Appendix for A Bayesian approach for Structure learning of Oscillating Regulatory Networks

April 24, 2015

1 Gibbs sampler for the Hierarchical Bayesian model.

Likelihood funcion for the LTI system.

$$
\begin{aligned}
\mathrm{p}\left(\mathbf{Q}_k|\sigma\right) &\propto \sigma_D^{-\frac{\mathrm{M}}{2}} \exp\left(-\frac{1}{2\sigma^{-2}}\mathrm{tr}\left(\mathbf{Q}_k^{\mathrm{T}}\mathbf{Q}_k\right)\right) \\
\prod_k \mathrm{p}\left(\mathbf{Q}_k|\sigma\right) &\propto \sigma_D^{-\frac{\mathrm{M+K}}{2}} \exp\left(-\frac{1}{2\sigma^{-2}}\sum_k \mathrm{tr}\left(\mathbf{Q}_k^{\mathrm{T}}\mathbf{Q}_k\right)\right) \\
\mathbf{Q}_k &= \left(\dot{\mathbf{X}}_k - \left[\begin{array}{cc}\mathbf{X}_k & \mathbf{U}_k\end{array}\right]\left[\begin{array}{c}\mathbf{A}^{\mathrm{T}} \\ \mathbf{C}^{\mathrm{T}}\end{array}\right]\right)
\end{aligned}
$$

Conditional distribution of $\left[\begin{array}{c}\mathbf{A}^{\mathrm{T}} \\ \mathbf{C}^{\mathrm{T}}\end{array}\right]$

Let's call $\mathbf{B} = \left[\begin{array}{c}\mathbf{A}^{\mathrm{T}} \\ \mathbf{C}^{\mathrm{T}}\end{array}\right]$ and $\mathbf{R}_k = \left[\begin{array}{cc}\mathbf{X}_k & \mathbf{U}_k\end{array}\right]$

$$
\begin{aligned}
\mathrm{p}\left(\left\{\dot{\mathbf{X}}_k\right\}, \left\{\mathbf{R}_\mathrm{k}\right\}|\mathbf{B}, \sigma_D\right) &\propto \prod_k \mathrm{p}\left(\dot{\mathbf{X}}_k|\mathbf{R}_k\mathbf{B}, \sigma\right)\mathrm{p}\left(\mathbf{B}|\mathbf{H}, \boldsymbol{\tau}\right) \\
&\propto \prod_\mathrm{k} \exp\left(-\frac{1}{2\sigma_D^2}\mathrm{Tr}\left(-2\left(\mathbf{R}_k^{\mathrm{T}}\dot{\mathbf{X}}_k\right)^{\mathrm{T}}\mathbf{B} + \mathbf{B}^{\mathrm{T}}\left(\mathbf{R}_k^{\mathrm{T}}\mathbf{R}_k\right)\mathbf{B}\right)\right) \\
&\propto \exp\left(-\frac{1}{2\sigma_D^2}\mathrm{Tr}\left(-2\boldsymbol{\eta}^{\mathrm{T}}\mathbf{B} + \mathbf{B}^{\mathrm{T}}\boldsymbol{\Psi}\mathbf{B}\right)\right) \\
\boldsymbol{\eta}^{\mathrm{T}} &= \left(\sum_\mathrm{k}\mathbf{R}_k^{\mathrm{T}}\dot{\mathbf{X}}_k\right) \\
\boldsymbol{\Psi} &= \left(\sum_\mathrm{k}\mathbf{R}_k^{\mathrm{T}}\mathbf{R}_k\right)
\end{aligned}
$$

Now by using the vectorization transformation for an arbitrary matrix $\mathbf{M}$, such that $\bar{\mathbf{M}} = \mathrm{vec}\left(\mathbf{M}\right)$ and properties

$$\begin{aligned}\mathrm{Tr}\left(\mathbf{M}^{\mathrm{T}}\mathbf{N}\right) &= \bar{\mathbf{M}}^{\mathrm{T}}\bar{\mathbf{N}}\\ \mathrm{vec}\left(\mathbf{M}\mathbf{N}\right) &= \left(\mathbf{I}\otimes\mathbf{M}\right)\bar{\mathbf{N}}\end{aligned}$$

Then we have

$$\begin{aligned}\mathrm{p}\left(\left\{\dot{\mathbf{X}}_k\right\},\left\{\mathbf{R}_{\mathrm{k}}\right\}|\mathbf{B},\sigma\right) &\propto \exp\left(-\frac{1}{2\sigma_D^2}\left(-2\mathrm{Tr}\left(\boldsymbol{\eta}^{\mathrm{T}}\mathbf{B}\right)+\mathrm{Tr}\left(\mathbf{B}^{\mathrm{T}}\boldsymbol{\Psi}\mathbf{B}\right)\right)\right)\\ &\propto \exp\left(-\frac{1}{2\sigma_D^2}\left(-2\bar{\boldsymbol{\eta}}^{\mathrm{T}}\bar{\mathbf{B}}+\bar{\mathbf{B}}^{\mathrm{T}}\mathrm{vec}\left(\boldsymbol{\Psi}\mathbf{B}\right)\right)\right)\\ &\propto \exp\left(-\frac{1}{2\sigma_D^2}\left(-2\bar{\boldsymbol{\eta}}^{\mathrm{T}}\bar{\mathbf{B}}+\bar{\mathbf{B}}^{\mathrm{T}}\left(\mathbf{I}\otimes\boldsymbol{\Psi}\right)\bar{\mathbf{B}}\right)\right)\end{aligned}$$

then by using vectorization we write the prior in canonical form

$$\begin{aligned}\mathrm{p}\left(\mathbf{B}|\mathbf{H},\boldsymbol{\tau}\right) &\propto \exp\left(-\frac{1}{2}\bar{\mathbf{B}}^{\mathrm{T}}\boldsymbol{\Gamma}\bar{\mathbf{B}}\right)\\ \boldsymbol{\Gamma} &= \mathrm{diag}\left(h_{ij}\tau_{ij}^2\right)\end{aligned}$$

finally multiplying by the gaussian with hypervariance given by the spike and slab prior and completing the square we get

$$\mathrm{p}\left(\left\{\dot{\mathbf{X}}_k\right\},\left\{\mathbf{R}_{\mathrm{k}}\right\}|\mathbf{B},\sigma\right)\mathrm{p}\left(\mathbf{B}|\mathbf{H},\boldsymbol{\tau}\right) \propto \exp\left(-\frac{1}{2\sigma^2}\left(-2\bar{\boldsymbol{\eta}}^{\mathrm{T}}\bar{\mathbf{B}}+\bar{\mathbf{B}}^{\mathrm{T}}\left(\mathbf{I}\otimes\boldsymbol{\Psi}\right)\bar{\mathbf{B}}+\bar{\mathbf{B}}^{\mathrm{T}}\sigma^2\boldsymbol{\Gamma}\bar{\mathbf{B}}\right)\right)$$

$$\mathrm{p}\left(\mathbf{B}|\left\{\dot{\mathbf{X}}_k\right\},\left\{\mathbf{R}_{\mathrm{k}}\right\},\mathbf{H},\boldsymbol{\tau},\sigma\right) \sim \mathcal{N}\left(\bar{\boldsymbol{\mu}},\sigma^2\boldsymbol{\Sigma}\right) \tag{1}$$

$$\bar{\boldsymbol{\mu}} = \boldsymbol{\Sigma}^{-1}\bar{\boldsymbol{\eta}}^{\mathrm{T}} \tag{2}$$

$$\boldsymbol{\Sigma}^{-\mathbf{1}} = \mathbf{I}\otimes\boldsymbol{\Psi}+\sigma^2\boldsymbol{\Gamma} \tag{3}$$

Spike and slab prior over the coefficients

$$\mathrm{p}\left(b_{ij}|h_{ij},\tau\right) \sim \mathcal{N}\left(0,h_{ij}\tau_{ij}^2\right) \tag{4}$$

$$\mathrm{p}\left(h_{ij}|w\right) \sim \left(1-w\right)\delta_{v0}+w\delta_1$$

Conditional distribution of τ^{-2}

$$\mathrm{p}\left(\tau_{ij}^{-2}|a_{ij},h_{ij}\right) \sim \mathrm{Gamma}\left(a_1+\frac{1}{2},a_2+\frac{b_{ij}^2}{2h_{ij}}\right) \tag{5}$$

Conditional distribution of w

$$\mathrm{p}\left(w|\mathbf{H}\right)\sim\mathrm{Beta}\left(c_1+\#\{h_{ij}=1\},c_2+\#\{h_{ij}=v_0\}\right) \tag{6}$$

Condtional distribution of σ_D^{-2}

$$\mathrm{p}\left(\sigma_D^{-2}|\left\{\mathbf{Q}_k\right\}\right)\sim\mathrm{Gamma}\left(b_1+\frac{\mathrm{MK}}{2},b_2+\frac{\sum_k\mathrm{tr}\left(\mathbf{Q}_k^{\mathrm{T}}\mathbf{Q}_k\right)}{2}\right) \tag{7}$$

Likelihood of the similarity scores

$$\mathrm{p}\left(s_{\mathrm{ij}}|\mathbf{H},\boldsymbol{\beta},\sigma_{seq}\right)\quad\propto\quad\sigma_{seq}^{-1/2}\exp\left(-\frac{1}{2\sigma_{seq}^2}\left(s_{ij}-\sum_l^{\mathrm{N}}h_{il}h_{jl}\beta_l\right)^2\right) \tag{8}$$

having vector$\bar{\mathbf{S}}=\left[s_{ij}\right]_{i<j}$ representing the off-diagonal elements of the upper triangular matrix of $\mathbf{S}$, vector$\bar{\mathbf{h}}_i$ representing the $i-th$ row vector of $\mathbf{H}$ and $\circ$ representing the elementwise product (Hadamard product). Then the distribution of the upper diagonal elements is:

$$\mathrm{p}\left(\bar{\mathbf{S}}|\mathbf{H},\boldsymbol{\beta},\sigma_{seq}\right)\quad\propto\quad\sigma_{seq}^{-\mathrm{D}/2}\exp\left(-\frac{1}{2\sigma_{seq}^2}\left(\bar{\mathbf{S}}-\left[\bar{\mathbf{h}}_i\circ\bar{\mathbf{h}}_j\right]_{i<j}\boldsymbol{\beta}\right)^{\mathrm{T}}\left(\bar{\mathbf{S}}-\left[\bar{\mathbf{h}}_i\circ\bar{\mathbf{h}}_j\right]_{i<j}\boldsymbol{\beta}\right)\right)$$

Conditional distribution for σ_{seq}^2

$$\begin{aligned}\mathrm{p}\left(\sigma_{seq}^{-2}|\bar{\mathbf{S}},\boldsymbol{\beta}\right)&\sim\mathrm{p}\left(\bar{\mathbf{S}}|\mathbf{H},\boldsymbol{\beta},\sigma_{seq}^{-2}\right)\mathrm{p}\left(\sigma_{seq}^{-2}\right)\\&\sim\mathrm{Gamma}\left(a_1+\frac{\mathrm{D}^2}{2},a2+\frac{\left(\bar{\mathbf{S}}-\left[\bar{\mathbf{h}}_i\circ\bar{\mathbf{h}}_j\right]_{i<j}\boldsymbol{\beta}\right)^{\mathrm{T}}\left(\bar{\mathbf{S}}-\left[\bar{\mathbf{h}}_i\circ\bar{\mathbf{h}}_j\right]_{i<j}\boldsymbol{\beta}\right)}{2}\right)\end{aligned} \tag{9}$$

Sampling $\boldsymbol{\beta}$

Instead of sampling from $\mathrm{p}\left(\beta_l|\cdot\right)$ we use the non-negative least square estimate of $\boldsymbol{\beta}$, by solving the quadratic form

$$\min_{\boldsymbol{\beta}}\left(\bar{\mathbf{S}}-\left[\bar{\mathbf{h}}_i\circ\bar{\mathbf{h}}_j\right]_{i<j}\boldsymbol{\beta}\right)^{\mathrm{T}}\left(\bar{\mathbf{S}}-\left[\bar{\mathbf{h}}_i\circ\bar{\mathbf{h}}_j\right]_{i<j}\boldsymbol{\beta}\right);\ s.t.\beta_l\geq 0 \tag{10}$$

Conditional distribution for h_{ij}

We define matrices$\mathbf{H}_{ij}^{v0}=\left[\bar{\mathbf{h}}_i\circ\bar{\mathbf{h}}_j\right]_{i<j}$ such that $h_{ij}=v0$, and $\mathbf{H}_{ij}^1=\left[\bar{\mathbf{h}}_i\circ\bar{\mathbf{h}}_j\right]_{i<j}$ such that $h_{ij}=1$.

$$\mathrm{p}\left(h_{ij}|h_{/ij},\boldsymbol{\beta},\sigma_{seq}^{-2},\mathbf{S},\mathbf{B},w,\tau,v_0\right)\quad=\quad P\left(\mathrm{h}_{ij}\mid w\right)P\left(a_{ij}\mid h_{ij},\tau_{ij}^2\right)\mathrm{p}\left(\bar{\mathbf{S}}|h_{/ij},h_{ij},\boldsymbol{\beta},\sigma_{seq}\right)$$

$$
\begin{aligned}
\mathrm{p}\left(h_{ij}|.\right) &\sim \frac{m_{v0}}{m_{v0}+m_1}\left(1-w\right)\delta_{v0}+\frac{m_1}{m_{v0}+m_1}w\delta_1 \qquad (11)\\
m_{v0} &= \sigma_{seq}^{-\mathrm{D}/2}v_0^{-1/2}\\
&\times \exp\left(-\frac{1}{2}\left(\frac{b_{ij}^2}{v_0\tau_{ij}^2}+\frac{1}{\sigma_{seq}^2}\left(\bar{\mathbf{S}}-\mathbf{H}_{ij}^{v0}\boldsymbol{\beta}\right)^{\mathrm{T}}\left(\bar{\mathbf{S}}-\mathbf{H}_{ij}^{v0}\boldsymbol{\beta}\right)\right)\right)\\
m_1 &= \sigma_{seq}^{-\mathrm{D}/2}\\
&\times \exp\left(-\frac{1}{2}\left(\frac{b_{ij}^2}{\tau_{ij}^2}+\frac{1}{\sigma_{seq}^2}\left(\bar{\mathbf{S}}-\mathbf{H}_{ij}^{1}\boldsymbol{\beta}\right)^{\mathrm{T}}\left(\bar{\mathbf{S}}-\mathbf{H}_{ij}^{1}\boldsymbol{\beta}\right)\right)\right)
\end{aligned}
$$

1.1 Algorithm

The Algorithm **??** presents the Gibbs sampler for parameter inference in the Hierarchical model. A brief overview of the inputs and outputs is also presented.

1.1.1 Program Inputs

Due that the DFT computed by the FFT algorithm yields complex numbers, a real representation of the coefficients is needed. We work with the so-called Real Discrete Fourier Transform. It consists of stacking the real over the imaginary part of the first $(M-1)/2$ FFT coefficients. We denote this $M \times N$ matrix $\mathbf{X}$. Figures (**??**) and (**??**) show plots for the DFT real and imaginary parts of the *A. thaliana* data-set, with a photoperiod of 12 hours. The magnitude and phase spectra is also shown, finally the RDFT is presented at the bottom of the picture.

Hyperparameters a_1 and a_2 are set to 1 so w is uniformly distributed in the interval $[0,1]$. Parameters b_1 and b_2 are set so the hypervariance $h_{ij}\tau^2$ has a continuos bi-modal distribution, according to the recommendations of [**?**], in which they set them to 5 and 50 respectively. Alternative parametrization of 50 and 500 was also tested yielding better results in some cases.

The hyperparameters c_1 and $c_2$forσ_D^{-2} are set to 0.001 and 0.001, this is a weak prior reflecting uncertainty about the linearity of the system. Hyperparamters d_1 and d_2 are set to 10 and 0.001, this parametrization required a manual tunning, as the scale parameter σ_{seq}^{-2} having a weak prior resulted in the effects of the sequence similarity model to banish. By modifying this prior we can give more “weight” to the sequence similarity clustering thus, the flexibility of the Hierarchical model.

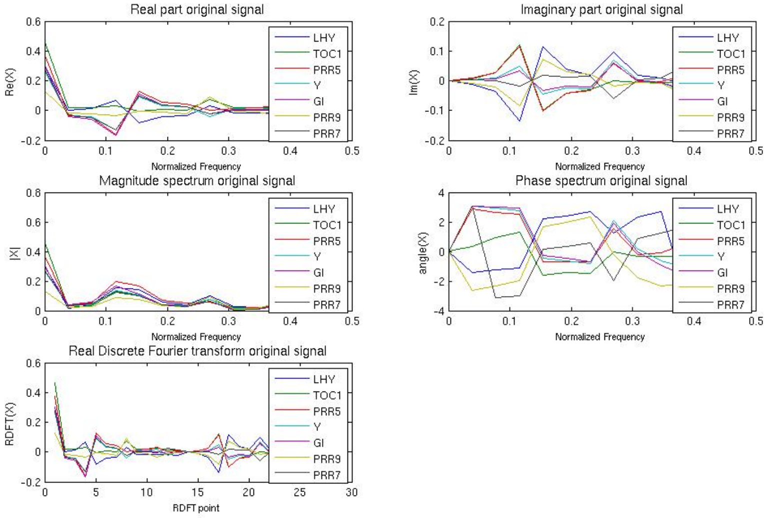


Figure 1: Spectra for the *A. thaliana* circadian clock simulation with a 12 hour photo-period

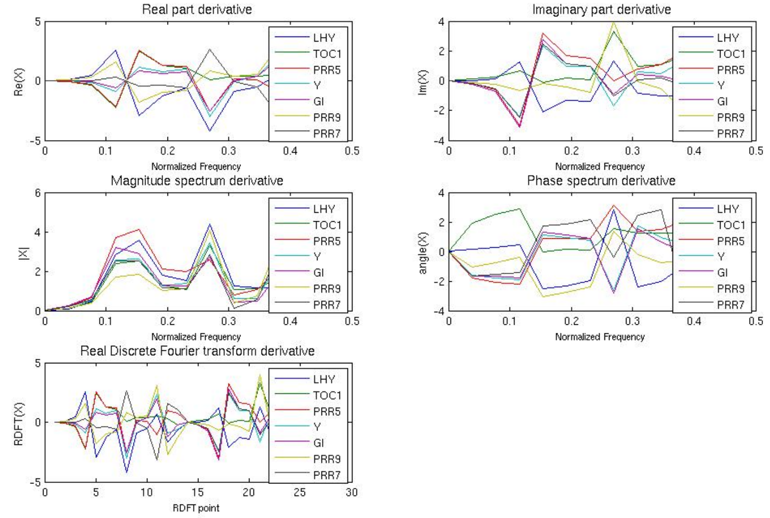


Figure 2: Derivative Spectra for the *A. thaliana* circadian clock simulation with a 12 hour photo-period

1.1.2 *A. thaliana* light inputs U

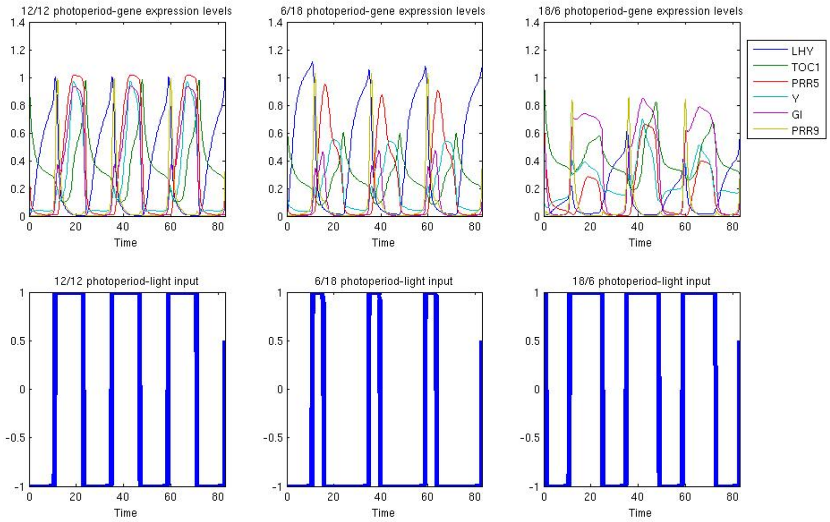


Figure 3: Three examples of light input for the *A. thaliana* circadian clock simulation.

1.1.3 Output

The Gibbs sampler presented in the previous section allows us to draw samples from the joint conditional distribution $p\left(\mathbf{H},\mathbf{A},\mathbf{C},\beta,w,\tau,\sigma_D,\sigma_s\middle|\left\{\mathbf{X_k}\right\}\right)$. The marginal distribution for each of the models parameters can be drawn from this joint distribution, and the expected value for each parameter equals to the average of the samples.

For example, figure illustrates the expected value for matrix $\mathbf{H}$ obtained from averaging over 1000 samples drawn from the marginal distribution $p\left(\mathbf{H}\middle|\left\{\mathbf{X_k}\right\}\right)$. This figure shows in dark red those elements with higher probability of a regulatory interaction under the model assumptions, except the diagonal elements, which represent the decay rates of the equation model. The AUPR were computed by thresholding the off-diagonal elements of this matrix for each data-set.

Algorithm 1 Algorithm for the DFT-based Spike and slab prior model with sequence similarity.

```
Inputs: $K$ time series of $M$ time-points for $N$ gene expression levels, encoded in matrices $\{x_k\}$. Prior hyper-parameters $a_1, a_2, b_1, b_2, c_1, c_2$. Optional similarity matrix $\mathbf{S}$.
Outputs: Joint conditional posterior distribution $p\left(\mathbf{H}, \mathbf{A}, \mathbf{C}, \beta, w, \tau, \sigma_D, \sigma_s | \left\{\mathbf{X_k}\right\}\right)$

1. Obtain the DFT of $\{x_k\}$ and the corresponding RDFT coefficient matrices $\{\mathbf{X_k}\}$
2. Compute the derivatives $\left\{\dot{\mathbf{X}}_\mathbf{k}\right\}$
3. Sample from the conditional distribution over the LTI coefficients, given in eq. (??)
4. Sample from the conditional distribution over$\tau^{-2}$ given by eq. (??)
5. Sample $\mathbf{H}$ from eq. (??), to account for the decay rates we set diagonal elements $h_{ii}$ to 1, and set the diagonal elements of matrix $\mathbf{A}$ to negative.
6. Sample $w$ from eq. (??)
7. Sample $\sigma_D$ from eq. (??)
8. OPTIONAL sample $\sigma_{seq}$ from eq. (??) and $\beta$ from the nonnegative least squares solution to equation (??).
9. Return to step 3

Note: A burn in period of 4000 samples is considered in the general purpose implementation of the model.
```

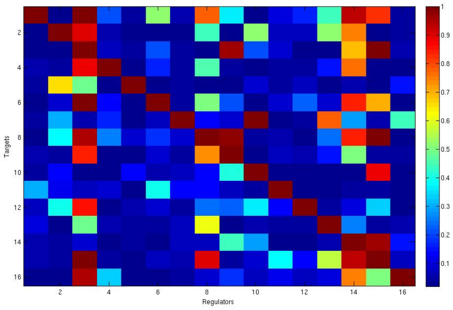
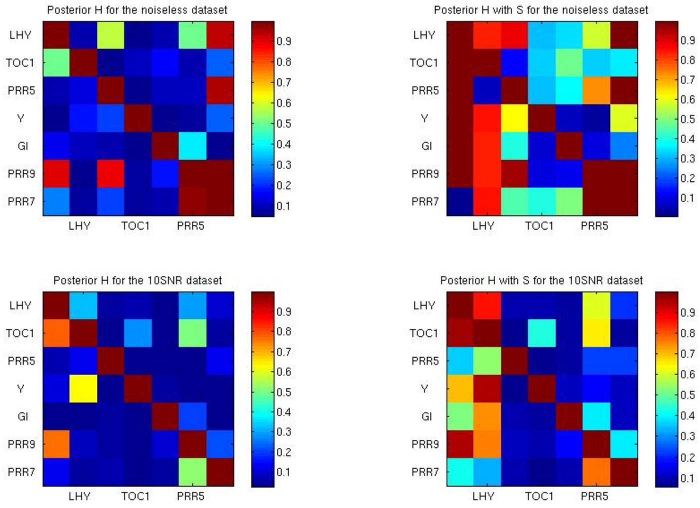


Figure 4: Heat-map representing the expected value for $p\left(\mathbf{H}|.\right)$ obtained by averaging the last 1000 samples. Rows represent targets and columns regulators. The diagonal indicates the decay parameters λ.

2 Heatmaps and PR curves for the experiments


Posterior H for the noiseless dataset

Posterior H with S for the noiseless dataset

Posterior H for the 10SNR dataset

Posterior H with S for the 10SNR dataset

Figure 5: Heatmaps representing the posterior probability for the *A. thaliana* circadian clock network

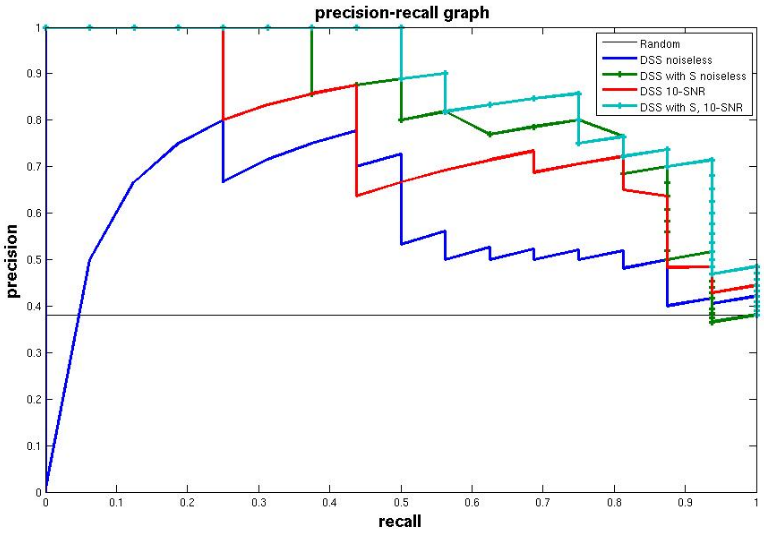


Figure 6: PR curves for the *A. thaliana* circadian clock network

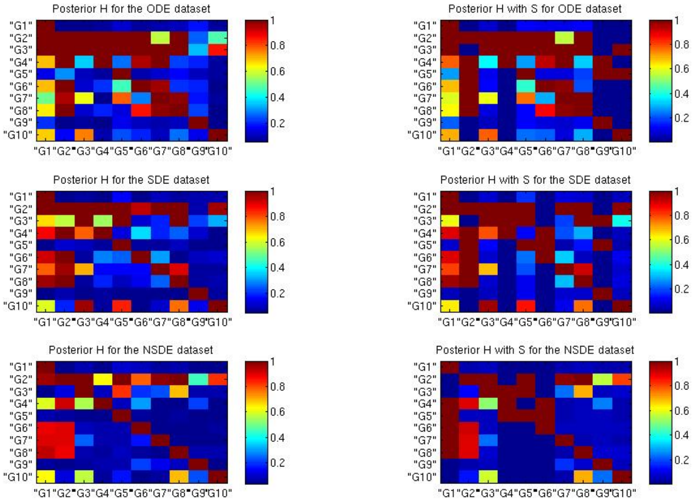


Figure 7: Heatmaps representing the posterior probability for the Dream4 Challenge network

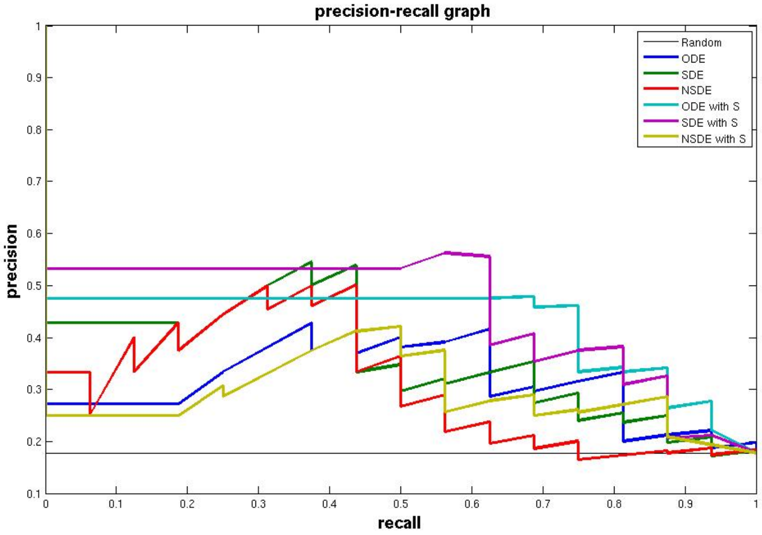


Figure 8: PR curves for the DREAM4 challenge network

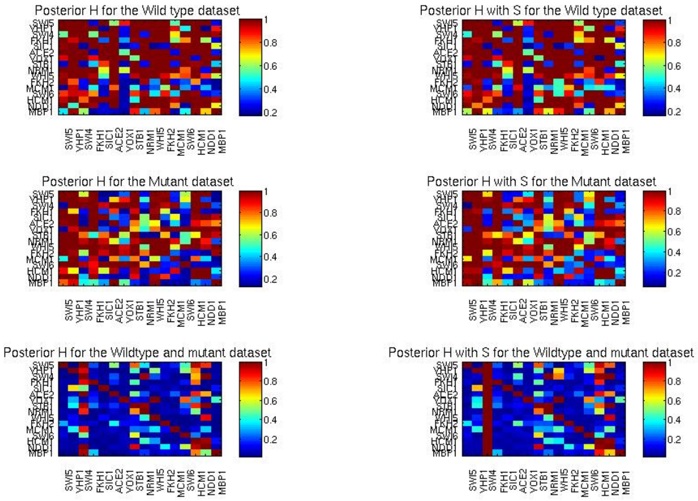


Figure 9: Heatmaps representing the posterior probability for the *s. cerevisiae* network

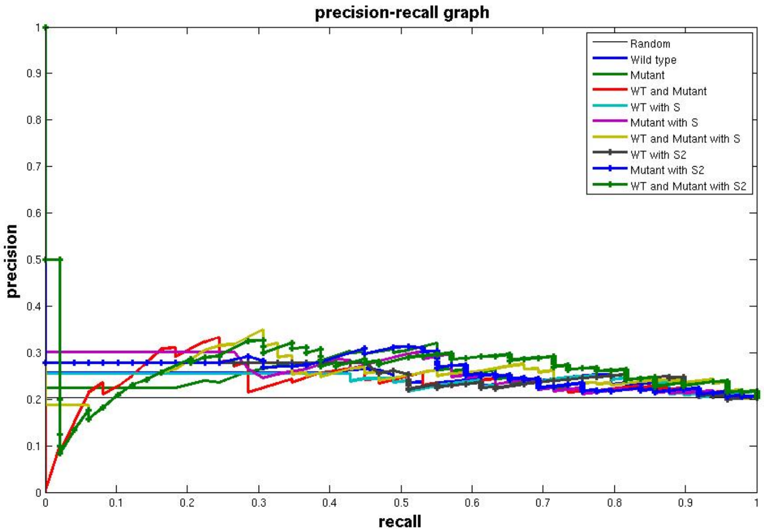


Figure 10: PR curves for the *s. cerevisiae* network

3 *A. thaliana* circadian clock

The arabidopsis thaliana circadian clock model as presented in, is shown in Fig. **??**.

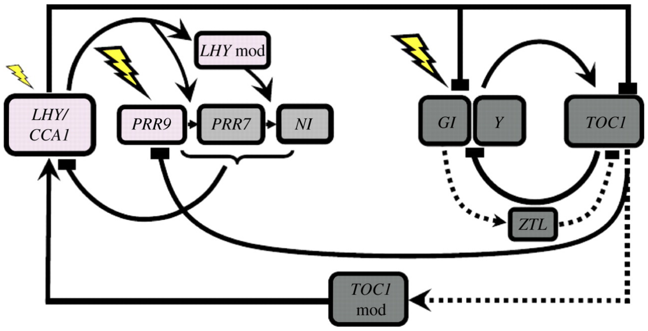


Figure 11: *A. thaliana* circadian clock model, transcriptional elements LHY, PRR9, PRR7, NI, Y, and TOC1. Post-transcriptional elements ZTL, TOC1mod and LHYmod. Light input is represented by a lighting symbol. Activating interactions are represented by solid line with arrows, repression by solid line with rectangles at the end, post transcriptional interactions are represented by dashed lines. .

4 DREAM4 network

The 10-node oscillatory network that was part of the DREAM4 challenge supplementary information data set is presented in Fig. **??**.

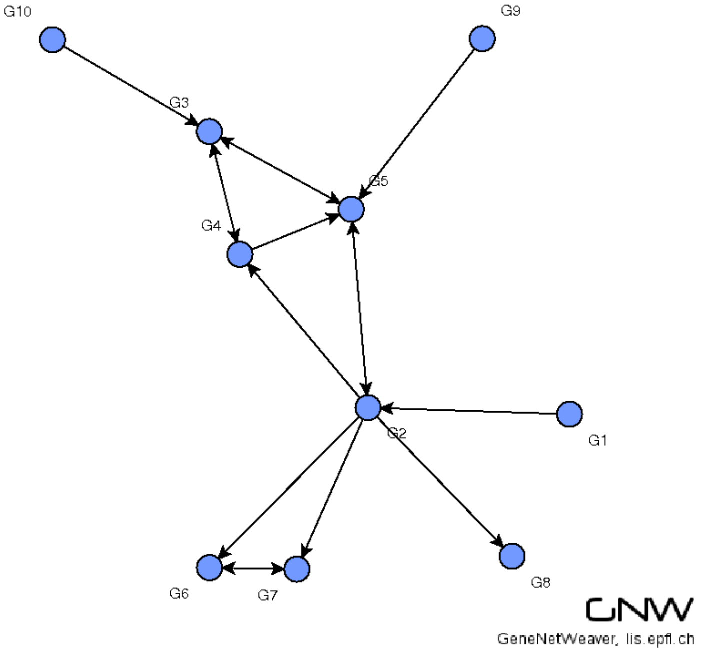
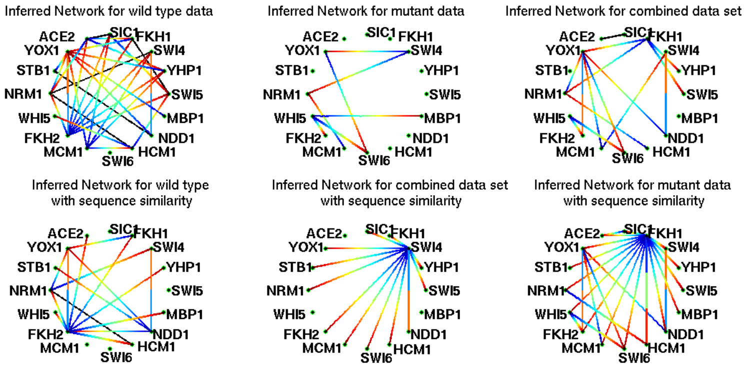


Figure 12: DREAM4 challenge network with 10 nodes, of those 3 are inputs, node G9 was subjected to a perturbation for half the time points.

5 *S. cerevisae* cell cycle network

In Fig. **??** the inferred networks after thresholding the value of $p\left(h_{ij}=1\right)$ are presented, the putative ground truth matrix is presented on Fig. **??**



Inferred Network for wild type data

Inferred Network for mutant data

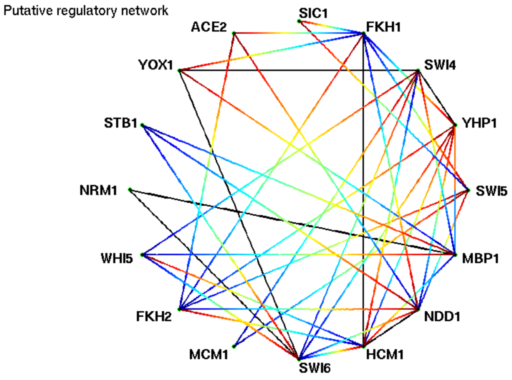
Inferred Network for combined data set

Inferred Network for wild type with sequence similarity

Inferred Network for combined data set with sequence similarity

Inferred Network for mutant data with sequence similarity

Figure 13: Inferred yeast networks for different data subsets with and without sequence information, edges go from blue (regulators) to red (targets)..



Putative regulatory network

Figure 14: Putative yeast regulatory network edges go from blue (regulators) to red (targets)..

The putative regulatory network was built based on these references.

- Regulation of SIC1 by SWI5 as in [?].
- Regulation of SWI4 by YHP1 as in [?].
- Regulation of YHP1, SWI4, YOX1 and HCM1 by SWI4 as in [?].
- Regulation of SWI5 and ACE2 [?]; YHP1 [?]; SIC1, YOX1 and HCM1 [?]; NDD1 [?]; by FKH1.
- Regulation of SWI6 and MBP1 by NRM1 as in [?].

- Regulation of SWI6, SWI4 and MBP1 by WHI5 as in [**?**].
- Regulation of SWI5, YHP1 and FKH1 [**?**]; ACE2 and NDD1 [**?**] by FKH2.
- Regulation of SWI4 and SWI5 by MCM1 as in [**?**].
- Regulation of SWI4, FKH1, YOX1, NRM1, HCM1 and NDD1 as in [**?**].
- Regulation of YHP1, FKH1, FKH2, WHI5 and NDD1 by HCM1 as in [**?**].
- Regulation of SWI5 and ACE2 [**?**]; YHP1 and HCM1 as in [**?**].
- Regulation of YHP1, FKH1, YOX1, NRM1 and HCM1 by MBP1 as in [**?**].
- Regulation of NDD1 [**?**]; SWI6 and MBP1 by the interaction of transcription factor MBF with STB1 [**?**].
- Regulation of SWI4 and SWI6 by its cobinding with MCM1p as in [**?**].

References

[Bahler, 2005] Bahler (2005), Cell-cycle control of gene expression in budding and fission yeast. *Annu Rev Genet*, **39** 69–94.

[Haase *et al*, 2014] Haase *et al* (2014),Topology and control of the cell-cycle-regulated transcriptional circuitry, *Genetics*, **196(1)**, 65–90.

[Ishwaran *et al.*, 2005] Ishwaran *et al.* (2005), Spike and slab variable selection: Frequentist and bayesian strategies, *The Annals of Statistics*, **33(2)**,730–773.

[Macisaac *et al*, 2006] Macisaac *et al* (2006), An improved map of conserved regulatory sites for Saccharomyces cerevisiae, *BMC Bioinformatics*, **7(113)**.

[Marbach *et al*, 2010] Marbach *et al* (2010), Revealing strengths and weaknesses of methods for gene network inference, *PNAS* **107(14)**, 6286–6291

[Ostrow *et al*, 2014] Ostrow *et al* (2014),Fkh1 and Fkh2 Bind Multiple Chromosomal Elements in the S. cerevisiae Genome with Distinct Specificities and Cell Cycle Dynamics, *PLoS ONE* **9(2)**.

[Pokhilko *et al.*, 2010] Pokhilko *et al.* (2010), Data assimilation constrains new connections and components in a complex, eukaryotic circadian clock model, *Molecular Systems Biology*, 6.

[Pramila *et al*, 2006] Pramila *et al* (2006),The Forkhead transcription factor Hcm1 regulates chromosome segregation genes and fills the S-phase gap in the transcriptional circuitry of the cell cycle, *Genes Dev*, **20(16)**, 2266–78.

[Stillman, 2013] Stillman (2013),Dancing the cell cycle two-step: regulation of yeast G1-cell-cycle genes by chromatin structure, *Trends Biochem Sci*, **38(9)**, 467–75.

[Venters *et al*, 2011] Venters *et al* (2011), A comprehensive genomic binding map of gene and chromatin regulatory proteins in *S*accharomyces, *Mol Cell*, **41(4)**, 480–92.

[Weiss, 2012] Weiss (2012), YeastBook - Cell Signaling and Development: Mitotic Exit and Separation of Mother and Daughter Cells, *Genetics* **192**, 1165–1202;