NETWORK INFERENCE VIA PROCESS MOTIFS FOR LAGGED CORRELATION IN LINEAR STOCHASTIC PROCESSES

Alice C. Schwarze, Sara M. Ichinaga, Bingni Brunton

A major challenge for causal inference from time-series data is the trade-off between computational feasibility and accuracy. Motivated by process motifs for lagged covariance in an autoregressive model with slow mean-reversion, we propose to infer networks of causal relations via pairwise edge measure (PEMs) that one can easily compute from lagged correlation matrices. We introduce a PEM with a correction for confounding factors (LCFC) and a PEM with a correction for reverse causation (LCRC).

We derive the respective correction terms from the contributions of process motifs (see Figure1 (a)) to a stochastic difference model

$$\mathbf{x}_t = (1-\Delta t)\,\mathbf{x}_{t-\Delta t} + \Delta t \sum_{k=1}^{p} \epsilon \mathbf{A}^{(k)} \mathbf{x}_{t-k\Delta t} + \sigma \Delta \mathbf{w}_t\,, \quad (1)$$

with time step Δt, coupling strength ϵ, lag-k adjacency matrix $\mathbf{A}^{(k)}$, noise strength σ and Gaussian white noise $\mathbf{w}_t$. This model interpolates between the discrete-time vector autoregression and the continuous-time Ornstein–Uhlenbeck process.

To demonstrate the performance of our PEMs, we consider linear stochastic processes on random networks and show that our proposed PEMs can infer networks accurate and efficiently. Specifically, our approach achieves higher accuracy than lag-1 correlation (LC), inverse covariance estimation (OUI), and a higher than or similar accuracy as results from Granger causality (GC), transfer entropy (TE), and convergent crossmapping (CM), but with much shorter computation time than any of these methods (see Figure1 (b–c)).

The theoretical framework underpinning our PEMs makes it possible to explain why some structural properties of networks (e.g. large network size and a large mean local anti-clustering coefficient[1]) pose challenges for PEM-based network inference (see Figure 1 (d)).

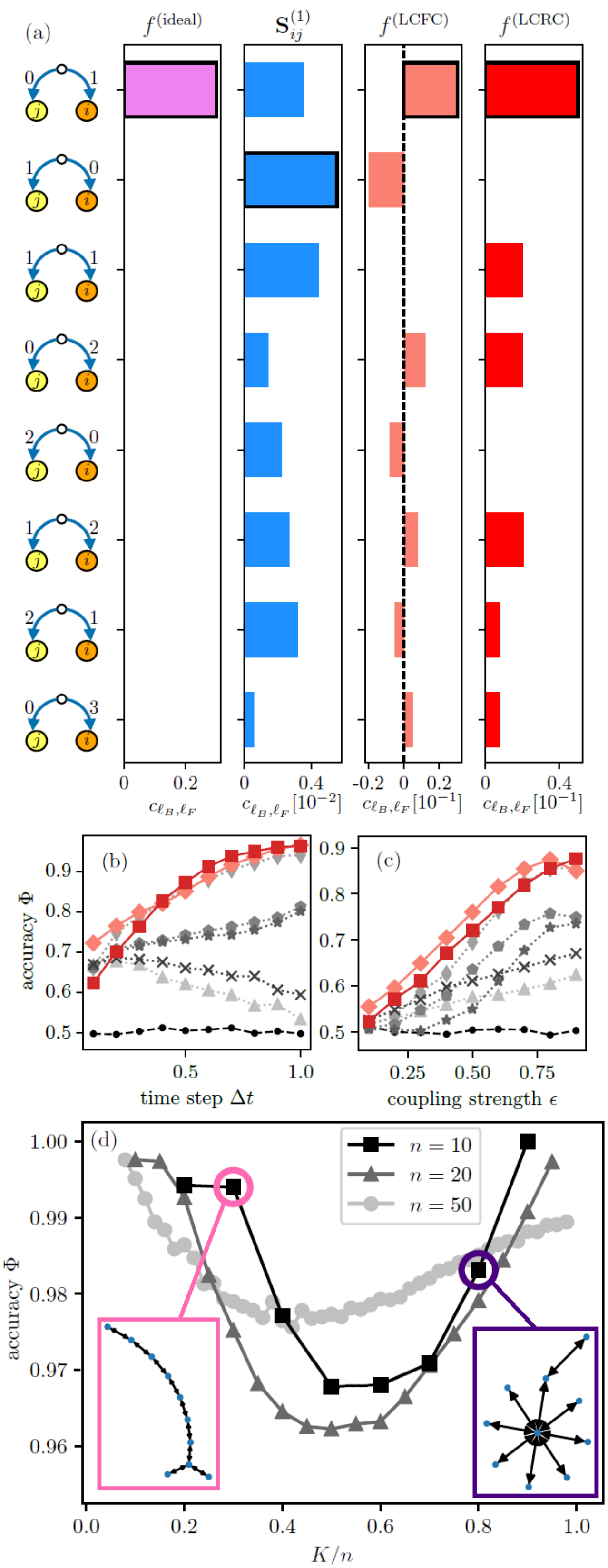

Figure 1: **Process motif contributions and inference accuracy.**. (a) Three process motifs and their contributions to an ideal PEM, lag-1 covariance $\mathbf{S}_{ij}^{(1)}$, and our proposed PEMs, $f^{(\text{LCFC})}$ and $f^{(\text{LCRC})}$; (b–c) inference accuracy with varying parameters of the stochastic-difference model; (d) inference accuracy with varying maximum degree K of a shooting-star graph with n nodes.

[1] We define the *local anti-clustering coefficient* of a node i with degree k_i and local clustering coefficient c_i as $\bar{c}_i := k_i(1-c_i)$