Stochastic properties of the plant circadian clock: *Electronic Supplementary Material*

Maria Luisa Guerriero[1], Alexandra Pokhilko[2], Aurora Piñas Fernández[2],
Karen J. Halliday[1,2], Andrew J. Millar[1,2] and Jane Hillston[1,3]

[1] Centre for Systems Biology at Edinburgh, University of Edinburgh,
C.H. Waddington Building, King's Buildings Campus, Mayfield Road, EH9 3JD, Edinburgh, UK

[2] School of Biological Sciences, University of Edinburgh,
King's Buildings Campus, Mayfield Road, EH9 3JH, Edinburgh, UK

[3] School of Informatics, University of Edinburgh,
Informatics Forum, 10 Crichton Street, EH8 9AB, Edinburgh, UK

A Experimental estimation of the system size

Our qPCR measurements showed that the clock components are expressed at similar relative levels: peak mRNA concentrations of the clock genes are estimated to lie within a ten-fold range (Supplementary Table 1), which demonstrates reasonably good correspondence with the deterministic model (Supplementary Figure 1). The absolute level of expression can be roughly estimated based on the only available data on mRNA level of the clock component ZTL, which was reported to reach the level of 7000 molecules per ng of total RNA [2]. Using the value of 450 μg of total RNA per g of fresh weight, routinely achieved in our lab in separations of *Arabidopsis* tissues, and the estimate of 25 millions of cells per g of fresh weight (personal communication with Ronan Sulpice), the peak expression level of ZTL is around 126 molecules per cell. Since ZTL is closer to the lower limit of expression of the clock genes (Supplementary Table 1), the average copy numbers of clock messengers can be estimated as several hundreds of molecules per cell. Although the levels of the clock proteins in *Arabidopsis* are largely unknown, our recent quantitative data on the clock proteins in *Ostreococcus tauri* showed peak abundance in the range of 150–400 molecules per cell, which is comparable to the above estimates for the *Arabidopsis* clock mRNA levels [16]. Our stochastic simulations with $\Omega = 100$ show good match to our qPCR data (Supplementary Figure 1).

B The Bio-PEPA language and its modelling and analysis tool

Bio-PEPA [7] is a stochastic process algebra, recently developed for the modelling and analysis of biochemical systems. In a Bio-PEPA model of a biochemical system, each molecular species (i.e. proteins, genes, mRNAs) is represented by a process, the state of the system at a given time is given by the current amount of the molecular species, and the result of the occurrence of a biochemical reaction is a change in the available amount of the involved molecules. Processes interact by means of shared action names representing reactions and specifying their role in the reaction (reactant, product, catalyser, inhibitor, etc.) and their stoichiometric coefficient for that reaction; the effect of a reaction occurrence is to decrease the amount of reactants and increase the amount of products. Species amounts in Bio-PEPA can either be

	LL		LD	
	expression level	phase (h)	expression level	phase (h)
CCA1	63.4 ± 4.27	2	417.1 ± 32.26	0
LHY	140 ± 25.87	2	670.2 ± 11.07	0
PRR9	25.9 ± 11.75	9	39.7 ± 5.83	4
PRR7	82.8 ± 21.81	9	54.2 ± 2.93	8
PRR5	144.9 ± 16.76	12	185.1 ± 32.43	10
GI	80 ± 9.63	12	106.6 ± 18.63	10
TOC1	20.2 ± 0.84	14	89.3 ± 22.08	10
LUX	40.9 ± 3.16	11	46.7 ± 4.57	10
ELF3	83.8 ± 20.09	12		
ELF4	101.1 ± 35.67	14		
ZTL	38.7 ± 21.19	0		

Table 1: The expression level of the clock genes at their maximum normalised to the level of actin. Gene expression was measured by real-time PCR from Col-4 13-day-old seedlings grown in 1/2 MS without sucrose. The seedlings were entrained for 7 days at 22°C in white light diurnal cycles LD 12:12. Then, the seedlings were transferred to experimental conditions of 17°C in either continuous white light (LL) or LD 12:12. The samples of seedlings grown in LL were collected every 2h during 3 days starting on third day of LL. The samples of seedlings grown in LD were collected every 2h on the fifth day at 17°C. To compare between gene concentrations we assigned the PCR instrument's cycle number crossing point (Cp = 20) to an arbitrary concentration of 1000 on a template linear regression curve (dilution curve) for each gene-specific PCR primer pair. The phase indicates the time of the maximum RNA level after dawn or predicted dawn. The standard deviation was calculated from 3 days in LL and 3 biological replicates in LD.
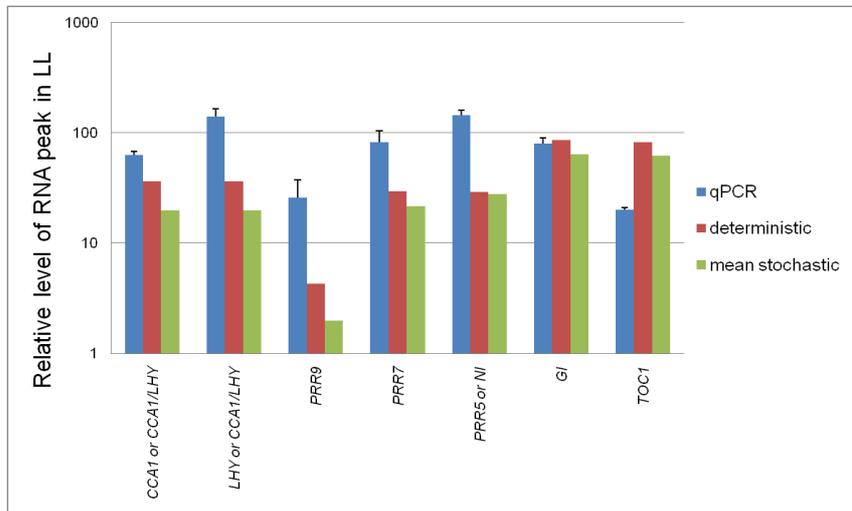

Figure 1: Relative levels of clock gene RNAs in models and experiments. The relative levels of the peaks in clock gene transcripts on the third day in constant light are shown for the qPCR data of Supplementary Table 1 (blue), for simulations of the deterministic model (red) and for the mean of 1000 runs of the stochastic model (green). Note that the qPCR data for all genes fall within a ten-fold range, with relatively low levels of PRR9 and TOC1. PRR9 levels are also low in the models, but TOC1 levels are higher.

concentrations (continuous semantics) or molecule counts (discrete semantics), hence allowing both numerical methods based on differential equations and stochastic analysis.

The syntax of a Bio-PEPA model is similar to that of a system of ODEs. Essentially, for each biochemical species, the modeller specifies the set of reactions in which the species is involved and the role of the species in each reaction; moreover, each reaction is associated with a kinetic law which specifies the rate of occurrence of that reaction. For a detailed presentation of Bio-PEPA syntax and semantics the reader is referred to [7]; here we illustrate the basic concepts using the following simple example.

A reaction $S \xrightarrow{E} P$ which converts a substrate molecule S into a product molecule P catalysed by an enzyme E is modelled in Bio-PEPA as

$$S \stackrel{def}{=} r_1 \downarrow$$
$$P \stackrel{def}{=} r_1 \uparrow$$
$$E \stackrel{def}{=} r_1 \oplus$$
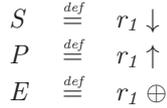
where r_1 is a name associated with the reaction, its kinetic law is defined by the Michaelis-Menten kinetics

$$r_1 = \frac{k_{\mathrm{cat}} \cdot E \cdot S}{k_{\mathrm{M}} + S}$$

and k_{cat} and k_{M} are the reaction kinetic constants.

This notation represents the fact that S, P and E are all involved in the occurrence of reaction r_1, and that the result of the occurrence of r_1 is to decrease the total amount of S molecules ($\downarrow$) and increase the total amount of P molecules ($\uparrow$); the role of the enzyme E is to speed up the reaction, but its amount is unaffected ($\oplus$).

A software framework for model development and analysis for Bio-PEPA, called the Bio-PEPA Eclipse Plug-in, is available from [3] and as part of the Systems Biology Software Infrastructure, available from [15]. It automatically performs a number of checks to identify syntactic errors and enables modellers to apply further static analysis — such as the identification of invariants, sources and sinks — with the purpose of detecting possible modelling errors, and dynamic time-series analysis, using stochastic simulation and the solution of differential equations. More information on the Bio-PEPA language and on the features of the tool and its import/export formats can be found in [3, 6, 8].

C   The Bio-PEPA model of the clock

The main difference between the ODE model in [13] and our Bio-PEPA model concerns the rescaling which was done in order to translate the molecular concentrations used in the original continuous deterministic model into molecular counts, which is needed in order to obtain a realistic discrete stochastic model. In the original ODE model the initial concentrations are given in arbitrary relative units, since their absolute values are not known, and consequently the scaling factor is also unknown and we estimated it experimentally as described in Appendix A of the Electronic Supplementary Material, and computationally as described in Section 3.1 of the main text.

We report here the definition of the species and kinetic laws of the Bio-PEPA model. Appendix F of the Electronic Supplementary Material contains the full Bio-PEPA model with kinetic parameters and initial state, and Appendix G is an equivalent representation of the Bio-PEPA model as saved in SBML format by the Bio-PEPA Eclipse Plug-in.

$$
\begin{aligned}
LHY\_mRNA &\stackrel{def}{=} cL\_m\_trscr{\uparrow} \; + \; cL\_m\_degr{\downarrow} \\
LHY\_prot &\stackrel{def}{=} cL\_trsl{\uparrow} \; + \; cL\_modif{\downarrow} \; + \; cL\_degr{\downarrow} \\
LHY\_prot\_modif &\stackrel{def}{=} cL\_modif{\uparrow} \; + \; cLm\_degr{\downarrow} \\
TOC1\_mRNA &\stackrel{def}{=} cT\_m\_trscr{\uparrow} \; + \; cT\_m\_degr{\downarrow} \\
TOC1\_prot &\stackrel{def}{=} cT\_trsl{\uparrow} \; + \; cT\_modif \oplus \; + cTm\_degr1{\downarrow} \; + \; cTm\_degr2{\downarrow} \\
TOC1\_prot\_modif &\stackrel{def}{=} cT\_modif{\uparrow} \; + \; cTm\_degr{\downarrow} \\
Y\_mRNA &\stackrel{def}{=} cY\_m\_trscr{\uparrow} \; + \; cY\_m\_degr{\downarrow} \\
Y\_prot &\stackrel{def}{=} cY\_trsl{\uparrow} \; + \; cY\_degr{\downarrow} \\
P &\stackrel{def}{=} cP\_trsl{\uparrow} \; + \; cP\_degr{\downarrow} \\
PRR9\_mRNA &\stackrel{def}{=} cP9\_m\_trscr{\uparrow} \; + \; cP9\_m\_degr{\downarrow} \\
PRR9\_prot &\stackrel{def}{=} cP9\_trsl{\uparrow} \; + \; cP9\_degr{\downarrow} \\
PRR7\_mRNA &\stackrel{def}{=} cP7\_m\_trscr{\uparrow} \; + \; cP7\_m\_degr{\downarrow} \\
PRR7\_prot &\stackrel{def}{=} cP7\_trsl{\uparrow} \; + \; cP7\_degr{\downarrow} \\
PRR5\_NI\_mRNA &\stackrel{def}{=} cNI\_m\_trscr{\uparrow} \; + \; cNI\_m\_degr{\downarrow} \\
PRR5\_NI\_prot &\stackrel{def}{=} cNI\_trsl{\uparrow} \; + \; cNI\_degr{\downarrow} \\
GI\_mRNA &\stackrel{def}{=} cG\_m\_trscr{\uparrow} \; + \; cG\_m\_degr{\downarrow} \\
GI\_prot &\stackrel{def}{=} cG\_trsl{\uparrow} \; + \; cG\_degr{\downarrow} \; + \\
& \quad cG\_cZTL\_assoc\_f{\downarrow} \; + \; cG\_cZTL\_assoc\_b{\uparrow} \\
ZTL &\stackrel{def}{=} cZTL\_trsl{\uparrow} \; + \; cZTL\_degr{\downarrow} \; + \\
& \quad cG\_cZTL\_assoc\_f{\downarrow} \; + \; cG\_cZTL\_assoc\_b{\uparrow} \\
GI{:}ZTL &\stackrel{def}{=} cG\_cZTL\_assoc\_f{\uparrow} \; + \; cG\_cZTL\_assoc\_b{\downarrow} \; + \; cZG\_degr{\downarrow}
\end{aligned}
$$

$$cL\_m\_trscr = \Omega \cdot \frac{(g_1 \cdot \Omega)^a}{(g_1 \cdot \Omega)^a + (PRR9\_prot + PRR7\_prot + PRR5\_NI\_prot)^a} \cdot$$
$$\left(n_0 \cdot light + \frac{q_1}{\Omega} \cdot light \cdot P + n_1 \cdot \frac{TOC1\_prot\_modif^b}{TOC1\_prot\_modif^b + (g_2 \cdot \Omega)^b} \right)$$

$$cL\_m\_degr = LHY\_mRNA \cdot (m_1 \cdot light + m_2 \cdot dark)$$

$$cL\_trsl = (p_1 \cdot light + p_2 \cdot dark) \cdot LHY\_mRNA$$

$$cL\_degr = m_3 \cdot LHY\_prot$$

$$cL\_modif = p_3 \cdot \frac{LHY\_prot^c}{LHY\_prot^c + (g_3 \cdot \Omega)^c} \cdot \Omega$$

$$cLm\_degr = m_4 \cdot LHY\_prot\_modif$$

$$cT\_m\_trscr = \Omega \cdot \left(n_2 \cdot \frac{Y\_prot^d}{Y\_prot^d + (g_4 \cdot \Omega)^d} + n_3 \right) \cdot \frac{(g_5 \cdot \Omega)^e}{(g_5 \cdot \Omega)^e + LHY\_prot^e}$$

$$cT\_m\_degr = TOC1\_mRNA \cdot m_5$$

$$cT\_trsl = p_4 \cdot TOC1\_mRNA$$

$$cT\_degr1 = (m_6 \cdot light + m_7 \cdot dark) \cdot TOC1\_prot \cdot (p_5 \cdot ZTL + GI{:}ZTL) / \Omega$$

$$cT\_degr2 = m_8 \cdot TOC1\_prot$$

$$cT\_modif = p_{15} \cdot \frac{TOC1\_prot^f}{TOC1\_prot^f + (g_6 \cdot \Omega)^f} \cdot \Omega$$

$$cTm\_degr = (m_{25} \cdot light + m_{26} \cdot dark) \cdot TOC1\_prot\_modif$$

$$cY\_m\_trscr = \Omega \cdot \left(\frac{q_2}{\Omega} \cdot light \cdot P + \frac{(g_7 \cdot \Omega)^s}{(g_7 \cdot \Omega)^s + TOC1\_prot^s} \cdot \frac{(g_{16} \cdot \Omega)^g}{(g_{16} \cdot \Omega)^g + LHY\_prot^g} \right) \cdot$$
$$(n_5 \cdot light + n_6 \cdot dark)$$

$$cY\_m\_degr = Y\_mRNA \cdot m_9$$

$$cY\_trsl = Y\_mRNA \cdot p_6$$

$$cY\_degr = m_{10} \cdot Y\_prot$$

$$cP\_trsl = p_7 \cdot dark \cdot (\Omega - P)$$

$$cP\_degr = m_{11} \cdot P \cdot light$$

$$cP9\_m\_trscr = \Omega \cdot \left(\frac{q_3}{\Omega} \cdot light \cdot P + \frac{(g_8 \cdot \Omega)^h}{(g_8 \cdot \Omega)^h + TOC1\_prot^h} \cdot \left(n_4 + n_7 \cdot \frac{LHY\_prot^i}{LHY\_prot^i + (g_9 \cdot \Omega)^i} \right) \right)$$

$$cP9\_m\_degr = m_{12} \cdot PRR9\_mRNA$$

$$cP9\_trsl = p_8 \cdot PRR9\_mRNA$$

$$cP9\_degr = (m_{13} \cdot light + m_{22} \cdot dark) \cdot PRR9\_prot$$

$$cP7\_m\_trscr = \Omega \cdot \left(n_8 \cdot \frac{(LHY\_prot + LHY\_prot\_modif)^j}{(g_{10} \cdot \Omega)^j + (LHY\_prot + LHY\_prot\_modif)^j} + n_9 \cdot \frac{PRR9\_prot^k}{(g_{11} \cdot \Omega)^k + PRR9\_prot^k} \right)$$

$$cP7\_m\_degr = m_{14} \cdot PRR7\_mRNA$$

$$cP7\_trsl = p_9 \cdot PRR7\_mRNA$$

$$cP7\_degr = PRR7\_prot \cdot (m_{15} \cdot light + m_{23} \cdot dark)$$

$$cNI\_m\_trscr = \Omega \cdot \left(n_{10} \cdot \frac{LHY\_prot\_modif^l}{(g_{12} \cdot \Omega)^l + LHY\_prot\_modif^l} + n_{11} \cdot \frac{PRR7\_prot^m}{(g_{13} \cdot \Omega)^m + PRR7\_prot^m} \right)$$

$$cNI\_m\_degr = m_{16} \cdot PRR5\_NI\_mRNA$$

$$cNI\_trsl = p_{10} \cdot PRR5\_NI\_mRNA$$

$$cNI\_degr = (m_{17} \cdot light + m_{24} \cdot dark) \cdot PRR5\_NI\_prot$$

$$cG\_m\_trscr = \Omega \cdot \left(\frac{q_4}{\Omega} \cdot light \cdot P + \frac{(g_{14} \cdot \Omega)^n}{(g_{14} \cdot \Omega)^n + TOC1\_prot^n} \cdot \frac{(g_{15} \cdot \Omega)^o}{(g_{15} \cdot \Omega)^o + LHY\_prot^o} \cdot n_{12} \cdot light \right)$$

$$cG\_m\_degr = GI\_mRNA \cdot m_{18}$$

$$cG\_trsl = p_{11} \cdot GI\_mRNA$$

$$cG\_degr = m_{19} \cdot GI\_prot$$

$$cZTL\_trsl = \Omega \cdot p_{14}$$

$$cZTL\_degr = m_{20} \cdot ZTL$$

$$cG\_cZTL\_assoc\_f = (p_{12} \cdot light \cdot ZTL \cdot GI\_prot) / \Omega$$

$$cG\_cZTL\_assoc\_b = p_{13} \cdot GI{:}ZTL \cdot dark$$

$$cZG\_degr = m_{21} \cdot GI{:}ZTL$$

The function *light* which is used in the kinetic laws for light-dependent reaction rates is defined by the following time-dependant function, which returns the value 1 in day-time and 0 during night-time. The function *dark* is defined as $dark = 1 - light$.

$$
\begin{aligned}
light =\ & light\_entrainment \cdot \mathbf{H}(entr\_time - \mathbf{t}) \\
+\ & light\_entrainment\_skeleton \cdot \mathbf{H}(entr\_time - \mathbf{t}) \\
+\ & light\_observation \cdot (1 - \mathbf{H}(entr\_time - \mathbf{t})) \\
+\ & light\_observation\_skeleton \cdot (1 - \mathbf{H}(entr\_time - \mathbf{t}))
\end{aligned}
$$

where $\mathbf{H}(x)$ is the Heaviside step function that returns 1 for $x > 0$ and 0 otherwise, the param-

eter *entr_time* specifies the number of hours of entrainment, the variable $\mathbf{t}$ is the simulation time and the functions *light_entrainment*, *light_entrainment_skeleton*, *light_observation* and *light_observation_skeleton* are defined as follows.

$$
\begin{aligned}
light\_entrainment = LD\_entr \cdot 0.5 \cdot \Bigg(&\left(1 + \tanh\left(\frac{\mathbf{t} - 24 \cdot \lfloor \mathbf{t}/24 \rfloor}{0.5}\right)\right) - \\
&\left(1 + \tanh\left(\frac{\mathbf{t} - 24 \cdot \lfloor \mathbf{t}/24 \rfloor - dusk\_entr}{0.5}\right)\right) + \\
&\left(1 + \tanh\left(\frac{\mathbf{t} - 24 \cdot \lfloor \mathbf{t}/24 \rfloor - 24}{0.5}\right)\right)\Bigg)
\end{aligned}
$$

$$
\begin{aligned}
light\_entrainment\_skeleton = LD\_entr\_skel \cdot 0.5 \cdot \Bigg(&\left(1 + \tanh\left(\frac{\mathbf{t} - 24 \cdot \lfloor \mathbf{t}/24 \rfloor}{0.5}\right)\right) - \\
&\left(1 + \tanh\left(\frac{\mathbf{t} - 24 \cdot \lfloor \mathbf{t}/24 \rfloor - dusk1\_entr\_skel}{0.5}\right)\right) + \\
&\left(1 + \tanh\left(\frac{\mathbf{t} - 24 \cdot \lfloor \mathbf{t}/24 \rfloor - dawn2\_entr\_skel}{0.5}\right)\right) - \\
&\left(1 + \tanh\left(\frac{\mathbf{t} - 24 \cdot \lfloor \mathbf{t}/24 \rfloor - dusk2\_entr\_skel}{0.5}\right)\right) + \\
&\left(1 + \tanh\left(\frac{\mathbf{t} - 24 \cdot \lfloor \mathbf{t}/24 \rfloor - 24}{0.5}\right)\right)\Bigg)
\end{aligned}
$$

$$
\begin{aligned}
light\_observation = \ &DD\_obs \cdot 0 \cdot \mathbf{H}\left(\mathbf{t} + 1\right) \\
&+ LL\_obs \cdot 1 \cdot \mathbf{H}\left(\mathbf{t} + 1\right) \\
&+ LD\_obs \cdot 0.5 \cdot \Bigg(\left(1 + \tanh\left(\frac{\mathbf{t} - 24 \cdot \lfloor \mathbf{t}/24 \rfloor}{0.5}\right)\right) - \\
&\qquad\left(1 + \tanh\left(\frac{\mathbf{t} - 24 \cdot \lfloor \mathbf{t}/24 \rfloor - dusk\_obs}{0.5}\right)\right) + \\
&\qquad\left(1 + \tanh\left(\frac{\mathbf{t} - 24 \cdot \lfloor \mathbf{t}/24 \rfloor - 24}{0.5}\right)\right)\Bigg)
\end{aligned}
$$

$$
\begin{aligned}
light\_observation\_skeleton = LD\_obs\_skel \cdot 0.5 \cdot \Bigg(&\left(1 + \tanh\left(\frac{\mathbf{t} - 24 \cdot \lfloor \mathbf{t}/24 \rfloor}{0.5}\right)\right) - \\
&\left(1 + \tanh\left(\frac{\mathbf{t} - 24 \cdot \lfloor \mathbf{t}/24 \rfloor - dusk1\_obs\_skel}{0.5}\right)\right) + \\
&\left(1 + \tanh\left(\frac{\mathbf{t} - 24 \cdot \lfloor \mathbf{t}/24 \rfloor - dawn2\_obs\_skel}{0.5}\right)\right) - \\
&\left(1 + \tanh\left(\frac{\mathbf{t} - 24 \cdot \lfloor \mathbf{t}/24 \rfloor - dusk2\_obs\_skel}{0.5}\right)\right) + \\
&\left(1 + \tanh\left(\frac{\mathbf{t} - 24 \cdot \lfloor \mathbf{t}/24 \rfloor - 24}{0.5}\right)\right)\Bigg)
\end{aligned}
$$

The definition of the light function reported above enables us to carry out easily experiments on various light conditions.

The variables LD_entr, LD_entr_skel, DD_obs, LL_obs, LD_obs and LD_obs_skel enable to change between the different light conditions. For instance, setting $LL\_obs = 1$ and $DD\_obs = LD\_obs = LD\_obs\_skel = 0$ indicates constant light, while setting $LD\_obs = 1$ and $LL\_obs = DD\_obs = LD\_obs\_skel = 0$ represents light/dark cycles.

The length of the photoperiod is set by the variables *dusk_obs* and *dusk_entr*, which define the time of the day (in hours) at which dusk occurs during the observation and entrainment phases, respectively; finally, the variables *dusk1_obs_skel*, *dawn2_obs_skel* and *dusk2_obs_skel* define the length of light and dark in the skeleton experiments (analogously for the entrainment phase).

Some of the reactions in the model are modelled via non-elementary reactions following Hill kinetic laws, in order to abstract from unknown details. The use of stochastic simulation algorithms for systems containing non-elementary reactions might be inappropriate in some cases, thus requiring non-elementary reactions to be decomposed into elementary reactions. This decomposition has been carried out, for instance, in [9]. The authors compared a stochastic model which contained non-elementary reactions described via Michaelis-Menten and Hill kinetics to a detailed version where all reactions had been decomposed into elementary steps and they show that the two models yielded similar results. Following these results, and considering that the compact model is much easier to handle, the same authors only consider the compact model with Michaelis-Menten and Hill kinetics in subsequent work [10]. Similarly, here we focus on the compact representation of our clock model. The validity of the stochastic approach for the analysis of this model is demonstrated by the good agreement observed between the stochastic simulation results and the experimental results. See [14, 4, 5, 11] for further discussion on the issue of the use of non-elementary reactions in stochastic models.

D Peaks/troughs identification and the computation of clock measures

In this section we shortly describe how we computed the distributions of phase, period and amplitude reported in the paper.

Although the concepts of phase, period and amplitude are basic, it is worth pointing out a couple of issues which often make their measurement tricky. First, experimental data are generally very noisy, they only refer to a few days of observation, and the same experiment is rarely repeated more than a couple of times due to the high costs. Moreover, noise in data can be of various types: variations due to differences in the observed cells/organisms, random variations due to system stochasticity, and experimental errors. Even when considering data obtained from simulations, where the only type of noise considered is the one due to stochasticity and despite the fact that simulations could be easily generated for extremely long observation periods and repeated multiple times, the stochastic variations could be so high that these measurements can still be problematic.

In the following, we illustrate the simple method we used in order to identify peaks and troughs in noisy data and to use these in computing the distribution of phase, period and amplitude in time-series data.

Especially when dealing with the noisy data coming from a single run, the first issue is to define what peaks and troughs are and how to distinguish them from random fluctuations. The method we have used here to find peaks is the following (analogously for troughs), assuming we have equally spaced time-series data points $x_0 \ldots x_{9600}$ for the time interval $[0, 960]$ hours.

1. Bin the time points into 1-hour-long slots and find the maximum in each hour $[max_0 \ldots max_{960}]$. This step smooths the data getting rid of the small fluctuations.

2. Find the local maxima in $[max_0 \ldots max_{960}]$: local maxima are data points $x_i \in [max_0 \ldots max_{960}]$ which are higher than a given number of their neighbours (the simplest case is $x_{i-1} < x_i > x_{i+1}$); we have also defined a minimum threshold that local maxima must pass in order to be considered peaks.

Based on the observation made in [1] that the normalised variation (i.e. the coefficient of variation) is minimal around peak time, we chose the peak time (rather than trough time) as the marker for phase. The period was then computed as the difference between consecutive peaks, and the amplitude was computed as difference between a peak and the following trough. on

E Additional results

In order to analyse the possible mechanisms of the damping experimentally observed in constant light (Figure 3(a) in the main text), we took two typical bioluminescence traces, measured in the same experiment for two different plants in [12]. The traces were slightly different in phase (Supplementary Figure 2(a)) and they differed also in the rate of signal decay, which could be related with differences in the levels of luciferin substrate that was sprayed onto plants. To reduce the difference, we subtracted the exponential decay of the mean (Supplementary Figures 2(b–c)). Interestingly, the resulting corrected traces had a similar high rate of damping in their rhythmic amplitude with a time constant of 1.4 days (Supplementary Figure 2(d)). This was substantially faster than the decay of the mean signal, which had time constants of 3.6 and 5.3 days and probably were related with the depletion of luciferin. This suggests that the fast damping of oscillations is related with another, luciferin-unrelated mechanism. Our model suggests that this mechanism is based on the desynchronisation of oscillations in various cells as described in the main text.

Supplementary Figures 3 and 4 report the phase, period and amplitude distributions and the autocorrelation function for the deterministic system. These results are discussed and compared with the results for the stochastic model in Section 3.2 and Section 3.5 of the main text, respectively.

Supplementary Figures 5, 6 and 7 report the results obtained in constant dark conditions (DD), analogous to ones discussed in Section 3.2 and Section 3.5 of the main text for LD and LL.

Finally, Supplementary Figure 8 reports the autocorrelation function for the system with $\Omega = 10$. These results are discussed and compared with the results shown for $\Omega = 100$ in Figure 9 in the main text.

References

[1] O.E. Akman, M.L. Guerriero, L. Loewe, and C. Troein. Complementary approaches to understanding the plant circadian clock. In *Proc. of FBTC*, EPTCS, pages 1–19, 2010.

[2] A. Baudry, S. Ito, Y.H. Song, A.A. Strait, T. Kiba, S. Lu, R. Henriques, J.L. Pruneda-Paz, N.H. Chua, E.M. Tobin, S.A. Kay, and T. Imaizumi. F-box proteins FKF1 and LKP2 act in concert with ZEITLUPE to control Arabidopsis clock progression. *Plant Cell*, 22(3):606–622, 2010.

[3] Bio-PEPA Home Page. `http://www.biopepa.org/`.

[4] R. Bundschuh, F. Hayot, and C. Jayaprakash. Fluctuations and Slow Variables in Genetic Networks. *Biophys. J.*, 84:1606–1615, 2003.

[5] Y. Cao, D.T. Gillespie, and L. Petzold. Accelerated Stochastic Simulation of the Stiff Enzyme-Substrate Reaction. *Journal of Chemical Physics*, 123:144917–144929, 2005.

[6] F. Ciocchetta, A. Duguid, S. Gilmore, M.L. Guerriero, and J. Hillston. The Bio-PEPA Tool Suite. In *Proc. of QEST'09*, pages 309–310. IEEE Computer Society, 2009.

[7] F. Ciocchetta and J. Hillston. Bio-PEPA: a Framework for the Modelling and Analysis of Biological Systems. *Theoretical Computer Science*, 410(33-34):3065–3084, 2009.

[8] A. Duguid, S. Gilmore, M.L. Guerriero, J. Hillston, and L. Loewe. Design and Development of Software Tools for Bio-PEPA. In *Proc. of WSC'09*, pages 956–967. IEEE Press, 2009.

[9] D. Gonze, J. Halloy, and A. Goldbeter. Deterministic versus Stochastic Models for Circadian Rhythms. *Journal of Biological Physics*, 28:637–653, 2002.
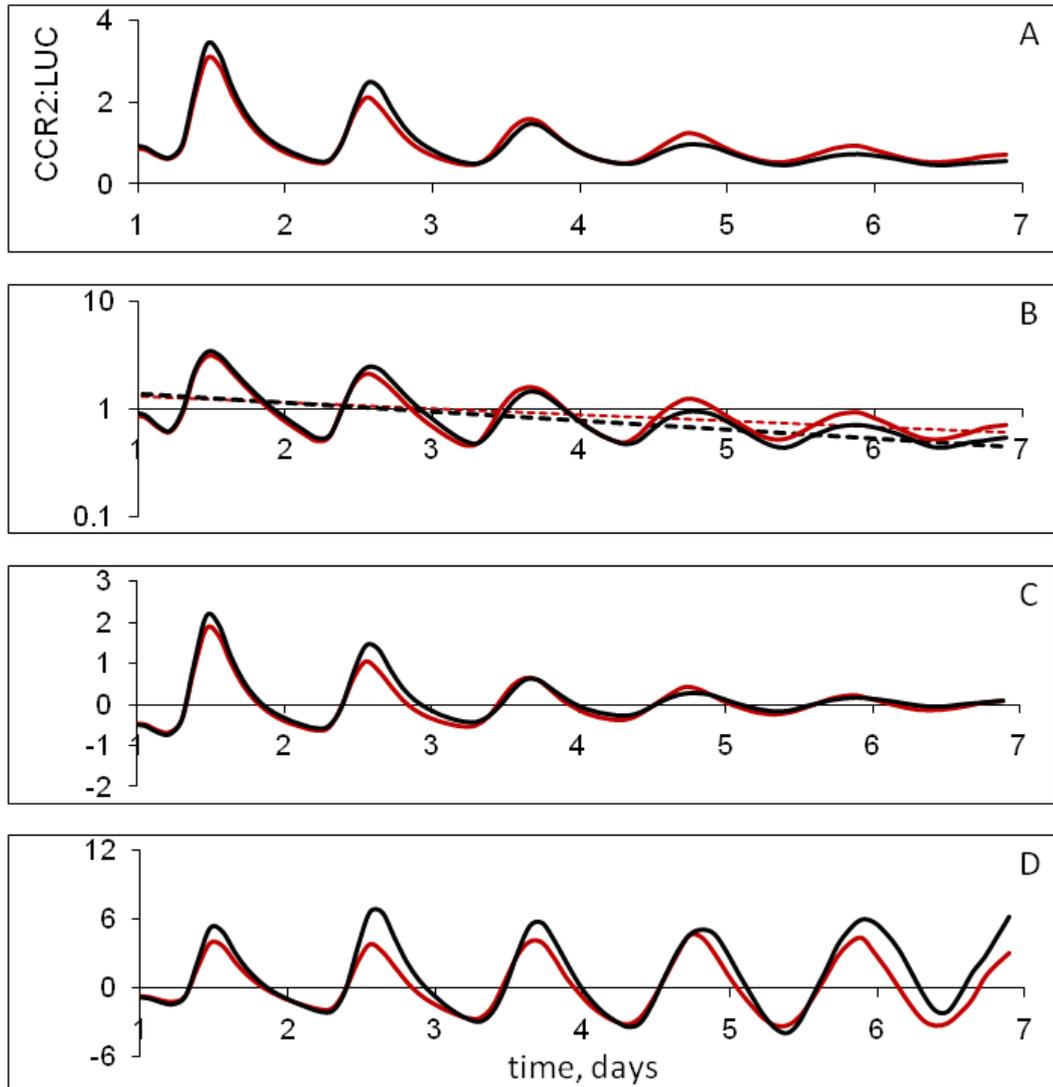
Figure 2: The decay of the luminescence and rhythmic damping have different time scales. (a) The bioluminescence of the clock marker CCR2:LUC for two individual plants was measured in the same experiment after release of the plants, entrained under LD 12:12 into constant light on day 1. (b) The same data on logarithmic scale with the straight lines showing exponential decay of the mean fluorescence. The time constants of the decay were different for the two plants (3.6 and 5.3 days). (c) The data after subtraction of the exponential decay shows quick damping of the circadian rhythm. (d) The data multiplied with a growing exponential (time constant 1.4 days) to determine the rate of damping.
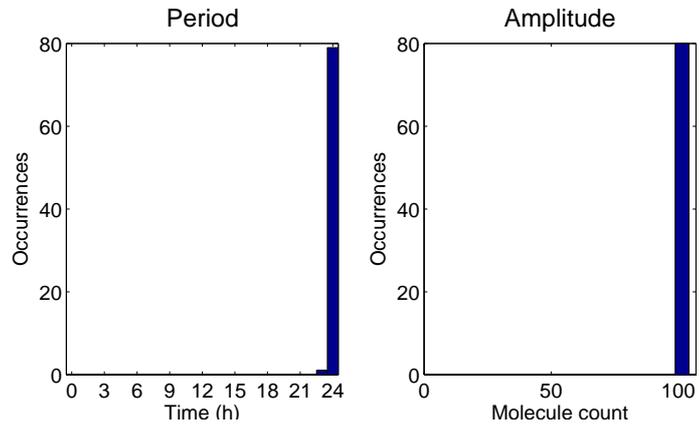
(a) LD 12:12
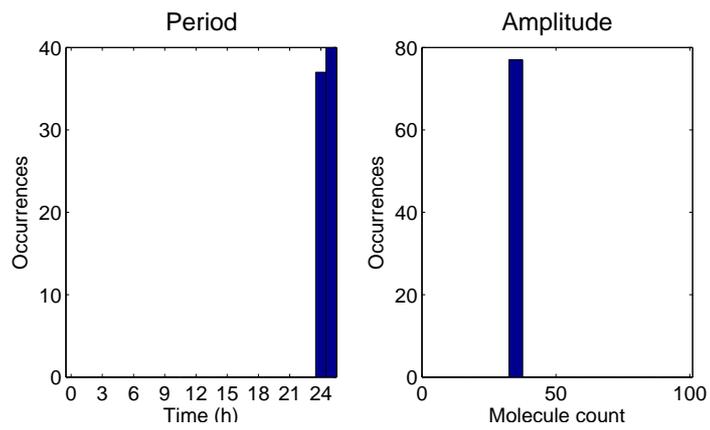
(b) LL

Figure 3: Distribution of period and amplitude of LHY mRNA oscillations over 80 days in the deterministic solution of the model: (a) LD 12:12 and (b) LL. In order to get rid of transient effects, the first 20 simulation days are discarded in both LD and LL, so the plots refer to values obtained between day 20 and day 100.
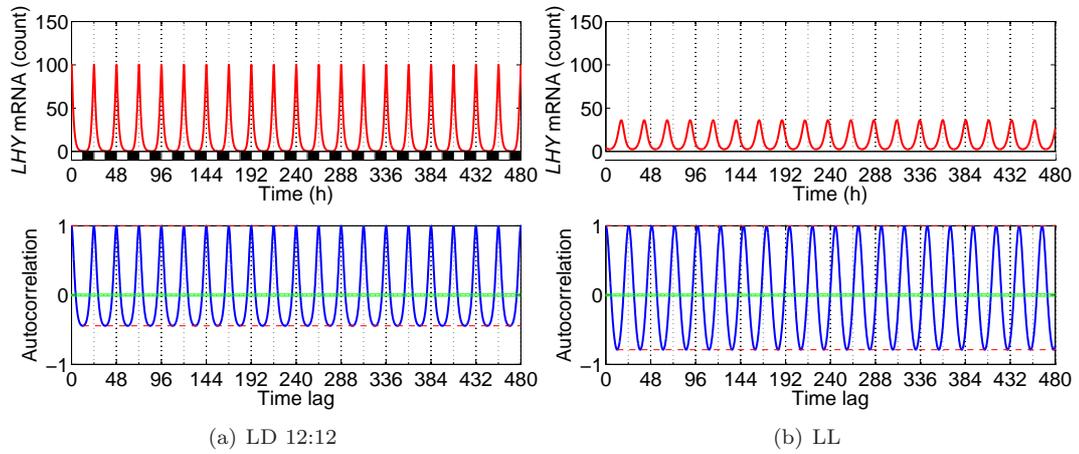
(a) LD 12:12   (b) LL

Figure 4: LHY mRNA time-series data of the deterministic behaviour and its autocorrelation function: (a) LD 12:12 and (b) LL. In the autocorrelation plots, the green horizontal lines show the 95% confidence band, while the black vertical line is the half-life of the mean autocorrelation function.
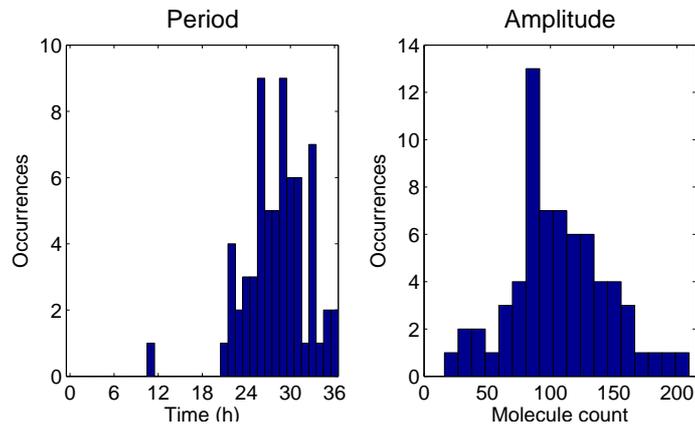

Figure 5: Distribution of period and amplitude of LHY mRNA oscillations over 80 days in a single stochastic run in DD.
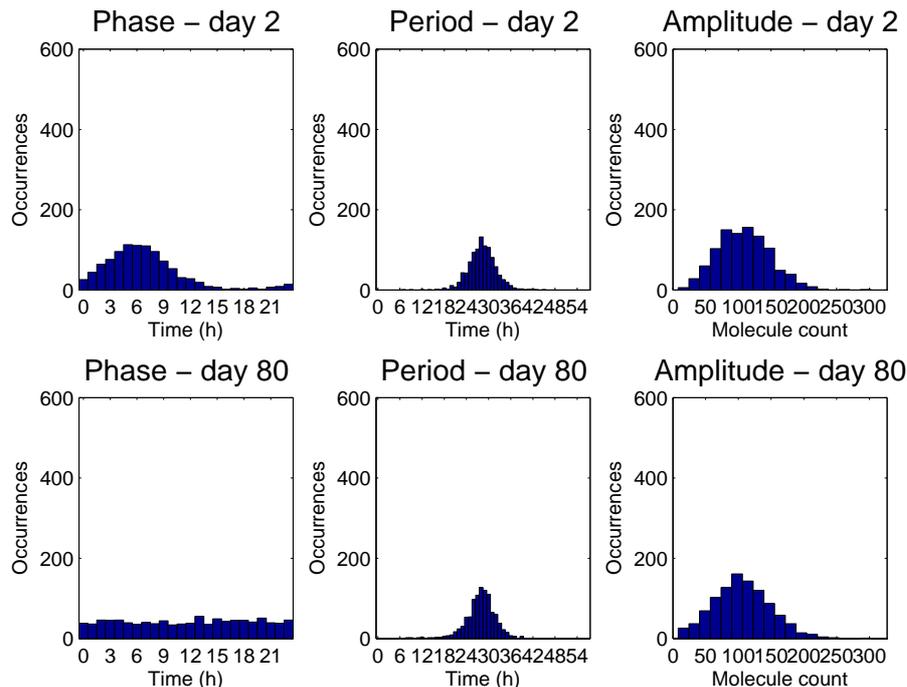
11

Figure 6: Time-dependent distribution of phase, period and amplitude of LHY mRNA oscillations in 1000 simulation runs in DD.

[10] D. Gonze, J. Halloy, J.-C. Leloup, and A. Goldbeter. Stochastic models for circadian rhythms: effect of molecular noise on periodic and chaotic behaviour. *Comptes Rendus Biologies*, 326(2):189–203, 2003.

[11] R. Grima. Noise-induced breakdown of the Michaelis-Menten equation in steady-state conditions. *Phys. Rev. Lett.*, 102:218103, 2009.

[12] J.C.W. Locke, L. Kozma-Bognar, P.D. Gould, B. Fehér, E. Kevei, F. Nagy, M.S. Turner, A. Hall, and A.J. Millar. Experimental validation of a predicted feedback loop in the multi-oscillator clock of *Arabidopsis thaliana*. *Mol. Syst. Biol.*, 2:59, 2006.

[13] A. Pokhilko, S.K. Hodge, K. Stratford, K. Knox, K.D. Edwards, A.W. Thomson, T. Mizuno, and A.J. Millar. Data assimilation constrains new connections and components in a complex, eukaryotic circadian clock model. *Mol. Syst. Biol.*, 6:416, 2010.

[14] C.V. Rao and A.P. Arkin. Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. *J. Chem. Phys.*, 118(11):4999–5010, 2003.

[15] The Systems Biology Software Infrastructure Home Page. `http://www.sbsi.ed.ac.uk/`.

[16] G. van Ooijen, L. Dixon, C. Troein, and A.J. Millar. Proteasome function is required throughout the 24-hour cycle. *Current Biology*, 2011. In press.

(a) Mean stochastic behaviour

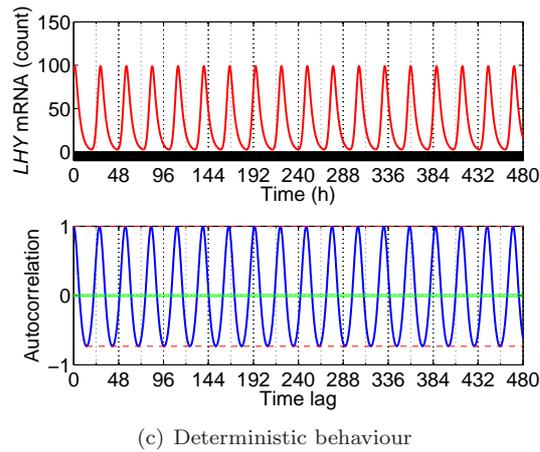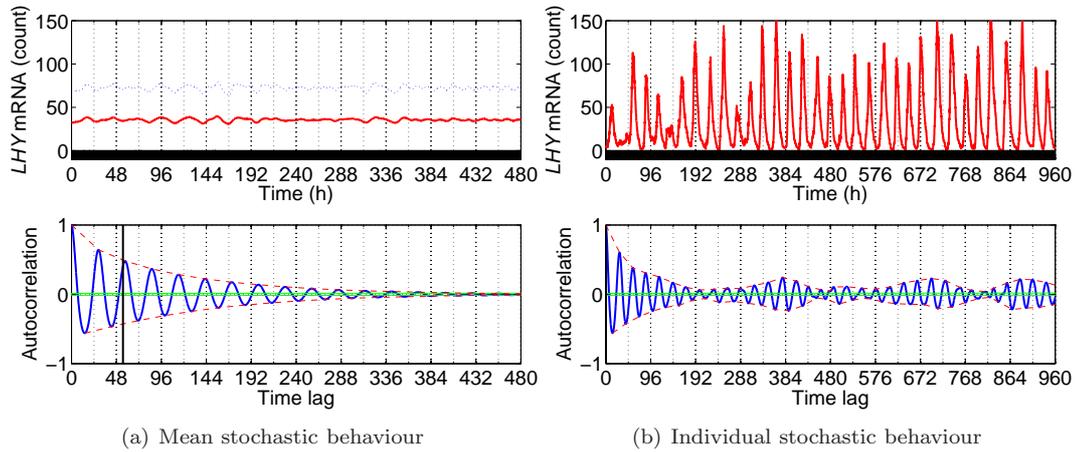(b) Individual stochastic behaviour

(c) Deterministic behaviour

Figure 7: LHY mRNA time-series data and its autocorrelation function in DD. Comparison of the autocorrelation of a single stochastic run, the mean autocorrelation of 1000 simulation runs, and of the deterministic behaviour. The first 15 simulation days were discarded when computing the autocorrelation function in order to get rid of transient effects. In the autocorrelation plots, the green horizontal lines show the 95% confidence band, while the black vertical line is the half-life of the mean autocorrelation function.

13

(a) LD 12:12, mean stochastic behaviour

(b) LD 12:12, individual stochastic behaviour

(c) LL, mean stochastic behaviour
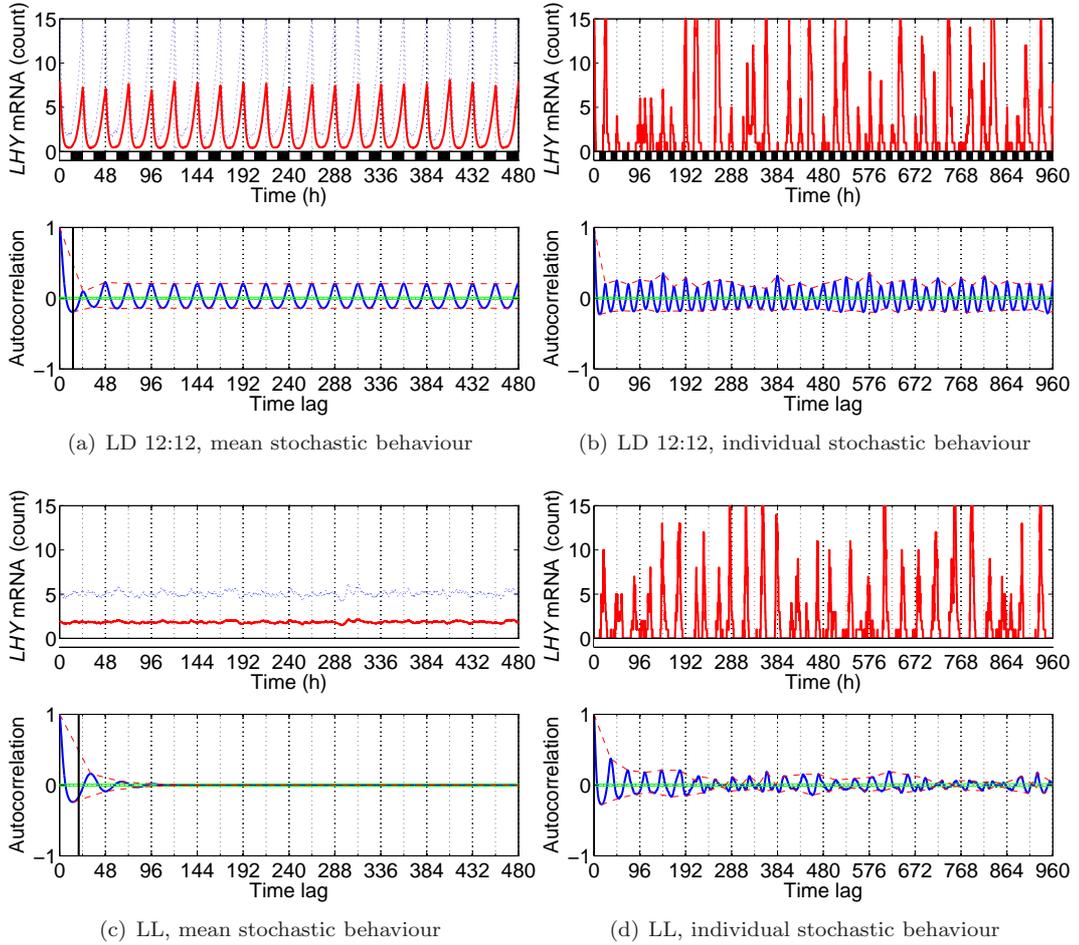
(d) LL, individual stochastic behaviour

Figure 8: LHY mRNA time-series data and its autocorrelation function for $\Omega = 10$: (a–b) LD 12:12 and (c–d) LL. Comparison of the autocorrelation of a single stochastic run and the mean autocorrelation of 1000 stochastic simulation runs. The first 15 simulation days were discarded when computing the autocorrelation function in order to get rid of transient effects. In the autocorrelation plots, the green horizontal lines show the 95% confidence band, while the black vertical line is the half-life of the autocorrelation function.