April 15, 2016

M.Inf.1809
Career-oriented Key Skills in a research oriented PW

Reconstructing spiking neural network connectivity in the presence of irregular dynamics

Name: **Dimitra Despoina Maoutsa**

MN: 21364423

Supervisor: Prof. Dr. Florentin Wörgötter
Prof. Dr. Marc Timme

Contents

Abstract

Knowledge of neural network synaptic connectivity is essential for understanding the underlying information processing in brain circuits. Although numerous connectivity inference methods exist that rely solely on spike train data, most of them effectively reveal only functional relationships between the activities of the neurons instead of anatomical (structural) ones. Recently we proposed a method for reconstruction of the synaptic (structural) connectivity of spiking neural networks from observed multiple spike trains without knowledge of an underlying specific model on how spikes are generated or how incoming signals are processed in a neuron. There we proposed a novel representation of the spike train recordings in high-dimensional spaces, termed event spaces. Although successful for networks with regular firing patterns, the method delivered non-optimal predictions in the presence of irregular spiking activity, especially for limited recording durations. Here, we present a short modification of the previously proposed method that effectively reveals the network synaptic connectivity unequivocally also for settings with irregular activity. In particular, we propose the identification of the directions in the event spaces that exhibit the smallest variance by employing a Minor Component Analysis on the observed events. Subsequently by considering only those minimum variance directions in the linear system of equations we may effortlessly identify the actual synaptic connectivity accurately. We demonstrate our results on simulated networks with known connectivity under various dynamical conditions and connectivity settings. The proposed modification may serve as a complement to our previously presented method and may promote the successful application of our reconstruction approach in experimental settings.

1. Introduction

Understanding how the brain operates and processes external stimuli comprises an overarching endevour that has engaged researchers for several decades. The brain may be construed as a large network where the different brain areas serve as nodes, while the inter-regional pathways connecting them may be considered as links. In turn each brain area may be decomposed into smaller networks where for these networks the individual neurons are considered nodes and the synapses projecting between them are viewed as links of the network [1, 2].

Commonly, networks are studied theoretically by decomposing them into their fundamental parts and interactions, and consequently investigating the behaviour of each constituent independently, as if it was isolated. By doing so one intends to identify the intrinsic dynamics of each constituent unit and define afterwards mathematical models that capture its function. Subsequently, one re-assembles all the parts together by coupling the individual models, formulating thereby a larger mathematical model that is supposed to capture the function and the key features of the network at hand [3]. Such approaches are commonly termed *forward or modelisation approaches* [4] and in the field of neuroscience have contributed to the understanding of how different activity patterns (e.g. synchronous/asynchronous or regular/irregular) emerge in networks of pulse coupled neurons [5, 6], or even how established anatomical connections in brain networks contribute to the manifestation of functional relationships [1].

However, such forward modeling of a network may be a rather complicated pursuit due to the very complexity of its structure or due to the complexity of the individual nodal dynamics [3]. For instance, the human brain comprises approximately 10^{12} neurons tightly interconnected through numerous synapses (approximately 10^4 for each neuron). These numbers render a forward approach intractable. In those settings *inverse approaches*, where one initiates the analysis from the observed measurements, may particularly help to gain insight on the function of the network at hand. However, inverse problems are particularly demanding; while in forward problems there is always a unique mapping from the model parameters to the measurements, inverse approaches may yield instead of a unique solution, a family of solutions that matches the observed measurements [4].

In particular the inference of neural network structure from multiple spike train recordings constitutes a rather challenging inverse problem [7]. The fundamental reason for the increased hardness of this problem may be attributed to the constrains that the current monitoring methods impose. In fact, with the current recording methods one may simultaneously observe the activity of an ensemble of neurons only *extracellularly*. That means that one may only observe when the natural state variable of each neuron, the membrane voltage, crosses the voltage threshold emitting thereby a spike. Therefore for the reconstruction of such networks one has to mitigate the limitation of partial access to the state variable of each neuron.

The existing methods that tackle the problem of the connectivity inference in neural networks may be classified in two broad categories: the model dependent approaches and the model independent ones. The methods that fall into the former category rely mainly on assumptions on some underlying neuronal model with either known [8] or unknown [9] physiological parameters and by employing some optimisation procedure, infer the actual synaptic structure of the network. On the other hand, the model independent frameworks employ mostly statistical [10] or information theoretic [11] measures to resolve the network structure. However these approaches result in predictions of functional connections (i.e. correlation in the activities of the neurons) instead structural ones.

In [12] we introduced a novel model-independent approach that, in contrast to the existent model-independent frameworks, succeeds to identify structural connections among the neurons. Although successful for networks with regular firing patterns, the method delivered non-optimal predictions in the presence of irregular spiking activity, especially for limited recording durations. Here, we present a short modification of the previously proposed method that effectively reveals the network synaptic connectivity unequivocally also for settings with irregular activity.

This report is structured as follows: In the next section 2 we present briefly the reconstruction method already introduced in [12], followed by a brief description of the mathematical tools we used to modify the reconstruction approach. In the succeeding section 3 we describe how we may modify the reconstruction framework to deliver accurate results also in the presence of irregular activity. Subsequently we present the efficiency of our approach in different network settings and finally in the last section 4 we offer a brief discussion.

2. Theoretical Background

2.1. Model-independent reconstruction of neural network connectivity

Let us assume that we have access to the extracellular recordings (multiple spike trains) of a network comprising N neurons, but we have no prior information regarding the established synaptic connections between the neurons and the underlying neuronal dynamics they follow. Would we be able to infer the network connectivity from the spike recordings alone under these conditions? Indeed in [12] we presented a framework that may help reveal the network interactions under these aforementioned conditions. We refer the reader to [12] for a detailed presentation of the method, while here we present the main outline of the approach.

Assuming a network of N neurons with unknown synaptic connections and underlying dynamics, we have only access to the extracellularly recorded activity of the neurons. Considering the reconstruction of the incoming connectivity of a unit i, we may define the following quantities:

- The m-th *inter-spike interval* of neuron i, denoted by $\Delta T^i_m \in \mathbb{R}$, is the duration of time between two consecutive firings of neuron i, and in particular between the m-th and $(m+1)$-th firing of i.

- The k-th *cross-spike interval* of neuron j with respect to the $m-th$ inter-spike interval of neuron i, denoted by $w^i_{j,k,m} \in \mathbb{R}$, is the duration between the onset of the m-th spike of neuron i and the k-th successive spike emitted from neuron j that occurs within the m-th inter-spike interval of i.

Although we do not have any prior information regarding the dynamics of neuron i, we may in general express the inter-spike intervals of neuron i as a function h_i that depends on the times that i receives its inputs from its presynaptic neurons (i.e. the neurons that project direct synapses to i). To that end, we may quantify the spike reception times with the cross-spike intervals of the presynaptic neurons of i with respect to the emitted spikes from neuron i. However, since the actual subset of the network that is directly connected to neuron i is unknown, we may consider the cross-spike intervals of the entire network with respect to i and define cross-spike intervals $w^i_{j,k,m}$ with $j \in \{1, 2, \ldots, N\} \setminus \{i\}$ for every spike that was emitted during the inter-spike intervals of i [Fig. 1(left)]. Therefore the m-th inter-spike interval of neuron i, ΔT^i_m, may be expressed as:

$$\Delta T^i_m = h_i(\Lambda^i W^i_m), \tag{1}$$

where the underlying neuronal dynamics are captured by the unknown function $h_i : \mathbb{R}^{N \times K^i} \to \mathbb{R}$ and Λ^i denotes the *explicit dependency matrix* of i (c.f [13]), a

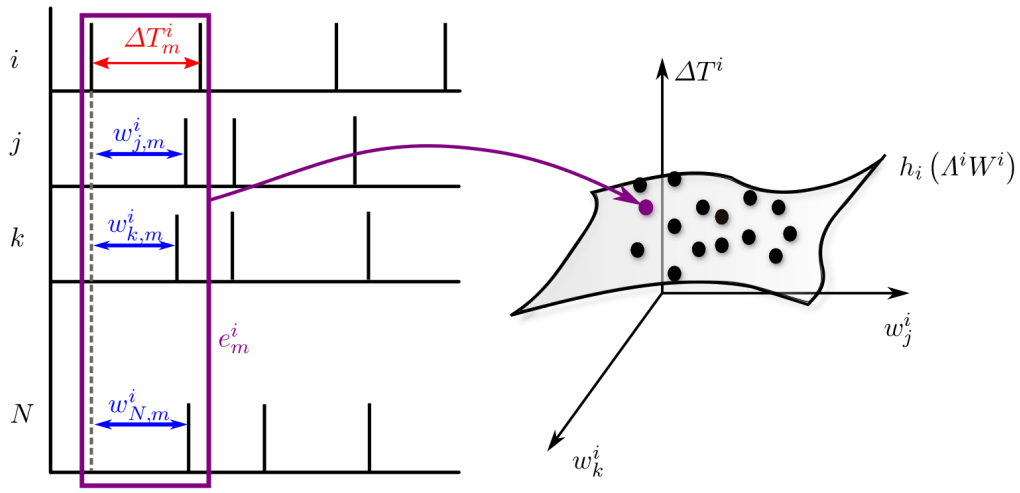

Figure 1: With the event space representation the network activity (left) **may be mapped onto discrete points in the event space** (right)**.** The events e^i_m constitute samplings of the (unknown) function h_i and are therefore expected to arise as a (possibly) non-linear manifold in the event space.

diagonal matrix that reflects the incoming connectivity of neuron i. W^i_m denotes the matrix of cross-intervals with respect to i and is structured as:

$$W^i_m = \begin{bmatrix} w^i_{1,1,m} & w^i_{1,2,m} & \cdots & w^i_{1,K^i,m} \\ w^i_{2,1,m} & w^i_{2,2,m} & \cdots & w^i_{2,K^i,m} \\ \vdots & \vdots & \ddots & \vdots \\ w^i_{N,1,m} & w^i_{N,2,m} & \cdots & w_{N,K^i,m} \end{bmatrix}. \tag{2}$$

Considering that the (unknown) function h_i maps the inputs of neuron i from the entire network to the observed durations of the inter-spike intervals of i, we may define a mapping that actually captures this relationship. Therefore, we may define an event as the concatenation of the cross-intervals of the entire network with respect to i with the associated inter-spike interval of neuron i, i.e.:

$$e^i_m := [vec\left(W^i_m\right), \Delta T^i_m]^\top \in \mathbb{R}^{(N\,K^i+1)}, \tag{3}$$

where the operator $vec : \mathbb{R}^{S\times R} \rightarrow \mathbb{R}^{S\,R}$ denotes the vectorization of a matrix [14], and subsequently define the *event space* as the space that is spanned from these events [Fig. 1].

Considering now that h_i is in general unknown, we may locally approximate this (probably) complicated function with a linear one that would be much simpler to handle. For that reason, we may select a reference event e^i_r in the event space,

that is closest to all other generated events

$$e_r^i := \min_{e_s^i} \sum_m \|e_m^i - e_s^i\|_2^2 \tag{4}$$

where $r, s \in \{1, 2, ..., M\}$, and approximate the function h_i locally around e_r^i as:

$$\Delta T_m^i = \Delta T_r^i + tr\left(\left(\frac{\partial h_i}{\partial W^i}\right)^\top \Lambda^i \left[W_m^i - W_r^i\right]\right), \tag{5}$$

where $tr(\cdot)$ denotes the trace operator [14] and $\left(\frac{\partial h_i}{\partial W^i}\right) \in \mathbb{R}^{N \times K^i}$ stands for the matrix derivative $\left(\frac{\partial h_i}{\partial W^i}\right) = \left(\frac{\partial h_i}{\partial W^i}\right)(\Lambda^i W_r^i)$.

Subsequently after some matrix operations described explicitly in [12] yield :

$$\Delta T_m^i = \Delta T_r^i + \sum_{k=1}^{K^i} \boldsymbol{\nabla} h_{i,k} \Lambda^i \left[\boldsymbol{w}_{k,m}^i - \boldsymbol{w}_{k,r}^i\right]. \tag{6}$$

By rewriting Eq. (6) in a matrix form we may obtain:

$$\Delta T_m^i = \Delta T_r^i + \boldsymbol{g}_i \Gamma^i \boldsymbol{y}_m^i, \tag{7}$$

where $\boldsymbol{g}_i := \left[\boldsymbol{\nabla} h_{i,1}, \boldsymbol{\nabla} h_{i,2}, \cdots, \boldsymbol{\nabla} h_{i,K^i}\right] \in \mathbb{R}^{NK^i}$ contains the gradients of h_i, $\Gamma^i \in \mathbb{R}^{N\,K^i \times N\,K^i}$ denotes a block diagonal matrix with Λ^i on the main diagonal and zeros elsewhere:

$$\Gamma^i := \begin{bmatrix} \Lambda^i & 0_{N\times N} & \cdots & 0_{N\times N} \\ 0_{N\times N} & \Lambda^i & \cdots & 0_{N\times N} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{N\times N} & 0_{N\times N} & \cdots & \Lambda^i \end{bmatrix}$$

and $\boldsymbol{y}_m^i := \left[\boldsymbol{w}_{1,m}^i - \boldsymbol{w}_{1,r}^i, \boldsymbol{w}_{2,m}^i - \boldsymbol{w}_{2,r}^i, \cdots, \boldsymbol{w}_{K^i,m}^i - \boldsymbol{w}_{K^i,r}^i\right]^\top \in \mathbb{R}^{N\,K^i}$.

Subsequently, by defining a particular distance metric, i.e. the Euclidean distance, we may select the M closest events (with respect to the defined distance metric) to the reference event e_r^i and construct a linear system of equations for those selected events

$$\boldsymbol{d}^i = \boldsymbol{g}_i \Gamma^i Y^i, \tag{8}$$

with the vector $\boldsymbol{d}^i := [\Delta T_1^i - \Delta T_r^i, \Delta T_2^i - \Delta T_r^i, \cdots, \Delta T_M^i - \Delta T_r^i] \in \mathbb{R}^M$ denoting the inter-spike interval differences between the selected and the reference event, and $Y^i := [\boldsymbol{y}_1^i, \boldsymbol{y}_2^i, \ldots, \boldsymbol{y}_M^i] \in \mathbb{R}^{N\,K^i \times M}$ containing the related cross-spike interval matrices.

The least squares solution of (Eq. 8) may be obtained by:

$$\boldsymbol{g}_i \Gamma^i = \boldsymbol{d}^i Y^{i\dagger}, \tag{9}$$

where $Y^{i\dagger}$ denotes the Moore-Penrose pseudo-inverse of Y^i , namely $Y^{i\dagger} := Y^{i\top}\left(Y^i Y^{i\top}\right)^{-1}$. Finally, to assess the strength of the synaptic couplings we defined the *connectivity characterisers* $\boldsymbol{a}^i \in \mathbb{R}^N$ as:

$$\boldsymbol{a}^i = I_{K^i} \boldsymbol{g}^i \Gamma^i, \tag{10}$$

where I_{K^i} denotes the matrix consisting of K^i replicas of the $N \times N$ identity matrix

$$I_{K^i} = [\mathbb{1}_{N\times N}, \mathbb{1}_{N\times N}, \cdots, \mathbb{1}_{N\times N}] \in \mathbb{R}^{N\times N\,K^i}.$$

2.2. From Singular Value Decomposition to Minor Component Analysis

2.2.1. Singular Value Decomposition

The Singular Value Decomposition (SVD) of a matrix $A \in \mathbb{R}^{n\times d}$ consists a framework that factorises the matrix A into the product

$$A = U \Sigma V^\top, \tag{11}$$

where the matrices $U \in \mathbb{R}^{n\times n}$ and $V \in \mathbb{R}^{d\times d}$ are orthonormal bases with $U^\top U = I_{n\times n}$ and $V^\top V = I_{d\times d}$, while the matrix $\Sigma \in \mathbb{R}^{n\times d}$ is a diagonal matrix, with the diagonal entries σ_i arranged with decreasing order of magnitude being commonly referred to as *singular values*. The vectors forming the columns of U and V are usually termed *left* and *right singular vectors*.

Assuming that the entries of A represent n points in a d-dimensional space, then the SVD yields an orthonormal basis composed by the column vectors of V (the right singular vectors), $V = [\mathbf{v}_1 \mathbf{v}_2 \dots \mathbf{v}_d]$, that solves the problem of *best least squares fit*, i.e. the singular vectors are aligned in such way in the d-dimensional space that the distances of the points to the basis vectors are minimised [15].

Computing the SVD of matrix A equals to performing an eigenvalue decomposition for the positive semi-definite matrices $AA^\top$ and $A^\top A$, with the eigenvectors of $A^\top A$ being the right singular vectors (columns of V) and the eigenvectors of $AA^\top$ the left singular vectors (the columns of U). The eigenvalues of both matrices $AA^\top$ and $A^\top A$ are the same and are equal to the square of the singular values contained in Σ.

2.2.2. Minor Component Analysis

The counterpart of the widely applied statistical method Principal Component Analysis [16] is the mostly unknown method *Minor Component Analysis*(MCA). Here one intends to identify the directions in the sample data space that exhibit the lowest variance. These directions essentially correspond to the eigenvectors that are associated to the smallest eigenvalues of the covariance matrix of the data matrix A.

One may employ SVD to perform MCA. Essentially the covariance matrix C of the data described by the matrix A (if we assume that A is centered (i.e. has zero column means)) can be diagonalised as:

$$C = VLV^\top \tag{12}$$

where V contains in its columns the eigenvectors of C, while the matrix L is diagonal matrix containing the eigenvalues λ_i of C in decreasing order. The eigenvectors (columns of V) are commonly termed *principal axes or directions*.

Returning to singular value decomposition of A (Eq. 11) we may rewrite C as:

$$C = \frac{1}{n-1} V\Sigma U^\top U\Sigma V^\top = V\frac{\Sigma^2}{n-1}V^\top. \tag{13}$$

By comparing Eq. 12 with Eq. 13 we may identify that the right singular vectors are essentially the principal axes (the eigenvectors of the covariance matrix) and that the eigenvalues of the covariance matrix are related to the singular values via the relationship:

$$\lambda_i = \frac{\sigma_i^2}{n-1}. \tag{14}$$

With that we have a direct relationship between the SVD and the principal components of the dataset represented by A. Subsequently if we want to identify the directions in the data space that exhibit the least variance, we may just identify the left singular vectors that are associated with the singular values of the smallest (non-zero) magnitude.

3. Results

3.1. Studying the spread of events in the event spaces

In [12] we presented and discussed the limitation of the proposed reconstruction method to reveal unequivocally weak synaptic connections in the presence of irregular dynamics (Coefficient of Variation $CV > 0.5$) for networks comprising both excitatory and inhibitory synapses. Although the (stronger) inhibitory synapses may be effortlessly recovered in the presence of irregular dynamics, the connectivity characterisers of the (weaker) excitatory connections were hardly separated from the characterisers of the absent synapses. However our method may be not limited to deliver accurate results only in the regular dynamics regime. Here, we present a small modification of our framework that results in accurate predictions also for networks with more irregularly firing neurons and serves as a complementary solution instead of a replacement of the previous approach.

As already discussed in [12] the variability of neural firing in networks exhibiting irregular dynamics generates events that span a more extended volume in the event space compared to networks that exhibit regular dynamics. For the reconstruction of the connectivity of those networks we proposed the employment in the linear system of equations (Eq. 8) only of those points that lay in the vicinity of the reference event (Eq 4). The underlying rationale for this selection strategy was that the linear approximation of the function h_i (Eq. 6) will be valid only in the proximity of the reference event, especially in the case of an underlying non linear function.

The selection criterion that we employed was based on the Euclidean distance of the events in the event space; the events that exhibited smaller Euclidean distance to the reference event were included in the linear system Eq. 8. However the high dimensionality of the event spaces may render the employment of the Euclidean distance imprecise. In the literature it has been widely argued that the problem of the selection of the nearest neighbors to a query point becomes ill defined already for spaces of dimensionality as low as $d = 15$ [17, 18].

Moreover, for the determination of the nearest events to the reference point we consider the distance of the events taking into account also the dimensions that are associated with the cross-spike intervals from the un-connected (and therefore non relevant) neurons to unit i, since we do not have prior knowledge of the subset of the neurons that are actually connected to i. Therefore, events that correspond to inter-spike intervals of i, during which i received identical inputs from its pre-synaptic neurons may be considered as distant because of the influence of the cross-spike intervals arising from the un-connected neurons. This caveat is particularly relevant for networks with irregular activity and sparse connectivities, where the majority of the dimensions in the event space may be construed as noise.

Subsequently, we observe the variance of the arising cross-spike intervals in

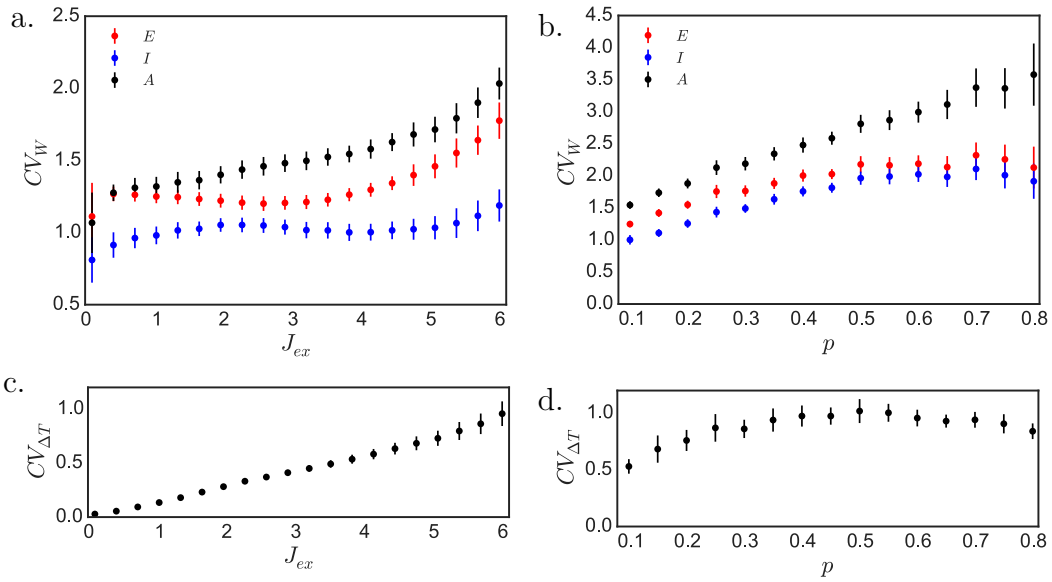

Figure 2: Cross-spike intervals of un-connected neurons exhibit consistently larger variance than cross-spike intervals of pre-synaptic neurons as the synaptic strength J_{ex} and the probability of connection increase. Coefficients of variation of cross-spike intervals CV_W arising from connected excitatory (E), connected inhibitory (I) and unconnected (A) neurons calculated from the first 2000 events monitored for each unit against **(a.)** connection strength and probability of connection $p = 0.1$, and against **(b.)** probability of connection in the network for $J_{ex} = 3.5$. **(c.,d.)** Coefficient of variation of inter-spike intervals $CV_{\Delta T}$ against connection strength and probability of connection. (Error bars denote standard deviation)

networks with known connectivities under different dynamical conditions. To that end, we simulate a network of $N = 100$ Leaky Integrate and Fire neurons (c.f. Appendix B) (with 80% excitatory and 20% inhibitory neurons) projecting synapses with probability of connection of $p = 0.1$ employing normalised realistic physiological neuronal parameters ($I_i^{ext} \in [1.4, 1.6]\ pA$, $\tau_{ref} = 2.0\ ms$, $\tau_i^m = 20.0\ ms$, $V_i^{th} = 20\ mV$). By varying the strength of the synapses (yet maintaining the ratio between excitatory and inhibitory strengths $J_{inh} = -5J_{ex}$) we can effectively modify the irregularity of firing in the network. For each network realisation we construct event spaces from the first $M = 2000$ inter-spike intervals generated from each neuron separately and consequently observe the Coefficients of Variation of the cross-spike intervals of the connected excitatory neurons, CV_{W_E}, of the connected inhibitory, CV_{W_I}, and of the unconnected neurons, CV_{W_A}.

Indeed, for weak synaptic strengths, where the CV of the inter-spike intervals is relatively small ($CV_{\Delta T} \leq 0.1$) [Fig. 2(c.)] , the CV of the cross-spike in-

tervals (CV_W) arising from unconnected and excitatory neurons are comparable [Fig. 2(a.)]. However, as the synaptic strength increases and therefore also the irregularity of firing elevates [Fig. 2(c.)], the variation of the cross-spike intervals defined with respect to connected units is consistently smaller than the variation of the cross-spike intervals of the unconnected neurons (blue and red marks in Fig. 2(a.) for connected against black marks for unconnected units).

3.2. Inference of synaptic connectivity with Minor Component Analysis

What if we could mitigate the influence of the highly variable cross-spike intervals of the unconnected neurons and try to identify the connectivity considering only the less variable dimensions of the event spaces? To that end, we may construct the linear system of equations (Eq. 8) employing the entire ensemble of monitored events and subsequently perform a Minor Component Analysis on the resulting matrix Y^i. That matrix is essentially composed of the cross-spike intervals differences of all the recorded events with the reference event, which geometrically corresponds to a simple translation of the entire cluster of events. Since the reference event is selected as the event with the median distance to the entire event ensemble, this translation may correspond to an approximate centering of the event manifold.

By performing a Singular Value Decomposition on the matrix Y^i (Eq. 8), we may obtain:

$$Y^i = U\Sigma V^\top, \tag{15}$$

where $U \in \mathbb{R}^{NK^i \times NK^i}$, $\Sigma \in \mathbb{R}^{NK^i \times M}$ and $V \in \mathbb{R}^{M \times M}$. Essentially the rows of matrix $U = [\mathbf{u}_1 \mathbf{u}_2 \dots \mathbf{u}_{NK^i}]^\top$ form an orthonormal basis that spans the range space of Y^i, where the vectors $\mathbf{u}_x$ are ordered with non-increasing variance, captured by the magnitude of the associated singular values, $\Sigma = diag[\sigma_1, \sigma_2, \dots \sigma_r, \mathbf{0}_{(NK^i - r)}]$, where $r = rank(Y^i)$. At the same time, the columns $\mathbf{q}_x$ of the product $\Sigma V^\top = [\mathbf{q}_1 \mathbf{q}_2 \dots \mathbf{q}_M]$ reflect the magnitude of the projection of each of the M events onto each direction of the orthonormal basis U.

Consequently, we propose the selection of s directions that exhibit the smallest variance in the event space and solve the system of equations (Eq. 8) by considering only those selected directions. For that purpose, we may eliminate the first $r - s$ singular values and obtain:

$$\tilde{\Sigma} = diag[\mathbf{0}_{r-s}, \sigma_{-s}, \sigma_{-s+1}, \dots, \sigma_{-1}]. \tag{16}$$

Consequently we obtain the modified matrix by $\tilde{Y}^i = U\tilde{\Sigma}V^\top$. Eventually, we may calculate the Moore-Penrose pseudo-inverse of $\tilde{Y}^i$, $\tilde{Y}^{i\dagger}$, to obtain an approximate

solution of the system of equations (Eq. 8):

$$\boldsymbol{g}_i \Gamma^i = \boldsymbol{d}^i \tilde{Y}^{i\dagger}. \tag{17}$$

Nevertheless, how many dimensions with minimum variance should we select? In other words, how could we define the optimal value of s? Systematical studies revealed that an effective strategy for the selection of s may consider the differences in magnitudes of consecutive singular values. In fact, we propose to identify the pair of consecutive singular values, σ_{r-s-1} and σ_{r-s}, that exhibits the larger magnitude difference and subsequently preserve the singular values that have magnitude equal or less than σ_{r-s}. In practice, we select s as:

$$\max_s \ (\sigma_{r-s-1} - \sigma_{r-s}), \quad s \in \{1, 2, \ldots, r-1\}. \tag{18}$$

Yet does this modified approach work? In order to assess the capabilities of this altered reconstruction strategy, we simulate a network of $N = 100$ neurons (80% excitatory, 20% inhibitory) with probability of connection $p = 0.1$ and synaptic strength for the excitatory synapses $J_{ex} = 3.5$ mV, while for the inhibitory ones $J_{inh} = -5J_{ex}$ with the same physiological parameters as mentioned previously. As a matter of fact, the network exhibits irregular activity, with $CV_{\Delta T} = 0.51$.

Subsequently, we reconstruct the network by employing both the approach presented in Section 2.1 (mentioned as *regular approach* hereafter) employing different extents of local samplings and the modified strategy involving the Minor Component Analysis. Indeed, as already presented in [12], the regular approach may hardly distinguish between the absent and the excitatory connections, independently of the number of employed equations in the linear system [Fig. 3(a.,b.,c.),Fig. 4(a.,b.)]. In contrast, the modified MCA approach effortlessly reveals the actual connectivity for both excitatory and inhibitory connections [Fig. 3(d.),Fig. 4(c.,d.)]. by considering only *two* singular vectors and the entire ensemble of recorded events. Indeed, by reconstructing the entire network and subsequently calculating the error between the actual and the inferred connections we find that the altered method errs only for very few connections by predicting excitatory synapses between unconnected pairs [Fig. 4(d.)]. In contrast, the regular approach successfully identifies only the inhibitory connections of the network [Fig. 4(b.)]. Thus, the updated reconstruction method effectively succeeds to identify the synaptic structure of the network.

To investigate the issue further, we simulated networks with the same settings as mentioned previously for three different recording durations ($T = \{80, 160, 240\}$s) for different dynamical conditions by varying the synaptic strength of the connections. For the reconstruction of the connectivities we employed both the regular and the presently proposed MCA method and observed the accuracy of the predictions as captured by the AUC score. For the regular approach we employed in the solution the $M = 2500$ closest events to the selected reference event.

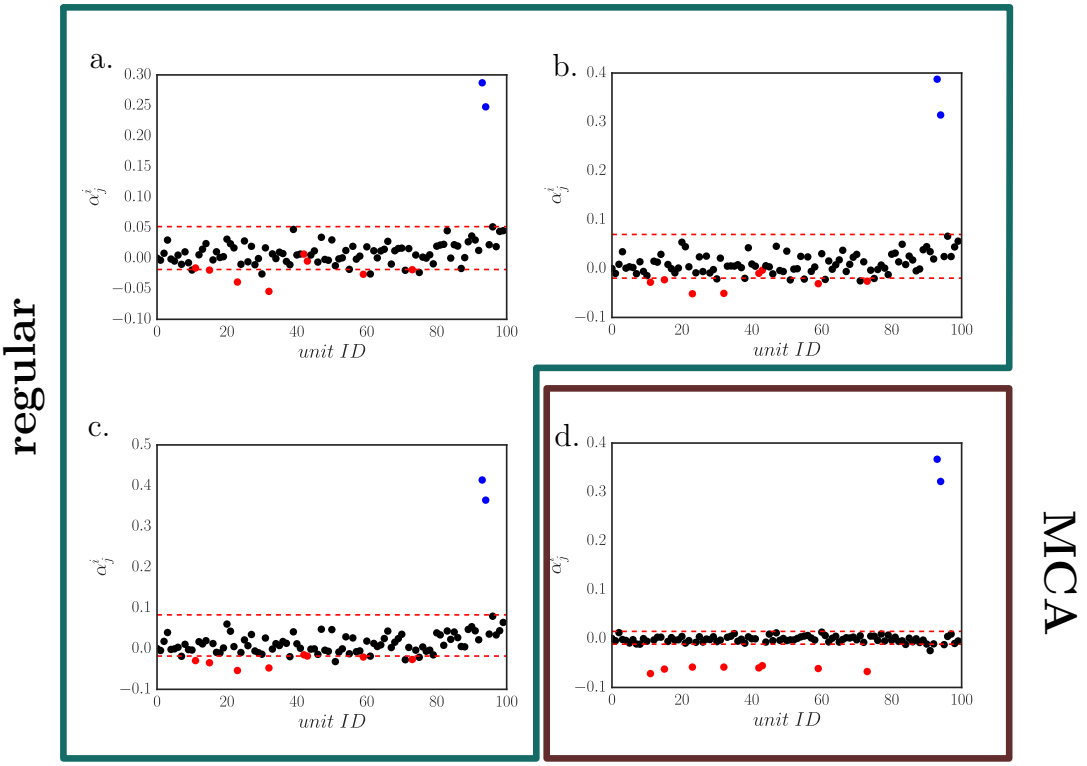

Figure 3: Minor component reveals accurately the connectivity for irregular dynamics. Reconstruction of incoming connectivity of a single neuron embedded in a network comprising $N = 100$ LIF neurons (80 %E - 20 %I) in the irregular activity ($J = -3.5\ mV$) regime simulated for $T = 160\ s$. **(a.,b.,c.)** Predicted connectivity characterisers resulting from local samplings of $M = \{2000, 2500, 3000\}$ events and **(d.)** from employment of all the events ($M^T = 7055$) considering only the two minor components. Blue and black dots indicate existing and absent connections, while the red line denotes the discrimination threshold as obtained by the Otsu method [19]. [c.f. Fig. 3.8, p.37 in [12]]

Indeed, as already observed in [12], as the synaptic connections become stronger and thereby the irregularity of firing increases, the regular approach starts to deliver non-accurate predictions resulting in AUC scores below 0.95 [Fig. 5(a.)]. Although the extension of the observation time induces a denser sampling of the event space and therefore results in more precise predictions, still as the CV_{ISI} exceeds 0.5 (approximately for $J_{ex} = 3.5$ mV, c.f. Fig. 2(c.)) the regular method fails to deliver accurate results. In contrast, the modified approach employing the MCA consistently succeeds to capture the actual synaptic structure of the network for synaptic strengths $J_{ex} < 6$ mV.

Surprisingly we notice that the modified method results in inaccurate connec-

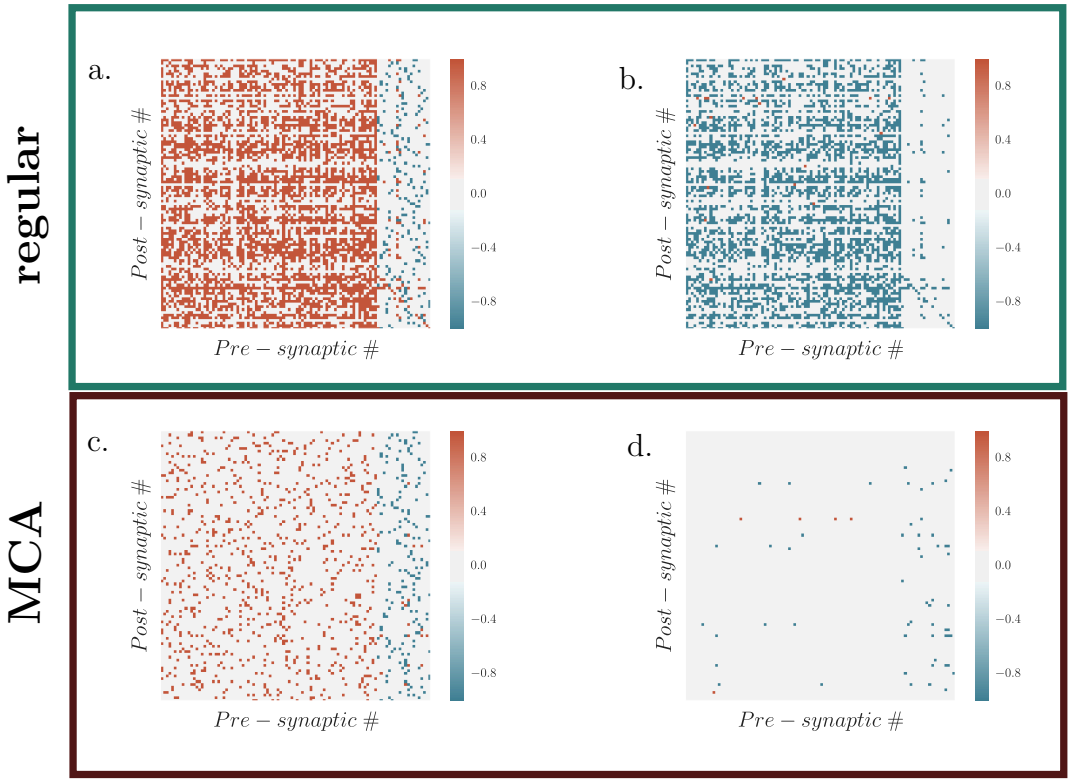

Figure 4: Reconstruction considering only the Minor Components successfully identifies the network structure, in contrast to the regular approach. (a.,b.) Resulting reconstruction and reconstruction error as delivered from the regular reconstruction method employing $M = 2500$ events in the linear system, and **(c.,d.)** resulting reconstruction and reconstruction error obtained from the the MCA modified approach for a network comprising $N = 100$ LIF neurons (80 % E - 20 % I) with probability of connection $p = 0.1$ and synaptic strength $J_{ex} = 3.5$, $J_{inh} = -5\, J_{ex}$ simulated for $T = 160\ s$.

tivity estimates for networks with very weak synapses ($J_{ex} \leq 1$), where the neurons fire almost periodically. As already demonstrated in Fig. 2(a.), in this setting the spread of the events along the dimensions that correspond to connected and unconnected neurons is almost indistinguishable and therefore a MCA-based reconstruction may fail per se. In these settings the connectivity may be inferred with the regular method.

Subsequently we study the accuracy of the predictions of the modified reconstruction method for networks with different connection densities. We simulate a network of $N = 100$ LIF neurons with $J_{ex} = 3.5$ and $J_{inh} = -5J_{ex}$ for $T = 160$s varying the probability of connection p. Interestingly we find that our approach delivers accurate results only for sparse networks (with $p < 0.5$) [Fig. 5(b.)]. However these results were intuitively expected since with increasing probability

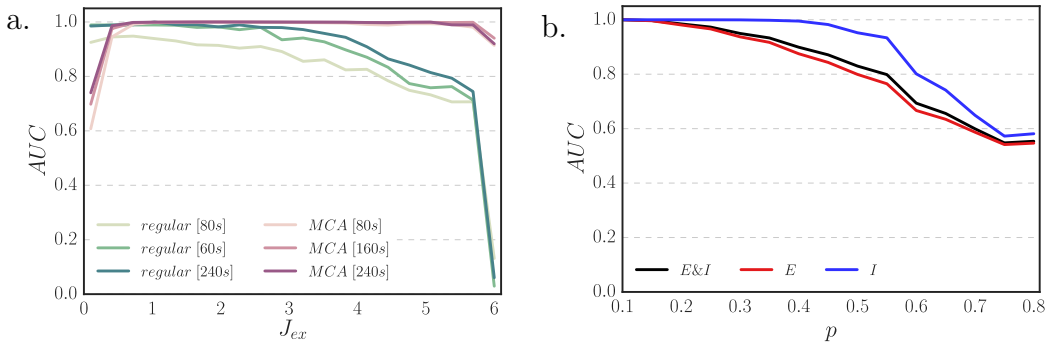

Figure 5: Modified reconstruction approach employing MCA effectively delivers accurate connectivity estimates for different connectivity strengths and sparse networks. Reconstruction accuracy (AUC score) of incoming connectivity of a network comprising $N = 100$ LIF neurons (80%E-20%I) against **(a.)** connection strength simulated for $T = \{80, 160, 240\}$ s and reconstructed employing the regular and the MCA modified approaches, and against **(b.)** probability of connection employing the MCA modified approach.

of connection also the number of relevant dimensions that we should consider increases. Therefore a reconstruction strategy that establishes a solution considering only a small subspace of the initial event space may fail to identify the exact synaptic structure. Still, the dimensions that were identified as relevant according to the criterion proposed in Eq. 18 were also for dense networks below 10. Therefore, an updated strategy for the identification of the relevant dimensions may replace the proposed criterion (Eq. 18) and actually result in accurate predictions.

4. Conclusion

Here we presented a modification of the model-independent reconstruction method introduced recently in [12]. That method, although effectively delivered accurate connectivity estimates for networks with regular activity patterns, failed to precisely infer the network structure when the constituent neurons fired irregularly. Here, we suggested a small adjustment that improves the inference results in settings with irregularly firing neurons. Specifically, we proposed the employment of Minor Component Analysis to determine the directions in the emerging event spaces along which the recorded events exhibit the minimum variance. By measuring the variance that arises in the event spaces along each dimension associated with either a connected or an unconnected neuron, we demonstrated that essentially in the presence of irregular dynamics the dimensions that correspond to connected neurons exhibit less variance. Based on this finding we employed in the linear system of equations only the directions of less variance and subsequently successfully recovered the network structure. We demonstrated the results of the updated approach for networks simulated under various dynamical conditions and different connectivity settings, indicating thereby also the limitations of this modified method.

Essentially, the shortcoming of the method presented in [12] lies partly on the selection strategy of the events that we employ in the linear system of equations that eventually yields the connectivity estimates. There we proposed to measure the proximity of events to the center of the approximation by measuring the Euclidean distance between the selected and the reference events. However the Euclidean distance in high dimensional spaces may be ill-defined [17, 18]. In the regularly firing networks, where the underlying function we approximate is presumably linear (c.f. [12]), the precise selection of events does not actually influence the resulting predictions. However, in the presence of irregular dynamics one might precisely select the events that will be employed in the formulation of the linear system, since in this setting the underlying approximated function is probably non-linear, rendering the linear approximation invalid for events that do not lie in the vicinity of the reference one. For these settings the modified approach delivers consistently accurate results for sparse network connectivities.

However, this updated framework may not be appropriate for every setting with irregularly firing neurons. Indeed, we demonstrated that for networks with dense connectivities the modified method performs slightly better than random guessing. These results were actually partially expected, since the elimination of the majority of dimensions in the event space that our method entails, hinders the identification of the contribution of the majority of the neurons in the network to the observed inter-spike intervals. Nevertheless, still the results obtained with this updated approach were comparable with the results delivered by the regular method.

Overall this complementary reconstruction framework by modifying the exist-

ing method to deliver more accurate results under biologically realistic conditions (i.e. irregularly firing neurons), may promote the application of our model-independent reconstruction method in actual experimental settings.

A. References

References

[1] Gustavo Deco, Viktor K Jirsa, and Anthony R McIntosh. Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nature Reviews Neuroscience*, 12(1):43–56, 2011.

[2] Olaf Sporns. *Networks of the Brain*. MIT press, 2011.

[3] Steven H Strogatz. Exploring complex networks. *Nature*, 410(6825):268–276, 2001.

[4] Albert Tarantola. *Inverse problem theory and methods for model parameter estimation*. siam, 2005.

[5] Nicolas Brunel and Vincent Hakim. Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural computation*, 11(7):1621–1671, 1999.

[6] Nicolas Brunel. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of computational neuroscience*, 8(3):183–208, 2000.

[7] Marc Timme and Jose Casadiego. Revealing networks from dynamics: an introduction. *Journal of Physics A: Mathematical and Theoretical*, 47(34):343001, 2014.

[8] Frank Van Bussel, Birgit Kriener, and Marc Timme. Inferring synaptic connectivity from spatio-temporal spike patterns. *Frontiers in computational neuroscience*, 5(3), 2011.

[9] Valeri A Makarov, Fivos Panetsos, and Oscar de Feo. A method for determining neural connectivity and inferring the underlying network dynamics using extracellular spike recordings. *Journal of neuroscience methods*, 144(2):265–279, 2005.

[10] Felipe Gerhard, Tilman Kispersky, Gabrielle J Gutierrez, Eve Marder, Mark Kramer, and Uri Eden. Successful reconstruction of a physiological circuit with known connectivity from spiking activity alone. *PLoS Comput Biol*, 9(7):e1003138, 2013.

[11] Olav Stetter, Demian Battaglia, Jordi Soriano, and Theo Geisel. Model-free reconstruction of excitatory neuronal connectivity from calcium imaging signals. *PLoS Comput Biol*, 8(8):e1002653, 2012.

[12] Dimitra Despoina Maoutsa. *Model-free reconstruction of synaptic connectivity from spike trains*. Master's thesis, Georg-August University of Goettignen, 2016.

[13] Jose Casadiego and Marc Timme. Network dynamics as an inverse problem. In *Mathematical Technology of Networks*, pages 39–48. Springer, 2015.

[14] Karim M Abadir and Jan R Magnus. *Matrix algebra*, volume 1. Cambridge University Press, 2005.

[15] Sudipto Banerjee and Anindya Roy. *Linear algebra and matrix analysis for statistics*. CRC Press, 2014.

[16] Lei Xu, Erkki Oja, and Ching Y Suen. Modified hebbian learning for curve and surface fitting. *Neural Networks*, 5(3):441–457, 1992.

[17] Kevin Beyer, Jonathan Goldstein, Raghu Ramakrishnan, and Uri Shaft. When is ñearest neighbormeaningful? In *Database theory ICDT'99*, pages 217–235. Springer, 1999.

[18] Charu C Aggarwal, Alexander Hinneburg, and Daniel A Keim. *On the surprising behavior of distance metrics in high dimensional space*. Springer, 2001.

[19] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296):23–27, 1975.

B. Leaky Integrate and Fire neuron

For the description of the state of the an isolated Leaky Integrate and Fire neuron (LIF) neuron through the trajectory of the membrane voltage $V_i(t)$, we may consider the neuronal membrane as a capacitor of capacitance C_i getting charged by the incoming external ionic currents $I_i(t)$:

$$C_i \frac{dV_i(t)}{dt} = I_i(t). \tag{19}$$

Furthermore, accounting for the membrane leakage due to ionic diffusion, we may attach a resistance R_i in parallel to the capacitor, yielding:

$$C_i \frac{dV_i(t)}{dt} = I_i(t) - \frac{V_i(t)}{R_i}. \tag{20}$$

Particularly, if we consider a constant external driving current $I_i(t) := I_i^{ext}$ and incorporate the membrane time constant $\tau_i^m := R_i C_i$, we may obtain the following equation that governs the dynamics of the sub-threshold membrane voltage V_i of the isolated LIF neuron:

$$\tau_i^m \frac{dV_i(t)}{dt} = -V_i(t) + R_i I_i^{ext}. \tag{21}$$

In equation Eq. (21), we have assumed that in the absence of the constant external input I_i^{ext} the neuron would have resided to a zero *resting potential*, $V_i^{rest} := 0$. Additionally, the model considers a spike being emitted artificially when the membrane voltage surpasses a certain threshold V_i^{th}, termed *spiking or voltage threshold*, and the membrane potential returns instantaneously to a lower value, the *reset potential* V_i^{reset}.

Without loss of generality, we may assume that at $t = 0$ the membrane voltage lies at the reset potential $V_i(t = 0) := V_i^{reset}$, obtaining eventually a solution of Eq. (21) in the form:

$$V_i(t) = I_i^{ext} R_i \left(1 - e^{-t/\tau_i^m}\right) + V_i^{reset} e^{-t/\tau_i^m}. \tag{22}$$

Embedded in a network, the LIF neuron will receive spikes arriving at its synapses from the set of its pre-synaptic neighbours, denoted by $Pre(i)$. In order to account for the network contributions, we may modify Eq. (21):

$$\tau_i^m \frac{dV_i(t)}{dt} = -V_i(t) + R_i I_i^{ext} + R_i s_i(t), \tag{23}$$

where

$$s_i(t) = \sum_{j \in Pre(i)} \sum_p J_{i,j} \delta\left(t - t_{j,p} - \tau_{i,j}\right) \tag{24}$$

represents the incoming spikes from the set of the pre-synaptic neurons $Pre(i)$ emitted at $t_{j,p}$ and received delayed by $\tau_{i,j}$. The synaptic efficacies captured by $J_{i,j}$ represent the magnitude and the type of the contribution delivered by a single synaptic event.