Abstract—We consider the problem of inferring the conditional independence graph (CIG) of improper complex-valued Gaussian vectors. A \( p \)-variate improper complex Gaussian graphical model associated with an undirected graph with \( p \) vertices is defined as the family of improper complex Gaussian distributions that obey the conditional independence restrictions implied by the edge set of the graph. For real random vectors, considerable body of paper exists where one first tests for exclusion of each edge from the saturated model, and then infers the CIG. Prior work on proper complex Gaussian graphical models is sparse, while that on improper complex Gaussian graphical models is non-existent. In this paper, we propose and analyze a generalized likelihood ratio test (GLRT) based edge exclusion test statistic for improper complex Gaussian graphical models. The null distribution of the test statistic is specified explicitly to allow analytical calculation of the test threshold. An alternative computationally fast version of the GLRT statistic is also derived. Simulation examples are presented to illustrate the proposed statistic.

Index Terms—Improper complex Gaussian graphical models, undirected graph, generalized likelihood ratio test.

I. INTRODUCTION

Graphical models provide a powerful tool for analyzing multivariate data [1]–[3]. They are statistical models “embodiment of a collection of marginal and conditional independences which may be summarized by means of a graph. Such models combine richness in modeling, clarity of interpretation and ease of analysis” [4]. In a typical setting of an undirected graphical model, the conditional dependency structure among \( p \) (real-valued) random variables \( x_1, x_2, \ldots, x_p \) is represented using an undirected graph \( G = (V, E) \), where \( V = \{1, 2, \ldots, p\} = [1, p] \) is the set of \( p \) nodes corresponding to the \( p \) random variables \( x_1, s, \ldots, x_p \), and \( E \subseteq [1, p] \times [1, p] \) is the set of undirected edges describing conditional dependencies among the components of \( x \). The graph \( G \) then is a conditional independence graph (CIG) of \( x \) where there is no edge between nodes \( i \) and \( j \) if and only if (iff) \( x_i \) and \( x_j \) are conditionally independent given the remaining \( p-2 \) variables \( x_{\ell}, \ell \in [1, p], \ell \neq i, \ell \neq j \) [2, p. 60]. Thus, edge \( \{i, j\} \in E \) iff \( x_i \) and \( x_j \) are conditionally dependent, and edge \( \{i, j\} \not\in E \) iff \( x_i \) and \( x_j \) are conditionally independent. Suppose real-valued \( x \) is multivariate Gaussian with positive-definite covariance matrix \( \Sigma \) and inverse covariance matrix (also known as precision matrix or concentration matrix) \( \Omega = \Sigma^{-1} \). Then \( \Omega_{ij} \), the \( (i, j) \)-th element of \( \Omega \), is zero iff \( x_i \) and \( x_j \) are conditionally independent [1, Proposition 5.2], [30, Relation (1.2)].

A CIG \( G \) and real multivariate Gaussian \( x \) specify a real-valued Gaussian graphical model for \( x \) where the distribution of \( x \) obeys the conditional independence restrictions of CIG \( G \) [1, Sec. 5.2], [2, p. 165]. Such models for real-valued \( x \) have been extensively studied, and found to be useful in a wide variety of applications [5]–[11]. For complex-valued \( x \), only the monograph [12], and more recently [13], have studied such models, where \( x \) is assumed to be proper complex. Proper complex-valued graphical models also arise in frequency-domain formulation of graphical modeling of real-valued time series [14]. Such models have been considered in [15] in low-dimensional settings (vector dimension \( p \ll N \), number of vector samples), and in [16]–[18] in high-dimensional settings (\( p \) comparable to \( N \)).

A significant application of graphical modeling of real-valued random vectors has been for analysis of functional magnetic resonance imaging (fMRI) data [19], [20], to provide insights into the functional connectivity of different brain regions [21]. It is stated in [22] that “In fMRI, the measured blood oxygen level dependent (BOLD) signal to detect neural activity is spatially encoded ...” The BOLD fluctuations are measured as a complex-valued fMRI signal over time in the spatial frequency-domain, then the k-space readout is reconstructed with the inverse Fourier transform (IFT). Before the statistical analysis of the fMRI data, the phase portion of the data is generally discarded, despite physiologically useful information contained in the phase.” It has been shown in [22], [23] that complex-valued fMRI data yields improved sensitivity in fMRI analysis. It turns out that fMRI data can be accurately modeled as improper complex-valued [24], [25]. Potential application to fMRI related problems is a motivation for investigation of improper complex Gaussian graphical models. Prior work on improper complex Gaussian graphical models is non-existent.

The results of [22], [23] show improvements in performance in detection of neural activity in regions-of-interest when both magnitude and phase fMRI data are used compared to using just the magnitude data. In a different application concerning independent component analysis of fMRI data, where the objective is to separate a linear mixture of independent signals into its independent components, [24], [25] show improved performance when both magnitude and phase fMRI data are used compared to usage of just the magnitude data. Although real-valued fMRI...
data-based graphical modeling has been investigated in [19], [20], no such application involving complex-valued fMRI data has been reported for graphical modeling. One possible reason for lack of such investigations could be that there are no complex-valued fMRI data sets available publicly; all public repositories (such as OpenNEURO https://openneuro.org/public/datasets) provide magnitude-only, i.e., real-valued, data.

Consider complex-valued random vectors \( x \) and \( z \). Let the superscripts \( \ast \) denote the complex conjugate, transpose and Hermitian (conjugate transpose) operations, respectively, and \( \mathbb{E} \) denote the expectation operation. Denote the cross-correlation matrix of \( x \) and \( z \) as \( R_{xz} = \mathbb{E}\{zx^{H}\} \). Random vector \( x \) is said to be circular (or circularly symmetric) if \( e^{i\theta}x \) has the same probability distribution as \( x \) for all real \( \theta \) [26], [27], p. 53. A complex-valued random vector \( x \) is said to be proper if \( x \) is uncorrelated with its complex conjugate, i.e., \( \mathbb{E}\{x(x^{H})\} = \mathbb{E}\{xx^{\ast}\} = \mathbb{E}\{x\} \mathbb{E}\{x^{\ast}\} \) [27], p. 35, [28]. A circular vector is proper but conversely is not necessarily true [27], p. 53. A complex zero-mean Gaussian vector \( x \) is proper if and only if it is circular [27], p. 53.

Edge exclusion tests for real Gaussian graphical models are in [1] (and others), and that for proper complex Gaussian graphical models are in [12], [13]. Prior work on improper complex Gaussian graphical models is non-existent. In this paper we exploit [1, Prop. 5.14] after casting improper complex Gaussian graphical models, as real Gaussian graphical models with multiple constraints.

In this paper, we propose and analyze a generalized likelihood ratio test (GLRT) based edge exclusion test statistic for improper complex Gaussian graphical models. In graphical model selection, one needs to decide if a given edge \( \{i, j\} \), for \( 1 \leq i < j \leq p \), in the associated graph \( G = (V, E) \), is in \( E \) or not in \( \mathbb{E} \). Since \( \Omega_{ij} = 0 \Leftrightarrow \{i, j\} \notin \mathbb{E} \) for real Gaussian graphical models and proper complex Gaussian graphical models, its estimate based on the random vector sample (or a related statistic) can be used to test if \( \{i, j\} \notin \mathbb{E} \); hence the term edge exclusion test [29]-[32]. There are \( p(p-1)/2 \) distinct edges in an undirected graph with \( p \) nodes. So for graphical model selection, one has to determine which set of edges out of total \( p(p-1)/2 \) edges belong to \( E \), which requires multiple testing based on control of overall significance level [31]. In this paper we focus on edge exclusion testing, and do not consider multiple testing issues.

The rest of the paper is organized as follows. In Section II-A we review some basic definitions and useful results from [1], [12] concerning simple undirected graphs, that we use later. In Section II-B we recall some existing results on edge exclusion tests for selection of real Gaussian graphical models and proper complex Gaussian graphical models. These tests do not apply to improper complex Gaussian graphical models, but their formulation offers a road map for the formulation of the multiple-edge binary hypothesis testing problem formulated in Section III, and solved in Section IV. In Section III-A we review improper complex Gaussian vectors, and in Section III-B, we define an improper complex Gaussian graphical model with \( p \) nodes in terms of an associated real Gaussian graphical model with \( 2p \) nodes, using an augmented real vector \( y \) comprised of \( \text{real}(x) \) and \( \text{imag}(x) \). A binary hypothesis testing problem is formulated in Section III-C exploiting the concept of “correct graph” [33], [34], and using the associated real Gaussian graphical model, where we test for exclusion/inclusion of four edges in the real Gaussian graphical model corresponding to a single edge in associated improper complex Gaussian graphical model. A GLRT is derived in Section IV-B based on [1, Proposition 5.14], which is discussed in Section IV-A. The null distribution of the test statistic is specified explicitly in Section IV-C, which allows for analytical calculation of the test threshold. A significant simplification of the GLRT statistic of Section IV-A is presented in Section V, and it requires \( O(p^3) \) flops compared to \( O(p^5) \) flops needed in Section IV-A. Simulation results are presented in Section VI.

A. Notation and Terminology

We use \( S \geq 0 \) and \( S > 0 \) to denote that symmetric (or Hermitian) \( S \) is positive semi-definite and positive definite, respectively. The sets of real and complex numbers are denoted by \( \mathbb{R} \) and \( \mathbb{C} \), respectively. For a square matrix \( A \), \( |A| \) and \( \text{etr}(A) \) denote the determinant and the exponential of the trace of \( A \), respectively, i.e., \( \text{etr}(A) = \exp(\text{tr}(A)) \). \( B_{k}(i,j;m) \) denotes the submatrix of the matrix \( B_{k} \) comprising its rows \( i \) through \( j \) and columns \( j \) through \( m \), \( B_{k}(i,j;m) \) is its \( (i,j) \)-th element, \( B_{ij} \) is the \( (i,j) \)-th element of \( B \), \( x_{i} \) is the \( i \)-th component of \( x \in \mathbb{C}^{p} \), and \( I \) is the identity matrix. The superscripts \( \ast, \mathbb{T} \) and \( H \) denote the complex conjugate, transpose and Hermitian (conjugate transpose) operations, respectively, and \( \mathbb{E} \) denotes the expectation operation. The real and imaginary parts of \( x \in \mathbb{C}^{p} \) are denoted by \( \text{real}(x) \) and \( \text{imag}(x) \), respectively.

For a set \( V, |V| \) denotes its cardinality, i.e., the number of elements in \( V \). The set of integers 1 through \( n \) is denoted by \([n]\), \( n \leq p \). For scalar \( x, x \sim \mathcal{N}(\mu, \Sigma) \) denotes a proper complex Gaussian vector with mean \( \mu \) and covariance \( \Sigma \). Also, for complex vector \( x \sim \mathcal{N}(\mu, \Sigma) \), \( |x| \) denotes the number of finite real number \( b \) such that \( \lim_{x \to b}|x|/|x| < 0 \).

The notation \( x \sim \mathcal{N}(\mu, \Sigma) \) denotes a real random vector \( x \) that is Gaussian with mean \( \mu \) and covariance \( \Sigma \). Also, for complex \( x \sim \mathcal{N}(\mu, \Sigma) \), \( |x| \) denotes a proper complex Gaussian vector with mean \( \mu \) and covariance \( \Sigma \). For scalar \( x, x \sim \mathcal{B}(\alpha, \beta) \) denotes a beta random variable with probability density function (pdf)

\[
f_{x}(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}, \quad 0 < x < 1,
\]

where \( \Gamma(\alpha) \) denotes the (complete) Gamma function \( (\alpha) := \int_{0}^{\infty} t^{\alpha-1}e^{-t} dt \), and we do not distinguish between a random variable and the value taken by it. The abbreviations w.p.1., w.r.t. and i.i.d. stand for with probability one, with respect to, and independent and identically distributed, respectively.

II. PRELIMINARIES AND BACKGROUND

In Section II-A, we review some basic definitions and useful results from [1, Chapter 2] and [12, Chapter 5] concerning simple undirected graphs, that we use later in this paper. In Section II-B we recall some existing results on edge exclusion tests
for selection of real Gaussian graphical models and proper complex Gaussian graphical models. These tests do not apply to improper complex Gaussian graphical models, but their formulation offers a road map for the formulation of the multiple-edge binary hypothesis testing problem formulated in Section III, and solved in Section IV.

A. Simple Undirected Graphs [1, Chapter 2], [12, Chapter 5]

A simple undirected graph $\mathcal{G}$ is a pair $(V, \mathcal{E})$, where $V$ is a finite set of elements called vertices (or nodes), and the set $\mathcal{E} \subseteq V \times V$ of unordered pairs (called edges) of distinct vertices. We denote the edge between vertices $\alpha$ and $\beta$ of $V$ as $\{\alpha, \beta\}$. The set $V \setminus A$ denotes the set of all elements in $V$ that are not in $A$.

A graph is said to be complete, or saturated, if $\mathcal{E} = V \times V$, i.e., if any given vertex is connected to every other vertex. A subset $A \subseteq V$ for $\mathcal{G} = (V, \mathcal{E})$ induces the subgraph $G_A = (A, \mathcal{E}_A)$ where $\mathcal{E}_A = \mathcal{E} \cap (A \times A)$, i.e., $\mathcal{E}_A$ contains exactly those edges in $\mathcal{E}$ which connect vertices from $A$. A subset is complete if it induces a complete subgraph.

A path of length $n$ in $\mathcal{G}$ is a sequence of distinct vertices $\alpha_0, \alpha_1, \ldots, \alpha_n$, where $\{\alpha_{j-1}, \alpha_j\} \in \mathcal{E}$ for all $j = 1, 2, \ldots, n$. Two subsets $A, B \subseteq V$ are said to be separated by $S \subseteq V$ if all paths from $A$ to $B$ go via $S$, i.e., the paths from a vertex in $A$ to a vertex in $B$ intersect $S$ at some vertex.

Definition 1 (Decomposition): A pair $(A, B)$ of subsets of $V$ of a simple undirected graph $\mathcal{G} = (V, \mathcal{E})$ is said to form a decomposition of $\mathcal{G}$ if $V = A \cup B$, $A \cap B$ is a complete subset of $V$, and $A \setminus B$ and $B \setminus A$ are separated by $A \cap B$.

Definition 2 (Clique): Let $\mathcal{G} = (V, \mathcal{E})$ be a simple undirected graph. A complete subset of $V$ which is maximal w.r.t. set inclusion is called a clique, i.e.,

\[
(C \text{ is complete and } C \subset C' \Rightarrow C' \text{ is not complete}) \\
\iff C \text{ is a clique ,}
\]

where $C, C' \subseteq V$.

Definition 3 (Decomposability): A simple undirected graph $\mathcal{G} = (V, \mathcal{E})$ is said to be decomposable if it is complete, or if there exists a decomposition formed by proper subsets $A$ and $B$ of $V$ into decomposable subgraphs $G_A$ and $G_B$.

Definition 4 is from [1, pp. 14–15], as specialized to this paper.

Definition 4: Perfect sequence and running intersection property. Let $B_1, \ldots, B_m$ be a sequence of subsets of $V$ of a simple undirected graph $\mathcal{G} = (V, \mathcal{E})$. Let

\[
H_j = B_1 \cup \cdots \cup B_j, \quad j \geq 1 \\
S_j = H_{j-1} \cap B_j, \quad j \geq 2.
\]

The sequence $B_1, \ldots, B_m$ is said to be perfect if the following conditions are fulfilled:

i) for all $i > 1$ there is a $j < i$ such that $S_i \subseteq B_j$;

ii) the sets $S_j$ are complete for all $i$.

The condition (i) is known as the running intersection property (RIP).

We conclude this section with Lemma 1, which is [1, Proposition 2.17] and is also in [12, Chapter 5].

Lemma 1: A simple undirected graph $\mathcal{G} = (V, \mathcal{E})$ is decomposable if and only if the cliques of $\mathcal{G}$ can be ordered (or, numbered) to form a perfect sequence.

1) Markov Properties on Undirected Graphs [1, Secs. 3.2, 5.2]: Given $x \in \mathbb{R}^p$ and associated simple undirected graph $\mathcal{G} = (V, \mathcal{E})$, a probability measure is said to obey the pairwise Markov property, relative to $\mathcal{G}$, if for any $\{j, k\} \notin \mathcal{E}$, $x_j$ and $x_k$ are conditionally independent [1, p. 32]. A probability measure is said to obey the global Markov property, relative to $\mathcal{G}$, if for any triple $(A, B, S)$ of disjoint subsets of $V$ such that $S$ separates $A$ and $B$ in $\mathcal{G}$, random vectors $x_A$ and $x_B$ are conditionally independent given $x_S$, where, for any $V \subseteq V$, we define $x_V = \{x_i\}_{i \in V} = \mathbf{1}$ column vector of dimension $|V|$, consisting of random variables associated with the nodes in $V$.

For real Gaussian graphical models, the pairwise Markov property implies the global Markov property [1, p. 131].

B. Gaussian Graphical Models: Real, and Proper Complex

A real Gaussian graphical model associated with a simple undirected graph $\mathcal{G} = (V, \mathcal{E})$ is defined as the family of $p$-variate real-valued Gaussian random vectors $x \in \mathbb{R}^p$, $p = |V|$, that obey the conditional independence restrictions implied by the edge set $\mathcal{E}$. In this paper, for $|V| = p$, we take $V = [1, p]$. For $x = [x_1 \, x_2 \, \cdots \, x_p]^\top$, in the corresponding graph $\mathcal{G}$, each variable $x_i$ is represented by a node $(i \in V)$, and associations between variables $x_i$ and $x_j$ are represented by edges between nodes $i$ and $j$ of $\mathcal{G}$. In a CIG $\mathcal{G} = (V, \mathcal{E})$ of $x$, there is no edge between nodes $i$ and $j$ iff $x_i$ and $x_j$ are conditionally independent given the remaining $p-2$ variables $x_{i\ell}$, $\ell \in [1, p], \ell \neq i, j$. Thus, edge $(i, j) \notin \mathcal{E}$ iff $x_i$ and $x_j$ are conditionally independent, and edge $(i, j) \in \mathcal{E}$ iff $x_i$ and $x_j$ are conditionally dependent. A CIG $\mathcal{G}$ and real multivariate Gaussian $x$ specify a real-valued Gaussian graphical model for $x$ where the distribution of $x$ obeys the conditional independence restrictions of CIG $\mathcal{G}$ [1, Sec. 5.2], [2, p. 165]. Suppose $x$ has positive-definite covariance matrix $\Sigma$ with inverse covariance matrix (also known as precision matrix or concentration matrix) $\Omega = \Sigma^{-1}$. Then $\Omega_{ij}$, the $(i, j)$-th element of $\Omega$, is zero iff $x_i$ and $x_j$ are conditionally independent [1, Proposition 5.2]. If $x \sim \mathcal{N}(\mathbf{0}, \Sigma)$ with $\Sigma > 0$, the probability density function (pdf) of $x$ is

\[
f_x(x) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp \left( -\frac{1}{2} x^\top \Sigma^{-1} x \right)
\]

\[
= \frac{1}{(2\pi)^{p/2}} \exp \left( -\frac{1}{2} x^\top \Omega x \right).
\]

Similarly, a proper complex Gaussian graphical model associated with an undirected graph $\mathcal{G} = (V, \mathcal{E})$ is defined as the family of $p$-variate proper complex-valued Gaussian random vectors $x \in \mathbb{C}^p$, $p = |V|$, that obey the conditional independence restrictions implied by the edge set $\mathcal{E}$. We take $x$ to be a complex, circularly symmetric (proper), Gaussian random vector, with zero mean and covariance $\Sigma$, i.e., $x \sim \mathcal{C}(\mathbf{0}, \Sigma)$, and we assume that $\Sigma$ is positive definite. Therefore, $\mathbb{E}\{x\} = 0$, $\mathbb{E}\{xx^\top\} = \Sigma \in \mathbb{C}^{p \times p}$,
and \(\mathbb{E}\{xx^\top\} = 0\) [27, p. 35]. The pdf of \(x\) is
\[
f_x(x) = \frac{1}{\pi^p|\Sigma|} \exp\left(-x^H \Sigma^{-1} x\right)
= \frac{|\Omega|}{\pi^p} \exp\left(-x^H \Omega x\right).
\]
(3)

It is fairly common to refer to proper complex Gaussian vectors as just complex Gaussian vectors [12]. However, since we are interested in the case where for \(x \in \mathbb{C}^p\), we do not necessarily have \(\mathbb{E}\{xx^\top\} = 0\), following [27], we explicitly distinguish between proper complex Gaussian vectors (\(\mathbb{E}\{xx^\top\} = 0\)) and improper complex Gaussian vectors (\(\mathbb{E}\{xx^\top\} \neq 0\)). A key result for proper complex Gaussian graphical models is that \(\Omega_{ij} = [\Omega]_{ij}\), the \((i,j)\)-th element of \(\Omega = \Sigma^{-1}\), is zero iff \(x_i\) and \(x_j\) are conditionally independent [12, Theorem 7.1].

1) **Edge Exclusion Tests for Real and Proper Complex Gaussian Graphical Models:** Suppose we are given \(N\) i.i.d. observations \(x(0), x(1), \ldots, x(N-1)\) of \(x\), where \(x(t) \in \mathbb{R}^p\) for real Gaussian graphical models and \(x(t) \in \mathbb{C}^p\) for proper complex Gaussian graphical models. In graphical model selection, one needs to decide if a given edge \(\{i, j\}\), for \(1 \leq i < j \leq p\), in the associated graph \(G = (V, \mathcal{E})\), is in \(\mathcal{E}\) or not in \(\mathcal{E}\). Since \(\Omega_{ij} = 0 \iff \{i, j\} \notin \mathcal{E}\), its estimate based on the random vector sample can be used to test if \(\{i, j\} \notin \mathcal{E}\); hence the term edge exclusion test [31], [32]. There are \(p(p - 1)/2\) distinct edges in an undirected graph with \(p\) nodes. So for graphical model selection, one has to determine which set of edges out of total \(p(p - 1)/2\) edges belong to \(\mathcal{E}\). If \(x(t) \sim \mathcal{N}_c(0, \Sigma)\), then the sample covariance matrix is \(\hat{\Sigma} = \frac{1}{N} \sum_{t=0}^{N-1} xx^\top(t)\), and an estimate of its inverse is \(\hat{\Omega} = \hat{\Sigma}^{-1}\). In [30], [31],
\[
L_R = \frac{|[\hat{\Omega}]_{ij}|}{\sqrt{|[\hat{\Omega}]_{ii}| |[\hat{\Omega}]_{jj}|}}
\]
(4)

(5)

where
\[
\mathbb{R}_+(\mathcal{G}) = \left\{ \Omega \in \mathbb{R}_+^{p \times p} \mid \forall i \neq j \in V : \{i, j\} \notin \mathcal{E} \right\}
= \left\{ \Omega : \impliedby \Omega_{ij} = 0 \right\}.
\]

The GLRT statistic in [1, Sec. 5.3.3] for problem (6) is
\[
L_{RG} = 1 - \frac{|[\hat{\Omega}]_{ij}|^2}{|[\hat{\Omega}]_{ii}| |[\hat{\Omega}]_{jj}|},
\]
where the null is rejected if (8) is small:
\[
L_{RG} \xrightarrow{h_0} h_t \xrightarrow{\tau} \tau.
\]

We see that the statistics (5) and (9) are equivalent.

Now consider proper complex Gaussian graphical models with \(N\) i.i.d. observations \(x(0), x(1), \ldots, x(N-1)\) of \(x \sim \mathcal{N}_c(0, \Sigma)\). Now the hypothesis testing problem is
\[
\mathcal{H}_0 : \Omega = \Omega(G) \in \mathbb{C}_+^{p\times p}, \quad \mathcal{G}' = (V, \mathcal{E}', \mathcal{E}' = \mathcal{E} \setminus \{i, j\}
\]
\[
\mathcal{H}_1 : \Omega = \Omega(G) \in \mathbb{C}_+^{p\times p}, \quad \mathcal{G} = (V, \mathcal{E}), \quad \mathcal{E} \text{ is saturated,}
\]
(10)

We first review improper complex Gaussian vectors in Section III-A, and then, in Section III-B, we define an improper complex Gaussian graphical model with \(p\) nodes in terms of an associated real Gaussian graphical model with \(2p\) nodes, using an augmented real vector \(Y\) comprised of \(\text{real}(x)\) and \(\text{imag}(x)\). A binary hypothesis testing problem is formulated in Section III-C exploiting the concept of “correct graph” [33], [34], and using the associated real Gaussian graphical model, where we test for exclusion/inclusion of four edges in the real Gaussian graphical model corresponding to a single edge in associated improper complex Gaussian graphical model.

### III. System Model

A. **Improper Complex Gaussian Vectors**

Given \(x = x_r + jx_i \in \mathbb{C}^p\), with real part \(x_r\) and imaginary part \(x_i\), define the augmented complex vector \(\mathbf{z}\) and the real
vector $\mathbf{y}$ as

$$\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \bar{\mathbf{x}} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} \mathbf{x}_r \\ \mathbf{x}_i \end{bmatrix}. \quad (14)$$

The pdf of an improper complex Gaussian $\mathbf{x}$ is defined in terms of that of the augmented $\mathbf{y}$ or $\mathbf{z}$ [27, Sec. 2.3.1]. Assume $E\{\mathbf{x}\} = \mathbf{0}$, and define $\mathbf{R}_{uv} = E\{\mathbf{uv}^\top\}$ for (zero-mean) $\mathbf{u}, \mathbf{v} \in \mathbb{R}^p$. Then we have $\mathbf{y} \sim \mathcal{N}_r(\mathbf{0}, \mathbf{R}_{zz})$ where

$$\mathbf{R}_{yy} = \begin{bmatrix} \mathbf{R}_{xixi} & \mathbf{R}_{xixr} \\ \mathbf{R}_{xirx} & \mathbf{R}_{xixi} \end{bmatrix}, \quad (15)$$

The augmented vectors $\mathbf{z}$ and $\mathbf{y}$ are related via

$$\mathbf{z} = \mathcal{T} \mathbf{y},$$

where

$$\mathcal{T} = \begin{bmatrix} I & \begin{bmatrix} J & 0 \\ 0 & -J \end{bmatrix} \\ I & \begin{bmatrix} J & 0 \\ 0 & -J \end{bmatrix} \end{bmatrix} \in \mathbb{C}^{(2p) \times (2p)} \quad (17)$$

and $\mathcal{T}$ is full-rank. For complex vectors, define the covariance matrix $\mathbf{R}_{yy} = E\{\mathbf{uv}^\top\}$, and the complementary covariance matrix $\mathbf{R}_{uu} = E\{\mathbf{uv}^\top\}$ [27, Sec. 2.2], for zero-mean $\mathbf{u}, \mathbf{v} \in \mathbb{C}^p$. Then we have

$$\mathbf{R}_{zz} = \begin{bmatrix} \mathbf{R}_{xx} & \mathbf{R}_{xz} \\ \mathbf{R}_{zx} & \mathbf{R}_{zz} \end{bmatrix} = \mathbf{R}_{zz}^H. \quad (18)$$

Since $\mathbf{y} \sim \mathcal{N}_r(\mathbf{0}, \mathbf{R}_{yy})$, its pdf is given by (assuming $\mathbf{R}_{yy} \succ 0$)

$$f_{\mathbf{y}}(\mathbf{y}) = \frac{1}{(2\pi)^{p/2} |\mathbf{R}_{yy}|^{1/2}} \exp \left(-\frac{1}{2} \mathbf{y}^\top \mathbf{R}_{yy}^{-1} \mathbf{y} \right). \quad (19)$$

Using (16), one can also express (19) as [27, Sec. 2.3.1]

$$f_{\mathbf{z}}(\mathbf{z}) := f_{\mathbf{y}}(\mathbf{z}) = \frac{1}{\pi^p |\mathbf{R}_{zz}|^{1/2}} \exp \left(-\frac{1}{2} \mathbf{z}^\top \mathbf{R}_{zz}^{-1} \mathbf{z} \right). \quad (20)$$

For proper $\mathbf{x}$, $\mathbf{R}_{xz} = 0$, and (20) reduces to (3).

In this paper, in order to exploit some developments in [1] pertaining to real Gaussian graphical models, we will use the representation $\mathbf{y}$ in (14) and its pdf (19) for improper complex Gaussian $\mathbf{x}$.

### B. Improper Complex Gaussian Graphical Model

An improper complex Gaussian graphical model associated with a simple undirected graph $\mathcal{G} = (V, \mathcal{E})$ is defined as the family of $p$-variate complex-valued improper Gaussian random vectors $\mathbf{x} \in \mathbb{C}^p$, $p = |V|$, $V = [1, p]$, that obey the conditional independence restrictions implied by the edge set $\mathcal{E}$. Similar to real Gaussian and proper complex Gaussian graphical models, the CIG $\mathcal{G}$ and multivariate complex-valued improper Gaussian $\mathbf{x}$ specify an improper complex Gaussian graphical model for $\mathbf{x}$ where the distribution of $\mathbf{x}$ obeys the conditional independence restrictions of CIG $\mathcal{G}$.

However, as discussed in Section III-A, since an improper $x_i$ is specified in terms of two random variables (real and imaginary parts of the random variable, or the variable and its complex conjugate), in an improper complex Gaussian graphical model, in fact, each node corresponds to two random variables $\text{real}(x_i) =: x_{jr}$ and $\text{imag}(x_i) =: x_{ji}$. In the notation of (20), conditional independence of improper $x_j$ and $x_k$ is equivalent to

$$f_{x_j,x_k|x_{jr},x_{ki}}(x_j,x_k|x_{jr},x_{ki}) = f_{x_j|x_{jr}}(x_j|x_{jr})f_{x_k|x_{ki}}(x_k|x_{ki}), \quad (21)$$

where $\bar{V} = V \setminus \{j, k\}$. Express the left-side of (21) as

$$f_{x_j|x_{jr},x_{ki}}(x_j|x_{jr},x_{ki}) = f_{x_{jr},x_{jr},x_{kr},x_{ki}}(x_{jr},x_{jr},x_{kr},x_{ki}|x_{jr},x_{ki}), \quad (22)$$

and similarly, express the right-side of (21) as ($l$ is either $j$ or $k$)

$$f_{x_l|x_{lr},x_{li}}(x_l|x_{lr},x_{li}) = f_{x_{lr},x_{lr},x_{lk},x_{li}}(x_{lr},x_{lr},x_{lk},x_{li}|x_{lr},x_{li}), \quad (23)$$

Since

$$f_{x_{jr},x_{jr},x_{kr},x_{ki}}(x_{jr},x_{jr},x_{kr},x_{ki}|x_{jr},x_{ki}) = \int \int f_{x_{jr},x_{jr},x_{kr},x_{ki}}(x_{jr},x_{ji},x_{kr},x_{ki}|x_{jr},x_{ki}) \, dx_{ji}, dx_{ki} \quad (24)$$

and

$$f_{x_{lr},x_{lr},x_{lk},x_{li}}(x_{lr},x_{lr},x_{lk},x_{li}|x_{lr},x_{li}) = \int \int f_{x_{lr},x_{lr},x_{lk},x_{lk}}(x_{lr},x_{li},x_{lk},x_{li}|x_{lr},x_{li}) \, dx_{li} \quad (25)$$

(21)–(25) imply that

$$f_{x_{jr},x_{kr},x_{jr}}(x_{jr},x_{kr}|x_{jr},x_{kr}) = f_{x_{jr},x_{jr},x_{lr},x_{lr}}(x_{jr},x_{jr},x_{lr},x_{lr}|x_{jr},x_{lr}). \quad (26)$$

Proceeding similarly, it then follows that

$$f_{x_{lr},x_{lr},x_{lr},x_{lr}}(x_{lr},x_{lr},x_{lr},x_{lr}|x_{lr},x_{lr}) = f_{x_{lr},x_{lr},x_{lr},x_{lr}}(x_{lr},x_{lr},x_{lr},x_{lr}|x_{lr},x_{lr}). \quad (27)$$

for any real scalars $u, v$ that satisfy $u \in \{\text{real}(x_j), \text{imag}(x_j)\}$, $v \in \{\text{real}(x_k), \text{imag}(x_k)\}$.

In order to exploit some results in [1] pertaining to real Gaussian graphical models for testing the validity of (27)–(28) for a given improper complex Gaussian graphical model, we will use the representation $\mathbf{y}$ in (14), and exploit a larger real Gaussian graphical model corresponding to the given improper complex Gaussian graphical model $\mathcal{G} = (\bar{V}, \mathcal{E})$. Consider a real Gaussian graphical model $\mathcal{G} = (\bar{V}, \mathcal{E})$ associated with improper complex Gaussian graphical model $\mathcal{G} = (V, \mathcal{E})$, where $\bar{V} = [1, 2p]$, vertex $j$ for $1 \leq j \leq p$ represents the real part of improper $x_j$, $\text{real}(x_j)$, and vertex $j + p$ represents $\text{imag}(x_j)$. For a given edge $\{j, k\}$, define the set of four edges $\mathcal{E}(jk)$ as

$$\mathcal{E}(jk) = \{\{j, k\}, \{j + p, k\}, \{j, k + p\}, \{j + p, k + p\}\}. \quad (29)$$

If the edge $\{j, k\} \notin \mathcal{E}$, then we have edges $\bar{\mathcal{E}}(jk) \cap \mathcal{E} = \emptyset$, implying relations (27)–(28) for all four possible values of pair $(u, v)$. Clearly, (21), i.e., $\{j, k\} \notin \mathcal{E}$, implies (27)–(28) for all four possible values of pairs $(u, v)$, which is equivalent to $\mathcal{E}(jk) \cap \mathcal{E} = \emptyset$. What about the converse? Here we invoke the Markov properties on undirected graphs, discussed in Section II-A1, following [1]. Consider subsets $A = \{j, j + p\}$, $B = \{k, k + p\}$ and $S = \bar{V} \setminus (A \cup B)$ of $\bar{V}$ for $j \neq k$, $j, k \in [1, p]$. These three sets are disjoint, and $S$ separates $A$ and $B$ when $\mathcal{E}(jk) \cap \mathcal{E} = \emptyset$.
Since pairwise Markov property implies global Markov property for real Gaussian graphical models ([1], p. 131), Section II-A1), $\mathcal{E}^{(jk)} \cap \hat{\mathcal{E}} = \emptyset$ implies that $x_S$ and $x_p$ are conditionally independent given $x_S$, which, in terms of the original improper complex Gaussian graphical model $\mathcal{G} = (V, \mathcal{E})$, implies that $\{j, k\} \not\in \mathcal{E}$.

Thus, we have established that given improper complex Gaussian graphical model $\mathcal{G} = (V, \mathcal{E})$ with $x \in \mathbb{C}^p$, and the associated real Gaussian graphical model $\tilde{\mathcal{G}} = (\tilde{V}, \tilde{\mathcal{E}})$ with $y = [\text{real}(x^T) \text{imag}(x^T)]^T \in \mathbb{R}^{2p}$,

$$\{j, k\} \not\in \mathcal{E} \Leftrightarrow \tilde{\mathcal{E}}^{(jk)} \cap \tilde{\mathcal{E}} = \emptyset.$$  

(30)

Let $R_{yy} = \mathbb{E}\{yy^T\} > 0$ and $\Omega = R^{-1}_{yy}$. Then, by [1, Prop. 5.2] pertaining to real Gaussian graphical models,

$$\{j, k\} \not\in \mathcal{E} \Leftrightarrow |\Omega|_{\ell m} = 0 \ \forall \{\ell, m\} \in \tilde{\mathcal{E}}^{(jk)}. \quad (31)$$

We state this result as Lemma 2.

**Lemma 2:** Consider an improper complex Gaussian graphical model $\mathcal{G} = (V, \mathcal{E})$ with $x \in \mathbb{C}^p$ and $V = [1, p]$, and the associated real Gaussian graphical model $\tilde{\mathcal{G}} = (\tilde{V}, \tilde{\mathcal{E}})$ with $y = [\text{real}(x^T) \text{imag}(x^T)]^T \in \mathbb{R}^{2p}$, $\tilde{V} = [1, 2p]$, where vertex $j$ for $1 \leq j \leq p$ represents the real part of improper $x_j$, real(x), and vertex $j + p$ represents imag($x_j$). Assume that $R_{yy} = \mathbb{E}\{yy^T\} > 0$. Then, $\forall j, k \in V, j \neq k$, $\{j, k\} \not\in \mathcal{E}$, i.e., $x_j$ and $x_k$ are conditionally independent given $x_{\bar{\mathcal{V}} \backslash \{j, k\}}$. In a binary hypothesis testing framework, we compare two competing models: $\tilde{\mathcal{G}} = (\tilde{V}, \tilde{\mathcal{E}})$ and $\tilde{\mathcal{G}}' = (\tilde{V}, \tilde{\mathcal{E}}')$ with $\tilde{\mathcal{E}}' = \tilde{\mathcal{E}} \backslash \tilde{\mathcal{E}}^{(jk)}$. What should $\tilde{\mathcal{E}}$ be in the real Gaussian graphical model, i.e., what should $\tilde{\mathcal{E}}$ be in the original improper complex Gaussian graphical model? Note that we do not know the true edge set $\mathcal{E}$. To this end, we first recall the concept of a correct graph from [33], [34], where it is defined for time series graphical models, but it applies here as well.

**Definition 5 (Correct Graph):** [33, Definition 2], [34, Definition 1]. Let $\mathcal{G} = (V, \mathcal{E})$, $\mathcal{E} = [1, p]$, denote the true graphical model for $x$ in $\mathbb{R}^p$ or $\mathbb{C}^p$. Then $\tilde{\mathcal{G}} = (\tilde{V}, \tilde{\mathcal{E}})$ is correct for $\mathcal{G} = (V, \mathcal{E})$ if $\tilde{\mathcal{E}} \subseteq \tilde{\mathcal{E}}$. Thus, if $\{j, k\} \in \mathcal{E}$, then we must have $\{j, k\} \in \hat{\mathcal{E}}$, and if $\{j, k\} \not\in \mathcal{E}$, then $\{j, k\} \not\in \hat{\mathcal{E}}$. Note that a saturated (complete) graph $\mathcal{G} = (V, \mathcal{E}_s)$ is a correct graph for any graphical model, where $\mathcal{E}_s$ denotes the set of all possible edges with vertices in $V$. This definition of a correct graph is central to the approach of [34], patterned after that of [33], to graphical modeling of real time series, and it applies in our case as well. To test if edge $\{j, k\} \not\in \mathcal{E}$, we consider if edge $\{j, k\} \not\in \mathcal{E}_s$ (null: correct model), or $\{j, k\} \in \mathcal{E}_s$ (alternative: incorrect model), where $(V, \mathcal{E}_s)$ is correct for any $(V, \mathcal{E})$. There are $L = p(p - 2)/2$ edges in $(V, \mathcal{E}_s)$. Number edge $\{j, k\}$ as $i = (j - 1)p + k, 1 \leq j < k \leq p$. Let $(V, \mathcal{E}^i)$ denote the graph where $\mathcal{E}^i = \mathcal{E}_s \backslash \{j, k\}$, with $i = (j - 1)p + k$, i.e., only one edge is missing from the saturated graph. Lemma 3 is a restatement of [34, Proposition 2].

**Lemma 3:** If the graph $(V, \mathcal{E}^i)$ is correct for edges corresponding to $i = i_1, i_2, \ldots, i_s$, and incorrect for all others, then the graphical model $\mathcal{G} = (V, \mathcal{E})$ for $x$ is the graph with only edges $\{i_1, i_2, \ldots, i_s\}$ missing.

Motivated by Lemma 3 and (6), and using Lemma 2, we consider the following hypothesis testing problem for testing the exclusion of single edge $\{j, k\}$ from $\mathcal{E}$:

$$\mathcal{H}_0 : \Omega = \Omega(\tilde{\mathcal{G}}) \in \tilde{R}_+ = (\tilde{V}^T \tilde{V}) \quad \forall j \neq k \in V : \tilde{\mathcal{E}}^{(jk)} \cap \tilde{\mathcal{E}} = \emptyset$$

$$\mathcal{H}_1 : \Omega = \Omega(\tilde{\mathcal{G}}) \in \tilde{R}_+ = (\tilde{V}^T \tilde{V}), \quad \tilde{\mathcal{E}} = \mathcal{E} \cup \tilde{\mathcal{E}}^{(jk)}$$

(32)

where

$$\tilde{R}_+ = \left\{ \Omega \in \mathbb{R}^{(2p) \times (2p)} : \forall j \neq k \in V : \tilde{\mathcal{E}}^{(jk)} \cap \tilde{\mathcal{E}} = \emptyset \right\}.$$  

(33)

It follows from Lemmas 2 and 3 that if $\mathcal{H}_0$ in (32) is true then $\{j, k\} \not\in \mathcal{E}$, else $\{j, k\} \in \mathcal{E}$. In the next section we derive a GLRT for testing problem (32).

**D. Relation to Prior Work**

As noted in Section I, real Gaussian graphical models associated with undirected graphs have been extensively investigated. These models are also called covariance selection models [1], [35], or Gaussian Markov random fields (GMRF) models [36]. Learning a Gaussian graphical model corresponds to learning which entries of the inverse covariance matrix are zero, and also estimating its nonzero entries. Dempster [35] appears to be the first to address this problem. One may classify prior work into two broad categories: low-dimensional models [1], [2], [7], [8], [19], [29]–[32], [35] where $p \ll N$, and high-dimensional models [3], [5], [6], [9]–[11], [20], [36], [37] where $p$ is of the order of $N$ (or larger).

Work on high-dimensional models relies crucially on the assumption that the non-zero elements of the inverse covariance matrix are sparse. Since $p$ is of the order of $N$ (or larger), the sample covariance matrix may not be invertible or may be quite ill-conditioned. Furthermore, testing for each possible edge for inclusion/exclusion in the graph is not practically feasible. So an emphasis is on devising approaches for estimating $\Omega$ in such a way that the resulting estimate of $\Omega$ is sparse. In [3], [9] a penalized regression (pseudo-likelihood) approach is used, whereas in [6], a penalized log-likelihood function is used with lasso-related penalties. Similar approaches include [10], [11]. Furthermore, these approaches, and in general, high-dimensional statistical methods, lack classical measures of uncertainty and statistical significance, such as confidence intervals and $p$-values, needed for classical hypothesis testing and estimator quality quantification.

This paper is concerned with the low-dimensional case, with no sparsity assumption or restriction. A focus in graphical
model selection has been to devise statistical tests for exclusion/inclusion of edges in the graph [1], [2], [29]–[32]. For a single edge, the edge exclusion test of [2, Sec. 6.8] is the same as that of [1, Sec. 5.3.3] (see (9)), but in order to calculate the test threshold to yield a specified significance level $\alpha$, [2, Sec. 6.8] uses an asymptotic $\chi^2$-distribution under the null hypothesis, whereas [1, Sec. 5.3.3] specifies an exact distribution under the null hypothesis. As noted in Section II-B1, an equivalent test statistic has also been used in [30], [31]. As already noted in Section I, there are $p(p-1)/2$ distinct edges in an undirected graph with $p$ nodes. Therefore, for graphical model selection, one has to determine which set of edges out of total $p(p-1)/2$ edges belong to $E$, which requires multiple testing based on control of overall significance level [31] over all tested edges. In this paper we focus on edge exclusion testing, and do not control of overall significance level [31] over all tested edges. By assumption, $\mathbb{E}\{x(t)\} = 0$, hence, $\mathbb{E}\{y(t)\} = 0$. The GLRT for problem (32) is given by

$$f_Y(Y) = \prod_{t=0}^{N-1} f_Y(t)(y(t)) = \prod_{t=0}^{N-1} \left(\frac{1}{(2\pi)^{p/2}|\mathbf{R}_yy|^{1/2}}\right) \exp\left(-\frac{1}{2}y^\top \mathbf{R}_yy\right) = \left|\Omega\right|^{-N/2}(2\pi)^{Np}\text{etr}\left(-\frac{1}{2}\Omega Y^\top Y\right).$$

The unknown in (35) is $\Omega$ which satisfies the respective restrictions in (32) under the two hypotheses. The GLRT for problem (32) is given by

$$\mathcal{L}(Y) = \sup_{\Omega} \mathcal{L}(Y|H_1)$$

$$\mathcal{L}(Y) = \sup_{\Omega} \mathcal{L}(Y|H_0).$$

The solution to (36)–(37) will be obtained by exploiting [1, Proposition 5.14], which we recall next.

A. GLRT for Decomposable Real Gaussian Graphical Models [1, Sec. 5.3.3]

Consider $x \in \mathbb{R}^p$ and the associated simple undirected graph $\mathcal{G} = (V,E)$ which is assumed to be decomposable. Let $\mathcal{G}_0 = (V,E_0)$ denote a decomposable submodel of $\mathcal{G}$ with $k \geq 1$ edges less than $\mathcal{G}$. By [1, Lemma 2.21], there is a sequence $\mathcal{G}_0 \subset \cdots \subset \mathcal{G}_k = \mathcal{G}$ of graphs that are decomposable and differ by one edge only. Let $e_i$ denote the edge that is in $\mathcal{G}_i$ but not in $\mathcal{G}_{i-1}$. Consider the hypothesis testing problem

$$\mathcal{H}_0 : \mathcal{G}_0 = (V,E_0) \quad \mathcal{H}_1 : \mathcal{G}_1 = (V,E_0) \quad \mathcal{H}_k : \mathcal{G}_k = (V,E_0),$$

given $N$ i.i.d. observations $x(t)$, $t = 0, 1, \ldots, N-1$, of zero-mean $x$. Define

$$X = [x(0) \ x(1) \cdots \ x(N-1)]^\top \in \mathbb{R}^{N \times p}.$$ 

The GLRT for problem (38) is given by

$$\mathcal{L}(X) = \frac{1}{\mathcal{H}_0} \tau$$
where
\[ \mathcal{L}(X) = \frac{\sup_{\Omega \in \Omega_+} f_{X|\mathcal{H}_1}(X|\mathcal{H}_1)}{\sup_{\Omega \in \Omega_+} f_{X|\mathcal{H}_0}(X|\mathcal{H}_0)} . \]

Under the above set-up, the solution (40)–(41) to problem (38) is specified in [1, Proposition 5.14], which we now state as Lemma 4 in the notation of this paper.

Lemma 4: The GLRT for the test (38) is given by (40) where
\[ \mathcal{L}(X) = \left( \frac{|\Omega_0|}{|\Omega|} \right)^{-N/2}, \]
\[ \hat{\Omega}_0 \text{ maximizes } f_{X|\mathcal{H}_0}(X|\mathcal{H}_0) \text{ under the constraint } \Omega(\mathcal{G}) \in \mathbb{R}_+ (\mathcal{G}_0), \text{ and } \Omega \text{ maximizes } f_{X|\mathcal{H}_1}(X|\mathcal{H}_1) \text{ under the constraint } \Omega(\mathcal{G}) \in \mathbb{R}_+ (\mathcal{G}). \]

The sequence of subsets \( B_1, B_2 \) of \( \tilde{V} \) is perfect since \( S_2 \) is complete and \( S_2 \subset B_1 \). The sets \( B_1 \) and \( B_2 \) are the cliques of \( \tilde{G}_1 \), which is decomposable by Lemma 1.

ii) Now remove edge \( \{j, k\} \) from \( \tilde{G}_2 \) to obtain \( \tilde{G}_3 = (\tilde{V}, \tilde{E}_\gamma \backslash \{(j, k), \{j, k\}\}) \). Define \( B_1 = \tilde{V}\backslash\{j, k\} \) and \( B_2 = \tilde{V}\backslash\{j\} \). With \( m = 2 \) in Definition 4, we have
\[ H_1 = B_1 = \tilde{V}\backslash\{j, k\}, \quad H_2 = B_1 \cup B_2 = \tilde{V} \]
\[ S_2 = H_1 \cap B_2 = \tilde{V}\backslash\{j\} . \]

The sequence of subsets \( B_1, B_2 \) of \( \tilde{V} \) is complete and \( S_2 \subset B_1 \). The sets \( B_1 \) and \( B_2 \) are the cliques of \( \tilde{G}_2 \), which is, therefore, decomposable by Lemma 1.

iv) Now remove edge \( \{j + p, k\} \) from \( \tilde{G}_2 \) to obtain \( \tilde{G}_3 = (\tilde{V}, \tilde{E}_\gamma \backslash \{(j, k), \{j, k\}\}) \). Define \( B_1 = \tilde{V}\backslash\{j, j + p\} \) and \( B_2 = \tilde{V}\backslash\{j\} \) and \( B_3 = \tilde{V}\backslash\{k, k + p\} \). With \( m = 3 \) in Definition 4, we have
\[ H_1 = B_1 = \tilde{V}\backslash\{j, j + p\}, \quad H_2 = B_1 \cup B_2 = \tilde{V}\backslash\{j\} \]
\[ H_3 = B_1 \cup B_2 \cup B_3 = \tilde{V} \]
\[ S_2 = H_1 \cap B_2 = \tilde{V}\backslash\{j, j + p, k\} \]
\[ S_3 = H_2 \cap B_3 = \tilde{V}\backslash\{j, k, k + p\} . \]

The sequence of subsets \( B_2 \) and \( B_3 \) are complete, \( S_2 \subset B_1 \), and \( S_3 \subset B_2 \). Hence the sequence of subsets \( B_1, B_2, B_3 \) of \( \tilde{V} \) is a perfect sequence. The sets \( B_1, B_2 \) and \( B_3 \) are all possible cliques of \( \tilde{G}_3 \), which is, therefore, decomposable by Lemma 1.

v) Finally, remove edge \( \{j + p, k + p\} \) from \( \tilde{G}_3 \) to obtain \( \tilde{G}_4 = (\tilde{V}, \tilde{E}_\gamma \backslash \{(j, k), \{j, k\}\}) \). Define \( B_1 = \tilde{V}\backslash\{j, j + p\} \) and \( B_2 = \tilde{V}\backslash\{k, k + p\} \). With \( m = 2 \) in Definition 4, we have
\[ H_1 = B_1 = \tilde{V}\backslash\{j, j + p\}, \quad H_2 = B_1 \cup B_2 = \tilde{V} \]
\[ S_2 = H_1 \cap B_2 = \tilde{V}\backslash\{j, j + p, k\} . \]

Since \( S_2 \) is complete and \( S_2 \subset B_1 \), the sequence \( B_1, B_2 \) is perfect. The sets \( B_1 \) and \( B_2 \) are all possible cliques of \( \tilde{G}_4 \), which is, therefore, decomposable by Lemma 1.

With the above results, we are ready to invoke Lemma 4 (i.e., [1, Proposition 5.14]). Define
\[ \tilde{G}_4 := \tilde{G}, \quad \tilde{G}_3 := (\tilde{V}, \tilde{E}_\gamma \backslash \{(j, k), \{j, k\}\}) \]
\[ \tilde{G}_2 := (\tilde{V}, \tilde{E}_\gamma \backslash \{(j, k), \{j, k\}\}) \]
\[ \tilde{G}_1 := (\tilde{V}, \tilde{E}_\gamma \backslash \{(j, k), \{j, k\}\}) \]

With reference to (32), then we have a sequence of graphs \( \tilde{G}_t = G_0 \subset G_1 \subset \cdots \subset G_4 = \tilde{G} \) that are decomposable, and differ by one edge only. With \( e_1 \) as in Lemma 4, we have
\[ e_1 = \{j + p, k\}, \quad e_2 = \{j + p, k\} \]
\[ e_3 = \{j, k + p\}, \quad e_4 = \{j, k\} . \]

Then with \( C_t \) as in Lemma 4, we have
\[ C_1 = \tilde{V}\backslash\{j, k\}, \quad C_2 = \tilde{V}\backslash\{j\} \]
\[ C_3 = \tilde{V}\backslash\{k\}, \quad C_4 = \tilde{V} \]

Thus, \( |C_1| = 2p - 2, |C_2| = 2p - 1, |C_3| = 2p - 1 \) and \( |C_4| = 2p - 1 \).

We also need the determinants of the ML estimates of \( \tilde{\Omega} \) under the two hypotheses. Define the sample covariance matrix (we now use \( \tilde{\Sigma} = R_{yy} \))
\[ \tilde{\Sigma} = \frac{1}{N} \mathbf{Y}^\top \mathbf{Y} = \frac{1}{N} \sum_{t=0}^{N-1} \mathbf{y}(t)\mathbf{y}^\top(t) . \]
Under $H_0$, the graph $\bar{G}' = \bar{G}_0$ is decomposable with cliques $\bar{V} \setminus \{j, j + p\}$ and separator $\bar{V} \setminus \{k, k + p\}$ (see step (v) of the procedure detailed above for finding the decomposable graphs and associated cliques). By [1, Proposition 5.9], the determinant of the ML estimate $\hat{\Omega}$ of $\bar{\Omega} = \bar{\Sigma}^{-1}$ is given by (noting that the mean of $z(t)$ is zero, by [1, Proposition 5.9]), we have
\[
|\hat{\Omega}| = \frac{|\hat{\Sigma}_V \setminus \{j, j + p, k + p\}|}{|\hat{\Sigma}_V \setminus \{j, j + p\}| |\hat{\Sigma}_V \setminus \{k, k + p\}|}.
\]

By [1, Proposition 5.9], the determinant of the ML estimate $\hat{\Omega}$ of $\bar{\Omega} = \bar{\Sigma}^{-1}$ under $H_0$ is determined by the cliques and separators of the graph $\bar{G}' = (\bar{V}, \bar{E}') = \bar{G}_0$. Noting that the mean of $z(t)$ is zero, by [1, Proposition 5.9], we have
\[
|\hat{\Omega}_0| = \frac{|\hat{\Sigma}_V \setminus \{j, j + p, k + p\}|}{|\hat{\Sigma}_V \setminus \{j, j + p\}| |\hat{\Sigma}_V \setminus \{k, k + p\}|}.
\]

Under $H_0$, the graph $\bar{G}' = \bar{G}_0$ is decomposable with cliques $\bar{V} \setminus \{j, j + p\}$ and separator $\bar{V} \setminus \{k, k + p\}$ (see step (v) of the procedure detailed above for finding the decomposable graphs and associated cliques). By [1, Proposition 5.9], the determinant of the ML estimate $\hat{\Omega}_0$ of $\bar{\Omega}_0 = \bar{\Sigma}^{-1}$ under $H_0$ is determined by the cliques and separators of the graph $\bar{G}' = (\bar{V}, \bar{E}') = \bar{G}_0$. Noting that the mean of $z(t)$ is zero, by [1, Proposition 5.9], we have
\[
|\hat{\Omega}_0| = \frac{|\hat{\Sigma}_V \setminus \{j, j + p, k + p\}|}{|\hat{\Sigma}_V \setminus \{j, j + p\}| |\hat{\Sigma}_V \setminus \{k, k + p\}|}.
\]

Under $H_0$, the graph $\bar{G}' = \bar{G}_0$ is decomposable with cliques $\bar{V} \setminus \{j, j + p\}$ and separator $\bar{V} \setminus \{k, k + p\}$ (see step (v) of the procedure detailed above for finding the decomposable graphs and associated cliques). By [1, Proposition 5.9], the determinant of the ML estimate $\hat{\Omega}_0$ of $\bar{\Omega}_0 = \bar{\Sigma}^{-1}$ under $H_0$ is determined by the cliques and separators of the graph $\bar{G}' = (\bar{V}, \bar{E}') = \bar{G}_0$. Noting that the mean of $z(t)$ is zero, by [1, Proposition 5.9], we have
\[
|\hat{\Omega}_0| = \frac{|\hat{\Sigma}_V \setminus \{j, j + p, k + p\}|}{|\hat{\Sigma}_V \setminus \{j, j + p\}| |\hat{\Sigma}_V \setminus \{k, k + p\}|}.
\]

Under $H_0$, the graph $\bar{G}' = \bar{G}_0$ is decomposable with cliques $\bar{V} \setminus \{j, j + p\}$ and separator $\bar{V} \setminus \{k, k + p\}$ (see step (v) of the procedure detailed above for finding the decomposable graphs and associated cliques). By [1, Proposition 5.9], the determinant of the ML estimate $\hat{\Omega}_0$ of $\bar{\Omega}_0 = \bar{\Sigma}^{-1}$ under $H_0$ is determined by the cliques and separators of the graph $\bar{G}' = (\bar{V}, \bar{E}') = \bar{G}_0$. Noting that the mean of $z(t)$ is zero, by [1, Proposition 5.9], we have
\[
|\hat{\Omega}_0| = \frac{|\hat{\Sigma}_V \setminus \{j, j + p, k + p\}|}{|\hat{\Sigma}_V \setminus \{j, j + p\}| |\hat{\Sigma}_V \setminus \{k, k + p\}|}.
\]

Under $H_0$, the graph $\bar{G}' = \bar{G}_0$ is decomposable with cliques $\bar{V} \setminus \{j, j + p\}$ and separator $\bar{V} \setminus \{k, k + p\}$ (see step (v) of the procedure detailed above for finding the decomposable graphs and associated cliques). By [1, Proposition 5.9], the determinant of the ML estimate $\hat{\Omega}_0$ of $\bar{\Omega}_0 = \bar{\Sigma}^{-1}$ under $H_0$ is determined by the cliques and separators of the graph $\bar{G}' = (\bar{V}, \bar{E}') = \bar{G}_0$. Noting that the mean of $z(t)$ is zero, by [1, Proposition 5.9], we have
\[
|\hat{\Omega}_0| = \frac{|\hat{\Sigma}_V \setminus \{j, j + p, k + p\}|}{|\hat{\Sigma}_V \setminus \{j, j + p\}| |\hat{\Sigma}_V \setminus \{k, k + p\}|}.
\]

Under $H_0$, the graph $\bar{G}' = \bar{G}_0$ is decomposable with cliques $\bar{V} \setminus \{j, j + p\}$ and separator $\bar{V} \setminus \{k, k + p\}$ (see step (v) of the procedure detailed above for finding the decomposable graphs and associated cliques). By [1, Proposition 5.9], the determinant of the ML estimate $\hat{\Omega}_0$ of $\bar{\Omega}_0 = \bar{\Sigma}^{-1}$ under $H_0$ is determined by the cliques and separators of the graph $\bar{G}' = (\bar{V}, \bar{E}') = \bar{G}_0$. Noting that the mean of $z(t)$ is zero, by [1, Proposition 5.9], we have
\[
|\hat{\Omega}_0| = \frac{|\hat{\Sigma}_V \setminus \{j, j + p, k + p\}|}{|\hat{\Sigma}_V \setminus \{j, j + p\}| |\hat{\Sigma}_V \setminus \{k, k + p\}|}.
\]
Proof: We are interested in determinants of certain matrices. Any elementary row or column operation on a matrix does not change its determinant. Just for this proof, in dealing with $\Sigma$ and $\Omega$, we move the elements associated with the four vertices $\{j, k, j+p, k+p\}$ to the top left corner of the respective matrices, in the order $j, j+p, k, k+p$. The so modified matrices are still denoted by $\hat{\Sigma}$ and $\hat{\Omega}$.

Consider the correspondence of the following conformably partitioned matrices

$$
\hat{\Omega} = \begin{bmatrix}
\hat{\Omega}_{(jk)} & \cdots & * \\
\vdots & \ddots & \vdots \\
\hat{\Sigma}_{(jk)} & \cdots & * \\
\end{bmatrix} = \Sigma^{-1}
$$

where $\Sigma_{(jk)}$ is defined using $\Sigma$ just as $\hat{\Omega}_{(jk)}$ is defined from $\hat{\Omega}$ in (66). For ease of notation, redefine some subsets of $\hat{\Sigma}$ as

$$
V_{(j):k} := \hat{\Sigma}_{(j, k+j+p, k+p)} \quad V_{(j):} := \{j, k, j+p, k+p\}
$$

In this notation, we rewrite (59)–(62) as

$$
\begin{align*}
\Sigma_{V(j):\{j, k, j+p, k+p\}} &= \Sigma_{V(j):k} \\
\Sigma_{\hat{V}(j):\{j, k, j+p, k+p\}} &= \Sigma_{\hat{V}(j):k}
\end{align*}
$$

From [1, Eqn. (B.2)] and (67), we have (recall the notation in (63))

$$
\Omega_{(j):k}^{-1} = \Sigma_{V(j):k} - \Sigma_{V(j):V(j):k} \Sigma_{V(j):V(j):k}^{-1} \Sigma_{V(j):V(j):k}.
$$

By [1, Proposition C.5], the right-side of (72) is the conditional covariance

$$
\Sigma_{V(j):V(j):k}^{-1} = E\{z_{V(j):k}(t)z_{V(j):k}^T(t)\} = \Sigma_{V(j):k} - \Sigma_{V(j):V(j):k} \Sigma_{V(j):V(j):k}^{-1} \Sigma_{V(j):V(j):k}.
$$

Thus

$$
\Omega_{(j):k}^{-1} = \Sigma_{V(j):V(j):k}.
$$

Similarly,

$$
\Omega_{(j):k}^{-1} = \Sigma_{V(j):V(j):k}.
$$

By [1, Eqn. (B.1)] concerning determinant of partitioned matrices, we have

$$
|\Sigma_V| = |\Sigma_{V\backslash\{j,j+p\}}| |\Sigma_{\{j,j+p\}}| |\Sigma_{\hat{V}\backslash\{j,j+p\}}| = |\Sigma_{V(j)}| |\Sigma_{V(j)}| |\Sigma_{V(j)}|.
$$

Similarly, we have

$$
|\Sigma_{V\backslash\{j,k+j+p,k+p\}}| = |\Sigma_{V(j)}| |\Sigma_{V(j)}| |\Sigma_{V(j)}|.
$$

Using notation (68), (76) and (77), we have

$$
|\Sigma_{V\backslash\{j,k+j+p,k+p\}}| |\Sigma_{\hat{V}}| = \frac{|\Sigma_{V(j)}| |\Sigma_{V(j)}|}{|\Sigma_{V(j)}| |\Sigma_{V(j)}|} = \frac{|\Sigma_{V(j)}| |\Sigma_{V(j)}|}{|\Sigma_{V(j)}| |\Sigma_{V(j)}|}.
$$

Since $\Omega_{(j):k}^{-1} = \Sigma_{V(j):V(j)}$ by (74), we have

$$
\Omega_{(j):k}^{-1} = \sum_{V(j):V(j)} [\Omega_{(j):k}^{-1}]_{1:2,1:2}.
$$

Assume that $N \geq 2p$, so that $\hat{\Sigma}^{-1}$ exists w.p.1. Lemma 8 then follows just as Lemma 7, by manipulations of $\Sigma^{-1}$ and submatrices of $\Sigma = \hat{\Sigma}$.

**Lemma 8:** An alternative expression for the test statistic of Theorem 1 is given by

$$
\sum_{V(j):\{j,k,j+p,k+p\}} \sum_{V(j):\{j,k,j+p,k+p\}} = \frac{|\Omega_{(j):k}^{-1}|_{1:2,1:2}}{|\Omega_{(j):k}^{-1}|_{1:2,1:2}}.
$$

where, with $\hat{\Omega}_{(j)}$ denoting the $(l,m)$-entry of $\hat{\Omega} = \hat{\Sigma}^{-1},$

$$
\hat{\Omega}_{(j):k} := \begin{bmatrix}
\hat{\Omega}_{(j):k} & \hat{\Omega}_{(j):k+p} \\
\hat{\Omega}_{(j):k+p} & \hat{\Omega}_{(j):k+p}
\end{bmatrix}
$$

and $|\hat{\Omega}_{(j):k}^{-1}|_{1:2,1:2}$ is a $2 \times 2$ submatrix of $\hat{\Sigma}^{-1}$ in its top left corner.

Using log-likelihood we have the following GLRT

$$
Y := \ln(L(Y)) = \frac{N}{2} \left[ \ln \left( |\hat{\Omega}_{(j):k}^{-1}|_{1:2,1:2} \right) \right]_{\mathcal{H}_1}^{\mathcal{H}_0} \tau_1
$$

where $\tau_1$ is picked to achieve a specified probability of false alarm (significance level) $P_{fa}$. We summarize the above results in Theorem 2.

**Theorem 2:** The GLRT for the test (32) is given by (83), where $\hat{\Omega}_{(j)}$ and $\hat{\Omega}_{(j):k}$ are given by (81) and (82), respectively. Under $\mathcal{H}_0$, the test statistic $Y$ has pdf (57), which can be approximated by (58) for $N \gg p$. 

**Remark 1:** Suppose that we need to carry out edge exclusion tests for all $p(p-1)/2$ edges in $E$. Then it is established easily that the computational complexity of our proposed alternative statistic is $O(p^3)$ versus $O(p^5)$ for the original expression of Theorem 1.
VI. NUMERICAL EXPERIMENTS

Given \( x(t) \in \mathbb{C}^p \), there are \( p(p - 1)/2 \) unordered pair of vertices in the associated CIG that may or may not be connected. So we have to perform at least \( p(p - 1)/2 \) binary hypothesis tests (a given edge is missing from the graph is the null, and the complete graph is the alternative hypothesis). Thus, we have a multiple testing problem where the main issue is how to control the overall significance level. Instead, we will use the ROC (receiver operating characteristic) averaged over all edges, as a performance measure. The ROC curve is the trade-off between the false-alarm rate \( P_{fa} \) and the detection probability \( P_d \). Here we follow [34, Sec. VII] where trade-off between average type I (false-alarm rate) and type II (miss probability) errors (over all edges) has been used as a performance measure (in the context of time series).

We consider two models for generating the synthetic data.

**Model 1:** We start with \( p \times p \Omega \) with all diagonal elements set to 1 and off-diagonal elements equal to 0.5. With probability \( q \) and independently, we set off-diagonal elements in the upper triangle of \( \Omega \) to zero (taking care to set the corresponding elements in lower triangle also to zero so that the resulting matrix \( \Omega \) is symmetric). Now set \( \Omega = \Omega + \beta \mathbf{I} \) with \( \beta \) picked to make \( \Omega \) positive definite. This choice of \( \Omega \) is similar to one of the examples in [37]. Since an off-diagonal element of \( \Omega \) is zero with probability \( q \), approximately 100\% of the entries of \( \Omega \) are null. A typical realization of the graph (for \( x \)) is in Fig. 1 for \( p = 40 \) and \( q = 0.7 \), where only the nodes that are connected to at least one another node, are shown. With \( \Phi \Phi^H = \Omega^{-1} \), we generate \( x = \Phi w \) with \( w \in \mathbb{C}^p \) as zero-mean, improper Gaussian with independent components. Let \( v \sim \mathcal{N}_p(0, \mathbf{I}) \). Then we set \( w_t \), the \( t \)th component of \( w \), as \( w_t = \text{real}(v_t) + j(0.9 \times \text{real}(v_t) + 0.2 \times \text{imag}(v_t)) \), yielding improper complex Gaussian \( w \). We generate \( N \) i.i.d. observations from \( x \), with \( q = 0.7 \).

**Model 2:** Following [34, Appendix], initially let \( A \in \mathbb{R}^{p \times p} \) have all zero entries. Then, for a fixed \( k \), elements in position \((i, j)\) of \( A \) for which \((i + j) \mod k = 1\), are randomly sampled from \( \mathcal{N} \sim (0, 1) \) (standard Gaussian) distribution. Set \( \Omega = A^T A + \beta \mathbf{I} \) with \( \beta = 0.01 \lambda_{max} \), where \( \lambda_{max} \) denotes the maximum eigenvalue of \( A^T A \). With \( \Phi \Phi^H = \Omega^{-1} \), we generate \( x = \Phi w \), with all other details as for Model 1. The choice of \( k \) controls the sparsity: choosing \( k = 5 \) makes 80\% entries of the \( \Phi \) matrix zero for \( p = 40 \), while choosing \( k = 20 \) makes 94\% entries of the \( \Phi \) matrix zero. A typical realization of the graph (for \( x \)) is in Fig. 2 for \( p = 40 \) and \( k = 5 \). In simulations we selected \( k \) randomly and uniformly over \( 5 \leq k \leq 20 \) in each run.

We apply the GLRT (83) coupled with Theorem 2 and Lemma 2, and pdf (57), to compute the test threshold for a given significance level \( \alpha \), to test each of \( p(p - 1)/2 \) edges.

The simulation results are shown in Figs. 3 and 4 for model 1, and in Figs. 5 and 6 for model 2, based on 1000 runs and averaged over all edges, with \( N = 100, 200 \) and \( 400 \), and \( p = 30 \) and 40 (\( y \in \mathbb{R}^{600} \) or \( \mathbb{R}^{800} \)), using a different randomly generated model in each run. It is seen from Figs. 3 and 5 that empirical \( \alpha \) tracks the design \( \alpha \) quite well, except for the case of \( p = 40, N = 100 \), where the \( 80 \times 80 \) data correlation matrix of augmented \( y \) is estimated from 100 samples. The estimated correlation matrix, though of full rank w.p.1, is ill-conditioned.
Fig. 4. Model 1, improper Gaussian, $p = 30$ or $40$, $N = 100$, 200 or 400. ROC curves.

Fig. 5. Model 2, improper Gaussian, $p = 30$ or $40$, $N = 100$, 200 or 400. Empirical significance level vs. design significance level.

Fig. 6. Model 2, improper Gaussian, $p = 30$ or $40$, $N = 100$, 200 or 400. ROC curves.

Fig. 7. Model 1, improper Gaussian, $p = 40$, $N = 200$. Empirical significance level vs. design significance level. Labels “proper” and “improper” refer to tests (13) and (83), respectively, and labels “real-1” and “real-2” refer to test (9) with threshold selected as in items (a) and (b), respectively, of Section VI-A.

Its inverse needed to implement (83) suffers from this ill-conditioning.

The ROC curves are shown in Figs. 4 and 6 for models 1 and 2, respectively, and they show that the performance improves with increasing sample size $N$ and decreasing number of nodes $p$, for a given value of $P_{fa}$ (i.e., significance level or type I error).

A. Applying Existing Methods

For a given edge in our improper complex Gaussian graphical model, we test for joint exclusion of four edges in an augmented real Gaussian graphical model. As noted in Section III-D, the single edge exclusion test (9) of [1, Sec. 5.3.3] can be used to test these four edges but only in a multiple testing set-up where one tests each of the four edges separately, one-by-one. For zero-mean $x$, under the null hypothesis, $L_{RG}$ in (9) is distributed as $B((N - 2p + 1)/2, 1/2)$. We applied multiple testing using (9) to data generated by model 1, based on two different methods of threshold selection:

a) Bonferroni method ([31] and [38, Sec. 10.7]) where, to achieve significance level $\alpha$ as in our proposed approach, we pick significance level $\alpha' = \alpha/4$ for each of four applications of the test (9). The overall significance level then has an upper bound $4\alpha' = \alpha$.

b) A naive approach where we set $\alpha' = \alpha$ for each of four applications of the test (9), and ignore the overall significance level which would be higher than $\alpha$.

We also applied test (13) for proper complex Gaussian graphical models to data generated by model 1, by arbitrarily assuming that $x$ is proper complex Gaussian. As shown in [13], under the null hypothesis, $-N \ln(L_{CG})$ is an exponential random variable with mean $N/(N - p + 1)$, provided $x$ is proper complex Gaussian. Here we do not need multiple testing.

The simulation results based on 1000 Monte Carlo runs, averaged over all edges, are shown in Figs. 7–9 for model 1 with...
$p = 40$ and $N = 200$. In these figures, labels “proper” and “improper” refer to tests (13) and (83), respectively, and labels “real-1” and “real-2” refer to test (9) with threshold selected as in items (a) and (b), respectively, in the preceding paragraph. It is seen from Fig. 7 that only the proposed approach delivers the design significance level. Test (13) cannot do so as it incorrectly assumes that $\mathbf{x}$ is proper. Test (9) based on Bonferroni method of multiple testing keeps the empirical significance level $\alpha$ below the upper bound of design value, as expected, but there is a wide gap between the design value and the empirical significance level. Test (9) based on selecting $\alpha' = \alpha$ and ignoring the overall significance level, yields an empirical significance level $\alpha$ that far exceeds the design value.

Because there is a discrepancy between empirical and design $P_{fa}$ (level $\alpha$) for all tests except for the proposed test (83), we show two ROC curves, $P_d$ versus empirical $P_{fa}$ in Fig. 8, and $P_d$ versus design $P_{fa}$ in Fig. 9. The performance of test (13) shown in Fig. 8 cannot be achieved since the test threshold corresponding to empirical $P_{fa}$ cannot be determined, either analytically or via simulations. If one uses the test threshold corresponding to design $P_{fa}$ which can be obtained, the performance shown in Fig. 9 for test (13) is significantly worse than that of the proposed test (83). Turning to test (9), with Bonferroni method based threshold selection, its performance (labeled “real-1”) shown in Fig. 9 is significantly worse than that of the proposed test (83). Its performance shown in Fig. 8 is better than that in Fig. 9, but the former cannot be achieved as the test threshold corresponding to empirical $P_{fa}$ cannot be determined, either analytically or via simulations. Performance of test (9) implemented via naive selection of test threshold (curves labeled “real-2”) shown in Fig. 9 appears to be quite good but as seen from Fig. 7, empirical $P_{fa}$ corresponding to the design $P_{fa}$ significantly exceeds the design value, and this fact accounts for its apparent good performance in Fig. 9. Finally, its performance shown in Fig. 8 cannot be achieved as the test threshold corresponding to empirical $P_{fa}$ cannot be determined. Thus, only the proposed test (83) and test (9) with Bonferroni method based threshold selection, can be implemented in practice, and comparing these two tests, as shown in Fig. 9, (83) significantly outperforms (9) in that for a given design $P_{fa}$, power of (83) is much higher than that of (9).

VII. CONCLUSION

We considered the problem of inferring the conditional independence graph of improper complex-valued multivariate Gaussian vectors. For real random vectors, considerable body of work exists where one first tests for exclusion of each edge from the saturated model, and then infers the CIG. Existing work on proper complex Gaussian graphical models is sparse, while that on improper complex Gaussian graphical models is non-existent. We proposed and analyzed a GLRT-based edge exclusion test statistic for improper complex Gaussian graphical models. The test statistic was also expressed in an alternative form, where for $p$-dimensional random vectors, the alternative statistic reduces the computational complexity from $O(p^3)$ to $O(p^2)$. Simulation examples were presented to illustrate the proposed statistic.

ACKNOWLEDGMENT

The author would like to thank three reviewers for their helpful comments and suggestions that significantly improved the paper.

REFERENCES


