Characterizing spreading dynamics of subsampled systems with non-stationary external input

Jorge de Heuvel$^1$, Jens Wilting$^1$, Moritz Becker$^{1,2}$, Viola Priesemann$^1$ and Johannes Zierenberg$^1$

$^1$ Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany, and
$^2$ Department of Computational Neuroscience, Third Institute of Physics – Biophysics, Georg-August-University, Göttingen, Germany

(Dated: May 5, 2020)

Many systems with propagation dynamics, such as spike propagation in neural networks and spreading of infectious diseases, can be approximated by autoregressive models. The estimation of model parameters can be complicated by the experimental limitation that one observes only a fraction of the system (subsampling) and potentially time-dependent parameters, leading to incorrect estimates. We show analytically how to overcome the subsampling bias when estimating the propagation rate for systems with certain non-stationary external input. This approach is readily applicable to trial-based experimental setups and seasonal fluctuations, as demonstrated on spike recordings from monkey prefrontal cortex and spreading of norovirus and measles.

Propagation dynamics in complex networks are often approximated by models with an autoregressive representation. Examples include affinity maturation in immune systems\cite{1}, reproductive dynamics of bacteria\cite{2,3} or humans\cite{4}, epidemiological disease spreading in a network of humans\cite{5,6}, neutron transport theory\cite{7} and collective cortical dynamics\cite{8,9}. The inference of propagation dynamics is often complicated. First, only a fraction of all system components can be observed experimentally (subsampling)\cite{10,11}. Second, the model parameters can be time-dependent (non-stationary), and specific time-dependent input rates can lead to signatures of criticality even for networks of uncorrelated units\cite{12,13}. In general, time-dependent input rates are ubiquitous for collective dynamics in neural networks, and are one source for seasonal fluctuations of infectious disease incidence\cite{14}.

The subsampling challenge is typically addressed for stationary model parameters. Recent progress has been made for equilibrium and non-equilibrium systems by explicitly modelling the hidden units\cite{15,16}. However, explicit knowledge about the hidden units cannot be guaranteed for real-world applications. A subsampling-invariant approach that does not require knowledge about the underlying model size was recently proposed\cite{12}. The authors showed that established estimators based on linear regression or Kalman filtering underestimate the propagation behaviour. They introduced a novel multistep regression (MR) estimator that is subsampling invariant by characterizing propagation dynamics through the systems autocorrelation time $\tau$. However, it does not consider time-dependent model parameters.

To tackle non-stationarities, recent approaches considered models with time-dependent parameters. Examples include Bayesian models based on Cox-processes\cite{17}, weighted least-squares\cite{18}, or expectation-maximization based on Kalman filtering\cite{19,20}. However, none of these methods consider the complication of subsampling, although real spreading processes are usually subsampled\cite{21,22}.

In this Letter, we derive an estimator for a subsampled process subject to a specific type of non-stationary external input, namely cyclostationary input. We first show that the subsampling-invariant MR estimator\cite{12} can be biased if the external input rate changes over time. We then analytically derive a generalization of the MR estimator that can overcome the bias in the case of cyclostationary input. This approach is subsampling invariant and readily applicable to two prevalent situations: First, to trial-based experiments, which are commonly found in neuroscience; Second, to periodic input rates, e.g. the seasonal fluctuations of infectious disease incidence\cite{20}. We demonstrate the applicability of our methodology on numerical data (testing robustness to relaxation of our assumptions) and on real-world experimental data.

We consider the class of stochastic processes with an autoregressive representation of first order. This includes widely-used processes, such as branching processes, Kesten processes, and AR(1) processes. Time evolves in discrete steps $\Delta t = 1$. Let $A^i_t$ denote the activity of a realization $i$ at time $t$, then the conditional expectation value over the ensemble of independent realizations is defined as

$$\langle A^i_{t+1} | A^i_t \rangle = m A^i_t + \langle h_t \rangle,$$

where $m$ is the time-independent mean offspring parameter and $\langle h_t \rangle$ is the average ensemble rate of a time-dependent input distribution. In the framework of spike propagation in neural networks, $m$ describes the average number of neurons that a single neuron subsequently activates and $\langle h_t \rangle$ describes the expected input rate at time $t$ from sensory modalities or other brain areas.

Note that the expectation values in Eq. (1) are defined over the ensemble of independent realizations (trials) of the stochastic process, e.g., $\langle h_t \rangle = \sum_i h^i_t$ (for the trial-average we drop the index that was summed over). For a general non-stationary external input, $\langle h_t \rangle$ cannot be defined unless one has multiple realizations from the same
FIG. 1. Unbiased estimation of internal autocorrelation time $\tau$ for a subsampled system with time-dependent input rate can be achieved after subtracting trial-ensemble average from activity. (a) Step function as example for a time-dependent input rate. (b) Subsampled activity $a^s_i$ of a branching process with constant internal autocorrelation time $\tau$ and non-stationary input rate shows non-stationary behavior in regime II and III. Colored lines are individual trials, black solid line is trial-ensemble average $\bar{a}_t$ (over 200 trials). (c) Linear regression slope estimates $r_k$ (blue dots) of time lag $k$ for process in (b) do not decay exponentially as expected from the process’ internal autocorrelation (red line), which makes an unbiased estimation of $\tau$ impossible. (d) Time series (b) corrected by subtracting trial-ensemble average: $\tilde{a}^s_i = a^s_i - \bar{a}_t$. (e) For corrected time series (d), the $r_k$ decay exponentially with $\tau$, such that $\tau$ can now be inferred without bias. Simulation parameters: Trial length $T = 10000$ steps, internal autocorrelation time $\tau = 20$ steps, number of trials $N = 200$, mean (fully sampled) baseline activity $\langle A_0 \rangle = 1000$, subsampling fraction $\alpha = 0.05$, relative step height $\langle h_{up} \rangle / \langle h_{down} \rangle = 2.6$, step duration $c = 200$ steps.

Time-dependent distributions $h^s_i \sim P(h_i)$. In nature, this is approximately realized by cyclostationarity, e.g., trial-based experiments or seasonal fluctuations. We make use of this to solve the problem even without knowledge of the precise realization of external inputs. In the following, we assume that the generation of offsprings is Poisson distributed with time-independent $m$, while the generation of external input is Poisson distributed with time-dependent rate $\langle h_i \rangle$.

Subsampling is incorporated as follows: We only require that the subsampled activity $a^s_i$ is on average linear in the full activity $A^s_i$, i.e., $\langle a^s_i | A^s_i \rangle = \alpha A^s_i$ (for details see Ref. [12]). For example, every spike or disease incidence is sampled with probability $p = \alpha$.

To estimate the spreading behaviour $m$ under subsampling and time-dependent external input rates, we follow the principle idea of the MR estimator [12]. We generalize Eq. (1) by recursive iteration to $k$ time steps:

$$\langle A^s_{i+k} | A^s_i \rangle = m^k A^s_i + \sum_{l=1}^{k} m^{k-l} \langle h_{i+l-1} \rangle \quad (2)$$

If the rate is time-independent ($\langle h_i \rangle = \langle h \rangle$), Eq. (2) implies that the original process $A^s_i$ has an exponential autocorrelation function

$$C(k) = m^k = \exp(-k \Delta t / \tau), \quad (3)$$

with the time lag $k$ in steps of $\Delta t$. The autocorrelation function relates the propagation dynamics ($m$) to an internal autocorrelation time $\tau = -\Delta t / \ln(m)$ and represents a measure of how long information persists in the activity [12]. For stationary processes the variance across trials is equal to the variance within trials $\langle \text{Var}_t(a^s_i) \rangle = \langle \text{Var}_t(a^s_i) \rangle = \langle \text{Var}_t(a^s_i) \rangle = 0$. While $b$ is biased under subsampling ($b < 1$ if $\alpha < 1$), the autocorrelation time $\tau$ is subsampling invariant and can be obtained by fitting Eq. (4) to the data [12].

For a time-dependent external input rate $\langle h_i \rangle$, however, the autocorrelation function is not time invariant, and if calculated does not necessarily decay exponentially (Fig. 1b-c). Consider, for example, a step-function external input rate. Linear regression applied to each regime independently would yield similar slopes (identical slopes for full activity $A^s_i$) but different offsets of linear regression (Supplemental Material, Fig. S2). Therefore, the naive application of the MR estimator fails even for full activity. This represents an issue for general time-dependent input.

In the following, we construct a reliable estimate of the internal autocorrelation time $\tau$ in the presence of cyclostationary external input rates. We focus our discussion on subsampled activity $a^s_i$, which includes the fully-sampled case ($\alpha = 1, b_i = 1$).
To correct the bias from cyclostationary external input $(\langle h_i \rangle$ is time-dependent but identical for each trial $i$), we introduce the following method: Given we have $N$ trials, with independent realizations $a_i$ of a subsampled linear autoregressive process, we calculate the time-dependent trial-ensemble average

$$\bar{a}_t = \frac{1}{N} \sum_{i=1}^{N} a_i$$

over all trials (not to be confused with an average calculated over all recorded times). Now, we correct for the non-stationarity of the original process by subtracting the trial-ensemble average (Fig. 4)

$$\tilde{a}_t = a_t - \bar{a}_t.$$  \hspace{1cm} (6)

Its linear regression slopes $r_k$ reveal the true internal autocorrelation time in their exponential decay (Fig. 5) for sufficiently large $N$ (see below and Supplemental Material Fig. S4).

From the corrected time series $\tilde{a}_t$, we can thus infer the unbiased autocorrelation time by applying the MR estimator \cite{32} (see Supplemental Material S.6 for the full derivation). To prove this, we reformulate Eq. (1) as simple linear regression at each time across trials, i.e.,

$$\tilde{r}_{k,t} = \text{Cov}(\tilde{a}_t, \tilde{a}_{t+k}) / \text{Var}(\tilde{a}_t).$$

For trial-ensemble corrected $\tilde{a}_t$, we find that the correction compensates the convolution in Eq. (2), such that $\tilde{r}_{k,t} = b_k \delta^{k}$ with time-independent decay but with time-dependent autocorrelation strength (Supplemental Material, Eq. (S23))

$$b_k = \alpha^2 \frac{\text{Var}(\tilde{a}_t)}{\text{Var}(\tilde{a}_t^2)} \approx \frac{1}{1 - (1 - \alpha^{-1}) F^{-1}},$$

where the relation to the (across-trial) Fano factor of the full activity $F_t = \text{Var}(A_t)/\langle A_t \rangle$ is strictly true only for binomial subsampling. However, we can show (Supplemental Material, Eq. (S25)-(S29)) that for the corrected time series direct application of Eq. (4) with the standard regression approaches yields an unbiased estimate of the internal autocorrelation time $\tau$ despite cyclostationary input and subsampling (for a proof of concept see Fig. S3).

In addition to the bias from subsampling or non-stationary input, there can be a bias from short trial length $T$ \cite{33} and from small trial number $N$. The short-trial bias can be avoided by estimating both covariance and variance as fluctuations around a global stationary mean (cf. “stationarymean” method in Ref \cite{32} with a detailed discussion). For all our analyses (experimental and numerical), we thus use the MR estimator toolbox \cite{32} with “stationarymean” method. In principle, this allows for an unbiased estimation down to $N = 10$ short trials (Fig. S4), while of course the variance of the results increases with decreasing $N$ (Supplemental Material, Sec. S.4 and S.7).

We tested the applicability of MR estimation for cyclostationary external input by increasing the level of realism for a numerical problem. The test case is a baseline rate $\langle h_0 \rangle$ plus step-function at onset time $t_{\text{step}}$ with step height $\Delta \langle h \rangle$ and step duration $c$. We consider three cases: i) perfect cyclostationarity across trials (Fig. 1 and Fig. S5 for an extreme example), ii) variation of onset time $t_{\text{step}} \sim N(T/2, \sigma_t)$ with $\Delta \langle h \rangle = \langle h_0 \rangle$ fixed (Fig. 2), and iii) variation of the step height $\Delta \langle h \rangle \sim N(\langle h_0 \rangle, \sigma_h)$ with $t_{\text{step}} = T/2$ fixed (Fig. 2). We generated $N = 200$ trials of branching processes with internal autocorrelation time $\tau$, trial duration $T = 1000 \tau$, and baseline activity $\langle A_0 \rangle = 5000$ such that $\langle h_0 \rangle = (1 - m) \langle A_0 \rangle$ ($m = \exp(-\Delta t/\tau)$, $\Delta t = 1$ step). This setup allows us to independently investigate variability in onset time and height of the input.

Variations in the onset time and step height do not hinder correct inference as long as the standard deviations are sufficiently low (Fig. 7). In our test case, variations in the onset time rarely affect the correct inference as long as the standard deviation $\sigma_t$ is below the magnitude of the autocorrelation time (Fig. 6). When $\sigma_t = O(\tau)$, the method still provides consistent estimates of the processes autocorrelation time. Moreover, the estimates improve for a given $\sigma/\tau$ with increasing autocorrelation time $\tau$. We observe, that the $5\%$ error bound scales as $\sigma_t \propto \tau^\gamma$ with $\gamma \approx 0.22(3)$. Similarly, variations in the step height barely affect the correct inference as long as the standard deviation $\sigma_h$ is below $\langle h_0 \rangle/5$ (Fig. 7). Again, the estimates improve with increasing autocorrelation time and the $5\%$ error bound scales as $\sigma_h/\Delta \langle h \rangle \propto \tau^\gamma$ with $\gamma \approx 0.4(1)$. We conclude that our method provides consistent results even after relaxation of perfect cyclostationarity.

We applied our method to two sets of experimental data. The first dataset consists of spiking activity in

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2}
\caption{Robustness of our estimate to variability in a non-stationary input (step function with step-duration $c = 10 \tau$). (a) Variability in the onset time $t_{\text{step}}$ with standard deviation $\sigma_t$. (b) Variability in the step height $\Delta \langle h \rangle$ with standard deviation $\sigma_h$. Color in matrix indicates relative error between the estimated autocorrelation time $\tau$ and the internal autocorrelation time $\tau$ of the branching process. The $5\%$ error bound was fitted (black lines) and scales as $\sigma \propto \tau^\gamma$. Simulation parameters: $T = 1000 \tau$ steps, $N = 300$, $\langle A_0 \rangle = 5000$, $\alpha = 0.01$.}
\end{figure}
pre-frontal cortex from a trial based short-term visual memory task on *macaque mulatta* [34] (about $N = 300$ trials each, see Supplemental Material Sec. S.11). In this dataset, the external input can be interpreted as sensory input from other brain areas to the investigated area. The second dataset are epidemiological case reports from the Robert-Koch-Institute [35] ($N = 18$ trials each, see Supplemental Material Sec. S.10). In the epidemiological dataset, the infections carried into the country via travel can be interpreted as non-stationary external input.

For the monkey data, we want to emphasize three findings: First, although the trial-ensemble average $\tau_i$ increases by a factor 3 (Fig. 3) the autocorrelation function hardly differs in most cases (Fig. 3b). Second, we find a systematic decrease of intrinsic timescales after correction, while for the majority of the recording sets the decrease was less than 10% (Fig. 3c). Third, a robustness test of our method with parameters adjusted to experimental scale (Fig. 3d, with experimentally realistic stimulus shape) indicates that our method yields less than 5% deviation from $\tau \geq 200$ ms despite stimulus onset variability with $\sigma_t < 50$ ms, which is a realistic constraint given the steep rise of typical ensemble responses within 30 ms–50 ms (Fig. 3a). To conclude, our method reveals intrinsic timescales in pre-frontal cortex between 57(4) ms and 345(26) ms with median 214 ms (compared to 239 ms if not corrected) from recordings covering the full task. Our results are consistent with previous results in pre-frontal areas of macaque (about 200 ms) confined to the stimulus foreperiod to approximate the resting state [36, 37].

In the example of disease spreading, our method accounts well for seasonal fluctuations (Fig. 3e–g). The weekly case number reports reveal a strong yearly periodicity, suggesting a year-wise separation into trials. The improvement due to trial-ensemble average correction is readily visible in the regression function $r_k$ (Fig. 3f). With the correction, the infectiousness estimate is higher than without (Fig. 3g). Norovirus: $\tau = 14(3)$ weeks, Measle: $\tau = 15(8)$ weeks). The disease results are in principle subject to additional uncertainty from the small number of trials (cf. Fig. S4), which are probably on the order of 10% and thus smaller than the error bars from the fits. Our results highlight that the correction by trial-ensemble average can reveal higher infectiousness of diseases, which might otherwise be underestimated due to seasonal fluctuations and other non-stationary effects, and that long-term recordings are necessary to reveal the intrinsic infectiousness of a disease.

In summary, we have presented a simple, subsampling-invariant estimate of the internal autocorrelation time for stochastic processes with an autoregressive representation subject to (approximate) cyclostationary external input. The key success of the presented approach (MR estimation with trial-ensemble average corrected time series) is the potential to disentangle the internal spreading from...
any hidden, but repetitive external input rate. Thereby, our approach solves the problem of apparent criticality due to non-stationary input rates [19] for repetitive stimulation protocols. We demonstrated the robustness of our approach to violations of perfect cyclostationarity for the external input rate; and we showed its applicability to real-world problems from neuroscience and epidemiology. In conclusion, we recommend the trial-ensemble average correction as best practise when approximating trial-based experiments with autoregressive models. A toolbox for the multistep-regression analysis is readily available [32].

We thank Matthias Munk for sharing his data. All authors acknowledge support by the Max Planck Society. Financial support was received from the Gertrud-Reemtsma-Stiftung (JW), the Joachim Herz Stiftung (JZ), and the German Ministry of Education and Research (BMBF) via the Bernstein Center for Computational Neuroscience (BCCN) Göttingen under Grant No. 01GQ1005B (MB, VP, JZ).

* Authors contributed equally
† johannes.zierenberg@ds.mpg.de  Authors contributed equally

[1] A. Nourmohammad, J. Otwinowski, M. Luksza, T. Mora, and A. M. Walczak, Molecular Biology and Evolution 36, 2184 (2019).
[2] B. H. Good and O. Hallatschek, Current Opinion in Microbiology 45, 203 (2018).
[3] D. G. Kendall, Journal of the Royal Statistical Society. Series B (Methodological) 11, 230 (1949).
[4] M. A. Muñoz, Reviews of Modern Physics 90, 031001 (2018).
[5] Y. Kozlovsky, I. Cohen, I. Golding, and E. Ben-Jacob, Physical Review E 59, 7025 (1999).
[6] I. M. del Puerto, M. González, C. Gutiérrez, R. Martínez, C. Minuesa, M. Molina, M. Mota, and A. Ramos, eds., Branching Processes and Their Applications, Lecture Notes in Statistics - Proceedings (Springer International Publishing, 2016).
[7] C. P. Farrington, M. N. Kanaan, and N. J. Gay, Biostatistics 4, 279 (2003).
[8] O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz, Journal of Mathematical Biology 28, 365 (1990).
[9] A. Pazy and P. H. Rabinowitz, Archive for Rational Mechanics and Analysis 51, 153 (1973).
[10] J. M. Beggs and D. Plenz, Journal of Neuroscience 23, 11167 (2003).
[11] C. Haldeman and J. M. Beggs, Physical Review Letters 94, 058101 (2005).
[12] J. Wilting and V. Priesemann, Nature Communications 9, 2325 (2018).
[13] J. Zierenberg, J. Wilting, V. Priesemann, and A. Levina, Physical Review E 101, 022301 (2020).
[14] J. P. Neto, F. P. Spitzner, and V. Priesemann, bioRxiv , 759613 (2019).
[15] A. Hagemann, J. Wilting, B. Samimizad, F. Mormann, and V. Priesemann, arXiv:2004.10642 [physics, q-bio] (2020).
[16] V. Priesemann, M. H. Munk, and M. Wibral, BMC Neuroscience 10, 40 (2009).
[17] T. L. Ribeiro, M. Copelli, F. Caixeta, H. Belchior, D. R. Chialvo, M. A. L. Nicolelis, and S. Ribeiro, PLOS ONE 5, e14129 (2010).
[18] A. Levina and V. Priesemann, Nature Communications 8, 15140 (2017).
[19] V. Priesemann and O. Shriki, PLOS Computational Biology 14, e1006081 (2018).
[20] J. E. Franke and A.-A. Yakubu, SIAM Journal on Applied Mathematics 66, 1563 (2006).
[21] B. Bravi, M. Oppe, and P. Sollich, Physical Review E 95 (2017).
[22] B. Dunn and Y. Roudi. Physical Review E 87, 022127 (2013).
[23] B. Bravi and P. Sollich, Journal of Statistical Mechanics: Theory and Experiment 2017, 063404 (2017).
[24] L. Bachschmid-Romano and M. Opper, Journal of Statistical Mechanics: Theory and Experiment 2014, P06013 (2014).
[25] B. Dunn and C. Battistin, Journal of Physics A: Mathematical and Theoretical 50, 124002 (2017).
[26] A. Das and I. R. Fiete, bioRxiv , 512053 (2019).
[27] C. Donner and M. Oppe, Journal of Machine Learning Research 19, 34 (2018).
[28] T. S. Rao, Journal of the Royal Statistical Society: Series B (Methodological) 32, 312 (1970).
[29] Z. Ghalamrou, Parameter Estimation for Linear Dynamical Systems, Tech. Rep. (Toronto: Department of Computer Science, University of Toronto, 1996).
[30] R. H. Shumway and D. S. Stoffer, Journal of Time Series Analysis 3, 253 (1982).
[31] L. Papoz, B. Balkau, and J. Lellouch, International Journal of Epidemiology 25, 474 (1996).
[32] Toolbox for the Multistep Regression Estimator, https://github.com/Priesemann-Group/mrestimator (2018).
[33] F. H. C. Marriott and J. A. Pope, Biometrika 41, 390 (1954).
[34] G. Pipa, E. S. Staedtler, E. F. Rodriguez, J. A. Waltz, L. Muckli, W. Singer, R. Goebel, and M. H. J. Munk, Frontiers in Integrative Neuroscience 2019-01-07.
[35] SurvStat@RKI 2.0, https://survstat.rki.de/, accessed: 2019-01-07.
[36] R. T. L. Ribeiro, M. Copelli, F. Caixeta, H. Belchior, D. R. Chialvo, M. A. L. Nicolelis, and S. Ribeiro, PLOS Computational Biology 3, 10.3389/neuro.07.025.2009 (2009).
[37] R. T. L. Ribeiro, M. Copelli, F. Caixeta, H. Belchior, D. R. Chialvo, M. A. L. Nicolelis, and S. Ribeiro, PLOS Computational Biology 3, 10.3389/neuro.07.025.2009 (2009).
[38] R. T. L. Ribeiro, M. Copelli, F. Caixeta, H. Belchior, D. R. Chialvo, M. A. L. Nicolelis, and S. Ribeiro, PLOS Computational Biology 3, 10.3389/neuro.07.025.2009 (2009).
SUPPLEMENTAL MATERIAL

S.1 List of notations

| Notation | Description |
|----------|-------------|
| $A^i_t$  | Fully sampled activity of trial $i$ of the autoregressive process at time step $t$ |
| $a^i_t$  | Subsampled activity of trial $i$ of the autoregressive process at time step $t$ |
| $\langle A^i_t \rangle$ | Expectation value over independent realizations for a given time (dropped $i$) |
| $[A^i]$  | Expectation value over time for a given realization (dropped $t$) |
| $m$      | Branching/offspring parameter |
| $H_t$    | External input at time step $t$ |
| $\langle h_t \rangle$ | Mean external input rate at time step $t$ |
| $T$      | Total length of a time series |
| $N$      | Number of trials |
| $\tau$   | Autocorrelation time |
| $\Delta t$ | Time step length (absolute) |
| $k$      | Relative time lag |
| $b$      | Subsampling induced correlation bias |
| $\alpha$ | Subsampling fraction |
| $F$      | Fano factor |
| $\tilde{r}_k$ | Linear regression slope estimate for time lag $k$ |
| $\tilde{s}_k$ | Linear regression offset estimate for time lag $k$ |
| $\text{Var}(\cdot)$ | Variance over index $i$ |
| $\text{Cov}(\cdot)$ | Covariance over index $i$ |

TABLE S1. List of notations.

S.2 Branching process

The branching process (BP) with immigration is a stochastic autoregressive process. Each realization of the process $i$ is described by a temporal evolution with time $t$. For realization $i$ at time $t$ there are $A^i_t$ units, of which each unit $j$ generates a random integer number of “offsprings” $y_{i,t,j}$ and all $y_{i,t,j} \in \mathbb{N}$ are independently and identically distributed with the mean $m$. Additionally, a time-dependent external input $H^i_t$, with mean $\langle h_t \rangle$ “immigrates” at each time step, where $\langle \cdot \rangle$ denotes the expectation value over independent realizations. The evolution of the total activity $A^i_t$ of the branching process is recursively given by

$$A^i_{t+1} = \sum_{j=1}^{A^i_t} y_{i,t,j} + H^i_t. \quad (S1)$$

**Autocorrelation time** The branching parameter $m$ is directly connected to the processes autocorrelation time $\tau$ by [12]

$$m = \exp(-\Delta t/\tau), \quad (S2)$$

given a time binning $\Delta t$, e.g., from simulation steps or data binning in experiments.

**Stationary BP** Assuming the mean of the branching parameter $m$ and the external input $\langle h_t \rangle$ are constant over time, i.e., $\langle h_t \rangle = \langle h \rangle$, we can derive the dynamics of the branching process. First, the expectation value of the time step $A^i_{t+1}$ given the activity $A^i_t$ [cf. Eq.(1)] becomes

$$\langle A^i_{t+1} | A^i_t \rangle = m A^i_t + \langle h \rangle. \quad (S3)$$

Recursive iteration of Eq. (S3) and identification of the geometric series yields the expectation value of the evolution over $k$ time steps

$$\langle A^i_{t+k} | A^i_t \rangle = m^k A^i_t + \langle h \rangle \frac{1-m^k}{1-m}. \quad (S4)$$

Now, we can separate the dynamics into three regimes: Subcritical for $m < 1$, critical for $m = 1$ and supercritical for $m > 1$. We find a stationary solution $\langle A_\infty \rangle$ for the subcritical case by iterating Eq. (S4) to $k \to \infty$ ($m \to 0$), such that

$$\langle A_\infty \rangle = \frac{\langle h \rangle}{1-m}. \quad (S5)$$

In the critical state, the mean activity shows linear growth due to $\langle h \rangle$, whereas the activity diverges exponentially in the supercritical state.

**BP with non-stationary input** Since our main manuscript addresses non-stationary external input, we here investigate a non-stationary branching process with a time-dependent external input $\langle h_t \rangle$. The stationary distribution in Eq. (S5) is no longer valid and by iterating Eq. (S3) with the time-dependent external input rate $\langle h_t \rangle$ we can derive the conditional expectation value of the activity after $k$ time steps:

$$\langle A^i_{t+k} | A^i_t \rangle = m^k A^i_t + \sum_{l=1}^{k} m^{k-l} \langle h_{t+l-1} \rangle. \quad (S6)$$

S.3 Subsampling

When only a fraction of the full system can be observed, this is defined as subsampling. Examples include electrophysiological recordings of neuronal activity in neuroscience or incomplete case reporting of infectious disease propagation. Naive analysis of the data neglecting the influence of subsampling can lead to severe misinterpretations of the system’s dynamics [12] [16] [18].

The theory and implications of subsampling for linear autoregressive processes have been described in detail in Ref. [12] and will here be recapitulated briefly. The time series $a_t$ is called a subsample of $A^i_t$, if

$$\langle a^i_t | A^i_t \rangle = \alpha A^i_t + \beta \quad (S7)$$

holds for all $t, j \in \mathbb{N}$ with constants $\alpha, \beta \in \mathbb{R}$. The subsample $a^i_t$ is constructed from the fully sampled time
series upon sampling and does not interfere with its evolution. We assume $\beta = 0$.

### S.4 Multistep regression estimation

To infer a network’s autocorrelation time $\tau$ and the branching parameter $m$ even under subsampling, Wilting & Priesemann developed the multistep regression (MR) estimator \[ \text{S.4} \]. It addresses the issue of classical estimators being biased under subsampling. The MR estimator is applicable to stationary autoregressive processes of first order only, giving misestimations when applied to a non-stationary autoregressive time series. The MR estimator works as follows: In a first step, we calculate via Eq. (S2) the branching parameter $\tau$ and offset $s_k$ for time lags $0 < k < k_{max}$ for the time steps $t < T - k$, by minimizing the sum of residuals

$$ R_k(\hat{r}_k, \hat{s}_k) = \sum_{i=1}^{N} \sum_{k=1}^{T-k} (a_{i+k}^1 - (\hat{r}_k \cdot a_i^1 + \hat{s}_k))^2. \tag{S8} $$

It can be shown \[ \text{[12]} \], that for stationary dynamics $\hat{r}_k$ converges in probability to

$$ \hat{r}_k \rightarrow bm^k, \tag{S9} $$

where $m^k$ is the slope between the fully sampled activity pairs and $b$ the bias in the slope estimation due to subsampling. More specifically, the linear regression slopes fulfill the relation

$$ \hat{r}_k = \alpha_2 \frac{\text{Var}_t(A^1_t)}{\text{Var}_t(a^1_t)} m^k = bm^k, \tag{S10} $$

where the notation $\text{Var}_t(\cdot)$ denotes the variance over independent realizations. The bias depends on the subsampling fraction $\alpha$ (see Eq. \[ S.7 \]), the variance of the full activity $A^1_t$ and the subsampled activity $a^1_t$ respectively. However, these are usually unknown. Then, in the second step of the estimator, the sum of residuals is minimized for

$$ R(\hat{b}, \hat{m}) = \sum_{k} (\hat{r}_k - \hat{b} \cdot \hat{m}^k)^2, \tag{S11} $$

where the two step estimation over various time lags $k$ allows us to infer the bias $b$, which remains unknown to classical linear-regression estimators, and the branching parameter $m$. The autocorrelation time $\tau$ can easily be calculated via Eq. \[ S.2 \]. The procedure is equivalent to the calculation of time series autocorrelation that has a decreased correlation strength $b$ in step 1 and fitting the exponential decay in step 2.

We used for all analyses the python toolbox \texttt{Mr. Estimator} \[ S.2 \] of the multistep regression estimator. The exponential function $f_{\exp}(x, \tau, b) = b \exp(-x/\tau)$ has been used for the purpose of this investigation as it addresses the pure exponential decay characteristic for the autocorrelation function of an autoregressive process. For the monkey dataset, the offset-exponential fit-function $f_{\exp}(x, \tau, b, c) = b \exp(-x/\tau) + c$ has been used.

Two estimation methods are implemented in the \texttt{Mr. Estimator} toolbox. The method \texttt{stationarymean} uses all trials combined to calculate the activity average $\bar{\pi}$, which is needed to calculate the linear regression slopes $\hat{r}_k$ numerically. The advantage of the method \texttt{stationarymean} is that the linear regression estimation is more robust if only short trials with few datapoints each are available. The method \texttt{trialseparated} calculates the activity average $\bar{\pi}$ and subsequently linear regression slopes for each trial independently and averages over all obtained regression slopes $\hat{r}_{k,i}$, see Ref. \[ S.24 \] for further details. In case the mean activity between trials varies significantly, the method \texttt{trialseparated} provides better estimation results for the regression slopes. However, when each trial is short but activity across trials is stationary (or as in our case cyclostationary) the method \texttt{stationarymean} corrects for short-trial biases \[ S.22 \]. We validate the trial ensemble average correction on both methods in Sec. S.8. For all estimations in the paper, the method \texttt{stationarymean} was used to correct for short-trial biases.

### S.5 Effect of non-stationary input on MR estimation

Assuming a branching process that is subject to a time-dependent external input rate $\langle h_t \rangle$, a naive application of the MR estimator gives a biased estimation $\hat{\tau}$. We will demonstrate this analytically in the following example of a step function, which can be generalized to arbitrary time-dependent external input rates.

Let $\{A^1_t\}_{t=0}^{T}$ be subject to a time-dependent external input rate $\langle h_t \rangle$ with a step function:

$$ \langle h_t \rangle = \begin{cases} h_1 & \forall t < t_{step} \\ h_2 & \forall t \geq t_{step} \end{cases} \tag{S12} $$

Now, one can divide the mean activity development into three regimes, as shown in Fig. \[ S.1 \]. Two stationary regimes I and III with different expectation values and one transient regime II right after the jump in the external input rate. The expectation values for the regimes can be derived from Eqs. \[ S.5 \] and \[ S.4 \].

$$ (A_t) \approx \begin{cases} h_1 \frac{1}{1-m} & \forall t < t_{step} \\ m^{1-t_{step}} h_1 \frac{1}{1-m} + h_2 \frac{1-m^{1-t_{step}}}{1-m} & \forall t \geq t_{step} \end{cases} \tag{S13} $$
Here, case 1 \((t < t_{\text{step}})\) and 3 \((t \gg t_{\text{step}})\) follow immediately from Eq. \((S5)\), while for case 2 \((t \geq t_{\text{step}})\) we assume the stationary solution of regime I \(\langle h_{t_{\text{step}}} \rangle \) at \(t_{\text{step}}\), insert this into Eq. \((S4)\) with \(\langle h \rangle = h_2\), and identify \(k = t - t_{\text{step}}\).

By applying the estimator to the different regimes separately, following the steps in Ref. \([12]\), one finds that the linear regression offset estimator \(s_k\) will take on different values due to \(\langle h_t \rangle\) in the different regimes:

\[
\begin{align*}
\hat{s}_{k,1} &\rightarrow h_1 \frac{1 - m^k}{1 - m} \quad \text{in regime I}, \\
\hat{s}_{k,II} &\rightarrow h_2 \frac{1 - m^k}{1 - m} \quad \text{in regime II}, \\
\hat{s}_{k,III} &\rightarrow h_2 \frac{1 - m^k}{1 - m} \quad \text{in regime III}.
\end{align*}
\]

Here we can clearly see that the least square estimation with Eq. \((S8)\) on the entire time series is influenced by the step function in \(\langle h_t \rangle\). As actually two different offsets would be treated as values of unity. Consequently, the estimation of \(\hat{r}_k\) will be biased by the time-dependence in the external input rate. To visualize that analytical example, the linear regression for a given time lag \(k\) for the regimes I to III separately and combined is visualized in Fig. \((S2)\).

S.6 Analytic derivation that linear regression on trial-ensemble average corrected time series allows to infer spreading dynamics

This section addresses the trial-ensemble average correction of non-stationary autoregressive time series of a BP subject to a non-stationary input with step-function rate \[\text{Eq. (S12)}\] can be divided in three regimes. Mean baseline activity \(\langle A_0 \rangle\) in regime I, the transient regime II with growing activity and a new mean activity regime III, compare Eq. \((S13)\).

First order, to realize a correct estimation of the processes spreading dynamics in terms of the branching parameter \(m\) and autocorrelation time \(\tau\) respectively and proves the validity analytically. We discuss subsampled systems with \(\langle A_0 \rangle = 1000\) in regime I, subsampling fraction \(\alpha = 1\), relative step height \(\langle h_{\text{up}} \rangle / \langle h_{\text{down}} \rangle = 10\).

When the external input \(H_t\) is drawn from the same time-dependent probability distribution, we can define a trial ensemble average

\[
\bar{a}_t = \frac{1}{N} \sum_{i=1}^{N} a^i_t,
\]

that averages over all trials for each time step and where \(N\) is the number of trials. For \(N \to \infty\) the trial ensemble average converges in probability to the time-dependent expectation value, thus \(\bar{a}_t \to \bar{a}_t\) at a given time step, which we assume for the following analytical derivations.

We start by defining the mean-corrected time series

\[
\tilde{a}^i_t = a^i_t - \langle a_t \rangle
\]

such that \(\langle \tilde{a}^i_t \rangle = 0\). We recall that the actual time evolution takes place in the original process \(A^i_t\).

Next, we show that the slopes \(r_{k,t}\) from linear regressions over mean-corrected subsampled activities from an ensemble of trials at time \(t\) and time \(t + k\) can be decomposed into a time-dependent correlation bias \(b_k\) and a \(k\)-dependent decay \(m^k\). For each time \(t\) we can solve the simple linear regression, Eq. \((S8)\), with

\[
r_{k,t} = \frac{\text{Cov}(\tilde{a}^i_t, \tilde{a}^i_{t+k})}{\text{Var}(\tilde{a}^i_t)},
\]

FIG. S1. Mean activity \(\langle A_t \rangle\) (bottom) of a BP subject to a non-stationary input \(\langle h_t \rangle\) (top) with step-function rate \[\text{Eq. (S12)}\] can be divided in three regimes. Mean baseline activity \(\langle A_0 \rangle\) in regime I, the transient regime II with growing activity and a new mean activity regime III, compare Eq. \((S13)\).

FIG. S2. Least square estimation of \(\hat{r}_k\) \[\text{Eq. (S8)}\] for BP subject to a non-stationary input with step-function rate \[\text{Eq. (S12)}\]. The lower left cloud results from data pairs \((A^i_t, A^i_{t+k})\) for \(t < t_{\text{step}}\) (I), the upper right from those for \(t \gg t_{\text{step}}\) (III). The colored lines represent fits of regimes I and II+III (black lines, no bias in slope) or I+II+III (red line, strong bias in slope), cf. Eq. \((S14)\). Simulation parameters: Trial length \(T = 100\) \(\tau\) steps, internal autocorrelation time \(\tau = 50\) steps, number of trials \(N = 1\), mean baseline activity \(\langle A_0 \rangle = 1000\) in regime I, subsampling fraction \(\alpha = 1\), relative step height \(\langle h_{\text{up}} \rangle / \langle h_{\text{down}} \rangle = 10\).
where the covariance is given by

\[ \text{Cov}_t(\hat{a}_t, \hat{a}_{t+k}) = \langle \hat{a}_t \hat{a}_{t+k} \rangle - \langle \hat{a}_t \rangle \langle \hat{a}_{t+k} \rangle, \quad (S18) \]

and \( \langle \cdot \rangle \) denotes the ensemble expectation value (where by convention the index \( i \) is dropped). Per construction \( \langle \hat{a}_t \rangle = \langle \hat{a}_{t+k} \rangle = 0 \), cf. Eq. \( (S16) \). We thus only need to calculate

\[ \langle \hat{a}_t \hat{a}_{t+k} \rangle = \langle (a_t - \langle a_t \rangle)(a_{t+k} - \langle a_{t+k} \rangle) \rangle \]

\[ = \langle a_t a_{t+k} \rangle - \langle a_t \rangle \langle a_{t+k} \rangle \]

\[ = \alpha^2 \langle (A_t A_{t+k}) - \langle A_t \rangle \langle A_{t+k} \rangle \rangle \]

\[ = \alpha^2 (m^k \langle A_t^2 \rangle + \langle A_t \rangle \sum_{l=1}^{k} m^{k-l} \langle h_{t+l-1} \rangle - m^k \langle A_t \rangle^2 - \langle A_t \rangle \sum_{l=1}^{k} m^{k-l} \langle h_{t+l-1} \rangle) \]

\[ = \alpha^2 m^k \text{Var}_t(A_t), \quad (S22) \]

where we used Eq. \( (S4) \). We thus find

\[ r_{k,t} = \frac{\alpha^2 \text{Var}_t(A_t^2)}{\text{Var}_t(A_t^2)} m^k = b_t m^k, \quad (S23) \]

with a time-dependent amplitude (or bias) \( b_t = \alpha^2 \text{Var}_t(A_t^2)/\text{Var}_t(A_t^2) \) and a purely \( k \)-dependent decay \( m^k \).

The time-dependent amplitude can be related to the Fano-factor of the original process. To see this, we note that per construction the trial-ensemble expectation value \( \langle \hat{a}_t \rangle = 0 \), such that \( \text{Var}_t(\hat{a}_t^2) = \langle \hat{a}_t^2 \rangle = \text{Var}_t(A_t) \). When the subsampling procedure can be described by a binomial distribution, where \( \text{Var}_t(A_t^2) = \alpha(1 - \alpha) A_t^2 \) we obtain from the law of total variance, \( \text{Var}_t(A_t^2) = \langle \text{Var}_t(\hat{a}_t^2) \rangle + \langle \text{Var}_t(\hat{a}_t^2) \rangle = \alpha(1 - \alpha) A_t^2 + \alpha \text{Var}_t(A_t) \). With the Fano factor \( F_t = \text{Var}_t(A_t)/\langle A_t \rangle \), the amplitude thus becomes

\[ b_t = \frac{\alpha^2 \text{Var}_t(A_t^2)}{\text{Var}_t(A_t^2)} = \frac{1}{1 - (1 - \alpha^{-1})F_t^{-1}}. \quad (S24) \]

Finally, we show that the time-dependent amplitude \( b_t \) still allows the application of the linear regression estimator \( r_k \) to the mean-corrected subsampled process as found in Eq. \( (S8) \) and Eq. \( (S11) \) despite cyclostationary external input. For this, it is important to notice that the minimization in the simple linear regression step, Eq. \( (S8) \), is solved by

\[ r_k = \frac{\text{Cov}_t(\hat{a}_t, \hat{a}_{t+k})}{\text{Var}_t(A_t^2)}, \quad (S25) \]

where both covariance and variance here run over trial ensemble \( i \) as well as time \( t \).

For the stationary mean method of the MR estimator \( [32] \), Eq. \( (S25) \) translates to

\[ r_k = \frac{1}{\sum_i (\langle \hat{a}_t \rangle - \langle a \rangle) (\hat{a}_{t+k} - \langle a \rangle)} \frac{\sum_i (\langle a \rangle - \langle a \rangle)^2}{\sum_i (\langle a \rangle - \langle a \rangle)^2}, \quad (S26) \]

where \( \langle a \rangle = \frac{1}{T} \sum_i \hat{a}_i \) by construction. We thus find

\[ r_k = \frac{1}{\sum_i (\hat{a}_t \hat{a}_{t+k})} = \frac{1}{\sum_i \text{Var}_t(A_t)} m^k = b_m^k, \quad (S27) \]

where we used Eq. \( (S23) \) and observe that an effective amplitude \( b = \sum b_i \text{Var}_i(A_t)/\sum \text{Var}_i(A_t) \) remains \( k \)-independent. With a \( k \)-independent amplitude, the second (fitting) step of the MR estimator, Eq. \( (S11) \) becomes unbiased.

For the trial-separated method of the MR estimator \( [32] \), Eq. \( (S25) \) translates to

\[ r_k = \frac{1}{N} \sum_i \frac{[\langle \hat{a}_t \rangle - \langle a \rangle)(\hat{a}_t \hat{a}_{t+k}) - \langle \hat{a}_t \rangle)}{\langle \hat{a}_t \rangle - \langle \hat{a}_t \rangle} \sum_i \frac{\text{Var}_t(A_t)}{\langle \hat{a}_t \rangle}, \quad (S28) \]

where \( [\cdot] \) denotes the time average and \( \langle \hat{a}_t \rangle = \frac{1}{T} \sum_i \hat{a}_i \approx 0 \). Because the trials are independent but identically distributed, we can assume that \( \text{Var}_t(\hat{a}_t) = \langle \hat{a}_t^2 \rangle \) is constant across trials, take it out of the sum, rearrange the double sum as in Eq. \( (S26) \), and find

\[ r_k = \frac{1}{\sum_i (\hat{a}_t \hat{a}_{t+k})} \frac{\sum_i b_i \text{Var}_t(A_t) m^k = b_m^k, \quad (S29) \]

where we used Eq. \( (S23) \) and observe that an effective amplitude \( b = \sum b_i \text{Var}_i(A_t)/T \text{Var}_i(A_t) \) remains \( k \)-independent. With a \( k \)-independent amplitude, the second (fitting) step of the MR estimator, Eq. \( (S11) \) becomes unbiased.

To summarize, we showed that the corrected time series \( \hat{a}_t \) enables the application of the MR estimator \( [12] [32] \) for an unbiased estimation of the internal dynamics \( m \) or equivalently \( \tau \) from subsampled data despite a time-dependent cyclostationary external input rate.

### S.7 Proof of concept

To verify the trial-ensemble average correction methodology, a numerical test was performed, see Fig. \( S3 \). More specifically, branching process trials with a time-dependent external input (step function) were simulated for various autocorrelation times \( \tau \) and subsampling fractions \( \alpha \). Both estimation methods stationary mean and trial-separated of the MR estimator toolbox were tested, see Sec. 5.4. For stationary mean (Fig. \( S3 \) a-c), the length of each trial was chosen as \( T = 100 \tau \), where \( N = 5000 \).
trials. For trialseparated (Fig. S3 d-f), the length of each trial was chosen as \( T = 5000 \tau \) (to avoid short-trial biases, cf. Sec.S4 and Refs [32, 33]), where \( N = 200 \) trials were simulated. In both cases, the relative height of the step function is \( \langle h_{\text{up}} \rangle / \langle h_{\text{down}} \rangle = 2 \). In Fig. S3 we plot independent estimates for each \( \tau \) (after trial-ensemble average correction) together with the fit error from the MR estimation toolbox [32]. One can clearly see that the results differ less than 1\% from the true \( \tau \), and moreover that about 2/3 of the results are correct within bootstrap errorbars as they should be. This demonstrates the general applicability of our method despite subsampling.

In addition, we checked the applicability of our methodology for small trial numbers \( N \) (Fig. S4). Again, we simulated branching processes with different autocorrelation times \( \tau \) and a time-dependent external input (step function), but now we fixed the trial length \( T = 100 \tau \) and the subsampling fraction \( \alpha = 0.01 \) and varied the number of trials \( N \). For each data point in Fig. S4 we generated 100 simulations with independent estimates of \( \tau \), and we plotted the mean and standard deviation (as errorbars). For stationarymean (Fig. S4 a-c), the standard deviation decreases with increasing trial number as expected (it starts with 10\% for \( N = 10 \), which would still enable one to quantify the order of magnitude, but falls to about 2\% for \( N \approx 200 \)), while the mean always coincides with the true \( \tau \). For trialseparated (Fig. S4 d-f), the variance also decreases for increasing \( N \), but the mean no longer coincides with the true \( \tau \). This is due to the before mentioned short-trial bias [32, 33]. Given the short trial length of \( T = 100 \tau \). This shows that for cyclostationary external input, the best choice of methods from the two above is the stationarymean.

S.8 Effect of time-dependent autocorrelation strength \( b_t \)

As we derived in Eq. (S24), regimes with rapid changes in the input rate leads to a variation of the Fano factor \( F_t \) and under subsampling subsequently the amplitude \( h_t \). More specifically, a rapid increase in \( \langle h_t \rangle \) decreases \( F_t \) and a rapid decrease in \( \langle h_t \rangle \) increases \( F_t \). These regimes of rapid change will be called transient regimes.

To test the influence of strong transients on the estimation of \( \hat{r}_k \) and subsequently \( \tau \), a branching process of extreme transients was simulated and analyzed with the MR estimator (Fig. S5). More specifically, a periodically recurring jump in the external input rate \( \langle h_t \rangle \) was implemented on a branching process with \( \tau = 200 \) steps (Fig. S5 a-f). The period of the external input was chosen as \( T_p = 10 \tau \) so that the up and down transients would cover five autocorrelation times each. This way approxi-
FIG. S4. Proof of concept B. Numerical validation of the trial-ensemble average correction for subsampled BP with time-dependent external input rate (step function as in Fig. S1) for various numbers of available trials \( N \) and autocorrelation times \( \tau \). Error bars denote standard deviation from 100 simulations with independent estimates via MR toolbox [32]. Simulation parameters (a-e): \( T = 100 \), \( \alpha = 0.01 \), \( \langle A_0 \rangle = 5000 \), \( \langle h_{\text{up}} \rangle / \langle h_{\text{down}} \rangle = 2 \). (a-c) Method stationarymean. The estimates lie in the range of 2% relative deviation (grey shading) for all simulated autocorrelation times \( \tau \) and subsampling fractions \( \alpha \). (d-f) Method trialseparated. The estimates suffer as expected from a short-trial bias [32, 33].

S.9 Numerical data

All branching process simulations were performed with C++, where the random number generator MT19937 was used to drive two Poisson distributions for the branching processes recurrent internal activation and the external input. Reproducible seeding was utilized. Subsampling from the full activity was implemented numerically by drawing from a Binomial distribution.

S.10 Epidemiological recordings

Case report data for measles and norovirus infections in Germany were obtained from the Robert-Koch-Institute [35]. Strong seasonal fluctuation and presumably non-stationarities with a period of 52 weeks motivated the investigation of the epidemiological case reports, preprocessed with the trial-ensemble average correction and the MR estimator. The case numbers were available with a weekly binning for 52+1 weeks per year from 2001 to 2018. Week 53 was omitted due to overlapping and only full years were used, thus ignoring the first recording year. The data was separated into trials representing one year each - in agreement with the 52 week periodicity of the fluctuations.

S.11 Spike data of *macaque mulatta*

The monkey experiments were performed according to the German Law for the Protection of Experimental Animals, and were approved by the Regierungspräsidium Darmstadt. The procedures also conformed to the regulations issued by the NIH and the Society for Neuroscience.

Spike data from electrophysiological recordings in the brain of *macaque mulatta* monkeys has been analyzed.
FIG. S5. Autocorrelation time \( \tau \) can be reliably inferred even for BP with strong periodically-alternating input rate (a-f) and periodically-increasing input rate (g-i). (a) Extreme example of a periodically-alternating input rate \( \langle h_t \rangle \). (b) Response of BP to external input with time-dependent rate (a) and internal autocorrelation time \( \tau = 200 \) steps, simulation length \( T = 1000 \tau \), subsampling fraction \( \alpha = 0.01 \), number of trials \( N = 200 \), relative step height \( \langle h_{\text{up}} \rangle / \langle h_{\text{down}} \rangle = 5 \) and period of input rate \( T_P = 10 \tau \). Colored lines are individual realizations, black line is trial-ensemble average. (c) Autocorrelation function of trial-ensemble average corrected time series, from which MR estimation yields correct internal autocorrelation time within errors. (d) Time-dependent Fano factor and time average (dashed line). (e) Time-dependent autocorrelation strength calculated from (d) via Eq. (7). (f) Time-dependent leading regression slope from (e) and \( m = e^{-1/\tau} \approx 0.995 \) together with time average (dashed line) shows that reliable estimates of the autocorrelation function can be obtained by averaging \( r_k \) over time. (g) Extreme example of a periodically-increasing input rate \( \langle h_t \rangle \). (h) Response of BP to external input with time-dependent rate (g) and internal autocorrelation time \( \tau = 200 \) steps, simulation length \( T = 1000 \tau \), subsampling fraction \( \alpha = 0.01 \), number of trials \( N = 200 \), step height \( \Delta \langle h \rangle = 100 \) and period of input rate \( T_P = 2.5 \tau \). Colored lines are individual realizations, black line is trial-ensemble average. (i) Time-dependent Fano factor and time average (dashed line) of (h). The autocorrelation function is similar to (c).

The dataset originates from a visual short-term memory experiment by Pipa et al. [34]. The monkeys were presented visual sample stimuli, which they had to remember for 3s. Afterwards test stimuli were shown, which the monkeys had to classify into matching and non-matching.

16 single-ended micro-electrodes and tetrodes in a 4x4 grid were placed in the lateral prefrontal cortex of the trained monkeys. The inter-electrode spacing was between 0.5 and 1 mm. The setup allowed a simultaneous activity recording of single units and field potentials at 1kHz, which was digitized and processed so that signal artefacts from licking and movement were rejected. The tetrode recordings were spike-sorted with the Spyke Viewer software, whereas micro-electrode data was processed with the Smart Spike Sorter by Nan-Hui Chen.

To convert the data into activity data, all simultaneous recordings were collapsed into one collective spike count and binned into \( \Delta t = 4\text{ms} \) time steps [12]. The 4ms binning represents the timescale in which spikes propagate from one neuron to the other, motivated by the autoregressive process as a model for neural activity propagation. The trials within the sets had slightly varying length, so they were cut off at the end to share the length of the shortest trial within the set. No temporal alignment was performed. For the MR estimation the maximum time lag was chosen as half the minimum trial length thus \( k_{\text{max}} = T_{\text{min}}/2 \). This way enough data is available for each regression step \( k \). For fitting, the offset-fit-function was used from the Mr. Estimator toolbox that is \( f_{\text{exp}}(x, \tau, b, c) = b \exp(-x/\tau) + c \), see Sec. S.4.

The maximum number of trials analyzed for each set was 300, where two sets had contained less than 300 trials, see Tab. [S2]
| Recording set | $\hat{\tau}$ in [ms] | # of trials |
|---------------|---------------------|-------------|
| C001          | 88 ± 2              | 300         |
| C002          | 130 ± 5             | 300         |
| C012          | 345 ± 26            | 300         |
| L001          | 175 ± 11            | 300         |
| L008b         | 239 ± 11            | 300         |
| L011b         | 232 ± 11            | 300         |
| L012b         | 270 ± 42            | 282         |
| L014b         | 205 ± 81            | 299         |
| 5115          | 214 ± 18            | 300         |
| 5117          | 57 ± 4              | 300         |
| 5144b         | 344 ± 105           | 300         |

TABLE S2. Detailed results of the spike data analysis of *macaque mulatta*. The estimated $\hat{\tau}$ are obtained with the novel trial-ensemble average correction.