Collective Spiking Dynamics in Cortical Networks

Dissertation
zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades
"Doctor rerum naturalium"
der Georg-August-Universität Göttingen

im Promotionsstudiengang Physik
der Georg-August University School of Science (GAUSS)
vorgelegt von

Jens Wilting

aus Köln

Göttingen, 2020

Betreuungsausschuss

Dr. Viola Priesemann, Max-Planck-Instiut für Dynamik und Selbstorganisation, Göttingen.
Prof. Dr. Florentin Wörgötter, 3. Physikalisches Institut, Georg-August-Universität Göttingen.
Prof. Dr. Hansjörg Scherberger, Deutsches Primatenzentrum, Göttingen.

Mitglieder der Prüfungskommission

Referentin: Dr. Viola Priesemann, Max-Planck-Instiut für Dynamik und Selbstorganisation, Göttingen.
Koreferent: Prof. Dr. Florentin Wörgötter, 3. Physikalisches Institut, Georg-August-Universität Göttingen.

Weitere Mitglieder der Prüfungskommission

Prof. Dr. Hansjörg Scherberger, Deutsches Primatenzentrum, Göttingen.
Prof. Dr. Theo Geisel, Max-Planck-Instiut für Dynamik und Selbstorganisation, Göttingen.
Prof. Dr. Stefan Klumpp, Institut für Dynamik komplexer Systeme, Georg-August-Universität Göttingen.
Prof. Dr. Peter Sollich, Institut für Theoretische Physik, Georg-August-Universität Göttingen.

Tag der mündlichen Prüfung: 24. September 2020.

Contents

Preface

This dissertation is the result of research at the Max-Planck-Institut for Dynamics and Self-Organization. The presented monograph majorily is a compilation of four publications:

1. J. Wilting & V. Priesemann. *Inferring collective dynamical states from widely unobserved systems.* Nature Communications **9**(1):2325 (2018), doi:10.1038/s41467-018-04725-4.

2. J. Wilting & V. Priesemann: Between perfectly critical and fully irregular: a reverberating model captures and predicts cortical spike propagation. *Cerebral Cortex* **29**(6):2759–2770 (2019), doi:10.1093/cercor/bhz049.

3. J. Zierenberg*, J. Wilting* & V. Priesemann: Homeostatic plasticity and external input shape neural network dynamics. *Physical Review X* **3**:031018 (2019), doi:10.1103/PhysRevX.8.031018.
   *These authors contributed equally.

4. J. Wilting, J. Dehning, J. Pinheiro Neto, L. Rudelt, M. Wibral, J. Zierenberg & V. Priesemann: Operating in a reverberating regime enables rapid tuning of network states to task requirements. *Frontiers in Systems Neuroscience*, **12**(November):55 (2018), doi:10.3389/fnsys.2018.00055.

On each of these publications, I am either first author or shared first author. Each of the publications has been published with open access and under the terms of a Creative Commons License (http://creativecommons.org/licenses/by/4.0/). The online supplementary material of publications 1 and 2 is presented in the appendix in this monograph. Part of the supplementary material of the first publication was previously included in my Bachelor Thesis in mathematics (Jens Wilting: Estimation of Branching Process Parameters. Bachelorarbeit in Mathematik, Georg-August-Universität Göttingen, 2015). The appendix of publication 3 was originally included within one manuscript. For consistency throughout this monograph, it has been separated from the main body and moved to a separate chapter in the appendix.

Acknowledgments

First of all, I want to thank Viola Priesemann. Her outstanding support for me, her trust in me, and her patience with me made my dissertation the success it was, and brought me where I am today. Even though my way may lead me out of science, I hope that our ways will not part. Of all members in Viola's group, I have to single out Johannes Zierenberg. Not only did we collaborate on most of our projects and share the authorship of many (upcoming) manuscripts, but he also grew a dear friend of mine. I hope that this friendship will extend throughout the different paths we will take in the future. I also want to thank Joao Pinheiro Neto, Lucas Rudelt, and Jonas Dehning. As group members of the very first hour, we were part of the amazing journey Viola's group took in the last few years. We shared many times a poster and even more times a beer at various conferences and retreats. These memories – from Mandarfen to Berlin to Lisbon – will not be forgotten. Last but not least, I want to thank Paul Spitzner, Jorge de Heuvel, Annika Hagemann, and Leonhard Leppin for a very successful and fun concluding year in Viola's group.

Beyond Viola's group, I want to thank all members of the Max-Planck-Institute for Dynamics and Self-Organization, who made may PhD studies a special time and enlightened me way beyond science. In particular, I want to mention Joscha Liedtke, Manuel Schottdorf, Rainer Engelken, and Frank Stollberg, who shared most of my time at the institute. As a matter of fact, I am colleagues with Joscha again at Robert Bosch GmbH and want to thank him for bringing this path to my attention. I want to give a special thanks to Theo Geisel. He not only accepted me at the institute — I am still convinced that this was mostly because he was looking for a skiing buddy — but supported and fostered me whenever necessary.

More than anyone else, I want to thank my whole family, whose backing I can always count on. My parents have never spared a pain to support me in every way imaginable. This support helped me to pursue my interests, foster my strengths, and made me the man I am. Likewise, my grandparents have gone through every length to help me achieve my dreams. They have always been an incredible source of support and motivation. Again, thank you to my family, for every opportunity you gave me and every door you opened.

Finally, it is my special desire to thank my partner Cathrin. Together, we have gone through some of the most challenging times of our lives, be it stressful examinations in medical studies, endless peer reviews, or the birth of our wonderful daughters Emilia and Sophie. Having the three of you by my side makes every challenge an adventure, every change a new chance, and every day a delight. I am looking forward to everything to come for us.

Chapter 1

Introduction

Cortical networks are confronted with ever-changing conditions, whether these are imposed on them by a natural environment, or induced by the actions of the subjects themselves. For example, when a predator is lurking for a prey it should detect the smallest movement in the bushes anywhere in the visual field, but as soon as the prey is in full view and the predator moves to strike, visual attention should focus on the prey. Optimal adaptation for these changing tasks requires a precise and flexible adjustment of input amplification and other properties within the local, specialized circuits of primary visual cortex: strong amplification of small input while lurking, but quenching of any irrelevant input when chasing. These are changes from one task to another. However, even the processing within a single task may require the joint contributions of networks with diverse computational properties. For example, listening to spoken language involves the integration of phonemes at the timescale of milliseconds to words and whole sentences lasting for seconds. A similar integration of information over multiple timescales is required for visual stimuli. Such temporal integration might be realized by a hierarchy of temporal receptive fields, a prime example of adaption to different processing requirements of each brain area (Murray *et al.*, 2014; Hasson *et al.*, 2015).

Basic network properties like sensitivity, amplification, and integration timescale optimize different aspects of computation, and hence a generic input-output relation can be used to infer signatures of the computational properties, and changes thereof (Kubo, 1957; Wilting and Priesemann, 2018a). Throughout this manuscript, we refer to *computation* capability

in the following two, high-level senses. First, the integration timescale determines the capability to process sequential stimuli. If small inputs are quenched away rapidly, the network may quickly be ready to process the next input. In contrast, networks that maintain input for long timescales may be slow at responding to novel input, but instead they can integrate information and input over extended time periods (Boedecker *et al.*, 2012; Del Papa *et al.*, 2017; Lazar, 2009; Bertschinger and Natschläger, 2004). This is at the heart of reservoir computing in echo state networks or liquid state machines (Buonomano and Merzenich, 1995; Maass *et al.*, 2002; Jaeger and Haas, 2004; Schiller and Steil, 2005; Jaeger *et al.*, 2007; Boedecker *et al.*, 2012). Second, the detection of small stimuli relies on a sufficient amplification (Douglas *et al.*, 1995). However, increased sensitivity to weak stimuli can lead to increased trial-to-trial variability (Gollo, 2017).

These examples show that local networks that are tuned to one task may perform worse at a different one, and there is no one-type-fits-all network for all environmental and computational demands. How does a neural network manage to both react quickly to new inputs when needed, but also maintain memory of the recent input, e.g. when a human listens to language? Did the brain evolve a large set of specialized circuits, or did it develop a manner to fine-tune its circuits quickly to the computational needs?

In order to understand how each cortical circuit or network processes its input, it would be desirable to first know its basic dynamical properties. For example, knowing which impact one additional spike has on the network (London *et al.*, 2010) would give insight into the amplifica-

tion of small stimuli (Douglas *et al.*, 1995; Suarez *et al.*, 1995; Miller, 2016). Knowing how much of cortical activity can be attributed to external activation or internal activation (Reinhold *et al.*, 2015) would allow to gauge how much of cortical activity is actually induced by stimuli, or rather internally generated, for example in the context of predictive coding (Rao and Ballard, 1999; Clark, 2013). Knowing the intrinsic network timescale (Murray *et al.*, 2014) would inform how long stimuli are maintained in the activity and can be read out for short term memory (Buonomano and Merzenich, 1995; Wang, 2002; Jaeger *et al.*, 2007; Lim and Goldman, 2013). However, not even these basic properties of cortical network dynamics are generally known with certainty.

To describe network responses, two contradicting hypotheses have competed for more than a decade, and are the subjects of ongoing scientific debate: One hypothesis suggests that collective dynamics are "asynchronous-irregular" (AI) (Burns and Webb, 1976; Softky and Koch, 1993; Stein *et al.*, 2005), i.e. neurons spike independently of each other and in a Poisson manner, which may reflect a balanced state (van Vreeswijk and Sompolinsky, 1996; Brunel, 2000). The other hypothesis proposes that neuronal networks operate at criticality (Beggs and Plenz, 2003; Levina *et al.*, 2007, 2009b; Muñoz, 2018; Beggs and Timme, 2012; Plenz and Niebur, 2014; Tkačik *et al.*, 2015; Humplik and Tkačik, 2017). Criticality is a particular state at a phase transition, characterized by high sensitivity and longrange correlations in space and time.

These hypotheses have distinct implications for the coding strategy of the brain. The typical balanced state minimizes redundancy (Barlow, 2012; Atick, 1992; Bell and Sejnowski, 1997; van Hateren and van der Schaaf, 1998; Hyvärinen and Oja, 2000), supports fast network responses (van Vreeswijk and Sompolinsky, 1996), and shows vanishing autocorrelation time or network timescale. In contrast, criticality in models optimizes performance in tasks that profit from extended reverberations of activity in the network (Bertschinger and Natschläger, 2004; Haldeman and Beggs, 2005; Kinouchi and

Copelli, 2006; Wang *et al.*, 2011; Boedecker *et al.*, 2012; Shew and Plenz, 2013; Del Papa *et al.*, 2017).

Surprisingly, there is experimental evidence for both AI and critical states in cortical networks, although both states are clearly distinct. Evidence for the AI state is based on characteristics of single neuron spiking, resembling a Poisson process, i.e. exponential inter spike interval (ISI) distributions and a Fano factor F close to unity (Burns and Webb, 1976; Tolhurst *et al.*, 1981; Vogels *et al.*, 1989; Softky and Koch, 1993; Gur *et al.*, 1997; de Ruyter van Steveninck *et al.*, 1997; Kara *et al.*, 2000; Carandini, 2004). Moreover, spike count cross-correlations (Ecker *et al.*, 2010b; Cohen and Kohn, 2011) are small. In contrast, evidence for criticality was typically obtained from a population perspective instead, and assessed neuronal avalanches, i.e. spatiotemporal clusters of activity (Beggs and Plenz, 2003; Pasquale *et al.*, 2008; Priesemann *et al.*, 2009; Friedman *et al.*, 2012; Tagliazucchi *et al.*, 2012; Shriki *et al.*, 2013), whose sizes are expected to be power-law distributed if networks are critical (Bak *et al.*, 1987). Deviations from powerlaws, typically observed for spiking activity in awake animals (Bédard *et al.*, 2006; Hahn *et al.*, 2010; Ribeiro *et al.*, 2010; Priesemann *et al.*, 2014), were attributed to subsampling effects (Priesemann *et al.*, 2009; Ribeiro *et al.*, 2010; Priesemann *et al.*, 2013; Girardi-Schappo *et al.*, 2013; Priesemann *et al.*, 2014; Ribeiro *et al.*, 2014; Levina and Priesemann, 2017). Hence, different analysis approaches provided evidence for one or the other hypothesis about cortical dynamics.

Evidence supports even more diversity of dynamical states across studies *in vivo* and *in vitro*. Collective spiking activity clearly differs between *in vitro* cultures and *in vivo* cortical networks. Cultures *in vitro* typically exhibit stretches of very little spiking activity, interrupted by strong bursts of highly synchronized or coherent activity (Robinson *et al.*, 1993; Van Pelt *et al.*, 2004; Chiappalone *et al.*, 2006; Wagenaar *et al.*, 2006; Orlandi *et al.*, 2013; Vardi *et al.*, 2016; Beggs and Plenz, 2003). In contrast, spiking activity recorded from cortex in awake animals *in vivo* lacks such pauses, and instead shows continuous, fluctuating activity. These fluctuations

show a dominant autocorrelation time that was proposed to increase hierarchically across cerebral cortex, from sensory to frontal areas (Murray *et al.*, 2014). Moreover, depending on experimental details such as brain area, species and vigilance state, one also observes evidence for asynchronous-irregular (AI) dynamics (Burns and Webb, 1976; Softky and Koch, 1993), oscillations (Gray *et al.*, 1989; Gray, 1994; Buzsaki, 2004), or strong fluctuations associated with criticality, bistability or up-and-down states (Breakspear, 2017; Priesemann *et al.*, 2009, 2013; Bellay *et al.*, 2015; Wilson, 2008; Stern *et al.*, 1997; Cossart *et al.*, 2003). These states differ not only in strength and structure of fluctuations, but also in synchrony among neurons, from uncorrelated to fully synchronized spiking. The observation of such a vast range of dynamic states is puzzling, considering that the dynamics of all networks presumably originate from similar single-neuron physiology and plasticity mechanisms.

In the past, insights about network properties have been strongly hampered by the inevitable limitations of spatial subsampling, i.e. the fact that only a tiny fraction of all neurons can be recorded experimentally with millisecond precision. Such spatial subsampling fundamentally limits virtually any recording and hinders inferences about the collective response of cortical networks (Priesemann *et al.*, 2009; Ribeiro *et al.*, 2010; Priesemann *et al.*, 2014; Ribeiro *et al.*, 2014; Levina and Priesemann, 2017). Because of subsampling, it has been impossible to conclusively distinguish AI and critical dynamics in cortical recordings, even though these states are fundamentally different on the population level.

This manuscript addresses the following questions: (i) How can one assess the dynamical state of cortical networks even under strong subsampling? (ii) What is the dynamical state of cortical networks *in vivo*? (iii) How can one explain the striking differences between dynamical states *in vivo* and *in vitro*? (iv) How can the dynamical state support rapid adaptation of computational requirements to changing environments? This manuscript is structured as follows. It is a compilation of four publications* (Wilting and

Priesemann, 2018a; Wilting *et al.*, 2018; Zierenberg *et al.*, 2018; Wilting and Priesemann, 2019b). Each of these manuscripts addresses one of the questions above.

In Chap. 2 (Wilting and Priesemann, 2018a), we show that subsampling leads to a strong overestimation of stability in a large class of time evolving systems, which include epidemic spread of infectious diseases (Farrington *et al.*, 2003), cell proliferation, evolution (see (Kimmel and Axelrod, 2015) and references therein), neutron processes in nuclear power reactors (Pazy and Rabinowitz, 1973), spread of bank-ruptcy (Filimonov and Sornette, 2012), evolution of stock prices (Mitov *et al.*, 2009), or the propagation of spiking activity in neural networks (Beggs and Plenz, 2003; Haldeman and Beggs, 2005). However, correct risk prediction is essential to timely initiate counter actions to mitigate the propagation of events. We introduce a novel estimator that allows correct risk assessment even under strong subsampling.

In Chap. 3 (Wilting and Priesemann, 2019b) we build on the subsampling-invariant approach presented in Chap. 2 in order to resolve the contradictory results about cortical dynamics. We establish an analytically tractable minimal model for *in vivo*-like activity and estimate the dynamical state of cortical activity. The model reproduces a number of dynamical properties of the network, which are experimentally accessible and enable us to validate our approach. We then predict a number of yet unknown network properties.

Both Chaps. 2 and 3 address a general readership and contain a minimum of mathematical formulae. The corresponding mathematical derivations are presented in the appendix in Chaps. A and B. These appendices also contain details on the used models and simulations, experimental setups, data analysis. Finally, they present supplementary figures with extended results.

In Chap. 4 (Zierenberg *et al.*, 2018), we propose that the input strength is the defining dif-

---

*On each of these publications, I am either first author or

ference between *in vitro* and *in vivo* dynamics. *In vitro* systems are completely isolated, whereas *in vivo* networks receive continuous input from sensory modalities and other brain areas. Under these different conditions, we propose that homeostatic plasticity is a sufficient mechanism to promote self-organization to a diverse set of dynamic states by mediating the interplay between external input rate and neural target spike rate. Thereby, our framework offers testable predictions for the emergence of characteristic but distinct network activity *in vitro* and *in vivo*. Chapter C presents additional derivations and extended results.

In Chap. 5 (Wilting *et al.*, 2018), we propose that cortex operates in a particular dynamic regime, the "reverberating regime" identified in Chap. 3, specifically because in this regime small changes in neural efficacy can tune computational properties over a wide range – a mechanism that we propose to call *dynamic adaptive computation*. In this regime a cortical circuit can interpolate between the asynchronous-irregular and the critical state.

In Chap. 6 we recapitulate the results and draw connections between them. We discuss how the presented results relate to other scientific work and how they advance the research in cortical networks. Finally, we outline future research questions that arise from the presented results, or are made possible through our findings.

Chapter 2

Inferring collective dynamical states from widely unobserved systems[†]

Abstract

When assessing spatially-extended complex systems, one can rarely sample the states of all components. We show that this spatial subsampling typically leads to severe underestimation of the risk of instability in systems with propagating events. We derive a subsampling-invariant estimator, and demonstrate that it correctly infers the infectiousness of various diseases under subsampling, making it particularly useful in countries with unreliable case reports. In neuroscience, recordings are strongly limited by subsampling. Here, the subsampling-invariant estimator allows to revisit two prominent hypotheses about the brain's collective spiking dynamics: asynchronous-irregular or critical. We identify consistently for rat, cat and monkey a state that combines features of both and allows input to reverberate in the network for hundreds of milliseconds. Overall, owing to its ready applicability, the novel estimator paves the way to novel insight for the study of spatially-extended dynamical systems.

[†]The content of this chapter is identical in wording and figures to the publication Wilting and Priesemann (2018a): J. Wilting & V. Priesemann. *Inferring collective dynamical states from widely unobserved systems.* Nature Communications 9(1):2325 (2018). The article is published under the terms of a Creative Common License (http://creativecommons.org/licenses/by/4.0/). The corresponding Appendix A of this monograph is identical in wording and figures to the online supplementary material of the same publication. To this publication, I contributed all presented analytical derivations, simulations of the models, and data analysis, as well as drafting all figures and creating all figures except for panel 2.1**a**. I drafted, wrote, and reviewed the manuscript jointly with V. Priesemann. V. Priesemann designed the study.

2.1 Introduction

How can we infer properties of a high-dimensional dynamical system if we can only observe a very small part of it? This problem of spatial subsampling is common to almost every area of research where spatially extended, time evolving systems are investigated. For example, in many diseases the number of reported infec-

tions may be much lower than the unreported ones (Papoz *et al.*, 1996), or in the financial system only a subset of all banks is evaluated when assessing the risk of developing system wide instability (Quagliariello, 2009) ("stress test"). Spatial subsampling is particularly severe when recording neuronal spiking activity, because the number of neurons that can be recorded with

ms precision is vanishingly small compared to the number of all neurons in a brain area (Priesemann *et al.*, 2009; Ribeiro *et al.*, 2010, 2014) (Fig. 2.1**a**).

Here, we show that subsampling leads to a strong overestimation of stability in a large class of time evolving systems (Sec. A.1), which include epidemic spread of infectious diseases (Farrington *et al.*, 2003), cell proliferation, evolution (see Kimmel and Axelrod (2015) and references therein), neutron processes in nuclear power reactors (Pazy and Rabinowitz, 1973), spread of bank-ruptcy (Filimonov and Sornette, 2012), evolution of stock prices (Mitov *et al.*, 2009), or the propagation of spiking activity in neural networks (Beggs and Plenz, 2003; Haldeman and Beggs, 2005) (Fig. 2.1**b**). However, correct risk prediction is essential to timely initiate counter actions to mitigate the propagation of events. We introduce a novel estimator that allows correct risk assessment even under strong subsampling. Mathematically, the evolution of all these systems is often approximated by a process with a 1st order autoregressive representation (PAR), e.g. by an AR(1), branching, or Kesten process (Fig. A.1, Sec. A.2). For these processes, we derive first the origin of the estimation bias and develop a novel estimator, which we analytically prove to be consistent under subsampling. We then apply the novel estimator to models and real-world data of disease and brain activity. To assure that a PAR is a reasonable approximation of the complex system under study, and to exclude contamination through potential non-stationarities, we included a set of automated, data-driven tests.

2.2   Model

In a PAR, the activity in the next time step, A_{t+1}, depends linearly on the current activity A_t. In addition, it incorporates external input, e.g. drive from stimuli or other brain areas, with a mean rate h, yielding the autoregressive representation

$$\langle A_{t+1}|A_t \rangle = m\, A_t + h, \tag{2.1}$$

For the mathematically inclined reader we recommend the detailed derivation in Secs. A.1 – A.4.

where $\langle \cdot | \cdot \rangle$ denotes the conditional expectation. The stability of A_t is solely governed by m, e.g. the mean number of persons infected by *one* diseased person (Heathcote, 1965). The activity is stationary if $m < 1$, while it grows exponentially if $m > 1$. The state $m = 1$ separates the stable from the unstable regime. Especially close to this transition, a correct estimate of m is vital to assess the risk that A_t develops a large, potentially devastating cascade or avalanche of events (e.g. an epidemic disease outbreak or an epileptic seizure), either generically or via a minor increase in m.

2.3   Results

2.3.1   Subsamling bias of established estimators

A conventional estimator (Heyde and Seneta, 1972; Wei and Winnicki, 1990) $\hat{m}_C$ of m uses linear regression of activity at time t and $t + 1$, because the slope of linear regression directly returns m owing to the autoregressive representation in Eq. (2.1). This estimation of m is consistent if the full activity A_t is known. However, under subsampling it can be strongly biased, as we show here. To derive the bias quantitatively, we model subsampling in a generic manner in our stochastic framework: We assume only that the subsampled activity a_t is a random variable that *in expectation* it is proportional to A_t, $\langle a_t|A_t \rangle = \alpha\, A_t + \beta$ with two constants α and β (Sec. A.3). This represents, for example, sampling a fraction α of all neurons in a brain area. Then the conventional estimator is biased by $m\,(\alpha^2 \text{Var}[A_t]/\text{Var}[a_t] - 1)$ (Corollary 6). The bias vanishes only when all units are sampled ($\alpha = 1$, Figs. 2.1**c**–**e**), but is inherent to subsampling and cannot be overcome by obtaining longer recordings.

Kalman filtering (Hamilton, 1994; Shumway and Stoffer, 1982; Ghahramani and Hinton, 1996), a state-of-the-art approach for system identification, cannot overcome the subsampling bias either, because it assumes Gaussian noise for both the evolution of A_t and the sampling process for generating a_t (see Sec. A.7). These assumptions
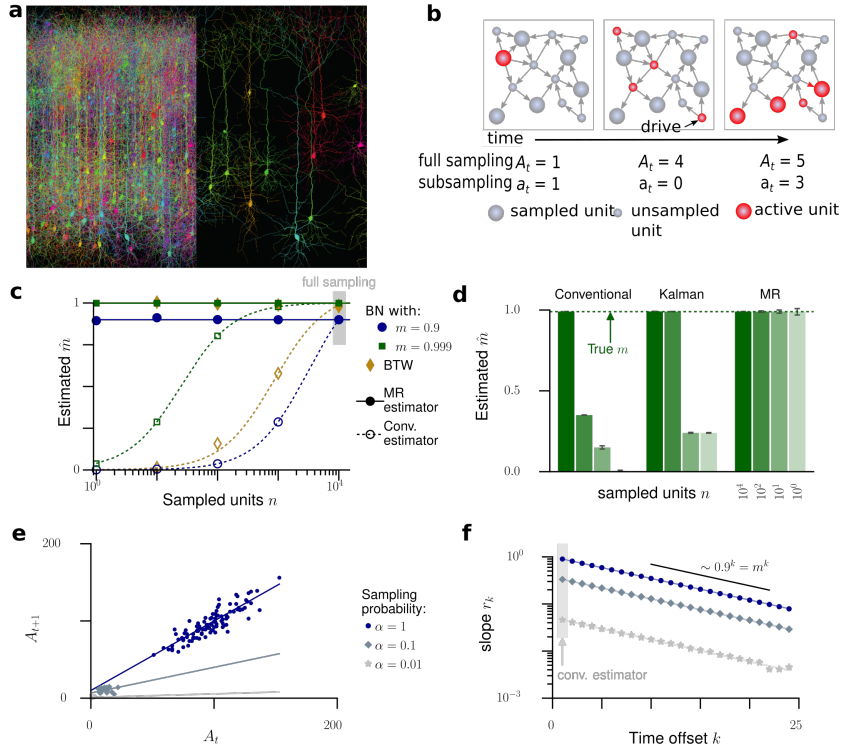
Figure 2.1: Spatial subsampling. a. In complex networks, such as the brain, often only a small subset of all units can be sampled (spatial subsampling); figure created using TREES (Cuntz *et al.*, 2010). **b**. In a branching network (BN), an active unit (e.g. a spiking neuron, infected individual, or defaulting bank) activates some of its neighbors in the next time step. Thereby activity can spread over the system. Units can also be activated by external drive. As the subsampled activity a_t may significantly differ from the actual activity A_t, spatial subsampling can impair inferences about the dynamical properties of the full system. **c**. In recurrent networks (BN, Bak-Tang-Wiesenfeld model (BTW)), the conventional estimator (empty symbols) substantially underestimates the branching ratio m when less units n are sampled, as theoretically predicted (dashed lines). The novel multistep regression (MR) estimator (full symbols) always returns the correct estimate, even when sampling only 10 or 1 out of all $N = 10^4$ units. **d**. For a BN with $m = 0.99$, the conventional estimator infers $\hat{m} = 0.37$, $\hat{m} = 0.1$ or $\hat{m} = 0.02$ when sampling 100, 10, or 1 units respectively. Kalman filtering based estimation returns approximately correct values under slight subsampling ($n = 100$), but is biased under strong subsampling. In contrast, MR estimation returns the correct $\hat{m}$ for any subsampling. **e**. MR estimation is exemplified for a subcritical branching process ($m = 0.9$, $h = 10$), where active units are observed with probability α. Under subsampling (gray), the regression slopes r_1 are smaller than under full sampling (blue). **f**. While conventional estimation of m relies on the linear regression r_1 and is biased under subsampling, MR estimation infers $\hat{m}$ from the exponential relation $r_k \propto m^k$, which remains invariant under subsampling.

15

are violated under typical subsampling conditions, when the values of a_t become too small, so that the central limit theorem is not applicable, and hence Kalman filtering fails (Figs. 2.1**d**, A.7). It is thus applicable to a much narrower set of subsampling problems and in addition requires orders of magnitude longer runtime compared to our novel estimator (Fig. A.7).

2.3.2 A novel, unbiased estimator

Our novel estimator takes a different approach than the other estimators (Sec. A.4). Instead of directly using the biased regression of activity at time t and $t+1$, we perform multiple linear regressions of activity between times t and $t + k$ with different time lags $k = 1, \dots, k_{\max}$. These return a collection of linear regression slopes r_k (note that r_1 is simply the conventional estimator $\hat{m}_C$). Under full sampling, one expects an exponential relation (Statman *et al.*, 2014) $r_k = m^k$ (Theorem 2). Under subsampling, however, we showed that all regressions slopes r_k between a_t and a_{t+k} are biased *by the same factor* $b = \alpha^2 \mathrm{Var}[A_t]/\mathrm{Var}[a_t]$ (Theorem 5). Hence, the exponential relation generalizes to

$$r_k = \alpha^2 \frac{\mathrm{Var}[A_t]}{\mathrm{Var}[a_t]} m^k = b\, m^k \qquad (2.2)$$

under subsampling. The factor b is, in general, not known and thus m cannot be estimated from any r_k alone. However, because b is constant, one does not need to know b to estimate $\hat{m}$ from regressing the collection of slopes r_k against the exponential model $b\, m^k$ according to Eq. (2.2). This result serves as the heart of our new multiple regression (MR) estimator (Fig. 2.1**f**, Figs. A.1, A.2, Corollary 3 and Theorem 5).

In fact, MR estimation is equivalent to estimating the autocorrelation time of subcritical PARs, where autocorrelation and regression r_k are equal: We showed that subsampling decreases the autocorrelation *strength* r_k, but the autocorrelation *time* τ is preserved. This is because the system itself evolves independently of the sampling process. While subsampling biases each regression r_k by decreasing the mutual dependence between subsequent observations (a_t, a_{t+k}), the temporal decay in $r_k \sim m^k =$

$e^{-k\,\Delta t/\tau}$ remains unaffected, allowing for a consistent estimate of m even when sampling only a single unit (Fig. 2.1**d**). Particularly close to $m = 1$ the autocorrelation time $\tau = -\Delta t / \log m$ diverges, which is known as critical slowing down (Scheffer *et al.*, 2012). Because of this divergence, MR estimation can resolve the distance to criticality in this regime with high precision.

The MR estimator is consistent under subsampling, because the system itself evolves independently of the sampling process: While subsampling biases each regression r_k by decreasing the mutual dependence between subsequent observations (a_t, a_{t+k}), the temporal decay in $r_k \sim m^k = e^{-k\,\Delta t/\tau}$ remains unaffected. Here, $\tau = -\Delta t / \log m$ refers to the autocorrelation time of stationary (subcritical) processes, where autocorrelation and regression r_k are equal, and Δt is the time scale of the investigated process. Thus for subcritical PARs, subsampling decreases the autocorrelation *strength* r_k, while the autocorrelation *time* τ is preserved. Making use of this result allows for a consistent estimate of m even when sampling only a single unit (Fig. 2.1**d**).

PARs are typically only a first order approximation of real world event propagation. However, their mathematical structure allowed for an analytical derivation of the subsampling bias and the consistent estimator. To show that the MR estimator returns correct results also for more complex systems, we applied it to more complex simulated systems: a branching network (Haldeman and Beggs, 2005) (BN) and the non-linear Bak-Tang-Wiesenfeld model (Bak *et al.*, 1987) (BTW). In contrast to generic PARs, these models (a) run on recurrent networks and (b) are of finite size. In addition, the second model shows (c) completely deterministic propagation of activity instead of the stochastic propagation that characterizes PARs, and (d) the activity of each unit depends on many past time steps, not only one. Both models approximate neural activity propagation in cortex (Beggs and Plenz, 2003; Haldeman and Beggs, 2005; Priesemann *et al.*, 2009; Ribeiro *et al.*, 2010; Priesemann *et al.*, 2013, 2014). For both models the numerical estimates of m were precisely biased as analytically predicted, although the models are only approximated by
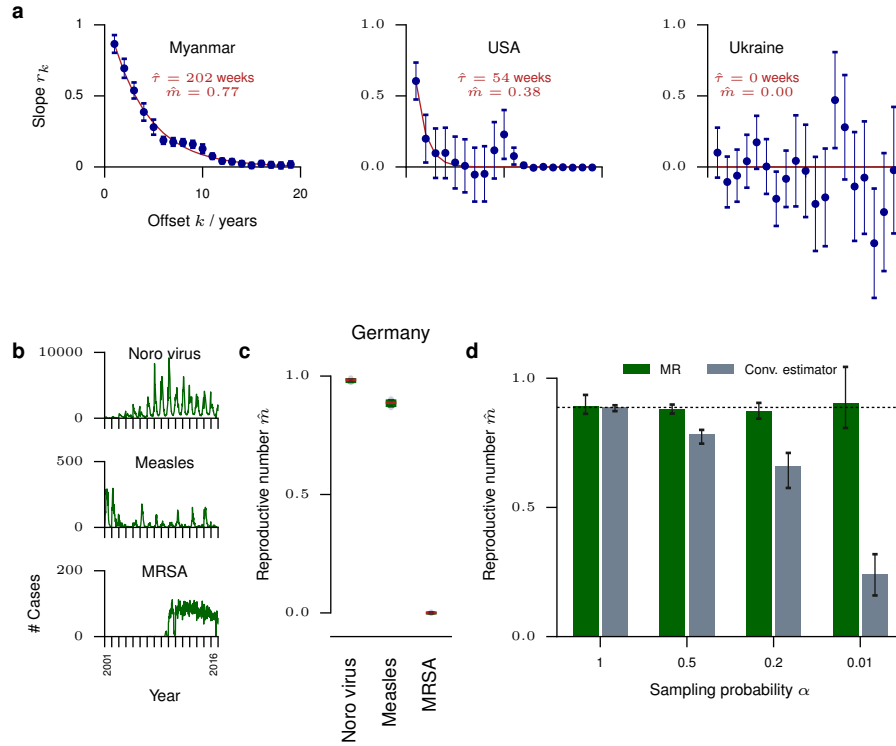
FIGURE 2.2: Disease propagation. In epidemic models, the reproductive number m can serve as an indicator for the infectiousness of a disease within a population, and predict the risk of large incidence bursts. We have estimated $\hat{m}$ from incidence time series of measles infections for 124 countries worldwide (Sec. A.9); as well as noroviral infection, measles, and invasive meticillin-resistant Staphylococcus aureus (MRSA) infections in Germany. **a.** MR estimation of $\hat{m}$ is shown for measles infections in three different countries. Error bars here and in all following figures indicate 1SD or the corresponding 16% to 84% confidence intervals if asymmetric. The reproductive numbers $\hat{m}$ decrease with the vaccination rate (Spearman rank correlation: $r = -0.342, p < 10^{-4}$). **b.** Weekly case report time series for norovirus, measles and MRSA in Germany. **c.** Reproductive numbers $\hat{m}$ for these infections. **d.** When artificially subsampling the measles recording (under-ascertainment), conventional estimation underestimates $\hat{m}_C$, while MR estimation still returns the correct value. Both estimator return the same $\hat{m}$ under full sampling.

a PAR (dashed lines in Fig. 2.1**c**, Eq. A.4). The bias is considerable: For example, sampling 10% or 1% of the neurons in a BN with $m = 0.9$ resulted in the estimates $\hat{m}_C = r_1 = 0.312$, or even $\hat{m}_C = 0.047$, respectively. Thus a process fairly close to instability ($m = 0.9$) is mistaken as Poisson-like ($\hat{m}_C = 0.047 \approx 0$) just because sampling is constrained to 1% of the units. Thereby the risk that systems may develop instabilities is severely underestimated.

MR estimation is readily applicable to subsampled data, because it only requires a sufficiently long time series a_t, and the assumption that in expectation a_t is proportional to A_t. Hence, in general it suffices to sample the system randomly,

without even knowing the system size N, the number of sampled units n, or any moments of the underlying process. Importantly, one can obtain a consistent estimate of m, even when sampling only a very small fraction of the system, under homogeneity even when sampling only one single unit (Figs. 2.1**c,d**, Fig. A.6). This robustness makes the estimator readily applicable to any system that can be approximated by a PAR. We demonstrate the bias of conventional estimation and the robustness of MR estimation at the example of two real-world applications.

2.3.3 Application to disease case reports.

We used the MR estimator to infer the "reproductive number" m from incidence time series of different diseases (Diekmann *et al.*, 1990). Disease propagation represents a nonlinear, complex, real-world system often approximated by a PAR (Earn, 2000; Brockmann *et al.*, 2006). Here, m determines the disease spreading behavior and has been deployed to predict the risk of epidemic outbreaks (Farrington *et al.*, 2003). However, the problem of subsampling or *under-ascertainment* has always posed a challenge (Papoz *et al.*, 1996; Hauri *et al.*, 2011).

As a first step, we cross-validated the novel against the conventional estimator using the spread of measles in Germany, surveyed by the Robert-Koch-Institute (RKI). We chose this reference case, because we expected case reports to be almost fully sampled owing to the strict reporting policy supported by child care facilities and schools (Hellenbrand *et al.*, 2003; Wichmann *et al.*, 2009), and to the clarity of symptoms. Indeed, the values for $\hat{m}$ inferred with the conventional and with the novel estimator, coincided (Fig. 2.2**d**, Sec. A.9). In contrast, after applying artificial subsampling to the case reports, thereby mimicking that each infection was only diagnosed and reported with probability $\alpha < 1$, the conventional estimator severely underestimated the spreading behavior, while MR estimation always returned consistent values (Fig. 2.2**d**). This shows that the MR estimator correctly infers the reproductive number m directly from subsampled time series, without the need to know the degree of under-ascertainment α.

Second, we evaluated worldwide measles case and vaccination reports for 124 countries provided by the WHO since 1980 (Fig. 2.2**a**, Sec. A.9), because the vaccination percentage differs in each country, and this is expected to impact the spreading behavior through m. The reproductive numbers $\hat{m}$ ranged between 0 and 0.93, and in line with our prediction clearly decreased with increasing vaccination percentage in the respective country (Spearman rank correlation: $r = -0.342, p < 10^{-4}$).

Third, we estimated the reproductive numbers for three diseases in Germany with highly different infectiousness: noroviral infection (Hauri *et al.*, 2011; Bernard *et al.*, 2014), measles, and invasive meticillin-resistant Staphylococcus aureus (MRSA, an antibiotic-resistant germ classically associated with health care facilities (Boucher and Corey, 2008), Figs. 2.2**b**,**c**), and quantified their propagation behavior. MR estimation returned the highest $\hat{m} = 0.98$ for norovirus, compliant with its high infectiousness (Teunis *et al.*, 2008). For measles we found the intermediate $\hat{m} = 0.88$, reflecting the vaccination rate of about 97%. For MRSA we identified $m = 0$, confirming that transmission is still minor in Germany (Köck *et al.*, 2011). However, a future increase of transmission is feared and would pose a major public health risk (DeLeo *et al.*, 2010). Such an increase could be detected by our estimator, even in countries where case reports are incomplete.

2.3.4 Reverberating spiking activity *in vivo*

We applied the MR estimator to cortical spiking activity *in vivo* to investigate two contradictory hypothesis about the collective spiking dynamics. One hypothesis suggests that the collective dynamics is "asynchronous irregular" (AI) (Burns and Webb, 1976; Softky and Koch, 1993; de Ruyter van Steveninck *et al.*, 1997; Ecker *et al.*, 2010a), i.e. neurons spike independently of each other and in a Poisson manner ($m = 0$), which may reflect a balanced state (van Vreeswijk and Sompolinsky, 1996; Brunel, 2000; Renart *et al.*, 2010). The other hypothesis suggests that neuronal networks operate at criticality ($m = 1$) (Beggs and Plenz, 2003; Priesemann *et al.*, 2009; Chialvo, 2010; Tkačik *et al.*, 2015; Humplik and Tkačik, 2017), thus in a particularly sensitive state close to a phase transition. These different hypotheses have distinct implications for the coding strategy of the brain: Criticality is characterized by long-range correlations in space and time, and in models optimizes performance in tasks that profit from long reverberation of the activity in the network (Haldeman and Beggs, 2005; Kinouchi and Copelli, 2006; Boedecker *et al.*, 2012; Shew and Plenz, 2013; Del Papa *et al.*, 2017). In contrast, the typical balanced state minimizes redundancy (Hyväri-
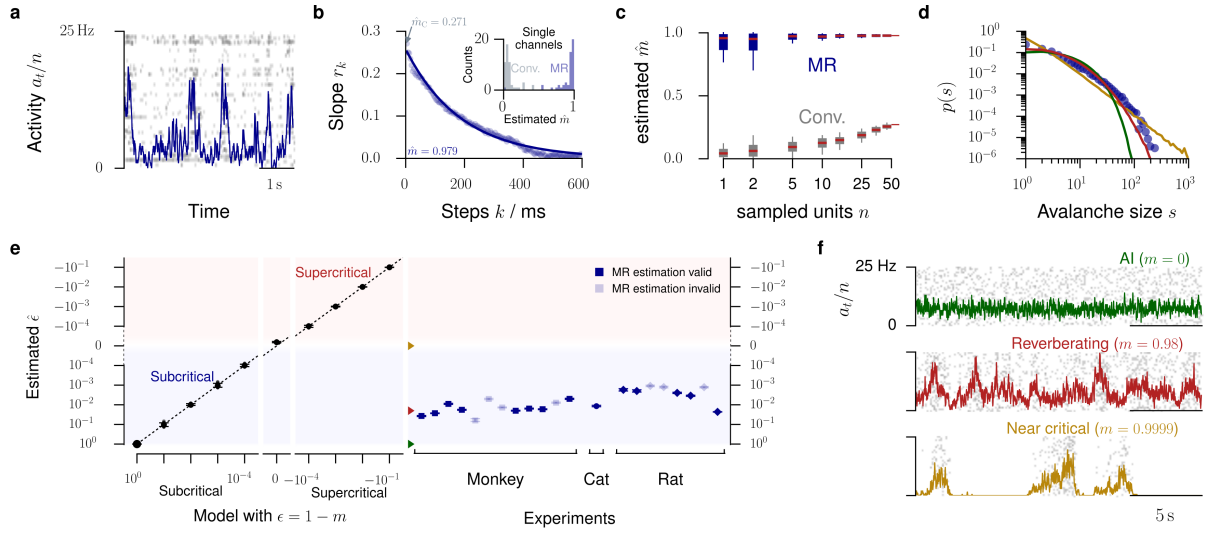
FIGURE 2.3: Animal spiking activity *in vivo*. In neuroscience, m denotes the mean number of spikes triggered by one spike. We estimated $\hat{m}$ from spiking activity recorded *in vivo* in monkey prefrontal cortex, cat visual cortex, and rat hippocampus. **a.** Raster spike plot and population rate a_t of 50 single units illustrated for cat visual cortex. **b.** MR estimation based on the exponential decay of the autocorrelation of r_k of a_t. Inset: Comparison of conventional and MR estimation results for single units (medians $\hat{m}_C = 0.057$ and $\hat{m} = 0.954$ respectively). **c** $\hat{m}$ estimated from from further subsampled cat recordings, estimated with the conventional and MR estimator. Error bars indicate variability over 50 randomly subsampled n out of the recorded 50 channels. **d** Avalanche size distributions for cat visual cortex (blue) and the networks with AI, reverberating and near-critical dynamics in panel **f**. **e.** For all simulations, MR estimation returned the correct distance to instability (criticality) $\epsilon = 1 - m$ (Sec. A.8). *In vivo* spike recordings from rat, cat, and monkey, clearly differed from critical ($\epsilon = 0$) and AI ($\epsilon = 1$) states (median $\hat{m} = 0.98$, error bars: 16% to 84% confidence intervals, note that some confidence intervals are too small to be resolved). Opaque symbols indicate that MR estimation was rejected (Fig. A.5, Sec. A.5). Green, red, and yellow arrows indicate ϵ for the dynamic states shown in panel **f**. **f.** Population activity and raster plots for AI activity, reverberating, *in vivo*-like, and near critical networks. All three networks match the recording from cat visual cortex with respect to number of recorded neurons and mean firing rate.

nen and Oja, 2000) and supports fast network responses (van Vreeswijk and Sompolinsky, 1996).

Analyzing *in vivo* spiking activity from Macaque monkey prefrontal cortex during a memory task, anesthetized cat visual cortex with no stimulus (Figs. 2.3**a**,**b**), and rat hippocampus during a foraging task (Sec. A.10) returned $\hat{m}$ to be between 0.963 and 0.998 (median $\hat{m} = 0.984$, Fig. 2.3**e**, Fig. A.5), corresponding to autocorrelation times between 100 ms and 2000 ms. This clearly suggests that spiking activity *in vivo* is neither AI-like ($m = 0$), nor consistent with a critical state ($m = 1$), but in a reverberating state that shows autocorrelation times of a few hundred milliseconds. We call the range of the dynamical states found in vivo *reverberating*, because input reverberates for a few hundred millisecond in the network, and therefore

enables integration of information (Murray *et al.*, 2014; Chaudhuri *et al.*, 2015; Jaeger and Haas, 2004). Thereby the reverberating state constitutes a specific narrow window between AI state, where perturbations of the firing rate are quenched immediately, and the critical state, in which perturbations can in principle persist infinitely long (for more details, see Wilting and Priesemann (2018a)).

We demonstrate the robustness to subsampling for the activity in cat visual cortex: we chose random subsets of n neurons from the total of 50 recorded single units. For any subset, even for single neurons, MR estimation returned about the same median $\hat{m}$ (Fig. 2.3**c**). In contrast, the conventional estimator misclassified neuronal activity by strongly underestimating $\hat{m}$: instead of $\hat{m} = 0.984$, it returned $\hat{m}_C =$

0.271 for the activity of all 50 neurons. This underestimation gets even more severe when considering stronger subsampling ($n < 50$, Fig. 2.3**c**). Ultimately, for single neuron activity, the conventional estimator returned $\hat{m}_C = 0.057 \approx 0$, which would spuriously indicate dynamics close to AI instead of the reverberating state (inset of Fig. 2.3**b**, Fig. 2.3**c** and Fig. A.6). The underestimation of $\hat{m}_C$ was present in all experimental recordings (r_1 in Fig. A.5).

On first sight, $\hat{m} = 0.984$ may appear close to the critical state, particularly as physiologically a 1.6% difference to $m = 1$ is small in terms of the effective synaptic strength. However, this seemingly small difference in single unit properties has a large impact on the *collective* dynamics and makes AI, reverberating, and critical states clearly distinct. This distinction is readily manifest in the fluctuations of the population activity (Fig. 2.3**f**). Furthermore, the distributions of avalanche sizes clearly differ from the power-law scaling expected for critical systems (Beggs and Plenz, 2003), but are well captured by a matched, reverberating model (Fig. 2.3**d**). Because of the large difference in the network dynamics, the MR estimator can distinguish AI, reverberating, and critical states with the necessary precision. In fact, the estimator would allow for 100 times higher precision when distinguishing critical from non-critical states, assuming *in vivo*-like subsampling and mean firing rate (sampling $n = 100$ from $N = 10^4$ neurons, Fig. 2.3**e**). With larger N, this discrimination becomes even more sensitive (detailed error estimates: Fig. A.4 and Sec. A.6). As the number of neurons in a given brain area is typically much higher than $N = 10^4$ in the simulation, finite size effects are not likely to account for the observed deviation from criticality $\epsilon = 1 - m \approx 10^{-2}$ *in vivo*, supporting that in rat, cat, and monkey the brain does *not* operate in a critical state. Still, additional factors like input or refractory periods may limit the maximum attainable m to quasi-critical dynamics on a Widom line (Williams-García *et al.*, 2014), which could in principle conform with our results.

2.4 Discussion

Most real-world systems, including disease propagation or cortical dynamics, are more complicated than a simple PAR. For cortical dynamics, for example, heterogeneity of neuronal morphology and function, non-trivial network topology, and the complexity of neurons themselves are likely to have a profound impact onto the population dynamics (Marom, 2010). In order to test for the applicability of a PAR approximation, we defined a set of conservative tests (Sec. A.5) and included only those time series, where the approximation by a PAR was considered appropriate. For example, we excluded all recordings that showed an offset in the slopes r_k, because this offset is, strictly speaking, not explained by a PAR and might indicate non-stationarities (Fig. A.3). Even with these conservative tests, we found the exponential relation $r_k = b\,m^k$ expected for PARs in the majority of real-world time series (Fig. A.5, Sec. A.9). This shows that a PAR is a reasonable approximation for dynamics as complex as cortical activity or disease propagation. With using PARs, we draw on the powerful advantage of analytical tractability, which allowed for valuable insight into dynamics and stability of the respective system. It is then a logical next step to refine the model by including additional relevant parameters (Eckmann *et al.*, 2007). However, the increasing richness of detail typically comes at the expense of analytical tractability.

By employing for the first time a consistent, quantitative estimation, we provided evidence that *in vivo* spiking population dynamics reflects a stable, fading reverberation state around $m = 0.98$ universally across different species, brain areas, and cognitive states. Because of its broad applicability, we expect that besides the questions investigated here, MR estimation can substantially contribute to the understanding of real-world dynamical systems in diverse fields of research where subsampling prevails.

Chapter 3

Between perfectly critical and fully irregular: a reverberating model captures and predicts cortical spike propagation[†]

Abstract

Knowledge about the collective dynamics of cortical spiking is very informative about the underlying coding principles. However, even most basic properties are not known with certainty, because their assessment is hampered by spatial subsampling, i.e. the limitation that only a tiny fraction of all neurons can be recorded simultaneously with millisecond precision. Building on a novel, subsampling-invariant estimator, we fit and carefully validate a minimal model for cortical spike propagation. The model interpolates between two prominent states: asynchronous and critical. We find neither of them in cortical spike recordings across various species, but instead identify a narrow "reverberating" regime. This approach enables us to predict yet unknown properties from very short recordings and for every circuit individually, including responses to minimal perturbations, intrinsic network timescales, and the strength of external input compared to recurrent activation – thereby informing about the underlying coding principles for each circuit, area, state and task.

3.1 Introduction

In order to understand how each cortical circuit or network processes its input, it would be desirable to first know its basic dynamical properties. For example, knowing which impact one additional spike has on the network (London *et al.*, 2010) would give insight into the amplification of small stimuli (Douglas *et al.*, 1995; Suarez *et al.*, 1995; Miller, 2016). Knowing how much of cortical activity can be attributed to external activa-

tion or internal activation (Reinhold *et al.*, 2015) would allow to gauge how much of cortical activity is actually induced by stimuli, or rather internally generated, for example in the context of predictive coding (Rao and Ballard, 1999; Clark, 2013). Knowing the intrinsic network timescale (Murray *et al.*, 2014) would inform how long stimuli are maintained in the activity and can be read out for short term memory (Buonomano and Merzenich, 1995; Wang, 2002; Jaeger *et al.*, 2007; Lim and Goldman, 2013). However, not even these basic properties of cortical network dynamics are generally known with certainty.

In the past, insights about these network properties have been strongly hampered by the inevitable limitations of spatial subsampling, i.e. the fact that only a tiny fraction of all neurons can be recorded experimentally with millisecond precision. Such spatial subsampling fundamentally limits virtually any recording and hinders inferences about the collective response of cortical networks (Priesemann *et al.*, 2009; Ribeiro *et al.*, 2010; Priesemann *et al.*, 2014; Ribeiro *et al.*, 2014; Levina and Priesemann, 2017).

To describe network responses, two contradicting hypotheses have competed for more than a decade, and are the subjects of ongoing scientific debate: One hypothesis suggests that collective dynamics are "asynchronous-irregular" (AI) (Burns and Webb, 1976; Softky and Koch, 1993; Stein *et al.*, 2005), i.e. neurons spike independently of each other and in a Poisson manner, which may reflect a balanced state (van Vreeswijk and Sompolinsky, 1996; Brunel, 2000). The other hypothesis proposes that neuronal networks operate at criticality (Beggs and Plenz, 2003; Levina *et al.*, 2007, 2009b; Muñoz, 2018; Beggs and Timme, 2012; Plenz and Niebur, 2014; Tkačik *et al.*, 2015; Humplik and Tkačik, 2017). Criticality is a particular state at a phase transition, characterized by high sensitivity and long-range correlations in space and time.

These hypotheses have distinct implications for the coding strategy of the brain. The typical balanced state minimizes redundancy (Barlow, 2012; Atick, 1992; Bell and Sejnowski, 1997; van Hateren and van der Schaaf, 1998; Hyvärinen and Oja, 2000), supports fast network re-

sponses (van Vreeswijk and Sompolinsky, 1996), and shows vanishing autocorrelation time or network timescale. In contrast, criticality in models optimizes performance in tasks that profit from extended reverberations of activity in the network (Bertschinger and Natschläger, 2004; Haldeman and Beggs, 2005; Kinouchi and Copelli, 2006; Wang *et al.*, 2011; Boedecker *et al.*, 2012; Shew and Plenz, 2013; Del Papa *et al.*, 2017).

Surprisingly, there is experimental evidence for both AI and critical states in cortical networks, although both states are clearly distinct. Evidence for the AI state is based on characteristics of single neuron spiking, resembling a Poisson process, i.e. exponential inter spike interval (ISI) distributions and a Fano factor F close to unity (Burns and Webb, 1976; Tolhurst *et al.*, 1981; Vogels *et al.*, 1989; Softky and Koch, 1993; Gur *et al.*, 1997; de Ruyter van Steveninck *et al.*, 1997; Kara *et al.*, 2000; Carandini, 2004). Moreover, spike count cross-correlations (Ecker *et al.*, 2010b; Cohen and Kohn, 2011) are small. In contrast, evidence for criticality was typically obtained from a population perspective instead, and assessed neuronal avalanches, i.e. spatiotemporal clusters of activity (Beggs and Plenz, 2003; Pasquale *et al.*, 2008; Priesemann *et al.*, 2009; Friedman *et al.*, 2012; Tagliazucchi *et al.*, 2012; Shriki *et al.*, 2013), whose sizes are expected to be power-law distributed if networks are critical (Bak *et al.*, 1987). Deviations from power-laws, typically observed for spiking activity in awake animals (Bédard *et al.*, 2006; Hahn *et al.*, 2010; Ribeiro *et al.*, 2010; Priesemann *et al.*, 2014), were attributed to subsampling effects (Priesemann *et al.*, 2009; Ribeiro *et al.*, 2010; Priesemann *et al.*, 2013; Girardi-Schappo *et al.*, 2013; Priesemann *et al.*, 2014; Ribeiro *et al.*, 2014; Levina and Priesemann, 2017). Hence, different analysis approaches provided evidence for one or the other hypothesis about cortical dynamics.

We here resolve the contradictory results about cortical dynamics, building on a subsampling-invariant approach presented in a companion study (Wilting and Priesemann, 2018a): (i) we establish an analytically tractable minimal model for *in vivo*-like activity, which can interpolate from AI to critical dynamics (Fig.
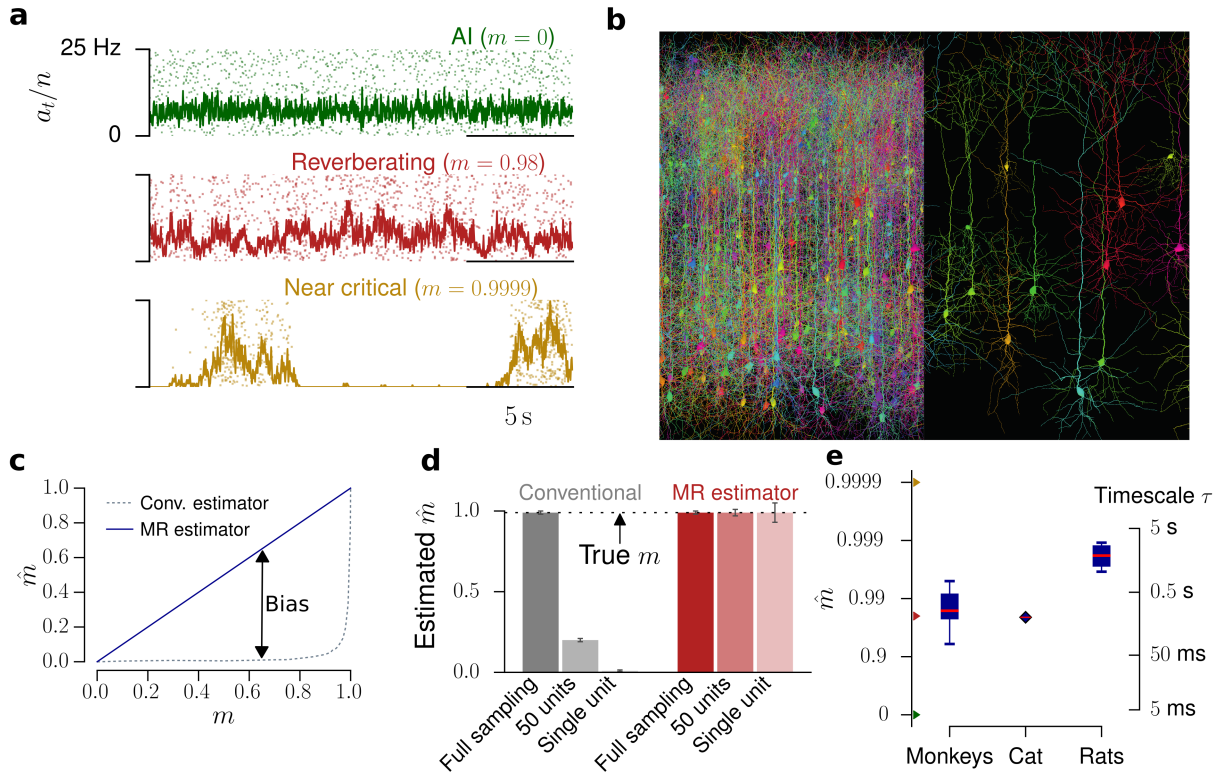
FIGURE 3.1: Reverberating versus critical and irregular dynamics under subsampling. a. Raster plot and population rate a_t for networks with different spike propagation parameters or neural efficacy m. They exhibit vastly different dynamics, which readily manifest in the population activity. **b.** When recording spiking activity, only a small subset of all neurons can be sampled with millisecond precision. This spatial subsampling can hinder correct inference of collective properties of the whole network; figure created using TREES (Cuntz *et al.*, 2010) and reproduced from Wilting and Priesemann (2018a). **c.** Estimated branching ratio $\hat{m}$ as a function of the simulated, true branching ratio m, inferred from subsampled activity (100 out of 10,000 neurons). While the conventional estimator misclassified m from this subsampled observation (gray, dotted line), the novel multistep regression (MR) estimator returned the correct values **d.** For a reverberating branching model with $m = 0.98$, the conventional estimator inferred $\hat{m} = 0.21$ or $\hat{m} = 0.002$ when sampling 50 or 1 units respectively, in contrast to MR estimation, which returned the correct $\hat{m}$ even under strong subsampling. **e.** Using the novel MR estimator, cortical network dynamics in monkey prefrontal cortex, cat visual cortex, and rat hippocampus consistently showed reverberating dynamics, with $0.94 < \hat{m} < 0.991$ (median $\hat{m} = 0.98$ over all experimental sessions, boxplots indicate median / 25% − 75% / 0% − 100% over experimental sessions per species). These correspond to intrinsic network timescales between 80 ms and 2 s.

3.1**a**); (ii) we estimate the dynamical state of cortical activity based on a novel, subsampling-invariant estimator (Wilting and Priesemann, 2018a) (Figs. 3.1**b − d**); (iii) the model reproduces a number of dynamical properties of the network, which are experimentally accessible and enable us to validate our approach; (iv) we predict a number of yet unknown network properties, including the expected number of spikes triggered by one additional spike, the intrinsic network timescale, the distribution

of the total number of spikes triggered by a single extra action potential, and the fraction of activation that can be attributed to afferent external input compared to recurrent activation in a cortical network.

3.2 Material and Methods

We analyzed *in vivo* spiking activity from Macaque monkey prefrontal cortex during a short term memory task (Pipa *et al.*, 2009),

anesthetized cat visual cortex with no stimulus (Blanche and Swindale, 2006; Blanche, 2009), and rat hippocampus during a foraging task (Mizuseki *et al.*, 2009b,a) (Chap. B). We compared the recordings of each experimental session to results of a minimal model of spike propagation, which is detailed in the following.

Minimal model of spike propagation

To gain an intuitive understanding of our mathematical approach, make a thought experiment in your favorite spiking network: apply one additional spike to an excitatory neuron, in analogy to the approach by London *et al.* (2010). How does the network respond to that perturbation? As a first order approximation, one quantifies the number of spikes that are directly triggered *additionally* in all postsynaptic neurons. This number may vary from trial to trial, depending on the membrane potential of the postsynaptic neurons. However, what interests us most is m, the *mean number of spikes triggered by the one extra spike*. Any of these triggered spikes can in turn trigger spikes in their postsynaptic neurons in a similar manner, and thereby the perturbation may cascade through the system.

In the next step, assume that perturbations are started continuously at rate h, for example through afferent input from other brain areas or sensory modalities. Together, this leads to the mathematical framework of a branching model (Harris, 1963; Heathcote, 1965; Pakes, 1971; Beggs and Plenz, 2003; Haldeman and Beggs, 2005; Ribeiro *et al.*, 2010; Priesemann *et al.*, 2013, 2014). This framework describes the number of active neurons A_t in discrete time bins of length Δt. Here, Δt should reflect the propagation time of spikes between neurons. Formally, each spike i at the time bin t excites a random number $Y_{t,i}$ of postsynaptic spikes, on average $m = \langle Y_{t,i} \rangle$. The activity A_{t+1}, i.e. the total number of spikes in the next time bin is then defined as the sum of the postsynaptic spikes of all current spikes A_t, as well as the input h_t:

$$A_{t+1} = \sum_{i=1}^{A_t} Y_{t,i} + h_t. \qquad (3.1)$$

This generic spiking model can generate dynamics spanning AI and critical states depending on the input (Zierenberg *et al.*, 2018), and hence is well suited to probe network dynamics *in vivo* (see Sec. B.3 for details). Most importantly, this framework enables us to infer m and other properties from the ongoing activity proper. Mathematical approaches to infer m are long known if the full network is sampled (Heyde and Seneta, 1972; Wei, 1991). Under subsampling, however, it is the novel estimator described in Wilting and Priesemann (2018a) that for the first time allows an unbiased inference of m, even if only a tiny fraction of neurons is sampled.

A precise estimate of m is essential, because the dynamics of the model is mainly governed by m (Fig. 3.1**a**). Therefore, after inferring m, a number of quantities can be analytically derived, and others can be obtained by simulating a branching model, which is constrained by the experimentally measured m and the spike rate.

Simulation

We simulated a branching model by mapping a branching process (Eq. (3.1) and Sec. B.3) onto a random network of $N = 10,000$ neurons in the annealed disorder limit (Haldeman and Beggs, 2005). An active neuron activated each of its $\kappa = 4$ postsynaptic neurons with probability $p = m/\kappa$. Here, the activated postsynaptic neurons were drawn randomly without replacement at each step, thereby avoiding that two different active neurons would both activate the same target neuron. The branching model is critical for $m = 1$ in the infinite-size limit, and subcritical (supercritical) for $m < 1$ ($m > 1$). We modeled input to the network at rate h by Poisson activation of each neuron at rate h/N. Subsampling (Priesemann *et al.*, 2009) was applied to the model by sampling the activity of n neurons only, which were selected randomly before the simulation, and neglecting the activity of all other neurons. Thereby, instead of the full activity A_t, only the subsampled activity a_t was considered for observation.

If not stated otherwise, simulations were run for $L = 10^7$ time steps (corresponding to ~11 h). Confidence intervals were estimated according

to Wilting and Priesemann (2018a) from $B = 100$ realizations of the model, both for simulation and experiments.

We compared the experimental recordings to three different models: AI, near-critical, and reverberating. All three models were set up to match the experiment in the number of sampled neurons n and firing rate $R = \langle a_t \rangle/(n \cdot \Delta t)$. The AI and near-critical models were set up with branching ratios of $m = 0$ or $m = 0.9999$, respectively. In addition, the reverberating model matched the recording in $m = \hat{m}$, where $\hat{m}$ was estimated from the recording using the novel subsampling-invariant estimator (see below). For all models, we chose a full network size of $N = 10^4$ and then determined the appropriate input $h = R \Delta t N (1 - m)$ in order to match the experimental firing rate. Exemplarily for the cat recording, which happened to represent the median $\hat{m}$, this yielded $\hat{m} = 0.98$, $n = 50$, and $R = 7.25$ Hz. From these numbers, $h = 290$, $h = 5.8$ and $h = 0.029$ followed for the AI, reverberating, and near-critical models, respectively.

In Fig. 3.2, the reverberating branching model was also matched to the length of the cat recording of 295 s. To test for stationarity, the cat recording and the reverberating branching model were split into 59 windows of 5 s each, before estimating m for each window. In Fig. 3.1**c**, subcritical and critical branching models with $N = 10^4$ and $\langle A_t \rangle = 100$ were simulated, and $n = 100$ units sampled.

Subsampling-invariant estimation of $\hat{m}$

Details on the analysis are found in Sec. B.2. For each experimental recording, we collected the spike times of all recorded units (both single and multi units) into one single train of population spike counts a_t, where a_t denotes how many neurons spiked in the tth time bin Δt. If not indicated otherwise, we used $\Delta t = 4$ ms, reflecting the propagation time of spikes from one neuron to the next.

From these experimental time series, we estimated $\hat{m}$ using the multistep regression (MR) estimator described in all detail in Wilting and

Priesemann (2018a). In brief, we calculated the linear regression slope $r_{k \Delta t}$, which describes the linear statistical dependence of a_{t+k} upon a_t, for different time lags $\delta t = k \Delta t$ with $k = 1, \dots, k_{\max}$. In our branching model, these slopes are expected to follow the relation $r_{\delta t} = b \cdot \hat{m}^{\delta t/\Delta t}$ (or $r_{k \Delta t} = b \cdot \hat{m}^k$), where b is an unknown parameter that depends on the higher moments of the underlying process and the degree of subsampling. However, it can be partialled out, allowing for an estimation of m without further knowledge about b. Throughout this study we chose $k_{\max} = 2500$ (corresponding to 10 s) for the rat recordings, $k_{\max} = 150$ (600 ms) for the cat recording, and $k_{\max} = 500$ (2000 ms) for the monkey recordings, assuring that $k_{\max} \Delta t$ was always in the order of multiple intrinsic network timescales. In order to test for the applicability of a MR estimation, we used a set of conservative tests (Wilting and Priesemann, 2018a). The exponential relation can be rewritten as an exponential autocorrelation function $r_{\delta t} = b m^{\delta t/\Delta t} = \exp(\ln m \, \delta t/\Delta t) = \exp(-\delta t/\tau)$, where the intrinsic network timescale τ relates to m as $m = \exp(-\Delta t/\tau)$. While the precise value of m depends on the choice of the bin size Δt and should only be interpreted together with the bin size ($\Delta t = 4$ ms throughout this study), the intrinsic network timescale is independent of Δt. Therefore, we report both values in the following.

3.3 Results

3.3.1 Reverberating spiking activity *in vivo*

We applied MR estimation to the binned population spike counts a_t of the recorded neurons of each experimental session across three different species (see methods). We identified a limited range of branching ratios *in vivo*: in the experiments $\hat{m}$ ranged from 0.963 to 0.998 (median $\hat{m} = 0.98$, for a bin size of $\Delta t = 4$ ms), which is only a narrow window in the continuum from AI ($m = 0$) to critical ($m = 1$). For these values of $\hat{m}$ found in cortical networks, the corresponding τ are between 100 ms and 2 s (median 247 ms, Fig. 3.1**e**, Fig. B.1). This clearly suggests that spiking activity *in vivo* is neither AI-like, nor consistent with a critical state. Instead, it is poised in a

regime that, unlike critical or AI, does not maximize one particular property alone but may flexibly combine features of both (Wilting *et al.*, 2018). Without a prominent characterizing feature, we name it the *reverberating regime*, stressing that activity reverberates (different from the AI state) at timescales of hundreds of milliseconds (different from a critical state, where they can persist infinitely).

3.3.2 Validity of the approach

There is a straight-forward verification of the validity of our phenomenological model: it predicts an exponential autocorrelation function $r_{\delta t}$ for the population activity a_t. We found that the activity in cat visual cortex (Figs. 3.2**a**,**a'**) is surprisingly well described by this exponential fit (Fig. 3.2**b**,**b'**). This validation holds to the majority of experiments investigated (14 out of 21, Fig. B.1).

A second verification of our approach is based on its expected invariance under subsampling: We further subsampled the activity in cat visual cortex by only taking into account spikes recorded from a subset n' out of all available n single units. As predicted (Fig. 3.2**c**), the estimates of $\hat{m}$, or equivalently of the intrinsic network timescale $\hat{\tau}$, coincided for any subset of single units if at least about five of the available 50 single units were evaluated (Fig. 3.2**c'**). Deviations when evaluating only a small subset of units most likely reflect the heterogeneity within cortical networks. Together, these results demonstrate that our approach returns consistent results when evaluating the activity of $n \geqslant 5$ neurons, which were available for all investigated experiments.

3.3.3 Origin of the activity fluctuations

The fluctuations found in cortical spiking activity, instead of being intrinsically generated, could in principle arise from non-stationary input, which could in turn lead to misestimation of m (Priesemann and Shriki, 2018). This is unlikely for three reasons: First, the majority of experiments passed a set of conservative tests that reject recordings that show any signature of common non-stationarities, as defined in Wilt-

ing and Priesemann (2018a). Second, recordings in cat visual cortex were acquired in absence of any stimulation, excluding stimulus-related nonstationarities. Third, when splitting the spike recording into short windows, the window-to-window variation of $\hat{m}$ in the recording did not differ from that of stationary *in vivo*-like reverberating models ($p = 0.3$, Figs. 3.2**d**,**d'**). For these reasons the observed fluctuations in the estimates likely originate from the characteristic fluctuations of collective network dynamics within the reverberating regime.

3.3.4 Timescales of the network and single units

The dynamical state described by m directly relates to an exponential autocorrelation function with an intrinsic network timescale $\tau = -\Delta t / \ln m$. Exemplarily for the cat recording, $m = 0.98$ implies an intrinsic network timescale of $\tau = 188\,\text{ms}$, with $\Delta t = 4\,\text{ms}$ reflecting the spike propagation time from one neuron to the next. While the autocorrelation function of the full network activity is expected to show an exponential decay (Fig. 3.3**a**, blue), this is different for the autocorrelation of single neurons – the most extreme case of subsampling. We showed that subsampling can strongly decrease the absolute values of the autocorrelation function for any non-zero time lag (Fig. 3.3**a**, gray). This effect is typically interpreted as a lack of memory, because the autocorrelation of single neurons decays at the order of the bin size (Fig. 3.3**a**, red). However, ignoring the value at $\delta t = 0$, the floor of the autocorrelation function still unveils the exponential relation. Remarkably, the autocorrelation function of single units in cat visual cortex displayed precisely the shape predicted under subsampling (compare Figs. 3.3**a** and **b**).

3.3.5 Established methods are biased to identifying AI dynamics

On the population level, networks with different m are clearly distinguishable (Fig. 3.1**a**). Surprisingly, single neuron statistics, namely interspike interval (ISI) distributions, Fano factors, conventional estimation of m, and the autocorrelation
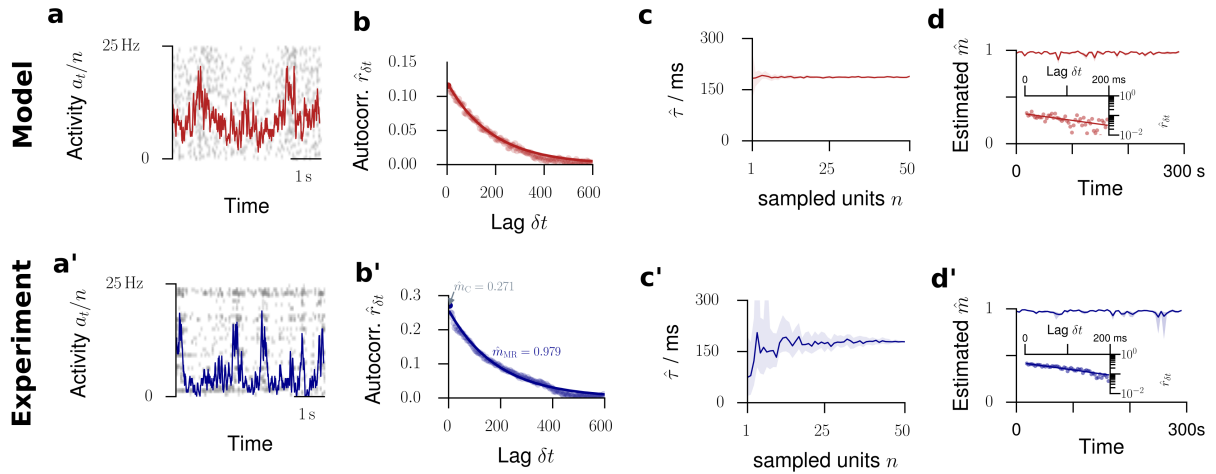
FIGURE 3.2: Validation of the model assumptions. The top row displays properties from a reverberating model, the bottom row spike recordings from cat visual cortex. **a/a'.** Raster plot and population activity a_t within bins of $\Delta t = 4\,\text{ms}$, sampled from $n = 50$ neurons. **b/b'.** Multistep regression (MR) estimation from the subsampled activity (5 min recording). The predicted exponential relation $r_{\delta t} \sim m^{\delta t/\Delta t} = \exp(-\delta t/\tau)$ provides a validation of the applicability of the model. The experimental data are fitted by this exponential with remarkable precision. **c/c'.** When subsampling even further, MR estimation always returns the correct timescale $\hat{\tau}$ (or $\hat{m}$) in the model. In the experiment, this invariance to subsampling also holds, down to $n \approx 10$ neurons (shaded area: 16% to 84% confidence intervals estimated from 50 subsets of n neurons). **d/d'.** The estimated branching parameter $\hat{m}$ for 59 windows of 5 s length suggests stationarity of m over the entire recording (shaded area: 16% to 84% confidence intervals). The variability in $\hat{m}$ over consecutive windows was comparable for experimental recording and the matched model ($p = 0.09$, Levene test). Insets: Exponential decay exemplified for one example window each.
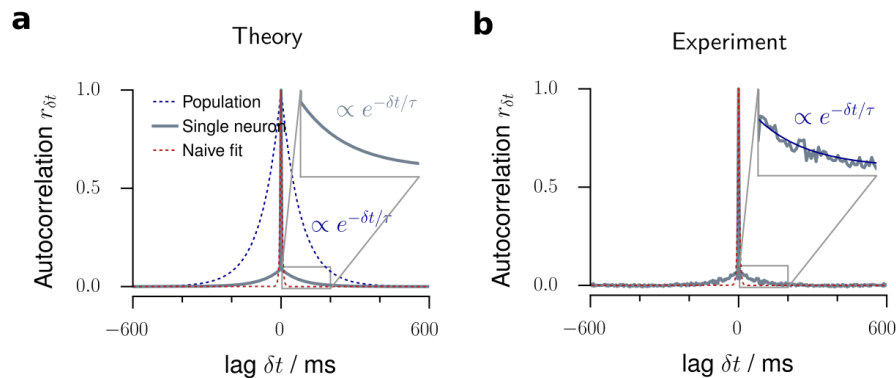
FIGURE 3.3: MR estimation and intrinsic network timescales. a. In a branching model, the autocorrelation function of the population activity decays exponentially with an intrinsic network timescale τ (blue dotted line). In contrast, the autocorrelation function for single neurons shows a sharp drop from $r_0 = 1$ at lag $\delta t = 0$ to the next lag $r_{\pm\Delta t}$ (gray solid line). We showed previously that this drop is a subsampling-induced bias. When ignoring the zero-lag value, the autocorrelation strength is decreased, but the exponential decay and even the value of the intrinsic network timescale τ of the network activity are preserved (inset). The red, dashed line shows a potential, naive exponential function, fitted to the single neuron autocorrelation function (gray). This naive fit would return a much smaller τ. **b.** The autocorrelation function of single neuron activity recorded in cat visual cortex (gray) precisely resembles this theoretical prediction, namely a sharp drop and then an exponential decay (blue, inset), which persists over more than 100 ms. A naive exponential fit (red) to the single neuron data would return an extremely short τ.

strength $r_{\delta t}$, all returned signatures of AI activity regardless of the underlying network dynamics, and hence these single-neuron properties don't serve as a reliable indicator for the network's dynamical state.

First, exponential interspike interval (ISI) distributions are considered a strong indicator of Poisson-like firing. Surprisingly, the ISIs of single neurons in the *in vivo*-like branching model closely followed exponential distributions as well. The ISI distributions were almost indistinguishable for reverberating and AI models (Figs. 3.4**a**,**a**', Fig. B.2). In fact, the ISI distributions are mainly determined by the mean firing rate. This result was further supported by coefficients of variation close to unity, as expected for exponential ISI distributions and Poisson firing (Fig. B.2).

Second, for both the AI and reverberating regime, the Fano factor F for single unit activity was close to unity, a hallmark feature of irregular spiking (Tolhurst *et al.*, 1981; Vogels *et al.*, 1989; Softky and Koch, 1993; Gur *et al.*, 1997; de Ruyter van Steveninck *et al.*, 1997; Kara *et al.*, 2000; Carandini, 2004) (Fig. 3.5**g**, analytical result: Eq. (B.9)). Hence it cannot serve to distinguish between these different dynamical states. When evaluating more units, or increasing the bin size to 4 s, the differences became more pronounced, but for experiments, the median Fano factor of single unit activity did not exceed $F = 10$ in any of the experiments, even in those with the longest reverberation (Figs. 3.4**b**,**b**', Fig. B.3). In contrast, for the full network the Fano factor rose to $F \approx 10^4$ for the *in vivo*-like branching model and diverged when approaching criticality (Fig. 3.5**g**, analytical result: Eq. (B.5)).

Third, conventional regression estimators (Heyde and Seneta, 1972; Wei, 1991) are biased towards inferring irregular activity, as shown before. Here, conventional estimation yielded a median of $\hat{m} = 0.057$ for single neuron activity in cat visual cortex, in contrast to $\hat{m} = 0.954$ returned by MR estimation (Fig. B.9).

Fourth, for the autocorrelation function of an experimental recording (Fig. 3.3**b**) the rapid decay of $r_{\delta t}$ prevails, and hence single neuron activity appears uncorrelated in time.

3.3.6 Cross-validation of model predictions

We compared the experimental results to an *in vivo*-like model, which was matched to each experiment only in the average firing rate, and in the inferred branching ratio $\hat{m}$. Remarkably, this *in vivo*-like branching model could predict statistical properties not only of single neurons (ISI and Fano factor, see above), but also pairwise and population properties, as detailed below. This prediction capability further underlines the usefulness of this simple model to approximate the default state of cortical dynamics.

First, the model predicted the activity distributions, $p(a_t)$, better than AI or critical models for the majority of experiments (15 out 21, Figs. 3.4**c**,**d**,**c**',**d**', Figs. B.5, B.6), both for the exemplary bin sizes of 4 ms and 40 ms. Hence, the branching models, which were only matched in their respective first moment of the activity distributions (through the rate) and first moment of the spreading behavior (through m), in fact approximated all higher moments of the activity distributions $p(a_t)$.

Likewise, the model predicted the distributions of neural avalanches, i.e. spatio-temporal clusters of activity (Figs. 3.4**e**,**f**,**e**',**f**', Figs. B.7, B.8). Characterizing these distributions is a classic approach to assess criticality in neuroscience (Beggs and Plenz, 2003; Priesemann *et al.*, 2014), because avalanche size and duration distributions, $p(s)$ and $p(d)$, respectively, follow power laws in critical systems. In contrast, for AI activity, they are approximately exponential (Priesemann and Shriki, 2018). The matched branching models predicted neither exponential nor power law distributions for the avalanches, but very well matched the experimentally obtained distributions (compare red and blue in Figs. 3.4**e**,**f**,**e**',**f**', Figs. B.7, B.8). Indeed, model likelihood (Clauset *et al.*, 2009) favored the *in vivo*-like branching model over Poisson and critical models for the majority experiments (18 out of 21, Fig. B.7). Our results here are consistent with those of spiking activity in awake animals, which typically do not display power laws (Priesemann *et al.*, 2014; Ribeiro *et al.*, 2010; Bédard *et al.*, 2006). In contrast, most evidence for criticality *in vivo*, in

Best examples

Typical examples
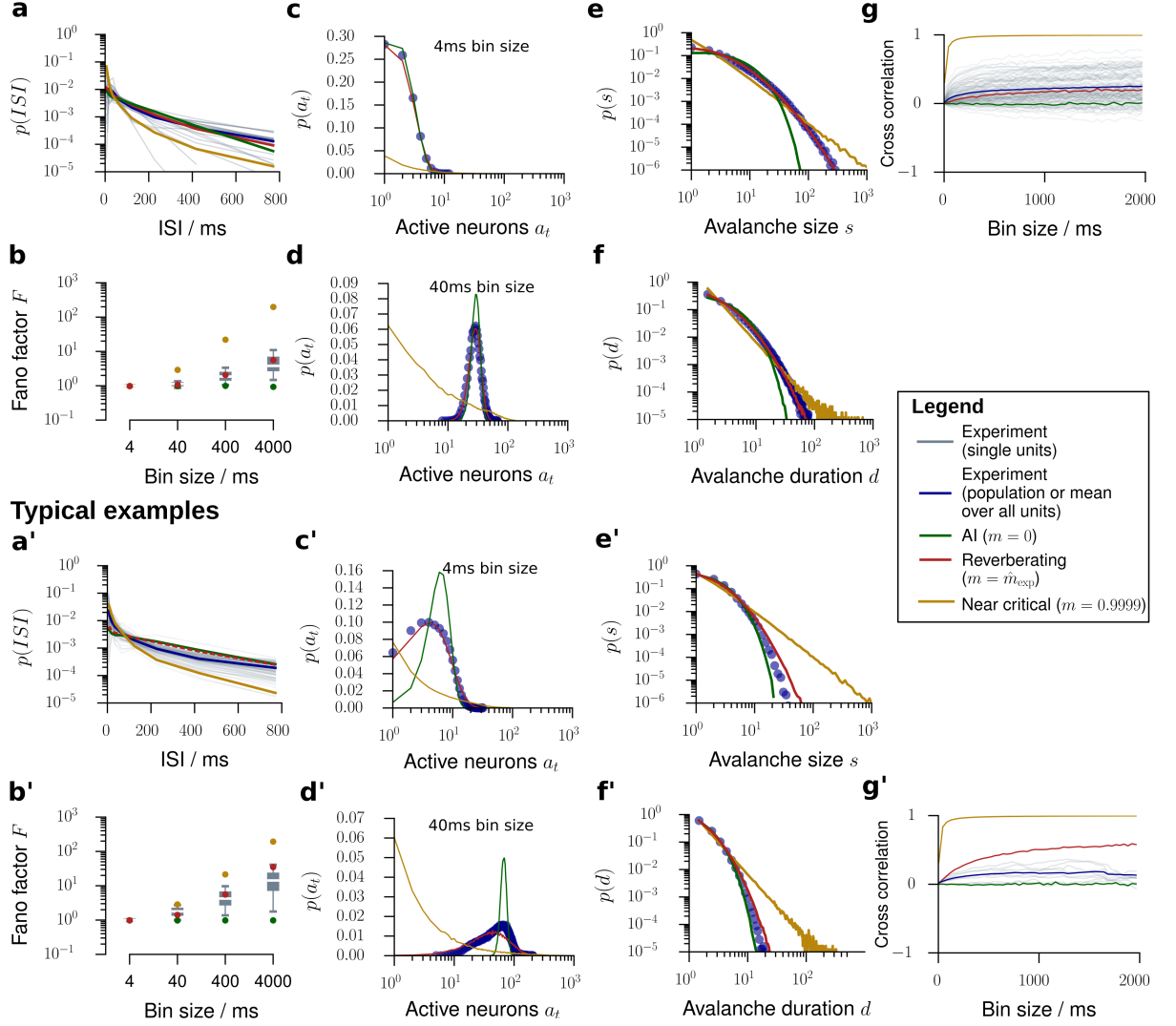
FIGURE 3.4: **Model validation for *in vivo* spiking activity.** We validated our model by comparing experimental results to predictions obtained from the *in vivo*-like, reverberating model, which was matched to the recording in the mean rate, inferred *m*, and number of recorded neurons. In general, the experimental results (gray or blue) were best matched by this reverberating model (red), compared to asynchronous-irregular (AI, green) and near-critical (yellow) models. From all experimental sessions, best examples (top) and typical examples (bottom) are displayed. For results from all experimental sessions see Figs. B.2 to B.9. **a/a'**. Inter-spike-interval (ISI) distributions. **b/b'**. Fano factors of single neurons for bin sizes between 4 ms and 4 s. **c/c'**. Distribution of spikes per bin $p(a_t)$ at a bin size of 4 ms. **d/d'**. Same as **c/c'** with a bin size of 40 ms. **e/e'**. Avalanche size distributions $p(s)$ for all sampled units. AI activity lacks large avalanches, near critical activity produces power-law distributed avalanches, even under subsampling. **f/f'**. Same as **e/e'**, but for the avalanche duration distributions $p(d)$. **g/g'**. Spike count cross-correlations (r_{sc}) as a function of the bin size.

particular the characteristic power-law distributions, has been obtained from *coarse* measures of neural activity (LFP, EEG, BOLD; see Priesemann *et al.* (2014) and references therein).

Last, the model predicted the pairwise spike count cross correlation r_{sc}. In experiments, r_{sc}

is typically between 0.01 and 0.25, depending on brain area, task, and most importantly, the analysis timescale (bin size) (Cohen and Kohn, 2011). For the cat recording the model even correctly predicted the bin size dependence of r_{sc} from $\bar{r}_{sc} \approx 0.004$ at a bin size of 4 ms (analytical result:

Eq. (B.12)) to $\bar{r}_{sc} \approx 0.3$ at a bin size of $2\,s$ (Fig. 3.4**g**). Comparable results were also obtained for some monkey experiments. In contrast, correlations in most monkey and rat recordings were smaller than predicted (Fig. 3.4**g**', Fig. B.4). It is very surprising that the model correctly predicted the cross-correlation even in some experiments, as m was inferred only from the *temporal* structure of the spiking activity alone, whereas r_{sc} characterizes spatial dependencies.

Overall, by only estimating the effective synaptic strength m from the *in vivo* recordings, higher-order properties like avalanche size distributions, activity distributions and in some cases spike count cross correlations could be closely matched using the generic branching model.

3.3.7 The dynamical state determines responses to small stimuli

After validating the model using a set of statistical properties that are well accessible experimentally, we now turn to making predictions about yet unknown properties, namely network responses to small stimuli. In the line of London *et al.* (2010), assume that on a background of spiking activity one single extra spike is triggered. This spike may in turn trigger new spikes, leading to a cascade of additional spikes Δ_t propagating through the network. A dynamical state with branching ratio m implies that *on average*, this perturbation decays with time constant $\tau = -\Delta t / \log m$. Similar to the approach in London *et al.* (2010), the evolution of the mean firing rate, averaged over a reasonable number of trials (here: 500) unveils the nature of the underlying spike propagation: depending on m, the rate excursions will last longer, the higher m (Figs. 3.5**a**,**b**,**c**, Fig. B.11). The perturbations are not deterministic, but show trial-to-trial variability which also increases with m (B.11**b**).

Unless $m > 1$, the theory of branching models ensures that perturbations will die out eventually after a duration d_Δ, having accumulated a total of $s_\Delta = \sum_{t=1}^d \Delta_t$ extra spikes in total. This perturbation size s_Δ and duration d_Δ follow specific distributions (Harris, 1963), which are determined by m: they are power law distributed in the crit-

ical state ($m = 1$), with a cutoff for any $m < 1$ (Fig. 3.5**f**, Figs. B.11**c**,**d**). These distributions imply a characteristic mean perturbation size $\langle s_\Delta \rangle$ (Fig. 3.5**d**), which diverges at the critical point. The variability of the perturbation sizes is also determined by m and also diverges at the critical point (inset of Fig. 3.5**d**, Fig. B.11**e**).

Taken together, these results imply that the closer a neuronal network is to criticality, the more sensitive it is to external perturbations, and the better it can amplify small stimuli. At the same time, these networks also show larger trial-to-trial variability. For typical cortical networks, we found that the response to one single extra spike will on average comprise between 20 and 1000 additional spikes in total (Fig. 3.5**e**).

3.3.8 The dynamical state determines network susceptibility and variability

Moving beyond single spike perturbations, our model gives precise predictions for the network response to continuous stimuli. If extra action potentials are triggered at rate h in the network, the network will again amplify these external activations, depending on m. Provided an appropriate stimulation protocol, this rate response could be measured and our prediction tested in experiments (Fig. B.11**g**). The susceptibility $\partial R / \partial h$ diverges at the critical transition and is unique to a specific branching ratio m. We predict that typical cortical networks will amplify a small, but continuous input rate by about a factor fifty (Fig. B.11**h**, red).

While the input and susceptibility determine the network's mean activity, the network still shows strong rate fluctuations around this mean value. The magnitude of these fluctuations in relation to the mean can be quantified by the network Fano factor $F = \text{Var}[A_t] / \langle A_t \rangle$ (Fig. 3.5**g**). This quantity cannot be directly inferred from experimental recordings, because the Fano factor of subsampled populations severely underestimates the network Fano factor, as shown before. We here used our *in vivo*-like model to obtain estimates of the network Fano factor: for a bin size of $\Delta t = 4\,\text{ms}$ it is about $F \approx 40$ and rises to $F \approx 4000$ for bin sizes of several seconds, highlighting that network fluctuations probably are much stronger
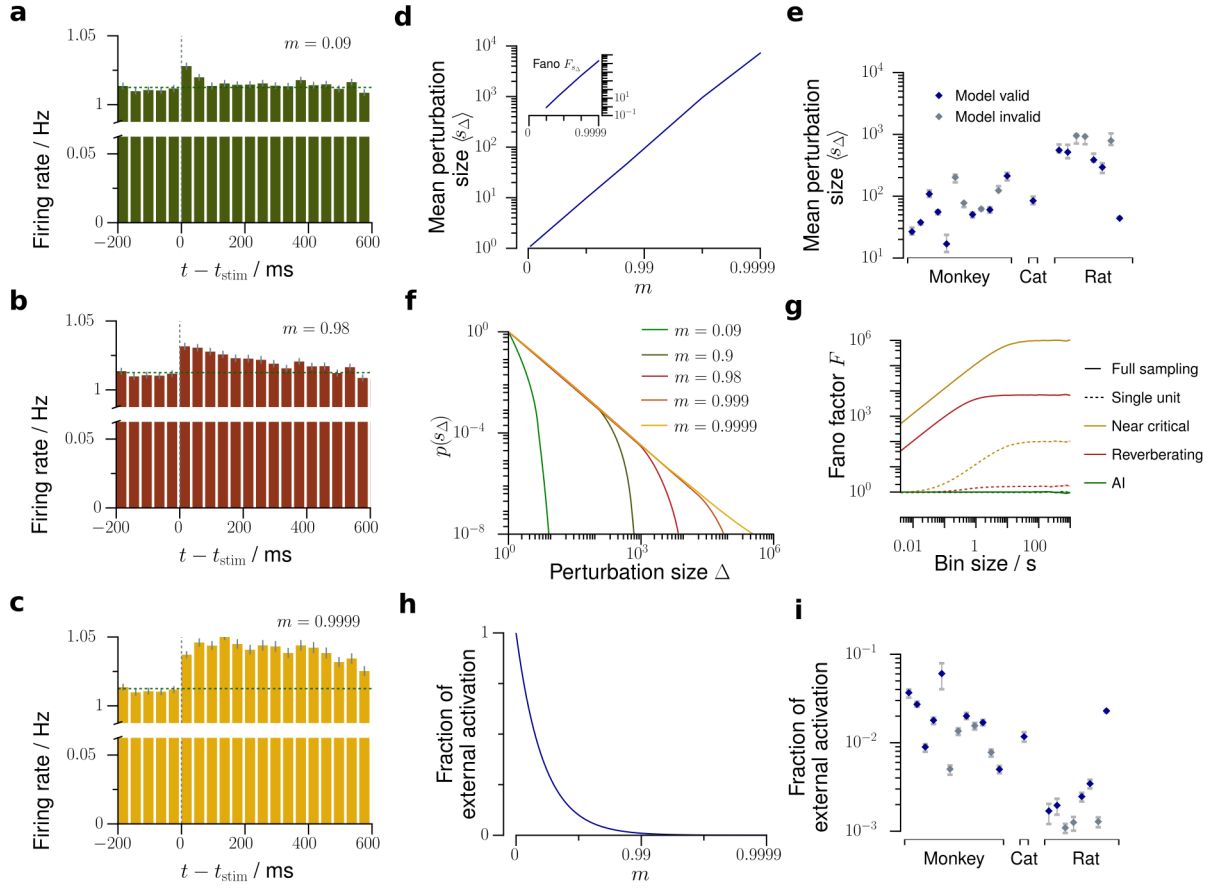
Figure 3.5: Predictions about network dynamics and propagation of perturbations. Using our *in-vivo*-like, reverberating model, we can predict several network properties, which are yet very complicated or impossible to obtain experimentally. **a – c.** In response to one single extra spike, a perturbation propagates in the network depending on the branching ratio *m*, and can be observed as a small increase of the average firing rate of the sampled neurons, here simulated for 500 trials (as in (London *et al.*, 2010)). This increase of firing rate decays exponentially, with the decay time τ being determined by *m*. The perturbation **a** is rapidly quenched in the asynchronous-irregular state, **b** decays slowly over hundreds of milliseconds in the reverberating state, or **c** persists almost infinitely in the critical state. **d.** The average perturbation size $\langle s_\Delta \rangle$ and Fano factor F_{s_Δ} (inset) increase strongly with *m*. **e.** Average total perturbation sizes predicted for each spike recording of mammalian cortex (errorbars: 5% – 95% confidence intervals). **f.** Distribution $p(s_\Delta)$ of total perturbation sizes s_Δ. The asynchronous-irregular models show approximately Poisson distributed, near critical models power-law distributed perturbation sizes. **g.** Bin size dependent Fano factors of the activity, here exemplarily shown for the asynchronous-irregular (*m* = 0, green), representative reverberating (*m* = 0.98, red), and near critical (*m* = 0.9999, yellow) models. While the directly measurable Fano factor of single neurons (dotted lines) underestimates the Fano factor of the whole network, the model allows to predict the Fano factor of the whole network (solid lines). **h.** The fraction of the externally generated spikes compared to all spikes in the network strongly decreases with larger *m*. **i.** Fraction of the externally generated spikes predicted for each spike recording of mammalian cortex (errorbars as in **e**).

than one would naively assume from experimental, subsampled spiking activity.

3.3.9 Distinguishing afferent and recurrent activation

Last, our model gives an easily accessible approach to solving the following question: given a

spiking neuronal network, which fraction of the activity $\langle A \rangle$ is generated by recurrent activation from within the network, and which fraction can be attributed to external, afferent excitation h? The branching model readily provides an answer: the fraction of external activation is $h/\langle A \rangle = 1 - m$ (Fig. 3.5**h**). In this framework, AI-like networks are completely driven by external input currents or noise, whereas reverberating networks generate a substantial fraction of their activity intrinsically. For the experiments investigated in this study, we inferred that between 0.1% and 7% of the activity are externally generated (median 2%, Fig. 3.5**i**).

While our model is quite simplistic given the complexity of neuronal network activity, keep in mind that "all models are wrong, but some are useful" (Box, 1979). Here, the model has proven to provide a good first order approximation to a number of statistical properties of spiking activity and propagation in cortex. Hence, it promises insight into cortical function because (i) it relies on simply assessing spontaneous cortical activity, (ii) it does not require manipulation of cortex, (iii) it enables reasonable predictions about sensitivity, amplification, and internal and external activation, (iv) this analysis is possible in an area specific, task- and state-dependent manner as only short recordings are required for consistent results.

3.4 Discussion

3.4.1 Our results resolve contradictions between AI and critical states

Our results for spiking activity *in vivo* suggest that network dynamics show AI-like statistics, because under subsampling the observed correlations are underestimated. In contrast, typical experiments that assessed criticality potentially overestimated correlations by sampling from overlapping populations (LFP, EEG) and thereby hampered a fine distinction between critical and subcritical states (Pinheiro Neto *et al.*, in prep). By employing for the first time a consistent, quantitative estimation, we provided evidence that *in vivo* spiking population dynamics

reflects a reverberating regime, i.e. it operates in a narrow regime around $m = 0.98$. This result is supported by the findings by (Dahmen *et al.*, 2016): based on distributions of covariances, they inferred that cortical networks operate in a regime below criticality. Given the generality of our results across different species, brain areas, and cognitive states, our results suggest self-organization to this reverberating regime as a general organization principle for cortical network dynamics.

3.4.2 The reverberating regime combines features of AI and critical state

At first sight, $\hat{m} = 0.98$ of the reverberating regime may suggest that the collective spiking dynamics is very close to critical. Indeed, physiologically a $\Delta m \approx 1.6\%$ difference to criticality ($m = 1$) is small in terms of the effective synaptic strength. However, this apparently small difference in single unit properties has a large impact on the collective *dynamical* fingerprint and makes AI, reverberating, and critical states clearly distinct: For example, consider the sensitivity to a small input, i.e. the susceptibility $\chi = \partial R / \partial h = \frac{1}{1-m}$. The susceptibility diverges at criticality, making critical networks overly sensitive to input. In contrast, states with $m \approx 0.98$ assure sensitivity without instability. Because this has a strong impact on network dynamics and putative network function, finely distinguishing between dynamical states is both important and feasible even if the corresponding differences in effective synaptic strength (m) appear small.

We cannot ultimately rule out that cortical networks self-organize as close as possible towards criticality, the platonic ideal being impossible to achieve for example due to finite-size, external input, and refractory periods. Therefore, the reverberating regime might conform with quasi-criticality (Williams-García *et al.*, 2014) or neutral theory (Martinello *et al.*, 2017). However, we deem this unlikely for two reasons. First, in simulations of finite-size networks with external input, we could easily distinguish the reverberating regime from states with $m = 0.9999$ (Wilting and Priesemann, 2018a), which are more than

one order of magnitude closer to criticality than any experiment we analyzed. Second, operating in a reverberating regime, which is between AI and critical, may combine the computational advantages of both states (Wilting *et al.*, 2018): the reverberating regime enables rapid changes of computational properties by small parameter changes, keeps a sufficient safety-margin from instability to make seizures sufficiently unlikely (Priesemann *et al.*, 2014), balances competing requirements (e.g. sensitivity and specificity (Gollo, 2017)), and may carry short term memory and allow to integrate information over limited, tunable timescales (Wang, 2002; Boedecker *et al.*, 2012). For these reasons, we consider the reverberating regime to be the explicit target state of self-organization. This is in contrast to the view of "as close to critical as possible", which still holds criticality as the ideal target.

3.4.3 More complex network models

Cortical dynamics is clearly more complicated than a simple branching model. For example, heterogeneity of single-neuron morphology and dynamics, and non-trivial network topology likely impact population dynamics. However, we showed that *statistics* of cortical network activity are well approximated by a branching model. Therefore, we interpret branching models as a *statistical* approximation of spike propagation, which can capture a fair extent of the complexity of cortical dynamics. By using branching models, we draw on the powerful advantage of analytical tractability, which allowed for basic insight into dynamics and stability of cortical networks.

In contrast to the branching model, doubly stochastic processes (i.e. spikes drawn from an inhomogeneous Poisson distribution) failed to reproduce many statistical features (Fig. B.10). We conjecture that the key difference is that doubly stochastic processes propagate the underlying firing rate instead of the actual spike count. Thus, propagation of the actual number of spikes (as e.g. in branching or Hawkes processes, Kossio *et al.* (2018)), not some underlying firing rate, seems to be integral to capture the statistics of cortical spiking dynamics.

Our statistical model stands in contrast to *generative* models, which generate spiking dynamics by physiologically inspired mechanisms. One particularly prominent example are networks with balanced excitation and inhibition (van Vreeswijk and Sompolinsky, 1996, 1997; Brunel, 2000), which became a standard model of neuronal networks (Hansel and van Vreeswijk, 2012). A balance of excitation and inhibition is supported by experimental evidence (Okun and Lampl, 2008). Our statistical model reproduces statistical properties of such networks if one assumes that the excitatory and inhibitory contributions can be described by an effective excitation. In turn, the results obtained from the well-understood estimator can guide the refinement of generative models. For example, we suggest that network models need to be extended beyond the asynchronous-irregular state (Brunel, 2000) to incorporate the network reverberations observed *in vivo*. Possible candidate mechanisms are increased coupling strength or inhomogeneous connectivity. Both have already been shown to induce rate fluctuations with timescales of several hundred milliseconds (Litwin-kumar and Doiron, 2012; Ostojic, 2014; Kadmon and Sompolinsky, 2015).

Because of the assumption of uncorrelated, Poisson-like network firing, models that study single neurons typically assume that synaptic currents are normally distributed. Our results suggest that one should rather use input with reverberating properties with timescales of a few hundred milliseconds to reflect input from cortical neurons *in vivo*. This could potentially change our understanding of single neuron dynamics, for example of their input-output properties.

3.4.4 Deducing network properties from the tractable model

Using our analytically tractable model, we could predict and validate network properties, such as avalanche size and duration, interspike interval, or activity distributions. Given the experimental agreement with these predictions, we deduced further properties, which are impossible or difficult to assess experimentally and gave insight into more complex questions about network responses: how do perturbations propagate within

the network, and how susceptible is the network to external stimulation?

One particular question we could address is the following: which fraction of network activity is attributed to external or recurrent, internal activation? We inferred that about 98% of the activity is generated by recurrent excitation, and only about 2% originates from input or spontaneous threshold crossing. This result may depend systematically on the brain area and cognitive state investigated: For layer 4 of primary visual cortex in awake mice, (Reinhold *et al.*, 2015) concluded that the fraction of recurrent cortical excitation is about 72%, and cortical activity dies out with a timescale of about 12 ms after thalamic silencing. Their numbers agree perfectly well with our phenomenological model: a timescale of τ = 12 ms implies that the fraction of recurrent cortical excitation is $m = e^{-\Delta t/\tau} \approx 72\%$, just as found experimentally. Under anesthesia, in contrast, they report timescales of several hundred milliseconds, in agreement with our results. These differences show that the fraction of external activation may strongly depend on cortical area, layer, and cognitive state. The novel estimator can in future contribute to a deeper insight into these differences, because it allows for a straight-forward assessment of afferent versus recurrent activation, simply from evaluating spontaneous spiking activity, without the requirement of thalamic or cortical silencing.

3.5 Acknowledgments

3.6 Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Chapter 4

Homeostatic plasticity and external input shape neural network dynamics[†]

Abstract

In vitro and *in vivo* spiking activity clearly differ. Whereas networks *in vitro* develop strong bursts separated by periods of very little spiking activity, *in vivo* cortical networks show continuous activity. This is puzzling considering that both networks presumably share similar single-neuron dynamics and plasticity rules. We propose that the defining difference between *in vitro* and *in vivo* dynamics is the strength of external input. *In vitro*, networks are virtually isolated, whereas *in vivo* every brain area receives continuous input. We analyze a model of spiking neurons in which the input strength, mediated by spike rate homeostasis, determines the characteristics of the dynamical state. In more detail, our analytical and numerical results on various network topologies show consistently that under increasing input, homeostatic plasticity generates distinct dynamic states, from bursting, to close-to-critical, reverberating and irregular states. This implies that the dynamic state of a neural network is not fixed but can readily adapt to the input strengths. Indeed, our results match experimental spike recordings *in vitro* and *in vivo*: the *in vitro* bursting behavior is consistent with a state generated by very low network input ($< 0.1\%$), whereas *in vivo* activity suggests that on the order of 1% recorded spikes are input-driven, resulting in reverberating dynamics. Importantly, this predicts that one can abolish the ubiquitous bursts of *in vitro* preparations, and instead impose dynamics comparable to *in vivo* activity by exposing the system to weak long-term stimulation, thereby opening new paths to establish an *in vivo*-like assay *in vitro* for basic as well as neurological studies.

4.1 Introduction

Collective spiking activity clearly differs between *in vitro* cultures and *in vivo* cortical networks (see examples in Fig. 4.1). Cultures *in vitro* typically exhibit stretches of very little spiking activity, interrupted by strong bursts of highly synchronized or coherent activity (Robinson *et al.*, 1993; Van Pelt *et al.*, 2004; Chiappalone *et al.*, 2006; Wagenaar *et al.*, 2006; Orlandi *et al.*, 2013; Vardi *et al.*, 2016; Beggs and Plenz, 2003). In contrast, spiking activity recorded from cortex in awake animals *in vivo* lacks such pauses, and instead shows continuous, fluctuating activity. These fluctuations show a dominant autocorrelation time that was proposed to increase hierarchically across cerebral cortex, from sensory to frontal areas (Murray *et al.*, 2014). Moreover, depending on experimental details such as brain area, species and vigilance state, one also observes evidence for asynchronous-irregular (AI) dynamics (Burns and Webb, 1976; Softky and Koch, 1993), oscillations (Gray *et al.*, 1989; Gray, 1994; Buzsaki, 2004), or strong fluctuations associated with criticality, bistability or up-and-down states (Breakspear, 2017; Priesemann *et al.*, 2009, 2013; Bellay *et al.*, 2015; Wilson, 2008; Stern *et al.*, 1997; Cossart *et al.*, 2003). These states differ not only in strength and structure of fluctuations, but also in synchrony among neurons, from uncorrelated to fully synchronized spiking. The observation of such a vast range of dynamic states is puzzling, considering that the dynamics of all networks presumably originate from similar single-neuron physiology and plasticity mechanisms.

One particular plasticity mechanism that regulates neural activity on a long time scale is homeostatic plasticity (Turrigiano *et al.*, 1998; Lissin *et al.*, 1998; O'Brien *et al.*, 1998; Turrigiano and Nelson, 2004; Davis, 2006; Williams *et al.*, 2013). Homeostatic plasticity can be implemented by a number of physiological candidate mechanisms, such as redistribution of synaptic efficacy (Markram and Tsodyks, 1996; Tsodyks and Markram, 1997), synaptic scaling (Turrigiano *et al.*, 1998; Lissin *et al.*, 1998; O'Brien *et al.*, 1998; Fong *et al.*, 2015), adaptation of membrane excitability (Davis, 2006; Pozo and Goda, 2010),

or through interactions with glial cells (De Pittà *et al.*, 2016; Virkar *et al.*, 2016). All these mechanisms have in common that they implement a slow negative feedback-loop in order to maintain a certain target spike rate and stabilize network dynamics. In general, they reduce (increase) excitatory synaptic strength or neural excitability if the spike rate is above (below) a target rate, allowing compensation against a potentially unconstrained positive feedback-loop through Hebbian-type plasticity (Bienenstock *et al.*, 1982; Miller and MacKay, 1994; Abbott and Nelson, 2000; Turrigiano and Nelson, 2000; Zenke *et al.*, 2013; Keck *et al.*, 2017; Zenke *et al.*, 2017). Recent results highlight the involvement of homeostatic plasticity in generating robust yet complex activity dynamics in recurrent networks (Naude *et al.*, 2013; Hellyer *et al.*, 2016; Gjorgjieva *et al.*, 2016).

To understand the physiological mechanisms behind this large set of dynamic states, different model networks have been proposed that reproduce one or a set of states. To name a few examples, deafferentiation in combination with homeostatic scaling can generate bursts (Frohlich *et al.*, 2008); the interplay between excitation and inhibition may lead to oscillations, synchronous-regular activity, or asynchronous-irregular activity (Wilson and Cowan, 1972; Vogels *et al.*, 2005; Fries *et al.*, 2007; Brunel, 2000), where switching between dynamic states can be induced by varying the input (Brunel, 2000; Lerchner and Latham, 2015; Muñoz, 2018); synaptic facilitation and depression promote regular and irregular network dynamics (Tsodyks *et al.*, 1998; Levina *et al.*, 2007, 2009a); plasticity at inhibitory synapses can stabilize irregular dynamics (Vogels *et al.*, 2011; Effenberger *et al.*, 2015), whereas specific types of structural (Bornholdt and Rohlf, 2000; Tetzlaff *et al.*, 2010) or synaptic (Levina *et al.*, 2007, 2009a; de Arcangelis *et al.*, 2006; Bonachela *et al.*, 2010; Costa *et al.*, 2015; Michiels van Kessenich *et al.*, 2016; Campos *et al.*, 2017; Hernandez-Urbina and Herrmann, 2017; Millman *et al.*, 2010; Muñoz, 2018; Del Papa *et al.*, 2017) plasticity foster strong temporal fluctuations characteristic for a critical state; last but not least, homeostasis is necessary to achieve stable dynamics in recurrent networks with spike-

timing dependent plasticity (STDP) or Hebbian-type synaptic plasticity (e.g. Lazar, 2009; Litwin-Kumar and Doiron, 2014; Zenke *et al.*, 2015; Miner and Triesch, 2016; Tosi and Beggs, 2017; Keck *et al.*, 2017). Overall, the dynamic state depends on all aspects: single-neuron properties, synaptic mechanisms, network topology, plasticity rules, and input characteristics. Recalling that the single-neuron properties, synaptic mechanisms, as well as plasticity rules are presumably very similar *in vitro* and *in vivo*, these factors are unlikely to explain the observed differences.

In this study, we propose that the input strength is the defining difference between *in vitro* and *in vivo* dynamics. *In vitro* systems are completely isolated, whereas *in vivo* networks receive continuous input from sensory modalities and other brain areas. Under these different conditions, we propose that homeostatic plasticity is a sufficient mechanism to promote self-organization to a diverse set of dynamic states by mediating the interplay between external input rate and neural target spike rate. Treating the external input as a control parameter in our theoretical framework, allows us to alter the network dynamics from bursting, to fluctuating, to irregular. Thereby, our framework offers testable predictions for the emergence of characteristic but distinct network activity *in vitro* and *in vivo*.

Based on our theory, we derive explicit experimental predictions and implications: (1) The direct relation between dynamic state, spike rate and input rate enables us to quantify the amount of input the neural network receives, e.g., in mildly anesthetized cat V1, we estimate an input rate of $\mathcal{O}(0.01\,\text{Hz/neuron})$. (2) This implies that about 2% of cortical activity in cat V1 are imposed by the input, whereas 98% are generated by recurrent activation from within the network. (3) Our results suggest that one can alter the dynamic state of an experimental preparation by altering the input strength. Importantly, we predict for *in vitro* cultures that increasing the input rate to about $\mathcal{O}(0.01\,\text{Hz/neuron})$ would be sufficient to induce *in vivo*-like dynamics.

4.2 Experimental observations

To demonstrate characteristic neural activity *in vitro* and *in vivo*, we analyzed exemplary recordings of spiking activity. Data sets included cultures of dissociated cortical neurons (Wagenaar *et al.*, 2006; Wagenaar), as well as hippocampus of foraging rats (Mizuseki *et al.*, 2009b,a) and visual cortex of mildly anesthetized cats (Blanche and Swindale, 2006; Blanche, 2009) (see Sec. C.1 for details). Note that all preparations were inevitably subsampled, as spikes were recorded only from a small number of all neurons. For illustrative purposes, we focus on the average (subsampled) spiking activity a_t and the (subsampled) avalanche-size distribution P_{sub} (see Appendix C.2 for details).

The spiking activity *in vitro* shows bursting behavior (Fig. 4.1), i.e., stretches of very low activity interrupted by periods of synchronized activity. The subsampled avalanche-size distributions $P_{\text{sub}}(s)$ exhibits partial power-law behavior resembling $P(s) \sim s^{-3/2}$ as expected from a critical branching process (Harris, 1963), and conjectured for the synchronous-irregular regime (Touboul and Destexhe, 2017). However, in addition $P_{\text{sub}}(s)$ also shows a peak at large avalanche sizes, which may indicate either finite-size effects, supercriticality, or characteristic bursts (Levina and Priesemann, 2017).

In contrast, the spiking activity *in vivo* shows fluctuating dynamics (Fig. 4.1). These have been described as reverberating dynamics, a dynamic state between critical and irregular dynamics (Wilting and Priesemann, 2018a), characterized by a finite autocorrelation time of a few hundred milliseconds. The subsampled avalanche-size distributions $P_{\text{sub}}(s)$ can be approximated by a power-law for small s but show a clear exponential tail. The tails indicate slightly subcritical dynamics (Priesemann *et al.*, 2014), especially because deviations in the tails are amplified under subsampling (Priesemann *et al.*, 2009, 2013; Levina and Priesemann, 2017).

In sum, the spiking activity and the corresponding avalanche-size distributions clearly differ between *in vitro* and *in vivo* recordings. Remarkably, however, the average neural firing
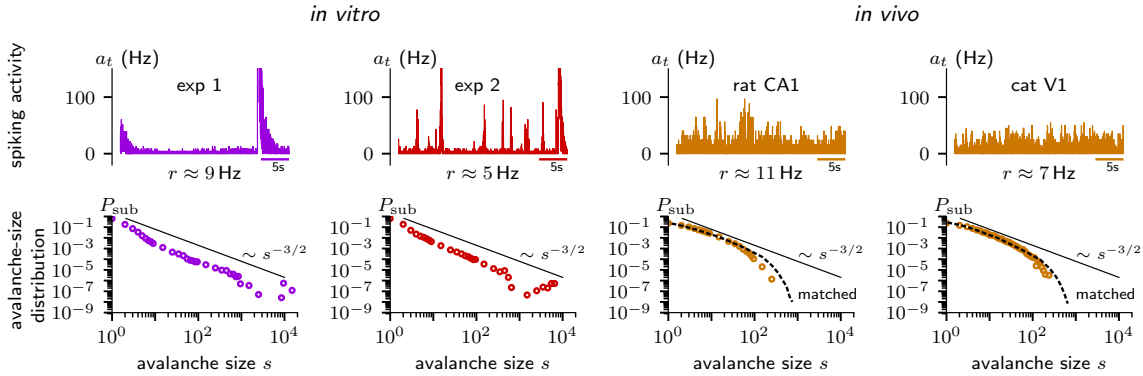
Figure 4.1: Examples of dynamic states observed in experiments. *In vitro* spike recordings are from cultures of dissociated rat cortical neurons (Wagenaar *et al.*, 2006). *In vivo* recordings are from right dorsal hippocampus in an awake rat during an open field task (Mizuseki *et al.*, 2009b) and from primary visual cortex in a mildly anesthetized cat (Blanche, 2009). Top row shows population spiking activity (Sec. C.2.1) from 30–60 single or multi-units ($\Delta t = 4$ ms) with average neural firing rate r; bottom row shows subsampled avalanche-size distributions (Sec. C.2.2). Solid lines indicate the power-law behavior $s^{-3/2}$ expected for a critical branching process. Dashed lines correspond to distributions obtained from branching networks matched to the experiments of rat CA1 ($\tau = 2$ s, $r^* = 11$ Hz, $h = 5.5 \times 10^{-3}$ Hz, $n = 31$, $N = 10^4$, $\Delta t = 1$ ms, $\tau_{\mathrm{hp}} = 10^5$ s) and cat V1 ($\tau = 0.2$ s, $r^* = 7$ Hz, $h = 3.5 \times 10^{-2}$ Hz, $n = 50$, $N = 10^4$, $\Delta t = 1$ ms, $\tau_{\mathrm{hp}} = 10^5$ s). For details and a definition of parameters see Sec. C.2.4.

rate r is similar across the different experimental setups.

4.3 Model

To investigate the differences between *in vitro* and *in vivo*, we make use of a branching network, which approximates properties of neural activity propagation. We extend the branching network by a negative feedback, which approximates homeostatic plasticity.

4.3.1 Branching network

In the brain, neurons communicate by sending spikes. The receiving neuron integrates its input, and if the membrane potential crosses a certain threshold, this neuron fires a spike itself. As long as a neuron does not fire, its time-varying membrane potential can be considered to fluctuate around some resting potential. In the following, we approximate the complex time-resolved process of action potential generation and transmission in a stochastic neural model with probabilistic activation.

Consider a network of size N. Each node

corresponds to an excitatory neuron, and spike propagation is approximated as a stochastic process at discrete time steps Δt. If a neuron, described by the state variable $s_{i,t} \in \{0, 1\}$, is activated, it spikes ($s_{i,t} = 1$), and immediately returns to its resting state ($s_{i,t+1} = 0$) in the next time step, unless activated again. Furthermore, it may activate post-synaptic neurons j with probability $p_{ij,t} = w_{ij}\alpha_{j,t}$, where $w_{ij} \in \{0, 1\}$ indicates whether two neurons are synaptically connected, and $\alpha_{j,t}$ is a homeostatic scaling factor. In addition, each neuron receives network-independent external input at rate h_i, representing external input from other brain areas, external stimuli, and importantly also spontaneous spiking of single neurons generated independently of pre-synaptic spikes (e.g. by spontaneous synaptic vesicle release, Kavalali (2014); Lenz *et al.* (2015)). The uncorrelated external input homogeneously affects the network at rate $h_i = h$, modeled as Poisson processes with an activation probability $1 - e^{-h\Delta t} \simeq h\Delta t$.

This model can be treated in the framework of a branching process (Harris, 1963), a linear process with a characteristic autocorrelation time τ (see below). The population activity is charac-

terized by the total number of spiking neurons, $A_t = \sum_{i=1}^{N} s_{i,t}$. Each spike at time t generates on average m postsynaptic spikes at time $t + 1$ such that on average $\mathbb{E}(A_{t+1}|A_t) = mA_t$, where m is called *branching parameter*. The branching parameter can be defined for each neuron individually: neuron i activates on average

$$m_{i,t} = \sum_{j=1}^{N} w_{ij} \alpha_{j,t} \qquad (4.1)$$

of its post-synaptic neurons (Haldeman and Beggs, 2005). This local branching parameter $m_{i,t}$ thus quantifies the impact of a spike in neuron i on the network. The network average (denoted in the following with a bar) of $m_{i,t}$ generates the (time-dependent) network branching parameter (Millman *et al.*, 2010)

$$\overline{m}_t = \frac{1}{N} \sum_{i=1}^{N} m_{i,t}. \qquad (4.2)$$

The external input generates on average $Nh\Delta t$ additional spikes per time step, resulting in a driven branching process (Heathcote, 1965; Pakes, 1971). The expected activity at time $t + 1$ is then $\mathbb{E}(A_{t+1}|A_t) = mA_t + Nh\Delta t$. For $m < 1$ the process is called *subcritical*, meaning that individual cascades of events will eventually die out over time. In this case, the temporal average (denoted in the following as $\langle \cdot \rangle$) of network activity A_t converges to a stationary distribution with average activity

$$\langle A \rangle = \frac{1}{T} \sum_{t=1}^{T} A_t \xrightarrow[T \to \infty]{} \frac{Nh\Delta t}{1 - m}. \qquad (4.3)$$

Considering a homogeneous neural spike rate $r_i = r = \langle A \rangle / N\Delta t$ this implies

$$r = \frac{h}{1 - m}. \qquad (4.4)$$

A constant mean spike rate r, which can be considered a biological constraint, is thus realized by adjusting either $m \in [0, 1)$ or $h \in [0, \infty)$.

The subcritical branching process ($m < 1$) is stationary with the autocorrelation function $C(l) = m^l$. The autocorrelation time can be identified by comparing with an exponential decay

$C(l) = e^{-l\Delta t/\tau}$, yielding (Wilting and Priesemann, 2018a)

$$\tau = -\Delta t / \ln(m), \qquad (4.5)$$

which diverges as $m \to 1$. At this divergence ($m = 1$) the process is *critical* and the activity A_t grows linearly in time with rate h. At criticality, assuming $h \to 0$, the number of events s in an avalanche triggered by a single seed event, is distributed according to a power law $P(s) \sim s^{-3/2}$ (Harris, 1963). For a non-vanishing h in the subcritical regime ($m < 1$), the avalanche-size distributions show a rapid decay, if they can be measured at all under persistent activity (Sec. C.2.2). Finally, for $m > 1$, the process is called *supercritical* and A_t can in principle grow to infinity. For a finite network, this of course is not possible and will manifest in a peak of the avalanche-size distribution at large avalanche sizes.

For the computational model, we consider a network of $N = 10^4$ neurons, which represents the size of *in vitro* cultures and *in vivo* cortical hypercolumns. The time step of the branching process has to reflect the causal signal propagation of the underlying physiological network. Since realistic propagation times of action potentials from one neuron to the next range from $1\,\mathrm{ms}$ to $4\,\mathrm{ms}$, we choose $\Delta t = 1\,\mathrm{ms}$. We consider three network topologies:

Directed Erdős-Rényi (ER) network: As a standard model of network topology, we consider a network with random directed connections. Each connection $w_{ij} = 1$ is added with probability p_{con}, excluding self-connections (i, i). Then, the degree distribution of outgoing as well as incoming connections follows a binomial distribution with average degree $\overline{k} = p_{\mathrm{con}}(N - 1) \simeq p_{\mathrm{con}}N$. We require $p_{\mathrm{con}} > \ln(N)/N$ to ensure that the graph is connected (Erdos and Rényi, 1960). The connectivity matrix w_{ij} is fixed throughout each simulation, such that averaging over simulations with different network realizations results in a quenched average. A cutout from an example graph is shown in Fig. 4.2**a**.

Spatially-clustered (SC) network: In order to consider a more detailed topology with dominant
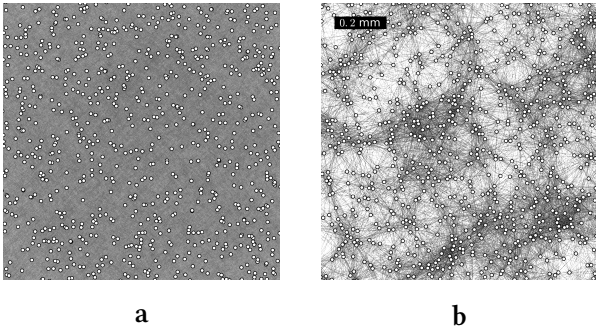
a   **b**

Figure 4.2: Cutouts of two random network topologies. (a) Subset of randomly spaced nodes in an Erdős-Rényi (ER) network with $p = 10^{-3}$. Note that connections cross the window also from neurons outside of the field of view, such that single connections cannot be distinguished visually in the sketch. (b) $1.4 \times 1.4\,\text{mm}^2$ subset of spatially-clustered (SC) topology generated by axonal-growth rules (Orlandi *et al.*, 2013; Hernández-Navarro *et al.*, 2017).

short-range connections, we follow Orlandi *et al.* who developed a model based on experimental observations of *in vitro* cultures (Orlandi *et al.*, 2013; Hernández-Navarro *et al.*, 2017). Neural somata are randomly placed as hard discs with radius $R_s = 7.5\,\mu\text{m}$, to account for the minimal distance between cell centers, on a $5 \times 5\,\text{mm}^2$ field. From each soma an axon grows into a random direction with a final length l given by a Rayleigh distribution $p(l) = (l/\sigma_l^2)\,\exp(-l^2/2\sigma_l^2)$ with $\sigma_l = 900\,\mu\text{m}$ and average axonal length $\bar{l} \simeq 1.1\text{mm}$. The axonal growth is a semiflexible path with segments of size $\Delta l = 10\,\mu\text{m}$ and orientation drawn from a Gaussian distribution relative to the previous segment with $\sigma_\theta = 15°$. A connection with another neuron is formed with probability 1/2 if the presynaptic axon intersects with the dendritic tree of a postsynaptic neuron (Wen *et al.*, 2009). The dendritic tree is modeled as a disc around the postsynaptic neuron with radius R_d drawn from a Gaussian distribution with mean $\bar{R}_d = 300\,\mu\text{m}$ and $\sigma_d = 20\mu\text{m}$. A cutout from an example graph is shown in Fig. 4.2**b**.

Annealed-average (AA) network: We consider in addition a network with k dynamically changing random connections (annealed average). The connections are distinguishable, exclude self-connections, and are redrawn every

time step. This model approximates the otherwise numerically expensive fully connected network (ER with $p_{\text{con}} = 1$) with a global m_t by choosing $\alpha_{j,t} = m_t/k$. In practice, we chose $k = 4$, which produces analogous dynamics to the fully-connected ($\bar{k} \approx 10^4$) network as long as $m_t < 4$.

Error bars are obtained as statistical errors from the fluctuations between independent simulations, which includes random network realizations $\{w_{ij}\}$ for ER and SC.

4.3.2 Homeostatic plasticity

In our model, homeostatic plasticity is implemented as a negative feedback, which alters the synaptic strength on the level of the postsynaptic neuron (the homeostatic scaling factor $\alpha_{j,t}$) to reach a target neural firing rate r_j^*. We consider a linear negative feedback with time constant τ_{hp}, which depends solely on the (local) activity of the postsynaptic neuron $s_{j,t}$

$$\Delta \alpha_{j,t} = (\Delta t\, r_j^* - s_{j,t})\left(\frac{\Delta t}{\tau_{\text{hp}}}\right), \qquad (4.6)$$

i.e., adapting a neuron's synaptic strength does not rely on information about the population activity A_t. Since $\alpha_{j,t}$ is a probability, we additionally demand $\alpha_{j,t} \geqslant 0$. Equation (4.6) considers homeostatic plasticity to directly couple to all postsynaptic synapses of any given neuron j. This can be implemented biologically as autonomous synaptic processes or somatic processes, such as translation and transcription. In order to further reduce complexity, we assume a uniform target rate $r_j^* = r^*$, while in fact experiments show a broad (log-normal) spike-rate distribution (Buzsáki and Mizuseki, 2014; Hengen *et al.*, 2016). Preliminary tests for a log-normal target rate distribution in ER networks ($p_{\text{con}} = 0.1$) showed consistent results. In our simulations, we typically consider a biologically motivated target rate $r^* = 1\,\text{Hz}$ and a homeostatic timescale of the order of an hour, $\tau_{\text{hp}} = 10^3\,\text{s}$.

4.4 Results

Including homeostatic plasticity in our model generates a broad range of dynamic states, depending on the external input. Figure 4.3
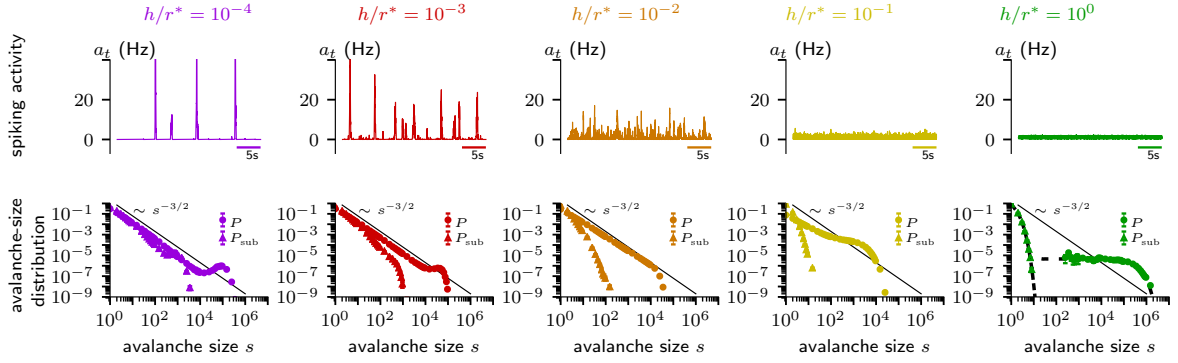
FIGURE 4.3: Homeostatic plasticity induces diverse dynamic states by regulating recurrent network interactions, mediating input rate h and target neural firing rate r^*. The annealed-average network reproduces bursting ($m > 1$, $h/r^* \leqslant 10^{-3}$, purple-red), fluctuating ($m \approx 0.99$, $h/r^* \approx 10^{-2}$ and $m \approx 0.9$, $h/r^* \approx 10^{-1}$, orange-yellow), and irregular ($m \approx 0$, $h/r^* = 1$, green) dynamics. The top row shows examplary spiking activity $a_t = A_t/N\Delta t$ (Sec. C.2.1); the bottom row shows avalanche-size distributions $P(s)$ ($n = N$, circles) and subsampled avalanche-size distributions $P_{\text{sub}}(s)$ ($n = 100$, triangles) averaged over 12 independent simulations (Sec. C.2.2). Solid lines show the power-law distribution $P(s) \propto s^{-3/2}$ (Harris, 1963), dashed lines show the analytical avalanche-size distribution of a Poisson process (Priesemann and Shriki, 2018). Parameters: $N = 10^4$, $\tau_{\text{hp}} = 10^3$ s, $r^* = 1$ Hz, $\Delta t = 1$ ms.

shows qualitatively representative results obtained for AA networks. For strong input ($h = \mathcal{O}(r^*)$), the network organizes itself into a dynamic state where neural firing is solely driven by the input fluctuations, resembling an asynchronous-irregular state (green). Here, temporal and pairwise spike count cross-correlations approach zero, and the avalanche-size distribution matches the analytic solution for a Poisson process (Priesemann and Shriki, 2018) shown as dashed lines. For weaker input ($h < r^*$) the system tunes itself towards fluctuating dynamics (orange-yellow). The average neural rate and sub-sampled avalanche-size distributions are qualitatively similar to reverberant *in vivo* dynamics with autocorrelation times of several hundred milliseconds (Fig. 4.1). In this regime, the temporal correlations increase when weakening the input, approaching close-to-critical dynamics, characterized by a power-law distribution $P(s) = s^{-3/2}$ (Harris, 1963), at the lower end of the regime. Decreasing the input even further ($h \ll r^*$) leads to bursting behavior, characterized by silent periods which are interrupted by strong bursts. These bursts are apparent as peak in the avalanche-size distribution at large avalanche sizes (purple-red). In this regime, the network steadily increases its synaptic strengths

during silent periods until a spike initiates a large burst, which in turn decreases the synaptic strengths drastically, and so on (cf. Sec. C.3). This regime captures the qualitative features of bursting *in vitro* dynamics (Fig. 4.1).

In the following, we derive a quantitative description of the three regimes sketched above. To quantify the dynamic state, we consider the temporal average of the branching parameter $m = \langle \overline{m} \rangle$, as well as the associated autocorrelation time τ of the population activity.

4.4.1 Mean-field solution

If we assume that τ_{hp} is sufficiently large (i.e. slow homeostatic plasticity), then $\Delta\alpha_j \approx 0$ and the dynamics of the network is fully determined by the approximately constant branching parameter $\overline{m}_t \approx m$. In this regime, (4.4) holds and combined with (4.5) and (4.6) we obtain the mean-field solution

$$m = 1 - h/r^* \quad \text{and} \quad \tau = -\Delta t/\ln(1 - h/r^*). \quad (4.7)$$

Hence, with decreasing input rate h, recurrent network activation (m) increases, i.e., perturbations cause a stronger network response and the autocorrelation time increases (Fig. 4.4, solid lines).
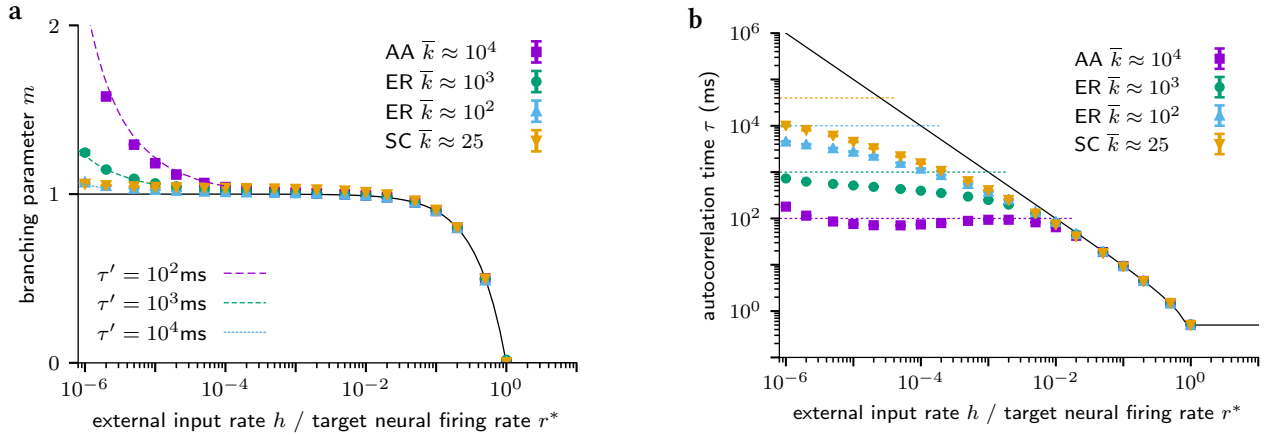
43

FIGURE 4.4: Quantitative distinction between dynamic states induced in neural networks of different topologies by homeostatic plasticity as a function of normalized input rate h/r^*. Data points are averages over 12 independent simulations ($N = 10^4$, $\tau_\alpha = 10^3$ s, $r^* = 1$ Hz, $\Delta t = 10^{-3}$ s) with connections generated according to annealed-average (AA), Erdős-Rényi (ER) or spatially-clustered (SC) topologies with average number of connections $\overline{k}$. Solid lines show the mean-field solution (4.7), dashed lines represent (semi-analytical) approximations of the bursting regime. **(a)** Branching parameter $m = \langle \overline{m} \rangle$ varies from irregular ($m \approx 0$), to fluctuating ($m \lessapprox 1$), to bursting ($m > 1$) dynamics. The behavior in the bursting regime strongly depends on the network timescale $\tau' = \tau_{hp}/\overline{k}$. **(b)** Integrated autocorrelation time of the network population activity (Sec. C.2.3) shows a crossover from irregular ($\tau = \mathcal{O}(\Delta t)$), over fluctuating ($\tau = -\Delta t / \ln[1 - h/r^*]$) to bursting ($\tau \approx \tau'$) dynamics.

In the light of this mean-field solution, we discriminate the three characteristic regimes as follows. First, we define the *input-driven regime* by $m \leqslant 0.5$ and $\tau \approx \Delta t$. Here, the network activity is dominated by input ($h = \mathcal{O}(r^*)$), and thus the dynamics follows the input statistics and becomes irregular. Second, we define the *fluctuating regime* for $0.5 < m < 1$ with a non-vanishing but finite autocorrelation time $\Delta t < \tau < \infty$. Here, the network maintains and amplifies input as recurrently generated fluctuations. In these two regimes the mean-field solution (4.7) matches numerical data on different network topologies (Fig. 4.4). Third, the mean-field solution predicts that in the limit $h \to 0$ the dynamics become critical with divergent autocorrelation time ($m \to 1$, $\tau \to \infty$). However, we observe a clear deviation from the mean-field solution, which defines the *bursting regime* with $m > 1$ and a finite autocorrelation time, as discussed below.

4.4.2 Bursting regime

Deviations from the mean-field solution (4.7) emerge when the assumption of "sufficiently large τ_{hp}" breaks down. We will derive a bound for τ_{hp}, below which the (rapid) homeostatic feedback causes notable changes of the network branching parameter $\overline{m}_t$ around its mean $m = \langle \overline{m} \rangle$, which in turn jeopardize the stability of the network dynamics.

To estimate the change of the network branching parameter, we first consider the change in local branching parameter $\Delta m_{i,t}$, which depends on each neurons out-degree $k_i = \sum_{j=1}^N w_{ij}$ and is given by

$$\Delta m_{i,t} = \sum_{j=1}^N w_{ij} \Delta \alpha_{j,t} = \left(k_i \Delta t\, r^* - \sum_{j=1}^N w_{ij} s_{j,t} \right) \left(\frac{\Delta t}{\tau_{hp}} \right).$$

On the network level, we make the assumption that the state of each neuron is approximated by the network average $s_{i,t} \approx A_t/N$, such that $\overline{\sum_{j=1}^N w_{ij} s_{j,t}} \approx \overline{k} \frac{A_t}{N}$. Then, the change in network

44

branching parameter can be approximated as

$$\Delta \overline{m}_t = \overline{\Delta m_t} \approx \left(\overline{k} \Delta t \, r^* - A_t \frac{\overline{k}}{N} \right) \left(\frac{\Delta t}{\tau_{\mathrm{hp}}} \right)$$
$$\approx \left(\Delta t \, r^* - \frac{A_t}{N} \right) \left(\frac{\Delta t}{\tau'} \right), \qquad (4.8)$$

where we have introduced an effective homeostatic network timescale $\tau' = \tau_{\mathrm{hp}}/\overline{k}$, for which (4.8) recovers the form of (4.6). Using τ' allows one to semi-analytically approximate the deviation of m from the mean-field solution (Fig. 4.4a, dashed lines, and Sec. C.3).

We next show that the stability of network dynamics requires the autocorrelation time of the dynamic process τ to be smaller than the timescale of homeostasis τ'. Stability demands that the homeostatic change in autocorrelation time $\Delta \tau$ is small compared to the autocorrelation time itself, i.e., $\Delta \tau \ll \tau$. We approximate $\Delta \tau$ by error propagation in (4.5), yielding

$$\Delta \tau \simeq \left| (\tau^2/\Delta t) e^{\Delta t/\tau} \right| \Delta m \simeq (\tau^2/\Delta t + \tau) \Delta m, \quad (4.9)$$

where we expanded the exponential for small $\Delta t/\tau$. For large τ, the leading term in (4.9) dominates and inserting (4.8) yields $\Delta \tau \simeq |\Delta t \, r^* - A_t/N| \left(\tau^2/\tau' \right)$. Thus, the dynamics can be described as a stationary branching process (mean-field solution) only as long as

$$\tau \ll \tau' \left| \Delta t \, r^* - A_t/N \right|^{-1}. \qquad (4.10)$$

Violation of (4.10) results in bursting behavior (Figs. 4.3 & C.1). For $A_t = \mathscr{O}(N)$ the right hand side of (4.10) is minimal, because $\Delta t \, r^* \ll 1$, which implies a maximal attainable autocorrelation time $\tau \simeq \tau' = \tau_{\mathrm{hp}}/\overline{k}$. This is in perfect agreement with the saturation of measured autocorrelation time in the bursting regime (Fig. 4.4b, dashed lines).

The transition from the fluctuating to the bursting regime occurs when the mean-field solution (4.7) equals the maximal attainable autocorrelation time, i.e., $\tau = -\Delta t/\ln(1 - h/r^*) \approx \tau'$. Hence, the transition occurs at $h/r^* \approx 1 - e^{-\Delta t/\tau'} \approx \Delta t/\tau'$. For even lower input rate, the dynamics become more and more bursty, and the avalanche-size distribution exhibits a peak at large avalanche sizes (Fig. 4.3 for $h/r^* < 10^{-2}$,

where $\tau' = 10^2$ ms, $\Delta t = 1$ ms). At the transition, the dynamics can be considered close-to-critical, because the (fully sampled) avalanche-size distribution is closest to a power-law with exponent $-3/2$.

4.4.3 Distributions of spiking activity

The different dynamical regimes imply characteristic distributions of neural network activity $P(a_t)$. Figure 4.5 shows an example of $P(a_t)$ for ER networks with $p_{\mathrm{con}} = 10^{-2}$, where the transition from fluctuating to bursting dynamics is expected for $h/r^* \approx \Delta t/\tau' = 10^{-4}$. In the irregular regime (green) $P(a_t)$ is a unimodal distribution. In the fluctuating regime (yellow-red), the peak in $P(a_t)$ shifts towards quiescence and the distribution develops a power-law tail with exponential cutoff, expected for a critical branching process. In the bursting regime (purple-blue), $P(a_t)$ is a bimodal distribution, reflecting network changes between quiescence and bursty activity. The position and sharpness of the high-activity maximum depend on the network connectivity and hence the heterogeneity in the single-neuron input.

4.4.4 Reproducing experimental results

Using the insight from our theory, we can reproduce experimental results. Spiking activity recorded *in vivo* resembles dynamics of the fluctuating regime. In this regime, the dynamic state is consistent for all topologies we considered (Fig. 4.4). Therefore, already a branching network on an AA topology suffices to quantitatively reproduce the avalanche-size distributions by matching model parameters with experimentally accessible estimates (Fig. 4.1 dashed lines). To match the branching network to recordings from cat V1 and rat CA1, we first estimated the spike rate r and autocorrelation time τ from the recordings of spiking activity (Wilting and Priesemann, 2018a); we then chose biologically plausible parameters for the network size N, the homeostatic timescale τ_{hp}, as well as the simulation time step Δt; and finally derived the external input h using (4.7) (for details see Sec. C.2.4). The resulting subsampled avalanche-size distri-
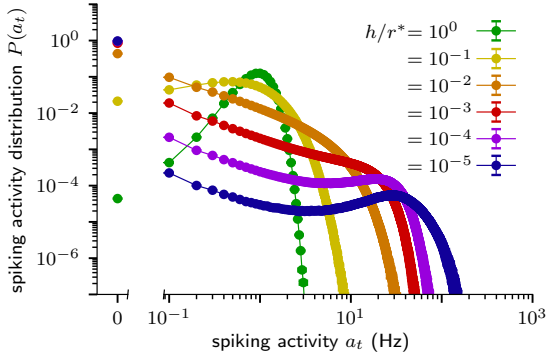
Figure 4.5: Distribution of spiking activity in weakly connected Erdős-Rényi networks ($p_{con} = 10^{-2}$, $\Delta t = 1\,ms$, $r^* = 1\,Hz$, $\tau' = 10^4\,ms$) averaged over 12 independent simulations. For irregular dynamics ($h/r^* \approx 10^0$) the distribution is clearly unimodal. For fluctuating dynamics ($10^{-4} < h/r^* < 10^0$) the distribution broadens and shifts the maximum towards quiescence. In addition towards the lower bound of the regime, the distribution develops a power-law tail with an exponential cutoff. At the crossover to bursting dynamics ($h/r^* \approx 10^{-4}$) the distribution becomes bimodal.

butions are in astonishing agreement with the experimental results, given the simplicity of our approach. Close inspection of the avalanche-size distribution for rat CA1 recordings still reveals small deviations from our model results. The deviations can be attributed to theta-oscillations in hippocampus, which result in subleading oscillations on an exponentially decaying autocorrelation function (Wilting and Priesemann, 2018a). While this justifies our approach to consider a single dominant autocorrelation time, theta oscillations slightly decorrelate the activity at short times and thereby foster premature termination of avalanches. Thus, the tail in the avalanche-size distribution is slightly shifted to smaller avalanche sizes (Fig. 4.1).

The *in vitro* results are qualitatively well matched by simulations in the bursting regime, with avalanche-size distributions showing a characteristic peak at large avalanche sizes (Fig. 4.3). It is difficult to quantitatively match a model to the data, because a number of parameters can only be assessed imprecisely. Most importantly, the autocorrelation time in the burst regime is not informative about the external in-

put rate h and depends on the average number of connections (Fig. 4.4). Likewise, the time-dependence of the branching parameter m_t cannot be assessed directly. Finally, system size and topology impact the network dynamics more strongly in this regime than in the fluctuating or input-driven regime. This yields a family of avalanche-size distributions with similar qualitative characteristics but differences in precise location and shape of the peak at large sizes.

4.5 Discussion

We propose the interplay of external input rate and target spike rate, mediated by homeostatic plasticity, as a neural mechanism for self-organization into different dynamic states (cf. sketch in Fig. 4.6). Using the framework of a branching process, we disentangled the recurrent network dynamics from the external input (e.g. input from other brain areas, external stimuli and spontaneous spiking of individual neurons). Our mean-field solutions, complemented by numeric results for generic spiking neural networks, show that for high input the network organizes into an input-driven state, while for decreasing input the recurrent interactions are strengthened, leading to a regime of fluctuating dynamics, resembling the reverberating dynamics observed *in vivo*. Decreasing the input further induces bursting behavior, known from *in vitro* recordings, due to a competition of timescales between homeostatic plasticity and the autocorrelation of population activity. Thereby our framework proposes a generic mechanism to explain the prominent differences between *in vivo* and *in vitro* dynamics.

Our theory suggests that also differences within the collective dynamic state observed *in vivo* can be explained by considering differences in input strength. For cortex, it was shown that layer 2/3 exhibits critical dynamics (Bellay *et al.*, 2015) and presumably deeper layers show reverberating dynamics (Priesemann *et al.*, 2014; Wilting and Priesemann, 2018a). We propose that this can be caused by different input strength: layer 2/3 is more recurrently connected, while layer 4 is the main target of thalamic input (Har-
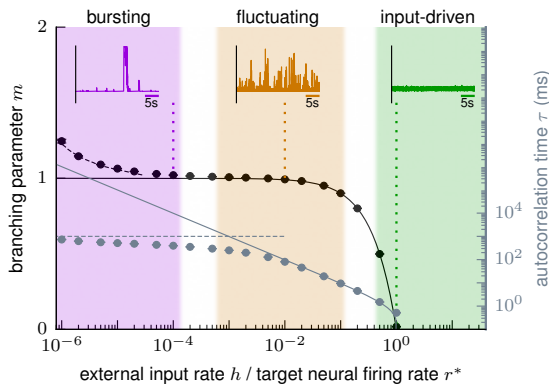
Figure 4.6: Sketch of bursting, fluctuating and input-driven network states, classified by the branching parameter and the autocorrelation time. We propose (solid lines) that homeostatic plasticity tunes the dynamic state depending on the ratio of external input rate (including spontaneous neural firing) and target neural firing rate. Data points and example activity traces stem from Erdős-Rényi networks ($N = 10^4$, $p = 10^{-1}$, $\tau_{hp} = 10^3$ s). In the bursting regime, the homeostatic timescale τ_{hp} influences the resulting dynamics (dashed lines).

ris and Mrsic-Flogel, 2013), hence receiving the stronger input. The dynamic state varies also across different cortical areas, where autocorrelation times of network activity reflect a hierarchical organization (Murray *et al.*, 2014; Hasson *et al.*, 2015): Cortical areas associated with higher-order function show a larger autocorrelation time. In the light of our results, a larger autocorrelation time implies less afferent input for the area in question. The hierarchical organization is further supported by our analysis of spiking activity *in vivo* (Fig. 4.1): the autocorrelation times in visual cortex ($\tau \approx 0.2$ s) and hippocampus ($\tau \approx 2$ s) precisely reflect that visual cortex is almost at the bottom, whereas hippocampus is at the top of the hierarchy of visual processing (Felleman and Van Essen, 1991).

Our theory provides an approach to experimentally infer the fraction of spikes generated recurrently within a network and generated by external input. For an average spike rate r, equation (4.7) implies $h/r = (1 − e^{−\Delta t/\tau})$. The external input rate can then be directly calculated from the autocorrelation time and by assuming a biologically plausible signal-propagation time, e.g.,

$\Delta t \approx 4$ ms. We estimate for recordings from visual cortex in mildly anesthetized cat that about 2% of the network activity is generated by the input, whereas the majority of 98% are generated recurrently within the network. From autocorrelation times measured across the cortical hierarchy (50 ms to 350 ms) in macaque monkey (Murray *et al.*, 2014), the fraction of spikes generated by external input decreases from $\sim 8\%$ to $\sim 1\%$ from lower to higher cortical areas. This is consistent with perturbation experiments in rat barrel cortex, where after triggering an extra spike the decay time of population rate was at least 50 ms (London *et al.*, 2010) indicating at most about 8% external input (for a detailed discussion see also Wilting and Priesemann (2018b)). Last, experiments on visual cortex of awake mice directly after thalamic silencing found a decay time of $\tau = 12(1)$ ms (Reinhold *et al.*, 2015), from which we would estimate about 70% recurrent activation. This is in perfect agreement with the experimentally measured 72(6)% of recurrent activation in the same study. This result thus validates our derived relation between h/r and τ.

One can interpret our findings in the light of up and down states (Wilson, 2008; Stern *et al.*, 1997; Holcman and Tsodyks, 2006; Millman *et al.*, 2010). Because the membrane potential was found to correlate with network activity (Cossart *et al.*, 2003; Vardi *et al.*, 2016), our results for the distribution of spiking activity in the bursting regime may correspond to the bimodal distributions of membrane potentials during up and down states (Fig 4.5). It has already been shown that negative feedback can stabilize up and down states (Holcman and Tsodyks, 2006; Millman *et al.*, 2010). In our theory, negative feedback leads to similar results in the low-input regime. Moreover, we predict that decreasing network input further, prolongs the quiescent periods or down states.

Our theory unifies previous numerical approaches of self-organization in neural networks, which typically considered a negative feedback mechanism but made very different choices on a (fixed) network input. For example, bursting dynamics have been generated by homeostatic build-up upon loss of network input (Frohlich

47

et al., 2008) or by self-organized supercriticality through dynamic neuronal gain (Costa et al., 2017). Adding weak input, self-organized criticality (Bak et al., 1987; Zapperi et al., 1995) has been achieved by local rewiring (Bornholdt and Rohlf, 2000; Tetzlaff et al., 2010) and synaptic depression (de Arcangelis et al., 2006; Levina et al., 2007, 2009a; Bonachela et al., 2010; Costa et al., 2015; Michiels van Kessenich et al., 2016; Campos et al., 2017; Hernandez-Urbina and Herrmann, 2017). In contrast, asynchronous-irregular network activity typically requires comparably strong input, assuming a balanced state (van Vreeswijk, 1996; Brunel, 2000; Renart et al., 2010), and a self-organized AI network state can be promoted by inhibitory plasticity (Vogels et al., 2011; Effenberger et al., 2015). While all these studies provide mechanisms of self-organization to one particular dynamic state, our theory highlights the role of input in combination with a negative feedback (Naude et al., 2013; Brunel, 2000; Lerchner and Latham, 2015; Muñoz, 2018) and provides a unifying mechanism of self-organization covering bursting, fluctuating and irregular dynamics.

From a broader perspective, we characterized driven systems with a negative feedback as a function of the input rate. The negative feedback compensates the input by regulating the system's self-activation to achieve a target activity. In this light of control theory, the bursting regime can be understood as resonances in a feedback loop, where feedback dynamics are faster than system dynamics (cf. Harnack et al. (2015)). This qualitative picture should remain valid for other connected graphs subject to external input with spatial and temporal correlations. In this case, however, we expect more complex network responses than predicted by our mean-field theory, which assumes self-averaging random networks subject to uncorrelated input.

Our results suggest that homeostatic plasticity may be exploited in experiments to generate *in vivo*-like dynamics in a controlled *in vitro* setup, in particular to abolish the ubiquitous bursts *in vitro*. Previous attempts to reduce bursting *in vitro* (Wagenaar, 2005) and in model-systems of epilepsy (Lian et al., 2003; Chiang et al., 2014; Co-

volan et al., 2014; Ladas et al., 2015) used short-term electrical and optical stimulation to attain temporal reduction in bursting. Alternatively, one can reduce bursting pharmacologically or by changing the calcium level, however, typically at the cost of changing single-neuron properties (Morefield et al., 2000; Shew et al., 2009; Penn et al., 2016). We propose a different approach, namely applying weak, global, long-term stimulation. Mediated by homeostasis, the stimulation should alter the effective synaptic strength, and thereby the dynamic state while preserving single-neuron dynamics. In particular, we predict that inducing in every neuron additional spikes with $h = \mathcal{O}(0.01\,\text{Hz})$ is sufficient to abolish the ubiquitous bursts *in vitro* and render the dynamics *in vivo*-like instead. If verified, this approach promises completely novel paths for drug studies. By establishing *in vivo*-like dynamics *in vitro*, fine differences between neurological disorders, which are otherwise masked by the ubiquitous bursts, can be readily identified. Altogether this would present a comparably cost-efficient, high-throughput, and well-accessible drug assay with largely increased sensitivity.

Acknowledgments

Chapter 5

Operating in a reverberating regime enables rapid tuning of network states to task requirements[†]

Abstract

Neural circuits are able to perform computations under very diverse conditions and requirements. The required computations impose clear constraints on their fine-tuning: a rapid and maximally informative response to stimuli in general requires decorrelated baseline neural activity. Such network dynamics is known as asynchronous-irregular. In contrast, spatio-temporal integration of information requires maintenance and transfer of stimulus information over extended time periods. This can be realized at criticality, a phase transition where correlations, sensitivity and integration time diverge. Being able to flexibly switch, or even combine the above properties in a task-dependent manner would present a clear functional advantage. We propose that cortex operates in a "reverberating regime" because it is particularly favorable for ready adaptation of computational properties to context and task. This reverberating regime enables cortical networks to interpolate between the asynchronous-irregular and the critical state by small changes in effective synaptic strength or excitation-inhibition ratio. These changes directly adapt computational properties, including sensitivity, amplification, integration time and correlation length within the local network. We review recent converging evidence that cortex *in vivo* operates in the reverberating regime, and that various cortical areas have adapted their integration times to processing requirements. In addition, we propose that neuromodulation enables a fine-tuning of the network, so that local circuits can either decorrelate or integrate, and quench or maintain their input depending on task. We argue that this task-dependent tuning, which we call "dynamic adaptive computation", presents a central organization principle of cortical networks and discuss first experimental evidence.

5.1 Introduction

Cortical networks are confronted with ever-changing conditions, whether these are imposed on them by a natural environment, or induced by the actions of the subjects themselves. For example, when a predator is lurking for a prey it should detect the smallest movement in the bushes anywhere in the visual field, but as soon as the prey is in full view and the predator moves to strike, visual attention should focus on the prey (Fig. 5.1C). Optimal adaptation for these changing tasks requires a precise and flexible adjustment of input amplification and other properties within the local, specialized circuits of primary visual cortex: strong amplification of small input while lurking, but quenching of any irrelevant input when chasing. These are changes from one task to another. However, even the processing within a single task may require the joint contributions of networks with diverse computational properties. For example, listening to spoken language involves the integration of phonemes at the timescale of milliseconds to words and whole sentences lasting for seconds. Such temporal integration might be realized by a hierarchy of temporal receptive fields, a prime example of adaption to different processing requirements of each brain area (Murray *et al.*, 2014; Hasson *et al.*, 2015, Fig. 5.1B).

Basic network properties like sensitivity, amplification, and integration timescale optimize different aspects of computation, and hence a generic input-output relation can be used to infer signatures of the computational properties, and changes thereof (Kubo, 1957; Wilting and Priesemann, 2018a). Throughout this manuscript, we refer to *computation* capability in the following two, high-level senses. First, the integration timescale determines the capability to process sequential stimuli. If small inputs are quenched away rapidly, the network may quickly be ready to process the next input. In contrast, networks that maintain input for long timescales may be slow at responding to novel input, but instead they can integrate information and input over extended time periods (Boedecker *et al.*, 2012; Del Papa *et al.*, 2017; Lazar, 2009;

Bertschinger and Natschläger, 2004). This is at the heart of reservoir computing in echo state networks or liquid state machines (Buonomano and Merzenich, 1995; Maass *et al.*, 2002; Jaeger and Haas, 2004; Schiller and Steil, 2005; Jaeger *et al.*, 2007; Boedecker *et al.*, 2012). Second, the detection of small stimuli relies on a sufficient amplification (Douglas *et al.*, 1995). However, increased sensitivity to weak stimuli can lead to increased trial-to-trial variability (Gollo, 2017).

These examples show that local networks that are tuned to one task may perform worse at a different one, and there is no one-type-fits-all network for every environmental and computational demand. How does a neural network manage to both react quickly to new inputs when needed, but also maintain memory of the recent input, e.g. when a human listens to language? Did the brain evolve a large set of specialized circuits, or did it develop a manner to fine-tune its circuits quickly to the computational needs? A flexible tuning of response properties would be desirable in the light of resource and space constraints. Indeed, in experiments one of the most prominent features of cortical responses is their strong dependence on cognitive state and context. For example, the cognitive state clearly impacts the strength, delay and duration of responses, the trial-to-trial variability, the network synchrony, and the cross-correlation between units (Kisley and Gerstein, 1999; Goard and Dan, 2009; Curto *et al.*, 2009; Marguet and Harris, 2011; Scholvinck *et al.*, 2015; Harris and Thiele, 2011; Kohn *et al.*, 2009; Poulet and Petersen, 2008; Massimini *et al.*, 2005; Priesemann *et al.*, 2013). Transitions between different cognitive states have been described by phase transitions (Steyn-Ross and Steyn-Ross, 2010; Steyn-Ross *et al.*, 2010; Galka *et al.*, 2010).

While a phase transitions can be very useful to realize cognitive state changes, we here want to emphasize a particular property of systems close to phase transitions: without actually crossing the critical point, already small changes in m can have a large impact on the network dynamics and function. Hence a classical phase transition may not be necessary for adaptation. In addition to the well-established phase transitions,

adaptation could be realized as a dynamic process that regulates the proximity to a phase transition and allows cortical networks to fine-tune their sensitivity, amplification, and integration timescale within one cognitive state, depending on the specific requirements. In order to allow efficient adaptation, cortical networks must evidently satisfy the following requirements. (i) The network properties are easily tunable to changing requirements, e.g. the required synaptic or neural changes should be small. (ii) The network is fully functional in its ground state, and also in the entire vicinity, i.e. the adaptive tuning does not destabilize or dysfunctionalize it. (iii) The network receives, modifies and transfers information according to its needs, e.g. it amplifies or quenches the input depending on task. (iv) The network's ground state in general should enable integration of input over any specific past window, as required by a given task.

We propose that cortex operates in a particular dynamic regime, the "reverberating regime", because in this regime small changes in neural efficacy can tune computational properties over a wide range – a mechanism that we propose to call *dynamic adaptive computation*. In this regime a cortical circuit can interpolate between two states described below, which both have been hypothesized to govern cortical dynamics and optimize different aspects of computation (Burns and Webb, 1976; Softky and Koch, 1993; Stein *et al.*, 2005; van Vreeswijk and Sompolinsky, 1996; Brunel, 2000; Beggs and Plenz, 2003; Beggs and Timme, 2012; Plenz and Niebur, 2014; Tkačik *et al.*, 2015; Humplik and Tkačik, 2017; Muñoz, 2018; Wilting and Priesemann, 2018a,b).

In the following, we recapitulate the computational properties of these two states and then identify recent converging evidence that in fact the reverberating regime governs cortical dynamics *in vivo*. We then show how specifically the reverberating regime can combine the computational properties of the two extreme states while maintaining stability and thereby satisfies all requirements for cortical network function postulated above. Finally, we outline future theoretical challenges and experimental predictions.

5.2 Hypotheses on cortical spiking dynamics

One hypothesis suggests that spiking statistics in the cortical ground state is asynchronous and irregular (Burns and Webb, 1976; Softky and Koch, 1993; Stein *et al.*, 2005), i.e. neurons spike independently of each other and in a Poisson manner (Fig. 5.1A). Such dynamics may be generated by a "balanced state", which is characterized by weak recurrent excitation compared to inhibition (van Vreeswijk and Sompolinsky, 1996; Brunel, 2000). The typical balanced state minimizes redundancy, has maximal entropy in its spike patterns, and supports fast network responses (Denève and Machens, 2016; van Vreeswijk and Sompolinsky, 1996). The other hypothesis proposes that neuronal networks operate at criticality (Bienenstock and Lehmann, 1998; Beggs and Plenz, 2003; Levina *et al.*, 2007; Beggs and Timme, 2012; Plenz and Niebur, 2014; Tkačik *et al.*, 2015; Humplik and Tkačik, 2017; Muñoz, 2018; Kossio *et al.*, 2018), and thus in a particularly sensitive state at a phase transition. This state is characterized by long-range correlations in space and time, and in models optimizes performance in tasks that profit from extended reverberations of input in the network (Bertschinger and Natschläger, 2004; Haldeman and Beggs, 2005; Kinouchi and Copelli, 2006; Wang *et al.*, 2011; Boedecker *et al.*, 2012; Shew and Plenz, 2013; Del Papa *et al.*, 2017). These two hypotheses, asynchronous-irregular and critical, can be interpreted as the two extreme points on a continuous spectrum of response properties to minimal perturbations, the first quenching any rate perturbation quickly within milliseconds, the other maintaining it for much longer. Hence the two hypotheses clearly differ already in the basic response properties they imply.

A general first approach to characterize the response properties of any dynamical system is based on linear response theory: When applying a minimal perturbation or stimulation, e.g. adding a single extra spike to neuron i, the basic response is characterized by m_i, the number of *additional* spikes triggered in all postsynaptic neurons (London *et al.*, 2010), which can be inter-
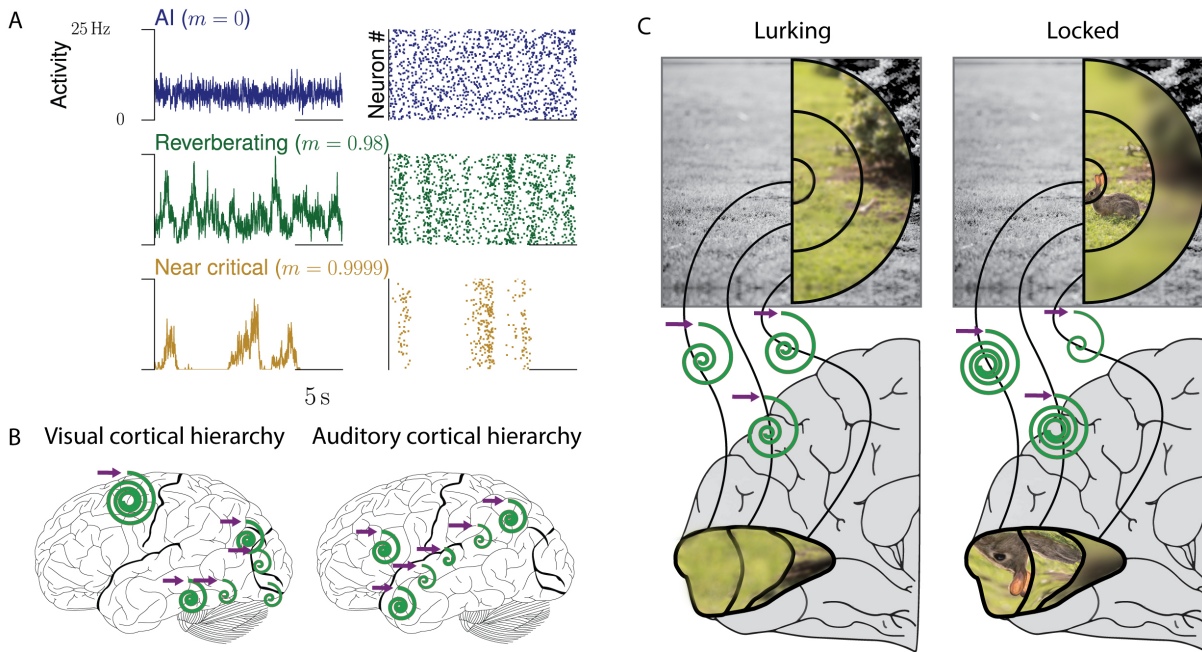
Figure 5.1: Collective dynamics of cortical networks. A. Examples of collective spiking dynamics representing either irregular and uncorrelated activity (blue), reverberations (green), or dynamics close to a critical state (yellow). Population spiking activity $a(t)$ and raster plots of 50 neurons are shown. **B.** Hierarchical organization of collective cortical dynamics. In primary sensory areas, input is maintained and integrated only for tens of milliseconds, whereas higher areas show longer reverberations and integration. The purple arrow represents any input to the respective area, the spirals the maintenance of the input over time (inspired from Hasson *et al.* (2015)). **C.** Dynamic adaptation of collective dynamics in local circuits. When a predator is lurking for prey, the whole field of view needs to be presented equally in cortex. Upon locking on prey, attention focuses on the prey. This could be realized by local adaptation of the network dynamics, which amplifies the inputs from the receptive fields representing the rabbit ("tuning in"), while quenching others ("tuning out").

preted as efficacy of the one neuron. If the efficacy is sufficiently homogeneous across neurons, then the average *neural efficacy m* represents a control parameter, and quantifies the impact of any single extra spike in a neuron on its postsynaptic neurons, and thus the basic network response properties to small input. In the next step, any of these triggered spikes in turn can trigger spikes in a similar manner, and thereby the small stimulation may cascade through the network. The network response may vary from trial to trial and from neuron to neuron, depending on excitation-inhibition ratio, synaptic strength, and membrane potential of the postsynaptic neurons. Thus m does not describe each single response, but the expected (average) response of the network, and thereby enables an assessment of the network's stability and computational properties. The magnitude of the neural ef-

ficacy m defines two different response regimes: If one spike triggers on average less than one spike in the next time step ($m < 1$), then any stimulation will die out in finite time. For $m > 1$, stimuli can be amplified infinitely, and $m = 1$ marks precisely the transition between stable and unstable dynamics (Figure 5.2B). In addition, m directly determines the amplification of the stimulus, the duration of the response, the intrinsic network timescale and the response variability, among others in the framework of autoregressive processes (Wilting and Priesemann, 2018a,b; Harris, 1963). Although details may depend on the specific process or model, many results presented in the following are qualitatively universal across diverse models that show a (phase) transition from stable to unstable, from ordered to chaotic, or from non-oscillatory to oscillatory activity. These include for example AR(1),
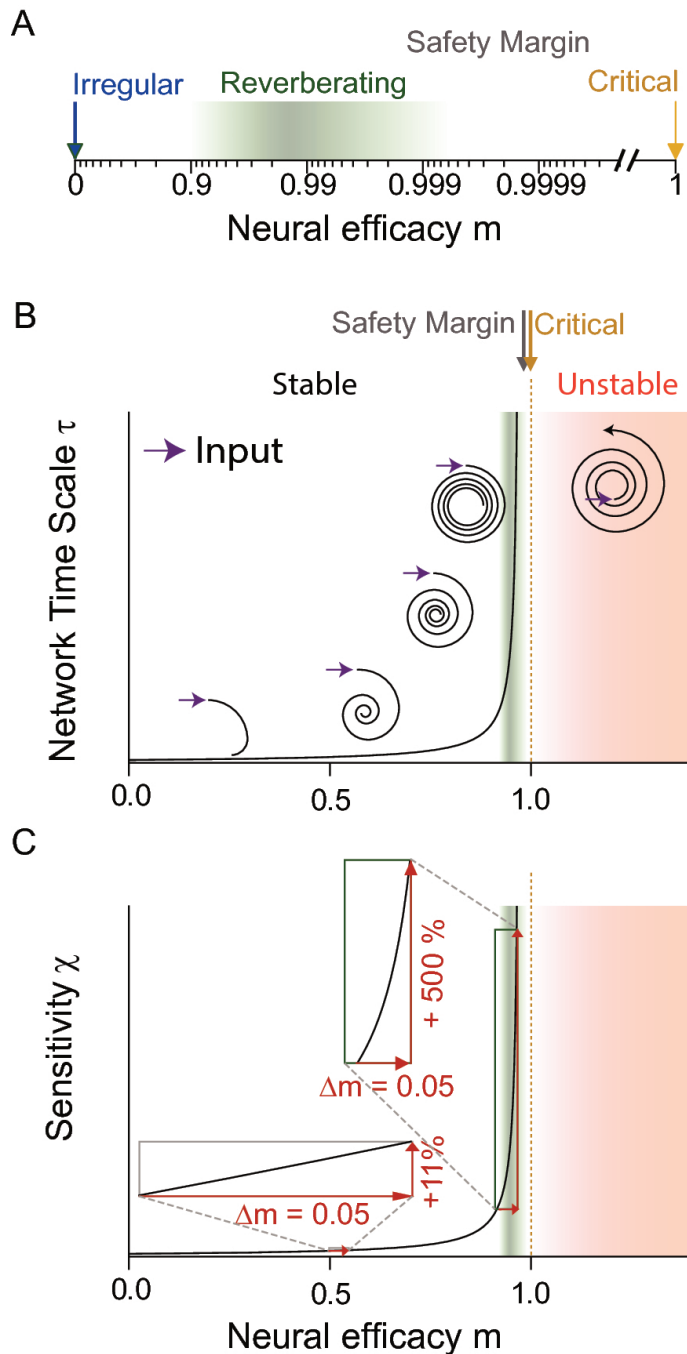
FIGURE 5.2: The neural efficacy m determines the average impact any spike has on the network. Depending on m, network dynamics can range from irregular ($m = 0$) to critical ($m = 1$) and unstable ($m > 1$) dynamics. **A.** In a logarithmic depiction of m, the "reverberating regime" (green) observed for cortex in vivo is well visible. It has clearly a larger m than the irregular state (blue), but maintains a safety margin to criticality (yellow) and the instability associated with the supercritical regime (red in **B** and **C**). **B,C.** Sketch to illustrate the divergence of dynamical and computational properties at a critical phase transition, at the example of the network timescale and the sensitivity, respectively. **B.** The network timescale determines how long input is maintained in the network. While any rate change is rapidly quenched close to the irregular state ($m = 0$), input "reverberates" in the network activity for increasingly long timescales when approaching criticality ($m = 1$). In the reverberating regime, the network timescale is tens to hundreds of milliseconds. For $m > 1$, input is amplified by the network, implying instability (assuming a supercritical Hopf bifurcation here for illustration). The reverberating regime keeps a sufficient safety margin from this instability. **C.** The reverberating regime found *in vivo* allows large tuning of the sensitivity by small changes of the neural efficacies (e.g. synaptic strength or excitation-inhibition balance), in contrast to states further away from criticality (insets).

Kesten, branching, and Ornstein-Uhlenbeck processes, systems that show a Hopf bifurcation, or systems at the transition to chaos (Huang and Doiron, 2017; Harris, 1963; Wilting and Priesemann, 2018a,b; Camalet *et al.*, 2000; Boedecker *et al.*, 2012). Hence the principle of dynamic computation detailed below can be implemented and exploited in very diverse types of neural networks, and thus presents a general framework.

The neural efficacy m is a statistical description of the effective recurrent activation, which takes into account both excitatory and inhibitory contributions. We here use it to focus on the mechanism of dynamic adaptive computation in an isolated setting, instead of including as many details as possible. We take this approach for two reasons. First, this abstraction allows to discuss possible generic principles for adaptation. Second, our approach enables us to assess the network state m and possible adaptation $m(t)$ from experiments. The abstract adaptation principles we consider here can be implemented by numerous physiological mechanisms, including top-down attention, adaptation of synaptic strengths or neuronal excitability, dendritic processing, disinhibition, changes of the local gain, or up-and-down states (Ramalingam *et al.*, 2013; Hirsch and Gilbert, 1991; London and Häusser, 2005; Karnani *et al.*, 2016; Piech *et al.*, 2013; Wilson, 2008).

Inferring the neural efficacy m experimentally is challenging, because only a tiny fraction of all neurons can be recorded with the required millisecond precision (Priesemann *et al.*, 2009; Ribeiro *et al.*, 2014; Levina and Priesemann, 2017). In fact, such spatial subsampling can lead to strong underestimation of correlations in networks, and subsequently of m (Wilting and Priesemann, 2018a,b). However, recently, a subsampling-invariant method has been developed that enables a precise quantification of m even from only tens of recorded neurons (Wilting and Priesemann, 2018a,b). Together with complementary approaches, either derived from the distribution of covariances or from a heuristic estimation, evidence is mounting that m is between $\approx$ 0.9 and $\approx$ 0.995, consistently for visual, somatosensory, motor and frontal cor-

tices as well as hippocampus (Wilting and Priesemann, 2018a,b; Priesemann *et al.*, 2014; Dahmen *et al.*, 2016). Hence, collective spiking activity is neither fully asynchronous nor critical, but in a *reverberating regime* between the two, and any input persists for tens to hundreds of milliseconds (Wilting and Priesemann, 2018a,b; Priesemann *et al.*, 2014; Murray *et al.*, 2014; Dahmen *et al.*, 2016; Hasson *et al.*, 2015). In more detail, both Dahmen and colleagues as well as Wilting & Priesemann estimated the neural efficacy to be about m = 0.98, ranging from about 0.9 to 0.995 when assessing spiking activity in various cortical areas (Figs. 5.1A, 5.2A). This magnitude of neural efficacy m implies intrinsic timescales of tens to hundreds of milliseconds. Such intrinsic timescales were directly estimated from cortical recordings in macaque by Murray and colleagues, who identified a hierarchical organization of timescales across somatosensory, medial temporal, prefrontal, orbitofrontal, and anterior cingulate cortex - indicating that every cortical area has adapted its response properties to its role in information processing (Murray *et al.*, 2014). These findings are also in agreement with the experiments by London and colleagues, who directly probed the neural efficacy by stimulating a single neuron in barrel cortex. They found the response to last at least 50 ms in that primary sensory cortex, implying $m \gtrsim 0.92$ (London *et al.*, 2010; Wilting and Priesemann, 2018a,b).

We will show that this reverberating regime with median m = 0.98 ($0.9 < m < 0.995$) observed *in vivo* is optimal in preparing cortical networks for flexible adaptation to a given task, and meets all the requirements for *dynamic adaptive computation* we postulated above.

5.3 Rapid adaption of functional properties

The reverberating regime allows tuning of computational properties by small, physiologically plausible changes of network parameters. This is because the closer a system is to criticality (m = 1), the more sensitive its properties are to small changes in the neural efficacy m, e.g. to any change of synaptic strength or excitation-

inhibition ratio (Fig.5.2C). In the reverberating regime the network can draw on this sensitivity: inducing small overall synaptic changes allows to adapt network properties to task requirements over a wide range. Assuming for example an AR(1), Ornstein-Uhlenbeck, or branching process, increasing m from 0.94 to 0.99 leads to a six-fold increase in the sensitivity of the network (Figure 5.2C). Here, the sensitivity $\partial r/\partial h \sim (1-m)^{-1}$ describes how the average network rate r responds to changes of the the input h (Wilting and Priesemann, 2018a). In contrast, the same absolute change from $m = 0.5$ to $m = 0.55$ only increases the sensitivity by about 11%. Similar relations apply to the network's amplification and intrinsic timescale (Wilting and Priesemann, 2018b). Thereby, any mechanism that increases or decreases the overall likelihood that a spike excites a postsynaptic neuron can mediate the change in neural efficacy m. Such mechanisms may act on the synaptic strength of many neurons in a given network, including neuromodulation that can change the response properties of a small population within a few hundred milliseconds, and homeostatic plasticity or long term potentiation or depression, which adapt the network over hours (Rang *et al.*, 2003; Turrigiano and Nelson, 2004; Zierenberg *et al.*, 2018). An alternative target could be the excitability of neurons, either rapidly by modulatory input, dendritic processing, disinhibition, changes of the local gain, or up-and-down states (Larkum, 2013; Ramalingam *et al.*, 2013; Hirsch and Gilbert, 1991; London and Häusser, 2005; Karnani *et al.*, 2016; Piech *et al.*, 2013; Wilson, 2008); or by changes of the intrinsic conductance properties over hours to days (Turrigiano *et al.*, 1994).

5.4 Retaining stability

While states even closer to the phase transition than $m = 0.98$ would imply even stronger sensitivity to changes in m, being too close to the phase transition comes with the risk of crossing over to instability ($m > 1$), because synapses are altered continuously by a number of processes, ranging from depression and facilitation to long-term plasticity. Hence, posing a system too close

to a critical phase transition may lead to instabilities and potentially causes epileptic seizures (Priesemann *et al.*, 2014; Meisel *et al.*, 2012; Wilting and Priesemann, 2018a). In the following, we estimate that the typical synaptic variability limits the precision of network tuning, and thereby defines an optimal regime for functional tuning that is about one percent away from the phase transition. Typically, single synapses exhibit about 50% variability in their strengths w, i.e. $\sigma_w \approx 0.5w$ over the course of hours and days (Statman *et al.*, 2014). If these fluctuations are not strongly correlated across synapses, the variance of the single neuron efficacy $m_i \approx k\langle w\rangle$ scales with the number k of outgoing synapses, $\mathrm{Var}[m_i] \approx k\,\mathrm{Var}[w]$ and gives

$$\sigma_{m_i} \approx 0.5\,m_i/\sqrt{k}. \qquad (5.1)$$

Assuming the network keeps a "safety margin" from instability ($m > 1$) of three standard deviations renders the network stable 99.9% of the time. The remaining, transient excursions into the unstable regime may be tolerable, because even in a slightly supercritical regime ($m \gtrsim 1$), runaway activity occurs only rarely (Harris, 1963; Zierenberg *et al.*, 2018). Thus assuming on average $k = \mathcal{O}(10,000)$ synapses per neuron (DeFelipe *et al.*, 2002) yields that a safety margin of about 1.5% from criticality is sufficient to establish stability. The safety margin can be even smaller if one assumes furthermore that the variability of m_i among neurons in a local network is not strongly correlated, because the stability of network dynamics is determined by the average neural efficacy $m = \langle m_i\rangle$, not by the individual efficacies. Furthermore, network structure might also contribute to stabilizing network activity (Kaiser and Simonotto, 2010). The resulting margin from criticality is compatible with the m observed *in vivo*.

The generic model reproduces statistical properties of networks where excitatory and inhibitory dynamics can be described by an effective excitation, e.g. because of a tight balance between excitation and inhibition (Sompolinsky *et al.*, 1988; van Vreeswijk and Sompolinsky, 1996; Ostojic, 2014; Kadmon and Sompolinsky, 2015; Huang and Doiron, 2017). In general, mod-

els should operate in a regime with $m < 1$ to maintain long-term stability and a safety-margin. However, transient and strong stimulus-induced activation is is key for certain types of computation, such as direction selectivity, or sub- and supra-linear summation, e.g. in networks with non-saturating excitation and feedback inhibition (Murphy and Miller, 2009; Lim and Goldman, 2013; Hennequin *et al.*, 2014; Rubin *et al.*, 2015; Hennequin *et al.*, 2017; Miller, 2016; Douglas *et al.*, 1995; Suarez *et al.*, 1995). These networks show transient instability (i.e. $m(t) > 1$) until inhibition stabilizes the activity. Whether such transient, large changes in $m(t)$ on a millisecond scale should be considered a "state" is an open question. Nonetheless, experimentally a time resolved $m(t)$ can be estimated with high temporal resolution, e.g. in experiments with a trial-based design. This measurement could then give insight into the state changes required for computation.

5.5 Balancing competing functional requirements

Besides the sensitivity, a number of other network properties also diverge or are maximized at the critical point and are hence equally tunable under dynamic adaptive computation. They include the spatial correlation length, amplification, active information storage, trial-to-trial variability, and the intrinsic network timescale (Sethna, 2006; Wilting and Priesemann, 2018a; Barnett *et al.*, 2013; Boedecker *et al.*, 2012; Harris, 1963). Some of these properties, which diverge at the critical point as $(1-m)^{-\beta}$ (with a specific scaling exponent β), are advantageous for a given task; others, in contrast, may be detrimental. For example, at criticality the trial-to-trial variability diverges and undermines reliable responses. Moreover, in the vicinity of the critical point convergence to equilibrium slows down (Scheffer *et al.*, 2012). Thus, network fine-tuning most likely is not based on optimizing one single network response property alone, but represents a trade-off between desirable and detrimental aspects. This tradeoff can be represented in the most simple case by a goal function

$$\Phi_\alpha = \Phi_+ - \alpha\Phi_- \propto (1-m)^{\beta_+} - \alpha'(1-m)^{\beta_-}, \quad (5.2)$$

which weighs the desired (Φ_+) and detrimental (Φ_-) aspects by a task dependent weight factor α, and might be called free energy in the sense of Friston (2010). Close to a phase transition, the desired and detrimental aspects diverge and depend on the critical scaling exponents β_+ and β_-. In this case, the normalization constants of Φ_+ and Φ_- are taken into account by the rescaled weight α'. Maximizing the goal function then yields an optimal neural efficacy m^*, which is here given by

$$m^* = 1 - \left(\frac{\alpha'\beta_-}{\beta_+}\right)^{-\frac{1}{\beta_+ - \beta_-}}. \quad (5.3)$$

This optimal neural efficacy (i) is in a subcritical regime unless $\alpha = 0$ (i.e. detrimental aspects do not matter) and (ii) depends on the weight α' and the exponents β_+ and β_-. In the simplified picture of branching processes, the resulting m^* determines a large set of response properties, which can thus only be varied simultaneously. An "ideal" network should combine the capability of dynamic adaptive computation with the ability to tune many response properties independently. To which extent such a network is conceivable at all and how it would have to be designed is an open question.

5.6 Fine-tuning temporal "windows of integration"

One particularly important network property is the intrinsic network timescale τ. In many processes this intrinsic network timescale emerges from recurrent activation and is connected to the neural efficacy as $\tau = -\Delta t/\log(m) \approx \Delta t/(1-m)$ (Wilting and Priesemann, 2018a), where Δt is a typical lag of spike propagation from the presynaptic to the postsynaptic neuron. Input reverberates in the network over this timescale τ, and can thereby enable short-term memory without any changes in synaptic strength (Fig. 5.2B). It has been proposed before that cortical computation relies on reverberating activity (Herz and Hopfield, 1995; Buonomano and Merzenich,

1995; Wang, 2002). Reverberations are also at the core of reservoir computing in echo state networks and liquid state machines (Maass *et al.*, 2002; Jaeger *et al.*, 2007; Boedecker *et al.*, 2012). Here, we extend on this concept and propose that cortical networks not only rely on reverberations, but specifically harness the reverberating regime in order to change their computational properties, in particular the specific τ, amplification, and sensitivity depending on needs.

We expect dynamic adaptive computation to fine-tune computational properties when switching from one task to the next, potentially mediated by neuromodulators, but we also expect that with development every brain area or circuit has developed computational properties that match their respective role in processing. Experimentally, evidence for a developmental or evolutionary tuning has been provided by Murray and colleagues, who showed that cortical areas developed a hierarchical organization as detailed above, with somatosensory areas showing fast responses ($\tau \approx 100$ms), and frontal slower ones ($\tau \approx 300$ms) (Murray *et al.*, 2014). This hierarchy indicates that the ground-state dynamics of cortical circuits is indeed precisely tuned, and it is hypothesized that the hierarchical organization provides increasingly larger windows for information integration for example across the visual hierarchy (Hasson *et al.*, 2008; Chen *et al.*, 2015; Hasson *et al.*, 2015; Badre and D'Esposito, 2009; Chaudhuri *et al.*, 2015). In addition to that backbone of hierarchical cortical organization, dynamic adaptive computation enables the fine-tuning of a given local circuit to specific task conditions. Indeed, experimental studies have shown that the response properties of cortical networks clearly change with task condition and cognitive state (Kisley and Gerstein, 1999; Goard and Dan, 2009; Curto *et al.*, 2009; Marguet and Harris, 2011; Scholvinck *et al.*, 2015; Harris and Thiele, 2011; Kohn *et al.*, 2009; Poulet and Petersen, 2008; Massimini *et al.*, 2005; Priesemann *et al.*, 2013). Relating these changes to specific functional task requirements remains a theoretical and experimental challenge for the future.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

All authors were involved in the conception and revision of this study. JW, JZ, and VP wrote the manuscript. JW, JZ, and VP calculated the presented derivations. JW, JPN, and VP drafted the figures.

Funding

Chapter 6

Discussion

This manuscript aimed at answering the following questions: (i) How can one assess the dynamical state of cortical networks event under strong subsampling? (ii) What is the dynamical state of cortical networks *in vivo*? (iii) How can one explain the striking differences between dynamical states *in vivo* and *in vitro*? (iv) How can the dynamical state support rapid adaptation of computational requirements to changing environments? Each of these questions was addressed in one of the previous chapters.

In Chap. 2, we showed that subsampling may lead to a severe underestimation of event propagation in complex systems. In our study, we explicitly demonstrated this underestimation in two different real-world applications, which have both been modeled in terms of propagating dynamics: epidemic spread of infectious diseases (Farrington *et al.*, 2003) and the propagation of spiking activity in neural networks (Beggs and Plenz, 2003; Haldeman and Beggs, 2005). In both systems, classical estimators underestimate event propagation. Beyond these two examples, our results may affect a multitude of further scientific disciplines, investigating, for example, cell proliferation, evolution (see (Kimmel and Axelrod, 2015) and references therein), neutron processes in nuclear power reactors (Pazy and Rabinowitz, 1973), spread of bank-ruptcy (Filimonov and Sornette, 2012) or evolution of stock prices (Mitov *et al.*, 2009).

We introduced a novel, multistep regression estimator (MR estimator) that overcomes the subsampling bias. This estimator is based on a set of multiple autoregressive analyses with increasing time lags. A subsampling-invariant estimate for event propagation can then be derived from the relative decay of the slopes of these autoregressive analyses. This method essentially corresponds to estimating the autocorrelation time of the underlying process, ignoring zero lag. In fact, we showed that (i) propagation dynamics is directly linked to a certain exponential autocorrelation time and (ii) this autocorrelation time is preserved under subsampling. The autocorrelation time is preserved, because the sampling process itself does not interfere with the underlying system evolution. Thereby, subsampling decreases the autocorrelation strength at different lags by the same factor, preserving the relative decay of the autocorrelation function.

The novel estimator overcomes limitations of previous estimators. In our stochastic framework, subsampling can essentially be understood as a filtering process. This idea of filtered observables has been used in Kalman filtering before. Expectation-Maximization based on Kalman filtering is state-of-the-art in system identification with noisy observations (Hamilton, 1994; Shumway and Stoffer, 1982; Ghahramani and Hinton, 1996). However, we showed that this method fails under strong subsampling, in contrast to our novel MR estimator. The MR estimator is more versatile for the estimation of subsampled propagation dynamics compared to methods based on Kalman filtering, because: (i) It puts no restrictions on the specific distributions of the underlying process, as long as it has an autoregressive representation. (ii) It puts very soft restrictions on the sampling filter and in particular is applicable even if the sampling filter is not Gaussian. (iii) The computational cost of multiple regressions is much smaller that Expectation-Maximization.

Our method complements other approaches to infer properties of collective dynamics from subsampled recordings. Levina and Priesemann (2017) used a similar, yet more restrictive stochastic description of subsampling. They constructed a scaling relation under subsampling, called subsampling scaling, which is different from the well-established finite-size scaling. Subsampling scaling provided a powerful method to identify criticality from avalanche size distributions, even under subsampling.

However, the extraction of avalanches from recordings of neural networks may be ambiguous. Subsequent avalanches are identified by pauses in the activity, but neural activity does not show a separation of timescales, which would allow for an unambiguous identification of avalanches. Instead, in awake cortex avalanches are continuously initialized, merge, and split up (Priesemann *et al.*, 2014). Hence, empty bins are only found if (i) the bin size tends towards zero, (ii) the system is heavily subsampled, or (iii) thresholding is applied to the activity (Del Papa *et al.*, 2017; Poil *et al.*, 2012; Priesemann *et al.*, 2014). This dependence on the experimental setting and analysis parameters renders the definition of avalanches ambiguous. As a consequence, the resulting avalanche distribution depends on the choice of the bin size, threshold, and the number of analyzed neurons or channels.

Our novel estimator overcomes this ambiguity, because it builds on a minimal model that does not require a separation of timescales. This minimal model of spike propagation is defined in terms of branching processes, which have been used to describe spiking dynamics in neural networks before (Beggs and Plenz, 2003; Haldeman and Beggs, 2005; Ribeiro *et al.*, 2010; Priesemann *et al.*, 2013, 2014). However, these previous models incorporated a separation of timescales, whereas we now explicitly include drive in the model. This allows the estimator to infer the dynamical state from ongoing neural activity. It can quantify the distance from criticality on a continuous spectrum (Wilting and Priesemann, 2018a), and is easy to apply to data: It requires only a few tens of multi-unit spike recordings, is subsampling-invariant, and does not require a

separation of timescales, or spike sorting.

Most real-world systems, including cortical dynamics, are more complicated than our minimal branching model. For cortical dynamics, for example, heterogeneity of neuronal morphology and function, non-trivial network topology, and the complexity of neurons themselves are likely to have a profound impact onto the population dynamics (Marom, 2010). However, we showed that *statistics* of cortical network activity are well approximated by a branching model. Therefore, we interpret branching models as a *statistical* approximation of spike propagation, which can capture a fair extent of the complexity of cortical dynamics. By using branching models, we draw on the powerful advantage of analytical tractability, which allowed for basic insight into dynamics and stability of cortical networks.

We used the novel estimator in order to infer propagation properties of collective spiking dynamics in cortical networks in Chap. 3. We identified a reverberating regime, which suggests that spiking activity *in vivo* is neither AI-like, nor consistent with a critical state. Instead, it is poised in a regime that, unlike critical or AI, does not maximize one particular property alone but may flexibly combine features of both (Wilting *et al.*, 2018). These results may resolve contradictions between AI and critical states. They suggest that network dynamics show AI-like statistics, because under subsampling the observed correlations are underestimated. In contrast, typical experiments that assessed criticality potentially overestimated correlations by sampling from overlapping populations (LFP, EEG) and thereby hampered a fine distinction between critical and subcritical states (Pinheiro Neto *et al.*, in prep).

The reverberating regime is supported by complementary approaches. First, Priesemann *et al.* (2014) and Dahmen *et al.* (2016) identified a similar reverberating regime, either derived from the distribution of covariances or from a heuristic estimation. Across all these studies, this regime is consistently found for visual, somatosensory, motor and frontal cortices as well as hippocampus (Wilting and Priesemann, 2018a, 2019b; Priesemann *et al.*, 2014; Dah-

men *et al.*, 2016). This reverberating regime implies intrinsic timescales, which allow any input to persist for tens to hundreds of milliseconds (Wilting and Priesemann, 2018a, 2019b; Priesemann *et al.*, 2014; Murray *et al.*, 2014; Dahmen *et al.*, 2016; Hasson *et al.*, 2015). Such intrinsic timescales were directly estimated from cortical recordings in macaque by Murray and colleagues, who identified a hierarchical organization of timescales across somatosensory, medial temporal, prefrontal, orbitofrontal, and anterior cingulate cortex - indicating that every cortical area has adapted its response properties to its role in information processing (Murray *et al.*, 2014). These findings are also in agreement with the experiments by London *et al.* (2010), who directly probed the neural efficacy by stimulating a single neuron in barrel cortex. They found the response to last at least 50 ms in that primary sensory cortex (London *et al.*, 2010; Wilting and Priesemann, 2018a, 2019b). Altogether, the generality of these results across different analysis techniques, species, brain areas, and cognitive states, suggests self-organization to this reverberating regime as a general organization principle for cortical network dynamics.

Using our analytically tractable model, we could link the dynamical state of the network to functional properties. Inferring the precise distance from criticality with our novel estimator allowed to predict and validate network properties, such as avalanche size and duration, interspike interval, or activity distributions for each network individually. These predictions agreed with experimental results. We hence deduced further properties, which are impossible or difficult to assess experimentally and gave insight into more complex questions about network responses: how do perturbations propagate within the network, and how susceptible is the network to external stimulation? First, we found that one additional spike elicits between 20 and 1000 additional spikes on average, thereby yielding characteristic amplification of small stimuli (Douglas *et al.*, 1995; Suarez *et al.*, 1995; London *et al.*, 2010; Miller, 2016). Second, we inferred that about 98% of cortical activity is generated by recurrent excitation, and only about 2% originates from in-

put or spontaneous threshold crossing (Reinhold *et al.*, 2015). This may indicate that a large proportion of spiking activity is generated in the context of predictive coding (Rao and Ballard, 1999; Clark, 2013). Third, we estimated that cortical spiking activity is self-similar over timescales between tens of milliseconds and a few seconds, indicating that stimuli are maintained in the activity and can be read out for short term memory over similar time spans (Buonomano and Merzenich, 1995; Wang, 2002; Jaeger *et al.*, 2007; Lim and Goldman, 2013; Murray *et al.*, 2014).

The reverberating regime may come with particular functional benefits (Wilting *et al.*, 2018; Wilting and Priesemann, 2019b). First, instead of solely maximizing singular properties, the reverberating regime can balance competing requirements, e.g. sensitivity vs. specificity (Gollo, 2017), quality of representation vs. integration time (Shriki and Yellin, 2016), or stimulus detection vs. discrimination (Tomen *et al.*, 2014; Clawson *et al.*, 2017). Second, the reverberating regime supports flexible adaption to task requirements, because small parameter changes can induce rapid, strong changes of computational properties (Wilting *et al.*, 2018). Third, the reverberating regime promotes amplification of small stimuli, while keeping a safety margin from the instability associated with supercriticality (Priesemann *et al.*, 2014). Last, it offers a substrate for short-term memory, as information about the input is maintained for well-defined time-spans ranging from tens of milliseconds to several seconds, which has been exploited in echo state networks (Maass *et al.*, 2002; Jaeger *et al.*, 2007; Boedecker *et al.*, 2012).

In Chap. 4 we proposed a mechanism to establish and maintain different dynamical states, ranging from asynchronous to the reverberating to critical to transiently supercritical states. Namely, we demonstrated that the interplay of external input rate and target spike rate, mediated by homeostatic plasticity, can promote self-organization into different dynamic states. Using the framework of driven branching processes, we disentangled the recurrent spike propagation from external input. We showed analytically and numerically that for high input the network or-

ganizes into an input-driven state. For decreasing input the recurrent interactions are strengthened and lead to a regime of fluctuating dynamics. This regime resembles the reverberating dynamics observed *in vivo*. Decreasing the input even further induced bursting behavior due to a competition of timescales between homeostatic plasticity and the autocorrelation of population activity. This bursting behavior resembles the bursting activity found in *in vitro* recordings. Thereby we proposed a generic mechanism to explain the prominent differences between *in vivo* and *in vitro* dynamics.

Our framework does not only explain differences between *in vivo* and *in vitro* dynamics, but suggests that also differences between collective dynamic states observed *in vivo* can be explained by differences in input strength. It was shown that cortical layer 2/3 exhibits critical dynamics (Bellay *et al.*, 2015), whereas presumably deeper layers show reverberating dynamics (Priesemann *et al.*, 2014; Wilting and Priesemann, 2018a). This difference agrees with our proposed framework: layer 2/3 is more recurrently connected, while layer 4 is the main target of thalamic input (Harris and Mrsic-Flogel, 2013), hence receiving the stronger input. The dynamic state also varies across different cortical areas: it was shown that intrinsic network timescales reflect a hierarchical organization (Murray *et al.*, 2014; Hasson *et al.*, 2015), where cortical areas associated with higher-order function show a larger autocorrelation time. As input presumably gets more and more filtered, i.e. weaker, from sensory modalities to higher order brain areas, this hierarchy aligns with our framework.

The hierarchy of intrinsic timescales provides first evidence for dynamic adaptive computation on an evolutionary scale. In Chap. 5 we proposed that cortex operates in the reverberating regime, particularly because in this regime small changes in neural efficacy can tune computational properties over a wide range. The hierarchy of intrinsic timescales indeed shows that with development every brain area or circuit has developed computational properties that match their respective role in processing, with somatosensory areas showing fast responses ($\tau \approx 100$ms),

and frontal slower ones ($\tau \approx 300$ms) (Murray *et al.*, 2014). This hierarchy indicates that the ground-state dynamics of cortical circuits is indeed precisely tuned, and it is hypothesized that the hierarchical organization provides increasingly larger windows for information integration for example across the visual hierarchy (Hasson *et al.*, 2008; Chen *et al.*, 2015; Hasson *et al.*, 2015; Badre and D'Esposito, 2009; Chaudhuri *et al.*, 2015). Hence, the hierarchy of intrinsic timescale provides evidence for dynamic adaptive computation on a developmental or evolutionary scale.

On top of this evolutionary backbone, we expect that dynamic adaptive computation modulates computational properties when switching from one task to the next. Thereby, dynamic adaptive computation enables the fine-tuning of a given local circuit to specific task conditions. One then expects that a modulation of the hierarchical backbone can be observed in various contexts, depending on vigilance state and task. Indeed, experimental studies have shown that the response properties of cortical networks clearly change with task condition and cognitive state (Kisley and Gerstein, 1999; Goard and Dan, 2009; Curto *et al.*, 2009; Marguet and Harris, 2011; Scholvinck *et al.*, 2015; Harris and Thiele, 2011; Kohn *et al.*, 2009; Poulet and Petersen, 2008; Massimini *et al.*, 2005; Priesemann *et al.*, 2013). There is also evidence that the power-law nature of avalanche size distributions changes with cognitive states, for example across the sleep-wake cycle (Ribeiro *et al.*, 2010; Priesemann *et al.*, 2013; Meisel *et al.*, 2017b,a) or under changes of consciousness (Tagliazucchi *et al.*, 2016; Bellay *et al.*, 2015; Fagerholm *et al.*, 2016, 2018; Fekete *et al.*, 2018; Lee *et al.*, 2019). Markers of criticality have been shown to transiently change depending on the behavioral state (Hahn *et al.*, 2017), attention (Fagerholm *et al.*, 2015), and stimuli (Arviv *et al.*, 2015; Yu *et al.*, 2017). These studies provided evidence that the brain is capable of flexible self-organization to different states. Precisely linking the changes in the reverberating regime to the specific task at hand remains a challenge for future work.

Where do we go from here? The series of publications presented in this monograph has laid

the foundation for consecutive research questions (Wilting and Priesemann, 2019a). Some of these are already being investigated in the work group of Viola Priesemann.

First, the subsampling problem illuminated in Chap. 2 is only one complication faced when assessing critical dynamics (Priesemann *et al.*, 2019). A different challenge is imposed by recordings from LFP, EEG, MEG, or BOLD, which are typically *coarse sampled*, i.e. the recorded activity represents the cumulative, filtered signal of many individual contributors. This coarse sampled signal will indeed produce the hallmark power laws in the distributions of avalanche sizes and durations if the underlying dynamics is critical (Yu *et al.*, 2014). However, the more interesting question is if power laws can even arise in these signals in the absence of criticality. Indeed, preliminary results by Pinheiro Neto *et al.* suggest that depending on the placement of electrodes, power laws can be found even if the underlying dynamics is not critical. Understanding the possible misinterpretations of coarse sampled avalanche size distributions may help shed light on the contradictory past results on the dynamical state of cortical network.

Second, the estimator presented in Chap. 2 is based on processes with an autoregressive representation, i.e. strictly speaking dimensionless processes. Strictly speaking, this class of processes represents mean field dynamics in an infinite population in the limit of vanishing activity. This representation is limited, as it (i) does not allow to explicitly model network topology and (ii) does not reflect *coalescence*, i.e. the fact hat two neurons might activate the same postsynaptic neuron simultaneously. In Chaps. 3 and 4, the dimensionless process was heuristically mapped onto a *branching network*, i.e. a network model with probabilistic activation. It is an open questions, how the two models map onto each other precisely, and if they experience the same type of phase transition at the critical point. These questions are addressed by Zierenberg *et al.* (2020b) and Zierenberg *et al.* (2020a).

Third, the establishment of a robust tool to infer the dynamical state motivates its application to experimental recordings. de Heuvel *et al.*

present an extension of the estimator, which allows to infer the dynamical state even from recordings with nonstationary input rate, as long as the experimental setup comprises similarly structured, repetitive trials. This extension is already included in an estimator toolbox (Spitzner *et al.*), which is at the heart of two experimental studies. Dehning *et al.* are probing a possible hierarchical organization of cortical areas, and the reorganization of this hierarchy under different states. Finding such a reorganization would be an indicator for the dynamic adaptive computation postulated in Chap. 5. Finally, Hagemann *et al.* are investigating possible systematic differences between epileptic and healthy hemispheres in spike recordings of epilepsy patients.

Overall, these studies demonstrate the usefulness of the multistep regression estimator, which can be helpful to analyze collective dynamics in cortical networks under a variety of experimental conditions. These studies represent steps towards a deeper understanding of collective cortical dynamics and its changes under different tasks. They form the basis for future research to refine this knowledge even further.

Appendix A

Supplementary material for "Inferring collective dynamical states from widely unobserved systems"[†]

A.1   Applicability of MR estimation

We here analytically derive the novel MR estimator for branching processes (BP) (Harris, 1963; Heathcote, 1965; Pakes, 1971). We expect that analogous derivations apply to any process with a first order autoregressive representation (PAR) (Ispány and Pap, 2010), because these processes fulfill Eq. (A.5). Beside BPs, PARs include autoregressive AR(1) processes, integer-valued autoregressive INAR(1) processes (Alzaid and Al-Osh, 1990) rounded integer-valued autoregressive RINAR(1) processes (Kachour and Yao, 2009), and Kesten processes (Kesten, 1973).

We emphasize that the MR estimator only requires the subsampled recording a_t of a system with full activity A_t conforming with the definition below. It is not necessary to know either the full system size, the number of subsampled units, nor any of the moments of the full process A_t.

A.2   Branching processes

In a branching process (BP) with immigration (Harris, 1963; Heathcote, 1965; Pakes, 1971) each unit i produces a random number $y_{t,i}$ of units in the subsequent time step. Additionally, in each time step a random number h_t of units immigrates into the system (drive). Mathematically, BPs are defined as follows (Harris, 1963; Heathcote, 1965): Let $y_{t,i}$ be independently and identically distributed non-negative integer-valued random variables following a law $\mathscr{Y}$ with mean $m = \langle \mathscr{Y} \rangle$ and variance $\sigma^2 = \mathrm{Var}[\mathscr{Y}]$. Further, $\mathscr{Y}$ shall be non-trivial, meaning it satisfies $\mathrm{P}[\mathscr{Y} = 0] > 0$ and $\mathrm{P}[\mathscr{Y} = 0] + \mathrm{P}[\mathscr{Y} = 1] < 1$. Likewise, let h_t be independently and identically distributed non-negative integer-valued random

variables following a law $\mathscr{H}$ with mean rate $h = \langle \mathscr{H} \rangle$ and variance $\xi^2 = \text{Var}[\mathscr{H}]$. Then the evolution of the BP A_t is given recursively by

$$A_{t+1} = \sum_{i=1}^{A_t} y_{t,i} + h_t, \tag{A.1}$$

i.e. the number of units in the next generation is given by the offspring of all present units and those that were introduced to the system from outside.

The stability of BPs is solely governed by the mean offspring m. In the subcritical state, $m < 1$, the population converges to a stationary distribution A_∞ with mean $\langle A_\infty \rangle = h/(1-m)$. At criticality ($m = 1$), A_t asymptotically exhibits linear growth, while in the supercritical state ($m > 1$) it grows exponentially. We will first show results that further specify the mean and variance of subcritical branching processes.

Theorem 1. *The stationary distribution of a subcritical BP satisfies*

$$\langle A_\infty \rangle = \frac{h}{1-m}, \qquad \text{Var}[A_\infty] = \frac{1}{1-m^2}\left(\xi^2 + \sigma^2 \frac{h}{1-m}\right),$$

where m, σ^2, h, and ξ^2 are defined as above.

Proof. The first result was stated before (Heathcote, 1965; Heyde and Seneta, 1972) and follows from taking expectation values of both sides of Eq. (A.1): $\langle A_{t+1} \rangle = m\langle A_t \rangle + h$. Because of stationarity $\langle A_{t+1} \rangle = \langle A_t \rangle = \langle A_\infty \rangle$ and the result follows easily. For the second result, observe that by the theorem of total variance, $\text{Var}[A_{t+1}] = \langle \text{Var}[A_{t+1}|A_t] \rangle + \text{Var}[\langle A_{t+1}|A_t \rangle]$, where $\langle \cdot \rangle$ denotes the expected value, and $A_{t+1}|A_t$ conditioning the random variable A_{t+1} on A_t. Because A_{t+1} is the sum of independent random variables, the variances also sum: $\text{Var}[A_{t+1}|A_t] = \sigma^2 A_t + \xi^2$. Using the result for $\langle A_\infty \rangle$ one then obtains

$$\text{Var}[A_{t+1}] = \xi^2 + \sigma^2 \frac{h}{1-m} + \text{Var}[mA_t + h] = \xi^2 + \sigma^2 \frac{h}{1-m} + m^2 \text{Var}[A_t]. \tag{A.2}$$

Again, in the stationary distribution $\text{Var}[A_{t+1}] = \text{Var}[A_t] = \text{Var}[A_\infty]$ and hence the stated result follows. $\square$

A.3 Subsampling

To derive the MR estimator for subsampled data, subsampling is implemented in a parsimonious way, according to the following definition:

Definition 1 (Subsampling). Let $\{A_t\}_{t\in\mathbb{N}}$ be a BP and $\{a_t\}_{t\in\mathbb{N}}$ a sequence of random variables. Then $\{a_t\}_{t\in\mathbb{N}}$ is called a subsampling of $\{A_t\}_{t\in\mathbb{N}}$ if it fulfills the following three conditions:

(i) Let $t', t \in \mathbb{N}$, $t' \neq t$. Then the conditional random variables* $(a_t|A_t = j)$ and $(a_{t'}|A_{t'} = l)$ are independent for any outcome $j, l \in \mathbb{N}$ of $A_t, A_{t'}$. If $A_t = A_{t'}$ then $(a_t|A_t = j)$ and $(a_{t'}|A_{t'} = j)$ are identically distributed.

(ii) Let $t \in \mathbb{N}$. Conditioning on a_t does not add further information to the process: The two random variables $(A_{t+1}|A_t = j, a_t = l)$ and $(A_{t+1}|A_t = j)$ are identically distributed for any $j, l \in \mathbb{N}$.

(iii) There are constants $\alpha, \beta \in \mathbb{R}$, $\alpha \neq 0$, such that $\langle a_t|A_t = j \rangle = \alpha j + \beta$ for all $t, j \in \mathbb{N}$.

Thus the subsample a_t is constructed from the full process A_t based on the three assumptions: (i) The sampling process does not interfere with itself, and does not change over time. Hence the realization of a subsample at one time does not influence the realization of a subsample at another time, and the conditional *distribution* of $(a_t|A_t)$ is the same as $(a_{t'}|A_{t'})$ if $A_t = A_{t'}$. However, even if $A_t = A_{t'}$, the subsampled a_t and $a_{t'}$ do not necessarily take the same value. (ii) The subsampling does not interfere with the evolution of A_t, i.e. the process evolves independent of the sampling. (iii) *On average* a_t is proportional to A_t up to a constant term.

It will be shown later, that the novel estimator is applicable to any time series a_t that was acquired from a BP conforming with this definition of subsampling. We will demonstrate possible applications at the hand of two examples:

A.3.1 Diagnosing infections with probability α.

For example, when a BP A_t represents the spread of infections within a population, each infection may be diagnosed with probability $\alpha \leqslant 1$, depending on the sensitivity of the test and the likelihood that an infected person consults a doctor. If each of the A_t infections is diagnosed independently of the others, then the number of diagnosed cases a_t follows a binomial distribution $a_t \sim \mathrm{Bin}(A_t, \alpha)$. Then $\langle a_t|A_t = j\rangle = \alpha j$ is given by the expected value of the binomial distribution. This implementation of subsampling conforms with the definition above, with the sampling probability α and the constant in (iii) being identical here.

A.3.2 Sampling a subset of system components.

In a different application, assume a high-dimensional system of interacting units that forms the substrate on which activation propagates. Often, the states of a subset of units are observed continuously, for example by placing electrodes that record the activity of the same set of neurons over the entire recording (Fig. 2.1**b**). This implementation of subsampling in finite size systems is mathematically approximated as follows: If n out of all N model units are sampled, the probability to sample a_t active units out of the actual A_t active units follows a hypergeometric distribution, $a_t \sim \mathrm{Hyp}(N, n, A_t)$. As $\langle a_t|A_t = j\rangle = jn/N$, this representation satisfies Def. 1 with $\alpha = n/N$. Choosing this special implementation of subsampling allows to evaluate $\mathrm{Var}[a_t]$ further in terms of A_t:

$$
\begin{aligned}
\mathrm{Var}[a_t] &= \langle \mathrm{Var}[a_t|A_t]\rangle + \mathrm{Var}[\langle a_t|A_t\rangle] \\
&= n\langle \frac{A_t}{N}\frac{N-A_t}{N}\frac{N-n}{N-1}\rangle + \mathrm{Var}[\frac{n}{N}A_t] \\
&= \frac{1}{N}\frac{n}{N}\frac{N-n}{N-1}\left(N\langle A_t\rangle - \langle A_t^2\rangle\right) + \frac{n^2}{N^2}\mathrm{Var}[A_t] \\
&= \frac{n}{N^2}\frac{N-n}{N-1}\left(N\langle A_t\rangle - \langle A_t\rangle^2\right) + \left(\frac{n^2}{N^2} - \frac{n}{N^2}\frac{N-n}{N-1}\right)\mathrm{Var}[A_t].
\end{aligned}
\tag{A.3}
$$

This expression precisely determines the variance $\mathrm{Var}[a_t]$ under subsampling from the properties $\langle A_t\rangle$ and $\mathrm{Var}[A_t]$ of the full process (which for BPs are known from Lemma 1), and from the parameters of subsampling n and N. Using Eq. (A.3), we could predict the linear regression slopes $\hat{r}_k$ under subsampling (Theorem 5, Eq. (A.17)) in more detail:

$$
r_k = \alpha^2 \frac{\mathrm{Var}[A_t]}{\mathrm{Var}[a_t]}m^k = \frac{n(N-1)\mathrm{Var}[A_t]}{(N-n)(N\langle A_t\rangle - \langle A_t\rangle^2) + (nN-N)\mathrm{Var}[A_t]}m^k =: b(N, n, \langle A_t\rangle, \mathrm{Var}[A_t])\,m^k. \tag{A.4}
$$

---

*Throughout this manuscript, the conditional random variable $(a_t|A_t = j)$ is to be read as "a_t given the realization $A_t = j$ of the random variable A_t".

The term $b = b(N, n, \langle A_t \rangle, \text{Var}[A_t])$ is constant when subsampling a given (stationary) system, and quantifies the factor by which $\hat{m}_C$ is biased when using the conventional estimate for m. It depends on N, n and the first two moments of A_t and is thus known for a BP. This relation was used for Fig. 2.1c.

A.4 MR estimation

We here derive an estimator for the mean offspring m based on the autoregressive representation of the BP,

$$\langle A_{t+1} | A_t = j \rangle = m j + h. \tag{A.5}$$

This novel estimator is based on multistep regressions (Statman *et al.*, 2014) (MR estimator), which generalize (A.5) to arbitrary time steps k. From iteration of Eq. (A.5), it is easy to see that

$$\langle A_{t+k} | A_t = j \rangle = m^k j + h \frac{1 - m^k}{1 - m}. \tag{A.6}$$

Definition 2 (Multistep regression estimator). Consider a subsampled BP $\{a_t\}$ of length T. Let $k_{\max} \in \mathbb{N}$, $k_{\max} \geq 2$. Then multistep regression (of $k_{\max}$-th order) estimates m in the following way:

1. For $k = 1, \dots, k_{\max}$, estimate the slope $\hat{r}_k$ and offset $\hat{s}_k$ of linear regression between the pairs $\{(a_t, a_{t+k})\}_{t=0}^{T-k}$, e.g. by least square estimation (Fig. 2.1e), i.e. by minimizing the residuals

$$R_k(\hat{r}_k, \hat{s}_k) = \sum_t \left(a_{t+k} - (\hat{r}_k \cdot a_t + \hat{s}_k) \right)^2. \tag{A.7}$$

2. Based on the relation (Statman *et al.*, 2014) $r_k = b \cdot m^k$, estimate $\hat{b}$ and $\hat{m}$ by minimizing the sum of residuals

$$R(\hat{b}, \hat{m}) = \sum_{k=1}^{k_{\max}} \left(\hat{r}_k - \hat{b} \cdot \hat{m}^k \right)^2, \tag{A.8}$$

with the collection of slopes $\{\hat{r}_k\}_{k=1}^{k_{\max}}$ obtained from step 1 (Fig. 2.1f).

Then $\hat{m}$ is the multistep regression (MR) estimate of the mean offspring m. For the application to experimental data, we further applied tests to identify nonstationarities (Sec. A.5).

We first prove that the MR estimator is consistent in the fully sampled case, and will then show the consistency under subsampling. First, we need the following result about the individual linear regression slopes $\hat{r}_k$ under full sampling:

Theorem 2. *The slope $\hat{r}_k$, obtained from A_t under full sampling, is a consistent estimator for m^k. If the process is subcritical, then the offset $\hat{s}_k$ is also a consistent estimator for $h \frac{1-m^k}{1-m}$.*

Remark. For $k = 1$, these results were already obtained by Heyde and Seneta (1972); Venkataraman (1982); Wei and Winnicki (1990), and details can be found in these sources. Based on their proofs, we here show the generalization to k timesteps.

Proof. Let $k \in \mathbb{N}$, $i \in \{0, \dots, k-1\}$. Construct a new random process by starting at time i and taking every k-th time step of the original process A_t. This new process is given by $A_{t'}^{(k,i)} = A_{i+k \cdot t'}$ with the index $t' \in \mathbb{N}$. Hence, the "time" t' of this new process relates to the time t of the old process as $t = i + k \cdot t'$.

For a time series of length T, let $r^{(k,i)}$ be the least square estimator for the slope and $\hat{s}^{(k,i)}$ the least square estimator for the intercept of linear regression on all pairs $(A_{t'+1}^{(k,i)}, A_{t'}^{(k,i)})$ from the time series $\{A_{t'}^{(k,i)}\}_{t'=0}^{\lfloor(T-1)/k\rfloor}$. We will derive that $r^{(k,i)}$ is a consistent estimator for m^k. According to Wei and Winnicki (1990), it is sufficient to show that the evolution of $A_{t'}^{(k,i)}$ can be rewritten as

$$A_{t'}^{(k,i)} = m^k \cdot A_{t'-1}^{(k,i)} + h\frac{1-m^k}{1-m} + \epsilon_{t'}^{(k,i)} \tag{A.9}$$

with a martingale difference sequence $\epsilon_{t'}^{(k,i)}$, as this is a stochastic regression equation. Hence, consider

$$\epsilon_{t'}^{(k,i)} = A_{t'}^{(k,i)} - m^k \cdot A_{t'-1}^{(k,i)} - h\frac{1-m^k}{1-m} = A_{i+kt'} - m^k \cdot A_{i+k\,(t'-1)} - h\frac{1-m^k}{1-m}. \tag{A.10}$$

We now show that $(\epsilon_{t'}^{(k,i)})_{t'\in\mathbb{N}}$ is a martingale difference sequence for all k. From iteration of Eq. (A.6), it is easy to see that

$$\langle A_{t'}^{(k,i)}|A_{t'-1}^{(k,i)} = j\rangle = \langle A_{kt'+i}|A_{kt'-k+i} = j\rangle = m^k j + h\frac{1-m^k}{1-m} \tag{A.11}$$

holds. Hence, $\langle\epsilon_{t'}^{(k,i)}|A_{t'-1}^{(k,i)} = j\rangle = 0$ for any j and $\{\epsilon_{t'}^{(k,i)}\}$ is indeed a martingale difference sequence. Therefore, $\{A_{t'}^{(k,i)}\}_{t'=0}^{\lfloor T/k\rfloor}$ satisfies a linear stochastic regression equation with slope m^k and intercept $h\frac{1-m^k}{1-m}$. The least square estimators return unbiased and consistent estimates for the slope and intercept in the subcritical case, i.e. the estimators converge in probability (Heyde and Seneta, 1972; Venkataraman, 1982; Wei and Winnicki, 1990):

$$\hat{r}^{(k,i)} \xrightarrow{\mathrm{P}} m^k \qquad \hat{s}^{(k,i)} \xrightarrow{\mathrm{P}} h\frac{1-m^k}{1-m}.$$

In the critical and supercritical cases, only $\hat{r}^{(k,i)} \xrightarrow{\mathrm{P}} m^k$ holds following (Wei and Winnicki, 1990). Hence, we obtain $\hat{r}_k \xrightarrow{\mathrm{P}} m^k$ for all m and $\hat{s}_k \xrightarrow{\mathrm{P}} h(1-m^k)/(1-m)$ if $m < 1$. $\qquad\square$

COROLLARY 3. *As least square estimation of $\hat{b}$ and $\hat{m}$ from minimizing the residual (A.8) is consistent, multistep regression is a consistent estimator for m under full sampling, $\hat{m} \xrightarrow{\mathrm{P}} m$.*

These results were obtained for BPs. However, the derivation is here only based on the autoregressive representation (A.5), motivation the following proposition:

CONJECTURE 4. *Multistep regression is a consistent estimator for m for any PAR satisfying Eq. (A.5).*

Numerical results for AR(1) and Kesten processes support this conjecture (Statman *et al.*, 2014) (Fig. A.1).

Next, we show that MR estimation is consistent in the subcritical case even if only the subsampled a_t is known:

THEOREM 5. *Let A_t be a PAR with $m < 1$ and a stationary limiting distribution A_∞ and let the PAR be started in the stationary distribution, i.e. $A_0 \sim A_\infty$. Let a_t be a subsampling of A_t. Multistep regression (MR) on the subsampled a_t is a consistent estimator of the mean offspring m.*

Proof. The existence of a stationary distribution A_∞ was shown by Heathcote (1965). The least square estimator for the slope of linear regression is also given by (Kenney and Keeping, 1962)

$$\hat{r}_k = \hat{\rho}_{a_t a_{t+k}} \frac{\hat{\sigma}_{a_t}}{\hat{\sigma}_{a_{t+k}}} \tag{A.12}$$

with the the estimated standard deviations $\hat{\sigma}_{a_t}$ and $\hat{\sigma}_{a_{t+k}}$ of a_t and a_{t+k} respectively. In the subcritical state, $\sigma_{a_t} = \sigma_{a_{t+k}}$ because of stationarity. Thus estimating the linear regression slope is equivalent to estimating the Pearson correlation coefficient $\hat{\rho}_{a_t a_{t+k}} = \hat{\rho}_{a_t}(k)$ (which is identical to the autocorrelation function of a_t). In the following, we calculate the Pearson correlation coefficient for the subsampled time series by evaluating $\langle a_t a_{t+k} \rangle$. We use the law of total expectation in order to express $\langle a_t a_{t+k} \rangle$ not in dependence of a_t, but in terms of A_t:

$$\langle a_t a_{t+k} \rangle = \langle \langle a_t a_{t+k} | A_t, A_{t+k} \rangle \rangle_{A_{t+k}, A_t}, \tag{A.13}$$

where the inner expectation value is taken with respect to the joint distribution of a_{t+k} and a_t, and the outer with respect to the joint distribution of A_{t+k} and A_t. Through conditioning on both A_t and A_{t+k}, $(a_t | A_t)$ and $(a_{t+k} | A_{t+k})$ become independent due to Def. 1. Hence, the joint distribution of $(a_t, a_{t+k} | A_t, A_{t+k})$ factorizes, and the expectation value factorizes as well. By definition, $\langle a_t | A_t = j \rangle = \alpha j + \beta$ and hence

$$\langle a_t a_{t+k} \rangle = \langle (\alpha A_{t+k} + \beta)(\alpha A_t + \beta) \rangle_{A_{t+k}, A_t} \tag{A.14}$$

Without loss of generality, we here show the proof for $\beta = 0$ which is easily extended to the general case. We express $\langle a_t a_{t+k} \rangle$ in terms of Eq. (A.6) using the law of total expectation again:

$$\begin{aligned}
\langle a_t a_{t+k} \rangle &= \alpha^2 \langle A_t A_{t+k} \rangle \\
&= \alpha^2 \langle \langle A_t A_{t+k} | A_t \rangle \rangle_{A_t} \\
&= \alpha^2 \langle A_t \left(m^k A_t + h \frac{1 - m^k}{1 - m} \right) \rangle_{A_t} \\
&= \alpha^2 \left(m^k \langle A_t^2 \rangle + (1 - m^k) \langle A_t \rangle^2 \right),
\end{aligned}$$

where the first expectation was taken with respect to the joint distribution of A_t and A_{t+k}. We then used that $\langle A_t^2 \rangle$ and $\langle A_t \rangle = h/(1 - m)$ exist, which follows from stationarity of the process. By a similar argument,

$$\langle a_{t+1} \rangle = \langle a_t \rangle = \langle \langle a_t | A_t \rangle \rangle_{A_t} = \alpha \langle A_t \rangle = \alpha \frac{h}{1 - m} \tag{A.15}$$

and combining these results the covariance is

$$\mathrm{Cov}[a_{t+k}, a_t] = \langle a_{t+k} a_t \rangle - \langle a_{t+k} \rangle \langle a_t \rangle = \alpha^2 \left(m^k \langle A_t^2 \rangle + (1 - m^k) \langle A_t \rangle^2 \right) - \alpha^2 \langle A_t \rangle^2 = \alpha^2 m^k \mathrm{Var}[A_t]. \tag{A.16}$$

Therefore, we find that the estimator $\hat{r}_k$ converges in probability:

$$\hat{r}_k \xrightarrow{\mathrm{p}} \rho_{a_t a_{t+k}} = \frac{\mathrm{Cov}[a_{t+k}, a_t]}{\mathrm{Var}[a_t]} = \alpha^2 \frac{\mathrm{Var}[A_t]}{\mathrm{Var}[a_t]} m^k. \tag{A.17}$$

Hence, the bias of of the conventional estimator $\hat{m}_{\mathrm{C}} = \hat{r}_1$ is precisely given by the factor $b = \alpha^2 \mathrm{Var}[A_t]/\mathrm{Var}[a_t]$. However, importantly the relation $\hat{r}_k = \hat{b} \hat{m}^k$ still holds for the subsampled a_t. Given a collection of multiple linear regressions $\hat{r}_1, \dots, \hat{r}_{k_{\max}}$, the least square estimation of $\hat{b}$ and $\hat{m}$ from minimizing the residual (A.8) yields a consistent estimator $\hat{m}$ for the mean offspring m even under subsampling and only requires the knowledge of a_t. $\qquad \square$

This proof also showed that the conventional estimator (Heyde and Seneta, 1972) is biased under sub-sampling:

COROLLARY 6. *Let $\{a_t\}$ be a subsampling of a subcritical PAR $\{A_t\}$. Then the conventional linear regression estimator $\hat{m}_C = \hat{r}_1$ by Heyde and Seneta (1972) is biased by $m(\alpha^2 \frac{\text{Var}[A_t]}{\text{Var}[a_t]} - 1)$. Equivalently, it is biased by the factor $\alpha^2 \frac{\text{Var}[A_t]}{\text{Var}[a_t]}$.*

A.4.1 Nonstationarity, criticality and supercriticality.

The consistency of the estimator in the fully sampled case is included in our proof of Lemma 2 and follows from the results by Heyde and Seneta (1972); Wei and Winnicki (1990). Our proof for the sub-sampled case (Theorem 5), in contrast, strictly requires stationarity ($A_t \sim A_\infty$ for any t) and the existence of the first two moments of A_t. We expect that the MR estimator is also consistent if the subcritical process is not started in the stationary distribution, $A_0 \nsim A_\infty$, because the results by (Heathcote, 1965) show that it will converge to this stationary distribution as $t \to \infty$ (Fig. A.2). Furthermore, numerical results suggest that the MR estimator is also consistent for critical and supercritical cases, where no stationary distribution exists (Fig. 2.3**d**).

A.5 Identifying common non-stationarities and Poisson activity.

In many types of analyses, non-stationarities in the time series can lead to wrong results, typically an overestimation of $\hat{m}$. We developed tests to exclude data sets with signatures of common non-stationarities. The different non-stationarities, their impact on the r_k and the rules for rejection of time series are outlined below.

First, *transient* increases of the drive h_t, e.g. in response to a stimulus, lead to a transient increase in $\langle A_t \rangle$. These transients induce correlations or anti-correlations, which prevail on long time scales (Figs. A.3**c,d**). The autocorrelation function is therefore better captured by an exponential with offset, $r_k = b_{\text{offset}} \cdot m_{\text{offset}}^k + c_{\text{offset}}$. If the residual of this exponential with offset R^2_{offset} was smaller than the residual of the MR model R^2_{exp} by a factor of two, $H_{\text{offset}} = (2 \cdot R^2_{\text{offset}} < R^2_{\text{exp}})$, then the data set was rejected. The factor two punishes for the differences in degree of freedom: The residuals of a model with two free parameters (exponential with offset) instead of one (exponential only) can only be smaller. Second, ramping of the drive can lead to overestimation of m (Fig. A.3**e**). The comparison of the two models with and without offset introduced above serves as a consistency check able to identify ramping: if the data are captured by a BP, both models should infer identical $\hat{m}$. Thus, a difference between $\hat{m}_{\text{exp}}$ and $\hat{m}_{\text{offset}}$ hints at the invalidity of MR estimation. Instead of $\hat{m}$, we compared the autocorrelation times $\hat{\tau}_{\text{offset}} = -\Delta t / \log \hat{m}_{\text{offset}}$ and $\hat{\tau}_{\text{exp}}$ obtained from both models, as the logarithmic scaling increases the sensitivity. If their relative difference was too large, then the data are inconsistent with a BP and MR estimation is invalid: $H_\tau = (|\tau_{\text{exp}} - \tau_{\text{offset}}| / \min\{\tau_{\text{exp}}, \tau_{\text{offset}}\} > 2)$.

Third, when a system changes between different states of activity, e.g. up and down states, the drive rate $\langle h_t \rangle$ may experience sudden jumps. These can lead to spurious autocorrelation (Fig. A.3**f**). To identify these trends resulting from non-stationary input h_t or from choosing too short data sets, we tested whether the sequence of r_k was fit better by a linear regression $r_k = q_1 k + q_2$ on the pairs (k, r_k), than by the exponential relation (A.8). If the residuals R^2_{lin} were smaller than R^2_{exp}: $H_{\text{lin}} = (R^2_{\text{lin}} < R^2_{\text{exp}})$, data were rejected.

Apart from non-stationarities, even Poisson activity ($m = 0$, $A_t = h_t$) with stationary rate may lead to a spurious overestimation of $\hat{m}$ as well: for *subsampled* branching processes of *finite* duration, the Poisson case and processes close to criticality ($m = 1$) can show very similar autocorrelation results, because the sequence of r_k is expected to be absolutely or almost flat, respectively. Moreover, for $m = 0$

H_{offset}	H_{τ}	H_{lin}	$H_{\bar{r}\leqslant 0}$	$(H_{q_1=0})$	interpretation	
×	×	×	×	–	BP with $m = \hat{m}$ explains data	MR estimation valid
✓	–	–	–	–		
–	✓	–	–	–	data not explained by BP	MR estimation invalid
–	–	✓	–	–		
–	–	–	✓	×		
–	–	–	✓	✓	Poisson activity ($m = 0$) explains data	MR estimation valid

TABLE A.1: **Consistency checks for MR estimation.** In order to assess if the results obtained from MR estimation are consistent with a BP with stationary parameters, we perform five tests (Sec. A.5). We discriminate the following cases in this order: A BP with $m = \hat{m}$ is only considered to explain the data, if the four tests H_{offset}, H_{τ}, H_{lin}, and $H_{\bar{r}\leqslant 0}$ are negative (×). If any of H_{offset}, H_{τ}, or H_{lin} is positive (✓), the data cannot be explained by a BP with any m, regardless of the other tests (–), and MR estimation is invalid. If $H_{\bar{r}\leqslant 0}$ is positive, the additional test $H_{q_1=0}$ becomes relevant: if it is negative, the data cannot be explained by a BP with any m. If it is also positive, the data are consistent with Poisson activity (BP with $m = 0$).

any solution on the manifold with $b = 0$ minimizes the residuals in Eq. (A.8). Hence, the estimator for $\hat{m}$ may yield any value depending on the initial conditions of the minimization scheme. To distinguish between $m = 0$ and $m > 0$, we used the fact that for $m = 0$, all slopes r_k are expected to be distributed around zero, $\langle r_k \rangle = 0$. In contrast, for processes with $m > 0$, all slopes are expected to be larger than zero $\langle r_k \rangle = b \cdot m^k > 0$. Thus to identify stationary Poisson activity, we tested (using a one-sided t-test) if the slopes obtained from the data were significantly larger than zero, yielding the p-value $p_{\bar{r}\leqslant 0}$ and the following test (Fig. A.3b): $H_{\bar{r}\leqslant 0} = (p_{\bar{r}\leqslant 0} \geqslant 0.1)$. The choice of the significance level should be guided by the severity of type I or II errors here: if it is set too liberal, Poisson activity may be mistaken for correlated activity, potentially even close-to-critical. On the other hand, if the significance level is too conservative, activity with long autocorrelation times may be spuriously considered Poissonian under strong subsampling (when b is small and all slopes only slightly differ from zero). For this study, we chose a significance level of $p_{\bar{r}\leqslant 0} < 0.1$ in order to not underestimate the risk of large activity cascades. To confirm candidates for Poisson activity identified through positive $H_{\bar{r}\leqslant 0}$, we assured that the r_k did not show a systematic trend, i.e. that linear regression of r_k as a function of k (see H_{lin} above) yielded slope zero: $H_{q_1=0} = (p_{q_1=0} \geqslant 0.05)$. The according significance level for this two sided test is then given by $p_{q_1\neq 0} < 0.05$.

We discriminate the following cases in the order indicated in Tab. A.1: $\hat{m}$ obtained from MR estimation is only valid if none of the tests (except $H_{q_1=0}$, which is ignored here) is positive. A positive result for any of H_{offset}, H_{τ}, or H_{lin} indicates non-stationarities, the data are not explained by a stationary BP, and MR estimation is invalid. If $H_{\bar{r}\leqslant 0}$ is positive, the data are potentially consistent with Poisson activity ($m = 0$). This is only the case if $H_{q_1=0}$ is also positive. If otherwise $H_{q_1=0}$ is negative, the Poisson hypothesis is also rejected and MR estimation invalid. This strategy correctly identified the validity of MR estimation for all investigated cases: stationary BPs with $m = 0.98$ and $m = 0.0$ were accepted, while nonstationary BPs with transient changes, ramping, or sudden jumps of the drive were excluded (Fig. A.3).

A.6 Variance of the estimates.

The distribution of $\hat{m}$ is consistent with a normal distribution $\mathcal{N}(m, \sigma_{\hat{m}}^2)$ centered around the true mean offspring m (Fig. A.4a; numerical results). The variance $\sigma_{\hat{m}}^2$ depends on the branching ratio m, the mean activity $\langle A_t \rangle$, the length L of the time series, and the sampling fraction α. Each of these factors affects $\sigma_{\hat{m}}^2$ mainly by changing the *effective length* of the time series, i.e. the number of non-zero entries $l = |\{A_t | A_t > 0\}|$. Thus, regardless of the actual time series length L or the mean activity $\langle A_t \rangle$, the

variance scales as a power-law in l, $\mathrm{Var}[\hat{m}] \propto l^{-\gamma}$ (Fig. A.4**b**). The exponent of this power-law depends on m. The closer to criticality the process is, the larger the exponent γ, i.e. the larger the benefit from longer time series length l. For $m = 0.99$, we found $\gamma \approx 3/2$. The performance of the estimator is in principle independent of the mean activity: Small $\langle A_t \rangle$ only affect the variance of the MR estimator through a potential decrease of l.

Similarly, the degree of subsampling only affects the variance of the estimator through a decrease of the effective length of a_t. While there may be a significant rise in $\sigma_{\hat{m}}^2$ when reducing the sampling fraction α, this increase can be explained by the coincidental decrease in l, as the rescaled variance $\sigma_{\hat{m}}^2 \cdot l^{\gamma}$ remains within one order of magnitude over four decades of the sampling fraction α (Fig. A.4**c**).

How does the variance change close to the critical transition? We found that the answer to this question highly depends on the specific choice of the parameters: if m is varied, one can either keep $\langle A_t \rangle$ or h constant, not both at the same time. If the mean activity $\langle A_t \rangle$ is fixed by choosing $h = \langle A_t \rangle (1 - m)$, then the variance of the process scales as $\mathrm{Var}[A_t] \propto 1/(1 - m)$ (Theorem 1). As $m \to 1$, the activity will inevitably get into a regime, where bursts of activity ($A_t > 0$) are disrupted by intermittent quiescent periods (A_t), thereby reducing l. In turn, the variance of the estimator increases as detailed before.

If however, the drive h is kept constant, we found that the variance scales linearly in the distance to criticality $\epsilon = 1 - m$ over at least 5 orders of magnitude of ϵ: $\sigma_{\hat{m}}^2 \propto \epsilon$ (Fig. A.4**d**). Thus, the variance decreases when approaching criticality, while the relative variance $\sigma_{\hat{m}}^2/\epsilon$ is constant. Note, however, that even though the standard deviation also decreases when approaching criticality ($\sigma_{\hat{m}} \propto \sqrt{\epsilon}$), the relative standard deviation increases ($\sigma_{\hat{m}}/\epsilon \propto 1/\sqrt{\epsilon}$).

For other measures of variation (e.g. quadratic (like the mean squared error MSE) and linear (like the inter-quartile range IQR)), we obtained scaling laws with the same exponents.

A.6.1 Confidence interval estimation.

We used a model based approach to estimate confidence intervals for both simulation and experimental data (for Figs. 2.1**c**,**d**, 2.2**c**,**d**, and 2.3**d**), because classical bootstrapping methods underestimate the estimator variance by treating all slopes r_k independently, while they are in fact dependent. We found that our model based approach constructs more conservative and representative confidence intervals.

For simulations, we simulated $B \in \mathbb{N}$ independent copies of the investigated model and applied MR estimation to each copy, yielding a collection of B independent estimates $\{\hat{m}^{(b)}\}_{b=1}^B$.

For experimental time series a_t with length L, mean activity $\langle a_t \rangle$, and number of sampled units n, MR estimation yields an estimate $\hat{m}$. We then simulated B copies of branching networks $\{A_t^{(b)}\}_{b=1}^B$ (for simulation details see Sec. A.8) with $N = 10,000$ units, $m = \hat{m}$ as inferred by MR estimation, and length L and rate $\langle a_t \rangle$ to match the data. The rate was matched by setting the drive to $h = \langle a_t \rangle (1 - \hat{m}) N/n$. Thereby, after subsampling n units, the mean activity of each resulting time series $a_t^{(b)}$ matched that of the original time series a_t, $\langle a_t^{(b)} \rangle = \langle a_t \rangle$. This procedure gives B copies of a BN that all match a_t in terms of the mean activity, the branching ratio, time series length, and number of sampled units. Applying MR estimation to these BNs yields a collection of B independent estimates $\{\hat{m}^{(b)}\}_{b=1}^B$. For both simulation and experimental data, the distribution of $\hat{m}$ and confidence intervals can be constructed from this collection.

A.7 Expectation maximization based on Kalman filtering

Kalman filtering is a method to predict the original time series A_t given a measurement a_t, defined for AR(1) processes and affine measurement transformation

$$A_{t+1} = m \cdot A_t + h_t$$
$$a_t = \alpha \cdot A_t + \beta_t \tag{A.18}$$

where h_t and β_t are independent Gaussian random variables $h_t \sim \mathcal{N}(h, \xi^2)$ and $\beta_t \sim \mathcal{N}(\beta, \zeta^2)$ and m and α constant real numbers. Assuming that $A_0 \sim \mathcal{N}(A, \psi)$, Kalman filtering infers the original time series $A_t | a_t, \mathcal{M}$ given a measured time series a_t and the known model $\mathcal{M} = (m, h, \xi^2, \alpha, \beta, \zeta^2, A, \psi)$. Based on an iterative expectation maximization algorithm which incorporates Kalman filtering (Hamilton, 1994; Shumway and Stoffer, 1982; Ghahramani and Hinton, 1996), the model parameters $\mathcal{M}$ can be estimated from a time series a_t. We used this algorithm to infer m. Because of the mutual dependence of the model parameters, we also needed to infer h, ξ^2, α, β, and ζ^2. In order to reduce the dimensionality of the maximization step, we disregarded A and ψ, as the influence of the initial value decreases if the time series gets long. For initial values, we chose $m = 0.5$ in the center of the range of interest for m, $h_t = \langle a_t \rangle \cdot (1 - m)$ (see Sec. A.2), $\xi = 0.1 \cdot h_t$, $\alpha = 1$, $\beta = 0$, and $\zeta = 0.1$. We further chose $A = \langle a_t \rangle$ and $\psi^2 = \mathrm{Var}[a_t]$ for the two model parameters that were not optimized.

We considered two termination criteria for the EM algorithm: First, it is recommended to restrict the EM algorithm to $10 - 20$ cycles in order to avoid overfitting, a common problem with likelihood-based fitting methods for multidimensional model parameters. Therefor we considered $\hat{m}$ inferred after 20 EM cycles. Second, we considered $\hat{m}$ after the results of two subsequent EM cycles did not differ by more than 0.01%.

We used the publicly available Python implementation of the Kalman EM algorithm, *pykalman*. All parameters were chosen as detailed above. The analysis was performed on a computer cluster, and reached runtimes of several days up to projected runtimes of weeks. In fact, this computational demand was a limiting factor in terms of widespread application. In contrast, MR estimation terminated within half a second on the same CPUs.

A.8 Simulations

A.8.1 Branching process.

We simulated BPs according to Eq. (A.1) in the following way: Realizations of the random numbers $y_{t,i}$ and h_t describing the number of offsprings, and the drive, were each drawn from a Poisson distribution: $y_{t,i} \sim \mathrm{Poi}(m)$ with mean m, and $h_t \sim \mathrm{Poi}(h)$ with mean h, respectively. Here, we used Poisson distributions as they allow for non-trivial offspring distributions with easy control of the branching ratio m by only one parameter. For the brain, one might assume that each neuron is connected to k postsynaptic neurons, each of which is excited with probability p, motivating a binomial offspring distribution with mean $m = kp$. As in cortex k is typically large and p is typically small, the Poisson limit is a reasonable approximation. For the performance of the MR estimator and the limit behavior of the BP, the particular form of the law Y is not important such that the special choice we made here does not restrict the generality of our results.

The mean rate $\langle A_t \rangle$ depends on m and h (Lemma 1). In the simulation we varied m and fixed $\langle A_t \rangle = 100$ by adjusting h accordingly if not stated otherwise. For subsampling the BP, each unit is observed independently with probability $p \leqslant 1$. Then a_t is distributed following a binomial distribution $\mathrm{Bin}(A_t, p)$, and subsampling is implemented by drawing a_t from A_t at each time step. As $\langle a_t \rangle = p A_t$, this implementation of subsampling satisfies the definition of stochastic subsampling with $\alpha = p$, $\beta = 0$.

A.8.2 Branching network.

In addition to the classical branching process, we also simulated a branching network model (BN) by mapping a branching process (Harris, 1963; Haldeman and Beggs, 2005) onto a fully connected network of $N = 10,000$ neurons. An active neuron activated each of its k postsynaptic neurons with probability $p = m/k$. Here, the activated postsynaptic neurons were drawn randomly without replacement at each step, thereby avoiding that two different active neurons would both activate the same target neuron. Similar to the BP, the BN is critical for $m = 1$ in the infinite size limit, and subcritical (supercritical) for $m < 1$ ($m > 1$). As detailed for the BP, h was adjusted to the choice of m to achieve $\langle A_t \rangle = 100$, which corresponds to a rate of 0.01 spikes per neuron and time step. Subsampling (Priesemann *et al.*, 2009) was applied to the model by sampling the activity of n neurons only, which were selected randomly before the simulation, and neglecting the activity of all other neurons.

A.8.3 Self-organized critical model.

The SOC neural network model we used here is the Bak-Tang-Wiesenfeld (BTW) model (Bak *et al.*, 1987). Translated to a neuroscience context, the model consisted of $N = 10,000$ (100×100) non-leaky integrate and fire neurons. A neuron i spiked if its membrane voltage $V_i(t)$ reached a threshold θ:

$$\text{If } V_i(t) > \theta, \ V_i(t + 1) = V_i(t) - 4. \tag{A.19}$$

Note that the choice of θ does not change the activity of the model at all, so we set $\theta = 0$ for convenience. The model neurons were arranged on a 2D lattice, and each neuron was connected locally to its four nearest neighbors with coupling strength $\alpha_{ij} = \alpha$:

$$V_i(t + 1) = V_i(t) + \sum_j \alpha_{ij} \delta(t - T_j) + h_i(t), \tag{A.20}$$

where T_j denotes the spike times of neuron j, and $h_i(t)$ is the Poisson drive to neuron i with mean rate h as defined for the BP above. Note that the neurons at the edges and corners of the grid had only 3 and 2 neighbors, respectively. This model is equivalent to the well-known Bak-Tang-Wiesenfeld model (Bak *et al.*, 1987) if $h \rightarrow 0$ and $\alpha = 1$. Subsampling (Priesemann *et al.*, 2009) was implemented in the same manner as for the BN.

A.8.4 Parameter choices.

If not stated otherwise, simulations were run for $L = 10^7$ time steps or until A_t exceeded 10^9, i.e. approximately half of the 32 bit integer range. If not stated otherwise, confidence intervals (Sec. A.6) were estimated from $B = 100$ samples, both for simulation and experiments.

In Figs. 2.1**c,d**, BNs and the BTW model were simulated with $N = 10^4$ units and $\langle A_t \rangle = 100$. In Fig. 2.1**e**, BPs were simulated with $m = 0.9$ and $\langle A_t \rangle = 100$.

In Fig. 2.3**c**, subcritical and critical BNs with $N = 10^4$ and $\langle A_t \rangle = 100$ were simulated, and $n = 100$ units sampled. Because of the non-stationary, exponential growth in the supercritical case, here BPs were simulated with $h = 0.1$ and units observed with probability $\alpha = 0.01$.

A.9 Epidemiological recordings

A.9.1 WHO data on measles worldwide.

Time series with yearly case reports for measles in 194 different countries are available online from the World Health Organization (WHO) for the years between 1980 and 2014. MR estimation was applied

to these time series. Because they contain very few data points and potential long-term drifts, we applied the consistency checks detailed above for every country (Tab. A.1). After these checks, 124 out of the 194 surveyed countries were accepted for MR analysis and included in our analysis. Yearly information on approximate vaccination percentages (measles containing vaccine dose 1, MCV1) for the same countries and time span are also available online from the WHO.

A.9.2 RKI data on norovirus, measles and MRSA in Germany.

For Germany, the Robert-Koch-Institute (RKI) surveys a range of infectious diseases on a weekly basis, including measles, norovirus, and invasive meticillin-resistant Staphylococcus aureus (MRSA). Case reports are available through their SURVSTAT@RKI server (Robert-Koch-Institute). Because of possible changes in report policies in the beginning of surveillance, we omitted the data from the first 6 months of each recording. Moreover, we omitted the incomplete week on the turn of the year, thus evaluating 52 full weeks in each year.

The MRSA recording showed a slow, small variation in the case reports that can be attributed to slow changes in the drive rates. To compensate for these slow drifts, we corrected the time series by subtracting a moving average over 3 years (156 weeks). We then applied MR estimation to the obtained time series. The recordings for measles and norovirus showed strong seasonal fluctuations of the case reports, resulting in a baseline oscillation of the autocorrelation function. We therefore used a modified model

$$r_k = b \cdot m^k + c \cdot \cos(2\pi k/T) \tag{A.21}$$

with a fixed period of $T = 52$ weeks, and estimated $\hat{m}$, $\hat{b}$, and $\hat{c}$ from minimizing the residual of this modified equation.

In order to obtain the naive estimates using the conventional linear regression estimator $\hat{m}_C = \hat{r}_1$, we used the following correction for seasonal fluctuations. Each incidence count a_t was normalized by the incidence counts from the same week, averaged over all years of recording ($\bar{a}_w = \langle a_{w+52 \cdot y} \rangle_y$ with the average taken over the years y for any week $w = 1, \dots, 52$), yielding the deseasonalized time series $a'_t = a_t/\bar{a}_{t \bmod 52}$. Linear regression was performed on this time series a'_t.

For Fig. 2.2**d**, subsampling was applied to the original time series assuming that every infection is diagnosed and reported with a probability α, yielding the binomial subsampling described in Sec. A.3. MR estimates were obtained from this subsampled time series according to Eq. (A.21), for the conventional estimator the subsampled time series was processed as described above.

A.10   Animal experiments

We evaluated spike population dynamics from recordings in rats, cats and monkeys. The rat experimental protocols were approved by the Institutional Animal Care and Use Committee of Rutgers University (Mizuseki *et al.*, 2009a,b). The cat experiments were performed in accordance with guidelines established by the Canadian Council for Animal Care (Blanche, 2009). 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. The spike recordings from the rats and the cats were obtained from the NSF-founded CRCNS data sharing website (Blanche and Swindale, 2006; Blanche, 2009; Mizuseki *et al.*, 2009a,b).

In rats the spikes were recorded in CA1 of the right dorsal hippocampus during an open field task. We used the first two data sets of each recording group (ec013.527, ec013.528, ec014.277, ec014.333, ec015.041, ec015.047, ec016.397, ec016.430). The data-sets provided sorted spikes from 4 shanks (ec013)

or 8 shanks (ec014, ec015, ec016), with 31 (ec013), 64 (ec014, ec015) or 55 (ec016) channels. We used both, spikes of single and multi units, because knowledge about the identity and the precise number of neurons is not required for the MR estimator. More details on the experimental procedure and the data-sets proper can be found in Mizuseki *et al.* (2009a,b).

For the spikes from the cat, neural data were recorded by Tim Blanche in the laboratory of Nicholas Swindale, University of British Columbia (Blanche, 2009). We used the data set pvc3, i.e. recordings in area 18 which contain 50 sorted single units (Blanche and Swindale, 2006). We used that part of the experiment in which no stimuli were presented, i.e., the spikes reflected spontaneous activity in the visual cortex of the anesthetized cat. Because of potential non-stationarities at the beginning and end of the recording, we omitted data before 25 s and after 320 s of recording. Details on the experimental procedures and the data proper can be found in Blanche (2009); Blanche and Swindale (2006).

The monkey data are the same as in Pipa *et al.* (2009); Priesemann *et al.* (2014). In these experiments, spikes were recorded simultaneously from up to 16 single-ended micro-electrodes ($\varnothing = 80\,\mu$m) or tetrodes ($\varnothing = 96\,\mu$m) in lateral prefrontal cortex of three trained macaque monkeys (M1: 6 kg ♀; M2: 12 kg ♂; M3: 8 kg ♀). The electrodes had impedances between 0.2 and 1.2 MΩ at 1 kHz, and were arranged in a square grid with inter electrode distances of either 0.5 or 1.0 mm. The monkeys performed a visual short term memory task. The task and the experimental procedure is detailed in Pipa *et al.* (2009). We analyzed spike data from 12 experimental sessions comprising almost 12.000 trials (M1: 4 sessions; M2: 5 sessions; M3: 3 sessions). 6 out of 12 sessions were recorded with tetrodes. Spike sorting on the tetrode data was performed using a Bayesian optimal template matching approach as described in Franke *et al.* (2010) using the "Spyke Viewer" software (Pröpper and Obermayer, 2013). On the single electrode data, spikes were sorted with a multi-dimensional PCA method (Smart Spike Sorter by Nan-Hui Chen).

A.10.1 Analysis.

For each recording, we collapsed the spike times of all recorded neurons into one single train of population spike counts a_t, where a_t denotes how many neurons spiked in the tth time bin Δt. We used $\Delta t = 4$ ms, reflecting the propagation time of spikes from one neuron to the next. Note that m scales with the bin size (bs) as $m(\text{bs} = k\Delta t) = m(\text{bs} = \Delta t)^k$, while the corresponding autocorrelation times are invariant under bin size changes. For Figs. 2.3**b** and A.6, we investigated single neuron activity by applying similar binning to the spike times of each neuron individually.

From these time series, we estimated $\hat{m}$ using the MR estimator with $k_{\max} = 2500$ (corresponding to 10 s) for the rat recordings, $k_{\max} = 150$ (600 ms) for the cat recording, and $k_{\max} = 500$ (2000 ms) for the monkey recordings, assuring that $k_{\max}$ was always in the order of multiple autocorrelation times. Experiments were excluded if the tests according to Sec. A.5 detected potential nonstationarities.

A.11   Code availability

Python code for MR estimation and branching process simulation is available from github and will be extended in the future.
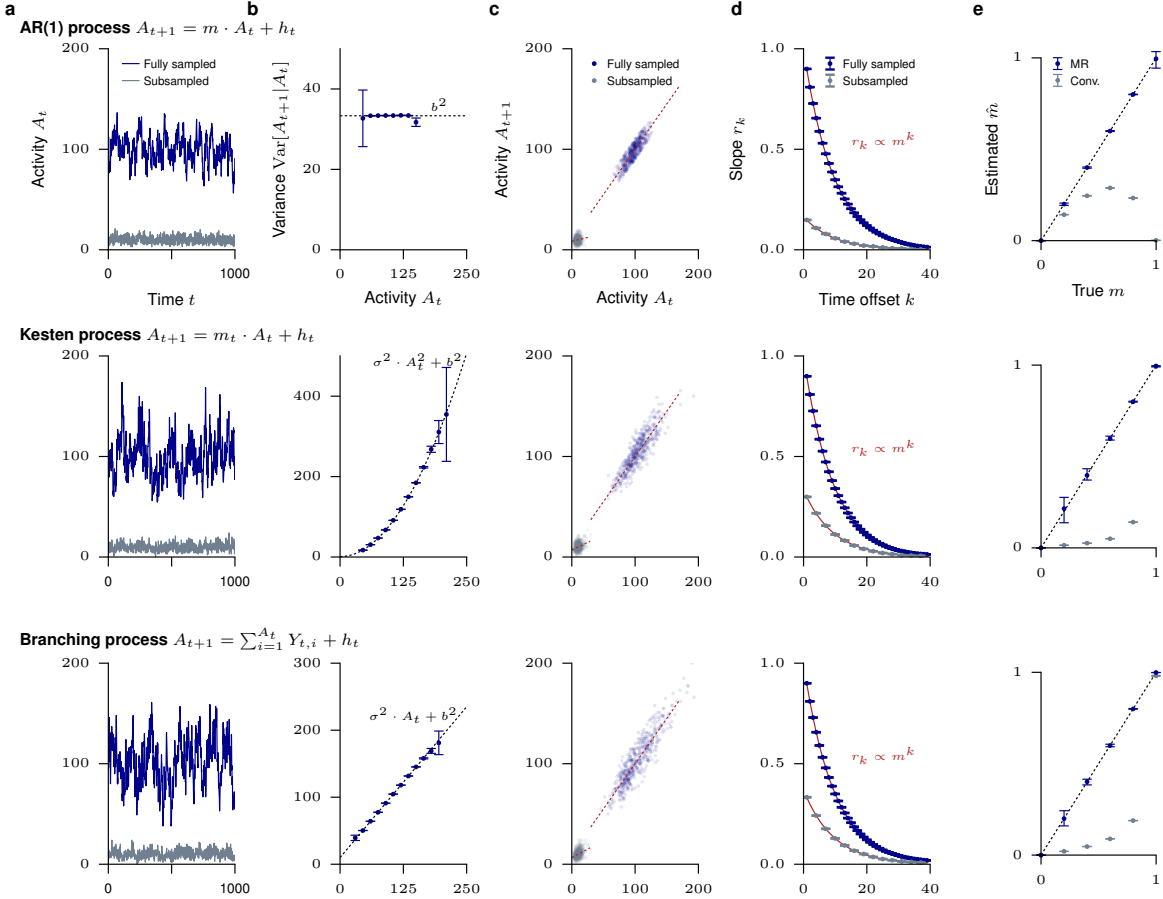
A.12 Supplementary Figures

FIGURE A.1: MR estimation for PARs. Although derived for branching processes (BPs), we conjectured that MR estimation is applicable to any process with a first order autoregressive representation (PAR). We here show exemplary results for three different classes of PARs: In AR(1) processes, additive noise h_t is drawn independently at each time step. Here, we considered a uniform distribution $h_t \sim \mathscr{U}(0, 2h)$. In a Kesten process, additive and multiplicative noise is drawn at each time step, both m_t and h_t being i.i.d. for all t. Here, $m_t \sim \mathscr{N}(m, \sigma^2)$ with $\sigma = m/10$ and $h_t \sim \mathscr{N}(h, b^2)$ with $b = h/10$ are normally distributed. In a BP, each unit i at time t generates $Y_{t,i}$ offspring, which are i.i.d. for all t and i. In addition, a random number h_t of units are introduced at each time step. Here, $Y_{t,i} \sim \text{Poi}(m)$ and $h_t \sim \text{Poi}(h)$ are Poisson distributed, σ^2 and b^2 denote the variances of $Y_{t,i}$ and h_t respectively. All three processes satisfy the first-order statistical recursion relation $\langle A_{t+1} | A_t \rangle = mA(t) + h$ (Eq. (A.5)). Parameters are chosen such that for all simulations the average activity is identical, $\langle A_t \rangle = 100$. **a.** Fully sampled and subsampled (binomial subsampling $a_t \sim \text{Bin}(A_t, \alpha)$ with $\alpha = 1/10$) time series are shown for $m = 0.9$ and $h = 10$. **b.** The three classes show the same first-order statistics according to Eq. (A.5). However, their second order statistics $\text{Var}[A_{t+1} | A_t]$ differ as indicated. **c.** Conventional linear regression underestimates $\hat{m}$ for all three processes under subsampling. **d.** MR estimation is applicable to all three processes under full sampling and subsampling, i.e. $r_k \propto m^k$ holds. **e.** While MR estimation returns consistent estimates of m even under subsampling, the conventional estimator underestimates $\hat{m}$ for all three processes.
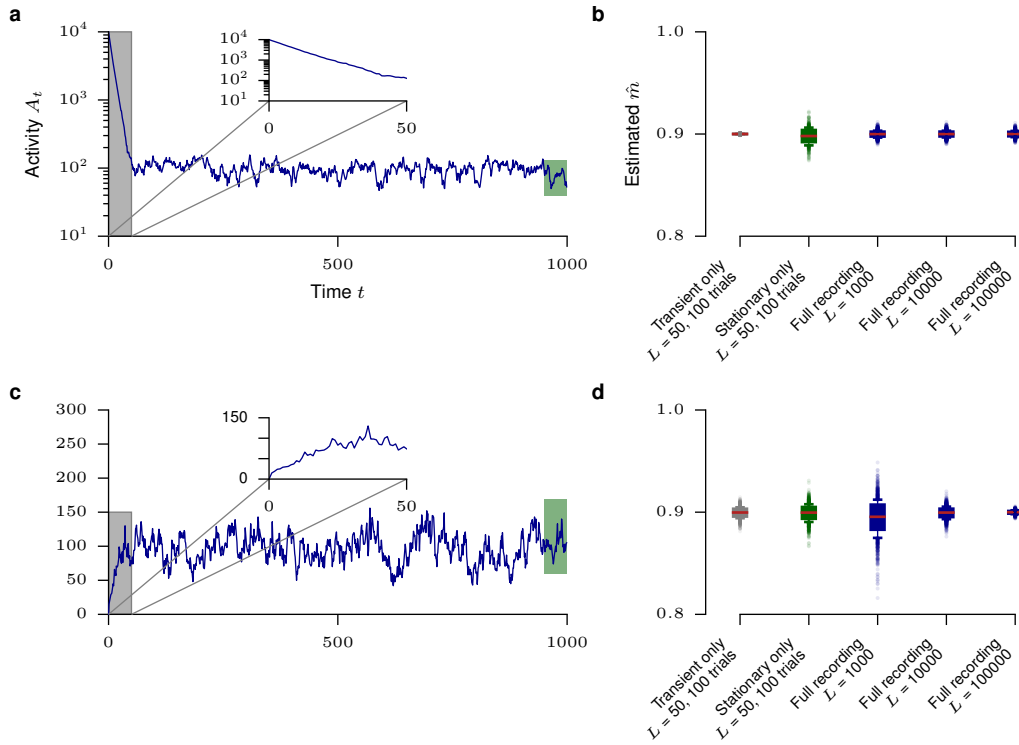
Figure A.2: MR estimation with transients. A branching process (BP) with $m = 0.9$ and expected activity $\langle A_t \rangle = 100$ is started far from the stationary distribution, namely with $A_0 = 10,000$ (top) or $A_0 = 0$ (bottom). Using MR estimation, $\hat{m}$ is inferred from: (i) only the first 50 data points of 100 independent trials, i.e. only transient parts of the activity in each trial (gray); (ii) 50 data points of 100 independent trials after the activity was allowed to relaxate to the stationary distribution in each trial (green); (iii) from one single trial comprising both transient and stationary parts, using 10^3, 10^4, or 10^5 time steps (blue). **a, c.** Activity A_t of one single trial of 10^3 time steps as a function of time t. Insets show magnified transient period where A_t converges to the stationary distribution. Shaded areas indicate transient (gray) and stationary (green) parts taken into account for estimates (i) and (ii) respectively. **b, d.** Boxplots (derived from 1000 independent realizations) for the result $\hat{m}$ of MR estimation, based on the data specified above.
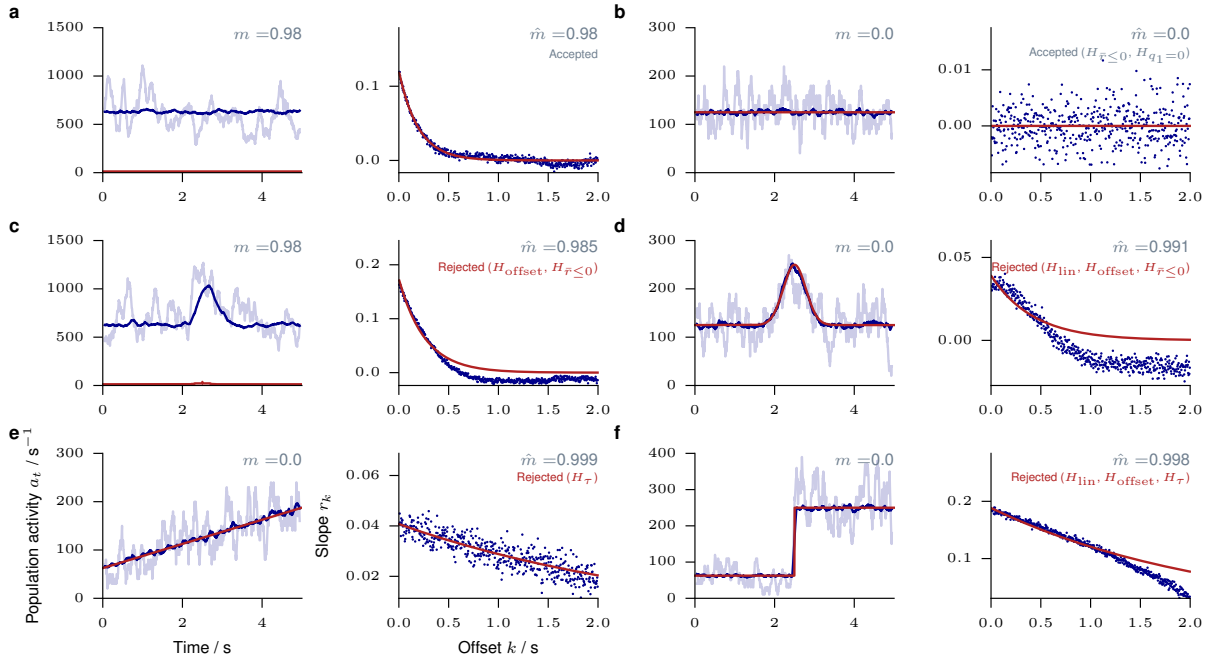
FIGURE A.3: Excluding nonstationary data. Each left panels shows the time series a_t of the activity from one single trial (light blue) and averaged activity from 100 trials (dark blue), recorded from $n = 50$ out of $N = 10^4$ neurons. Each right panels shows the corresponding MR estimation from one single trial. We investigated the following, generic cases for the temporal evolution of the drive rate $\langle h_t \rangle$: **a, b**. The drive is stationary ($\langle h_t \rangle$ identical for all t, red), so are the mean rates $\langle a_t \rangle$. **c, d**. The drive exhibits a transient increase centered around half of the simulation time. The mean rate $\langle a_t \rangle$ is therefore also time-dependent and follows the temporal evolution of $\langle h_t \rangle$. **e**. The drive shows a linear increase over the simulation. **f**. The drive exhibits a step function after half the simulation. Nonstationarities (**c** – **f**) typically lead to an overestimation of $\hat{m}$, which is particularly severe if the underlying dynamics is Poissonian ($m = 0$). The tests defined in Sec. A.5 (see Tab. A.1) were able to exclude time series where the investigated nonstationarities were present, while accepting the stationary cases **a**, **b**.
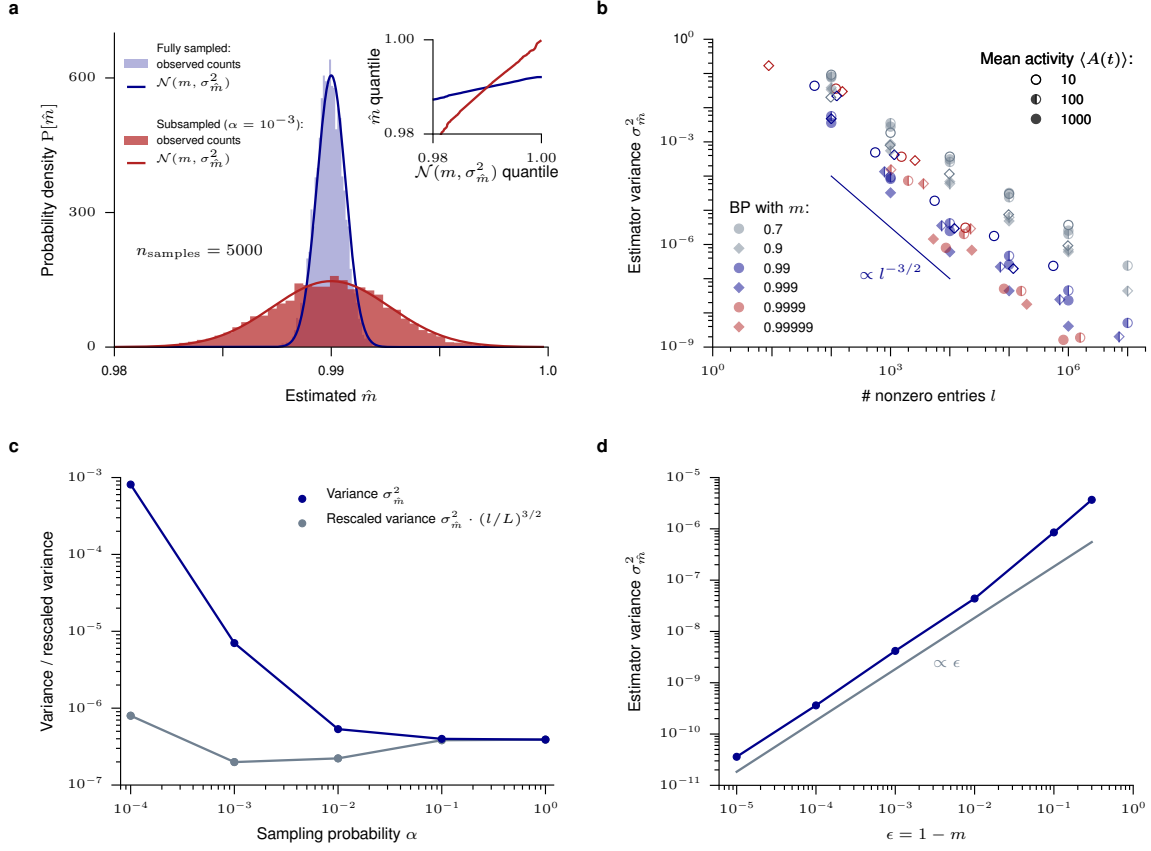
Figure A.4: Variance of the MR estimates. This figure shows numerical result for the distribution and variability of the estimate $\hat{m}$ as a function of multiple parameters. **a.** Distribution of the estimate $\hat{m}$, estimated from 5000 independent copies of a branching process (BP) with $m = 0.99$, $\langle A_t \rangle = 100$ and length $L = 10^5$: normalized histograms of the probability of estimating $\hat{m}$ for full sampling (blue) and binomial subsampling with $\alpha = 0.001$ (red), together with normal distributions $\mathcal{N}(m, \hat{\sigma}_{\hat{m}}^2)$. Inset: Q-Q-plot for the quantiles of $\mathcal{N}(m, \hat{\sigma}_{\hat{m}}^2)$ and the quantiles of the estimated $\hat{m}$ under both samplings. The estimated $\hat{m}$ are found to be distributed normally in both cases (fully sampled: $r^2 = 0.9995$, subsampled: $r^2 = 0.998$). **b.** The variance $\sigma_{\hat{m}}^2$ of the estimate $\hat{m}$ is estimated from 100 independent copies of a BP. Results for different m, mean activities $\langle A_t \rangle$ and time series lengths L are plotted as a function of the effective time series length $l = |\{A_t | A_t > 0\}|$, the number of nonzero entries. For any given m, the variance of $\hat{m}$ shows algebraic scaling $\sigma_{\hat{\epsilon}}^2 \propto l^\gamma$. The exponent of this scaling depends on m, with higher γ the closer m is to unity. Hence, the benefit from longer time series is larger the closer a system is to criticality. Importantly, the variance does not directly depend on the mean activity $\langle A_t \rangle$, this number only influences the accuracy of MR estimation via the potential change in l. **c.** The variance of the estimate $\hat{m}$ is estimated from 100 independent copies of a BP with $m = 0.99$, $\langle A_t \rangle = 100$, and $L = 10^5$ and plotted as a function of the sampling probability α under binomial subsampling. While the variance appears to increase dramatically under stronger subsampling, this increase can be attributed to the according decrease of the effective time series length l. After rescaling by $(l/L)^{3/2}$ (cf. panel **b**), the rescaled variance remains within one order of magnitude over four orders of magnitude in α. Hence, the accuracy of the estimator is not directly influenced by the degree of subsampling. **d.** The variance $\sigma_{\hat{m}}^2$ is estimated from 100 independent copies of a BP with $m = 0.99$, $h = 1$, and $L = 10^5$ and plotted as function of the distance to criticality $\epsilon = 1 - m$. The variance is found numerically to scale as $\sigma_{\hat{m}}^2 \propto \epsilon$, hence the standard deviation scales as $\sigma_{\hat{m}} \propto \sqrt{\epsilon}$. Similar scaling results were found for other linear (like the interquartile range) and quadratic (like the mean squared error) measures of variation.
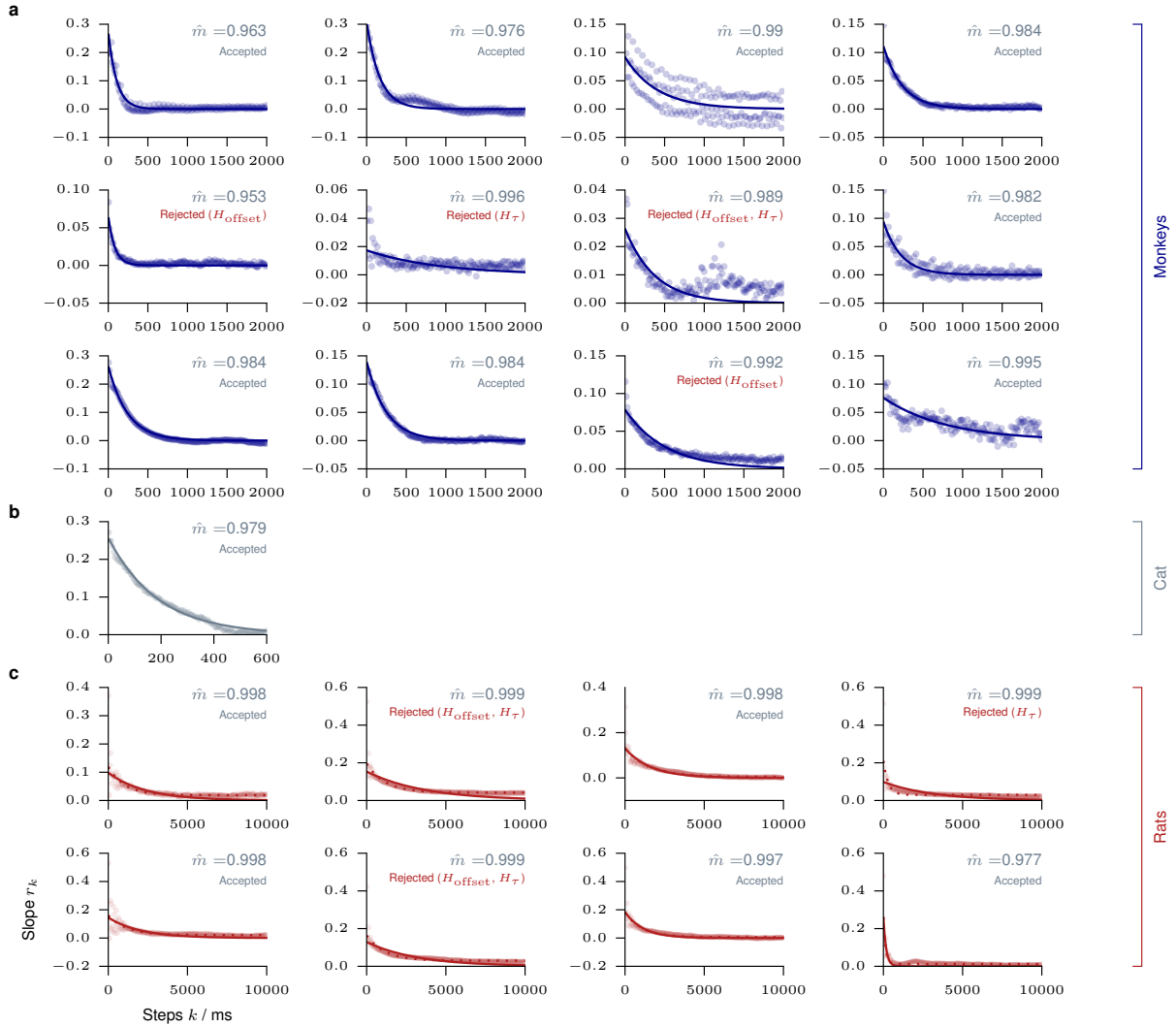
Figure A.5: MR estimation for individual animals. MR estimation is shown for every individual animal (see A.10). The consistency checks are detailed in Sec. A.5 (see Tab. A.1). **a**. Data from monkey prefrontal cortex during an working memory task. The third panel shows a oscillation of r_k with a frequency of 50 Hz, corresponding to measurement corruption due to power supply frequency. **b**. Data from anesthetized cat primary visual cortex. **c**. Data from rat hippocampus during a foreaging task. In addition to a slow exponential decay, the slopes r_k show the ϑ-oscillations of 6 – 10 Hz present in hippocampus. Dashed lines indicate results for an exponential model with offset, solid lines results for the model without offset (compare Sec. A.5).
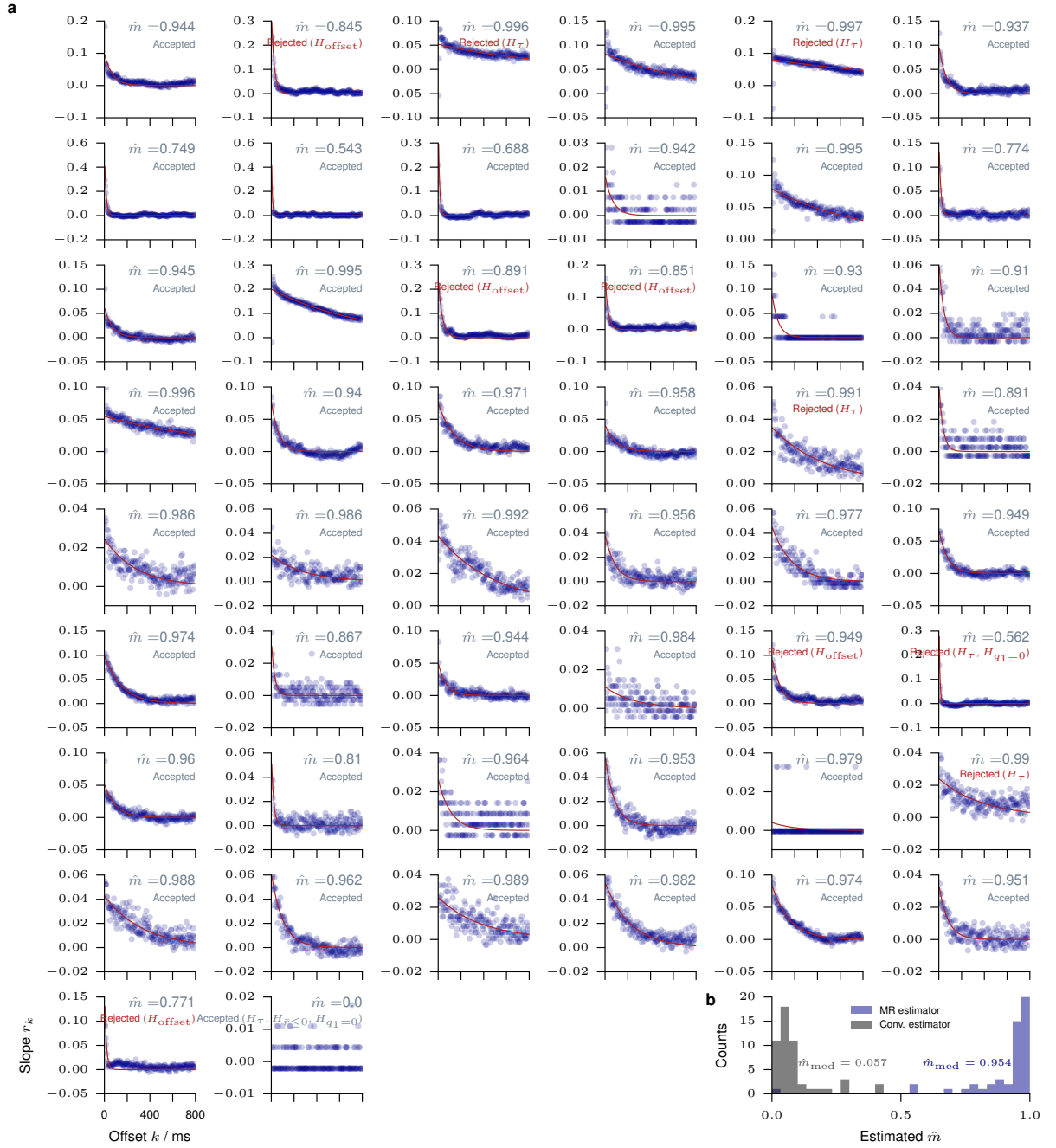
Figure A.6: MR estimation from single neuron activity (cat). MR estimation is used to estimate $\hat{m}$ from the activity a_t of a single neurons in cat visual cortex. **a.** Each panel shows MR estimation for one of the 50 recorded neurons. Autocorrelations decay rapidly in some neurons, but long-term correlations are present in the activity of most neurons. The consistency checks are detailed in Sec. A.5 (see Tab. A.1). **b.** Histogram of the single neuron branching ratios $\hat{m}$, inferred with the conventional estimator and using MR estimation. The difference between these estimates demonstrates the subsampling bias of the conventional estimator, and how it is overcome by MR estimation.

83

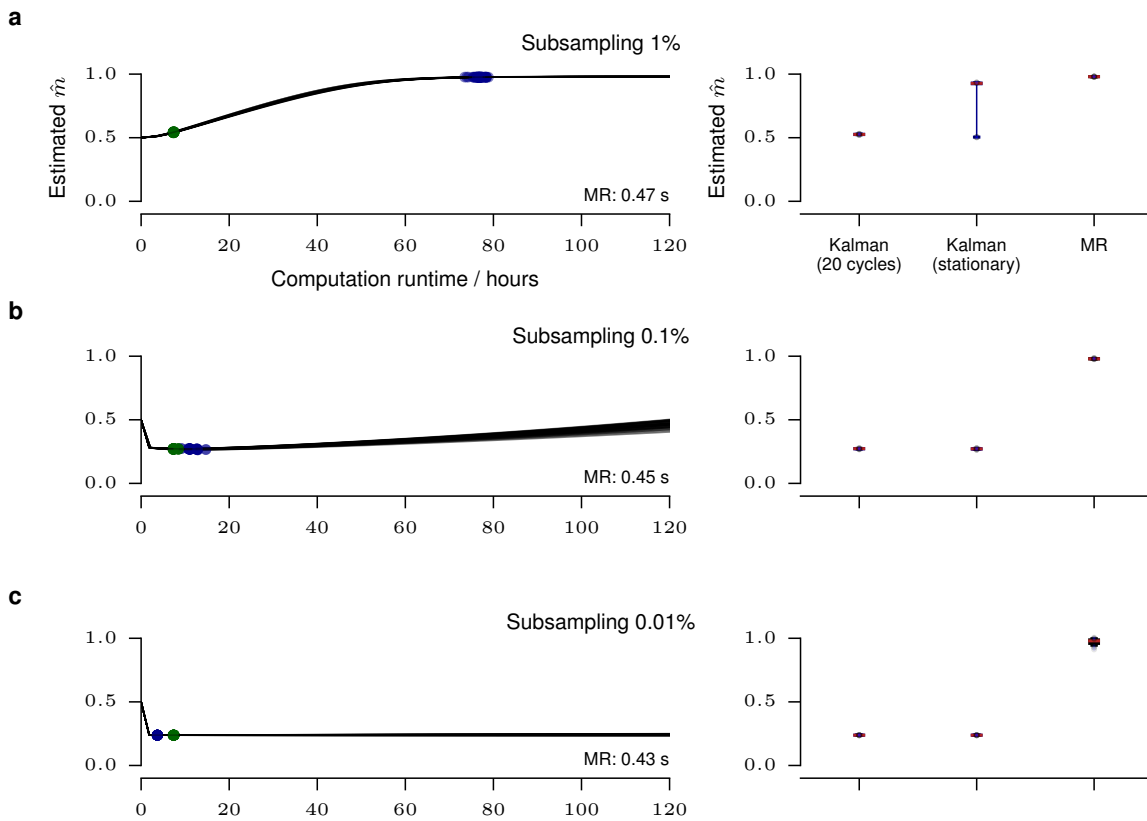**Figure A.7: Kalman EM estimation.** Expectation maximization (EM) based on Kalman filtering and MR estimation are used to infer $\hat{m}$ from BPs with $m = 0.99$ and different degrees of subsampling. Left column: inferred $\hat{m}$ as a function of the EM runtime for 100 independent copies of the BP. The EM algorithm is terminated after 20 cycles (green dots) or after the inferred $\hat{m}$ changed only marginally (blue dots, see Sec. A.7). The median runtime of MR estimation for the same BPs is also indicated. Right column: estimated $\hat{m}$ for all three methods. **a.** Under 1% subsampling, the EM algorithm converged after runtimes of about 80 h, compared to 0.43 s for MR estimation. **b.** Under 0.1% subsampling, $\hat{m}$ inferred by the EM algorithm reaches a steady state after 10 h, but is severely biased. The slow rise of $\hat{m}$ might lead to a convergance to the proper m after several weeks of projected runtime (ignoring common termination criteria). **c.** Under 0.01% subsampling, $\hat{m}$ inferred by the EM algorithm converge to a biased value. In contrast, MR estimation returns a correct $\hat{m}$ in all three cases, and outperforms the EM algorithm by a factor of 10^5 to 10^6 in terms of the runtime.

Appendix B

Supplementary material for "Between perfectly critical and fully irregular: a reverberating model captures and predicts cortical spike propagation"†

† The content of this chapter is identical in wording and figures to the online supplementary material of the publication Wilting and Priesemann (2019b): J. Wilting & V. Priesemann: Between perfectly critical and fully irregular: a reverberating model captures and predicts cortical spike propagation. *Cerebral Cortex* **29**(6):2759–2770 (2019). It has some overlap with Chap. A, as both chapters correspond to the online supplementary material of two independent publications, respectively.

B.1 Experiments

We evaluated spike population dynamics from recordings in rats, cats and monkeys. The rat experimental protocols were approved by the Institutional Animal Care and Use Committee of Rutgers University (Mizuseki *et al.*, 2009a,b). The cat experiments were performed in accordance with guidelines established by the Canadian Council for Animal Care (Blanche, 2009). 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. The spike recordings from the rats and the cats were obtained from the NSF-founded CRCNS data sharing website (Blanche and Swindale, 2006; Blanche, 2009; Mizuseki *et al.*, 2009a,b).

B.1.1 Rat experiments.

In rats the spikes were recorded in CA1 of the right dorsal hippocampus during an open field task. We used the first two data sets of each recording group (ec013.527, ec013.528, ec014.277, ec014.333, ec015.041, ec015.047, ec016.397, ec016.430). The data-sets provided sorted spikes from 4 shanks (ec013) or 8 shanks (ec014, ec015, ec016), with 31 (ec013), 64 (ec014, ec015) or 55 (ec016) channels. We used both, spikes of single and multi units, because knowledge about the identity and the precise number of neurons is not required for the MR estimator. More details on the experimental procedure and the data-sets proper can be found in Mizuseki *et al.* (2009a,b).

B.1.2 Cat experiments.

Spikes in cat visual cortex were recorded by Tim Blanche in the laboratory of Nicholas Swindale, University of British Columbia (Blanche, 2009). We used the data set pvc3, i.e. recordings of 50 sorted single units in area 18 (Blanche and Swindale, 2006). We used that part of the experiment in which no stimuli were presented, i.e., the spikes reflected spontaneous activity in the visual cortex of the anesthetized cat. Because of potential non-stationarities at the beginning and end of the recording, we omitted data before 25 s and after 320 s of recording. Details on the experimental procedures and the data proper can be found in Blanche (2009); Blanche and Swindale (2006).

B.1.3 Monkey experiments.

The monkey data are the same as in Pipa *et al.* (2009); Priesemann *et al.* (2014). In these experiments, spikes were recorded simultaneously from up to 16 single-ended micro-electrodes ($\varnothing = 80\,\mu$m) or tetrodes ($\varnothing = 96\,\mu$m) in lateral prefrontal cortex of three trained macaque monkeys (M1: 6 kg ♀; M2: 12 kg ♂; M3: 8 kg ♀). The electrodes had impedances between 0.2 and 1.2 MΩ at 1 kHz, and were arranged in a square grid with inter electrode distances of either 0.5 or 1.0 mm. The monkeys performed a visual short term memory task. The task and the experimental procedure is detailed in Pipa *et al.* (2009). We analyzed spike data from 12 experimental sessions comprising almost 12.000 trials (M1: 5 sessions; M2: 4 sessions; M3: 3 sessions). 6 out of 12 sessions were recorded with tetrodes. Spike sorting on the tetrode data was performed using a Bayesian optimal template matching approach as described in Franke *et al.* (2010) using the "Spyke Viewer" software (Pröpper and Obermayer, 2013). On the single electrode data, spikes were sorted with a multi-dimensional PCA method (Smart Spike Sorter by Nan-Hui Chen).

B.2 Analysis

B.2.1 Temporal binning.

For each recording, we collapsed the spike times of all recorded neurons into one single train of population spike counts a_t, where a_t denotes how many neurons spiked in the tth time bin Δt. If not indicated otherwise, we used $\Delta t = 4$ ms, reflecting the propagation time of spikes from one neuron to the next.

B.2.2 Multistep regression estimation of $\hat{m}$.

From these time series, we estimated $\hat{m}$ using the MR estimator described in Wilting and Priesemann (2018a). For $k = 1, \dots, k_{\max}$, we calculated the linear regression slope $r_{k\Delta t}$ for the linear statistical dependence of a_{t+k} upon a_t. From these slopes, we estimated $\hat{m}$ following the relation $r_{\delta t} = b \cdot \hat{m}^{\delta t/\Delta t}$, where b is an (unknown) parameter that depends on the higher moments of the underlying process and the degree of subsampling. However, for an estimation of m no further knowledge about b is required.
Throughout this study we chose $k_{\max} = 2500$ (corresponding to 10 s) for the rat recordings, $k_{\max} = 150$ (600 ms) for the cat recording, and $k_{\max} = 500$ (2000 ms) for the monkey recordings, assuring that $k_{\max}$ was always in the order of multiple intrinsic network timescales (i.e., autocorrelation times).
In order to test for the applicability of a MR estimation, we used a set of conservative tests (Wilting and Priesemann, 2018a), which found the expected exponential relation $r_{\delta t} = b m^{\delta t/\Delta t}$ in the majority of experimental recordings (14 out of 21, Fig. B.1).

B.2.3 Avalanche size distributions.

Avalanche sizes were determined similarly to the procedure described in Priesemann *et al.* (2009, 2014). Assuming that individual avalanches are separated in time, let $\{t_i\}$ indicate bins without activity, $a_{t_i} = 0$.

The size s_i of one avalanche is defined by the integrated activity between two subsequent bins with zero activity:

$$s_i = \sum_{t=t_i}^{t_{i+1}} a_t.$$ (B.1)

From the sample $\{s_i\}$ of avalanche sizes, avalanche size distributions $p(s)$ were determined using frequency counts. For illustration, we applied logarithmic binning, i.e. exponentially increasing bin widths for s.

For each experiments, these empirical avalanche size distributions were compared to avalanche size distributions obtained in a similar fashion from three different matched models (see below for details). Model likelihoods $l(\{s_i\})|m)$ for all three models were calculated following Clauset *et al.* (2009), and we considered the likelihood ratio to determine the most likely model based on the observed data.

B.2.4 ISI distributions, Fano factors and spike count cross-correlations.

For each experiment and corresponding reverberating branching model (subsampled to a single unit), ISI distributions were estimated by frequency counts of the differences between subsequent spike times for each channel.

We calculated the single unit Fano factor $F = \mathrm{Var}[a_t]/\langle a_t \rangle$ for the binned activity a_t of each single unit, with the bin sizes indicated in the respective figures. Likewise, single unit Fano factors for the reverberating branching models were calculated from the subsampled and binned time series.

From the binned single unit activities a_t^1 and a_t^2 of two units, we estimated the spike count cross correlation $r_{\mathrm{sc}} = \mathrm{Cov}(a_t^1, a_t^2)/\sigma_{a_t^1}\sigma_{a_t^2}$. The two samples a_t^1 and a_t^2 for the reverberating branching models were obtained by sampling two randomly chosen neurons.

B.3 Branching processes

In a branching process (BP) with immigration (Harris, 1963; Heathcote, 1965; Pakes, 1971) each unit i produces a random number $y_{t,i}$ of units in the subsequent time step. Additionally, in each time step a random number h_t of units immigrates into the system (drive). Mathematically, BPs are defined as follows (Harris, 1963; Heathcote, 1965): Let $y_{t,i}$ be independently and identically distributed non-negative integer-valued random variables following a law Y with mean $m = \langle Y \rangle$ and variance $\sigma^2 = \mathrm{Var}[Y]$. Further, Y shall be non-trivial, meaning it satisfies $\mathrm{P}[Y = 0] > 0$ and $\mathrm{P}[Y = 0] + \mathrm{P}[Y = 1] < 1$. Likewise, let h_t be independently and identically distributed non-negative integer-valued random variables following a law H with mean rate $h = \langle H \rangle$ and variance $\xi^2 = \mathrm{Var}[H]$. Then the evolution of the BP A_t is given recursively by

$$A_{t+1} = \sum_{i=1}^{A_t} y_{t,i} + h_t,$$ (B.2)

i.e. the number of units in the next generation is given by the offspring of all present units and those that were introduced to the system from outside.

The stability of BPs is solely governed by the mean offspring m. In the subcritical state, $m < 1$, the population converges to a stationary distribution A_∞ with mean $\langle A_\infty \rangle = h/(1-m)$. At criticality ($m = 1$), A_t asymptotically exhibits linear growth, while in the supercritical state ($m > 1$) it grows exponentially. We will now derive results for the mean, variance, and Fano factor of subcritical branching processes. Following previous results, taking expectation values of both sides of Eq. (A.1) yields $\langle A_{t+1} \rangle = m\langle A_t \rangle + h$.

Because of stationarity $\langle A_{t+1} \rangle = \langle A_t \rangle = \langle A_\infty \rangle$ and the mean activity is given by

$$\langle A_\infty \rangle = \frac{h}{1-m}. \tag{B.3}$$

In order to derive an expression for the variance of the stationary distribution, observe that by the theorem of total variance, $\text{Var}[A_{t+1}] = \langle \text{Var}[A_{t+1} | A_t] \rangle + \text{Var}[\langle A_{t+1} | A_t \rangle]$, where $\langle \cdot \rangle$ denotes the expected value, and $A_{t+1} | A_t$ conditioning the random variable A_{t+1} on A_t. Because A_{t+1} is the sum of independent random variables, the variances also sum: $\text{Var}[A_{t+1} | A_t] = \sigma^2 A_t + \xi^2$. Using the previous result for $\langle A_\infty \rangle$ one then obtains

$$\text{Var}[A_{t+1}] = \xi^2 + \sigma^2 \frac{h}{1-m} + \text{Var}[mA_t + h] = \xi^2 + \sigma^2 \frac{h}{1-m} + m^2 \text{Var}[A_t].$$

Again, in the stationary distribution $\text{Var}[A_{t+1}] = \text{Var}[A_t] = \text{Var}[A_\infty]$ which yields

$$\text{Var}[A_\infty] = \frac{1}{1-m^2} \left(\xi^2 + \sigma^2 \frac{h}{1-m} \right), \tag{B.4}$$

The Fano factor $F_{A_t} = \text{Var}[A_t]/\langle A_t \rangle$ is easily computed from (B.3) and (B.4):

$$F_{A_t} = \frac{\xi^2}{h(1+m)} + \frac{\sigma^2}{1-m^2}. \tag{B.5}$$

Interestingly, the mean rate, variance, and Fano factor all diverge when approaching criticality (given a constant input rate h): $\langle A_\infty \rangle \to \infty$, $\quad \text{Var}[A_\infty] \to \infty$, and $F_{A_t} \to \infty$ as $m \to 1$.

These results were derived without assuming any particular law for Y or H. Although the limiting behavior of BPs does not depend on it (Harris, 1963; Heathcote, 1965; Pakes, 1971), fixing particular laws allows to simplify these expressions further.

We here chose Poisson distributions with means m and h for Y and H respectively: $y_{t,i} \sim \text{Poi}(m)$ and $h_t \sim \text{Poi}(h)$. We chose these laws for two reasons: (1) Poisson distributions allow for non-trivial offspring distributions with easy control of the branching ratio m by only one parameter. (2) For the brain, one might assume that each neuron is connected to k postsynaptic neurons, each of which is excited with probability p, motivating a binomial offspring distribution with mean $m = kp$. As in cortex k is typically large and p is typically small, the Poisson limit is a reasonable approximation. Choosing these distributions, the variance and Fano factor become

$$\text{Var}[A_t] = h/((1-m)^2(1+m)),$$
$$F_{A_t} = 1/(1-m^2). \tag{B.6}$$

Both diverge when approaching criticality ($m = 1$).

B.4 Subsampling

A general notion of subsampling was introduced in Wilting and Priesemann (2018a). The subsampled time series a_t is constructed from the full process A_t based on the three assumptions: (i) The sampling process does not interfere with itself, and does not change over time. Hence the realization of a subsample at one time does not influence the realization of a subsample at another time, and the conditional *distribution* of $(a_t | A_t)$ is the same as $(a_{t'} | A_{t'})$ if $A_t = A_{t'}$. However, even if $A_t = A_{t'}$, the subsampled a_t and

$a_{t'}$ do not necessarily take the same value. (ii) The subsampling does not interfere with the evolution of A_t, i.e. the process evolves independent of the sampling. (iii) *On average a_t is proportional to A_t up to a constant term*, $\langle a_t | A_t \rangle = \alpha A_t + \beta$.

In the spike recordings analyzed in this study, the states of a subset of neurons are observed by placing electrodes that record the activity of the same set of neurons over the entire recording. This implementation of subsampling translates to the general definition in the following manner: If n out of all N neurons are sampled, the probability to sample a_t active neurons out of the actual A_t active neurons follows a hypergeometric distribution, $a_t \sim \text{Hyp}(N, n, A_t)$. As $\langle a_t | A_t = j \rangle = j n / N$, this representation satisfies the mathematical definition of subsampling with $\alpha = n/N$. Choosing this special implementation of subsampling allows to derive predictions for the Fano factor under subsampling and the spike count cross correlation. First, evaluate $\text{Var}[a_t]$ further in terms of A_t:

$$
\begin{aligned}
\text{Var}[a_t] &= \langle \text{Var}[a_t | A_t] \rangle + \text{Var}[\langle a_t | A_t \rangle] \\
&= n \langle \frac{A_t}{N} \frac{N - A_t}{N} \frac{N - n}{N - 1} \rangle + \text{Var}[\frac{n}{N} A_t] \\
&= \frac{1}{N} \frac{n}{N} \frac{N - n}{N - 1} \left(N \langle A_t \rangle - \langle A_t^2 \rangle \right) + \frac{n^2}{N^2} \text{Var}[A_t] \\
&= \frac{n}{N^2} \frac{N - n}{N - 1} \left(N \langle A_t \rangle - \langle A_t \rangle^2 \right) + \left(\frac{n^2}{N^2} - \frac{n}{N^2} \frac{N - n}{N - 1} \right) \text{Var}[A_t].
\end{aligned}
\tag{B.7}
$$

This expression precisely determines the variance $\text{Var}[a_t]$ under subsampling from the properties $\langle A_t \rangle$ and $\text{Var}[A_t]$ of the full process, and from the parameters of subsampling n and N. We now show that the Fano factor approaches and even falls below unity under strong subsampling, regardless of the underlying dynamical state m. In the limit of strong subsampling ($n \ll N$) Eq. (A.3) yields:

$$
\text{Var}[a_t] \approx \frac{n}{N^2} \left(N \langle A_t \rangle - \langle A_t \rangle^2 \right) + \frac{n^2 - n}{N^2} \text{Var}[A_t].
\tag{B.8}
$$

Hence the subsampled Fano factor is given by

$$
F_{a_t} = \frac{\text{Var}[a_t]}{\langle a_t \rangle} \approx 1 - \frac{\langle A_t \rangle}{N} + \frac{n - 1}{N} \frac{\text{Var}[A_t]}{\langle A_t \rangle} = 1 - \frac{\langle A_t \rangle - (n - 1) F_{A_t}}{N}.
\tag{B.9}
$$

Interestingly, when sampling a single unit ($n = 1$) the Fano factor of that unit becomes completely independent of the Fano factor of the full process:

$$
F_{a_t} = 1 - \langle A_t \rangle / N = 1 - \langle a_t \rangle / n = 1 - R,
\tag{B.10}
$$

where $R = \langle a_t \rangle / n$ is the mean rate of a single unit.

Based on this implementation of subsampling, we derived analytical results for the cross-correlation between the activity of two units on the time scale of one time step. The pair of units is here represented by two independent samplings a_t and $\tilde{a}(t)$ of a BP A_t with $n = 1$, i.e. each represents one single unit. Because both samplings are drawn from identical distributions, their variances are identical and hence the correlation coefficient is given by $r_{\text{sc}} = \text{Cov}(a_t, \tilde{a}(t)) / \text{Var}[a_t]$. Employing again the law of total expectation and using the independence of the two samplings, this can be evaluated:

$$
\text{Cov}(a_t, \tilde{a}(t)) = \langle \langle a_t \tilde{a}(t) | A_t \rangle \rangle_{A_t} - \langle \langle a_t | A_t \rangle \rangle_{A_t}^2 = \frac{1}{N^2} \text{Var}[A_t],
\tag{B.11}
$$

with the first inner expectation being taken over the joint distribution of a_t and $\tilde{a}(t)$. Using Eq. (B.8), one easily obtains

$$r_{\mathrm{sc}} = \frac{\mathrm{Var}[A_t]}{N\langle A_t\rangle - \langle A_t\rangle^2} = \frac{F_{A_t}}{N - \langle A_t\rangle} = \frac{F_{A_t}}{N\,(1 - R)} \tag{B.12}$$

with the mean single unit rate $R = \langle A_t\rangle/N$. For subcritical systems, the Fano factor F_{A_t} is much smaller than N, and the rate is typically much smaller than 1. Therefore, the cross-correlation between single units is typically very small.
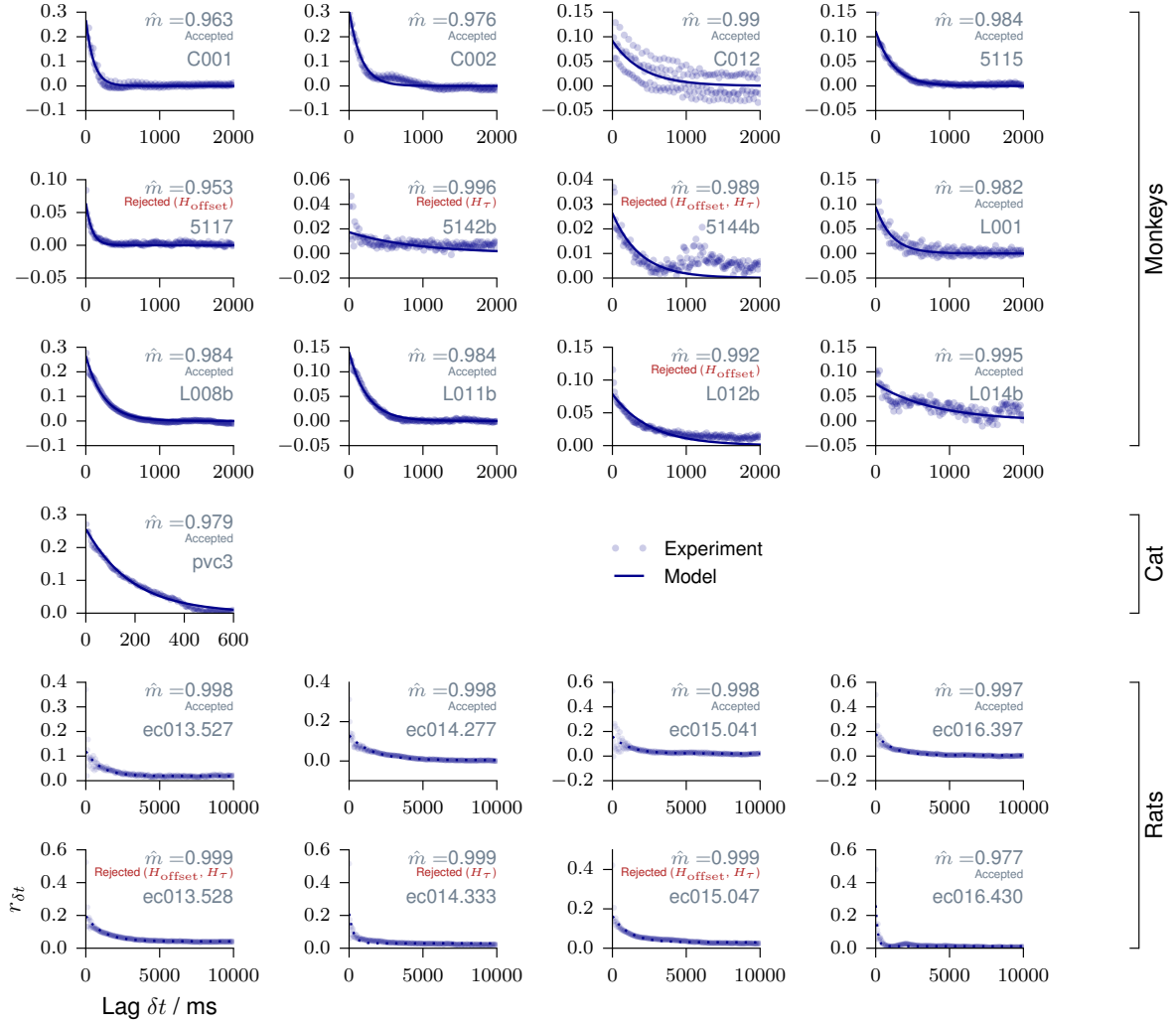
B.5   Supplementary Figures

Figure B.1: MR estimation for individual recording sessions. Reproduced from Wilting and Priesemann (2018a). MR estimation is shown for every individual animal. The consistency checks are detailed in Wilting and Priesemann (2018a). Data from monkey were recorded in prefrontal cortex during an working memory task. The third panel shows a oscillation of r_k with a frequency of 50 Hz, corresponding to measurement corruption due to power supply frequency. Data from anesthetized cat were recorded in primary visual cortex. Data from rat were recorded in hippocampus during a foraging task. In addition to a slow exponential decay, the slopes r_k show the ϑ-oscillations of 6 – 10 Hz present in hippocampus.
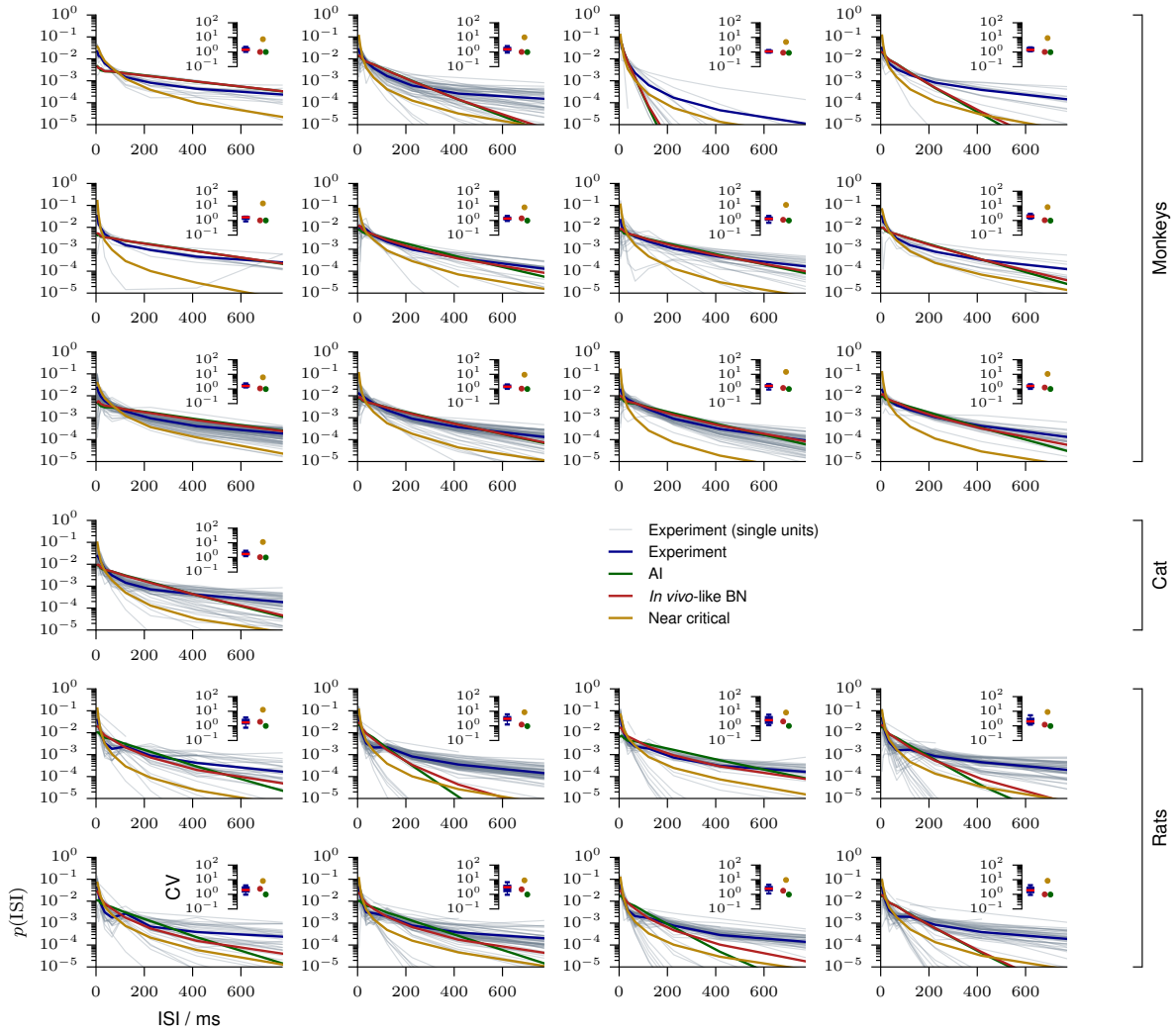
Figure B.2: Interspike interval distribution for individual recording sessions. Interspike interval (ISI) distributions are shown for individual units of each recording (gray), for the average over units of each recording (blue), as well as for the matched models, either AI (green), *in vivo*-like (red), or near critical (yellow). The insets show the corresponding coefficients of variation (CV). For every experiment AI and *in vivo*-like models are virtually indistinguishable by the ISI distributions.
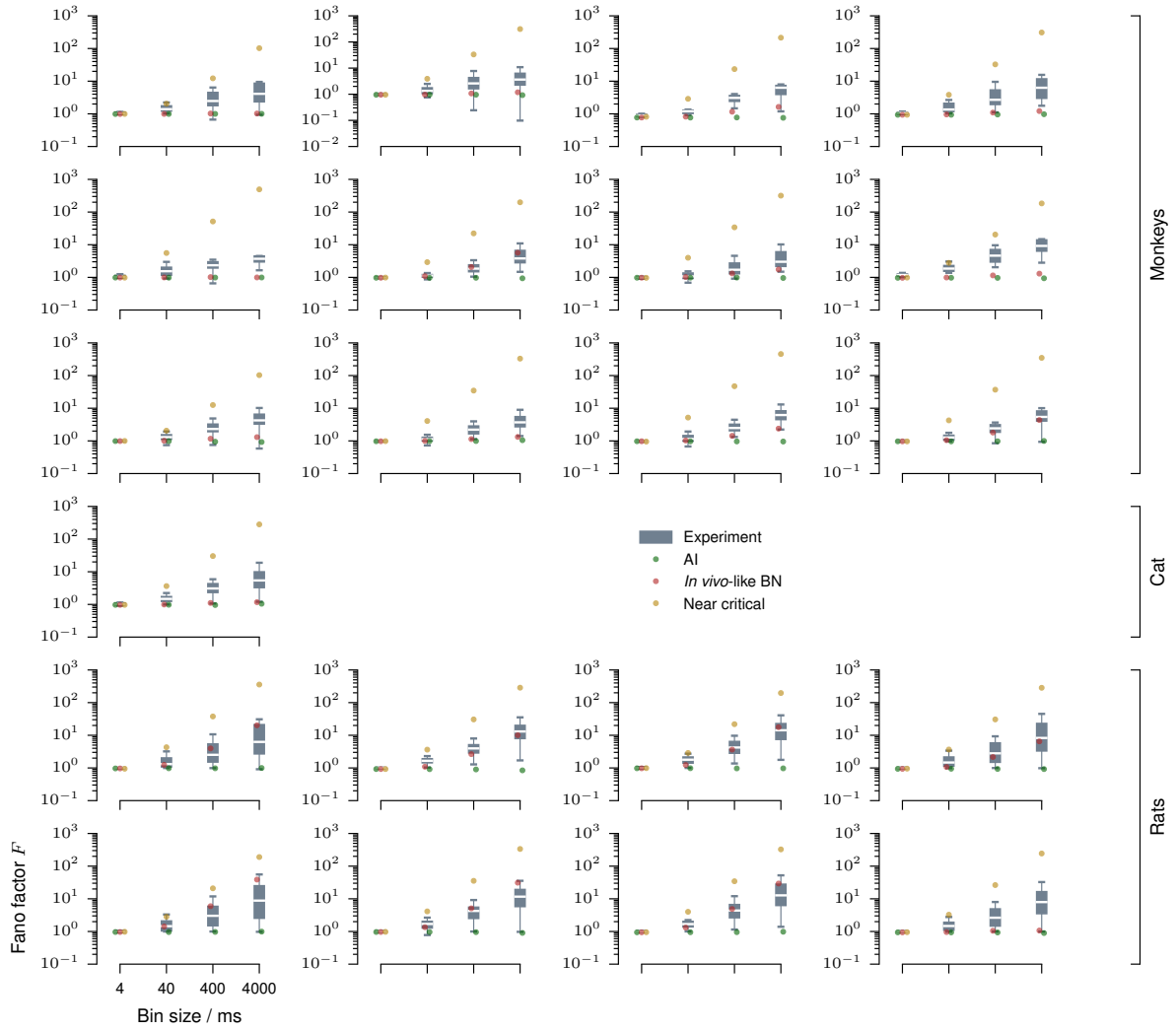
FIGURE B.3: **Fano factors for individual recording sessions.** Fano factors are shown for individual single or multi units of every recording (gray boxplots, median / 25% – 75%, 2.5% – 97.5%), as well as for the matched models, either AI (green), *in vivo*-like (red), or near critical (yellow).
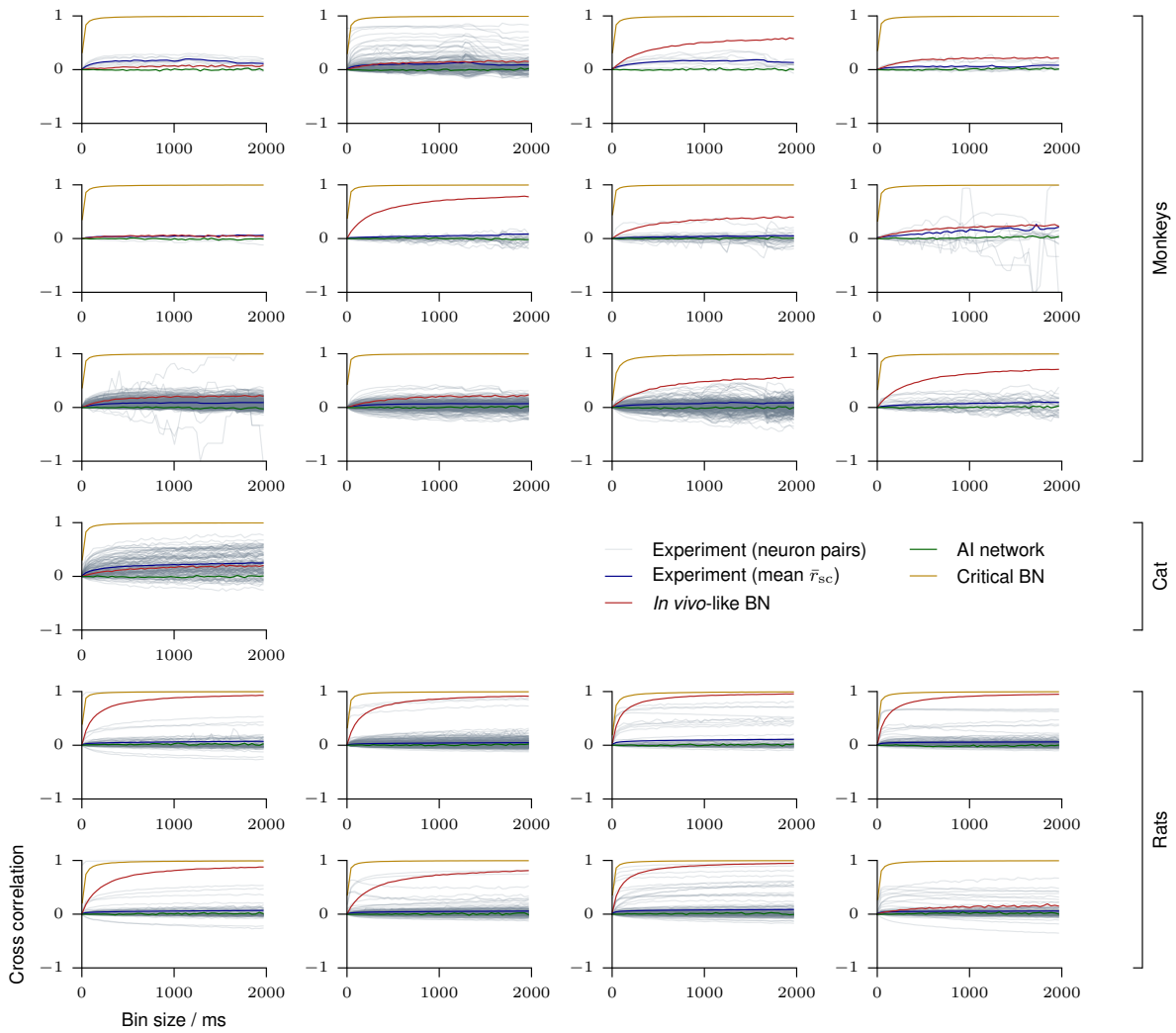
FIGURE B.4: Cross correlations for individual recording sessions. Spike count cross correlations (r_{sc}) are shown for every neuron pair (gray) and the ensemble average (blue) of each recording, for bin sizes from 1 ms to 2s. Cross correlations are also shown for the matched models, either AI (green), *in vivo*-like (red), or near critical (yellow).

94

Figure B.5: Activity distributions (4 ms bin size). Activity distributions are shown for every recording for a bin size of 4 ms (blue). Activity distributions for the matched models, either AI (green), *in vivo*-like (red), or near critical (yellow) are also shown. The color of the asterisk indicates which of the three models yielded the highest likelihood for the data following Clauset *et al.* (2009).

Figure B.6: Activity distributions (40 ms bin size). Activity distributions are shown for every recording, for a bin size of 40 ms (blue). Activity distributions for the matched models, either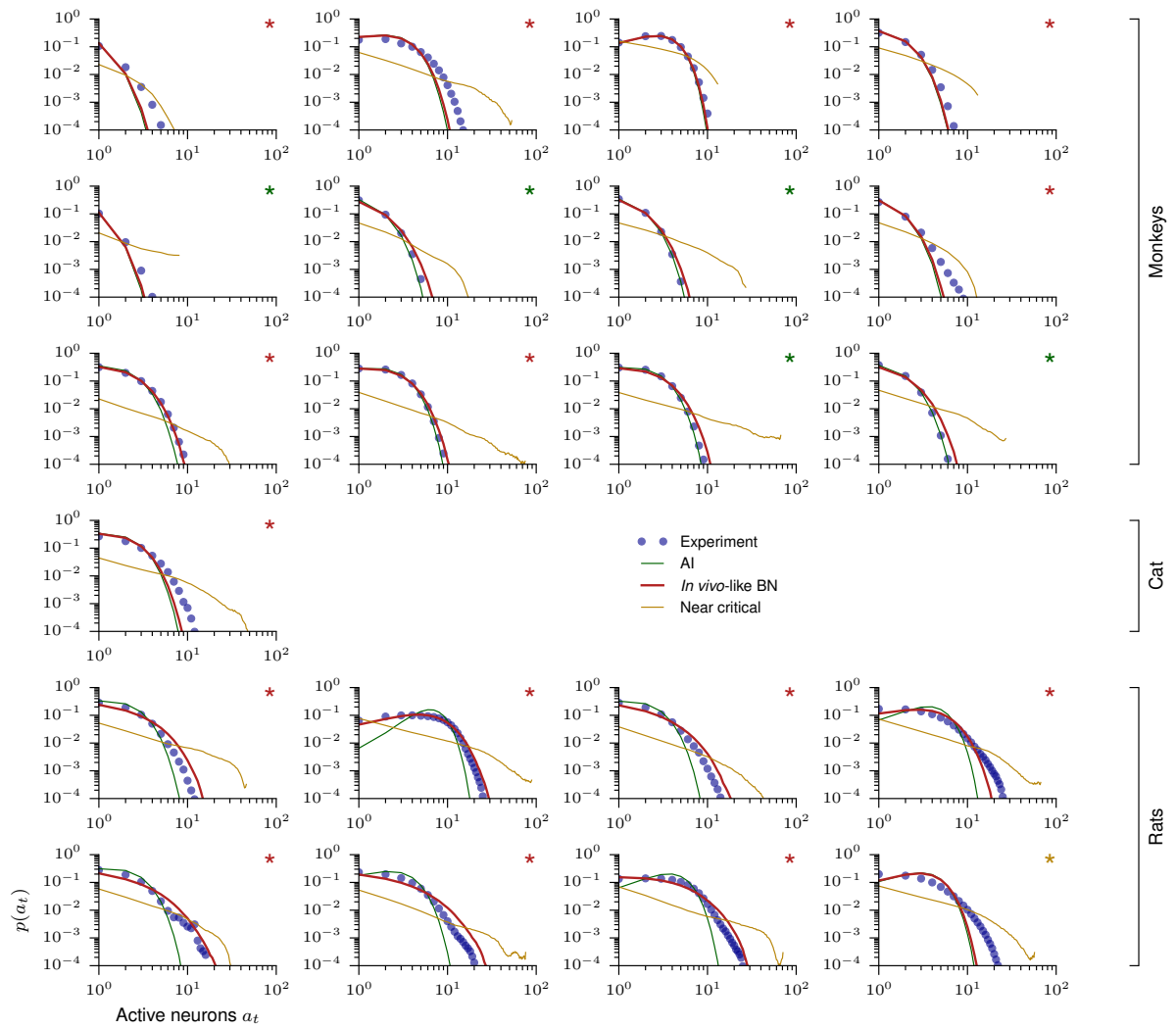 AI (green), *in vivo*-like (red), or near critical (yellow) are also shown. The color of the asterisk indicates which of the three models yielded the highest likelihood for the data following Clauset *et al.* (2009).

FIGURE B.7: Avalanche size distribution for individual recording sessions. Avalanche size distributions are shown for every recording (blue) and for matched models, either AI (green), *in vivo*-like (red), or near critical (yellow). The color of the asterisk indicates which of the three models yielded the highest likelihood for the data following Clauset *et al.* (2009).

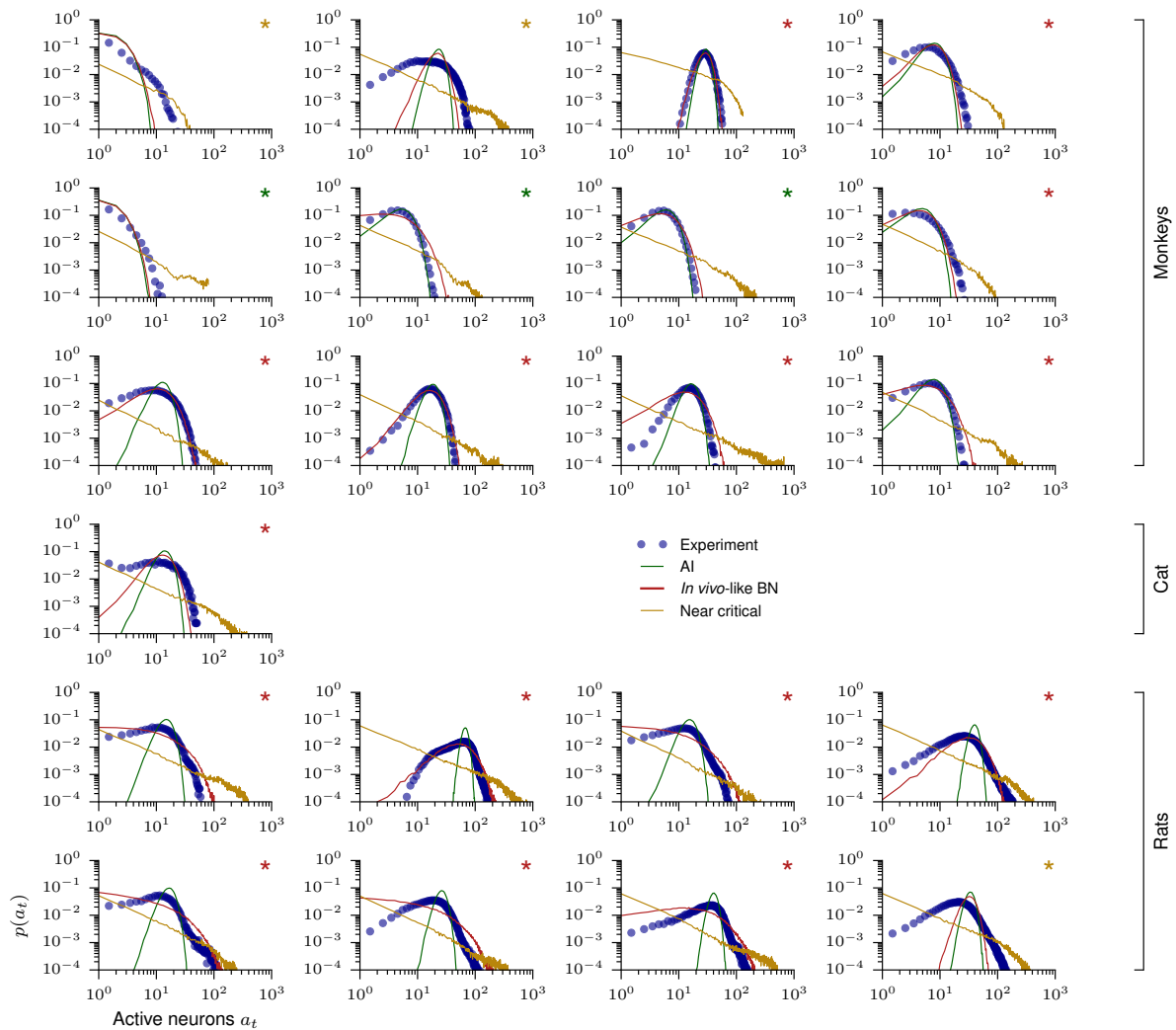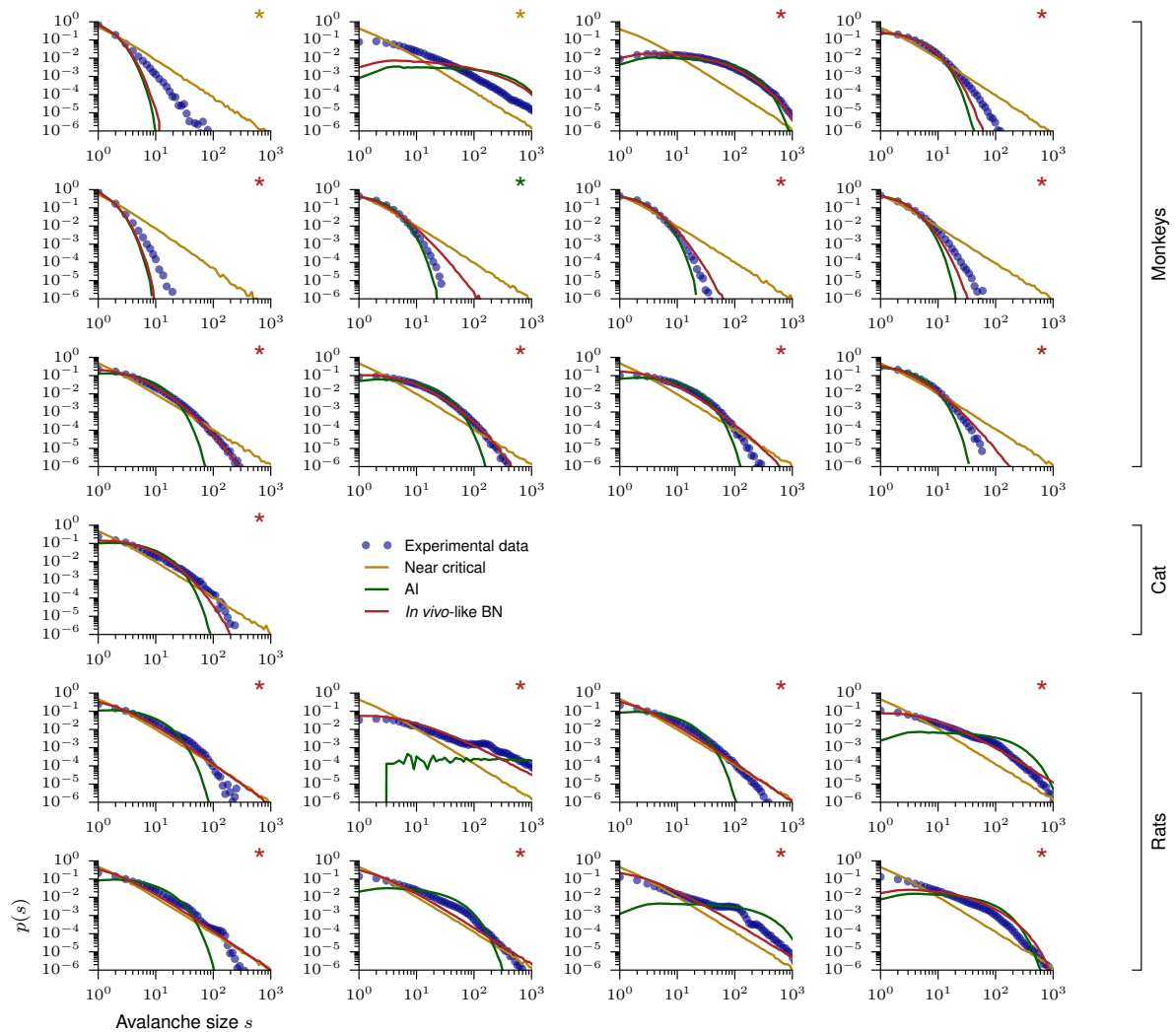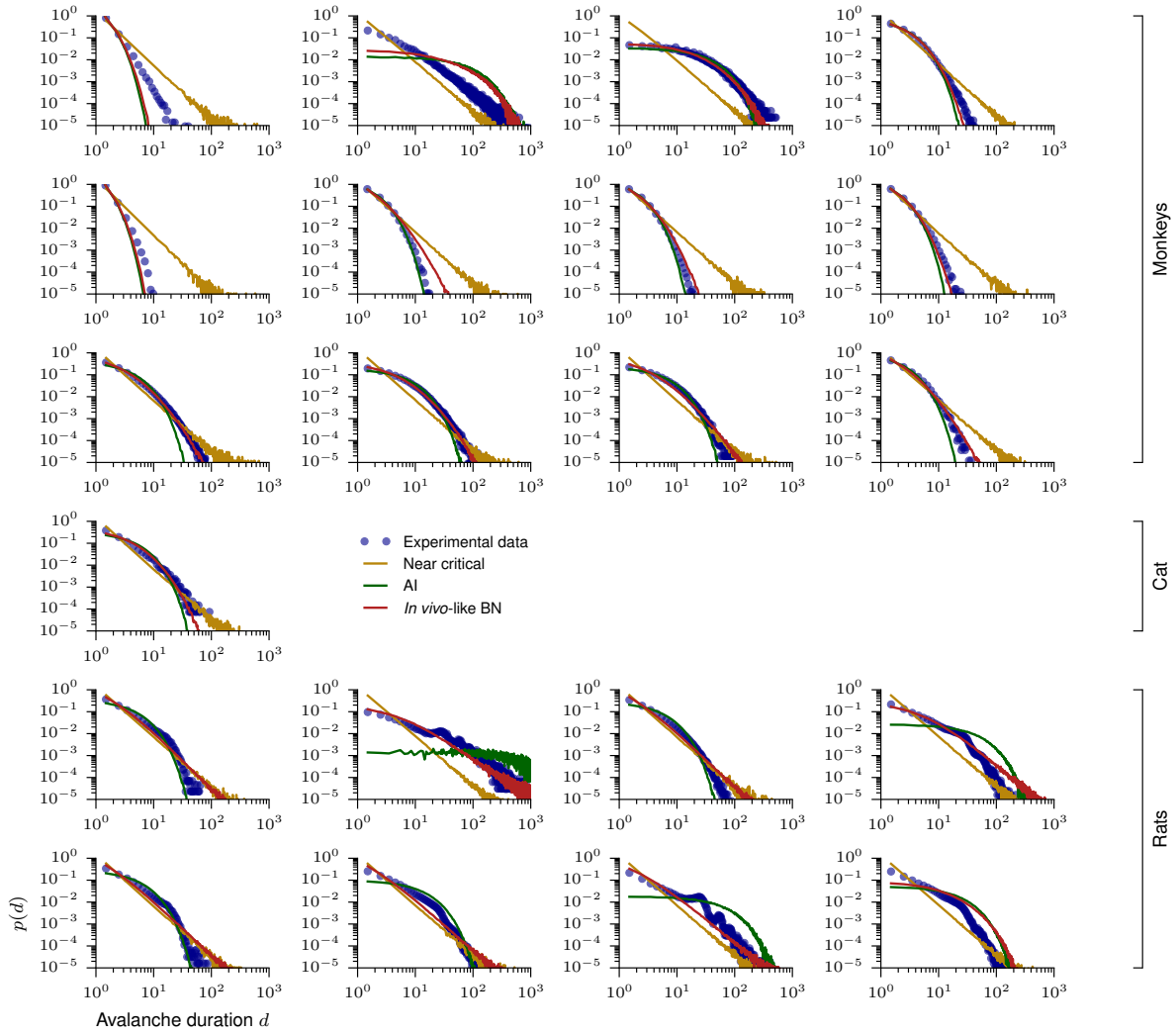**Figure B.8: Avalanche duration distribution for individual recording sessions.** Avalanche duration distributions are shown for every recording (blue) and for matched models, either AI (green), *in vivo*-like (red), or near critical (yellow).
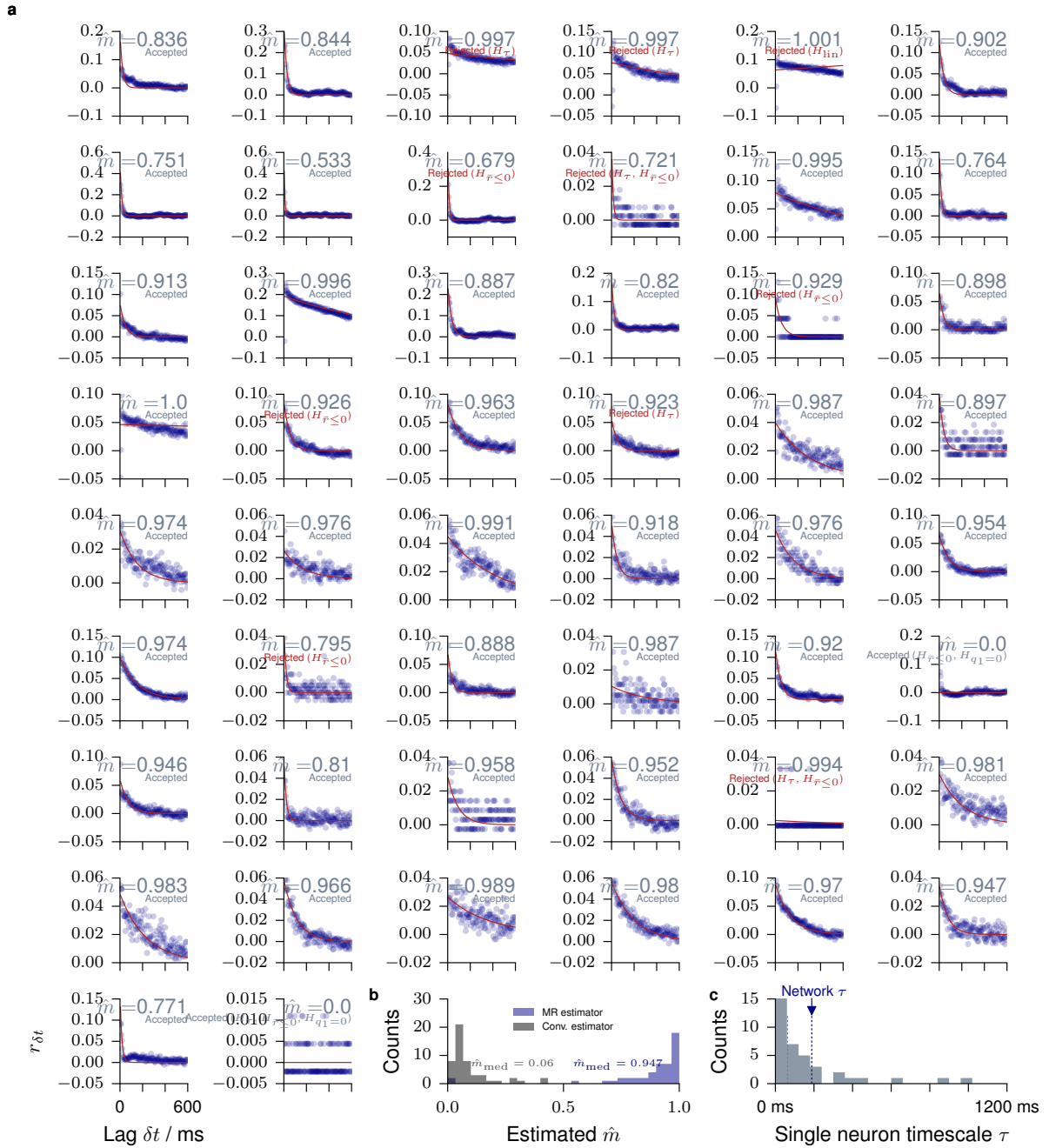
FIGURE B.9: MR estimation from single neuron activity (cat). Modified from Wilting and Priesemann (2018a). MR estimation is used to estimate $\hat{m}$ from the activity a_t of a single units in cat visual cortex. **a.** Each panel shows MR estimation for one of the 50 recorded units. Autocorrelations decay rapidly in some units, but long-term correlations are present in the activity of most units. The consistency checks are detailed in Wilting and Priesemann (2018a). **b.** Histogram of the single unit branching ratios $\hat{m}$, inferred with the conventional estimator and using MR estimation. The difference between these estimates demonstrates the subsampling bias of the conventional estimator, and how it is overcome by MR estimation. **c.** Histogram of single unit timescales with their median (gray dotted line) and the timescale of the dynamics of the whole network (blue dotted line).
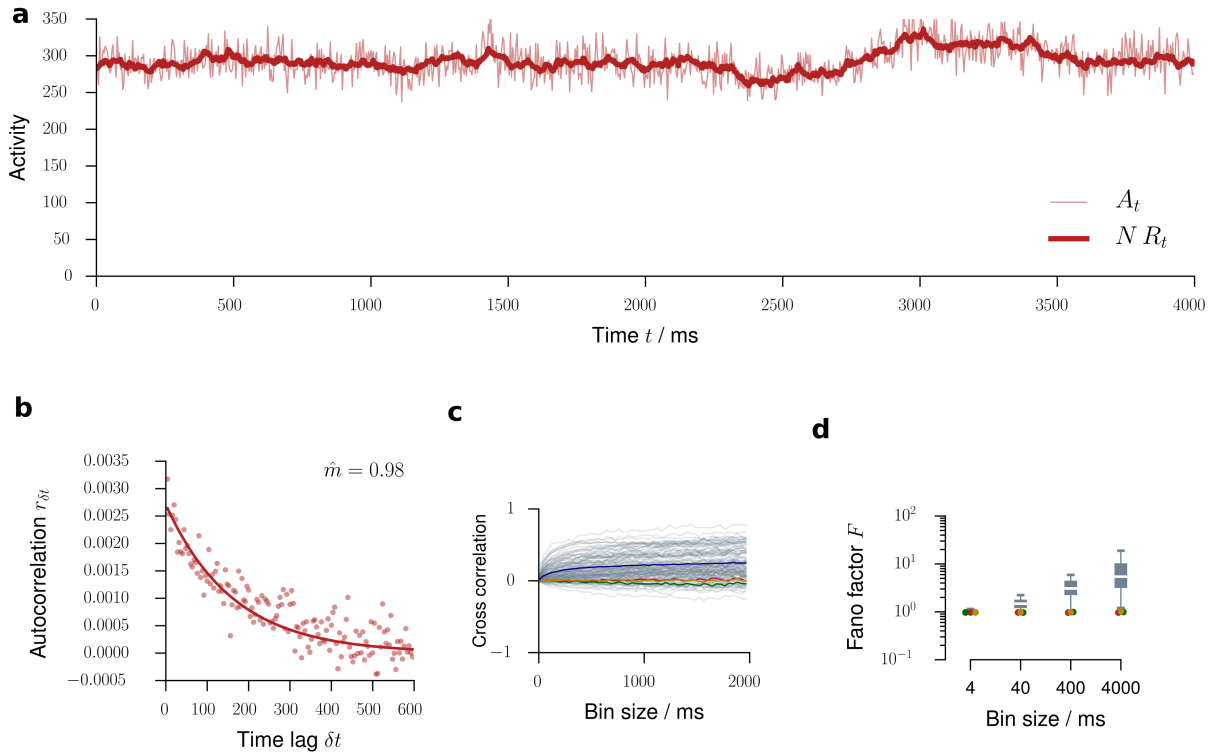
FIGURE B.10: Doubly stochastic model. Instead of a branching model, we here matched a doubly stochastic process to the data. The rates evolved according to $R_{t+1} = mR_t + h_t$ where h_t is drawn from a Poisson distribution. The actual activity is then drawn from a Poisson distribution according to $A_t \sim \text{Poi}(NR_t)$. Here, results for the experiment in cat visual cortex are shown. **a** Time evolution of R_t and A_t. As the activity is not fed back into the evolution of R_t, the second step effectively adds measurement noise to the underlying process. **b** The subsampled activity (50 out of 10,000, as in the branching models) shows the expected autocorrelation function. **c** Any of the doubly stochastic processes underestimated the spike count cross correlations. **d** Any of the doubly stochastic processes underestimated the single unit Fano factors.
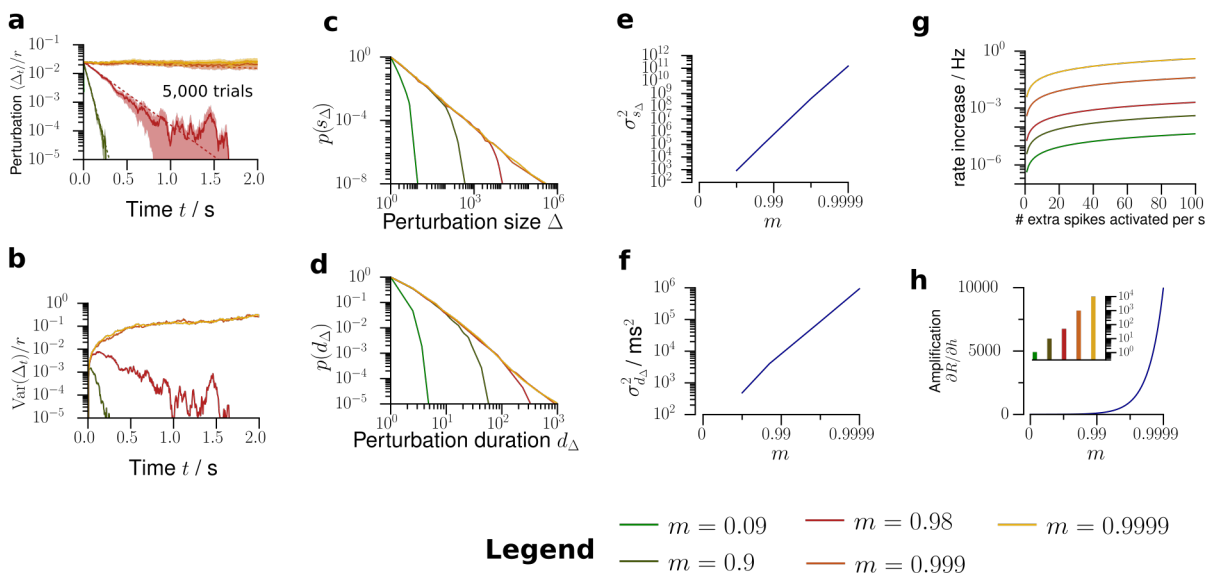
Figure B.11: Further predictions about network activity. **a.** The model predicts that the perturbation decays exponentially with decay time $\tau = -\Delta t / \log m$. **b** The variance across trials of the perturbed firing rate has a maximum, whose position depends on m. **c.** Depending on m, the model predicts the distributions for the total number of extra spikes s_Δ generated by the network following a single extra spike. **d.** Likewise, the model predicts distributions of the duration d of these perturbations. **e.** Variance of the total perturbation size as a function of m. **f.** Variance of the total perturbation duration as a function of m. **g.** Increase of the network firing rate as a function of the rate of extra neuron activations for different m. **h.** Amplification (susceptibility) dr/dh of the network as a function of the branching ratio m.

Appendix C

Appendix for "Homeostatic plasticity and external input shape neural network dynamics"[†]

C.1   Experimental details

C.1.1   Dissociated dense cultures of cortical rat neurons:

The spike-time data from dissociated cortical rat neurons of mature dense cultures was recorded by Wagenaar *et al.* (2006) and was obtained freely online (Wagenaar). The experimental setup uses multi-electrode-arrays (MEA) with $n = 59$ electrodes. Cortical cells were obtained from dissecting the anterior part of the cortex of Wistar rat embryos (E18), including somatosensory, motor, and association areas. For details, we refer to Wagenaar *et al.* (2006). Measurements were performed every day *in vitro* (DIV). We here focus on the dense case with 50 000 cells plated initially with a density of $2.5(1.5) \times 10^3$ cells/mm^2 at 1 DIV, which is compatible with standard *in vitro* experiments in the field that claim to observe critical dynamic behavior. We selected the representative recordings 8-2-34 (exp 1) and 7-2-35 (exp 2) at mature age (34/35 DIVs) for Fig. 4.1.

C.1.2   Rat hippocampus:

The spiking data from rats were recorded by Mizuseki *et al.* (2009b,a) with experimental protocols approved by the Institutional Animal Care and Use Committee of Rutgers University. The data was obtained from the NSF-founded CRCNS data sharing website (Mizuseki *et al.*, 2009a). The spikes were recorded in CA1 of the right dorsal hippocampus during an open field task. Specifically, we used the data set ec013.527 with sorted spikes from 4 shanks with $n = 31$ channels. For details we refer to Mizuseki *et al.* (2009b,a).

C.1.3 Primary visual cat cortex:

The spiking data from cats were recorded by Tim Blanche in the laboratory of Nicholas Swindale, University of British Columbia, in accordance with guidelines established by the Canadian Council for Animal Care (Blanche and Swindale, 2006; Blanche, 2009). The data was obtained from the NSF-founded CRCNS data sharing website (Blanche, 2009). Specifically, we used the data set pvc3 with recordings of $n = 50$ sorted single units (Blanche and Swindale, 2006) in area 18. For details we refer to Blanche (2009); Blanche and Swindale (2006). We confined ourselves to the experiments where no stimuli were presented such that spikes reflect the spontaneous activity in the visual cortex of mildly anesthetized cats. In order to circumvent potential non-stationarities at the beginning and end of the recording, we omitted the initial 25 s and stopped after 320 s of recording (Wilting and Priesemann, 2018a).

C.2 Analysis details

C.2.1 Spiking activity:

In order to present the spiking activity over time, we partition the time axis of experimental or numerical data into discrete bins of size Δt. For the time-discrete simulations the time bin naturally matches the time step. For experimental data we set $\Delta t = 4$ ms. In each time bin we count the total number of spikes A_t and normalize with the number of neurons N to obtain the average spiking activity $a_t = A_t / N\Delta t$. Note that experimental preparations were inevitably subsampled, as spikes were recorded only from a small number of all neurons.

C.2.2 Avalanche-size distribution:

We define the avalanche size s as the number of spikes enclosed along the discrete time axis by bins with zero activity (Beggs and Plenz, 2003). To test for criticality in terms of a branching process, one compares $P(s)$ to the expected $P(s) \sim s^{-3/2}$. This is a valid approach in the limit $h \to 0$, where avalanches can be clearly identified, and for fully sampled systems (Priesemann *et al.*, 2014). However, experiments are limited to record only from n out of N neurons. As a result, the distributions for subsampled activity $P_{\mathrm{sub}}(s)$ differ due to subsampling bias (Priesemann *et al.*, 2009, 2013). Therefore, we numerically measure both full ($n = N$) and subsampled ($n < N$) avalanche-size distributions to qualitatively compare $P(s)$ to the theory and $P_{\mathrm{sub}}(s)$ to experimental data.

C.2.3 Integrated autocorrelation time:

We measure the autocorrelation time of spiking activity a_t in terms of the integrated autocorrelation time τ_{int}, for details see, e.g., Janke (2002). In brief, we sum over the normalized autocorrelation function $C(l) = Cov[a_t, r_{t+l}]/Var[a_t]$ until the sum converges. Following conventions, we define $\tau_{\mathrm{int}} = \Delta t[\frac{1}{2} + \sum_{l=1}^{l_{\max}} C(l)]$, where $l_{\max}$ is self-consistently obtained as the minimal $l_{\max} > 6\,\tau_{\mathrm{int}}(l_{\max})$.

C.2.4 Reproducing experimental results:

We use a branching network with AA topology subject to homeostatic plasticity to quantitatively reproduce *in vivo* subsampled avalanche-size distributions. We chose networks of size $N = 10^4$ with sufficiently large homeostatic timescale $\tau_{\mathrm{hp}} = 10^5$ s. The following model parameters can be obtained from experimentally measured values: In the chosen recordings, we measured the average rate ($r_{\mathrm{cat}} \approx 7$ Hz and $r_{\mathrm{rat}} \approx 11$ Hz) as well as the subsampling corrected branching parameter (Wilting and Priesemann, 2018a) ($m_{\mathrm{cat}} \approx 0.98$ and $m_{\mathrm{rat}} \approx 0.997$ for $\Delta t = 4$ ms). In fact, the branching parameter is not suitable to identify the input rate via (4.7), because it refers to a process in discrete time steps. Since we are treating
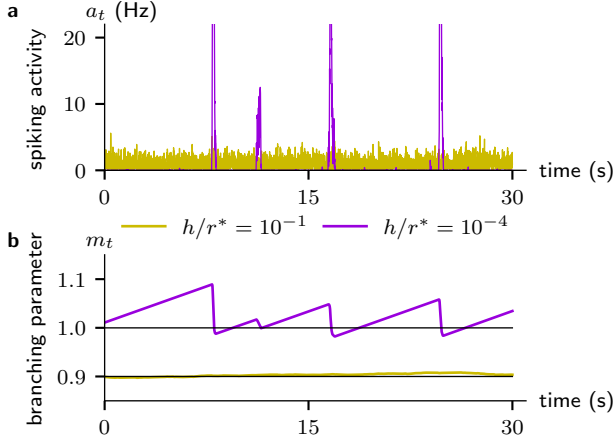
a a_t (Hz)

spiking activity

$h/r^* = 10^{-1}$ $h/r^* = 10^{-4}$

b m_t

branching parameter

FIGURE C.1: Temporal fluctuations in an annealed-average network with homeostatic plasticity subject to different external input rates. (**a**) Spiking activity a_t shows small fluctuations for large input rates (yellow) and bursts for small input rates (purple), cf. Fig. 4.3. (**b**) Branching parameter fluctuates around predicted value (black horizontal lines) and develops distinct saw-tooth pattern for small input rates.

a continuous process, the invariant quantity is the autocorrelation time ($\tau_{\text{cat}} \approx 0.2\,\text{s}$ and $\tau_{\text{rat}} \approx 1.6\,\text{s}$). According to our theory, we can then calculate the input rate per neuron $h = (1 - \exp(-\Delta t/\tau))r$. In order to avoid convergence effects, we need to choose a sufficiently small time step $\Delta t = 1\,\text{ms}$ of signal propagation (resulting in $h_{\text{cat}} \approx 3.5 \times 10^{-2}\,\text{Hz}$ and $h_{\text{rat}} \approx 5.5 \times 10^{-3}\,\text{Hz}$), while we record in time bins of $4\,\text{ms}$ to match the analysis of the experiments. Subsampled avalanche-size distributions are estimated by randomly choosing $n < N$ neurons, where we approximated n by the number of electrodes or channels ($n_{\text{cat}} = 50$ and $n_{\text{rat}} = 31$).

C.3   Approximating the dynamic state in the bursting regime

We showed in Sec. 4.4 that decreasing the external input to recurrent networks with homeostatic plasticity leads to bursting behavior (Fig. C.1**a**). This is directly related to the network branching parameter $m_t = \overline{m}_t$ no longer showing small fluctuations around the predicted value but instead exhibiting a prominent saw-tooth pattern (Fig. C.1**b**), a hallmark of the homeostatic buildup in the long pauses with no input.

We here show a semi-analytical approximation of the network branching parameter in the bursting regime. For sufficiently small external input we may assume separation of timescales, i.e., every externally induced spike drives one avalanche with periods of silence in between. Let us first consider the periods of silence, i.e., no activity per site. This holds during the entire growth period T such that (4.8) yields

$$m_t - m_{t-T} = (\Delta t r^*)\frac{T}{\tau'}. \tag{C.1}$$

The situation becomes more involved within the bursts, where the behavior of m_t is nonlinear. Consider an external spike that triggers an avalanche at $t = s$ which ends at $t = e$. Due to the separation of timescales we can assume $A_s = 1$. There are two possible scenarios: (i) The avalanche dies out before a burst can develop and (ii) the input triggers a proper burst with a macroscopic activation.

We first estimate the probability that an avalanche dies out before a burst develops. For $\tau_{\text{hp}} \gg \Delta t$ we approximate $m_t \approx m_s = \text{const}$. Then, the probability of ultimate extinction θ can be calculated as the solution of $\theta = \Pi(\theta)$ with $\Pi(\theta)$ the probability generating function (Harris, 1963). In the onset phase, the branching process is described by a Poisson process per event with mean m_s, such that $\Pi(\theta) = e^{-m_s(1-\theta)}$. We are thus looking for a solution of

$$\theta = e^{-m_s(1-\theta)}, \tag{C.2}$$

which can be rewritten to

$$-m_s\theta e^{-m_s\theta} = -m_s e^{-m_s}. \tag{C.3}$$

Figure C.2: Inter-burst-interval (IBI) distribution for annealed-average networks averaged over 12 independent simulations. Intervals are measured as times between proper burst onsets ($a_t > 20r^*$). Dashed lines show the exponential inter-spike-interval distribution of the Poisson external drive.
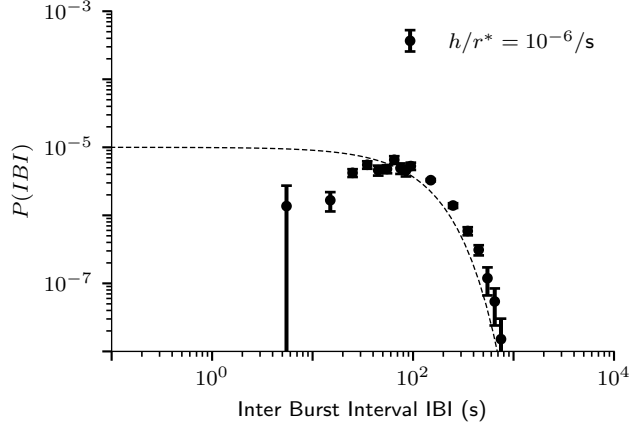
We identify the Lambert-W function $W(z)e^{W(z)} = z$ (Corless *et al.*, 1996) with $W(z) = -m_s\theta = W(-m_s e^{-m_s})$ and find for the probability that no burst develops

$$p_{\text{no–burst}}(m_s) = \theta = -\frac{1}{m_s}W\left(-m_s e^{-m_s}\right). \tag{C.4}$$

If a proper burst develops, the strong activity diminishes m_t until the burst dies out again. We cannot analytically estimate the branching parameter m_e after burst end, but we can use a deterministic numerical approximation to obtain $m_e(m_s)$. Instead of stochastically generating new (discrete) events according to some distribution $P(m_t)$ with average m_t, we approximate the branching process as deterministic (continuous) evolution $A_{t+1} = m_t A_t$. For a finite network, we need to consider convergence effects when one neuron is activated by two or more neurons at the same time. In the absence of external input, this introduces for an AA (i.e. approximating fully connected) network the activity-dependent branching parameter (Zierenberg *et al.*, 2020b,a)

$$m_t(A_t) = \frac{N}{A_t}\left(1 - \left(1 - \frac{m_t}{N}\right)^{A_t}\right), \tag{C.5}$$

which we need to consider for the activity propagation within the burst, i.e., $A_{t+1} = m_t(A_t)A_t$. In addition, we introduce an upper bound $A_t \leqslant N$. The upper limit on A_t puts a lower bound on Δm_t according to (4.8) and thus extends the duration of avalanches. Evolving $m_{t+1} = m_t + (\Delta t r^* - A_t/N)(\Delta t/\tau')$, with $m_{t+1} \geqslant 0$, we iterate until $A_e < 1$. This is a quick and numerically robust iterative scheme to estimate $m_e(m_s)$.

Putting everything together, we numerically approximate the average network branching parameter m under homeostatic plasticity in the bursting regime of low external input for an AA network. For this, we sample the external spikes (drive) as 10^4 inter-drive intervals T_s from an exponential distribution $P(T) = (1/hN)e^{-T/hN}$, corresponding to N Poisson processes with rate h. The remaining part can be interpreted as an event-based sampling with approximate transformations: Starting with $m_0 = 0$, we evolve m_t for each inter-drive interval T_s according to (C.1). If $m_t > 1$, we keep m_t with probability $p_{\text{no–burst}}(m_t)$ or else initiate a burst by setting $m_t = m_e(m_t)$. Afterwards we continue evolving m_t.

C.4 Characteristic duration of inter-burst-intervals in burst regime

In the bursting regime of low external input, the spiking activity suggests a characteristic time between bursts. In order to test for periodicity, we analyzed the distribution of inter-burst-intervals (IBI), where intervals are measured as the time between two consecutive burst onsets, defined as a spiking activity

$a_t > 20r^*$. We find (Fig. C.2) that large IBI are suppressed by the exponentially distributed inter-drive intervals (dashed lines), while short IBI are suppressed by the probability $p_{\text{no-burst}}(m)$ that a given external spike does not trigger a proper burst (Appendix C.3). This gives rise to a characteristic duration of inter-burst-intervals in the burst regime, although the dynamics are not strictly periodic.

References

L. F. Abbott and S. B. Nelson. Synaptic plasticity: Taming the beast. *Nature Neuroscience*, **3**(S11):1178 (2000).

A. A. Alzaid and M. Al-Osh. An Integer-Valued pth-Order Autoregressive Structure (INAR(p)) Process. *Journal of Applied Probability*, **27**(2):314 (1990).

O. Arviv, A. Goldstein, and O. Shriki. Near-Critical Dynamics in Stimulus-Evoked Activity of the Human Brain and Its Relation to Spontaneous Resting-State Activity. *Journal of Neuroscience*, **35**(41):13927 (2015).

J. J. Atick. Could information theory provide an ecological theory of sensory processing? *Network: Computation in neural systems*, **3**(2):213 (1992).

D. Badre and M. D'Esposito. Is the rostro-caudal axis of the frontal lobe hierarchical? *Nature Reviews Neuroscience*, **10**(9):659 (2009).

P. Bak, C. Tang, and K. Wiesenfeld. Self-organized criticality: An explanation of the 1/f noise. *Physical Review Letters*, **59**(4):381 (1987).

H. B. Barlow. Possible Principles Underlying the Transformations of Sensory Messages. In W. A. Rosenblith (ed.), *Sensory Communication*, 217—234. The MIT Press (2012).

L. Barnett, J. T. Lizier, M. Harré, A. K. Seth, and T. Bossomaier. Information flow in a kinetic ising model peaks in the disordered phase. *Physical Review Letters*, **111**(17):1 (2013).

C. Bédard, H. Kröger, and a. Destexhe. Does the 1/f frequency scaling of brain signals reflect self-organized critical states? *Physical Review Letters*, **97**(11):1 (2006).

J. M. Beggs and D. Plenz. Neuronal Avalanches in Neocortical Circuits. *Journal of Neuroscience*, **23**(35):11167 (2003).

J. M. Beggs and N. Timme. Being critical of criticality in the brain. *Frontiers in Physiology*, **3 JUN**(June):1 (2012).

A. J. Bell and T. J. Sejnowski. The 'independent components' of natural scenes are edge filters. *Vision Research*, **37**(23):3327 (1997).

T. Bellay, A. Klaus, S. Seshadri, and D. Plenz. Irregular spiking of pyramidal neurons organizes as scale-invariant neuronal avalanches in the awake state. *eLife*, **4**(JULY 2015):1 (2015).

H. Bernard, D. Werber, and M. Höhle. Estimating the under-reporting of norovirus illness in Germany utilizing enhanced awareness of diarrhoea during a large outbreak of Shiga toxin-producing E. coli O104:H4 in 2011 – a time series analysis. *BMC Infectious Diseases*, **14**(1):1 (2014).

N. Bertschinger and T. Natschläger. Real-Time Computation at the Edge of Chaos in Recurrent Neural Networks. *Neural Computation*, **16**(7):1413 (2004).

E. Bienenstock and D. Lehmann. Regulated Criticality in the Brain? *Advances in Complex Systems*, **01**(04):361 (1998).

L. Bienenstock, N. Cooper, and W. Munro. Theory for the Development of Neuron Selectivity: Orientation Specificity and Binocular Interaction in Visual Cortex . *Journal of Neuroscience Methods*, **2**(1):32 (1982).

T. Blanche. Multi-neuron recordings in primary visual cortex (2009).

T. J. Blanche and N. V. Swindale. Nyquist interpolation improves neuron yield in multiunit recordings. *Journal of Neuroscience Methods*, **155**(1):81 (2006).

J. Boedecker, O. Obst, J. T. Lizier, N. M. Mayer, and M. Asada. Information processing in echo state networks at the edge of chaos. *Theory in Biosciences*, **131**(3):205 (2012).

J. A. Bonachela, S. de Franciscis, J. J. Torres, and M. A. Muñoz. Self-organization without conservation: Are neuronal avalanches generically critical? *Journal of Statistical Mechanics*, **2010**(02):P02015 (2010).

S. Bornholdt and T. Rohlf. Topological Evolution of Dynamical Networks. *Physical Review Letters*, **84**(26):6114 (2000).

H. W. Boucher and G. R. Corey. Epidemiology of Methicillin-Resistant Staphylococcus aureus. *Clinical Infectious Diseases*, **46**(S5):S344 (2008).

G. E. P. Box. Robustness in the strategy of scientific model building. In R. L. Launer and G. N. Wilkinson (eds.), *Robustness in Statistics*, vol. 1, 201–236. Academic Press (1979).

M. Breakspear. Dynamic models of large-scale brain activity. *Nature Neuroscience*, **20**(3):340 (2017).

D. Brockmann, L. Hufnagel, and T. Geisel. The scaling laws of human travel. *Nature*, **439**(7075):462 (2006).

N. Brunel. Dynamics of networks of randomly connected excitatory and inhibitory spiking neurons. *Journal of Physiology Paris*, **94**(5-6):445 (2000).

D. Buonomano and M. Merzenich. Temporal information transformed into a spatial code by a neural network with realistic properties. *Science*, **267**(5200):1028 (1995).

B. D. Burns and A. C. Webb. The Spontaneous Activity of Neurones in the Cat's Cerebral Cortex. *Proceedings of the Royal Society B: Biological Sciences*, **194**(1115):211 (1976).

G. Buzsaki. Neuronal Oscillations in Cortical Networks. *Science*, **304**(5679):1926 (2004).

G. Buzsáki and K. Mizuseki. The log-dynamic brain: how skewed distributions affect network operations. *Nature Reviews Neuroscience*, **15**(4):264 (2014).

S. Camalet, T. Duke, F. Julicher, and J. Prost. Auditory sensitivity provided by self-tuned critical oscillations of hair cells. *Proceedings of the National Academy of Sciences*, **97**(7):3183 (2000).

J. a. G. F. Campos, A. d. A. Costa, M. Copelli, and O. Kinouchi. Correlations induced by depressing synapses in critically self-organized networks with quenched dynamics. *Physical Review E*, **95**(4):042303 (2017).

M. Carandini. Amplification of trial-to-trial response variability by neurons in visual cortex. *PLoS Biology*, **2**(9):e264 (2004).

R. Chaudhuri, K. Knoblauch, M. A. Gariel, H. Kennedy, and X. J. Wang. A Large-Scale Circuit Mechanism for Hierarchical Dynamical Processing in the Primate Cortex. *Neuron*, **88**(2):419 (2015).

J. Chen, U. Hasson, and C. J. Honey. Processing Timescales as an Organizing Principle for Primate Cortex. *Neuron*, **88**(2):244 (2015).

D. R. Chialvo. Emergent complex neural dynamics. *Nature Physics*, **6**(10):744 (2010).

C.-C. Chiang, T. P. Ladas, L. E. Gonzalez-Reyes, and D. M. Durand. Seizure Suppression by High Frequency Optogenetic Stimulation Using In Vitro and In Vivo Animal Models of Epilepsy. *Brain Stimulation*, **7**(6):890 (2014).

M. Chiappalone, M. Bove, A. Vato, M. Tedesco, and S. Martinoia. Dissociated cortical networks show spontaneously correlated activity patterns during in vitro development. *Brain Research*, **1093**(1):41 (2006).

A. Clark. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, **36**(03):181 (2013).

A. Clauset, C. Rohilla Shalizi, and M. E. J Newman. Power-Law Distributions in Empirical Data. *SIAM Review*, **51**(4):661 (2009).

W. P. Clawson, N. C. Wright, R. Wessel, and W. L. Shew. Adaptation towards scale-free dynamics improves cortical stimulus discrimination at the cost of reduced detection. *PLoS Computational Biology*, **13**(5):1 (2017).

M. R. Cohen and A. Kohn. Measuring and interpreting neuronal correlations. *Nature Neuroscience*, **14**(7):811 (2011).

R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. On the LambertW function. *Advances in Computational Mathematics*, **5**(1):329 (1996).

R. Cossart, D. Aronov, and R. Yuste. Attractor dynamics of network UP states in the neocortex. *Nature*, **423**(6937):283 (2003).

A. A. Costa, L. Brochini, and O. Kinouchi. Self-organized supercriticality and oscillations in networks of stochastic spiking neurons. *Entropy*, **19**(8):1 (2017).

A. D. A. Costa, M. Copelli, and O. Kinouchi. Can dynamical synapses produce true self-organized criticality? *Journal of Statistical Mechanics: Theory and Experiment*, **2015**(6):P06004 (2015).

L. Covolan, A.-C. G. de Almeida, B. Amorim, *et al.* Effects of Anterior Thalamic Nucleus Deep Brain Stimulation in Chronic Epileptic Rats. *PLoS ONE*, **9**(6):e97618 (2014).

H. Cuntz, F. Forstner, A. Borst, and M. Häusser. One Rule to Grow Them All: A General Theory of Neuronal Branching and Its Practical Application. *PLoS Computational Biology*, **6**(8):e1000877 (2010).

C. Curto, S. Sakata, S. Marguet, V. Itskov, and K. D. Harris. A Simple Model of Cortical Dynamics Explains Variability and State Dependence of Sensory Responses in Urethane-Anesthetized Auditory Cortex. *Journal of Neuroscience*, **29**(34):10600 (2009).

D. Dahmen, M. Diesmann, and M. Helias. Distributions of covariances as a window into the operational regime of neuronal networks. *Preprint at arXiv:160504153* (2016).

G. W. Davis. HOMEOSTATIC CONTROL OF NEURAL ACTIVITY: From Phenomenology to Molecular Design. *Annual Review of Neuroscience*, **29**(1):307 (2006).

L. de Arcangelis, C. Perrone-Capano, and H. J. Herrmann. Self-Organized Criticality Model for Brain Plasticity. *Physical Review Letters*, **96**(2):028107 (2006).

J. de Heuvel, J. Wilting, M. Becker, V. Priesemann, and J. Zierenberg. Characterizing spreading dynamics of subsampled systems with non-stationary external input. *In revision.*

M. De Pittà, N. Brunel, and A. Volterra. Astrocytes: Orchestrating synaptic plasticity? *Neuroscience*, **323**:43 (2016).

R. R. de Ruyter van Steveninck, G. D. Lewen, S. P. Strong, R. Koberle, and W. Bialek. Reproducibility and Variability in Neural Spike Trains. *Science*, **275**(5307):1805 (1997).

J. DeFelipe, L. Alonso-Nanclares, and J. I. Arellano. Microstructure of the neocortex: Comparative aspects. *Journal of Neurocytology*, **31**(3-5 SPEC. ISS.):299 (2002).

J. Dehning, N. M. Dotson, S. J. Hoffman, C. M. Gray, and V. Priesemann. Hierarchy and task-dependence of intrinsic timescales across primate cortex. *In preparation.*

B. Del Papa, V. Priesemann, and J. Triesch. Criticality meets learning: Criticality signatures in a self-organizing recurrent neural network. *PLoS ONE*, **12**(5):1 (2017).

F. R. DeLeo, M. Otto, B. N. Kreiswirth, and H. F. Chambers. Community-associated meticillin-resistant Staphylococcus aureus. *The Lancet*, **375**(9725):1557 (2010).

S. Denève and C. K. Machens. Efficient codes and balanced networks. *Nature Neuroscience*, **19**(3):375 (2016).

O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz. On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology*, **28**(4):365 (1990).

R. Douglas, C. Koch, M. Mahowald, K. Martin, and H. Suarez. Recurrent excitation in neocortical circuits. *Science*, **269**(5226):981 (1995).

D. J. Earn. A Simple Model for Complex Dynamical Transitions in Epidemics. *Science*, **287**(5453):667 (2000).

A. S. Ecker, P. Berens, G. a. Keliris, M. Bethge, N. K. Logothetis, and A. S. Tolias. Decorrelated Neuronal Firing in Cortical Microcircuits. *Science*, **327**(5965):584 (2010a).

A. S. Ecker, P. Berens, G. A. Keliris, M. Bethge, N. K. Logothetis, and A. S. Tolias. Supporting Online Material for Decorrelated Neuronal Firing in Cortical Microcircuits. *Science*, **327**(584):26 (2010b).

J. P. Eckmann, O. Feinerman, L. Gruendlinger, E. Moses, J. Soriano, and T. Tlusty. The physics of living neural networks. *Physics Reports*, **449**(1-3):54 (2007).

F. Effenberger, J. Jost, and A. Levina. Self-organization in Balanced State Networks by STDP and Homeostatic Plasticity. *PLoS Computational Biology*, **11**(9):1 (2015).

P. Erdos and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, **5**:17 (1960).

E. D. Fagerholm, M. Dinov, T. Knöpfel, and R. Leech. The characteristic patterns of neuronal avalanches in mice under anesthesia and at rest: An investigation using constrained artificial neural networks. *PLoS ONE*, **13**(5):1 (2018).

E. D. Fagerholm, R. Lorenz, G. Scott, *et al.* Cascades and Cognitive State: Focused Attention Incurs Subcritical Dynamics. *Journal of Neuroscience*, **35**(11):4626 (2015).

E. D. Fagerholm, G. Scott, W. L. Shew, *et al.* Cortical Entropy, Mutual Information and Scale-Free Dynamics in Waking Mice. *Cerebral Cortex*, **26**(10):3945 (2016).

C. P. Farrington, M. N. Kanaan, and N. J. Gay. Branching process models for surveillance of infectious diseases controlled by mass vaccination. *BioStatistics (Oxford, England)*, **4**(2):279 (2003).

T. Fekete, D. B. Omer, K. O'Hashi, A. Grinvald, C. van Leeuwen, and O. Shriki. Critical dynamics, anesthesia and information integration: Lessons from multi-scale criticality analysis of voltage imaging data. *NeuroImage*, **183**(November 2017):919 (2018).

D. J. Felleman and D. C. Van Essen. Distributed Hierarchical Processing in the Primate Cerebral Cortex. *Cerebral Cortex*, **1**(1):1 (1991).

V. Filimonov and D. Sornette. Quantifying reflexivity in financial markets: Toward a prediction of flash crashes. *Physical Review E*, **85**(5):056108 (2012).

M. F. Fong, J. P. Newman, S. M. Potter, and P. Wenner. Upward synaptic scaling is dependent on neurotransmission rather than spiking. *Nature Communications*, **6**:1 (2015).

F. Franke, M. Natora, C. Boucsein, M. H. J. Munk, and K. Obermayer. An online spike detection and spike classification algorithm capable of instantaneous resolution of overlapping spikes. *Journal of Computational Neuroscience*, **29**(1-2):127 (2010).

N. Friedman, S. Ito, B. a. W. Brinkman, *et al.* Universal Critical Dynamics in High Resolution Neuronal Avalanche Data. *Physical Review Letters*, **108**(20):208102 (2012).

P. Fries, D. Nikolić, and W. Singer. The gamma cycle. *Trends in Neurosciences*, **30**(7):309 (2007).

K. Friston. The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, **11**(2):127 (2010).

F. Frohlich, M. Bazhenov, and T. J. Sejnowski. Pathological Effect of Homeostatic Synaptic Scaling on Network Dynamics in Diseases of the Cortex. *Journal of Neuroscience*, **28**(7):1709 (2008).

A. Galka, K. Wong, and T. Ozaki. Generalized state-space models for modeling nonstationary EEG time-series. In *Modeling Phase Transitions in the Brain*, 27–52. Springer New York, New York, NY (2010).

Z. Ghahramani and G. E. Hinton. *Parameter estimation for linear dynamical systems.* Technical Report, University of Toronto (1996).

M. Girardi-Schappo, O. Kinouchi, and M. H. R. Tragtenberg. Critical avalanches and subsampling in map-based neural networks coupled with noisy synapses. *Physical Review E*, **88**(2):1 (2013).

J. Gjorgjieva, J. F. Evers, and S. J. Eglen. Homeostatic Activity-Dependent Tuning of Recurrent Networks for Robust Propagation of Activity. *Journal of Neuroscience*, **36**(13):3722 (2016).

M. Goard and Y. Dan. Basal forebrain activation enhances cortical coding of natural scenes. *Nature Neuroscience*, **12**(11):1444 (2009).

L. L. Gollo. Coexistence of critical sensitivity and subcritical specificity can yield optimal population coding. *Journal of The Royal Society Interface*, **14**(134):20170207 (2017).

C. M. Gray. Synchronous oscillations in neuronal systems: Mechanisms and functions. *Journal of Computational Neuroscience* **1**:11 (1994).

C. M. Gray, P. König, A. K. Engel, and W. Singer. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, **338**(6213):334 (1989).

M. Gur, A. Beylin, and S. DM. Response variability of neurons in primary visual cortex (V1) of alert monkeys. TL - 17. *Journal of Neuroscience*, **17**(8):2914 (1997).

A. Hagemann, J. Wilting, B. Samimizad, F. Mormann, and V. Priesemann. No evidence that epilepsy impacts criticality in pre-seizure single-neuron activity of human cortex. *In revision.*

G. Hahn, T. Petermann, M. N. Havenith, *et al.* Neuronal avalanches in spontaneous activity in vivo. *Journal of Neurophysiology*, **104**(6):3312 (2010).

G. Hahn, A. Ponce-Alvarez, C. Monier, *et al.* Spontaneous cortical activity is transiently poised close to criticality. *PLoS Computational Biology*, **13**(5):1 (2017).

C. Haldeman and J. Beggs. Critical Branching Captures Activity in Living Neural Networks and Maximizes the Number of Metastable States. *Physical Review Letters*, **94**(5):058101 (2005).

J. D. Hamilton. *Time series analysis*, vol. 2. Princeton university press Princeton (1994).

D. Hansel and C. van Vreeswijk. The Mechanism of Orientation Selectivity in Primary Visual Cortex without a Functional Map. *Journal of Neuroscience*, **32**(12):4049 (2012).

D. Harnack, M. Pelko, A. Chaillet, Y. Chitour, and M. C. van Rossum. Stability of Neuronal Networks with Homeostatic Regulation. *PLoS Computational Biology*, **11**(7):1 (2015).

K. D. Harris and T. D. Mrsic-Flogel. Cortical connectivity and sensory coding. *Nature*, **503**(7474):51 (2013).

K. D. Harris and A. Thiele. Cortical state and attention. *Nature Reviews Neuroscience*, **12**(9):509 (2011).

T. E. Harris. *The Theory of Branching Processes.* Springer Berlin (1963).

U. Hasson, J. Chen, and C. J. Honey. Hierarchical process memory: Memory as an integral component of information processing. *Trends in Cognitive Sciences*, **19**(6):304 (2015).

U. Hasson, E. Yang, I. Vallines, D. J. Heeger, and N. Rubin. A Hierarchy of Temporal Receptive Windows in Human Cortex. *Journal of Neuroscience*, **28**(10):2539 (2008).

A. M. Hauri, H.-J. Westbrock, H. Claus, *et al.* Electronic Outbreak Surveillance in Germany: A First Evaluation for Nosocomial Norovirus Outbreaks. *PLoS ONE*, **6**(3):e17341 (2011).

C. R. Heathcote. A Branching Process Allowing Immigration. *Journal of the Royal Statistical Society B*, **27**(1):138 (1965).

W. Hellenbrand, A. Siedler, A. Tischer, *et al.* Progress toward measles elimination in Germany. *Journal of infectious diseases*, **187 Suppl**(Suppl 1):S208 (2003).

P. J. Hellyer, B. Jachs, C. Clopath, and R. Leech. Local inhibitory plasticity tunes macroscopic brain dynamics and allows the emergence of functional brain networks. *NeuroImage*, **124**:85 (2016).

K. B. Hengen, A. Torrado Pacheco, J. N. McGregor, S. D. Van Hooser, and G. G. Turrigiano. Neuronal Firing Rate Homeostasis Is Inhibited by Sleep and Promoted by Wake. *Cell*, **165**(1):180 (2016).

G. Hennequin, E. J. Agnes, and T. P. Vogels. Inhibitory Plasticity: Balance, Control, and Codependence. *Annual Review of Neuroscience*, **40**(1):557 (2017).

G. Hennequin, T. P. Vogels, and W. Gerstner. Optimal control of transient dynamics in balanced networks supports generation of complex movements. *Neuron*, **82**(6):1394 (2014).

L. Hernández-Navarro, J. G. Orlandi, B. Cerruti, E. Vives, and J. Soriano. Dominance of Metric Correlations in Two-Dimensional Neuronal Cultures Described through a Random Field Ising Model. *Physical Review Letters*, **118**(20):208101 (2017).

V. Hernandez-Urbina and J. M. Herrmann. Self-organized Criticality via Retro-Synaptic Signals. *Frontiers in Physics*, **4** (2017).

A. V. M. Herz and J. J. Hopfield. Earthquake Cycles and Neural Reverberations: Collective Oscillations in Systems with Pulse-Coupled Threshold Elements. *Physical Review Letters*, **75**(6):1222 (1995).

C. C. Heyde and E. Seneta. Estimation Theory for Growth and Immigration Rates in a Multiplicative Process. *Journal of Applied Probability*, **9**(2):235 (1972).

J. Hirsch and C. Gilbert. Synaptic physiology of horizontal connections in the cat's visual cortex. *Journal of Neuroscience*, **11**(6):1800 (1991).

D. Holcman and M. Tsodyks. The Emergence of Up and Down States in Cortical Networks. *PLoS Comput Biol*, **2**(3):e23 (2006).

C. Huang and B. Doiron. Once upon a (slow) time in the land of recurrent neuronal networks…. *Current Opinion in Neurobiology*, **46**:31 (2017).

J. Humplik and G. Tkačik. Probabilistic models for neural populations that naturally capture global coupling and criticality. *PLoS Computational Biology*, **13**(9):1 (2017).

A. Hyvärinen and E. Oja. Independent component analysis: Algorithms and applications. *Neural Networks*, **13**:411 (2000).

M. Ispány and G. Pap. Critical branching processes with immigration. In M. G. Velasco, I. M. Puerto, R. Martínez, M. Molina, M. Mota, and A. Ramos (eds.), *Lecture Notes in Statistics*, vol. 197, 135–146. Springer Berlin Heidelberg (2010).

H. Jaeger and H. Haas. Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication. *Science*, **304**(5667):78 (2004).

H. Jaeger, W. Maass, and J. Principe. Special issue on echo state networks and liquid state machines. *Neural Networks*, **20**(3):287 (2007).

W. Janke. Statistical analysis of simulations: Data correlations and error estimation. *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*, **10**:423 (2002).

M. Kachour and J. F. Yao. First-order rounded integer-valued autoregressive (RINAR(1)) process. *Journal of Time Series Analysis*, **30**(4):417 (2009).

J. Kadmon and H. Sompolinsky. Transition to chaos in random neuronal networks. *Physical Review X*, **5**(4):1 (2015).

M. Kaiser and J. Simonotto. Limited spreading: How hierarchical networks prevent the transition to the epileptic state. In *Modeling Phase Transitions in the Brain*, 99–116. Springer New York, New York, NY (2010).

P. Kara, P. Reinagel, and R. C. Reid. Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. *Neuron*, **27**(3):635 (2000).

M. M. Karnani, J. Jackson, I. Ayzenshtat, *et al.* Opening Holes in the Blanket of Inhibition: Localized Lateral Disinhibition by VIP Interneurons. *Journal of Neuroscience*, **36**(12):3471 (2016).

E. T. Kavalali. The mechanisms and functions of spontaneous neurotransmitter release. *Nature Reviews Neuroscience*, **16**(1):5 (2014).

T. Keck, M. Hübener, and T. Bonhoeffer. Interactions between synaptic homeostatic mechanisms: an attempt to reconcile BCM theory, synaptic scaling, and changing excitation/inhibition balance. *Current Opinion in Neurobiology*, **43**:87 (2017).

J. F. Kenney and E. S. Keeping. Linear regression and correlation. *Mathematics of Statistics*, **1**:252 (1962).

H. Kesten. Random difference equations and Renewal theory for products of random matrices. *Acta Mathematica*, **131**(1):207 (1973).

M. Kimmel and D. E. Axelrod. *Branching Processes in Biology*, *Interdisciplinary Applied Mathematics*, vol. 19. Springer New York, New York, NY (2015).

O. Kinouchi and M. Copelli. Optimal dynamical range of excitable networks at criticality. *Nature Physics*, **2**(5):348 (2006).

M. a. Kisley and G. L. Gerstein. Trial-to-trial variability and state-dependent modulation of auditory-evoked responses in cortex. *Journal of Neuroscience*, **19**(23):10451 (1999).

R. Köck, A. Mellmann, F. Schaumburg, A. W. Friedrich, F. Kipp, and K. Becker. The epidemiology of methicillin-resistant Staphylococcus aureus (MRSA) in Germany. *Deutsches Ärzteblatt international*, **108**(45):761 (2011).

A. Kohn, A. Zandvakili, and M. A. Smith. Correlations and brain states: from electrophysiology to functional imaging. *Current Opinion in Neurobiology*, **19**(4):434 (2009).

F. Y. K. Kossio, S. Goedeke, B. van den Akker, B. Ibarz, and R.-M. Memmesheimer. Growing Critical: Self-Organized Criticality in a Developing Neural System. *Physical Review Letters*, **121**(5):058301 (2018).

R. Kubo. Statistical-mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems. *Journal of the Physical Society of Japan*, **12**(6):570 (1957).

T. P. Ladas, C.-C. Chiang, L. E. Gonzalez-Reyes, T. Nowak, and D. M. Durand. Seizure reduction through interneuron-mediated entrainment using low frequency optical stimulation. *Experimental Neurology*, **269**:120 (2015).

M. Larkum. A cellular mechanism for cortical associations: An organizing principle for the cerebral cortex. *Trends in Neurosciences*, **36**(3):141 (2013).

A. Lazar. SORN: a Self-organizing Recurrent Neural Network. *Frontiers in Computational Neuroscience*, **3**(October):1 (2009).

H. Lee, D. Golkowski, D. Jordan, *et al.* Relationship of critical dynamics, functional connectivity, and states of consciousness in large-scale human brain networks. *NeuroImage*, **188**(December 2018):228 (2019).

M. Lenz, S. Platschek, V. Priesemann, *et al.* Repetitive magnetic stimulation induces plasticity of excitatory postsynapses on proximal dendrites of cultured mouse CA1 pyramidal neurons. *Brain Structure and Function*, **220**(6):3323 (2015).

A. Lerchner and P. E. Latham. A unifying framework for understanding state-dependent network dynamics in cortex. *Preprint at arXiv:151100411 Q-Bio* (2015).

A. Levina, J. Herrmann, and T. Geisel. Erratum: Phase Transitions towards Criticality in a Neural System with Adaptive Interactions [Phys. Rev. Lett. 102, 118110 (2009)]. *Physical Review Letters*, **103**(6):069901 (2009a).

A. Levina, J. Herrmann, and T. Geisel. Phase Transitions towards Criticality in a Neural System with Adaptive Interactions. *Physical Review Letters*, **102**(11):118110 (2009b).

a. Levina, J. M. Herrmann, and T. Geisel. Dynamical synapses causing self-organized criticality in neural networks. *Nature Physics*, **3**(12):857 (2007).

A. Levina and V. Priesemann. Subsampling scaling. *Nature Communications*, **8**(May):15140 (2017).

J. Lian, M. Bikson, C. Sciortino, W. C. Stacey, and D. M. Durand. Local Suppression of Epileptiform Activity by Electrical Stimulation in Rat Hippocampus *In Vitro*. *Journal of Physiology*, **547**(2):427 (2003).

S. Lim and M. S. Goldman. Balanced cortical microcircuitry for maintaining information in working memory. *Nature Neuroscience*, **16**(9):1306 (2013).

D. V. Lissin, S. N. Gomperts, R. C. Carroll, *et al.* Activity differentially regulates the surface expression of synaptic AMPA and NMDA glutamate receptors. *Proceedings of the National Academy of Sciences of the United States of America*, **95**(12):7097 (1998).

A. Litwin-kumar and B. Doiron. Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nature Neuroscience*, **15**(11):1498 (2012).

A. Litwin-Kumar and B. Doiron. Formation and maintenance of neuronal assemblies through synaptic plasticity. *Nature Communications*, **5**(1):5319 (2014).

M. London and M. Häusser. Dendritic Computation. *Annual Review of Neuroscience*, **28**(1):503 (2005).

M. London, A. Roth, L. Beeren, M. Häusser, and P. E. Latham. Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex. *Nature*, **466**(7302):123 (2010).

W. Maass, T. Natschläger, and H. Markram. Real-Time Computing Without Stable States: A New Framework for Neural Computation Based on Perturbations. *Neural Computation*, **14**(11):2531 (2002).

S. L. Marguet and K. D. Harris. State-Dependent Representation of Amplitude-Modulated Noise Stimuli in Rat Auditory Cortex. *Journal of Neuroscience*, **31**(17):6414 (2011).

H. Markram and M. Tsodyks. Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature*, **382**(6594):807 (1996).

S. Marom. Neural timescales or lack thereof. *Progress in Neurobiology*, **90**(1):16 (2010).

M. Martinello, J. Hidalgo, A. Maritan, S. Di Santo, D. Plenz, and M. A. Muñoz. Neutral theory and scale-free neural dynamics. *Physical Review X*, **7**(4):1 (2017).

M. Massimini, F. Ferrarelli, R. Huber, S. K. Esser, H. Singh, and G. Tononi. Breakdown of Cortical Effective Connectivity During Sleep. *Science*, **309**(5744):2228 (2005).

C. Meisel, K. Bailey, P. Achermann, and D. Plenz. Decline of long-range temporal correlations in the human brain during sustained wakefulness. *Scientific Reports*, **7**(1):11825 (2017a).

C. Meisel, A. Klaus, V. V. Vyazovskiy, and D. Plenz. The Interplay between Long- and Short-Range Temporal Correlations Shapes Cortex Dynamics across Vigilance States. *Journal of Neuroscience*, **37**(42):10114 (2017b).

C. Meisel, A. Storch, S. Hallmeyer-Elgner, E. Bullmore, and T. Gross. Failure of adaptive self-organized criticality during epileptic seizure attacks. *PLoS Computational Biology*, **8**(1):e1002312 (2012).

L. Michiels van Kessenich, L. de Arcangelis, and H. J. Herrmann. Synaptic plasticity and neuronal refractory time cause scaling behaviour of neuronal avalanches. *Scientific Reports*, **6**(1):32071 (2016).

K. D. Miller. Canonical computations of cerebral cortex. *Current Opinion in Neurobiology*, **37**:75 (2016).

K. D. Miller and D. J. C. MacKay. The Role of Constraints in Hebbian Learning. *Neural Computation*, **6**(1):100 (1994).

D. Millman, S. Mihalas, A. Kirkwood, and E. Niebur. Self-organized criticality occurs in non-conservative neuronal networks during 'up' states. *Nature Physics*, **6**(10):801 (2010).

D. Miner and J. Triesch. Plasticity-Driven Self-Organization under Topological Constraints Accounts for Non-random Features of Cortical Synaptic Wiring. *PLOS Computational Biology*, **12**(2):e1004759 (2016).

G. K. Mitov, S. T. Rachev, Y. S. Kim, and F. J. Fabozzi. Barrier Option Pricing by Branching Processes. *International Journal of Theoretical and Applied Finance*, **12**(07):1055 (2009).

K. Mizuseki, A. Sirota, E. Pastalkova, and G. Buzsáki. Multi-unit recordings from the rat hippocampus made during open field foraging (2009a).

K. Mizuseki, A. Sirota, E. Pastalkova, and G. Buzsáki. Theta Oscillations Provide Temporal Windows for Local Circuit Computation in the Entorhinal-Hippocampal Loop. *Neuron*, **64**(2):267 (2009b).

S. Morefield, E. Keefer, K. Chapman, and G. Gross. Drug evaluations using neuronal networks cultured on microelectrode arrays. *Biosensors and Bioelectronics*, **15**(7-8):383 (2000).

M. A. Muñoz. Colloquium : Criticality and dynamical scaling in living systems. *Reviews of Modern Physics*, **90**(3):031001 (2018).

B. K. Murphy and K. D. Miller. Balanced Amplification: A New Mechanism of Selective Amplification of Neural Activity Patterns. *Neuron*, **61**(4):635 (2009).

J. D. Murray, A. Bernacchia, D. J. Freedman, *et al.* A hierarchy of intrinsic timescales across primate cortex. *Nature Neuroscience*, **17**(12):1661 (2014).

J. Naude, B. Cessac, H. Berry, and B. Delord. Effects of Cellular Homeostatic Intrinsic Plasticity on Dynamical and Computational Properties of Biological Recurrent Neural Networks. *Journal of Neuroscience*, **33**(38):15032 (2013).

R. J. O'Brien, S. Kamboj, M. D. Ehlers, K. R. Rosen, G. D. Fischbach, and R. L. Huganir. Activity-dependent modulation of synaptic AMPA recdptor accumulation. *Neuron*, **21**(1989):1067 (1998).

M. Okun and I. Lampl. Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nature Neuroscience*, **11**(5):535 (2008).

J. G. Orlandi, J. Soriano, E. Alvarez-Lacalle, S. Teller, and J. Casademunt. Noise focusing and the emergence of coherent activity in neuronal cultures. *Nature Physics*, **9**(9):582 (2013).

S. Ostojic. Two types of asynchronous activity in networks of excitatory and inhibitory spiking neurons. *Nature Neuroscience*, **17**(4):594 (2014).

A. G. Pakes. Branching Processes with Immigration. *Journal of Applied Probability*, **8**(1):32 (1971).

L. Papoz, B. Balkau, and J. Lellouch. Case counting in epidemiology: Limitations of methods based on multiple data sources. *International Journal of Epidemiology*, **25**(3):474 (1996).

V. Pasquale, P. Massobrio, L. L. Bologna, M. Chiappalone, and S. Martinoia. Self-organization and neuronal avalanches in networks of dissociated cortical neurons. *Neuroscience*, **153**(4):1354 (2008).

A. Pazy and P. Rabinowitz. On a branching process in neutron transport theory. *Archive for Rational Mechanics and Analysis*, **51**(2):153 (1973).

Y. Penn, M. Segal, and E. Moses. Network synchronization in hippocampal neurons. *Proceedings of the National Academy of Science USA*, **113**(12):3341 (2016).

V. Piech, W. Li, G. N. Reeke, and C. D. Gilbert. Network model of top-down influences on local gain and contextual interactions in visual cortex. *Proceedings of the National Academy of Sciences*, **110**(43):E4108 (2013).

J. Pinheiro Neto, P. F. Spitzner, and V. Priesemann. A unified picture of neuronal avalanches arises from the understanding of samplingeffects. *In review (Preprint at arXiv:191009984)*.

G. Pipa, E. S. Städtler, E. F. Rodriguez, *et al.* Performance- and stimulus-dependent oscillations in monkey pre-frontal cortex during short-term memory. *Frontiers in Integrative Neuroscience*, **3**(October):25 (2009).

D. Plenz and E. Niebur (eds.). *Criticality in Neural Systems*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany (2014).

S.-S. Poil, R. Hardstone, H. D. Mansvelder, and K. Linkenkaer-Hansen. Critical-State Dynamics of Avalanches and Oscillations Jointly Emerge from Balanced Excitation/Inhibition in Neuronal Networks. *Journal of Neuroscience*, **32**(29):9817 (2012).

J. F. A. Poulet and C. C. H. Petersen. Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice. *Nature*, **454**(7206):881 (2008).

K. Pozo and Y. Goda. Unraveling mechanisms of homeostatic synaptic plasticity. *Neuron*, **66**(3):337 (2010).

V. Priesemann, A. Levina, and J. Wilting. Assessing Criticality in Experiments. In E. N. Tomen, M. Herrmann, and U. Ernst (eds.), *The Functional Role of Critical Dynamics in Neural Systems*, 199–232. Springer, Cham (2019).

V. Priesemann, M. H. J. Munk, and M. Wibral. Subsampling effects in neuronal avalanche distributions recorded in vivo. *BMC Neuroscience*, **10**:40 (2009).

V. Priesemann and O. Shriki. Can a time varying external drive give rise to apparent criticality in neural systems? *PLOS Computational Biology*, **14**(5):e1006081 (2018).

V. Priesemann, M. Valderrama, M. Wibral, and M. Le Van Quyen. Neuronal avalanches differ from wakefulness to deep sleep–evidence from intracranial depth recordings in humans. *PLoS Computational Biology*, **9**(3):e1002985 (2013).

V. Priesemann, M. Wibral, M. Valderrama, *et al.* Spike avalanches in vivo suggest a driven, slightly subcritical brain state. *Frontiers in Systems Neuroscience*, **8**(June):108 (2014).

R. Pröpper and K. Obermayer. Spyke Viewer: a flexible and extensible platform for electrophysiological data analysis. *Frontiers in Neuroinformatics*, **7**(November):26 (2013).

M. Quagliariello. *Stress-testing the banking system: methodologies and applications*. Cambridge University Press, New York (2009).

N. Ramalingam, J. N. J. McManus, W. Li, and C. D. Gilbert. Top-Down Modulation of Lateral Interactions in Visual Cortex. *Journal of Neuroscience*, **33**(5):1773 (2013).

H. P. Rang, M. M. Dale, J. M. Ritter, and P. K. Moore. *Pharmacology*. Churchill Livingstone, Edinburgh, 5th ed. (2003).

R. P. N. Rao and D. H. Ballard. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, **2**(1):79 (1999).

K. Reinhold, A. D. Lien, and M. Scanziani. Distinct recurrent versus afferent dynamics in cortical visual processing. *Nature Neuroscience*, **18**(12):1789 (2015).

A. Renart, J. de la Rocha, P. Bartho, *et al.* The Asynchronous State in Cortical Circuits. *Science*, **327**(5965):587 (2010).

T. L. Ribeiro, M. Copelli, F. Caixeta, *et al.* Spike Avalanches Exhibit Universal Dynamics across the Sleep-Wake Cycle. *PLoS ONE*, **5**(11):e14129 (2010).

T. L. Ribeiro, S. Ribeiro, H. Belchior, F. Caixeta, and M. Copelli. Undersampled critical branching processes on small-world and random networks fail to reproduce the Statistics of spike avalanches. *PloS one*, **9**(4):e94992 (2014).

Robert-Koch-Institute. SurvStat@RKI 2.0 ().

H. P. Robinson, M. Kawahara, Y. Jimbo, K. Torimitsu, Y. Kuroda, and A. Kawana. Periodic synchronized bursting and intracellular calcium transients elicited by low magnesium in cultured cortical neurons. *Journal of Neurophysiology*, **70**(4):1606 (1993).

D. B. Rubin, S. D. VanHooser, and K. D. Miller. The stabilized supralinear network: A unifying circuit motif underlying multi-input integration in sensory cortex. *Neuron*, **85**(2):402 (2015).

M. Scheffer, S. R. Carpenter, T. M. Lenton, *et al.* Anticipating critical transitions. *Science*, **338**:344 (2012).

U. D. Schiller and J. J. Steil. Analyzing the weight dynamics of recurrent learning algorithms. *Neurocomputing*, **63**(SPEC. ISS.):5 (2005).

M. L. Scholvinck, A. B. Saleem, A. Benucci, K. D. Harris, and M. Carandini. Cortical State Determines Global Variability and Correlations in Visual Cortex. *Journal of Neuroscience*, **35**(1):170 (2015).

J. Sethna. *Statistical mechanics: entropy, order parameters, and complexity*. Oxford University Press, Oxford (2006).

W. L. Shew and D. Plenz. The functional benefits of criticality in the cortex. *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry*, **19**(1):88 (2013).

W. L. Shew, H. Yang, T. Petermann, R. Roy, and D. Plenz. Neuronal avalanches imply maximum dynamic range in cortical networks at criticality. *Journal of Neuroscience*, **29**(49):15595 (2009).

O. Shriki, J. Alstott, F. Carver, *et al.* Neuronal avalanches in the resting MEG of the human brain. *Journal of Neuroscience*, **33**(16):7079 (2013).

O. Shriki and D. Yellin. Optimal Information Representation and Criticality in an Adaptive Sensory Recurrent Neuronal Network. *PLoS Computational Biology*, **12**(2):1 (2016).

R. H. Shumway and D. S. Stoffer. An approach to time series smoothing and forecasting using the EM algorithm. *Journal of Time Series Analysis*, **3**(4):253 (1982).

W. R. Softky and C. Koch. The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *Journal of Neuroscience*, **13**(1):334 (1993).

H. Sompolinsky, A. Crisanti, and H. J. Sommers. Chaos in random neural networks. *Physical Review Letters*, **61**(3):259 (1988).

F. P. Spitzner, J. Dehning, J. Wilting, *et al.* Mr. Estimator, a toolbox to determine intrinsic timescales from sub-sampled spiking activity. *In preparation* ().

A. Statman, M. Kaufman, A. Minerbi, N. E. Ziv, and N. Brenner. Synaptic Size Dynamics as an Effectively Stochastic Process. *PLoS Computational Biology*, **10**(10):e1003846 (2014).

R. B. Stein, E. R. Gossen, and K. E. Jones. Neuronal variability: noise or part of the signal? *Nature Reviews Neuroscience*, **6**(5):389 (2005).

E. A. Stern, A. E. Kincaid, and C. J. Wilson. Spontaneous Subthreshold Membrane Potential Fluctuations and Action Potential Variability of Rat Corticostriatal and Striatal Neurons In Vivo. *Journal of Neurophysiology*, **77**(4):1697 (1997).

A. Steyn-Ross and M. Steyn-Ross. *Modeling Phase Transitions in the Brain.* Springer New York, New York, NY (2010).

D. Steyn-Ross, M. Steyn-Ross, M. Wilson, and J. Sleigh. Phase transitions in single neurons and neural populations: Critical slowing, anesthesia, and sleep cycles. In *Modeling Phase Transitions in the Brain*, 1–26. Springer New York, New York, NY (2010).

H. Suarez, C. Koch, and R. Douglas. Modeling direction selectivity of simple cells in striate visual cortex within the framework of the canonical microcircuit. *Journal of Neuroscience*, **15**(10):6700 (1995).

E. Tagliazucchi, P. Balenzuela, D. Fraiman, and D. R. Chialvo. Criticality in large-scale brain fmri dynamics unveiled by a novel point process analysis. *Frontiers in Physiology*, **3**(February):1 (2012).

E. Tagliazucchi, D. R. Chialvo, M. Siniatchkin, *et al.* Large-scale signatures of unconsciousness are consistent with a departure from critical dynamics. *Journal of The Royal Society Interface*, **13**(114):20151027 (2016).

C. Tetzlaff, S. Okujeni, U. Egert, F. Wörgötter, and M. Butz. Self-organized criticality in developing neuronal networks. *PLoS Computational Biology*, **6**(12):e1001013 (2010).

P. F. Teunis, C. L. Moe, P. Liu, *et al.* Norwalk virus: How infectious is it? *Journal of Medical Virology*, **80**(8):1468 (2008).

G. Tkačik, T. Mora, O. Marre, *et al.* Thermodynamics and signatures of criticality in a network of neurons. *Proceedings of the National Academy of Sciences*, **112**(37):11508 (2015).

D. J. Tolhurst, J. A. Movshon, and I. D. Thompson. The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast. *Experimental Brain Research*, **41**(3-4):414 (1981).

N. Tomen, D. Rotermund, and U. Ernst. Marginally subcritical dynamics explain enhanced stimulus discriminability under attention. *Frontiers in Systems Neuroscience*, **8**(August):1 (2014).

Z. Tosi and J. Beggs. Cortical Circuits from Scratch: A Metaplastic Architecture for the Emergence of Lognormal Firing Rates and Realistic Topology. *Preprint at arXiv:1706.00133* (2017).

J. Touboul and A. Destexhe. Power-law Statistics and universal scaling in the absence of criticality. *Physical Review E*, **95**(1):1 (2017).

M. Tsodyks, K. Pawelzik, and H. Markram. Neural Networks with Dynamic Synapses. *Neural Computation*, **10**(4):821 (1998).

M. V. Tsodyks and H. Markram. The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proceedings of the National Academy of Science*, **94**(2):719 (1997).

G. Turrigiano, L. F. Abbott, and E. Marder. Activity-dependent changes in the intrisic properties of cultured neurons. *Science*, **264**(May):974 (1994).

G. G. Turrigiano, K. R. Leslie, N. S. Desai, L. C. Rutherford, and S. B. Nelson. Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature*, **391**(6670):892 (1998).

G. G. Turrigiano and S. B. Nelson. Hebb and homeostasis in neuronal plasticity. *Current Opinion in Neurobiology*, **10**(3):358 (2000).

G. G. Turrigiano and S. B. Nelson. Homeostatic plasticity in the developing nervous system. *Nature Reviews Neuroscience*, **5**(2):97 (2004).

J. H. van Hateren and A. van der Schaaf. Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society of London B: Biological Sciences*, **265**(1394):359 (1998).

J. Van Pelt, M. Corner, P. Wolters, W. Rutten, and G. Ramakers. Longterm stability and developmental changes in spontaneous network burst firing patterns in dissociated rat cerebral cortex cell cultures on multielectrode arrays. *Neuroscience Letters*, **361**(1-3):86 (2004).

C. van Vreeswijk. Partial synchronization in populations of pulse-coupled oscillators. *Physical Review E*, **54**:5522 (1996).

C. van Vreeswijk and H. Sompolinsky. Chaos in Neuronal Networks with Balanced Excitatory and Inhibitory Activity. *Science*, **274**(5293):1724 (1996).

C. van Vreeswijk and H. Sompolinsky. Irregular firing in cortical circuits with inhibition/excitation balance. In Bower (ed.), *Computational Neuroscience: Trends in Reserach, 1997*, 209–213. Plenum Press, New York (1997).

R. Vardi, A. Goldental, S. Sardi, A. Sheinin, and I. Kanter. Simultaneous multi-patch-clamp and extracellular-array recordings: Single neuron reflects network activity. *Scientific Reports*, **6**(1):36228 (2016).

K. N. Venkataraman. A Time Series Approach to the Study of the Simple Subcritical Galton-Watson Process with Immigration. *Advances in Applied Probability*, **14**(1):1 (1982).

Y. S. Virkar, W. L. Shew, J. G. Restrepo, and E. Ott. Feedback control stabilization of critical dynamics via resource transport on multilayer networks : How glia enable learning dynamics in the brain. *Physical Review E* **94**:042310 (2016).

R. Vogels, W. Spileers, and G. A. Orban. The response variability of striate cortical neurons in the behaving monkey. *Experimental Brain Research*, **77**(2):432 (1989).

T. P. Vogels, K. Rajan, and L. F. Abbott. Neural Network Dynamics. *Annual Reviews Neuroscience* **28**:357 (2005).

T. P. Vogels, H. Sprekeler, F. Zenke, C. Clopath, and W. Gerstner. Inhibitory Plasticity Balances Excitation and Inhibition in Sensory Pathways and Memory Networks. *Science*, **334**(6062):1569 (2011).

D. A. Wagenaar. http://neurodatasharing.bme.gatech.edu/development-data/ html/index.html.

D. A. Wagenaar. Controlling Bursting in Cortical Cultures with Closed-Loop Multi-Electrode Stimulation. *Journal of Neuroscience*, **25**(3):680 (2005).

D. a. Wagenaar, J. Pine, and S. M. Potter. Searching for plasticity in dissociated cortical cultures on multi-electrode arrays. *Journal of Negative Results in Biomedicine*, **5**:16 (2006).

X. J. Wang. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*, **36**(5):955 (2002).

X. R. Wang, J. T. Lizier, and M. Prokopenko. Fisher Information at the Edge of Chaos in Random Boolean Networks. *Artificial Life*, **17**(4):315 (2011).

C. Wei and J. Winnicki. Estimation of the Means in the Branching Process with Immigration. *The Annals of Statistics*, **18**(4):1757 (1990).

C. Z. Wei. Convergence rates for the critical branching process with immigration. *Statistica Sinica*, **1**:175 (1991).

Q. Wen, A. Stepanyants, G. N. Elston, A. Y. Grosberg, and D. B. Chklovskii. Maximization of the connectivity repertoire as a statistical principle governing the shapes of dendritic arbors. *Proceedings of the National Academy of Science*, **106**(30):12536 (2009).

O. Wichmann, A. Siedler, D. Sagebiel, *et al.* Further efforts needed to achieve measles elimination in Germany: Results of an outbreak investigation. *Bulletin of the World Health Organization*, **87**(2):108 (2009).

A. Williams, T. O'Leary, and E. Marder. Homeostatic Regulation of Neuronal Excitability. *Scholarpedia*, **8**(1):1656 (2013).

R. V. Williams-García, M. Moore, J. M. Beggs, and G. Ortiz. Quasicritical brain dynamics on a nonequilibrium Widom line. *Physical Review E*, **90**(6):062714 (2014).

C. Wilson. Up and down states. *Scholarpedia*, **3**(6):1410 (2008).

H. R. Wilson and J. D. Cowan. Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons. *Biophysical Journal*, **12**(1):1 (1972).

J. Wilting, J. Dehning, J. Pinheiro Neto, *et al.* Operating in a Reverberating Regime Enables Rapid Tuning of Network States to Task Requirements. *Frontiers in Systems Neuroscience*, **12**(November):55 (2018).

J. Wilting and V. Priesemann. Inferring collective dynamical states from widely unobserved systems. *Nature Communications*, **9**(1):2325 (2018a).

J. Wilting and V. Priesemann. On the ground state of spiking network activity in mammalian cortex. *Preprint at arXiv:180407864* (2018b).

J. Wilting and V. Priesemann. 25 years of criticality in Neuroscience — established results, open controversies, novel concepts. *Current Opinion in Neurobiology*, **58**:105 (2019a).

J. Wilting and V. Priesemann. Between Perfectly Critical and Fully Irregular: A Reverberating Model Captures and Predicts Cortical Spike Propagation. *Cerebral Cortex*, **29**(6):2759 (2019b).

S. Yu, A. Klaus, H. Yang, and D. Plenz. Scale-Invariant Neuronal Avalanche Dynamics and the Cut-Off in Size Distributions. *PLoS ONE*, **9**(6):e99761 (2014).

S. Yu, T. L. Ribeiro, C. Meisel, *et al.* Maintained avalanche dynamics during task-induced changes of neuronal activity in nonhuman primates. *eLife*, **6**:1 (2017).

S. Zapperi, K. B. Lauritsen, and H. E. Stanley. Self-Organized Branching Processes: Mean-Field Theory for Avalanches. **75**(22):4071 (1995).

F. Zenke, E. J. Agnes, and W. Gerstner. Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. *Nature Communications*, **6**(1):6922 (2015).

F. Zenke, W. Gerstner, and S. Ganguli. The temporal paradox of Hebbian learning and homeostatic plasticity. *Current Opinion in Neurobiology*, **43**:166 (2017).

F. Zenke, G. Hennequin, and W. Gerstner. Synaptic Plasticity in Neural Networks Needs Homeostasis with a Fast Rate Detector. *PLoS Comput Biol*, **9**(11):e1003330 (2013).

J. Zierenberg, J. Wilting, and V. Priesemann. Homeostatic Plasticity and External Input Shape Neural Network Dynamics. *Physical Review X*, **8**(3):031018 (2018).

J. Zierenberg, J. Wilting, V. Priesemann, and A. Levina. Description of spreading dynamics by microscopic network models and macroscopic branching processes can differ due to coalescence. *Physical Review E*, **101**(2):022301 (2020a).

J. Zierenberg, J. Wilting, V. Priesemann, and A. Levina. Tailored ensembles of neural networks optimize sensitivity to stimulus Statistics. *Physical Review Research*, **2**(1):013115 (2020b).