popper, "Breaking LTE on layer two," in *Proc. IEEE Symp. Secur. Privacy (SP)*, May 2019, pp. 1121–1136.

[5] S. F. Mjølsnes and R. F. Olimid, "Easy 4G/LTE IMSI catchers for non-programmers," in *Computer Network Security*, J. Rak, J. Bay, I. Kotenko, L. Popyack, V. Skormin, and K. Szczypiorski, Eds. Cham, Switzerland: Springer, 2017, pp. 235–246.

[6] R. M. Rao, S. Ha, V. Marojevic, and J. H. Reed, "LTE PHY layer vulnerability analysis and testing using open-source SDR tools," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Oct. 2017, pp. 744–749.

[7] R. Borgaonkar, L. Hirschi, S. Park, and A. Shaik, "New privacy threat on 3G, 4G, and upcoming 5G AKA protocols," *Privacy Enhancing Technol.*, vol. 2019, no. 3, pp. 108–127, 2019.

[8] A. Shaik, R. Borgaonkar, S. Park, and J.-P. Seifert, "New vulnerabilities in 4G and 5G cellular access network protocols: Exposing device capabilities," in *Proc. 12th Conf. Secur. Privacy Wireless Mobile Netw.*, New York, NY, USA, 2019, pp. 221–231.

[9] A. Shaik, R. Borgaonkar, N. Asokan, V. Niemi, and J.-P. Seifert, "Practical attacks against privacy and availability in 4G/LTE mobile communication systems," in *Proc. 23rd Annu. Netw. Distrib. Syst. Secur. Symp. (NDSS)*, 2016, pp. 1–16.

[10] I. Palamà, F. Gringoli, G. Bianchi, and N. B. Melazzi, "The diverse and variegated reactions of different cellular devices to IMSI catching attacks," in *Proc. 14th Int. Workshop Wireless Netw. Testbeds, Experim. Eval. Characterization*, New York, NY, USA, Sep. 2020, pp. 80–86.

[11] R. P. Jover, "LTE security, protocol exploits and location tracking experimentation with low-cost software radio," 2016, *arXiv:1607.05171*.

[12] S. R. Hussain, O. Chowdhury, S. Mehnaz, and E. Bertino, "LTEInspector: A systematic approach for adversarial testing of 4G LTE," in *Proc. Netw. Distrib. Syst. Secur. Symp.*, 2018, pp. 1–15.

[13] C. Yu, S. Chen, Z. Cai, and J. Díaz-Verdejo, "LTE phone number catcher: A practical attack against mobile privacy," *Secur. Commun. Netw.*, vol. 2019, Sep. 2019, Art. no. 7425235.

[14] *Study on 5G Security Enhancements Against False Base Stations*, document ETSI 3GPP TR 33.809 V0.11.0 (2020-10), 3GPP, Oct. 2020.

[15] *Study on NR Positioning Support*, document ETSI TR 138 855 V16.0.0, 3GPP, Dec. 2019.

[16] R. Poisel, *Modern Communications Jamming Principles and Techniques* (Artech House Intelligence and Information Operations Series). Norwood, MA, USA: Artech House, 2011.

[17] D. Margaria, B. Motella, M. Anghileri, J.-J. Floch, I. Fernandez-Hernandez, and M. Paonni, "Signal structure-based authentication for civil GNSSs: Recent solutions and perspectives," *IEEE Signal Process. Mag.*, vol. 34, no. 5, pp. 27–37, Sep. 2017.

[18] L. Heng, J. J. Makela, A. D. Domínguez-García, R. B. Bobba, W. H. Sanders, and G. X. Gao, "Reliable GPS-based timing for power systems: A multi-layered multi-receiver architecture," in *Proc. Power Energy Conf. at Illinois (PECI)*, Champaign, IL, USA, Feb. 2014, pp. 1–7.

[19] R. Morales-Ferre, P. Richter, E. Falletti, A. de la Fuente, and E. S. Lohan, "A survey on coping with intentional interference in satellite navigation for manned and unmanned aircraft," *IEEE Commun. Surveys Tuts.*, vol. 22, pp. 249–291, 2020.

[20] D. Orlando, I. Palama, S. Bartoletti, G. Bianchi, and N. B. Melazzi, "Design and experimental assessment of detection schemes for air interface attacks in adverse scenarios," *IEEE Wireless Commun. Lett.*, vol. 10, no. 9, pp. 1989–1993, Sep. 2021.

[21] M. Lichtman, T. Czauski, S. Ha, P. David, and J. H. Reed, "Detection and mitigation of uplink control channel jamming in LTE," in *Proc. IEEE Mil. Commun. Conf.*, Oct. 2014, pp. 1187–1194.

[22] R. Di Pietro and G. Oligeri, "Jamming mitigation in cognitive radio networks," *IEEE Netw.*, vol. 27, no. 3, pp. 10–15, May/Jun. 2013.

[23] F. M. Aziz, J. S. Shamma, and G. L. Stuber, "Jammer-type estimation in LTE with a smart jammer repeated game," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 7422–7431, Aug. 2017.

[24] K. Firouzbakht, G. Noubir, and M. Salehi, "On the performance of adaptive packetized wireless communication links under jamming," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3481–3495, Jul. 2014.

[25] O. A. Topal, S. Gecgel, E. M. Eksioglu, and G. Karabulut Kurt, "Identification of smart jammers: Learning-based approaches using wavelet preprocessing," *Phys. Commun.*, vol. 39, Apr. 2020, Art. no. 101029.

[26] J. Vinogradova, E. Björnson, and E. G. Larsson, "Detection and mitigation of jamming attacks in massive MIMO systems using random matrix theory," in *Proc. IEEE 17th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2016, pp. 1–5.

[27] H. L. Van Trees, *Optimum Array Processing (Detection, Estimation, and Modulation Theory, Part IV)*. Hoboken, NJ, USA: Wiley, 2002.

[28] M. Liyanage, I. Ahmad, A. B. Abro, A. Gurtov, and M. Ylianttila, *A Comprehensive Guide to 5G Security*, 1st ed. Hoboken, NJ, USA: Wiley, 2017.

[29] Y. Arjoune and S. Faruque, "Smart jamming attacks in 5G new radio: A review," in *Proc. 10th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2020, pp. 1010–1015.

[30] Y. Chen, W. Trappe, and R. P. Martin, "Detecting and localizing wireless spoofing attacks," in *Proc. 4th Annu. IEEE Commun. Soc. Conf. Sensor, Mesh Ad Hoc Commun. Netw.*, Jun. 2007, pp. 193–202.

[31] J. Yang, Y. Chen, and W. Trappe, "Detecting spoofing attacks in mobile wireless environments," in *Proc. 6th Annu. IEEE Commun. Soc. Conf. Sensor, Mesh Ad Hoc Commun. Netw.*, Jun. 2009, pp. 1–9.

[32] P. Jokar, N. Arianpoo, and V. C. M. Leung, "Spoofing detection in IEEE 802.15.4 networks based on received signal strength," *Ad Hoc Netw.*, vol. 11, no. 8, pp. 2648–2660, Nov. 2013.

[33] L. Xiao, Y. Li, G. Han, G. Liu, and W. Zhuang, "PHY-layer spoofing detection with reinforcement learning in wireless networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10037–10047, Dec. 2016.

[34] H. Akhlaghpasand, S. M. Razavizadeh, E. Björnson, and T. T. Do, "Jamming detection in massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 2, pp. 242–245, Apr. 2018.

[35] W. Xu, C. Yuan, S. Xu, H. Q. Ngo, and W. Xiang, "On pilot spoofing attack in massive MIMO systems: Detection and countermeasure," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 1396–1409, 2021.

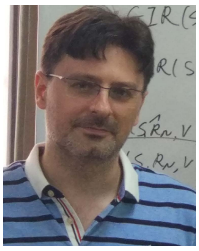
[36] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, vol. 2, P. Hall, Ed. Upper Saddle River, NJ, USA: Prentice-Hall, 1998.

[37] M. Basseville and I. Nikiforov, *Detection of Abrupt Changes: Theory and Application*, (Information and System Sciences Series). Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

[38] C. Truong, L. Oudre, and N. Vayatis, "Selective review of offline change point detection methods," *Signal Process.*, vol. 167, Feb. 2020, Art. no. 107299.

[39] J. H. Sullivan and W. H. Woodall, "Change-point detection of mean vector or covariance matrix shifts using multivariate individual observations," *IIE Trans.*, vol. 32, no. 6, pp. 537–549, Jun. 2000.

[40] A. Sen and M. S. Srivastava, "On tests for detecting change in mean," *Ann. Statist.*, vol. 3, no. 1, pp. 98–108, Jan. 1975.
[41] P. Addabbo, S. Han, F. Biondi, G. Giunta, and D. Orlando, "Adaptive radar detection in the presence of multiple alternative hypotheses using Kullback–Leibler information criterion—Part I: Detector designs," *IEEE Trans. Signal Process.*, vol. 69, pp. 3730–3741, 2021.
[42] S. Dwivedi *et al.*, "Positioning in 5G networks," *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 38–44, Dec. 2021.
[43] S. Bartoletti *et al.*, "Positioning and sensing for vehicular safety applications in 5G and beyond," *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 15–21, Nov. 2021.
[44] I. Lapin, G. Seco-Granados, O. Renaudin, F. Zanier, and L. Ries, "Joint delay and phase discriminator based on ESPRIT for 5G NR positioning," *IEEE Access*, vol. 9, pp. 126550–126563, 2021.
[45] L. J. Gleser, "Aspects of multivariate statistical theory," *Technometrics*, vol. 26, no. 2, pp. 191–192, May 1984.
[46] P. Stoica and Y. Selen, "Model-order selection: A review of information criterion rules," *IEEE Signal Process. Mag.*, vol. 21, no. 4, pp. 36–47, Jul. 2004.
[47] J. A. del Peral-Rosado *et al.*, "Physical-layer abstraction for hybrid GNSS and 5G positioning evaluations," in *Proc. IEEE 90th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2019, pp. 1–6.
[48] E. Rastorgueva-Foi, M. Costa, M. Koivisto, K. Leppanen, and M. Valkama, "User positioning in mmW 5G networks using beam-RSRP measurements and Kalman filtering," in *Proc. 21st Int. Conf. Inf. Fusion (FUSION)*, Jul. 2018, pp. 1–7.
[49] L. Childs, *A Concrete Introduction to Higher Algebra* (Undergraduate Texts in Mathematics). New York, NY, USA: Springer 2008.
[50] H. Lütkepohl, *Handbook Matrices*. Hoboken, NJ, USA: Wiley, 1997.
[51] P. Stoica and Y. Selén, "Cyclic minimizers, majorization techniques, and the expectation-maximization algorithm: A refresher," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 112–114, Jan. 2004.



Danilo Orlando (Senior Member, IEEE) was born in Gagliano del Capo, Italy, in August 1978. He received the Dr. Eng. degree (Hons.) in computer engineering and the Ph.D. degree (Hons.) in information engineering from the University of Salento (formerly University of Lecce), Lecce, Italy, in 2004 and 2008, respectively. From July 2007 to July 2010, he was with the University of Cassino, Cassino, Italy, engaged in a research project on algorithms for track-before-detect of multiple targets in uncertain scenarios. From September to November 2009, he was a Visiting Scientist with the NATO Undersea Research Centre, La Spezia, Italy. From September 2011 to April 2015, he was with Elettronica S.p.A. engaged as a System Analyst in the field of electronic warfare. In May 2015, he joined the Università degli Studi Niccolò Cusano, Rome, Italy, where he is currently an Associate Professor. In 2007, he has held visiting positions with the Department of Avionics and Systems, ENSICA (now Institut Supérieur de l'Aéronautique et de l'Espace, ISAE), Toulouse, France, and from 2017 to 2019, he was with the Chinese Academy of Science, Beijing, China. He is the author or coauthor of more than 150 scientific publications in international journals, conferences, and books. His main research interests include statistical signal processing with more emphasis on adaptive detection and tracking of multiple targets in multisensor scenarios. He was a Senior Area Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He is currently an Associate Editor for the IEEE OPEN JOURNAL ON SIGNAL PROCESSING, *EURASIP Journal on Advances in Signal Processing*, and *Remote Sensing* (MDPI).

Stefania Bartoletti (Member, IEEE) received the Laurea degree *(summa cum laude)* in electronics and telecommunications engineering and the Ph.D. degree in information engineering from the University of Ferrara, Italy, in 2011 and 2015, respectively. She is currently a Researcher with the Institute of Electronics, Computer and Telecommunication Engineering (IEIIT), National Research Council of Italy (CNR). She was a Marie Skłodowska-Curie Global Fellow within the Horizon 2020 European Framework for a research project with the Wireless Information & Network Science Laboratory, Massachusetts Institute of Technology (MIT) and the University of Ferrara, from 2016 to 2019. Her research interests include theory and experimentation of wireless networks for passive localization and physical behavior analysis. She was a recipient of the 2016 Paul Baran Young Scholar Award of the Marconi Society. She served as the Chair of the TPC for the IEEE ICC and Globecom Workshops on Advances in Network Localization and Navigation (ANLN) from 2017 to 2021, and as a reviewer for numerous IEEE journals and international conferences. She is Associate Editor of the IEEE COMMUNICATIONS LETTERS.

Ivan Palamà was born in Rome, Italy, in 1996. He received the master's degree *(cum laude)* in ICT and internet engineering from the University of Rome "Tor Vergata", Italy, in October 2020. He is currently pursuing the Ph.D. degree in electronic engineering. He has been a CNIT Researcher since January 2018.

Giuseppe Bianchi is currently a Full Professor in networking with the University of Rome Tor Vergata, Rome, Italy. He has coordinated six large-scale EU projects. His research activities include wireless networks (an area where he has carried out pioneering research work on WLAN modeling and assessment), programmable network systems, security monitoring and vulnerability assessment, and traffic modeling and control, and is documented in about 280 peer-reviewed international journal articles and conference papers, accounting for more than 20000 citations. He has been (or still is) an Editor for several journals in his field, including IEEE/ACM TRANSACTIONS ON NETWORKING, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT, and *Computer Communications* (Elsevier).

Nicola Blefari Melazzi is currently a Professor in telecommunications with the University of Roma Tor Vergata and the Director of CNIT, a non-profit consortium of 38 Italian universities.