Computational framework for global sensitivity analysis of high-dimensional and poorly connected parameter spaces

Ž Pušnik, L Magdevska, M Mraz, N Zimic, and M Moškon*

Computational Biology Group, Faculty of Computer and Information Science, University of Ljubljana, Slovenia,
**Correspondence: miha.moskon@fri.uni-lj.si*

Keywords: Genetic algorithms, Morris sensitivity analysis, orthogonal sampling, clustering.

1 Introduction

Sensitivity analysis methods are widely applied to systems as well as to synthetic biology since they are able to provide the system's robustness, can guide parameter estimation, experimental design and model simplification, and last but not least, can serve as a model validation tool [1]. These methods can be roughly divided in two groups, i.e. *local* and *global methods* [2, 3]. While local methods are easy to apply, they yield system's sensitivity to perturbations of single parameter at a time in the neighbourhood of the nominal parameter values. Local sensitivity analysis methods prove to be inefficient when the parameter values describing the dynamics of the system exhibit large variations or when parameter values are missing or only partially known, which are usually the cases when dealing with biological systems [4]. Global sensitivity analysis methods on the other hand tend to investigate the whole space of possible parameter values applying different sampling techniques in a combination with computer simulations [3, 4]. In some cases, however, the subset of parameter values, so called *viable parameter space*, for which a biological system maintains *desired qualitative behaviour*, e.g., sustained oscillations, is very limited [5]. *Quantitative sensitivity analysis* assessment, which describes for example fluctuations of oscillation amplitudes and periods, is in such cases focused to viable parameter regions only. Global methods prove to be inefficient when viable solution space is small in comparison to the whole solutions space and especially when dealing with high dimensional models [5]. *Glocal methods*, which combine global approaches to identify viable parameter regions and local approaches to inspect the viable regions, have already been reported on the field of robustness analysis [5, 6]. These methods, however, still reflect certain limitations. While method described in [5] is unable to identify viable regions that are not connected and reside far from each other, method described in [6] is limited to only convex parameter regions. Moreover, its complexity scales exponentially with the number of model dimensions. As an alternative, sensitivity analysis based on sparse polynomial approximations of high dimensional models was introduced in [7]. However, the method proves to be efficient only for relatively small feasible parameter ranges.

Herein, we introduce a computational framework that is able to perform the global sensitivity analysis of high-dimensional and poorly connected viable parameter regions. Each of the steps comprising the framework is described in Section 2, while Section 3 discusses the results of the framework and concludes the manuscript.

2 Computational framework

2.1 Generating viable solutions

Viable parameter spaces are composed of viable solutions, for which desired simulation response of observed model is obtained. While brute force investigation of parameter space exhibits exponential dependence on the number of dimensions [5] we identify viable solutions using optimisation metaheuristics. More precisely we apply *genetic algorithms* (GAs), which mimic natural evolution in order to exhaustively examine the landscape of possible solution (see Figure 1).

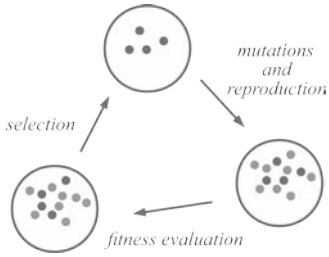

Figure 1: Generating viable solutions with genetic algorithms.

2.2 Clustering the solutions

Clustering is performed to identify different, potentially distant and poorly connected regions within the viable parameter space (see Figure 2). We perform the clustering using *k-means method*, which generates *k* parameter regions on the basis of distance between the parameter values describing each of the solutions.

2.3 Sampling

While the solutions obtained with GAs are biased and may have different numbers of representatives within different parameter regions, resampling is performed on each of the clusters. We apply *orthogonal sampling* [9] to generate the same number of samples in each cluster (see Figure 3).

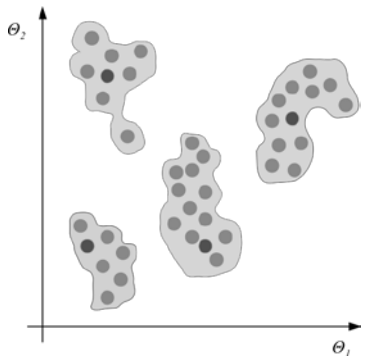

Figure 2: Identifying viable parameter regions in two-dimensional parameter space using clustering.

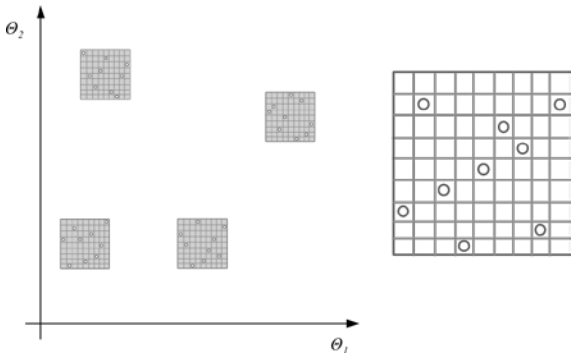

Figure 3: Resampling the viable parameter regions with orthogonal sampling.

2.4 Sensitivity analysis

Each of the generated samples presents a single nominal value around which the *Morris sensitivity analysis* [1, 3, 10] is performed (see Figure 4). Sensitivity analysis yields the average mean and standard deviation of partial effects for each parameter and for each cluster. The sensitivity is assessed in qualitative as well as in quantitative aspects. While qualitative aspect evaluates the sensitivity of solution viability (e.g., sensitivity of oscillatory behaviour), quantitative aspect approximates the variability of viable solutions (e.g., sensitivity of oscillation amplitudes and periods).

2.5 Combining results

In the last step of the framework average means and standard deviations of partial effects are combined for all clusters. We also observe the standard deviations of partial effects between clusters, which can be used to evaluate the effects different regions might have on the sensitivity values.

3 Results and conclusion

Described computational framework can be efficiently applied to the sensitivity assessment even when dealing with poorly connected, non-convex and high dimensional parameter spaces. We successfully applied the framework to the design and analysis of genetic *master-slave D flip-flop*, for which several unconnected viable parameter regions were

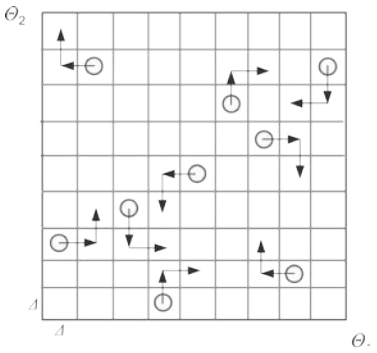

Figure 4: Graphical representation of Morris sensitivity assessment in two-dimensional parameter space, where circles represent samples generated in preceding step and Δ describes the parameter perturbation step.

identified. Results obtained with the framework complied with other analyses, such as *evaluation of entropies of optimal kinetic parameter values*. We were able to use the framework to distinguish the most robust topologies and identify the parameter regions with the best dynamical properties.

References

[1] Z. Zi. "Sensitivity analysis approaches applied to systems biology models", *IET Systems Biology*, **5**, pp. 336-346, (2011).

[2] B. Taylor, T. J. Lee, J. S. Weitz. "A guide to sensitivity analysis of quantitative models of gene expression dynamics", *Methods*, **62**, pp. 109-120, (2013).

[3] A. Saltelli, S. Tarantola, F. Campolongo, M. Ratto: *Sensitivity analysis in practice: A guide to assessing scientific models*, John Wiley & Sons, Ltd, 2004.

[4] S. Marino, I. B. Hogue, C. J. Ray, D. E. Kirschner. "A Methodology For Performing Global Uncertainty And Sensitivity Analysis In Systems Biology", *Journal of Theoretical Biology*, **254**, pp. 178-196, (2008).

[5] E. Zamora-Sillero, M. Hafner, A. Ibig, J. Stelling, A. Wagner. "Efficient characterization of high-dimensional parameter spaces for systems biology", *BMC Systems Biology*, **5** (2011).

[6] M. Hafner, H. Koeppl, M. Hasler, A. Wagner. "*Glocal* robustness analysis and model discrimination of circadian oscillators", *PLoS Computational Biology*, **5**, e1000534, (2009).

[7] C. Schillings, M. Sunnåker, J. Stelling, C. Schwab. "Efficient characterization of parametric uncertainty of complex (bio)chemical networks", *PLoS Computational Biology*, **11**, e1004457, (2015).

[8] J. Sun, J. M. Garibaldi, C. Hodgman. "Parameter estimation using metaheuristics in systems biology: A comprehensive review", *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **9**, pp. 185-202, (2012).

[9] A. G. Garcia. "Orthogonal sampling formulas: A unified approach". *SIAM Review*, **42**, pp. 499-512, (2000).

[10] M. Morris. "Factorial Sampling Plans for Preliminary Computational Experiments". *Technometrics*, **33**, pp. 161-174. (1991).