We thank the reviewers for insightful comments and address the major issues they raised.

Reviewer 1 **Weaknesses:** (a) LDFA-H can be applied to spike count data in its current form, and we will change our discussion to clarify; we can also demonstrate its use in a spike count simulation. The misleading statement in our discussion should have said, more specifically, that we believe a better approach (for future work) would be to add an additional Poisson layer to the model, for reasons discussed in Vinci et al (2018) and related references. (b) The Kronecker product assumption does not force within-area interaction into the latent vector. Through simulations we have investigated spatio-temporal interactions under our Kronecker product assumption, and in our revision we will show how to handle them in the supplementary material.

Reviewer 2 **Weaknesses:** Although we have been unable to find a good way to make direct comparisons with other methods, the method here is reasonably fast and we will note that a single fit of the real data took less 19 seconds on a 2.2 GHz single-processor MacBook. **Correctness:** The definition of information flow in our real data section is intuitive, yet it is also subject to judgement. Because neural time series can be treated as vector auto-regressive, in the supplementary material we will provide partial χ^2 as another way to characterize information flow. In our revision, we also plan to add figures showing $\hat{\Pi}_f^{12}$ to further reveal the dynamic connectivity between two areas. Notice that $\hat{\Pi}_f^{12}$ is banded, and we are calculating $I_{f,out}(t)$ starting from $100ms$ (the bandwidth). The quantity $I_{f,out}(t)$ does not necessarily increase with t because it involves summation of non-zero entries only within the band. **Feedback:** We will add figures for PFC in the supplementary material. In addition, we will also add within-area connectivity for comparison with cross-area connectivity.

Reviewer 3 **Weaknesses:** (a) LDFA-H allows latent connectivity to be non-stationary, and thus, unlike GPFA, its structure can evolve across time; in this sense it is dynamic. It deals explicitly with the high dimensionality of non-stationary time series within areas by assuming a sparse, banded, Kronecker product structure. We will make this clearer. (b) The reviewer makes the important point that LDFA-H assumes the correlation across areas occurs in matched latent variables, so that differing components remain independent across areas, and the reviewer notes this may be restrictive. A key observation is that even in cases when this assumption is incorrect, the method can identify cross-area correlation, but, as we have now shown in a simple simulation, the situation is somewhat subtle. We simulated Z_1', Z_2' from Σ_1', Σ_2' according to our model with one area leading the other area for Z_1' and the reverse lead-lag relationship for Z_2'. When then combined the latents in the form $Z_1 = 0.6 * Z_1' + 0.4 * Z_2'$, $Z_2 = 0.4 * Z_1' + 0.6 * Z_2'$, and re-computed Σ_1 and Σ_2 for each factor. The corresponding Π_f^{12} indicates (intuitively) *both* lead-lag relationships. To demonstrate recovery we generated data from Z_1 and Z_2 and applied our method. It produced factors that separately pull out the two lead-lag relationships which, if combined, tell the ground-truth story about lead-lag in Z. The reviewer is thus correct that interpretation requires care, and we will say so in our revision, including the simulation results in the supplementary material. Furthermore, in future work we plan to generalize the approach to avoid this restrictive assumption, and will say so.

Reviewer 4 **Weaknesses:** (a) We should have included discussion of additional related work, including the articles the reviewer mentions. In our revision we will explain that none of those methods is directly comparable to ours. For example, we agree it could be beneficial to apply a spectral approach (though in sliding windows to deal with non-stationarity), but that is complementary to ours in the sense that it aims for connectivity among oscillatory components: our view is that pulses of activity often have temporal profiles that are better described in the time domain. (b) Concerning key assumptions, please see above. (c) In our revision we will articulate the main novel purpose and ability of our approach: again, it can find connectivity for pulses of activity that are not accurately described by stationary processes. (d) We will add details about the anonymous work. **Clarity:** We will introduce an example in section 2. **Feedback:** we will discuss the extension to multiple groups, which is straightforward.