A Comparison of Neural Network-based CRN Abstraction Methods: StochNetV2 and DeepCME

Ivan Prigarin

Master's Project WS21/22

Motivation and Aims

- Research questions:
  - How well does StochNetV2 [1] perform on additional examples presented in the DeepCME [2] paper?
  - Is it possible to use StochNetV2 to accurately estimate CRN species concentration moments?
  - Is it possible to use StochNetV2 to accurately estimate CRN parameter sensitivities?
  - If so, what are the advantages and disadvantages of using StochNetV2 for these tasks when compared to DeepCME?

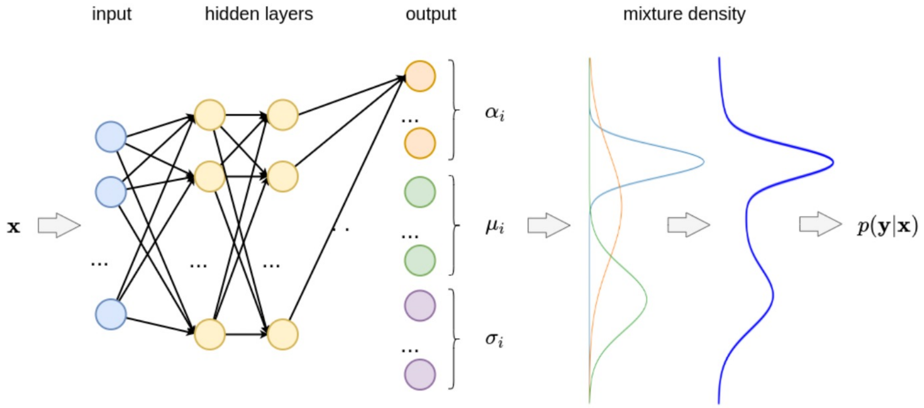
[1] StochNetV2: A Tool for Automated Deep Abstractions for Stochastic Reaction Networks, Denis Repin, Nhat-Huy Phung, Tatjana Petrov, QEST 2020
[2] DeepCME: A deep learning framework for computing solution statistics of the chemical master equation by Ankit Gupta, Christoph Schwab and Mustafa Khammash, *PLoS Computational Biology 2021*

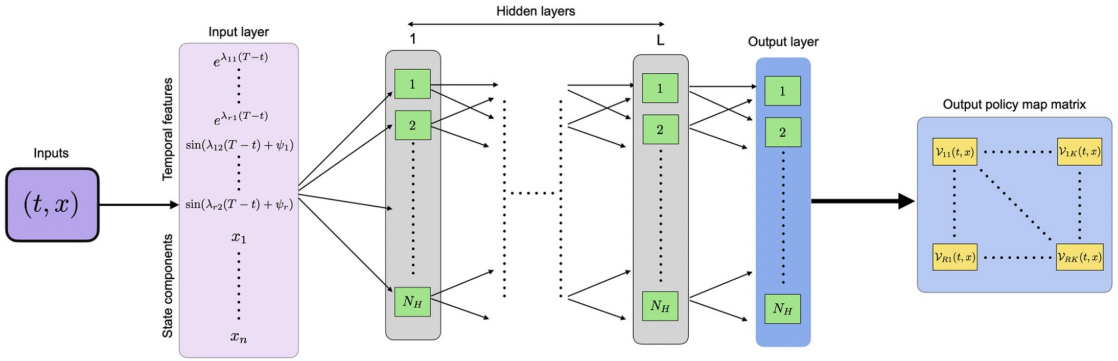
Overview of the methods

	StochNetV2	DeepCME
Training data	CRN trajectories generated with a stochastic simulation	
Inputs	A state of a learned CRN at time t_0, containing the species concentrations and, optionally, CRN parameters	
Outputs	The species concentrations at each time $t_i = t_{i-1} + time\ step$, with the first time being $t_1 = t_0 + time\ step$, until the specified $end\ time$	The 1st and 2nd moments of the n^{th} species of the learned CRN, as well as its parameter sensitivities, at the specified $end\ time$
Neural Network type	Mixture Density Network	Reinforcement Learning
Architecture	Manually defined, or dynamically optimised via the Neural Architecture Search component	Manually defined

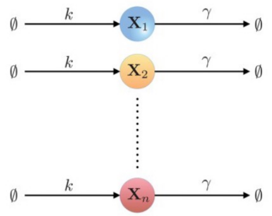
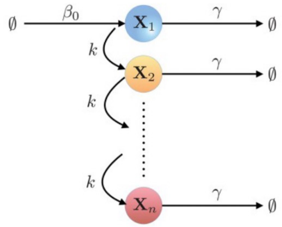
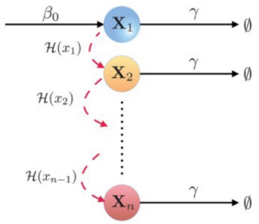
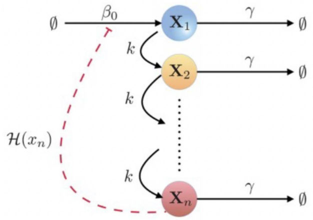
Overview of the methods

StochNetV2	DeepCME
Architecture	

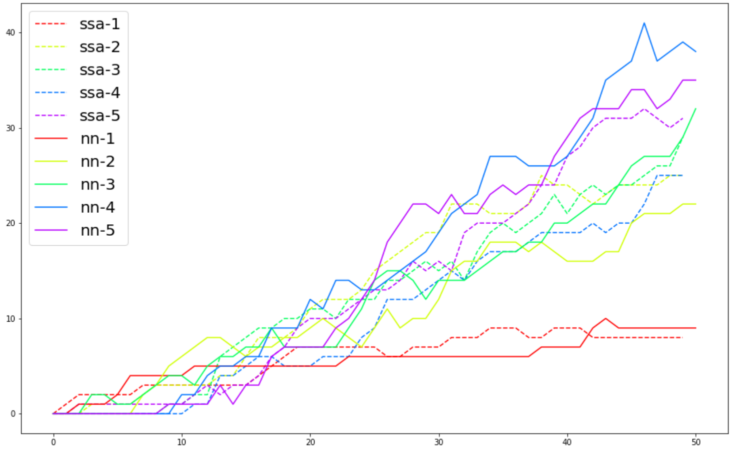
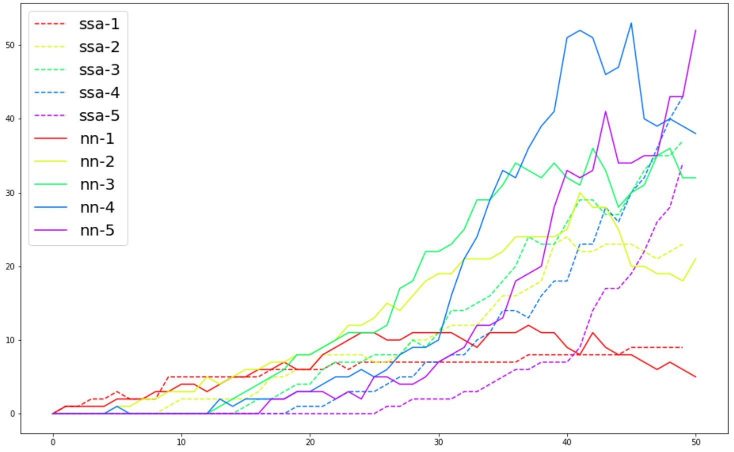
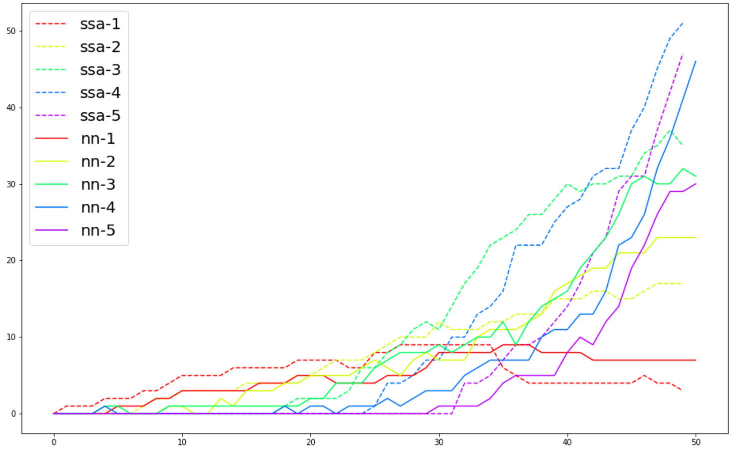




Example CRNs

Name	Reactions	Illustration
Independent birth death network **(BD)**	2 linear reactions	
Linear signalling cascade **(LSC)**	3 linear reactions	
Nonlinear signalling cascade **(NSC)**	1 propensity function with 4 parameters, 2 linear reactions	
Linear signalling cascade with feedback **(LSCF)**	1 propensity function with 4 parameters, 3 linear reactions	

Example CRNs – simulated trajectories

StochNetV2 vs. Gillespie, 5 species		
LSC	NSC5	LSCF5
		

Solid lines – trajectories generated with a trained StochNetV2 neural network.
Dashed lines – trajectories generated with Gillespie.

Note: the “birth death network” is not shown here in the interest of space, considering the relative simplicity of the CRN.

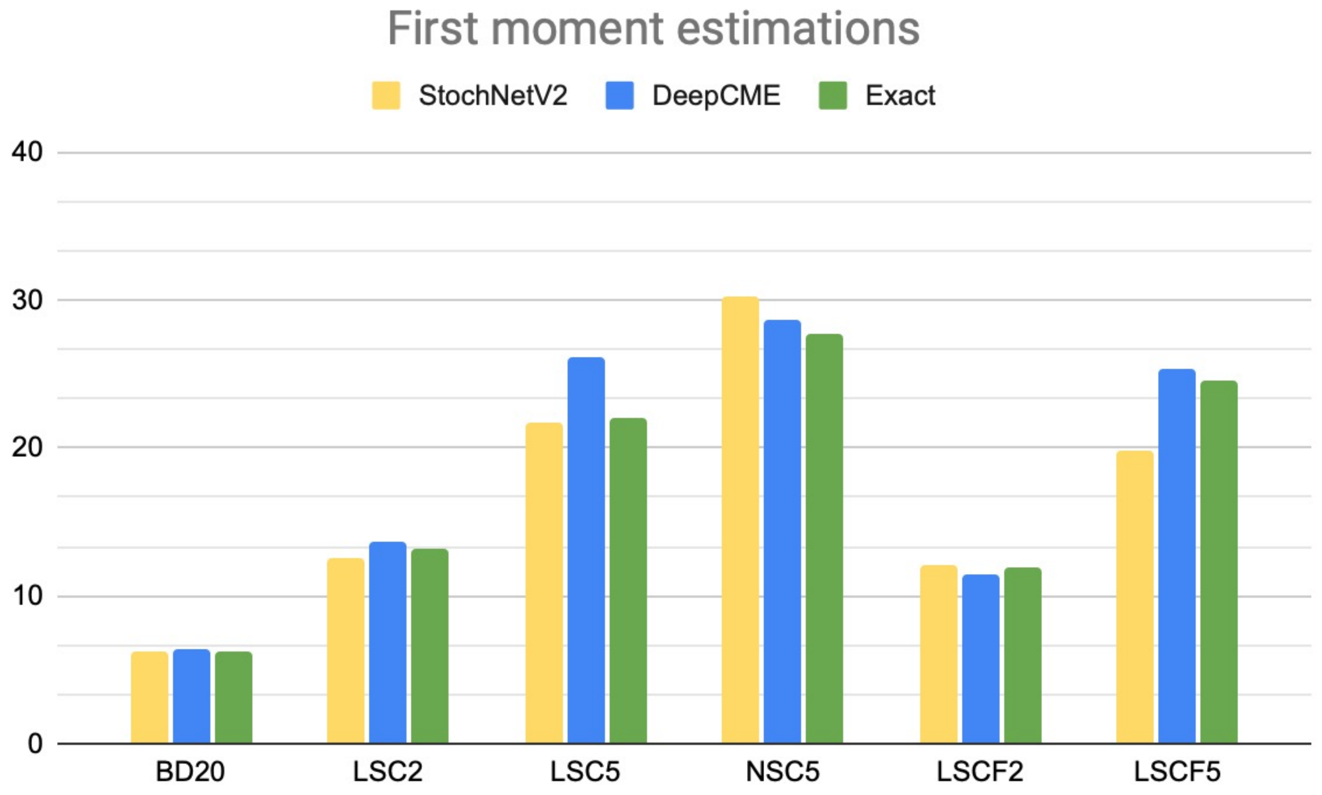
Moment Estimation

- In the corresponding publication, DeepCME focuses on estimating the 1st and 2nd moments of the concentration of species X_n at time T (i.e. the average number of species X_n at time T, as well as the magnitude of the root-mean-square fluctuations about this average).

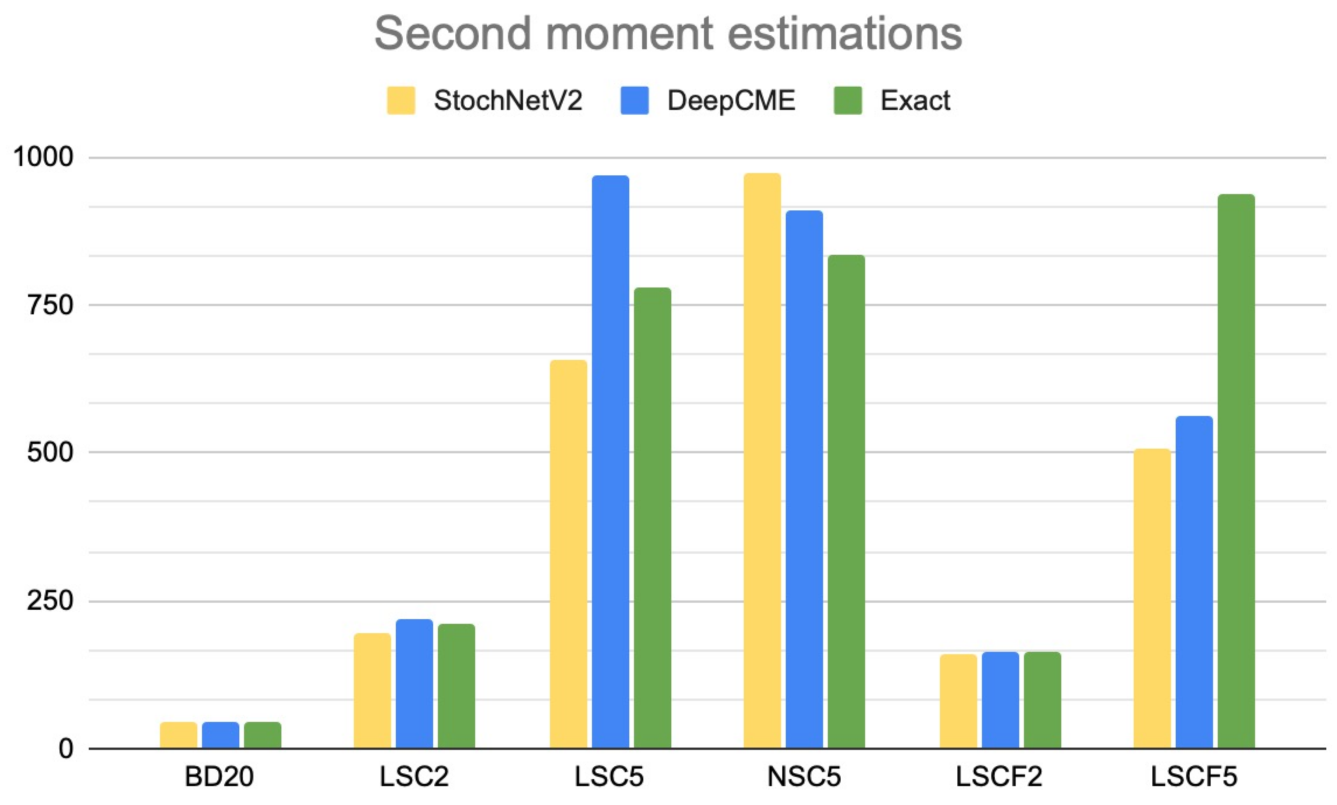
- Using the arrays of traces produces by StochNetV2, we are able to iteratively estimate these quantities for each CRN being abstracted:
  1. For a given CRN, we use the trained StochNetV2 neural network to generate a large number of trajectories using the default parameters;
  2. We iterate through the trajectories, and compute the average of the concentrations of species X_n at $end\ time$, as well as the average of the squares of these concentrations;
  3. The more trajectories that were generated, the smaller the standard deviation of the moment estimates.

- In the following slides, the moment estimations computed for StochNetV2 are displayed alongside the results from abstracting the same CRNs using DeepCME (including the *exact* values for the moments), for select CRNs.

Moment Estimation – first moment



Moment Estimation – second moment



Moment Estimation - findings

- We were able to verify that it is possible to estimate the 1st and 2nd moments of species X_n for a given CRN using the trajectories generated by a correspondingly trained StochNetV2 neural network.

- The achieved results were on par with the estimations obtained from DeepCME, at times presenting a higher degree of accuracy when compared with the exact values, and at other times producing slightly less accurate estimations.

- With this, we have demonstrated that StochNetV2 is able to produce CRN trajectories that successfully capture the temporal dynamics described by the network, and that are able to yield estimates of statistical quantities, such as the species concentration moments considered here, that are sufficiently accurate.

Parameter Sensitivities

- DeepCME defines the sensitivity of the output function on the model parameters as

$$S_\theta(g, T) = \frac{\partial}{\partial \theta}\mathbb{E}(g(X_\theta(T)))$$

- By varying the CRN parameters during the generation of training trajectories for StochNetV2, the corresponding neural network is able to learn the parameter dependencies present in the CRN.

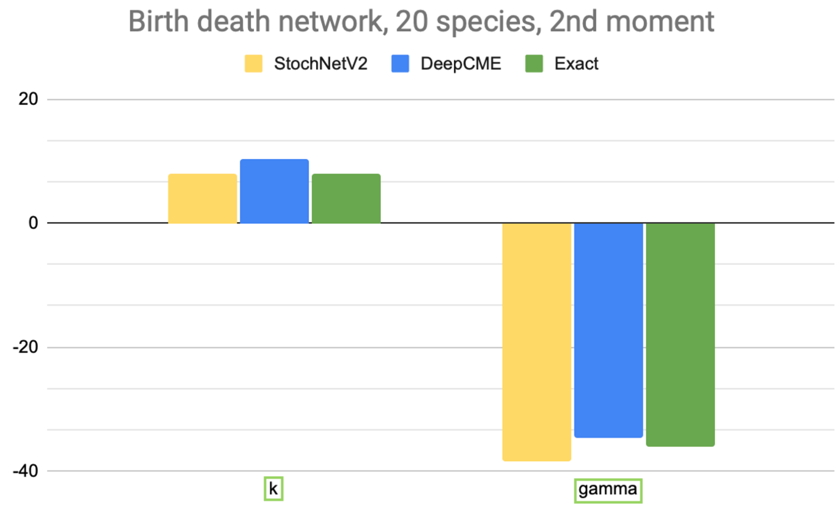
- By using the moment estimations shown before, as well as the above fact, we are able to estimate the sensitivities of the moment estimations with regards to the CRN parameters:
    1. Given a CRN and its parameters, we generate an array of initial states, where the value for one parameter is continuously shifted by a small amount relative to the default value, and all the other parameters are kept fixed at their default values;
    2. The initial states are fed as input into the trained StochNetV2 neural network corresponding to the CRN, and the moment values are estimated;
    3. The average change for each moment is the computed over the shifts of the parameter value;
    4. The process is repeated for each parameter of the CRN.

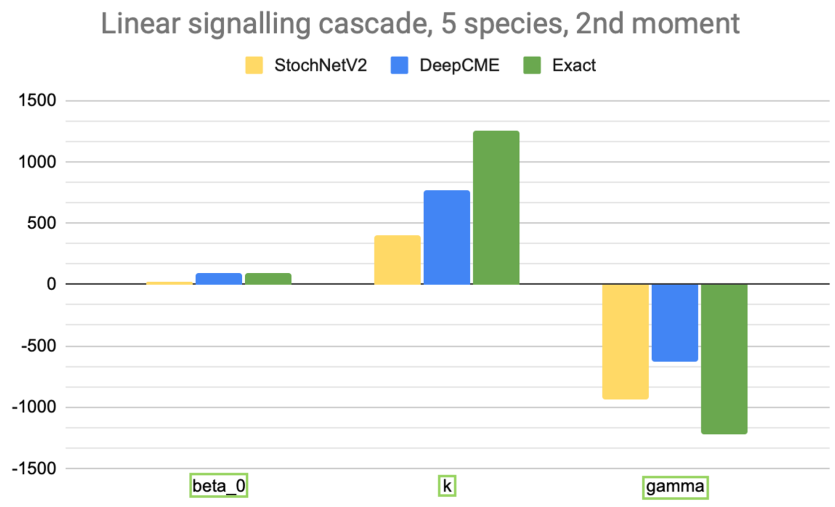
- As before, the parameter sensitivity estimations, as well as the exact values, presented in the following charts were computed by applying StochNetV2 and DeepCME to each example CRN.

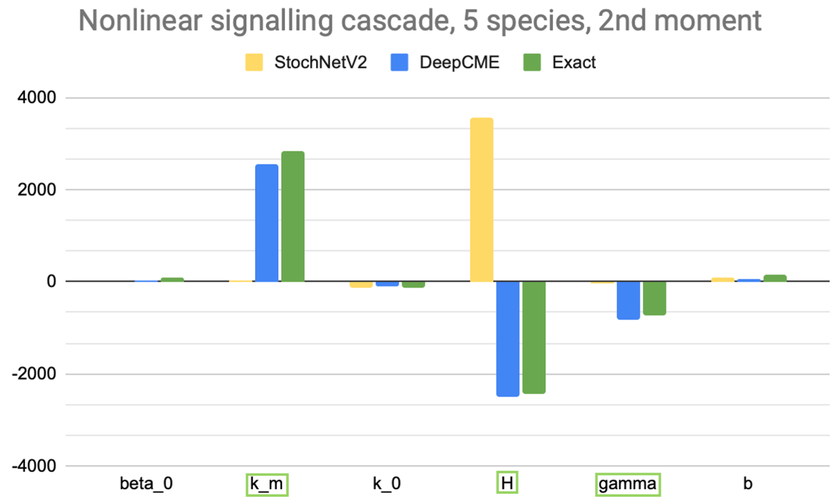
Parameter Sensitivities – first moment

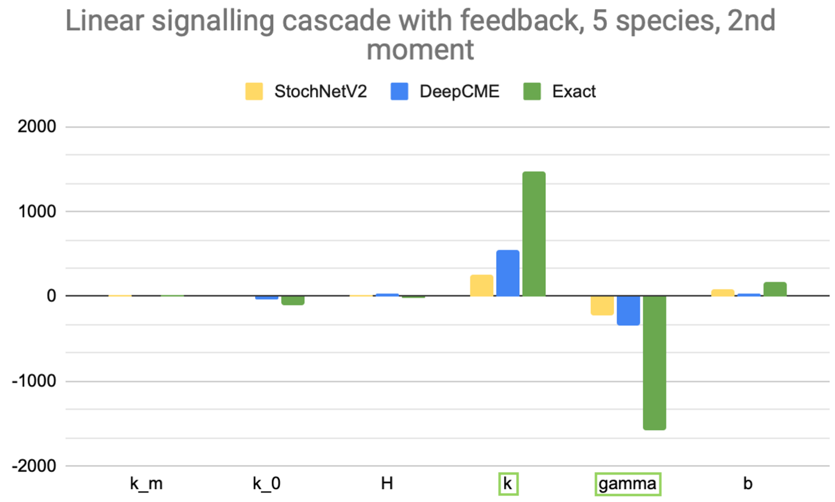
Note: the parameters presented in the DeepCME paper are marked with a green rectangle.
Others were omitted due to relatively small sensitivity estimations, thus being uninteresting.

Parameter Sensitivities – second moment



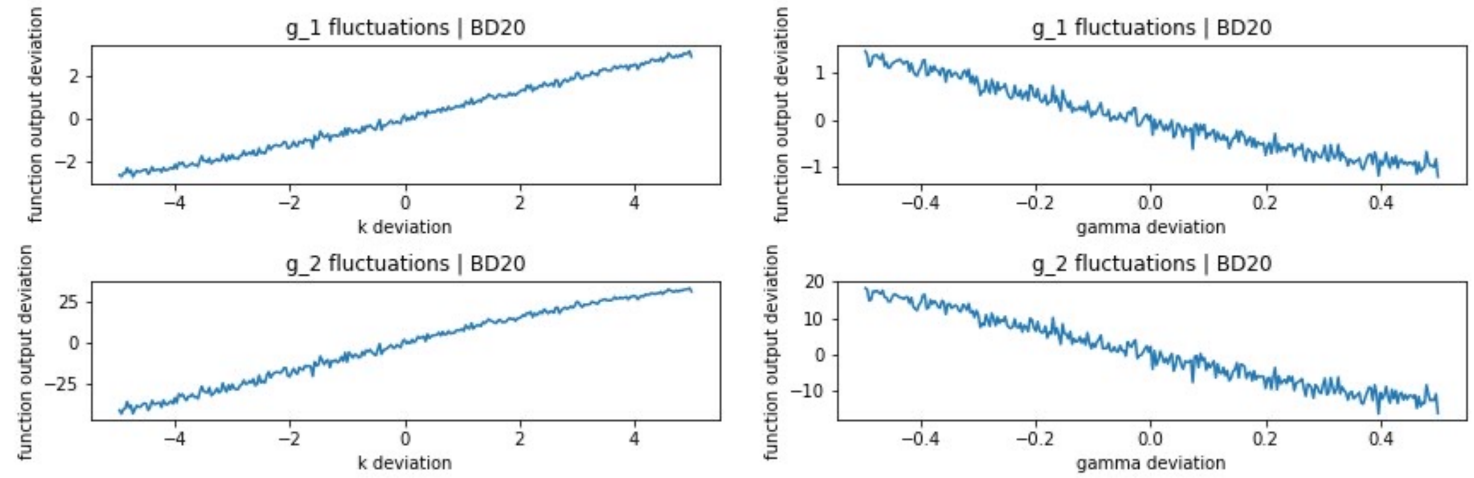




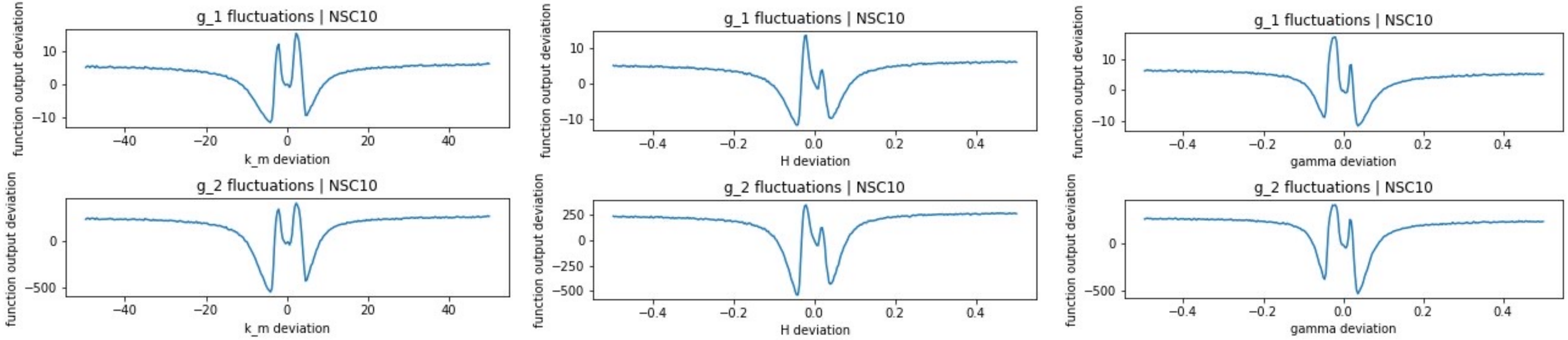


Note: the parameters presented in the DeepCME paper are marked with a green rectangle.
Others were omitted due to relatively small sensitivity estimations, thus being uninteresting.

Parameter Sensitivities - fluctuations

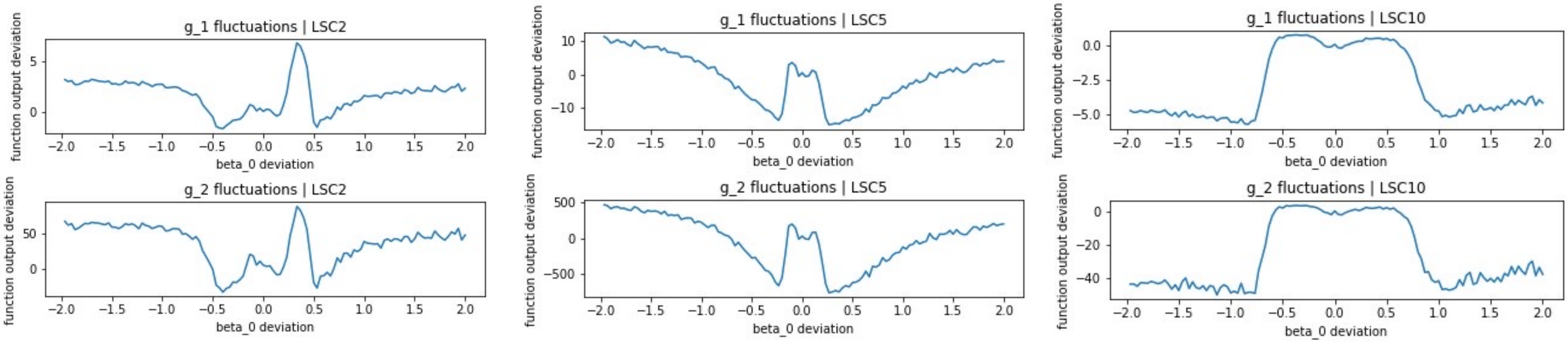


- Besides the sensitivity quantities, we also plotted the moment estimation values together with each parameter to see how the outputs fluctuate in relation to small parameter changes.
- The **birth death network** graphs seen on the left present a clearly linear dependency, while the **nonlinear signalling cascade** graphs seen below signify a much more complex learned parameter dependency.
- Note: *g_1* and *g_2* refer to the first and second moment estimations respectively.



Parameter Sensitivities - fluctuations

- Here we can see the moment estimation fluctuations for the three species configurations of the **linear signalling cascade** network.



Parameter Sensitivities - findings

- It is evident that the CRN parameter dependencies learned by StochNetV2 during training do indeed affect how the simulated network evolves over time.
- The task of estimating parameter sensitivities, though, carries a lot of layered uncertainty, since **the results depend on another estimated quantity**, the species concentration moments.
- In certain cases, this leads to results that either significantly deviate from what is expected (*nonlinear signalling cascade*), follow the correct trajectory but have the wrong scale (*linear signalling cascade with feedback*), or match the expected values relatively accurately (*birth death network, linear signalling cascade*).
- In order to investigate the learned dependencies in more detail, we plotted the **fluctuations in the moment estimations with respect to the changes in each parameter**. This revealed that in all but one case, the parameter dependencies are non-linear around the default parameter value, **eventually becoming non-existent** when the parameter value is shifted far enough from the default.
- Potentially, this signifies that **the neural network might be overfitting**, resulting in poor generalisation when presented with parameter values too dissimilar from what the network encountered during training.

Conclusion and outlook

- We were able to answer the research questions defined for this project:
  - StochNetV2 was successfully applied to the example CRNs presented in the DeepCME paper, producing trajectories that **accurately followed the temporal dynamics** seen when performing stochastic simulations of the CRNs using the Gillespie algorithm;
  - Moreover, we have shown that the trajectories generated with a trained StochNetV2 neural network **can be used to accurately estimate statistical quantities of CRNs**, such as the species concentration moments. We compared these estimations with the values produced by DeepCME to validate this claim.
  - Additionally, we investigated whether the generated trajectories **can be used for estimating how the above moment estimations depend on the CRN parameters**. Our findings demonstrated that **it is possible, though not highly reliable**, due to the nature of estimating sensitivities of estimated values. The StochNetV2 neural networks are able to learn the parameter dependencies for a given CRN, but evidently struggle to generalise to the greater parameter value space.
  - DeepCME, on the other hand, **is built to output estimations of the above quantities**, as opposed to generating trajectories for a given CRN. It expectedly outperforms StochNetV2 in terms of how accurately these values are estimated, though requires additional configuration and training when needing to estimate other statistics. StochNetV2 neural networks only need to be trained once, and the generated trajectories can be analysed to compute any number of CRN statistics iteratively.
- In terms of the next steps, we would like to investigate the various components of StochNetV2 and DeepCME in more detail, and assess the potential combinations that could benefit each method, or result in a new method altogether.