Technische Hochschule Nürnberg

Fusing Mathematical Modelling, Biomedical Data and Learning Algorithms for Computational Oncology

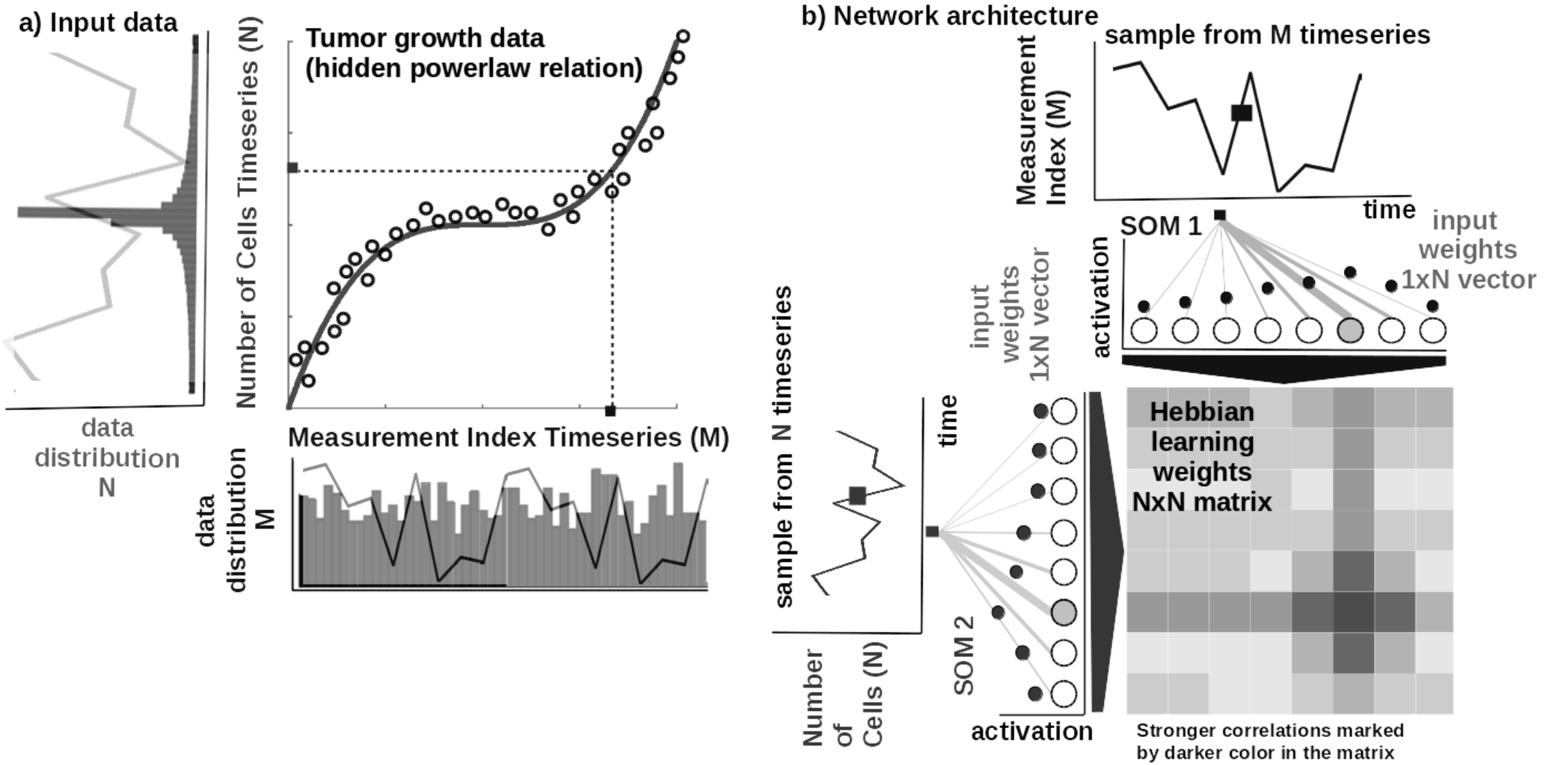
Computational Modelling of Complex Biological Systems
Workshop, 23.06.2023, University of Surrey, UK

Cristian Axenie, Prof. Dr. Eng. Sc. M.Sc.

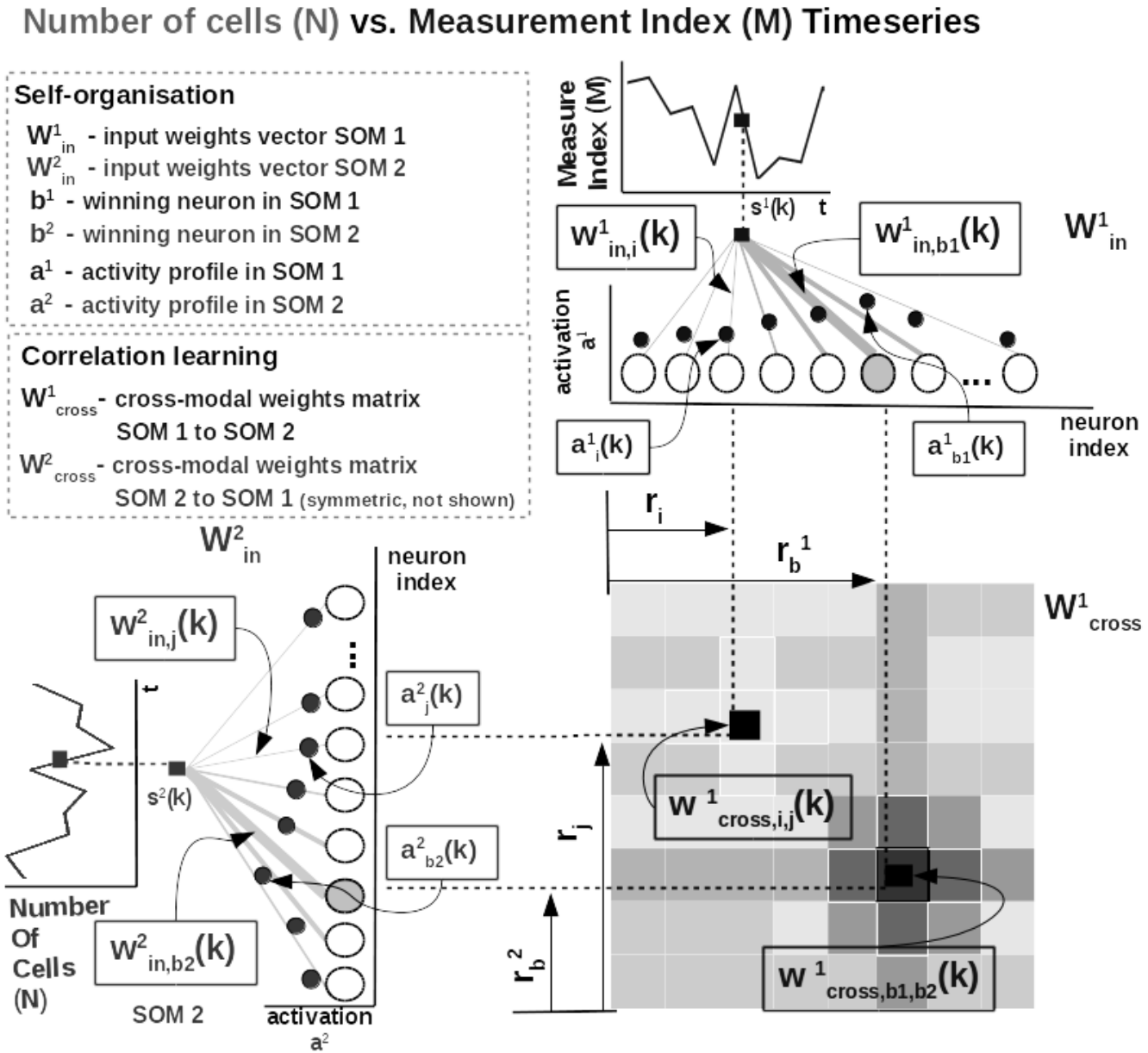
A Framework for Mathematical and Computational Oncology

Interacting Computational Maps

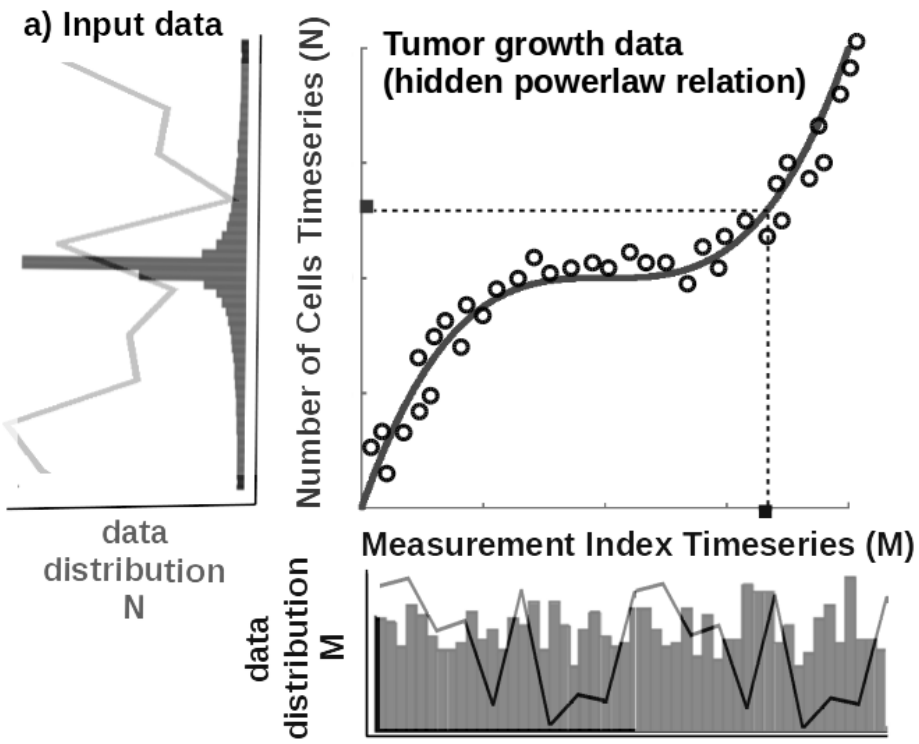
Core model

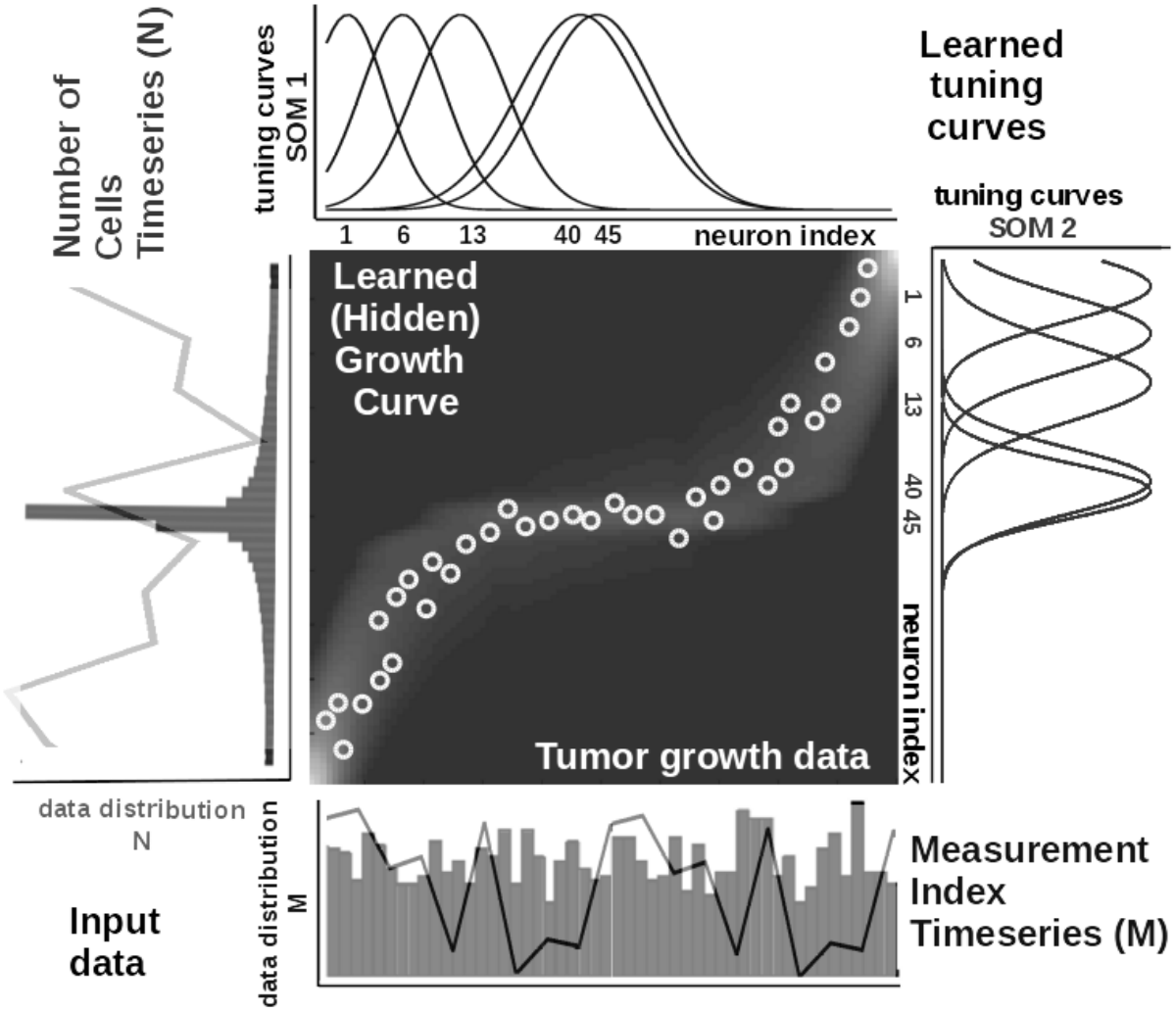

a) Input data
data distribution N
Number of Cells Timeseries (N)
Tumor growth data
(hidden powerlaw relation)
Measurement Index Timeseries (M)
data distribution M
b) Network architecture
sample from M timeseries
Measurement Index (M)
time
SOM 1
input weights 1xN vector
activation
sample from N timeseries
time
Number of Cells (N)
SOM 2
activation
Hebbian learning weights NxN matrix
Stronger correlations marked by darker color in the matrix

Core model internals

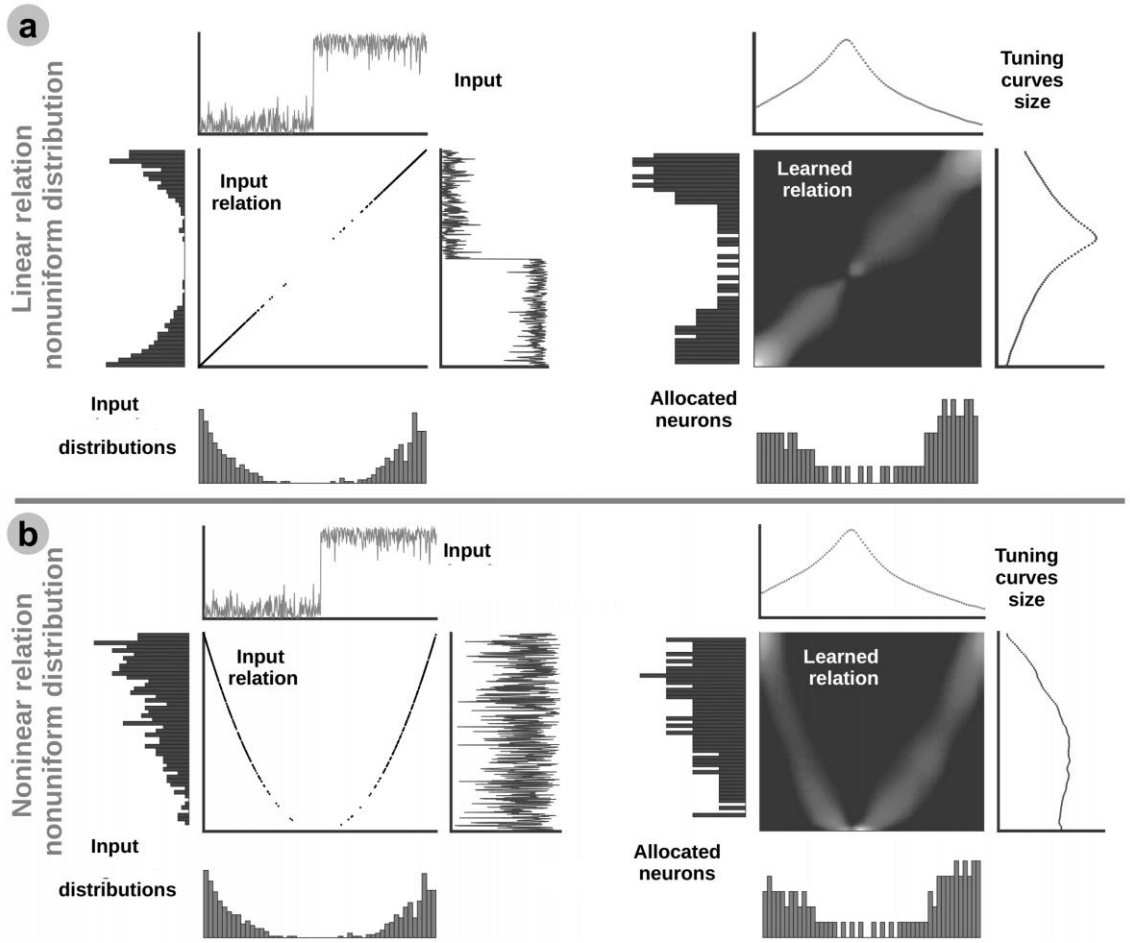


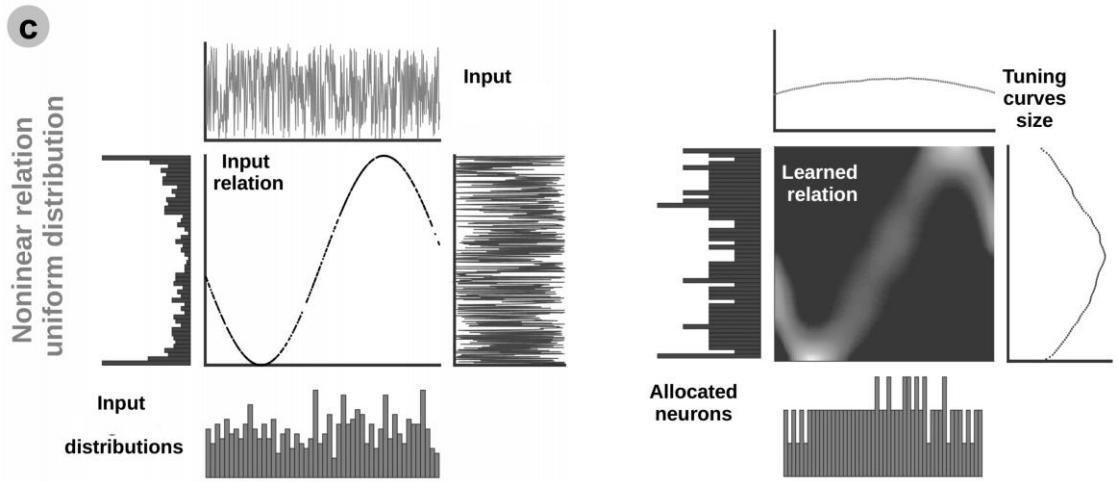
Learning capabilities I



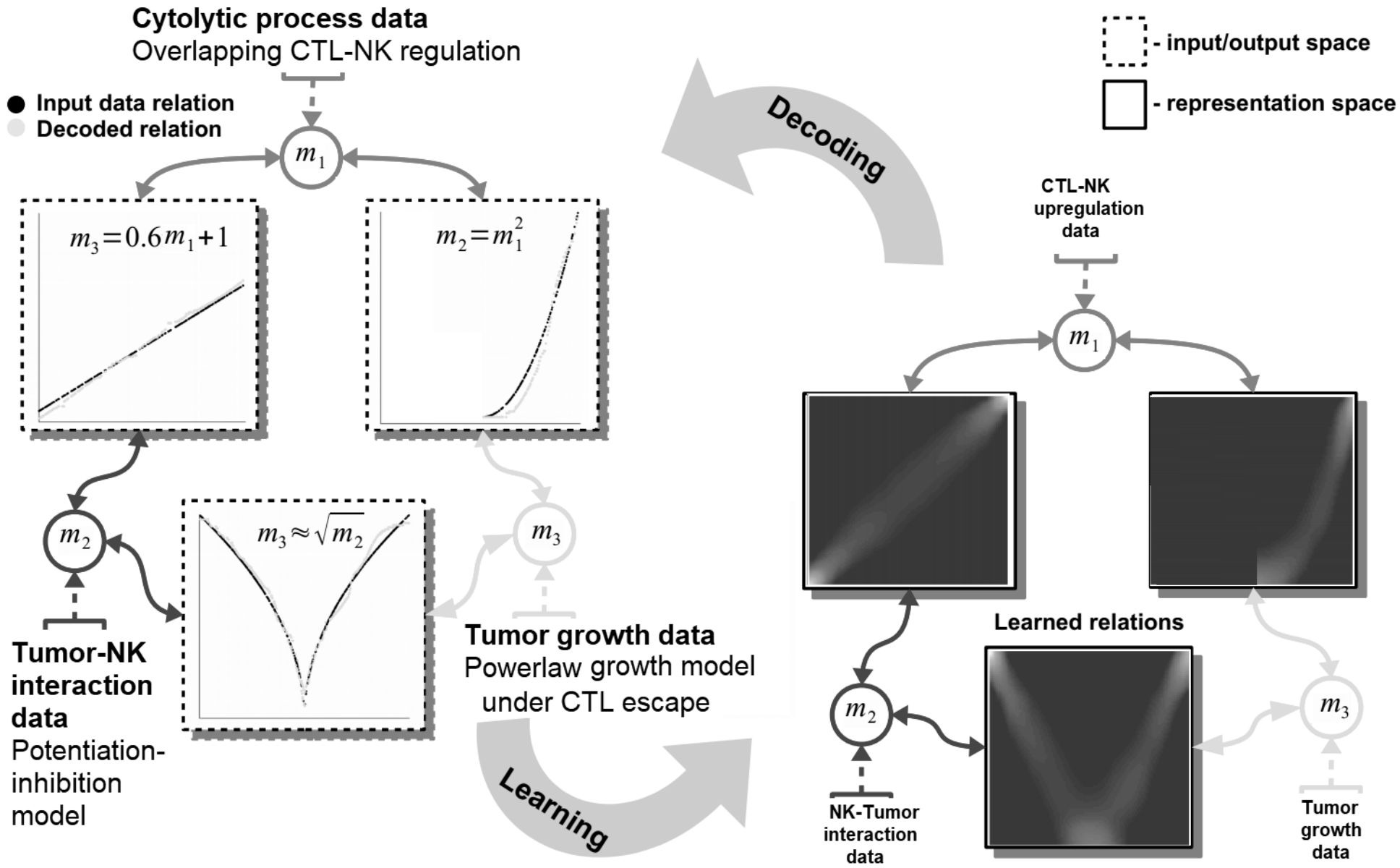


Learning capabilities II





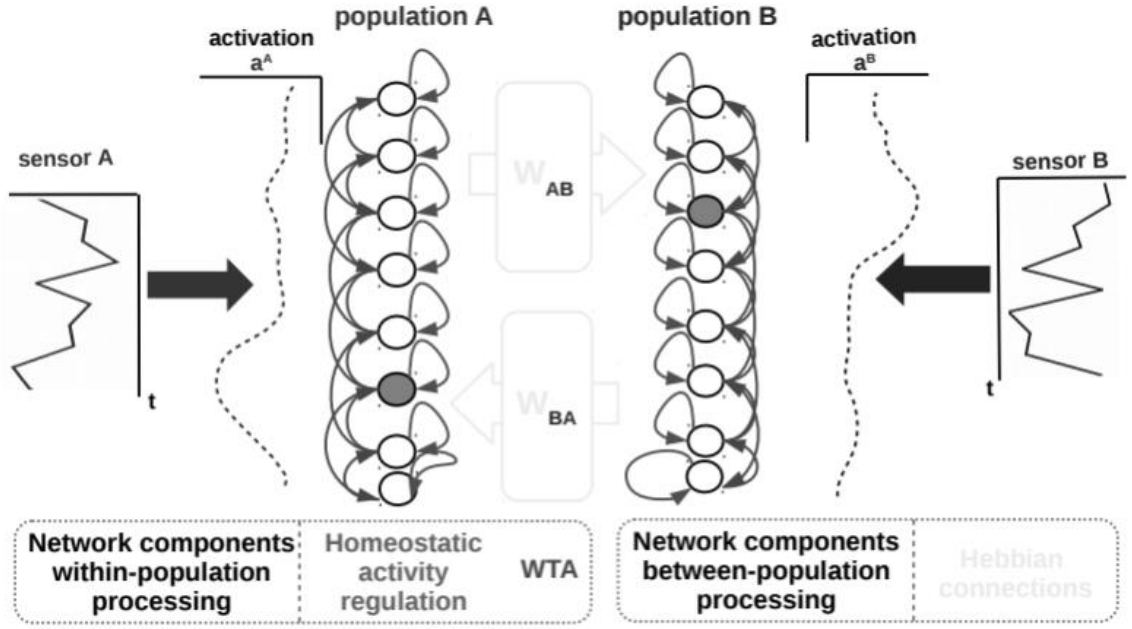
Extensibility

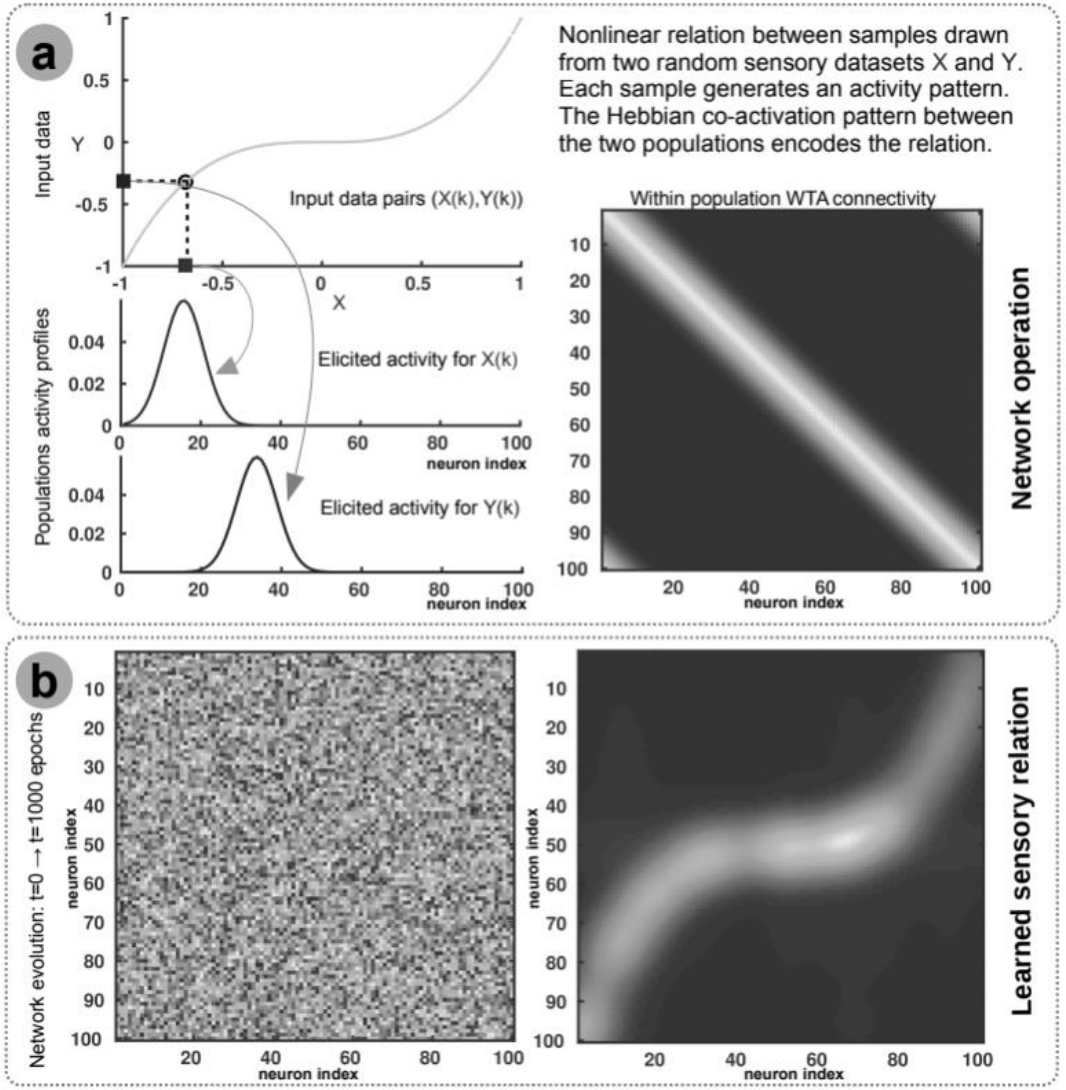


Comparable systems I

Cook et al.

Cook M., Jug F., Krautz C., Steger A. (2010). “Unsupervised Learning of Relations,” in International Conference on Artificial Neural Networks (Springer;), 164–173. 10.1007/978-3-642-15819-3_21

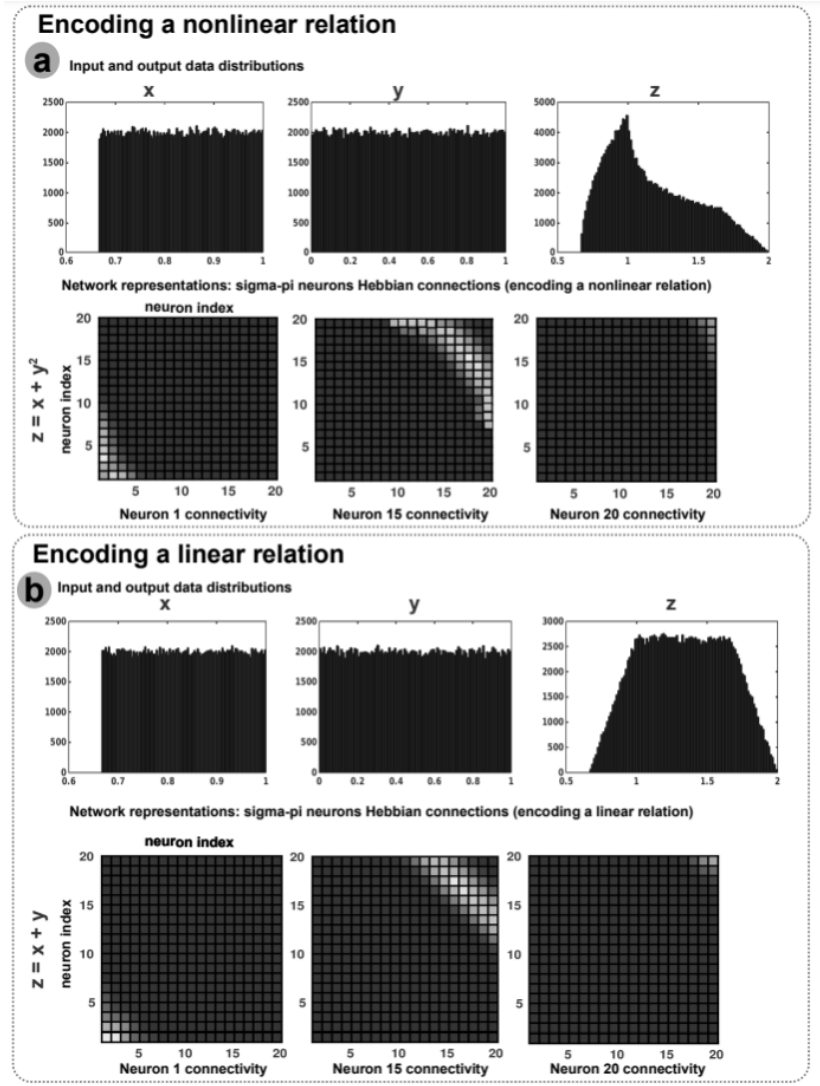


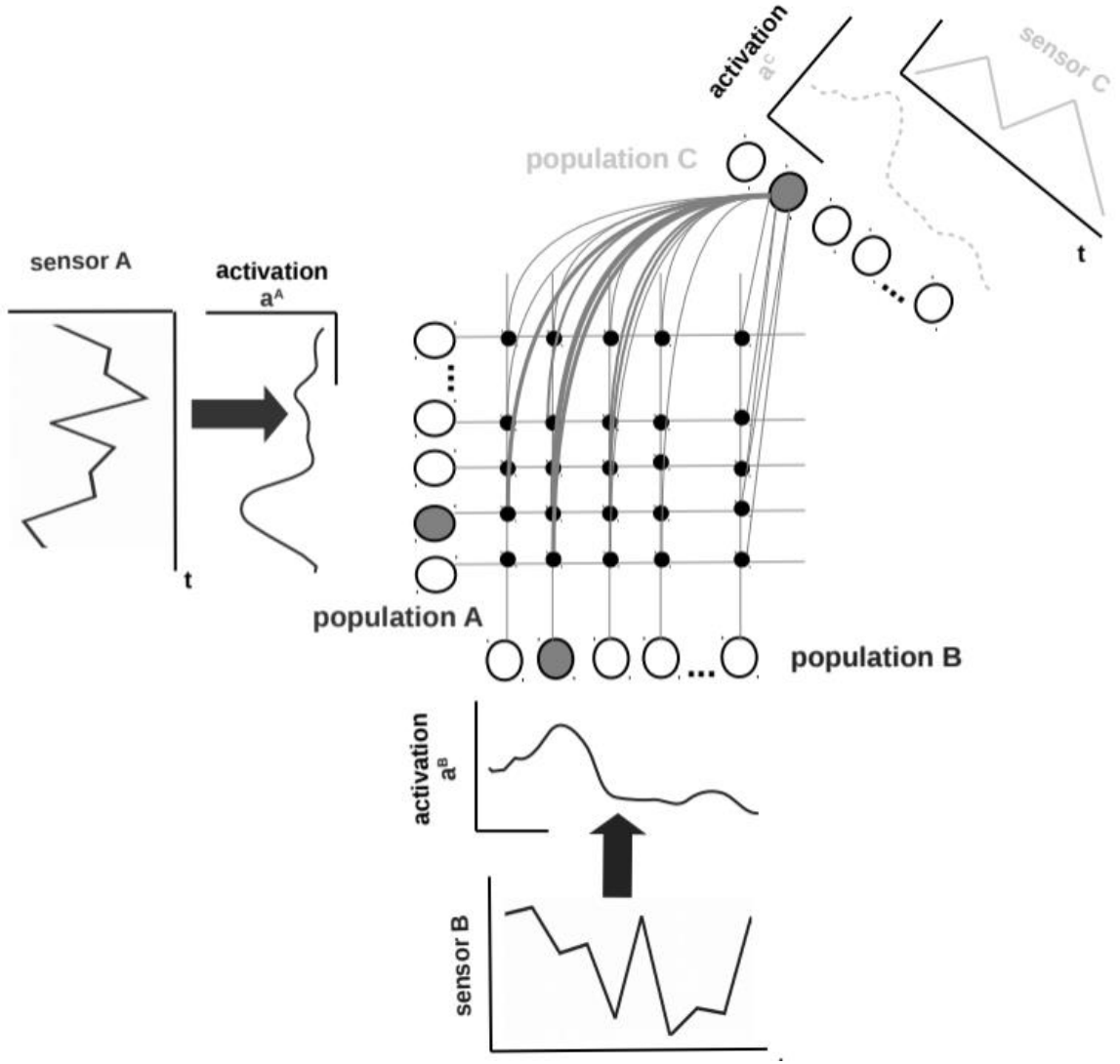


Comparable systems II

Weber and Wermter

Weber C., Wermter S. (2007). A Self-Organizing Map of Sigma-Pi Units. Neurocomputing 70, 2552–2560. 10.1016/j.neucom.2006.05.014

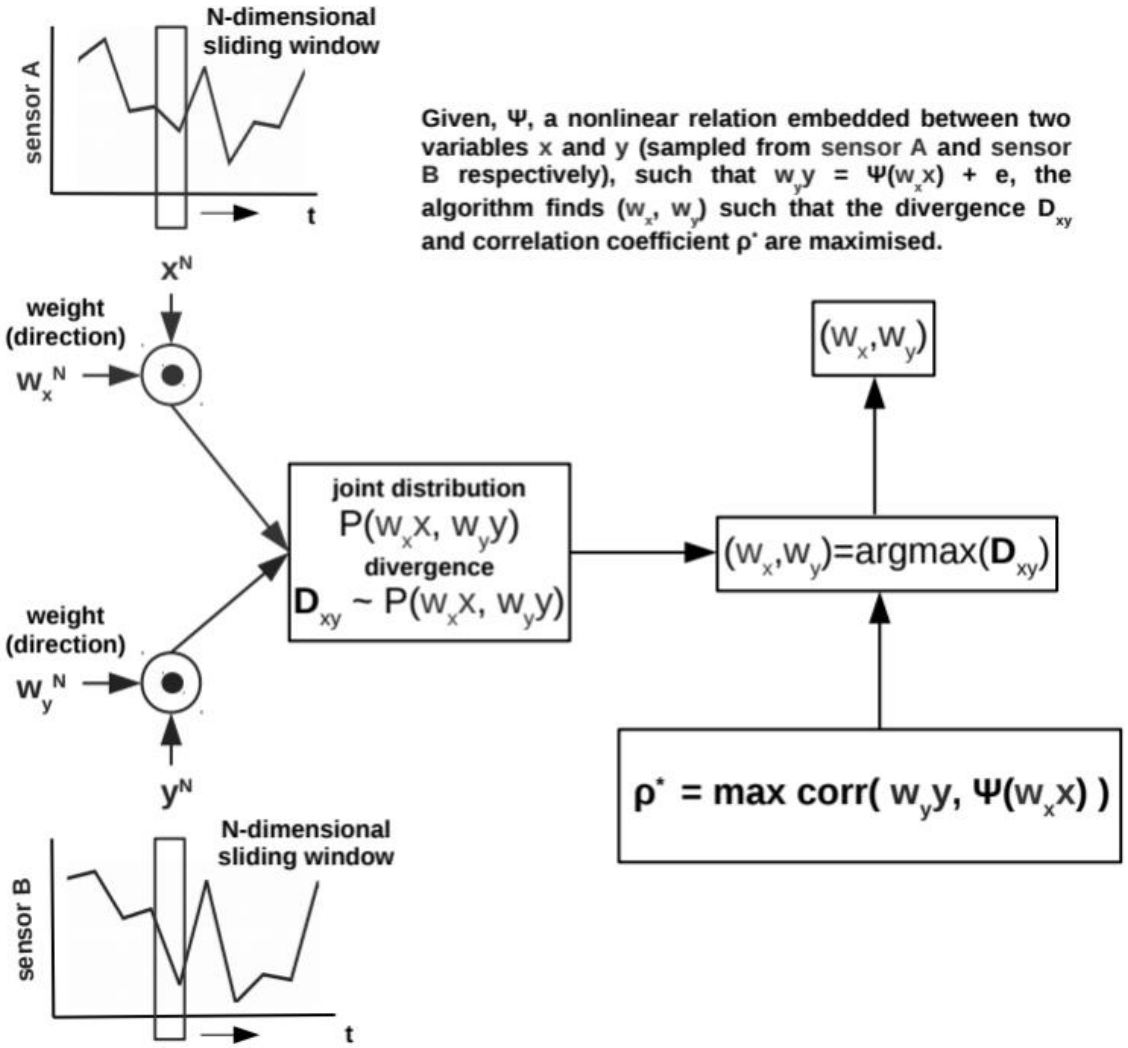


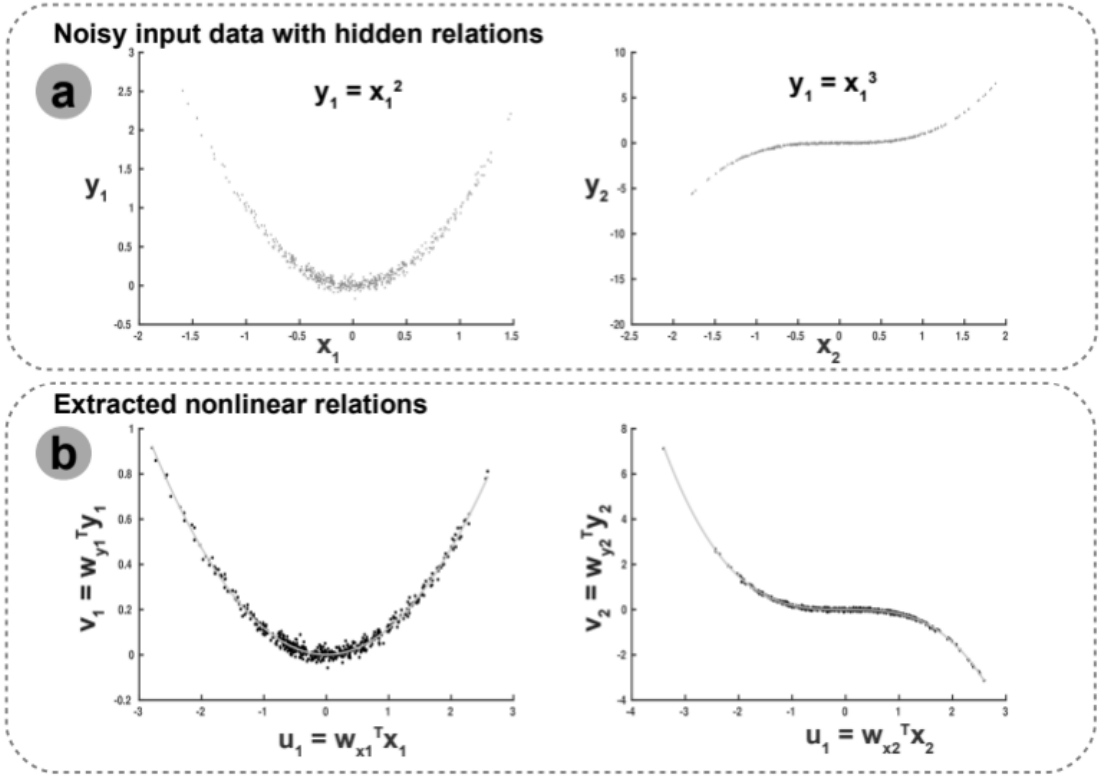


Comparable systems III

Mandal and Cichocki

Mandal A., Cichocki A. (2013). Non-linear Canonical Correlation Analysis Using Alpha-Beta Divergence. *Entropy* 15, 2788–2804. 10.3390/e15072788

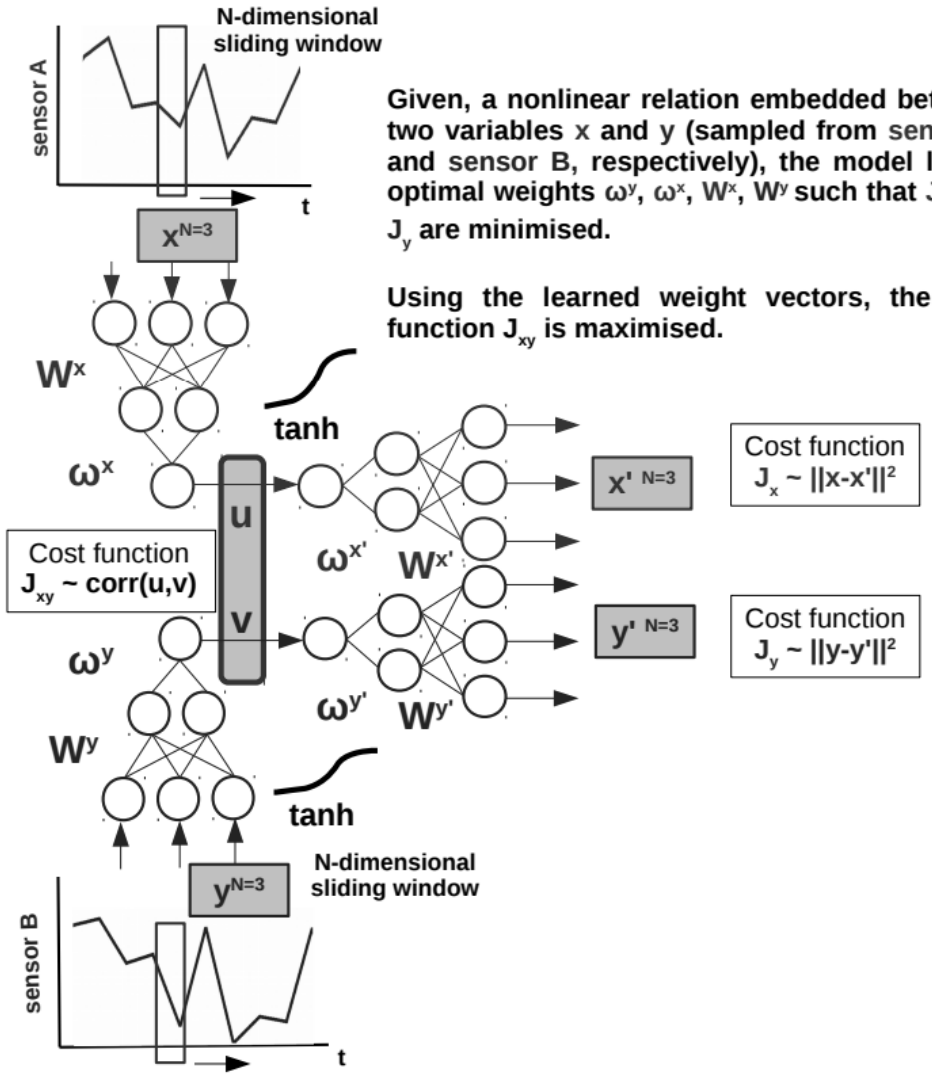




Comparable systems IV

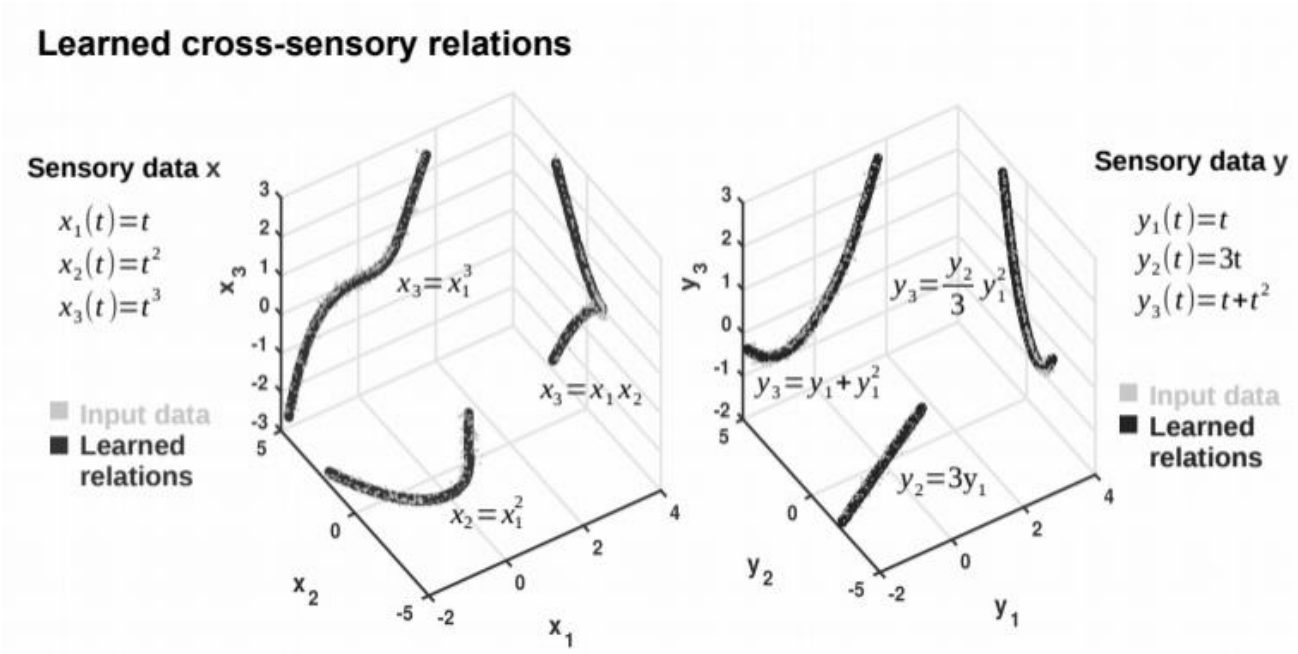
Champion et al.

Champion K., Lusch B., Kutz J. N., Brunton S. L. (2019). Data-driven Discovery of Coordinates and Governing Equations. *Proc. Natl. Acad. Sci. USA* 116, 22445–22451. 10.1073/pnas.1906995116



Given, a nonlinear relation embedded between two variables x and y (sampled from sensor A and sensor B, respectively), the model learns optimal weights ω^y, ω^x, W^x, W^y such that J_x and J_y are minimised.

Using the learned weight vectors, the cost function J_{xy} is maximised.



Experiments, evaluation, and results

TABLE 1 | Description of the datasets used in the experiments.

	Experimental dataset setup			
Dataset	**Cancer type**	**Data type**	**Data points**	**Data freq.**
1	Breast[1] (MDA-MB-231 cell line)	Fluorescence imaging	7	2×/week
2	Breast[2] (MDA-MB-435 cell line)	Digital caliper	14	2×/week
3	Lung[3]	Caliper	10	7×/week
4	Leukemia[4]	Microscopy	23	7×/week
5	Breast[5] (MCF-7 cell line)	Microscopic imaging	8	1×/week
6	Breast[6] (LM2-4LUC + cell line)	Digital caliper	10	3×/week
7	Breast[7] (stage 2/3 cancers)	Functional magnetic resonance imaging	5	1×/week
8	Breast[8] (ductal carcinoma *in situ*)	Histopathology	5	1×/week

[1]*Dataset from the study by Rodallec et al. (2019)*
[2]*Dataset from the study by Volk et al. (2011)*
[3]*Dataset from the study by Benzekry et al. (2019)*
[4]*Dataset from the study by Simpson-Herren and Lloyd (1970)*
[5]*Dataset from the study by Tan et al. (2015)*
[6]*Dataset from the study by Mastri et al. (2019)*
[7]*Dataset from the study by Yee et al. (2020)*
[8]*Dataset from the study by Edgerton et al. (2011)*

TABLE 2 | Evaluation metrics for data-driven relation learning systems. We consider N—number of measurements, σ—standard deviation of data, p—number of parameters of the model.

Metric	**Equation**
SSE	$\Sigma_{i=1}^{N}\left(\frac{y^i-y_m^i}{\frac{\sigma_i}{\sigma}}\right)$
RMSE	$\sqrt{\frac{SSE}{N-p}}$
sMAPE	$\frac{1}{N}\Sigma_{i=1}^{N}\left(2\frac{\lvert y^i-y_m^i\rvert}{(\lvert y^i\rvert+\lvert y_m^i\rvert)}\right)$
AIC	$N\,ln\left(\frac{SSE}{N}\right)+2p$
BIC	$N\,ln\left(\frac{SSE}{N}\right)+ln\,(N)p$

Learning Growth Patterns of Preinvasive Breast Cancer

We fed the systems with DCIS data from Edgerton et al. (2011), namely, time series of nutrient diffusion penetration length within the breast tissue (L), ratio of cell apoptosis to proliferation rates (A), and radius of the breast tumor (R).

The study by Edgerton et al. (2011) postulated that the value of R depends on A and L following a "master equation" whose predictions are consistent with nearly 80% of in situ tumors identified by mammographic screenings.

$$A = 3\frac{L}{R}\left(\frac{1}{tanh\left(\frac{R}{L}\right)} - \frac{L}{R}\right)$$

TABLE 3 | Evaluation of the data-driven relation learning systems.

	Evaluation metrics		
Dataset/system	**SSE**	**RMSE**	**sMAPE**
Breast (DCIS), Edgerton et al. (2011)			
Cook et al.	56.321	0.4867	0.5901
Weber et al.	59.879	0.5099	0.6512
Mandal et al.	62.346	0.5617	0.6800
Champion et al.	58.645	0.4721	0.6054
Our system	54.216	0.4656	0.5734

Learning Unperturbed Tumor Growth Curves Within and Between Cancer Types

In the second task, we evaluated the systems on learning unperturbed (i.e., growth without treatment) tumor growth curves.

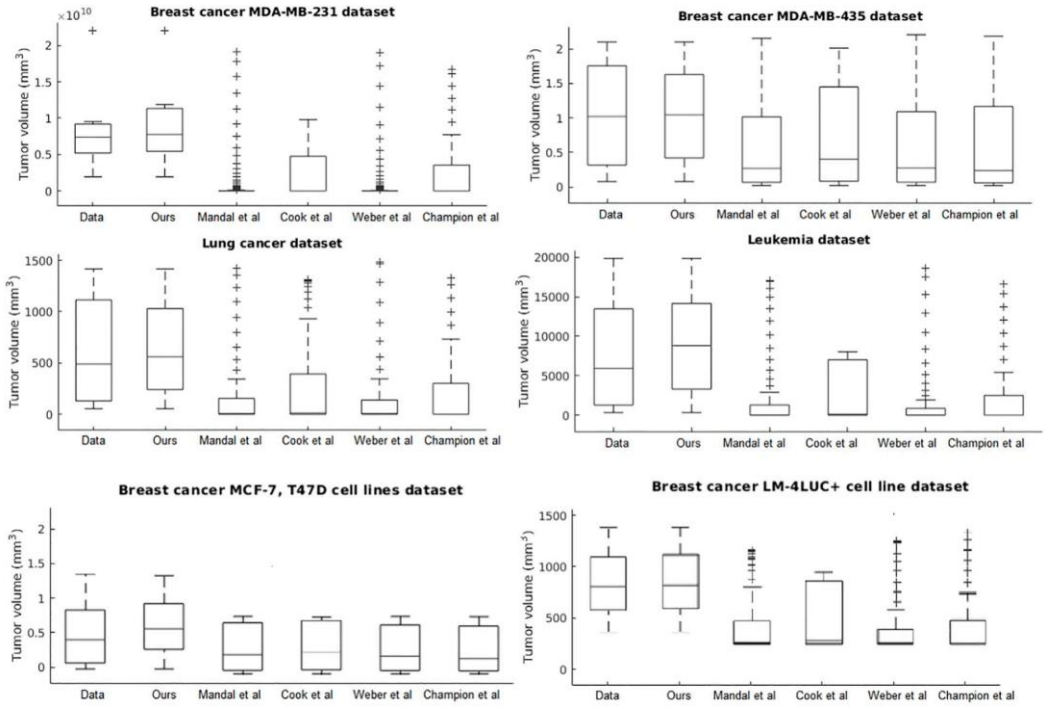


FIGURE 5 | Evaluation of the data-driven relation learning systems on tumor growth: summary statistics.

TABLE 4 | Evaluation of the data-driven relation learning systems on tumor growth curve extraction.

	Evaluation metrics (smaller value is better)					
Dataset/system	**SSE**	**RMSE**	**sMAPE**	**AIC**	**BIC**	**Rank3**
Breast4 cancer						
Cook et al.	7,009.6	37.4423	1.7088	52.3639	52.2557	2
Weber et al.	8,004.9	44.7350	1.7088	55.2933	55.1310	5
Mandal et al.	7,971.8	39.9294	1.7088	53.2643	53.1561	4
Champion et al.	6,639.1	40.7403	1.4855	53.9837	53.8215	3
Our system	119.3	4.1285	0.0768	19.8508	19.8508	1
Breast5 cancer						
Cook et al.	0.2936	0.1713	0.1437	−40.5269	−39.5571	4
Weber et al.	0.2315	0.1604	0.1437	−41.3780	−39.9233	2
Mandal et al.	0.3175	0.1782	0.1437	−39.5853	−38.6155	5
Champion et al.	0.2699	0.1732	0.1512	−39.5351	−38.0804	3
Our system	0.0977	0.0902	0.0763	−57.7261	−57.7261	1
Breast6 cancer						
Cook et al.	3.0007	0.7071	1.0606	50.1322	51.2887	2
Weber et al.	3.2942	0.8116	1.6626	56.4133	55.1915	5
Mandal et al.	3.1908	0.7292	1.3506	53.2643	52.5421	4
Champion et al.	3.4772	0.8339	1.1288	53.9837	53.7775	3
Our system	0.7668	0.3096	0.2615	19.3208	19.1298	1
Breast7 cancer						
Cook et al.	45.6031	2.3875	1.2216	−40.0084	−39.9975	4
Weber et al.	56.0738	2.8302	1.8346	−41.2345	−39.1234	2
Mandal et al.	53.2428	2.5797	1.4816	−39.5853	−37.1260	5
Champion et al.	54.7189	2.7958	1.5086	−39.1234	−38.0664	3
Our system	0.2008	0.1417	0.0364	−57.1221	−57.6112	1
Lung cancer						
Cook et al.	44.5261	2.2243	1.5684	19.3800	20.1758	2
Weber et al.	54.1147	2.6008	1.5684	23.5253	24.7190	5
Mandal et al.	53.2475	2.4324	1.5684	21.3476	22.1434	4
Champion et al.	50.6671	2.5166	1.5361	22.8012	23.9949	3
Our system	3.6903	0.5792	0.2121	−12.0140	−12.0140	1
Leukemia						
Cook et al.	223.7271	3.2640	1.6368	56.3235	58.5944	2
Weber et al.	273.6770	3.6992	1.6368	62.9585	66.3649	5
Mandal et al.	259.9277	3.5182	1.6368	59.7729	62.0439	4
Champion et al.	248.5784	3.5255	1.6001	60.7461	64.1526	3
Our system	35.2541	1.2381	0.3232	9.8230	9.8230	1

Notes: 3—Calculated as best in 3/5 metrics; 4—MDA-MB-231 cell line; 5—MDA-MB-435 cell line; 6—MCF-7, T47D cell line; 7—LM2-4LUC + cell line.

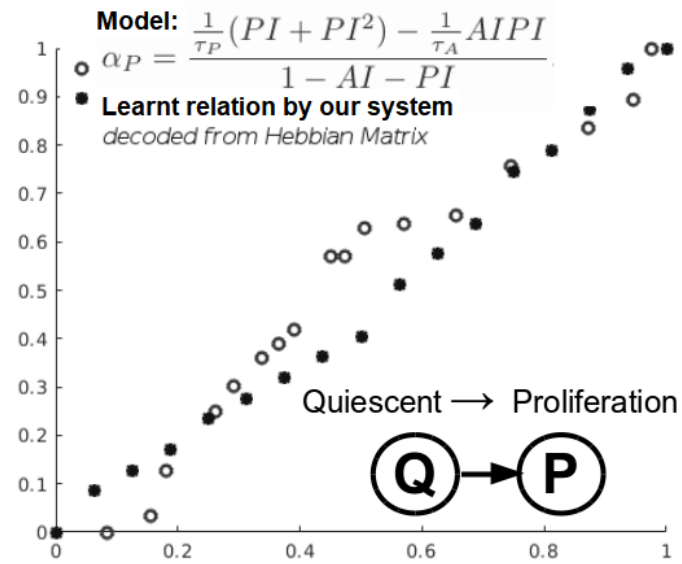
Extracting Tumor Phenotypic Stage Transitions

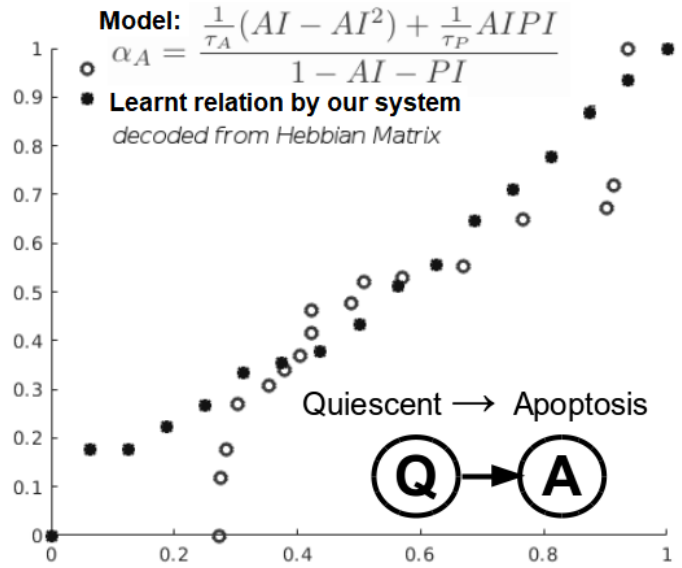
For this experiment, we considered the study of 17 breast cancer patients in the study by Edgerton et al. (2011). In this instantiation, we focus on a simplified three-state phenotypic model (i.e., containing P, Q, A states). Each of the data-driven relation learning systems is fed with time series of raw immunohistochemistry and morphometric data.

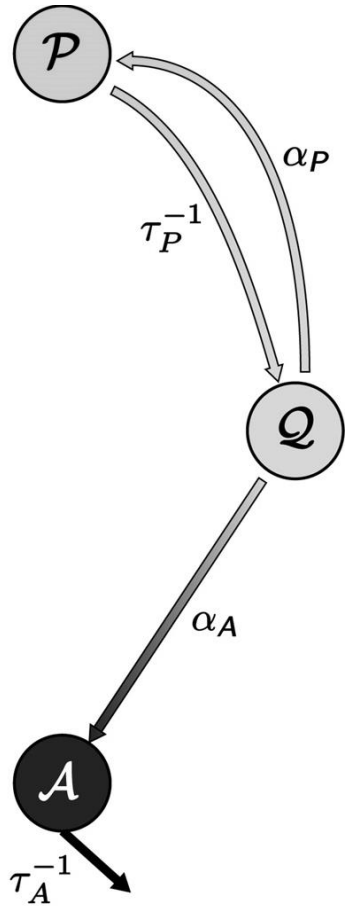
$$\alpha_P = \frac{\frac{1}{\tau_P}(PI + PI^2) - \frac{1}{\tau_A}AIPI}{1 - AI - PI}$$

$$\alpha_A = \frac{\frac{1}{\tau_A}(AI - AI^2) + \frac{1}{\tau_P}AIPI}{1 - AI - PI}$$

where, τ_P is the cells cycle time, τ_A cells apoptosis time , PI proliferation index and AI apoptosis index.







Extracting Tumor Phenotypic Stage Transitions

For this experiment, we considered the study of 17 breast cancer patients in the study by Edgerton et al. (2011). In this instantiation, we focus on a simplified three-state phenotypic model (i.e., containing P, Q, A states). Each of the data-driven relation learning systems is fed with time series of raw immunohistochemistry and morphometric data.

$$\alpha_P = \frac{\frac{1}{\tau_P}(PI + PI^2) - \frac{1}{\tau_A} AIPI}{1 - AI - PI}$$

$$\alpha_A = \frac{\frac{1}{\tau_A}(AI - AI^2) + \frac{1}{\tau_P} AIPI}{1 - AI - PI}$$

where, τ_P is the cells cycle time, τ_A cells apoptosis time , PI proliferation index and AI apoptosis index.

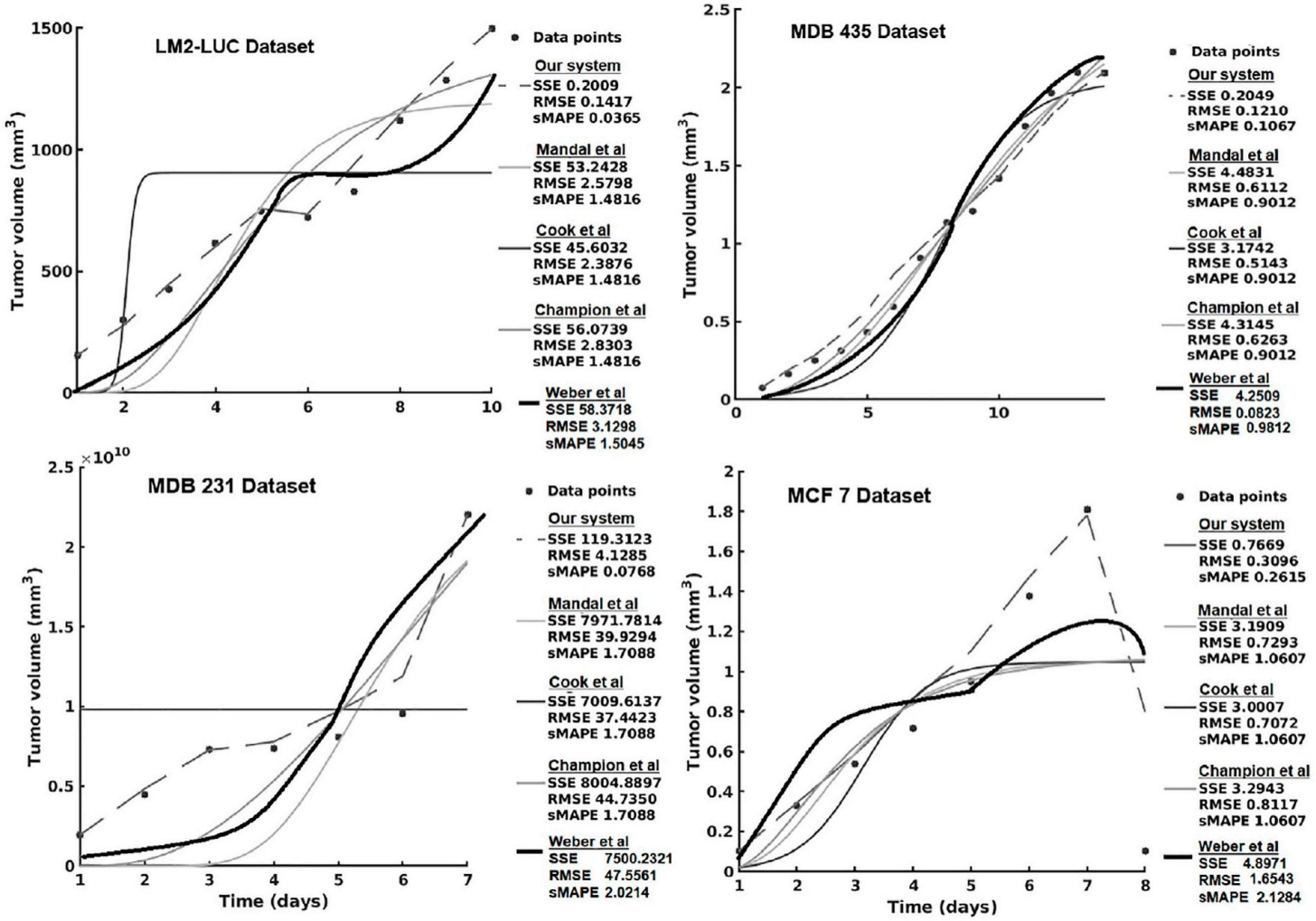
TABLE 5 | Evaluation of the data-driven relation learning systems for extracting phenotypic transitions relations.

	Evaluation metrics		
State transition/system	**SSE**	**RMSE**	**sMAPE**
Quiescent(Q) to proliferation(P) transition relation			
Cook et al.	0.820	0.240	0.190
Weber et al.	0.865	0.294	0.196
Mandal et al.	0.904	0.320	0.214
Champion et al.	0.845	0.274	0.189
Our system	0.750	0.210	0.172
Quiescent(Q) to apoptosis(A) transition relation			
Cook et al.	0.421	0.162	0.140
Weber et al.	0.484	0.178	0.154
Mandal et al.	0.490	0.182	0.151
Champion et al.	0.441	0.166	0.147
Our system	0.398	0.153	0.131

Simultaneously Extracting Drug-Perturbed Tumor Growth and Drug Pharmacokinetics

For this task, we present the experimental results of all the evaluated systems and consider (1) accuracy in learning the chemotherapy-perturbed tumor growth model and (2) accuracy in learning the pharmacokinetics of the chemotoxic drug (i.e., paclitaxel) dose.

For the tumor growth curve extraction, we considered four cell lines of breast cancer (i.e., MDA-MB-231, MDA-MB-435, MCF-7, LM2-LUC + cell lines).



Simultaneously Extracting Drug-Perturbed Tumor Growth and Drug Pharmacokinetics

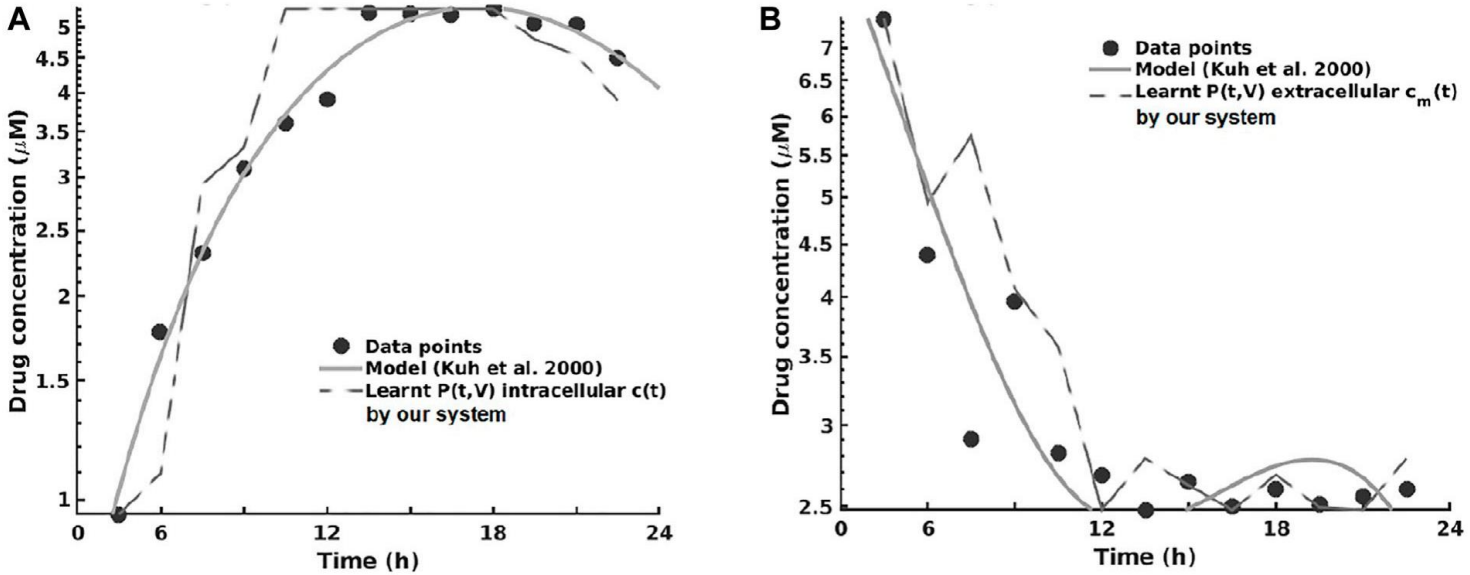
In our study, we use the data from the computational model of **Paclitaxel pharmacokinetics** of Kuh et al. 2000 [8], due to its wide use in **breast cancer chemotherapy schemes**.

The model describes the factors that determine the kinetics of **Paclitaxel uptake**, **binding**, and **efflux** from cells

$$\frac{dc(t)}{dt} = \left[\frac{-A + \sqrt{A^2 + 4K_{d,m}c_m(t)}}{2} - \frac{-B + \sqrt{B^2 + 4(1+NSB)K_{d,c}c(t)}}{2(1+NSB)}\right]\frac{CL_f}{V_{onecell}} - k_{cellnumber}c(t)$$

$$\frac{dc_m(t)}{dt} = \left[\frac{-A + \sqrt{A^2 + 4K_{d,m}c_m(t)}}{2} - \frac{-B + \sqrt{B^2 + 4(1+NSB)K_{d,c}c(t)}}{2(1+NSB)}\right]\frac{CL_f ICN e^{k_{cellnumber}t}}{V_m}$$

Learnt relations:



Simultaneously Extracting Drug-Perturbed Tumor Growth and Drug Pharmacokinetics

In our study, we use the data from the computational model of **Paclitaxel pharmacokinetics** of Kuh et al. 2000 [8], due to its wide use in **breast cancer chemotherapy schemes**. The model describes the factors that determine the kinetics of **Paclitaxel uptake**, **binding**, and **efflux** from cells.

Learnt relations:

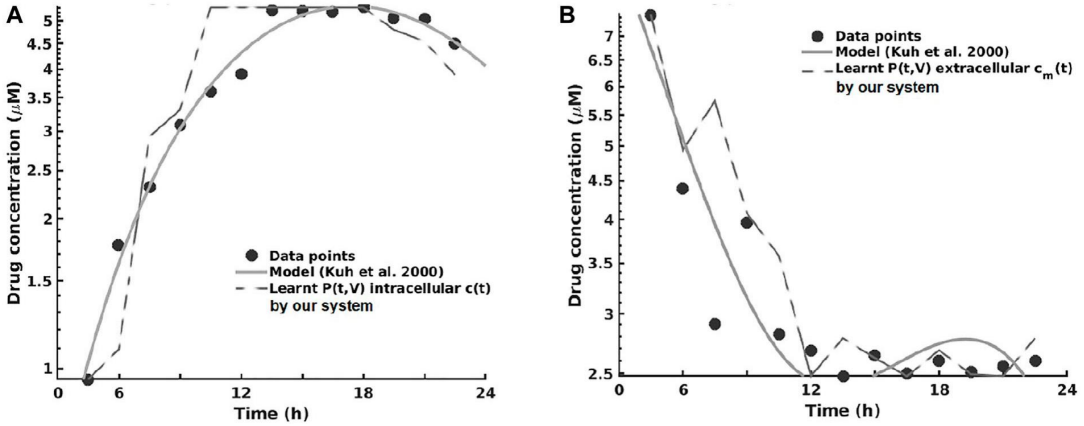


TABLE 7 | Evaluation of the data-driven relation learning systems for pharmacokinetics extraction.

	Evaluation metrics		
Pharmacokinetics data/system	**SSE**	**RMSE**	**sMAPE**
Intracellular paclitaxel			
Cook et al.	0.6234	0.2812	0.1487
Weber et al.	0.7212	0.3689	0.1794
Mandal et al.	0.6743	0.3046	0.1602
Champion et al.	0.6539	0.2607	0.1500
Our system	0.5960	0.2141	0.1403
Extracellular paclitaxel			
Cook et al.	0.5676	0.2341	0.1213
Weber et al.	0.6674	0.2891	0.1289
Mandal et al.	0.6128	0.2974	0.1366
Champion et al.	0.5790	0.2633	0.1156
Our system	0.5484	0.2054	0.1068