# Anomaly Detection on Whole-Brain Functional Imaging of Neuronal Activity using Graph Scan Statistics

Youngser Park
Center for Imaging Science
Johns Hopkins University
Baltimore, MD, 21218
youngser@jhu.edu

Heng Wang
Dept. of Applied Mathematics
& Statistics
Johns Hopkins University
Baltimore, MD, 21218
hwang82@jhu.edu

Tobias Nöbauer
Research Institute of
Molecular Pathology (IMP)
Max. F. Perutz Laboratories
(MFPL) & Research Platform
for Quantum Phenomena and
Nanoscale Biological Systems
(QuNaBioS)
University of Vienna
Vienna, Austria
tobias.noebauer@imp.ac.at

Alipasha Vaziri
Research Institute of
Molecular Pathology (IMP)
Max. F. Perutz Laboratories
(MFPL) & Research Platform
for Quantum Phenomena and
Nanoscale Biological Systems
(QuNaBioS)
University of Vienna
Vienna, Austria
vaziri@imp.ac.at

Carey E. Priebe
Dept. of Applied Mathematics
& Statistics
Johns Hopkins University
Baltimore, MD, 21218
cep@jhu.edu

## ABSTRACT

The ability to detect anomalies and/or change-points in a dynamic network or a time series of graphs is an increasingly important task in many applications of the emerging discipline of data mining. We consider scan statistics and demonstrate the successful detection of known but also more "hidden" events of neuronal activities as well as motor response associated signatures in three-dimensional time-series of functional calcium intensity data. We validate our discoveries using persistence analysis by varying a set of input parameters.

## 1. INTRODUCTION

Anomaly and/or change-point detection is an increasingly important problem in many applications of data mining. Dynamic network data are often readily observed, with vertices denoting entities and time evolving edges signifying relationships between entities, and thus considered as a time-series of graphs which is a natural framework for investigation [1, 2].

In this paper, we consider the problem of detecting "state change" of a time-series of neuronal activity data obtained via three-dimensional calcium imaging in larval zebrafish using scan statistics. Scan statistics are commonly used in signal processing to detect a local signal in an instantiation of

some random field [3]. The idea is to scan over a small time or spatial window of the data and calculate some locality statistic for each window. The maximum of these locality statistics is known as the scan statistic. Large values of the scan statistic suggests existence of nonhomogeneity, for example, a local region with significantly excessive activity.

We first convert the multivariate time-series into unweighted time-series of graphs using correlation coefficients [8]. We then calculate scan statistics of graphs and identify change points. We further validate our methodology using persistence analysis. We conclude with discussion and suggestions for future work.

## 2. ZEBRAFISH DATASET

The original data we use in this paper is a simultaneous whole-brain neuronal activity data at near single cell resolution obtained using Light-Field Deconvolution Microscopy (LFDM) combined with GCaMP as a calcium reporter [6]. The data consist of periods of spontaneous neuronal activity and sequences of different types of olfactory stimulations. More concretely, the raw data consists of a multivariate time series $D$ of dimension $n \times m$ where $n(= 5379)$ is the number of spatial filters which largely overlaps with single zebrafish neurons and $m(= 5000)$ is the number of total time frames across 250 seconds. Each single time frame is approximately $1/20$ of a second, which means that the data was collected in 20 Hz. $D_{i,j}$ records a measure of activity level of the neuron $i$ at time frame $j$ by computing fluorescence traces of spatial filters divided by their means. Figure 1 depicts (a) a maximum-intensity projection of the raw volumetric data along $z$ axis overlaid with the locations of 5379 spatial filters and (b) a heatmap of $5379 \times 5000$ matrix, $D$.

During the data collection process over time, a lab scientist created an underlying change-point occurring at $16^{th}$ second, by giving an olfactory stimulus to the zebrafish.
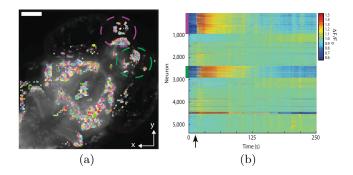
**Figure 1: (a) Maximum-intensity projection (MIP) of a light-field deconvolved volume showing the $xy$ plane. In total, 5,379 filters were automatically identified, most of which correspond to individual neurons. (b) Extracted intensity signal of $Ca^{2+}$ fluorescence using spatial filters shown in (a). Each row shows a time-series heatmap. Color bars near $y$-axis denote encircled regions in (a), which include the olfactory epithelium, olfactory bulb and telencephalon. The arrow at $\sim$16 seconds denotes the addition of an aversive odor. Scale bar is $100\ \mu m$. (Images courtesy of [6]; permission obtained from the publisher.)**

This stimulus lasts about 2 seconds. The overall changes in the neuronal population response based on this stimulus is shown in Figure 1 (b). Additionally, during data pre-processing, some spurious edge neurons were removed. The cleaned data used in following experiments is a multivariate time series $D$ where $n = 5105$ and $m = 5000$.

# 3. CONSTRUCTION OF TIME SERIES OF ASSOCIATION GRAPHS

Using the multivariate time series $D$ and two parameters – window size $\Delta$ and edge threshold $\theta$, we construct a time series of unweighted graphs $\{G_t\}_{t=1}^T$ with a coarser resolution by following three steps below:

1. First, we split the data across time into chunks and let each chunk contain data across $\Delta$ time frames out of $m(= 5000)$ frames. Note that we also let adjacent chunks share overlapping $\frac{\Delta}{2}$ steps so that contiguous chunks are dependent. Hence, we consider $t = 1, \ldots, \frac{2m}{\Delta} - 1$ chunks in total and each chunk represents $\frac{\Delta}{20}$ seconds in real-time duration (i.e., $\Delta$ time steps in the resolution of $D$). For example, if $\Delta = 50$, the first chunk is $[1, 50]$, the second chunk is $[26, 75]$, the third chunk is $[51, 100]$ and so on. The interval of $t$-th time chunk is $[t_1, t_2]$ in original $m$ time steps, where $t_1 = \frac{(t-1)\Delta}{2} + 1$ and $t_2 = \frac{(t-1)\Delta}{2} + \Delta$.

2. For each time stamp $t$, each neuron then has $\Delta$ data samples of measurements of activity levels between $[t_1, t_2]$ in original $D$. Based on $\Delta$ samples in $[t_1, t_2]$, we construct an association matrix $\rho^{(t)}$ for time stamp $t$ where $\rho_{i,j}^{(t)}$ denotes the absolute value of sample Pearson correlation coefficient between neuron $i$ and neuron

$j$:

$$\rho_{i,j}^{(t)} = \frac{|\sum_{k=t_1}^{t_2}(D_{i,k} - \bar{D}_i)(D_{j,k} - \bar{D}_j)|}{\sqrt{\sum_{k=t_1}^{t_2}(D_{i,k} - \bar{D}_i)^2}\sqrt{\sum_{k=t_1}^{t_2}(D_{j,k} - \bar{D}_j)^2}}$$

where $\bar{D}_i = \frac{1}{\Delta}\sum_{k=t_1}^{t_2} D_{i,k}$.

3. So far, $\{\rho^{(t)}\}_{t=1}^T$ can be seen as a time series of weighted adjacency matrices where $T = \frac{2m}{\Delta} - 1$. However, we consider only unweighted graphs in this paper and thus set a threshold $\theta$ on all entries of $\rho$ to convert a weighted adjacency matrix to an unweighted adjacency matrix, denoted by $A^{(t)}$. Specifically, for any pair of vertices $u$ and $v$ in temporal graphs $\{G_t\}_{t=1}^T$, $(u, v)$ is connected if and only if $\rho_{u,v}^{(t)} > \theta$, i.e., $A_{u,v}^{(t)} = 1 \iff \rho_{u,v}^{(t)} > \theta$.

# 4. SCAN STATISTICS AND ANOMALOUS COMMUNITY IDENTIFICATION

After the construction of temporal graphs $\{G_t\}_{t=1}^T$, we are able to apply our anomalous community detection technique on $\{G_t\}_{t=1}^T$. The goal is to find change-points at which a subgroup of neurons shows excessive increase of activity.

## 4.1 Locality-based Scan Statistics

Suppose we are given a time-series of graphs $\{G_t\}_{t\geq 1}$ where $V(G_t)$ is independent of $t$, i.e., the graphs $G_t$ are constructed on the same vertex set $V$. We now define a locality statistic on $\{G_t\}$. For a given $t$, let $\Psi_{t;k}(v)$ be defined for all $k \geq 1$ and $v \in V$ by

$$\Psi_{t;k}(v) = |E(\Omega(N_k(v; G_t))|.$$

$\Psi_{t;k}(v)$ counts the number of edges in the subgraph of $G_t$ induced by $N_k(v; G_t)$, the vertices $u$ at a distance at most $k$ from $v$ in $G_t$. In a slight abuse of notation, we let $\Psi_{t;0}(v)$ denote the degree of $v$ in $G_t$. The statistic $\Psi_t$ was first introduced in [7], and further investigated in [5, 10].

Let $J_{t,t';k}$ be the locality statistic $\Psi_{t';k}$, where for ease of exposition, the index $t$ is a dummy index when $J_{t,t';k} = \Psi_{t';k}$. We now define two kind of normalized statistics for $J_{t,t';k}$, a vertex-dependent normalization and a temporal normalization.

For a given $\tau \geq 0$ and $v \in V$, we define the *vertex-dependent normalization* $\widetilde{J}_{t,\tau;k}(v)$ of $J_{t,t';k}(v)$ by

$$\widetilde{J}_{t,\tau;k}(v) = \begin{cases} J_{t,t;k}(v) & \tau = 0 \\ J_{t,t;k}(v) - \hat{\mu}_{t,\tau;k}(v) & \tau = 1 \\ (J_{t,t;k}(v) - \hat{\mu}_{t,\tau;k}(v))/\hat{\sigma}_{t,\tau;k} & \tau > 1 \end{cases},$$

where $\mu_{t,\tau;k}$ and $\sigma_{t,\tau;k}$ are defined as

$$\hat{\mu}_{t,\tau;k}(v) = \frac{1}{\tau}\sum_{s=1}^{\tau} J_{t,t-s;k}(v),$$

$$\hat{\sigma}_{t,\tau;k}^2(v) = \frac{1}{\tau - 1}\sum_{s=1}^{\tau}(J_{t,t-s;k}(v) - \hat{\mu}_{t,\tau;k}(v))^2.$$

As mentioned in [7], this normalization is necessary to standardize the scales of the raw locality statistics $J_{t,t';k}(v)$. Otherwise, a noiseless vertex in the past

who has dramatically increasing communications at the current time $t$ would be inconspicuous because there might exist a "talkative" vertex who keeps an even higher but unchanged communication level throughout time.

We then consider the maximum of these vertex-dependent temporal normalization for all $v \in V$, i.e., we define a $M_{\tau,k}(t)$ by

$$M_{\tau,k}(t) = \max_v(\widetilde{J}_{t,\tau;k}(v)).$$

Finally, for a given $l \geq 0$, we define the *temporal normalization* of $M_{\tau,k}(t)$ for a $l > 0$ by

$$S_{\tau,\ell,k}(t) = \begin{cases} M_{\tau,k}(t) & l = 0 \\ M_{\tau,k}(t) - \tilde{\mu}_{\tau,\ell,k}(t) & l = 1 \\ (M_{\tau,k}(t) - \tilde{\mu}_{\tau,\ell,k}(t))/\tilde{\sigma}_{\tau,\ell,k}(t) & l > 1 \end{cases},$$

where $\tilde{\mu}_{\tau,\ell,k}$ and $\tilde{\sigma}_{\tau,\ell,k}$ are defined as

$$\tilde{\mu}_{\tau,\ell,k}(t) = \frac{1}{\ell}\sum_{s=1}^{\ell} M_{\tau,k}(t-s),$$

$$\tilde{\sigma}^2_{\tau,\ell,k}(t) = \frac{1}{\ell-1}\sum_{s=1}^{\ell}\left(M_{\tau,k}(t-s) - \tilde{\mu}_{\tau,\ell,k}(t)\right)^2.$$

The motivation behind temporal normalization is to perform smoothing for the statistics $M_{\tau,k}$, similar to how smoothing is performed in time series analysis [7]. Large values of the smoothed statistic indicates an anomaly where there is an excessive increase in activity among a subset of vertices.

## 4.2 Experiments

In our settings, we need to specify five parameters $(\theta, \tau, \ell, k, \Delta)$ as inputs of detection algorithm. In all experiments below, we select the neighborhood distance $k = 1$ because the graph order ($n = 5105$) is in a moderate scale. We fix $\theta = 0.8$ in this section and the effect of varying $\theta$ will be discussed in the next section.

After truncating first $\tau + \ell$ time steps for dual normalizations, Figure 2 depicts $S_{\tau,\ell,1}(t;\Psi)$ given $(\tau,\ell) = (10,10)$ and $(5,5)$ and $\Delta \in \{25,50\}$. In this experiment, detections are defined as time $t$ for which $M_{\tau,1}(t)$ achieves a value greater than ten standard deviations above its mean, i.e., $S_{\tau,\ell,1}(t;\Psi) > 5$. In this figure, the first thing we can notice is that the stimulus point at $t^\star = 16$ for $\Delta = 50$ is not strong enough to be claimed as a detection, mainly due to a sequence of normalizations. It is detected with a smaller window size as shown when $\Delta = 25$. There are several other noticeable anomalies including $t^\star = \{59, 78, 218\}$, which are pink-marked with arrows. These time points coincided with initiation of motor responses by the animal such as a slight eye movement happening at $59^{th}$ second, and tentative tail movements happening at $78^{th}$ and $218^{th}$ seconds.

For each known $t^\star$, the responsible spatial filters $N_1[v^\star; G_{t^\star}]$ are plotted in Figure 3. For example, when $t^\star = 16$, there are $|N| = 778$ associated neurons highlighted in red, which largely agree with the ones in the encircled regions in Figure 1 (a). For comparison, $N_1[v^\star; G_{t^\star-1}]$
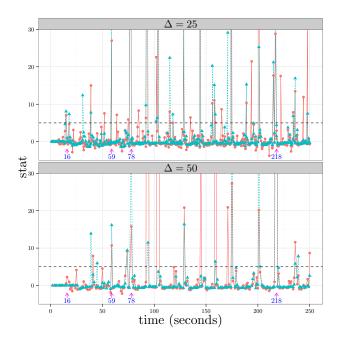


**Figure 2:** $S_{\tau=5,\ell=5,1}(t;\Psi)$ **(red circles) and** $S_{\tau=10,\ell=10,1}(t;\Psi)$ **(green triangles), the temporally-normalized standardized scan statistics using** $\Delta = 25$ **(top panel) and** $\Delta = 50$ **(bottom panel), in time series of zebrafish association-graphs across 250 seconds. Anomaly detection is identified if** $S_{\tau,\ell,1}(t;\Psi) > 5$ **(dashed horizontal line).** $t^\star = 16^{th}$ **second is an underlying change-point at which zebrafish is given a odor stimulus and this stimulus lasts for 2 seconds.** $t^\star = 59^{th}, 78^{th}, 218^{th}$ **seconds are underlying change-points caused by zebrafish eye movement or tail movements. Summary of selected change-points (pink-marked with arrows) are provided in Table 1 and the responsible spatial filters at selected change-points are identified in Figure 3.**

and $N_1[v^\star; G_{t^\star+1}]$ for $t^\star = 16$ are also depicted in Figure 4. According to Figures 2 and 3, our findings are summarized in the following Table 1. Note that four ground truths of change-points observed, i.e., $t^\star = \{16, 59, 78, 218\}$, are correctly identified by $S_{\tau,\ell,1}(t;\Psi)$ in most cases. This demonstrates the efficacy and practicality of our methodology.

Figure 5 depicts histograms of the responsible spatial filters (red points) in Figure 3 for each of the four anomalies, in terms of the neuron ids of the data. It is interesting to point out that even though $t^\star = 16$ was not detected by our statistics for a certain choice of $(\tau,\ell,\Delta)$, it is apparent that there are two clusters of neurons; one in the first 1000 neurons and the other in $[2000, 3000]$, as shown in Figure 1 (b). Similarly, the detection $t^\star = 59$ has associated neurons *not* associated with any of the other three anomalies near 4500-ish which corresponds to the third narrow band in Figure 1 (b), which turned out to be related to eye movement as mentioned earlier.

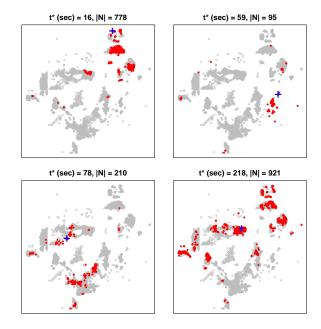Our results suggest that this methodology prompts us

Figure 3: For each $t^\star \in \{16, 59, 78, 218\}$, the responsible spatial filters $N_1[v^\star; G_{t^\star}]$ are visualized in red when $S_{\tau,\ell,1}(t; \Psi)$ is employed for detection with $(\tau, \ell, \theta, \Delta) = (5, 5, 0.8, 50)$. All neurons are spatially located according to their $(x, y)$ coordinates. The blue "+" symbol denotes $v^\star = \arg\max_v(\widetilde{J}_{t^\star, \tau; 1}(v))$, the center of anomalous community. "$|N|$" denotes the number of responsible neurons in $N_1[v^\star; G_{t^\star}]$. For example, when $t^\star = 16$, there are $|N| = 778$ associated neurons largely localized in the top right portion of the plot, which largely agree with the ones in the encircled regions in Figure 1 (a).
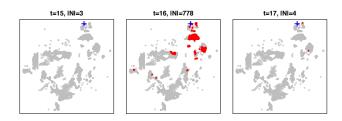


Figure 4: For $t^\star = 16$, the responsible spatial filters $N_1[v^\star; G_{t^\star-1}]$, $N_1[v^\star; G_{t^\star}]$, and $N_1[v^\star; G_{t^\star+1}]$ are visualized in red when $S_{\tau,\ell,1}(t; \Psi)$ is employed for detection with $(\tau, \ell, \theta, \Delta) = (5, 5, 0.8, 50)$. Notice the difference of $|N_1|$ at $t^\star = 16$ compared to before and after.

to look more closely at the data and find these anomalous events, otherwise we would not have paid attention. This confirms that the methodology finds events that could otherwise remain undiscovered.

## 5.   DETECTION PERSISTENCE ANALYSIS

In the previous section, $S_{\tau,\ell,1}(t; \Psi)$ performs well on anomaly detection of neuron associations when $(\tau, \ell, \theta) =$

| $t^\star$ | anomaly | $\tau = \ell = 5$ | | $\tau = \ell = 10$ | |
|---|---|---|---|---|---|
| | | $\Delta = 25$ | $\Delta = 50$ | $\Delta = 25$ | $\Delta = 50$ |
| 16 | odor | ✓ | × | ✓ | NA |
| 59 | eye | ✓ | ✓ | ✓ | ✓ |
| 78 | tail | ✓ | ✓ | ✓ | ✓ |
| 218 | tail | ✓ | ✓ | ✓ | ✓ |

Table 1: Summary results of anomaly detection on $\{G_t\}_{t=1}^T$ by employing $S_{\tau,\ell,1}(t; \Psi)$. Anomaly is identified if $S_{\tau,\ell,1}(t; \Psi) > 5$. "✓" and "×" denotes the success and failure of detection respectively. NA is applicable in the case that $t^\star \leq \tau + \ell$ while $\{G_t\}_{t=1}^{\tau+\ell}$ are truncated for vertex-standardization and temporal-normalization.
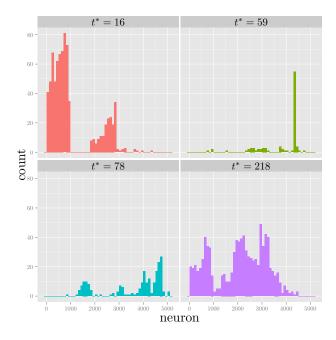


Figure 5: Histograms of the responsible spatial filters (red points) in Figure 3 for each of the four anomalies, in terms of the neuron ids of the data. For $t^\star = 16$, it is apparent that there are two clusters of neurons; one in the first 1000 neurons and the other in $[2000, 3000]$, as shown in Figure 1 (b). Similarly, the detection $t^\star = 59$ has associated neurons *not* associated with any of the other three anomalies near 4500-ish which corresponds to the third narrow band in Figure 1 (b).

$(10, 10, 0.8)$ or $(\tau, \ell, \theta) = (5, 5, 0.8)$ using $\Delta = 50$. Under above two particular settings of $(\tau, \ell, \theta)$, the practicality of $S_{\tau,\ell,1}(t; \Psi)$ is validated by ground truth – identified anomalies are in fact triggered by tentative zebrafish eye or tail movements. However, there is a question remaining unanswered: How does $S_{\tau,\ell,1}(t; \Psi)$ perform with other selections of $(\tau, \ell, \theta)$ and are detections still persistent with varying parameters?

In this section, to eliminate the effect of parameter selection on performance evaluation, we investigate the persistence of detections by calculating $\{S_{\tau,\ell,1}(t; \cdot)\}_{t=1}^T$, fixing all-but-one parameters and varying the remain-
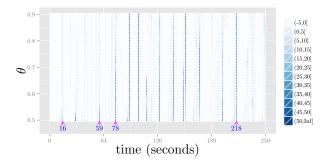
**Figure 6: Persistent plot with respect to the thresholding parameter $\theta$ by fixing $(\tau, \ell, \Delta) = (5, 5, 50)$ and letting $\theta$ range from $0.5$ to $0.9$ with step size $0.01$. Values of $S_{\tau,\ell,k}(t; \Psi)$ at all entries are quantitatively displayed by colors in legend.**



**Figure 7: Persistence plot with respect to the vertex-dependent normalization window $\tau$ by fixing $(\theta, \ell) = (0.8, 2)$(upper), $(\theta, \ell) = (0.8, 5)$(middle), $(\theta, \ell) = (0.8, 10)$(lower) and letting $\tau$ range from $2$ to $10$ with step size $1$. Values of $S_{\tau,\ell,1}(t; \Psi)$ at all entries are quantitatively displayed by colors in legend.**



**Figure 8: Persistence plot with respect to the temporal normalization window $\ell$ by fixing $(\theta, \tau) = (0.8, 2)$(upper), $(\theta, \tau) = (0.8, 5)$(middle), $(\theta, \tau) = (0.8, 10)$(lower) and letting $\tau$ range from $2$ to $10$ with step size $1$. Values of $S_{\tau,\ell,1}(t; \Psi)$ at all entries are quantitatively displayed by colors in legend.**

ing target parameter. A persistence plot can be obtained where $x$−axis is time (in seconds) and $y$−axis is continuous values of the target parameter. The darkness/color at $(x, y)$ entry is proportional to the scale of values of $S_{\tau,\ell,1}(t; \Psi)$ with $t = x$ and target parameter $= y$. If a range of values of $S_{\tau,\ell,1}(t; \Psi)$ across the target parameter (i.e., $y$−axis) are large at some $t = t^{\star}$, we claim that detection at $t^{\star}$ is *persistent* with respect to the target parameter when $S_{\tau,\ell,1}(t; \Psi)$ is used as test statistic. An ideal scenario would be that all detections identified in previous section, $t^{\star} = \{16, 59, 78, 218\}$, are persistent with respect to $\theta$, $\tau$, $\ell$ respectively.

Figure 6 is a persistent plot with respect to $\theta$ by fixing $(\tau, \ell) = (5, 5)$ and letting $\theta$ range from $0.5$ to $0.9$ with step size $0.01$. The superiority of persistence at a particular $t^{\star}$ here is quantified by $\sum_{\theta=0.5}^{0.9} \mathbf{1}_{\{S_{5,5,1}(t^{\star}; \Psi) > 5\}}$, the cumulative counts of alarmed detections across varying $\theta$. We can see that all the detections at $t^{\star} = \{16, 59, 78, 129\}$, identified in previous section, show clear persistences with respect to $\theta$ using the underlying locality statistics, as it gets stronger for smaller $\theta$ for the first two $t^{\star}$s.

Similarly, Figures 7 and 8 are persistence plots with respect to $\tau$ and $\ell$ with underlying scan statistics $S_{\tau,\ell,1}(t; \Psi)$ respectively. The superiority of persistence at a particular $t^{\star}$ is quantified by $\sum_{\tau=2}^{10} \mathbf{1}_{\{S_{\tau,\ell,1}(t^{\star}; \Psi) > 5\}}$ and $\sum_{\ell=2}^{10} \mathbf{1}_{\{S_{\tau,\ell,1}(t^{\star}; \Psi) > 5\}}$, the cumulative counts of identified detections across varying $\tau$ or $\ell$. We can see that detections at $t^{\star} = \{16, 78, 129, 218\}$ are persistent with respective to $\tau$ regardless of selection of $\ell$ and persistent with respective to $\ell$ regardless of selection of $\tau$. Note that as $\tau$ and $\ell$ increase, the signal or consistency of detection, quantified through values of scan statistics $S_{\tau,\ell,1}(t; \Psi)$, is often weakened in both figures. This is reasonable because the signal is going to be smoothed out if there is a large degree of temporal normalization. Thus, this observation will not have influence on our conclusion about the persistent detections with respect to $\tau$ or $\ell$.

## 6. CONCLUSIONS

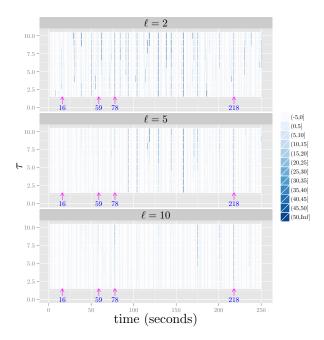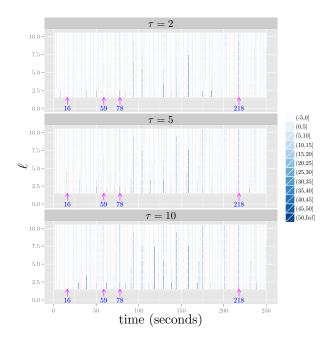We have demonstrated via illustrative experimental results for a time series of graphs derived from functional cal-

cium imaging data that scan statistics successfully detect known and hidden anomalous neuronal activities. We also validated our discoveries using persistence analyses by varying a set of input parameters.

An important extension of this work will be to investigate a strategy of choice of parameters. A thorough simulation study with a suite of planted anomalies with different characteristics would be necessary.

Another extension will be to time-series of weighted graphs, where the correlation coefficient matrix can be used as it is so that the parameter $\theta$ can be eliminated.

Interesting aspect will be to develop principled statistical pattern recognition and machine learning methods for clustering neurons and/or neuron states based on multiple disparate data sets, for example, calcium imaging neuronal *activity* data (like the zebrafish data used in this paper), optogenetic *behavior* data [9], EM *connectome* data [4], and etc. The clustering results will provide the basis for determining when the anomalies occur and/or how the structure of neural circuits relates to their function.

# 7. ACKNOWLEDGMENTS

# 8. REFERENCES

[1] C. C. Aggarwal, Y. Zhao, and P. S. Yu. Outlier detection in graph streams. In *Data Engineering (ICDE), 2011 IEEE 27th International Conference on*, pages 399–409. IEEE, 2011.

[2] L. Akoglu, H. Tong, and D. Koutra. Graph based anomaly detection and description: a survey. *Data Mining and Knowledge Discovery*, 29(3):626–688, 2014.

[3] J. Glaz, J. Naus, and S. Wallenstein. *Scan statistics*. Springer Series in Statistics. Springer-Verlag, New York, New York, NY, 2001.

[4] T. Ohyama, C. M. Schneider-Mizell, R. D. Fetter, J. V. Aleman, R. Franconville, M. Rivera-Alba, B. D. Mensh, K. M. Branson, J. H. Simpson, J. W. Truman, A. Cardona, and M. Zlatic. A multilevel multimodal circuit enhances action selection in Drosophila. *Nature*, 520(7549):633–639, Apr. 2015.

[5] Y. Park, C. E. Priebe, and A. Youssef. Anomaly Detection in Time Series of Graphs using Fusion of Graph Invariants. *IEEE Journal of Selected Topics in Signal Processing*, 7(1):67–75, Feb. 2013.

[6] R. Prevedel, Y.-G. Yoon, M. Hoffmann, N. Pak, G. Wetzstein, S. Kato, T. Schrödel, R. Raskar, M. Zimmer, E. S. Boyden, and A. Vaziri. Simultaneous whole-animal 3D imaging of neuronal activity using light-field microscopy. *Nature Methods*, 11(7):727–730, May 2014.

[7] C. E. Priebe, J. M. Conroy, D. J. Marchette, and Y. Park. Scan statistics on enron graphs. *Computational and Mathematical Organization Theory*, 3:229–247, 2005.

[8] T. Schrödel, R. Prevedel, K. Aumayr, M. Zimmer, and A. Vaziri. Brain-wide 3D imaging of neuronal activity in Caenorhabditis elegans with sculpted light. *Nature Methods*, 10(10):1013–1020, Oct. 2013.

[9] J. T. Vogelstein, Y. Park, T. Ohyama, R. A. Kerr, J. W. Truman, C. E. Priebe, and M. Zlatic. Discovery of brainwide neural-behavioral maps via multiscale unsupervised structure learning. *Science*, 344(6182):386–392, Apr. 2014.

[10] H. Wang, M. Tang, Y. Park, and C. E. Priebe. Locality Statistics for Anomaly Detection in Time Series of Graphs. *IEEE Transactions on Signal Processing*, 62(3):703–717, Feb. 2014.