OKVAR-Boost: a novel boosting algorithm to infer nonlinear dynamics and interactions in gene regulatory networks (Supplementary material)

Néhémy Lim, Yasin Şenbabaoğlu, George Michailidis and Florence d'Alché-Buc

1 Insights on OKVAR model: matrix-valued kernel properties

The matrix-valued function K presented in Eq. (4) is the Hadamard product of two kernels K_0 and K_1, each of them satisfying the properties of a matrix-valued kernel introduced by Senkene and Tempel'man, 1973 : (1), $\forall(\mathbf{x},\mathbf{x}') \in \mathbb{R}^p \times \mathbb{R}^p$, $K(\mathbf{x},\mathbf{x}') = K(\mathbf{x}',\mathbf{x})^T$ and (2): $\forall m \in \mathbb{N}$, $\forall \{(\mathbf{x}_i,\mathbf{y}_i)\}_{i=1\dots m} \subseteq \mathbb{R}^p \times \mathbb{R}^p$,

$$\sum_{i,j=1}^{m} \langle K(\mathbf{x}_i,\mathbf{x}_j)\mathbf{y}_i, \mathbf{y}_j\rangle_{\mathbb{R}^p} \geq 0 \ .$$

Due to a theorem proven in Caponnetto *et al.*, 2008, as a Hadamard product of two matrix-valued kernels, it is also a matrix-valued kernel.

2 Estimation of W_m

Empirical partial correlations[1] r_{ij}'s are computed for each pair of variables (i,j) conditional on all other variables from data projected on the subspace defined by $\mathcal{S}_m$. The r_{ij}'s can be computed from the inverse of $\hat{\Sigma}$, the empirical covariance matrix as follows: $r_{ij} = \frac{-\hat{\Sigma}^{-1}_{ij}}{\sqrt{\hat{\Sigma}^{-1}_{ii}\hat{\Sigma}^{-1}_{jj}}}$. If we assume that the variables are distributed according to a multivariate normal distribution, r_{ij} is zero if, and only if, states i and j are conditionally independent given the other variables. We carry out a statistical test based on the null hypothesis $H_0 : r_{ij} = 0$ (no partial correlation between i and j) vs $H_1 : r_{ij} \neq 0$. The test statistic is a Fisher's z-transform of the partial correlation : $z(r_{ij}) = \frac{1}{2}\ln\left(\frac{1+r_{ij}}{1-r_{ij}}\right)$. The null hypothesis H_0 is rejected with significance level α if : $\sqrt{(N-1)-(k-2)-3}\cdot|z(r_{ij})| > \Phi^{-1}\left(1-\frac{\alpha}{2}\right)$ where Φ is the cumulative distribution function of a standard normal distribution $\mathcal{N}(0,1)$. We define W_m as follows : $w^{(m)}_{ij} = 1$ if H_0 is rejected, 0 otherwise and obtain B_m as the Laplacian of W_m, e.g. as $B_m = D_m - W_m$ where D_m is the degree matrix of W_m.

3 Jacobian expression for a base model h_m

$$\begin{aligned}
J^{(m)}_{ij}(t) &= \sum_{k=0}^{N-2}\sum_{\ell=1}^{p} c^{(m)}_{k,\ell}\frac{\partial(K^{(m)}(\mathbf{x}_k,\mathbf{x}_t)_{i\ell})}{\partial(\mathbf{x}_t)_j} \\
&= \sum_{k=0}^{N-2}\sum_{\ell=1}^{p} c^{(m)}_{k,\ell}\frac{\partial}{\partial(\mathbf{x}_t)_j}\left(b^{(m)}_{i\ell}\exp(-\gamma_0||\mathbf{x}_k-\mathbf{x}_t||^2)\ \exp\left(-\gamma_1(x_{ki}-x_{t\ell})^2\right)\right) \\
&= \sum_{k=0}^{N-2}\sum_{\ell=1}^{p} b^{(m)}_{i\ell}c^{(m)}_{k,\ell}\left[\exp(-\gamma_0||\mathbf{x}_k-\mathbf{x}_t||^2)\frac{\partial}{\partial(\mathbf{x}_t)_j}\ (\exp\left(-\gamma_1(x_{ki}-x_{t\ell})^2\right)\right. \\
&\quad \left. + \exp\left(-\gamma_1(x_{ki}-x_{t\ell})^2\right)\frac{\partial}{\partial(\mathbf{x}_t)_j}(\exp(-\gamma_0||\mathbf{x}_k-\mathbf{x}_t||^2)\right]
\end{aligned}$$

Finally, we get the expression for the instantaneous Jacobian of h_m:

[1] We omit the index m for the sake of clarity

$$
\begin{aligned}
J_{ij}^{(m)}(t) &= 2b_{ij}^{(m)}\gamma_1 \sum_{k=0}^{N-2} c_{k,j}^{(m)} \exp(-\gamma_0||\mathbf{x}_k - \mathbf{x}_t||^2)(x_{ki} - x_{tj}) \exp\left(-\gamma_1(x_{ki} - x_{tj})^2\right) \\
&\quad +2\gamma_0 \sum_{k=0}^{N-2}\sum_{\ell=1}^{p} c_{k,\ell}^{(m)} b_{i\ell}^{(m)} (x_{kj} - x_{tj}) \exp(-\gamma_1(x_{ki} - x_{t\ell})^2) \exp(-\gamma_0||\mathbf{x}_k - \mathbf{x}_t||^2)
\end{aligned}
$$

Note that when γ_0 is close to 0, $K^{(m)}(\mathbf{x}_k, \mathbf{x}_t)_{ij} \approx b_{ij}^{(m)} \exp\left(-\gamma_1(x_{ki} - x_{tj})^2\right)$ and the Jacobian coefficient $J_{ij}^{(m)}(t) \approx 2b_{ij}^{(m)}\gamma_1 \sum_k c_{k,j}^{(m)}(x_{ki} - x_{tj}) \exp\left(-\gamma_1(x_{ki} - x_{tj})^2\right)$, meaning that the value b_{ij} is central to impose the zero's to the model. The average Jacobian of h_m writes as follows:

$$
J_{ij}^{(m)} = \frac{1}{N-1} \sum_{t=0}^{N-2} J_{ij}^{(m)}(t)
$$

4 Block-instability criterion for model selection

The BIS criterion is defined from a given time series and for a choice of λ_1 and λ_2. It measures the empirical mean of the difference using the Frobenius norm between the Jacobian matrices $J(H_{b,1})$ and $J(H_{b,2})$ computed from a pair of models $(H_{b,1}, H_{b,2})$ built from the b^{th} pair of block-bootstrapped subsamples.

$$
BIS(\lambda_1, \lambda_2; \mathbf{x}_0^{N-1}) = \frac{1}{B} \sum_{b=1}^{B} \parallel J(H_{b,1}) - J(H_{b,2}) \parallel^2 \tag{1}
$$

where $H_{b,1}$ (resp. $H_{b,2}$) is the autoregressive model built from the block sample $(b, 1)$ (resp. $(b, 2)$) drawn from a single time series $\mathbf{x}_0, \ldots, \mathbf{x}_{N-1}$.

5 Tables

Table S1: Average-degree, density, and modularity for DREAM3 networks.

Size10	Ecoli1	Ecoli2	Yeast1	Yeast2	Yeast3
Average degree	2.2	3.0	2.0	5.0	4.4
Density	0.244	0.333	0.222	0.556	0.489
Modularity	0.016 (2)	0 (1)	0.260 (3)	0 (1)	0 (1)
Size100	Ecoli1	Ecoli2	Yeast1	Yeast2	Yeast3
Average degree	2.5	2.38	3.32	7.78	11.02
Density	0.025	0.024	0.033	0.079	0.111
Modularity	0.643 (6)	0.661 (7)	0.681 (8)	0.328 (6)	0.088 (14)

The total degree of a node (gene) is the sum of its in- and out-degrees, while the average across genes gives the **average-degree** for the entire network. The **density** of a network is the ratio of the number of edges in the network to the maximum possible number of edges. The calculated *modularity* index corresponds to the optimal number of modules for that network (numbers given in parentheses).

Table S2: AUROC and AUPR for OKVAR-Boost, run on size-10 Ecoli1 network. Average ± Standard Deviation values are computed on using combinations of 1 up to 4 time series.

Number of time series	1	2	3	4
AUROC	0.665 ± 0.088	0.696 ± 0.101	0.715 ± 0.041	0.853
AUPR	0.272 ± 0.081	0.273 ± 0.137	0.270 ± 0.092	0.583

Table S3: AUROC and AUPR for OKVAR-Boost ($\lambda_1 = 0.01, \lambda_2 = 0.1$ selected by *Block-Stability*), LASSO and Äijö's algorithm run on the IRMA network. All the results are obtained using either the four switch-off time series or the five switch-on time series. The numbers in **boldface** are the maximum values for each column.

	Switch-off		Switch-on	
	AUROC	AUPR	AUROC	AUPR
OKVAR-Boost	0.807	0.807	**1**	**1**
LASSO	0.500	0.253	0.583	0.474
Äijö	**0.875**	**0.848**	0.838	0.836

6 Figures

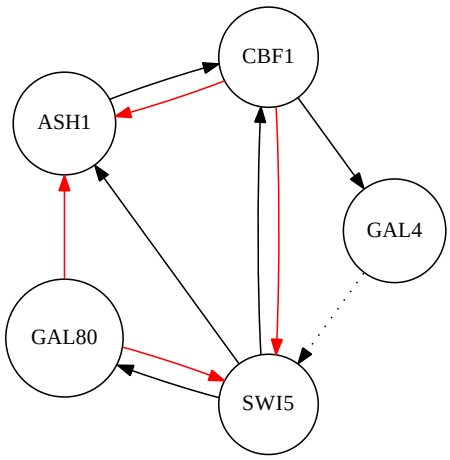

Figure S1: Inferred IRMA network from OKVAR-Boost algorithm using switch-off time series ($\lambda_1 = 0.01, \lambda_2 = 0.1$). Solid, dotted and red lines correspond respectively to correct, missing and false edges.

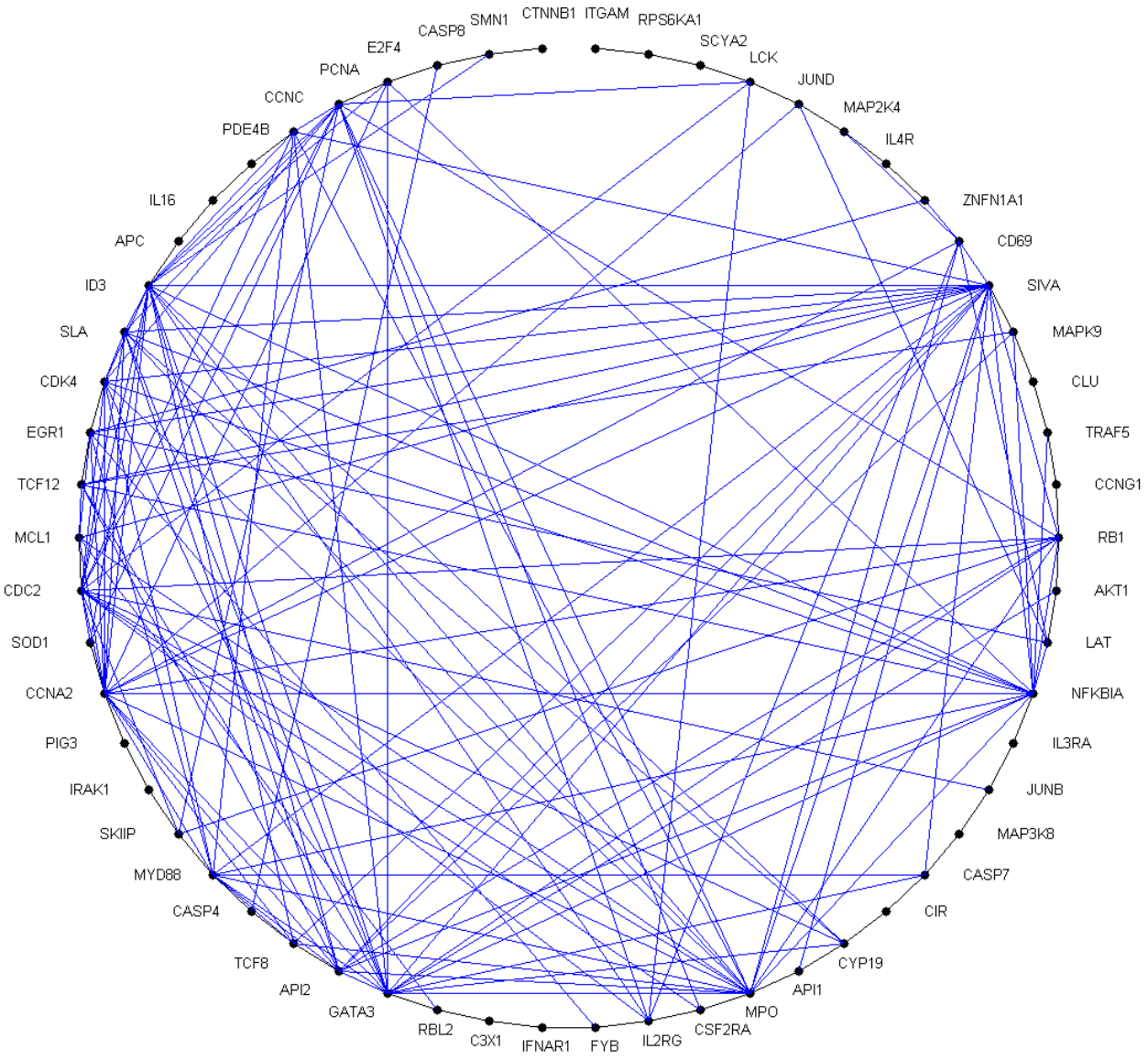

Figure S2: Reconstructed T-cell activation regulatory network using OKVAR-Boost ($\lambda_1 = \lambda_2 = 1$, consensus threshold= 0.01).