SUPPLEMENT TO "TORUS GRAPHS FOR MULTIVARIATE PHASE COUPLING ANALYSIS"

By Natalie Klein[‡,§,*], Josue Orellana[¶,§,*], Scott Brincat[‖], Earl K. Miller[‖], and Robert E. Kass[‡,¶,§,†]

Department of Statistics and Data Science, Carnegie Mellon University[‡], Machine Learning Department, Carnegie Mellon University[§], Center for the Neural Basis of Cognition[¶], and Department of Brain and Cognitive Science, Massachusetts Institute of Technology[‖]

S1. Proof of Theorem 2.1 (Torus graph model). To derive the appropriate first- and second-order sufficient statistics that correspond to first circular moments and to circular covariances, we first represent the angles as unit modulus complex random variables (where i is the imaginary unit):

$$Z_j = e^{iX_j}.$$

The first circular moment is defined as

$$E[Z_j] = R_j e^{i\mu_j}$$

where μ_j is the mean direction and R_j is the resultant length, so the corresponding complex first order sufficient statistic for a single observation is simply

$$z_j = \cos(x_j) + i\sin(x_j)$$

which may be described as a real-valued sufficient statistic vector

$$\mathbf{S}_j^1(x_j) = [\cos(x_j), \sin(x_j)]^T.$$

When considering second-order interactions between complex variables, there are two types of covariance (Schreier and Scharf, 2010, Ch. 2.2). Rotational covariance between Z_j and Z_k is described by

$$E[(e^{iX_j} - R_j e^{i\mu_j})\overline{(e^{iX_k} - R_k e^{i\mu_k})}] = E[e^{i(X_j - X_k)}] - R_j R_k e^{i(\mu_j - \mu_k)},$$

[*] N.K. and J.O. contributed equally to this work.

[†] To whom correspondence should be addressed. E-mail: kass@stat.cmu.edu

Keywords and phrases: graphical models, circular statistics, network analysis

where $\overline{Z_k}$ is the complex conjugate, while reflectional covariance is described by

$$E[(e^{iX_j} - e^{i\mu_j})(e^{iX_k} - e^{i\mu_k})] = E[e^{i(X_j+X_k)}] - R_j R_k e^{i(\mu_j+\mu_k)}.$$

This shows that in addition to the first-order statistics, we additionally need two more complex sufficient statistics to describe the second-order behavior:

$$\begin{aligned} e^{i(x_j-x_k)} &= \cos(x_j - x_k) + i\sin(x_j - x_k) \\ e^{i(x_j+x_k)} &= \cos(x_j + x_k) + i\sin(x_j + x_k). \end{aligned}$$

These may be collected into a real-valued vector

$$\mathbf{S}^2_{jk}(x_j, x_k) = [\cos(x_i - x_j),\ \sin(x_i - x_j),\ \cos(x_i + x_j),\ \sin(x_i + x_j)]^T.$$

Therefore, the canonical exponential family distribution given the first circular moments of each variable and the complete second-order interactions (rotational and reflectional) between each variable coincides with that given in Klein et al. 2019, Equation 2.3:

$$p(\mathbf{x}) \propto \exp\left\{\sum_{j=1}^{d} \boldsymbol{\phi}_j^T \mathbf{S}^1_j(x_j) + \sum_{j<k} \boldsymbol{\phi}_{jk}^T \mathbf{S}^2_{jk}(x_j, x_k)\right\} \tag{S1.1}$$

$$= \exp\left\{\sum_{j=1}^{d} \boldsymbol{\phi}_j^T \begin{bmatrix}\cos(x_j)\\ \sin(x_j)\end{bmatrix} + \sum_{j<k} \boldsymbol{\phi}_{jk}^T \begin{bmatrix}\cos(x_j - x_k)\\ \sin(x_j - x_k)\\ \cos(x_j + x_k)\\ \sin(x_j + x_k)\end{bmatrix}\right\}. \tag{S1.2}$$

Thus, the torus graph model is maximum entropy with respect to constraints on the expected values of the sufficient statistics, that is, the circular first moments and complex covariances Wainwright et al. (2008). We note that, similar to the multivariate Gaussian distribution, the torus graph model only contains sufficient statistics for circular first moments and covariances, but it does not contain sufficient statistics corresponding to the second circular moment of a single angle X_j (that is, it does not include interactions of the form $Z_j Z_j$); such a model was recently explored in Navarro, Frellsen and Turner (2017).

The maximum entropy motivation for this model also offers some intuition for interpretation of the parameters; in particular, we see that the subvector $\boldsymbol{\phi}_{jk,1:2}$ corresponds to rotational covariance while the subvector $\boldsymbol{\phi}_{jk,3:4}$ corresponds to reflectional covariance. However, the magnitude of each parameter is difficult to interpret directly because it depends not only on the

covariance but also the marginal concentration of each variable (which is related to the resultant lengths R_j and R_k) as well as the sum or difference of the mean directions.

S2. Reparameterization to compare to previous work. While the canonical exponential family form in Equation S1.2 is useful for understanding the maximum entropy constraints of the model and for deriving score matching estimators, it does not immediately appear similar to previous work in multivariate circular statistics (such as the *sine model*). To obtain another form that is easier to compare to previous work, we begin with an alternate parameterization that is similar to the *sine model*, then show how it can be transformed into our parameterization. Crucially, this transformation can also be reversed to potentially aid in interpretation of parameters.

Consider the *mean-centered parameterization*

$$p(\mathbf{x};\boldsymbol{\theta}) \propto \exp\left\{\sum_{j=1}^{d} \kappa_j \cos(x_j-\mu_j) + \sum_{j<k} \begin{bmatrix}\lambda_{jk}^{cc}\\ \lambda_{jk}^{cs}\\ \lambda_{jk}^{sc}\\ \lambda_{jk}^{ss}\end{bmatrix}^T \begin{bmatrix}\cos(x_j-\mu_j)\cos(x_k-\mu_k)\\ \cos(x_j-\mu_j)\sin(x_k-\mu_k)\\ \sin(x_j-\mu_j)\cos(x_k-\mu_k)\\ \sin(x_j-\mu_j)\sin(x_k-\mu_k)\end{bmatrix}\right\}$$

where the parameters are $\boldsymbol{\theta} = [\boldsymbol{\mu}, \boldsymbol{\kappa}, \boldsymbol{\lambda^{cc}}, \boldsymbol{\lambda^{cs}}, \boldsymbol{\lambda^{sc}}, \boldsymbol{\lambda^{ss}}]$ with the interpretation that $\mu_j \in [0, 2\pi)$ is the mean direction of x_j, $\kappa_j > 0$ is the marginal concentration of x_j, and the λ parameters control interactions between angles.

In the univariate terms, we use trigonometric sum and difference formulas to rewrite

$$\kappa_j \cos(x_j-\mu_j) = \kappa_j \cos(\mu_j)\cos(x_j) + \kappa_j \sin(\mu_j)\sin(x_j)$$

so that in the parameterization of Equation S1.2,

$$\boldsymbol{\phi}_j = \begin{bmatrix}\kappa_j \cos(\mu_j)\\ \kappa_j \sin(\mu_j)\end{bmatrix}.$$

Therefore, we can clearly calculate $\boldsymbol{\phi}_j$ for given κ_j and μ_j, and also, given $\boldsymbol{\phi}_j$, we have (using the Pythagorean theorem and definition of tangent)

$$\kappa_j = \sqrt{\boldsymbol{\phi}_{j,1}^2 + \boldsymbol{\phi}_{j,2}^2}$$
$$\mu_j = \arctan\left(\frac{\boldsymbol{\phi}_{j,2}}{\boldsymbol{\phi}_{j,1}}\right).$$

Thus we have demonstrated a diffeomorphism between the two parameterizations for the marginal terms.

Similarly, for the pairwise coupling terms, we consider without loss of generality the pair $\{X_j, X_k\}$ but for simplicity drop subscripts on the λ parameters; using trigonometric sum and difference identities and simplifying, we find the pairwise coupling term is

$$\frac{1}{2}\begin{bmatrix} (\lambda^{cc}+\lambda^{ss})\cos(\mu_j-\mu_k)+(\lambda^{cs}-\lambda^{sc})\sin(\mu_j-\mu_k) \\ (\lambda^{sc}-\lambda^{cs})\cos(\mu_j-\mu_k)+(\lambda^{cc}+\lambda^{ss})\sin(\mu_j-\mu_k) \\ (\lambda^{cc}-\lambda^{ss})\cos(\mu_j+\mu_k)+(-\lambda^{cs}-\lambda^{sc})\sin(\mu_j+\mu_k) \\ (\lambda^{cs}+\lambda^{sc})\cos(\mu_j+\mu_k)+(\lambda^{cc}-\lambda^{ss})\sin(\mu_j+\mu_k) \end{bmatrix}^T \begin{bmatrix} \cos(x_j-x_k) \\ \sin(x_j-x_k) \\ \cos(x_j+x_k) \\ \sin(x_j+x_k) \end{bmatrix}$$

so that it is straightforward to calculate $\boldsymbol{\phi}_{jk}$ given $\mu_j, \mu_k, \kappa_j, \kappa_k$, and the four λ parameters.

The λ parameters may be recovered as follows, where for brevity we use $\mu^-_{jk} = \mu_j - \mu_k$ and $\mu^+_{jk} = \mu_j + \mu_k$:

$$\begin{aligned}
\lambda^{cc} &= \boldsymbol{\phi}_{jk,1}\cos(\mu^-_{jk}) + \boldsymbol{\phi}_{jk,2}\sin(\mu^-_{jk}) + \boldsymbol{\phi}_{jk,3}\cos(\mu^+_{jk}) + \boldsymbol{\phi}_{jk,4}\sin(\mu^+_{jk}) \\
\lambda^{cs} &= -\boldsymbol{\phi}_{jk,2}\cos(\mu^-_{jk}) + \boldsymbol{\phi}_{jk,1}\sin(\mu^-_{jk}) + \boldsymbol{\phi}_{jk,4}\cos(\mu^+_{jk}) - \boldsymbol{\phi}_{jk,3}\sin(\mu^+_{jk}) \\
\lambda^{sc} &= \boldsymbol{\phi}_{jk,2}\cos(\mu^-_{jk}) - \boldsymbol{\phi}_{jk,1}\sin(\mu^-_{jk}) + \boldsymbol{\phi}_{jk,4}\cos(\mu^+_{jk}) - \boldsymbol{\phi}_{jk,3}\sin(\mu^+_{jk}) \\
\lambda^{ss} &= \boldsymbol{\phi}_{jk,1}\cos(\mu^-_{jk}) + \boldsymbol{\phi}_{jk,2}\sin(\mu^-_{jk}) - \boldsymbol{\phi}_{jk,3}\cos(\mu^+_{jk}) - \boldsymbol{\phi}_{jk,4}\sin(\mu^+_{jk}).
\end{aligned}$$

This shows we have a diffeomorphism between the parameterizations.

Theorem 4.2.2 of Kass and Vos (2007) states that a subfamily of a regular exponential family is itself a (lower-dimensional) regular exponential family if and only if the subspace of the natural parameter space corresponding to the subfamily is an affine subspace of the natural parameter space. In the mean-centered parameterization, the *sine model* has parameter constraints $\lambda^{cc} = \lambda^{cs} = \lambda^{sc} = 0$. Given the equations above, the *sine model* corresponds to a restriction of the natural parameter space of the torus graph density of Equation S1.2:

$$\boldsymbol{\phi}_{jk} = \frac{1}{2}\lambda^{ss}[\cos(\mu^-_{jk}),\ \sin(\mu^-_{jk}),\ -\cos(\mu^+_{jk}),\ -\sin(\mu^+_{jk})]^T. \tag{S2.1}$$

This implies that the pairwise interactions must follow a specific structured form in the *sine model*, where the magnitude of interactions is governed by λ^{ss} and the following relationship between the parameters is observed (regardless of μ_j and μ_k):

$$\boldsymbol{\phi}^2_{jk,1} + \boldsymbol{\phi}^2_{jk,2} = \boldsymbol{\phi}^2_{jk,3} + \boldsymbol{\phi}^2_{jk,4}. \tag{S2.2}$$

Because of the nonlinear relationship between the parameters, the subspace corresponding to the *sine model* is not an affine subspace of the natural parameter space. Therefore, the *sine model* is not itself a regular exponential family. On the other hand, the uniform marginal model and the phase difference model both are defined by setting components of the natural parameter to zero, as is the phase difference model with uniform marginals, so each of these families is itself a regular exponential family. This proves Klein et al. 2019, Theorem 3.1.

S3. Proof of Corollary 2.1.1 (Torus graph properties).

1. Because exponential family models are maximum entropy models subject to constraints on the expected values of the sufficient statistics (Wainwright et al., 2008), the torus graph is the maximum entropy model subject to constraints on the first circular moments and complex covariances between angles (following the derivation in Section S1 that relates the sufficient statistics to circular first moments and to complex covariances).
2. The torus graph density is positive and continuous on $[0, 2\pi)^d$ and factorizes into pairwise interaction terms as shown in Equation S1.2. By the Hammersley-Clifford theorem (Lauritzen, 1996), the random variables X_j and X_k are conditionally independent given all other variables if and only if $\boldsymbol{\phi}_{jk} = \mathbf{0}$.

S4. Derivations of phase differences in torus graph models. First, we state the Harmonic Addition Theorem which will be very useful throughout this set of derivations (see Weisstein (2017) for proof).

THEOREM S4.1 (Harmonic Addition Theorem). *The weighted sum of cosine functions with the same period and arbitrary phase shifts is the cosine function*

$$\sum_{i=1}^{n} a_i \cos(x - \delta_i) = A \cos(x - \Delta)$$

where

$$
\begin{aligned}
b_x &= \sum_{i=1}^{n} a_i \cos(\delta_i) \\
b_y &= \sum_{i=1}^{n} a_i \sin(\delta_i) \\
A &= \sqrt{b_x^2 + b_y^2} \\
\Delta &= \arctan\left(\frac{b_y}{b_x}\right).
\end{aligned}
$$

Throughout these derivations, when we use the arctangent function $\arctan(\cdot)$, it is understood that the angular value is chosen to fall in the same interval as the random variables (in this case, $[0, 2\pi)$, though other intervals such as $[-\pi, \pi)$ could be chosen). We will use the notation $\boldsymbol{\phi}_{jk} = [\alpha_{jk}, \beta_{jk}, \gamma_{jk}, \delta_{jk}]^T$ to refer to elements of the pairwise coupling parameter vector.

For the bivariate torus graph model, we derive the distribution of phase differences to compare with bivariate phase coupling measures, which depend on the distribution of phase differences. Let $\theta = X_1 - X_2$ be a random variable with support $(-2\pi, 2\pi)$ (as $X_1 \in [0, 2\pi)$ and $X_2 \in [0, 2\pi)$) and let $p_{X_1,X_2}(x_1, x_2)$ denote the bivariate phase difference model density. Applying the change of variables $\theta = X_1 - X_2$ and trigonometric identity $\sin(\theta) = \cos\left(\theta - \frac{\pi}{2}\right)$ yields

$$
\begin{aligned}
p_{\theta,X_2}(\theta, x_2) &= p_{X_1,X_2}(\theta + x_2, x_2) \\
&\propto \exp\left\{\kappa_1 \cos(x_2 - (\mu_1 - \theta)) + \kappa_2 \cos(x_2 - \mu_2)\right\} \times \\
&\quad \exp\left\{\alpha_{12} \cos(\theta) + \beta_{12} \cos\left(\theta - \tfrac{\pi}{2}\right)\right\}.
\end{aligned}
$$

Applying Theorem S4.1 to each factor,

$$
p_{X_1,X_2}(\theta + x_2, x_2) \propto \exp\left\{A_1 \cos(\theta - \Delta_1)\right\} \exp\left\{A_2 \cos(x_2 - \Delta_2)\right\}
$$

where

$$
\begin{aligned}
A_1 &= \sqrt{\alpha_{12}^2 + \beta_{12}^2} \\
\Delta_1 &= \arctan\left(\tfrac{\beta_{12}}{\alpha_{12}}\right) \\
A_2(\theta) &= \sqrt{\kappa_1^2 + \kappa_2^2 + 2\kappa_1\kappa_2 \cos(\theta - (\mu_1 - \mu_2))} \\
\Delta_2(\theta) &= \arctan\left(\tfrac{\kappa_1 \sin(\mu_1 - \theta) + \kappa_2 \sin(\mu_2)}{\kappa_1 \cos(\mu_1 - \theta) + \kappa_2 \cos(\mu_2)}\right),
\end{aligned}
$$

and we use the notation $A_2(\theta), \Delta_2(\theta)$ to indicate that these are functions of θ.

To obtain the marginal density of θ, we need to integrate over x_2 and also wrap the resulting distribution back to the support $[0, 2\pi)$ (though we could choose a different support of length 2π, such as $[-\pi, \pi)$, if desired). Notice that $p_{X_1,X_2}(\theta + x_2, x_2)$ has constraints $X_1, X_2 \in [0, 2\pi)$, implying that $0 \leq \theta + X_2 < 2\pi$ so $-\theta \leq X_2 < 2\pi - \theta$. This means that when $\theta < 0$, $X_2 \in [-\theta, 2\pi)$ and when $\theta > 0$, $X_2 \in [0, 2\pi - \theta)$, so the marginal distribution of θ is defined piecewise:

$$p_\theta(\theta) \propto \begin{cases} \mathbb{1}_{(\theta \in [-2\pi, 0])} g(\theta) \int_{-\theta}^{2\pi} \exp\{A_2(\theta)\cos(x_2 - \Delta_2(\theta))\}\, dx_2 \\ \mathbb{1}_{(\theta \in [0, 2\pi])} g(\theta) \int_{0}^{2\pi-\theta} \exp\{A_2(\theta)\cos(x_2 - \Delta_2(\theta))\}\, dx_2 \end{cases}$$

where $g(\theta) = \exp\{A_1 \cos(\theta - \Delta_1)\}$.

Define $W = \theta \pmod{2\pi}$ to be the wrapped version of θ so that $W \in [0, 2\pi)$ has the wrapped distribution

$$\begin{aligned} p_W(w) &= p_\theta(w) + p_\theta(w - 2\pi) \\ &\propto g(w) \int_0^{2\pi - w} \exp\{A_2(w)\cos(x_2 - \Delta_2(w))\}\, dx_2 \\ &\quad + g(w - 2\pi) \int_{2\pi - w}^{2\pi} \exp\{A_2(w - 2\pi)\cos(x_2 - \Delta_2(w - 2\pi))\}\, dx_2 \end{aligned}$$

Using the fact that g, A_2, and Δ_2 are 2π-periodic functions and the definition of I_0 (the modified Bessel function of the first kind), we obtain

$$p_W(w) \propto g(w) \int_0^{2\pi} \exp\{A_2(w)\cos(x_2 - \Delta_2(w))\}\, dx_2 = g(w) I_0(A_2(w)).$$

Thus, $g(w)$ is the direct coupling term and $f(w) = I_0(A_2(w))$ is the marginal concentration term.

The derivation for phase differences from the trivariate torus graph model uses similar techniques. Consider a trivariate torus graph with $\kappa_1 = \kappa_2 = \kappa_3 = 0$ for simplicity; applying trigonometric identities yields

$$\begin{aligned} p(x_1, x_2, x_3) &\propto \exp\left\{ \sum_{(i,j)\in E} \begin{bmatrix} \alpha_{ij} \\ \beta_{ij} \\ \gamma_{ij} \\ \delta_{ij} \end{bmatrix}^T \begin{bmatrix} \cos(x_i - x_j) \\ \sin(x_i - x_j) \\ \cos(x_i + x_j) \\ \sin(x_i + x_j) \end{bmatrix} \right\} \\ &= \exp\left\{ \sum_{(i,j)\in E} \begin{bmatrix} \alpha_{ij} \\ \beta_{ij} \\ \gamma_{ij} \\ \delta_{ij} \end{bmatrix}^T \begin{bmatrix} \cos(x_i - x_j - 0) \\ \cos(x_i - x_j - \pi/2) \\ \cos(x_i + x_j - 0) \\ \cos(x_i + x_j - \pi/2) \end{bmatrix} \right\} \end{aligned}$$

where $E = \{(1,2),\ (1,3),\ (2,3)\}$. Apply the variable transformation $\theta_{12} = x_1 - x_2$ and expand the expression above:

$$\begin{aligned} p(\theta_{12}, x_2, x_3) \propto \exp &\left\{ \begin{bmatrix} \alpha_{12} \\ \beta_{12} \\ \gamma_{12} \\ \delta_{12} \end{bmatrix}^T \begin{bmatrix} \cos(\theta_{12} - 0) \\ \cos(\theta_{12} - \pi/2) \\ \cos(\theta_{12} + 2x_2 - 0) \\ \cos(\theta_{12} + 2x_2 - \pi/2) \end{bmatrix} \right\} \times \\ &\exp \left\{ \begin{bmatrix} \alpha_{13} \\ \beta_{13} \\ \gamma_{13} \\ \delta_{13} \end{bmatrix}^T \begin{bmatrix} \cos(x_1 - x_3 - 0) \\ \cos(x_1 - x_3 - \pi/2) \\ \cos(x_1 + x_3 - 0) \\ \cos(x_1 + x_3 - \pi/2) \end{bmatrix} \right\} \times \\ &\exp \left\{ \begin{bmatrix} \alpha_{23} \\ \beta_{23} \\ \gamma_{23} \\ \delta_{23} \end{bmatrix}^T \begin{bmatrix} \cos(x_2 - x_3 - 0) \\ \cos(x_2 - x_3 - \pi/2) \\ \cos(x_2 + x_3 - 0) \\ \cos(x_2 + x_3 - \pi/2) \end{bmatrix} \right\} \end{aligned}$$

To get the marginal distribution of θ_{12}, we need to integrate out other variables, which is not tractable analytically for the full torus graph model, so we consider the phase difference model which corresponds to setting $\gamma = \delta = 0$ for all pairs. This has the effect of making the density depend only on phase differences.

$$\begin{aligned} p(\theta_{12}, x_2, x_3) \propto \exp &\left\{ \begin{bmatrix} \alpha_{12} \\ \beta_{12} \end{bmatrix}^T \begin{bmatrix} \cos(\theta_{12} - 0) \\ \cos(\theta_{12} - \pi/2) \end{bmatrix} \right\} \times \\ &\exp \left\{ \begin{bmatrix} \alpha_{13} \\ \beta_{13} \end{bmatrix}^T \begin{bmatrix} \cos(x_1 - x_3 - 0) \\ \cos(x_1 - x_3 - \pi/2) \end{bmatrix} \right\} \times \\ &\exp \left\{ \begin{bmatrix} \alpha_{23} \\ \beta_{23} \end{bmatrix}^T \begin{bmatrix} \cos(x_2 - x_3 - 0) \\ \cos(x_2 - x_3 - \pi/2) \end{bmatrix} \right\}. \end{aligned}$$

Similarly to the bivariate case, we apply Theorem S4.1 to each factor; the first factor, $g(\theta_{12}) = \exp\{A_1 \cos(\theta_{12} - \Delta_1)\}$, has the same form as in the bivariate case and represents the direct coupling between X_1 and X_2. The second factor may also be written as

$$\exp\{A_{13} \cos(x_1 - x_3 - \Delta_{13}\}$$

where $A_{13} = \sqrt{\alpha_{13}^2 + \beta_{13}^2}$ and $\Delta_{13} = \arctan(\beta_{13}/\alpha_{13})$, and the third factor may also be written as

$$\exp\{A_{23} \cos(x_2 - x_3 - \Delta_{23}\},$$

where $A_{23} = \sqrt{\alpha_{23}^2 + \beta_{23}^2}$ and $\Delta_{23} = \arctan(\beta_{23}/\alpha_{23})$. Next we apply Theorem S4.1 again to combine the second and third factors:

$$p(\theta_{12}, x_3) \propto \exp\left\{A_1 \cos(\theta_{12} - \Delta_1)\right\} \times \exp\left\{A_3(\theta_{12}) \cos(x_3 - \Delta_3)\right\},$$

where $A_3(\theta_{12})$ is

$$A_3(\theta_{12}) = \Bigg[\Big(A_{13}\cos(x_1 - \Delta_{13}) + A_{23}\cos(x_2 - \Delta_{23})\Big)^2 + \Big(A_{13}\sin(x_1 - \Delta_{13}) + A_{23}\sin(x_2 - \Delta_{23})\Big)^2\Bigg]^{1/2}.$$

Because we will integrate out x_3, the form of Δ_3 is not important (it will not affect the integral on this circular domain).

Expanding the squares, simplifying, and using a trigonometric sum identity yields

$$A_3(\theta_{12}) = \sqrt{\alpha_{13}^2 + \beta_{13}^2 + \alpha_{23}^2 + \beta_{23}^2 + 2t\cos(\theta_{12} - u)} \tag{S4.1}$$

where

$$t = \sqrt{(\alpha_{13}^2 + \beta_{13}^2)(\alpha_{23}^2 + \beta_{23}^2)}$$
$$u = \Delta_{13} - \Delta_{23} = \arctan\left(\frac{\beta_{13}}{\alpha_{13}}\right) - \arctan\left(\frac{\beta_{23}}{\alpha_{23}}\right).$$

Similar to the bivariate case, we integrate over x_3 and wrap the resulting distribution to obtain the marginal distribution of the wrapped phase difference $W = \theta_{12} \pmod{2\pi}$:

$$\begin{aligned} p(w) &\propto \exp\left\{A_1 \cos\left(w - \Delta_1\right)\right\} \int_0^{2\pi} \exp\left\{A_3(w)\cos\left(x_3 - \Delta_3\right)\right\} dx_3 \\ &\propto \exp\left\{A_1 \cos\left(w - \Delta_1\right)\right\} I_0(A_3(w)) \end{aligned}$$

where $A_1 = \sqrt{\alpha_{12}^2 + \beta_{12}^2}$, $\Delta_1 = \arctan(\beta_{12}/\alpha_{12})$, and $A_3(w)$ is given in Equation S4.1. Thus, $g(w)$ is the direct coupling term, and $h(w)$ is the indirect coupling term.

S5. Proof of Theorem 5.1 (Score matching estimators for torus graphs). Let $p_{\mathbf{X}}(\mathbf{x})$ be the unknown d-dimensional circular data density and $p(\mathbf{x}; \boldsymbol{\phi}) = \frac{1}{Z(\boldsymbol{\phi})} q(\mathbf{x}; \boldsymbol{\phi})$ be a d-dimensional model density with parameter

vector $\boldsymbol{\phi} \in \mathbb{R}^m$. Define the log model density gradient $\boldsymbol{\psi} : [0, 2\pi)^d \to \mathbb{R}^d$ as $\boldsymbol{\psi}(\mathbf{x}; \boldsymbol{\phi}) = \nabla_{\mathbf{x}} \log q(\mathbf{x}; \boldsymbol{\phi})$; similarly, let $\boldsymbol{\psi}_{\mathbf{X}}(\mathbf{x}) = \nabla_{\mathbf{x}} \log p_{\mathbf{X}}(\mathbf{x})$.

To prove Klein et al. 2019, Theorem 5.1, make the following regularity assumptions:

A. For all $i \in \{1, ..., d\}$, $\boldsymbol{\psi}(\mathbf{x}; \boldsymbol{\phi})$ is differentiable with respect to $\mathbf{x}_i$ on $[0, 2\pi)$.
B. For all $\boldsymbol{\phi}$, $E_{\mathbf{x}}\left[||\boldsymbol{\psi}(\mathbf{x}; \boldsymbol{\phi})||^2\right]$ and $E_{\mathbf{x}}\left[||\boldsymbol{\psi}_{\mathbf{X}}(\mathbf{x})||^2\right]$ are finite.

These assumptions clearly hold for torus graphs as the log density is comprised of finite linear combinations of sine and cosine functions of $\mathbf{x}$, each of which is infinitely differentiable with derivatives bounded within $[-1, 1]$. Note that we need one less assumption than the original formulation of score matching in Hyvärinen (2005) due to the circular nature of the density.

Proof of Theorem 5.1. First, we show that the score matching objective function only depends on the unknown data density through an expectation.

Expanding the squared difference gives

$$\begin{aligned} J(\boldsymbol{\phi}) = &\int_0^{2\pi} p_{\mathbf{X}}(\mathbf{x}) \left[\tfrac{1}{2}||\nabla_{\mathbf{x}} \log p_{\mathbf{X}}(\mathbf{x})||_2^2\right] \, d\mathbf{x} \\ &+ \int_0^{2\pi} p_{\mathbf{X}}(\mathbf{x}) \left[\tfrac{1}{2}||\nabla_{\mathbf{x}} \log q(\mathbf{x}; \boldsymbol{\phi})||_2^2\right] \, d\mathbf{x} \\ &- \int_0^{2\pi} p_{\mathbf{X}}(\mathbf{x}) [\nabla_{\mathbf{x}} \log q(\mathbf{x}; \boldsymbol{\phi})]^T [\nabla_{\mathbf{x}} \log p_{\mathbf{X}}(\mathbf{x})] \, d\mathbf{x}. \end{aligned}$$

The first term does not depend on $\boldsymbol{\phi}$ and the second term is already in terms of an expectation over the data density, so we focus now on the third term (call it A):

$$\begin{aligned} A &= -\int_0^{2\pi} p_{\mathbf{X}}(\mathbf{x}) \left[\sum_{i=1}^d \boldsymbol{\psi}_i(\mathbf{x}; \boldsymbol{\phi}) \boldsymbol{\psi}_{\mathbf{X},i}(\mathbf{x})\right] \, d\mathbf{x} \\ &= -\sum_{i=1}^d \int_0^{2\pi} \left[\int_0^{2\pi} p_{\mathbf{X}}(\mathbf{x}) \boldsymbol{\psi}_i(\mathbf{x}; \boldsymbol{\phi}) \boldsymbol{\psi}_{\mathbf{X},i}(\mathbf{x}) \, d\mathbf{x}_i\right] \, d\mathbf{x}_{-i} \\ &= -\sum_{i=1}^d \int_0^{2\pi} \left[\int_0^{2\pi} p_{\mathbf{X}}(\mathbf{x}) \frac{\partial}{\partial \mathbf{x}_i} \log p_{\mathbf{X}}(\mathbf{x}) \boldsymbol{\psi}_i(\mathbf{x}; \boldsymbol{\phi}) \, d\mathbf{x}_i\right] \, d\mathbf{x}_{-i} \\ &= -\sum_{i=1}^d \int_0^{2\pi} \left[\int_0^{2\pi} \frac{\partial}{\partial \mathbf{x}_i} p_{\mathbf{X}}(\mathbf{x}) \boldsymbol{\psi}_i(\mathbf{x}; \boldsymbol{\phi}) \, d\mathbf{x}_i\right] \, d\mathbf{x}_{-i}. \end{aligned}$$

Applying integration by parts, the inner integral becomes

$$p_{\mathbf{X}}(\mathbf{x})\boldsymbol{\psi}_i(\mathbf{x};\boldsymbol{\phi})\Big|_{\mathbf{x}_i=0}^{\mathbf{x}_i=2\pi} - \int_0^{2\pi} p_{\mathbf{X}}(\mathbf{x})\frac{\partial}{\partial \mathbf{x}_i}\left(\boldsymbol{\psi}_i(\mathbf{x};\boldsymbol{\phi})\right)\, d\mathbf{x}_i.$$

Notice that because the variables are circular on $[0, 2\pi)$,

$$p_{\mathbf{X}}(\mathbf{x})\big|_{\mathbf{x}_i=0} = p_{\mathbf{X}}(\mathbf{x})\big|_{\mathbf{x}_i=2\pi}$$
$$\boldsymbol{\psi}_i(\mathbf{x};\boldsymbol{\phi})\big|_{\mathbf{x}_i=0} = \boldsymbol{\psi}_i(\mathbf{x};\boldsymbol{\phi})\big|_{\mathbf{x}_i=2\pi}$$

Therefore, $p_{\mathbf{X}}(\mathbf{x})\boldsymbol{\psi}_i(\mathbf{x};\boldsymbol{\phi})\big|_{\mathbf{x}_i=0}^{\mathbf{x}_i=2\pi} = 0$, so A becomes

$$\begin{aligned} A &= -\sum_{i=1}^{d}\int_0^{2\pi}\left[-\int_0^{2\pi} p_{\mathbf{X}}(\mathbf{x})\frac{\partial}{\partial \mathbf{x}_i}\left(\boldsymbol{\psi}_i(\mathbf{x};\boldsymbol{\phi})\right)\, d\mathbf{x}_i\right]\, d\mathbf{x}_{-i} \\ &= \int_0^{2\pi} p_{\mathbf{X}}(\mathbf{x})\left[\sum_{i=1}^{d}\frac{\partial}{\partial \mathbf{x}_i}\left(\boldsymbol{\psi}_i(\mathbf{x};\boldsymbol{\phi})\right)\right]\, d\mathbf{x} \end{aligned}$$

Therefore, the score matching objective is

$$\begin{aligned} J(\boldsymbol{\phi}) &= C + \int_0^{2\pi} p_{\mathbf{X}}(\mathbf{x})\left[\tfrac{1}{2}||\boldsymbol{\psi}(\mathbf{x};\boldsymbol{\phi})||^2\right]\, d\mathbf{x} \\ &\quad + \int_0^{2\pi} p_{\mathbf{X}}(\mathbf{x})\left[\sum_{i=1}^{d}\frac{\partial}{\partial \mathbf{x}_i}\left(\boldsymbol{\psi}_i(\mathbf{x};\boldsymbol{\phi})\right)\right]\, d\mathbf{x} \\ &= C + E_{\mathbf{x}}\left\{\tfrac{1}{2}||\boldsymbol{\psi}(\mathbf{x};\boldsymbol{\phi})||^2 + \sum_{i=1}^{d}\frac{\partial}{\partial \mathbf{x}_i}\left(\boldsymbol{\psi}_i(\mathbf{x};\boldsymbol{\phi})\right)\right\} \end{aligned} \tag{S5.1}$$

where C does not depend on $\boldsymbol{\phi}$ and may be ignored without affecting the minima of the objective function. This coincides with the form of score matching given in Hyvärinen (2005) except with the integral over the circular domain $[0, 2\pi)$.

Next, we show the explicit form of the score matching estimator for torus graphs. As shown in Forbes and Lauritzen (2015); Yu, Drton and Shojaie (2018), for exponential families, this score matching estimator is quadratic in the parameters. Specifically, the torus graph density in Klein et al. 2019, Theorem 2.1 has a log density of the form

$$\log q(\mathbf{x};\boldsymbol{\phi}) = \boldsymbol{\phi}^T\mathbf{S}(\mathbf{x})$$

where $\boldsymbol{\phi}$ are vectors of length $m = 2d^2$ (the number of sufficient statistics).

Therefore,

$$\boldsymbol{\psi}(\mathbf{x};\boldsymbol{\phi}) = \boldsymbol{\phi}^T\mathbf{D}(\mathbf{x})$$

where the Jacobian $\mathbf{D}(\mathbf{x})$ is $m \times d$ with i,jth element $\frac{\partial}{\partial \mathbf{x}_j}\mathbf{S}_i$. Thus the first term inside the expectation in the score matching objective of Equation S5.1 may be written

$$\tfrac{1}{2}||\boldsymbol{\psi}(\mathbf{x};\boldsymbol{\phi})||^2 = \tfrac{1}{2}\boldsymbol{\phi}^T\mathbf{D}(\mathbf{x})\mathbf{D}(\mathbf{x})^T\boldsymbol{\phi} \equiv \tfrac{1}{2}\boldsymbol{\phi}^T\boldsymbol{\Gamma}(\mathbf{x})\boldsymbol{\phi}.$$

The elements of $\mathbf{D}(\mathbf{x})$ correspond to partial derivatives of the sufficient statistics with respect to the data. The derivatives of the univariate sufficient statistics $\mathbf{S}^1$ are given by

$$\frac{\partial}{\partial x_\ell}\cos(x_j) = \begin{cases} -\sin(x_j), & \ell = j \\ 0, & \ell \neq j \end{cases},$$
$$\frac{\partial}{\partial x_\ell}\sin(x_j) = \begin{cases} \cos(x_j), & \ell = j \\ 0, & \ell \neq j \end{cases}.$$

Similarly, the derivatives of the pairwise sufficient statistics $\mathbf{S}^2$ may be calculated as

$$\frac{\partial}{\partial x_\ell}\cos(x_j - x_k) = \begin{cases} -\sin(x_j - x_k), & \ell = j \\ \sin(x_j - x_k), & \ell = k \\ 0, & \ell \notin \{j,k\} \end{cases},$$
$$\frac{\partial}{\partial x_\ell}\sin(x_j - x_k) = \begin{cases} \cos(x_j - x_k), & \ell = j \\ -\cos(x_j - x_k), & \ell = k \\ 0, & \ell \notin \{j,k\} \end{cases},$$
$$\frac{\partial}{\partial x_\ell}\cos(x_j + x_k) = \begin{cases} -\sin(x_j + x_k), & \ell \in \{j,k\} \\ 0, & \ell \notin \{j,k\} \end{cases},$$
$$\frac{\partial}{\partial x_\ell}\sin(x_j + x_k) = \begin{cases} \cos(x_j + x_k), & \ell \in \{j,k\} \\ 0, & \ell \notin \{j,k\} \end{cases}.$$

Now we show that the second term inside the expectation in Equation S5.1 may be written simply in terms of the sufficient statitsics. Notice that the ith element of the gradient may be written in terms of columns of the Jacobian:

$$\boldsymbol{\psi}_i(\mathbf{x};\boldsymbol{\phi}) = \boldsymbol{\phi}^T[\mathbf{D}(\mathbf{x})]_{\cdot,i}$$

so that

$$\frac{\partial}{\partial \mathbf{x}_i}\left(\boldsymbol{\psi}_i(\mathbf{x};\boldsymbol{\phi})\right) = \boldsymbol{\phi}^T\frac{\partial}{\partial \mathbf{x}_i}[\mathbf{D}(\mathbf{x})]_{\cdot,i}.$$

Therefore, the second term inside the expectation in the score matching objective may be written

$$\sum_{i=1}^{d} \frac{\partial}{\partial \mathbf{x}_i} \left(\boldsymbol{\psi}_i(\mathbf{x};\boldsymbol{\phi})\right) = \boldsymbol{\phi}^T \left[\sum_{i=1}^{d} \frac{\partial}{\partial \mathbf{x}_i} [\mathbf{D}(\mathbf{x})]_{\cdot,i}\right] \equiv \boldsymbol{\phi}^T \mathbf{H}(\mathbf{x})$$

where

$$\mathbf{H}(\mathbf{x}) = [\mathbf{S}^1(\mathbf{x}),\, 2\mathbf{S}^2(\mathbf{x})]^T.$$

This relation holds because all nonzero elements of $\mathbf{D}(\mathbf{x})$ come from derivatives of sines and cosines; due to the relations $\frac{d}{dx}\cos(x) = -\sin(x)$ and $\frac{d}{dx}\sin(x) = \cos(x)$, taking the derivative again essentially converts the elements back to sufficient statistics. □

S6. Proof of Theorem 2.2 (Conditional distributions in torus graphs). We prove that the distribution of one angle conditional on the other angles is von Mises as stated in Klein et al. 2019, Theorem 2.2, enabling the use of Gibbs sampling for drawing samples from the distribution. We will use the notation $\boldsymbol{\phi}_{jk} = [\alpha_{jk}, \beta_{jk}, \gamma_{jk}, \delta_{jk}]^T$ to refer to elements of the pairwise coupling parameter vector.

PROOF. Let $c_{ij}^- = \cos(x_i - x_j)$, $c_{ij}^+ = \cos(x_i + x_j)$, $s_{ij}^- = \sin(x_i - x_j)$, and $s_{ij}^+ = \sin(x_i + x_j)$. Factor the torus graph density into terms containing X_k and not containing X_k:

$$\begin{aligned} p(\mathbf{x};\boldsymbol{\phi}) =& C(\boldsymbol{\phi}) \exp\left\{\sum_{i\neq k} \kappa_i \cos(x_i - \mu_i) + \sum_{i<j,j\neq k} \begin{bmatrix}\alpha_{ij}\\ \beta_{ij}\\ \gamma_{ij}\\ \delta_{ij}\end{bmatrix}^T \begin{bmatrix}c_{ij}^-\\ s_{ij}^-\\ c_{ij}^+\\ s_{ij}^+\end{bmatrix}\right\} \times \\ & \exp\left\{\kappa_k \cos(x_k - \mu_k)\right\} \times \\ & \exp\left\{\sum_{i<k} \begin{bmatrix}\alpha_{ik}\\ \beta_{ik}\\ \gamma_{ik}\\ \delta_{ik}\end{bmatrix}^T \begin{bmatrix}c_{ik}^-\\ s_{ik}^-\\ c_{ik}^+\\ s_{ik}^+\end{bmatrix} + \sum_{i>k} \begin{bmatrix}\alpha_{ki}\\ \beta_{ki}\\ \gamma_{ki}\\ \delta_{ki}\end{bmatrix}^T \begin{bmatrix}c_{ki}^-\\ s_{ik}^-\\ c_{ik}^+\\ s_{ik}^+\end{bmatrix}\right\}. \end{aligned} \tag{S6.1}$$

Let the first factor (including the normalization constant) be denoted by $g(\mathbf{x}_{-d};\boldsymbol{\phi})$ and the second and third factors be denoted by $f(\mathbf{x};\boldsymbol{\phi})$. Then the conditional distribution is

$$p(x_k|\mathbf{x}_{-k};\boldsymbol{\phi}) = \frac{g(\mathbf{x}_{-k};\boldsymbol{\phi}) f(\mathbf{x};\boldsymbol{\phi})}{g(\mathbf{x}_{-k};\boldsymbol{\phi}) \int f(\mathbf{x};\boldsymbol{\phi})\, dx_k} = \frac{f(\mathbf{x};\boldsymbol{\phi})}{\int f(\mathbf{x};\boldsymbol{\phi})\, dx_k}.$$

Applying trigonometric identities to the third factor of Equation S6.1 and simplifying, we have

$$\begin{aligned}
&\exp\left\{\sum_{i<k}\begin{bmatrix}\alpha_{ik}\\ \beta_{ik}\\ \gamma_{ik}\\ \delta_{ik}\end{bmatrix}^T\begin{bmatrix}c_{ik}^-\\ s_{ik}^-\\ c_{ik}^+\\ s_{ik}^+\end{bmatrix}+\sum_{i>k}\begin{bmatrix}\alpha_{ki}\\ \beta_{ki}\\ \gamma_{ki}\\ \delta_{ki}\end{bmatrix}^T\begin{bmatrix}c_{ki}^-\\ s_{ik}^-\\ c_{ik}^+\\ s_{ik}^+\end{bmatrix}\right\}\\
&=\exp\left\{\sum_{i<k}\begin{bmatrix}\alpha_{ik}\\ \beta_{ik}\\ \gamma_{ik}\\ \delta_{ik}\end{bmatrix}^T\begin{bmatrix}\cos(x_k-x_i)\\ \cos(x_k-x_i+\pi/2)\\ \cos(x_k+x_i)\\ \cos(x_k+x_i-\pi/2)\end{bmatrix}\right\}\times\\
&\qquad\exp\left\{\sum_{i>k}\begin{bmatrix}\alpha_{ki}\\ \beta_{ki}\\ \gamma_{ki}\\ \delta_{ki}\end{bmatrix}^T\begin{bmatrix}\cos(x_k-x_i)\\ \cos(x_k-x_i-\pi/2)\\ \cos(x_k+x_i)\\ \cos(x_k+x_i-\pi/2)\end{bmatrix}\right\}\\
&=\exp\left\{\sum_{i\neq k}\begin{bmatrix}\alpha_{ik}\\ \beta_{ik}\\ \gamma_{ik}\\ \delta_{ik}\end{bmatrix}^T\begin{bmatrix}\cos(x_k-x_i)\\ \cos(x_k-x_i+\operatorname{sgn}(i-k)\pi/2)\\ \cos(x_k+x_i)\\ \cos(x_k+x_i-\pi/2)\end{bmatrix}\right\}
\end{aligned}$$

where, with slight abuse of notation, we let, for instance, α_{ik} denote either α_{ik} if $i<k$ or α_{ki} if $i>k$, and $\operatorname{sgn}(\cdot)$ is the signum function. Now we see $f(\mathbf{x};\boldsymbol{\phi})$ is a sum of cosine functions with argument x_k, so applying Theorem S4.1, we have

$$f(\mathbf{x};\boldsymbol{\phi})=\exp(A\cos(x_k-\Delta))$$

where $A=\sqrt{b_x^2+b_y^2}$, $\Delta=\arctan(b_y/b_x)$, defined as

$$\begin{aligned}
b_x&=\sum_m L_m\cos(V_m)\\
b_y&=\sum_m L_m\sin(V_m)\\
L&=\left[\kappa_k,\boldsymbol{\alpha}_{\cdot,k},\boldsymbol{\beta}_{\cdot,k},\boldsymbol{\gamma}_{\cdot,k},\boldsymbol{\delta}_{\cdot,k}\right]=[\kappa_k,\boldsymbol{\phi}_{\cdot k}]\\
V&=[\mu_k,\mathbf{x}_{-k},\mathbf{x}_{-k}+\operatorname{sgn}(i-k)\tfrac{\pi}{2},-\mathbf{x}_{-k},-\mathbf{x}_{-k}+\tfrac{\pi}{2}]\\
&=[\mu_k,\mathbf{x}_{-k},\mathbf{x}_{-k}+\mathbf{h}\tfrac{\pi}{2},-\mathbf{x}_{-k},-\mathbf{x}_{-k}+\tfrac{\pi}{2}],
\end{aligned}$$

with, for example, $\boldsymbol{\alpha}_{\cdot,k}$ denoting all α parameters involving index k and $\mathbf{h}_j=-1$ if $j<k$ and $\mathbf{h}_j=1$ otherwise. Then since $\int f(\mathbf{x};\boldsymbol{\phi})\,dx_k=2\pi I_0(A)$

we find that the conditional density is von Mises with concentration A and mean Δ. □

S7. Measures of positive and negative circular dependence. The dependence between two circular variables can be measured using a correlation coefficient, ρ_c, analogous to the Pearson correlation coefficient for linear analysis Jammalamadaka and Sarma (1988),

$$\rho_c = \frac{E\{\sin(X_i - \mu_i)\sin(X_j - \mu_j)\}}{\sqrt{\mathrm{Var}(\sin(X_i - \mu_i))\mathrm{Var}(\sin(X_j - \mu_j))}}$$

where μ represents a mean circular direction. Using variance properties and trigonometric identities we have

$$\rho_c = \frac{E\{\cos(X_i - X_j - (\mu_i - \mu_j)) - \cos(X_i + X_j - (\mu_i + \mu_j))\}}{2\sqrt{E\{\sin^2(X_i - \mu_i)\}E\{\sin^2(X_j - \mu_j)\}}}. \tag{S7.1}$$

The first component of the numerator measures the positive correlations from the concentration of $X_i - X_j - (\mu_i - \mu_j)$ and the second component measures the negative (or *reflectional*) correlations from the concentration of $X_i - (-X_j) - (\mu_i - (-\mu_j))$. Analogous to the real-valued data, it is important to note that both positive and negative circular correlations are possible and both are needed to fully define circular dependence between two variables.

The numerator of Equation S7.1 can be rewritten as

$$E\left\{\begin{bmatrix}\cos(X_i - X_j)\\ \sin(X_i - X_j)\end{bmatrix}^T \begin{bmatrix}\alpha\\ \beta\end{bmatrix} - \begin{bmatrix}\cos(X_i + X_j)\\ \sin(X_i + X_j)\end{bmatrix}^T \begin{bmatrix}\gamma\\ \delta\end{bmatrix}\right\}$$

where $\alpha = \cos(\mu_i - \mu_j)$, $\beta = \sin(\mu_i - \mu_j)$, $\gamma = \cos(\mu_i + \mu_j)$, $\delta = \sin(\mu_i + \mu_j)$. This shows that the dependence between angles may be decomposed into a four-term linear combination involving the sines and cosines of the phase differences and phase sums, where the phase difference terms correspond to the positive correlation and the phase sum terms correspond to the negative correlation. This corresponds to the phase sum and phase difference terms that appear in the torus graph density, reinforcing the interpretation of the different pairwise coupling parameters as reflecting positive and negative rotational dependence. To illustrate the distinction between positive and negative rotational dependence, we show bivariate torus graphs with positive, negative, or both kinds of dependence in Figure S1, and show what trial-to-trial rotational and reflectional covariance would look like in Figure

S2. For the case of uniform marginal distributions, the circular correlation coefficient becomes Jammalamadaka and Sengupta (2001):

$$\rho_c = \frac{R_{Xi-Xj} - R_{Xi+Xj}}{2\sqrt{E\{\sin^2(X_i - \mu_i)\}E\{\sin^2(X_j - \mu_j)\}}}$$

where $R_{Xi-Xj} \equiv |E\{\exp(\mathbf{i}(X_i - X_j))\}|$ corresponds to the positive correlation and $R_{Xi+Xj} \equiv |E\{\exp(\mathbf{i}(X_i + X_j))\}|$ corresponds to the negative correlation. The theoretical Phase Locking Value (PLV), for which an estimator is given in Klein et al. 2019, Equation 4.1, is equal to R_{Xi-Xj}. This shows that PLV is similar to a measure of positive circular correlation under the assumption of uniform marginal distributions (when the denominator of the circular correlation coefficient would be equal to 1).

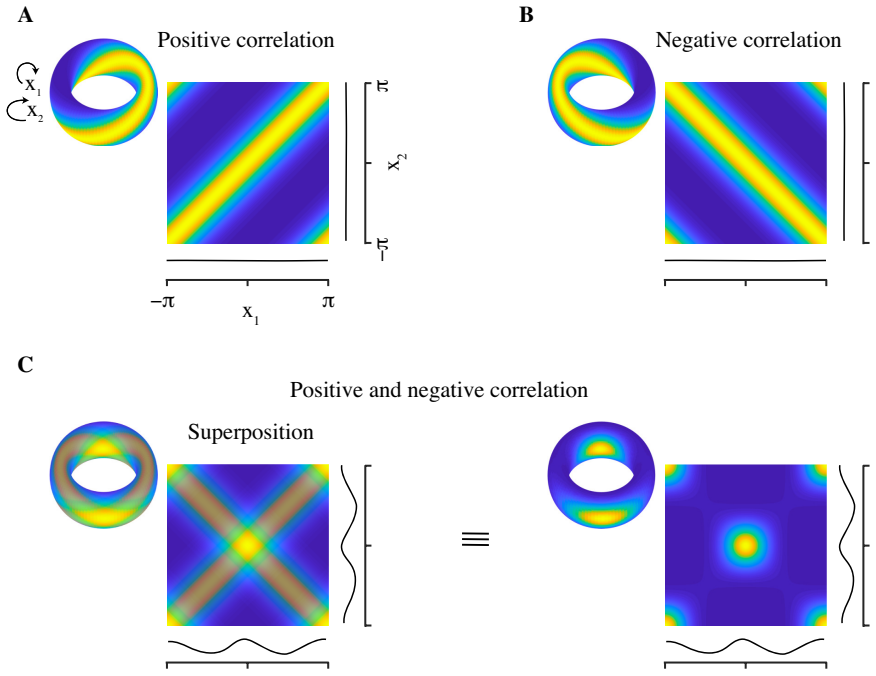

FIG S1. *Bivariate torus graph densities with uniform marginal distributions shown on the torus and flattened on* $[-\pi, \pi]$ *under positive, negative, or both kinds of circular covariance. (A) Coupling parameters chosen to induce only positive correlation (coupling based on phase differences). (B) Coupling parameters chosen to induce only negative correlation (coupling based on phase sums). (C) Equal amounts of positive and negative coupling result in a distribution with two isotropic modes; the superposition figure (left) is shown to provide intuition about the resulting distribution (right).*

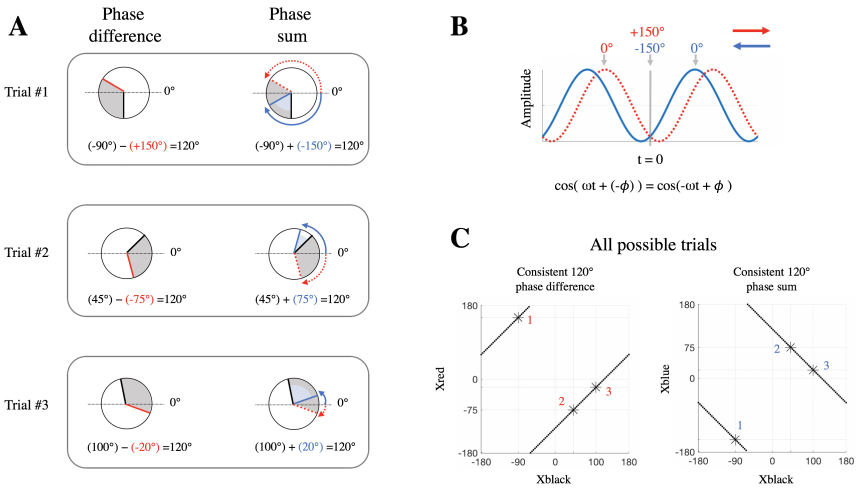

FIG S2. *Intuition about rotational and reflectional dependence. A) Illustration of* 120° *rotational (positive) and* 120° *reflectional (negative) dependence across three hypothetical observations. Rotational dependence implies a consistent phase offset between oscillations across trials (shaded gray angles) while reflectional dependence implies a consistent phase sum (shaded blue angles) that corresponds to a consistent phase offset between oscillations after one of the oscillations has been reflected with respect to* 0°*. B) Example of an oscillation (blue) and its reflection with respect to* 0° *(dashed red). The reflected signal is leading in time, i.e.* $\phi = 150°$*, whereas the blue signal is lagging in time, i.e.* $\phi = -150°$*. This demonstrates that we can think of the reflected signal as moving across time in the opposite direction. In neural data this phenomenon could arise, for example, if there was bidirectional communication. C) The lines indicate dependence between phase angles used in panel A, with rotational dependence on the left and reflectional dependence on the right and trials shown in panel A marked with stars. When phases have uniform marginal distributions across trials, but exhibit phase coupling, positive dependence is observed in the bivariate relationship on the left as a line with fixed orientation at* 45°*; rotation produces a shift along the anti-diagonal. On the other hand, negative dependence is observed in the bivariate relationship on the right as a line with fixed orientation at* 135°*; reflection produces a shift along the diagonal. In the current example, the respective shifts are at* 120°*; see Figure S1 to compare with the case of* 0°*. Weaker dependence blurs the relationship line but does not change the orientation.*

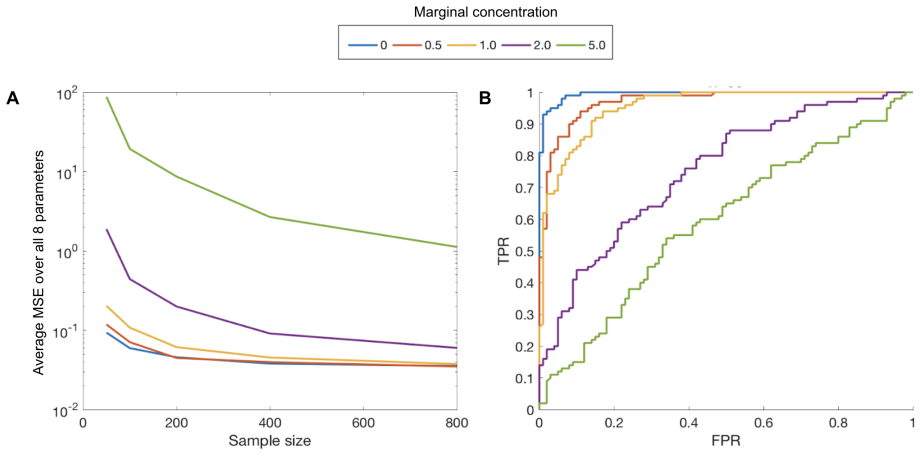

FIG S3. *In data simulated from a bivariate torus graph, the average MSE over all parameters is shown in panel (A) as a function of sample size and marginal concentration. The MSE is higher overall when marginal concentration is high. (B) ROC curves averaged over 200 simulations, half of which had no edge and half of which had an edge between the variables, as a function of marginal concentration for a fixed sample size (N=50), suggesting that structure recovery is also diminished when high marginal concentration is present.*

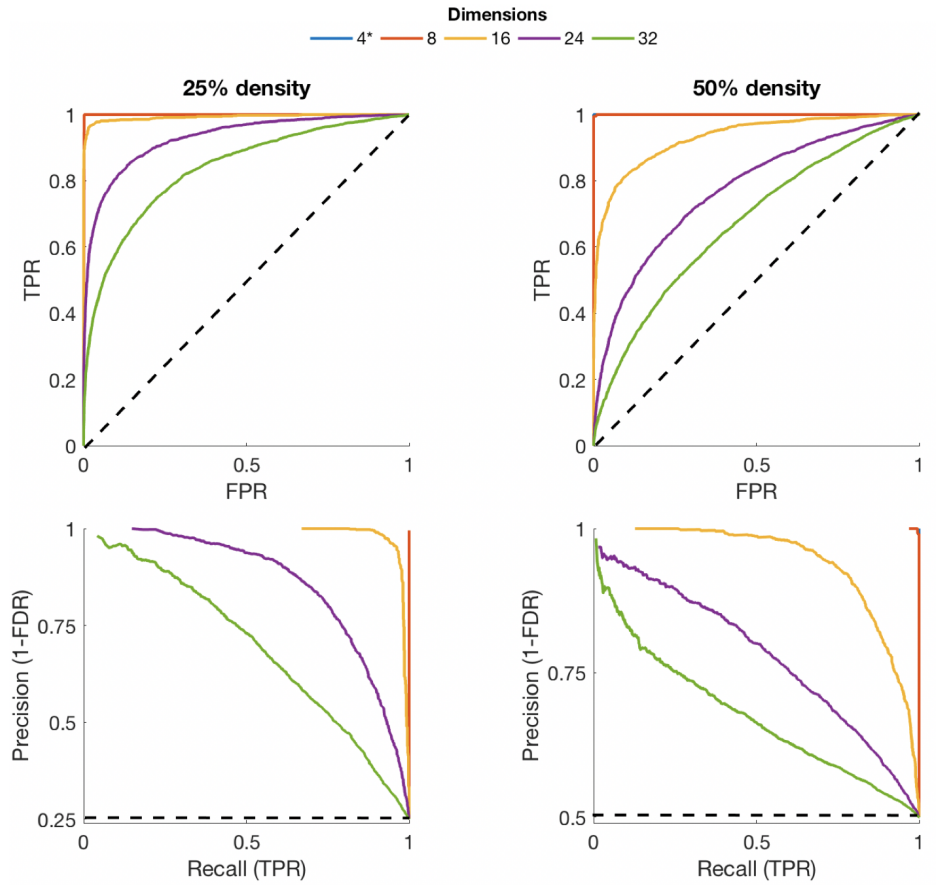

FIG S4. *Further detail on the simulation results shown in Figure 5 of Klein et al. (2019) for a sample size of 840 (matching the real LFP data). Top row: ROC curves colored by dimension for two different underlying edge densities (averaged across 30 simulations). Bottom row: averaged precision curves corresponding to the same densities as the top row.*

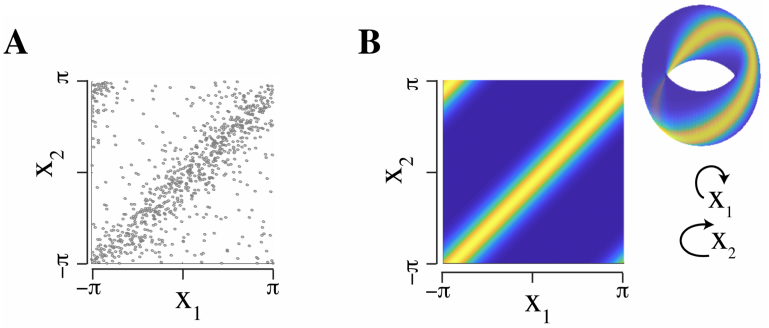

FIG S5. *(A) Scatter plot of data from two channels in dentate gyrus (DG). The angles follow a pattern of positive dependence similar to the simulated data of Figure 1 Klein et al. (2019), which was used to demonstrate the need for circular wrapping when modeling dependent phase angles. (B) Fitted torus graph density on the plane and torus.*

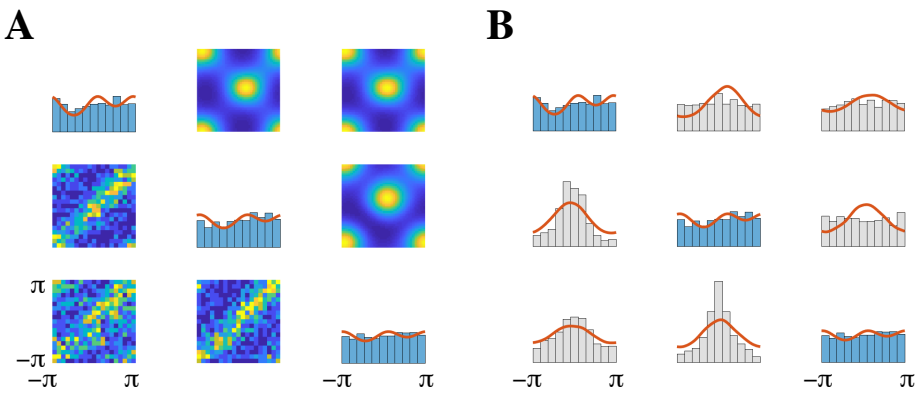

FIG S6. *Similar to Klein et al. (2019), Figure 9, but using the sine model as the theoretical distribution. The sine model fails to accurately fit this data set, which is evident in the bivariate dependence and sufficient statistics. (A) Along the diagonal are the marginal distributions of the three phase angles, where the real data is represented by blue histograms and the theoretical marginal densities from the sine model are overlaid as solid red traces. Two-dimensional distributions (off-diagonal) show bivariate relationships, with theoretical densities above the diagonal and real data represented using two-dimensional histograms below the diagonal. The multimodal behavior of the sine model is apparent in the two-dimensional distributions, which do not appear to match the real data. (B) Plots along the diagonal same as panel A. Below the diagonal are distributions of pairwise phase differences and above the diagonal are distributions of pairwise phase sums, represented by histograms for the real data and by solid red density plots for the theoretical torus graph model. In contrast to the torus graph model, the sine model fails to accurately capture the distributions of the sufficient statistics from these data.*

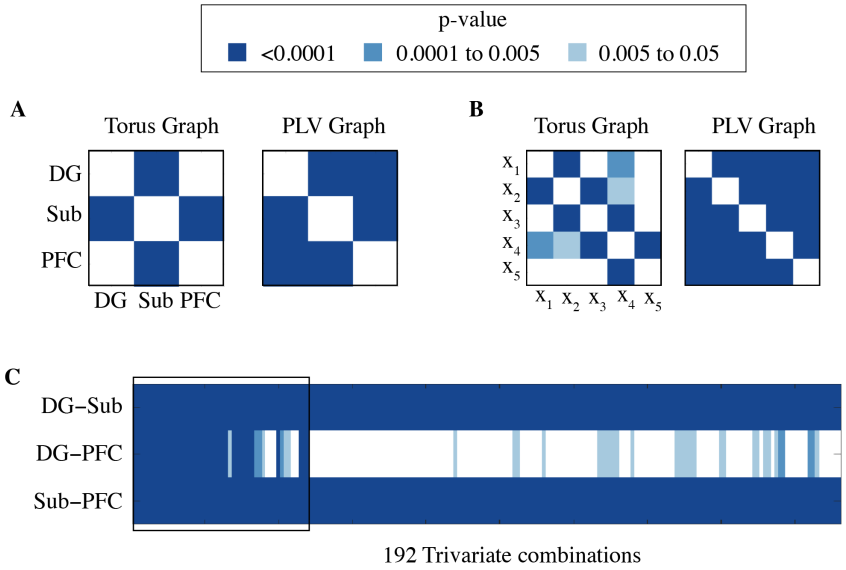

FIG S7. *(A) Adjacency matrices for the three-dimensional LFP analysis with entries colored by p-value. PLV p-values are very small for all connections, while torus graph p-values reflect finer structure, such as PFC-Sub coupling that is apparently more salient than PFC-DG coupling. The adjacency matrix for the trivariate network is a representative combination of channels reflecting the pattern that dominates in all trivariate combinations. (B) Same as (A) but for the five-dimensional LFP analysis, where the torus graph reveals nearest-neighbor structure along the linear probe in CA3 that PLV misses. (C) Edgewise p-values for each edge in each possible trivariate graph (composed of each combination of electrodes from each of the three areas). Note that channels 6 and 7 (boxed region) of DG are on the border between DG and CA3 and may be picking up signals from CA3; omitting these channels gives stronger evidence of an overall lack of connections between DG and PFC (corresponding to the adjacency matrix in (A)).*

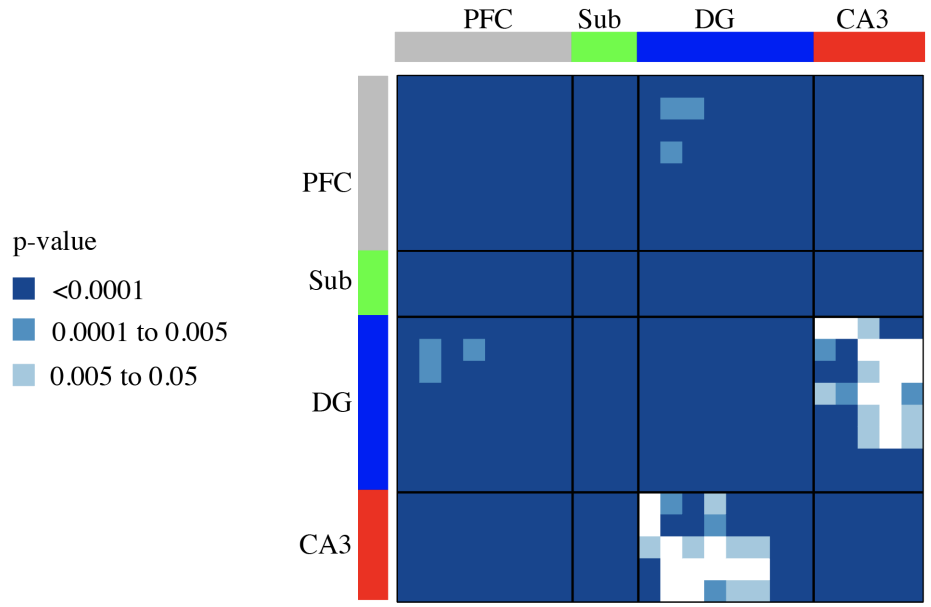

FIG S8. *Adjacency matrix for PLV graph with edgewise p-values determined using Rayleigh's test of uniformity on the circle for each pairwise phase difference. Entries are colored by p-value and, compared to the torus graph adjacency matrix (Klein et al., 2019, Figure 8.C), there is very little noticeable structure in the graph even for very small p-value thresholds (aside from a lack of edges between CA3 and DG).*

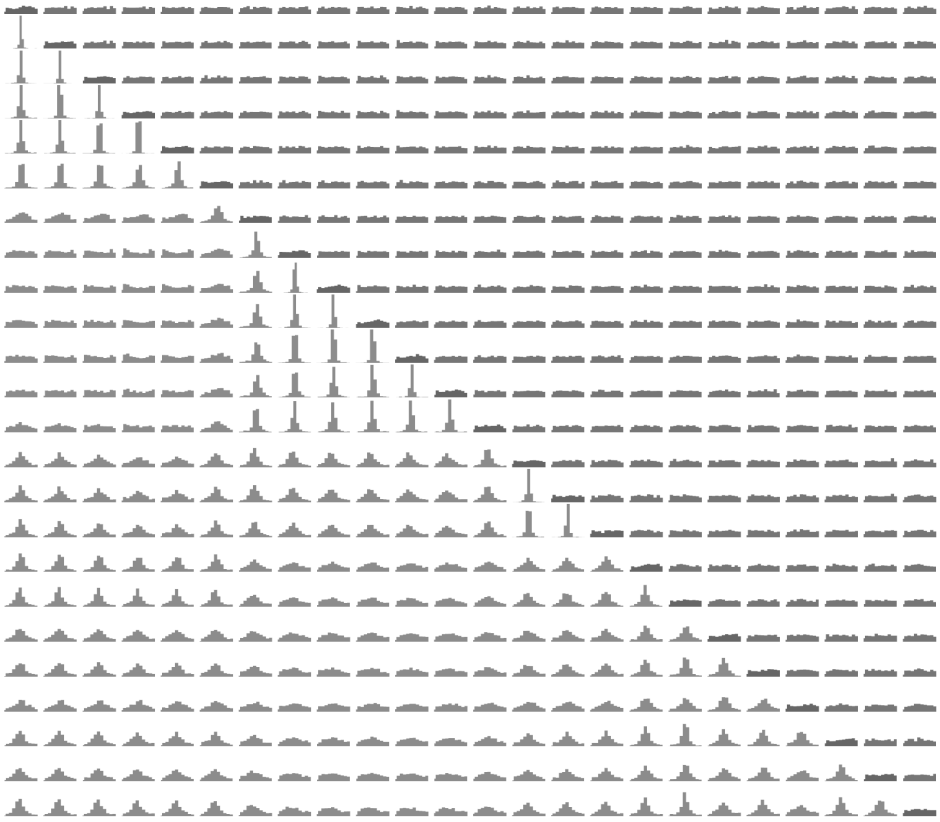

FIG S9. *For the 24-dimensional real LFP data, the diagonal shows univariate histograms which appear to have low concentration in all cases. Below the diagonal are histograms of phase differences between pairs of angles, showing some highly concentrated distributions suggesting rotational dependence; above the diagonal are histograms of phase sums, showing very little concentration, suggesting there is not strong evidence for reflectional dependence in these data.*

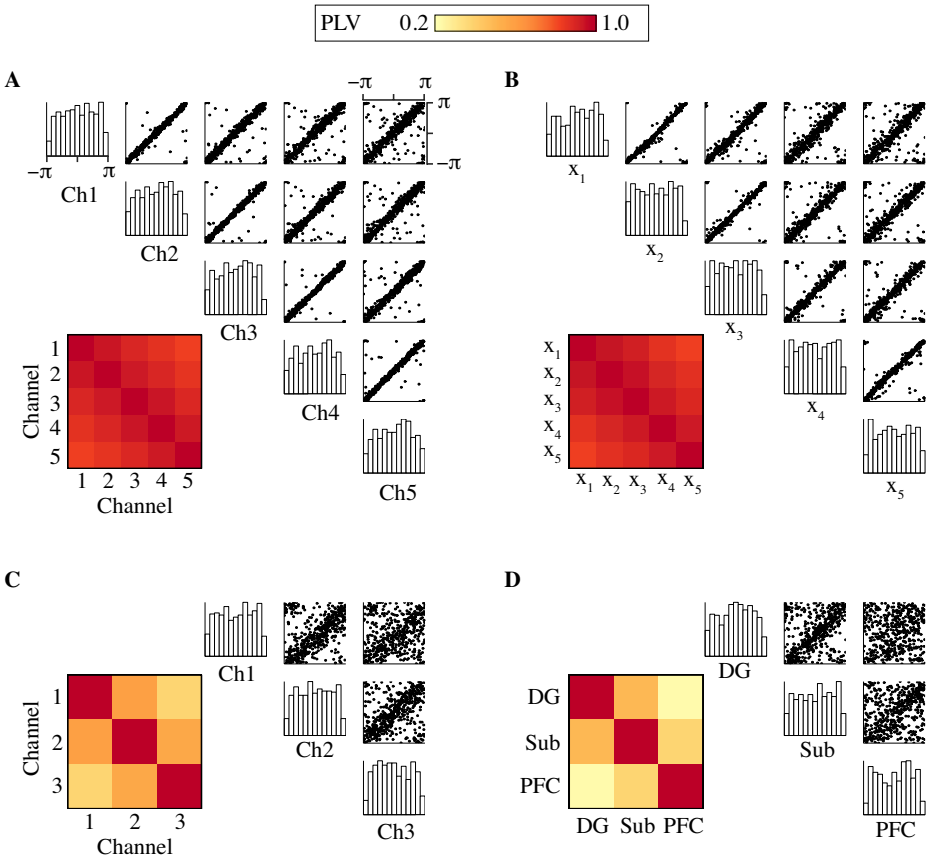

FIG S10. *Examples of simulated and real data to demonstrate the validity of the simulation process. Upper right: histograms and pairwise scatter plots, bottom left: estimated PLV matrices (color scale 0.2 to 1, with red indicating higher PLV values). (A) Simulated 5-channel data with linear probe structure. (B) Real 5-channel data from CA3. (C) Simulated 3-channel data. (D) Real 3-channel data from separate regions (DG, Sub, and PFC).*

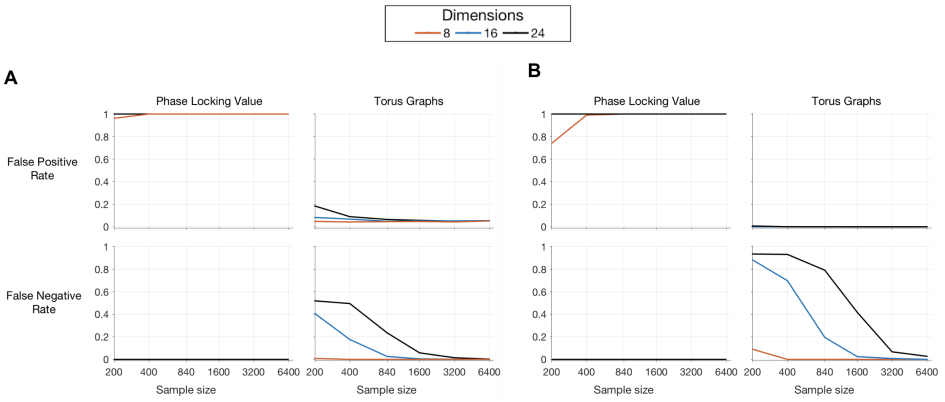

FIG S11. *Investigation of False Positive Rate (FPR) and False Negative Rate (FNR) for graphs of varying dimensions as sample size increases. (A) FPR and FNR for PLV (left) and torus graphs (right) using an alpha level of 0.05 for the edgewise hypothesis tests with no Bonferroni correction for the number of edges. PLV has high FPR for all sample sizes while torus graphs control the FPR; on the other hand, PLV has low FNR, but torus graphs is more conservative and for low sample sizes may be missing some edges. (B) Same as A, but with Bonferroni correction.*

References.

Forbes, P. G. and Lauritzen, S. (2015). Linear estimating equations for exponential families with application to Gaussian linear concentration models. *Linear Algebra and its Applications* **473** 261–283.

Hyvärinen, A. (2005). Estimation of non-normalized statistical models by score matching. *Journal of Machine Learning Research*.

Jammalamadaka, S. R. and Sarma, Y. (1988). A correlation coefficient for angular variables. *Statistical theory and data analysis II* 349–364.

Jammalamadaka, S. R. and Sengupta, A. (2001). *Topics in circular statistics* **5**. World Scientific.

Kass, R. E. and Vos, P. W. (2007). *Geometrical foundations of asymptotic inference*. John Wiley & Sons.

Klein, N., Orellana, J., Brincat, S., Miller, E. K. and Kass, R. E. (2019). Torus Graphs for Multivariate Phase Coupling Analysis.

Lauritzen, S. L. (1996). *Graphical models* **17**. Clarendon Press.

Navarro, A. K., Frellsen, J. and Turner, R. E. (2017). The multivariate generalised von Mises distribution: inference and applications. In *AAAI* 2394–2400.

Schreier, P. J. and Scharf, L. L. (2010). *Statistical signal processing of complex-valued data: the theory of improper and noncircular signals*. Cambridge university press.

Wainwright, M. J., Jordan, M. I. et al. (2008). Graphical models, exponential families, and variational inference. *Foundations and Trends® in Machine Learning* **1** 1–305.

Weisstein, E. W. (2017). Harmonic Addition Theorem. From MathWorld—A Wolfram Web Resource. Last visited on 2/6/2017.

Yu, S., Drton, M. and Shojaie, A. (2018). Graphical models for non-negative data using generalized score matching. *arXiv preprint arXiv:1802.06340*.

N. Klein
J. Orellana
R. E. Kass
Carnegie Mellon University
5000 Forbes Ave, Pittsburgh, PA 15213
E-mail: neklein@stat.cmu.edu
josue@cmu.edu
kass@stat.cmu.edu

S. Brincat
E. K. Miller
The Picower Institute for Learning and Memory
Department of Brain and Cognitive Science
Massachusetts Institute of Technology
77 Massachusetts Ave, Cambridge, MA 02139
E-mail: sbrincat@mit.edu
ekmiller@mit.edu