Appendix 1

Computing the Adaptive Cycle

Consider a system $\mathcal{V}$ of interacting agents, each of which is described by a finite time series of states $(x_1, \ldots, x_T)$. We consider each time series as finite realization of a stationary Markov process of order k. Assuming that every effective interaction among the agents leads to a transfer of information between them, transfer entropy as defined by Thomas Schreiber (Schreiber 2000) can be used to quantify interactions within the system. Precisely, let $(x_1, \ldots, x_T)$ and $(y_1, \ldots, y_T)$ be the time series of agents X and Y, respectively. Let k and l, respectively, denote their estimated Markov orders. In the following, this quantity will be called *history length*. Setting a *window size* w_t ($\max\{k,l\}+1 \leq w_t \leq T$), transfer entropy at time point t ($\max\{k,l\}+1 \leq t \leq T$) can be estimated via

$$\tilde{T}^t_{Y\to X} = \sum_{i=t-w_t+1+\max\{k,l\}}^{t-1} \log\left(\frac{\tilde{p}\left(x_{i+1}|x_i^{(k)}, y_i^{(l)}\right)}{\tilde{p}\left(x_{i+1} \,\middle|\, x_i^{(k)}\right)}\right), \tag{1}$$

with the probabilities/densities $\tilde{p}$ being estimated on basis of data in the time window, i.e. $(x_{t-w_t+1}, \ldots, x_t)$ and $(y_{t-w_t+1}, \ldots, y_t)$. Note that, if $T < 15$, two additional data points are interpolated between two original data points (x_i, x_{i+1}) each via a piecewise cubic spline before the estimation, thereby increasing the stability of the results. In this case, this fact has to be respected in the choice of the history length.

We use the Kraskov-Stögbauer-Grassberger estimator as being incorporated in the JIDT toolkit (Lizier 2014). Note that, in the estimation procedure, a certain amount of random Gaussian noise is added to the original data in order to guarantee reliability of the estimator. In the following, we call this quantity *noise level*. Simultaneously to the estimation, we conduct a significance test being provided by the JIDT toolkit. Only results passing a certain *level of significance* are taken into account.

We repeat this procedure with all pairs of components at time t. Considering the system's components as nodes, the transfer entropy $\tilde{T}^t_{Y\to X}$ as weight of edge $e_{Y\to X}$ at time t, we gain a weighted, directed graph as inferred model of interaction at time t. Setting $\tilde{T}^t = \sum \tilde{T}^t_{Y\to X}$, we define the system's *potential* at the given time as

$$P = -\sum_{(Y,X)\in\mathcal{V}\times\mathcal{V}} \tilde{T}^t_{Y\to X} \cdot \log_2\left(\frac{\tilde{T}^t_{Y\to X}}{\tilde{T}^t}\right)$$

and the system's *connectedness* at the given time as

$$C = \sum_{(Y,X)\in\mathcal{V}\times\mathcal{V}} \tilde{T}^t_{Y\to X} \cdot \log_2\left(\frac{\tilde{T}^t_{Y\to X} \cdot \tilde{T}^t}{\sum_{X'\in\mathcal{V}} \tilde{T}^t_{Y\to X'} \cdot \sum_{Y'\in\mathcal{V}} \tilde{T}^t_{Y'\to X}}\right).$$

Denote by A the graph's adjacency matrix, by D_{out} and D_{in} its directed degree matrices. Setting a *standardization constant* c, we define

$$L_{out} = c \cdot D_{out}^{-\frac{1}{2}}\left(D_{out} - A\right), \quad \text{and} \quad L_{in} = c \cdot \left(D_{in} - A\right) D_{in}^{-\frac{1}{2}}.$$

as the graph's directed Laplacian matrices. Let L_{in} and L_{out} be the Laplacian matrices of the system's information network. We define the smallest non-trivial real part of the eigenvalues of L_{out} and L_{in},

$$R = \min \{\mathcal{R}\, \sigma : \sigma \in \mathrm{Spec}(L_{out}) \cup \mathrm{Spec}(L_{in}), \sigma \neq 0\},$$

as the system's *resilience*.

Given time series of abundances of length T for each component, we can estimate a sequence of interaction networks for time points $w_1, w_1+1, \ldots, T$. This allows us to determine the development of the three systemic variables during this period.