Biostatistics (2010), **1**, 1, *pp.* 1–11
doi:10.1093/biostatistics/bio000

Supplementary material to Inferring the topology of a time-invariant non-linear sparse gene regulatory network using fully Bayesian spline autoregression

Edward R. Morrissey

Systems Biology Centre,
University of Warwick, Coventry House, CV4 7AL, Coventry, UK.

MIGUEL A. JUÁREZ*

Systems Biology Centre,
University of Warwick, Coventry House, CV4 7AL, Coventry, UK.
m.a.juarez@warwick.ac.uk

KATHERINE J. DENBY

WHRI and Systems Biology Centre,
University of Warwick, Coventry House, CV4 7AL, Coventry, UK.

NIGEL J. BURROUGHS

Systems Biology Centre,
University of Warwick, Coventry House, CV4 7AL, Coventry, UK.

1. THE MODEL

For completeness, we repeat the model and the prior specification. Stacking the bases and the coefficients into $X^t = \{X_1^t, \ldots, X_G^t\} \in \mathbb{R}^{MG}$ and $\boldsymbol{\beta}_g = \{\boldsymbol{\beta}_{1g}, \ldots, \boldsymbol{\beta}_{Gg}\} \in \mathbb{R}^{MG}$, respectively, we can express the model as $y_g^{t+1} = \mu_g + X^t \boldsymbol{\beta}_g + \varepsilon_g^t$ and after further stacking the equations over time we have,

$$\boldsymbol{y}_g = \boldsymbol{\mu}_g + \mathcal{X}\, \boldsymbol{\beta}_g + \boldsymbol{\varepsilon}_g \,, \qquad g = 1, \ldots, G \,, \tag{1.1}$$

where $\boldsymbol{\mu}_g = \mu_g \boldsymbol{\iota}_T'$, with $\boldsymbol{\iota}_T$ a row vector of ones of size T and $\mathcal{X} = \{X^1, X^2, \ldots, X^T\}'$ a bases matrix of size $[T \times MG]$.

Precisions. We use conjugate, iid gamma priors, $\mathrm{Ga}(\lambda_g \mid a_\lambda, b_\lambda)$, on the gene preci-

*To whom correspondence should be addressed.

sions, $\boldsymbol{\lambda} = \{\lambda_1, \ldots, \lambda_G\}$,

$$\pi(\boldsymbol{\lambda}) = \prod_{g=1}^{G} \frac{b_\lambda^{a_\lambda}}{\Gamma[a_\lambda]} \lambda_g^{a_\lambda - 1} \exp[-b_\lambda \lambda_g] \ . \tag{1.2}$$

Constant term. An independent Gaussian prior, $\mathrm{N}(\boldsymbol{\mu} \mid \mathbf{0}, \tau_\mu I)$, for the gene-specific constant, $\boldsymbol{\mu} = \{\mu_1, \ldots, \mu_G\}$

$$\pi(\boldsymbol{\mu}) = \left(\frac{\tau_\mu}{2\pi}\right)^{G/2} \exp\left[-\frac{\tau_\mu}{2} \boldsymbol{\mu}' \boldsymbol{\mu}\right] \ . \tag{1.3}$$

Network structure. We provide two alternatives for modelling the network topology. The first is to define the *overall network connectivity*, ρ, as $\mathrm{P}\left[\gamma_{jg} = 1\right] = \rho$ and complement it with a Beta prior, $\mathrm{Be}(\rho \mid a_\rho, b_\rho)$. The full specification is then,

$$\pi(\gamma_{jg} \mid \rho) = \rho^{\gamma_{jg}} (1-\rho)^{1-\gamma_{jg}} \ , \quad g, j = 1, \ldots, G \ , \tag{1.4}$$

$$\pi(\rho) = [\mathrm{B}(a_\rho, b_\rho)]^{-1} \ \rho^{a_\rho - 1}(1-\rho)^{b_\rho - 1} \quad 0 < \rho < 1 \ . \tag{1.5}$$

Alternatively, we can accommodate *parent-wise connectivity*, $\mathrm{P}[\gamma_{jg} = 1] = \rho_j$, by letting

$$\pi(\gamma_{jg} \mid \rho_j) = \rho_j^{\gamma_{jg}} (1-\rho_j)^{1-\gamma_{jg}} \ , \quad g = 1, \ldots, G \ , \tag{1.6}$$

$$\pi(\rho_j) = [\mathrm{B}(a_\rho, b_\rho)]^{-1} \ \rho_j^{a_\rho - 1}(1-\rho_j)^{b_\rho - 1} \quad j = 1, \ldots, G \ . \tag{1.7}$$

Spline Coefficients. We use the prior on the coefficients $\boldsymbol{\beta}_{jg}$ to shrink them towards the origin specifying a second order Markov process prior

$$\pi(\boldsymbol{\beta}_{jg} \mid \tau_{jg}) = \mathrm{N}(\boldsymbol{\beta}_{jg} \mid \mathbf{0}, \tau_{jg} K) \ . \tag{1.8}$$

Where,

$$K_{M,M-2} = 1 \ , \qquad K_{M-2,M-1} = -4 \ , \qquad K_{M,M} = 1 \ ,$$
$$K_{M,M-1} = -2 \ , \qquad K_{M-1,M-1} = 5 \ ;$$

and for all $i, j \in \{3, \ldots M-2\}$,

$$K_{i,j} = \begin{cases} 0 & |i-j| > 2 \\ -4 & |i-j| = 1 \\ 1 & |i-j| = 2 \\ 6 & |i-j| = 0 \end{cases} .$$

Commonly, the two remaining coefficients are given an improper prior, $\pi(\beta_1, \beta_2) \propto 1$.

We discuss conditions for posterior propriety in Section 1.1.
Smoothness parameters. An inverted Pareto prior, $\text{Ip}(\cdot \mid a_\tau, b_\tau)$:

$$\pi(\tau_{jg} \mid a_\tau, b_\tau) = \frac{a_\tau}{b_\tau}\left(\frac{\tau_{jg}}{b_\tau}\right)^{a_\tau - 1}, \qquad \tau_{jg} \leqslant b_\tau\,,\ a_\tau > 0\,. \tag{1.9}$$

1.1 *Posterior propriety*

Fahrmeir and Kneib (2009) discuss conditions for posterior propriety using the covariance structure, K, arising from the second order Markov process prior and different alternatives for the smoothing parameters within the context of structured additive models. We provide a result justifying the extension of this prior for GRN inference used in the paper. The proof is standard and therefore omitted.

Theorem 1 Consider the longitudinal data set $Y = \{y_g^t\}$, consisting of $g = 1, \ldots, G$ genes measured at times $t = 1, \ldots, T$, modelled as (1.1) and with prior given by (1.2)–(1.9). Let $\mathcal{K}_g = \text{blkdiag}\,[\tau_{1g}K, \tau_{2g}K, \ldots, \tau_{G^\star g}K]$ and $\Psi_g = \mathcal{X}_g'\mathcal{X}_g + \mathcal{K}_g$. Where $\mathcal{X}_g$ is the design sub-matrix conformable to $G_g^\star$, the number of parents of gene g. Then, the posterior distribution of $\{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_G, \boldsymbol{\lambda}\}$ is proper if Ψ_g is positive definite for every g and $M \times G_g^\star < T$.

Given that in most of our applications we will only have a limited number of time measurements compared to the number of genes, this leads to an improper posterior if the prior was not proper, since the number of parents for any given gene only needs to exceed T/M. To construct a proper prior we supply (1.8) with an independent specification for the first two coefficients,

$$\pi(\beta_1, \beta_2) = \text{N}(\beta_1 \mid 0, k_1)\ \text{N}(\beta_2 \mid 0, k_2)\,. \tag{1.10}$$

Including these into the covariance structure we have

$$K_{1,1} = (1 + k_1/\tau),\ K_{1,2} = -2,\ K_{1,3} = 1,\ K_{2,2} = (5 + k_2/\tau) \text{ and } K_{2,3} = 1,$$

for the appropriate smoothness parameter, τ.

2. MCMC SCHEME

Combining the likelihood with the prior and letting Θ denote all the model parameters we obtain,

$$\pi(\Theta \mid \mathcal{X}, Y) \propto \left[\prod_g \text{N}_T\big(\boldsymbol{y}_g \mid \boldsymbol{\mu}_g + \mathcal{X}\,\widetilde{\boldsymbol{\beta}}_g, \lambda_g I_T\big)\right] \times$$

$$\left[\prod_g \pi(\boldsymbol{\beta}_g \mid \boldsymbol{\tau}_g)\pi(\boldsymbol{\tau}_g)\pi(\lambda_g)\pi(\mu_g)\pi(\boldsymbol{\gamma}_g \mid \rho_g)\pi(\rho_g)\right],$$

where I_T is the identity matrix of size T, $\boldsymbol{\gamma}_g = \{\gamma_{1g}, \ldots, \gamma_{Gg}\}$ and $\boldsymbol{\tau}_g = \{\tau_{1g}, \ldots, \tau_{Gg}\}$. As there is no closed form expression for the posterior numerical methods are needed. We propose a Metropolis-within-Gibbs scheme which is drafted below.

Precisions The full conditional of λ_g, $g = 1, \ldots, G$ is given by

$$\pi(\lambda_g \mid \text{—}) \propto \lambda_g^{T/2+a_\lambda-1} \exp\left[-\lambda_g\left(b_\lambda + \frac{1}{2}\,\boldsymbol{e}_g'\,\boldsymbol{e}_g\right)\right]$$

which is the kernel of a gamma distribution, with $\boldsymbol{e}_g = \boldsymbol{y}_g - \boldsymbol{\mu}_g - \mathcal{X}\widetilde{\boldsymbol{\beta}}_g$.

Constant term μ_g is conditionally Gaussian, with mean and precision

$$m_g = \frac{\bar{\boldsymbol{y}}_g - \bar{\mathcal{X}}\widetilde{\boldsymbol{\beta}}_g}{\lambda_g + \tau_\mu/T} \qquad \text{and} \qquad \tau'_\mu = \tau_\mu + T\,\lambda_g\,,$$

respectively, where $\bar{\boldsymbol{y}}_g = T^{-1}\sum_t y_g^t$ and $\bar{\mathcal{X}} = T^{-1}\sum_t \mathcal{X}$.

Connectivity The full conditionals for the gene-wise connectivity, ρ_g, are obtained as

$$\pi(\rho_g \mid \text{—}) \propto \rho_g^{S_g+a_\rho-1}(1-\rho_g)^{G+b_\rho-S_g-1} \text{ with } S_g = \sum_i^G \gamma_{gi}\,,$$

and are sampled from a $\text{Be}(\rho_g \mid S_g + a_\rho, G + b_\rho - S_g)$, for $g = 1, \ldots, G$.
The overall connectivity, ρ, is sampled from a $\text{Be}\big(\rho \mid S + a_\rho, G^2 + b_\rho - S\big)$, with $S = \sum_{g=1}^G S_g$.

Smoothness parameters When the corresponding link is on, the full conditional is given by

$$\pi(\tau_{jg} \mid \text{—}) \propto \tau_{jg}^{(M-2)/2+a_\tau-1} \exp\left[-\tau_{jg}\,\frac{1}{2}\widetilde{\boldsymbol{\beta}}'_{jg}K\,\widetilde{\boldsymbol{\beta}}_{jg}\right], \qquad 0 < \tau_{jg} < b_\tau$$

and can be sampled from a truncated gamma distribution (Damien and Walker, 2001; Gentle, 2003) with parameters $\left\{(M-2)/2 + a_\tau, \widetilde{\boldsymbol{\beta}}'_{jg}K\,\widetilde{\boldsymbol{\beta}}_{jg}/2\right\}$. An observation is drawn from the prior when the link is off.

Spline Coefficients and link probabilities The update of the spline coefficients and indicator variables is performed as a block. Specifically, the update of a given indicator variable γ_{jg} and all the coefficients of the regression for gene g, $\boldsymbol{\beta}_g$, are

performed simultaneously. In practice, as the regression is sparse, only a few links are actually present drastically reducing this computation. At every iteration, the individual link indicator γ_{jg} is turned on (off) if it is off (on) and the associated coefficient, $\boldsymbol{\beta}_g \in \mathbb{R}^{MG}$, for present links (on) is proposed from the joint conditional. Schematically we have,

$$\gamma : 0 \to 1 \qquad \text{and} \qquad \boldsymbol{\beta} : \boldsymbol{\beta}^a \to \boldsymbol{\beta}^b$$

with acceptance probability

$$\alpha = \min \left\{ \frac{\pi\left(\widetilde{\boldsymbol{\beta}}^b\right)}{\pi\left(\widetilde{\boldsymbol{\beta}}^a\right)} \frac{q(\boldsymbol{\beta}^a \mid \gamma^a) q(\gamma^a)}{q(\boldsymbol{\beta}^b \mid \gamma^b) q(\gamma^b)},\ 1 \right\} ,$$

where the subscripts have been omitted for clarity. γ is proposed symmetrically, thus $q(\gamma^a)/q(\gamma^b) = 1$. For $q(\boldsymbol{\beta} \mid \gamma)$ we use the proposal

$$q(\boldsymbol{\beta} \mid \text{—}) = N(\mu_{\boldsymbol{\beta}}, \Sigma_{\boldsymbol{\beta}})$$

with

$$\Sigma_{\boldsymbol{\beta}} = \left[\lambda_g \mathcal{X}_g' \mathcal{X}_g + \Upsilon_g\right] \qquad \text{and} \qquad \mu_{\boldsymbol{\beta}} = \lambda_g (\boldsymbol{y}_g - \boldsymbol{\mu}_g) \mathcal{X}_g \Sigma_{\boldsymbol{\beta}}^{-1} ,$$

where Υ_g is the block diagonal penalty (precision) matrix, calculated by multiplying each block in $\mathcal{K}_g$ times the corresponding τ_{jg}. Note that, as only the coefficients with non zero indicator variable are updated, $\mathcal{X}_g$, $\boldsymbol{y}_g$ and Υ_g are adjusted to only include the appropriate elements. Substituting this in the Hastings ratio gives

$$\frac{\rho}{1-\rho}\ \tau_0 \tau_{jg}^{(M-2)/2} \frac{\exp\left[\frac{1}{2} \mu_{\boldsymbol{\beta}}^b \Sigma_{\boldsymbol{\beta}}^{-1b} \mu_{\boldsymbol{\beta}}^b\right]}{\exp\left[\frac{1}{2} \mu_{\boldsymbol{\beta}}^a \Sigma_{\boldsymbol{\beta}}^{-1a} \mu_{\boldsymbol{\beta}}^a\right]} \frac{\left|\Sigma_{\boldsymbol{\beta}}^b\right|^{1/2}}{\left|\Sigma_{\boldsymbol{\beta}}^a\right|^{1/2}} .$$

The opposite move (switching an indicator variable off) can be performed using the reciprocal of the ratio above.

In order to enforce the identifiability restriction, at each step we calculate $\bar{m}_g = \boldsymbol{\iota}_T \times \left[\mathcal{X}\ \widetilde{\boldsymbol{\beta}}_g\right]$, for every gene, subtract it from the splines and add it to the constant term, μ_g.

Our sampler exploits the conditional independence structure of the model. We constructed a parallel scheme where the calculation for each parent is assigned to a CPU-node, these communicating only when the overall connectivity is updated and for sample recording. Gains in computation times can potentially be up to n-fold, with n the number of CPU-nodes used.

2.1 *Improving convergence*

The last move in Section 2 leads to a dramatic decrease in autocorrelation of the Markov chain, compared to a Gibbs move. Indeed, a common approach in these cases is to use a full Gibbs specification, with a full conditional Bernoulli distribution on the γ_{jg} and a full conditional Gaussian for the coefficients $\boldsymbol{\beta}_{jg}$. The latter requires the introduction of a so-called pseudo-prior which needs to be tuned to improve the mixing of the chain (Dellaportas *and others*, 2000; Ntzoufras, 2002; O'Hara and Sillanpää, 2009). In order to assess the gains in mixing, we implemented a full Gibbs sampler for the linear model. When chain mixing is compared, the advantage of our MH update becomes apparent as illustrated in Figure 1, obtained by running both samplers on the non-linear synthetic data described in Section 3.1. The top panels plot the number of times that link was switched during the MCMC run against the posterior probability of the link being present. One would expect that links with probabilities around $1/2$ would change more often, as in Figure 1b. However, the Gibbs strategy tends to mix more slowly, as shown in Figure 1a. Although the MH step is more computationally demanding, the benefit brought about by the improved mixing of the chain, quantified by the reduction in autocorrelation (ACF), offsets this cost easily (compare Figure 1c with Figure 1d). Given that the parameter space of the splines model is much larger than the linear one, the benefits of using this move, compared to the full Gibbs alternative are expected to be even greater.

3. ILLUSTRATIONS AND APPLICATIONS

In all our applications, we include a slight modification of the structure of the network topology to that described in Section 1. We know from the context that each gene has a decay term, corresponding to mRNA decay. We include this information in the prior by fixing $\gamma_{gg} = 1$. As we also know that this decay is close to linear, we set the shape of the inverted Pareto prior for these smoothing parameters to thrice the value used for the rest, *i.e.* $a_{\tau_{gg}} = 3 \times a_{\tau_{ij}}, i \neq j$.

The splines and linear models were fitted using the overall and gene-wise connectivity specifications. Throughout, 13 bases were used, i.e. splines of degree 3 with 10 evenly spaced knots. Prior parameters were set to $\{a_\rho, b_\rho\} = \{1/2, 1/2\}$, $\{a_\lambda, b_\lambda\} = \{2, 0.01\}$, $\tau_\mu = 1/4$, $\tau_0 = 0.25$ and $\{a_\tau, b_\tau\} = \{1.5, 10^4\}$. We ran two parallel chains of length 10^5, dropping the first 10^4 steps and then recording every tenth draw. We performed some sensitivity analyses, varying a_τ from 1 (uniform prior) up to 3, setting $a_\rho = b_\rho = 1, 2$ and using flatter versions of the prior for λ by setting $a_\lambda = 1, 0.1$, without finding noteworthy differences. Convergence was assessed by comparing both chains graphically and by formal tests using the CODA package (Plummer *and others*, 2006).

3.1 *Discrete time synthetic networks*

In order to assess the network topology recovery power of our model, we produced two synthetic, first order autoregressive processes. One has only linear and the second a number of non-linear (S-shaped) relations. In the non-linear case, all the functional relations were produced using Hill functions, except for the self-interactions which are linear. In both cases we set $G = 16$, $T = 40$, and $\rho \approx 0.1$.

The models with gene-wise and overall connectivity produced almost indistinguishable estimations for the network topology and thus we report the results for the simpler model only. In Figure 2a we plot the marginal posterior and prior distributions of the model precisions, λ_g (for a selection of the genes only, to avoid clutter). We also performed a sensitivity analysis on ρ, fixing the prior parameters $a_\rho = b_\rho = 2$. As shown in Figure 2b, the posterior was practically unaffected by this change.

We use the AUC and MxE for model fit comparison. The AUC is the area under the ROC curve, plotted in Figure 3 below. It is apparent that network topology prediction from both models is almost identical when fitting the linear interactions network (left panel). When non-linearities are present, the splines model achieves a better network retrieval, according to the ROC curve (right panel).

3.2 *Biological* GRN*: the plant Circadian Clock*

Most organisms have the ability to track time even in the absence of external input (*e.g.* light). This ability allows the organism to anticipate and prepare for future events, thus enabling it to optimise the interaction with the environment. In some cases, such as in *Arabidopsis*, diurnal period tracking is achieved via a regulatory network that oscillates with a period close to 24 hrs. This period then propagates through one or more of the core genes of the clock to target genes responsible for other biological processes (reviews can be found in Harmer, 2009; Más, 2008; McClung, 2006). The circadian clock is of central importance and has been extensively studied both experimentally and through mathematical modelling. It has recently been reported to regulate up to 90% of the *Arabidopsis* genome under some environmental conditions (Michael *and others*, 2008). While the circadian clock is able to maintain oscillations without the need of light, it is known that the period is modified by light exposure, allowing it to adapt to shorter and longer daylight hours.

3.2.1 *Differential Equation Data.* We generated data from Locke *and others* (2006) using COPASI (Hoops *and others*, 2006) fixing the light source to be permanently on. The data was then subsampled, logged and standardised. The resulting data set has 50 time points with a time spacing of 1 hr, see Figure 4 a. Given that simultaneous measurement of multiple proteins is currently very hard, usually only mRNA is available for network inference. For this reason, although the ODE model outputs protein concentration and location, we used only the mRNA data. We expect the data generated from this model

to be a reasonable reflection of experimental data, not only because it is a continuous time model with nonlinear interactions, but it also reflects realistic sampling regimes and interaction intensities.

3.2.2 *Experimental Data.* Whole leaves were harvested every 2 hrs for 48 hrs, with four biological replicates at each time point. To reduce variability, the same leaf (the 7th leaf to emerge) was harvested for each sample. This means that the same plant was not monitored over the entire time series but leaves of 96 distinct plants grown in identical conditions were sampled (four at each of the 24 time points). Full genome expression profiles of these leaves were generated using CATMA arrays (Sclep *and others*, 2007). Data processing and normalisation of the time series was carried out using a pipeline based on the R package MAANOVA (Wu *and others*, 2003). Given that the replicates showed some outliers we use the median of the four bioreplicates as the observed series. We use the same genes as those that appear in the ODE model, leaving some freedom to choose which genes to use for the ambiguous nodes. For the two genes that represent pairs, we selected those that showed least variability across replicates (LHY, PRR7). To represent X we chose the genes amongst the candidates in the current working model that showed the strongest signal-to-noise ratio: LUX and ELF4. For Y we chose PRR5 as it had a stronger signal than GI. Traces of the selected genes are depicted in Figure 4 b.

One way to assess the accuracy of the splines inferred network topology is to restrict our analysis to the most extensively studied genes: PRR7, LHY, TOC1 and (less so) PRR5. We can reasonably assume that the connections between these links are known. As can be seen in Figure 4b, there are 6 connections amongst these genes. The splines model predicts 5 connections, of which 3 are correctly predicted. Both of the incorrect predictions appear for genes that are missing a link in the inferred networks, indicating that the model has found the wrong parent rather than overfitting with more parents than necessary. Neither candidate gene for X (LUX and ELF4) regulates LHY, which would be evidence supporting the hypothesis that one of them is the unknown gene. On the other hand, these genes were proposed to be X as they are known to be involved in the clock and have some effect on the system, so the predictions can serve as a working hypotheses for determining the role they play within the network.

4. COMPUTATION TIMES

Our algorithm took 2.7 hrs to run 10^5 iterations with the nonlinear synthetic data ($G = 16$, $T = 40$, $\rho \approx 0.1$) and scaling is likely to be quadratic in the number of genes and number of potential parents, and linear in the number of time points; for instance, fitting a microarray gene expression data set —not shown— with $G = 30$ and $T = 37$ ($\hat{\rho} \approx 0.15$) took 20 hrs for the same run length. Thus, for data sets with a large number of genes a parallel algorithm is available which reduces computation time approximately linearly in the number of cpu-nodes; for instance, using 31 cpus the former data set took

28.6 mins and the latter 3 hrs.

REFERENCES

DAMIEN, P. AND WALKER, S. G. (2001). Sampling truncated Normal, Beta and Gamma densities. *Journal of Computational and Graphical Statistics* **10**, 206–215.

DELLAPORTAS, P., FOSTER, J. J. AND NTZOUFRAS, I. (2000). Bayesian variable selection using the Gibbs sampling. In: Dey, D. K., Ghosh, S. K. and Mallick, B. K. (editors), *Generalized linear models: a Bayesian perspective*. New York: Marcel Dekker, pp. 273–286.

FAHRMEIR, L. AND KNEIB, T. (2009). Property of posteriors in structured additive regression models: Theory and empirical evidence. *Journal of Statistical Planning and Inference* **139**, 843–859.

GENTLE, J. E. (2003). *Random number generation and Monte Carlo methods*, 2 edition. New York: Springer-Verlag.

HARMER, S. L. (2009). The circadian system in higher plants. *Annual Review of Plant Biology* **60**, 357–377.

HOOPS, S., SAHLE, S., GAUGES, R., LEE, C., PAHLE, J., SIMUS, N., SINGHAL, M., XU, L., MENDES, P. AND KUMMER, U. (2006). COPASI —a COmplex PAthway SImulator. *Bioinformatics* **22**, 3067–3074.

LOCKE, J. C. W., KOZMA-BOGNÁR, L., GOULD, P. D., FEHÉR, B., KEVEI, E., NAGY, F., TURNER, M. S., HALL, A. AND MILLAR, A. J. (2006). Experimental validation of a predicted feedback loop in the multi-oscillator clock of *Arabidopsis thaliana*. *Molecular Systems Biology* **2**, 59.

MÁS, P. (2008). Circadian clock function in *Arabidopsis thaliana*: time beyond transcription. *Trends in Cell Biology* **18**, 273–281.

MCCLUNG, C. R. (2006). Plant circadian rhythms. *The plant Cell* **18**, 792–803.

MICHAEL, T. P., MOCKLER, T. C., BRETON, G., MCENTEE, C., BYER, A., TROUT, J. D., HAZEN, S. P., SHEN, R., PRIEST, H. D., SULLIVAN, C. M., GIVAN, S. A., YANOVSKY, M., HONG, F., KAY, S. A. *and others*. (2008). Network discovery pipeline elucidates conserved time-of-day-specific cis-regulatory modules. *PLoS Genetics* **4**, e14.

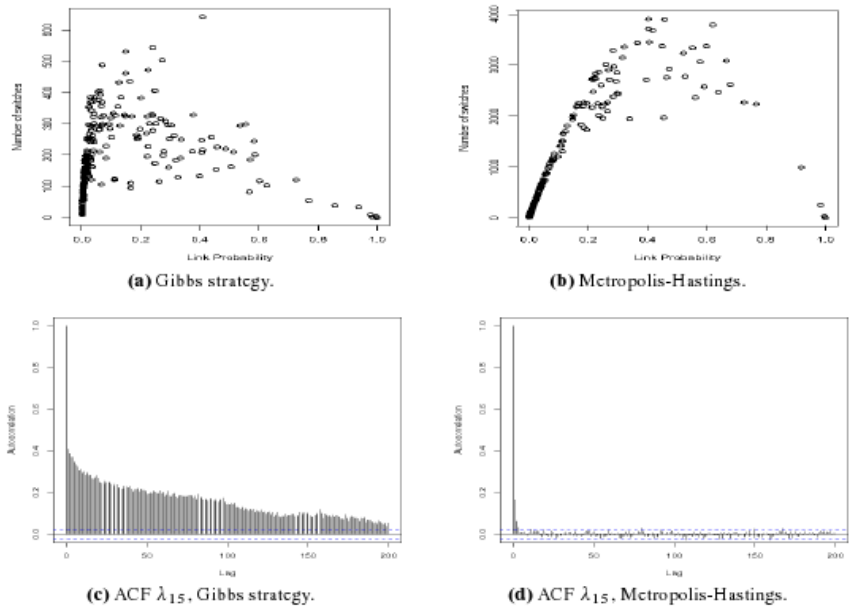
NTZOUFRAS, I. (2002). Gibbs variable selection using BUGS. *Journal of Statistical Software* **7**(7), 1–19.

O'HARA, R. B. AND SILLANPÄÄ, M. J. (2009). A review of Bayesian variable selection methods: What, how and which. *Bayesian Analysis* **4**, 85–118.

PLUMMER, M., BEST, N., COWLES, K. AND VINES, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News* **6**, 7–11.

SCLEP, G., ALLEMEERSCH, J., LIECHTI, R., DEMEYER, B., BEYNON, J., BHALERAO, R., MOREAU, Y., NIETFELD, W., RENOU, J. P., REYMOND, P., KUIPER, M. T. R. *and others*. (2007). CATMA, a comprehensive genome-scale resource for silencing and transcript profiling of *Arabidopsis* genes. *BMC Bioinformatics* **8**, 400.

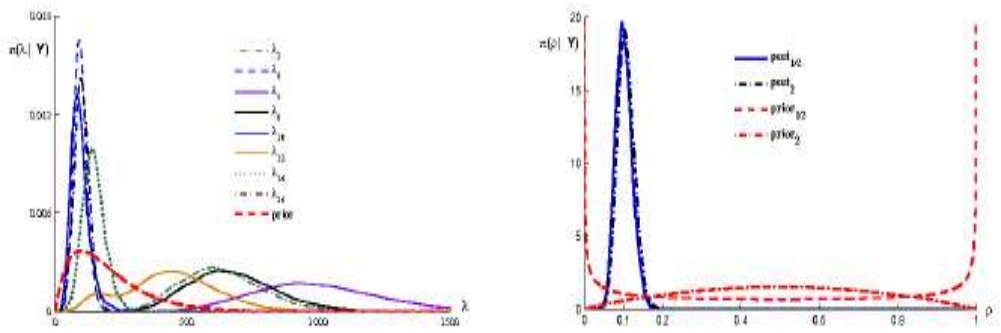
WU, HAO, KERR, M. K., CUI, XIANGQIN AND CHURCHILL, G. A. (2003). MAANOVA: A software package for the analysis of spotted cDNA microarray experiments. In: Parmigiani, G., Garrett, E. S., Irizarry, R. A. and Zeger, S. L. (editors), *The Analysis of Gene Expression Data: Methods and Software*. New York: Springer, pp. 313–341.



(a) Gibbs strategy. **(b)** Metropolis-Hastings.

(c) ACF λ_{15}, Gibbs strategy. **(d)** ACF λ_{15}, Metropolis-Hastings.

Fig. 1. Chain mixing comparison of the Gibbs and MH strategies. Top panels plot the number of state changes of a link during the MCMC run against its posterior probability. The bottom panels show the autocorrelation function (ACF) for a single link's precision (gene 15).



(a) Precision posteriors. **(b)** Connectivity posterior.

Fig. 2. Marginal posterior distributions calculated when fitting the splines model to the non-linear synthetic data set. (a) The posterior for a selection of the gene precisions, λ_g and the corresponding prior. (b) The posterior of the overall-connectivity, ρ, from two different priors, $a_\rho = b_\rho = 1/2$ and 2. In both panels priors are depicted by the thick (red) dashed lines.

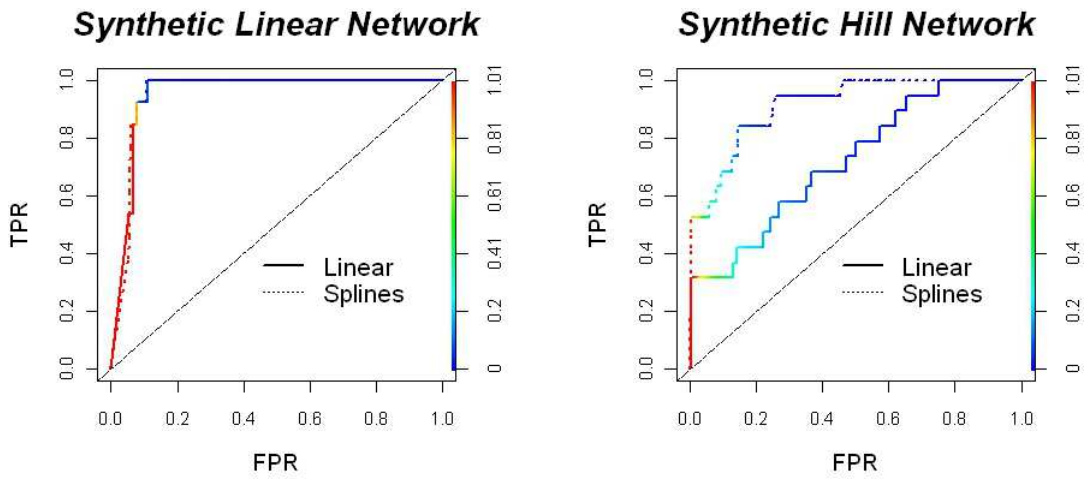


Fig. 3. ROC curves from synthetic networks. On the left panel, the ROC curves from the linear synthetic network obtained when fitting the linear AR(1) and the splines models. The right panel depicts the ROC from the non-linear, synthetic network.

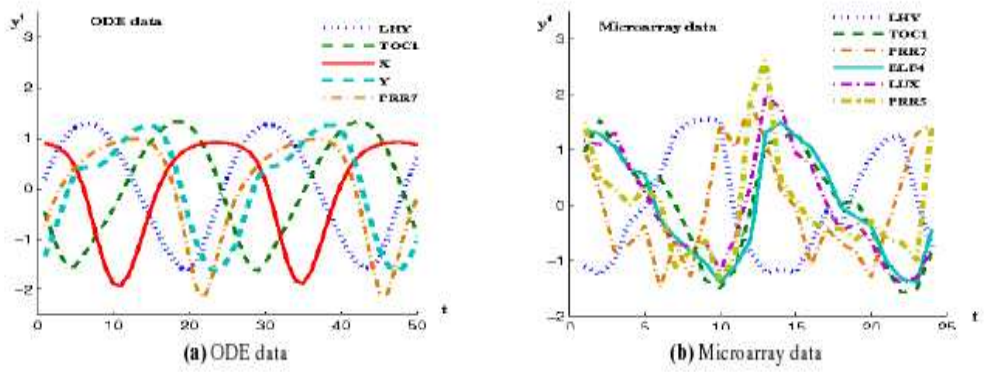


Fig. 4. Time traces of the ODE model and experimental data for the Circadian Clock in *Arabidopsis thaliana*. (a) Data simulated with the ODE clock of Locke *and others* (2006). (b) Gene expression profiles of *Arabidopsis* leaves. Data sets are standardised.