Modelling Biological Clocks with Bio-PEPA: Stochasticity and Robustness for the *Neurospora crassa* Circadian Network

Ozgur E. Akman[1], Federica Ciocchetta[2], Andrea Degasperi[3], and Maria Luisa Guerriero[4]

[1] Centre for Systems Biology at Edinburgh, The University of Edinburgh, Edinburgh EH9 3JU, Scotland, UK
[2] The Microsoft Research - University of Trento Centre for Computational and Systems Biology, Trento, Italy
[3] Laboratory for Foundations of Computer Science, The University of Edinburgh, Edinburgh EH8 9AB, Scotland, UK
[4] Department of Computing Science, The University of Glasgow, Glasgow G12 8QQ, Scotland, UK

Abstract. Circadian clocks are biochemical networks, present in nearly all living organisms, whose function is to regulate the expression of specific mRNAs and proteins to synchronise rhythms of metabolism, physiology and behaviour to the 24 hour day/night cycle. Because of their experimental tractability and biological significance, circadian clocks have been the subject of a number of computational modelling studies.

In this study we focus on the simple circadian clock of the fungus *Neurospora crassa*. We use the Bio-PEPA process algebra to develop both a stochastic and a deterministic model of the system. The light on/off mechanism responsible for entrainment to the day/night cycle is expressed using discrete time-dependent events in Bio-PEPA.

In order to validate our model, we compare it against the results of previous work which demonstrated that the deterministic model is in agreement with experimental data. Here we investigate the effect of stochasticity on the robustness of the clock's function in biological timing. In particular, we focus on the variations in the phase and amplitude of oscillations in circadian proteins with respect to different factors such as the presence/absence of a positive feedback loop, and the presence/absence of light. The time-dependent sensitivity of the model with respect to some key kinetic parameters is also investigated.

1   Introduction

Circadian clocks are oscillatory gene networks developed by living organisms in order to adapt to the 24-hour day/night cycle. In general, the biochemical mechanisms regulating circadian rhythms are robust enough for approximately 24 hour oscillations to persist over a range of constant lighting and temperature conditions. Exposure to periodic external stimuli (e.g. light/dark or temperature cycles) has the effect of resetting these free-running oscillations so as to establish stable phase relationships with the forcing stimulus (circadian entrainment). This enables cyclic changes in the environment to

be anticipated, such as seasonal variations in the length of day (photoperiod) [1]. Circadian rhythms are present in nearly all eukaryotes, from mammals and plants, to insects and fungi. There is now detailed experimental data showing that these rhythms can be produced by networks of multiple, interlocked positive and negative feedback loops in which the protein product of a gene modulates expression of either its own transcript or that of another target gene in the network [2].

Several mathematical models have been proposed in recent years to describe the specific oscillation-generating mechanisms in a range of different organisms. These include the fruit fly *Drosophila melanogaster* [3, 4], the plant *Arabidopsis thaliana* [5, 6] and the mouse *Mus musculus* [7, 8]. Here we focus on the fungus *Neurospora crassa*, which possesses one of the most comprehensively studied circadian networks [9]. In recent years, a number of mathematical models of the *Neurospora* clock have been developed, including continuous-deterministic models that are described in terms of ordinary differential equations (ODEs) [10–14], as well as discrete-stochastic models [15, 16]. Such models have been used successfully to explore the relationship between the architecture of the *Neurospora* circadian network and the robustness of its function in biological timing.

Within this theme, our aim in this work is to investigate the effect of stochastic fluctuations on the performance of the *Neurospora* clock. While deterministic models are good approximations of real biochemical systems when the number of molecules is sufficiently high, at low copy numbers the effect of random fluctuations becomes significant and so stochasticity needs to be taken into account to obtain a faithful representation of the real biochemical system [17]. To explore the effect of these fluctuations on circadian timing in *Neurospora*, we implement a discrete-stochastic version of a continuous ODE model previously developed to investigate the entrainment of the clock by light and temperature [13, 18]. We use the ODE representation of this clock to validate our stochastic model and to highlight the differences between deterministic and stochastic representations of the network. In particular, where previous stochastic studies have concentrated mainly on the unforced (free-running) *Neurospora* clock, modelling entrainment as a weak modulation of transcription [15, 16], here we investigate how stochasticity affects the robustness of circadian oscillations for a more realistic model which explicitly incorporates elements of the light-signalling pathway [13, 18]. We exploit discrete time-dependent events to represent light/dark cycles and analyse the behaviour of the system under different light conditions and in the absence of a core feedback loop. As part of this analysis, we use a novel sensitivity analysis method to determine the time within the circadian cycle at which a given phase marker is most responsive to parameter variations.

We use Bio-PEPA [19, 20] as our modelling language. Bio-PEPA is a process algebra recently developed for modelling biochemical systems. Among its key strengths as a language for systems biology is the fact that it is equipped with different semantics, enabling both continuous-deterministic and discrete-stochastic representations of the same model description to be automatically generated. Another important feature of Bio-PEPA is that it permits the definition of generic rate laws. This allows the specification of complex kinetic formulae, such as those used in the ODE representation of the original *Neurospora* model (see Sect. 4 below). In addition, time-dependent events

can be easily incorporated, enabling periodic external stimuli such as light/dark cycles to be represented in a straightforward manner.

The rest of the paper is structured as follows. The circadian clock of *Neurospora crassa* and the Bio-PEPA model of the clock are described in Sect. 2 and Sect. 4, respectively. Bio-PEPA is introduced in Sect. 3. In Sect. 5 the simulation and analysis results are presented. Finally, in Sect. 6 we report some concluding remarks.

2   The Circadian Clock of *Neurospora Crassa*

Neurospora exhibits a 22 hour rhythm in asexual spore formation (conidiation) when grown in constant darkness (DD). The conidiation rhythm is a key clock output which can be entrained by both light and temperature [21]. In natural 24 hour cycles of alternating light and dark (LD), the phase of entrainment (judged by the time of conidiation onset) coincides with the middle of the night in both long and short days, providing a simple, biologically relevant measure of circadian function [22, 13, 18].

The core, multi-loop genetic oscillator believed to underlie the conidiation rhythm is formed by the rhythmic gene *frequency (frq)* and the constitutively expressed gene *white collar-1 (wc-1)* [9]. The protein product of the *white collar-1* gene, WC-1, comprises the positive element of a central negative feedback loop, activating transcription of *frq*. The protein product of the *frq* gene, FRQ, is the negative element of the loop, interacting with *frq*-bound WC-1 to inhibit *frq* expression [23, 24]. In addition to its role as a transcriptional inhibitor, FRQ positively regulates expression of WC-1, giving a positive feedback loop that interlocks with the central loop [25]. Light entrains the clock by promoting the binding of a flavin chromophore to WC-1, resulting in a light-activated form which enhances *frq* transcription [24].
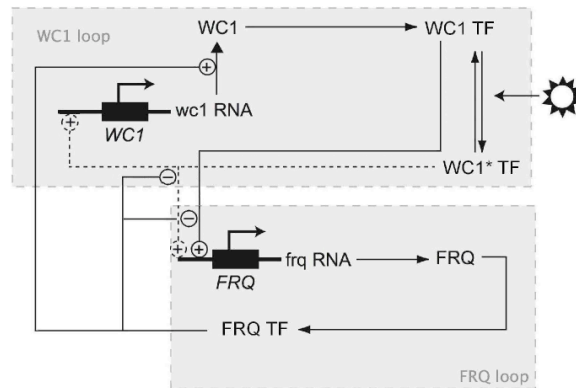
Fig. 1. A schematic representation of the gene network underlying the model of the *Neurospora* clock. WC-1* represents light-activated WC-1. The dashed lines indicate light-dependent gene-protein interactions.

3

A network diagram for the model of the core oscillator that we consider here is shown in Fig. 1. For the ODE representation of the model presented in [13, 18], the repressive action of FRQ transcription factor (hereafter called *PF*) on *frq*-bound WC-1 transcription factor (*PW*) and *frq*-bound light-activated WC-1 transcription factor (*PWL*) was assumed to occur through a noncompetitive inhibition process modelled using Hill kinetics. Hill kinetics were also used to describe the upregulation of WC-1 translation by *PF* as well as the light-dependent increase in the transcription of *wc-1* mRNA (*MW*), necessary to simulate loss-of-function *wc-1* mutants [13]. Michaelis-Menten kinetics were used to describe enzyme-mediated degradation of mRNA and protein, while the conversion of *PW* to *PWL* was modelled as a reversible first order mass-action reaction $PW \rightleftharpoons PWL$ with a light-dependent forward rate. The light input to the ODEs took the form of a smoothly differentiable function that switches rapidly between 0 and 1 at dawn ($t = t_{dawn}$), and from 1 back to 0 at dusk ($t = t_{dusk}$), modelling the lighting protocol commonly used in circadian experiments.

In order to obtain oscillatory behaviour, a delay was introduced into the central negative feedback loop by assuming that just-translated FRQ protein (*E1F*) is modified into a second intermediate protein (*E2F*) before being converted into transcription factor [13]. The conversion processes $MF \rightarrow E1F \rightarrow E2F \rightarrow PF$, which include translation of FRQ from *frq* mRNA (*MF*), were each modelled as first order mass-action reactions. Similarly, a delay was introduced into the positive feedback loop by introducing two intermediate WC-1 protein species (*E1W* and *E2W*), and describing the conversions $MW \rightarrow E1W \rightarrow E2W \rightarrow PW$ with first order kinetics.

The ODE representation of the model comprises 9 equations with 34 kinetic parameters. The parameters were fitted to gene and protein expression time series in DD and LD using a bipartite optimisation method developed for high-dimensional computational biology models [5]. This technique combines a random parameter search with simulated annealing to minimise a qualitative cost function that assesses the goodness-of-fit of the model to key experimental data sets [5, 13]. For the ODE model, the best parameter set was taken to be that yielding the smallest cost function score following the application of the optimisation scheme to 50 million randomly distributed points in the 34-dimensional parameter space [18]. This optimal parameter set yielded a good fit to each of the target time series, and also reproduced the variation in entrainment phase with photoperiod observed experimentally [18]. We use the same parameter set here for the Bio-PEPA representation of the model detailed in Sect. 4.

3 Bio-PEPA

In this section we give a short description of Bio-PEPA [19, 20], a language that has recently been developed for the modelling and analysis of biological systems. The main components of a Bio-PEPA system are the *species components*, describing the behaviour of each species, and the *model component*, describing the interactions between the species and initial amounts.

The syntax of the Bio-PEPA components is defined as:

$$S ::= (\alpha, \kappa) \text{ op } S \mid S + S \mid C \quad \text{with op} = \downarrow \mid \uparrow \mid \oplus \mid \ominus \mid \odot \qquad P ::= P \bowtie_{\mathcal{L}} P \mid S(x)$$

4

where S is the *species component* and P is the *model component*. In the prefix term (α, κ) op S, κ is the *stoichiometry coefficient* of species S in reaction α, and the *prefix combinator* "op" represents the role of S in the reaction. Specifically, $\downarrow$ indicates a *reactant*, $\uparrow$ a *product*, $\oplus$ an *activator*, $\ominus$ an *inhibitor* and $\odot$ a generic *modifier*. We can use the shorthand notations (α, κ) op and α op for (α, κ) op S and $(\alpha, 1)$ op S, respectively.

The operator "+" expresses the choice between possible actions, and the constant C is defined by an equation $C \stackrel{def}{=} S$. The process $P \underset{\mathcal{L}}{\bowtie} Q$ denotes synchronisation between components P and Q; the set $\mathcal{L}$ determines those activities on which the operands are forced to synchronise, with $\bowtie$ denoting a synchronisation on all common action types. In the model component $S(x)$, the parameter $x \in \mathbb{R}$ represents the initial concentration (or the number of molecules in a discrete-stochastic setting). The reader is referred to [19] for further details on the language and its semantics.

Recently Bio-PEPA has been extended to incorporate events [26], constructs that represent changes in the system due to some triggering conditions. This allows biochemical perturbations to the system to be represented, such as the timed introduction of reagents or the modulation of system components by external stimuli. A Bio-PEPA event has the form (*id, trigger, event_assignment, delay*), where *id* is the event name, *trigger* is a mathematical expression involving the components of the Bio-PEPA model and/or time, *event_assignment* is a list of assignments causing some changes to elements in the system, and *delay* is either 0 (*immediate events*) or a positive real value (*delayed events*).

A Bio-PEPA system representing a biochemical network consists of a set of sequential components, a model component, and context (defining information such as kinetics rates, parameters, locations, and events). Its formal definition is the following:

Definition 1. *A Bio-PEPA system $\mathcal{P}$ is a 8-tuple $\langle t, \mathcal{L}, \mathcal{N}, \mathcal{K}, \mathcal{F}_R, Comp, P, Events \rangle$, where: t is time, $\mathcal{L}$ is the set of locations, $\mathcal{N}$ is the set of (auxiliary) information for the species, $\mathcal{K}$ is the set of parameters, $\mathcal{F}_R$ is the set of functional rates, Comp is the set of species components, P is the model component and Events is the set of events.*

Bio-PEPA offers a formal representation of biochemical systems, on which different kinds of analysis can be carried out, through the defined mappings into continuous-deterministic and discrete-stochastic modelling languages. The Bio-PEPA language is supported by software tools which automatically process Bio-PEPA models and generate other representations in forms suitable for different kinds of analysis [19, 27]. In particular, the generated simulation model can be executed using the Dizzy simulation tool [28], in which both stochastic simulation algorithms and differential equation solvers are implemented.

4   The Bio-PEPA Model of the Circadian Clock

In the following we provide an overview of the Bio-PEPA model for the circadian clock described in Sect. 2. The full model is reported in the Supplementary Material.

The clock is characterised by robust entrainment to light/dark cycles. Light entrains the clock by modulating general kinetic laws different from mass-action that abstract complex sequences of more elementary steps [13, 18]. These features can be easily

represented in Bio-PEPA using events and functional rates. From the Bio-PEPA description of the clock we can derive both the model for stochastic simulation and the related system of differential equations.

In order to derive a stochastic model, the continuous concentration values of ODE models must be translated into discrete numbers of molecules. In general, assuming concentrations are expressed in *molars* (M), the initial amounts must be multiplied by the factor $N_A \cdot V$ (where N_A is the Avogadro number and V is the volume of the compartment in which the reactions take place), and the kinetic parameters must be rescaled accordingly (see [29] for details). For the *Neurospora* model, as the compartment size and absolute concentration values are not known to any great accuracy, we follow the approach used in [15] and introduce a generic scaling factor Ω that modulates the number of molecules. Specifically, concentrations are turned into discrete numbers of molecules by multiplying them by Ω, and the kinetic parameters are also rescaled by Ω (see the Supplementary Material for details).

Reactions and kinetic laws. Each reaction is associated with an action type and a functional rate, expressing the kinetic law. For instance, the transcription of *MF* under upregulation by *PWL* and inhibition by *PF* is represented by the action type *transcription_MF_by_PWL* and the kinetic law:

$$transcription\_MF\_by\_PWL \;=\; \frac{a_1 \cdot PWL^n}{(1 + (PF/b_1)^g) \cdot (PWL^n + b_2{}^n)}$$

Species. Each biological species is abstracted by a Bio-PEPA species component. Below we report the definition of *PF*; the other species are described similarly.

$$PF \stackrel{def}{=} (transcription\_MF\_by\_PW, 1) \ominus \; + \; (transcription\_MF\_by\_PWL, 1) \ominus \; +$$
$$(transformation\_E2F\_to\_PF, 1) \uparrow \; + \; (degradation\_PF, 1) \downarrow \; +$$
$$(translation\_E1W\_by\_PF, 1) \oplus$$

PF is involved in five reactions: it is an inhibitor of the transcription of *MF* with and without the influence of light (first line), a product of the transformation from *E2F* to *PF*, a reactant in the degradation of *PF* (second line) and an activator of the translation of *E1W* (last line). Note the use of shorthand notation in the definition of *PF*.

The full system is described in terms of the model component

$$MF(mf0) \underset{*}{\bowtie} E1F(e1f0) \underset{*}{\bowtie} E2F(e2f0) \underset{*}{\bowtie} PF(pf0) \underset{*}{\bowtie} MW(mw0) \underset{*}{\bowtie}$$
$$PW(pw0) \underset{*}{\bowtie} PWL(pwl0) \underset{*}{\bowtie} E1W(e1w0) \underset{*}{\bowtie} E2W(e2w0)$$

where the values in parenthesis are the initial values for the species.

Events. Entrainment by light/dark cycles is represented by events in Bio-PEPA. In the initial state the system is in dark conditions and, therefore, the transformation from the protein *PW* to the form activated by light (*PWL*) is not possible. This is represented by setting the kinetic parameter r_1 for the transformation reaction $PW \rightarrow PWL$ equal to 0. At dawn, the reaction is suddenly activated and therefore

r_1 is reset to its maximum value 5.1759. At dusk the reaction is deactivated again by resetting r_1 to 0. This periodic sequence of parameter changes is represented by the following set of immediate events

$$Events = [(dawn_i; t = t_{dawn} \cdot i; r_1 = 5.1759; 0),$$
$$(dusk_i; t = t_{dusk} \cdot i; r_1 = 0; 0), \ i = 1, 2, \ldots, D \ ]$$

where D is the number of simulated days, and t_{dawn} and t_{dusk} are the times of the day at which dawn and dusk occur, respectively. By changing the values of t_{dawn} and t_{dusk} we can simulate the effect of changing the photoperiod ($t_{dusk} - t_{dawn}$), a key entrainment parameter for the *Neurospora* clock [22, 18]. Here we focus on two conditions: constant darkness (DD) and alternating 12-hour cycles of light and dark (12:12 LD).

5 Model Analysis

In this section we present the validation of our model against the original ODE representation and we illustrate some analysis results. We use a version of the Dizzy simulator [28] developed at the University of Edinburgh [30], which extends the tool with sensitivity analysis techniques and additional simulation methods. The time-dependent events in the Bio-PEPA model are translated into time-dependent reaction rates in the Dizzy model (defined in terms of the step function *theta*, which is predefined in Dizzy), and we use the Gibson-Bruck stochastic simulation algorithm [31]. The choice of this algorithm is due to its efficiency (in simulation time) with respect to other stochastic simulators and to the fact it supports time-dependent rates.

5.1 Validation of the Model

As a preliminary step, we validate the Bio-PEPA model by comparing it against the original deterministic representation [18, 13]. In Fig. 2(a) and Fig. 2(b) we show the comparison for the DD system and for 12:12 LD cycles respectively. In each graph we plot three time-series: the behaviour of the original model (dashed lines), the solution of the system of ODEs generated by the Bio-PEPA model (solid lines), and the average behaviour over 10 stochastic simulation runs with scaling factor $\Omega = 10000$ (points). The variables plotted are the clock outputs *frq* mRNA (MF), *wc-1* mRNA (MW), total FRQ protein ($FP = E1F + E2F + PF$), and total WC-1 protein ($WP = E1W + E2W + PW + PWL$).

The scaling factor $\Omega = 10000$ was chosen with the purpose of having a reasonably high number of molecules to minimise the effect of stochastic fluctuations [29]. Consequently, in both DD and LD systems, the time-series resulting from the deterministic and the stochastic analysis of the Bio-PEPA models are in very close agreement, with the stochasticity almost unnoticeable, despite the small number of simulation runs. Comparing the deterministic time-series generated by our model against that of the original model, we observe that in the DD system they are in perfect agreement, while a small difference can be observed in the LD system (especially for WP): the reason

7

(a) Simulations in DD.   (b) Simulations in LD.
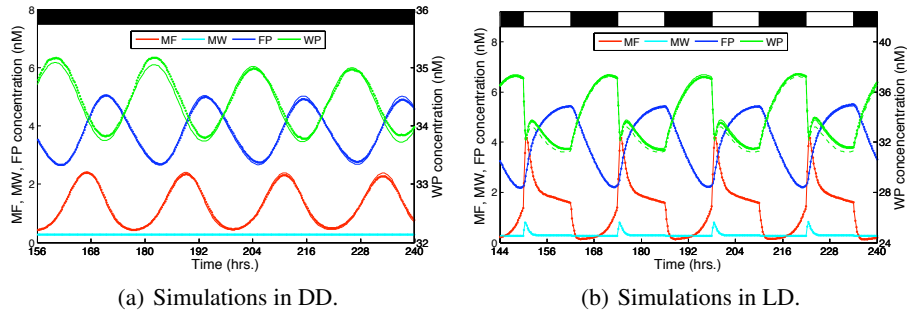
Fig. 2. Comparisons of the original ODE model of the *Neurospora* clock with the deterministic and stochastic Bio-PEPA representations. Black bars represent lights-off and white bars lights-on.

for this difference resides in the different ways in which the light switch is modelled: a smooth function in the original ODE model versus discrete events in the Bio-PEPA model. A similar agreement was obtained for different photoperiods (results not shown).

5.2 Effect of the Scaling Factor Ω on Stochasticity

Higher values of the scaling factor Ω correspond to larger molecular populations in the stochastic model, yielding smaller stochastic fluctuations [29]. We have seen in the previous section that for $\Omega = 10000$ the stochasticity is reduced to such a point that even with a small number of simulations runs, the average simulation behaviour is near-identical to the deterministic behaviour. As a consequence, the higher the scaling factor, the more regular the circadian oscillations will be, whereas we expect the effects of noise to be more evident with a smaller scaling factor. Figure 3 shows the average oscillations in total FRQ protein FP for different values of Ω (10, 100 and 1000). We observe that for $\Omega = 10$ the average behaviour of the stochastic system differs significantly from the FP oscillation in the deterministic system, yielding unstable oscillations that are inconsistent with the stable cycling of FRQ observed experimentally [9]. By contrast, for $\Omega = 100$ and $\Omega = 1000$, regular oscillating dynamics are obtained. We also note that the average oscillations for $\Omega = 100$ and $\Omega = 1000$ are very close to the deterministic solution, indicating that increasing Ω in this range only affects the variability about the average. We consider a scaling factor $\Omega = 1000$ in the remainder of the work. Similar results (with slightly higher variability) were obtained in all cases with $\Omega = 100$.

5.3 Investigating the Role of Positive Feedback

In this section, we study how the positive feedback loop of the *Neurospora* network affects its stochastic behaviour. A previous analysis of the deterministic model identified the upregulation of WC-1 production by FRQ transcription factor, controlled by the parameter a_7, as a key regulatory process in the light entrainment of the clock [18].
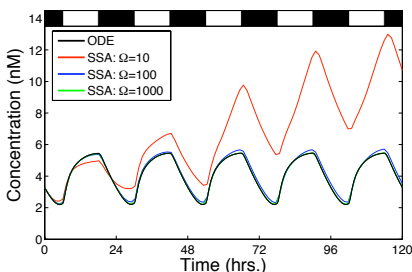
Fig. 3. The effect of increasing Ω on the LD system. Plotted are time-series showing oscillations in total FRQ protein FP averaged over 1000 stochastic simulations for Ω=10, 100 and 1000 (coloured lines). Each time series has been normalised by Ω to enable comparison with the FP oscillations obtained from the deterministic model (black line).
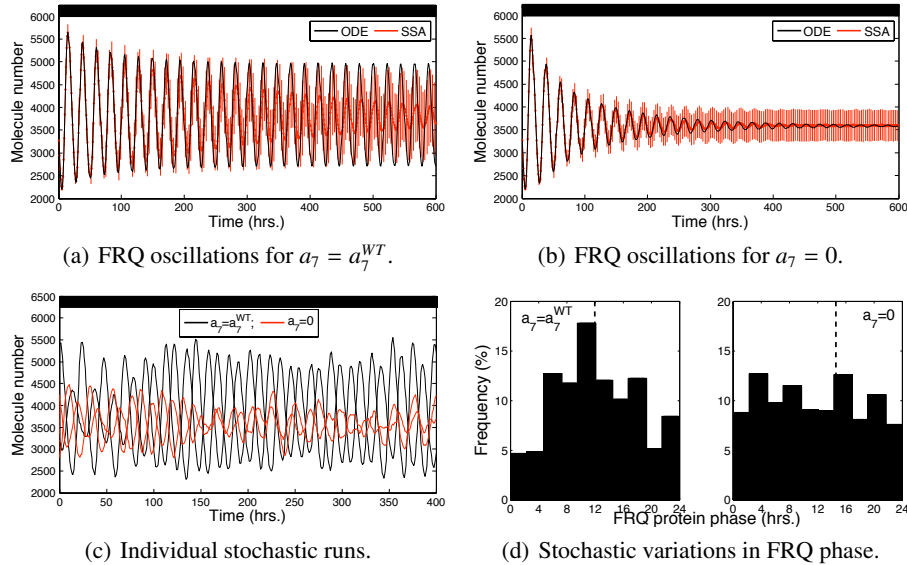
With a_7 at its nominal value a_7^{WT} = 2.4695, the ODE model yields the correct experimental responses to changes in photoperiod, with the phase of total FRQ protein FP tracking the middle of the night in both long and short days [18]. Knocking out positive feedback by reducing a_7 to 0 destroys self-sustained oscillations in DD by pushing the deterministic model through a supercritical Hopf bifurcation [18]. This is in agreement with the loss of free-running conidiation rhythms reported in experiments [32]. The destruction of the DD limit cycle has a significant effect on the light responses of the model, yielding a system that is unable to respond to changes in photoperiod during long days. This suggests a role for the positive loop in promoting robustness against seasonal photoperiod changes [18]. Here, we compare the behaviour of the stochastic and deterministic models for DD and 12:12 LD cycles in the presence and absence of the positive loop, focusing on the resulting changes to the FRQ oscillation in each case.

The Effect of Removing Positive Feedback on the DD System. Figure 4(a) shows the difference between the deterministic and stochastic behaviour for the unperturbed network in DD. While the ODE model exhibits self-sustained FRQ oscillations, the average oscillation generated by the stochastic system damps to a constant value. This is a consequence of the individual realisations of the stochastic model going out of phase with each other, as can be seen in Fig. 4(c) and Fig. 4(d). This phase diffusion in the free-running system, characterised by a phase distribution spanning the full circadian cycle, agrees with previous stochastic analysis of circadian models [15].

Setting the positive feedback strength a_7 to 0 yields damped FRQ oscillations in the ODE model, as a_7 is below the Hopf bifurcation value (Fig. 4(b)). Individual realisations of the stochastic model, however, are still oscillatory, albeit with smaller amplitudes compared to the unperturbed network (Fig. 4(c)). Again, the average FRQ oscillation damps to a constant value as a consequence of phase diffusion.

The persistence of self-sustained oscillations when positive feedback is removed demonstrates that stochasticity can introduce greater robustness against modifications to the network architecture. This finding is consistent with models of the mammalian

clock for which simulated knockouts that are arrhythmic in ODE implementations can become rhythmic when stochasticity is incorporated [8].

(a) FRQ oscillations for $a_7 = a_7^{WT}$.

(b) FRQ oscillations for $a_7 = 0$.

(c) Individual stochastic runs.

(d) Stochastic variations in FRQ phase.

Fig. 4. Changes to the FRQ protein oscillation FP resulting from the removal of positive feedback in DD. Black lines in (a) and (b) denote the solution of the deterministic system, red points the average of 1000 stochastic simulations and red lines the corresponding standard deviations. In (d), the phase of FRQ protein was taken as the time at which FP has decreased to its half-maximum value over the interval $576 \leq t \leq 600$. The histograms show the distribution of this phase marker over all 1000 runs of the stochastic model. Dotted lines denote the phase of FRQ in the corresponding deterministic systems.

The Effect of Removing Positive Feedback on the LD System. Finally, we consider the 12:12 LD system and examine the effect of setting a_7 to 0 on the oscillatory behaviour of the model. Comparing Fig. 4 and Fig. 5, it is clear that for both the unperturbed system and the positive loop knockout, entrainment regularises the dynamics, markedly reducing the variability of oscillations compared to the free-running system (similar findings were reported for a model of the *Drosophila* clock in [15]). In both cases, there is relatively little phase diffusion, as evidenced by phase distributions that are concentrated about their corresponding deterministic values (Fig. 5(c)). Interestingly, although removal of positive feedback shifts the mean value of FRQ phase, consistent with the analysis of the deterministic system ([18]), the variation about the mean is unaffected (the standard deviation is 0.3720 for the unperturbed system and 0.3238 for the loop knockout). This demonstrates that the positive loop is able to buffer the

10

clock against environmental variations (seasonal changes in photoperiod) without degrading its robustness to stochastic fluctuations in the chemical reactions comprising the oscillatory mechanism.

(a) $a_7 = a_7^{WT}$.

(b) $a_7 = 0$.

(c) Stochastic variations in FRQ phase.
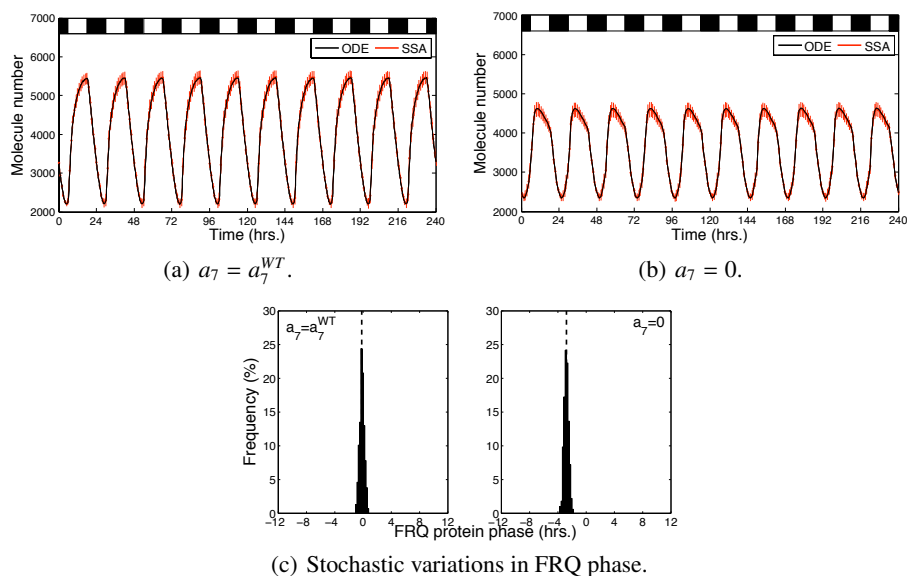
Fig. 5. Changes to the FRQ protein oscillation FP resulting from the removal of positive feedback in 12:12 LD cycles. Black lines in (a) and (b) denote the solution of the deterministic system, red points the average of 1000 stochastic simulations and red lines the corresponding standard deviations. The phase of the FRQ oscillation in (c) was taken as the time at which FP decreases to its half-maximum value, identified recently as a molecular correlate of conidiation onset [22, 13, 18]. As in Fig. 4, the phase marker was computed over the interval $576 \leq t \leq 600$ for all stochastic realisations of the model. The dotted lines denote FRQ phase for the ODE model.

5.4 Sensitivity Analysis for the 12:12 LD System

Sensitivity Analysis (SA) aims to identify the relationships between the inputs and outputs of mathematical models of biochemical networks [33]. A key goal is the production of Sensitivity Indices (SI) that quantify these relationships, revealing which factors are the most influential with respect to model outcome. The most widespread SA method is "one-at-a-time" (OAT). Given a mathematical model with parameters set to those considered the most likely (also called *nominal parameters*), each parameter is perturbed individually by a fixed value or by a percentage of its nominal value, and the change in the output(s) of interest measured. OAT has seen widespread use in ODE models of biochemical interactions; this has included circadian networks for which a standard

11

approach has been to compute the sensitivities of period and amplitude over one cycle of the oscillation [16, 11, 12].

In [34] this method has been extended to stochastic models. In this case, the output at a given time is not just a value representing the amount of a species as in ODEs; it is, instead, a set of possible values, obtained from independent stochastic simulations. The SA extension has been obtained by substituting the difference between perturbed and nominal output values employed in the traditional approach with a difference measure based on the density distribution surface of the output, estimated with a suitable number of simulations. An estimate of this density distance based on stochastic simulations can be obtained using histogram distance, as originally presented in [35]. This stochastic version of OAT therefore applies when one is interested in observing the change in the distribution of the amount of a particular species at a given time.

Here we apply the traditional approach to the means of the stochastic simulations and also consider the extended approach, based on histogram distance. These are indeed complementary, as the former does not incorporate any notion of stochastic variability, while the latter quantifies the likelihood of having the same distribution in both the perturbed and unperturbed systems. Moreover, a feature of the histogram distance is that its value will always be 2 when there is no overlapping of the distributions, making the traditional approach still necessary to determine sensitivities for large displacements.

Figure 6 summarises the result of the local sensitivity analysis obtained by changing a subset of parameters predicted to have a significant effect on entrained phase. Each parameter was incremented by 10%, and the results shown are averages over 1000 simulation runs for $\Omega = 100$ and $\Omega = 1000$. It can clearly be seen that both the density- and average-based sensitivity measures vary significantly over the circadian cycle for the parameters considered. For both measures, the most sensitive parameters are a_4, d_1 and d_3 representing the maximum rates of light-independent *wc-1* transcription, *frq* degradation and *wc-1* degradation respectively. All 3 parameters yield maximum sensitivities with respect to the average-based measure around dusk, when FP is close to its peak value. By contrast, maximum sensitivities with respect to the density-based measure occur at the time when FP decreases to its half-maximum value, a molecular correlate of conidiation onset [22, 13, 18]. This demonstrates that in terms of the average FRQ oscillation, the marker of entrained phase most responsive to evolutionary parameter variations is peak FRQ phase. However, when stochastic variations in the FRQ waveform are considered, the most responsive marker is the phase of the FRQ half-maximum. As it is the latter which correlates with physiological entrained phase for the *Neurospora* clock, this analysis suggests that stochastic fluctuations in FRQ expression may have been an important contributing factor to the selection of the half-maximum as a phase marker in nature.

As previously mentioned, average-based SA computes differences in terms of the average FP value, while density-based SA also considers the stochastic variability. Therefore, the differences in the sensitivities when using different scaling factors can only be captured consistently by using density-based SA. For instance, similar average-based sensitivities are obtained for $\Omega = 100$ and $\Omega = 1000$ (cf. Figs. 6(a) and 6(c)), while the density-based sensitivities are in general smaller for $\Omega = 100$ as a conse-

(a) Average-based sensitivity: $\Omega = 100$.

(b) Density-based sensitivity: $\Omega = 100$.

(c) Average-based sensitivity: $\Omega = 1000$.

(d) Density-based sensitivity: $\Omega = 1000$.
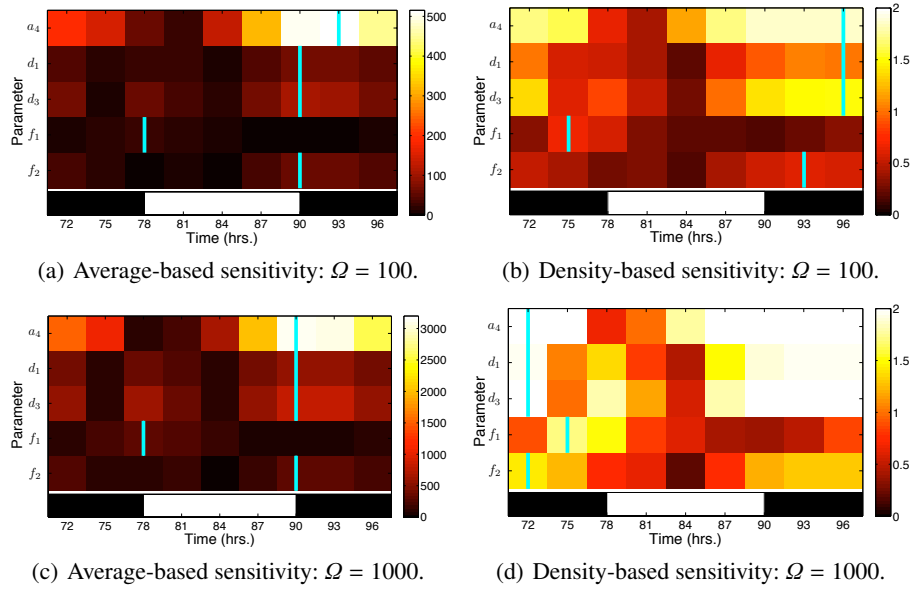
Fig. 6. Local sensitivity to parameter variation. Sensitivities were computed every 3 hours over one circadian cycle ($72 \leq t \leq 96$). The color gradient represents the difference in the amount of FP between the nominal and modified parameter sets in each case (with increasing sensitivity going from black to white). In all panels, the light blue lines denote the cycle phase at which maximum sensitivity is attained for a given parameter. Because of the periodic behaviour of the system, qualitatively similar sensitivities are obtained over other 24h intervals (see Fig. S1 in the Supplementary Material).

13

quence of the higher variance causing greater overlapping in the probability distributions of FP (cf. Figs. 6(b) and 6(d)).

6   Conclusions

In this work we presented and studied a stochastic model of the circadian clock in *Neurospora crassa* under two different light conditions: constant darkness and 12:12 light/dark cycles. We used Bio-PEPA as our modelling language. This language allowed us to represent in a straightforward manner two features of the system: complex kinetic laws and time-dependent events representing cyclic light/dark conditions. The model was validated against an existing ODE representation describing key behaviours observed in laboratory experiments, including the variation of entrained phase with photoperiod. We presented some analysis results illustrating the differences between deterministic and stochastic representations of the clock. In particular, we investigated the effect of removing the positive feedback loop, previously identified as a significant factor in the determination of entrained phase. We found that while removal of positive feedback destroys self-sustained free-running (DD) rhythms in the deterministic system, oscillations with significant amplitude persist when stochastic fluctuations are considered, demonstrating the greater robustness of the oscillatory mechanism in the stochastic model. In addition, we showed that knocking out the loop has little effect on the stochastic variability of entrained phase for a given photoperiod, suggesting that positive feedback can be used to tune the phase-photoperiod relationship without introducing greater variation due to noise amplification.

Finally, we considered sensitivity analysis techniques in order to identify the most influential parameters on the circadian function of the 12:12 LD system. We focused on the variations in FRQ expression resulting from perturbations to 5 putatively sensitive parameters, applying a local sensitivity method at different time points within the 24-hour cycle. By using a novel stochastic sensitivity measure based on histogram distance, we found that the FRQ waveform is maximally sensitive at the time it reaches is half-maximum level, a molecular correlate for conidiation onset. We commented that this implicates stochasticity as a potential factor in the selection of this seemingly complex phase marker by evolution, rather than the phase of peak FRQ expression that is predicted to be maximally sensitive when variations in average FRQ level are considered. We conclude that while the local method we used only focuses around a specific point in the parameter space, it can still be informative, giving an idea about the impact of parameter changes on the behaviour of the system. In the future, we plan to apply some global methods in order to explore the full parameter space (or a meaningful subset of it) and to quantify the relationships between different parameters.

The use of stochastic simulation with our model merits some discussion, as it is characterised by some non-elementary reactions with complex kinetic laws, abstracting sets of interactions whose details are unknown. The use of Gillespie's stochastic simulation algorithm (or its variants, such as Gibson-Bruck [31]) in the case of general kinetic laws has been discussed by several authors [36–38]. Rao and Arkin [36] showed that this approach is valid for some specific kinetic laws, such as Michaelis-Menten and competitive inhibition. On the other hand, in [37] the authors demonstrated

that this extension of Gillespie's algorithm is not always appropriate. Here, we applied stochastic simulation paying particular attention to the interpretation of the simulation results and to their validation: in Sect. 5 we showed that the behaviour we obtain using our stochastic model is in agreement with the known behaviour of the system, and therefore we conclude that in this case the use of stochastic simulation is appropriate.

Acknowledgements

Supplementary Material

The Bio-PEPA source files, the generated Dizzy files, some further information on the model and supplementary figure S1 are available online on the Bio-PEPA web page `http://homepages.inf.ed.ac.uk/jeh/Bio-PEPA/ACDG_CMSB09_Supplement.zip`.

References

1. Dunlap, J., Loros, J., DeCoursey, P.: Chronobiology: Biological Timekeeping. Sinauer, Sunderland (2003)
2. Young, M., Kay, S.: Time zones: a comparative genetics of circadian clocks. Nat Rev Genet **2**(9) (2001) 702–15
3. Tyson, J., Hong, C., Thron, C., Novak, B.: A simple model of circadian rhythms based on dimerization and proteolysis of PER and TIM. J Biophys **77** (1999) 2411–2417
4. Ueda, H., Hagiwara, M., Kitano, H.: Robust oscillations within the interlocked feedback model of *Drosophila* circadian rhythm. J Theor Biol **210** (2001) 401–406
5. Locke, J., Kozma-Bognar, L., Gould, P., Fehér, B., Kevei, E., Nagy, F., Turner, M., Hall, A., Millar, A.: Experimental validation of a predicted feedback loop in the multi-oscillator clock of *Arabidopsis thaliana*. Mol Sys Biol **2** (2006) 59
6. Zeilinger, M., Farré, E., Taylor, S., Kay, S., Doyle, F.: A novel computational model of the circadian clock in *Arabidopsis* that incorporates PRR7 and PRR9. Mol Sys Biol **2**(60) (2006)
7. Leloup, J., Goldbeter, A.: Toward a detailed computational model for the mammalian circadian clock. Proc Natl Acad Sci USA **100** (2003) 7051–7056
8. Forger, D., Peskin, C.: Model based conjectures on mammalian clock controversies. Theor Biol **230**(4) (2004) 533–539
9. Loros, J., Dunlap, J.: Genetic and molecular analysis of circadian rhythms in *Neurospora*. Annu Rev Physiol **63** (2001) 757–794
10. Leloup, J., Gonze, D., Goldbeter, A.: Limit cycle models for circadian rhythms based on transcriptional regulation in *Drosophila* and *Neurospora*. J Biol Rhythms **14**(6) (1999) 433–448

11. Francois, P.: A model for the *Neurospora* circadian clock. Biophys J **88**(4) (2005) 2369–2383
12. Ruoff, P., Loros, J., Dunlap, J.: The relationship between FRQ-protein stability and temperature compensation in the *Neurospora* circadian clock. Proc Natl Acad Sci USA **102**(49) (2005) 17681–6
13. Akman, O., Locke, J., Tang, S., Carré, I., Millar, A., Rand, D.: Isoform switching facilitates period control in the *Neurospora crassa* circadian clock. Mol Sys Biol **4** (2008) 64
14. Hong, C., Jolma, I., Loros, J., Dunlap, J., Ruoff, P.: Simulating dark expressions and interactions of *frq* and *wc-1* in the *Neurospora* circadian clock. Biophys J **94**(4) (2008) 1221–32
15. Gonze, D., Halloy, J., Goldbeter, A.: Robustness of circadian rhythms with respect to molecular noise. Proc Natl Acad Sci USA **99**(2) (2002) 673–8
16. Smolen, P., Baxter, D., Byrne, J.: Reduced models of the circadian oscillators in *Neurospora crassa* and *Drosophila melanogaster* illustrate mechanistic similarities. OMICS **7**(4) (2003) 337–54
17. McAdams, H., Arkin, A.: Stochastic mechanisms in gene expression. Proc Natl Acad Sci USA **94**(3) (1997) 814–9
18. Akman, O., Rand, D., Brown, P., Millar, A.: Robustness from flexibility in the fungal circadian clock. Submitted (2009)
19. Ciocchetta, F., Hillston, J.: Bio-PEPA: a Framework for the Modelling and Analysis of Biochemical Networks. Theoretical Computer Science (2009, to appear)
20. Ciocchetta, F., Hillston, J.: Bio-PEPA: an extension of the process algebra PEPA for biochemical networks. In: Proc. of FBTC'07. Volume 194 of ENTCS. (2008) 103–117
21. Vitalini, M., de Paula, R., Park, W., Bell-Pedersen, D.: The rhythms of life: circadian output pathways in *Neurospora*. J Biol Rhythms **21**(6) (2006) 432–44
22. Merrow, M., Boesl, C., Ricken, J., Messerschmitt, M., Goedel, M., Roenneberg, T.: Entrainment of the *Neurospora* circadian clock. Chronobiol Int **23**(1-2) (2006) 71–8
23. Merrow, M., Franchi, L., Dragovic, Z., Gorl, M., Johnson, J., Brunner, M., Macino, G., Roenneberg, T.: Circadian regulation of the light input pathway in *Neurospora crassa*. EMBO J **20**(3) (2001) 307–15
24. Froehlich, A., Loros, J., Dunlap, J.: Rhythmic binding of a WHITE COLLAR-containing complex to the *frequency* promoter is inhibited by FREQUENCY. Proc Natl Acad Sci USA **100**(10) (2003) 5914–9
25. Cheng, P., Yang, Y., Liu, Y.: Interlocked feedback loops contribute to the robustness of the *Neurospora* circadian clock. Proc Natl Acad Sci USA **98**(13) (2001) 7408–13
26. Ciocchetta, F.: Bio-PEPA with events. Transactions on Computational Systems Biology (2009, to appear)
27. Bio-PEPA Home Page: `http://www.biopepa.org/`
28. Dizzy Home Page: `http://magnet.systemsbiology.net/software/Dizzy`
29. Gillespie, D.: Exact stochastic simulation of coupled chemical reactions. J Phys Chem **81**(25) (1977) 2340–2361
30. Dizzy Edinburgh version: `http://homepages.inf.ed.ac.uk/stg/software/Dizzy/`
31. Gibson, M., Bruck, J.: Efficient Exact Stochastic Simulation of Chemical Systems with Many Species and Many Channels. The Journal of Chemical Physics **104** (2000) 1876–1889
32. Schafmeier, T., Káldi, K., Diernfellner, A., Mohr, C., Brunner, M.: Phosphorylation-dependent maturation of *Neurospora* circadian clock protein from a nuclear repressor toward a cytoplasmic activator. Cell **20**(3) (2006) 297–306
33. Heinrich, R., Schuster, S.: The Regulation of Cellular Systems. Chapman & Hall (1996)
34. Degasperi, A., Gilmore, S.: Sensitivity Analysis of Stochastic Models of Bistable Biochemical Reactions. In: Formal Methods for Computational Systems Biology (SFM'08). Volume 5016 of LNCS. Springer-Verlag (2008) 1–20
35. Cao, Y., Petzold, L.: Accuracy limitations and the measurements of errors in the stochastic simulation of chemically reacting systems. J. Comput. Phys. **212**(1) (2006) 6–24

36. Rao, C.V., Arkin, A.P.: Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. J Chem Phys **118**(11) (2003) 4999–5010
37. Bundschuh, R., Hayot, F., Jayaprakash, C.: Fluctuations and Slow Variables in Genetic Networks. Biophys. J. **84** (2003) 1606–1615
38. Cao, Y., Gillespie, D., Petzold, L.: Accelerated Stochastic Simulation of the Stiff Enzyme-Substrate Reaction. Journal of Chemical Physics **123** (2005) 144917–144929