Bridging the Gap between Random Differential Equations and Structural Causal Models

Stephan Bongers
Informatics Institute
University of Amsterdam
The Netherlands
S.R.Bongers@uva.nl

Joris M. Mooij
Informatics Institute
University of Amsterdam
The Netherlands
J.M.Mooij@uva.nl

Random differential equations (RDEs) provide a natural extension of ordinary differential equations to the stochastic setting and have been widely accepted as an important mathematical tool in modeling and analysis of numerous processes in physics, finance, biochemistry and engineering systems (Bunke, 1972; Soong, 1973; Sobczyk, 1991; Rupp and Neckel, 2013). The processes described by these systems are not only of stochastic nature, but are also of causal nature, in the sense that the random differential equations encode how the time-varying internal or external processes change other processes over time. Although in principle random differential equations could be used for modeling causal relationships between such processes, infering such causal models from data is often difficult. A significant practical drawback of this modeling class is that obtaining time series data with sufficiently high temporal resolution is often costly, impractical or even impossible. Another issue is that if one has only access to a subset of the system's processes, for example due to practical limitations on the measurability of some of the processes, then in general there does not have to exist an RDE on this subset of processes that could be estimated. A similar issue arises when the RDE contains exogenous latent confounding processes.

Structural causal models (SCMs), also known as structural equation models, are another well-studied causal modeling tool and have been widely applied in the genetics, economics, engineering and social sciences (Pearl, 2009; Spirtes et al., 2000; Bollen, 1989). One of the advantages of SCMs over other causal modeling tools is that they have the ability to deal with cyclic causal relationships (Spirtes, 1995; Hyttinen et al., 2012; Mooij et al., 2011; Forré and Mooij, 2017; Bongers et al., 2018). In particular, recent work has shown how one can apply the (general) directed global Markov property (Forré and Mooij, 2017), how one can deal with marginalization and how one can causally interpret these models in the cyclic setting (Bongers et al., 2018).

Over the years, several attempts have been made to interpret these structural causal models that include cyclic causal relationships. They can be derived from an underlying discrete-time or continuous-time dynamical system (Fisher, 1970; Iwasaki and Simon, 1994; Dash, 2005; Lacerda et al., 2008; Mooij et al., 2013; Rubenstein et al., 2018). All these methods assume that the dynamical system under consideration converges to a single static equilibrium, with the exception of the analysis by Fisher (1970), who assumes that observations are time averages of a dynamical system, and the recent work of Rubenstein et al. (2018), who has shown that dynamic asymptotic behaviour of ordinary differential equations can be captured by dynamic structural causal models. This single static equilibrium assumption gives rise to a more parsimonious description of the causal relationships of the equilibrium solutions and ignores the complicated but decaying transient dynamics of the dynamical system. The assumption that the system has to equilibrate to a single static equilibrium is rather strong and limits the applicability of the theory, as many dynamical systems have multiple equilibrium solutions.

In this work, we relax this condition and capture, under certain equilibrium convergence assumptions, every random equilibrium solution of the RDE in an SCM. Conversely, we show that under suitable conditions, every solution of the SCM corresponds to a sample-path solution of the RDE. Intuitively, the idea is that in the limit when time tends to infinity the random differential equations converge exactly to the structural equations of the SCM, and hence in this limit every sample-path solution of the RDE is encoded as a solution of the structural equations. This is illustrated for the basic enzyme reaction in Figure 1. There we show how the randomness of the initial conditions evolve over time and is captured in the constructed SCM at equilibrium. Moreover, we show that this construction is compatible with interventions under similar equilibrium convergence assumptions, as is illustrated at the bottom part of Figure 1. We like to stress

that our construction automatically captures the stochastic behavior of the RDE in the associated SCM. It can deal with randomness in the initial conditions, the coefficients and via the random inhomogenous part (captured as noise in the SCM), thereby significantly extending the work by Mooij et al. (2013) who only considers the deterministic setting.

The advantage of SCMs over RDEs is that by not modeling the transient random dynamics of the RDE, one arrives at a more compact representation for learning and prediction purposes of random systems that have reached equilibrium. Another advantage is that the equilibrium solutions of the RDE can be studied by statistical tools applicable to SCMs. For example, one can marginalize over a subset of the system's variables and get an even more parsimonious representation that preserves the causal semantics (Bongers et al., 2018). This is illustrated in Figure 2 for the example of a damped coupled harmonic oscillator, where we marginalize over the momentum variables. Moreover, we can apply the (general) directed global Markov property to the equilibrium solutions of the SCM by using d-separation (and σ-separation) (Bongers et al., 2018). This is also illustrated in Figure 2, where we perform d- and σ-separation on the equilibrium solutions of the intervened model and see that the position variables Q_1 and Q_5 are independent given the position variable Q_3 which we held fixed. This enables the study of stochastic and causal behavior of the equilibrium solutions of the RDE in terms of SCMs and hence this sheds some new light on the concept of causality as expressed within the framework of structural causal models. Yet another advantage is that it is easier to deal with confounders within the framework of SCMs, as we only need to model the equilibrium distribution of these confounders, and don't need to model their dynamics.

In summary, we built a bridge between the world of random differential equations and the world of structural causal models. This allows us to study a plethora of physical and engineering systems subject to time-varying random disturbances within the framework of structural causal models. We naturally extend the work of Mooij et al. (2013) to the stochastic setting, which allows us to address both cycles and confounders. In particular, we relaxed the condition that the dynamical system has to equilibrate to a single static equilibrium, such that it can deal with multiple equilibrium solutions, and show that if an RDE is sufficiently regular all equilibrium sample-path solutions of the RDE are described by the solutions of the associated SCM, while preserving the causal semantics.

Acknowledgements

Stephan Bongers and Joris Mooij were supported by NWO, the Netherlands Organization for Scientific Research (VIDI grant 639.072.410).

References

Bollen, K. (1989). *Structural Equations with Latent Variables*. John Wiley & Sons, New York, USA.

Bongers, S., Peters, J., Schölkopf, B., and Mooij, J. M. (2018). Theoretical aspects of cyclic structural causal models. *arXiv.org preprint*, arXiv:1611.06221v2 [stat.ME].

Dash, D. (2005). Restructuring dynamic causal systems in equilibrium. In *Proceedings of the Tenth International Workshop on Artificial Intelligence and Statistics (AISTATS 2005)*.

Fisher, F. M. (1970). A Correspondence Principle for Simultaneous Equation Models. *Econometrica*, 38(1):73–92.

Forré, P. and Mooij, J. M. (2017). Markov properties for graphical models with cycles and latent variables. *arXiv.org preprint*, arXiv:1710.08775 [math.ST].

Hyttinen, A., Eberhardt, F., and Hoyer, P. O. (2012). Learning linear cyclic causal models with latent variables. *The Journal of Machine Learning Research*, 13(1):3387–3439.

Iwasaki, Y. and Simon, H. A. (1994). Causality and model abstraction. *Artificial Intelligence*, 67(1):143–194.

Lacerda, G., Spirtes, P. L., Ramsey, J., and Hoyer, P. O. (2008). Discovering cyclic causal models by independent components analysis. In *Proceedings of the Twenty-Fourth Conference on Uncertainty in Artificial Intelligence (UAI 2008)*.

Mooij, J. M., Janzing, D., Heskes, T., and Schölkopf, B. (2011). On causal discovery with cyclic additive noise models. In *Advances in Neural Information Processing Systems (NIPS 2011)*, pages 639–647.

Mooij, J. M., Janzing, D., and Schölkopf, B. (2013). From ordinary differential equations to structural causal models: the deterministic case. In Nicholson, A. and Smyth, P., editors, *Proceedings of the 29th Annual Conference on Uncertainty in Artificial Intelligence (UAI-13)*, pages 440–448. AUAI Press.

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*. Cambridge University Press, New York, USA, 2nd edition.

Rubenstein, P. K., Bongers, S., Schölkopf, B., and Mooij, J. M. (2018). From deterministic ODEs to dynamic

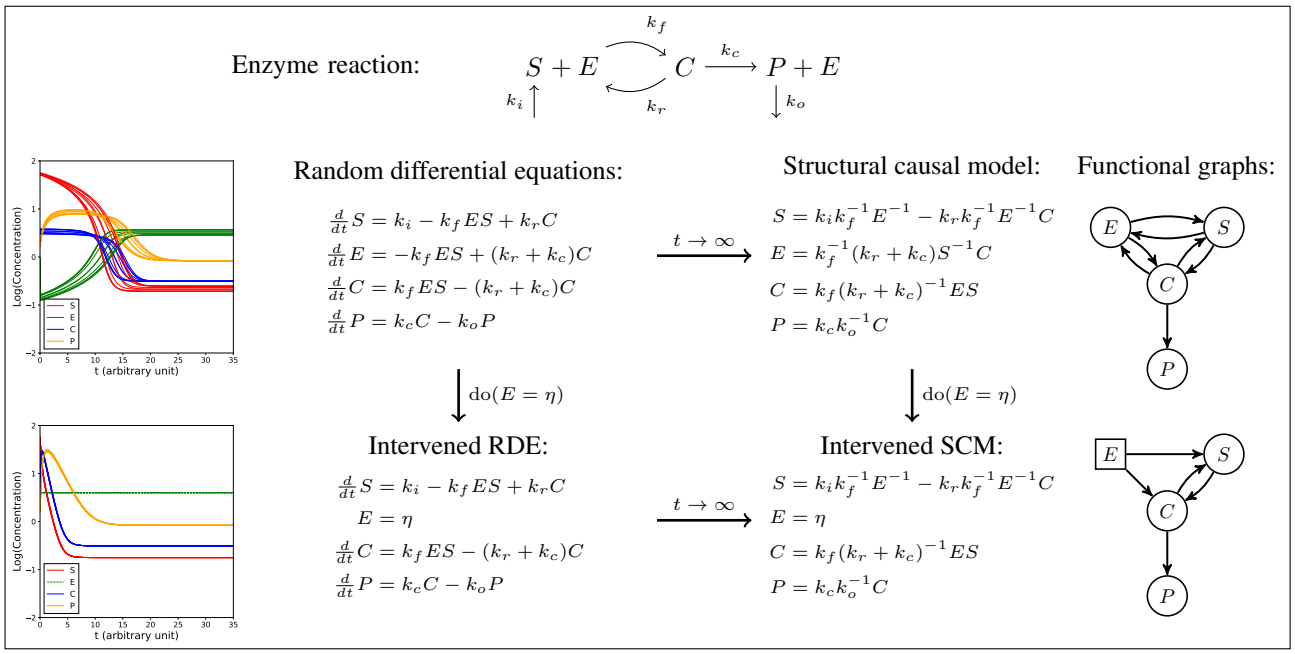

Figure 1: *Example that illustrates how structural causal models can be used to describe the equilibrium solutions of random differential equations (in this case, an enzyme reaction) and how these change under external interventions (in this case, keeping the enzyme concentration E at a fixed value η). The diagram is commutative.*

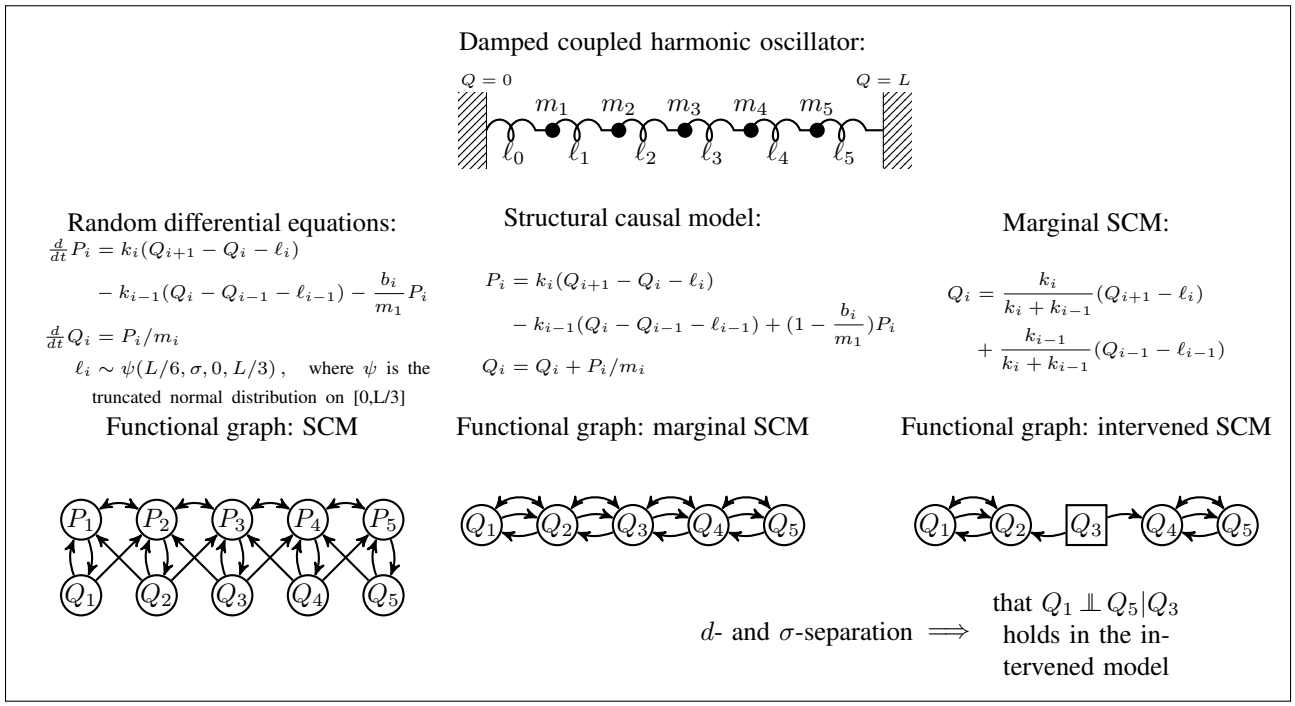

Figure 2: *Example that illustrates that the equilibrium solutions of the RDE (in this case, of a damped coupled harmonic oscillator) can be studied by statistical tools applicable to SCMs (in this case, we first perform a marginalization over the momentum variables, and in turn perform d- and σ-separation on the intervened graph to conclude that $Q_1 \perp\!\!\!\perp Q_5 | Q_3$).*

structural causal models. In *Proceedings of the 34th Annual Conference on Uncertainty in Artificial Intelligence (UAI-18)*.

Spirtes, P. (1995). Directed cyclic graphical representations of feedback models. In *Proceedings of the Eleventh conference on Uncertainty in Artificial Intelligence (UAI 1995)*, pages 491–498.

Spirtes, P., Glymour, C., and Scheines, R. (2000). *Causation, Prediction, and Search*. The MIT Press, Cambridge, Massachusetts, 2nd edition.