Mathematical derivation

February 24, 2020

1 Cross-correlation

The normalized pairwise cross-correlation of the time traces x_{i1}, x_{j1} of two nodes with means $\mu(\cdot)$ and standard deviations $\sigma(\cdot)$ is given by:

$$R_{ij} = \int_0^T \frac{\big(x_{i1}(t) - \mu(x_{i1})\big)\big(x_{j1}(t) - \mu(x_{j1})\big)}{\sigma(x_{i1})\sigma(x_{j1})} dt \tag{1}$$

Above cross-correlation for all pairs of nodes gives us a cross-correlation matrix $\boldsymbol{R}$. The mean of all its entries is the network cross-correlation c:

$$c = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} R_{ij} \tag{2}$$

2 Mathematical formulation of the optimal control problem

2.1 Optimization problem: general formulation

In this section the mathematical foundations for controlling the network dynamics are laid. We use nonlinear sparse optimal control. All below derivations are similar to the ones in references (?) where sparse optimal control was applied to a system of partial differential equations with diffusive coupling.

First we define our minimization problem. For this we use a more general expression for our network system N d-dimensional nodes. It is given by following system of dN stochastic differential equations:

$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{h}(\mathbf{x}(t)) + \sigma(\mathbf{A} \otimes \mathbf{G})\mathbf{x}(t) + (\mathbf{B} \otimes \mathbf{K})\mathbf{u}(t) + \eta(\mathbf{I_N} \otimes \mathbf{D})\boldsymbol{\xi}(t) \tag{3}$$

with the state vector $\mathbf{x} = (\mathbf{x_1}, ..., \mathbf{x_N})$ and $\mathbf{x_i} = (x_{i1}, \ldots, x_{id})$. $\otimes$ denotes the Kronecker product. The local node dynamics are $\mathbf{h}(\mathbf{x}) = (\mathbf{h}(\mathbf{x_1}), ..., \mathbf{h}(\mathbf{x_N}))$

with $\mathbf{h}(\mathbf{x_i}) = (h_1(\mathbf{x_i}), \ldots, h_d(\mathbf{x_i}))$. The coupling term consists of the $N \times N$-dimensional adjacency matrix $\mathbf{A}$ and the $d \times d$-dimensional local coupling scheme $\mathbf{G}$. The control term consists of the control vector $\mathbf{u} = (\mathbf{u_1}, ..., \mathbf{u_N})$ with $\mathbf{u_i} = (u_{i1}, \ldots, u_{id})$, the diagonal $N \times N$ dimensional control matrix $\mathbf{B}$ and the $d \times d$-dimensional local control scheme $\mathbf{K}$. The noise term consists of the noise intensity η and the Kronecker product of an N-dimensional identity matrix $\mathbf{A}$ and the $d \times d$-dimensional local noise scheme $\mathbf{G}$. The stochastic process is given by ξ. Following boundary conditions are given:

$$\mathbf{x}(t = 0) = \mathbf{x_0} \tag{4}$$

When inserting the local dynamics of a FitzHugh-Nagumo oscillator for $\mathbf{h}(\mathbf{x}(t))$ given in equation (**??**), setting $\mathbf{B} = \mathbf{I_N}$ and $\mathbf{G} = \mathbf{K} = \mathbf{D} = [[1, 0], [0, 0]]$ equation (3) is equivalent to equation (**??**) in above section.

Since our system is stochastic and the control we want to apply is deterministic, the optimal control should minimize the mean cost functional over many noise realizations R. The angle brackets denote the mean over realizations. The mean functional that is to be minimized is defined as:

$$\langle F(\mathbf{x}(\mathbf{u}), \mathbf{u})\rangle = \int_0^T \langle f(\mathbf{x}(\mathbf{u}), \mathbf{u})\rangle dt \tag{5}$$

Our goal is to find the optimal control $\overline{\mathbf{u}}$ that minimizes the above mean functional and for which following variational inequality holds for every dimension i:

$$\langle F(\mathbf{x}(u_i), u_i)\rangle - \langle F(\mathbf{x}(\overline{u}_i), \overline{u}_i)\rangle \geq 0 \tag{6}$$

With this inequality we ensure, that at the optimal solution $\overline{u}$ the functional has a minimum (and not a saddle). We linearize around the optimal solution and substitute $(u_i - \overline{u}_i)$ with δu.

$$\langle F(\mathbf{x}(u_i), u_i)\rangle - \langle F(\mathbf{x}(\overline{u}_i), \overline{u}_i)\rangle \approx \frac{\partial}{\partial u_i}\langle F(\mathbf{x}(u_i), u_i)\rangle\bigg|_{\overline{u}_i} \delta u_i \geq 0 \tag{7}$$

Demanding that above inequality holds for all directions δu_i as well as $-\delta u_i$ leads us to the stronger necessary condition for optimal solutions:

$$\langle F(\mathbf{x}(u_i), u_i)\rangle - \langle F(\mathbf{x}(\overline{u}_i), \overline{u}_i)\rangle = 0 \tag{8}$$

Similarly we can write for the right side of equation (5)

$$\int_0^T \Big[\langle f(\mathbf{x}(u_i), u_i)\rangle - \langle f(\mathbf{x}(\overline{u}_i), \overline{u}_i)\rangle\Big] dt \approx \int_0^T \frac{\partial}{\partial u_i}\langle f(\mathbf{x}(u_i), u_i)\rangle\bigg|_{\overline{u}_i} (u_i - \overline{u}_i) dt = 0 \tag{9}$$

Since above equation (9) should hold for all directions i, we write:

$$\int_0^T \nabla_{\mathbf{u}} \langle f(\mathbf{x}(\overline{\mathbf{u}}), \overline{\mathbf{u}}) \rangle \circ (\mathbf{u} - \overline{\mathbf{u}}) dt = 0 \tag{10}$$

with $\circ$ denoting the elementwise multiplication (Schur product). The cost functional can be defined as a sum over one part depending on the control and not on the network state and a second part depending only on the network state:

$$\begin{gathered} \langle F(\mathbf{x}(\mathbf{u}), \mathbf{u}) \rangle = F^u(\mathbf{u}) + \langle F^x(\mathbf{x}(\mathbf{u})) \rangle, \\ \int_0^T \langle f(\mathbf{x}(\mathbf{u}), \mathbf{u}) \rangle dt = \int_0^T \Big[f^u(\mathbf{u}) + \langle f^x(\mathbf{x}(\mathbf{u})) \rangle \Big] dt \end{gathered} \tag{11}$$

Applying the chain rule to the derivative of the functional in equation (10) gives

$$\begin{gathered} \int_0^T \nabla_{\mathbf{u}} \langle f(\mathbf{x}(\overline{\mathbf{u}}), \overline{\mathbf{u}}) \rangle \circ (\mathbf{u} - \overline{\mathbf{u}}) dt = \\ \int_0^T \Big[\nabla_u f^u(\overline{\mathbf{u}}) + \langle D_u^T(\mathbf{x}(\overline{\mathbf{u}})) \nabla_{\mathbf{x}} f^x(\mathbf{x}(\overline{\mathbf{u}})) \rangle \Big] \circ (\mathbf{u} - \overline{\mathbf{u}}) dt \end{gathered} \tag{12}$$

with $D_u(\mathbf{x}(\overline{\mathbf{u}}))$ being the Jacobian matrix of the state $\mathbf{x}$ with respect to the control vector $\overline{\mathbf{u}}$.

In order to calculate the above derivative we need to solve the jacobian $D_u(\mathbf{x}(\overline{\mathbf{u}}))$. How the network state depends on the control is not trivial. Therefore, in the following the Lagrange Formalism is applied to derive an alternative expression for the gradient of the functional that is finally given in equation (24) and solve the above optimization problem.

2.2 Lagrange method

Formulation of the Lagrange function for the functional (5) that is to be minimized under the constrains given in equation (3) with boundary conditions given in equation (4)

$$\begin{gathered} \langle \mathcal{L}(\mathbf{x}(t), \mathbf{u}(t)) \rangle = \\ \langle F(\mathbf{x}, \mathbf{u}) - \int_0^T \Big[\frac{d}{dt}\mathbf{x} - \mathbf{h}(\mathbf{x}) - \sigma(\mathbf{A} \otimes \mathbf{G})\mathbf{x} - (\mathbf{B} \otimes \mathbf{K})\mathbf{u} - \eta(\mathbf{I_N} \otimes \mathbf{G})\boldsymbol{\xi} \Big]^T \boldsymbol{\phi} \; dt \rangle = \\ \langle \int_0^T \Big[f(\mathbf{x}, \mathbf{u}) - \Big(\frac{d}{dt}\mathbf{x} - \mathbf{h}(\mathbf{x}) - \sigma(\mathbf{A} \otimes \mathbf{G})\mathbf{x} - (\mathbf{B} \otimes \mathbf{K})\mathbf{u} - \eta(\mathbf{I_N} \otimes \mathbf{G})\boldsymbol{\xi} \Big)^T \boldsymbol{\phi} \Big] dt \rangle \end{gathered} \tag{13}$$

with the dN-dimensional vector of Lagrange multipliers $\boldsymbol{\phi}(t) = (\boldsymbol{\phi_1}(t), ..., \boldsymbol{\phi_N}(t))$ and $\boldsymbol{\phi_i}(t) = (\phi_{i1}(t), \ldots, \phi_{id}(t))$. For the Lagrange method an similar application as in [?] is used. Furthermore, Tröltzsch et al. use a different derivation method, leading to similar results for a system reaction-diffusion equations [**?**].

Similar to setting up the variational inequality for the optimization problem in equations (6)-(10),for the mean Lagrange function we derive following variational inequalities for the optimal control:

$$\langle \int_0^T \nabla_{\mathbf{u}} \mathcal{L}(\mathbf{x}(\overline{\mathbf{u}}), \overline{\mathbf{u}}) \circ \boldsymbol{\delta u} \ dt \rangle = 0 \tag{14}$$

and for the corresponding network state

$$\langle \int_0^T \nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}(\overline{\mathbf{u}}), \overline{\mathbf{u}}) \circ \boldsymbol{\delta x} \ dt \rangle = 0 \tag{15}$$

Here the system is being linearized around the optimal solution. Inserting the Lagrangian defined in equation (13) to the second variational inequality given in equation (15) and applying partial integration and the chain rule we obtain:

$$\begin{aligned} &\langle \nabla_x \mathcal{L}(\boldsymbol{x}(\overline{\mathbf{u}})) \circ \boldsymbol{\delta x} \rangle = \\ &\langle \int_0^T \nabla_x f(\boldsymbol{x}(\overline{\mathbf{u}}), \overline{\mathbf{u}}) \circ \boldsymbol{\delta x} \ dt - \int_0^T \nabla_x \big(\frac{d}{dt} \boldsymbol{x}(\overline{\boldsymbol{u}})\big)^T \boldsymbol{\phi} \circ \boldsymbol{\delta x} \ dt \\ &- \int_0^T \nabla_x \Big[\Big(\mathbf{h}(\boldsymbol{x}(\overline{\mathbf{u}})) - \sigma(\mathbf{A} \otimes \mathbf{G}) \boldsymbol{x}(\overline{\mathbf{u}}) - (\mathbf{B} \otimes \mathbf{K}) \overline{\mathbf{u}} - \eta(\mathbf{I_N} \otimes \mathbf{G}) \boldsymbol{\xi} \Big)^T \boldsymbol{\phi} \Big] \circ \boldsymbol{\delta x} \ dt \rangle = \\ &\langle \int_0^T \Big[\Big(\nabla_x f^x(\boldsymbol{x})\big) + \frac{\partial}{\partial t} \boldsymbol{\phi} + \big(D_x(\mathbf{h}) + \sigma(\mathbf{A} \otimes \mathbf{G}) \Big)^T \boldsymbol{\phi} \Big] \circ \boldsymbol{\delta x} \ dt \\ &- \boldsymbol{\phi}(T) \circ \boldsymbol{\delta x}(T) \rangle = \mathbf{0} \end{aligned} \tag{16}$$

We first consider the boundary term and ask for

$$\boldsymbol{\phi}(T) \circ \boldsymbol{\delta x}(T) = \mathbf{0} \tag{17}$$

in every realization. Since $\boldsymbol{\delta x}(T)$ could be unequal 0, we must have following boundary condition for the vector of Lagrange multipliers

$$\boldsymbol{\phi}(T) = \mathbf{0} \tag{18}$$

When omitting the integral, the remaining term of equation (16) gives us an dN-dimensional system of so called adjoint states $\phi(t)$ given by a linear ODE with periodic coefficients

$$-\frac{d}{dt} \boldsymbol{\phi}(t) = \big(D_x(\mathbf{h}) + \sigma(\mathbf{A} \otimes \mathbf{G}) \big)^T \boldsymbol{\phi}(t) + \nabla_x f^x(\boldsymbol{x})) \tag{19}$$

with the Jacobian matrix $D_x(\mathbf{h})$ of the local dynamics with respect to the network state $\mathbf{x}$. The system of adjoint states is obtained for every realization by solving equation (19) backwards in time with the final time conditions (18).

Inserting the Lagrangian defined in equation (13) to the first variational inequality given in equation (14) we obtain:

$$\begin{aligned} \langle \nabla_u \mathcal{L}(\boldsymbol{x}(\overline{\mathbf{u}}), \overline{\mathbf{u}}) \circ \boldsymbol{\delta u} \rangle = \\ \langle \int_0^T \nabla_u f(\boldsymbol{x}(\overline{\mathbf{u}}), \overline{\mathbf{u}}) \circ \boldsymbol{\delta u}\ dt - \int_0^T \frac{d}{dt}[D_u(\boldsymbol{x}(\overline{\mathbf{u}})]^T \boldsymbol{\phi} \circ \boldsymbol{\delta u}\ dt - \\ \int_0^T \nabla_u \Big[\Big(\mathbf{h}(\boldsymbol{x}(\overline{\mathbf{u}}) - \sigma(\mathbf{A} \otimes \mathbf{G}) \boldsymbol{x}(\overline{\mathbf{u}} - (\mathbf{B} \otimes \mathbf{K}) \overline{\mathbf{u}} \Big)^T \boldsymbol{\phi} \Big] \circ \boldsymbol{\delta u}\ dt \rangle = \mathbf{0} \end{aligned} \tag{20}$$

When inserting the variational inequality of the functional (10), the first term of equation (20) vanishes. Furthemore, applying partial integration and the chain rule, above equation simplifies to

$$\begin{aligned} \langle \nabla_u \mathcal{L}(\boldsymbol{x}(\overline{\mathbf{u}}), \overline{\mathbf{u}}) \circ \boldsymbol{\delta u} \rangle = \\ \langle \int_0^T \big(D_u(\boldsymbol{x}(\overline{\mathbf{u}})) \big)^T \frac{d}{dt} \boldsymbol{\phi} \circ \boldsymbol{\delta u}\ dt \\ + \int_0^T \Big[\big(D_u(\boldsymbol{x}(\overline{\mathbf{u}})) \big)^T \Big(\nabla_x \big(\mathbf{h}(\boldsymbol{x}(\overline{\mathbf{u}})) + \sigma(\mathbf{A} \otimes \mathbf{G}) \big)^T \phi \Big) + (\mathbf{B} \otimes \mathbf{K})^T \phi \Big] \circ \boldsymbol{\delta u} dt \rangle = \\ \langle \int_0^T \Big[\big(D_u(\boldsymbol{x}(\overline{\mathbf{u}})) \big)^T \big(\frac{d}{dt} + D_x(\mathbf{h}) + \sigma(\mathbf{A} \otimes \mathbf{G}) \big)^T \phi + (\mathbf{B} \otimes \mathbf{K})^T \phi \Big] \circ \boldsymbol{\delta u}\ dt \rangle = \mathbf{0} \end{aligned} \tag{21}$$

The boundary term of the partial integration vanishes due to the boundary conditions of $\boldsymbol{\phi}$ and $\mathbf{x}$.

Inserting the equations of the system of adjoint states (19) to equation (21) we obtain

$$\langle \int_0^T \Big(-D_u^T(\boldsymbol{x}(\overline{\mathbf{u}})) \nabla_x f^x(\boldsymbol{x}(\overline{\mathbf{u}})) + (\mathbf{B} \otimes \mathbf{K})^T \boldsymbol{\phi} \Big) \circ \boldsymbol{\delta u}\ dt \rangle = \mathbf{0} \tag{22}$$

Above equation can be used to define an alternative expression for the gradient of the functional given in equation (12). Inserting equation (22) to equation (12) leads to an expression for the gradient of the functional that depends on the adjoint states

$$\int_0^T \Big(\nabla_u f^u(\overline{\mathbf{u}}) + (\mathbf{B} \otimes \mathbf{K})^T \langle \boldsymbol{\phi} \rangle \Big) \circ \boldsymbol{\delta u}\ dt = \mathbf{0} \tag{23}$$

For the derivation of the optimal control a time-dependent gradient of the functional is needed and therefore a stronger necessary optimality condition is imposed:

$$\nabla_u f^u(\overline{\mathbf{u}}) + (\mathbf{B} \otimes \mathbf{K})^T \langle \boldsymbol{\phi} \rangle = \mathbf{0} \tag{24}$$

for all times $t \in (0, T)$. This equation is similar to the equation derived by Tröltzsch et al. [**?**] for a system of reaction-diffusion equations.

References