We thank all reviewers for their time in reviewing our manuscript and their feedback on our work. We apologize for the various formatting issues in the references; these are now fixed, along with typos and other linguistics mishaps [R1-4]. If accepted, we will move the discussion concerning the Parra and Tobar (2017) paper in the main text [R1, R3], as well as the phase-shift interpretation of the Hilbert transform [R3].

Reviewer 1 — The authors do not provide any code for their GPFADS method. I presume that code will be made available upon acceptance.

Yes, code will be made available online upon paper release in the form of a python library which is under preparation.

Reviewer 2 — Additional discussion on where this method could fail or would not be a good method would have been useful.

Yes, we will add more discussion on the various theoretical limitations arising from the model, including the implications of the Gaussian process assumption (second-order non-reversibility as opposed to non-reversibility in higher-order moments; see also answer to Reviewer #3), and of the specific ways in which non-reversibility is introduced in the kernels. Additionally, we will discuss the limitations of the simple noise model we have worked with. For single-trial spiking data, for instance, we would expect that the model would work better if it included Poisson (as opposed to Gaussian) observations; this is next on our list of extensions.

Reviewer 3 — I found the notion of reversibility quite confusing. The paper defines it as "the probability of immediately returning to an initial state must be small", but this is not the standard definition e.g. https://en.wikipedia.org/wiki/Time_reversibility.

We will rewrite this part of the paper to improve on clarity. Our definition of reversibility indeed follows the definition based on detailed balance in the "Stochastic processes" section of the wikipedia page referenced by the Reviewer. We deem a process $x(t)$ reversible if for any pair of times t and s and any two vectors a and b,

$$p(x(t) = a, x(s) = b) = p(x(t) = b, x(s) = a). \tag{1}$$

If $x(t)$ is a zero-mean, stationary Gaussian process (as assumed in this paper), then it is entirely defined by its space-time covariance function, such that the detailed balance condition above becomes a time-reversal symmetry condition for the temporal cross-covariances. Specifically, a stationary GP is reversible if for any two time points t and s, the covariance matrix $\langle x(t)x(s)^T \rangle$ is symmetric.

The classical pendulum is reversible in the sense of a time reverse trajectory of a solution is a valid solution, and also all solutions are periodic and so indeed return to their initial state. What does 'immediately returning' mean?

Thanks, we will remove this confusing definition. Concerning the pendulum, the dynamics of the angle (as an observation) are indeed fully reversible. However, the dynamics of the system, considering its full state $(\theta, \dot{\theta})$, are highly non-reversible: oscillations in θ arise from near-circular state trajectories in the $(\theta, \dot{\theta})$ plane that evolve clockwise, but never counter-clockwise.

Reviewer 4 — Are you sure the expression of the non-reversibility index in Eq. 6 is correct ? [...] Expansion of Eq. B.(22) is not Eq. B.(23) [...] My intuition is [...], otherwise, as given it is zero if K is an odd function.

We have doubled checked, and Eq. 6 is indeed correct. The expression is simplified using the fact that (by stationarity) $K(-\tau) = K(\tau)^T$ – we will add this point to the paragraph preceding the equation. (Also, just to clarify, $K(\cdot)$ is a covariance function and can never be odd.)

I didn't find the proof of Eq. 8 in the supplementary

Thank you for pointing out this oversight, this will be added. In short, Eq. 7 is an orthogonal decomposition, such that the sum of squares in $K(\cdot)$ (as a matrix-valued function) is equal to the sum of squared weights in the decomposition (i.e. the sum of λ^2). Moreover, since the two sums in Eq. 7 separately decompose the numerator and denominator in Eq. 6 (uniqueness of the symmetric/skew-symmetric decomposition of a matrix-valued function), Eq. 8 follows.

Why do you call the decomposition in Eq. 7 "Kronecker", any reference?

Equation 7 defined the space-time covariance $K(\tau)$ as a function of the time-lag τ. Since each term in the sum is the product of a spatial component and a temporal component, any Gram matrix instantiating the kernel at a discrete set of time points is a sum of Kronecker products. We will explain the origin of this terminology in the main text.

Could you mention in the main text that the decomposition comes from a generalized SVD and give a reference for this mathematical result? Could you give a reference for the Heywood cases (l. 225)?

We will add references for these in the text. These will include: C. Van Loan, Journal of comp. and applied mathematics. (2000), Crane et al, SIAM J. Numer. Anal. (2020) and Martin, J.K. et al, Psychometrika 40, 505-517, (1975).