Abstract

High-dimensional neural recordings across multiple brain regions can be used to establish functional connectivity with good spatial and temporal resolution. We designed and implemented a novel method, Latent Dynamic Factor Analysis of High-dimensional time series (LDFA-H), which combines (a) a new approach to estimating the covariance structure among high-dimensional time series (for the observed variables) and (b) a new extension of probabilistic CCA to dynamic time series (for the latent variables). Our interest is in the cross-correlations among the latent variables which, in neural recordings, may capture the flow of information from one brain region to another. In simulations, we demonstrate that LDFA-H outperforms existing methods in the sense that it captures target factors even when noise correlation dominates cross-region correlation. We applied our method to local field potential (LFP) recordings from 192 electrodes in Prefrontal Cortex (PFC) and visual area V4 during a memory-guided saccade task. The results capture multiple lead-lag dependencies between PFC and V4, each associated with a different spatial distribution mode.

1 Introduction

New electrode arrays for recording electrical activity generated by large networks of neurons have created great opportunities, but also great challenges for statistical machine learning (e.g. Steinmetz et al., 2018). A key task is to develop methods that can track the flow of information across brain regions at millisecond precision. Local Field Potentials (LFPs) are signals that represent the bulk activity in relatively small volumes of tissue (Buzsaki et al., 2012; Einevoll et al., 2013) and they have been shown to correlate substantially with the BOLD fMRI signal (Logothetis et al., 2001; Magri et al., 2012), which is used for brain imaging studies. Typical LFP data sets may have dozens to hundreds of time series in each of two or more brain regions, recorded simultaneously across many experimental repetitions, or trials. We develop here a method that can uncover multidimensional interactions among two or more groups of high-dimensional time series, we demonstrate its effectiveness, and we apply it to LFP recordings from prefrontal cortex (PFC) and visual area V4.

A popular technique for analyzing multidimensional neural time series is Gaussian Process Factor Analysis (GPFA Yu et al., 2009). GPFA uses Gaussian processes to define a set of latent factors that map linearly to the multidimensional observations, which are also assumed to be driven by noise, analogously to the classical factor analysis model. Thus, one idea for solving the time-series group
interaction problem would be to extend GPFA to two or more groups, where the main interest would be the cross-group interactions. Our approach uses probabilistic CCA to carry out such an extension, but the framework allows far richer spatiotemporal dependencies than is typically assumed in GPFA, and it has been built to handle high-dimensional problems. Here, “spatial” dependence refers to dependence among the various observational time series and, in the neural context, this results from the spatial arrangement of the electrodes, each of which records one of the time series. In the usual setup of GPFA, two reasonable simplifications are made: first, the observation noise is assumed to be white and, second, the latent Gaussian processes are stationary. Our approach relaxes these assumptions: We allow the observation noise to have spatiotemporal dependence, which we have found more realistic, and we let the latent processes be non-stationary, so their dependence can evolve dynamically, which is important in our applications because cross-process dependence describes the sudden flow of information from one brain region to another during short epochs. We thus call our method LDFA-H. These generalizations come at a cost: we now have a high-dimensional time series problem within each brain region together with a high-dimensional covariance structure. We solve these high-dimensional problems by imposing sparsity of the dominant effects, building on related, but as yet unpublished work (Anonymous, 2020) that treats the high-dimensional covariance structure in the context of observational white noise, and by incorporating banded covariance structure as in Bickel and Levina (2008). In a simulation study, based on realistic synthetic time series, we verify recovery of cross-region structure even when some of our assumptions are violated, and even in the presence of high noise. We then apply the method to 192 LFP time series recorded simultaneously from both Prefrontal Cortex (PFC) and visual area V4, during a memory task, and find time-varying cross-region dependencies.

2 Latent Dynamic Factor Analysis of High-dimensional time series

We treat the case of two groups of time series observed, repeatedly, \( N \) times. Let \( X_{1,t} \in \mathbb{R}^{p_1} \) and \( X_{2,t} \in \mathbb{R}^{p_2} \) be \( p_1 \) and \( p_2 \) recordings at time \( t \) in each of the two groups, for \( t = 1, \ldots, T \). As in Yu et al. (2009), we assume that a \( q \)-dimensional latent factor \( Z_{1,t} \in \mathbb{R}^q \) drives each group, here, each brain region, according to the linear relationship

\[
X_{f,t} = \mu_{f,t} + \beta_f \cdot Z_{f,t} + \epsilon_{f,t},
\]

for brain region \( f = 1, 2 \), where \( \mu_{f,t} \in \mathbb{R}^{p_f} \) are mean vectors, \( \beta_f \in \mathbb{R}^{p_f \times q} \) are matrices of constant factor loadings, and \( \epsilon_{f,t} \in \mathbb{R}^{p_f} \) are errors centered at zero (independently of the latent vectors \( Z \)). We are interested in the pairwise cross-group dependencies of the latent vectors \( Z_{1,t} \) and \( Z_{2,t} \), for \( f = 1, \ldots, q \). As in our related work (Anonymous, 2020), we assume that the time series of these latent vectors follows a multivariate normal distribution

\[
\begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix} \sim \text{MVN}(0, \Sigma_f), \quad f = 1, \ldots, q,
\]

where \( \Sigma_f \) describes all of their simultaneous and lagged dependencies, both within and between the two vectors. We assume the \( N \) sets of random vectors \((\epsilon, Z)\) are independent and identically distributed. We let \( P_f \) be the correlation matrix corresponding to \( \Sigma_f \), and write its inverse as

\[
\Pi_f = P_f^{-1} = \begin{pmatrix} \Pi_f^{11} & \Pi_f^{12} \\ \Pi_f^{21} & \Pi_f^{22} \end{pmatrix}
\]

where \( \Pi_f^{11} \) and \( \Pi_f^{22} \) are the scaled auto-precision matrices and \( \Pi_f^{12} \) is the scaled cross-precision matrix. We now assume finite-range partial autocorrelation and cross-correlation for \((Z_{1,t}, Z_{2,t})\), so that \( \Pi_f^{11} \), \( \Pi_f^{22} \) and \( \Pi_f^{12} \) in Equation (3) have a banded structure. Specifically, for \( f, k, l = 1, 2 \), we assume there is a value \( h_{k,l}^f \) such that \( \Pi_f^{11} \) is a \((2h_{k,1}^f + 1)\)-diagonal matrix. Because our goal is to address the cross-region connectivity and lead-lag relationship, we are particularly interested in the estimation of \( \Pi_f^{12} \) for each latent factor \( f = 1, \ldots, q \). Note that the non-zero elements \( \Pi_f^{12}(t,s) \) determine associations between the latent pair \( Z_{1,t} \) and \( Z_{2,t} \), which are simultaneous when \( t = s \) and lagged when \( t \neq s \). Finally, we model the noise in Eq. (1) as a Gaussian random vector

\[
\text{Vec}(\epsilon^k) = (\epsilon_{1,k}^k; \epsilon_{2,k}^k; \ldots; \epsilon_{T,k}^k) \sim \text{MVN}(0, \Phi^k), \quad k = 1, 2,
\]
where we allow $\Phi^k$ to have non-zero off-diagonal elements to account for within-group spatiotemporal dependence. We assume $\Phi^k$ can be written in Kronecker product form

$$\Phi^k = \Phi_T^k \otimes \Phi_S^k, \quad k = 1, 2, \tag{5}$$

where $\Phi_T^k$ and $\Phi_S^k$ are the temporal and spatial components of $\Phi^k$, as is often assumed for spatiotemporal matrix-normal distributions, e.g., (Dawid [1981]. Although this is a strong approximation, implying, for instance, that the auto-correlation of every $X_i^k$ is proportional to $\Phi_T^k$, we regard $\Phi_k$ as a nuisance parameter: our primary interest is $\Sigma_f$ in Eq. (2). We also assume an auto-regressive process of order at most $h_k^k$, so that $\Gamma_k = (\Phi_T^k)^{-1}$ is a $(2h_k^k + 1)$-diagonal matrix. In our simulation we show that we can recover $\Sigma_f$ accurately even when the Kronecker product and bandedness assumptions fail to hold.

The model in Equations (1)-(5) generalizes other known models. First, when $q = 1$, and $Z^1 = Z^2$ remains constant over time, in the noiseless case ($\epsilon_k = 0$), it reduces to the probabilistic CCA model of Bach and Jordan (2005). Thus, model (1)-(5) can be viewed as a denoising, multi-level and dynamic version of probabilistic CCA. Second, when $k = 1$, the Gaussian processes are stationary, and the $\epsilon$ vectors are white noise, (1)-(5) reduces to GPFA (Yu et al. [2009]). Thus, (1)-(5) is a two-group, nonstationary extension of GPFA that allows for within-group spatio-temporal dependence.

**Identifiability and sparsity constraints** Despite the structure imposed on $\Phi_k$ in Eq. (5), parameter identifiability issues remain. Our model in Eqs. (1), (2) and (4) induces the marginal distribution of the observed data $(X^1, X^2)$:

$$(X_{1:i}^1; X_{1;i}^2; \ldots; X_{2:T}^2) \sim N((\mu_{1;1}^1; \mu_{1;2}^1; \ldots; \mu_{2;T}^2), S) \tag{6}$$

where $S$ is the marginal covariance matrix given by:

$$S = \begin{bmatrix} \Phi_T^1 \otimes \Phi_S^1 & 0 \\ 0 & \Phi_T^2 \otimes \Phi_S^2 \end{bmatrix} + \sum_{f=1}^{q} \begin{bmatrix} \Sigma_{f11}^1 \otimes (\beta_1^f \beta_1^f \top) + \Sigma_{f12}^2 \otimes (\beta_2^f \beta_2^f \top) \\ \Sigma_{f21}^1 \otimes (\beta_1^f \beta_1^f \top) + \Sigma_{f22}^2 \otimes (\beta_2^f \beta_2^f \top) \end{bmatrix}. \tag{7}$$

The family of parameters

$$\theta(\alpha^1, \alpha^2) = \left\{ \Sigma_1^{\{\alpha^1, \alpha^2\}}, \ldots, \Sigma_q^{\{\alpha^1, \alpha^2\}}, \Phi_T^1 - \sum_{f=1}^{q} \alpha_f^1 \beta_f^1 \beta_f^1 \top, \Phi_S^1 - \sum_{f=1}^{q} \alpha_f^2 \beta_f^2 \beta_f^2 \top, \Phi_T^2, \Phi_S^2, \beta_1, \beta_2, \mu_1, \mu_2 \right\} \tag{8}$$

where $\Sigma_f^{\{\alpha^1, \alpha^2\}} = \left\{ \Sigma_f + \left[ \alpha_f^1 \Phi_T^1 \mid 0 \mid 0 \mid 0 \mid \alpha_f^2 \Phi_S^1 \right] \right\}$, induce the same marginal distribution in Eq. (6), for all $\alpha^1, \alpha^2 \in \mathbb{R}^2$ (notice that $\theta = \theta(0,0) = \{ \Sigma_1, \ldots, \Sigma_q, \Phi_T^1, \Phi_S^1, \Phi_T^2, \Phi_S^2, \beta_1, \beta_2, \mu_1, \mu_2 \}$ is the original parameter). Preliminary analysis of LFP data indicated that strong cross-region dependence occurs relatively rarely. We therefore resolve this lack of identifiability by choosing the solution given by maximizing the likelihood with an L1 penalty, under the assumption that the inverse cross-correlation matrix $\Pi_f^2$ is a sparse $(2h_k^k + 1)$-diagonal matrix.

**Latent Dynamic Factor Analysis of High-dimensional time series (LDFA-H)** Given $N$ simultaneously recorded pairs of neural time series $\{X^1[n], X^2[n]\}_{n=1, \ldots, N}$, the maximum penalized likelihood estimator (MPLE) of the inverse correlation matrix of the latent variables solves

$$\left( \Pi_1, \ldots, \Pi_q \right) = \text{argmin } -\frac{1}{N} \sum_{n=1}^{N} l(\theta; X^1[n], X^2[n]) + \sum_{f=1}^{q} \sum_{k,l=1}^{2} \| \Lambda_f^{kl} \otimes \Pi_f^{kl} \|_1 \tag{9}$$

s.t. $\Gamma_k^k$ is $(2h_k^k + 1)$-diagonal,

where the log-likelihood is

$$l(\theta; X^1, X^2) = -\log \det S - (X_{1;i}^1 - \mu_{1;i}^1; \ldots; X_{2;T}^2 - \mu_{2;T}^2)^\top S^{-1}(X_{1;i}^1 - \mu_{1;i}^1; \ldots; X_{2;T}^2 - \mu_{2;T}^2), \tag{10}$$
We note that the simulation does not satisfy some of the model assumptions in Section 2. The noise \( N \) we simulated

With the noise correlation and latent factor correlation disentangled, the M-step reduces to easy

(friedman et al., 2007) and the minimization with respect to \( \Phi \)

The magnitudes of cross-region correlation and within-region noise auto-correlation are quantified by

The bandwidths are chosen using domain knowledge and preliminary data analyses. We determine

the determinant of each matrix, known as the generalized variance (Sengupta, 2004); their logarithms

all pairs of simulated time series, are shown in Fig. 1b, for four levels of white noise contamination.

correlations. The resulting temporal noise correlation matrices, found by averaging correlations over

experimental data analyzed in Section 3.2, first permuted to remove cross-region correlations, then

and correlation matrix \( k \)

3 Results

One major novelty of our method is its accounting of auto-correlated noise in neural time series

to better estimate cross-region associations in CCA type analysis. This is illustrated in Section 3.1

based on simulated data. Then in Section 3.2, we apply LDFA-H to experimental data to examine the

lead-lag relationships across two brain areas and the spatial distribution of factor loadings.

3.1 LDFA-H retrieves cross-correlations even when noise auto-correlations dominate

We simulated \( N = 1000 \) i.i.d. neural time series \( X^k \) of duration \( T = 50 \) from Eq. (1) for brain regions \( k = 1, 2 \). The latent time series \( Z^k \) were generated from Eq. (2) with \( q = 1 \) pair of factors and correlation matrix \( P_1 \) depicted in Fig. 1a. The noise \( \epsilon^k \) was taken to be the \( N = 1000 \) trials of the experimental data analyzed in Section 3.2 first permuted to remove cross-region correlations, then contaminated with white noise to modulate the strength of noise correlation relative to cross-region correlations. The resulting temporal noise correlation matrices, found by averaging correlations over all pairs of simulated time series, are shown in Fig. 1b for four levels of white noise contamination. The magnitudes of cross-region correlation and within-region noise auto-correlation are quantified by the determinant of each matrix, known as the generalized variance (Sengupta, 2004); their logarithms are provided atop the panels in Fig. 1a and Fig. 1b. Generalized variance ranges from 0 (identical signals) to 1 (independent signals). Other simulation details are in Appendix B.

We note that the simulation does not satisfy some of the model assumptions in Section 2. The noise vectors \( \epsilon^k \) are not matrix-variate distributed as in Eqs. (4) and (5) and the derived \( \Gamma^k_f \) does not satisfy

with \( S \) defined in Eq. (7), and the constraints are

\[
\Lambda_{f,(t,s)}^{kl} = \begin{cases} 
\infty, & (t,s) : |t-s| > h_f^{kl}, \\
\lambda_f, & (t,s) : 0 < |t-s| \leq h_f^{kl}, \quad k \neq l, \\
0, & \text{otherwise.}
\end{cases}
\]

(11)

for factor \( f = 1, \ldots, q \) and brain region \( k = 1, 2 \). The first constraint forces the corresponding \( \Pi_f^{k1} \) to zero and thus imposes a banded structure for \( \Pi_f^{kj} \), and the second assigns the same sparsity constraint \( \lambda_f \) on the off-diagonal elements of \( \Pi_f^{j2} \). Finally, to make calibration of tuning parameters computationally feasible, we use the same bandwidth to latent precision and noise precision within a region and to latent precision across regions, respectively, and the same sparsity parameter:

\[
h_f^{kl} = h_c^{k} = h_{auto}, \quad h_f^{j2} = h_{cross} \quad \text{and} \quad \lambda_f = \lambda_{cross},
\]

for each factor \( f = 1, \ldots, q \) and region \( k = 1, 2 \). The remaining hyperparameters become the temporal bandwidths \( h_{auto} \) and \( h_{cross} \); the sparsity penalty \( \lambda_{cross} \); and the number of latent factors \( q \).

The bandwidths are chosen using domain knowledge and preliminary data analyses. We determine

the remaining parameters by 5-fold cross-validation (CV).

Solving Eq. (3) requires \( S^{-1} \). Because it is not available analytically and a numerical approximation is computationally prohibitive, we solve Eq. (3) using an EM algorithm (Dempster et al., 1977). Let \( \theta^{(r)} \) be the parameter estimate at the \( r \)-th iteration. We consider the data \( \{X^T[n], X^2[n]\}_{n=1,\ldots,N} \) to be incomplete observations of \( \{X^1[n], Z^1[n], X^2[n], Z^2[n]\}_{n=1,\ldots,N} \). In the E-step, we estimate the conditional mean and covariance matrix of each \( \{Z^1[n], Z^2[n]\} \) with respect to \( \{X^1[n], X^2[n]\} \) and \( \theta^{(r)} \). Given these sufficient statistics, the MPLE decomposes into two separate minimizations of

1. the negative log-likelihood of \( \Sigma_f \), w.r.t. the latent factor model (Eq. (2)) and
2. the negative log-likelihood of \( \Phi^k_S, \Phi^k_T, \Phi^k_f, \beta^1, \beta^2, \mu^1, \mu^2 \) w.r.t. the observation model (Eqs. (1) and (4)).

With the noise correlation and latent factor correlation disentangled, the M-step reduces to easy sub-problems. For example, the minimization with respect to \( \Sigma_f \) is a graphical Lasso problem (friedman et al., 2007) and the minimization with respect to \( \Phi^k_S \) and \( \Phi^k_T \) is a maximum likelihood estimation of a matrix-variate distribution (David, 1981). We thus obtain an affordable M-step, and alternating E and M-steps produces a solution to the MPLE problem. We derive the full formulations in Appendix A. Code is provided at https://github.com/AutoAnonymous/ldfa_anon.
Figure 1: Simulation settings. (a) True correlation matrix $P_1$ for latent factors $Z_1$ and $Z_2$; from model in Eq. (2); close-up of the cross-correlation matrix; corresponding precision matrix $\Pi_1 = P_1^{-1}$; and close-up of cross-precision matrix $\Pi_{12}$ (Eq. (3)). Matrix axes represent the duration, $T = 50$ ms, of the time series. Factors $Z_1$ and $Z_2$ are associated in two epochs: $Z_2$ precedes $Z_1$ by 7 ms from $t = 13$ to 19 ms, and $Z_1$ precedes $Z_2$ by 7 ms from $t = 33$ to 42 ms. (b) Noise auto-correlation matrices (Eq. (5)) for pairs of simulated time series at four strength levels. Log det in (a) and (b) measure correlation strengths.

Figure 2: Simulation results: LDFA-H cross-precision matrix estimates. Estimates of $\Pi_{12}$, shown in the right-most panel of Fig. 1a, using LDFA-H, for the four noise auto-correlation strengths shown in Fig. 1b. LDFA-H identified the true cross-area connections at all noise strengths.

a banded structure as in Eq. (9). Also, the latent auto-correlations (Fig. 1a) are not banded as assumed in Eq. (9).

We applied LDFA-H with $q = 1$ factor, $h_{\text{cross}} = 10, h_{\text{auto}}$ equal to the maximum order of the auto-correlations in the 2000 observed simulated time series, and $\lambda_{\text{cross}}$ determined by 5-fold CV. Fig. 2 shows LDFA-H cross-precision matrix estimates corresponding to the four level of noise correlation shown in Fig. 1a. They closely match the true $\Pi_{12}$ shown in the right most panel of Fig. 1a.

We also applied five other methods to estimate cross-region connections in the simulated data. They include the popular averaged pairwise correlation (APC); correlation of averaged signals (CAS); and CCA (Hotelling, 1992), applied to the $NT$ observed pairs of multivariate random vectors $\{X_{1,i}, X_{2,j}\}_{n, i, [N] \times [T]}$ to estimate the cross-correlation matrix between the canonical variables; as well as DKCCA (Rodu et al., 2018) and Method A (Anonymous, 2020). The first four methods do not explicitly provide cross-precision matrix estimates, so we display their cross-correlation matrix estimates in Fig. 3 along with LDFA-H cross-correlation estimates in the last row. It is clear that only LDFA-H successfully recovered the true cross-correlations shown in the second panel of Fig. 1a at all auto-correlated noise levels.

3.2 Experimental Data Analysis from Monkey Saccade Task

We now report the analysis of LFP data in areas PFC and V4 of a monkey during an eye saccade task. One trial of the experiment consisted of four stages: (i) fixation: the animal fixated at the center of
Figure 3: **Simulation results: cross-correlation matrix estimates.** Estimates of $\Sigma_{12}^{12}$ using (a) averaged pairwise correlation (APC), (b) correlation of averaged signal (CAS), (c) canonical correlation analysis (CCA, Hotelling [1992]), (d) dynamic kernel CCA (DKCCA, Rodu et al. [2018]), (e) Method A (Anonymous [2020]), and (f) LDFA-H under four noise correlation levels. Only LDFA-H successfully recovered the true cross-correlation at all noise auto-correlation strengths.

The screen; (ii) cue: a cue appeared on the screen randomly at one of eight locations; (iii) delay: the animal had to remember the cue location while maintaining eye fixation; (iv) choice: the monkey made a saccade to the remembered cue location. We focused our analysis on the 500 ms delay period, when the animal both processed cue information and prepared a saccade. LFP data were recorded for $N=1000$ trials by two 96-electrode Utah arrays each implanted in PFC and V4, $\beta$ band-passed filtered, and down-sampled from 1 kHz to 100 Hz.
Figure 4: Experimental data results for the top 4 factors. (a) Factor loadings, rescaled between -1 and 1, plotted against the electrode coordinates (µm) of the V4 Utah array. Factors have different spatial modes over the physical space of the Utah array. \( \log_{10} \| \Sigma_f \|_F^2 \), written atop the panels, measures the strength of each factor. Notice that the strength of the first factor is over 100 order larger than the second largest factor. (b) Dynamic information flow from V4 \( \rightarrow \) PFC (blue) and PFC \( \rightarrow \) V4 (orange). In and out flows seem to peak either at the beginning or at the end of the delay period, and different couplings of the two flows may indicate different communication modes between V4 and PFC.

We applied LDFA-H using \( h_{\text{auto}} = h_{\text{cross}} = 10 \), corresponding to 100 ms (at 100 Hz); the LFP \( \beta \)-power envelopes have frequencies between 12.5 Hz to 30 Hz, and \( h_{\text{auto}} = 10 \) just enables the slowest filtered signal to complete one full oscillation period. The other tuning parameters were determined by 5-fold CV over \( \lambda_{\text{cross}} \in \{0.0001, 0.001, 0.01, 0.1\} \) and \( q \in \{5, 10, 15, 20, 25, 30\} \), yielding optimal values \( \lambda_{\text{cross}} = 0.01 \) and \( q = 10 \). The fitted factors were ranked based on the Frobenius norms of their covariance matrices \( \| \Sigma_f \|_F^2 \); norms are plotted versus \( f \) in decreasing order in Fig. C.1 and \( \log_{10} \| \Sigma_f \|_F^2 \) of the four most dominant factors are provided atop each panel in Fig. 4a. Factor loadings (slightly smoothed over space) for the 96 V4 electrodes are shown in Fig. 4a for the top four factors (first four columns of the estimate of \( \beta_k \) in Eq. (9), with area \( k = 1 \) being V4), arranged spatially according to electrode positions on the Utah array. The factors have different spatial modes over the physical space of the Utah array. For example, the dominant first factor has positive weights concentrated along a vertical strip on the left of the array, especially in the mid-to-upper left, and negative weights along a vertical strip to the right, separated by roughly 2000 microns.

We also summarized, for each factor \( f \), the temporal information flow at time \( t \) from V4 to PFC and to V4 from PFC with

\[
I_{f,\text{out}}(t) = \sum_{t' \geq t} \left| \hat{\Pi}^2_{f, (t,t')} \right| \quad \text{and} \quad I_{f,\text{in}}(t) = -\sum_{t' < t} \left| \hat{\Pi}^2_{f, (t,t')} \right|,
\]

respectively, where \( \hat{\Pi} \) is the inverse correlation matrix estimate in Eq. (9). Figure 4b displays smoothed \( I_{f,\text{out}}(t) \) and \( I_{f,\text{in}}(t) \) as functions of \( t \in [100, 400] \) ms for the top four factors. Lead-lag relationships between V4 and PFC change dynamically over time, and the information flow tends to peak either at the beginning of the delay period, when the animal must remember the cue, or at the end, when it must make a saccade decision. We observe the strongest information flow from V4 to PFC in the dominant first factor, when the animal needs to process the visual signal from V4 during the delay period. We also observe different information flow patterns: (1) asymmetric (first factor): information in and out avoid conflicting with each other and peak at different times; (2) parallel (second factor): information in and out are parallel and oscillate over time; (3) symmetric (third and fourth factors): information in and out are symmetric, meaning they are activated and suppressed simultaneously.
4 Conclusion

To identify dynamic interactions across brain regions we have developed LDFA-H, a nonstationary, multi-group extension of GPFA that allows for within-group spatio-temporal dependence among high-dimensional neural recordings. Although we treated the two-group case, and applied it to interactions across two brain regions, several groups can be handled with obvious, and straightforward modifications. The approach could, in principle, be applied to many different types of time series, but it has some special features: first, like all methods based on sparsity, it assumes a small number of large effects are of primary interest; second, it uses repetitions, here, repeated trials, to identify time-varying dependence; third, because the within-group spatio-temporal structure is not of interest, the method can remain useful even with some modest within-group model misspecification.

We applied LDFA-H to LFP data, while GPFA has been applied mainly to neural spike count data. In the analysis of spike counts, we have been struck by the strong attenuation of effects due to Poisson-like noise, as discussed in Vinci et al. (2018) and references therein. A version of LDFA-H built for Poisson-like counts, or for point processes, could be the subject of additional research. It may also be advantageous to model spatial dependence explicitly, perhaps based on physical distance between electrodes, analogously to what was done in Vinci et al. (2018), and there may be important simplifications available in the temporal structure as well. In addition, it would be helpful to include statistical inferences for assessing effects. In the future, we hope to pursue these possible directions, and refine the application of this promising approach to the analysis of high-dimensional neural data.
Broader Impact

While progress in understanding the brain is improving life through research, especially in mental health and addiction, in no case is any brain disorder well understood mechanistically. Faced with the reality that each promising discovery inevitably reveals new subtleties, one reasonable goal is to be able to change behavior in desirable ways by modifying specific brain circuits and, in animals, technologies exist for circuit disruptions that are precise in both space and time. However, to determine the best location and time for such disruptions to occur, with minimal off-target effects, will require far greater knowledge of circuits than currently exists: we need good characterizations of interactions among brain regions, including their timing relative to behavior. The over-arching aim of our research is to provide methods for describing the flow of information, based on evolving neural activity, among multiple regions of the brain during behavioral tasks. Such methods can lead to major advances in experimental design and, ultimately, to far better treatments than currently exist.

Acknowledgments and Disclosure of Funding

Use unnumbered first level headings for the acknowledgments. All acknowledgments go at the end of the paper before the list of references. Moreover, you are required to declare funding (financial activities supporting the submitted work) and competing interests (related financial activities outside the submitted work). More information about this disclosure can be found at: https://neurips.cc/Conferences/2020/PaperInformation/FundingDisclosure

Do not include this section in the anonymized submission, only in the final paper. You can use the ack environment provided in the style file to automatically hide this section in the anonymized submission.

References


A  EM-algorithm to fit LDFA-H (Section 2)

Initialization  Let \( \hat{\theta}^{(0)} = \{ \hat{\Sigma}_1^{(0)}, \ldots, \hat{\Sigma}_q^{(0)}, \hat{\Phi}_S^{(0)}, \hat{\Phi}_T^{(1,0)}, \hat{\Phi}_T^{(2,0)}, \hat{\beta}_1^{(0)}, \hat{\beta}_2^{(0)}, \hat{\mu}_1^{(0)}, \hat{\mu}_2^{(0)} \} \) be the initial parameter value. Since the MPLE objective function for LDFA-H given in Eq. (9) is not guaranteed convex, an EM-algorithm may find a local minimum according to a choice of the initial value. Hence a good initialization is crucial to a successful estimation. Here we suggest an initialization by a canonical correlation analysis (CCA).

Let \( \{ X_1^{(n)}, X_2^{(n)} \}_{n=1, \ldots, N} \) be \( N \) simultaneously recorded pairs of neural time series. We can view them as \( NT \) recorded pairs of multivariate random vectors \( \{ X_1^{(n, t)}, X_2^{(n, t)} \}_{(n, t) \in [N] \times [T]} \). We obtain \( \hat{\beta}_1^{(0)} \) and \( \hat{\beta}_2^{(0)} \) by CCA as follows:

\[
\hat{\beta}_1^{(0)}, \hat{\beta}_2^{(0)} = \underset{\beta_1 \in \mathbb{R}^p, \beta_2 \in \mathbb{R}^q}{\text{argmax}} \ \frac{\beta_1^T S^{12} \beta_2}{\sqrt{\beta_1^T S^{11} \beta_1} \sqrt{\beta_2^T S^{22} \beta_2}} \tag{A.1}
\]

where

\[
S^{11} = \frac{1}{NT} \sum_{n, t} (X_1^{(n, t)} - \frac{1}{NT} \sum_{n, t} X_1^{(n, t)}) (X_1^{(n, t)} - \frac{1}{NT} \sum_{n, t} X_1^{(n, t)})^T
\]
\[
S^{22} = \frac{1}{NT} \sum_{n, t} (X_2^{(n, t)} - \frac{1}{NT} \sum_{n, t} X_2^{(n, t)}) (X_2^{(n, t)} - \frac{1}{NT} \sum_{n, t} X_2^{(n, t)})^T \tag{A.2}
\]
\[
S^{12} = \frac{1}{NT} \sum_{n, t} (X_1^{(n, t)} - \frac{1}{NT} \sum_{n, t} X_1^{(n, t)}) (X_2^{(n, t)} - \frac{1}{NT} \sum_{n, t} X_2^{(n, t)})^T.
\]

According to the equivalence between CCA and probabilistic CCA shown by A. Anonymous, it gives an estimate of the first latent factors

\[ \hat{Z}_1^{(0)}[n] = \hat{\beta}_1^{(0)} X_1^{(n)} \tag{A.3} \]

for \( n = 1, \ldots, N \) and \( k = 1, 2 \). The initial second latent factors \( \hat{Z}_2^{(0)} \) and the corresponding factor loading \( \hat{\beta}_2^{(0)} \) is similarly set by the second pair of canonical variables, and so on. Then we assign the empirical covariance matrix of \( \{ \hat{Z}_1^{(0)}[n], \hat{Z}_2^{(0)}[n] \}_{n \in [N]} \) to the initial latent covariance matrix \( \hat{\Sigma}_n^{(0)} \) for \( f = 1, \ldots, q \) and the matrix-variate normal estimate \( \{ \hat{\mu}_f^{(0)}[n] := \hat{X}_f^{(n)}[n] - \hat{\beta}_f^{(0)} \hat{Z}_2^{(0)}[n] \}_{n \in [N]} \) to \( \hat{\Phi}_S^{(0)} \) and \( \hat{\Phi}_T^{(0)} \) for \( k = 1, 2 \). Along \( \hat{\mu}_f^{(0)} := \frac{1}{N} \sum_{n=1}^{N} X_f^{(n)} \), the above parameters comprises the initial parameter set \( \hat{\theta}^{(0)} \).

However, we cannot run an E-step on the above parameter set because \( \hat{\Phi}_f^{(0)} \) is not invertible. We instead pick one of its unidentifiable parameter sets \( \hat{\theta}^{(0)} \cdot \{ \alpha, \gamma \} \), defined in Eq. (8), with all \( \hat{\Phi}_f^{(0)} \)'s and \( \hat{\Sigma}_n^{(0)} \)'s invertible. Specifically, we take

\[ \alpha_f^{\frac{k}{2}} = \frac{1}{2} \lambda_{\text{min}} \left( \hat{\Sigma}_f^{(0)} \right)^{1/2} \left[ \begin{array}{c} \hat{\Phi}_T^{(0)} \\ 0 \end{array} \right]^{-1} \left( \hat{\Sigma}_f^{(0)} \right)^{1/2} \tag{A.4} \]

for \( f = 1, \ldots, q \) and \( k = 1, 2 \) where \( \lambda_{\text{min}}(A) \) is the smallest eigenvalue of symmetric matrix \( A \). Henceforth, we note \( \hat{\theta}^{(0)} \cdot \{ \alpha, \gamma \} \) by \( \hat{\theta}^{(0)} \). For \( t = 1, 2, \ldots \), we iterate the following E-step and M-step until convergence.

Another promising initialization is by finding time \( (t, s) \) on which the canonical correlation between \( X_1^{(t)} \) and \( X_2^{(s)} \) maximizes. i.e., we initialize \( \hat{\beta}_1^{(0)} \) and \( \hat{\beta}_2^{(0)} \) by

\[
\hat{\beta}_1^{(0)}, \hat{\beta}_2^{(0)} = \underset{\beta_1 \in \mathbb{R}^p, \beta_2 \in \mathbb{R}^q}{\text{argmax}} \ \frac{\beta_1^T S^{12}_{(t, s)} \beta_2^T}{\sqrt{\beta_1^T S^{11}_{(t, s)} \beta_1^T} \sqrt{\beta_2^T S^{22}_{(s, s)} \beta_2^T}} \text{ such that } |t - s| < h_{\text{cross}}. \tag{A.5}
\]
where

\[ S^{11}_{(t,t)} = \frac{1}{N} \sum_{n,t} (X^1_{t,n} - \frac{1}{N} \sum_n X^1_{t,n})(X^1_{t,n} - \frac{1}{N} \sum_n X^1_{t,n})^T \]

\[ S^{22}_{(s,s)} = \frac{1}{N} \sum_{n,s} (X^2_{s,n} - \frac{1}{N} \sum_n X^2_{s,n})(X^2_{s,n} - \frac{1}{N} \sum_n X^2_{s,n})^T \]

\[ S^{12}_{(t,s)} = \frac{1}{N} \sum_{n,t} (X^1_{t,n} - \frac{1}{N} \sum_n X^1_{t,n})(X^2_{s,n} - \frac{1}{N} \sum_n X^2_{s,n})^T. \]  

(A.6)

for \((t, s) \in [T] \times [T] \). Then the other parameters are initialized as above. We can even take an ensemble approach in which we fit LDFA-H on different initialized values and pick the estimate with the minimum cost function (Eq. (9)).

Now, for \( r = 1, 2, \ldots \), we alternate an E-step and an M-step until the target parameter \( \Pi_f \) convergences.

**E-step** Given \( \hat{\theta} := \hat{\theta}^{(r-1)} \) from the previous iteration, the conditional distribution of latent factors \( Z^1[n] \) and \( Z^2[n] \) with respect to observed data \( X^1[n] \) and \( X^2[n] \) on trial \( n = 1, \ldots, N \) follows

\[ (Z^1_{1,n}; Z^1_{2,n}; \ldots; Z^2_{q,n}; [n]) \mid X^1[n], X^2[n] \sim \text{MVN} \left( m_{Z|X}^{(r)} [n], V_{Z|X}^{(r)} \right), \]  

(A.7)

where

\[ V_{Z|X}^{(r)} = \begin{pmatrix} V_{Z_1,Z_1|X}^{(r)} & \cdots & V_{Z_1,Z_q|X}^{(r)} \\ \vdots & \ddots & \vdots \\ V_{Z_q,Z_1|X}^{(r)} & \cdots & V_{Z_q,Z_q|X}^{(r)} \end{pmatrix} = \begin{pmatrix} W_{Z_1,Z_1|X}^{(r)} & \cdots & W_{Z_1,Z_q|X}^{(r)} \\ \vdots & \ddots & \vdots \\ W_{Z_q,Z_1|X}^{(r)} & \cdots & W_{Z_q,Z_q|X}^{(r)} \end{pmatrix}^{-1} \]  

(A.8)

and

\[ m_{Z|X}^{(r)} [n] = \begin{pmatrix} m_{Z_1|X}^{(r)}[n]; m_{Z_2|X}^{(r)}[n]; \ldots; m_{Z_q|X}^{(r)}[n] \end{pmatrix} \]

\[ = V_{Z|X}^{(r)} \left( \hat{\beta}^1 \hat{\Gamma}_X X^1[n]\hat{\Gamma}_X^T; \hat{\beta}^2 \hat{\Gamma}_X X^2[n]\hat{\Gamma}_X^T; \ldots; \hat{\beta}_q \hat{\Gamma}_X X^2[n]\hat{\Gamma}_X^T \right) \]  

(A.9)

given

\[ W_{Z_f,Z_g|X}^{(r)} = \begin{pmatrix} \hat{\beta}_f \hat{\Gamma}_X \hat{\Gamma}_X^T \hat{\beta}_g \hat{\Gamma}_X^T & 0 \\ 0 & \hat{\beta}_f \hat{\Gamma}_X \hat{\Gamma}_X^T \hat{\beta}_g \hat{\Gamma}_X^T \end{pmatrix} + I_{(f=g)} \hat{\Omega}_f, \quad \hat{I}_{(f=g)} = \begin{cases} 1, & f = g \\ 0, & \text{o.w.} \end{cases} \]  

(A.10)

for \( f, g = 1, \ldots, q \).

**M-step** We find \( \hat{\theta}^{(r)} \) which maximize the conditional expectation of the penalized likelihood under the same constraints in Eq. (9), i.e.

\[ \hat{\theta}^{(r)} = \arg\min \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}_{Z[n]|X[n], \hat{\theta}^{(r-1)}} \left[ \log p(X^1[n], X^2[n], Z^1[n], Z^2[n]; \hat{\theta}^{(r-1)}) \right] \]

\[ + \sum_{k=1}^{q} \sum_{k=1}^{2} \| \Lambda_f^{k_l} \odot \Pi_f^{k_l} \|_1 \text{ s.t. } \hat{\Gamma}_f^{k_l} \text{ is } (2h_e^{k_l} + 1)\text{-diagonal} \]  

(A.11)

where \( p \) is the probability density function of our model in Eqs. (1), (4) and (5) and the expectation \( \mathbb{E}_{Z[n]|X[n], \hat{\theta}^{(r-1)}} \) follows the conditional distribution in Eq. (A.7). Taking a block coordinate descent approach, we solve the optimization problem by alternating M1 - M4.

M1: With respect to latent precision matrices \( \Omega_f \), Eq. (A.11) reduces to a graphical Lasso problem,

\[ \hat{\Omega}_f^{(r)} = \arg\min_{\Omega_f} \left\{ -\log \det(\Omega_f) + \text{tr} \left( \Omega_f \left( V_{Z_f|X}^{(r)} + \mathbb{E}_{Z_f|X}[m_{Z_f|X}^{(r)}] \right) \right) \right\} \]  

(A.12)
We simulated realistic data with known cross-region connectivity as follows. Simulating 
\[ \hat{\theta}^{(r)} \] where 
\[ \Gamma = B \) Simulation details (Section 3)

M2: With respect to \( \Gamma_k \), Eq. (A.11) reduces to an estimation of matrix-variate normal model (Zhou 2014). The estimation problem can be formulated as

\[
\hat{\Gamma}^{(r)}_S = \frac{1}{T} \left( \hat{E} \left[ m^{(r)}_s | X \right] m^{(r) \top}_s | X \right] + \sum_{f,g=1}^q \text{tr}(V^{(r)} f g \Gamma^{(r)} S \beta_f \beta_g \Gamma^{(r)} T) \right) \tag{A.13}
\]

and

\[
\hat{\Gamma}^{(r)}_T = \underset{\Gamma_T}{\text{argmin}} \left\{ -\log \det(\Gamma_T^{(r)}) + \frac{1}{p_k} \text{tr} \left( \Gamma_T^{(r)} \left( \sum_{f,g=1}^q (\beta_f^{(r)} \Gamma_T^{(r)} \beta_g^{(r)}) V^{(r)} f g \right) \right) + \hat{E} \left[ m^{(r) \top}_s | X \Gamma_T^{(r)} \hat{\theta}^{(r)}_s \right] \right\} \tag{A.14}
\]

s.t. \( \hat{\Gamma}^{(r)}_T \) is \( (2h_k^T + 1) \)-diagonal

for each \( k = 1,2 \) where \( m^{(r)}_{s} | X = X_k - \beta_k m^{(r)}_{s} | X - \mu_k \) and \( \hat{E}[A] \) is the empirical mean of a random matrix \( A \). The estimation of \( \Gamma_T^{(r)} \) under the bandedness constraint is tractable with modified Cholesky factor decomposition approach with bandwidth \( h_k + 1 \) using the procedure by Bickel and Levina (2008).

M3: With respect to \( \beta_k \), Eq. (A.11) reduces to a quadratic program

\[
\hat{\beta}^{(r)} = \underset{\beta_k}{\text{argmax}} \sum_{t,s} \Gamma_{T,(t,s)} \beta_k \left( V \beta_k \left( Z^{(r)}_{t,s} | X + \hat{\text{Cov}}[m^{(r)}_{s} | X, m^{(r)}_{s} | X] \right) \right) - 2 \sum_{t,s} \Gamma_{T,(t,s)} \text{tr} \left( \hat{\Gamma}_S \beta_k \hat{\text{Cov}}[X_k^{(r)} | X, m^{(r)}_{s} | X] \right) \tag{A.15}
\]

where \( \Gamma_{T,(t,s)} \) is the \( (t,s) \) entry in \( \Gamma_T^{(r)} \) and \( \hat{\text{Cov}}(A,B) \) is the empirical covariance matrix between random vectors \( A \) and \( B \). The analytic form of the solution is given by

\[
\beta_k = \left( \sum_{t,s} \Gamma_{T,(t,s)} \left( V \beta_k \left( Z^{(r)}_{t,s} | X + \hat{\text{Cov}}[m^{(r)}_{s} | X, m^{(r)}_{s} | X] \right) \right) \right)^{-1} \left( \sum_{t,s} \Gamma_{T,(t,s)} \hat{\text{Cov}}[m^{(r)}_{s} | X, X_k^{(r)}] \right) \tag{A.16}
\]

M4: With respect to \( \mu_k \), it is straightforward that Eq. (A.11) yields

\[
\hat{\mu}^{(r)}_s = \hat{E} \left[ X_k - \sum_{f=1}^q \beta_f^{(r)} m^{(r) \top}_f | X \right].
\]

B Simulation details (Section 3)

We simulated realistic data with known cross-region connectivity as follows. Simulating \( q = 1 \) pair of latent time-series \( Z^k \) from Equation (2), we introduced an exact ground-truth for the inverse cross-correlation matrix \( \Pi^T \) by setting:

\[
\Pi_1 = \left[ \left( (P_{i,0}^{(1)})^{-1} \right) 0 \right] \left( (P_{i,0}^{(2)})^{-1} \right) \right] + \left[ \begin{array}{c} \Pi_{11}^{(1)} \Pi_{12}^{(1)} \Pi_{11}^{(2)} \Pi_{12}^{(2)} \end{array} \right] \tag{B.1}
\]

where \( D^1 \) and \( D^2 \) are diagonal matrices with elements \( D_{(t,t)}^1 = \sum_s \Pi_{11}^{(1)}(t,s) \) and \( D_{(s,s)}^2 = \sum_s \Pi_{11}^{(2)}(s,s) \), which ensures that the matrix on the right hand side is positive definite. The matrix on the left hand side contains the auto-precision matrices of the two latent time series, with elements simulated from the squared exponential function:

\[
P_{1,0}^{(k)} = \left[ \exp \left( -\epsilon_k (t-s)^2 \right) \right]_{t,s} + \lambda I_T, \tag{B.2}
\]

with \( \epsilon^1 = 0.105 \) and \( \epsilon^2 = 0.142 \), chosen to match the observed LFPs autocorrelations in the experimental dataset (Section 3.2). We added the regularizer \( \lambda I_T, \lambda = 1 \), to render \( P^{kk} \) invertible.
Figure C.1: Squared Frobenius norms of covariance matrix estimates, $\hat{\Sigma}_f$, for all factors $f = 1, \ldots, 10$. Notice that the amplitudes of the top four factors dominate the others.

We designed the true inverse cross-correlation matrix $\Pi^{12}$ to induce lead-lag relationship between $Z_1$ and $Z_2$ in two epochs as depicted in the right-most panel of Fig. 1a. Specifically, the elements of $\Pi^{12}$ were set:

$$\Pi^{12}_{(t,s)} = \begin{cases} -r, & \text{where } Z_1^1, t \text{ and } Z_2^1, s \text{ partially correlate,} \\ 0, & \text{elsewhere,} \end{cases}$$

where the association intensity $r = 0.6$ was chosen to match our cross-correlation estimate in the experimental data (Section 3.2). Finally, we rescaled $P_1 = \Pi^{12}_1$ to have diagonal elements equal to one. The corresponding factor loading vector $\beta_1^k$ was randomly generated from standard multivariate normal distribution and then scaled to have $\|\beta_1^k\|_2 = 1$.

We generated the noise $\epsilon^k$ from the $N = 1000$ trials of the experimental data analyzed in Section 3.2. First, we permuted the trials in one region to remove cross-region correlations. Let $\{Y_1^k[n], Y_2^k[n]\}_{n=1,\ldots,N}$ be the permuted dataset. Then we contaminated the dataset with white noise to modulate the strength of noise correlation relative to cross-region correlations, i.e.

$$\epsilon^k_{i,t} = Y^k_{i,t} - \mu^k_{i,t} + \eta^k_{i,t}, \quad \eta^k_{i,t} \sim \text{MVN} \left( 0, \lambda \hat{\text{Cov}}[Y^k_{i,t}] \right), \quad \text{and} \quad \mu^k_{i,t} = \hat{E}[Y^k_{i,t}]$$

where $\hat{E}[Y^k_{i,t}]$ and $\hat{\text{Cov}}[Y^k_{i,t}]$ were the empirical mean and covariance matrix of $Y^k_{i,t}$, respectively, for $k = 1, 2, t = 1, \ldots, T$. The noise auto-correlation level was modulated by $\lambda \in \{2.78, 1.78, 0.44, 0.11\}$. We also obtained $\Sigma_1$ by scaling $P_1$ so that $\Sigma_{1,(t,s)}^{kk} = \beta_1^k \top S^k t \beta_1^k$. Putting all the pieces together, we generated observed time series by Eq. (1).

C Experimental data analysis details (Section 3.2)

We investigated the strength of each factor, which is characterized by $\Sigma_f$, in Fig. C.1. Notice that the strength decreases fast initially and becomes relatively slower starting from the fifth factor. Therefore, we pick the top 4 factors in our result section.

We also re-formulate the definition of information flow in the context of vector auto-regressive model. For the latent factor $f$ in V4 at time $t$, consider the full regression model using the full history of latent variables in both area, $Z_{1,t}^f \sim Z_{1,1:t-1}^f + Z_{1,1:t-1}^2$ vs the reduced model using history of latent variables in V4 only, $Z_{1,t}^f \sim Z_{1,1:t}^f$. The partial $R^2$ summarizes the contribution of PFC history to V4, thus can be viewed as information flow from V4 to PFC at time $t$. Dynamic information flow from V4 to PFC is defined similarly. The results are shown in Fig. C.2. Even from a different perspective, we reach to a similar conclusion as Fig. 4b.
Figure C.2: **Information flow by partial $R^2$ for the top four factors.** In this figure, we characterize dynamic information flow in terms of partial $R^2$. We show dynamic information flow from V4 $\rightarrow$ PFC (blue) and PFC $\rightarrow$ V4 (orange). In and out flows seem to peak at either the beginning or the end of the delay period, stronger V4 $\rightarrow$ PFC is identified, and different couplings of the two flows are also observed under this new definition. This figure echos with Fig. 4b.