Pruning neural network models for gene regulatory dynamics using data and domain knowledge

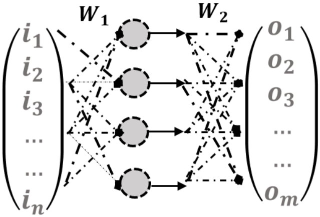
Intekhab Hossain†, Jonas Fischer†, Rebekka Burkholz*, John Quackenbush*

Harvard T.H. Chan School of Public Health
Max Planck Institute for Informatics
Helmholtz Center CISPA for Information Security

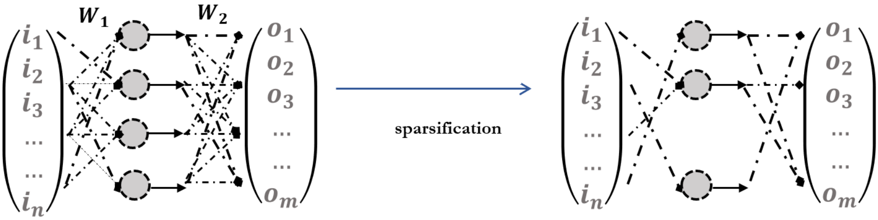
†, * equal contribution

Friday, December 13, 2024

Sparsifying neural networks

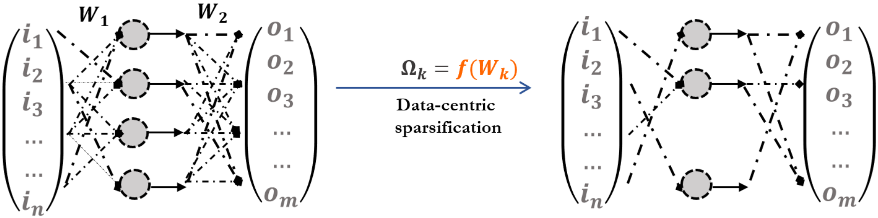


Sparsifying neural networks



- ☆ Identify the most important circuits/activation patterns
- ☆ Sparsified model has much fewer parameters; easier to interpret

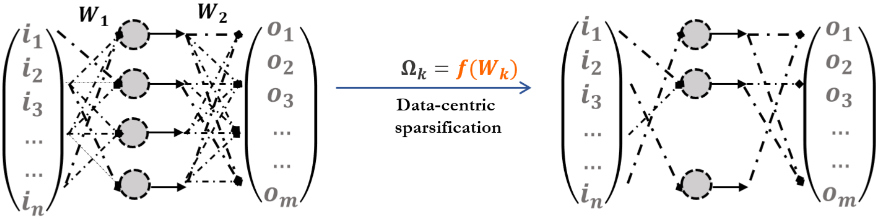
Current strategies may not yield meaningful interpretations



$\Omega_k = f(W_k)$

Data-centric sparsification

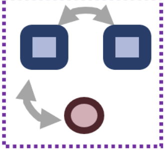
Towards meaningful interpretations



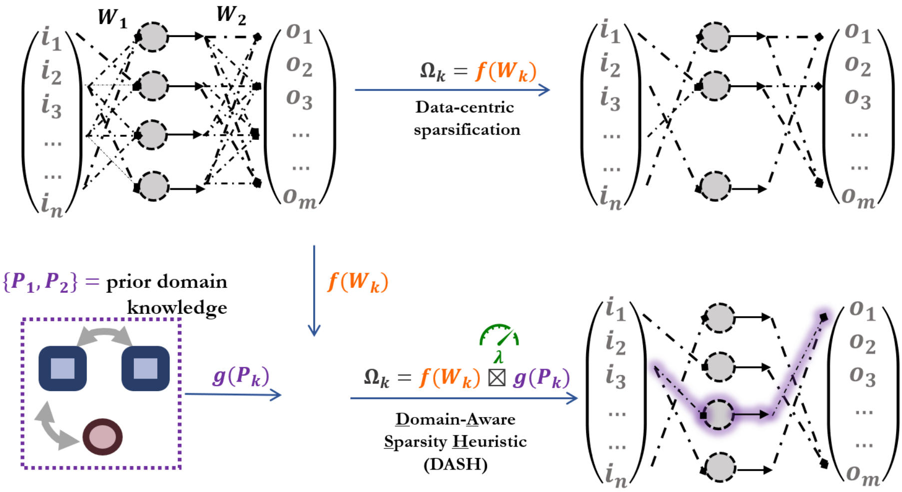
$\Omega_k = f(W_k)$

Data-centric sparsification

$\{P_1, P_2\}$ = prior domain knowledge



DASH: sparsifies towards a meaningful structure



$\Omega_k = f(W_k)$

Data-centric sparsification

$\{P_1, P_2\}$ = prior domain knowledge

$g(P_k)$

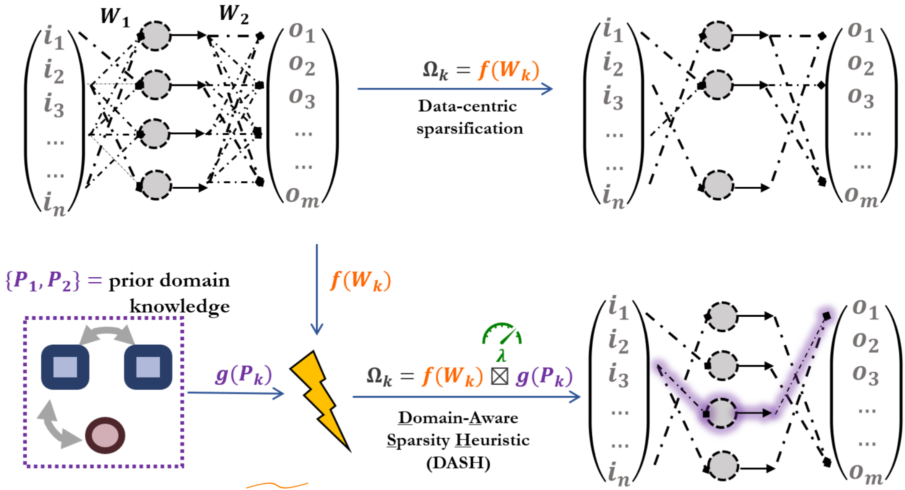
$f(W_k)$

λ

$\Omega_k = f(W_k) \boxtimes g(P_k)$

Domain-Aware Sparsity Heuristic (DASH)

DASH: simple calculations



$\Omega_k = f(W_k)$

Data-centric sparsification

$f(W_k)$

$\{P_1, P_2\}$ = prior domain knowledge

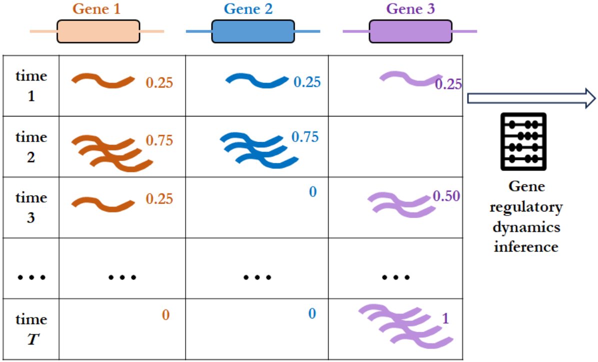
$g(P_k)$

$\Omega_k = f(W_k) \boxtimes g(P_k)$

Domain-Aware Sparsity Heuristic (DASH)

- ☆ $$\Omega_1^{(t)} := (1-\lambda_1)\left|\widetilde{W_1^{(t)}}\right| + \lambda_1\left|\mathtt{PInv}_L\left(\Omega_1^{(t-1)\top}\right)\cdot P_1\right|$$
- ☆ $$\Omega_2^{(t)} := (1-\lambda_2)\left|\widetilde{W_2^{(t)}}\right| + \lambda_2\left|P_2\cdot\mathtt{PInv}_R\left(\Omega_1^{(t)}\right)\right|$$

Gene regulatory dynamics **inference**



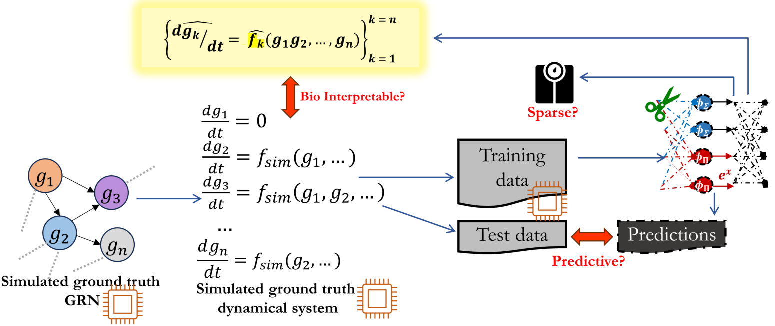
	Gene 1	Gene 2	Gene 3
time 1	0.25	0.25	0.25
time 2	0.75	0.75	
time 3	0.25	0	0.50
...	...	...	...
time T	0	0	1

$$\widehat{dg_1}/_{dt} = \boldsymbol{f}_1(g_1, g_2, g_3)\ ??$$

$$\widehat{dg_2}/_{dt} = \boldsymbol{f}_2(g_1, g_2, g_3)\ ??$$

$$\widehat{dg_3}/_{dt} = \boldsymbol{f}_3(g_1, g_2, g_3)\ ??$$

In silico benchmarking pipeline for sparsification



$$\left\{ \widehat{dg_k/dt} = \widehat{f_k}(g_1 g_2, \dots, g_n) \right\}_{k=1}^{k=n}$$

Bio Interpretable?

$$\frac{dg_1}{dt} = 0$$

$$\frac{dg_2}{dt} = f_{sim}(g_1, \dots)$$

$$\frac{dg_3}{dt} = f_{sim}(g_1, g_2, \dots)$$

…

$$\frac{dg_n}{dt} = f_{sim}(g_2, \dots)$$

Simulated ground truth dynamical system

Simulated ground truth GRN

Training data

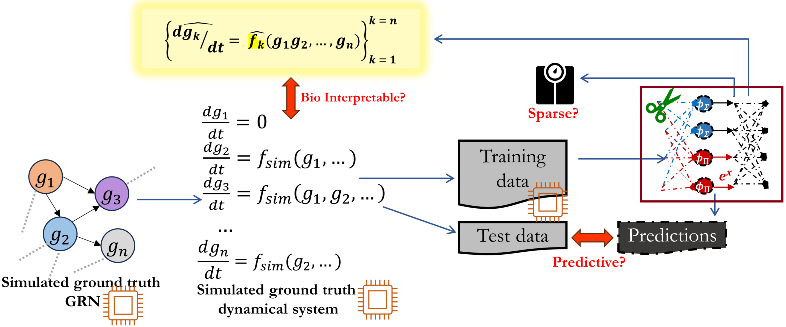
Test data

Predictions

Sparse?

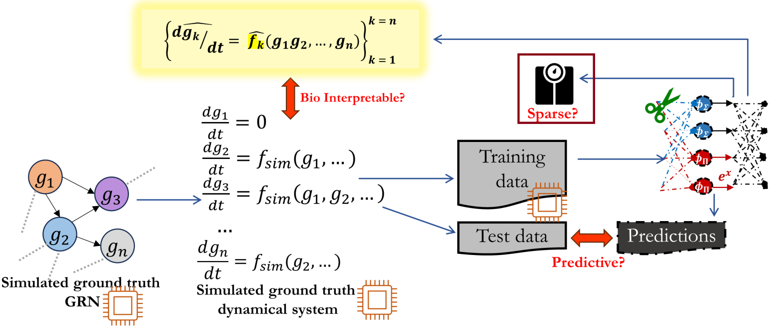
Predictive?

In silico benchmarking pipeline for sparsification



- ✫ Base Model = PHOENIX (NeuralODE), Hossain *et al*, Genome Biology, 2024

In silico benchmarking pipeline for sparsification



$$\left\{ \widehat{dg_k/dt} = \widehat{f_k}(g_1 g_2, \ldots, g_n) \right\}_{k=1}^{k=n}$$

$$\frac{dg_1}{dt} = 0$$

$$\frac{dg_2}{dt} = f_{sim}(g_1, \ldots)$$

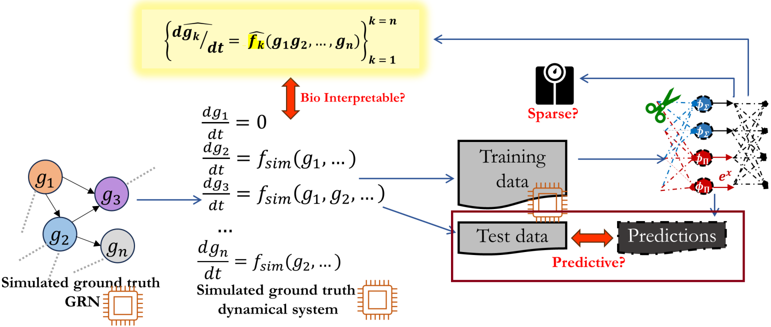
$$\frac{dg_3}{dt} = f_{sim}(g_1, g_2, \ldots)$$

$$\ldots$$

$$\frac{dg_n}{dt} = f_{sim}(g_2, \ldots)$$

- ☆ Base Model = PHOENIX (NeuralODE), Hossain *et al*, Genome Biology, 2024

In silico benchmarking pipeline for sparsification



$$\left\{\widehat{dg_k/dt} = \widehat{f_k}(g_1 g_2, \dots, g_n)\right\}_{k=1}^{k=n}$$

$$\frac{dg_1}{dt} = 0$$

$$\frac{dg_2}{dt} = f_{sim}(g_1, \dots)$$

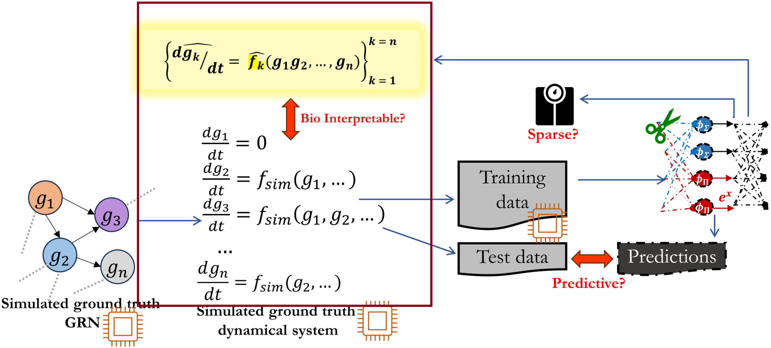
$$\frac{dg_3}{dt} = f_{sim}(g_1, g_2, \dots)$$

$$\dots$$

$$\frac{dg_n}{dt} = f_{sim}(g_2, \dots)$$

☆ Base Model = PHOENIX (NeuralODE), Hossain *et al*, Genome Biology, 2024

In silico benchmarking pipeline for sparsification



$$\left\{\widehat{{dg_k}/{dt}} = \widehat{f_k}(g_1 g_2, \dots, g_n)\right\}_{k=1}^{k=n}$$

Bio Interpretable?

$$\frac{dg_1}{dt} = 0$$
$$\frac{dg_2}{dt} = f_{sim}(g_1, \dots)$$
$$\frac{dg_3}{dt} = f_{sim}(g_1, g_2, \dots)$$
$$\dots$$
$$\frac{dg_n}{dt} = f_{sim}(g_2, \dots)$$

Simulated ground truth
GRN

Simulated ground truth
dynamical system

Training data

Test data

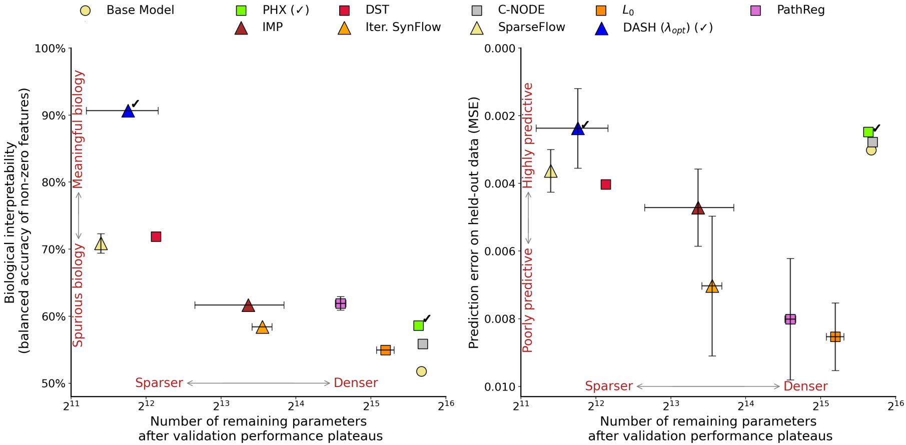
Predictions

Predictive?

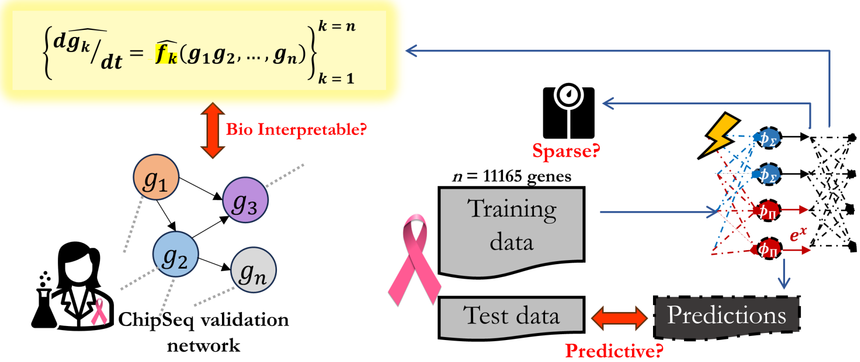
Sparse?

- ☆ Base Model = PHOENIX (NeuralODE), Hossain *et al*, Genome Biology, 2024

DASH is biologically interpretable



Real data application: sparse dynamics for breast cancer



$$\left\{ {\widehat{dg_k}}/{dt} = \widehat{f}_k(g_1 g_2, \dots, g_n) \right\}_{k=1}^{k=n}$$

Bio Interpretable?

ChipSeq validation network

Sparse?

n = 11165 genes

Training data

Test data

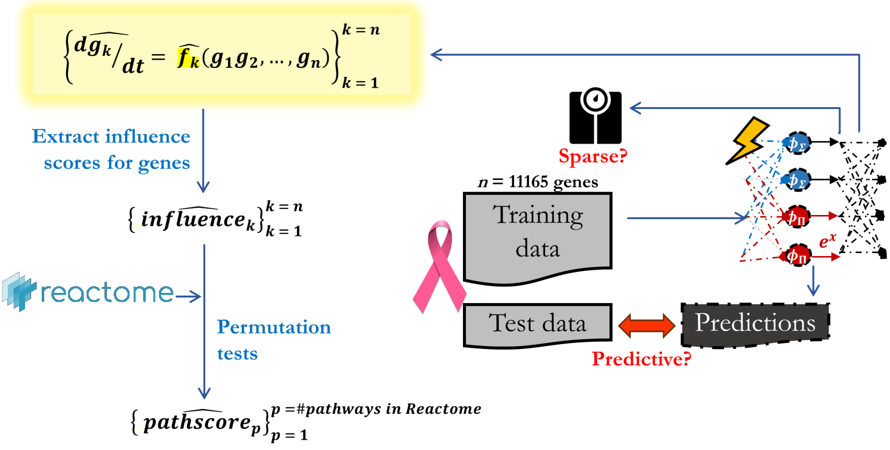
Predictions

Predictive?

Real data application: sparse dynamics for breast cancer

Strategy	Sparsity	Bio. Interp.	Test MSE (10^{-5})
None/Baseline	0.03%	49.99%	7.78
PathReg	14.09%	50.24%	7.92
PINN	0.11%	49.99%	7.82
DST	67.02%	50.42%	7.78
Iter. SynFlow	91.93%	49.37%	7.78
SparseFlow	**95.70%**	49.70%	**7.76**
DASH	92.71%	**95.69%**	**7.76**

Sparse models to discover breast cancer pathways



$$\left\{ \widehat{dg_k/dt} = \widehat{f_k}(g_1 g_2, \ldots, g_n) \right\}_{k=1}^{k=n}$$

Extract influence scores for genes

$$\left\{ \widehat{influence_k} \right\}_{k=1}^{k=n}$$

reactome

Permutation tests

$$\left\{ \widehat{pathscore_p} \right\}_{p=1}^{p=\#pathways\ in\ Reactome}$$

Sparse?

$n = 11165$ genes

Training data

Test data

Predictions

Predictive?

Sparse models to discover breast cancer pathways