Supplemental Materials: Oscillating neural circuits: Phase, amplitude, and the complex normal distribution

S1 Theorems and Proofs

S1.1 Proof of Theorem 1

We provide a more precise statement of Theorem 1 and then prove it. To start, we recall some preliminaries about exponential families — see Chapter 8 of Barndorff-Nielsen (2014) for additional details. Let $\mathcal{X}$ be a measurable space with a σ-finite measure ν. An exponential family is a collection $\mathcal{D}$ of distributions P for which the density can be written in the form

$$p(x) = a(P)b(x)\exp\left\{\langle\theta(P), t(x)\rangle\right\}, \quad x \in \mathcal{X},$$

for some functions $a : \mathcal{D} \to \mathbb{R}_+$, $b : \mathcal{X} \to \mathbb{R}_+$, $\theta : \mathcal{D} \to \mathbb{R}^k$, and $t : \mathcal{X} \to \mathbb{R}^k$. We call $\theta(P)$ and $t(x) = (t_1(x), \ldots, t_k(x))$ the parameter and sufficient statistics, respectively, of the distribution.

We say the parametrization θ is affinely independent if there exists no affine combination of the entries $\theta_i : \mathcal{D} \to \mathbb{R}$ that results in the zero function. If θ is affinely independent, then $t(x)$ is minimal sufficient. Let $\Theta = \{\theta(P) \in \mathbb{R}^k : P \in \mathcal{D}\}$ be the parameter space of the family. We now state two fundamental properties of exponential families, mentioned in Theorem 1, which are *fullness* and *regularity*.

Definition S1. An exponential family $\mathcal{D}$ is

1. *full* if the parameter space Θ contains all parameters that induce integrable densities. That is,
$$\Theta = \{\theta' \in \mathbb{R}^k : \int b(x)\exp\left(\langle\theta', t(x)\rangle\right) d\nu(x) < \infty\}, \quad \text{and}$$
2. *regular* if the parameter space Θ is open in $\mathbb{R}^k$ under some parametrization θ and sufficient statistics t.

We now formally define the classes $\mathcal{D}_1, \mathcal{D}_2$ and $\mathcal{D}_3$, which correspond to the sets of complex normal distributions, proper complex normal distributions, and circularly symmetric complex normal distributions. Let $\mathcal{D}_1$ be the collection of distributions over $\mathbb{C}^d$ with densities

$$p(x) \propto \exp\left\{-\frac{1}{2}((\mathfrak{Re}x, \mathfrak{Im}x) - \mu)^\top \Sigma^{-1}((\mathfrak{Re}x, \mathfrak{Im}x) - \mu)\right\}, \quad x \in \mathbb{C}^d, \tag{S1}$$

for some vector $\mu \in \mathbb{R}^{2d}$ and positive definite matrix $\Sigma \in \mathbb{R}^{2d\times 2d}$. The family $\mathcal{D}_1$ represents the entire family of complex normal distributions which can be parametrized by a complex mean vector $m \in \mathbb{C}^d$, complex covariance matrix $\Gamma \in \mathbb{C}^{d\times d}$, and complex pseudo-covariance matrix $C \in \mathbb{C}^{d\times d}$. We denote each distribution by $\mathcal{CN}(m, \Gamma, C)$. Let $\mathcal{D}_2$ be the collection of distributions over $\mathbb{C}^d$ with densities of the same form but for some vector $\mu \in \mathbb{R}^{2d}$ and positive definite matrix $\Sigma \in \mathbb{R}^{2d\times 2d}$ satisfying

$$\Sigma = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}, \tag{S2}$$

where $A \in \mathbb{R}^{d\times d}$ is a symmetric matrix and $B \in \mathbb{R}^{d\times d}$ is an anti symmetric matrix (i.e., $B = -B^\top$). The pseudo-covariance matrix C under the distributions in $\mathcal{D}_2$ is the zero matrix. The distributions in $\mathcal{D}_2$ constitute the family of proper complex normal distributions, denoted by $\mathcal{CN}(m, \Gamma, 0)$.

The equivalence between the class of distributions of the form given by (S1) and (S2) and proper complex normal distributions is a consequence of the structure imposed on the real-valued covariance matrix Σ when the pseudo-covariance matrix is the zero matrix; see Andersen et al. (1995) and Picinbono (1996). Finally, we define $\mathcal{D}_3$ to be the collection of distributions satisfying all the aforementioned conditions and $\mu = 0$. This represents the family of circularly symmetric complex normal distributions, denoted by $\mathcal{CN}(0, \Gamma, 0)$. Theorem 1 is restated as follows.

Theorem S2. *$\mathcal{D}_1, \mathcal{D}_2$ and $\mathcal{D}_3$ are all full and regular exponential families.*

Proof. Denoting $\widetilde{x} = (\mathfrak{Re}x, \mathfrak{Im}x)$, we rewrite the density of $P \in \mathcal{D}_1$ (given in eq. S1) as

$$\begin{aligned} p(x) &\propto \exp\left\{-\frac{1}{2}(\widetilde{x}-\mu)^\top \Sigma^{-1}(\widetilde{x}-\mu)\right\} \\ &\propto \exp\left\{-\frac{1}{2}\widetilde{x}^\top \Sigma^{-1}\widetilde{x} + \mu^\top \Sigma^{-1}\widetilde{x}\right\} \\ &\propto \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}\widetilde{x}\widetilde{x}^\top\right) + \mu^\top \Sigma^{-1}\widetilde{x}\right\}. \end{aligned}$$

The parameter and sufficient statistics are clearly $\theta_1(P) = ((\Sigma^{-1})_{ij} : i \le j) \oplus \mu^\top \Sigma^{-1}$ and $t_1(x) = (\widetilde{x}_i \widetilde{x}_j : i \le j) \oplus \widetilde{x}$, where $\oplus$ denotes the vector concatenation operator. Because $t_1(x)$ is a collection of distinct monomials of orders 1 and 2, the sufficient statistics are affinely independent and therefore minimal. Given the parametrization and minimal sufficient statistics, $\mathcal{D}_1$ is known to be full and regular (Barndorff-Nielsen, 2014, pp. 116–117).

Now consider the family $\mathcal{D}_2$, where Σ has the form given in (S2). As we prove in Lemma S6,

$$\Sigma^{-1} = \begin{pmatrix} A & B \\ -B & A \end{pmatrix},$$

where $A \in \mathbb{R}^{d\times d}$ is a symmetric matrix and $B \in \mathbb{R}^{d\times d}$ is an anti symmetric matrix (i.e. $B = -B^\top$). Then, for $P \in \mathcal{D}_2$, the density (S1) can be rewritten as

$$\begin{aligned} p(x) &\propto \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\begin{pmatrix} A & B \\ -B & A \end{pmatrix}\widetilde{x}\widetilde{x}^\top\right) + \mu^\top \Sigma^{-1}\widetilde{x}\right\} \\ &\propto \exp\left\{-\frac{1}{2}\left(\operatorname{tr}\left(A\ \mathfrak{Re}x\mathfrak{Re}x^\top\right) + \operatorname{tr}\left(A\ \mathfrak{Im}x\mathfrak{Im}x^\top\right) + 2\operatorname{tr}\left(B\ \mathfrak{Re}x\mathfrak{Im}x^\top\right)\right) + \mu^\top \Sigma^{-1}\widetilde{x}\right\} \\ &\propto \exp\left\{-\frac{1}{2}\left[\operatorname{tr}\left\{A\ (\mathfrak{Re}x\mathfrak{Re}x^\top + \mathfrak{Im}x\mathfrak{Im}x^\top)\right\} + \operatorname{tr}\left\{B\ (\mathfrak{Re}x\mathfrak{Im}x^\top - \mathfrak{Im}x\mathfrak{Re}x^\top)\right\}\right] + \mu^\top \Sigma^{-1}\widetilde{x}\right\}. \end{aligned}$$

The third line follows from the fact that $B = -B^\top$ (one consequence of which is that $B_{ii} = 0$). We can easily see that the parameter and sufficient statistics are $\theta_2(P) = (A_{ij} : i \le j) \oplus (B_{ij} : i < j) \oplus \mu^\top \Sigma^{-1}$ and $t_2(x) = (\mathfrak{Re}x_i\mathfrak{Re}x_j + \mathfrak{Im}x_i\mathfrak{Im}x_j : i \le j) \oplus (\mathfrak{Re}x_i\mathfrak{Im}x_j - \mathfrak{Im}x_i\mathfrak{Re}x_j : i < j) \oplus \widetilde{x}$. In addition, $t_2(x)$ is a collection of distinct monomials of orders 1 and 2, so the sufficient statistics are affinely independent and therefore minimal.

Let $p_{\theta'}(x) \propto \exp\{\theta'^\top t_2(x)\}$ for $\theta' \in \mathbb{R}^{d^2+2d}$. To see that $\mathcal{D}_2$ is full, we must show that, for every θ' for which $\int_x \exp\{t_2(x)^\top \theta'\}dx < \infty$, the distribution corresponding to $p_{\theta'}(x)$ lies in $\mathcal{D}_2$.

For any such θ', it is easy to see from that we could construct a symmetric A, an antisymmetric B, and vector μ such that

$$
\begin{aligned}
p_{\theta'}(x) &\propto \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\begin{pmatrix} A & B \\ -B & A \end{pmatrix}\widetilde{x}\widetilde{x}^\top\right) + \mu^\top \begin{pmatrix} A & B \\ -B & A \end{pmatrix}\widetilde{x}\right\} \\
&= \exp\left\{-\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}\widetilde{x}\widetilde{x}^\top\right) + \mu^\top\Sigma^{-1}\widetilde{x}\right\}
\end{aligned}
\tag{S3}
$$

and

$$
\Sigma^{-1} = \begin{pmatrix} A & B \\ -B & A \end{pmatrix}. \tag{S4}
$$

The fact that (S3) has the same form as the densities in $\mathcal{D}_1$ but also $\int p^{\theta'}(x)dx < \infty$ implies that Σ^{-1} must be positive definite since, otherwise, the integration over any eigenvector with a non positive eigenvalue will diverge to ∞ by well-known properties of the multivariate normal distribution. The matrix Σ is also positive definite and has the form in (S2) due to Lemma S6. There exists $P \in \mathcal{D}_2$ such that the density is identical to $p_{\theta'}(x)$. Therefore, $\mathcal{D}_2$ is full.

To see that $\mathcal{D}_2$ is regular, we need to show that $\Theta_2 = \{\theta_2(P) : P \in \mathcal{D}_2\}$ is open. We can construct a linear injection $f : \mathbb{R}^{d^2+2d} \to \mathbb{R}^{d(2d+1)+2d}$ such that

$$
f((A_{ij} : i \le j) \oplus (B_{ij} : i < j) \oplus m) = (S_{ij} : i \le j) \oplus m,
$$

where

$$
S = \begin{pmatrix} A & B \\ -B & A \end{pmatrix}.
$$

We note that $f(\theta_2(P)) = \theta_1(P)$ for $P \in \mathcal{D}_2$, where $\theta_1(P)$ is well-defined as $\mathcal{D}_2 \subset \mathcal{D}_1$.

Assume, towards a contradiction, that the set Θ_2 is not open in $\mathbb{R}^{d^2+2d}$. Then, there exists some $\theta_0 \in \Theta_2$ such that for every ϵ there exists another point $\theta' \in \mathbb{R}^{d^2+2d}$ such that $\theta' = (A'_{ij} : i \le j) \oplus (B'_{ij} : i < j) \oplus m \notin \Theta_2$ but $||\theta' - \theta_0||_2 < \epsilon$. For

$$
S' = \begin{pmatrix} A' & B' \\ -B' & A' \end{pmatrix},
$$

we therefore have that S' is not positive definite, and therefore, $f(\theta') \notin \Theta_1$. However, $||f(\theta') - f(\theta_0)||_2 = ||\theta' - \theta_0||_2 \le \epsilon$. Since ϵ is arbitrary, our previous assumption would imply that Θ_0 is not open. This contradicts the fact that $\mathcal{D}_1$ is regular. Therefore, Θ_2 must be open and so $\mathcal{D}_2$ is regular.

Finally, we want to show the same properties for $\mathcal{D}_3$. For each distribution P in this family, the density has the form in (S1) with $\mu = 0$, i.e.,

$$
p(x) \propto \exp\left(-\frac{1}{2}\left[\operatorname{tr}\left\{A\ (\mathfrak{Re}x\mathfrak{Re}x^\top + \mathfrak{Im}x\mathfrak{Im}x^\top)\right\} + \operatorname{tr}\left\{B\ (\mathfrak{Re}x\mathfrak{Im}x^\top - \mathfrak{Im}x\mathfrak{Re}x^\top)\right\}\right]\right).
$$

The parameter and sufficient statistics are $\theta_2(P) = (A_{ij} : i \le j) \oplus (B_{ij} : i < j)$ and $t_2(x) = (\mathfrak{Re}x_i\mathfrak{Re}x_j + \mathfrak{Im}x_i\mathfrak{Im}x_j : i \le j) \oplus (\mathfrak{Re}x_i\mathfrak{Im}x_j - \mathfrak{Im}x_i\mathfrak{Re}x_j : i < j)$, which are minimal. The fullness and regularity of $\mathcal{D}_3$ follow as corollaries of those of $\mathcal{D}_2$. □

S1.2 The Multivariate Generalized von Mises Distribution

The mGvM distribution generalizes the torus graphs distribution by containing second moment terms such as $(\cos\Theta_i)^2, (\sin\Theta_i)^2$, and $(\cos\Theta_i)(\sin\Theta_i)$ (Navarro et al., 2017).

Definition S3 (mGvM distribution)**.** A random vector $\Theta_d : \Omega \to [0, 2\pi)^d$ has a mGvM distribution, denoted $\Theta_d \sim \mathcal{MGVM}(\nu, \psi)$, for a $2d$-dimensional vector ν and $2d \times 2d$ symmetric matrix ψ if

$$p_\Theta(\theta) \propto \exp\left(\nu^\top \begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix} + \begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix}^\top \psi \begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix}\right).$$

By constraining the parameters of the mGvM distribution, we observe that the TG distribution is a special case of the mGvM distribution.

Lemma S4 (Relationship between the mGvM and TG distributions)**.** *If* $\Theta \sim \mathcal{MGVM}(\nu, \psi)$ *and* $\psi_{ii} = \psi_{i+d,i+d}$ *for* $1 \le i \le d$*, and* $\psi_{i,i+d} = 0$ *for* $1 \le i \le d$*, then* $\Theta \sim \mathcal{TG}(\eta)$ *for*

$$\eta_{ij} = 2(\psi_{ij}, \psi_{i,j+d}, \psi_{i+d,j}, \psi_{i+d,j+d})^\top \quad \textit{and} \quad \eta_{ii} = (\nu_i, \nu_{i+d}).$$

Proof. Assume that $\Theta \sim \mathcal{MGVM}(\nu, \psi)$ for the conditions stated in the lemma. Then

$$\begin{aligned}
p_\Theta(\theta) &\propto \exp\left\{\nu^\top \begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix} + \begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix}^\top \psi \begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix}\right\} \\
&= \exp\Bigg\{\sum_i \begin{pmatrix}\nu_i & \nu_{i+d}\end{pmatrix} \begin{pmatrix}\cos\theta_i \\ \sin\theta_i\end{pmatrix} + \sum_i \{(\cos\theta_i)^2\psi_{ii} + (\sin\theta_i)^2\psi_{i+d,i+d}\} + \\
&\qquad \sum_{i\neq j} \big(\psi_{i,j}\cos\theta_i\cos\theta_j + \psi_{i,j+d}\cos\theta_i\sin\theta_j + \\
&\qquad\qquad \psi_{i+d,j}\sin\theta_i\cos\theta_j + \psi_{i+d,j+d}\sin\theta_i\sin\theta_j\big)\Bigg\} \\
&= \exp\Bigg\{\sum_i \begin{pmatrix}\nu_i & \nu_{i+d}\end{pmatrix} \begin{pmatrix}\cos\theta_i \\ \sin\theta_i\end{pmatrix} + \sum_i \psi_{ii} \\
&\qquad + \sum_{i<j} 2\begin{pmatrix}\psi_{ij} & \psi_{i,j+d} & \psi_{i+d,j} & \psi_{i+d,j+d}\end{pmatrix} \begin{pmatrix}\cos\theta_i\cos\theta_j \\ \cos\theta_i\sin\theta_j \\ \sin\theta_i\cos\theta_j \\ \sin\theta_i\sin\theta_j\end{pmatrix}\Bigg\} \\
&\propto \exp\left\{\sum_i \begin{pmatrix}\nu_i & \nu_{i+d}\end{pmatrix} \begin{pmatrix}\cos\theta_i \\ \sin\theta_i\end{pmatrix} + \sum_{i<j} 2\begin{pmatrix}\psi_{ij} & \psi_{i,j+d} & \psi_{i+d,j} & \psi_{i+d,j+d}\end{pmatrix} \begin{pmatrix}\cos\theta_i\cos\theta_j \\ \cos\theta_i\sin\theta_j \\ \sin\theta_i\cos\theta_j \\ \sin\theta_i\sin\theta_j\end{pmatrix}\right\} \\
&= \exp\left\{\sum_i \eta_i^\top \begin{pmatrix}\cos\theta_i \\ \sin\theta_i\end{pmatrix} + \sum_{i<j} \eta_{ij}^\top \begin{pmatrix}\cos\theta_i\cos\theta_j \\ \cos\theta_i\sin\theta_j \\ \sin\theta_i\cos\theta_j \\ \sin\theta_i\sin\theta_j\end{pmatrix}\right\},
\end{aligned}$$

where the final line has the form of the TG distribution, with the stated values of η. □

Further, by conditioning the angles of the complex normal distribution on the amplitudes, we achieve an mGvM distribution for any parameterization of the normal distribution.

Lemma S5 (Complex normal and mGvM distributions)**.** *Assume $X \sim \mathcal{CN}(m, \Gamma, C)$. By definition, there exist μ and Σ such that $(\mathfrak{Re}X, \mathfrak{Im}X) \sim \mathcal{N}(\mu, \Sigma)$. Let $\Theta_i = \arg X_i$ denote the angle of X_i, and let $R_i = |X_i|$ denote the magnitude of X_i. Then $\Theta \mid R = r \sim \mathcal{MGVM}(\nu, \psi)$, where*

$$\nu = \begin{pmatrix} r \\ r \end{pmatrix} \odot (\Sigma^{-1}\mu) \quad \textit{and} \quad \psi = -\frac{1}{2}\Sigma^{-1} \odot \begin{pmatrix} r \\ r \end{pmatrix} \begin{pmatrix} r \\ r \end{pmatrix}^\top.$$

Proof. Let $Z = (\mathfrak{Re}X, \mathfrak{Im}X)$. By a change of variables, note that

$$p_Z(r_1 \cos\theta_1, \ldots, r_d \cos\theta_d, r_1 \sin\theta_1, \ldots, r_d \sin\theta_d) \prod_i r_i = p_{\Theta,R}(\theta, r).$$

Then

$$\begin{aligned}
p_{\Theta|R}(\theta \mid r) = \frac{p_{\Theta,R}(\theta, r)}{p_R(r)} &\propto p_{\Theta,R}(\theta, r) \propto p_Z(r_1 \cos\theta_1, \ldots, r_d \cos\theta_d, r_1 \sin\theta_1, \ldots, r_d \sin\theta_d) \\
&\propto \exp\left\{-\frac{1}{2}\left(\begin{pmatrix} r \odot \cos\theta \\ r \odot \sin\theta \end{pmatrix} - \mu\right)^\top \Sigma^{-1} \left(\begin{pmatrix} r \odot \cos\theta \\ r \odot \sin\theta \end{pmatrix} - \mu\right)\right\} \\
&\propto \exp\left\{\begin{pmatrix} r \odot \cos\theta \\ r \odot \sin\theta \end{pmatrix}^\top \Sigma^{-1}\mu - \frac{1}{2}\begin{pmatrix} r \odot \cos\theta \\ r \odot \sin\theta \end{pmatrix} \Sigma^{-1} \begin{pmatrix} r \odot \cos\theta \\ r \odot \sin\theta \end{pmatrix}\right\} \\
&= \exp\left\{\begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}^\top \left((\Sigma^{-1}\mu) \odot \begin{pmatrix} r \\ r \end{pmatrix}\right) - \frac{1}{2}\begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \left(\Sigma^{-1} \odot \begin{pmatrix} r \\ r \end{pmatrix}\begin{pmatrix} r \\ r \end{pmatrix}^\top\right) \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}\right\} \\
&= \exp\left\{\nu^\top \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} + \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}^\top \psi \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}\right\}
\end{aligned}$$

for the values of ν and ψ given in the lemma. □

S1.3 Proof of Theorem 2

Proof. Note that $(\mathfrak{Re}X, \mathfrak{Im}X) \sim \mathcal{N}(\mu, \Sigma)$. By Lemma S5,

$$\Theta \mid R = r \sim \mathrm{MGVM}(\nu, \psi) = \mathcal{MGVM}\left(\begin{pmatrix} r \\ r \end{pmatrix} \odot (\Sigma^{-1}\mu), -\frac{1}{2}\Sigma^{-1}\begin{pmatrix} r \\ r \end{pmatrix}\begin{pmatrix} r \\ r \end{pmatrix}^\top\right).$$

If $\psi_{ii} = \psi_{i+d,i+d}$ and $\psi_{i,i+d} = 0$, then Lemma S4 applies and the corollary follows. By construction, $\Sigma^{-1}_{ii} = \Sigma^{-1}_{i+d,i+d}$ and $\Sigma^{-1}_{i,i+d} = 0$, and hence

$$\begin{aligned}
\psi_{ii} &= \left\{-\frac{1}{2}\Sigma^{-1}\begin{pmatrix} r \\ r \end{pmatrix}\begin{pmatrix} r \\ r \end{pmatrix}^\top\right\}_{ii} = -\frac{1}{2}r_i r_i \Sigma^{-1}_{ii} \\
&= \frac{-1}{2}r_i r_i \Sigma^{-1}_{i+d,i+d} = \left\{-\frac{1}{2}\Sigma^{-1}\begin{pmatrix} r \\ r \end{pmatrix}\begin{pmatrix} r \\ r \end{pmatrix}^\top\right\}_{i+d,i+d} = \psi_{i+d,i+d}
\end{aligned}$$

and

$$\psi_{i,i+d} = \left\{-\frac{1}{2}\Sigma^{-1}\begin{pmatrix} r \\ r \end{pmatrix}\begin{pmatrix} r \\ r \end{pmatrix}^\top\right\}_{i,i+d} = -\frac{1}{2}r_{i,i+d}\Sigma^{-1}_{i,i+d} = 0,$$

which meets the requirement. □

S1.4 Structure of Proper Matrices

We need the following auxiliary result for the proof of Theorem 3, that the propriety of the complex normal distribution induces constraints on the conditional TG distribution.

Lemma S6. *Assume that W is a $2d \times 2d$ positive semidefinite, symmetric matrix such that, for a partition*

$$W = \begin{pmatrix} W_{11} & W_{12} \\ W_{12}^\top & W_{22} \end{pmatrix},$$

that W_{12} is antisymmetric (i.e., $W_{12} = -W_{12}^\top$) and $W_{11} = W_{22}$. Assume that the inverses of each block of W (i.e., W_{11}^{-1} and W_{12}^{-1}) exist. Then W^{-1} is a symmetric matrix for which $(W^{-1})_{12} = -(W^{-1})_{12}^\top$ and $(W^{-1})_{11} = (W^{-1})_{22}$.

Proof. Let W be a symmetric matrix that can be written as

$$W = \begin{pmatrix} A & B \\ -B & A \end{pmatrix}.$$

Using the block-matrix inversion formula which leverages the Schur complement of a matrix (see Zhang 2006),

$$\begin{aligned} W^{-1} &= \begin{pmatrix} (A - BA^{-1}(-B))^{-1} & -(A - BA^{-1}(-B))^{-1}BA^{-1} \\ -A^{-1}(-B)(A - BA^{-1}(-B))^{-1} & A^{-1} + A^{-1}(-B)(A - BA^{-1}(-B))^{-1}BA^{-1} \end{pmatrix} \\ &= \begin{pmatrix} (A + BA^{-1}B)^{-1} & -(A + BA^{-1}B)^{-1}BA^{-1} \\ A^{-1}B(A + BA^{-1}B)^{-1} & A^{-1} - A^{-1}B(A + BA^{-1}B)^{-1}BA^{-1} \end{pmatrix}. \end{aligned}$$

There are two important aspects of this expression to note. First, the upper right and lower left blocks are antisymmetric (in other words, $W^{-1}_{i,j+d} = -W^{-1}_{j,i+d}$ for $1 \le i \le d$ and $1 \le j \le d$), so

$$\begin{aligned} -(A + BA^{-1}B)^{-1}BA^{-1} &= -(A + BA^{-1}B)^{-1}(AB^{-1})^{-1} \\ &= -((AB^{-1})(A + BA^{-1}B))^{-1} = -(AB^{-1}A + B)^{-1} \end{aligned}$$

and

$$\begin{aligned} \{-(A + BA^{-1}B)^{-1}BA^{-1}\}^\top &= \{-(AB^{-1}A + B)^{-1}\}^\top \\ &= \{-(AB^{-1}A + B)^\top\}^{-1} = \{-(A(B^\top)^{-1}A + B^\top)\}^{-1} \\ &= (AB^{-1}A + B)^{-1} = (A + BA^{-1}B)^{-1}BA^{-1}. \end{aligned}$$

Further, the top left and bottom right blocks of A^{-1} are equal ($A^{-1}_{ij} = A^{-1}_{i+d,j+d}$ for $1 \le i \le d, 1 \le j \le d$). To see this, first note that

$$\begin{aligned} &A^{-1} - A^{-1}B(A + BA^{-1}B)^{-1}BA^{-1} \\ &= A^{-1}(A + BA^{-1}B)(A + BA^{-1}B)^{-1} - A^{-1}B(A + BA^{-1}B)^{-1}BA^{-1} \\ &= (I + A^{-1}BA^{-1}B)(A + BA^{-1}B)^{-1} - A^{-1}B(A + BA^{-1}B)^{-1}BA^{-1} \\ &= (A + BA^{-1}B)^{-1} + A^{-1}BA^{-1}B(A + BA^{-1}B)^{-1} - A^{-1}B(A + BA^{-1}B)^{-1}BA^{-1}. \end{aligned}$$

Therefore, we have the desired equality if $A^{-1}BA^{-1}B(A+BA^{-1}B)^{-1} = A^{-1}B(A+BA^{-1}B)^{-1}BA^{-1}$. To show that this is the case, note that

$$
\begin{aligned}
A^{-1}BA^{-1}B(A+BA^{-1}B)^{-1} &= (B^{-1}AB^{-1}A)^{-1}(A+BA^{-1}B)^{-1} \\
&= \{(A+BA^{-1}B)B^{-1}AB^{-1}A\}^{-1} \\
&= (AB^{-1}AB^{-1}A+A)^{-1}
\end{aligned}
$$

and that

$$
\begin{aligned}
A^{-1}B(A+BA^{-1}B)^{-1}BA^{-1} &= (B^{-1}A)^{-1}(A+BA^{-1}B)^{-1}(AB^{-1})^{-1} \\
&= \{AB^{-1}(A+BA^{-1}B)(B^{-1}A)\}^{-1} \\
&= (AB^{-1}AB^{-1}A+A)^{-1}.
\end{aligned}
$$

Therefore, W^{-1} has the desired form. □

We also need the following auxiliary result relating the parametrization of the TG distribution $\mathcal{TG}(\eta)$ to the parametrization $\mathcal{TG}(\eta(\phi))$ as specified in (2) and (3).

Lemma S7. *Assume that* $\Theta \sim \mathcal{TG}(\eta)$. *Then* $\Theta \sim \mathcal{TG}\in(\eta(\phi))$ *for*

$$
\phi_i = \eta_i \quad \textit{and} \quad \phi_{ij} = \frac{1}{2}(\eta_{ij,1}+\eta_{ij,4}, -\eta_{ij,2}+\eta_{ij,3}, \eta_{ij,1}-\eta_{ij,4}, \eta_{ij,2}+\eta_{ij,3}),
$$

where $\eta_{ij,k}$ *denotes the* k*th component of* η_{ij}.

Proof. From basic properties of trigonometric functions

$$
\phi_{ij}^\top \begin{pmatrix} \cos(\theta_i-\theta_j) \\ \sin(\theta_i-\theta_j) \\ \cos(\theta_i+\theta_j) \\ \sin(\theta_i+\theta_j) \end{pmatrix} = \eta_{ij}^\top \begin{pmatrix} \cos\theta_i\cos\theta_j \\ \cos\theta_i\sin\theta_j \\ \sin\theta_i\cos\theta_j \\ \sin\theta_i\sin\theta_j \end{pmatrix} = \frac{1}{2}\eta_{ij}^\top \begin{pmatrix} \cos(\theta_i+\theta_j)+\cos(\theta_i-\theta_j) \\ \sin(\theta_i+\theta_j)-\sin(\theta_i-\theta_j) \\ \sin(\theta_i+\theta_j)+\sin(\theta_i-\theta_j) \\ \cos(\theta_i-\theta_j)-\cos(\theta_i+\theta_j) \end{pmatrix}
$$

and hence

$$
\phi_{ij} = \frac{1}{2}(\eta_{ij,1}+\eta_{ij,4}, -\eta_{ij,2}+\eta_{ij,3}, \eta_{ij,1}-\eta_{ij,4}, \eta_{ij,2}+\eta_{ij,3}).
$$

The proof then follows by comparing the pdfs. □

S1.5 Proof of Theorem 3

Proof. Let $\Sigma = \text{cov}\{(\mathfrak{Re}X, \mathfrak{Im}X), (\mathfrak{Re}X, \mathfrak{Im}X)\}$. From Picinbono (1996),

$$
\Sigma_{11} = \frac{1}{2}\mathfrak{Re}\Gamma, \quad \Sigma_{12} = -\frac{1}{2}\mathfrak{Im}\Gamma, \quad \Sigma_{21} = \frac{1}{2}\mathfrak{Im}\Gamma, \text{ and } \quad \Sigma_{22} = \frac{1}{2}\mathfrak{Re}\Gamma,
$$

so Σ is of the the form specified in Lemma S6 as is Σ^{-1}. Note that Σ^{-1} meets the conditions of Theorem 2 since $(\Sigma^{-1})_{ii} = (\Sigma^{-1})_{i+d,i+d}$ and $(\Sigma^{-1})_{i,i+d} = (\Sigma^{-1})_{i+d,i} = 0$. The second fact follows from the observation that the diagonal of any antisymmetric matrix must be zero. Therefore, $\Theta \mid R = r \sim \mathcal{TG}(\eta)$, where

$$\eta_{ij} = \frac{-1}{2} r_i r_j (\Sigma^{-1}_{i,j}, \Sigma^{-1}_{i,j+d}, \Sigma^{-1}_{i+d,j}, \Sigma^{-1}_{i+d,j+d}).$$

Therefore,

$$\begin{aligned}\phi_{ij} &= \frac{1}{2}(\eta_{ij,1} + \eta_{ij,4}, -\eta_{ij,2} + \eta_{ij,3}, \eta_{ij,1} - \eta_{ij,4}, \eta_{ij,2} + \eta_{ij,3}) \\ &= \frac{1}{2}\left(-\frac{1}{2} r_i r_j\right)(2(\Sigma^{-1})_{ij}, -2(\Sigma^{-1})_{i,j+d}, 0, 0),\end{aligned}$$

which is exactly the form desired (i.e., $\phi_{ij,3} = \phi_{ij,4} = 0$). All parameters are zero except for those multiplied by statistics of the form $\cos(\theta_i - \theta_j)$ or $\sin(\theta_i - \theta_j)$. Clearly, these terms are invariant to a circular shift since $(\theta_i + \epsilon) - (\theta_j + \epsilon) = \theta_i - \theta_j$ for all ϵ. Therefore, Θ is circularly symmetric. □

S1.6 Proof of Corollary 4

Proof. We have the same setup as in Theorem 3, so ϕ_{ij} has the same form. By Theorem 2, $\eta_i = r_i((\Sigma^{-1}\mu)_i, (\Sigma^{-1}\mu)_{i+d})$ and, by Lemma S7, $\eta_i = \phi_i$. Therefore, $\mu = \mathrm{E}((\mathfrak{Re} X, \mathfrak{Im} X)) = 0$ and $\phi_i = 0$. □

S1.7 Proof of Theorem 6

Proof. Let $\mathcal{I} = [d] \setminus \{i, j\}$. We note that

$$\begin{aligned}&\operatorname{argmin}_{Y \in \mathcal{L}(\widetilde{X}_\mathcal{I})} \mathrm{E}\{(Y - X_i)^H (Y - X_i)\} = \operatorname{argmin}_{Y \in \mathcal{L}(\widetilde{X}_\mathcal{I})} \left\{\mathrm{E}(Y^H Y) - \mathrm{E}(Y^H X_i) - \mathrm{E}(X_i^H Y)\right\} \\ &= \operatorname{argmin}_{\alpha + \beta^H \widetilde{X}_\mathcal{I}} E\left\{(\alpha + \beta^H \widetilde{X}_\mathcal{I})^H (\alpha + \beta^H \widetilde{X}_\mathcal{I}) - \alpha^H X_i - \widetilde{X}_\mathcal{I}^H \beta X_i - X_i^H \alpha - X_i^H \beta^H \widetilde{X}_\mathcal{I}\right\}.\end{aligned}$$

It turns out that $\beta^* = \operatorname{cov}(\widetilde{X}_\mathcal{I})^{-1} \operatorname{cov}(\widetilde{X}_\mathcal{I}, X_i)$ and $\alpha^* = \mathrm{E}(X_i) - \beta^{*H} \mathrm{E}(\widetilde{X}_\mathcal{I})$ are the optimal parameters for the above optimization problem. Therefore

$$\begin{aligned}&\operatorname{cov}\{X_i - \operatorname{proj}_{\mathcal{L}(\widetilde{X}_\mathcal{I})}(X_i), X_j - \operatorname{proj}_{\mathcal{L}(\widetilde{X}_\mathcal{I})}(X_j)\} \\ &\quad = \operatorname{cov}(X_i - [\mathrm{E}(X_i) + \operatorname{cov}(X_i, \widetilde{X}_\mathcal{I}) \operatorname{var}(\widetilde{X}_\mathcal{I})^{-1}\{\widetilde{X}_\mathcal{I} - \mathrm{E}(\widetilde{X}_\mathcal{I})\}], \\ &\qquad\quad X_j - [\mathrm{E}(X_j) + \operatorname{cov}(X_j, \widetilde{X}_\mathcal{I}) \operatorname{var}(\widetilde{X}_\mathcal{I})^{-1}\{\widetilde{X}_\mathcal{I} - \mathrm{E}(\widetilde{X}_\mathcal{I})\}]) \\ &\quad = \operatorname{cov}(X_i, X_j) - \operatorname{cov}(X_i, \widetilde{X}_\mathcal{I}) \operatorname{var}(\widetilde{X}_\mathcal{I})^{-1} \operatorname{cov}(\widetilde{X}_\mathcal{I}, X_j).\end{aligned}$$

Let $\widetilde{X}_{\{i,j\}} = (X_i, \overline{X}_i, X_j, \overline{X}_j)^\top$. Consider the following block matrix, which is a version of $\widetilde{\Gamma}$ with rearranged rows and columns

$$U = \begin{pmatrix} \operatorname{var}(\widetilde{X}_{\{i,j\}}) & \operatorname{cov}(\widetilde{X}_{\{i,j\}}, \widetilde{X}_\mathcal{I}) \\ \operatorname{cov}(\widetilde{X}_\mathcal{I}, \widetilde{X}_{\{i,j\}}) & \operatorname{var}(\widetilde{X}_\mathcal{I}) \end{pmatrix}.$$

We can use the Schur complement to note that, if we invert U, the resulting upper-left block is given by $A = \operatorname{var}(\widetilde{X}_{\{i,j\}}) - \operatorname{cov}(\widetilde{X}_{\{i,j\}}, \widetilde{X}_\mathcal{I}) \operatorname{var}(\widetilde{X}_\mathcal{I})^{-1} \operatorname{cov}(\widetilde{X}_\mathcal{I}, \widetilde{X}_{\{i,j\}})$. Since U is just a rearranged version of $\widetilde{\Gamma}$, we can observe that $\widetilde{\Gamma}^{-1}$ will contain the entries of A. In particular, note that $\widetilde{\Gamma}_{ij} = \operatorname{cov}(X_i, X_j) = U_{1,3}$. Therefore,

$$\widetilde{\Gamma}^{-1}_{i,j} = A_{1,3} = \operatorname{cov}(X_i, X_j) - \operatorname{cov}(X_i, \widetilde{X}_{\mathcal{I}}) \operatorname{var}(\widetilde{X}_{\mathcal{I}})^{-1} \operatorname{cov}(\widetilde{X}_{\mathcal{I}}, X_j)$$
$$= \operatorname{cov}\{X_i - \operatorname{proj}_{\mathcal{L}(\widetilde{X}_{\mathcal{I}})}(X_i), X_j - \operatorname{proj}_{\mathcal{L}(\widetilde{X}_{\mathcal{I}})}(X_j)\},$$

and it immediately follows that

$$P_{ij} = (\widetilde{\Gamma}^{-1}_{ii})^{-1/2} \widetilde{\Gamma}^{-1}_{ij} (\widetilde{\Gamma}^{-1}_{jj})^{-1/2} = \rho(X_i, X_j \mid X_{[d]\setminus\{i,j\}}),$$

as desired. A very similar argument establishes that $P_{i,j+d} = \widetilde{\rho}(X_i, X_j \mid X_{[d]\setminus\{i,j\}})$.

□

S1.8 Proof of Theorem 7

Proof. In the proof of Theorem 6 we show that

$$\operatorname{cov}\{X_i - \operatorname{proj}_{\mathcal{L}(\widetilde{X}_{\mathcal{I}})}(X_i), X_j - \operatorname{proj}_{\mathcal{L}(\widetilde{X}_{\mathcal{I}})}(X_j)\} = \operatorname{cov}(X_i, X_j) - \operatorname{cov}(X_i, \widetilde{X}_{\mathcal{I}}) \operatorname{var}(\widetilde{X}_{\mathcal{I}})^{-1} \operatorname{cov}(\widetilde{X}_{\mathcal{I}}, X_j).$$

We use this result below. Let

$$J_d = \begin{pmatrix} I_d & \iota I_d \\ I_d & -\iota I_d \end{pmatrix},$$

where I_d is the d-dimensional identity matrix. Note that $J_d^{-1} = J_d^H$ and $\widetilde{X} = J_d \cdot (\mathfrak{Re}X, \mathfrak{Im}X)^\top$. Then

$$\operatorname{cov}(X_i, X_j \mid X_{\mathcal{I}}) = \operatorname{cov}\left\{(1 \;\; \iota)\begin{pmatrix}\mathfrak{Re}X_i \\ \mathfrak{Im}X_i\end{pmatrix}, (1 \;\; \iota)\begin{pmatrix}\mathfrak{Re}X_j \\ \mathfrak{Im}X_j\end{pmatrix} \middle| X_{\mathcal{I}}\right\}$$
$$= (1 \;\; \iota) \operatorname{cov}\left\{\begin{pmatrix}\mathfrak{Re}X_i \\ \mathfrak{Im}X_i\end{pmatrix}, \begin{pmatrix}\mathfrak{Re}X_j \\ \mathfrak{Im}X_j\end{pmatrix} \middle| X_{\mathcal{I}}\right\} \begin{pmatrix}1 \\ -\iota\end{pmatrix}.$$

Recall that $X \sim \mathcal{CN}(m, \Gamma, C)$ if $(\mathfrak{Re}X, \mathfrak{Im}X) \sim \mathcal{N}(\mu, \Sigma)$. Based on the Schur complement formula, the conditional covariance of a real-valued Gaussian distribution is

$$\begin{aligned}
&\operatorname{cov}(X_i, X_j \mid X_{\mathcal{I}}) \\
&\quad= (1 \;\; \iota) \operatorname{cov}\left\{\begin{pmatrix}\mathfrak{Re}X_i \\ \mathfrak{Im}X_i\end{pmatrix}, \begin{pmatrix}\mathfrak{Re}X_j \\ \mathfrak{Im}X_j\end{pmatrix} \middle| \begin{pmatrix}\mathfrak{Re}X_{\mathcal{I}} \\ \mathfrak{Im}X_{\mathcal{I}}\end{pmatrix}\right\} \begin{pmatrix}1 \\ -\iota\end{pmatrix} \\
&\quad= (1 \;\; \iota) \operatorname{cov}\left\{\begin{pmatrix}\mathfrak{Re}X_i \\ \mathfrak{Im}X_i\end{pmatrix}, \begin{pmatrix}\mathfrak{Re}X_j \\ \mathfrak{Im}X_j\end{pmatrix}\right\} \begin{pmatrix}1 \\ -\iota\end{pmatrix} \\
&\qquad - (1 \;\; \iota) \operatorname{cov}\left\{\begin{pmatrix}\mathfrak{Re}X_i \\ \mathfrak{Im}X_i\end{pmatrix}, \begin{pmatrix}\mathfrak{Re}X_{\mathcal{I}} \\ \mathfrak{Im}X_{\mathcal{I}}\end{pmatrix}\right\} \left(\operatorname{var}\begin{pmatrix}\mathfrak{Re}X_{\mathcal{I}} \\ \mathfrak{Im}X_{\mathcal{I}}\end{pmatrix}\right)^{-1} \operatorname{cov}\left\{\begin{pmatrix}\mathfrak{Re}X_{\mathcal{I}} \\ \mathfrak{Im}X_{\mathcal{I}}\end{pmatrix}, \begin{pmatrix}\mathfrak{Re}X_j \\ \mathfrak{Im}X_j\end{pmatrix}\right\} \begin{pmatrix}1 \\ -\iota\end{pmatrix}.
\end{aligned}$$

Similarly,

$$\operatorname{var}(\widetilde{X}_{\mathcal{I}}) = \operatorname{var}\left\{J_{|\mathcal{I}|}\begin{pmatrix}\mathfrak{Re}X_{\mathcal{I}} \\ \mathfrak{Im}X_{\mathcal{I}}\end{pmatrix}\right\} = J_{|\mathcal{I}|} \operatorname{var}\left\{\begin{pmatrix}\mathfrak{Re}X_{\mathcal{I}} \\ \mathfrak{Im}X_{\mathcal{I}}\end{pmatrix}\right\} J_{|\mathcal{I}|}^H$$

and

$$\operatorname{cov}(X_i, \widetilde{X}_{\mathcal{I}}) = \operatorname{cov}\left\{(1 \;\; \iota)\begin{pmatrix}\mathfrak{Re}X_i \\ \mathfrak{Im}X_i\end{pmatrix}, J_{|\mathcal{I}|}\begin{pmatrix}\mathfrak{Re}X_{\mathcal{I}} \\ \mathfrak{Im}X_{\mathcal{I}}\end{pmatrix}\right\} = (1 \;\; \iota) \operatorname{cov}\left\{\begin{pmatrix}\mathfrak{Re}X_i \\ \mathfrak{Im}X_i\end{pmatrix}, \begin{pmatrix}\mathfrak{Re}X_{\mathcal{I}} \\ \mathfrak{Im}X_{\mathcal{I}}\end{pmatrix}\right\} J_{|\mathcal{I}|}^H.$$

We then see that

$$\operatorname{var}(\widetilde{X}_{\mathcal{I}})^{-1} = \left\{ J_{|\mathcal{I}|} \operatorname{var} \begin{pmatrix} \Re\mathfrak{e} X_{\mathcal{I}} \\ \Im\mathfrak{m} X_{\mathcal{I}} \end{pmatrix} J_{|\mathcal{I}|}^H \right\}^{-1} = J_{|I|} \left\{ \operatorname{var} \begin{pmatrix} \Re\mathfrak{e} X_{\mathcal{I}} \\ \Im\mathfrak{m} X_{\mathcal{I}} \end{pmatrix} \right\}^{-1} J_{|I|}^H$$

since $J_d^H = J_d^{-1}$. Therefore,

$$\begin{aligned}
&\operatorname{cov}\left\{ X_i - \operatorname{proj}_{\mathcal{L}(\widetilde{X}_{\mathcal{I}})}(X_i), X_j - \operatorname{proj}_{\mathcal{L}(\widetilde{X}_{\mathcal{I}})}(X_j) \right\} \\
&\quad = \operatorname{cov}(X_i, X_j) - \operatorname{cov}(X_i, \widetilde{X}_{\mathcal{I}}) \{\operatorname{var}(\widetilde{X}_{\mathcal{I}})\}^{-1} \operatorname{cov}(\widetilde{X}_{\mathcal{I}}, X_j) \\
&\quad = (1 \quad \iota) \operatorname{cov}\left\{ \begin{pmatrix} \Re\mathfrak{e} X_i \\ \Im\mathfrak{m} X_i \end{pmatrix}, \begin{pmatrix} \Re\mathfrak{e} X_j \\ \Im\mathfrak{m} X_j \end{pmatrix} \right\} \begin{pmatrix} 1 \\ -\iota \end{pmatrix} \\
&\qquad - (1 \quad \iota) \operatorname{cov}\left\{ \begin{pmatrix} \Re\mathfrak{e} X_i \\ \Im\mathfrak{m} X_i \end{pmatrix}, \begin{pmatrix} \Re\mathfrak{e} X_{\mathcal{I}} \\ \Im\mathfrak{m} X_{\mathcal{I}} \end{pmatrix} \right\} J_{|\mathcal{I}|}^H J_{|\mathcal{I}|} \left\{ \operatorname{var} \begin{pmatrix} \Re\mathfrak{e} X_{\mathcal{I}} \\ \Im\mathfrak{m} X_{\mathcal{I}} \end{pmatrix} \right\}^{-1} J_{|\mathcal{I}|}^H J_{|\mathcal{I}|} \\
&\qquad\qquad \times \operatorname{cov}\left\{ \begin{pmatrix} \Re\mathfrak{e} X_{\mathcal{I}} \\ \Im\mathfrak{m} X_{\mathcal{I}} \end{pmatrix}, \begin{pmatrix} \Re\mathfrak{e} X_i \\ \Im\mathfrak{m} X_i \end{pmatrix} \right\} \begin{pmatrix} 1 \\ -\iota \end{pmatrix} \\
&\quad = (1 \quad \iota) \operatorname{cov}\left\{ \begin{pmatrix} \Re\mathfrak{e} X_i \\ \Im\mathfrak{m} X_i \end{pmatrix}, \begin{pmatrix} \Re\mathfrak{e} X_j \\ \Im\mathfrak{m} X_j \end{pmatrix} \right\} \begin{pmatrix} 1 \\ -\iota \end{pmatrix} \\
&\qquad - (1 \quad \iota) \operatorname{cov}\left\{ \begin{pmatrix} \Re\mathfrak{e} X_i \\ \Im\mathfrak{m} X_i \end{pmatrix}, \begin{pmatrix} \Re\mathfrak{e} X_{\mathcal{I}} \\ \Im\mathfrak{m} X_{\mathcal{I}} \end{pmatrix} \right\} \left\{ \operatorname{var} \begin{pmatrix} \Re\mathfrak{e} X_{\mathcal{I}} \\ \Im\mathfrak{m} X_{\mathcal{I}} \end{pmatrix} \right\}^{-1} \operatorname{cov}\left\{ \begin{pmatrix} \Re\mathfrak{e} X_{\mathcal{I}} \\ \Im\mathfrak{m} X_{\mathcal{I}} \end{pmatrix}, \begin{pmatrix} \Re\mathfrak{e} X_j \\ \Im\mathfrak{m} X_j \end{pmatrix} \right\} \begin{pmatrix} 1 \\ -\iota \end{pmatrix} \\
&= \operatorname{cov}(X_i, X_j \mid X_{\mathcal{I}}).
\end{aligned}$$

Letting $j = i$, we can see that, likewise, the conditional variance and the variance of the terms $X_i - \operatorname{proj}_{\mathcal{L}(X_{\mathcal{I}})}$ are the same. Thus, the partial correlation and the conditional correlation are the same. We can apply a very similar strategy to show that the pseudo partial correlation and the pseudo conditional correlation are the same. □

S1.9 Proof of Theorem 8

Proof. Let $\mathcal{I} = [d] \setminus \{i, j\}$. Assume $\operatorname{cov}(X_i, X_j \mid X_{\mathcal{I}}) = \operatorname{pcov}(X_i, X_j \mid X_{\mathcal{I}}) = 0$. Suppose that $J_d = \begin{pmatrix} I_d & \iota I_d \\ I_d & -\iota I_d \end{pmatrix}$. Then

$$\begin{aligned}
&J_1 \operatorname{cov}\left\{ \begin{pmatrix} \Re\mathfrak{e} X_i \\ \Im\mathfrak{m} X_i \end{pmatrix}, \begin{pmatrix} \Re\mathfrak{e} X_j \\ \Im\mathfrak{m} X_j \end{pmatrix} \middle| X_{\mathcal{I}} \right\} J_1^H \\
&\quad = \operatorname{cov}\left(\widetilde{X}_i, \widetilde{X}_j \middle| X_{\mathcal{I}} \right) = \begin{pmatrix} \operatorname{cov}(X_i, X_j \mid X_{\mathcal{I}}) & \operatorname{pcov}(X_i, X_j \mid X_{\mathcal{I}}) \\ \operatorname{pcov}(X_j, X_i \mid X_{\mathcal{I}}) & \operatorname{cov}(X_j, X_i \mid X_{\mathcal{I}}) \end{pmatrix} = 0
\end{aligned}$$

and so

$$\operatorname{cov}\left\{ \begin{pmatrix} \Re\mathfrak{e} X_i \\ \Im\mathfrak{m} X_i \end{pmatrix}, \begin{pmatrix} \Re\mathfrak{e} X_j \\ \Im\mathfrak{m} X_j \end{pmatrix} \middle| X_{\mathcal{I}} \right\} = 0$$

since J_1 is invertible (namely $J_1^{-1} = J_1^H$). By properties of the real normal distribution, this implies that

$$\begin{pmatrix} \Re\mathfrak{e} X_i \\ \Im\mathfrak{m} X_i \end{pmatrix} \perp\!\!\!\perp \begin{pmatrix} \Re\mathfrak{e} X_j \\ \Im\mathfrak{m} X_j \end{pmatrix} \middle| X_{\mathcal{I}} \implies X_i \perp\!\!\!\perp X_j \mid X_{\mathcal{I}},$$

as desired. The other direction is trivially true from the definition of covariance. □

S2 Examples Comparing PLV, Amplitude Correlation, and Complex Correlation

It is helpful to reconsider the single-frequency signals $X_i(t) = R_i \exp\{\iota(\Theta_i + 2\pi\omega_0 t)\}$ discussed in Section 4.1. Writing the purely angular factor in the rotational phase correlation (which is the factor estimated by PLV) as

$$\mathrm{PLV}_{12}(\omega) = \left|\mathrm{E}\left(e^{\iota(\Theta_1-\Theta_2)}\right)\right|,$$

in the case that $(R_1, R_2) \perp\!\!\!\perp (\Theta_1, \Theta_2)$, (4) and (5) give

$$f_{ii}(\omega) = \mathrm{E}(R_i^2) = \mathrm{var}(R_i) + \mathrm{E}(R_i)^2$$

and

$$|f_{12}(\omega)| = \mathrm{E}(R_1R_2)\left|\mathrm{E}\{e^{\iota(\Theta_1-\Theta_2)}\}\right| = (\mathrm{cov}(R_1, R_2) + \mathrm{E}(R_1)\,\mathrm{E}(R_2))\mathrm{PLV}_{12}(\omega).$$

Letting $\mathrm{ICV}(R_i) = \mathrm{E}(R_i)/\sqrt{\mathrm{var}(R_i)}$ be the inverse coefficient of variation of R_i for $i \in \{1, 2\}$, the coherence, which is the absolute value of the complex correlation, at frequency ω is

$$\begin{aligned}\tau_{12}(\omega) = |\mathrm{corr}\{X_1(t), X_2(t)\}| &= \frac{\mathrm{cov}(R_1, R_2) + \mathrm{E}(R_1)\,\mathrm{E}(R_2)}{\sqrt{(\mathrm{var}(R_1) + \mathrm{E}(R_1)^2)(\mathrm{var}(R_2) + \mathrm{E}(R_2)^2)}}\mathrm{PLV}_{12}(\omega)\\ &= \frac{\mathrm{corr}(R_1, R_2) + \mathrm{ICV}(R_1)\mathrm{ICV}(R_2)}{\sqrt{(1 + \mathrm{ICV}(R_1)^2)(1 + \mathrm{ICV}(R_2)^2)}}\mathrm{PLV}_{12}(\omega).\end{aligned}$$

Thus, in this independent case, if the amplitude is highly concentrated so that the ICV is large, coherence and PLV will be roughly equal, but if the ICV is small or moderate, they will differ.

We can observe instances in which the PLV provides stronger evidence of an association than complex correlation. For instance, assume that $\Theta_1 \sim \mathcal{U}[-\pi, \pi)$ and $\Theta_2 = \Theta_1 + \epsilon$ (where ϵ is a von Mises random variable concentrated around zero), and let $R_1 = R_2 = |\Theta_1 - \Theta_2|$. In this case, PLV is large, but the complex covariance will be relatively small due to the specific form of R_1 and R_2.

Finally, there are numerous cases in which complex covariance (and thus coherence in the stationary, band-pass-filtered case) may reveal information about associations not apparent by examining either PLV or amplitude correlation. We show some of these cases in Table S1. Overall, coherence and PLV can assess phase interaction differently, and in many applications the situations discussed here may usefully inform the choice of one over the other.

Explanation	Setting	Result
PLV, amplitude covariance and complex means are zero, but complex covariance is nonzero	$\Theta_1, \Theta_2 \overset{iid}{\sim} \mathcal{U}[-\pi, \pi)$ $R_1 = \Theta_2$ $R_2 = \Theta_1$	$\mathrm{PLV}(\Theta_1, \Theta_2) = 0$ $\mathrm{cov}(R_1, R_2) = 0$ $\mathrm{E}(X_1) = \mathrm{E}(X_2) = 0$ $\lvert\mathrm{cov}(X_1, X_2)\rvert > 0$ $\lvert\mathrm{pcov}(X_1, X_2)\rvert > 0$
PLV, amplitude covariance and complex means are zero, but complex covariance is nonzero	$\Theta_1, \Theta_2 \overset{iid}{\sim} \mathcal{U}[-\pi, \pi)$ $R_1 = \Theta_1 + \Theta_2 \mod 2\pi$ $R_2 = \Theta_1 - \Theta_2 \mod 2\pi$	$\mathrm{PLV}(\Theta_1, \Theta_2) = 0$ $\mathrm{cov}(R_1, R_2) = 0$ $\mathrm{E}(X_1) = \mathrm{E}(X_2) = 0$ $\lvert\mathrm{cov}(X_1, X_2)\rvert > 0$
PLV, and amplitude covariance are zero, but complex means and complex covariance are nonzero	$\Theta_1, \Theta_2 \overset{iid}{\sim} \mathcal{U}[-\pi, \pi)$ $R_1 = \Theta_1 + \Theta_2$ $R_2 = \Theta_1 - \Theta_2 + 2\pi$	$\mathrm{PLV}(\Theta_1, \Theta_2) = 0$ $\mathrm{cov}(R_1, R_2) = 0$ $\mathrm{E}(X_1) \neq 0, \quad \mathrm{E}(X_2) \neq 0$ $\lvert\mathrm{cov}(X_1, X_2)\rvert > 0$
PLV, amplitude covariance are complex covariance are zero but complex means are nonzero	$\Theta_1, \Theta_2 \overset{iid}{\sim} \mathcal{U}[-\pi, \pi)$ $R_1 = \Theta_1$ $R_2 = \Theta_2$	$\mathrm{PLV}(\Theta_1, \Theta_2) = 0$ $\mathrm{cov}(R_1, R_2) = 0$ $\mathrm{E}(X_1) \neq 0, \quad \mathrm{E}(X_2) \neq 0$ $\lvert\mathrm{cov}(X_1, X_2)\rvert = 0$ $\lvert\mathrm{pcov}(X_1, X_2)\rvert = 0$

Table S1: Examples of various ways in which complex covariance and means can capture associations not captured by PLV or amplitude correlation. In all cases, the variables are independent unless otherwise stated. We define $X_1 = R_1 \exp(\iota\Theta_1)$ and $X_2 = R_2 \exp(\iota\Theta_2)$.

S3 Supplemental Figures

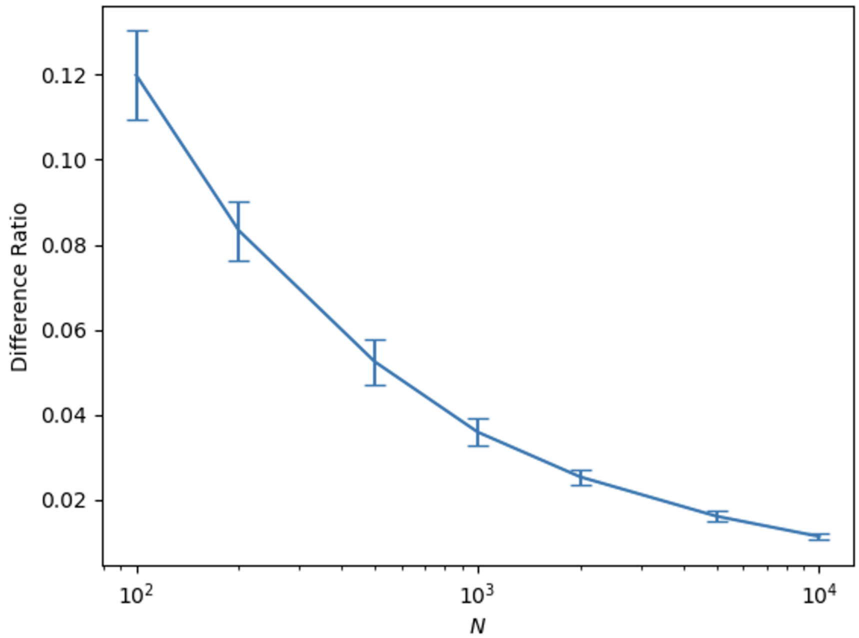

Figure S1: Estimation error as a function of sample size. To construct a simulation similar to the data analyzed in Section 6, we first fit a model to the dataset described in Section 6 and then simulate synthetic datasets by using those estimated parameters. We then estimate parameters for the synthetic datasets for a variety of values of N, the size of the simulated dataset. We then compute the estimation error as a function of N. For any matrix A, let $|A| = \sum_{ij} |A|_{ij}$, let $\widehat{\Gamma}$ denote the estimated latent covariance matrix, and let Γ denote the true latent covariance matrix. Our metric for estimation error is the ratio $|\widehat{\Gamma} - \Gamma|/|\Gamma|$. Error bars denote 95% confidence intervals for the error.

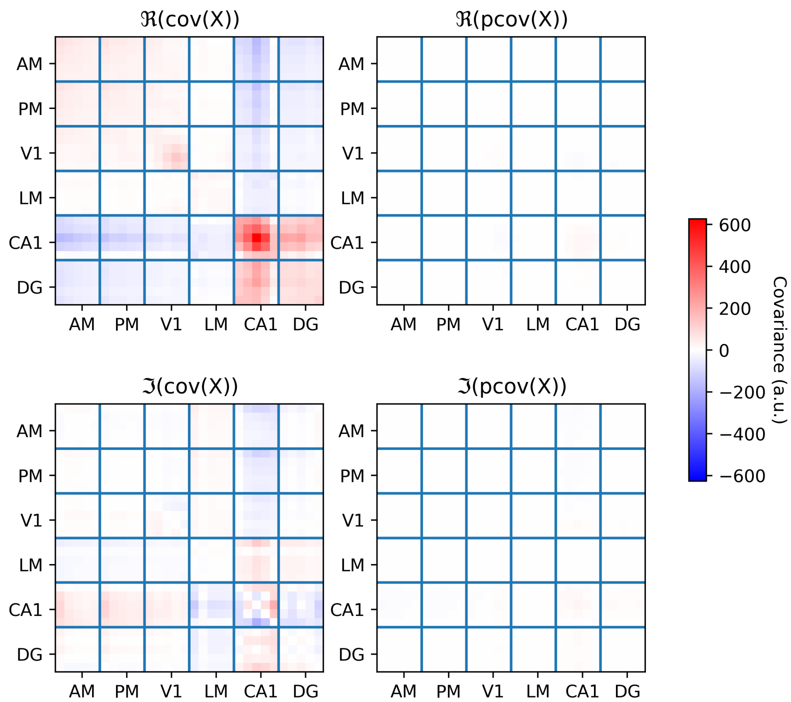

Figure S2: Elements of the covariance and pseudo-covariance matrices for the LFP data. Both the covariance and pseudo-covariance matrices are plotted, denoted by $\text{cov}(X)$ and $\text{pcov}(X)$, respectively (units are arbitrary). In a proper complex normal distribution, the matrix $\text{pcov}(X)$ is exactly equal to zero. In this figure, we observe that the entries in both $\mathfrak{Re}\{\text{pcov}(X)\}$ and $\mathfrak{Im}\{\text{pcov}(X)\}$ are much smaller than those in $\mathfrak{Re}\{\text{cov}(X)\}$ and $\mathfrak{Im}\{\text{cov}(X)\}$, which suggests that the assumption of propriety may be reasonable.