

Statistical properties of a filtered Poisson process with additive random noise: distributions, correlations and moment estimation

A Theodorsen^{1,3}, O E Garcia¹ and M Rypdal²

¹ Department of Physics and Technology, UiT The Arctic University of Norway, NO-9037 Tromsø, Norway

² Department of Mathematics and Statistics, UiT The Arctic University of Norway, NO-9037 Tromsø, Norway

E-mail: audun.theodorsen@uit.no, odd.erik.garcia@uit.no and martin.rypdal@uit.no

Received 5 January 2017, revised 24 March 2017

Accepted for publication 27 March 2017

Published 19 April 2017



Abstract

Filtered Poisson processes are often used as reference models for intermittent fluctuations in physical systems. Such a process is here extended by adding a noise term, either as a purely additive term to the process or as a dynamical term in a stochastic differential equation. The lowest order moments, probability density function, auto-correlation function and power spectral density are derived and used to identify and compare the effects of the two different noise terms. Monte-Carlo studies of synthetic time series are used to investigate the accuracy of model parameter estimation and to identify methods for distinguishing the noise types. It is shown that the probability density function and the three lowest order moments provide accurate estimations of the model parameters, but are unable to separate the noise types. The auto-correlation function and the power spectral density also provide methods for estimating the model parameters, as well as being capable of identifying the noise type. The number of times the signal crosses a prescribed threshold level in the positive direction also promises to be able to differentiate the noise type.

Keywords: filtered Poisson process, shot noise process, reference model, Monte-Carlo estimation, scrape-off layer turbulence, plasma turbulence

(Some figures may appear in colour only in the online journal)

1. Introduction

Intermittent fluctuations are found in a variety of physical systems such as atmospheric winds [1], astrophysical plasmas [2, 3], fission chambers [4], diodes and electric circuits [5, 6] and magnetic confinement experiments [7–11], as well as in fields such as finance [12] and physiology [13]. In several such systems, treatments of intermittent effects as a superposition of random variables has been fruitful, see for instance [14, 15]. In this contribution, we will focus on a particular

reference model for intermittent fluctuations, the filtered Poisson process (FPP) (also called a generalized shot noise process) [4, 5, 12, 13]. This model consists of a superposition of uncorrelated pulses with a uniform pulse shape and randomly distributed pulse amplitudes, arriving according to a Poisson process [16, 17]. The FPP has been considered by e.g. [18–22]. In particular, the FPP with exponentially distributed pulse amplitudes and an exponential pulse shape can be shown to be Gamma distributed [12, 16–18].

This contribution is primarily motivated by turbulent flows in the scrape-off layer of magnetically confined plasmas. Evidence points towards these fluctuations being caused

³ Author to whom any correspondence should be addressed.

by filamentary structures with excess particles and heat compared to the ambient plasma moving radially outwards, transporting particles and heat through the scrape-off layer towards main chamber walls [8, 10, 11]. Recently, the statistical properties of the intermittent fluctuations have been revealed by exceptionally long data time series from numerous tokamak experiments, and exhibit similar behavior for a wide range of machine parameters [23–28]. The fluctuations have skewed and flattened probability distribution functions (PDFs) resembling Gamma distributions. Conditional averaging reveals that large-amplitude fluctuations have an exponential pulse shape, exponentially distributed maximum amplitudes and exponentially distributed waiting times between pulses, consistent with a Poisson process [23–28]. The auto-correlation functions and power spectral densities (PSDs) have a universal shape in accordance with the model for a range of plasma parameters and radial locations in the SOL [29, 30]. The underlying assumptions as well as the predictions of this model are thus in good agreement with experimental measurements.

While the Gamma distribution describes large-amplitude fluctuations in the scrape-off layer well, it often fails to describe fluctuations around the mean value of the signal. By adding an independent, normally distributed random variable to the process, the PDF of the resulting process is a convolution of a Gamma distribution and a normal distribution. This result has been shown to be in excellent agreement with probe measurements from the Alcator C-Mod and KSTAR tokamaks [26, 31]. Adding a new random variable to the process may also change other statistical properties of the process, such as its moments, auto-correlation function, PSD and the rate at which the process crosses a prescribed threshold level. These changes depend on the second-order statistics of the normally distributed process.

In this contribution, we will extend the reference FPP model by adding normally distributed noise in two different ways, either as a purely additive term to the process, modeling measurement noise or other processes unconnected to the dynamics of the FPP, or as a dynamical noise term in the stochastic differential equation for the reference model, resembling an Ornstein–Uhlenbeck process. We will mainly consider how these different noise terms affect the PDF of the resulting signal, its auto-correlation function and its PSD. Additionally, we will consider the rate at which the processes cross a certain threshold level in the positive direction. The goal of this contribution is to find methods for discriminating the two types of noise and to identify reliable methods for estimating the model parameters in a given realization of the process.

This contribution is organized as follows: in section 2, the pure FPP is considered. In section 3, the two types of noise are considered. The lowest order moments and the PDF of the FPP with additional noise are discussed in section 4, and the PSD and auto-correlation function of this process are discussed in section 5. In order to differentiate the types of noise and to compare different methods of parameter estimation, Monte-Carlo studies of synthetic data are presented in section 6. Section 7 concludes the contribution.

In appendix A, a list of symbols and the most important analytical results of this contribution are collected. Appendix B contains derivations relating to the PSDs and auto-correlation functions discussed.

2. Filtered Poisson process

In this section, we present the FPP to be analyzed in this contribution. This process is constructed as a super-position of K pulses arriving in a time interval $[0, T]$:

$$\Phi_K(t) = \sum_{k=1}^{K(T)} A_k \varphi\left(\frac{t - t_k}{\tau_d}\right), \quad (1)$$

where the pulse duration time τ_d is taken to be the same for all pulses. The pulse amplitudes A_k are taken to be exponentially distributed with mean value $\langle A \rangle$,

$$P_A(A; \langle A \rangle) = \frac{1}{\langle A \rangle} \exp\left(-\frac{A}{\langle A \rangle}\right), \quad (2)$$

where $A \geq 0$ and $\langle \bullet \rangle$ here and in the following indicates the average over all random variables.

As an idealization of a pulse with a fast rise and an exponential decay, we use the one-sided pulse form

$$\varphi(\eta) = \Theta(\eta) \exp(-\eta), \quad (3)$$

where Θ is the Heaviside step function and η is a dimensionless variable. Compared to a two-sided exponential pulse function, this simplification does not affect the moments or PDF of the process, simplifies the auto-correlation and power spectra (see [27] for the auto-correlation of this process with finite growth) and allows the formulation of the process by a simple stochastic differential equation as described below.

The pulses are assumed to arrive according to a Poisson process with constant rate (see for example [19], [20, chapter 4.1] or [32, p 562]). Thus the (non-negative) number of arrivals $K(T)$ is Poisson distributed,

$$P_K(K; T, \tau_w) = \frac{1}{K!} \left(\frac{T}{\tau_w}\right)^K \exp\left(-\frac{T}{\tau_w}\right), \quad (4)$$

where we have taken the mean value to be $\langle K \rangle = T/\tau_w$. It can be shown that the waiting time between consecutive pulses is exponentially distributed with mean value τ_w [20, p 135], while the K arrival times t_k are independent and uniformly distributed on the interval $[0, T]$ [20, p 140]. The ratio between pulse duration time and average time between pulses,

$$\gamma = \frac{\tau_d}{\tau_w}, \quad (5)$$

is in the following referred to as the *intermittency parameter*.

While the FPP is a continuous process, any experimental data or synthetic realization of the process is discrete. As the time resolution of numerical data may be important for the noise processes, we also introduce the normalized time step,

$$\theta = \frac{\Delta t}{\tau_d}, \quad (6)$$

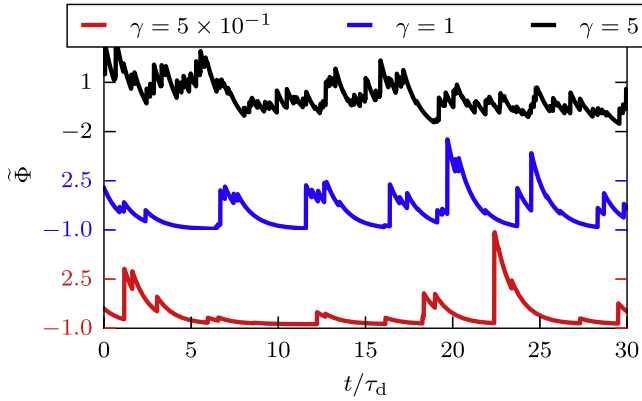


Figure 1. Synthetically generated pure filtered Poisson processes with $\theta = 10^{-2}$ and various values of γ .

where Δ_t is the time step for synthetically generated signals. Details on the synthetically generated signals will be discussed in section 6. Some realizations of $\Phi_K(t)$ for various values of γ are presented in figure 1, where we have used the normalization

$$\tilde{\Phi} = \frac{\Phi - \langle \Phi \rangle}{\Phi_{\text{rms}}}. \quad (7)$$

Here and in the following, we will use a tilde to denote a normalized variable with zero mean and unit standard deviation. For $\gamma < 1$, pulses arrive rarely and the signal spends large amounts of time close to zero value, resulting in a strongly intermittent signal. For $\gamma > 1$, pulses overlap and the signal begins to resemble random and symmetric fluctuations around a mean value.

It can be shown that the stationary PDF of the random variable Φ is a Gamma distribution [16, 17];

$$P_{\Phi}(\phi; \gamma, \langle A \rangle) = \frac{1}{\langle A \rangle \Gamma(\gamma)} \left(\frac{\phi}{\langle A \rangle} \right)^{\gamma-1} \exp\left(-\frac{\phi}{\langle A \rangle}\right), \quad (8)$$

where $\phi > 0$ due to the non-negative pulse amplitudes and pulse functions. The lowest order moments of Φ are the mean $\langle \Phi \rangle = \gamma \langle A \rangle$, the standard deviation $\Phi_{\text{rms}} = \gamma^{1/2} \langle A \rangle$, the skewness $S_{\Phi} = 2/\gamma^{1/2}$ and the flatness $F_{\Phi} = 3 + 6/\gamma$, giving a parabolic relationship between the skewness and flatness moments of Φ [16, 17],

$$F_{\Phi} = 3 + 3S_{\Phi}/2. \quad (9)$$

In section 4 it will be shown how additive noise alters this relationship.

It is possible to write the FPP as a stochastic differential equation. In the case of a one-sided exponential pulse function, it takes the form [33]

$$\tau_d \frac{d\Phi_K}{dt} = -\Phi_K + \sum_{k=1}^K A_k \delta\left(\frac{t-t_k}{\tau_d}\right), \quad (10)$$

where δ is the Dirac delta function. The model described by equation (10) can be seen as a train of δ -pulses, arriving at times t_k with amplitudes A_k . These pulses pass through a filter causing an exponential decay with characteristic decay time τ_d , giving the process its name.

We remark that equation (10) can be written as

$$\mathcal{L}\Phi_K = f_K(t), \quad (11)$$

where \mathcal{L} is the linear operator $\mathcal{L} = 1 + \tau_d d/dt$ and the forcing is given by

$$f_K(t) = \sum_{k=1}^K A_k \delta\left(\frac{t-t_k}{\tau_d}\right). \quad (12)$$

A Green function for the operator \mathcal{L} , that is any function fulfilling $\mathcal{L}G(t; s) = \delta(t-s)$, is given by

$$G(t-s) = \exp\left(-\frac{t-s}{\tau_d}\right) \Theta\left(\frac{t-s}{\tau_d}\right). \quad (13)$$

The solution of equation (11) is then the convolution

$$\Phi_K(t) = \int_{-\infty}^{\infty} ds G(t-s) f_K(s), \quad (14)$$

where $G(t-s)$ can be interpreted as the filter which $f_K(t)$ passes through. Since both G and f_K are non-zero only for positive arguments, the integration in equation (14) could be taken over the interval $[0, t]$, and this is done in section 6. Here, however, we take the integration limits to infinity in order to remain consistent with the Fourier transforms, discussed below and in the appendices.

In appendix B, the PSD of the FPP is shown to be

$$S_{\Phi}(\omega) = \Phi_{\text{rms}}^2 \frac{2\tau_d}{1 + \tau_d^2 \omega^2} + 2\pi \langle \Phi \rangle^2 \delta(\omega). \quad (15)$$

The mean value of the signal gives a zero frequency contribution to the PSD, while the fluctuations around the mean value give rise to a Lorentzian power spectrum. Since the pulses are uncorrelated, there is no explicit dependence on the average waiting time τ_w , apart from contributing to the value of $\langle \Phi \rangle$ and Φ_{rms} . Moreover, we see from equations (B5) and (B11) that the Poisson point process $f_K(t)$ provides the zero frequency contribution as well as a flat contribution independent from the frequency due to the lack of correlation between the pulses, while the Lorentzian spectrum comes entirely from the filter $G(t)$.

The auto-correlation and PSD are Fourier transform pairs under the Fourier transform over the entire real line. Thus we readily obtain the auto-correlation function

$$\begin{aligned} R_{\Phi}(\tau) &= \mathcal{F}^{-1}[S_{\Phi}](\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega e^{i\omega\tau} S_{\Phi}(\omega) \\ &= \Phi_{\text{rms}}^2 \exp\left(-\frac{|\tau|}{\tau_d}\right) + \langle \Phi \rangle^2, \end{aligned} \quad (16)$$

where $\mathcal{F}^{-1}[\bullet](\tau)$ denotes the inverse Fourier transform. We see that the time dependence of the auto-correlation function comes entirely from the Green function $G(t)$, as is expected, since the pulses are uncorrelated and thus the average time between pulses does not appear explicitly in the auto-correlation function.

It can be shown that the PSD and auto-correlation function of the normalized variable $\tilde{\Phi}$ are given by

$$S_{\tilde{\Phi}}(\omega) = \frac{S_{\Phi}(\omega) - 2\pi \langle \Phi \rangle^2 \delta(\omega)}{\Phi_{\text{rms}}^2} = \frac{2\tau_d}{1 + \tau_d^2 \omega^2}, \quad (17)$$

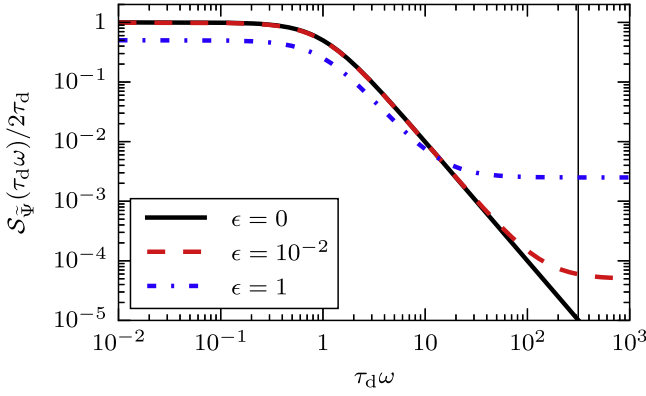


Figure 2. Comparison of power spectral densities with $\theta = 10^{-2}$ and various values of ϵ . The solid line denotes both $S_{\tilde{\Phi}}$ and $S_{\tilde{\Delta}}$ while the broken lines denote $S_{\tilde{\Omega}}$. The vertical line gives the Nyquist frequency.

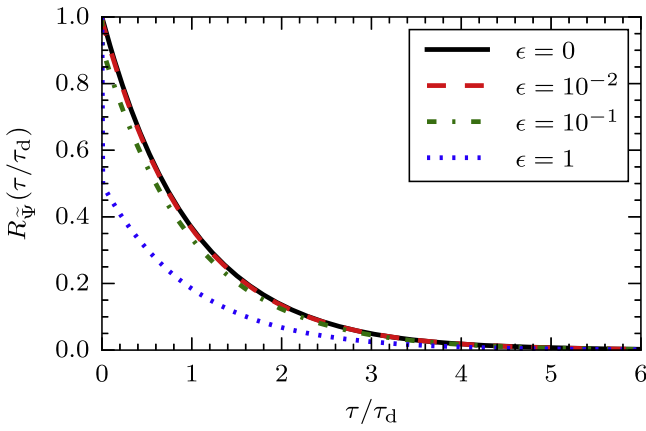


Figure 3. Comparison of auto-correlation functions with $\theta = 10^{-2}$ and various values of ϵ . The solid line denotes both $R_{\tilde{\Phi}}$ and $R_{\tilde{\Delta}}$ while the broken lines denote $R_{\tilde{\Omega}}$.

$$R_{\tilde{\Phi}}(\tau) = \frac{R_{\Phi}(\tau) - \langle \Phi \rangle^2}{\Phi_{\text{rms}}^2} = \exp\left(-\frac{|\tau|}{\tau_d}\right). \quad (18)$$

These expressions are presented in figures 2 and 3, respectively. The solid line in figure 2 represents the PSD of $\tilde{\Phi}$, while the solid line in figure 3 represents the auto-correlation function of $\tilde{\Phi}$. The other elements in these figures will be discussed further in section 5. From equation (18), it is evident that the e-folding time of $R_{\tilde{\Phi}}$ is the pulse duration time τ_d . This corresponds to $\tau_d \omega = 1$ in figure 2, giving the approximate frequency where the PSD changes from a flat spectrum to power law behavior.

3. Additive noise

In this section, we will expand the FPP with two forms of additional noise terms, which will be referred to as either observational (Ω) or dynamical (Δ) noise. In the following, the specific realization of the FPP (that is, the pulse amplitudes, arrival times and number of pulses for a particular realization) plays no role. Thus, for simplicity of notation, the subscript K will be suppressed in the following.

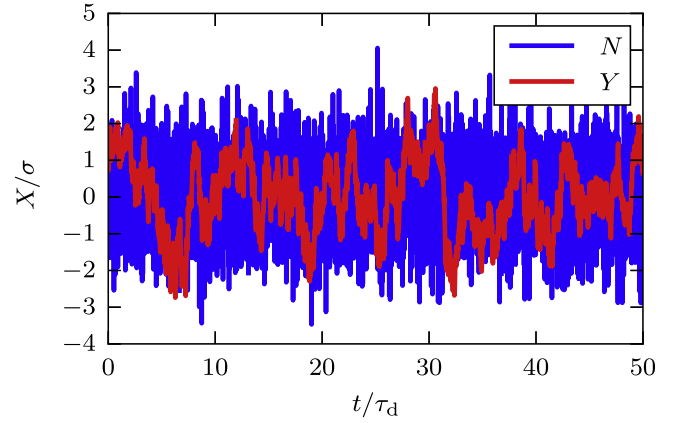


Figure 4. Synthetically generated observational and dynamic noise terms for $\theta = 10^{-2}$.

Observational noise denotes noise unconnected to the FPP. Thus, a noise term is simply added to a realization of the process $\Phi(t)$;

$$\Omega(t) = \Phi(t) + \sigma N(t), \quad (19)$$

where $N(t)$ is a normally distributed process with vanishing mean and unit standard deviation and σ is the noise intensity parameter, effectively describing the standard deviation of the noise process.

In the case of dynamical noise, the noise term is added as random forcing in the stochastic differential equation (10), and is therefore connected to the pulses:

$$\tau_d d\Delta = \left[-\Delta(t) + \sum_{k=1}^K A_k \delta\left(\frac{t-t_k}{\tau_d}\right) \right] dt + \sqrt{2\tau_d} \sigma dW \quad (20)$$

where $W(t)$ is the Wiener process. It is possible to solve equation (20) in the same way as equation (10) was solved, giving

$$\Delta(t) = \Phi(t) + \sigma Y(t), \quad (21)$$

where

$$Y(t) = \sqrt{2/\tau_d} \int_{-\infty}^{\infty} G(t-s) dW(s) \quad (22)$$

is an Ornstein–Uhlenbeck process, and the random variable Y is normally distributed with zero mean and unit standard deviation.

We also introduce a parameter ϵ , describing the relative noise level. Using X as a collective symbol for both σN and σY , we define ϵ as

$$\epsilon = \left(\frac{X_{\text{rms}}}{\Phi_{\text{rms}}} \right)^2 = \frac{\sigma^2}{\gamma \langle A \rangle^2}. \quad (23)$$

Although $N(t)$ and $Y(t)$ have the same probability distributions, they exhibit very different dynamical behavior, as illustrated by realizations of the processes presented in figure 4. While $N(t)$ fluctuates rapidly on the sampling time scale around the zero value, $Y(t)$ wanders around the zero value, with finite temporal correlations. This is quantified in appendix B.2, where the auto-correlation functions and PSDs

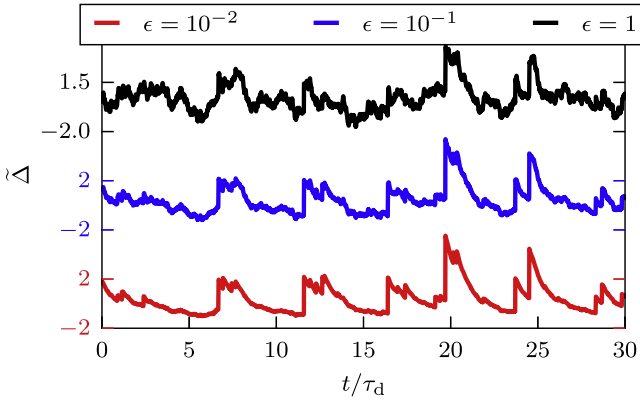


Figure 5. Synthetically generated filtered Poisson processes with dynamical noise, $\gamma = 1$, $\theta = 10^{-2}$ and various values of ϵ .

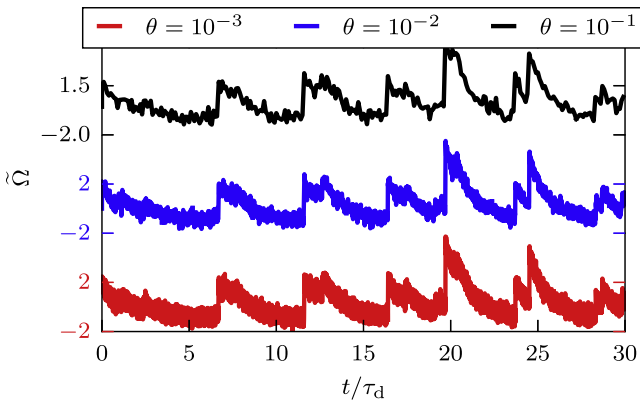


Figure 6. Synthetically generated filtered Poisson processes with observational noise, $\gamma = 1$, $\epsilon = 10^{-1}$ and various values of θ .

of N and Y are derived. Note that both N and Y are independent of the process Φ .

Realizations of the processes Φ , Δ and Ω for different parameters γ , ϵ and θ are presented in figures 1, 5 and 6. Realizations of the pure FPP for various γ are presented in figure 1. For $\gamma < 1$, the signal is very intermittent and spends large amounts of time close to zero. For $\gamma > 1$, pulse overlap is much more significant, washing out the intermittent features. As γ becomes very large, the signal resembles symmetric fluctuations around the mean value, and it can be shown that in the limit $\gamma \rightarrow \infty$, the FPP has a probability distribution resembling a normal distribution [5, 16, 17, 19].

Realizations of the FPP with dynamical noise for various ϵ are presented in figure 5. For very small ϵ , there is very little difference between the pure process and the process with dynamical noise. For larger ϵ , the noise process plays a larger role, concealing all but the largest pulses.

The effect of changing θ is clearest for the FPP with observational noise. Realizations of this process for various values of θ are presented in figure 6. Here we see that for large θ , the process resembles a pure FPP. For small θ the noise process dominates, even though its rms-value is $1/10$ the rms-value of the FPP in this case. A smaller time step means more data points in a given time interval and thus more chances for large values of the noise process. In contrast, the

pure FPP itself is much less sensitive to changes in time resolution, the primary effect being that separate pulses may be counted as one in the computation, if they are close enough. The FPP with dynamical noise is also less sensitive to θ due to the exponential damping the noise.

4. Moments and distributions

In this section, we present the lowest order moments and the PDF of the FPP with additive noise and describe how the model parameters can be estimated from these for a realization of the process. The probability density function of the stationary process Φ is given in equation (8), and both N and Y are normally distributed with vanishing mean and unit standard deviation. Thus, we write

$$\Psi = \Phi + X, \quad (24)$$

where the random variable X has the probability density function

$$P_X(x; \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right). \quad (25)$$

Note that X has the same distribution as both σN and σY , while Ψ has the same distribution as both Ω and Δ . In other words, the PDF of the FPP with additional noise is the same for both types of noise considered here. In this section, we assume continuous random variables X and Φ . It can be shown that the convergence of the moments of Φ to their true values depends on θ . A discussion on this and the estimation of the moments of Φ in general is given in [17, 34]. Under the assumption that Φ and X are independent, and using that $\langle X \rangle = 0$, the lowest order moments of Ψ are readily calculated as

$$\langle \Psi \rangle = \langle \Phi + X \rangle = \gamma \langle A \rangle, \quad (26a)$$

$$\Psi_{\text{rms}}^2 = \langle [(\Phi + X) - \langle \Phi \rangle]^2 \rangle = \gamma \langle A \rangle^2 + \sigma^2, \quad (26b)$$

$$S_\Psi = \frac{\langle [(\Phi + X) - \langle \Phi \rangle]^3 \rangle}{\Psi_{\text{rms}}^3} = \frac{2\gamma \langle A \rangle^3}{(\gamma \langle A \rangle^2 + \sigma^2)^{3/2}}, \quad (26c)$$

$$F_\Psi = \frac{\langle [(\Phi + X) - \langle \Phi \rangle]^4 \rangle}{\Psi_{\text{rms}}^4} = \frac{3}{(\gamma \langle A \rangle^2 + \sigma^2)^2} \times [\gamma \langle A \rangle^4 (\gamma + 2) + 2\gamma \langle A \rangle^2 \sigma^2 + \sigma^4]. \quad (26d)$$

Using equation (23), we have the moments

$$\langle \Psi \rangle = \gamma \langle A \rangle, \quad (27a)$$

$$\Psi_{\text{rms}}^2 = (1 + \epsilon) \gamma \langle A \rangle^2, \quad (27b)$$

$$S_\Psi = \frac{2}{(1 + \epsilon)^{3/2} \gamma^{1/2}}, \quad (27c)$$

$$F_\Psi = 3 + \frac{6}{(1 + \epsilon)^2 \gamma}. \quad (27d)$$

Also in this case, we can find a parabolic relation between the skewness and flatness moments,

$$F_\Psi = 3 + \frac{3}{2} (1 + \epsilon) S_\Psi^2. \quad (28)$$

The effect of additional noise is to increase the pre-factor in the parabolic relationship.

The model parameters γ and ϵ can be estimated from the moments in a several different ways. Using equations (27a) and (27b) to eliminate $\langle A \rangle$. The relative fluctuation level $\Psi_{\text{rms}}/\langle \Psi \rangle$ is then related to the model parameters by

$$\frac{\Psi_{\text{rms}}}{\langle \Psi \rangle} = \sqrt{\frac{1 + \epsilon}{\gamma}}, \quad (29)$$

clearly showing how the additional noise amplifies the fluctuation level. As the estimators for lower order moments are more accurate than those for higher order moments, it is reasonable to assume the most accurate estimators for the model parameters come from using the lowest order moments, $\Psi_{\text{rms}}/\langle \Psi \rangle$ and S_{Ψ} :

$$\epsilon = \left(\frac{2}{S_{\Psi}} \right)^{1/2} \left(\frac{\Psi_{\text{rms}}}{\langle \Psi \rangle} \right)^{1/2}, \quad (30a)$$

$$\gamma = \left(\frac{2}{S_{\Psi}} \right)^{1/2} \left(\frac{\langle \Psi \rangle}{\Psi_{\text{rms}}} \right)^{3/2}. \quad (30b)$$

In experimental fluctuation data time series there can sometimes be reasons not to trust the mean value of a signal, for example due to externally imposed low frequency noise or trends [23, 24]. Such problems typically do not affect the large-amplitude fluctuations, leaving the higher order moments trustworthy. In this case, using the normalization in equation (7) and observing that $S_{\tilde{\Psi}} = S_{\Psi}$ and $F_{\tilde{\Psi}} = F_{\Psi}$, we have

$$\epsilon = \frac{2(F_{\tilde{\Psi}} - 3)}{3 S_{\tilde{\Psi}}^2} - 1, \quad (31a)$$

$$\gamma = \frac{27 S_{\tilde{\Psi}}^4}{2(F_{\tilde{\Psi}} - 3)^3}. \quad (31b)$$

In section 6, it will be shown that estimating the parameters from equation (30) is preferable to using equation (31), in cases where $\Psi_{\text{rms}}/\langle \Psi \rangle$ is reliable.

The probability density function of a sum of two independent random variables Φ and X is a convolution of their respective probability density functions [32]:

$$P_{\Psi}(\psi; \gamma, \langle A \rangle, \sigma) = \int_{-\infty}^{\infty} d\phi P_{\Phi}(\phi; \gamma, \langle A \rangle) P_X(\psi - \phi; \sigma). \quad (32)$$

As a consistency check, it should be noted that in the limit $\epsilon \rightarrow 0$, P_{Ψ} should be the probability density function of a pure FPP. This is indeed the case, as in this limit $\sigma \rightarrow 0$, and by definition,

$$\lim_{\sigma \rightarrow 0} P_X(\psi - \phi; \sigma) = \delta(\psi - \phi),$$

and the Gamma distribution of the FPP without noise is recovered. Thus, in the following, we take the case $\epsilon = 0$ to signify a FPP without additive noise.

The expression for the probability density function of $\tilde{\Psi}$ is given in the appendix, equation (A9). In order to illustrate

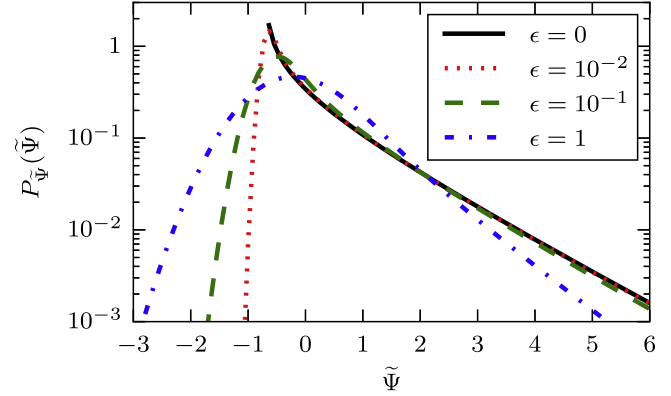


Figure 7. Probability density function of the random variable Ψ for $\gamma = 1/2$ and various values of ϵ .

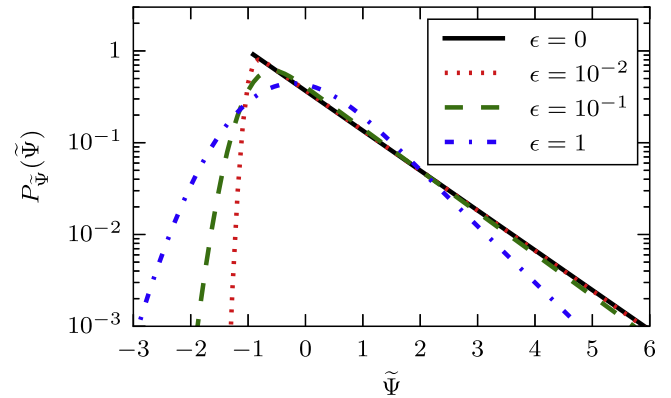


Figure 8. Probability density function of the random variable Ψ for $\gamma = 1$ and various values of ϵ .

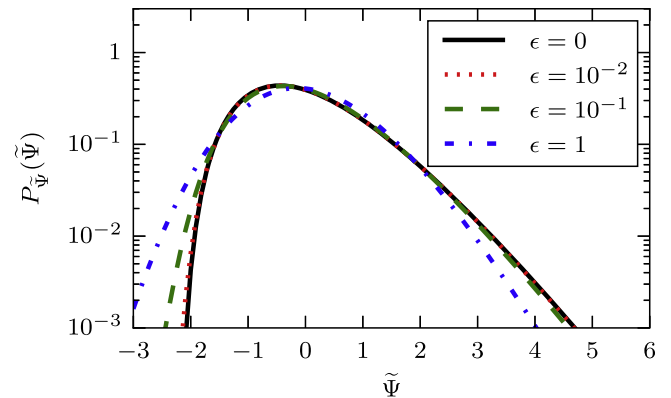


Figure 9. Probability density function of the random variable Ψ for $\gamma = 5$ and various values of ϵ .

the effect of pulse overlap and additional noise, the PDF for $\tilde{\Psi}$ is shown for various values of ϵ and $\gamma = 1/2, 1$ and 5 in figures 7–9, respectively. Clearly, as ϵ increases beyond unity, the probability distribution changes towards a normal distribution. For $\epsilon = 0$, the random variable Ψ is non-negative, causing an abrupt halt in the distribution for $\gamma \leq 1$. This jump does not exist for $\epsilon > 0$. Thus negative values for Ψ , or equivalently, $\tilde{\Psi} < -\gamma^{1/2}$, signifies a deviation from a pure FPP.

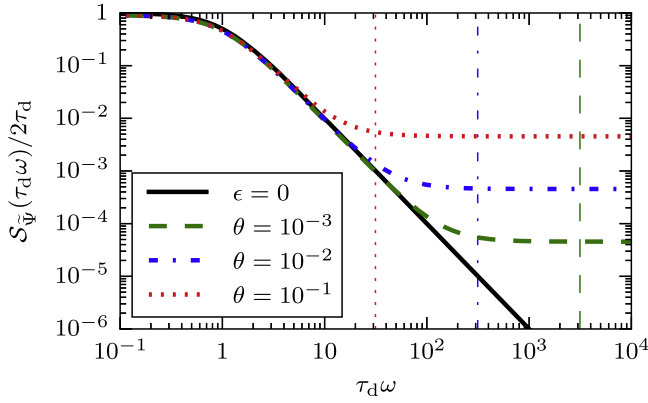


Figure 10. Comparison of power spectral densities with $\epsilon = 10^{-2}$ and various values of θ . The solid line denotes $\mathcal{S}_{\tilde{\Phi}}$ and $\mathcal{S}_{\tilde{\Delta}}$ while the broken lines denote $\mathcal{S}_{\tilde{\Omega}}$. The vertical lines give the Nyquist frequency of the corresponding expressions.

5. Spectra and correlations

The PSD of the sum of two independent random variables is the sum of their respective PSDs. The power spectra for the FPP with additional noise are derived in appendix B. With equations (15) and (B15), the PSD of the FPP with dynamical noise is therefore

$$\begin{aligned} \mathcal{S}_{\Delta}(\omega) &= \mathcal{S}_{\tilde{\Phi}}(\omega) + \sigma^2 \mathcal{S}_Y(\omega) \\ &= \Delta_{\text{rms}}^2 \frac{2\tau_d}{1 + \tau_d^2 \omega^2} + 2\pi \langle \Delta \rangle^2 \delta(\omega), \end{aligned} \quad (33)$$

where $\langle \Delta \rangle$ and Δ_{rms} are given by equations (27a) and (27b), respectively. We have also used the relation $\sigma^2 = \gamma \epsilon \langle A \rangle^2$. Note that the spectrum in equation (33) is identical to that for the pure FPP. Thus, we conclude that the auto-correlation functions of Φ and Δ have the same functional shape as well, up to the difference in the first two moments.

The PSD of the FPP with observational noise is obtained from equations (15) and (B23):

$$\mathcal{S}_{\tilde{\Omega}}(\omega) = \Omega_{\text{rms}}^2 \frac{2\tau_d}{1 + \epsilon} \left[\frac{1}{1 + \tau_d^2 \omega^2} + \frac{\epsilon}{2} \theta \right] + 2\pi \langle \Omega \rangle^2 \delta(\omega). \quad (34)$$

This function is qualitatively different from the PSDs of Φ and Δ , although it converges to both in the limit $\epsilon \rightarrow 0$. These differences are now explored in detail.

Using the normalizations in equations (7) and (17), we can list the PSDs of the rescaled signals $\tilde{\Phi}$, $\tilde{\Omega}$ and $\tilde{\Delta}$ as

$$\mathcal{S}_{\tilde{\Phi}}(\omega) = \mathcal{S}_{\tilde{\Delta}}(\omega) = 2\tau_d \frac{1}{1 + \tau_d^2 \omega^2}, \quad (35a)$$

$$\mathcal{S}_{\tilde{\Omega}}(\omega) = \frac{2\tau_d}{1 + \epsilon} \left[\frac{1}{1 + \tau_d^2 \omega^2} + \frac{\epsilon}{2} \theta \right]. \quad (35b)$$

The PSD of $\tilde{\Psi}$ is presented in figures 2 and 10. In figure 2, the difference between $\tilde{\Phi}$ or $\tilde{\Delta}$ and $\tilde{\Omega}$ is presented. Higher ratio of noise signal to FPP decreases the value of the PSD for low

frequencies and causes a transition from a power law spectrum to a constant spectrum at higher frequencies. The Nyquist frequency ω_N for $\theta = 10^{-2}$ is indicated by the vertical line,

$$\omega_N = 2\pi \frac{1}{2\Delta_t} = \frac{\pi}{\tau_d \theta}, \quad (36)$$

and shows that for small ϵ , this transition happens at too high frequencies to be reliably observed. In figure 10, the PSD of $\tilde{\Omega}$ is presented for $\epsilon = 10^{-1}$ and various values of θ . For low frequencies, the difference between $\mathcal{S}_{\tilde{\Phi}}(\omega)$ and $\mathcal{S}_{\tilde{\Omega}}(\omega)$ is too small to be of practical use, while the effect of noise for high frequencies can only be observed without aliasing for very low θ .

The auto-correlation function of Φ is given in equation (16). Since the PSD of the rescaled processes $\tilde{\Phi}$ and $\tilde{\Delta}$ are identical, we conclude that their auto-correlation functions are as well,

$$R_{\tilde{\Phi}}(\tau) = R_{\tilde{\Delta}}(\tau) = \exp\left(-\frac{|\tau|}{\tau_d}\right), \quad (37)$$

where the auto-correlation of a rescaled random process is given in equation (18). With the auto-correlation function of $N(t)$ from equation (B18), we have that the auto-correlation function of $\tilde{\Omega}$ is

$$R_{\tilde{\Omega}}(\tau) = \frac{1}{1 + \epsilon} \left[\exp\left(-\frac{|\tau|}{\tau_d}\right) + \epsilon \left(1 - \frac{|\tau|}{\theta \tau_d}\right) \Theta\left(1 - \frac{|\tau|}{\theta \tau_d}\right) \right]. \quad (38)$$

Note that $R_{\tilde{\Omega}}(0) = 1$ as required by the normalization, while for correlation times longer than the sampling time, the correlation function of the normalized variable $\tilde{\Omega}$ has a value of $1/(1 + \epsilon)$ times the correlation function of $\tilde{\Delta}$. Thus, equation (38) can also be written as

$$R_{\tilde{\Omega}}(\tau) = \begin{cases} 1, & \tau = 0, \\ \frac{1}{1 + \epsilon} \exp\left(-\frac{|\tau|}{\tau_d}\right), & |\tau| \geq \Delta_t. \end{cases} \quad (39)$$

If there is no appreciable drop from $\tau = 0$ to $\tau = \Delta_t$, and the correlation functions overlap for $\tau \geq \Delta_t$, the observational noise is negligible.

As expected, in the limit $\epsilon \rightarrow 0$, $R_{\tilde{\Omega}}$ approaches $R_{\tilde{\Phi}}$. In practice, the auto-correlation function should be better at revealing the presence of observational noise than the PSD, as this difference is largest for small time lags, where the auto-correlation function is the most accurate. Still, $R_{\tilde{\Delta}}(\Delta_t) = \exp(-\theta)$ and $R_{\tilde{\Omega}}(\Delta_t) = \exp(-\theta)/(1 + \epsilon)$, a difference which may be difficult to verify in practice for small, but appreciable ϵ (say, $\epsilon \sim 10^{-1}$). An example of this is seen in figure 3. In this figure, the auto-correlation function of $\tilde{\Psi}$ is presented as a function of τ/τ_d for $\theta = 10^{-2}$ and various values of ϵ . For large ϵ , both the initial drop and the reduced value of $R_{\tilde{\Omega}}$ compared to $R_{\tilde{\Delta}}$ is clear, but this is not the case for $\epsilon \leq 10^{-1}$. The behavior of the auto-correlation of $\tilde{\Omega}$ for $\epsilon = 10^{-1}$ and varying θ is shown in figure 11. All functions fulfill $R_{\tilde{\Phi}}(0) = 1$, and for $\tau > \Delta_t$, all $R_{\tilde{\Omega}}$ have the same value.

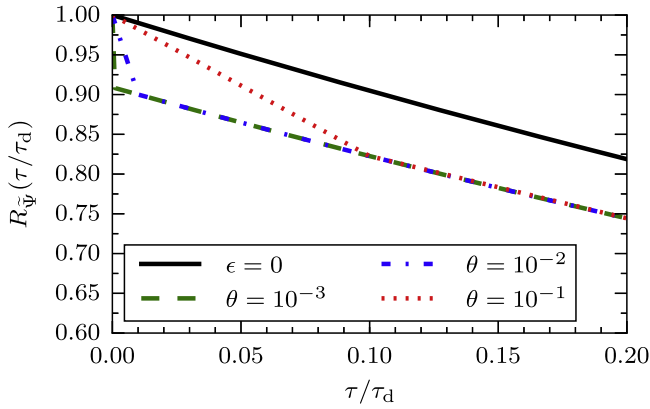


Figure 11. Comparison of auto-correlation functions with $\epsilon = 10^{-1}$ and various values of θ . The solid line denotes both R_{Φ} and R_{Δ} , while the broken lines denote R_{Γ} .

6. Parameter estimation

In this section, we present results from numerical analysis using synthetically generated time series. The time series are generated following the convolutions in equations (14) and (22), with integration limits $[0, t]$. The convolutions are performed by a fast Fourier transform numerical convolution. All random numbers are generated using a Mersenne Twister. The time array is constructed as $t_m = m\Delta_t$, with $m = 0, 1, \dots, M - 1$ and $T = \Delta_t M$. We have set $\tau_d = 1$, so we vary θ by varying Δ_t . The K pulse amplitudes are drawn from an exponential distribution with $\langle A \rangle = 1$, the K arrival times correspond to K integers $\{m_k\}_{k=1}^K$ uniformly distributed on $[0, M - 1]$, giving

$$f_K(t) \rightarrow f_K[m] = \sum_{k=1}^K A_k \delta_{mm_k}, \quad (40)$$

where δ_{ab} is the Kronecker delta function for integers a and b . $N[m]$ is an array of M independent and identically distributed normal variables with vanishing mean and unit standard deviation, while the discrete version of dW is an equally shaped array of independent and identically distributed normal variables with vanishing mean and standard deviation $\Delta_t^{1/2}$.

6.1. Reliability of parameter estimation

The aim of this section is to numerically test and verify results from the prior sections. We present analysis of parameter estimation for synthetic time series with the goal of separating the two types of noise, where the analysis is performed on 1000 time series of each type, each of length $M = 10^6$ data points, with parameters close to experimental values (see [26–28]); $\gamma = 2$, $\epsilon = 5 \times 10^{-2}$ and $\theta = 2 \times 10^{-2}$. This gives time series of duration $T/\tau_d = 2 \times 10^4$.

Given a time series, we can estimate γ and ϵ by comparing the estimated PDF for the time series to the PDF in equation (A9), or by comparing the sample estimate of the moments to the moments in equations (30) or (31). These two methods do not discriminate the nature of the noise. However, comparing equations (37) and (39), we see that equation (39)

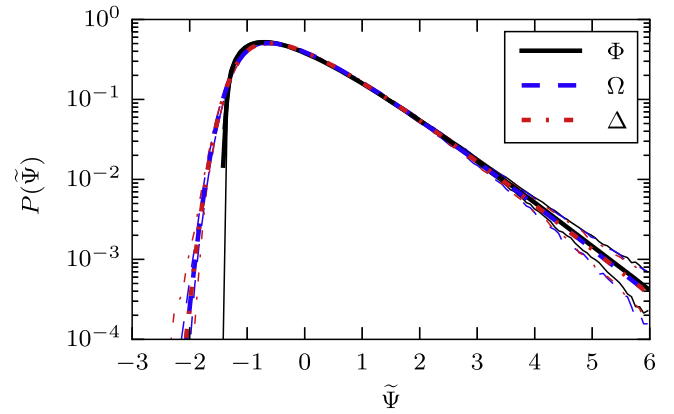


Figure 12. Mean probability density functions (thick lines) and the maximal deviations from the mean (thin lines) of synthetically generated signals for $\gamma = 2$, $\epsilon = 5 \times 10^{-2}$, $\theta = 2 \times 10^{-2}$, and 10^3 samples, each with 10^6 data points.

gives the auto-correlation function of a FPP with observational noise for $\epsilon > 0$ and the auto-correlation function of a FPP with dynamical noise or without noise for $\epsilon = 0$. Thus equation (39) can potentially separate observational from dynamical noise. In addition, it gives us an estimate of θ . The same considerations hold for the PSDs in equation (35a) as compared to equation (35b). In all the following cases where we fit synthetic data to an analytical function, the initial values are given as the true parameter values of the process.

The estimated PDFs of the synthetic signals are presented in figure 12. The thick lines give the average PDF for all synthetic signals, and the thin lines indicate the maximal deviation from this mean value. Visually, the only reliable difference between the signals is the elevated tail for negative values of $\tilde{\Delta}$ and $\tilde{\Omega}$ compared to $\tilde{\Phi}$. In figure 13, we present PDFs for the estimated parameters $\hat{\gamma}$ and $\hat{\epsilon}$ from the moments and the PDF of the processes. Here and in the following, the hat symbol $\hat{\cdot}$ indicates an estimated value. The values in figures 13(a) and (b) were obtained by estimating the relative fluctuation level and skewness of the synthetic data and using equation (30). The values in figures 13(c) and (d) were obtained by estimating the skewness and flatness and using equation (31). In figures 13(e) and (f), the values were obtained by fitting the function in equation (A9) to the estimated PDF of the synthetic data with a least squares routine under the constraints $\hat{\gamma} \geq 10^{-1}$ and $\hat{\epsilon} \geq 10^{-6}$ to ensure convergence. The rms-value of the distributions of $\hat{\gamma}$ and $\hat{\epsilon}$ are presented in tables 1 and 2, respectively.

As is evident from figure 13, using the three lowest order moments or the PDF of the signal to estimate the parameters is far preferable to using the estimated skewness and flatness moments of the signal. Figures 13(c) and (d) have far broader distributions than the other methods of parameter estimation. Due to the significant overlap between the distributions in figure 13(d), the presence or absence of noise can be difficult to determine reliably. For the underlying parameters used here, table 1 shows that the three lowest order moments give a better estimate for γ than the PDF, while table 2 shows the reverse for ϵ . This result is likely highly dependent on the

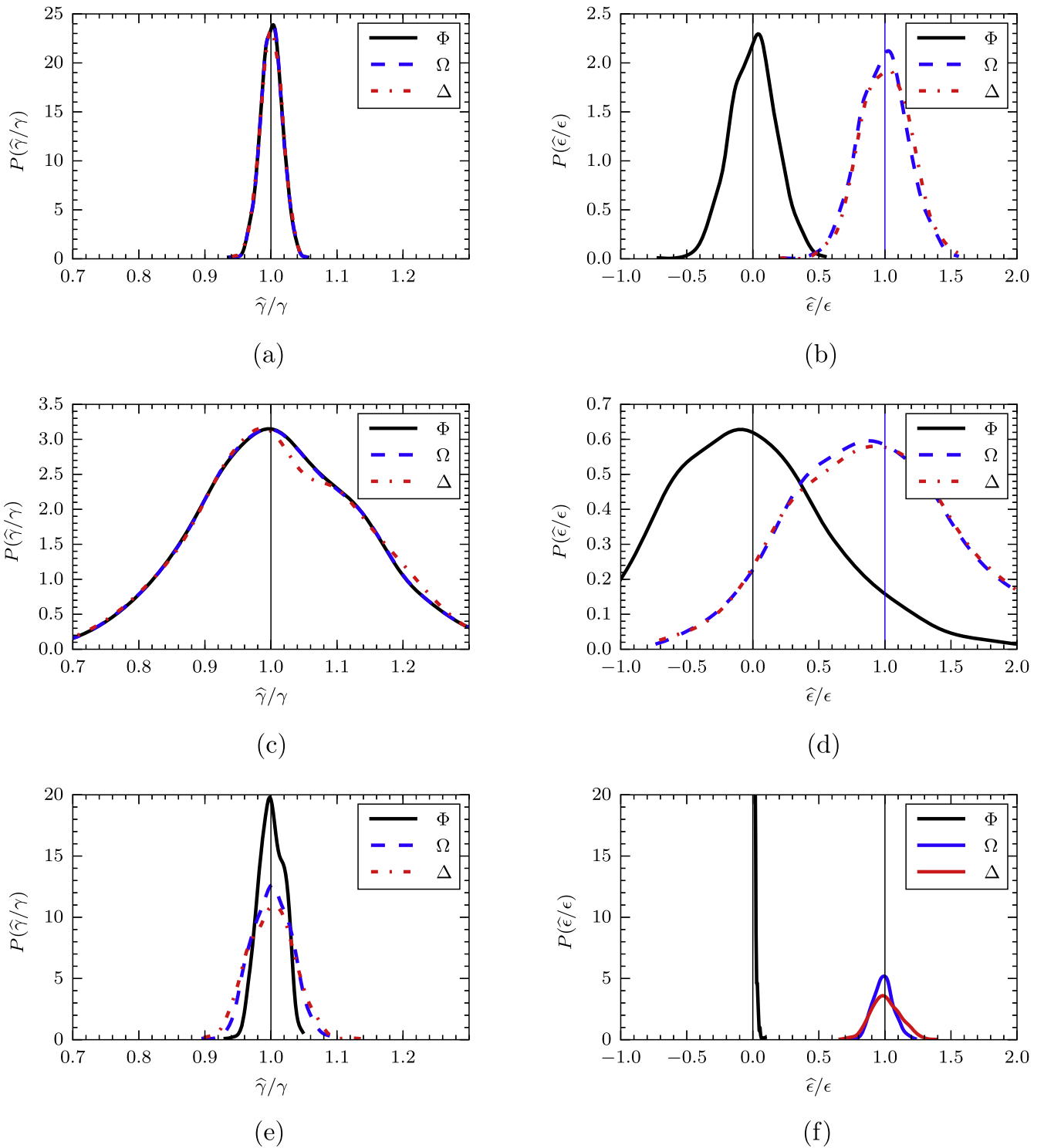


Figure 13. PDFs of the intermittency parameter $\hat{\gamma}/\gamma$ and the noise level parameter ϵ estimated from the moments and the PDF of synthetically generated signals. The thin vertical lines give the true parameters. (a) and (b) give PDFs from the relative fluctuation level and skewness. (c) and (d) give PDFs from the skewness and flatness moments. (e) and (f) give PDFs from fitting to the estimated PDF.

estimation methods employed, as well as the properties of the time series in question. Thus, the authors recommend a full Monte Carlo analysis, as presented here, in order to determine errors in parameter estimation. As expected, the type of noise cannot be determined from the moments or PDF of the signal.

The auto-correlation functions of the synthetic signals are presented in figure 14(a), while the PSDs are presented in

figure 14(b). Again the thick lines give the average result from all signals, while the thin lines give the maximal deviation from the average. The PSD was computed using Welch's method with 2^{10} data points window size, 50% overlap and a Hanning window. In both cases, the observational noise gives a clear visual difference compared to the pure FPP and the same process with dynamical noise, having

Table 1. Standard deviation of estimated γ -values of synthetically generated signals for $\gamma = 2$, $\epsilon = 5 \times 10^{-2}$, $\theta = 2 \times 10^{-2}$, and 10^3 samples, each with 10^6 data points.

$(\hat{\gamma}/\gamma)_{\text{rms}}$	$\Psi_{\text{rms}}/\langle\Psi\rangle, S_{\Psi}$	S_{Ψ}, F_{Ψ}	$P(\tilde{\Psi})$
Φ	0.016	0.13	0.019
Δ	0.017	0.13	0.030
Ω	0.016	0.13	0.035

Table 2. Standard deviation of estimated ϵ -values of synthetically generated signals for $\gamma = 2$, $\epsilon = 5 \times 10^{-2}$, $\theta = 2 \times 10^{-2}$, and 10^3 samples, each with 10^6 data points.

$(\hat{\epsilon}/\epsilon)_{\text{rms}}$	$\Psi_{\text{rms}}/\langle\Psi\rangle, S_{\Psi}$	S_{Ψ}, F_{Ψ}	$P(\tilde{\Psi})$
Φ	0.17	0.64	0.011
Δ	0.19	0.69	0.077
Ω	0.19	0.67	0.11

an elevated tail as predicted in equation (35b). Note also the slight lifting of the tail in figure 14(b) for Φ and Δ . This is most likely caused by roundoff errors near the Nyquist frequency. Changing the window size or the windowing function does not correct the problem. Direct computation of the periodogram using a fast Fourier transform does not have this problem, although it presents other problems for parameter estimation. The effect of this lifting of the tail on parameter estimation is discussed below.

In order to estimate θ and ϵ , we have fitted the auto-correlation function of the synthetic signals to the discrete version of equation (39):

$$R[m] = \frac{1}{1 + \epsilon} \exp(-\theta m), \quad m \geq 1, \quad (41)$$

where we ignore the $m = 0$ contribution, since this equals unity for all processes discussed. Due to the uncertainties in the auto-correlation function for large time lags, only the first 50 time steps are used. The PSD of the synthetic signals is fitted to the function in equation (35b). In both cases, a nonlinear least-squares fit routine with the true values as initial values was used. In figures 14(c) and (e) the PDFs of $\hat{\theta}$ and $\hat{\epsilon}$ from the auto-correlation are presented, respectively. As θ is the same for all three classes of signal, the overlapping PDFs are no surprise, and we note the small variation around the true value ($\hat{\theta}_{\text{rms}} = 0.015\theta$ for all processes). Concerning $\hat{\epsilon}$, the auto-correlation function does indeed separate the observational noise from the dynamical noise. For the observational noise, the rms-value is low, $\hat{\epsilon}_{\text{rms}} = 0.046\epsilon$, and the auto-correlation function can be used to estimate the noise level. The PDFs of $\hat{\theta}$ and $\hat{\epsilon}$ from the PSD are presented in figures 14(d) and (f), respectively. Again the rms-values of all parameters is small, 1%–2% of the original parameters, with the exception of the ϵ -parameter in figure 14(f) for Φ and Δ , with $\hat{\epsilon}_{\text{rms}} = 0.0012\epsilon$. In these figures, we have a clear bias in both $\hat{\theta}$ and $\hat{\epsilon}$. This is most likely created by the lifting of the tail of the PSD estimate, as discussed above. An elevated tail signifies observational

noise, so an artificial elevated tail leads to an over-estimation of ϵ , as seen in figure 14(f). This leads to a bias in the estimation of θ , as seen in figure 14(d). Eliminating this bias by restricting the fit range is not recommended, as this will compromise the accuracy in separating the PSDs. As this bias is reproducible, it does not present a significant problem for parameter estimation, and the PSD could be used as a sanity check for the auto-correlation function.

6.2. Level crossing statistics

Another measure which would intuitively separate the two types of noise is the number of upwards crossings above a certain threshold level per unit time, or the rate of level crossings. This quantity has been explored in [6, 17, 35–37] for the pure FPP, in [38, 39] for a Gamma distributed random process and in [40, 41] for atmospheric plasma. The rate of level crossings above a threshold Ψ as a function of the threshold is presented in figure 15 for the synthetic data discussed in the previous section. The thick lines give the mean values for the given threshold, while the thin lines represent the minimal and maximal value for all synthetic time series generated. The threshold is in units of signal rms-value above signal mean value. In agreement with intuition, the FPP with observational noise crosses the threshold much more frequently than the two others due to the rapid fluctuations around the mean value of the pure FPP at any amplitude triggering spurious crossings. The difference between the pure FPP and the FPP with dynamical noise is largest for small threshold values, where the number of threshold crossings is largest. Note that while Φ has its maximum number of level crossings for $\tilde{\Psi} = 0$, this is down-shifted for the processes with noise, since the noise does not contribute to the mean value of the process. While we know of no theoretical estimate for the rate of level crossings for a FPP with noise, this value can still be found from synthetic signals, generated by estimating γ and ϵ from the PDF of a measurement signal and θ from its auto-correlation function. Comparing the true rate of level crossings to the rate of level crossings for synthetic signals with different types of noise could separate the noise types.

7. Discussion and conclusions

Motivated by previous analysis of measurement data from magnetically confined plasmas, we have here investigated a FPP with normally distributed noise included as either observational or dynamical noise. The PDF, PSD and auto-correlation function of the pure FPP have been presented. These have also been extended to include noise, showing how the Gamma probability distribution of the FPP moves towards a normal distribution as the influence of noise increases and how dynamical noise gives the same auto-correlation and PSD as the pure FPP. The model including noise has a PDF in accordance with recent measurements [26, 31] and a parabolic relation between the skewness and flatness moments, as found in a variety of experiments on magnetized plasmas [23, 24, 26–

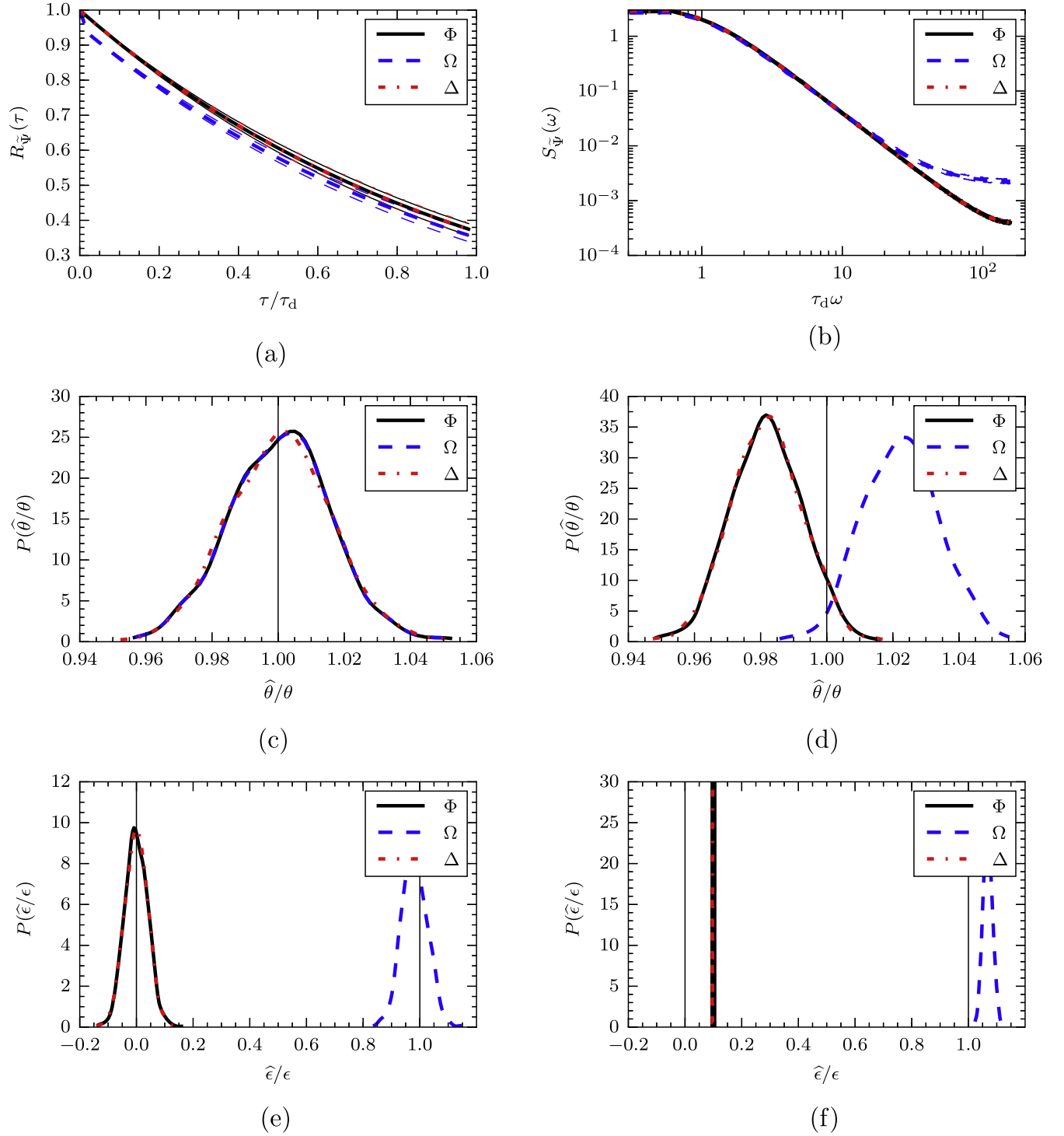


Figure 14. Results from the auto-correlation function and power spectral density of synthetically generated signals with $\gamma = 2$, $\epsilon = 5 \times 10^{-2}$, $\theta = 2 \times 10^{-2}$ and samples with 10^6 data points. The thin vertical lines give the true parameters. (a) and (b) give mean auto-correlation and power spectral density (thick lines) and maximal deviations from the mean (thin lines). (c) and (e) give PDFs from fitting to the auto-correlation function. (d) and (f) give PDFs from fitting to the power spectral density.

28, 31, 42]. In several of these investigations, the comparison made to the parabolic relation for the pure FPP could be improved by including the noise parameter [23, 24, 27, 31]. The model also has an approximately exponentially decaying auto-correlation function and a power law PSD, in accordance with measurements from the scrape-off layer of a range of

fusion plasmas [26–28, 43–45]. While measurement noise may always be present to some extent, the degree of noise can be quantified by the power spectrum. Any additional noise present in the PDF may then be due to dynamical noise.

Using synthetically generated time series with experimentally relevant parameters, we have shown that the PDF of

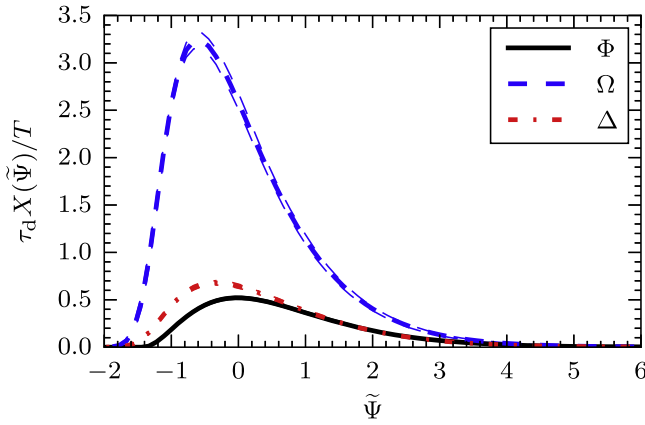


Figure 15. Mean rate of level crossings (thick lines) and the maximal deviation from the mean (thin lines) for synthetically generated time series with $\gamma = 2$, $\epsilon = 5 \times 10^{-2}$ and $\theta = 2 \times 10^{-2}$.

the FPP with additional noise is indeed capable of separating a process with noise from a noise-less process, and the sample variance of the estimated parameters provide an indication of the error of estimation. In addition, it was shown that estimating the intermittency parameter and the noise parameter from the estimated relative fluctuation level and skewness of the process is comparable in accuracy to estimating the parameters from the estimated PDF of the process, while using the estimated skewness and flatness gives a much worse estimate. It was furthermore shown that the auto-correlation function can provide an estimate of the characteristic decay time of the FPP, and is also capable of separating the types of noise. In the case of observational noise, the auto-correlation function can also provide an estimate of the noise level, but can not do this in the case of dynamical noise. The PSD has comparable capabilities to the auto-correlation function, but care must be taken in its use, in particular in the case of marginal sampling frequency. Lastly, the rate of level crossings was proposed as another way to differentiate the types of noise. This method takes advantage of the visual differences in the noise demonstrated in figure 4, and was shown to be capable of detecting both the presence of noise and differentiating the types of noise.

It should be noted that the results presented here were not a parameter scan, and some of the conclusions may depend on the model parameters and especially the duration of the signal. For instance, from figure 6, it is evident that the rate of level crossings for $\tilde{\Omega}$ is dependent on θ . In applying the methods from this contribution to experimental data, one should carry out a full Monte Carlo analysis as presented here, using parameters estimated from the PDF and auto-correlation function of the data set.

In this contribution, we have presented an extension of a reference model for intermittent fluctuations in physical systems. Its main application will be for analysis of fluctuation data time series from probe measurements in the scrape-off layer of magnetically confined plasmas, although it is sufficiently general to be applicable elsewhere. The previous analysis of experimental data demonstrate that significant noise levels can be present, and the extension of the FPP

presented here is indeed necessary in order to reliably estimate the underlying parameters of the process. While the noise itself may be insignificant for the large-amplitude fluctuations, the inclusion of the noise parameter has a large effect on the rate of level crossings and may affect the correlation functions and spectra of the process. This model has no built-in long range correlations, interaction between pulses or other non-stationary effects, meaning it could serve as a useful null hypothesis for investigations of such effects.

Acknowledgments

This work was supported with financial subvention from the Research Council of Norway under grant 240510/F20. Discussions with H L Pécseli and F C Geisler are gratefully acknowledged.

Appendix A. List of symbols and results

A.1. Time series

-
- Φ The filtered Poisson process.
 - N Normally distributed, uncorrelated noise with vanishing mean and unit standard deviation.
 - Y An Ornstein–Uhlenbeck process with vanishing mean and unit standard deviation.
 - X A collective term for either σN or σY , where σ is a noise intensity parameter.
 - Ω Denotes $\Phi + \sigma N$.
 - Δ Denotes $\Phi + \sigma Y$.
 - Ψ Denotes $\Phi + X$, it is a collective term for both Δ and Ω .
-

We use the normalization

$$\bar{\bullet} = \frac{\bullet - \langle \bullet \rangle}{\bullet_{\text{rms}}}. \quad (\text{A1})$$

A.2. Moments and probability density functions

The PDF of Φ is a Gamma distribution with shape parameter γ and scale parameter $\langle A \rangle$, given by equation (8). Denoting the mean of the process Ψ by $\langle \Psi \rangle$, its standard deviation by Ψ_{rms} , its skewness by S_{Ψ} and its flatness by F_{Ψ} , we have

$$\langle \Psi \rangle = \gamma \langle A \rangle, \quad (\text{A2})$$

$$\Psi_{\text{rms}}^2 = (1 + \epsilon) \gamma \langle A \rangle^2, \quad (\text{A3})$$

$$S_{\Psi} = \frac{2}{(1 + \epsilon)^{3/2} \gamma^{1/2}}, \quad (\text{A4})$$

$$F_{\Psi} = 3 + \frac{6}{(1 + \epsilon)^2 \gamma}. \quad (\text{A5})$$

where $\epsilon = X_{\text{rms}}^2 / \Phi_{\text{rms}}^2$. Setting $\epsilon = 0$ in the above equation gives the moments of Φ .

With the PDFs of Φ and X given in equations (8) and (25), and using $\sigma^2 = \gamma \epsilon \langle A \rangle^2$, we have

$$P_{\Psi}(\psi; \gamma, \langle A \rangle, \epsilon) = 2^{-\gamma/2} (\gamma \epsilon)^{\gamma/2-1} \langle A \rangle^{-1} \times \exp\left(-\frac{\psi^2}{2\gamma \epsilon \langle A \rangle^2}\right) \times \left\{ \frac{(\gamma \epsilon)^{1/2}}{2^{1/2} \Gamma((1+\gamma)/2)} \times M\left(\frac{\gamma}{2}, \frac{1}{2}; \frac{1}{2\epsilon} \left(\frac{\psi}{\gamma^{1/2} \langle A \rangle} - \gamma^{1/2} \epsilon\right)^2\right) + \frac{\gamma^{1/2}}{\Gamma(\gamma/2)} \left(\frac{\psi}{\gamma^{1/2} \langle A \rangle} - \gamma^{1/2} \epsilon\right) \times M\left(\frac{1+\gamma}{2}, \frac{3}{2}; \frac{1}{2\epsilon} \left(\frac{\psi}{\gamma^{1/2} \langle A \rangle} - \gamma^{1/2} \epsilon\right)^2\right) \right\}, \quad (\text{A6})$$

where $M(a, b; z)$ is the confluent hypergeometric function of the first kind, for parameters a and b and argument z [46]. Using the normalization in equation (7), we have that

$$\tilde{\Psi} = (1 + \epsilon)^{-1/2} \left(\frac{\Psi}{\gamma^{1/2} \langle A \rangle} - \gamma^{1/2} \epsilon \right). \quad (\text{A7})$$

We then have

$$P_{\tilde{\Psi}}(\tilde{\psi}) = (1 + \epsilon)^{1/2} \gamma^{1/2} \langle A \rangle \times P_{\Psi}(\gamma^{1/2} \langle A \rangle [(1 + \epsilon)^{1/2} \tilde{\psi} + \gamma^{1/2} \epsilon]), \quad (\text{A8})$$

giving

$$P_{\tilde{\Psi}}(\tilde{\psi}; \gamma, \epsilon) = \left(\frac{\gamma}{2}\right)^{\gamma/2} \epsilon^{\gamma/2-1} (1 + \epsilon)^{1/2} \times \exp\left(-\frac{[(1 + \epsilon)^{1/2} \tilde{\psi} + \gamma^{1/2} \epsilon]^2}{2\epsilon}\right) \times \left\{ \frac{\epsilon^{1/2}}{2^{1/2} \Gamma((1+\gamma)/2)} \times M\left(\frac{\gamma}{2}, \frac{1}{2}; \frac{1}{2\epsilon} [(1 + \epsilon)^{1/2} \tilde{\psi} + (1 - \epsilon)\gamma^{1/2} \epsilon]^2\right) + \frac{(1 + \epsilon)^{1/2} \tilde{\psi} + (1 - \epsilon)\gamma^{1/2} \epsilon}{\Gamma(\gamma/2)} \times M\left(\frac{1+\gamma}{2}, \frac{3}{2}; \frac{1}{2\epsilon} [(1 + \epsilon)^{1/2} \tilde{\psi} + (1 - \epsilon)\gamma^{1/2} \epsilon]^2\right) \right\}. \quad (\text{A9})$$

This expression is independent of $\langle A \rangle$ due to the normalization of Ψ . Comparing this distribution to a realization of the process gives γ and ϵ as fit parameters. In the limit of vanishing ϵ , this expression reduces as expected to a Gamma distribution for Ψ .

A.3. Power spectral densities

The PSD of a random process $\Phi(t)$ is defined as

$$\mathcal{S}_{\Phi}(\omega) = \lim_{T \rightarrow \infty} \langle |\mathcal{F}_T[\Phi](\omega)|^2 \rangle, \quad (\text{A10})$$

where

$$\mathcal{F}_T[\Phi_K](\omega) = \frac{1}{\sqrt{T}} \int_0^T dt \exp(-i\omega t) \Phi(t) \quad (\text{A11})$$

is the Fourier transform of the random variable over the domain $[0, T]$. Analytical functions which fall rapidly enough to zero (such as the Green function G given in equation (13)) have the Fourier transform

$$\mathcal{F}[G](\omega) = \int_{-\infty}^{\infty} ds G(s) \exp(-i\omega s) \quad (\text{A12})$$

and the inverse transform

$$G(\tau) = \mathcal{F}^{-1}[\mathcal{F}[G](\omega)](\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \exp(i\omega\tau) \mathcal{F}[G](\omega). \quad (\text{A13})$$

We thus have the relevant PSDs,

$$\mathcal{S}_{\Phi}(\omega) = \Phi_{\text{rms}}^2 \frac{2\tau_d}{1 + \tau_d^2 \omega^2} + 2\pi \langle \Phi \rangle^2 \delta(\omega), \quad (\text{A14})$$

$$\mathcal{S}_{\Delta}(\omega) = \Delta_{\text{rms}}^2 \frac{2\tau_d}{1 + \tau_d^2 \omega^2} + 2\pi \langle \Delta \rangle^2 \delta(\omega), \quad (\text{A15})$$

$$\mathcal{S}_{\Omega}(\omega) = \Omega_{\text{rms}}^2 \frac{2\tau_d}{1 + \epsilon} \left[\frac{1}{1 + \tau_d^2 \omega^2} + \frac{\epsilon}{2} \theta \right] + 2\pi \langle \Omega \rangle^2 \delta(\omega). \quad (\text{A16})$$

A.4. Auto-correlation functions

The auto-correlation of a random variable Φ is the inverse Fourier transform of its PSD, $R_{\Phi}(\tau) = \mathcal{F}^{-1}[\mathcal{S}_{\Phi}(\omega)](\tau)$, where the inverse Fourier transform is given by equation (A13). We thus have

$$R_{\Phi}(\tau) = \Phi_{\text{rms}}^2 \exp\left(-\frac{|\tau|}{\tau_d}\right) + \langle \Phi \rangle^2, \quad (\text{A17})$$

$$R_{\Delta}(\tau) = \Delta_{\text{rms}}^2 \exp\left(-\frac{|\tau|}{\tau_d}\right) + \langle \Delta \rangle^2, \quad (\text{A18})$$

$$R_{\Omega}(\tau) = \frac{\Omega_{\text{rms}}^2}{1 + \epsilon} \left[\exp\left(-\frac{|\tau|}{\tau_d}\right) + \epsilon \left(1 - \frac{|\tau|}{\theta\tau_d}\right) \Theta\left(1 - \frac{|\tau|}{\theta\tau_d}\right) \right] + \langle \Omega \rangle^2. \quad (\text{A19})$$

Appendix B. Derivation of the PSD and auto-correlation function

Results for the autocorrelation function and PSD of a white noise process or Ornstein–Uhlenbeck process are numerous in the literature, see for instance [32, 47, 48]. The same results for FPP are also readily available, see e.g. [5, 17, 20, 22]. For completeness, we present full derivations in this appendix.

B.1. PSD of the FPP

To find the PSD of the FPP, we start from equation (14), and take the Fourier transform

$$\begin{aligned}\mathcal{F}_T[\Phi_K](\omega) &= \frac{1}{\sqrt{T}} \int_0^T dt \exp(-i\omega t) \Phi_K(t) \\ &= \frac{1}{\sqrt{T}} \int_0^T dt \exp(-i\omega t) \int_{-\infty}^{\infty} ds G(s) f_K(t-s).\end{aligned}\quad (\text{B1})$$

where we have exchanged the functions in the convolution given by equation (14). A change of variables $u(t) = t - s$ gives

$$\begin{aligned}\mathcal{F}_T[\Phi_K](\omega) &= \int_{-\infty}^{\infty} ds G(s) \exp(-i\omega s) \\ &\quad \times \frac{1}{\sqrt{T}} \int_{-s}^{T-s} du f_K(u) \exp(-i\omega u).\end{aligned}\quad (\text{B2})$$

Note that $G(s)$ is only non-zero for positive s and is negligible after a few τ_d . Moreover, since no pulses arrive for negative times, $f_K(u) = 0$ for $u < 0$. Assuming $T/\tau_d \gg 1$, we can therefore approximate the limits of the second integral in equation (B2) as $u \in [0, T]$, and the two integrals become independent. This gives

$$\mathcal{F}_T[\Phi_K](\omega) = \mathcal{F}[G](\omega) \mathcal{F}_T[f_K](\omega), \quad (\text{B3})$$

where the Fourier transform of $G(s)$ is given in equation (A12). The PSD of the stationary process Φ is thus

$$\begin{aligned}\mathcal{S}_\Phi(\omega) &= \lim_{T \rightarrow \infty} \langle |\mathcal{F}_T[\Phi_K](\omega)|^2 \rangle \\ &= \langle |\mathcal{F}[G](\omega)|^2 \rangle \lim_{T \rightarrow \infty} \langle |\mathcal{F}_T[f_K](\omega)|^2 \rangle,\end{aligned}\quad (\text{B4})$$

where $\mathcal{S}_\Phi(\omega)$ is independent of K , since the average is over all random variables. The Fourier transform of the Green functions $\mathcal{F}[G](\omega)$ is easily computed as $(i\omega + 1/\tau_d)^{-1}$, giving

$$\langle |\mathcal{F}[G](\omega)|^2 \rangle = \frac{\tau_d^2}{1 + \tau_d^2 \omega^2}. \quad (\text{B5})$$

We also readily find the Fourier transform of the forcing,

$$\mathcal{F}_T[f_K](\omega) = T^{-1/2} \sum_{k=1}^K A_k \exp(-i\omega t_k). \quad (\text{B6})$$

Multiplying this expression with its complex conjugate and averaging over all random variables gives

$$\begin{aligned}\langle |\mathcal{F}_T[f_K](\omega)|^2 \rangle &= \sum_{K=0}^{\infty} P_K(K; T, \tau_w) \frac{1}{T} \sum_{k=1}^K \sum_{l=1}^K \\ &\quad \times \int_0^T \frac{dt_l}{T} \dots \int_0^T \frac{dt_K}{T} \\ &\quad \times \int_0^{\infty} dA_1 P_A(A_1) \dots \int_0^{\infty} \\ &\quad \times dA_K P_A(A_K) A_k A_l \exp(i\omega(t_l - t_k)).\end{aligned}\quad (\text{B7})$$

In this equation, there are K terms where $k = l$ and $K(K - 1)$ terms where $k \neq l$, for which all events are independent.

Summing over all these terms, we have

$$\begin{aligned}\langle |\mathcal{F}_T[f_K](\omega)|^2 \rangle &= \sum_{K=0}^{\infty} P_K(K; T, \tau_w) \left[\frac{K}{T} \langle A^2 \rangle \right. \\ &\quad \left. + \frac{K(K-1)}{T^3} \langle A \rangle^2 \int_0^T dt \int_0^T ds \exp(i\omega(t-s)) \right],\end{aligned}\quad (\text{B8})$$

giving

$$\begin{aligned}\langle |\mathcal{F}_T[f_K](\omega)|^2 \rangle &= \sum_{K=0}^{\infty} P_K(K; T, \tau_w) \left[2 \frac{K}{T} \langle A \rangle^2 \right. \\ &\quad \left. + \frac{2K(K-1)}{T^3} \langle A \rangle^2 \frac{1 - \cos(\omega T)}{\omega^2} \right]\end{aligned}\quad (\text{B9})$$

where we have used that for an exponentially distributed variable, $\langle A^n \rangle = n! \langle A \rangle^n$. Thus, averaging over all K and using $\langle K \rangle = T/\tau_w$ and $\langle K(K-1) \rangle = T^2/\tau_w^2$ gives

$$\langle |\mathcal{F}_T[f_K](\omega)|^2 \rangle = \frac{2}{\tau_w} \langle A \rangle^2 + \frac{2}{\tau_w^2} \langle A \rangle^2 \frac{1 - \cos(\omega T)}{T\omega^2}. \quad (\text{B10})$$

The second term in this equation resembles a Dirac delta function in the limit $T \rightarrow \infty$. With the appropriate normalization $\int_{-\infty}^{\infty} d\omega \delta(\omega) = 1$ this gives

$$\lim_{T \rightarrow \infty} \langle |\mathcal{F}_T[f_K](\omega)|^2 \rangle = \frac{2}{\tau_w} \langle A \rangle^2 + \frac{2\pi}{\tau_w^2} \langle A \rangle^2 \delta(\omega), \quad (\text{B11})$$

which together with equation (B5) gives the PSD of the FPP as

$$\begin{aligned}\mathcal{S}_\Phi(\omega) &= 2\gamma \langle A \rangle^2 \frac{\tau_d}{1 + \tau_d^2 \omega^2} + 2\pi\gamma^2 \langle A \rangle^2 \delta(\omega) \\ &= \Phi_{\text{rms}}^2 \frac{2\tau_d}{1 + \tau_d^2 \omega^2} + 2\pi \langle \Phi \rangle^2 \delta(\omega).\end{aligned}\quad (\text{B12})$$

B.2. PSD and auto-correlation function of the noise processes

Since $Y(t)$ is constructed as a convolution in the same way as the FPP, we have an analog of equation (B5) for this process:

$$\mathcal{S}_Y(\omega) = \langle |\mathcal{F}[G](\omega)|^2 \rangle \lim_{T \rightarrow \infty} \langle |\mathcal{F}_T[dW](\omega)|^2 \rangle. \quad (\text{B13})$$

With equation (B5) and the relation

$$\begin{aligned}\lim_{T \rightarrow \infty} \langle |\mathcal{F}_T[dW](\omega)|^2 \rangle &= \lim_{T \rightarrow \infty} \frac{2}{\tau_d T} \\ \left\langle \int_0^T dW(t) \exp(-i\omega t) \int_0^T dW(s) \exp(i\omega s) \right\rangle &= \frac{2}{\tau_d},\end{aligned}\quad (\text{B14})$$

we have the PSD for the Ornstein–Uhlenbeck process,

$$\mathcal{S}_Y(\omega) = \frac{2\tau_d}{1 + \tau_d^2 \omega^2}. \quad (\text{B15})$$

The PSD of $Y(t)$ is therefore Lorentzian with the same parameter as for the PSD of $\Phi(t)$, implying that the auto-correlation function of $Y(t)$ is also an exponentially decaying function with the same rate as the auto-correlation

function of $\Phi(t)$:

$$R_Y(\tau) = \exp\left(-\frac{|\tau|}{\tau_d}\right). \quad (\text{B16})$$

For observational noise we require that $N(t)$ is a unit-less variable. One way of realizing such a process is by using integrated increments of the Wiener process:

$$N(t) = \frac{1}{\Delta_t^{1/2}} \int_t^{t+\Delta_t} dW(s), \quad (\text{B17})$$

where Δ_t is the sampling time (in this case, each sample $N[n] = N(n\Delta_t)$, $n = 0, 1, 2, \dots$ is normally and independently, identically distributed with zero mean and unit standard deviation). In this case, the most direct route to the PSD is via the auto-correlation function. We find that

$$R_N(\tau) = \left(1 - \frac{|\tau|}{\Delta_t}\right) \Theta(\Delta_t - |\tau|), \quad (\text{B18})$$

giving

$$S_N(\omega) = \int_{-\infty}^{\infty} d\tau \exp(-i\omega\tau) R_N(\tau) = 2 \frac{1 - \cos(\Delta_t\omega)}{\Delta_t\omega^2}. \quad (\text{B19})$$

Using the normalized sampling time θ from equation (6), this expression can be written as

$$S_N(\omega) = 2 \frac{\tau_d}{\theta} \frac{1 - \cos(\theta \tau_d \omega)}{\tau_d^2 \omega^2}. \quad (\text{B20})$$

For small θ , the cosine function can be expanded around 0, and we have

$$\lim_{\theta \rightarrow 0} \frac{S_N(\omega)}{\tau_d \theta} = 1. \quad (\text{B21})$$

In this limit, $N(t)$ approaches white noise, which has a flat power spectrum. Since we have demanded $N_{\text{rms}} = \int_{-\infty}^{\infty} d\omega S_N(\omega)^2 = 1$, the higher resolution granted by $\theta \rightarrow 0$ means the total power of N is divided among a greater number of frequencies, reducing the power per frequency.

Since Δ_t is the smallest time we will observe, it makes more sense to use the i.i.d. random sequence $N[n]$ and to translate the auto-correlation to discrete time:

$$R_N[n] = \begin{cases} 1, & n = 0, \\ 0, & n \geq 1. \end{cases} \quad (\text{B22})$$

The discrete Fourier transform corresponding to $\mathcal{F}[\bullet](\omega)$ is then

$$S_N(\omega) = \Delta_t \sum_{n=-\infty}^{\infty} \exp(-i\omega n) R_N[n] = \tau_d \theta. \quad (\text{B23})$$

Intuitively, this corresponds to the case where we do not see the effects of θ , that is the regime $\theta \ll 1$. Since this is also the spectral density we will observe for a realization of the process $N(t)$, we will use equation (B23) when discussing the PSD of $\Omega(t)$.

References

- [1] Narashima R, Kumar S R, Prabhu A and Kailas S V 2007 *Phil. Trans. R. Soc.* **365** 841
- [2] Robinson P A 1995 *Phys. Plasmas* **2** 1466
- [3] Kristensen L, Casanova M, Courtney M and Troen I 1991 *Bound.-Layer Meteorol.* **55** 91
- [4] Elter Z, Jammes C, Pázit I, Pál L and Filliatre P 2015 *Nucl. Instrum. Methods Phys. Res. A* **774** 60
- [5] Rice S O 1944 *Bell Syst. Tech. J.* **23** 282
- [6] Rice S O 1945 *Bell Syst. Tech. J.* **24** 46
- [7] Endler M, Giannone L, McCormick K, Niedermeyer H, Rudyj A, Theimer G, Tsois N and the ASDEX Team 1995 *J. Nucl. Mater.* **220** 293
- [8] Carreras B A 2005 *J. Nucl. Mater.* **337** 315
- [9] Liewer P C 1985 *Nucl. Fusion* **25** 543
- [10] D'Ippolito D A, Myra J R and Zweben S J 2011 *Phys. Plasmas* **18** 060501
- [11] Zweben S J, Boedo J A, Grulke O, Hidalgo C, LaBombard B, Maqueda R J, Scarin P and Terry J L 2007 *Plasma Phys. Controlled Fusion* **49** 1
- [12] Jang J-W 2004 *J. Risk Insur.* **71** 201
- [13] Fesce R, Segal J R and Hurlbut W P 1986 *J. Gen. Physiol.* **88** 25
- [14] Hole M J, Robinson P A and Cairns I H 2004 *Phys. Rev. E* **70** 036619
- [15] Hole M J, Robinson P A and Cairns I H 2006 *Phys. Plasmas* **13** 012101
- [16] Garcia O E 2012 *Phys. Rev. Lett.* **108** 265001
- [17] Garcia O E, Kube R, Theodorsen A and Pécseli H L 2016 *Phys. Plasmas* **23** 052308
- [18] Bondesson L 1982 *Adv. Appl. Prob.* **14** 855
- [19] Pécseli H L 2000 *Fluctuations in Physical Systems* (Cambridge: Cambridge University Press)
- [20] Parzen E 1999 *Stochastic Processes* (Philadelphia, PA: Society for Industrial and Applied Mathematics)
- [21] Kotler G, Kontorovich V, Lyandres V and Primak S 1999 *Signal Process.* **75** 79
- [22] Lowen S B and Teich M C 2005 *Fractal-Based Point Processes* (New York: Wiley)
- [23] Garcia O E, Fritzner S M, Kube R, Cziegler I, LaBombard B and Terry J L 2013a *Phys. Plasmas* **20** 082309
- [24] Garcia O E, Cziegler I, Kube R, LaBombard B and Terry J L 2013b *J. Nucl. Mater.* **438** S180
- [25] Graves J P, Horacek J, Pitts R A and Hopcraft K I 2005 *Plasma Phys. Controlled Fusion* **47** L1
- [26] Kube R, Theodorsen A, Garcia O E, LaBombard B and Terry J L 2016 *Plasma Phys. Control. Fusion* **58** 054001
- [27] Theodorsen A, Garcia O E, Horacek J, Kube R and Pitts R A 2016 *Plasma Phys. Control. Fusion* **58** 044006
- [28] Garcia O E, Horacek J, Pitts R A, Nielsen A H, Fundamenski W, Naulin V and Rasmussen J J 2007 *Nucl. Fusion* **47** 667
- [29] Theodorsen A, Garcia O E, Kube R, LaBombard B and Terry J L 2017 Relationship between frequency power spectra and intermittent, large-amplitude bursts in the Alcator C-Mod scrape-off layer *Nucl. Fusion* submitted (arXiv:1702.01104)
- [30] Garcia O E and Theodorsen A 2017 *Phys. Plasmas* **24** 032309
- [31] Garcia O E, Kube R, Theodorsen A, Bak J-G, Hong S-H, Kim H-S, Pitts R A and the KSTAR Project Team 2016 *Nucl. Mater. Energy* at press (<https://doi.org/10.1016/j.nme.2016.11.008>)
- [32] Stark H and Woods J W 2012 *Probability, Statistics and Random Processes for Engineers* ed M J Horton 4th edn (Hoboken, NJ: Pearson Education)
- [33] Kontorovich V and Lyandres V 1996 *Signal Process.* **51** 121

- [34] Kube R and Garcia O E 2015 *Phys. Plasmas* **22** 012502
- [35] Theodorsen A and Garcia O E 2016 *Phys. Plasmas* **23** 040702
- [36] Biermé H and Desolneux A 2012 *J. Appl. Prob.* **49** 100
- [37] Bar-David I and Nemirovsky A 1972 *IEEE Trans. Inf. Theory* **18** 27
- [38] Barakat R 1988 *J. Opt. Soc. Am. A* **5** 1244
- [39] Hopcraft K I, Ingrey P C and Jakeman E 2007 *Phys. Rev. E* **76** 031134
- [40] Sato H, Pécseli H L and Trulsen J 2012 *J. Geophys. Res.* **117** A03329
- [41] Fattorini L, Fredriksen Å, Pécseli H L, Riccardi C and Trulsen J K 2012 *Plasma Phys. Control. Fusion* **54** 085017
- [42] Sattin F *et al* 2009 *Plasma Phys. Control. Fusion* **51** 055013
- [43] Cziegler I, Terry J L, Hughes J W and LaBombard B 2010 *Phys. Plasmas* **17** 056120
- [44] Dewhurst J M, Hnat B, Ohno N, Dendy R O, Masuzaki S, Morisaki T and Komori A 2008 *Plasma Phys. Control. Fusion* **50** 095013
- [45] Tanaka H, Ohno N, Asakura N, Tsuji Y, Kawashima H, Takamura S, Uesugi Y and the JT-60U Team 2009 *Nucl. Fusion* **49** 065017
- [46] Muller K E 2001 *Numer. Math.* **90** 179
- [47] Uhlenbeck G E and Ornstein L S 1930 *Phys. Rev.* **36** 823
- [48] Schuss Z 2010 *Theory and Applications of Stochastic Processes: An Analytical Approach* (Berlin: Springer)