

Supplementary information for Boolean network inference and robust solution identification applied to human embryology

Célia Messaoudi^{1,*}, Laurent David^{2,3,4}, Jérémie Bourdon¹, Carito Guziolowski^{1,*}

September 8, 2026

¹ Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, F-44000 Nantes, France

² Nantes Université, CHU Nantes, Inserm, CR2TI, F-44000 Nantes, France

³ Nantes Université, CHU Nantes, CNRS, Inserm, BioCore, F-44000 Nantes, France

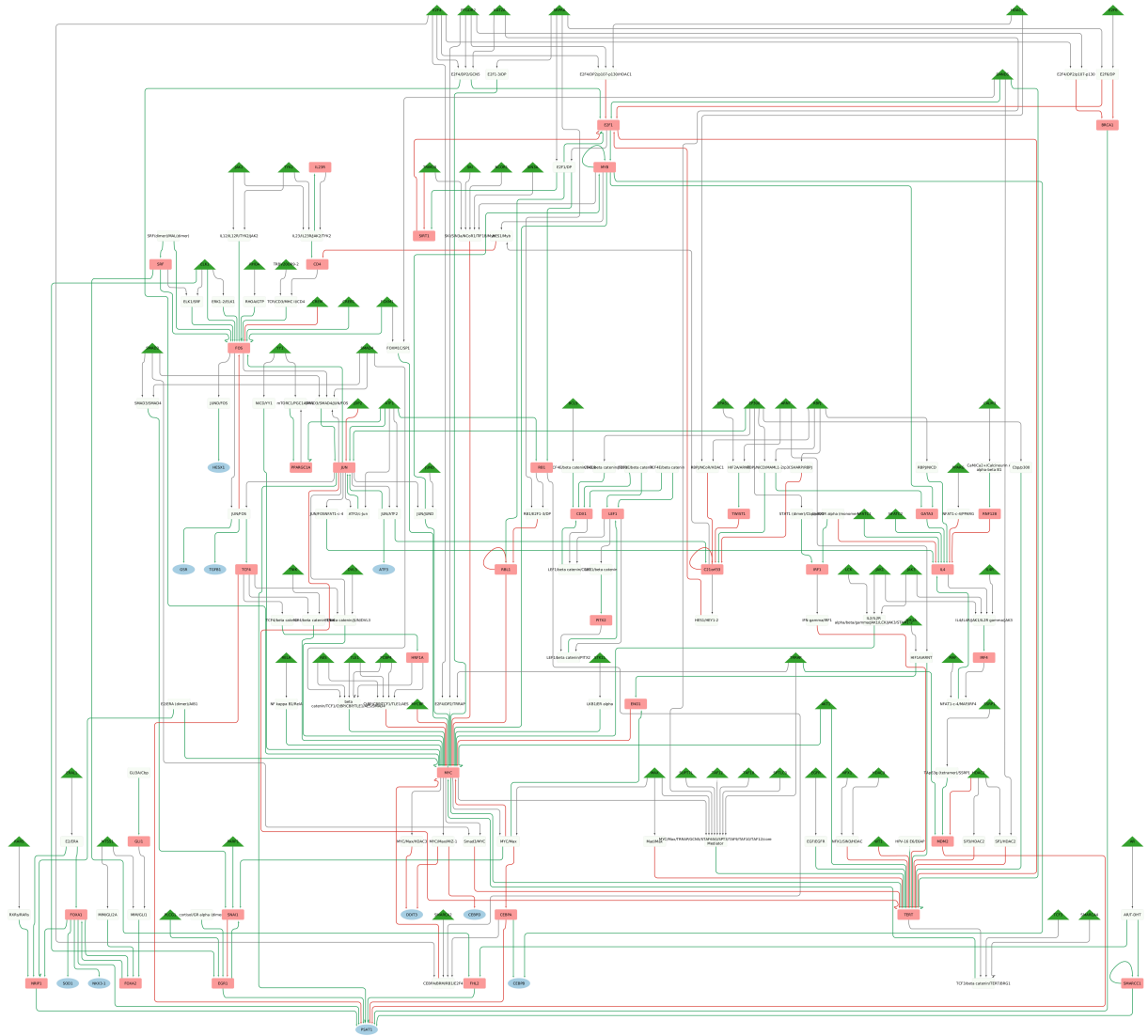
⁴ CHU Nantes, Service de Médecine et Biologie de la Reproduction, F-44000 Nantes, France

* Corresponding authors: celia.messaoudi@ls2n.fr, carito.guziolowski@ls2n.fr

Contents

1	Supplementary Figure 1. Strictly reduced PKN	3
2	Supplementary Figure 2. Linking Accuracy and Jaccard Index	4
3	Supplementary Figure 3. Results for pGMM scenario	5
4	Supplementary Figure 5. Results for $k = 13$ input-intermediate control genes	6
5	Supplementary Figure 4. Comparing solutions close and far from the convergence point	7

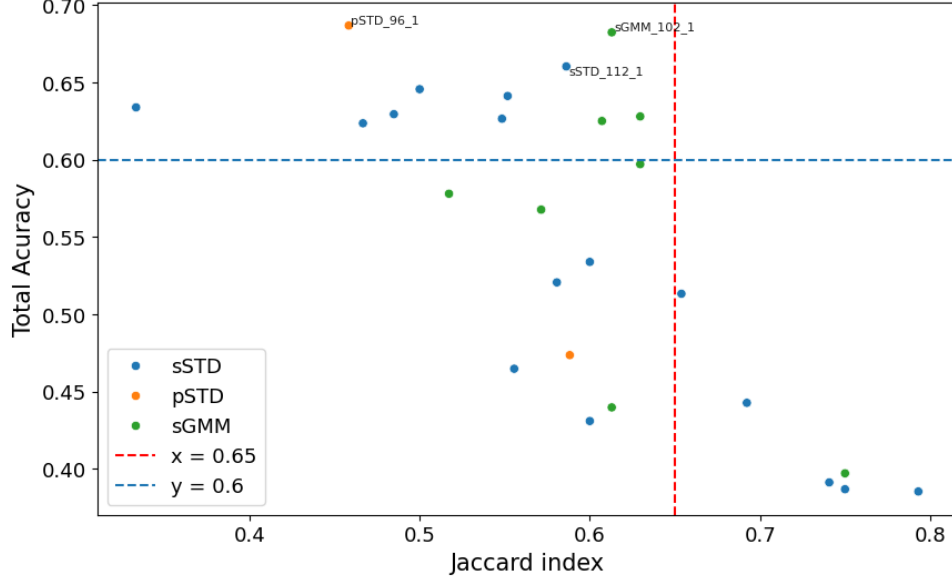
1 Supplementary Figure 1. Strictly reduced PKN



Supplementary Figure S1: Strictly reduced PKN. Prior knowledge network reconstructed using Pathway Commons v14, a list of transcription factors involved in human embryonic development and a strict reduction function. *Green triangles:* inputs *Red rectangles:* intermediates *White rectangles:* complexes *Blue circles:* readouts *Red arrows:* inhibition *Green arrows:* activation *Gray arrows:* part of.

2 Supplementary Figure 2. Linking Accuracy and Jaccard Index

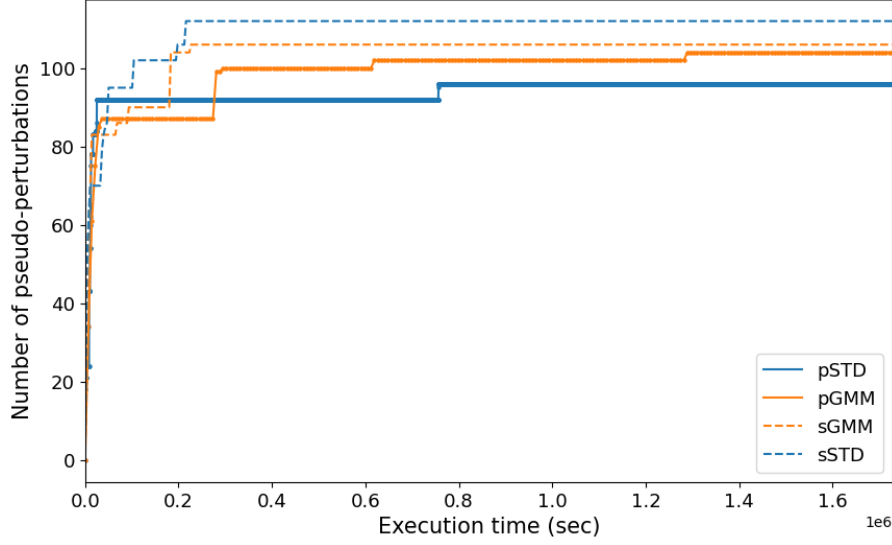
We plotted the total accuracy and the Jaccard index (JI) for the 27 near-optimal solutions presented in the main study, along with their respective thresholds on Fig. S2, to better visualize the relationship between the two metrics. The 11 best solutions, listed in Table 3 of the main study, are all located in the upper-left section of the figure. Notably, none of the best-performing solutions has a JI above the 0.65 threshold. This can be explained by the fact that highly similar BNs provide less distinction between cells, making cell classification more difficult.



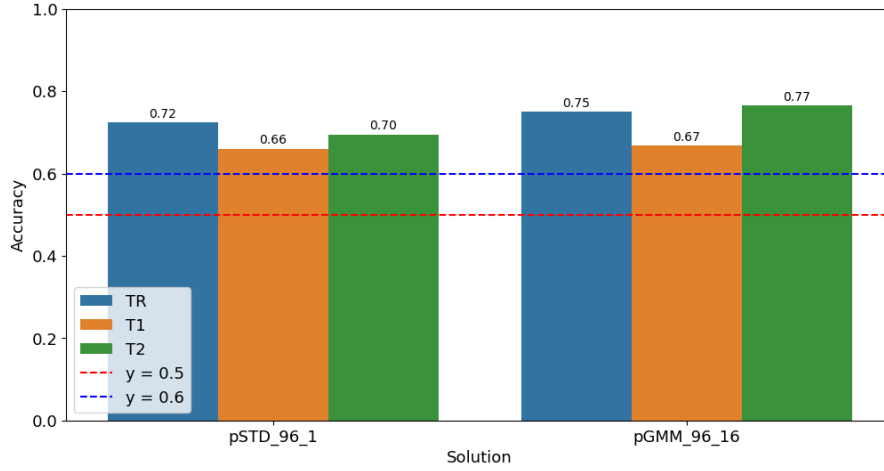
Supplementary Figure S2: The best solutions have a lower Jaccard index. The best solutions are situated in the upper-left part of the graph and the three best ones are annotated. The blue dashed line marks $y = 0.6$ corresponding to the threshold applied to the total accuracy of the best solutions in the main study. The red dashed line marks $x = 0.65$ which corresponds to the recommended upper threshold for Jaccard index. Each scenario is plotted with a different color.

3 Supplementary Figure 3. Results for pGMM scenario

We focus now on comparing STD and GMM binarizations when using a permissive PKN reduction. The results are seen in Fig. S3a: 96 pseudo perturbations were found for the pSTD scenario and 104 for the pGMM one. The dashed lines represent the dynamic of the pseudo-perturbations search for the sSTD and sGMM scenarios, presented in the main study, and serve as a reference for comparisons with the permissive scenarios. We note that less pseudo-perturbations are found when using a permissively reduced PKN than a strictly reduced one. We also assessed the cell classifying performance of solutions with 96 pseudo-perturbations. The pSTD approach yielded 2 equivalent solutions while the pGMM one yielded 52 at that stage. Figure S3b show the results for the best solution of each scenario with 96 pseudo-perturbations. *pGMM_96_16* performs better than *emphpSTD_96_1* on all three datasets. The difference is particularly important for TR and T2 sets where the pGMM solution reaches 0.75 and 0.77, respectively, whereas the pSTD solution achieves 0.72 and 0.70.



(a) Convergence of the pseudo-perturbations search for all new scenarios

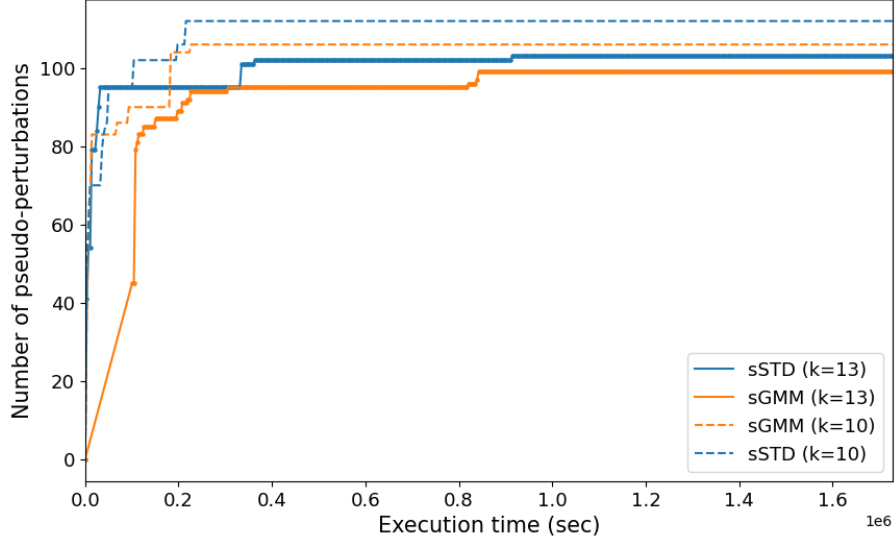


(b) Accuracy of optimal solutions using permissive PKN reduction

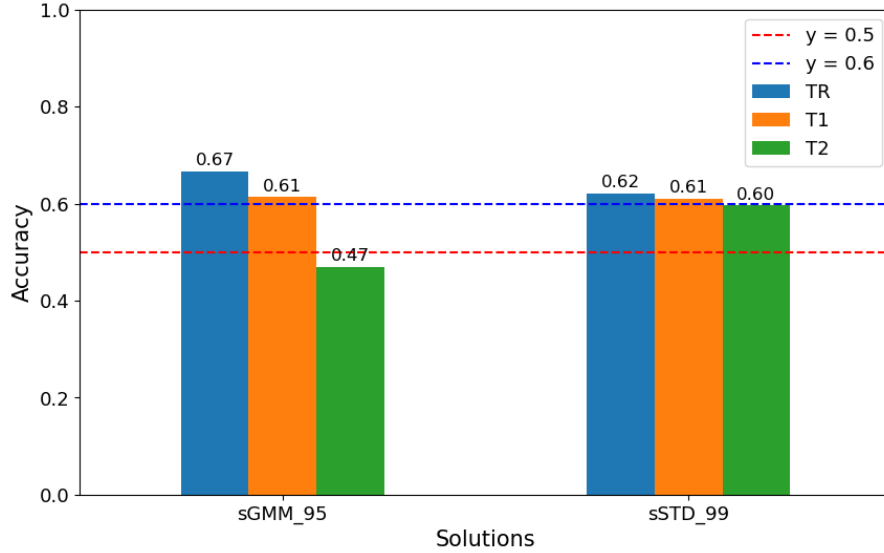
Supplementary Figure S3: Comparing pSTD and pGMM. (a) Number of pseudo-perturbations found across time for strict scenarios (sSTD and sGMM) in dashed lines and permissive scenarios (pGMM and pSTD) in solid lines. Computation was run for 20 days on a computer cluster comprising 4992 CPUs and 22 To of RAM. The pSTD results were computed on an older computer cluster for 20 days. 96 pseudo-perturbations were found for the pSTD case and 104 for the pGMM one. (b) Accuracy scores shown for the best solution with 96 pseudo-perturbations of each scenario using the permissively reduced PKN. The results correspond to the performance of one solution in each scenario. The red dashed line marks $y = 0.5$ which corresponds to the performance of a random classifier. The blue dashed line marks $y = 0.6$ corresponding to the threshold applied to the total accuracy of the best solutions in the main study. **Training set:** cells selected for the pseudo-perturbations **Testing set 1:** redundant cells **Testing set 2:** unselected cells **PKN:** Prior Knowledge Network

4 Supplementary Figure 5. Results for $k = 13$ input-intermediate control genes

As seen on figure S4, the initial pseudo-perturbation search phase is slower than the one for $k = 10$ (fig. S4a) and less pseudo-perturbations are found with $k = 13$. 99 pseudo-perturbations are found for sSTD and 95 for sGMM with $k = 13$, whereas 112 and 106 were found for sSTD and sGMM, respectively, with $k = 10$. This happens because the problem is more difficult with $k = 13$. Indeed, finding pairs of cells with the same expression for k genes increases in difficulty as k increases. The solutions have good accuracies but not better than the solutions obtained with $k = 13$. Therefore, choosing a lower value for k seems more reasonable.



(a) Convergence of the pseudo-perturbations search ($k = 13$)

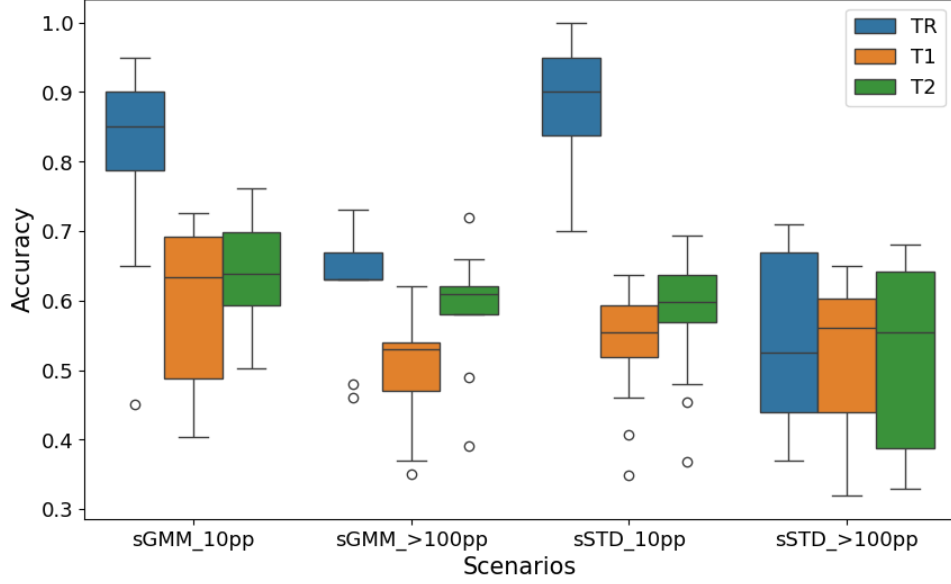


(b) Accuracy of optimal solutions.

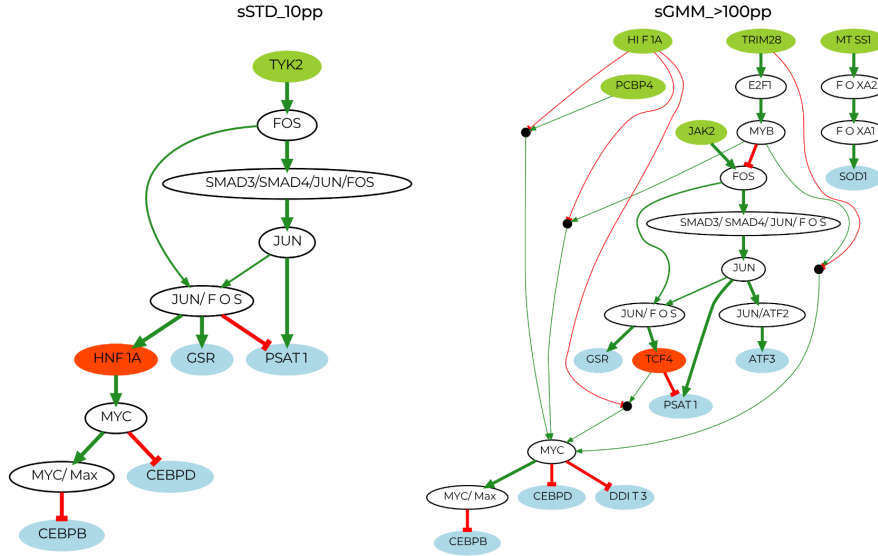
Supplementary Figure S4: Stable and exhaustive solutions. (a) Number of pseudo-perturbations found across time for 13 input-intermediate control genes (i.e. $k = 13$) in solid lines and for 10 input-intermediate genes (i.e. $k = 10$) in dashed lines. Computation was run for 20 days on a computer cluster comprising 4992 CPUs and 22 To of RAM. The pSTD results were computed on an older computer cluster for 20 days. For $k = 13$, 99 pseudo-perturbations were found for the sSTD solution and 95 for the sGMM one. (b) Accuracy scores computed for solutions closest to the optimum i.e. found at the 20 computation days cutoff. The results correspond to the performance of one of the optimal solutions in each scenario. The red dashed line marks $y = 0.5$ which corresponds to the performance of a random classifier. The blue dashed line marks $y = 0.6$ corresponding to the threshold applied to find the best results in the main study. **Training set (TR)**: cells selected for the pseudo-perturbations **Testing set 1 (T1)**: redundant cells **Testing set 2 (T2)**: unselected cells

5 Supplementary Figure 4. Comparing solutions close and far from the convergence point

Figure S5 compares the accuracies obtained for the different solution sets and scenarios. We sampled two sets of 20 solutions comprising ten pseudo-perturbations, one for sSTD and one for sGMM, from millions of possible solutions. The near-optimal solutions are the ones presented in the main study: 9 sGMM solutions and 16 sSTD ones. The solutions found further from the optimum (*sSTD_10pp* and *sGMM_10pp*) seem to perform better than their counterparts closer to the optimum (*sSTD_>100pp* and *sGMM_>100pp*). Particularly, the *_10pp* solutions achieve higher accuracies for their training sets. However, it is important to note that TR comprises only ten pseudo-perturbations for these solutions hence the high scores. Additionally, the *_10pp* solutions are only samples of the millions of others possible at this stage and may therefore not be representative of this larger solution space. In addition, using only ten pseudo-perturbations for the inference of BNs offers a limited representation of the diversity of gene expression patterns in our dataset. The BNs inferred are also smaller with very few I/O behaviours, showing little diversity in the proposed mechanisms. For these reasons, we focus on solutions closer to the optimum, which form a smaller and more stable set and are therefore easier to manage and analyze.



(a) Boxplot of accuracies for various scenarios



(b) Examples of mature TE BNFs

Supplementary Figure S5: Solutions closer to the optimum are more reliable. (a) A sample of 20 solutions among millions was used for *sSTD_10pp* and another for *sGMM_10pp*. *sGMM_>100pp* and *sSTD_>100pp* comprise the solutions presented in the main article i.e. 9 and 16 respectively. A lower number of pseudo-perturbations yields acceptable performances but the number of solutions possible is too big to exhaustively explore. (b) Example of mature TE BNFs. The BNF shown for *sGMM_10pp* is small and exhibits only one I/O behaviour whereas the one shown for *sGMM_>100pp* is bigger, has more logical gates and exhibits 8 I/O behaviours. **pp**: pseudo-perturbations **Training set**: cells selected for the pseudo-perturbations **Testing set 1**: redundant cells **Testing set 2**: unselected cells