

# Radio pulsar glitches as a state-dependent Poisson process

W. Fulgenzi<sup>1,\*</sup>, A. Melatos<sup>1,†</sup>, B. D. Hughes<sup>2</sup>,

<sup>1</sup>*School of Physics, The University of Melbourne, Parkville, VIC 3010, Australia*

<sup>2</sup>*School of Mathematics and Statistics, The University of Melbourne, Parkville, VIC 3010, Australia*

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

Gross–Pitaevskii simulations of vortex avalanches in a neutron star superfluid are limited computationally to  $\lesssim 10^2$  vortices and  $\lesssim 10^2$  avalanches, making it hard to study the long-term statistics of radio pulsar glitches in realistically-sized systems. Here, an idealised, mean-field model of the observed Gross–Pitaevskii dynamics is presented, in which vortex unpinning is approximated as a state-dependent, compound Poisson process in a single random variable, the spatially-averaged crust–superfluid lag. Both the lag-dependent Poisson rate and the conditional distribution of avalanche-driven lag decrements are inputs into the model, which is solved numerically (via Monte-Carlo simulations) and analytically (via a master equation). The output statistics are controlled by two dimensionless free parameters:  $\alpha$ , the glitch rate at a reference lag, divided by the spin-down time-scale, divided by the critical lag for unpinning; and  $\beta$ , the minimum fraction of the lag that can be restored by a glitch. The system evolves naturally to a self-regulated stationary state, whose properties are determined by  $\alpha/\alpha_c(\beta)$ , where  $\alpha_c(\beta) \approx \beta^{-1/2}$  is a transition value. In the regime  $\alpha \gtrsim \alpha_c(\beta)$ , one recovers qualitatively the power-law size and exponential waiting-time distributions observed in many radio pulsars and Gross–Pitaevskii simulations. For  $\alpha \ll \alpha_c(\beta)$ , the size and waiting-time distributions are both power-law-like, and a correlation emerges between size and waiting time until the next glitch, contrary to what is observed in most pulsars. Comparisons with astrophysical data are restricted by the small sample sizes available at present, with  $\leq 35$  events observed per pulsar.

**Key words:** pulsars: general – stars: neutron – stars: rotation

## 1 INTRODUCTION

More than 400 rotational glitches have been detected during nearly five decades of long-term monitoring and timing of rotation-powered radio pulsars (Hobbs et al. 2004; Melatos et al. 2008; Espinoza et al. 2011; Yu et al. 2013).<sup>1</sup> With this volume of data, it is now feasible to analyse meaningfully the statistics of individual pulsars, seven of which have glitched more than 10 times. Most pulsars exhibit a broad distribution of glitch sizes consistent with a power law, whose index differs among individual pulsars, and an exponential distribution of inter-glitch waiting times (Melatos et al. 2008). Some pulsars, however, including Vela and PSR J0537–6910, glitch quasi-periodically with unimodal size and waiting-time distributions (Melatos et al. 2008; Shannon et al. 2016). A recent nonparametric analysis hints at an intermediate class of behaviour, e.g. in PSR J1341–6220 (Ashton et al. 2017; Howitt et al. 2017). In all

but one pulsar (PSR J0537–6910), no correlation is observed between the size of a glitch and the waiting time until the next one. Recent analyses of the Crab pulsar suggest the existence of a minimum glitch size of  $\approx 0.05 \mu\text{Hz}$  (Espinoza et al. 2014) and changes in the mean glitch rate on a decadal time-scale (Lyne et al. 2015).

Numerous physical mechanisms have been proposed to explain pulsar glitches, including crust-quakes (Ruderman 1969; Smoluchowski 1970; Ruderman 1991; Larson & Link 2002; Jones 2003; Middleditch et al. 2006), core quakes (Takatsuka & Tamagaki 1988), superfluid vortex avalanches (Anderson & Itoh 1975; Cheng et al. 1988; Warszawski & Melatos 2011, 2013), magnetospheric state changes (Keith et al. 2013; Kerr et al. 2016) and hydrodynamical instabilities (Mastrano & Melatos 2005; Peralta et al. 2006; Melatos & Peralta 2007; Glampedakis & Andersson 2009; Sidery et al. 2010; Sourie et al. 2017). The reader is referred to Haskell & Melatos (2015) for a contemporary review. Historically, theoretical attention has focussed on the microphysics of the vortex unpinning model, e.g. *ab initio* calculations of the nuclear pinning potentials (Donati & Pizzochero 2004; Seveso et al. 2016), nuclear lattice geometry (Jones 1991), vortex motion near a pinning site (Sedrakian 2005), long-range vortex hopping

\* E-mail: w.fulgenzi@student.unimelb.edu.au

† E-mail: amelatos@unimelb.edu.au

<sup>1</sup> Electronic access to up-to-date glitch catalogues is available at the following locations on the World Wide Web: <http://www.jb.man.ac.uk/pulsar/glitches/gTable.html> and <http://www.atnf.csiro.au/people/pulsar/psrcat/glitchTbl.html>

(Haskell & Melatos 2016), and the role of mutual friction (Haskell 2016) and super-conductivity (Haskell et al. 2013).

Quantitative predictive theories that describe the long-term glitch statistics generated by the vortex avalanche model are rare (Warszawski & Melatos 2008; Melatos & Warszawski 2009; Warszawski & Melatos 2013; Haskell 2016). In the simplest physical scenario, where unpinning events are threshold-driven but spatially uncorrelated, the central limit theorem predicts size and waiting-time distributions much narrower than those observed (Warszawski & Melatos 2013). *Correlated unpinning* is therefore an essential ingredient. Numerical simulations based on the Gross-Pitaevskii equation (GPE) have identified two specific knock-on mechanisms that cause spatially correlated vortex unpinning (Warszawski & Melatos 2011; Warszawski et al. 2012).<sup>2</sup> GPE simulations generate size and waiting-time distributions, which are broadly consistent with observational data. For example, the largest GPE simulations published so far produce a power-law-like size distribution over  $\sim 1.5$  decades (Warszawski & Melatos 2011; Melatos et al. 2015). We emphasise, however, that the simulations are restricted computationally to unrealistically small systems featuring  $\lesssim 10^2$  vortices (c.f.  $\sim 10^{18}$  in a pulsar),  $\lesssim 10^2$  avalanches, and  $\lesssim 10$  pinning sites between vortices (c.f.  $\sim 10^9$  in a pulsar). There is an urgent need to supplement GPE studies with other theoretical approaches, which incorporate the lessons from the GPE output into a faster computational framework capable of addressing larger, more realistic systems.

In this paper, we present a mean-field vortex avalanche model, in which the angular velocity lag between the pulsar’s superfluid interior and rigid crust fluctuates according to a state-dependent Poisson process. The model does not track vortex motion and knock-on; it parametrises the avalanche physics using published GPE results (Warszawski & Melatos 2011). The computational savings are re-invested into longer runs (to generate a larger sample of events), an extended dynamic range (to capture the scale invariance of the size distribution), and a systematic survey of the range of glitch behaviour as a function of key astrophysical control parameters (e.g. spin-down torque), all of which are out of reach with contemporary GPE simulations. The paper is structured as follows. In Section 2, we summarise measurements of glitch statistics and recent theoretical work on vortex avalanches. In Section 3, we present and motivate physically the equation of motion governing the mean-field model. In Sections 4 and 5, we compute the glitch statistics arising from the model in the short and long term, shorter than and comparable to the spin-down time-scale respectively. In Section 6, we interpret the numerical results in Sections 4 and 5 analytically in terms of a master equation for the system.

## 2 VORTEX AVALANCHES

We begin by reviewing briefly the indirect evidence that neutron star glitches are caused by superfluid vortex avalanches.

<sup>2</sup> Temporal correlations (“memory”) play a similar role, e.g., in the coherent noise model (Melatos & Warszawski 2009).

The evidence is drawn from observations (statistical data; Section 2.1) and theory (GPE simulations; Section 2.2).

### 2.1 Observed glitch statistics

Pulsar glitch statistics are limited by sample size. Seven pulsars have been observed to glitch  $\geq 10$  times; four have been observed to glitch  $\geq 20$  times. The most prolific, PSR J1740–3015, has glitched 33 times. Notwithstanding these relatively small numbers, the probability density function (PDF) of glitch sizes in individual pulsars is consistent with a power law in three objects that have glitched more than ten times: PSR J0534–2200, PSR J1331–6220 and PSR J1740–3015. Glitch sizes span up to four decades in individual objects, e.g.,  $10^{-10} \lesssim \Delta\nu/\nu \lesssim 10^{-6}$  in PSR J0534+2200, where  $\Delta\nu$  is the step increment in the spin frequency  $\nu$ . Across the whole population, the smallest observed glitch has  $\Delta\nu/\nu = 8 \times 10^{-11}$  (PSR J1824–2452), and the largest has  $\Delta\nu/\nu = 6.5 \times 10^{-5}$  (PSR J1647–4552). Scale invariant event sizes are a feature of avalanche processes in far-from-equilibrium systems. In self-organised critical systems, for example, a global driver adds stress slowly to the system, and the stress is released spasmodically in local stick-slip interactions (Bak et al. 1987; Jensen 1998). Magnetic flux tube cascades in type-II superconductors epitomize this behaviour (Field et al. 1995; Altschuler et al. 2004).

In pulsars whose glitch size PDFs resemble power laws, the waiting-time PDF is approximately exponential, with  $p(\Delta t) \propto \exp(-\bar{\lambda}\Delta t)$ , where  $\bar{\lambda}$  is the average glitch rate. The relationship between  $\bar{\lambda}$  and a pulsar’s spin-down age  $\tau_c$  remains unclear. Melatos et al. (2008) analysed data from seven pulsars with  $0.35 \text{ yr}^{-1} \leq \bar{\lambda} \leq 2.6 \text{ yr}^{-1}$  and found no correlation between  $\bar{\lambda}$  and  $\tau_c$  (see Figure 10 of the latter paper). Espinoza et al. (2011) found that  $\bar{\lambda}$  for the most frequent glitchers decreases monotonically with age (see Figure 9 of the latter paper). McKenna & Lyne (1990) first proposed the glitch activity parameter, which is the fractional increase in  $\nu$  due to glitches per annum and therefore includes information about both  $\bar{\lambda}$  and sizes. They found that the most active pulsars are middle-aged, with  $10^4 \lesssim \tau_c/(1 \text{ yr}) \lesssim 10^5$ .

Three objects do not exhibit the scale-invariant behaviour described above. In Vela (PSR J0835–4510) and PSR J0537–6910, the size and waiting-time PDFs are unimodal, i.e., peaked at characteristic scales (Middleditch et al. 2006; Melatos et al. 2008). PSR J1341–6220 displays signs of intermediate behaviour between the unimodal and scale-invariant cases; its classification will be clarified, as more data become available (Howitt et al. 2017). A unimodal waiting-time PDF corresponds to quasi-periodic behaviour, which is sometimes interpreted in terms of a global threshold mechanism: each event empties fully the global stress “reservoir,” which then replenishes fully over a characteristic time-scale, before the next event occurs. Self-organised critical systems can generate quasi-periodic events of roughly equal sizes in the regime where they are rapidly driven (Jensen 1998; Melatos et al. 2008). However, a global threshold implies a correlation between the size of a glitch and the waiting time until the next glitch. Such a correlation is observed in PSR J0537–6910 (Middleditch et al. 2006) but not in PSR J0834–4511 or PSR J1341–6220. Moreover, no correlation

between the size of a glitch and the time since the previous glitch is observed in any object (Peralta 2006). The reason for the widespread lack of correlation remains unknown.

The minimum observed size approaches the resolution  $\Delta\nu/\nu \sim 10^{-11}$  inferred from detection simulations (Janssen & Stappers 2006). Therefore, it is unclear whether the observed minimum is fundamental or an observational artifact. Espinoza et al. (2014) reviewed 29 years of data for the Crab and found that the number of events with  $\Delta\nu \lesssim 0.05 \mu\text{Hz}$  is significantly less than expected from extrapolating downwards a power-law PDF. Whether or not this implies a fundamental minimum size in the Crab, however, depends on the form of the PDF, which remains uncertain because of the relatively small number of events. It will be interesting to search for a shortfall in events with  $\Delta\nu \lesssim 0.05 \mu\text{Hz}$  in other objects, as more data become available.

## 2.2 Gross–Pitaevskii simulations

When interpreted in terms of vortex unpinning, the data described in Section 2.1 imply that vortex motion must be *correlated* during avalanches. Consider the standard, idealised model in which a glitch occurs when  $N_u$  (out of a total of  $N_v$ ) vortices unpin and move radially outward a distance  $\Delta r \approx \xi \lambda_F$ , which is comparable to the Feynmann distance,  $\lambda_F = 1.3 \times 10^{-2} (\nu/1 \text{ Hz})^{-1/2} \text{ cm}$ , i.e., we have  $\xi \approx 1$ .<sup>3</sup> Assuming angular momentum conservation between the rigid crust (moment of inertia  $I_c$ ) and the superfluid (nominal moment of inertia  $I_s$ ), one obtains

$$\frac{\Delta\nu}{\nu} = \frac{I_s N_u \lambda_F \xi}{I_c N_v R_*}, \quad (1)$$

where  $R_*$  is the stellar radius. Hence even the smallest glitches involve  $N_u = 1.6 \times 10^{13} (\Delta\nu/10^{-11}\nu) (I_s/I_c)^{-1} \xi^{-1} (\nu/1 \text{ Hz})^{3/2}$  vortices. Uncorrelated motion of so many vortices leads, from the central limit theorem, to narrow size and waiting-time PDFs with dispersions  $N_u^{-1/2}$  times the mean, which are nothing like what is observed.

The knock-on mechanism triggering correlated motion remained uncertain until the advent of GPE simulations recently (Warszawski et al. 2012). The GPE is a time-dependent, nonlinear Schrodinger equation for  $\psi$ , the order parameter of the neutron condensate. For systems in contact with a thermal bath, it takes the phenomenological form (Gardiner et al. 2002)

$$(i - \gamma)\psi t = -\frac{1}{2}\nabla^2\psi + (V + |\psi|^2)\psi - \Omega_c \hat{L}_z \psi + i\gamma\mu\psi. \quad (2)$$

Equation (2) is written in dimensionless form; the units are defined below equation (1) in Melatos et al. (2015) and are not relevant here. The term involving the coupling coefficient  $\gamma$  on the left-hand side represents coupling to thermal excitations. On the right-hand side,  $V$  denotes the sum of

the trap and pinning potentials,  $\mu$  is the chemical potential,  $\hat{L}_z$  is the angular momentum operator projected along the rotation ( $z$ -)axis, and  $|\psi|^2\psi$  describes the bosonic self-attraction. Detailed accounts of the physics underlying the above model, its limitations and how it is solved, are provided in Warszawski & Melatos (2011).

In the neutron-star context, GPE simulations are idealised: they treat the neutron superfluid as a dilute, zero-temperature, Bose-Einstein condensate in a decelerating trap containing a regular lattice of localized pinning sites, even though realistic neutron superfluids are not dilute and generally contain multiple neutral and charged components coupled by mutual friction and entrainment (Andersson & Comer 2006). Nonetheless, the simulations display clearly the operation of two knock-on mechanisms: proximity knock-on, in which an unpinned vortex moves past its pinned neighbours and unpins them by mutual repulsion (Magnus force); and acoustic knock-on, where sound waves emitted by an unpinned vortex, as it moves, trigger further unpinning at a distance. For the largest simulations involving  $\approx 2 \times 10^2$  vortices, the size PDF approaches a power law over four decades,  $10^{-2} \lesssim \Delta\nu/\nu \lesssim 10^{-6}$  [see Figure 12 in Warszawski & Melatos (2011)], with index between  $-1.0$  and  $-1.25$ , depending on the pinning strength, pinning site density and spin-down rate (Warszawski & Melatos 2011). The waiting-time PDF is fitted well by an exponential curve. The mean waiting time decreases monotonically, as the spin-down torque increases, while the mean jump size varies in a complicated non-monotonic way as the torque is adjusted [see Tables 7 and 9 in Warszawski & Melatos (2011)]. The simulations do not exhibit a global ‘reservoir effect,’ i.e., no correlation between size and waiting time is observed, consistent with other data.

Despite the above successes, the GPE simulations are limited in their applicability, because their dynamic range is much smaller than in a realistic pulsar. The simulations typically contain  $\lesssim 10^2$  vortices, compared to  $\sim 10^{18}$  in a pulsar. The simulation typically runs for about one spin-down time-scale, during which it generates  $\lesssim 50$  glitches; in contrast, a pulsar generates  $\gtrsim 10^3$  glitches per spin-down time-scale. In a pulsar, up to  $\sim 10^{10}$  pinning sites separate neighbouring vortices, compared to  $\sim 10$  in the simulations. The latter property especially raises a fundamental question in a neutron star: does an unpinned vortex travel far enough to unpin another vortex via proximity knock-on, or does it repin before it gets the chance? If re-pinning forestalls knock-on, superfluid vorticity is expelled gradually by vortex creep (Alpar et al. 1989; Link & Epstein 1991; Link et al. 1993; Link 2014), instead of impulsively via avalanches. Recently, single-vortex calculations have confirmed that a vortex typically travels one inter-vortex spacing ( $\approx \lambda_F$ ) before re-pinning for a range of plausible neutron star conditions, determined typically by the mutual friction coefficient, and how close the system approaches the critical unpinning threshold (Haskell & Melatos 2016). This is consistent with the GPE output, where the maximum glitch size is independent of  $n_{\text{pin}}/n_v$  for  $n_{\text{pin}} \gtrsim n_v$ , where  $n_v \approx \lambda_F^{-2}$  is the vortex density, and  $n_{\text{pin}}$  is the pinning site density. The mean waiting time decreases inversely with  $n_v/n_{\text{pin}}$ , while the mean size does not change monotonically, suggesting that the typical distance travelled by a vortex in a glitch is a complex

<sup>3</sup> The expression for  $\lambda_F$  follows from the quantisation of circulation ( $\oint \mathbf{v} \cdot d\mathbf{l}$ , where  $\mathbf{v}$  is the superfluid velocity, and  $d\mathbf{l}$  is a line element), and the minimisation of free energy, which induces vortices to arrange themselves quasi-equidistantly in a uniform Abrikosov lattice.

function of  $N_u$  and  $n_{\text{pin}}$ , even though it is always comparable to  $\lambda_F$  (Warszawski & Melatos 2011).

### 3 STATE-DEPENDENT POISSON PROCESS

In this section we formulate an idealised model of glitch activity as a stochastic process caused by vortex avalanches, which embodies the key elements of the astronomical observations and GPE simulations summarised in Section 2. We write down a stochastic equation of motion for the fluctuating crust-superfluid lag in Section 3.1, and propose physically motivated PDFs for the unobservable sizes (power law; Section 3.2) and waiting times (Poisson; Section 3.3) of the vortex avalanches driving the lag. A Monte-Carlo algorithm for evolving the model numerically in dimensionless variables is presented in Sections 3.4 and 3.5. It is important to recognise that the vortex avalanche statistics prescribed in Sections 3.2 and 3.3 are *not* the same as the observable statistics of the glitch sizes and waiting times predicted by the model. Specifically, the vortex avalanches are modelled as having power-law sizes and Poissonian waiting times, broadly consistent with the GPE results. But, because the Poisson rate is state-dependent, and the crust-superfluid lag self-regulates by reverting to the mean (and to a stationary PDF), the observable glitch activity may look nothing like the underlying vortex avalanches. Examples of the diverse variety of possible observable behaviours, derived from numerical simulations, are presented in Section 4.

#### 3.1 Equation of motion

Let  $X$  be a stochastic variable equal to the globally-averaged lag between the angular speeds of the crust,  $\Omega_c$ , and the superfluid interior,  $\Omega_s \geq \Omega_c$ , viz,

$$X(t) = \Omega_s(t) - \Omega_c(t). \quad (3)$$

In reality  $\Omega_s$  varies throughout the star, according to the local density and configuration of superfluid vortices. By averaging  $\Omega_s$  spatially over the stellar volume, we make a mean-field approximation, which does not explicitly track cooperative mechanisms like vortex knock-on (Warszawski & Melatos 2011; Warszawski et al. 2012). The approximation is justified by GPE simulations, which show that the vortices self-adjust to maintain a roughly uniform Abrisokov lattice at all times, even when the pinning potentials are of random strengths (Warszawski & Melatos 2013). The scale-invariance of the vortex avalanches, which depends critically on spatial correlations and knock-on, is captured *implicitly* but incompletely by the choice of jump distribution (e.g. power-law) in Section 3.2.

The rigid crust, which is strongly coupled to the star's magnetic field, spins down deterministically due to magnetic braking, interrupted by instantaneous spin-up events  $\Delta\Omega_c^{(i)}$ . The approximation that glitches occur instantaneously is accurate in an astrophysical context. While the typical pulsar spin-down time-scale is  $\tau_c \gtrsim 10^3$  yr, individual glitches occur over  $\leq 30$  seconds; their rise times are normally un-

resolved by current measurements (McCulloch et al. 1990; Epstein & Baym 1992; Dodson et al. 2007). We write

$$\Omega_c(t) = \Omega_c(0) - \frac{N_{\text{em}}t}{I_c} + \sum_{i=1}^{N(t)} \Delta\Omega_c^{(i)}, \quad (4)$$

where  $N_{\text{em}}$  is the magnetic torque,  $I_c$  is the crust's moment of inertia, and  $N(t)$  is the number of glitches that occur up to time  $t$ . We discuss  $N(t)$  further in Section 3.3. For now, we simply note that  $N(t)$  is an integer-valued random variable arising from a counting process.

The angular speed of the superfluid, which is weakly coupled to the crust, remains constant until a glitch occurs, whereupon it spins down instantaneously by an amount  $\Delta\Omega_s^{(i)}$ :

$$\Omega_s(t) = \Omega_s(0) + \sum_{i=1}^{N(t)} \Delta\Omega_s^{(i)}. \quad (5)$$

Equation (5) neglects vortex creep, i.e. the gradual decrease of  $\Omega_s$  between glitches due to unpinning by thermal activation and quantum tunnelling. As a first approximation, creep modifies  $\dot{\Omega}_s$  ( $\neq 0$ ) and hence  $\dot{X}$  between glitches. Hence it can be incorporated by reducing the effective torque  $N_{\text{em}}$ . The reduction factor depends on the mean rate of angular momentum exchange; its evaluation lies outside the scope of the present work. One can also add an additional white-noise Langevin force to equation (5) to model the fluctuation in the vortex creep rate about its mean value (Daly & Porporato 2006). A Langevin analysis of this kind is difficult and likewise lies outside the scope of this paper.

The dynamic feedback between the crust and superfluid is governed by angular momentum conservation, viz.

$$I_s \Delta\Omega_s^{(i)} = -I_c \Delta\Omega_c^{(i)}. \quad (6)$$

Upon combining (3)–(6), we obtain

$$X(t) = X(0) + \frac{N_{\text{em}}t}{I_c} - \sum_{i=1}^{N(t)} \Delta X^{(i)}, \quad (7)$$

where we absorb the moment-of-inertia factors into  $\Delta X^{(i)}$  according to the equation

$$\Delta X^{(i)} = \frac{(I_c + I_s)\Delta\Omega_s^{(i)}}{I_c}. \quad (8)$$

The classic vortex unpinning model of radio pulsar glitches in the literature is epitomised by (7).

The first term in (7) is an astrophysically irrelevant initial condition, the second is deterministic, and the third is stochastic. In reality the braking torque decreases as  $N_{\text{em}} \propto \Omega_c^3$ , but glitch waiting times typically satisfy  $\lesssim 10^{-3}\tau_c$ , and individual glitches change  $\Omega_c$  by a factor of  $10^{-4}$  at most (see Section 2), so the approximation  $N_{\text{em}} = \text{constant}$  is accurate. In Section 5, we investigate the implications of the quasistatic decrease in  $N_{\text{em}}$  during a pulsar's lifetime. <sup>4</sup>

<sup>4</sup> Equation (7), with  $N_{\text{em}} = \text{constant}$ , is a special case of a broader class of stochastic processes, whose instantaneous *deter-*

Equations (4)–(7) assume that angular momentum is exchanged exclusively between the crust and superfluid. This assumption has been challenged recently by some theoretical calculations of the effective moment of inertia of the neutron superfluid, which predict that the angular momentum reservoir is too small to account for the observed glitches (Andersson et al. 2012), although other calculations have concluded that the shortfall only arises for certain equations of state (Piekarewicz et al. 2014).

### 3.2 Avalanche size

Let  $\eta(x|y) dx$  be the conditional probability of jumping to  $x \leq X \leq x + dx$ , given that the system is at  $X = y$ , when a glitch occurs. The form of  $\eta$  is determined by the internal avalanche physics, i.e., the collective process whereby vortices knock on and unpin. In GPE simulations of vortex avalanches and studies of other self-organised critical systems, including sand-piles and earthquakes, the duration of an avalanche is much shorter than the time-scale over which stress builds up, so it is natural that  $\eta(x|y)$  should be independent of other control parameters like the spin-down rate.

We impose two physical restrictions on  $\eta(x|y)$  as follows. Every glitch reduces  $X$ , so we require  $\eta(x|y) = 0$  for  $x > y$ . We also assume that no glitch can ever cause  $X$  to be negative,<sup>5</sup> implying  $\eta(x|y) = 0$  for  $x < 0$ . These two restrictions yield the fundamental normalisation condition for  $\eta$ ,

$$1 = \int_0^y dx \eta(x|y). \quad (9)$$

GPE results indicate that the PDF of the jumps  $\Delta X^{(i)}$  is consistent with a power law, independent of  $y$ , except near zero lag, with

$$\eta(x|y) \propto (y - x)^{-\delta}, \quad (10)$$

and  $\delta \approx 1.5$ . Near zero lag, however, a lower cut-off on  $\Delta X^{(i)}$  is required in order to ensure  $X \geq 0$  and keep the power law from diverging in (9). The cut-off scales with  $y$  in the fashion explained in Section 4.

The PDF of  $\Delta X^{(i)}$  is, in general, conditionally dependent on the value of  $X$  immediately before the  $i$ -th glitch occurs. Hence  $\Delta X^{(i)}$ ,  $\Delta X^{(i+1)}$ ,  $\Delta X^{(i+2)}$ ... are not identically distributed, even after the system reaches stationarity. Moreover, it is necessary to distinguish clearly between  $\eta$ , an input into the model determined by the avalanche physics, and the PDF of  $\Delta\Omega_c$ , an output of the model determined jointly by  $\eta$  and  $X$ , which develops as the system evolves. Throughout the rest of this paper, we refer to  $\eta$  as the conditional jump distribution, and to  $p(\Delta\Omega_c)$  as the PDF of the observed glitch sizes.

*ministic* evolution depends on the current system state. In Appendix E we include this generalisation for completeness.

<sup>5</sup> Conceivably, one can imagine a non-mean-field scenario, where a large amount of lag builds up in a small region of the pulsar, and a large avalanche occurs which temporarily leaves  $X$  negative. However, such a scenario is highly unlikely for  $N_v > N_u \gg 1$ .

### 3.3 Avalanche rate

The probability of an avalanche occurring in an infinitesimal time interval depends on the crust-superfluid lag at that instant. We model avalanche triggering as a variable-rate Poisson process with instantaneous rate  $\lambda(X)$ . As it is a function of the random variable  $X$ ,  $\lambda$  itself is a random variable. The assumption that  $X$  alone determines  $\lambda$  involves a significant mean-field simplification. In reality, the distribution of vortices in local patches, not the spatially averaged lag, controls the likelihood of an avalanche being triggered locally. Avalanches are triggered at a single pinning site (and then propagate by knock-on), because the avalanche duration is much less than the waiting time between avalanches, as GPE simulations demonstrate (Warszawski & Melatos 2013).

The rate  $\lambda(X)$  is a monotonically increasing function of  $X$  up to a critical lag  $X_c$ , at which an avalanche is certain to occur, because the Magnus force on all the vortices exceeds the pinning force. In this paper we model this behaviour phenomenologically with the rate function

$$\lambda(X) = \lambda_0 (1 - X/X_c)^{-1}, \quad (11)$$

where  $\lambda_0$  is related to the avalanche rate at some reference lag, e.g.,  $\lambda_0 = \lambda(0)$ . The value of  $X_c$  is a complex function of the pinning site density and potential, which remain approximately constant from one glitch to the next, as well as the superfluid density and vortex alignment. Link & Epstein (1991) expressed the critical, or breakaway, lag at a single pinning site as

$$\begin{aligned} X_c &= \frac{F_{\max}}{\rho_s \kappa l R_*} \\ &= 8 \times 10^{-2} \\ &\times \left( \frac{F_{\max}}{\text{keV fm}^{-1}} \right) \left( \frac{\rho_s}{10^{13} \text{ gcm}^{-3}} \right)^{-1} \left( \frac{l}{10^2 \text{ fm}} \right)^{-1} \text{ rad s}^{-1}, \end{aligned} \quad (12)$$

where  $F_{\max}$  is the maximum pinning force per site,  $\rho_s$  is the superfluid mass density,  $\kappa$  is the quantum of circulation, and  $l$  is the pinning-site separation; see Tables 1 and 2 in Link & Epstein (1991). We stress that a lot of complicated and uncertain physics goes into  $F_{\max}$  in (12), including the form of the nuclear pinning potential (Alpar et al. 1989; Donati & Pizzochero 2006; Avogadro et al. 2008; Seveso et al. 2016), the role of vortex tension and tangled vorticity in regulating single- versus multi-site breakaway (Link & Epstein 1991), and collective vortex-vortex interactions, which are pivotal to the propagation of avalanches, as GPE simulations demonstrate (Warszawski et al. 2012).

The rate function  $\lambda(X)$  influences the inter-glitch waiting-time PDF in the following way. For a Poisson process whose rate function  $\lambda(t)$  evolves *deterministically* as a function of time, the waiting time  $\Delta t$  until the next event, given that we begin observing the system at time  $t$ , has PDF  $p(\Delta t; t)$  given by (Ross 2006)

$$p(\Delta t; t) = \lambda(t + \Delta t) \exp \left[ - \int_t^{t+\Delta t} dt' \lambda(t') \right]. \quad (13)$$

In equation (7),  $X(t)$  and therefore also  $\lambda[X(t)]$  evolve de-

terministically between glitches. If a glitch occurs at time  $t$  and sets the lag to  $X(t)$ , then until the next glitch we have  $X(t') = X(t) + N_{\text{em}}(t' - t)/I_c$  for  $t \leq t' \leq t + \Delta t$ . Hence, given the value of  $X(t)$  immediately after the previous glitch, the PDF of the waiting time until the next glitch is

$$p[\Delta t|X(t)] = \lambda \left[ X(t) + \frac{N_{\text{em}}\Delta t}{I_c} \right] \exp \left\{ - \int_t^{t+\Delta t} dt' \lambda[X(t')] \right\}. \quad (14)$$

We find *a posteriori* that the results depend weakly on the functional form of  $\lambda$ . Its key features are: (i) it increases monotonically with  $X$ ; and (ii) it guarantees unpinning at the critical lag with  $\lambda(X) \rightarrow \infty$  as  $X \rightarrow X_c$ . Equation (11) allows for a non-zero avalanche rate at zero lag due to quantum-mechanical tunnelling and/or thermal activation (Link et al. 1993), but this is not a necessary feature of the model. In Section 4 we find that this feature has observable consequences under a limited set of circumstances. Equation (11) qualitatively resembles the phenomenological Arrhenius law used widely in previous glitch studies (Hänggi et al. 1990; Chevalier 1993; Warszawski & Melatos 2013), except that the Arrhenius rate remains bounded even as it becomes very large for  $X \rightarrow X_c$ .<sup>6</sup>

Cox (1955) first introduced the doubly-stochastic Poisson process, in which the rate function  $\lambda$  is a random variable. Cox processes appear in many contexts, for example in the pricing of financial derivatives and insurance risk (Basu 1999; Dassios & Jang 2003), where random, catastrophic events drastically alter the underlying instantaneous rate of an observable counting process, as well as in models of rainfall (Ramesh et al. 2012). However, the state-dependent Poisson process (7) differs from a Cox process in one fundamental respect. In a Cox process, the rate variable  $\lambda$  evolves randomly according to an external “environment” and is independent of  $N(t)$ . In our model, the rate is a function of the occurrence times and sizes of the previous events, introducing a complex feedback mechanism. An equivalent state-dependent Poisson formalism has been explored as a model for solar flare statistics (Wheatland & Glukhov 1998; Wheatland 2008, 2009; Kanazir & Wheatland 2010). Daly & Porporato (2006) modelled forest fires as a state-dependent Poisson process, where every jump fully resets the “stress” (unburnt fuel) to zero. They developed a formalism for calculating the waiting-time PDF for general  $\eta(x|y)$  (Daly & Porporato 2007), an issue that we pick up in Section 6.

<sup>6</sup> To check the sensitivity of the results for the functional form of  $\lambda(X)$ , we repeat a subset of the simulations in Section 4 for the rate function  $\lambda(X) = 2\lambda_0 \tan(\pi X/2X_c)$ , which increases monotonically and diverges at  $X = X_c$ , just like  $\lambda(X) = \lambda_0/(1 - X/X_c)$ . We find that the key PDFs for the waiting times and glitch sizes are broadly unaffected: their shapes and the variation of their moments and modes as functions of  $I_c X_c \lambda_0/N_{\text{em}}$  are similar to what follows from (11). Some differences in detail are observed, which arise mainly because  $\lim_{X \rightarrow 0} \lambda(X)$  is different for the two rate functions, not because the tangent and hyperbola have different shapes.

### 3.4 Dimensionless variables

The maximum lag  $X_c$  defines a natural scale. We replace  $X$  by its dimensionless counterpart  $\tilde{X} = X/X_c$ , with  $0 \leq \tilde{X} \leq 1$ . We also introduce a dimensionless time variable,  $\tilde{t} = N_{\text{em}}t/(X_c I_c)$ . Physically,  $\Delta\tilde{t} = 1$  corresponds to the time required for the electromagnetic spin down to bring the system from zero lag to the critical lag. With the above replacements, the dimensionless equation of motion reads

$$\tilde{X}(\tilde{t}) = \tilde{X}(0) + \tilde{t} - \sum_{i=0}^{N(\tilde{t})} \Delta\tilde{X}^{(i)}. \quad (15)$$

Throughout the rest of the paper, we drop the tildes for clarity of presentation.

The phenomenological rate function (11) takes the dimensionless form

$$\lambda(x) = \frac{\alpha}{1-x} \quad (16)$$

when normalised as above. This introduces a key control parameter of the model:

$$\alpha = \frac{I_c X_c \lambda_0}{N_{\text{em}}}. \quad (17)$$

The value of  $\alpha$  in any given pulsar is a complex function of the pinning physics, as discussed in Sections 2.2 and 3.3, and the spin-down rate.

### 3.5 Monte-Carlo automaton

Equation (15) can be simulated numerically by executing the following simple, Monte-Carlo automaton to update the state variable  $X(t)$  at time  $t$ .

1. Draw a random  $\Delta t$ , the waiting time to the next glitch, using (14).
2. Update and record the lag,  $X(t) + \Delta t$ , after the crust spins down deterministically for time  $\Delta t$ .
3. Draw a random avalanche size  $\Delta X$  from the conditional jump distribution  $\eta$ .
4. Update and record the new system state,  $X(t + \Delta t) = X(t) + \Delta t - \Delta X$ .
5. Return to step 1 and repeat.

Random variables are drawn using the standard inverse cumulative distribution method (Press et al. 2007). We collect and record the values of  $\Delta t$ ,  $\Delta X$  (and hence  $\Delta\Omega_c$ ) for post-processing and statistical analysis.

## 4 SHORT-TERM DYNAMICS

In this section, we compute the glitch statistics generated by the model in Section 3 on time-scales that are short compared to the spin-down time-scale  $\tau_c$ , so that the system’s control parameters remain constant to a good approximation. For the calculations in this section, we work with a conditional jump distribution of the form

$$\eta(x|y) = \left[ \int_0^y d\xi g(\xi, y) \right]^{-1} g(y - x, y), \quad (18)$$

which automatically satisfies (9). Motivated by (10), we choose  $g$  to be

$$g(\xi, y) = \xi^{-1.5} H(\xi - \beta y). \quad (19)$$

Equation (19) is not derived from the avalanche microphysics or from observational data. It is a trial function which captures the main results of the GPE simulations (see Section 2.2). The Heaviside function  $H(y - x - \beta y)$  kills the divergence of the inverse power-law conditional jump distribution in the limit  $x \rightarrow y$ . The dimensionless scaling constant  $\beta$ , with  $0 < \beta < 1$ , determines the minimum avalanche size. If we have  $\beta = 0.1$ , for example, the minimum size is 0.1 times the crust-superfluid lag immediately before the glitch occurs. A smaller value of  $\beta$  produces a wider range of possible avalanche sizes. In the limit  $\beta \rightarrow 1$ , the minimum avalanche size approaches the maximum avalanche size, which always equals the pre-glitch lag, so  $\eta(x|y)$  approaches the delta function  $\delta(x - y)$ , and every jump restores the lag to zero.

#### 4.1 Representative time series

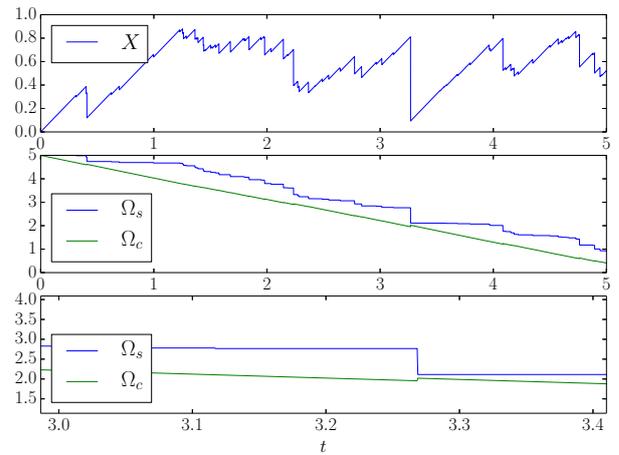
Figure 1 displays how the crust-superfluid lag  $X(t)$  evolves according to (15) for a single, typical realisation of random avalanches for the illustrative case  $\alpha = 5$ ,  $\beta = 10^{-2}$ .

The top panel of Figure 1 exemplifies how  $X(t)$  approaches stationarity. Few glitches occur for  $t \lesssim 1$ , as  $\lambda(X)$  is small far from the asymptote at  $X = 1$ . This initial transient is irrelevant astrophysically. At  $t \gtrsim 1$  the system reaches a stationary state, where the avalanche rate and sizes self-adjust to balance the secular spin down on average. The lag fluctuates around a mean value of  $\langle X \rangle \approx 0.72$ . Each glitch causes a vertical drop in  $X(t)$ . The deterministic spin down between glitches has a gradient of  $X'(t) = 1$  as in (15). The lower two panels show how  $\Omega_s$  and  $\Omega_c$  evolve in the same simulation; the bottom panel zooms in to magnify the interval  $2.2 \leq t \leq 3$ . For the superfluid,  $\Omega_s(t)$  traces a piece-wise constant curve (i.e. pinned vortices), interrupted by vertical decrements whenever a glitch occurs. For the crust,  $\Omega_c(t)$  decreases steadily, interrupted by instantaneous spin-up events which conserve angular momentum according to (6). We always have  $\Omega_s(t) > \Omega_c(t)$  as discussed in Section 3.2. The crust and superfluid jumps are related through (6) by  $\Delta\Omega_c = -I_s \Delta X / (I_s + I_c)$ , with  $I_c = 10I_s$  in our illustrative example. We note that, of the three quantities plotted in Figure 1, only  $\Omega_c$  is observable astronomically.

#### 4.2 Size and waiting-time distributions

The steady-state PDFs of the observed glitch sizes,  $p(\Delta\Omega_c)$ , and waiting times,  $p(\Delta t)$ , are both astronomically observable quantities measured during a typical pulsar monitoring and timing campaign (Melatos et al. 2008; Espinoza et al. 2011). To construct  $p(\Delta\Omega_c)$  and  $p(\Delta t)$  from the stochastic equation (15), we run a simulation like Figure 1 to produce  $N = 10^7$  glitches and begin sampling at  $t \gg 1$ , once the system reaches stationarity.

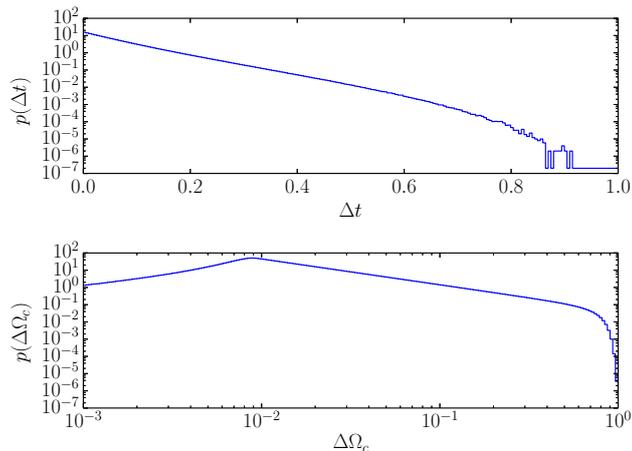
The top panel of Figure 2 displays  $p(\Delta t)$  on log-linear axes for the simulation in Figure 1 for  $6.6 \times 10^2 \leq t \leq 6.6 \times 10^5$ . The PDF is close to exponential for  $\Delta t \lesssim 0.6$ , with  $p(\Delta t) \propto \exp(-\bar{\lambda}\Delta t)$  and  $\bar{\lambda} = 13.8$ . The bottom panel



**Figure 1.** Representative evolution of the dimensionless crust-superfluid lag  $X(t) = \Omega_s - \Omega_c$  (top panel) and its dimensionless components  $\Omega_s(t)/X_c$  (blue curve) and  $\Omega_c(t)/X_c$  (green curve) for a single random realization of vortex avalanches. The bottom panel magnifies a particular large glitch at  $t \approx 3.27$ . Time (horizontal axis) is measured in units of  $X_c I_c N_{\text{em}}$ . Parameters:  $\alpha = 5$ ,  $\beta = 10^{-2}$ . Initial conditions:  $\Omega_c(0) = \Omega_s(0) = 5$ .

shows the PDF of glitch sizes  $p(\Delta\Omega_c)$  on a log-log scale. It is a power-law over a sizeable portion of the domain  $10^{-2} < \Delta\Omega_c < 1$ , with index 1.5, equal to the index of the conditional jump distribution  $\eta$ . The Heaviside cut-off in  $\eta$  is blurred out in  $p(\Delta\Omega_c)$ , because  $\eta$  is convolved with the spread in  $X$  immediately before a glitch occurs. Both  $p(\Delta t)$  and  $p(\Delta\Omega_c)$  roll over for large  $\Delta t$  and  $\Delta\Omega_c$ , because the hard wall at  $X = 1$  in  $\lambda(X)$  places upper bounds  $\Delta X < 1$  and  $\Delta t < 1$ . The mean inter-glitch waiting time is  $\langle \Delta t \rangle \approx 0.0667 \approx \bar{\lambda}^{-1}$ . The statistics in Figure 3 are qualitatively consistent with pulsars like PSR J0534–2200, PSR J1331–6220 and PSR J1740–3015, which exhibit exponential waiting times and power-law sizes (see Section 2.1). Note that  $\Omega_c$  decreases by  $\lesssim 1\%$  over a typical timing campaign lasting  $\sim 40$  yr, so the PDFs of  $\Delta\Omega_c$  and  $\Delta\Omega_c/\Omega_c$  have the same shape to a good approximation.

We intentionally sample a single simulation for many glitches to produce Figure 2 instead of running an ensemble of simulations and sampling at a fixed time  $t$ . There is an observable difference between these two approaches, so it is important to clearly state what sampling procedure is used. We discuss this subtle point more fully in Section 6. If we sample an ensemble at fixed  $t_0$ , taking  $\Delta t$  to be the length of the time interval we “land in” and  $\Delta\Omega_c$  to be the size of the next glitch, then our distribution is implicitly biased. The interval  $T^{(i)} < t_0 < T^{(i+1)}$ , where  $T^{(i)}$  is the epoch of the glitch immediately preceding  $t_0$ , is longer on average than if we sample every interval  $\Delta t$  in a single simulation, because we are more likely to “land in” a longer inter-glitch interval than a shorter one. This is a manifestation of the inspection (“bus-stop”) paradox in renewal theory (Cox 1970). For similar but less obvious reasons,  $p(\Delta\Omega_c)$  is biased as well. Of the two methods, a single simulation rather than an ensemble is better suited to pulsar data, where we are interested in the PDFs of waiting times and sizes for individual pulsars over long time-scales (but still short compared to  $\tau_c$ ). An



**Figure 2.** Representative PDFs of inter-glitch waiting times  $p(\Delta t)$  (top panel) and glitch sizes  $p(\Delta\Omega_c)$  (bottom panel) for the single simulation in Figure 1 extended to  $N = 10^8$  glitches. Waiting times and sizes are measured in units of  $I_c X_c / N_{\text{em}}$  and  $X_c(I_c + I_s) / I_s$ , respectively. Parameters:  $\alpha = 5, \beta = 10^{-2}$ .

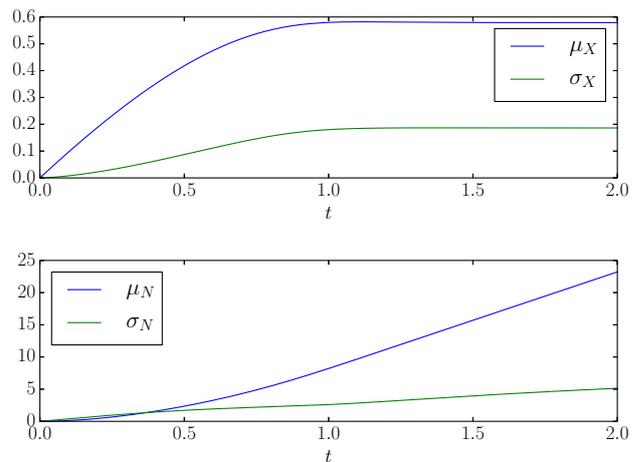
ensemble is better suited to compute the PDF of  $X$ , which evolves continuously. We do this in Section 4.3.

### 4.3 Moments

In order to compute the moments of  $X$  and  $N$  as functions of time, we carry out an ensemble of  $S = 10^6$  simulations like the one displayed in Figure 1. At some instant  $t$  we sample  $X$  and  $N$  from every simulation. Figure 3 shows the time-evolution of the ensemble mean  $\mu$  and standard deviation  $\sigma$  of  $X(t)$  and  $N(t)$  (the subscript indicates the random variable). The top panel clearly shows that  $\mu_X$  and  $\sigma_X$  increase steadily before plateauing at stationarity. At stationarity, the PDF  $p(t, X)$  does not change with time. In general, the system reaches stationarity for  $t \geq 1 - X(0)$ , whereupon we find  $\mu_N \propto t$  and  $\sigma_N \propto t^{1/2}$ , as for a constant-rate Poisson process. The proportionality constant for  $\mu_N(t)$ , by definition, equals  $\langle \Delta t \rangle^{-1}$ . In Figure 3, we find  $\langle \Delta t \rangle^{-1} = 15.0 = \langle \lambda \rangle$ , where  $\langle \lambda \rangle$  is computed by observing all values of  $\lambda(X)$  in the ensemble at time  $t$  and taking their average. In Appendix A, we prove analytically that the intuitive relationship  $\langle \Delta t \rangle^{-1} = \langle \lambda \rangle$  holds, so that the term “average glitch rate” can be used unambiguously to refer either to the ensemble mean of  $\lambda$  or to the inverse of the mean waiting time of a single simulation. We note that  $\lambda(\langle X \rangle) = 12.0 \neq \langle \lambda(X) \rangle = 15.0$ , i.e. the instantaneous avalanche rate at the average lag does not equal the average avalanche rate.

### 4.4 Spin-down torque

The control parameter  $\alpha$  is not observable, as it involves  $\lambda_0$  and  $X_c$ . We expect it to vary from one pulsar to the next in inverse proportion to the observable spin-down rate,  $\dot{\Omega}_c \approx N_{\text{em}} / I_c$ , and to a lesser extent as a function of the star’s density and temperature, which control the pinning and unpinning physics in  $X_c$  and  $\lambda_0$  respectively.



**Figure 3.** Mean and standard deviation of the crust-superfluid lag  $X(t)$  (top panel) and the number of glitches  $N(t)$  (bottom panel) during the approach towards equilibrium and beyond. The moments are computed from an ensemble of  $S = 10^6$  simulations. In the top panel,  $X$  is measured in units of  $X_c$ . In both panels,  $t$  is measured in units of  $I_c X_c / N_{\text{em}}$ .  $N$  is an integer. Parameters:  $\alpha = 5, \beta = 10^{-2}$ . Initial conditions:  $X(0) = 0$ .

#### 4.4.1 Transition from small to large $\alpha$

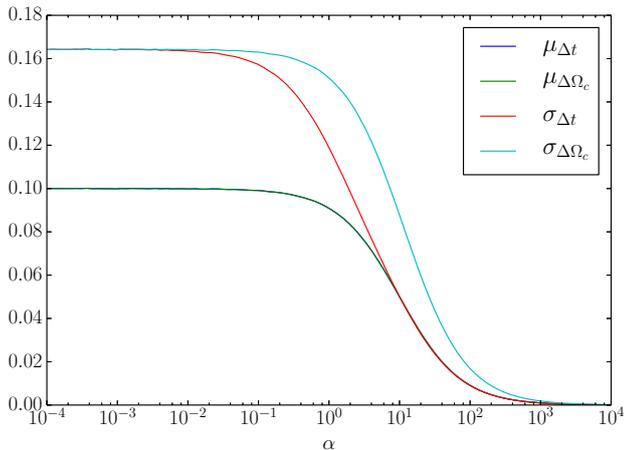
The parameter  $\alpha$  quantifies the ratio of the “natural” avalanche rate at a reference lag (i.e.  $\lambda_0$ ) to the electromagnetic spin-down rate (scaled by  $X_c$ ). We expect it to influence the size and waiting-time PDFs. If  $\alpha$  is small the system tends to be driven to  $X \approx 1$  before each glitch. If  $\alpha$  is large, avalanches occur at smaller  $X$ . This reduces the mean waiting time, because the system reaches smaller  $X$  more quickly. It reduces the glitch sizes because  $\Delta X$  is always less than  $X$  just before an avalanche.

Figure 4 shows how the means and standard deviations of the observed glitch sizes and waiting times depend on  $\alpha$  for  $\beta = 10^{-2}$ . The plot is drawn on log-linear axes. We see that  $\mu_{\Delta\Omega_c}$  (green curve) and  $\mu_{\Delta t}$  (dark blue curve) overlap for all  $\alpha$ , i.e. they are equal up to a multiplicative constant, with  $\mu_{\Delta\Omega_c} = N_{\text{em}} I_s \mu_{\Delta t} / [I_c(I_c + I_s)]$ , when the dimensions are temporarily restored. The equality of the means reflects the fact that  $\langle X \rangle$  is constant under stationary conditions, so the mean glitch rate times the mean glitch size must, on average, exactly balance  $N_{\text{em}} / I_c$  in (7). The standard deviations are not equal, however, except in the regime  $\alpha \ll 1$  (see Section 4.4.4); we see from Figure 2 that  $p(\Delta t)$  and  $p(\Delta\Omega_c)$  are not identical.

A qualitative change takes place at the intermediate value  $\alpha \approx \alpha_c(\beta) \sim 1$ , where  $\alpha_c$  depends on  $\beta$  as discussed below. For  $\alpha \lesssim \alpha_c$ , the mean and standard deviation asymptote to constant values. This corresponds to the system approaching  $X \approx 1$  immediately before every glitch. If one always has  $X \approx 1$  before a glitch, then the PDF of avalanche sizes is given by  $\eta(1 - \Delta X|1)$  from (10). Hence the mean glitch size is given by

$$\langle \Delta\Omega_c \rangle \approx \frac{X_c I_s}{I_s + I_c} \int_0^y d(\Delta X) \eta(1 - \Delta X|1) \quad (20)$$

$$\approx \frac{X_c I_s \beta^{1/2}}{I_s + I_c}, \quad (21)$$



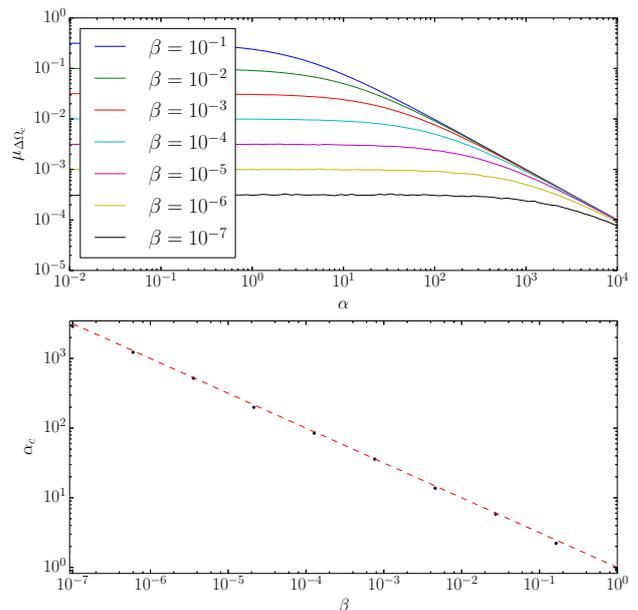
**Figure 4.** Mean and standard deviation of  $\Delta t$  (dark blue and red curves respectively) and  $\Delta\Omega_c$  (green and cyan curves respectively) as functions of  $\alpha$  ( $10^{-4} \leq \alpha \leq 10^4$ ) on log-linear axes. The means of both variables overlap over the whole domain. In order to maintain a common scale on the vertical axis,  $\Delta\Omega_c$  is measured in units of  $X_c(I_c + I_s)/I_s$ , and  $\Delta t$  is measured in units of  $I_c X_c/N_{\text{em}}$ . For each value of  $\alpha$ ,  $\Delta\Omega_c$  and  $\Delta t$  are sampled from a single simulation featuring  $N = 10^8$  glitches after the system reaches stationarity. Parameter:  $\beta = 10^{-2}$ .

from (18) and (19) and with dimensions temporarily restored.<sup>7</sup> This gives  $\langle \Delta\Omega_c \rangle \approx 0.1 X_c I_s / (I_c + I_s)$  for  $\beta = 10^{-2}$ , in agreement with Figure 4.

The limiting behaviour for large  $\alpha \gg \alpha_c(\beta)$  is best visualised on a log-log plot (not shown here). As  $\alpha$  increases, the mean and standard deviation of  $\Delta t$  and  $\Delta\Omega_c$  decrease as  $\alpha^{-1}$ . Physically, the instantaneous avalanche rate is large even at small  $X$  values. As  $\alpha$  increases, the system becomes more likely to glitch at smaller  $X$  values. As  $\lambda(X) \approx \alpha$  is approximately constant for  $X \ll 1$ , the average avalanche rate is approximately proportional to  $\alpha$ . The standard deviation of  $\Delta t$  approaches the mean in the limiting regime  $\alpha \gg \alpha_c$ , consistent with  $p(\Delta t)$  approaching an exponential curve.

#### 4.4.2 Transition versus minimum avalanche size

The value of  $\alpha$  at which the transition in Section 4.4.1 occurs depends on the minimum avalanche size, parametrised by  $\beta$ . In the limit  $\beta \rightarrow 1$ , every glitch resets  $X(t)$  to 0. Therefore smaller avalanches are allowed, as  $\beta$  decreases, given any particular pre-glitch value of  $X$ . Viewed another way, as  $\beta$  increases, smaller values of  $\alpha$  are required to drive the system to  $X \approx 1$  before every glitch. This is confirmed in the top panel of Figure 5, which graphs  $\mu_{\Delta\Omega_c}$  as a function of  $\alpha$  for seven values of  $\beta$ . Not only does the maximum value of  $\mu_{\Delta\Omega_c}$  (as  $\alpha \rightarrow 0$ ) increase as  $\beta^{1/2}$ , as required by (19), but so does the value of  $\alpha$  where the transition occurs. Let us arbitrarily define  $\alpha_c$  as the value of  $\alpha$  where  $\mu_{\Delta\Omega_c}$  reaches



**Figure 5.** (Top panel.) Mean glitch size  $\mu_{\Delta\Omega_c}$  as a function of  $\alpha$  for  $\beta = 10^{-1}$  (blue curve),  $10^{-2}$  (green curve),  $10^{-3}$  (red curve),  $10^{-4}$  (cyan curve),  $10^{-5}$  (magenta curve),  $10^{-6}$  (yellow curve) and  $10^{-7}$  (black curve). The vertical axis is scaled by  $X_c(I_s + I_c)/I_s$ . We define  $\alpha_c$  as the value of  $\alpha$  where  $\mu_{\Delta\Omega_c}(\alpha, \beta)$  (top panel) reaches half its maximum value, i.e.  $\mu_{\Delta\Omega_c}(\alpha, \beta) = \sqrt{\beta}/2$ . (Bottom panel.)  $\alpha_c$  versus  $\beta$ . In both panels,  $\mu_{\Delta\Omega_c}$  is obtained for each pair  $(\alpha, \beta)$  by sampling  $N = 10^6$  glitches after the system reaches stationarity.

half its maximum. In the bottom panel of Figure 5, we show the  $\beta$  dependence of  $\alpha_c$  explicitly. An empirical fit yields

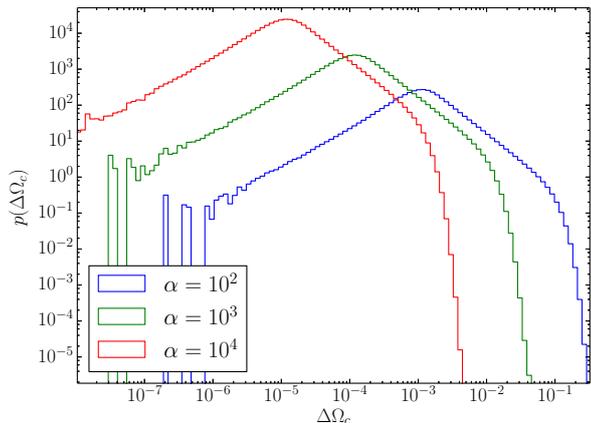
$$\alpha_c(\beta) \approx \beta^{-1/2}. \quad (22)$$

#### 4.4.3 PDFs for slow spin-down: $\alpha \gtrsim \alpha_c(\beta)$

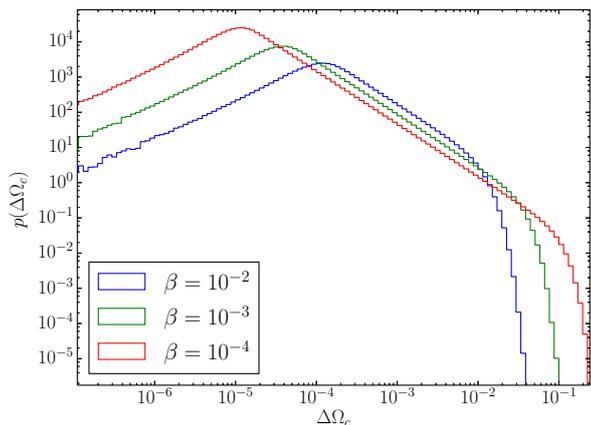
Let us look in more detail at the regime  $\alpha \gtrsim \alpha_c$ . The lag where glitches typically occur decreases, as  $\alpha$  increases, reducing  $\Delta\Omega_c$  and  $\Delta t$ . Figure 6 shows  $p(\Delta\Omega_c)$  for three  $\alpha$  values covering the range  $10^2 \leq \alpha \leq 10^4$ , with  $\beta = 10^{-2}$ , plotted on log-log axes. The data are sampled from a single simulation as in Figure 2, once the system reaches stationarity. We see that  $p(\Delta\Omega_c)$  broadly maintains the same functional form for all  $\alpha$ . All three curves are power laws over approximately two decades, with index  $\approx 1.5$ , just as  $\eta$  is a power law over two decades for  $\beta = 10^{-2}$ . Interestingly, below the lower cut-off, we find  $p(\Delta\Omega_c) \propto \Delta\Omega_c$ , independent of  $\alpha$ . The upper cut-off of  $\Delta\Omega_c$  scales as  $\alpha^{-1}$ . It is also set by the value of  $\beta$ , which determines the number of decades over which  $\eta$  is a power law.

The  $\beta$ -dependence of  $p(\Delta\Omega_c)$  is shown in Figure 7 for  $\beta = 10^{-2}, 10^{-3}, 10^{-4}$ , and  $\alpha = 10^3$ . In each case, the number of decades over which  $p(\Delta\Omega_c)$  is a power law is the same as for  $\eta$  and equals  $-\log(\beta)$ . Interestingly, although  $\beta$  sets only the minimum avalanche size, it controls the maximum observed  $\Delta\Omega_c$  as much as the minimum observed  $\Delta\Omega_c$ . This is because a lower value of  $\beta$  reduces the mean avalanche

<sup>7</sup> More generally, if we always have  $X = y$  immediately before a glitch, we obtain  $\langle \Delta\Omega_c \rangle = (\beta y)^{1/2}$  in units of  $X_c I_s / (I_s + I_c)$ .



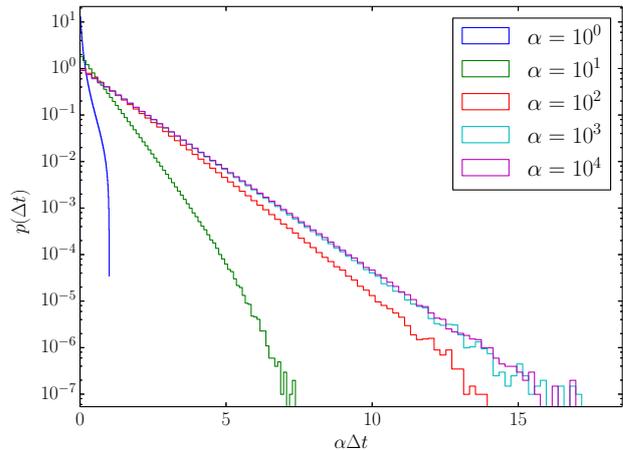
**Figure 6.** PDF of glitch sizes  $p(\Delta\Omega_c)$  for slow spin down  $\alpha \gtrsim \alpha_c(\beta)$  with  $\alpha = 10^2$  (blue curve),  $10^3$  (green curve) and  $10^4$  (red curve), where  $\Delta\Omega_c$  is measured in units of  $X_c(I_s + I_c)/I_s$ . For each value of  $\alpha$ ,  $\Delta\Omega_c$  is sampled over the course of a single simulation featuring  $N = 10^8$  glitches after the system reaches stationarity. Parameter:  $\beta = 10^{-2}$ .



**Figure 7.** PDF of glitch sizes  $p(\Delta\Omega_c)$  for slow spin down  $\alpha \gtrsim \alpha_c(\beta)$  with  $\beta = 10^{-2}$  (blue curve),  $10^{-3}$  (green curve), and  $10^{-4}$  (red curve). For each value of  $\beta$ ,  $\Delta\Omega_c$  is sampled over the course of a single simulation featuring  $N = 10^8$  glitches, after the system reaches stationarity. Parameter:  $\alpha = 10^3$ .

size, increases the value of  $X$  where glitches are likely to occur, and hence increases the maximum lag that can be erased in a single glitch. Interestingly, the shape of the PDF below the lower cut-off is independent of both  $\beta$  and  $\alpha$  for  $\alpha \gtrsim \alpha_c$ .

The effect of changing  $\alpha$  on  $p(\Delta t)$  is shown in Figure 8. In order to keep multiple curves on the same graph, we plot  $\alpha\Delta t$  on the horizontal axis. We see that for  $\alpha \geq 10^2$ , the distribution has an exponential shape. The upper cut-off  $\Delta t$  scales as  $\alpha^{-1}$ . In this regime, the system glitches at such small  $X$  that  $\lambda(X)$  in (11) is approximately constant, and the average avalanche rate is proportional to  $\alpha$ . As  $\alpha$  decreases, the dependence of  $p(\Delta t)$  on  $\alpha$  becomes more complicated, as dynamic feedback emerges between glitch sizes and waiting times.



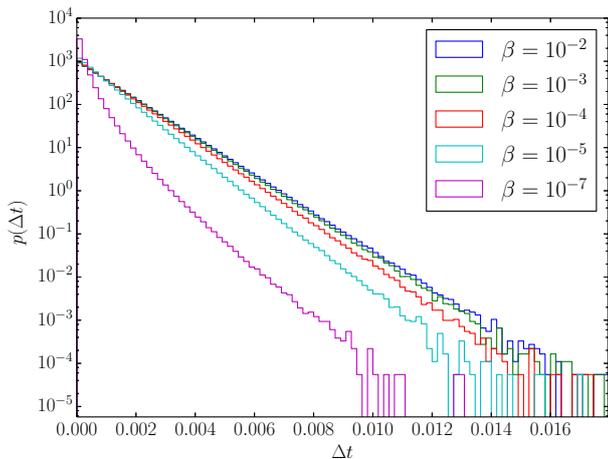
**Figure 8.** PDF of waiting times  $p(\Delta t)$  for slow spin down  $\alpha \gtrsim \alpha_c(\beta)$  with  $\alpha = 10^0$  (blue curve),  $10^1$  (green curve),  $10^2$  (red curve),  $10^3$  (cyan curve) and  $10^4$  (violet curve).  $\Delta t$  is measured in units of  $I_c X_c / N_{\text{em}}$ . In order to keep the plots on the same scale, we scale the horizontal axis by  $\alpha$  for each plot. The vertical spacing between the curves arises from normalisation. For each value of  $\alpha$ ,  $\Delta t$  is sampled over the course of a single simulation featuring  $N = 10^8$  glitches, after the system reaches stationarity. Parameter:  $\beta = 10^{-2}$ .

The effect of the minimum avalanche size parameter  $\beta$  on  $p(\Delta t)$  is encapsulated by Figure 9. For  $\beta \lesssim 10^{-4}$ ,  $p(\Delta t)$  diverges from an exponential and approaches a power law. This behaviour is evident most clearly at  $\beta = 10^{-7}$ . Note that, on the  $\beta = 10^{-4}$  curve in Figure 5, the system transitions to the  $\alpha \lesssim \alpha_c$  regime at  $\alpha_c \approx 10^3$ . Likewise, in the top panel of Figure 5,  $\mu_{\Delta\Omega_c}$  and hence  $\mu_{\Delta t}$  do not change with  $\beta$  for  $\alpha \ll \alpha_c$ . In Figure 9,  $p(\Delta t)$  maintains the same exponential functional form for all  $\beta \gtrsim 10^{-4}$ . In other words, the entire PDF is independent of  $\beta$  in this regime, not just  $\langle \Delta t \rangle$ .

#### 4.4.4 PDFs for rapid spin-down: $\alpha \lesssim \alpha_c(\beta)$

We now ask how the observable PDFs change as we move to the small- $\alpha$  regime. Figure 10 shows the size PDF  $p(\Delta\Omega_c)$  for five values of  $\alpha$  in the range  $10^{-4} \leq \alpha \leq 10$  and  $\beta = 10^{-2}$ . The range covers the full suite of behaviour seen in the moments of  $\Delta\Omega_c$  and  $\Delta t$  in Figure 5; it brackets  $\alpha_c(\beta) \approx \beta^{1/2} = 10^{-1}$ . The vertical axis is scaled by  $X_c(I_c + I_s)/I_s$  so that the PDF equals the avalanche size distribution  $p(\Delta X)$ . For  $\alpha \geq 10^{-4}$ , the lag is close to  $X \approx 1$  before every glitch. The spread in  $X$  before every glitch is small, so that  $p(\Delta X)$  is approximately  $\eta(1 - \Delta X|1)$ . As  $\alpha$  increases,  $\eta$  is convolved with the increasing spread in  $X$  before a glitch occurs, so that the Heaviside cut-off in  $p(\Delta X)$  becomes blurred. At  $\alpha = 10^1$ ,  $p(\Delta\Omega_c)$  is a power law over two decades, but the value of the maximum and minimum change with  $\alpha$ . We also find  $p(\Delta\Omega_c) \propto \Delta\Omega_c^{3/2}$  below the lower cut-off.

Unlike  $p(\Delta\Omega_c)$ , which is always a power law,  $p(\Delta t)$  changes its functional form when we move to the small- $\alpha$  regime. Figure 11 displays  $p(\Delta t)$  for the same values of  $\alpha$  and  $\beta$  as in Figure 10. For  $\alpha = 10^{-4}$ ,  $p(\Delta t)$  is approximately a power law of the same form as  $p(\Delta\Omega_c)$ . Physically, the lag reaches  $X \approx 1$  before every glitch, so the dimen-



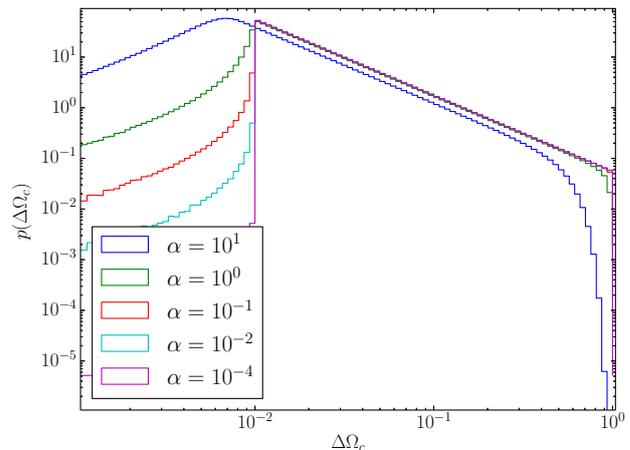
**Figure 9.** PDF of waiting times  $p(\Delta t)$  for slow spin down  $\alpha \gtrsim \alpha_c(\beta)$  with  $\beta = 10^{-2}$  (blue curve),  $10^{-3}$  (green curve),  $10^{-4}$  (red curve),  $10^{-5}$  (cyan curve), and  $10^{-7}$  (magenta curve).  $\Delta t$  is measured in units of  $N_{\text{em}}/(I_c X_c)$ . For each value of  $\beta$ ,  $\Delta t$  is sampled over the course of a single simulation featuring  $N = 10^8$  glitches, after the system reaches stationarity. Parameter:  $\alpha = 10^3$ .

sionless forward waiting time until the next glitch approximately equals the dimensionless size of the previous glitch. (This suggests the existence of a correlation between size and forward waiting-time, which we test in Section 4.5.) As  $\alpha$  increases, the spread in  $X$  when a glitch occurs widens, and  $p(\Delta t)$  turns into an exponential. For  $10^0 \leq \alpha < 10^{-2}$ , there is an interesting mixture of power-law and exponential components. For  $\alpha \ll 10^{-4}$ , the value of  $\alpha$  has no influence on  $p(\Delta t)$ , unlike in the regime  $\alpha \gtrsim \alpha_c$ .

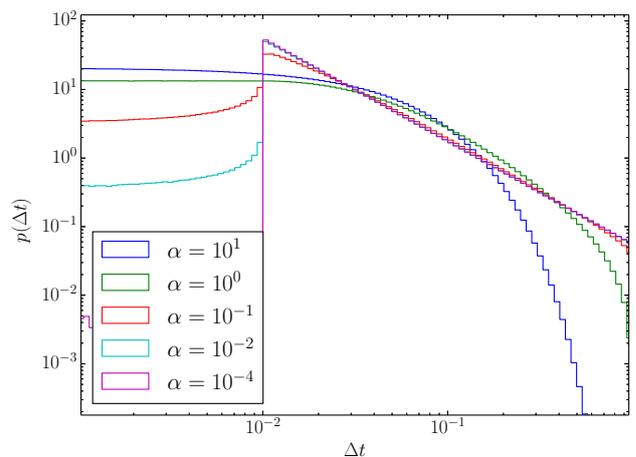
The minimum avalanche size (through  $\beta$ ) determines the range over which  $p(\Delta\Omega_c)$  and  $p(\Delta t)$  are power laws. For  $\alpha \ll \alpha_c$ ,  $p(\Delta\Omega_c)$  and  $p(\Delta t)$  are practically identical. Figure 12 displays both PDFs for  $\beta = 10^{-2}$  and  $\beta = 10^{-4}$  and  $\alpha = 10^{-3}$ . We rescale both random variables in order to reveal the connection between the distributions:  $\Delta t$  is scaled by  $N_{\text{em}}/I_c$  and  $\Delta\Omega_c$  is scaled by  $X_c(I_c + I_s)/I_s$ . Both distributions are power laws over three (two) decades for  $\beta = 10^{-3}$  ( $\beta = 10^{-2}$ ).

#### 4.5 Size-waiting-time correlation

The classic vortex unpinning model predicts a correlation between glitch sizes and waiting times, because the weakly coupled superfluid plays the role of an angular momentum reservoir. According to this argument, the larger the glitch, the more completely the reservoir is emptied, and the more time must pass until the next glitch, before  $X(t)$  builds up sufficiently. That is, one expects a correlation between the sizes and forward waiting times. By the same token, the longer the waiting time that elapses, the more completely the reservoir is replenished, and the larger is the maximum possible  $\Delta X^{(i)}$  for the next glitch. That is, one also expects a correlation between sizes and backward waiting times, especially if glitches empty the reservoir fully. In reality, however, no such correlations are observed, except in PSR J0537–6910 (see Section 2.1), whose sizes are correlated with forward waiting times (Middleditch et al. 2006).



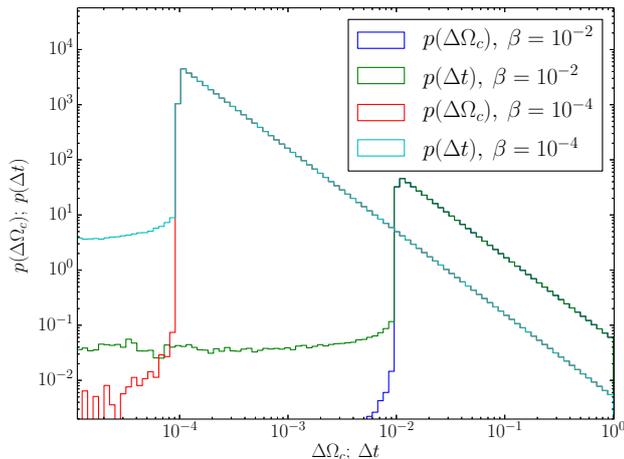
**Figure 10.** PDF of glitch sizes  $p(\Delta\Omega_c)$  for rapid spin down  $\alpha \lesssim \alpha_c(\beta)$  with  $\alpha = 10$  (blue curve),  $10^0$  (green curve),  $10^{-1}$  (red curve),  $10^{-2}$  (cyan curve) and  $10^{-4}$  (magenta curve).  $\Delta\Omega_c$  is measured in units of  $I_s/[X_c(I_c + I_s)]$ . For each value of  $\beta$ ,  $\Delta\Omega_c$  is sampled over the course of a single simulation featuring  $N = 10^8$  glitches, after the system reaches stationarity. Parameter:  $\beta = 10^{-2}$ .



**Figure 11.** PDF of waiting times  $p(\Delta t)$  for rapid spin down  $\alpha \lesssim \alpha_c(\beta)$  with  $\alpha = 10$  (blue curve),  $10^0$  (green curve),  $10^{-1}$  (red curve),  $10^{-2}$  (cyan curve) and  $10^{-4}$  (magenta curve).  $\Delta t$  is measured in units of  $N_{\text{em}}/(I_c X_c)$ . For each value of  $\beta$ ,  $\Delta t$  is sampled over the course of a single simulation featuring  $N = 10^8$  glitches, after the system reaches stationarity. Parameter:  $\beta = 10^{-2}$ .

The general lack of correlation is normal in self-organised critical systems, where most jumps empty a small fraction of the stress reservoir (Jensen 1998; Melatos et al. 2008). No pulsars exhibit a correlation with backward waiting times in data collected to date.

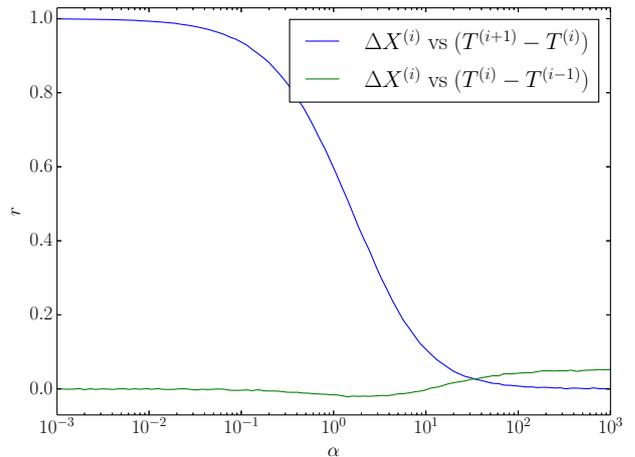
What does our model say about these correlations? The answer depends on  $\alpha$ . In Figure 13, we compute the Pearson correlation coefficient,  $r$ , for both the forward [ $\Delta\Omega_c$  versus  $T^{(i+1)} - T^{(i)}$ ; blue curve] and backward [ $\Delta\Omega_c$  versus  $T^{(i)} - T^{(i-1)}$ ; green curve] waiting times as a function of  $\alpha$ , where  $T^{(i)}$  is the epoch of the  $i$ -th glitch. To produce the plots, we run a simulation generating  $N = 10^6$  glitches



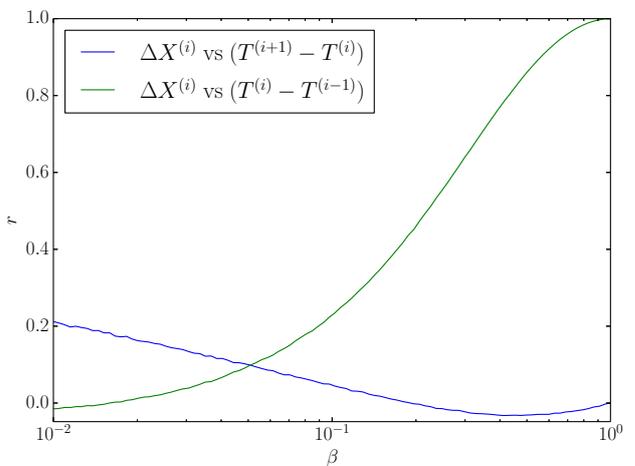
**Figure 12.** Functional forms of the PDFs of glitch sizes  $p(\Delta\Omega_c)$  and waiting times  $p(\Delta t)$  for rapid spin down  $\alpha \lesssim \alpha_c(\beta)$  with  $\beta = 10^{-4}$  (red and cyan curves respectively) and  $\beta = 10^{-2}$  (dark blue and green curves respectively). In order to maintain a common scale on the vertical axis,  $\Delta\Omega_c$  is measured in units of  $I_s/[X_c(I_c + I_s)]$ , and  $\Delta t$  is measured in units of  $N_{\text{em}}/(I_c X_c)$ . For each value of  $\beta$ ,  $\Delta t$  and  $\Delta\Omega_c$  are sampled over the course of a single simulation featuring  $N = 10^8$  glitches, after the system reaches stationarity. Parameter:  $\alpha = 10^{-3}$ .

for each value of  $\alpha$ . The Pearson coefficient is a measure of correlation (Press et al. 2007):  $r = 1$  ( $r = -1$ ) indicates that the two sets are perfectly correlated along a line with positive (negative) gradient, whereas  $r = 0$  indicates no correlation. We see that the forward correlation asymptotes to 1 for  $\alpha \ll \alpha_c$  but quickly drops toward zero for  $\alpha \gtrsim \alpha_c$ . Physically, for small  $\alpha$ , the spin-down rate is high compared to  $\lambda_0$ . Hence, after a glitch of dimensionless size  $\Delta X$ , the system waits for  $\Delta t \approx \Delta X$  (until  $X \approx 1$ ) before glitching again. Interestingly, the same does not hold for the backward correlation, which is approximately zero for all  $\alpha \lesssim \alpha_c$  and only increases slightly at  $\alpha \gtrsim \alpha_c$ . For small  $\alpha$ , the system glitches at  $X \approx 1$  so the avalanche size distribution is always  $p(\Delta x) \approx \eta(1 - \Delta X|1)$ , independent of the size of the previous glitch, and there is zero backward correlation. As  $\alpha$  increases, the maximum value of  $X$  where glitches occur decreases. A glitch occurs at a higher value of  $X$ , on average, after a longer waiting-time interval than a shorter one, so a correlation begins to emerge for large  $\alpha$ . The correlation remains weak even for  $\alpha \gg \alpha_c$ , however, because there is always a statistical spread in glitch sizes over  $-\log(\beta)$  decades. In the limit  $\beta \rightarrow 1$ , the jump distribution  $\eta(x|y)$  approaches a delta function  $\delta(y - x)$ , so that every glitch restores the system stress to zero. In this case, there is perfect correlation between glitch sizes and backward waiting times, because the time since the last glitch exactly equals the instantaneous crust-superfluid lag released in the next glitch. Figure 14 shows the emergence of a correlation with backward waiting times as  $\beta \rightarrow 1$ .

In summary, a broad conditional jump distribution with  $\beta \ll 1$ , of the form observed in GPE simulations, destroys the correlation between sizes and backward waiting-times for all  $\alpha$ . Sure enough, correlations with backward waiting times are not observed in any pulsar. A broad  $\eta(x|y)$  does allow a significant correlation between sizes and forward waiting



**Figure 13.** Correlation between sizes and forward/backward waiting times: Pearson coefficient,  $r$ , as a function of the control parameter  $\alpha$ . For each value of  $\alpha$ , we sample  $\Delta t$  and  $\Delta X$  over the course of a single simulation featuring  $N = 10^6$  glitches, after the system reaches stationarity. For the forward waiting times, we compute the correlations between  $\Delta X^{(i)}$  and  $T^{(i+1)} - T^{(i)}$ , where  $\Delta X^{(i)}$  and  $T^{(i)}$  are the dimensionless size and epoch of the  $i$ -th avalanche, respectively. For the backward waiting times, we compute the correlation between  $\Delta X^{(i)}$  and  $T^{(i)} - T^{(i-1)}$ . Parameter:  $\beta = 10^{-2}$ .



**Figure 14.** Correlation between sizes and forward/backward waiting times: Pearson coefficient,  $r$ , as a function of the control parameter  $\beta$ . For each value of  $\beta$ , the forward and backward correlations are obtained as in Figure 13. Parameter:  $\alpha = 5$

times provided that one has  $\alpha \lesssim \alpha_c$ . A correlation with forward waiting times is observed in one pulsar only, namely PSR J0537–6910.

#### 4.6 Summary

The results in Sections 4.1–4.5 can be interpreted in a unified way as follows. Given a conditional jump distribution  $\eta$ , whose lower cut-off is set by  $\beta$ , there are three regions of  $\alpha$  which exhibit different behaviour.

For  $\alpha \gg \alpha_c(\beta)$ , the system is driven slowly compared to the reference avalanche rate  $\lambda_0$ . The system self-adjusts so

that avalanches occur near  $X \approx 0$ , where  $\lambda(X) \approx \alpha$  is essentially constant. Hence the waiting-time PDF  $p(\Delta t)$  reduces to that of a constant-rate Poisson process with exponential waiting times, as in Figure 8. As  $\lambda(X)$  is approximately constant, the conditional jump distribution and hence  $\beta$  exert little influence on  $p(\Delta t)$ , as shown in Figure 9. The value of  $\alpha$  controls the observed upper cut-off of  $\Delta\Omega_c$  and  $\Delta t$ .

For  $\alpha \ll \alpha_c(\beta)$ ,  $X$  comes up against a hard wall with  $\lambda(X) \rightarrow \infty$  as  $X \rightarrow 1$ . The lag returns to  $X \approx 1$  before every glitch. After a glitch, a time interval proportional to the previous  $\Delta\Omega_c$  elapses before the next glitch, introducing a correlation, and causing  $p(\Delta t)$  to approach  $p(\Delta X) \approx \eta(1 - \Delta X|1)$ . The strong correlation that emerges between glitch sizes and forward waiting times in this regime is not observed in most pulsars. The correlation between sizes and backward waiting times remains weak, because of the broad conditional avalanche distribution assumed in the model and observed in pulsars.

In the intermediate regime  $\alpha \approx \alpha_c(\beta)$ , with  $\alpha$  spanning about two decades, glitches occur over a wide range of  $X$  values, where  $\eta$  and  $\lambda$  vary substantially. This introduces a complex dynamic feedback between glitch sizes and waiting times. From an astrophysical perspective, this regime is especially interesting. Without an analytic solution, however, it is hard to make general, quantitative statements about the behaviour in this regime. We know of no such solution for  $\eta$  given by (18) and (19). We return to the issue in Section 6, where we present analytic solutions for the special case of separable conditional jump distributions.

## 5 LONG-TERM DYNAMICS

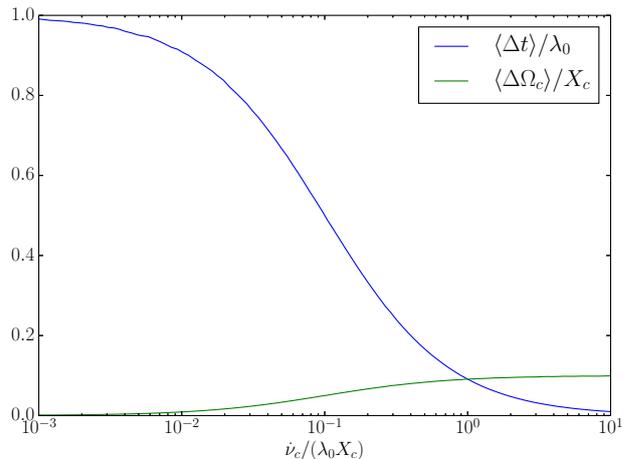
How does a pulsar's glitch activity evolve, as the pulsar ages? The results in Section 4 can be applied directly to answer this question. The time-scale on which the stochastic process in (7) reaches stationarity, identified as  $I_c X_c / N_{\text{em}} \sim 2X_c \tau_c / \Omega_c$  in Section 4, is much shorter than the spin-down time-scale,  $\tau_c$ , because one has  $X_c \ll \Omega_c$  typically (see Section 3.3). For example, the typical inter-glitch time-scale is  $\sim 1$  yr for glitch sizes resolvable by current timing experiments, while the spin-down time-scale satisfies  $\tau_c \sim 10^3$  yr. Hence the control parameters  $\alpha$  and  $\beta$  vary quasistatically over the life-time of a pulsar; the system self-adjusts to remain in the stationary state considered in Section 4 at all times.

### 5.1 Glitch history of an individual pulsar

The control parameter  $\alpha$  varies chiefly because  $N_{\text{em}}$  decreases, as the pulsar ages. It also varies with temperature, as the pulsar cools, through  $X_c$  and  $\lambda_0$ , but this is a weaker scaling, and the cooling history of neutron stars is uncertain, so for now we treat  $X_c$  and  $\lambda_0$  as constants. For the simple magnetic-dipole spin-down law  $\Omega_c \propto \Omega_c^3$ , we have

$$\Omega_c(t) = \Omega_c(0) \left[ 1 + \frac{t}{\tau_c(0)} \right]^{-1/2}, \quad (23)$$

where  $t = \tau_c$  is the pulsar's age,  $\Omega_c(0)$  and  $\dot{\Omega}_c(0)$  are the pulsar angular speed and spin-down rate at birth respectively,



**Figure 15.** Long-term evolution of a glitching pulsar: mean waiting time  $\langle \Delta t \rangle$  and glitch size  $\langle \Delta \Omega_c \rangle$  as a function of the instantaneous crust spin-down rate,  $\dot{\nu}_c$ . For each value of  $\dot{\nu}_c$ , we sample  $\Delta t$  and  $\Delta \Omega_c$  from a single simulation featuring  $N = 10^6$  glitches after the system reaches stationarity. Parameter:  $\beta = 10^{-2}$ .

and we define  $\tau_c(0) = \Omega_c(0) / [2\dot{\Omega}_c(0)]$ . With the approximation  $\Omega_c \ll \Omega_c(0)$ , we find  $\tau_c \approx 2\Omega_c / |\dot{\Omega}_c|$ , and hence

$$\alpha = \frac{2\lambda_0 X_c \tau_c}{\Omega_c}. \quad (24)$$

Observations indicate a modified braking law with  $\dot{\Omega}_c \propto \Omega_c^n$  and  $2 \leq n \leq 3$ , (Melatos 1997; Archibald et al. 2016), but the results do not change qualitatively when this is taken into account.

The second control parameter,  $\beta$ , is independent of  $\tau_c$ . The ratio of the minimum to maximum avalanche size is an intrinsic property of the avalanche knock-on physics. The natural time- and frequency-scales used to make (7) dimensionless also scale with  $\tau_c$  as  $\alpha/\lambda_0 \propto \tau_c$  and  $X_c = \text{constant}$ , respectively. For  $\Omega_c(0) \gg \Omega_c(t)$ , we find  $\dot{\Omega}_c \propto \tau_c^{-3/2}$  and hence  $\alpha \propto \tau_c^{3/2}$ , when holding  $X_c$  and  $\lambda_0$  constant.

In Figure 15 we plot the evolution of the mean waiting time and glitch size versus  $\dot{\nu}_c = (2\pi)^{-1}\Omega_c$ . We note that the small- $\dot{\Omega}_c$  behaviour in Figure 15, in which the average glitch rate asymptotes to a maximum value, is not a fundamental feature of the model, but rather a consequence of the specific choice of  $\lambda(X)$  in (11), which gives  $\lambda(X) \rightarrow \lambda_0$  for  $X \rightarrow 0$  when dimensions are restored.

### 5.2 Activity parameter

An immediate implication of the model is that, in the stationary state, we have  $\langle \Delta X \rangle = \langle \Delta t \rangle$ , where  $\Delta X$  and  $\Delta t$  refer to the dimensionless scaled variables defined in (15), and the angle brackets denote a long-time average over a single simulation, with  $\alpha$  and  $\beta$  held constant, as throughout Section 4. This result is clear from Figure 4, and is proved analytically for arbitrary  $\lambda$  and  $\eta$  in Appendix B; the system always self-adjusts to generate glitches of the correct size and frequency to accommodate the driver  $\dot{\Omega}_c(t)$ . Restoring

dimensions temporarily, the scaling relations in (15) convert  $\langle \Delta X \rangle = \langle \Delta t \rangle$  to

$$\frac{\langle \Delta \nu_c \rangle}{\langle \Delta t \rangle} = \frac{I_c \dot{\nu}_c}{I_c + I_s}. \quad (25)$$

Equation (25) is closely related to the glitch activity parameter, which was defined by McKenna & Lyne (1990) as

$$A = \sum_{i=1}^{N(T)} \frac{\Delta \nu_c^{(i)}}{T}, \quad (26)$$

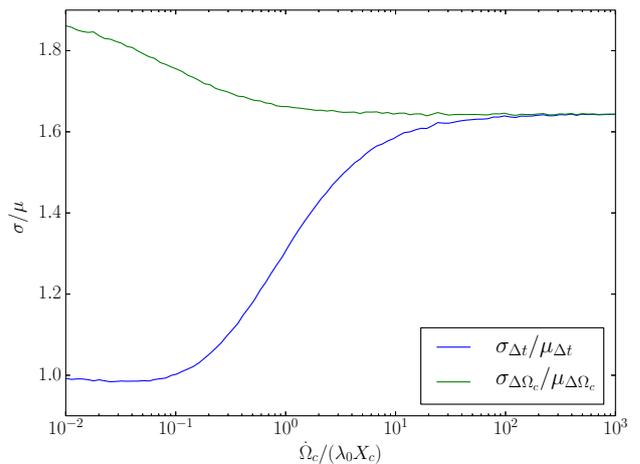
where  $T$  is the total observation time and  $N(T)$  is the number of glitches during that time. As  $N(T)$  grows, the activity parameter  $A$  approaches the left-hand side of (25). The same quantity is referred to in Espinoza et al. (2011) as  $\dot{\nu}_{\text{glitch}}$ .

We note that equation (25) is an immediate consequence of any model that assumes that glitches release stress from a finite reservoir of angular momentum, and that the crust-superfluid lag fluctuates around an equilibrium value. It does not require a mean-field approximation or the assumption that glitch activity occurs as a state-dependent Poisson process; i.e. it is not particular to the model in this paper. Ruderman et al. (1998) also derived expression (25); see equation (27) of the latter paper. Link et al. (1999) noted that, if the superfluid is allowed to transfer angular momentum to the crust between glitches also, then (25) becomes an inequality,  $\dot{\nu}_{\text{glitch}}/\dot{\nu}_c \leq I_s/(I_s + I_c)$ . From the activity parameter for Vela, they concluded  $I_s/(I_s + I_c) \geq 1.4\%$ , while Lyne et al. (2000) found  $\approx 10^{-2}$  for the same ratio. Accurate quantitative testing of (25) is constrained by our inability to directly measure  $I_s/(I_s + I_c)$  for individual pulsars. Some theoretical predictions put the ratio at  $10^{-2}$  (Piekarewicz et al. 2014), while recent theoretical work has supported the possibility of larger  $I_s$  (Chamel 2012). If  $I_s/(I_c + I_s)$  is approximately the same throughout the population, however, the plot of  $\log(\dot{\nu}_{\text{glitch}})$  versus  $\log(\dot{\nu}_c)$  should have unit slope.

Observational data agree with the above prediction (Lyne et al. 2000; Espinoza et al. 2011). This is reflected in an approximately monotonic decrease in glitch activity with characteristic age  $\tau_c \approx 2\nu_c/\dot{\nu}_c$ . Within a group of young pulsars, however, including the Crab, PSR J1833–1034, PSR J1119–6127, and PSR J1846–0258, all of which have  $\tau_c < 10$  kyr,  $A$  is lower than for middle-aged pulsars with  $10^3 \text{ yr} \leq \tau_c \leq 10^4 \text{ yr}$  (McKenna & Lyne 1990; Roy et al. 2012). McKenna & Lyne (1990) suggested that younger stars are hotter, so most of the superfluid angular momentum transfers smoothly to the crust via vortex creep, which is powered by thermal activation of pinned vortices, instead of impulsive glitches. This issue is not considered by our model, which ignores vortex creep. We note that the measured value of  $\langle \Delta \nu_c \rangle$  over  $\sim 40$  yr of timing observations may underestimate the true long-term average, if the left-hand side of (25) is dominated by large, rare glitches (e.g. from an underlying power-law distribution with a fat tail), which have not yet been observed.

### 5.3 Waiting-time moments

As a pulsar transitions from rapid to slow spin down, the waiting-time PDF  $p(\Delta t)$  changes its functional form, as described in Section 4.4. The standard deviation of  $p(\Delta t)$



**Figure 16.** Ratio of standard deviation to mean of glitch sizes (green curve) and waiting times (blue curve) as a function of  $\dot{\Omega}_c$ . We assume that  $\lambda_0$  and  $X_c$  remain constant throughout spin down, so  $\dot{\Omega}_c \propto \alpha^{-1}$ . For each value of  $\alpha$  we sample  $\Delta t$  and  $\Delta \Omega_c$  from a single simulation featuring  $N = 10^6$  glitches after the system reaches stationarity. Parameter:  $\beta = 10^{-2}$ .

equals its mean in the rapid spin-down regime (corresponding to an exponential PDF) but equals the standard deviation of the avalanche size PDF in the slow spin-down regime (see Figure 4). In Figure 16 we plot this behaviour explicitly. We assume  $\lambda_0$  and  $X_c$  remain constant, implying  $\alpha \propto \dot{\Omega}_c^{-1}$ . As the pulsar spins down and moves from right to left across the plot, the ratio of mean to standard deviation of waiting times drops over approximately two decades from  $\alpha \approx \alpha_c \approx 10^2$  to  $\alpha \approx 10^0$ .

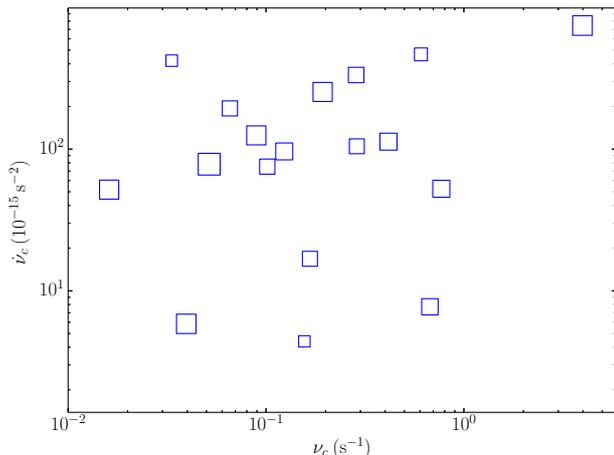
How does the behaviour in Figure 16 compare with actual pulsars? In Figure 17, we plot the standard deviation divided by the mean of  $\Delta t$  for 18 pulsars with five or more observed glitches. The size of each point is proportional to  $\sigma_{\Delta t}/\mu_{\Delta t}$ . The pulsars are shown in their position on the standard  $\nu_c$ - $\dot{\nu}_c$  diagram. Data are taken from the ATNF and Jodrell Bank pulsar catalogues (Espinoza et al. 2011; Manchester et al. 2005) (see footnote 1). No clear trend is visible; further analysis is required, as more data become available. We note that  $\alpha$  is a function of  $\dot{\Omega}_c$ ,  $X_c$  and  $\lambda_0$ , but not of  $\Omega_c$ , as the lag is determined by  $\Omega_s - \Omega_c$ .

## 6 MASTER EQUATION

The results in Sections 4 and 5 demonstrate the complex behaviour generated by (15). Given a rate function  $\lambda(X)$  and conditional jump distribution  $\eta$ , the system self-adjusts by changing the PDF of the superfluid-crust lag  $X$ , until a stationary state is reached. In this section, we develop an analytic model of this stochastic evolutionary process based on a master equation for the lag PDF.

### 6.1 Equation of motion

Let  $p(t, x) dx$  be the probability that the crust-superfluid lag falls in the range  $x \leq X(t) \leq x + dx$  at time  $t$ . From



**Figure 17.** Scatter plot of the ratio of the standard deviation to mean of waiting times, over spin frequency  $\nu_c$  versus spin frequency derivative  $\dot{\nu}_c$  for 18 pulsars with five or more observed inter-glitch intervals (six or more observed glitches). The area of each scatter point is linearly proportional to  $\sigma_{\Delta t}/\mu_{\Delta t}$ . Data are taken from the Jodrell Bank catalogue and ATNF pulsar catalogue (see footnote 1). The plot is drawn on log-log axes.

(15),  $p(t, x)$  evolves according to the dimensionless master equation

$$p(t, x)t = -p(t, x)x - \lambda(x)p(t, x) + \int_x^1 dy p(t, y)\lambda(y)\eta(x|y). \quad (27)$$

The first term on the RHS of (27) equals the rate at which probability exits the interval  $(x, x + dx)$  due to the deterministic spin down of the crust. The second term equals the rate at which probability exits the interval  $(x, x + dx)$ , when a glitch moves the system away from  $x$  at time  $t$ . The third term is the rate that probability enters the interval  $(x, x + dx)$ , when a glitch moves the system from  $y$  to  $(x, x + dx)$  at time  $t$ . [Daly & Porporato \(2007\)](#) and [Wheatland \(2009\)](#) considered (27) for a more general form of deterministic forcing, which we treat in the appendices, and allowed  $X(t)$  to assume positive and negative values.

From numerical simulations (see Section 4.1), we know that the system evolves to a stationary state with  $p(t, x) = p(x)$  for  $t \gtrsim 1$ . Solving (27) analytically for  $t \lesssim 1$  to describe the approach to stationarity lies outside scope of this paper. In the stationary state, we have

$$0 = -\frac{dp(x)}{dx} - \lambda(x)p(x) + \int_x^1 dy p(y)\lambda(y)\eta(x|y). \quad (28)$$

The only physically relevant solutions of (28) are non-negative functions  $p(x)$  with unit integral over the interval  $0 < x < 1$  and which are differentiable in the interior of the interval. Although we select  $\lambda(x) = \alpha/(1-x)$  as an exemplar, the general analysis only assumes the following conditions: (i)  $\lambda(x) \geq 0$  (i.e. the instantaneous avalanche rate cannot be negative); (ii)  $\Lambda(x) \rightarrow \infty$  as  $x \rightarrow 1$  from below, with

$$\Lambda(x) = \int_0^x d\xi \lambda(\xi) \quad (29)$$

(i.e. the integrated avalanche rate diverges as the critical lag is approached); and (iii)  $\eta(x|1) \neq 0$  in some finite  $x$  interval.

## 6.2 Boundary conditions

Relevant, self-consistent solutions of (28) satisfy the boundary conditions

$$p(0) = 0 = p(1). \quad (30)$$

It may be noted that integration of (27) produces the same boundary conditions for the time-dependent problem also. Equation (30) holds for the following reason. From condition (iii) in Section 6.1, we see that

$$\langle \lambda \rangle = \int_0^1 dx \lambda(x)p(x). \quad (31)$$

must be finite. If we integrate (28) over  $0 < x < 1$  and interchange the orders of integration in the double integral, we find  $p(1) - p(0) + \langle \lambda \rangle = \langle \lambda \rangle$ , and hence  $p(0) = p(1)$ . However condition (ii) and the finite  $\langle \lambda \rangle$  enforce  $p(1) = 0$ ; the divergence  $\lambda(X) \rightarrow \infty$  as  $X \rightarrow 1$  keeps the system at  $X < 1$  always.

In Section 6.3 we examine the PDFs of observable quantities, which we are able to relate to  $p(x)$ . Specific solutions for a special class of conditional jump PDFs  $\eta(x|y)$  are discussed in Section 6.4.

## 6.3 Observable PDFs

Neither  $X(t)$  nor  $p(x)$  can be observed directly. The observables are the inter-glitch waiting times, with PDF  $p(\Delta t)$ , and the glitch sizes, with PDF  $p(\Delta\Omega_c)$ . For notational simplicity, we work in this section with the avalanche size PDF  $p(\Delta x)$ , where  $p(\Delta x)d\Delta x$  is the probability that  $\Delta X$  is in the interval  $(\Delta x, \Delta x + d\Delta x)$ . We note that the glitch size distribution  $p(\Delta\Omega_c)$  can be obtained from this by scaling by a moment-of-inertia factor, viz.  $\Delta X = -I_s \Delta\Omega_c / (I_s + I_c)$ . Clearly  $p(x)$  influences  $p(\Delta t)$  and  $p(\Delta x)$  because it determines the distribution of lags at which the system is likely to glitch. If the system lag equals  $y$  immediately *after* the last avalanche, then, from (14), the conditional PDF of the waiting time to the next avalanche is

$$p(\Delta t|y) = \lambda(y + \Delta t)e^{-\Lambda(y+\Delta t)}e^{\Lambda(y)}, \quad (32)$$

where  $\Lambda(y)$  is defined in (29) above. Equation (32) follows from the standard formula (14) for the waiting time of a Poisson process with deterministic rate  $\lambda(t)$ . In order to obtain  $p(\Delta t)$ , we marginalise (32) over the distribution of  $y$  values immediately after an avalanche. Crucially the latter distribution is not the stationary PDF of  $X$ . Likewise, if instead the system lag equals  $y$  immediately *before* an avalanche, the conditional PDF of the avalanche size is

$$p(\Delta x|y) = \eta(y - \Delta x|y). \quad (33)$$

To obtain the avalanche size distribution, we marginalise (33) over the distribution of  $y$  immediately before a glitch. Again, the latter distribution is not the stationary PDF of  $X$ .

To make further progress, we follow [Daly & Porporato \(2007\)](#) and define two new PDFs:  $p_e(x)dx$ , the probability that the lag lies in  $(x, x + dx)$  at the end of a complete interval of deterministic evolution, immediately before a glitch occurs; and  $p_s(x)dx$ , the probability that the lag lies in  $(x, x + dx)$  at the beginning of a complete interval of deterministic evolution, immediately after a glitch occurs. With these definitions, the waiting-time PDF is

$$p(\Delta t) = \int_0^{1-\Delta t} dy p_s(y) \lambda(y + \Delta t) e^{-\Lambda(y+\Delta t)} e^{\Lambda(y)}, \quad (34)$$

and the avalanche size PDF is

$$p(\Delta x) = \int_{\Delta x}^1 dy p_e(y) \eta(y - \Delta x | y). \quad (35)$$

How do  $p_s(x)$ ,  $p_e(x)$  and  $p(x)$  relate? The latter is the time-dependent PDF  $p(t, x)$  calculated in the stationary regime, where  $p(t, x) = p(x)$  is independent of  $t$ . In contrast, it makes no sense to ask the value of  $p_s$  or  $p_e$  at time  $t$ ;  $p_s$  and  $p_e$  have meaning only at the specific values of  $t$  where  $N(t)$  jumps by one due to a glitch. We discuss this issue more in [Section 7](#). The three PDFs  $p_s(x)$ ,  $p_e(x)$  and  $p(x)$  can be related mathematically in the following way. Under stationary conditions,  $p_s(x)dx$  is the probability that a glitch takes the system from the interval  $(y, y + dy)$  to  $(x, x + dx)$ , integrated over all  $y$ :

$$p_s(x) = \int_x^1 dy p_e(y) \eta(x|y). \quad (36)$$

By the same token,  $p_e(x)dx$  is the probability that the system evolves deterministically for a time interval  $\Delta t = y - x$  from  $(y, y + dy)$  to  $(x, x + dx)$ , integrated over all  $y$ :

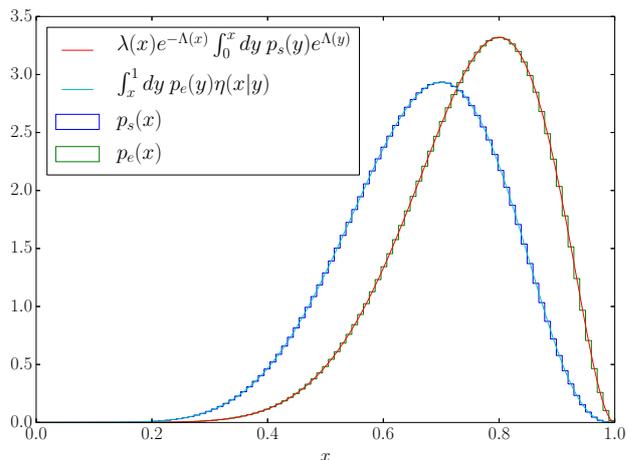
$$p_e(x) = \lambda(x) e^{-\Lambda(x)} \int_0^x dy p_s(y) e^{\Lambda(y)}. \quad (37)$$

Equation (37) follows directly from (32). As a consistency check, it is easy to show that (36) and (37) conserve normalisation, by integrating both sides over  $0 \leq x \leq 1$  and changing the order of integration.

Figure 18 verifies equations (36) and (37). First, we obtain  $p_s(x)$  (dark blue curve) and  $p_e(x)$  (green curve) numerically by running a single simulation involving  $10^8$  glitches. From  $p_e(x)$ , we predict  $p_s(x)$  (cyan curve) according to (36). Similarly, from  $p_s(x)$  we predict  $p_e(x)$  (red curve) according to (37). The four curves overlap in pairs, proving that (36) and (37) correctly describe the relation between  $p_s$  and  $p_e$ . For this example, we choose  $\eta(x|y) = (\delta + 1)x^\delta y^{-\delta-1}$ , with  $\delta = 7$ , and  $\lambda$  as in [Section 4](#) with  $\alpha = 3$ . The choice of  $\delta = 7$  is arbitrary and has been made for illustrative purposes only. The agreement between  $p_s(x)$  generated straight from the simulations (dark blue curve) and from using (36) (cyan curve) is preserved irrespective of the choice of  $\lambda$  and  $\eta$ . The same is true for  $p_e(x)$ .

Equations (36) and (37) together define both  $p_s(x)$  and  $p_e(x)$  uniquely up to normalisation constants. We substitute one equation into the other to obtain an integral equation for each PDF involving only itself:

$$p_s(x) = \int_{y=x}^{y=1} dy \eta(x|y) \lambda(y) e^{-\Lambda(y)} \int_{z=0}^{z=y} dz e^{\Lambda(z)} p_s(z), \quad (38)$$



**Figure 18.** Numerical verification of equations (36) and (37). We generate the PDFs  $p_s(x)$  (dark blue curve) and  $p_e(x)$  (green curve) by sampling a single simulation of  $N = 10^8$  glitches after stationarity. From  $p_s(x)$  we make a prediction (red curve) of  $p_e(x)$  from (36). From  $p_e(x)$  we make a prediction (cyan curve) of  $p_s(x)$  from (37). The curves overlap in pairs. We use  $\eta(x|y) = (\delta + 1)x^\delta y^{-\delta-1}$  and  $\lambda$  as in [Section 4](#). Parameters:  $\delta = 7$ ,  $\alpha = 3$ .

$$p_e(x) = \lambda(x) e^{-\Lambda(x)} \int_{y=0}^{y=x} dy e^{\Lambda(y)} \int_{z=y}^{z=1} dz \eta(y|z) p_e(z). \quad (39)$$

In [Appendix B](#) we detail the general properties of and relations between  $p_s(x)$ ,  $p_e(x)$  and  $p(x)$ . We show that if the stationary master equation (28) has a unique non-negative normalizable solution, then it follows that we have

$$p_e(x) = \frac{\lambda(x)p(x)}{\langle \lambda \rangle}, \quad (40)$$

in accord with [Daly & Porporato \(2007\)](#). Observationally, the normalisation constant  $\langle \lambda \rangle$  defined in (31) is the average  $\lambda$  obtained by sampling an ensemble of simulations at some fixed time  $t \gtrsim 1$ .

#### 6.4 Separable $\eta$

In order to generate practical results from the master equation and gain a fuller understanding of its physical implications, it is useful to develop a tool kit of analytical solutions, which can be manipulated easily and cross-checked against Monte-Carlo simulations. To this end, we temporarily restrict attention to separable conditional jump distributions. From the fundamental normalisation condition (9), they have the form

$$\eta(x|y) = \frac{\varphi(x)}{\Phi(y)}, \quad (41)$$

with

$$\Phi(y) = \int_0^y d\xi \varphi(\xi). \quad (42)$$

#### 6.4.1 Analytic solution

When  $\eta(x|y)$  is separable, the master equation can be solved analytically to obtain the stationary PDF of any quantity of interest. For example, one can show that  $X$  is distributed as

$$p(x) = C\Phi(x)e^{-\Lambda(x)}, \quad (43)$$

where

$$C^{-1} = \int_0^1 dx \Phi(x)e^{-\Lambda(x)} \quad (44)$$

is a normalisation constant. It is simple to show by direct substitution that (43) satisfies (28). The proof that this solution is unique up to normalisation, under specific conditions to be satisfied by  $\lambda$  and  $\eta$ , is set out in Appendix C. We note that the solution (43) for general  $\varphi(x)$  contains the solution of the forest-fire model of [Daly & Porporato \(2006\)](#), where every jump resets the system stress to zero, with the specific choice  $\varphi(x) = \delta(x)$ .

#### 6.4.2 Observable PDFs

What do the distributions discussed in Sections 3–5 look like, when  $\eta$  is separable?  $X_s$  and  $X_e$ , which are not observable, have distributions

$$p_s(x) = \frac{C}{\langle \lambda \rangle} \varphi(x) e^{-\Lambda(x)} \quad (45)$$

and

$$p_e(x) = \frac{C}{\langle \lambda \rangle} \Phi(x) \lambda(x) e^{-\Lambda(x)} \quad (46)$$

respectively. Substituting (45) and (46) into the expressions (34) and (35) for the observed waiting times and avalanche sizes yields

$$p(\Delta x) = \frac{C}{\langle \lambda \rangle} \int_{\Delta x}^1 dy \lambda(y) e^{-\Lambda(y)} \varphi(y - \Delta x) \quad (47)$$

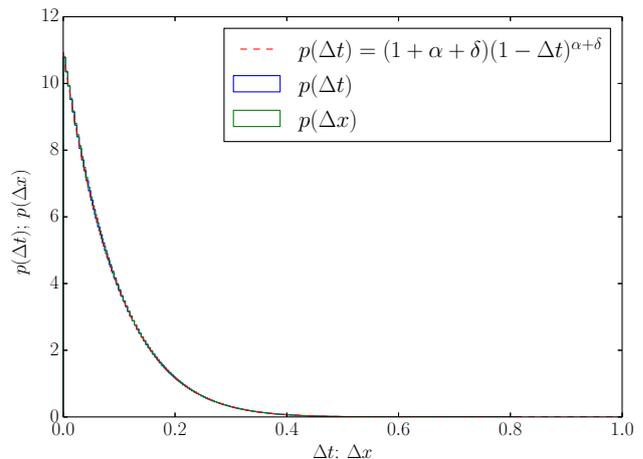
and

$$p(\Delta t) = \frac{C}{\langle \lambda \rangle} \int_0^{1-\Delta t} dy \lambda(y + \Delta t) e^{-\Lambda(y + \Delta t)} \varphi(y). \quad (48)$$

Changing the integration variable in (47) to  $u = y + \Delta x$  leads to (48). This is an interesting result: when  $\eta$  is separable, the observed waiting-time PDF has the same functional form as the size PDF. This generally fails to hold for a non-separable conditional jump distribution except in the regime  $\alpha \gtrsim \alpha_c$ , e.g. we find power-law sizes and exponential waiting times in Section 4.

#### 6.4.3 Illustrative special case: $\varphi(x) = x^\delta$

The conditional jump distribution defined by (18) and (19) and used in Sections 3–5 is not separable, so it does not relate directly to the solution (43). We know of no analytic solution for the general non-separable case. Nevertheless we can



**Figure 19.** Avalanche waiting-time and size distributions  $p(\Delta t)$  (blue curve) and  $p(\Delta x)$  (green curve) for the separable conditional jump distribution  $\eta(x|y) = (\delta + 1)x^\delta y^{-\delta-1}$ . The variables are sampled from a single simulation of  $N = 10^8$  glitches. The PDFs agree with each other and the analytic expression (D1) (red dotted line). Parameters:  $\alpha = 3$ ,  $\delta = 7$ .

study qualitatively the behaviour in Sections 3–5 by analogy with a similar yet separable jump distribution. We take  $\varphi(x) = x^\delta$  as an example, with  $\delta > 0$ . For large  $\delta$ ,  $\varphi(x) = x^\delta$  sharply favours small jumps but still allows glitches that reset the stress to zero, as in Section 4. There is, however, an important difference. The requirement of separability means we are unable to set a non-zero minimum avalanche size conditional on the value of  $X$  when a glitch occurs, tuned with the parameter  $\beta$  in Section 4.

Figure 19 displays the waiting time and avalanche size PDFs  $p(\Delta x)$  and  $p(\Delta t)$  obtained by simulating  $N = 10^8$  glitches, alongside the analytic expression for  $p(\Delta t)$  and  $p(\Delta x)$ , for  $\varphi(x) = x^\delta$ , viz.

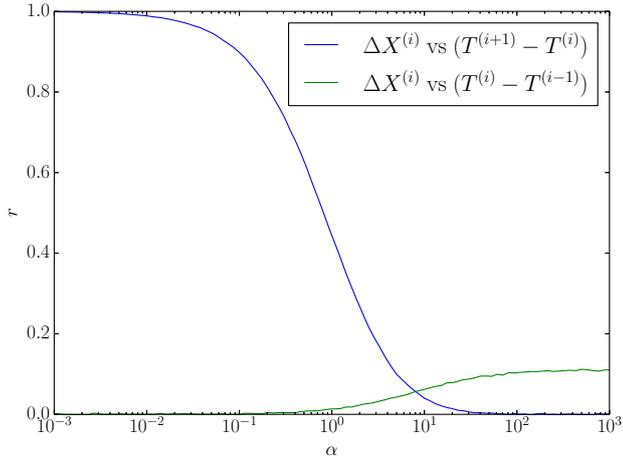
$$p(\Delta t) = (1 + \alpha + \delta)(1 - \Delta t)^{\alpha + \delta}. \quad (49)$$

Equation (49) is derived in Appendix D. The three curves agree with one another. Physically, the  $\alpha$  and  $\delta$  dependence of (49) can be explained as follows. A larger  $\delta$  favours smaller jumps (qualitatively similar to reducing  $\beta$  in Section 4), which forces the system to self-adjust, so that  $p(x)$  is squeezed closer to the upper bound at  $X = 1$ . Increasing  $\alpha$  has an opposite effect: it makes avalanches more likely to occur at smaller  $X$  values, pushing  $p(x)$  toward zero.

The waiting-time (or equivalently avalanche-size) PDF (49) yields the following result for the means:

$$\langle \Delta t \rangle = \frac{1}{\alpha + \delta + 2} = \langle \Delta X \rangle. \quad (50)$$

This is qualitatively the same behaviour as that observed for the non-separable  $\eta$  used in Sections 4 and 5. Now the transition between the rapid and slow spin-down regimes occurs at  $\alpha_c \approx \delta$ . (Here too  $\alpha_c$  is a function of the avalanche size parameter.) Slow spin down corresponds to  $\alpha \gg \delta$ , where (50) yields  $\langle \Delta X \rangle = \langle \Delta t \rangle \approx (\alpha + 2)^{-1}$ . Hence both  $\langle \Delta t \rangle$  and  $\langle \Delta X \rangle$  decrease like  $\alpha^{-1}$ , just as in Section 4 (illustrated in



**Figure 20.** Correlation between sizes and forward/ backward waiting times: Pearson coefficient,  $r$ , as a function of the control parameter  $\alpha$ , for separable  $\eta$  with  $\varphi(x) = x^\delta$ . For each value of  $\alpha$ , we sample  $\Delta t$  and  $\Delta X$  over the course of a single simulation of  $N = 10^6$  glitches, after the system reaches stationarity. We then calculate the pair-wise correlation between the two sets for each  $\alpha$ . Parameter:  $\delta = 7$ .

Figure 4). Rapid spin down corresponds to  $\alpha \ll \delta$  and hence  $\langle \Delta X \rangle \approx (\delta + 2)^{-1}$ . This equals the mean glitch size,

$$\int_0^1 d(\Delta x) \eta(1 - \Delta x|1)\Delta x = \frac{1}{\delta + 2}, \quad (51)$$

assuming  $X_e \approx 1$ , as in (20).

#### 6.4.4 Size-waiting time correlation for $\varphi(x) = x^\delta$

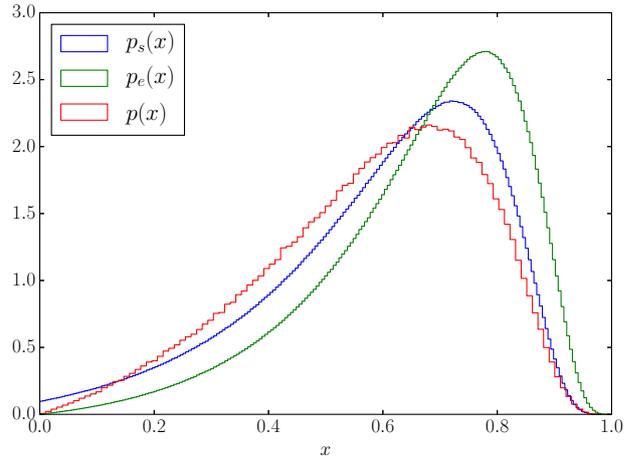
Figure 20 displays the correlation between glitch sizes and forward and backward waiting-times for  $\varphi(x) = x^\delta$ . The figure is constructed from Monte-Carlo output, but it is also possible to compute the correlations from the covariance analytically (outside the scope of this work). We note that the presence of a correlation depends strongly on  $\alpha$ , even though the PDFs  $p(\Delta t)$  and  $p(\Delta x)$  are identical, when  $\eta$  is separable as in Figure 19. In other words, the PDFs are equal, because the system self-adjusts, not because glitch sizes are correlated with forward waiting times, as in the  $\alpha \ll \alpha_c(\beta)$  regime in Section 4. The explanation for the qualitative behaviour seen in Figure 20 is the same as for Figure 13.

### 6.5 Lag-dependent spin down

For the sake of completeness, we extend the results of Section 6 to the case of a general lag-dependent spin-down torque in Appendix E.

## 7 SAMPLING SIZE BIAS

We now consider more carefully how we sample the different random variables used throughout this paper:  $X$ ,  $N$ ,  $\Delta t$ ,  $\Delta\Omega_c$  ( $\propto \Delta X$ ),  $X_e$  and  $X_s$ . Both  $X$  and  $N$  differ fundamentally from the other variables, because they are defined



**Figure 21.** Stationary PDFs  $p_s(x)$ ,  $p_e(x)$  and  $p(x)$  for the  $\lambda$  and power-law  $\eta$  used throughout Section 4. Both  $p_e(x)$  and  $p_s(x)$  are sampled from a single simulation of  $N = 10^8$  glitches after stationarity. We obtain  $p(x)$  by sampling an ensemble of  $S = 10^6$  simulations at  $t = 8$ . Parameters:  $\beta = 10^{-2}$ ,  $\alpha = 5$ .

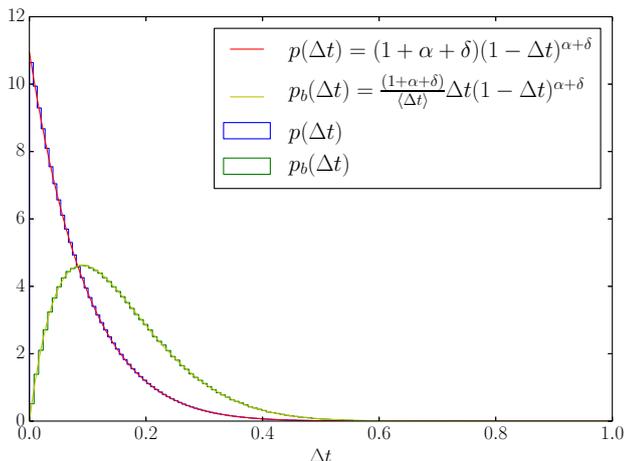
continuously for all  $t$ . The other variables are defined at discrete instants, when  $N(t)$  jumps by one due to a glitch. We therefore sample  $X(t)$  and  $N(t)$  by running an ensemble of simulations and observe each simulation at some time  $t_0$ . The other variables are sampled at every glitch in a single simulation once it reaches stationarity.

As an example of the subtle consequences of sampling bias, consider the ordering of the first moments  $\langle X \rangle$ ,  $\langle X_s \rangle$ , and  $\langle X_e \rangle$ . Figure 21 shows  $p(x)$ ,  $p_e(x)$  and  $p_s(x)$  for the non-separable power-law  $\eta$  (18) and (19) from Section 4. Note that  $p_s(x)$  is skewed toward larger values than  $p(x)$ . By integrating the curves, we find  $\langle X_s \rangle = 0.600$ ,  $\langle X \rangle = 0.579$ ,  $\langle X_e \rangle = 0.667$ , and hence  $\langle X \rangle < \langle X_s \rangle < \langle X_e \rangle$ . This seems wrong:  $X_s$ , the value of  $X$  immediately after a glitch, ought to be smaller, on average, than  $X$  measured at some arbitrary time  $t$ . However,  $X$  is sampled from an ensemble at fixed time  $t_0$ , whereas  $X_s$  and  $X_e$  are drawn from a different ensemble (every glitch event in a single simulation). If we observe the system at time  $t$ , we are more likely to land in a longer inter-glitch time interval than a smaller one, in which the (biased)  $X_s$  and  $X_e$  are spaced farther apart.

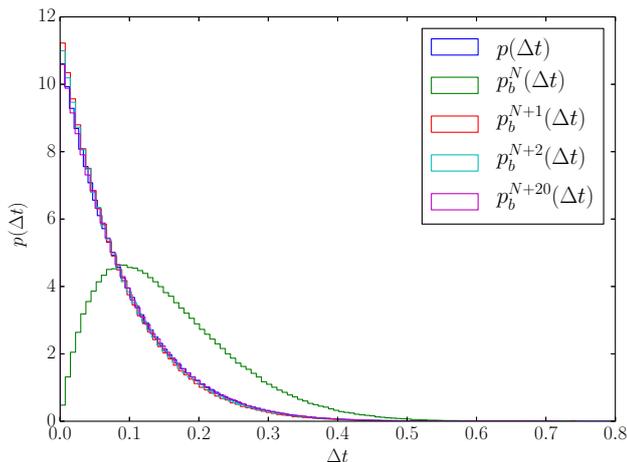
Sampling bias has an important observable consequence, even though  $p_s$  and  $p_e$  cannot be observed. We can obtain an ensemble distribution of waiting times by choosing a time  $t_0 \gg 1$  and measuring the length of the interval  $T^{(i+1)} - T^{(i)}$  for every simulation in the ensemble, where  $T^{(i)}$  is the time of the glitch immediately before  $t_0$ . We denote the interval of these lengths  $p_b(\Delta t)$ . For each simulation in the ensemble, the waiting-time intervals are distributed as  $p(\Delta t)$ , and we are proportionally more likely to randomly land in an interval in proportion to its length  $\Delta t$  (Cox 1970). The biased distribution  $p_b(\Delta t)$  is therefore

$$p_b(\Delta t) = \frac{p(\Delta t)\Delta t}{\langle \Delta t \rangle}, \quad (52)$$

where  $p(\Delta t)$  is the unbiased, “true” waiting time distribution obtained by sampling  $\Delta t$  over the course of a single simulation, and  $\langle \Delta t \rangle$ , the mean (unbiased) waiting time, is



**Figure 22.** The biased and unbiased waiting-time PDFs,  $p_b(\Delta t)$  and  $p(\Delta t)$ .  $p(\Delta t)$  is obtained by sampling  $\Delta t$  after the system reaches stationarity from a single simulation of  $N = 10^8$  glitches.  $p_b(\Delta t)$  is obtained by observing an ensemble of  $S = 10^6$  simulations at time  $t_0 = 8$  and sampling the  $\Delta t$  in each simulation that  $t_0$  falls in. Parameters:  $\delta = 7$ ,  $\alpha = 3$ .



**Figure 23.** The effect of sampling bias for waiting times. The unbiased waiting time PDF  $p(\Delta t)$  (blue curve) is obtained by sampling  $\Delta t$  after the system reaches stationarity from a single simulation of  $N = 10^8$  glitches. The PDFs  $p_b^{N+k}(\Delta t)$  are obtained by observing an ensemble of simulations at  $t_0 = 8$ , and sampling the width of the inter-glitch interval  $k$  glitches later. For  $k = 0$  (green curve), we sample the width of the interval we land in. For  $k = 1$  (red curve),  $k = 2$  (cyan curve) and  $k = 20$  (violet curve), we sample the length of the interval one, two and 20 glitches later respectively. Parameters:  $\delta = 7$ ,  $\alpha = 3$ .

required for normalisation. In Figure 22, we plot  $p(\Delta t)$  and  $p_b(\Delta t)$  from simulations for separable  $\eta$  with  $\varphi(x) = x^\delta$ , along with the analytic results. The agreement is excellent.

How does biasing affect pulsar timing measurements? Figure 23 provides an indication of the answer. First, we obtain  $p(\Delta t)$  by sampling  $\Delta t$  from a single simulation of  $N = 10^8$  glitches (blue curve). The other curves are obtained by running an ensemble of  $S = 10^6$  simulations, and sampling at a fixed time  $t_0 = 8$ . Let  $p_b^N(\Delta t)$  (green curve) be the PDF of  $T^{(N+1)} - T^{(N)}$ , where  $T^{(N)}$  is the time of

the last glitch before  $t_0$ . This is just the biased distribution  $p_b(\Delta t)$  plotted in Figure 22. Similarly,  $p_b^{N+1}$  is the PDF of  $T^{(N+2)} - T^{(N+1)}$ ,  $p_b^{N+2}$  (cyan curve) is the PDF of lengths of the interval immediately after that, and  $p_b^{N+20}$  (violet curve) is the PDF of the length of the interval 20 glitches after  $t_0$ . Only  $p_b^N(\Delta t)$  differs substantially  $p(\Delta t)$ . The PDFs  $p_b^{N+1}$  and  $p_b^{N+2}$  are slightly more skewed toward smaller  $\Delta t$ , with a larger intercept and sharper drop-off. This is a result of having a state-dependent Poisson rate, so that after a large interval, the next interval is slightly more likely to be smaller. By comparison, there is no visible difference between  $p_b^{N+20}(\Delta t)$  and  $p(\Delta t)$ .

The implications for pulsar measurements are as follows. If a glitch has not yet been observed, we are essentially making a biased measurement by looking at the pulsar at a random time and are more likely to be in a longer inter-glitch interval. Any attempt to place an upper bound on the average glitch rate by considering how long the pulsar has been observed without glitching must therefore take biasing into account. We do not know when the previous glitch occurred, so we do not measure the width of the biased interval [with PDF  $p_b^N(\Delta t)$  in Figure 23]. After the first glitch, the PDF of the time to the next glitch is described, qualitatively, by  $p_b^{N+1}$  in Figure 23, which closely approximates the unbiased waiting time PDF  $p(\Delta t)$ .<sup>8</sup> It is worth examining pulsars with a large number of glitches to determine whether the first observed inter-glitch waiting time interval is different, on average, from the following ones.

## 8 CONCLUSIONS

In this paper, we present a mean-field vortex avalanche model for pulsar glitches, in which the crust-superfluid lag fluctuates according to a state-dependent Poisson process. Using a conditional jump distribution inspired by results of published GPE simulations, we predict the observed waiting-time and size statistics, as a function of two key physical parameters:  $\beta$ , which sets maximum fraction of stress that can be released in a glitch, and  $\alpha$ , which is proportional to the avalanche rate at a reference lag divided by the electromagnetic spin-down rate. We find empirically that the system evolves to equilibrium. In the fast spin-down regime  $\alpha \lesssim \alpha_c(\beta)$ , the waiting-time PDF is approximately exponential, with an upper cut-off that scales as  $\alpha^{-1}$ , while the size PDF is power-law-like over  $-\log[\beta]$  decades. In the slow spin-down regime  $\alpha \gtrsim \alpha_c(\beta)$ , the waiting-time PDF approximately equals the glitch size PDF due to correlation between sizes and forward waiting times that is not observed in most pulsars. If the avalanche size distribution is sufficiently broad, spanning at least one decade, the model does not predict a significant correlation between sizes and backward waiting times, consistent with observations.

In order to study the intermediate regime  $\alpha \approx \alpha_c(\beta)$  where feedback between the avalanche rate and sizes is strongest, we develop an analytic solution in the special case

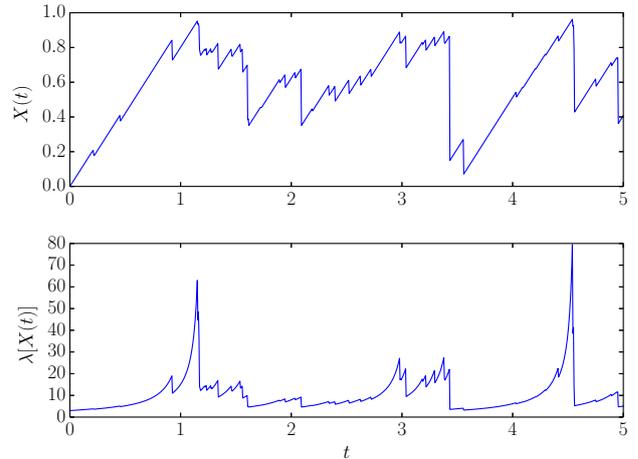
<sup>8</sup> Biasing occurs also in renewal processes (including the constant-rate Poisson process), which have inter-glitch intervals that are independent and identically distributed. However, in this case, one has  $p_b^{N+k}(\Delta t) = p(\Delta t)$  for all  $k > 0$ , where  $k$  is an integer.

where the conditional jump distribution  $\eta(x|y)$  is separable, and discuss its properties. We derive analytic expressions for the observed waiting-time and size PDFs. We find that, when  $\eta(x|y)$  is separable, the sizes and waiting times have the same distribution, a result that does not hold in the non-separable case in general. We also find qualitative similarities between the non-separable and separable cases, including how average sizes, waiting times and size-waiting-time correlations depend on key physical parameters.

An important consequence of the model is that the avalanche rate fluctuates over a wide range, even when the stochastic process is stationary. Figure 24 displays a representative evolution of  $\lambda[X(t)]$  with  $\alpha = 3$  and  $\beta = 10^{-2}$ . In the stationary regime  $t \gtrsim 1$ ,  $\lambda[X(t)]$  varies between  $3\lambda_0$  and  $80\lambda_0$ , when dimensions are restored. This is counter-intuitive:  $X$  fluctuates by a maximum of  $\approx 30\%$  from the mean, but  $\lambda$  varies by much more due to the asymptote in the avalanche rate as  $X \rightarrow 1$ . However,  $\lambda[X(t)]$  is not astrophysically observable. The question therefore arises: do fluctuations in  $\lambda[X(t)]$  have observable consequences? In Section 4, we find  $p(\Delta t) = \langle \lambda \rangle \exp(-\langle \lambda \rangle \Delta t)$  for  $\alpha \lesssim \alpha_c(\beta)$ . This is the same PDF as for a constant-rate Poisson process with rate  $\langle \lambda \rangle$ . Hence  $p(\Delta t)$  alone cannot distinguish between a constant-rate Poisson process and the state-dependent rate considered in this model for  $\alpha \lesssim \alpha_c(\beta)$ .

A recent analysis of 45 years of glitch data from the Crab (Lyne et al. 2015) found an unusual degree of clustering of 20 glitch events between June 1996 and August 2006. The clustering does not affect  $p(\Delta t)$ , which by construction records the time between consecutive glitch events and is not sensitive to the location of groups of glitches. Considering only glitches from 1996 until 2006, Lyne et al. (2015) sampled the time interval between each glitch and every other glitch occurring after it. They compared the distribution obtained to simulations of a constant-rate Poisson process and found significant disagreement (see Figure 6 of the latter paper). It is worth investigating whether the inclusion of a state-dependent rate, as done in this paper, permits greater clustering than a constant-rate Poisson process. In Figure 24, we see that, qualitatively, the events tend to be clustered near a series of small jumps, when  $X$  is large. This question deserves to be examined quantitatively in future work, using mathematical tools from the study of spatial point processes. We also investigate the subtle effect of sampling bias (“bus-stop paradox”) on observed waiting-time PDFs in Section 7.

Another important recent discovery is that the Crab appears to have a minimum glitch size above detection limits; the number of glitches with  $\Delta\nu \leq 0.05 \mu\text{Hz}$  is substantially below what is expected for a power law (c.f. a maximum observed glitch size of  $\Delta\nu = 1.48 \mu\text{Hz}$ ) (Espinoza et al. 2014). More data are needed to confirm the existence of a minimum  $\Delta\nu$  and search for it in other pulsars. Nevertheless, it is easily accommodated by the model in this paper, if we set  $\beta \approx 0.05/1.48$ . We emphasise strongly that the model does not provide any physical insight concerning the existence of a minimum  $\Delta\nu$  and the reason behind the implied value of  $\beta$ ; it simply accommodates the phenomenon in a natural way. Again, more data and enhanced analysis tools are needed to search for a minimum  $\Delta\nu$  in other pulsars. If these future experiments find that  $\beta$  takes on a common value across a number of objects, they will motivate future first-principles



**Figure 24.** Rate variability: representative time series of the crust-superfluid lag  $X(t)$  (top panel) and the resulting evolution of the instantaneous avalanche rate  $\lambda[X(t)]$  (bottom panel), during the approach to stationarity and beyond. Parameters:  $\alpha = 5$ ,  $\beta = 10^{-2}$ .

theoretical studies of vortex avalanches, e.g. with GPE simulations (Warszawski & Melatos 2011), to understand why the scale-invariant character of the avalanche process breaks down at  $\Delta\nu \approx 0.05 \mu\text{Hz}$ .

We emphasise in closing that comparisons with astrophysical data are restricted by the small sample sizes available at present. To date, there are 482 events in 168 objects recorded in the Jodrell Bank catalogue (see footnote 1). Of these objects, seven have glitched between 10 and 35 times. More data continue to flow from ongoing radio telescope timing campaigns.

## ACKNOWLEDGEMENTS

The authors thank Nathan Ross from the University of Melbourne for a helpful discussion on random jump processes.

Parts of this research were conducted by the Australian Research Council Centre of Excellence for Gravitational Wave Discovery (OzGrav) through project number CE170100004.

## APPENDIX A: PROOF OF $\langle \lambda \rangle^{-1} = \langle \Delta t \rangle$

In this appendix, we prove the intuitive result that under stationary conditions the ensemble average of  $\lambda(X)$  equals the reciprocal of the mean waiting-time interval.

As defined in (31),  $\langle \lambda \rangle$  is the ensemble expectation value of  $\lambda$ , obtained by sampling an ensemble of glitch sequences at some large time  $t$ , once each sequence is stationary. The mean waiting-time interval is given by

$$\langle \Delta t \rangle = \int_0^1 d(\Delta t) p(\Delta t) \Delta t, \quad (\text{A1})$$

where  $p(\Delta t)$ , the PDF of inter-glitch waiting times, is obtained by sampling the intervals between a large number of

glitches drawn from a single sequence. Inserting (34) into (A1), we obtain

$$\langle \Delta t \rangle = \int_0^1 d(\Delta t) \Delta t \int_0^{1-\Delta t} dy p_s(y) \lambda(y+\Delta t) e^{-\Lambda(y+\Delta t)} e^{\Lambda(y)}. \quad (\text{A2})$$

Noting  $\lambda(x) = \Lambda'(x)$ , we can write

$$\lambda(y+\Delta t) e^{-\Lambda(y+\Delta t)} = -\frac{\partial}{\partial(\Delta t)} e^{-\Lambda(y+\Delta t)} \quad (\text{A3})$$

and hence

$$\langle \Delta t \rangle = -\int_0^1 d(\Delta t) \Delta t \int_0^{1-\Delta t} dy (\Delta t) e^{-\Lambda(y+\Delta t)} p_s(y) e^{\Lambda(y)} \quad (\text{A4})$$

$$= -\int_0^1 dy p_s(y) e^{\Lambda(y)} \int_0^{1-y} d(\Delta t) \Delta t (\Delta t) e^{-\Lambda(y+\Delta t)}. \quad (\text{A5})$$

Integrating the inner integral by parts and noting that we have  $e^{-\Lambda(y+\Delta t)} \leq 1$  and  $e^{-\Lambda(y+\Delta t)} \rightarrow 0$  as  $\Delta t \rightarrow 1-y$  (from the divergence of  $\Lambda(x)$  as  $x \rightarrow 1$  from below), we find

$$\langle \Delta t \rangle = \int_0^1 dy p_s(y) e^{\Lambda(y)} \int_0^{1-y} d(\Delta t) e^{-\Lambda(y+\Delta t)} \quad (\text{A6})$$

$$= \int_0^1 dy p_s(y) e^{\Lambda(y)} \int_y^1 d\xi e^{-\Lambda(\xi)} \quad (\text{A7})$$

$$= \int_0^1 d\xi e^{-\Lambda(\xi)} \int_0^\xi dy p_s(y) e^{\Lambda(y)}. \quad (\text{A8})$$

However, from equation (37) we have

$$\int_0^\xi dy p_s(y) e^{\Lambda(y)} = \frac{e^{\Lambda(\xi)} p_e(\xi)}{\lambda(\xi)} \quad (\text{A9})$$

and hence

$$\langle \Delta t \rangle = \int_0^1 d\xi \frac{p_e(\xi)}{\lambda(\xi)}. \quad (\text{A10})$$

Finally, inserting  $\langle \lambda \rangle p_e(x) = \lambda(x) p(x)$  from (40) into (A10), we obtain

$$\langle \Delta t \rangle = \int_0^1 d\xi \frac{p(\xi)}{\langle \lambda \rangle} \quad (\text{A11})$$

$$= \frac{1}{\langle \lambda \rangle}. \quad (\text{A12})$$

## APPENDIX B: GENERAL PROPERTIES OF

$p_s(x)$ ,  $p_e(x)$ ,  $p(x)$

In this appendix, we prove several general properties of  $p_s$ ,  $p_e$  and  $p$ , as well as the relations between these three PDFs. We also prove the equality (in dimensionless units) of the mean avalanche waiting time and size.

We define the integrated avalanche rate  $\Lambda(x)$  in such a way as to satisfy  $\Lambda'(x) = \lambda(x) \geq 0$ ,  $\Lambda(0) = 0$ , and  $\Lambda(x) \rightarrow \infty$  as  $x \rightarrow 1$  from below. If we write  $p_e(x) = \lambda(x)g(x)$ , then the divergence of  $\Lambda(x)$  as  $x \rightarrow 1$  enforces the condition

$g(1) = 0$  in order to maintain the integrability of  $p_e(x)$ . In addition, we have  $p_e(0) = 0$  from equation (37). Eliminating  $p_e$  in favour of  $g$  in equation (39), multiplying by  $e^{\Lambda(x)}$  and differentiating, we obtain

$$g'(x) + \lambda(x)g(x) = \int_x^1 dz g(z) \lambda(z) \eta(x|z). \quad (\text{B1})$$

Hence  $g$  satisfies the same integrodifferential equation (28) that  $p$  does, and the same homogeneous boundary conditions (30). If the solution of the integrodifferential equation with these boundary conditions is unique up to a normalization constant, then we are forced to conclude that  $g$  is a constant multiple of  $p$ , and it follows that one has

$$p_e(x) = \frac{\lambda(x)p(x)}{\langle \lambda \rangle}. \quad (\text{B2})$$

Daly & Porporato (2007) found (B2) by arguing that  $p_e(x)$  ‘‘can be obtained by considering the probability of being in  $x+dx$  and (independently) jumping during a time interval  $dt$  from that level, i.e.,  $p(x)\lambda(x)dxdt$ , and then normalising it by the total probability of jumping during the interval  $dt$ ’’ (verbatim quote).

Combining (36), (37) and (B2), we find that  $p(x)$  and  $p_s(x)$  are related via

$$p_s(x) = \frac{1}{\langle \lambda \rangle} \int_x^1 dy \lambda(y) \eta(x|y) p(y), \quad (\text{B3})$$

and

$$p(x) = \langle \lambda \rangle e^{-\Lambda(x)} \int_0^x dy p_s(y) e^{\Lambda(y)}. \quad (\text{B4})$$

From equations (B2), (B3) and the master equation (28), we arrive at the important result

$$\frac{p'(x)}{\langle \lambda \rangle} = p_s(x) - p_e(x). \quad (\text{B5})$$

Finally we establish the equality of the mean avalanche size and the mean waiting time between glitches. To do this we multiply the master equation (28) by  $x$  and integrate over  $0 < x < 1$ . After a little algebra, including integration by parts of  $xp'(x)$ , an interchange of orders of integration in the double integral, and the renaming of one of the dummy integration variables, we find

$$1 = \int_0^1 dy \lambda(y) p(y) \left[ y - \int_0^y dx x \eta(x|y) \right]. \quad (\text{B6})$$

We can rewrite (B6) as

$$1 = \int_0^1 dy \lambda(y) p(y) \int_0^y dx (y-x) \eta(x|y). \quad (\text{B7})$$

Because  $y-x$  corresponds to the avalanche size when the glitch occurs at  $X_e = y$ , the inner integral equals the mean of  $\Delta X$ , conditional on the system glitching at  $X_e = y$ , which we denote by  $\langle \Delta X | X_e = y \rangle$ . Combining with the relation (B2) between  $p$  and  $p_e$ , we find

$$1 = \langle \lambda \rangle \int_0^1 dy p_e(y) \langle \Delta X | X_e = y \rangle \quad (\text{B8})$$

$$= \langle \lambda \rangle \langle \Delta X \rangle. \quad (\text{B9})$$

Finally, inserting  $\langle \Delta t \rangle = \langle \lambda \rangle^{-1}$  from (A11), we obtain

$$\langle \Delta t \rangle = \langle \Delta X \rangle. \quad (\text{B10})$$

Equation (B10) proves, for general  $\lambda$  and  $\eta$ , the equality of the mean dimensionless waiting times and sizes. This matches the results from numerical simulations for the special choices of  $\lambda$  and  $\eta$  in Section 4, illustrated in Figure 4.

By integrating (B5) over  $0 \leq x \leq 1$  and inserting (B9), we confirm the intuitive result

$$\langle X_s \rangle = \langle X_e \rangle - \langle \Delta X \rangle. \quad (\text{B11})$$

### APPENDIX C: SOLUTION OF THE MASTER EQUATION WITH SEPARABLE $\eta(x|y)$

In this appendix we show that the stationary master equation (28) with a separable conditional jump distribution given by (41) and (42) has the solution (43) and (44), provided that

$$\Lambda(x) \rightarrow \infty \text{ as } x \rightarrow 1 \text{ from below,} \quad (\text{C1})$$

with

$$\Lambda(x) = \int_0^x ds \lambda(s). \quad (\text{C2})$$

We also establish that under the condition (C1) this solution is the only non-negative solution that satisfies the boundary conditions

$$p(0) = 0 = p(1). \quad (\text{C3})$$

Let us begin by rearranging the stationary master equation (28) into the form

$$p'(x) + \lambda(x)p(x) = \int_x^1 dy p(y) \lambda(y) \eta(x|y) \quad (\text{C4})$$

and define

$$p_*(x) = \Phi(x) e^{-\Lambda(x)}. \quad (\text{C5})$$

The limits  $\Phi(x) \rightarrow 0$  and  $\Lambda(x) \rightarrow 0$  as  $x \rightarrow 0$  imply  $p_*(0) = 0$ . The condition (C1) also implies

$$0 = \lim_{x \rightarrow 1} e^{-\Lambda(x)}. \quad (\text{C6})$$

As  $\Phi(1)$  is finite, we have  $p_*(1) = 0$ .

It is easily verified by direct substitution, integration by parts and use of the result (C6), that  $Cp_*(x)$  is a solution of the integrodifferential equation (C4) and boundary conditions (C3) for any constant  $C$  given  $\eta(x|y) = \varphi(x)/\Phi(y)$ .

Choosing  $C$  according to (44) returns the correct normalisation to make  $p(x)$  a PDF.

More generally, we seek a solution of the form

$$p(x) = r(x)p_*(x), \quad (\text{C7})$$

where  $r(x)$  is non-negative. We show that the boundary conditions (C3) imply  $r'(x) = 0$  for  $0 < x < 1$ , making  $r(x) = \text{constant}$  the only solution.

Inserting the trial solution (C7) into (C4), we find after a little algebra

$$\begin{aligned} [r'(x)\Phi(x) + r(x)\varphi(x)]e^{-\Lambda(x)} &= \varphi(x) \int_x^1 dy r(y) \lambda(y) e^{-\Lambda(y)} \\ &= -\varphi(x) \int_x^1 dy r(y) y \left[ e^{-\Lambda(y)} \right]. \end{aligned} \quad (\text{C8})$$

The requirement  $p(x) \rightarrow 0$  as  $x \rightarrow 1$  and the observation that  $\Phi(x)$  is nonzero in this limit imply  $r(y)e^{-\Lambda(y)} \rightarrow 0$  as  $y \rightarrow 1$ . Consequently on integrating (C8) by parts we find

$$r'(x)\Phi(x)e^{-\Lambda(x)} = \varphi(x) \int_x^1 dy r'(y)e^{-\Lambda(y)}. \quad (\text{C10})$$

We can rewrite (C10) as

$$0 = x \left[ \Phi(x) \int_x^1 dy r'(y)e^{-\Lambda(y)} \right], \quad (\text{C11})$$

so that we have

$$A = \Phi(x) \int_x^1 dy r'(y)e^{-\Lambda(y)}, \quad (\text{C12})$$

for some constant  $A$  in the domain  $0 < x < 1$ . If we are able to conclude that  $A = 0$  holds, it follows immediately that we have  $r'(y) = 0$  for  $0 < y < 1$ .

If  $A$  is nonzero, then the integral in equation (C12) diverges as  $x \rightarrow 0$  from above, given  $\Phi(x) \rightarrow 0$  as  $x \rightarrow 0$ . Hence for all fixed  $a$  with  $0 < a < 1$  we have

$$A = \lim_{x \rightarrow 0} \Phi(x) \int_x^a dy r'(y)e^{-\Lambda(y)}. \quad (\text{C13})$$

However, we also have  $\Lambda(y) \rightarrow 0$  and  $e^{-\Lambda(y)} \rightarrow 1$  as  $y \rightarrow 0$ , implying

$$A = \lim_{x \rightarrow 0} \Phi(x) \int_x^a dy r'(y), \quad (\text{C14})$$

that is,

$$r(x) \sim -\frac{A}{\Phi(x)}. \quad (\text{C15})$$

To stop the solution becoming negative, we require  $A \leq 0$ . Moreover, as  $x \rightarrow 0$  from above we have

$$p(x) = r(x)\Phi(x)e^{-\Lambda(x)} \rightarrow |A|, \quad (\text{C16})$$

and the boundary condition  $p(0) = 0$  implies  $A = 0$ . Thus we confirm that one has  $r'(x) = 0$  throughout the interval, as previously claimed.

#### APPENDIX D: ANALYTIC PDF FOR SEPARABLE, POLYNOMIAL $\eta(x|y)$

For a separable conditional jump distribution of the form (41), with  $\varphi(x) = x^\delta$ , where  $\delta$  is a non-negative constant, the stationary PDFs given by (43), (45) and (46) equate to

$$p(x) = Cx^{\delta+1}(1-x)^\alpha, \quad (\text{D1})$$

$$p_s(x) = \frac{C}{\langle \lambda \rangle} x^\delta (1-x)^\alpha, \quad (\text{D2})$$

and

$$p_e(x) = \frac{C}{\langle \lambda \rangle} x^{\delta+1} (1-x)^{\alpha-1}, \quad (\text{D3})$$

with  $C = \Gamma(\alpha + \delta + 3) [\Gamma(\alpha + 1)\Gamma(\delta + 2)]^{-1}$ , where  $\Gamma$  denotes the usual gamma function. Equations (D2) and (D3) can be used to integrate (47) and (48) analytically to obtain the observed waiting-time PDF,

$$p(\Delta t) = (1 + \alpha + \delta)(1 - \Delta t)^{\alpha+\delta}. \quad (\text{D4})$$

Equation (D4) is equal to the size PDF  $p(\Delta x)$ , as  $\eta(x|y)$  is separable. We find from appropriate integrals based on (D1) and the identities (A12) and (B10) that the following moments satisfy

$$\langle \Delta X \rangle = \langle \Delta t \rangle = \langle \lambda \rangle^{-1} = \frac{1}{\alpha + \delta + 2}. \quad (\text{D5})$$

The results (D1)–(D4) are plotted and discussed in Section 6.4. The choice  $\varphi(x) = x^\delta$  favours small jumps, consistent with the power-law avalanche sizes observed in GPE simulations, and serves as a semi-quantitative approximation to the nonseparable  $\eta(x|y)$  given by (18) and (19) and used in the Monte Carlo simulations in Sections 4 and 5.

#### APPENDIX E: GENERAL DETERMINISTIC SPIN DOWN

For the sake of completeness, we extend Section 6 to the case of a general, lag-dependent deterministic spin-down law. In this case, (7) no longer holds, and the equation of motion is instead (Daly & Porporato 2007)

$$X(t)t = f(X) - \sum_{i=1}^{N(t)} \Delta X^{(i)} \delta [t - T^{(i)}], \quad (\text{E1})$$

where  $f(X)$  is the instantaneous spin-down rate, and  $\Delta X^{(i)}$  and  $T^{(i)}$  are the size and epoch of the  $i$ -th avalanche, respectively. Equation (E1) reduces to (7) for the special case  $f(X) = 1$  relevant to pulsars. We assume that the instantaneous spin-down rate does not depend explicitly on time.

The stationary master equation associated with (E1) takes the form [cf. (28)]

$$0 = -x[p(x)f(x)] - \lambda(x)p(x) + \int_x^1 dy p(y)\lambda(y)\eta(x|y). \quad (\text{E2})$$

#### E1 Observable PDFs

We obtain the waiting-time and size PDFs as follows. For the case of trivial forcing,  $f(X) = 1$ , when a glitch at time  $t$  sets the lag to  $X(t)$ , the lag evolves as  $X(t+\Delta t) = X(t) + \Delta t$  until the next glitch. For general  $f(X)$ , the equation of motion in the deterministic phase between two glitches is given by the first-order differential equation

$$Xt = f(X), \quad (\text{E3})$$

which yields

$$\Delta t = \int_{X(t)}^{X(t+\Delta t)} \frac{d\xi}{f(\xi)}. \quad (\text{E4})$$

Equation (E4) must be inverted to obtain an expression for  $X(t + \Delta t)$  as a function of  $X(t)$  and  $\Delta t$ . We denote this as

$$X(t + \Delta t) = g[\Delta t, X(t)]. \quad (\text{E5})$$

With this definition, inverting the integration terminals in (E4) yields the relation  $X(t) = g[-\Delta t, X(t + \Delta t)]$ . We stress that equations (E3)–(E5) hold only for the deterministic evolution between two glitches. As an example, if we choose  $f(X) = 1 - X$ , so that the forcing slows to zero as we approach the critical lag, (E5) becomes

$$g[\Delta t, X(t)] = 1 - [1 - X(t)]e^{-\Delta t}. \quad (\text{E6})$$

Now let  $X(t)$  equal  $y$  at the start of a deterministic phase of motion, immediately after a glitch at time  $t$ . The conditional waiting-time PDF (13) for the time until the next glitch is

$$p[\Delta t | X(t) = y] = \lambda[X(t + \Delta t)] e^{-\int_t^{t+\Delta t} dt' \lambda[X(t')]} \quad (\text{E7})$$

Changing the variable of integration in (E7) through (E3) and inserting the relation  $X(t + \Delta t) = g(\Delta t, y)$  from (E5), we obtain

$$p[\Delta t | X(t) = y] = \lambda[g(\Delta t, y)] e^{-\Psi[g(\Delta t, y)]} e^{\Psi(y)}, \quad (\text{E8})$$

where  $\Psi(x)$  is defined as

$$\Psi(x) = \int_0^x d\xi \frac{\lambda(\xi)}{f(\xi)}. \quad (\text{E9})$$

At stationarity the value of  $X(t) = y$  immediately after a glitch has PDF  $p_s(y)$ . The observed waiting time PDF is therefore obtained by marginalising (E7) over  $y$ :

$$p(\Delta t) = \int_0^{g(-\Delta t, 1)} dy p_s(y) \lambda[g(\Delta t, y)] e^{-\Psi[g(\Delta t, y)]} e^{\Psi(y)}. \quad (\text{E10})$$

The upper terminal of (E10),  $g(-\Delta t, 1)$ , gives the largest value of  $X_s = y$ , such that the stress does not exceed the critical lag, if the system evolves for time  $\Delta t$  before the next glitch occurs. Daly & Porporato (2007) took an alternative route to obtain the waiting-time PDF, by solving the stationary master equation (E2) in the absence of jumps, and

obtaining a survivor equation for the time until the total probability vanishes. The expression for the avalanche size PDF is unchanged from (35):

$$p(\Delta x) = \int_{\Delta x}^1 dy p_e(y) \eta(y - \Delta x | y). \quad (\text{E11})$$

## E2 Lag PDFs

We now derive relations between  $p_s(x)$ ,  $p_e(x)$  and  $p(x)$ .

The generalisation of equation (36) is

$$p_e(x) = \frac{\lambda(x)}{f(x)} e^{-\Psi(x)} \int_0^x dy p_s(y) e^{\Psi(y)}. \quad (\text{E12})$$

Equations (37) and (40) remain unchanged:

$$p_s(x) = \int_x^1 dy p_e(y) \eta(x | y), \quad (\text{E13})$$

$$p_e(x) = \frac{p(x) \lambda(x)}{\langle \lambda \rangle}. \quad (\text{E14})$$

It follows from (E17)–(E14) that the time-independent master equation (E2) can be written as

$$\frac{1}{\langle \lambda \rangle} x [p(x) f(x)] = p_s(x) - p_e(x). \quad (\text{E15})$$

By a trivial extension of the argument after (B5) in Appendix B, the relation between the mean avalanche rate and size is now

$$\langle f \rangle = \langle \lambda \rangle \langle \Delta X \rangle, \quad (\text{E16})$$

where  $\langle f \rangle = \int_0^1 d\xi p(\xi) f(\xi)$  is the ensemble expectation value of the instantaneous spin-down rate.

## E3 Separable $\eta(x|y)$

If we now restrict attention to the special case where  $\eta$  is separable, as defined by (41) and (42), the master equation (E2) is satisfied by

$$p(x) = \frac{C \Phi(x) e^{-\Psi(x)}}{f(x)}, \quad (\text{E17})$$

where  $C$  is a normalisation constant,

$$C^{-1} = \int_0^1 dx \frac{\Phi(x) e^{-\Psi(x)}}{f(x)}. \quad (\text{E18})$$

The other stationary PDFs are given by

$$p_s(x) = \frac{C \varphi(x) e^{-\Psi(x)}}{\langle \lambda \rangle}, \quad (\text{E19})$$

$$p_e(x) = \frac{C \lambda(x) \Phi(x) e^{-\Psi(x)}}{\langle \lambda \rangle f(x)}. \quad (\text{E20})$$

We use (E6) and (E7) to obtain explicit expressions for the waiting-time and size PDFs:

$$p(\Delta t) = \frac{C}{\langle \lambda \rangle} \int_0^{g(-\Delta t, 1)} dy \frac{\lambda[g(\Delta t, y)]}{f(y)} \varphi(y) e^{-\Psi[g(\Delta t, y)]}, \quad (\text{E21})$$

$$p(\Delta x) = \frac{C}{\langle \lambda \rangle} \int_{\Delta x}^1 dy \frac{\lambda(y)}{f(y)} \varphi(y - \Delta x) e^{-\Psi(y)}. \quad (\text{E22})$$

## E4 Relating $\lambda(x)$ and $f(x)$

We conclude this section by noting the connection between  $\lambda(x)$  and  $f(x)$ . In (E2), we can substitute a solution of the form  $p(x) = q(x)/f(x)$  to obtain

$$0 = -q(x)x - \gamma(x)q(x) + \int_0^1 dy \eta(x|y) \gamma(y)q(y), \quad (\text{E23})$$

with  $\gamma(x) = \lambda(x)/f(x)$ . Referring to (27), we see that  $q(x)$  is itself the solution to a master equation with trivial deterministic evolution and an avalanche rate function  $\gamma(X) = \lambda(X)/f(X)$ .

What does this mean, physically? Consider the example  $\lambda(X) = \alpha/(1-X)$  from (11) with  $f(X) = 1$  from Section 4. We obtain the same equation for  $q(x)$ , using the substitution leading to (E23), from a constant rate avalanche process,  $\lambda(X) = 1$ , and deterministic forcing  $f(X) = (1-X)/\alpha$ . Although the avalanche rate is constant, a glitch is more likely to occur in the interval  $(x, x+dx)$  near the critical lag (where the forcing is slower) than far away, because the system spends more time there. In the limit  $X \rightarrow 1$  we have  $f(X) \rightarrow 0$ , meaning that near the critical lag the system is hardly spinning down at all. We stress that the PDFs  $p$  and  $q$  are the distributions of two completely different physical variables. All observable quantities, including avalanche sizes and inter-glitch waiting times, must be computed from  $p(x)$ . As an example of the observable differences, we note that  $f(X) = (1-X)/\alpha$  with  $\lambda = 1$  does not enforce a maximum cut-off of inter-glitch waiting times. In contrast, in Sections 3–6, we find  $\Delta t < 1$  always for  $\lambda(X) = \alpha/(1-X)$  and  $f(X) = 1$ .

## REFERENCES

- Alpar M. A., Cheng K. S., Pines D., 1989, *ApJ*, **346**, 823  
 Altshuler E., et al., 2004, *Phys. Rev. B*, **70**, 140505  
 Anderson P. W., Itoh N., 1975, *Nature*, **256**, 25  
 Andersson N., Comer G. L., 2006, *Classical and Quantum Gravity*, **23**, 5505  
 Andersson N., Glampedakis K., Ho W. C. G., Espinoza C. M., 2012, *Physical Review Letters*, **109**, 241103  
 Archibald R. F., et al., 2016, *ApJ*, **819**, L16  
 Ashton G., Prix R., Jones D. I., 2017, preprint, ([arXiv:1704.00742](https://arxiv.org/abs/1704.00742))  
 Avogadro P., Barranco F., Broglia R. A., Vigezzi E., 2008, *Nuclear Physics A*, **811**, 378  
 Bak P., Tang C., Wiesenfeld K., 1987, *Phys. Rev. Lett.*, **59**, 381  
 Basu S., 1999, PhD thesis, London School of Economics  
 Chamel N., 2012, *Phys. Rev. C*, **85**, 035801

- Cheng K. S., Pines D., Alpar M. A., Shaham J., 1988, *ApJ*, **330**, 835
- Chevalier E., 1993, *ApJ*, **414**, L113
- Cox D. R., 1955, *Journal of the Royal Statistical Society. Series B (Methodological)*, **17**, 129
- Cox D., 1970, *Renewal Theory*. Methuen science paperbacks, Methuen
- Daly E., Porporato A., 2006, *Phys. Rev. E*, **74**, 041112
- Daly E., Porporato A., 2007, *Phys. Rev. E*, **75**, 011119
- Dassios A., Jang J.-W., 2003, *Finance and Stochastics*, **7**, 73
- Dodson R., Lewis D., McCulloch P., 2007, *Ap&SS*, **308**, 585
- Donati P., Pizzochero P. M., 2004, *Nuclear Physics A*, **742**, 363
- Donati P., Pizzochero P. M., 2006, *Physics Letters B*, **640**, 74
- Epstein R. I., Baym G., 1992, *ApJ*, **387**, 276
- Espinoza C. M., Lyne A. G., Stappers B. W., Kramer M., 2011, *MNRAS*, **414**, 1679
- Espinoza C. M., Antonopoulou D., Stappers B. W., Watts A., Lyne A. G., 2014, *MNRAS*, **440**, 2755
- Field S., Witt J., Nori F., Ling X., 1995, *Physical Review Letters*, **74**, 1206
- Gardiner C. W., Anglin J. R., Fudge T. I. A., 2002, *Journal of Physics B: Atomic, Molecular and Optical Physics*, **35**, 1555
- Glampedakis K., Andersson N., 2009, *Physical Review Letters*, **102**, 141101
- Hänggi P., Talkner P., Borkovec M., 1990, *Reviews of Modern Physics*, **62**, 251
- Haskell B., 2016, *MNRAS*, **461**, L77
- Haskell B., Melatos A., 2015, *International Journal of Modern Physics D*, **24**, 1530008
- Haskell B., Melatos A., 2016, *MNRAS*, **461**, 2200
- Haskell B., Pizzochero P. M., Seveso S., 2013, *ApJ*, **764**, L25
- Hobbs G., Lyne A. G., Kramer M., Martin C. E., Jordan C., 2004, *Monthly Notices of the Royal Astronomical Society*, **353**, 1311
- Howitt G. A. W., Melatos A., Delaigle A., Hall P., 2017, Nonparametric estimation of the size and waiting time distributions of radio pulsar glitches
- Janssen G. H., Stappers B. W., 2006, *A&A*, **457**, 611
- Jensen H. J., 1998, *Self-organized criticality: emergent complex behavior in physical and biological systems*, 1 edn. Vol. 13, Cambridge University Press
- Jones P. B., 1991, *ApJ*, **373**, 208
- Jones P. B., 2003, *ApJ*, **595**, 342
- Kanazir M., Wheatland M. S., 2010, *Sol. Phys.*, **266**, 301
- Keith M. J., Shannon R. M., Johnston S., 2013, *MNRAS*, **432**, 3080
- Kerr M., Hobbs G., Johnston S., Shannon R. M., 2016, *MNRAS*, **455**, 1845
- Larson M. B., Link B., 2002, *Monthly Notices of the Royal Astronomical Society*, **333**, 613
- Link B., 2014, *ApJ*, **789**, 141
- Link B. K., Epstein R. I., 1991, *ApJ*, **373**, 592
- Link B., Epstein R. I., Baym G., 1993, *ApJ*, **403**, 285
- Link B., Epstein R. I., Lattimer J. M., 1999, *Physical Review Letters*, **83**, 3362
- Lyne A. G., Shemar S. L., Smith F. G., 2000, *MNRAS*, **315**, 534
- Lyne A. G., Jordan C. A., Graham-Smith F., Espinoza C. M., Stappers B. W., Weltevrede P., 2015, *MNRAS*, **446**, 857
- Manchester R. N., Hobbs G. B., Teoh A., Hobbs M., 2005, *AJ*, **129**, 1993
- Mastrano A., Melatos A., 2005, *MNRAS*, **361**, 927
- McCulloch P. M., Hamilton P. A., McConnell D., King E. A., 1990, *Nature*, **346**, 822
- McKenna J., Lyne A. G., 1990, *Nature*, **343**, 349
- Melatos A., 1997, *MNRAS*, **288**, 1049
- Melatos A., Peralta C., 2007, *ApJ*, **662**, L99
- Melatos A., Warszawski L., 2009, *ApJ*, **700**, 1524
- Melatos A., Peralta C., Wyithe J. S. B., 2008, *ApJ*, **672**, 1103
- Melatos A., Douglass J. A., Simula T. P., 2015, *ApJ*, **807**, 132
- Middleditch J., Marshall F. E., Wang Q. D., Gotthelf E. V., Zhang W., 2006, *ApJ*, **652**, 1531
- Peralta C. A., 2006, PhD thesis, School of Physics, University of Melbourne
- Peralta C., Melatos A., Giacobello M., Ooi A., 2006, *ApJ*, **651**, 1079
- Piekarewicz J., Fattoyev F. J., Horowitz C. J., 2014, *Phys. Rev. C*, **90**, 015803
- Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 2007, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, 3 edn. Cambridge University Press, New York, NY, USA
- Ramesh N. I., Onof C., Xie D., 2012, *Advances in Water Resources*, **45**, 58
- Ross S. M., 2006, *Introduction to Probability Models*, Ninth Edition. Academic Press, Inc., Orlando, FL, USA
- Roy J., Gupta Y., Lewandowski W., 2012, *MNRAS*, **424**, 2213
- Ruderman M., 1969, *Nature*, **223**, 597
- Ruderman M., 1991, *ApJ*, **382**, 587
- Ruderman M., Zhu T., Chen K., 1998, *ApJ*, **492**, 267
- Sedrakian A., 2005, *Phys. Rev. D*, **71**, 083003
- Seveso S., Pizzochero P. M., Grill F., Haskell B., 2016, *Monthly Notices of the Royal Astronomical Society*, **455**, 3952
- Shannon R. M., Lentati L. T., Kerr M., Johnston S., Hobbs G., Manchester R. N., 2016, *MNRAS*, **459**, 3104
- Sidery T., Passamonti A., Andersson N., 2010, *Monthly Notices of the Royal Astronomical Society*, **405**, 1061
- Smoluchowski R., 1970, *Physical Review Letters*, **24**, 923
- Sourie A., Chamel N., Novak J., Oertel M., 2017, *MNRAS*, **464**, 4641
- Takatsuka T., Tamagaki R., 1988, *Progress of Theoretical Physics*, **79**, 274
- Warszawski L., Melatos A., 2008, *MNRAS*, **390**, 175
- Warszawski L., Melatos A., 2011, *MNRAS*, **415**, 1611
- Warszawski L., Melatos A., 2013, *MNRAS*, **428**, 1911
- Warszawski L., Melatos A., Berloff N. G., 2012, *Phys. Rev. B*, **85**, 104503
- Wheatland M. S., 2008, *ApJ*, **679**, 1621
- Wheatland M. S., 2009, *Sol. Phys.*, **255**, 211
- Wheatland M. S., Glukhov S., 1998, *ApJ*, **494**, 858
- Yu M., et al., 2013, *MNRAS*, **429**, 688