Universität Potsdam

Igor Goychuk | Andriy Goychuk

Stochastic Wilson–Cowan models of neuronal network dynamics with memory and delay

New Journal of Physics

The open access journal at the forefront of physics

PAPER

Stochastic Wilson–Cowan models of neuronal network dynamics with memory and delay

Igor Goychuk[1] and Andriy Goychuk[2]

[1]   Institute for Physics and Astronomy, University of Potsdam, Karl-Liebknecht-Str. 24/25, D-14476 Potsdam-Golm, Germany
[2]   Faculty of Physics, Ludwig Maximilian University of Munich, Munich, Germany

E-mail: igoychuk@uni-potsdam.de

Abstract

We consider a simple Markovian class of the stochastic Wilson–Cowan type models of neuronal network dynamics, which incorporates stochastic delay caused by the existence of a refractory period of neurons. From the point of view of the dynamics of the individual elements, we are dealing with a network of non-Markovian stochastic two-state oscillators with memory, which are coupled globally in a mean-field fashion. This interrelation of a higher-dimensional Markovian and lower-dimensional non-Markovian dynamics is discussed in its relevance to the general problem of the network dynamics of complex elements possessing memory. The simplest model of this class is provided by a three-state Markovian neuron with one refractory state, which causes firing delay with an exponentially decaying memory within the two-state reduced model. This basic model is used to study critical avalanche dynamics (the noise sustained criticality) in a balanced feedforward network consisting of the excitatory and inhibitory neurons. Such avalanches emerge due to the network size dependent noise (mesoscopic noise). Numerical simulations reveal an intermediate power law in the distribution of avalanche sizes with the critical exponent around −1.16. We show that this power law is robust upon a variation of the refractory time over several orders of magnitude. However, the avalanche time distribution is biexponential. It does not reflect any genuine power law dependence.

1. Introduction

Network complexity pervades biology and medicine [1], and the human organism can be considered as an integrated complex network of different physiological systems [2] such as circulatory and respiratory systems, visual system, digestive and endocrine systems, etc, which are coordinated by autonomic and central nervous systems including the brain. The dynamics of the sleep-wake transitions during the sleep of humans and other mammals [3–5] presents one important examples of such complex dynamics featured by randomly repeating active periods with scale-invariant power-law distributions of their size and duration, which reflects a self-organized criticality or SOC [1, 4–7] in such complex network systems. In turn, functioning of the human brain presents in essence a network activity of coupled and interrelated neurons surrounded by glia cells. The immense complexity of this subject matter [8] does not exclude, but rather invites thinking in terms of simple physical modeling approaches, see e.g. in [3, 5], since even very simple physical models can display very complex behavior. The models of critical dynamical phenomena such as SOC [1, 6, 7] are especially important in this respect [4, 5]. Physical modeling can help to discriminate the physical and biological complexity from the complexity of mental processes, the 'form within' [8], which is mediated but not determined in fine features by the background physical processes. The recently discovered complexity of the critical brain dynamics [9–11] in principle does not have anything in common with the complexity of mental processes as it is already displayed by organotypic networks of neurons formed by cortical slices on a multi-electrode array [9]. Such a physical complexity is essentially the complexity of crude matter that got self-organized following physical laws. It thus

belongs to statistical physics or system biophysics. Physical models such as SOC are especially important and helpful here.

The Wilson and Cowan model [12] presents one of the well-established models of neuronal network dynamics [13]. It incorporates individual elements in a simplest possible fashion as two-state stochastic oscillators with one quiescent state and one excited state, and random transitions between these two states which are influenced by the mutual coupling among the network elements. The model has been introduced in the deterministic limit of huge amounts of coupled elements in complete neglect of the intrinsic mesoscopic noise and became immensely popular with the years [13], being used e.g. to describe neuronal oscillations in visual cortex within a mean-field approximation [1, 14]. Recently, the previously neglected mesoscopic noise effects were incorporated in this model for a finite-size network [15, 16]. Such a noise has been shown to be very important, in particular, for the occurrence of the critical avalanche dynamics [15] absent in the deterministic Wilson–Cowan model and also for the emergence of oscillatory noisy dynamics [16]. At first glance, such a noisy dynamics can look like a chaotic deterministic one. Deterministic chaos influenced by the noise can also be a natural feature of a higher-dimensional dynamics, beyond the original two-dimensional (2D) Wilson–Cowan mean field model. Indeed, deterministic chaos has been found in the brain dynamics some time ago [1, 17]. However, it cannot be described within the memoryless Wilson–Cowan model because the minimal dimension for chaos is three [18].

Stochastic mesoscopic noise effects due to a finite number of elements in finite size systems attract substantial attention over several decades, especially with respect to chemical reactions on the mesoscale [19, 20], being especially pertinent to the physico-chemical processes in living cells [21]. In particular, such intrinsic noise can cause and optimize spontaneous spiking (coherence resonance [22, 23]) in the excitable clusters of ionic channels in cell membranes, which are globally coupled through the common membrane potential, or the response of such systems to periodic external signals (stochastic resonance [24]) within a stochastic Hodgkin–Huxley model [25]. Finite-size networks of globally coupled bistable stochastic oscillators were also considered without relation to the Wilson–Cowan model [26, 27], including non-Markovian memory effects [23, 28–31].

In this paper, we consider a class of higher-dimensional generalizations of the stochastic Wilson–Cowan model aimed to incorporate non-Markovian memory effects in the dynamics of individual neurons. Such effects are caused by the existence of a refractory period or inactivated state from which the neuron cannot be excited immediately. First, we discuss a general class of such models. Then, we apply the simplest two-state non-Markovian model of this class, embedded as a three-state Markovian model with one inactivated state, to study a critical avalanche dynamics in a balanced network of the excitatory and inhibitory neurons within a mean-field approximation. Here, we restrict ourselves to the simplest example of a fully connected network with all-to-all coupling of its elements. In particular, we derive the power law exponents characterizing the critical self-organized dynamics of the network from the precise numerical simulations done with the dynamic Monte Carlo, or Gillespie algorithm. We also compare these results with similar results obtained within approximate stochastic Langevin dynamics, or, equivalently, within a diffusional approximation to the discrete state dynamics. Here, we reveal a profound difference.
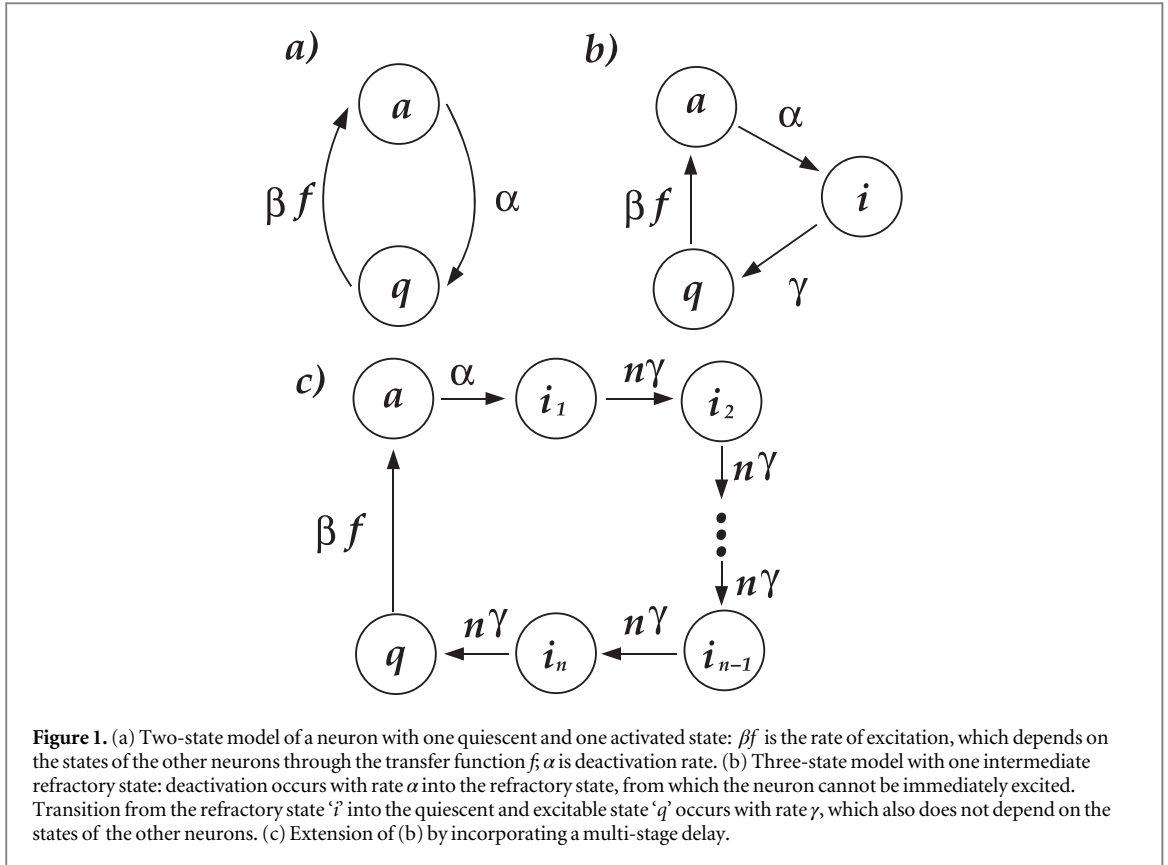
2. The model and theory

2.1. Stochastic models of single neurons

Let us depart from the Markovian model of a neuron possessing one activated or excited state 'a', and a quiescent state 'q', see figure 1(a). The excitation of the neuron occurs with the rate βf, where β is a rate constant and f is a dimensionless transfer function which depends on the states of the other neurons, and will be discussed below. Deactivation occurs with the rate α, which does not depend on states of other neurons. Let us assume for a while that f is not explicitly time dependent. From the point of view of the theory of continuous time random walks (CTRWs) or renewal processes [32, 33], such a two state neuron can be completely characterized by the residence time distributions (RTDs) in its two states, $\psi_a(t)$, and $\psi_q(t)$, correspondingly (assuming that no correlations between the residence time intervals is present—the renewal or semi-Markovian assumption). RTDs define completely the trajectory realizations of such a renewal process. In the Markovian case, the RTDs are strictly exponential, $\psi_a(t) = \alpha \exp(-\alpha t)$, and $\psi_q(t) = \nu \exp(-\nu t)$, where we denoted $\nu = \beta f$. Then, such a trajectory description corresponds to the Markovian balance or master equations for the probabilities to populate the states 'a' and 'q', $u(t)$ and $q(t)$, respectively. Due to the probability conservation, $u(t) + q(t) = 1$,

$$\dot{u}(t) = -\alpha u(t) + \beta f [1 - u(t)]. \tag{1}$$

The memory effect due to delay of a new excitation event after the neuron comes into the quiescent state, or the existence of some refractory period τ_d, can be captured within the trajectory description by a non-exponential

Figure 1. (a) Two-state model of a neuron with one quiescent and one activated state: βf is the rate of excitation, which depends on the states of the other neurons through the transfer function f; α is deactivation rate. (b) Three-state model with one intermediate refractory state: deactivation occurs with rate α into the refractory state, from which the neuron cannot be immediately excited. Transition from the refractory state 'i' into the quiescent and excitable state 'q' occurs with rate γ, which also does not depend on the states of the other neurons. (c) Extension of (b) by incorporating a multi-stage delay.

RTD $\psi_q(t)$. This transforms the corresponding master equation into a generalized master equation (GME) with memory, where the term $\beta f q(t)$ is replaced by $\int_{t_0}^{t} K(t - t') q(t') dt'$, with a memory kernel $K(t)$. Here, t_0 is the starting time, $t_0 = 0$, if not a different one is explicitly stated. Hence, equation (1) is replaced by

$$\dot{u}(t) = -\alpha u(t) + \int_{0}^{t} K(t - t')[1 - u(t')] dt'. \tag{2}$$

In the CTRW theory it is well-known how the memory kernel $K(t)$ and the RTD $\psi_q(t)$ are related [33, 34] (see also appendix of [35]). Namely, their Laplace-transforms (denoted as $\tilde{F}(s) = \int_{0}^{\infty} \exp(-st) F(t) dt$, for any function $F(t)$) are related as

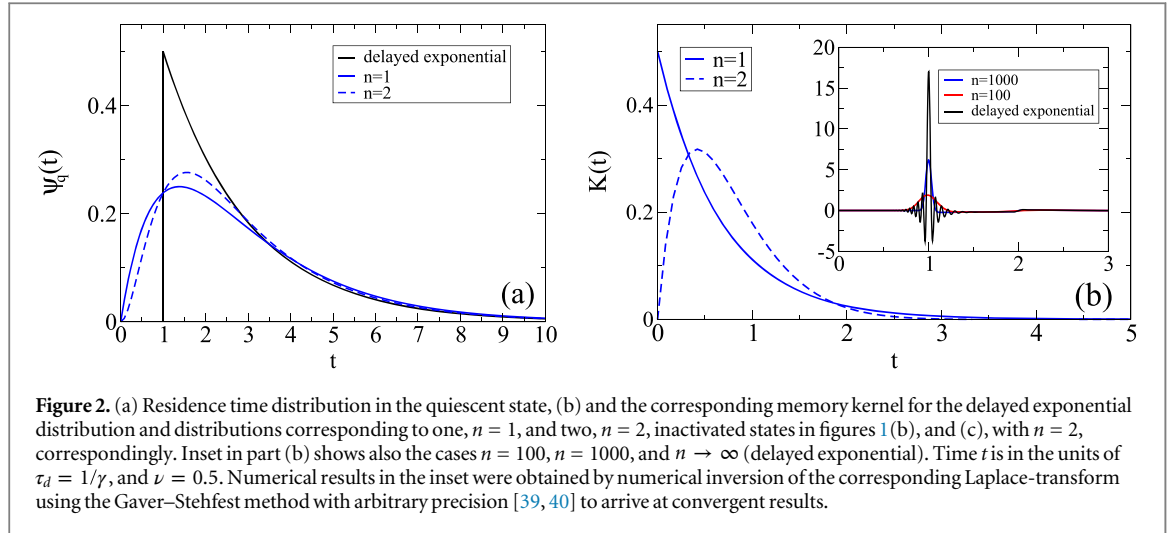$$\tilde{K}(s) = \frac{s \tilde{\psi}_q(s)}{1 - \tilde{\psi}_q(s)}. \tag{3}$$

In neurosciences, a delayed exponential, or delayed Poissonian model is popular [36]. It is featured by the absolute refractory period τ_d, i.e. $\psi_q(t) = 0$, for $0 \leqslant t < \tau_d$, and the exponential distribution, $\psi_q(t) = \nu \exp[-\nu(t - \tau_d)]$, for $t \geqslant \tau_d$, see in figure 2. This model corresponds to $\tilde{\psi}_q(s) = \exp(-\tau_d s) \nu/(s + \nu)$ and the memory kernel $\tilde{K}(s) = \nu s/[(\nu + s) \exp(\tau_d s) - \nu]$. The numerical inverse Laplace transform of this memory kernel is depicted in the inset of figure 2(b). Notice that it does not correspond to the memory kernel $K(t) = \nu_r \delta(t - \tau_d)$, which would correspond to the master equation with the deterministic delay [37]

$$\dot{u}(t) = -\alpha u(t) + \beta f_r \Big[1 - u(t - \tau_d) \Big]. \tag{4}$$

However, this memory kernel is strongly peaked at $t = \tau_d$, and can thus be approximated, with $\nu_r = \lim_{s \to 0} \tilde{K}(s) = \nu/(1 + \nu \tau_d)$, which is the inverse mean time of the delayed Poissonian distribution $\psi_q(t)$. In the corresponding Markovian approximation, equation (1) is restored with a renormalized transfer function

$$f_r = \frac{f}{1 + \beta \tau_d f}. \tag{5}$$

This is the simplest way to account for the delay effects. Obviously, any delay should suppress excitability within this approximation, because $f_r < f$. However, suppression of the excitability of the inhibitory neurons may enhance the excitability of the whole network consisting of both excitatory and inhibitory neurons. Hence, possible effects are generally non-trivial even in this approximation. Moreover, equation (5) makes it immediately clear that the delay effects are generally expected to be very substantial for $\beta \tau_d \geqslant 1$.

Figure 2. (a) Residence time distribution in the quiescent state, (b) and the corresponding memory kernel for the delayed exponential distribution and distributions corresponding to one, $n = 1$, and two, $n = 2$, inactivated states in figures 1(b), and (c), with $n = 2$, correspondingly. Inset in part (b) shows also the cases $n = 100$, $n = 1000$, and $n \to \infty$ (delayed exponential). Time t is in the units of $\tau_d = 1/\gamma$, and $\nu = 0.5$. Numerical results in the inset were obtained by numerical inversion of the corresponding Laplace-transform using the Gaver–Stehfest method with arbitrary precision [39, 40] to arrive at convergent results.

2.1.1. The simplest non-Markovian model and its Markovian embedding

It is well-known that in many cases non-Markovian CTRW dynamics can be embedded as some Markovian dynamics in a higher-dimensional, possibly infinite dimensional space [38]. Given a non-trivial form of the memory kernel for the delayed exponential distribution of the quiescent times, we can ask the question: What is the simplest non-Markovian model and the corresponding Markovian embedding to account for the memory effects? From the point of view of GME, it is $K(t) = \kappa \exp(-rt)$, i.e. an exponentially decaying memory kernel. The corresponding memory kernel with $\kappa = \nu\gamma$, and $r = \nu + \gamma$ corresponds to $\psi_q(t)$, which is the time convolution of two exponential distributions, $\psi_q^{(0)}(t) = \nu \exp(-\nu t)$, and $\psi_i(t) = \gamma \exp(-\gamma t)$. It corresponds to a compound state '$\mathbf{q} = q \cup i$' in figure 1(b). Indeed, the Laplace transform of the corresponding compound distribution is just the product of the Laplace-transforms of two exponential distributions, i.e., $\tilde{\psi}_q(s) = \nu\gamma/[(s + \nu)(s + \gamma)]$. By equation (3) it corresponds precisely to the stated exponential memory kernel. The corresponding $\psi_q(t) = \nu\gamma [\exp(-\nu t) - \exp(-\gamma t)]/(\nu - \gamma)$ has a maximum at the most probable time interval $t_{max} = \ln(\nu/\gamma)/(\nu - \gamma)$, see figure 2(a), reflecting the most probable stochastic time delay. This simplest non-Markovian model with memory allows, however, for a very simple Markovian embedding by introduction of an intermediate refractory state 'i' shown in figure 1(b), with the population probability $x(t)$ and the exponential RTD given above. It has the mean refractory time $\tau_d := \langle \tau \rangle = \int_0^\infty \psi_i(\tau)\,d\tau = 1/\gamma$, and the relative standard deviation, or the coefficient of variation, $C_V := \sqrt{\langle \tau^2 \rangle - \langle \tau \rangle^2}/\langle \tau \rangle = 1$. Using the conservation law, $u + q + x = 1$, the corresponding master equations can be written either as

$$\dot{u} = -\alpha u + \nu(1 - u - x),$$
$$\dot{x} = \alpha u - \gamma x, \tag{6}$$

or as

$$\dot{u} = -\alpha u + \nu q, \tag{7a}$$

$$\dot{q} = -(\gamma + \nu)q + \gamma(1 - u). \tag{7b}$$

From (7b) follows

$$q(t) = e^{-(\gamma+\nu)t}q(0) + \gamma \int_0^t e^{-(\gamma+\nu)(t-t')}[1 - u(t')]\,dt'. \tag{8}$$

After substitution of this equation into (7a) one obtains indeed equation (2) with the discussed exponential memory kernel provided that $q(0) = 0$. The latter condition is natural because every sojourn in the compound quiescent state '$\mathbf{q}$' starts from the substate 'i' (resetting memory of this neuron to zero), and $q(t) + x(t)$ is the probability of the compound quiescent state within the two-state non-Markovian reduction of the three-state Markovian problem. Here, one can also see the origin of a profound problem with the description of the whole network dynamics of interacting non-Markovian renewal elements as a hyper-dimensional renewal process. Obviously, the behavior of the whole network cannot be considered as a renewal process, because after each and every de-excitation event only one element is reset. Then it starts with zero memory, while all others keep their memory until they are reset. Hence, any Gillespie type simulation of the whole network of interacting non-Markovian elements must account for the 'age' of each network element separately. Markovian embedding

allows to circumvent this problem and dramatically accelerate simulations within the mean-field approximation, see below.

The considered three-state Markovian cyclic model presents one of the fundamental kinetic models in biophysics. It provides, in particular, a paradigm for non-equilibrium steady state cycling. For example, the cyclic kinetics of an enzyme E, which binds a substrate molecule S, converts it to a product P, and releases it afterwards can be represented as a three-state cycle, $E \to ES \to EP \to E$. This model was used e.g. in [29] for an excitable unit. Furthermore, three-state non-Markovian models can be used with a non-exponential distribution $\psi_i(t)$. For example, if to use the deterministically delayed $\psi_i(t) = \delta(t - \tau_d)$, and exponential $\psi_q(t)$ within the three-state cyclic model, then one obtains the delayed exponential distribution within the two-state reduced model, which was discussed above. In addition, $\psi_a(t)$ can also be non-exponential. For $\psi_a(t) = \delta(t - 1/\alpha)$ in the excited state of the three-state non-Markovian model, one obtains the model used in [28, 29].

2.1.2. Markovian embedding with many substates

One can also introduce many delayed substates as shown in figure 1(c). Within the three-state non-Markovian model this can be considered as having one delayed state '*i*' characterized by the special Erlangian distribution [32], $\psi_i(t) = n\gamma (n\gamma t)^{n-1} \exp(-n\gamma t)/(n-1)!$, with the Laplace-transform $\tilde{\psi}_i(s) = 1/(1 + s/(n\gamma))^n$ reflecting the corresponding multiple convolution. Such a non-Markovian three-state model has been considered in [30], and a non-Markovian two-state model with the Erlangian distribution of the quiescent times has been studied in [31]. The compound quiescent state corresponding to the model in [30] is characterized by $\tilde{\psi}_q(s) = \nu/[(1 + s/(n\gamma))^n (\nu + s)]$. The mean delay time is the same $\tau_d = 1/\gamma$ for any n, and the coefficient of variation becomes ever smaller with increasing n, $C_V = 1/\sqrt{n}$, i.e. the distribution of the refractory times becomes ever more sharpened. The Laplace-transformed memory kernel is $\tilde{K}(s) = \nu s/[(1 + s/(n\gamma))^n (\nu + s) - \nu]$. Some corresponding $\psi_q(t)$ and $K(t)$ are shown in figure 2. Already for $n = 2$, the memory kernel starts to show a peaked structure. Notice that in the limit $n \to \infty$, the above delayed exponential (or Poissonian) model immediately follows with $\tau_d = 1/\gamma$. For any n, the inverse mean time in the quiescent state is given by βf_r with f_r in (5). Increasing n yields an ever better approximation for the delayed Poissonian model. However, it can be considered as a useful model in itself. The corresponding Markovian embedding master equation reads (with q excluded by the probability conservation law):

$$
\begin{aligned}
\dot{u} &= -\alpha u + \beta f \left(1 - u - \sum_{i=1}^{n} x_i \right), \\
\dot{x}_1 &= \alpha u - n\gamma x_1, \\
\dot{x}_j &= n\gamma (x_{j-1} - x_j), \quad j = 2, ..., n,
\end{aligned}
\tag{9}
$$

with, $x_j(0) = 0$, for $j = 2, \ldots, n$, initially. With a different initial condition, the corresponding GME obtained upon projection of the multi-dimensional dynamics onto the subspace of u and q variables will contain a dependence on this initial condition in the subspace of hidden Markovian variables. On the level of non-Markovian dynamics this can be accounted for by a different choice of the RTD $\psi_q^{(0)}(t)$ for the first sojourn in the quiescent state. It depends on how long this state has been populated before the dynamics started [32]. The GME (2), (3) corresponds to the particular choice, $\psi_q^{(0)}(t) = \psi_q(t)$.

We mention in passing also that it is straightforward to consider a power-law distributed delay, both within a semi-Markovian model and within an approximate finite-dimensional Markovian embedding. Also a stochastic model for bursting neurons can be introduced immediately. We shall not, however, consider these possibilities in the present work.

2.2. Network of neurons within the mean field dynamics

Following Wilson and Cowan, we consider a network of N_e excitatory and N_i inhibitory neurons, with the probabilities of neurons to be in their excited states $u_i(t)$ and $v_i(t)$, correspondingly. The neuron k can influence the neuron l and possibly itself ($k = l$) by excitation, or inhibition with the coupling constants, $w_{ee}^{lk} > 0$, $w_{ie}^{lk} > 0$ for the excitatory neuron k, and $-w_{ei}^{lk} < 0$, $-w_{ii}^{lk} < 0$, for the inhibitory neuron k. The absolute value of the coupling constant reflects the synaptic strength. Each excitatory neuron l thus obtains an averaged input $s_l = (1/N_e^l) \sum_k w_{ee}^{lk} u_k - (1/N_i^l) \sum_p w_{ei}^{lp} v_p + h_e^l$, and the inhibitory neuron p receives the input $s_p = (1/N_e^p) \sum_l w_{ie}^{pl} u_l - (1/N_i^p) \sum_k w_{ii}^{pk} v_k + h_i^p$, where N_e^l and N_i^l, etc is the number of inputs which the lth neuron obtains from the excitatory and inhibitory neurons, correspondingly. The constants h_e^l and h_i^p serve to fix the spontaneous spiking rates, $\beta_l f(h_e^l)$, $\beta_p f(h_i^p)$, in the absence of coupling, $w_{ee}^{lk} \to 0$, $w_{ii}^{lk} \to 0$, $w_{ei}^{lk} \to 0$, $w_{ie}^{lk} \to 0$. Coupling can either enhance, or suppress these rates. Phenomenologically, this is accounted for by the transfer function $f(s)$, which we assume to be the same for

all neurons. Some common biophysically motivated popular choices of the transfer function $f(s)$ are

$$f(s) = \tanh(s)\theta(s), \tag{10}$$

where $\theta(s)$ is the Heaviside step function, and

$$f(s) = 1/[1 + \exp(-s)]. \tag{11}$$

Both are bounded as $0 \leqslant f \leqslant 1$. Evidently, this is a very rich model even for the simplest two-state model of neurons, in the absence of memory effects. The simplest further approximation to describe the collective dynamics of neurons is to invoke the mean field approximation [14]. It is equivalent to assuming that all the coupling constants like w_{ee}^{lk}, etc, thresholds h_e^l, etc, and rates β_l, α_l, are equal within a subpopulation, $w_{ee}^{lk} = w_{ee}$, $h_e^l = h_e$, $\beta_e^l = \beta_e$, or $\beta_i^l = \beta_i$, etc. Furthermore, one can introduce the occupation numbers of the excited neurons in each population, $u(t) = (1/N_e)\sum_{i=1}^{N_e} u_i(t)$, and $v(t) = (1/N_i)\sum_{i=1}^{N_i} v_i(t)$, and consider the dynamics of these variables. They present the fractions of neurons which are excited.

We restrict our treatment in the rest of this paper to the simplest two state non-Markovian model within the three state Markovian embedding and introduce the occupation numbers of neurons, $x(t) = (1/N_e)\sum_{i=1}^{N_e} x_i(t)$, and $y(t) = (1/N_i)\sum_{i=1}^{N_i} y_i(t)$, in the corresponding delayed states. Then, in the deterministic limit N_e, $N_i \to \infty$, one obtains a four-dimensional nonlinear dynamics

$$
\begin{aligned}
\dot{u} &= -\alpha_e u + \beta_e f\left(w_{ee} u - w_{ei} v + h_e\right)(1 - u - x), \\
\dot{x} &= \alpha_e u - \gamma_e x, \\
\dot{v} &= -\alpha_i v + \beta_i f\left(w_{ie} u - w_{ii} v + h_i\right)(1 - v - y), \\
\dot{y} &= \alpha_i v - \gamma_i y.
\end{aligned}
\tag{12}
$$

Notice that unlike the original 2D mean-field Wilson–Cowan dynamics in the deterministic limit, the considered four-dimensional dynamics can in principle be chaotic, for some parameters (which remains an open question). Dynamical chaos might emerge already when only one sort of neurons, e.g. inhibitory neurons, exhibits delayed dynamics, since the minimal dimension for nonlinear chaotic dynamics is three. Then, in the macroscopic deterministic limit

$$
\begin{aligned}
\dot{u} &= -\alpha_e u + \beta_e f\left(w_{ee} u - w_{ei} v + h_e\right)(1 - u), \\
\dot{v} &= -\alpha_i v + \beta_i f\left(w_{ie} u - w_{ii} v + h_i\right)(1 - v - y), \\
\dot{y} &= \alpha_i v - \gamma_i y.
\end{aligned}
\tag{13}
$$

However, we shall not investigate the possibility of a deterministic chaos emerging due to a delay within the minimal extensions of the Wilson–Cowan model in the present work, but rather focus on the mesoscopic intrinsic noise effects caused by finite N_e and N_i. Then, the occupational numbers are random variables (at any fixed time t).

2.2.1. Langevin dynamics

For a very large number of neurons, one can account for the mesoscopic noise effect within the Langevin dynamics, or the diffusional approximation of the discrete state birth-and-death process describing the evolution of the network. This procedure is standard, by analogy with the stochastic theory of chemical reactions [20]. Since we have only direct 'reactions' like $q \to a$, $a \to i$, $i \to q$, for two types of neurons, one must introduce six variables and six independent zero-mean white Gaussian noise sources $\xi_i(t)$, $\langle \xi_i(t)\xi_j(t')\rangle = \delta_{ij}\delta(t - t')$. Stochastic dynamics is, however, effectively four-dimensional because of two probability conservation laws, which allow to exclude two variables out of six:

$$
\begin{aligned}
\dot{u} &= -\alpha_e u + \beta_e f\left(w_{ee} u - w_{ei} v + h_e\right)(1 - u - x) \\
&\quad - \frac{1}{\sqrt{N_e}}\sqrt{\alpha_e u}\,\xi_1(t) + \frac{1}{\sqrt{N_e}}\sqrt{\beta_e f\left(w_{ee} u - w_{ei} v + h_e\right)(1 - u - x)}\,\xi_2(t), \\
\dot{x} &= \alpha_e u - \gamma_e x + \frac{1}{\sqrt{N_e}}\sqrt{\alpha_e u}\,\xi_1(t) - \frac{1}{\sqrt{N_e}}\sqrt{\gamma_e x}\,\xi_3(t),
\end{aligned}
$$

Table 1. Transitions and rates.

i	Transition	Rate r_i
1	$(N, M, N_1, M_1) \rightarrow (N-1, M, N_1+1, M_1)$	$r_1 = N\alpha_e$
2	$(N, M, N_1, M_1) \rightarrow (N, M-1, N_1, M_1+1)$	$r_2 = M\alpha_i$
3	$(N, M, N_1, M_1) \rightarrow (N+1, M, N_1, M_1)$	$r_3 = (N_e - N - N_1)\beta_e f(w_{ee}N/N_e - w_{ei}M/N_i + h_e)$
4	$(N, M, N_1, M_1) \rightarrow (N, M+1, N_1, M_1)$	$r_4 = (N_i - M - M_1)\beta_i f(w_{ie}N/N_e - w_{ii}M/N_i + h_i)$
5	$(N, M, N_1, M_1) \rightarrow (N, M, N_1-1, M_1)$	$r_5 = \gamma_e N_1$
6	$(N, M, N_1, M_1) \rightarrow (N, M, N_1, M_1-1)$	$r_6 = \gamma_i M_1$

$$
\dot{v} = -\alpha_i v + \beta_i f\left(w_{ie}u - w_{ii}v + h_i\right)(1 - v - y)
$$
$$
\quad - \frac{1}{\sqrt{N_i}}\sqrt{\alpha_i v}\,\xi_4(t) + \frac{1}{\sqrt{N_i}}\sqrt{\beta_i f\left(w_{ie}u - w_{ii}v + h_i\right)(1 - v - y)}\,\xi_5(t),
$$
$$
\dot{y} = \alpha_i v - \gamma_i y + \frac{1}{\sqrt{N_i}}\sqrt{\alpha_i v}\,\xi_4(t) - \frac{1}{\sqrt{N_i}}\sqrt{\gamma_i y}\,\xi_6(t). \tag{14}
$$

In the limit $N_e, N_i \rightarrow \infty$, the deterministic description in equation (12) is restored. The noise is multiplicative and the Langevin equations must be Ito-interpreted, as it is always the case if the Langevin dynamics results from the standard diffusional approximation of a birth-and-death process, or chemical master equation [20]. Notice that such a Langevin stochastic description can become problematic, if any of the variables u, v, x, y becomes temporally zero or one. Even if some of the noise terms do vanish at the boundaries, where the corresponding rates vanish, the others do not, when a particular boundary is hit. Hence, the occupational numbers can in principle become temporally negative, or larger than one. This unphysical feature is produced by the standard diffusional approximation. However, this problem can be fixed in the numerical simulations by introduction of the corresponding reflecting boundaries and taking sufficiently small integration time steps, as done e.g. in [25, 41] for stochastic Hodgkin–Huxley equations.

2.3. Exact stochastic simulations

Within the mean-field approximation of Markovian dynamics, it suffices to count the numbers of neurons in the corresponding activated, N and M, and refractory, N_1 and M_1 states. Then, we are dealing with a random walk on a four-dimensional lattice (N, M, N_1, M_1) with the discrete variables N, and N_1 taking values in the range from zero to N_e, and the variables M and M_1 in the range from zero to N_i, so that also $0 \leqslant N + N_1 \leqslant N_e$ and $0 \leqslant M + M_1 \leqslant N_i$. From the site (N, M, N_1, M_1) six different transitions are possible. They are enlisted in table 1 with the corresponding transition rates. The master equation governing this birth-and-death process can be readily written. However, it is bulky and not very insightful. For this reason, it is not presented here. The corresponding stochastic process can be easily simulated with the dynamical Monte Carlo or Gillespie algorithm [19], which is exact. Namely, on each step one draws two random numbers. The first one, τ, is drawn from the exponential distribution characterized by the total rate $r_\Sigma = \sum_{i=1}^{6} r_i$. It gives a random time interval at which the network state is updated. Given a uniformly distributed random variable ζ_1, $0 \leqslant \zeta_1 \leqslant 1$, $\tau = (1/r_\Sigma)\ln(1/\zeta_1)$. Then, one of the transitions in table 1 is chosen in accordance with its probability $p_i = r_i/r_\Sigma$. For this, one generates a uniformly distributed random variable ζ_2 bounded as $0 \leqslant \zeta_2 \leqslant 1$. If $0 < \zeta_2 < p_1$, then the first transition is chosen. If $p_1 \leqslant \zeta_2 < p_1 + p_2$, then the second transition is chosen, etc, i.e. in accordance with the length of the corresponding interval p_i, $\sum_{i=1}^{6} p_i = 1$.

Notice that an attempt to generalize this scheme towards a non-Markovian renewal walk on a four-dimensional lattice to account for the memory in the inactivated state is logically inconsistent because in such a case accomplishing each step would mean reset, or renewal of *all* neurons, and not the only one which actually makes transition. However, each non-Markovian element has its individual memory. In a direct simulation of the network of non-Markovian elements one must therefore consider them individually, even within the mean-field approximation. Then, one has to consider CTRW on a hyper-dimensional lattice of huge dimensionality, which will dramatically slow down simulations imposing computational restrictions on the maximal size of the network. Of course, beyond the mean field approximation one must also simulate each element separately. Here, a direct semi-Markovian approach can be preferred. In this work, we restrict ourselves to the mean-field dynamics within the Markovian embedding framework, which allows for exact simulations of very large networks within a reasonable computational time.
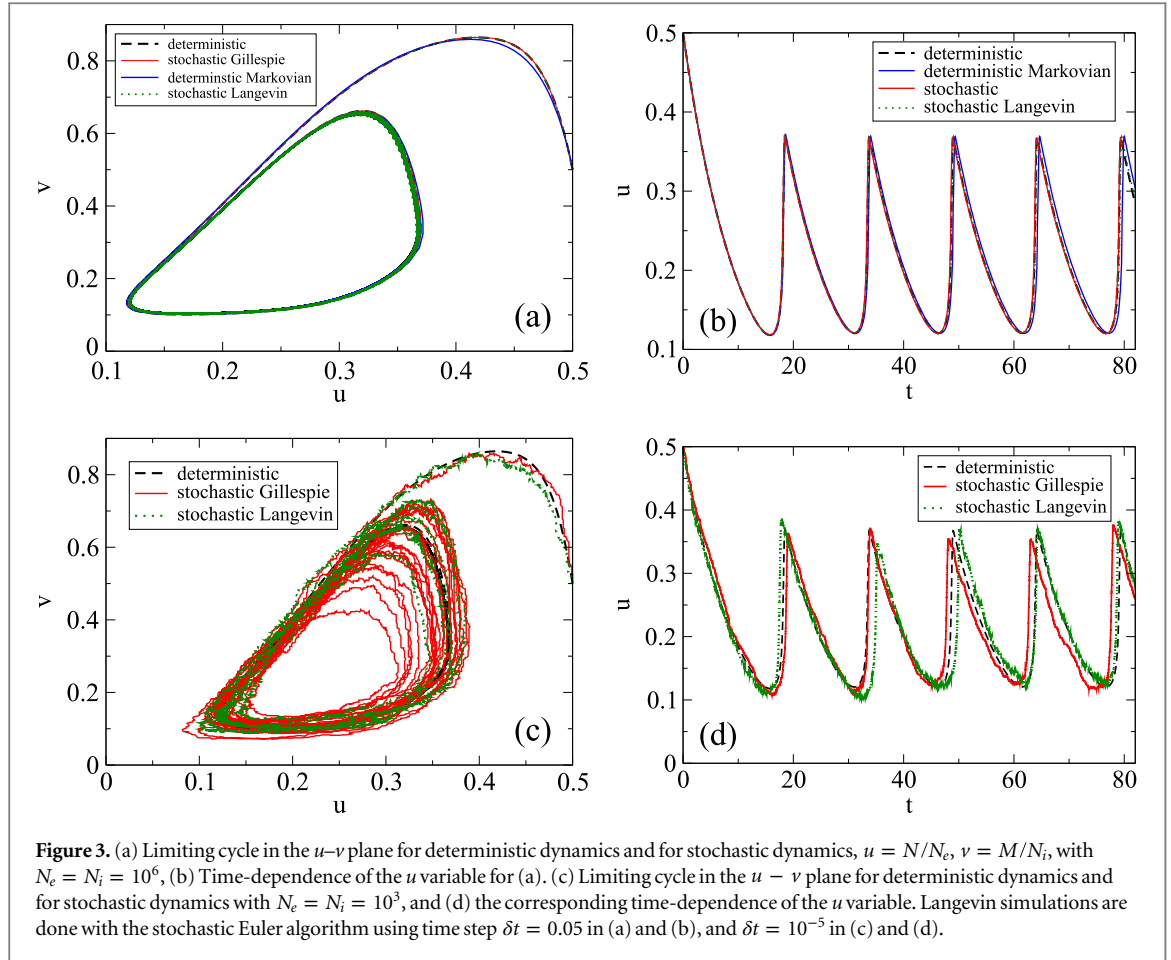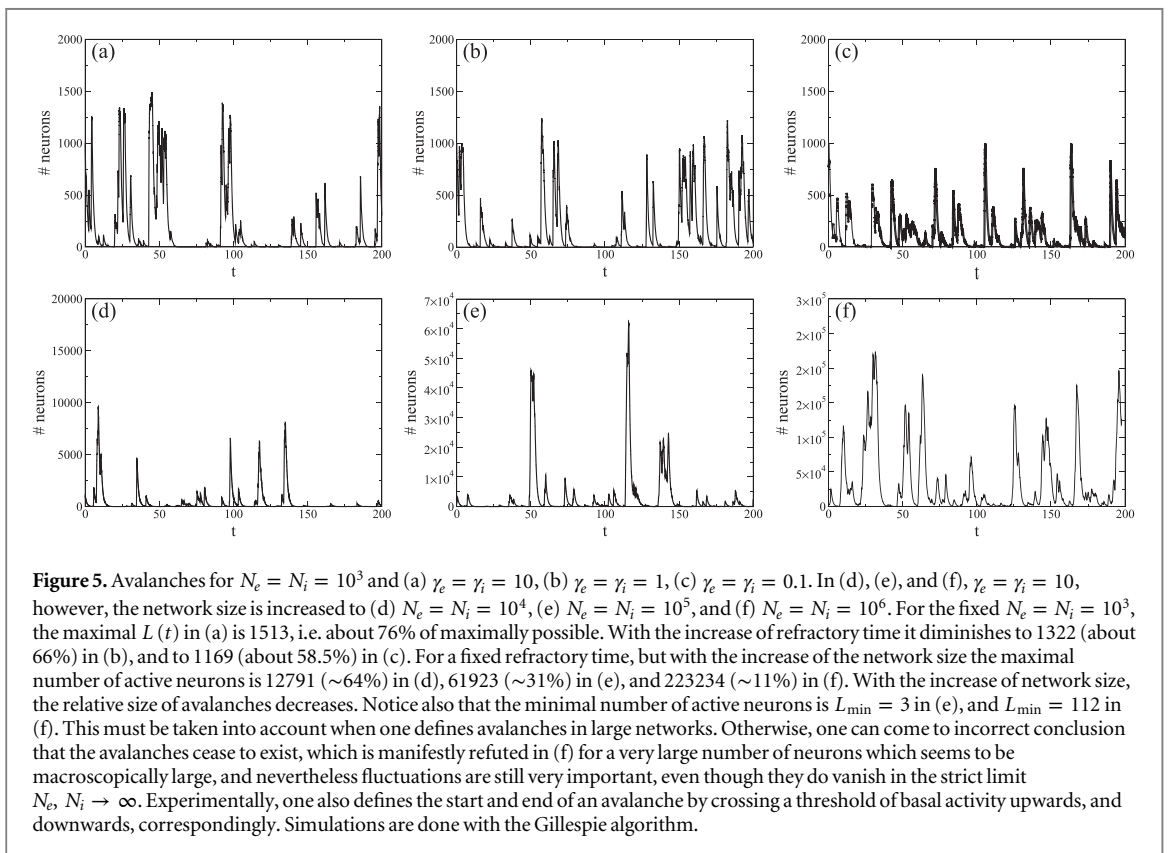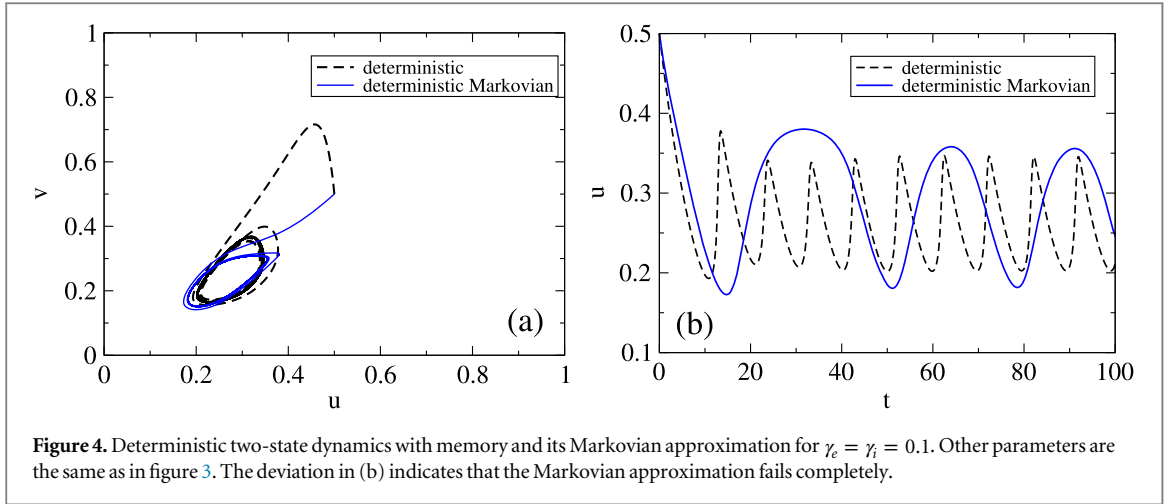
Figure 3. (a) Limiting cycle in the u–v plane for deterministic dynamics and for stochastic dynamics, $u = N/N_e$, $v = M/N_i$, with $N_e = N_i = 10^6$, (b) Time-dependence of the u variable for (a). (c) Limiting cycle in the $u - v$ plane for deterministic dynamics and for stochastic dynamics with $N_e = N_i = 10^3$, and (d) the corresponding time-dependence of the u variable. Langevin simulations are done with the stochastic Euler algorithm using time step $\delta t = 0.05$ in (a) and (b), and $\delta t = 10^{-5}$ in (c) and (d).
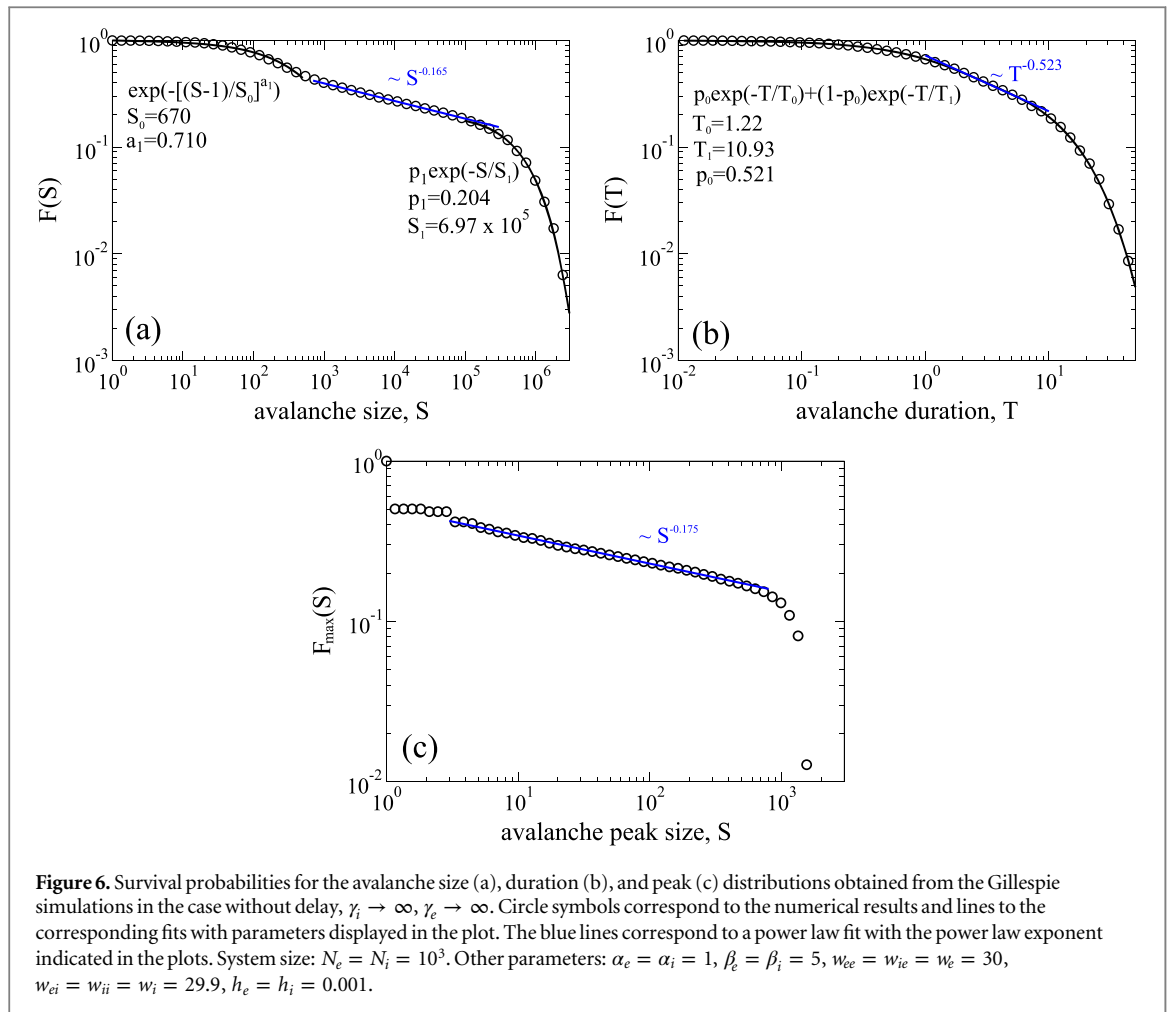
3. Results and discussion

3.1. Oscillatory dynamics of neuronal network

First, we test our stochastic simulations done with XPPAUT [42] against nonlinear deterministic dynamics for a very large network size with $N_e = 10^6$ and $N_i = 10^6$. For this, departing from the parameter set in [16] (the case without delay) we use a set of parameters, where an oscillatory dynamics emerges: $\alpha_e = 0.1$, $\alpha_i = 0.2$, $\beta_e = 1$, $\beta_i = 2$, $\gamma_e = 10$, $\gamma_i = 10$, $w_{ee} = 32$, $w_{ei} = 32$, $w_{ie} = 28$, $w_{ii} = 2$, $h_e = -3.8$, $h_i = -9$, and the transfer function in equation (11). Time is in milliseconds and the rate constants are in inverse milliseconds. The difference is barely detectable in figures 3(a), (b), where we present the results of stochastic simulations done both with the exact Gillespie algorithm and within the approximate Langevin dynamics. However, stochastic effects become immediately seen in figures 3(c), (d), where we reduced the number of neurons to $N_e = 10^3$ and $N_i = 10^3$. We also compare in figures 3(a), (b), the results for the considered dynamics and its two-variable Markovian approximation given by the standard Wilson–Cowan model in which, however, the transfer functions are renormalized in accordance with equation (5), where the parameter $\beta\tau_d$ is replaced by $\beta_e/\gamma_e = 0.1$ and $\beta_i/\gamma_i = 0.2$, correspondingly. The deviations are visible, but small. However, the differences become very pronounced for small $\gamma_e = \gamma_i = 0.1$ corresponding to the mean refractory period $\tau_d = 1/\gamma_{e,i}$ of 10 ms. Then, the Markovian approximation fails completely, see in figure 4, especially in part (b), revealing that neither the form of the oscillations, not their period are reproduced even approximately. Especially remarkable is that contrary to intuition the increase of the refractory period of a single neuron does not increase the period of oscillations, as the Markovian approximation predicts, but rather makes it smaller—the tendency is opposite! Therefore, non-Markovian memory effects generally do matter and one should take such effects seriously into account. With a small further decrease of γ_e, γ_i to $\gamma_e = \gamma_i \approx 0.08873$ with $\tau_d \approx 11.270$ the oscillations are terminated by a supercritical Hopf bifurcation (not shown). Interestingly, the Markovian approximation also predicts such a termination, but at a slightly larger critical value $\gamma_e = \gamma_i \approx 0.0987$ with critical $\tau_d \approx 10.132$. This makes clear that the phase transitions between the quiescent network and the network undergoing synchronized oscillations are possible with respect to the length of the refractory period used as a control parameter.

Figure 4. Deterministic two-state dynamics with memory and its Markovian approximation for $\gamma_e = \gamma_i = 0.1$. Other parameters are the same as in figure 3. The deviation in (b) indicates that the Markovian approximation fails completely.

Figure 5. Avalanches for $N_e = N_i = 10^3$ and (a) $\gamma_e = \gamma_i = 10$, (b) $\gamma_e = \gamma_i = 1$, (c) $\gamma_e = \gamma_i = 0.1$. In (d), (e), and (f), $\gamma_e = \gamma_i = 10$, however, the network size is increased to (d) $N_e = N_i = 10^4$, (e) $N_e = N_i = 10^5$, and (f) $N_e = N_i = 10^6$. For the fixed $N_e = N_i = 10^3$, the maximal $L(t)$ in (a) is 1513, i.e. about 76% of maximally possible. With the increase of refractory time it diminishes to 1322 (about 66%) in (b), and to 1169 (about 58.5%) in (c). For a fixed refractory time, but with the increase of the network size the maximal number of active neurons is 12791 (~64%) in (d), 61923 (~31%) in (e), and 223234 (~11%) in (f). With the increase of network size, the relative size of avalanches decreases. Notice also that the minimal number of active neurons is $L_{\min} = 3$ in (e), and $L_{\min} = 112$ in (f). This must be taken into account when one defines avalanches in large networks. Otherwise, one can come to incorrect conclusion that the avalanches cease to exist, which is manifestly refuted in (f) for a very large number of neurons which seems to be macroscopically large, and nevertheless fluctuations are still very important, even though they do vanish in the strict limit $N_e, N_i \to \infty$. Experimentally, one also defines the start and end of an avalanche by crossing a threshold of basal activity upwards, and downwards, correspondingly. Simulations are done with the Gillespie algorithm.
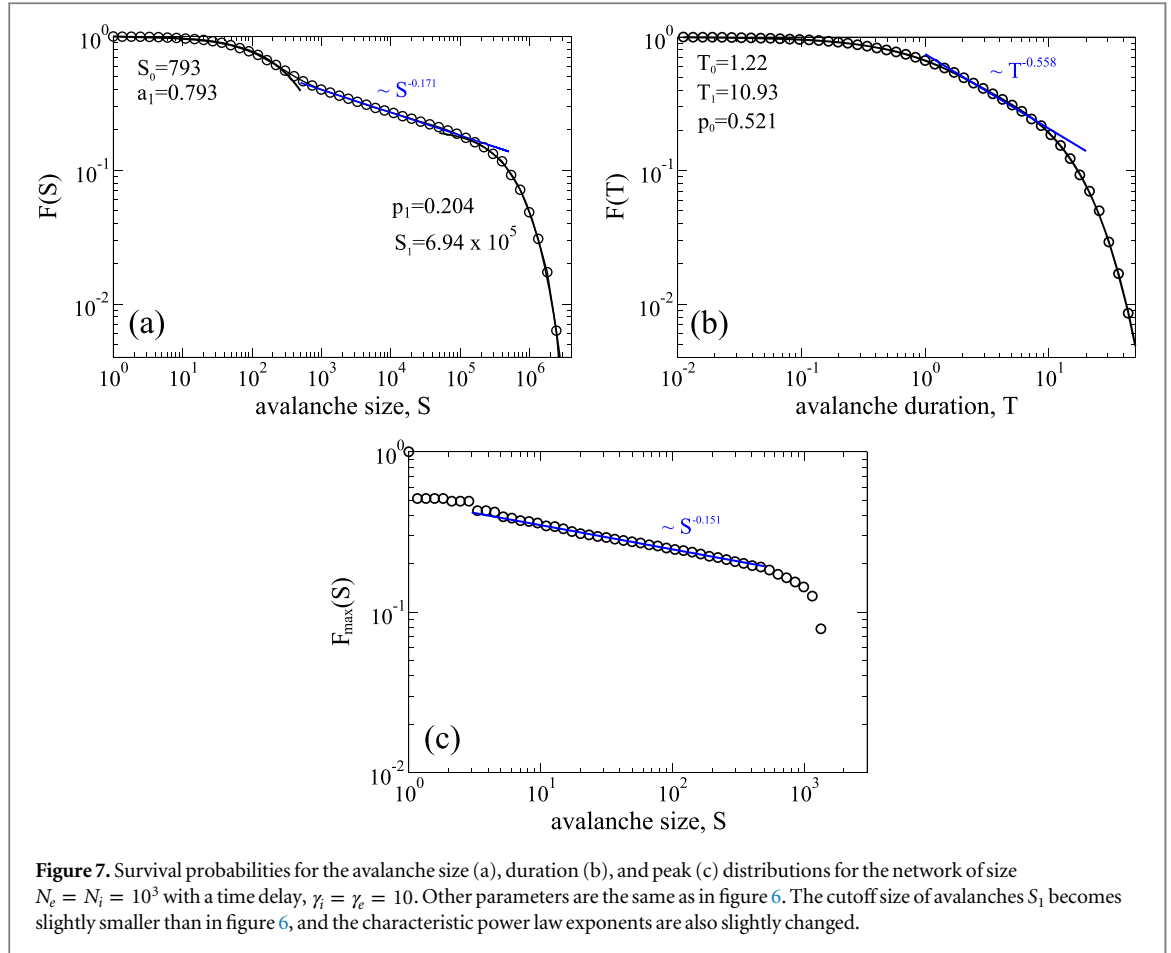
3.2. Noise-induced critical avalanche dynamics

In the remainder, we investigate the influence of memory effects on the avalanche dynamics. As it has been shown in [15], in order to have avalanche dynamics the excitatory and inhibitory processes should be nearly balanced, and the network should have a so-called feedforward structure. Then, one can achieve a sort of SOC state [6, 7] sustained due to intrinsic mesoscopic fluctuations. Very different from other SOC models, here fluctuations play a major role and in the deterministic limit avalanches disappear, i.e. they are of mesoscopic nature. The nullclines of 2D deterministic dynamics in the absence of memory effects, $\dot{u} = 0$ and $\dot{v} = 0$, should cross at a very small angle in the u–v plane, so that fluctuations can produce large amplitude outbursts of the u and v variables moving synchronously but randomly, i.e. the subpopulations of excitatory and inhibitory neurons are synchronized exhibiting stochastic dynamics at the same time [15]. In the same spirit, we choose $\alpha_e = \alpha_i = 1$, $\beta_e = \beta_i = 5$, $w_{ee} = w_{ie} = w_e = 30$, $w_{ei} = w_{ii} = w_i = 29.9$, so that $w_e - w_i \ll w_e + w_i$, and overall excitation slightly dominates over inhibition. Furthermore, we choose $h_e = h_i = 0.001$, and the transfer function in equation (10), as in [15]. The rates γ_e, γ_i and the number of neurons were varied. Large $\gamma_e = \gamma_i = 10$

Figure 6. Survival probabilities for the avalanche size (a), duration (b), and peak (c) distributions obtained from the Gillespie simulations in the case without delay, $\gamma_i \to \infty$, $\gamma_e \to \infty$. Circle symbols correspond to the numerical results and lines to the corresponding fits with parameters displayed in the plot. The blue lines correspond to a power law fit with the power law exponent indicated in the plots. System size: $N_e = N_i = 10^3$. Other parameters: $\alpha_e = \alpha_i = 1$, $\beta_e = \beta_i = 5$, $w_{ee} = w_{ie} = w_e = 30$, $w_{ei} = w_{ii} = w_i = 29.9$, $h_e = h_i = 0.001$.

corresponds to a small refractory time of 0.1 ms, whereas $\gamma_e = \gamma_i = 0.1$ corresponds to a profound delay with $\tau_d = 10$ ms, so that the individual spiking rate of neurons cannot exceed 100 Hz being limited by the refractory period. Typical avalanche dynamics is shown in figure 5 for $L(t) = N(t) + M(t)$ with $N_e = N_i = 10^3$, and (a) $\gamma_e = \gamma_i = 10$, (b) $\gamma_e = \gamma_i = 1$, (c) $\gamma_e = \gamma_i = 0.1$. Furthermore, in figures 5(d)–(f), we show the influence of an increasing number of neurons on the avalanche dynamics. The following tendencies are clear. First, the increase of the refractory period reduces the maximal amount of neurons involved in spiking, from about 76% in figure 5(a) to 58.5% in figure 5(c). Such a tendency is already expected from the renormalization of the transfer function in the Markovian approximation, cf equation (5). However, this tendency is in fact much weaker since $\beta\tau_d = 50$ in the part (c), and the renormalization argumentation would predict almost complete suppression of avalanches for such a delay. Even more astonishing is that avalanches still did not vanish even for a very large $N_i = N_e = 10^6$, see figure 5(f). This is very different from the oscillatory dynamics of a network of the same size, cf figures 3(a) and (b), which is practically deterministic. Of course, with increasing network size, the relative amplitude of avalanches becomes ever smaller, and there also emerges a minimal number of neurons excited, i.e. the network activity never goes down to zero. However, this is also so in the real neuronal dynamics. Such a dominance of mesoscopic fluctuations in a system of millions of elements with a special (feedforward) structure of coupling is really astonishing. This is a feature of some critical state, as we know from statistical physics.

To statistically characterize the avalanche size distribution and their duration we proceed in accordance with the procedure outlined in [15]. It reflects, in part, also the experimental procedure [10]. Namely, we first discretize the time series with a time bin of the size $\langle \Delta t \rangle$, which corresponds to the averaged interspike time distance in a particular simulation. Then, an avalanche is defined by its start, when the spiking activity crosses some threshold level L_{thr}, and its end, when the network activity drops to ($L_{\text{thr}} = 0$) or below $L_{\text{thr}} > 0$ after some time, which defines the avalanche duration. The size is defined as the sum of the number of neurons active in each time bin during the avalanche. It is also defined experimentally in such a way. In essence, the size S of an avalanche is the integral of the network activity in figure 5 over the time during each avalanche period divided by the time bin width. Of course, as also in experiments the critical exponents discussed below depend both on the time bin width and on the basal level of neuronal activity L_{thr}. However, this dependence is weak for a truly critical dynamics. By doing statistical analysis, we first find the survival probability $F(S)$, or, equivalently, the

Figure 7. Survival probabilities for the avalanche size (a), duration (b), and peak (c) distributions for the network of size $N_e = N_i = 10^3$ with a time delay, $\gamma_i = \gamma_e = 10$. Other parameters are the same as in figure 6. The cutoff size of avalanches S_1 becomes slightly smaller than in figure 6, and the characteristic power law exponents are also slightly changed.

cumulative probability $1 - F(S)$ from the numerical data. Then, the distribution density follows as $p(S) = -\mathrm{d}F(S)/\mathrm{d}S$.

Let us start from the case without any time delay, $\tau_d = 0$. The survival probability for the avalanche size distribution $F(S)$ is shown in figure 6(a), for the time bin $\langle \Delta t \rangle = 0.00643597$ and $L_{\mathrm{thr}} = 0$. It shows three characteristic features: (1) an initial Weibull distribution, $F(S) = \exp(-[(S-1)^{a_1}/S_0^{a_1}])$, with $a_1 \approx 0.71$ and $S_1 \approx 670$; (2) an intermediate power law $F(S) \propto S^{a_2}$, with $a_2 \approx -0.165$, and (3) an exponential tail $F(S) = p_1 \exp(-S/S_1)$ with $p_1 \approx 0.204$ and $S_2 = 6.97 \times 10^5$. The size distribution $p(S)$ is, therefore, initially approximately a power law with negative exponent $a_1 - 1 \approx -0.29$, followed by a power law with negative exponent $a = a_2 - 1 \approx -1.165$. The latter one extends over approximately two size decades and ends with an exponential tail characterized by a cutoff size, S_1. The corresponding survival probability $F(T)$ for the avalanche durations T is shown in figure 6(b). It can be well fitted by a sum of two exponentials

$$F(T) = p_0 \exp\left(-T/T_0\right) + \left(1 - p_0\right)\exp\left(-T/T_1\right). \tag{15}$$

However, it also seems to display an intermediate power law over about one time decade, from 1 to 10 ms, with the power exponent $b_1 \approx -0.523$, and the cutoff time $T_1 \approx 10.93$ ms. Hence, the probability distribution $p(T) = -\mathrm{d}F(T)/\mathrm{d}T$ also appears to reflect an intermediate power law $p(T) \propto T^b$ with $b = b_1 - 1 \approx -1.523$. Interestingly, the duration of avalanches in experiments with organotypic cortical neuronal systems has a similar cutoff time of about 10–20 ms, with a maximal avalanche duration of about 40–80 ms, which is restricted by the period of γ-oscillations [10]. The intermediate power law also extends over about one time decade in the experiments. However, the experimental power law exponent is different, $b_{\mathrm{exp}} \approx -2$. It should be mentioned in this respect that the time bin in the experiments is also very different, $\Delta t \sim 1$–4 ms. One electrode measures in experiments a contribution of many neurons. In fact, coarse graining over some unknown ΔL should be done. The experimental size exponent a is also different, $a_{\mathrm{exp}} \approx -1.5$. It is not, however, the goal of this paper to provide a model fully consistent with the experimental observations, which are subject of ongoing research work and some controversy in the literature [43]. In this respect, a bi-exponential dependence can be perceived as a power law over one intermediate time decade, as our fit also shows.

As an additional characteristics of the avalanches size, one can also consider the maximal number of neurons activated at once during an avalanche, or the avalanche peak with a distribution density $p_{\mathrm{max}}(S) = -\mathrm{d}F_{\mathrm{max}}(S)/\mathrm{d}S$. The corresponding survival probability, $F_{\mathrm{max}}(S)$, also exhibits a power law,
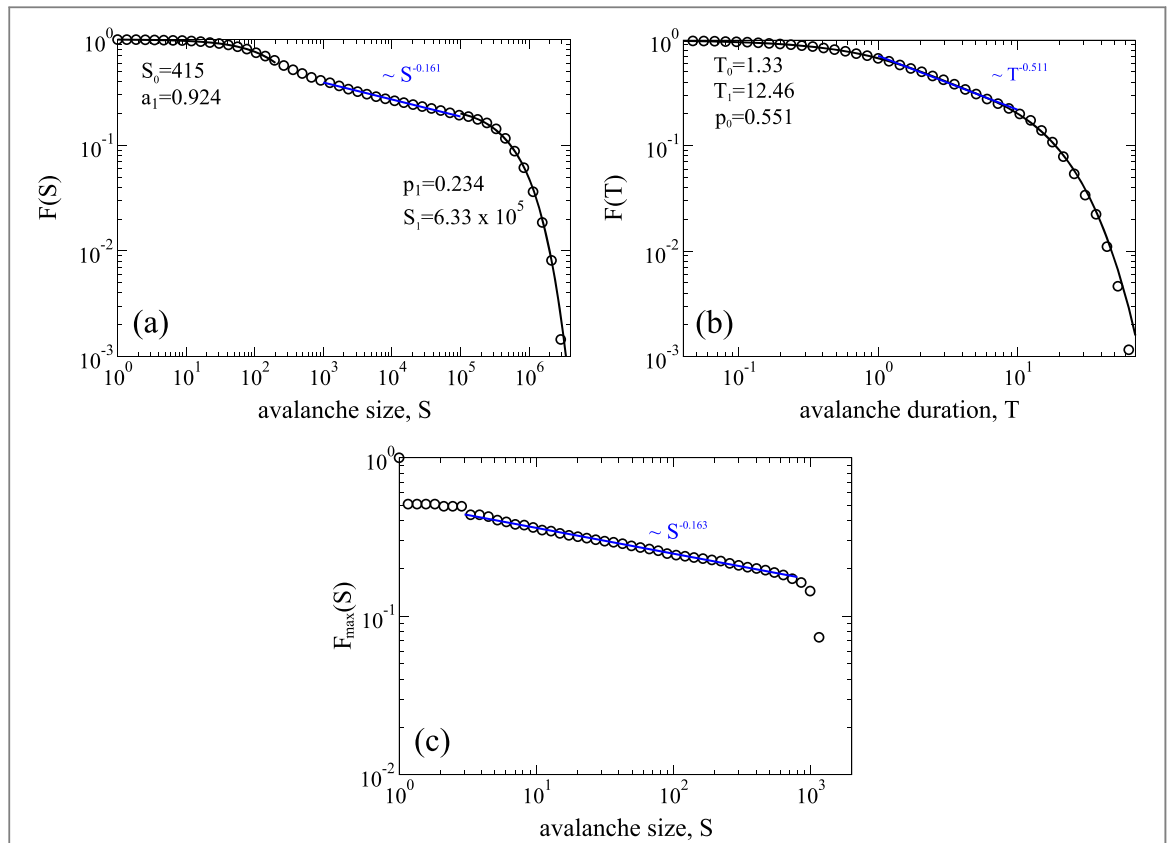
Figure 8. The time delay is increased further with respect to the one in figure 7. Here, $\gamma_i = \gamma_e = 1$. The other parameters remain the same. The cutoff size of avalanches, $S_1 = 6.33 \times 10^5$, is now visibly smaller than one without delay, $S_1 = 6.97 \times 10^5$ in figure 6. The cutoff time $T_1 = 12.46$ is increased with respect to $T_1 = 10.93$ in figures 6 and 7, i.e. the avalanches last longer. The power law exponents here deviate slightly in the opposite direction from the one in figure 7. They become closer to the case without delay in figure 6. This indicates that the time delay does not affect significantly the power law exponents.
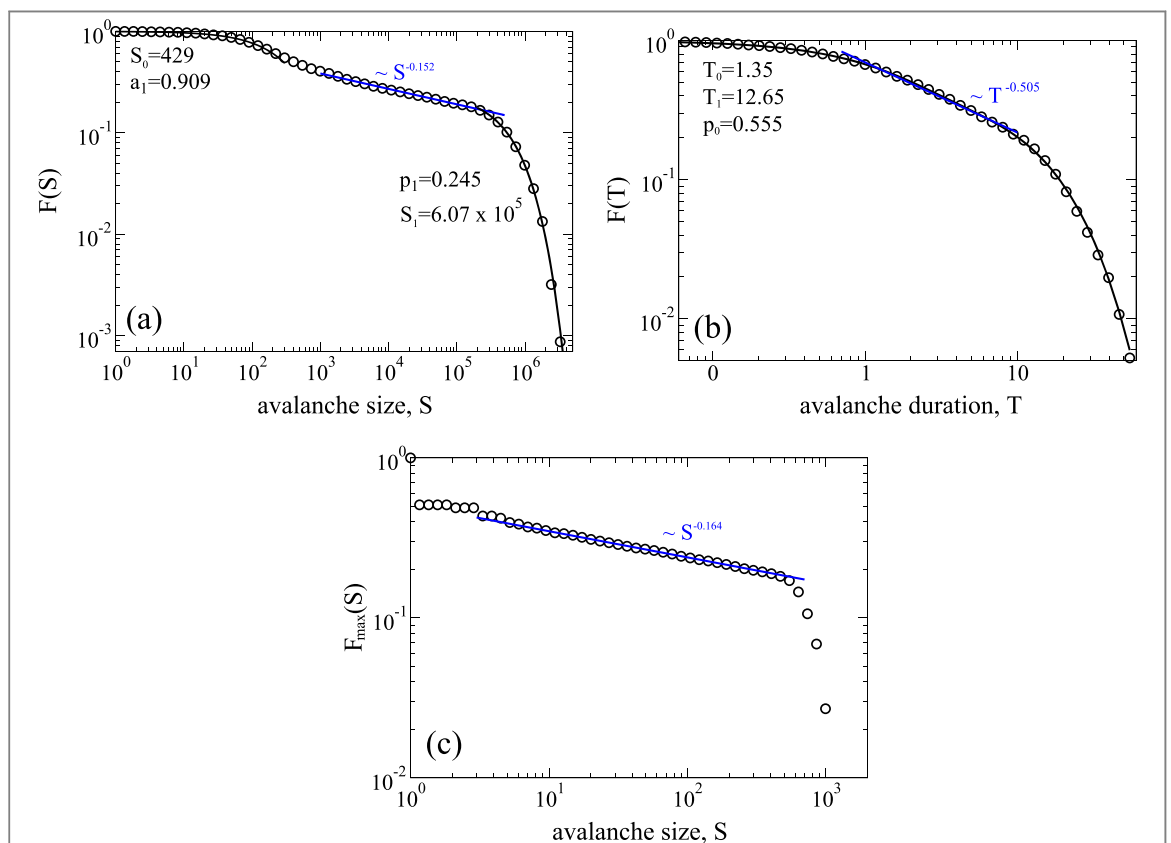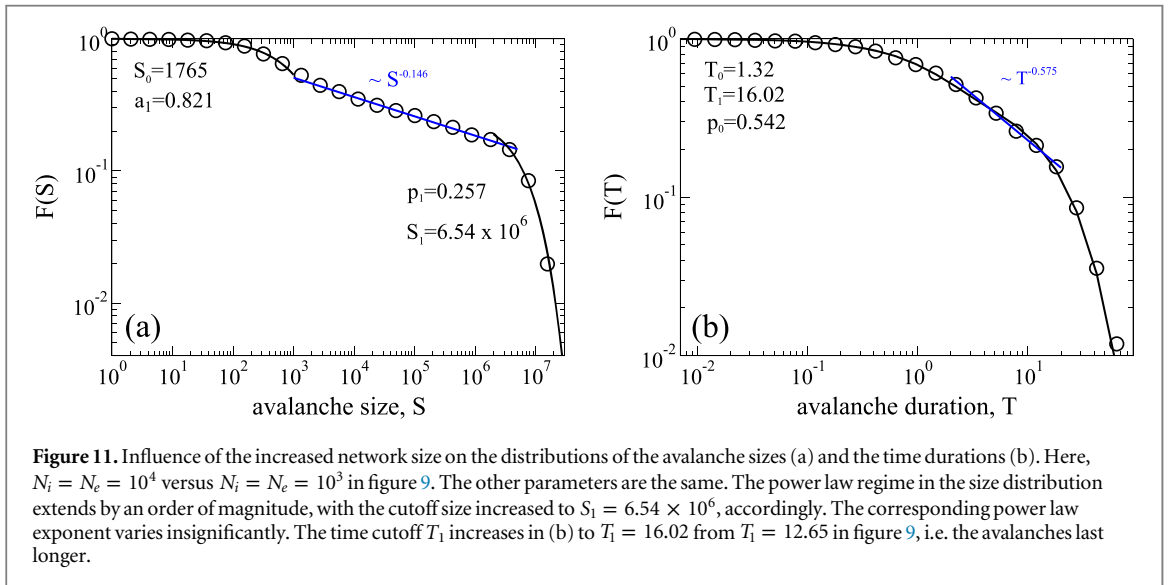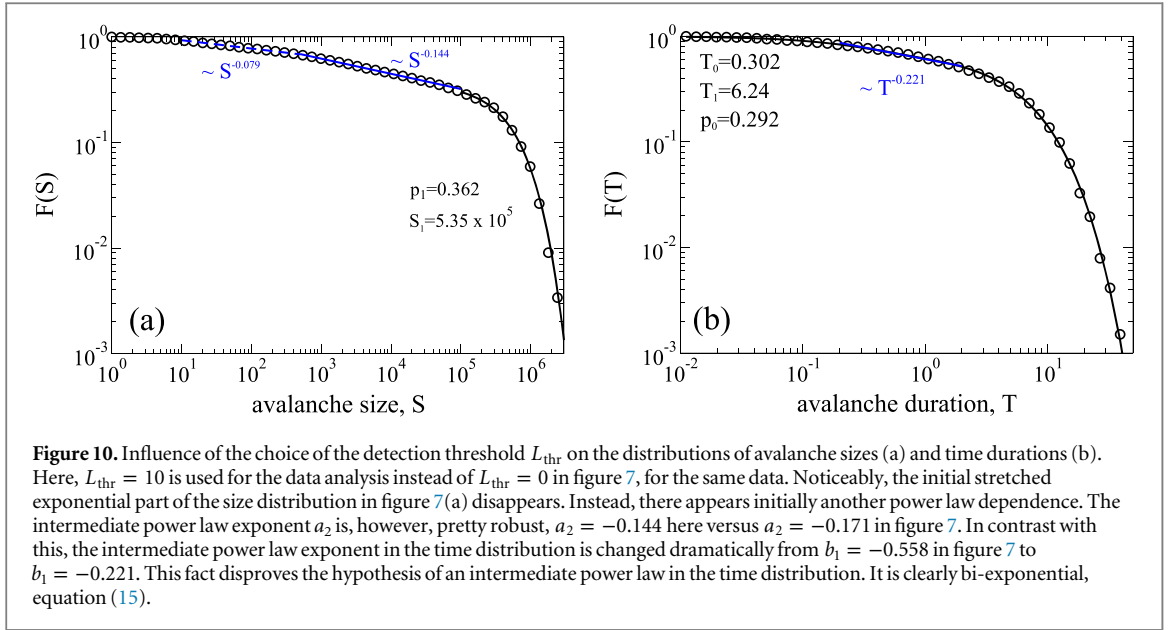

Figure 9. Influence of a further increase of the time delay by an order of magnitude, $\tau_d = 10$, on the distributions depicted in figures 7 and 8. Here, $\gamma_i = \gamma_e = 0.1$, and the other parameters are not changed. S_1 drops further to $S_1 = 6.07 \times 10^5$, and T_1 slightly increases to $T_1 = 12.65$. The power law exponents exhibit, however, merely some fluctuations without any systematic trend in figures 6–9.

Figure 10. Influence of the choice of the detection threshold L_{thr} on the distributions of avalanche sizes (a) and time durations (b). Here, $L_{thr} = 10$ is used for the data analysis instead of $L_{thr} = 0$ in figure 7, for the same data. Noticeably, the initial stretched exponential part of the size distribution in figure 7(a) disappears. Instead, there appears initially another power law dependence. The intermediate power law exponent a_2 is, however, pretty robust, $a_2 = -0.144$ here versus $a_2 = -0.171$ in figure 7. In contrast with this, the intermediate power law exponent in the time distribution is changed dramatically from $b_1 = -0.558$ in figure 7 to $b_1 = -0.221$. This fact disproves the hypothesis of an intermediate power law in the time distribution. It is clearly bi-exponential, equation (15).

Figure 11. Influence of the increased network size on the distributions of the avalanche sizes (a) and the time durations (b). Here, $N_i = N_e = 10^4$ versus $N_i = N_e = 10^3$ in figure 9. The other parameters are the same. The power law regime in the size distribution extends by an order of magnitude, with the cutoff size increased to $S_1 = 6.54 \times 10^6$, accordingly. The corresponding power law exponent varies insignificantly. The time cutoff T_1 increases in (b) to $T_1 = 16.02$ from $T_1 = 12.65$ in figure 9, i.e. the avalanches last longer.

$F_{max}(S) \propto S^{a_2'}$, with $a_2' \approx -0.175$, in figure 6(c). Hence, $p_{max}(S) \propto S^{a'}$, with $a' = a_2' - 1 \approx -1.175$, which only slightly differs from $a \approx -1.165$ meaning that the avalanche size is roughly proportional to its peak. However, the cutoff of $F_{max}(S)$ is super-exponentially sharp, because the maximal number of neurons involved in an avalanche at the same time is restricted by the total number of neurons in the network. Furthermore, F_{max} reveals a very large portion of avalanches whose peak does not exceed 10, which explains the initial stretched exponential dependence in figure 6(a). Strictly speaking, this part of the size distribution (with $L_{thr} = 0$) reflects a background or basal noise, where neurons practically do not interact with each other, and there are no avalanches of spontaneously increased activity, which are characterized by a power law distribution.

Next, we like to clarify how robust these features are for networks with a time delay. For this, we study the influence of the mean delay time by decreasing the rates $\gamma_e = \gamma_i$ from 10 through 1.0 to 0.1 in figures 7–9, respectively. The mean delay time increases, accordingly, from 0.1 through 1.0 to 10 ms. Even though the parameters of the distributions do change, these changes are not dramatical. In particular, the corresponding critical size exponent a changes from -1.165 (no delay), to -1.152, -1.171 and -1.161, respectively. Accordingly, the critical time exponent b changes from -1.523 (no delay) to -1.505, -1.558, and -1.511, respectively. Such changes are not statistically significant, and one cannot detect any systematic tendency upon a variation of τ_d. The point is that these exponents are also changed a bit, if we use e.g. $2\langle \Delta t \rangle$, or $3\langle \Delta t \rangle$ for the time bin (not shown). They also depend on the threshold L_{thr}. In this respect, if to change L_{thr} from 0 to 10, the initial stretched exponential part of the size distribution practically disappears. However, there appears an initial power law instead, see in figure 10. Remarkably, the intermediate power law exponent remains rather robust. It is
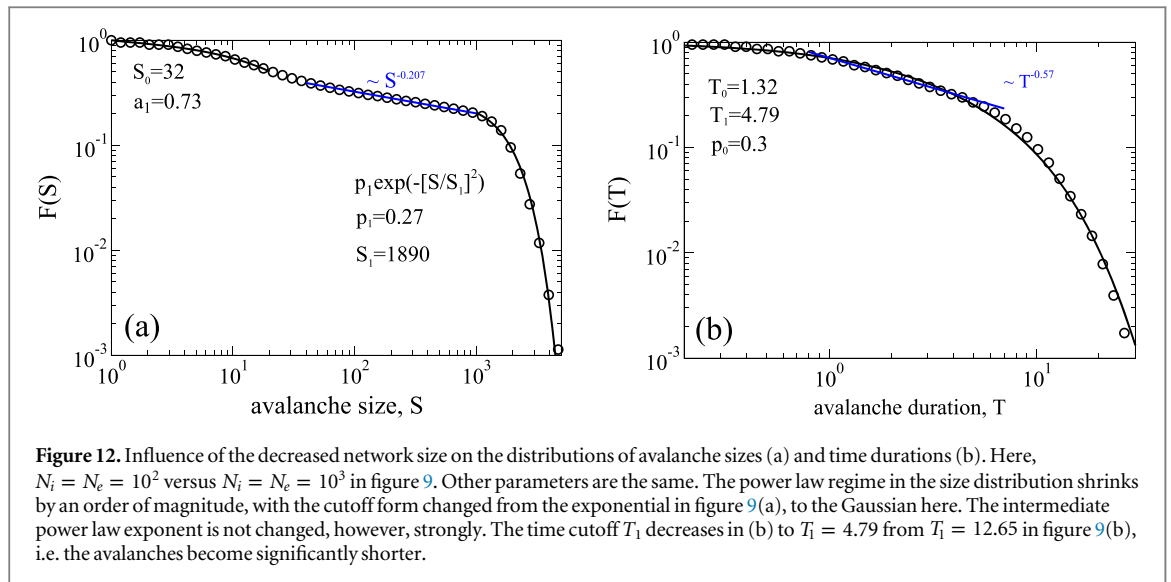
Figure 12. Influence of the decreased network size on the distributions of avalanche sizes (a) and time durations (b). Here, $N_i = N_e = 10^2$ versus $N_i = N_e = 10^3$ in figure 9. Other parameters are the same. The power law regime in the size distribution shrinks by an order of magnitude, with the cutoff form changed from the exponential in figure 9(a), to the Gaussian here. The intermediate power law exponent is not changed, however, strongly. The time cutoff T_1 decreases in (b) to $T_1 = 4.79$ from $T_1 = 12.65$ in figure 9(b), i.e. the avalanches become significantly shorter.
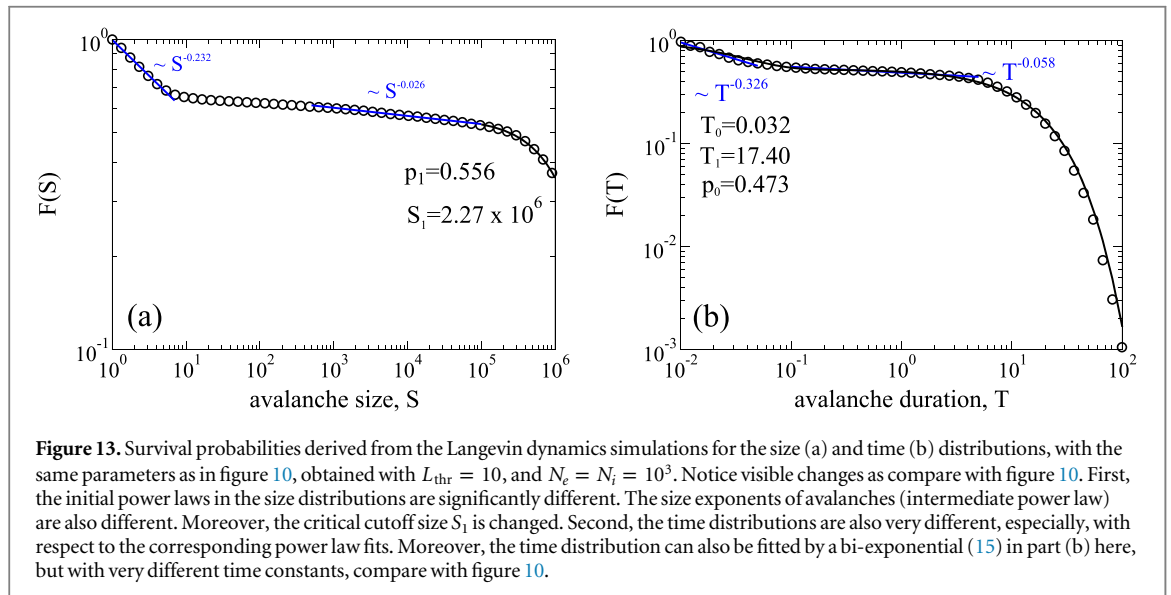
Figure 13. Survival probabilities derived from the Langevin dynamics simulations for the size (a) and time (b) distributions, with the same parameters as in figure 10, obtained with $L_{thr} = 10$, and $N_e = N_i = 10^3$. Notice visible changes as compare with figure 10. First, the initial power laws in the size distributions are significantly different. The size exponents of avalanches (intermediate power law) are also different. Moreover, the critical cutoff size S_1 is changed. Second, the time distributions are also very different, especially, with respect to the corresponding power law fits. Moreover, the time distribution can also be fitted by a bi-exponential (15) in part (b) here, but with very different time constants, compare with figure 10.

changed from $a = -1.171$ in figure 7(a) to $a = -1.144$ in figure 10(a). This is a small variation. Notice, however, that the results in figure 10(b) in fact reject the hypothesis that there is an intermediate power law in the time distribution of avalanches. First, the power law region changes from larger to smaller times, and also (more important!) the corresponding time exponent changes from $b_1 = -0.558$ in figure 7(b) to $b_1 = -0.221$ in figure 10(b). Clearly, such a strong influence of the choice of L_{thr} on the 'power law' exponent b makes it clear that this is not a power law. In fact, the time distribution is clearly biexponential.

Though plausible until this point, it remains, however, strictly speaking, still not quite clear if a is indeed a critical exponent. If true, the extension of the power law domain of the whole size-distribution and the cutoff size S_1 should increase with the system size accordingly. Indeed, if we increase the system size tenfold keeping the other parameters the same as in figure 9, the power law domain in the size distributions also increases by an order of size magnitude, see in figure 11(a). The time cutoff T_1 also increases in figure 11(b), i.e. the avalanches become longer. Also with decreasing the system size tenfold the power law domain shrinks accordingly in size, see figure 12(a), and the avalanches become essentially shorter, as indicated by the decreased cutoff time T_1 in figure 12(b). Such scaling dependencies on the system size are typical in experiments. From this we can conclude that the size exponent a is indeed a critical exponent. However, within the considered model the avalanches do gradually vanish with an increase of the system size. Therefore, the adjective 'critical' should be used also with respect to the exponent a with some reservations. We consider a rather atypical SOC model, even though the exponent a is by chance close to that of the sandpile model [6, 7]. It should also be noticed that the initial distributions of the sizes and times and the tail functional dependencies can be sensitive to both the system size

and the choice of the threshold L_{thr}. For example, the size distribution exhibits a Gaussian tail in figure 12(a), for a small system size. The intermediate power law in the size distribution is, however, rather robust, with a being in the range $[-1.207, -1.144]$ for the data presented, with the average $\langle a \rangle \approx -1.164$.

3.2.1. Langevin dynamics of avalanches
Within the Langevin approximation of the discrete state dynamics, the avalanches look very similar to the ones depicted in figure 5. However, their statistics is very different. We performed the corresponding Langevin simulations for the same parameters as in figures 7, 10 with the time step taken to be the same $\langle \Delta t \rangle = 0.00618608$ as the time bin used to produce the results in figures 7, 10. We also used $L_{thr} = 10$ to analyze the data, as in figure 10. The results shown in figure 13 reveal similar intermediate power laws both in the size and the time distributions yielding $a_L \approx -1.026$, $b_L \approx -1.058$. However, these results differ essentially from the results obtained within the exact dynamic Monte Carlo simulations, compare figure 13 with figure 10. This indicates that the Gauss–Langevin or diffusional approximation of the genuine discrete state dynamics can deliver incorrect results for the fluctuation-induced avalanche dynamics. This fact makes any analytical theory for the numerical results presented in this work especially challenging. It is almost hopeless to develop such a theory for the discrete state avalanche dynamics within the studied model. Multi-dimensional birth-and-death processes are very difficult for any analytical treatment. Within the Langevin dynamics approximation, or the equivalent multi-dimensional Fokker–Planck equation description an analytical treatment is more feasible. However, such a theory will not help to understand the critical features of the discrete state dynamics, as our numerical results imply.

4. Summary and conclusions

In this paper, we studied a generalization of the stochastic Wilson–Cowan model of neuronal network dynamics aimed to incorporate a refractory period delay on the level of individual elements. Considered as stochastic bistable elements such model neurons exhibit non-Markovian dynamics with memory, which can be characterized by a non-exponential RTD in the resting state of the neuron (semi-Markovian description), or, alternatively, by the related memory kernel within a generalized master equation description. Such a non-Markovian description generally allows for a Markovian embedding by enlarging the dynamical space upon introduction of new state variables. The simplest two-state non-Markovian model with an exponentially decaying memory kernel can be embedded as a three state cyclic Markovian model, where the refractory period is exponentially distributed. Multi-state Markovian embedding also allows to treat a special Erlangian distribution of the refractory periods, which can be sharply peaked at a characteristic delay time. Moreover, models of bursting neurons and neurons with a power law distributed memory can, in principle, be considered in this generic Markovian embedding setup. The approach of Markovian embedding is especially suitable to treat the mean-field dynamics of the network, which presents a Markovian renewal process in the enlarged space of collective network variables. This is the simplest kind of network, where all the elements are virtually connected in all-to-all fashion. In this respect, the mean field dynamics of a network of non-Markovian renewal elements does not represent a renewal process in the reduced space of non-Markovian collective variables. Then, all the elements must be treated individually, keeping trace of their individual memory. The methodology of Markovian embedding allows to circumvent this problem for the mean-field dynamics.

  In the Wilson–Cowan model, two different sorts of interacting neurons are considered, excitatory and inhibitory. We focused on the simplest non-Markovian generalization of this model, where the observed two-state non-Markovian dynamics of a single neuron is embedded as a three state cyclic Markovian process. The corresponding nonlinear mean-field dynamics is four-dimensional. It has two dimensions more than in the original model. Moreover, it is stochastic and includes mesoscopic fluctuations due to a finite network size. For a sufficiently large system size, stochastic dynamics can be described within a Langevin equation approximation following the so-called chemical Langevin equation approach, with the noise terms vanishing in the deterministic limit of infinite size. We also exactly simulated the underlying dynamics as a continuous time Markovian random walk on a four-dimensional lattice using the well-known dynamical Monte Carlo (Gillespie) algorithm. The results of both stochastic approaches agree well with the deterministic dynamics within an oscillatory regime for a very large number of elements (several millions). Here, we showed that non-Markovian effects can be very essential. In particular, even deterministic dynamics with an exponentially decaying memory in the space of observable variables can be very different from the dynamics obtained within the Markovian approximation utilizing a delay-renormalized transfer function—the simplest approach to account for the delay effects. However, already the simplest approach allows to describe a dynamical phase transition from the silent network to coherent nonlinear oscillations of synchronized neurons upon a change of the delay period. This important feature is absent in the original Wilson–Cowan model.

In more detail, we investigated the avalanche dynamics in a critically balanced network, where the processes of excitation and inhibition nearly compensate each other in the deterministic limit, where no avalanches are possible within the model considered. Mesoscopic noise fluctuations make, however, avalanches possible even in large networks with millions of neurons, where the deterministic description becomes completely inadequate, very differently from the oscillatory dynamics in such large networks. This result goes beyond the results in [15], where the avalanches cease to exist for already several tens of thousands of elements. Even though a large delay should suppress avalanches by a transfer function renormalization if to think within the Markovian approximation, in reality the suppression is much weaker. Moreover, it turns out that the power law characterizing the distribution of avalanche sizes is very robust with respect to variation of both the delay period, and the system size, over several orders of magnitude, as well as the choice of the avalanche threshold. The latter fact proves that this is a real power law originated due to critical dynamics. It is characterized by a power-law exponent around $a \sim -1.16$ which is similar to the size exponent of the critical sandpile dynamics (though the both models are not really comparable). However, it is different from the critical exponent -1.62 found in [15], though for different network parameters. The distribution of the avalanche durations is, however, biexponential. We disproved that it presents a power law within our model, even though it can look like a power law over one time decade, as in experiments. In this respect, experiments [9–11] seem to reveal a real power law with the critical size exponent $a_{\exp} \sim -1.5$ because its range extends with the growing system size. However, the experimental power law in the time duration does not show this important property. As a matter of fact, it extends over merely one time decade being restricted by the period of γ-oscillations. Any power law extending over one time or spatial decade can be fitted by a sum of just two exponentials, as we also show in this work for the time distribution. A further research is, therefore, required to clarify the nature of the apparent power law feature in the avalanche time distribution for the observed neuronal avalanches.

Also very important is that the Langevin or diffusional approximation does change the critical exponents of the studied avalanche dynamics. Especially, the critical Langevin size exponent is different, $a_L \sim -1.03$. This feature should be kept in mind while doing diffusional approximations in other models of critical dynamics. It may deliver incorrect results even for a large number of elements.

We believe that the results of this work have methodological value and can be extended onto the dynamics of other networks with delay. They can also serve as a basis for further investigations of the role of non-Markovian memory effects in the dynamics of Wilson–Cowan type neuronal networks, including networks of bursting neurons, and networks with non-trivial topology, which we plan to investigate in the future.

Acknowledgments

References

[1] Sole R and Goodwin B 2000 *Signs of Life: How Complexity Pervades Biology* (New York: Basic Books)
[2] Bashan A, Bartsch R P, Kantelhardt J W, Havlin S and Ivanov P Ch 2012 *Nat. Commun.* **3** 702
[3] Lo C-C, Amaral L A N, Havlin S, Ivanov P Ch, Penzel T, Peter J-H and Stanley H E 2002 *Europhys. Lett.* **57** 625
[4] Lo C-C, Chou T, Penzel T, Scammell T, Strecker R E, Stanley H E and Ivanov P Ch 2004 *Proc. Natl Acad. Sci. USA* **101** 17545
[5] Lo C-C, Bartsch R P and Ivanov Ch 2013 *Europhys. Lett.* **102** 10008
[6] Bak P, Tang C and Wiesenfeld K 1987 *Phys. Rev. Lett.* **59** 381
[7] Kertesz J and Kiss L B 1990 *J. Phys. A: Math. Gen.* **23** L433
[8] Pribram K H 2013 *The Form Within: My Point of View* (Westport: Prospecta Press)
[9] Beggs J M and Plenz D 2003 *J. Neurosci.* **23** 11167
[10] Plenz D 2012 *Eur. Phys. J. Spec. Top.* **205** 259
[11] Beggs J M 2008 *Phil. Trans. R. Soc.* A **366** 329
[12] Wilson H R and Cowan J D 1972 *Biophys. J.* **12** 1
     Wilson H R and Cowan J D 1973 *Kybernetik* **13** 55
[13] Destexhe A and Sejnowski T J 2009 *Biol. Cybern.* **101** 1
[14] Schuster H G and Wagner P 1990 *Biol. Cybern.* **64** 77
[15] Benayon M, Cowan J D, van Drongelen W and Wallace E 2010 *PLoS Comput. Biol.* **7** e1000846
[16] Wallace E, Benayon M, van Drongelen W and Cowan J D 2011 *PLoS One* **6** e14804
[17] Babloyanz A and Destexhe A 1986 *Proc. Natl Acad. Sci. USA* **83** 3513
[18] Ott E 1993 *Chaos in Dynamical Systems* (Cambridge: Cambridge University Press)
[19] Gillespie D T 1976 *J. Comput. Phys.* **22** 403
[20] Gillespie D T 2000 *J. Chem. Phys.* **113** 297
[21] Phillips R, Kondev J, Theriot J and Garcia H G 2013 *Physical Biology of the Cell* 2nd edn (London: Garland Science)
[22] Pikovsky A S and Kurths J 1997 *Phys. Rev. Lett.* **78** 775
[23] Linder B, Garcia-Ojalvo J, Neiman A and Schimansky-Geier L 2004 *Phys. Rep.* **392** 321
[24] Gammaitoni L, Hänggi P, Jung P and Marchesoni F 1998 *Rev. Mod. Phys.* **70** 223

[25] Schmid G, Goychuk I and Hänggi P 2001 *Europhys. Lett.* **56** 22
[26] Turalska M, Lukovic M, West B J and Grigolini P 2009 *Phys. Rev.* E **80** 021110
[27] Pinto I L D, Escaff D, Harbola U, Rosas A and Lindenberg K 2014 *Phys. Rev.* E **89** 052143
[28] Nikitin A, Neda Z and Viscek T 2001 *Phys. Rev. Lett.* **87** 024101
[29] Prager T, Naundorf B and Schimansky-Geier L 2003 *Physica* A **325** 176
[30] Prager T, Falcke M, Schimansky-Geier L and Zaks M A 2007 *Phys. Rev.* E **76** 011118
[31] Kouvaris N, Müller F and Schimansky-Geier L 2010 *Phys. Rev.* E **82** 061124
[32] Cox D E 1962 *Renewal Theory* (London: Methuen)
[33] Hughes B D 1995 *Random Walks and Random Environments* vol 1, 2 (Oxford: Clarendon)
[34] Kenkre V M, Montroll E W and Shlesinger M F 1973 *J. Stat. Phys.* **9** 45
[35] Goychuk I 2004 *Phys. Rev.* E **70** 016109
[36] Koch Ch 1999 *Biophysics of Computation, Information Processing in Single Neurons* (New York: Oxford University Press)
[37] Tsimring L S and Pikovsky A 2001 *Phys. Rev. Lett.* **87** 250602
[38] Haus J W and Kehr K W 1987 *Phys. Rep.* **150** 263
[39] Stehfest H 1970 *Commun. ACM* **13** 47
     Stehfest H 1970 *Commun. ACM* **13** 624 (erratum)
[40] Goychuk I and Hänggi P 2006 *Chem. Phys.* **324** 160
[41] Goychuk I 2014 *Commun. Comput. Inf. Sci.* **438** 325
[42] Ermentrout B 2002 *Simulating, Analyzing, and Animating Dynamical Systems: A Guide to XPPAUT for Researchers and Students* (Philadelphia: SIAM)
[43] Touboul J and Destexhe A 2010 *PLoS One* **5** e8982