How to find a unicorn: a novel model-free, unsupervised anomaly detection method for time series

Zsigmond Benkő[1,2], **Tamás Bábel**[1,3], **and Zoltán Somogyvári**[1,3,*]

[1]Wigner Research Center for Physics, Department of Computational Sciences, Budapest, H-1121, Hungary
[2]School of Doctoral Studies, Semmelweis University, Budapest, H-1085, Hungary
[3]Neuromicrosystems Ltd., Budapest, H-1113, Hungary
[*]somogyvari.zoltan@wigner.hu

ABSTRACT

Recognition of anomalous events is a challenging but critical task in many scientific and industrial fields, especially when the properties of anomalies are unknown. In this paper, we present a new anomaly concept called "unicorn" or unique event and present a new, model-independent, unsupervised detection algorithm to detect unicorns. The Temporal Outlier Factor (TOF) is introduced to measure the uniqueness of events in continuous data sets from dynamic systems. The concept of unique events differ significantly from traditional outliers in many aspects: while repetitive outliers are no longer unique events, a unique event is not necessarily outlier in either pointwise or collective sense, it does not necessarily fall out from the distribution of normal activity. We examined the performance of our algorithm in recognizing unique events on different types of simulated data sets with anomalies and compared it with the standard Local Outlier Factor (LOF). TOF had superior performance compared to LOF even in recognizing traditional outliers and also recognized unique events that LOF did not. Benefits of the unicorn concept and the new detection method was illustrated by example data sets from very different scientific fields. In cases where the unique event is already known, our algorithm successfully recognized them: the gravitational waves of a black hole merger on LIGO detector data and the signs of respiratory failure on ECG data series. Furthermore, unique events were found on the LIBOR data set of the last 30 years.

Significance statement

Anomalies in time series are rare and abnormal patterns that can be signs of transient, but significant changes, and therefore their automatic detection is often critical. This is especially difficult when we do not know which parameter of the anomalous pattern differs from normal activity. We have developed a new anomaly detection method that measures the uniqueness of events in time series and based on this, finds special, unique patterns that we have named "unicorns". We have shown that this approach, in addition to finding the anomalies that traditional methods do, also recognizes anomalies that they do not. This is demonstrated on data sets from different fields, from gravitational waves through ECG to economic indicators.

Introduction

Anomalies in time series are rare and non-typical patterns that deviate from normal observations and may indicate a transiently activated mechanism different from the generating process of normal data. Accordingly, recognition of anomalies is often important or critical, invoking interventions in various industrial and scientific applications.

Anomalies can be classified according to various aspects[1]. According their appearance, these non-standard observations can appear as single points falling out from the standard amplitude range (point outliers), however it can take place that amplitudes were in the normal range, but the measured values do not fit into some context (contextual outliers) or the combination of values forms an anomaly (collective outliers). Thus, in case of point outliers, a single points is enough to distinguish between normal and anomalous states, while in case of collective anomalies a pattern of multiple observations is required to recognize the outliers. According to their generation process, two characteristic examples of extreme events are black swans[2] and dragon kings[3]. Black swans are generated by a powerlaw process and they are usually unpredictable by nature. In contrast, the dragon king, such as stock market crashes, occurs after a phase transition and generated by different mechanism from normal samples, therefore it may be predicted. Either black swans and dragon kings are extreme events easily recognisable post-hoc (retrospectively), but not all the anomalies are so effortless to detect. When the amplitude of the event does not fall out from the data distribution, even post-hoc detection can be a troublesome procedure.

Although, not easy to define exactly what an anomaly is, according to most of the definitions, their two key features are the rarity and dissimilarity from the normal data. Most, if not all the outlier detection algorithms approach the anomalies from the dissimilarity point of view. They search for the most distant, most deviant points, while rarity is not in the main focus. In contrast, we approach from the other end: we quantify the rarity of a state, largely independently from the dissimilarity.

Here we introduce a new type of anomaly, the unique event, which is not an outlier in the classical sense of the word: it does not necessarily lie out from the background distribution, neither point-wise, nor collectively. A unique event is defined as a unique pattern which appears only once during the system's investigated history. Based on their hidden nature and uniqueness one could call these unique events as "unicorns" and add them to the strange zoo of anomalies. Note, that unicorns can be traditional outliers appeared only once, but can be patterns, that do not necessarily differ from the normal population in either of their parameters.

But how do you find something you've never seen before, and the only thing you know about is that it only appeared once? The answer would be straightforward for discrete patterns, but for continuous variables, where none of the states are exactly the same, it is challenging to distinguish the really unique states from a dynamical point of view.

Classical supervised, semi-supervised and unsupervised strategies have been used to detect anomalies[1,4,5] and recently deep learning techniques[6–8] was applied to detect extreme events of complex systems[9]. Supervised outlier detection techniques can be applied to identify anomalies, when labeled training data is available for both normal and outlier classes. Semi-supervised techniques also utilize labeled training data, but this is limited to the normal or the outlier class. When the mechanism causing an anomaly is well known and simple, then model based pattern matching techniques can be applied to detect specific anomalies with best results[10]. However when the background is less well known or the system is too complex to get analytical results (or to run detailed simulations), even specific type of anomalies are hard to detect with model -based techniques, because the waveforms are unknown. In cases, when no tractable models or training data is available, model-free unsupervised outlier detection techniques can be applied to detect unexpected events from time series. The logic of this Exploratory Data Analysis (EDA) is reversed compared to the conventional modeling framework: first one could detect events from time-series, then the detected parts can be interpreted by domain experts. Several detection methods can be run in unsupervised mode, few of the most simple methods are the k nearest neighbors (kNN) distance-based anomaly detection techniques. A widely used example for this approach is the standard Local Outlier Factor (LOF) algorithm[11], which was also adapted to time series data[12].

To adapt collective outlier-detection to time series data, nonlinear time series analysis provides the

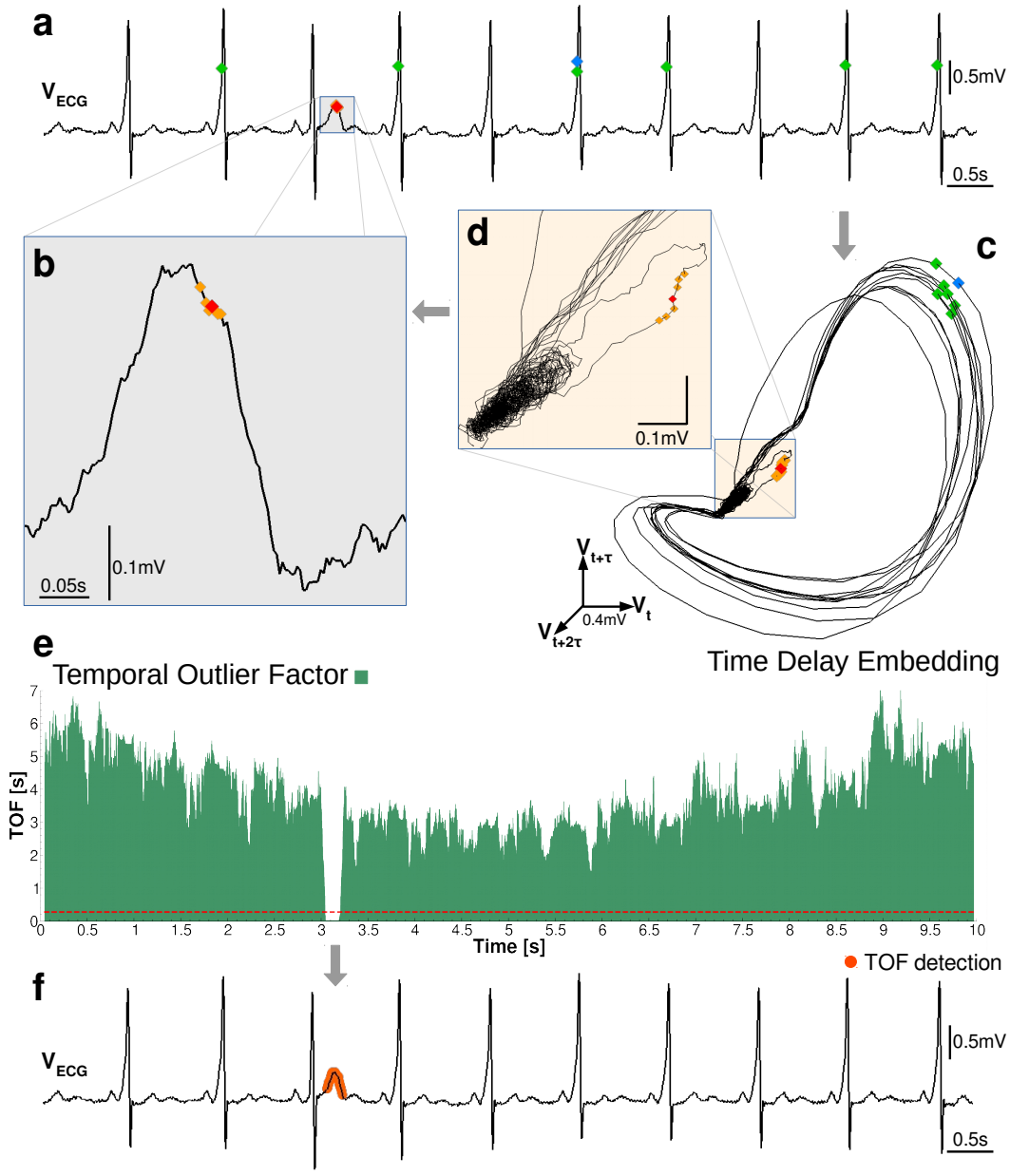


Figure 1. Schema of our unique event detection method and the Temporal Outlier Factor (TOF). **a** An ECG time series from a patient with Wolff-Parkinson-White Syndrome, a strange and unique t-wave zoomed on the graph **b**. **c** The reconstructed attractor in the 3D state space by time delay embedding (E=3, $\tau = 0.011s$). Two example states (red and blue diamonds) and their 6 nearest states space neighbors (orange and green diamonds respectively) are shown. The system returned several times back to the close vicinity of the blue the state, thus the green diamonds are evenly distributed in time, on graph **a**. In contrast, the orange state space neighbors of the red point (zoomed on graph **d**) are close to the red point in time as well on graph **b**. These low temporal distances are show that the red point marks a unique event. **e** TOF measures the temporal dispersion of the k nearest state space neighbors ($k = 20$). Red dashed line is the threshold $\theta = 0.28s$. Low values of TOF below the threshold mark the unique events, denoted by orange dots on the original ECG data on graph **f**.

possibility to generate the multivariate state space from scalar observations. The dynamical state of the system can be reconstructed from scalar time series[13] by taking the temporal context of each point according to Takens' embedding theorem[14]. This can be done via time delay embedding:

$$X(t) = [x(t), x(t+\tau), x(t+2*\tau), \ldots . x(t+(E-1)*\tau)] \quad (1)$$

where $X(t)$ is the reconstructed state at time t, $x(t)$ is the scalar time series. The procedure has two parameters: the embedding delay (τ) and the embedding dimension (E). If E is sufficiently big ($E > 2*d$) compared to the dimension of the attractor (d) and some mild conditions met[14], then the embedded (reconstructed) space is topologically equivalent with the system's state space. Moreover, it has been shown that overembedding can tackle challenges posed by nonstationary signals, if $E > 2*(d+p)$, where p is the number of slowly changing nonstationary parameters[15]. In the overembedded state space, different values of p maps into different state space domains.

As a consequence of Takens' theorem, small neighborhoods around points in the reconstructed state-space also form neighborhoods in the original state space, therefore a small neighborhood around a point represents nearly similar states. This topological property has been leveraged to perform nonlinear prediction[16], noise filtering[17,18] and causality analysis[19–22]. Naturally, time delay embedding can be introduced as a preprocessing step before outlier detection (with already existing methods i.e. LOF) to create the contextual space for collective outlier detection from time series.

Besides the spatial information preserved in reconstructed state space, temporal relations in small neighborhoods can contain clues about the dynamics. For example recurrence time statistics were applied to discover nonstationary time series[23,24], to measure attractor dimensions[25–27] and to detect changes in dynamics[28,29].

In the followings we introduce a new method leveraging nonlinear time series analysis and temporal information to quantify uniqueness, which forms an effective tool to find unicorns.

Results

Temporal Outlier Factor

We present a new model-free unsupervised anomaly detection algorithm to detect unicorns (unique events) based on a dynamical systems' approach to the time series. The algorithm builds on nonlinear time series analysis techniques, such as time delay embedding[14] and upgrades time-recurrence based non-stationarity detection methods[23] by defining a local measure of uniqueness for each point.

The key question in unicorn search is how to measure the uniqueness of a state, as this is the only attribute of a unique event. The simplest possible definition would be, that a unique state is that visited only once in the time series. The problem with this definition is that in case of continuous valued observations, almost every state is visited only once, since the probability of stepping exactly into an already visited point is zero. Thus, a different strategy should be applied to find the unicorns. Our approach is based on measuring the temporal dispersion of the state-space neighbors. If state space neighbors are separated by large time intervals, that means, that the system returned to the same state time-to-time. In contrast, if all the state space neighbors are temporal neighbors as well, that shows, that the system never returned into that state again. This concept is shown on an example ECG data series from a patient with Wolff-Parkinson-White (WPW) Syndrome (Fig. 1). The WPW syndrome is due to an aberrant atrio-ventricular connection in the heart. Its diagnostic signs are shortened PR-interval and appearance of the delta wave, a slurred upstroke of the QRS complex. However, for our representational purpose, we have chosen a data segment, which contained one strange T-wave with uniquely high amplitude (Fig. 1 a).

To quantify the uniqueness on a given time series, the Temporal Outlier Factor (TOF) is calculated in the following steps (SFig. 3):

Firstly, we reconstruct the system's state by time delay embedding (Eq. 1), resulting a manifold, topologically equivalent to the attractor of the system (Fig. 1 c-d and SFig. 3 b).

Secondly, the kNNs in the state space are searched for at each time instances on the attractor. Two examples are shown on Fig. 1 c: a red and a blue diamond and their 6 nearest neighbors marked by orange and green diamonds respectively.

Thirdly, the Temporal Outlier Factor (TOF) is computed from the time indices of the kNN points (SFig. 3 c):

$$\mathrm{TOF}(t) = \sqrt[q]{\frac{\sum_{i=1}^{k} |t - t_i|^q}{k}}. \tag{2}$$

Where t is the time index of the sample point ($X(t)$) and t_i is the time index of the i-th nearest neighbor in reconstructed state-space. Where $q \in \mathscr{R}^+$, in our case we use $q = 2$ (Fig. 1 e).

As a final step for identifying unicorns, a proper threshold θ should be defined for TOF (Fig. 1 e, dashed red line), to mark unique events (orange dots, Fig. 1 f).

TOF measures an expected temporal distance of the kNN neighbors in reconstructed state-space locally at each sample (Eq. 2) and has time dimension. A high or medium value of TOF implies that neighboring points in state-space were not close in time, therefore the investigated part of state-space was visited several different occasions by the system. In our example, green diamonds on (Fig. 1 c) mark states which were the closest points to the blue diamond in the state space, but were evenly distributed in time, on Fig. 1 a. Thus the state marked by the blue diamond was not a unique state, the system returned there several times.

However a small value of TOF implies that neighboring points in state-space were also close in time, therefore this part of the space was visited only once by the system. On Fig. 1 c and d orange diamonds mark the closest states to the red diamond and they are also close to the red diamond in time, on the (Fig. 1 b). This results in low value of TOF in the state marked by the red diamond and means that it was a unique state never visited again. Thus, small TOF values feature the uniqueness of sample points in state-space, and can be interpreted as an outlier factor. Correspondingly, TOF values exhibit a clear breakdown at time interval of the anomalous T-wave (Fig. 1 f).

The number of neighbors (k) used during the estimation procedure sets the possible minimal TOF value:

$$\mathrm{TOF}_{\min} = \sqrt{\frac{\sum_{i=-\lfloor k/2 \rfloor}^{\lfloor k/2 \rfloor + k \bmod 2} i^2}{k}} \Delta t, \tag{3}$$

Where $\lfloor k/2 \rfloor$ is the integer part of $k/2$, $\bmod$ is the modulo operator and Δt is the sampling period.

The approximate possible maximum of TOF is determined by the length (T) and neighborhood size (k) of the embedded time series:

$$\mathrm{TOF}_{\max} = \sqrt{\frac{\sum_{i=0}^{k-1} (T - i\Delta t)^2}{k}} \tag{4}$$

TOF shows a time-dependent mean baseline and variance (Fig. 1 e, SFig. 1) which can be computed if the time indices of the nearest points are evenly distributed along the whole time series. The approximate

mean baseline is a square-root-quadratic expression, it has the lowest value in the middle and highest value at the edges (see exact derivation for continuous time limit and $q = 1$ in Appendix, SFig. 1-2):

$$\sqrt{\left\langle \mathrm{TOF}_{\mathrm{noise}}(t)^2 \right\rangle} = \sqrt{t^2 - tT + \frac{T^2}{3}} \tag{5}$$

$$\mathrm{VAR}\left(\mathrm{TOF}^2_{\mathrm{noise}}(t)\right) = \frac{1}{k}\left(\frac{t^5 + (T-t)^5}{5T} - \left(t^2 - tT + \frac{T^2}{3}\right)^2\right) \tag{6}$$

Based on the above considerations, imposing a threshold θ on TOF_k has a straightforward meaning: it sets a maximum detectable event length (M) or vice versa:

$$\theta = \sqrt{\frac{\sum_{i=0}^{k-1}(M - i\Delta t)^2}{k}} \quad \Big| \quad k\Delta t \overset{!}{\leq} M \tag{7}$$

Where in the continuous limit, the threshold and the event length becomes equivalent:

$$\lim_{\Delta t \to 0} \theta(M) = M \tag{8}$$

Also, the parameter k sets a necessary detection-criteria on the minimal length of the detectable events: only the events longer than $k\Delta t$ may be detected. This property comes from the requirement, that there must be at least k neighbors within the unique dynamic regime of the anomaly.

We compare our method to the standard Local Outlier Factor (LOF) metrics (see Methods and Appendix). The main purpose of the comparison is not to show that our method is superior to the LOF in outlier detection. In contrast, we want to present the fundamental differences between the traditional outlier concept and the unicorns.

The difference between the anomaly-concepts are further emphasized by the fact, that the first steps of the LOF and TOF algorithms are parallel: The LOF uses the time-delay embedding as well, as a preprocessing step to define a state-space and also searches for the kNNs in the state-space for each time instances. As a key difference, LOF calculates the distance of the actual point in the state-space from their nearest neighbors and normalizes it with the mean distance of those nearest neighbors from their nearest neighbors. LOF values around 1 are considered the signs of normal behavior, while higher LOF values mark the outliers. LOF concentrates to the spatial distances in the state-space, while TOF considers the temporal distance of the state-space neighbors (Fig. 1).

Evaluation of performance on simulated data series

Simulated data series with discrete temporal dynamics We simulated two datasets with deterministic chaotic discrete-time dynamics generated by a logistic map[30] ($N = 2000$, $100 - 100$ instances each) and inserted variable-length ($l = 20 - 200$ step) outlier-segments into the time series at random times (Fig. 2 a-b). Two types of outliers were used in these simulations, the first type was generated from a tent-map dynamics (Fig. 2 a) and the second type was simply a linear segment with low gradient (Fig. 2 b) for simulation details see Methods. The tent map demonstrates the case, when the underlying dynamics is changed for a short interval, but it generates a very similar periodic or chaotic oscillatory activity (depending on the parameters) than the original dynamics. This type of anomaly is hard to distinguish by

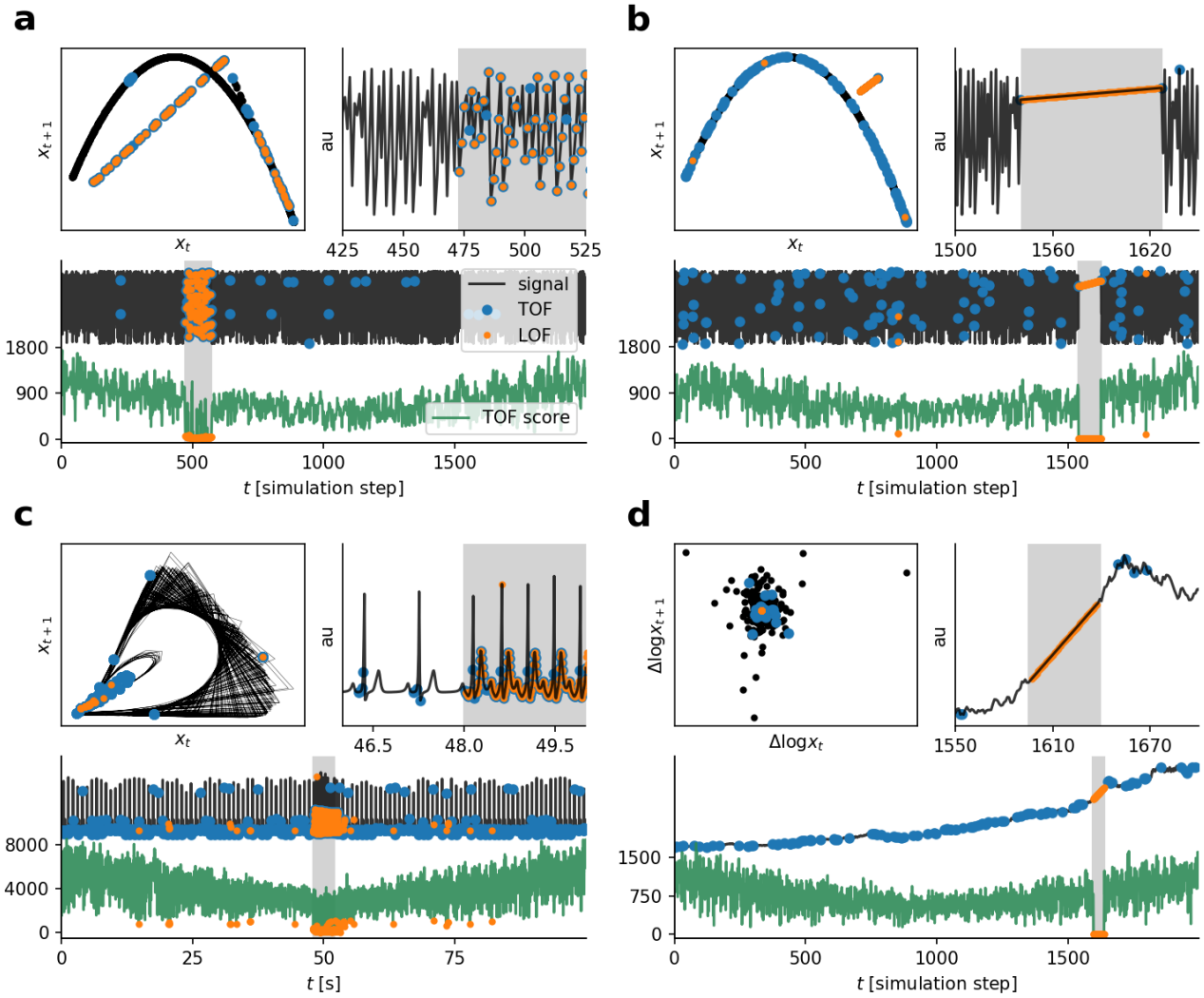


Figure 2. Simulated time series with anomalies of different kinds. a Logistic map time series with tent-map anomaly. **b** Logistic map time series with linear anomaly. **c** Simulated ECG time series with tachycardia. **d** Random walk time series with linear anomaly, where the TOF was measured on the discrete time derivative ($\Delta log x_t$). Each subplot shows an example time series of the simulations (black) in arbitrary units and in three forms: Top left the return map, which is the results of the 2D time delay embedding, and defines the dynamics of the system or its the 2D projection. Bottom: Full length of the simulated time series (black) and the corresponding TOF values (green) shaded areas show anomalous sections. Top right: Zoom to the onset of the anomaly. In all graphs outliers detected by TOF and LOF are marked by orange and blue dots respectively. While in **a** and **b** cases the anomalies form clear collective outliers, **d** shows an example where the unique event is clearly not an outlier neither pointwise nor collectively, it is in the centre of the distribution. Both LOF and TOF detected well the anomalies in cases **a** and **c**, but only TOF was able to detect the anomalies in **b** and **d** cases.

naked eye. In contrast, a linear outlier is easy to identify for a human observer but not for many traditional outlier detecting algorithms. The linear segment is a collective outlier and all the points of it represents a state which were visited only once during the whole data sequence, therefore they are unique events as well.

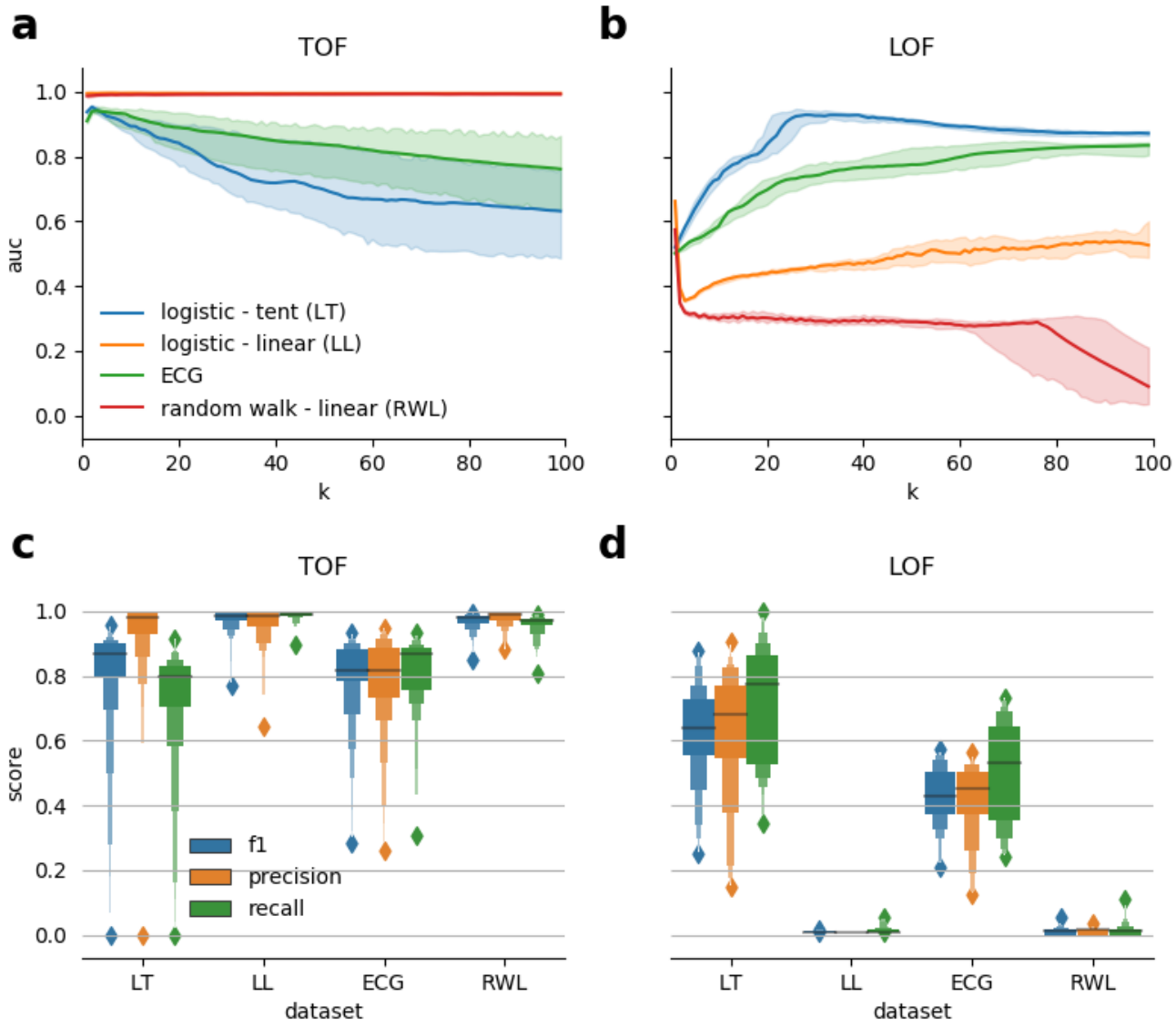


Figure 3. Performance evaluation of TOF and LOF on simulated datasets. a and **b**: Median ROC AUC score and the median absolut difference for TOF (**a**) and LOF (**b**) are showed in the function of neighborhood size (k). TOF showed the best results for small neighborhoods. In contrast, LOF showed better results for larger neighborhoods in the case of logistic map and ECG datasets, but did not reach reasonable performance on linear outliers. **c** TOF: Median F_1 score, precision and recall values showed very good precision scores on all datasets. The recall was very high for the linear anomalies and slightly lower for logistic map - tent map anomaly dataset. **d** LOF: Median F_1 shows very low-valued metrics on datasets with linear anomaly and mediocre values on tent-map anomaly and simulated ECG time series.

Simulated ECG datasets with tachycardia. As a continuous deterministic dynamics with realistic features, we simulated electrocorticograms with short tachycardic periods where beating frequency was higher (Fig. 2 c). The simulations were carried out according to the model of Rhyzhii & Ryzhii[31], where the three heart pacemakers and muscle responses were modelled as a system of nonlinear differential equations (see Methods). We generated 100 seconds of ECG and randomly inserted $2-20$ seconds long

dataset	TOF		LOF	
	k	AUC	k	AUC
logmap tent	2	0.953 ± 0.027	28	0.928 ± 0.075
logmap linear	6	0.996 ± 0.004	1	0.662 ± 0.016
sim ECG tachy	3	0.941 ± 0.030	99	0.834 ± 0.053
randwalk linear	70	0.993 ± 0.007	1	0.573 ± 0.012

Table 1. Detection performance on simulations in terms of maximal ROC AUC scores and the optimal neighborhood parameter k. Maximal median ROC AUC values and the corresponding median absolute difference ranges are shown. LOF detected tent map outliers and tachycardia with reasonable reliability but TOF outperformed LOF for all data series. Linear outliers can not be detected by the LOF method at all, while TOF detected them almost perfectly. TOF reached its maximal performace mostly for low k values, only random walk with linear outlier required larger neighborhood to compute, while LOF required larger k for optimal performance on those two data series, on which it worked reasonably.

dataset	TOF			LOF		
	F_1	precision	recall	F_1	precision	recall
logmap tent	0.869 ± 0.055	0.979 ± 0.031	0.797 ± 0.069	0.640 ± 0.130	0.681 ± 0.158	0.777 ± 0.175
logmap linear	0.986 ± 0.014	0.985 ± 0.022	0.991 ± 0.004	0.009 ± 0.002	0.009 ± 0.000	0.009 ± 0.004
sim ECG tachy	0.816 ± 0.095	0.818 ± 0.129	0.868 ± 0.052	0.431 ± 0.107	0.452 ± 0.079	0.533 ± 0.205
randwalk linear	0.980 ± 0.011	0.991 ± 0.013	0.973 ± 0.009	0.014 ± 0.010	0.019 ± 0.000	0.012 ± 0.013

Table 2. Performance evaluation by F_1, precision and recall scores on simulations. Median scores and median absolute differences are shown. In case of TOF, $k = 4$ is used, while for LOF, the k resulted the best ROC AUC were used. TOF is characterized by high precision and recall for all datasets. As a result, very high F_1 scores were reached for all datasets. In comparison, LOF produced mediocre F_1 scores for tent and tachycardia and very low F_1 scores for linear outliers.

faster heart-rate segments, corresponding to tachycardia ($n = 100$ realizations).

Deterministic anomaly on stochastic background dynamics. Takens' time delay embedding theorem is valid for time series generated by deterministic dynamical systems, but not for stochastic ones. In spite of this, we investigated the applicability of time delay embedded temporal and spatial outlier detection on stochastic signals with deterministic dynamics as outliers. We established a dataset of multiplicative random walks ($n = 100$ instances, $T = 2000$ steps each) with randomly inserted variable length linear outlier segments ($l = 20 - 200$, see Methods). As a preprocessing step, to make the random walk data series stationary, we took the log-difference of time series as it is general in case of economic data series (Fig. 2 d).

Performance measures and dependency on neigbourhood size. To detect anomalies we applied TOF and LOF on the datasets ($E = 3$, $\tau = 1$) and measured detection-performance by area under receiver operating characteristic curve (ROC AUC), F_1 score, precision and recall metrics. (Fig. 3 and Table 1). F_1

score is especially useful to evaluate detection performance in case of highly unbalanced datasets as it is our case.

Fig. 3 a and b show, the performance of the two methods in terms of median ROC AUC for ($n = 100$) realizations and its dependency on the neighborhood size ($k = 1 - 100$) that used for the calculations.

Two types of behavior were observed among the four experimental setup: First, the linear anomalies were almost perfectly detected by the TOF, with no significant dependency on the k neighborhood, while LOF was not able to detect these anomalies at all (ROC AUC was close to 0.5) again independently from the k neighborhood. The second group is formed by the logistic-tent map dataseries and the simulated ECG with tachycardia. Both TOF and LOF resulted in reasonably high ROC AUC values on these anomalies, however with different k dependency. TOF reached the highest scores for low neighborhoods while LOF required higher k for its optimal performance. As we compared the achieved ROC AUC scores of the two methods for each cases at their optimal neighborhood size, we can conclude, that TOF slightly outperformed LOF even in this second group of data series as well (Table 1).

To further evaluate the components of the performances and the type of errors of these algorithms, F_1 score, precision and recall were computed for both TOF and LOF (Fig. 3 c, d and Table 2). As the TOF showed best performance with lower k neighborhood sizes on logmap-tent and simulated ECG-tachycardia values and showed no significant dependency with linear anomalies, the F_1 scores were calculated at a fixed $k = 4$ neighborhood forming a simplex in the 3 dimensional embedding space[19]. In contrast, as LOF showed stronger dependency on neighborhood size, thus the optimal neighborhood sizes were used to F_1 score calculations. Finally, thresholds corresponding to $M = 110$ timesteps and the top 5.5% percentile were used for TOF and LOF respectively, which is the expectation of the simulated anomaly length.

For the linear outlier datasets, TOF performed almost perfectly with high precision and recall, hence with high F_1 score as well. In contrast LOF performed poorly in all these measures. On the logmap-tent dataset TOF showed good F_1 score due to very high precision and slightly lower recall score, while LOF reached a mediocre F_1-score. On the simulated ECG-tachycardia data series, TOF produced reasonably high precision, recall and F_1 score, while LOF produced low quality measures in all means.

dataset	Density		LOF	
	Normal	Anomaly	*Normal*	Anomaly
logmap tent	95.759 ± 12.070	11.606 ± 1.146	1.039 ± 0.010	3.424 ± 1.990
logmap linear	95.190 ± 9.305	97.413 ± 51.289	1.040 ± 0.012	1.398 ± 0.451
sim ECG tachy	10146 ± 2227	168.370 ± 38.699	1.106 ± 0.022	1.264 ± 0.227
randwalk linear	197.919 ± 3.866	52590 ± 61527	1.623 ± 0.661	1.872 ± 0.920

Table 3. State space densities and LOF values within normal and anomalous activity. Median and median absolute difference of the points density and LOF values in the reconstructed state space are shown, calculated from the distance of the 20 nearest neighbors. The density of the anomaly was significantly lower than the density generated by normal activity in two cases: the tent map anomaly in logistic backgroud and the tachycardia within the normal heart rhythm and resulted in higher LOF values of anomalies in these two cases too. While the density of the linear anomaly segments were not significantly different from the logistic background, the linear anomalies generated much higher density than the normal random walk time series after detrending. Correspondingly, LOF values were not significantly higher in these two cases within the anomaly than the normal activity.

Based on these simulations we can conclude, that there are anomalous events, like tent map on logistic background or high frequency tachycardic events, which could be found by both TOF and LOF based methods, while other types of unique events, like the linear sequences can only be found by TOF method. Table 3 shows, that the tent map and the tachycardia produce lower density, thus more dispersed points in the state space, presumably made them detectable by the LOF. In contrast, linear segments resulted similar density of points than the normal logistic activity, or higher density than the random walk background. In the latter case, the applied preprocessing step, the detrending via differentiation of the logarithm, which made the data series stationary, increased drastically the state space density of the anomaly. While LOF counted only the low density sets as outliers, TOF was able to find the unique events independently of their density. To sum it up, TOF has reached better performance to detect anomalies in all the investigated cases.

TOF detects unicorns

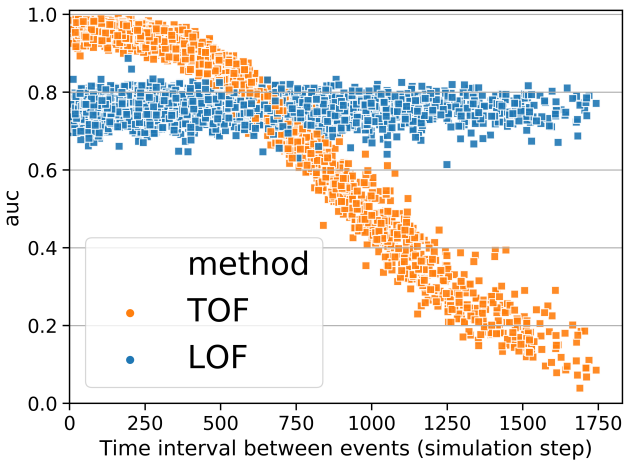


Figure 4. TOF detects unique events. Detection performance measured by ROC AUC as a function of the minimum inter event interval between two inserted tent-map outlier segments. Below 300 step inter event intervals, TOF found outliers with good performance, however for greater than 300 steps the algorithm found fewer outliers. In contrast, LOF's performance remained constant over the whole IEI range.

To show that TOF enables to detect only unique events, additional simulations were carried out, where two, instead of one, tent-map outlier segments were inserted into the logistic map simulations. We detected outliers by TOF and LOF and subsequently ROC AUC values were analysed as a function of the inter event interval (IEI, Fig. 4) of the outlier segments. LOF performed independent of IEI, but TOF's performance showed strong IEI-dependence. Highest TOF ROC AUC values were found at small IEI-s and AUC was decreasing with higher IEI. Also the variance of ROC AUC values was increasing with IEI. This result showed, that the TOF algorithm can detect only unique events: if the two outlier event is close to each other enough, they can be considered as one unique event together. In this case, the TOF can detect it with higher precision, compared to LOF. However if they are farther away than the time limit determined by the detection threshold, then the detection performance decreasing rapidly.

The results also showed, that anomalies can be found by the TOF only if they are alone, a second appearance decreases the detection rate significantly. Based on these results we can conclude, that the TOF method can be applied to real world datasets to reveal unicorns in them.

Application examples on real-world data series

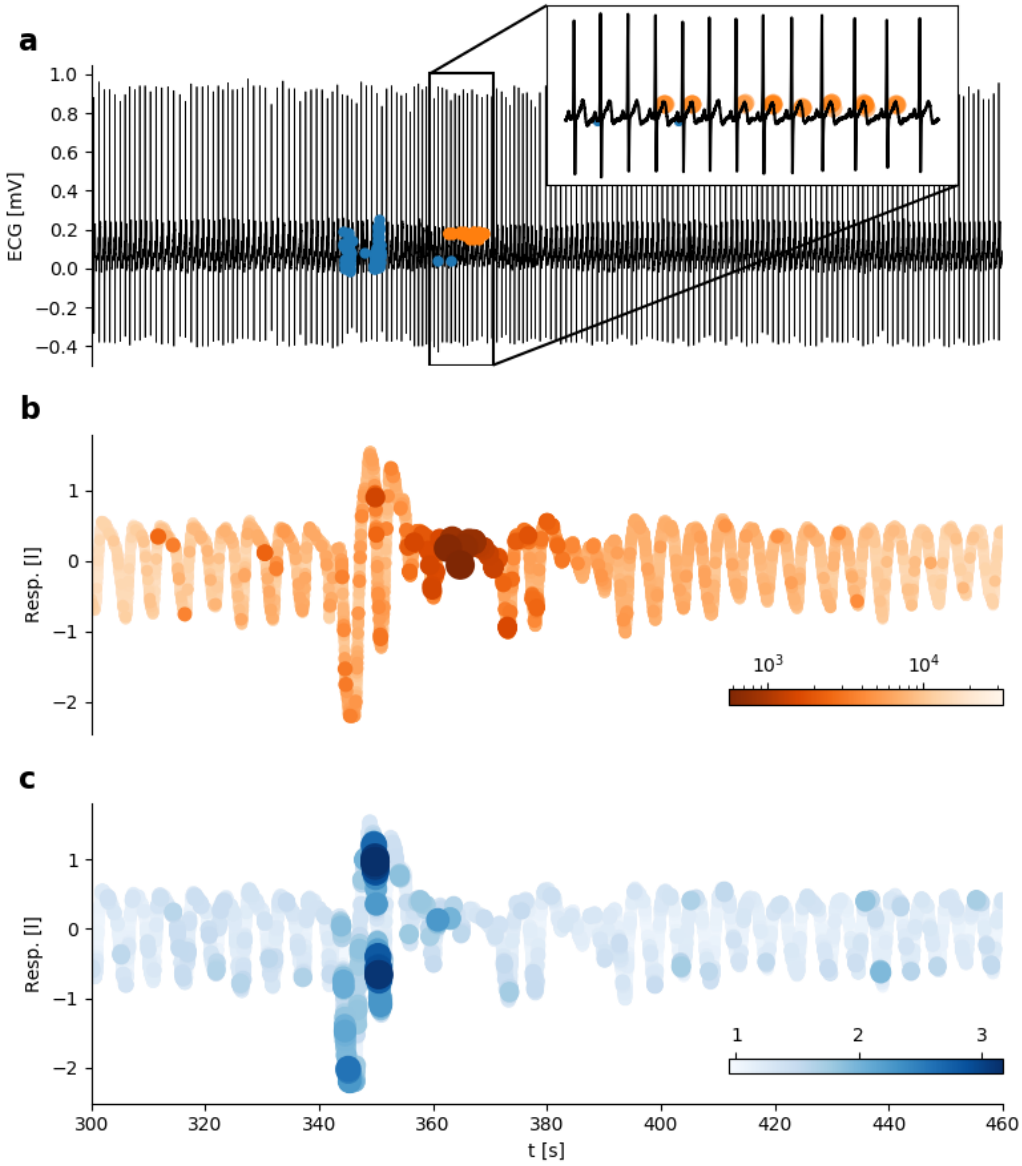


Figure 5. Detecting hypopnea with arousal on ECG by TOF and LOF. a ECG time series with unique events detected by TOF (orange, $E = 3, \tau = 0.02\,\text{s}, k = 11, M = 5\,\text{s}$) and outliers detected by LOF (blue, $E = 7, \tau = 0.02\,\text{s}, k = 100$, threshold$= 0.5\%$). The inset shows a more detailed pattern of detections: unique behavior mainly appears on the T waves. **b** Air-flow time series with coloring corresponds to the TOF score at each sample. Low values mark the anomaly. After a normal period, the breathing gets irregular and almost stops, then after arousal the breathing pattern becomes normal again. TOF finds the period, when the breathing activity almost stops. **c** Air-flow time series with coloring corresponds to the LOF score at each sample. Higher LOF values mark the outliers. LOF finds irregular breathing preceding the hypopnea.

Detecting hypopnea events on ECG time series. We applied TOF on the MIT-BIH Polysomnographic Database's[32,33] ECG measurements to detect hypopnea events. Multichannel recordings were

taken on 250 Hz sampling frequency, and the ECG and respiratory signal of the first recording was selected for further analysis ($n = 40000$ data points 1600 secs).

While the respiratory signal clearly showed the hypopnea, there were no observable changes on the parallel ECG signal.

We applied time delay embedding with $E_{\mathrm{TOF}} = 3$, $E_{\mathrm{LOF}} = 7$ and $\tau = 0.02\,\mathrm{s}$ according to the first zerocrossing of the autocorrelation function (SFig. 5). TOF successfully detected hypopnea events in ECG time series, curiously the unique behavior was found mostly during T-waves while the breathing activity was almost shut down (Fig. 5, $k = 11$, $M = 5$ s). In contrast, LOF was sensitive to the increased and irregular breathing before hypopnea ($k = 200$, threshold$= 0.5$ %). This example shows that this new method could be useful for biomedical signal processing and sensor data analysis.

Detecting gravitational waves. As a second example of real world datasets, we analyzed gravitational wave detector time series around the GW150914 merger event[10] (Fig. 6). The LIGO Hanford detector's signal (4096 Hz) were downloaded from the GWOSC database[34].

A 12 s long segment of strain data around the GW150914 merger event were selected for further analysis. As a preprocessing, the signal were bandpass-filtered (50-300 Hz). Time delay embedding were carried out with embedding delay of 8 time-steps (1.953 ms) and embedding dimension of $E = 6$ and $E = 11$ for TOF and LOF respectively We set the parameters of the algorithms as follows: $k = 12, M =$ 146.484 ms for TOF and $k = 100$, threshold$= 0.5\%$ for LOF (SFig. 6).

Both TOF and LOF detected the merger event, however TOF detected selectively the period when the chirp of the spiralling mergers was the loudest (Fig. 6 b, c).

London InterBank Offer Rate dataset We also applied TOF and LOF on the London InterBank Offer Rate (LIBOR) dataset. As a preprocessing, discrete time derivative was calculated to eliminate global trends, then TOF ($E = 3, \tau = 1, k = 5, M = 30$ month) and LOF ($E = 3, \tau = 1, k = 30$, threshold$= 18.86\,\%$) was applied on the derivative (SFig. 7-8). TOF found the uprising period prior to the 2008 crisis and the slowly rising period from 2012 to nowdays as outlier segments. LOF detected several points, but no informative pattern emerges from the detections (Fig. 7). While in this case the ground-truth was not known, the two highlighted periods show specific patterns of monotonous growth. Moreover, the fact that both two periods were detected by the TOF shows that both dynamics are unique, therefore different from each other during the two periods.

Discussion

In this paper we introduced a new concept of anomalous event called unicorn and we have defined the Temporal Outlier Factor to quantify this concept. Unicorns are the unique states of the system, which were visited only once. A new anomaly concept can be valid only if a proper detection algorithm is provided. We demonstrated that TOF is a model-free, non-parametric, domain independent anomaly detection tool, which can detect unicorns. TOF measures the temporal dispersion of state space neighbors for each points. If state space neighbors are temporal neighbors as well, then the system has never returned into that state, therefore it is a unique event. ie. a unicorn.

The unicorns are not just outliers in the usual sense, they are conceptually different. As an example for their inherently different behavior, one can consider a simple linear data series: As all of its points are visited only once and the system never returned into neither of them, all the points of that line will be unique events. While this property may seem to be counter-intuitive, this ensures that our algorithm finds unique events regardless of their other properties, such as amplitude or frequency. This example also shows, that the occurrence of unique events are not necessarily rare: actually, all the points of a time series can be unique. This property clearly differs from other anomaly concepts: most of them assume that there

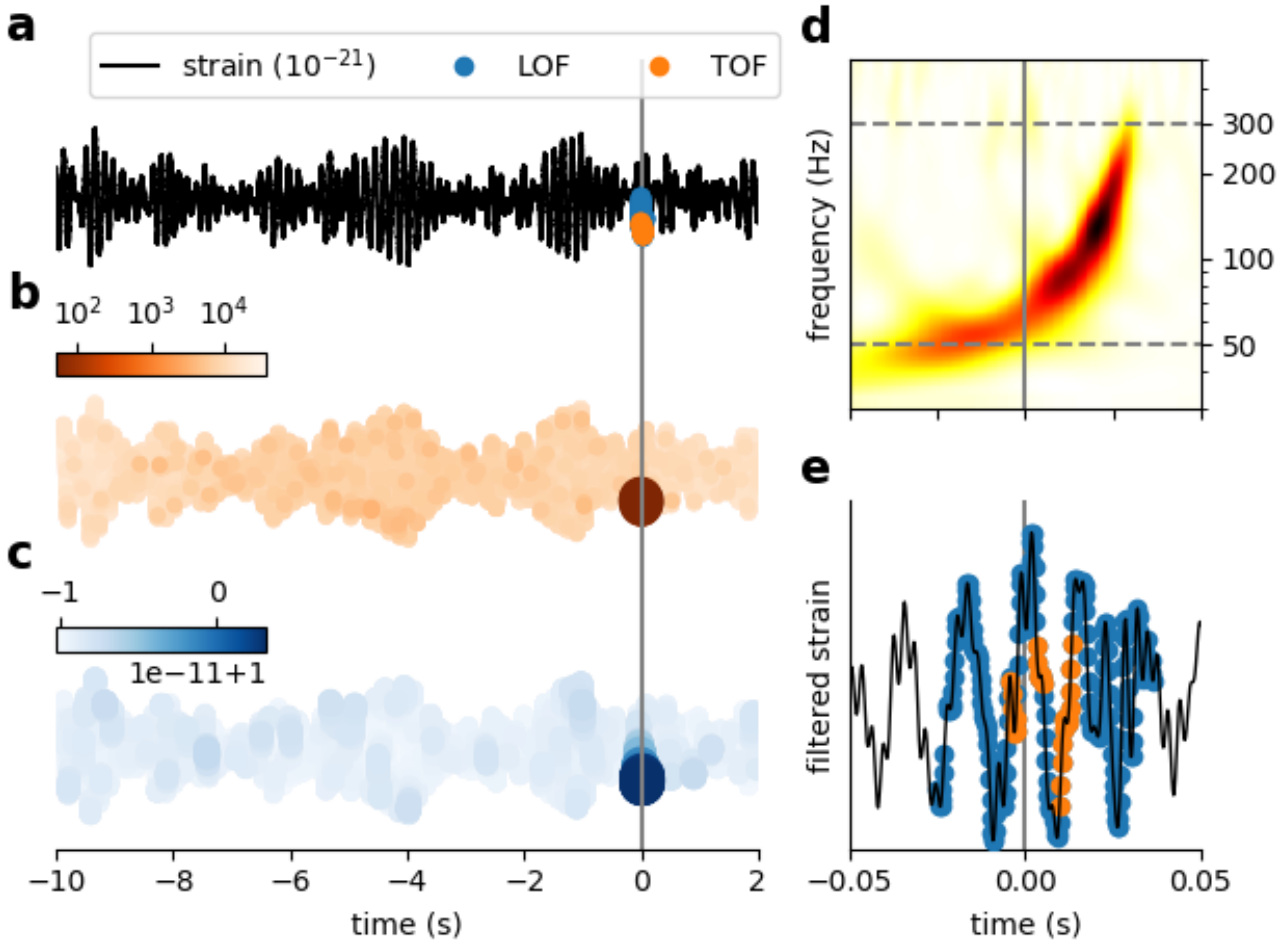


Figure 6. Detection of the GW150914 event on LIGO open data with LOF and TOF. a Strain time series (black) from Hanford detector around GW150914 event (grey) with LOF (blue) and TOF (orange) detections. TOF score values (**b**) and LOF scores (**c**) are mapped to the time series (orange and blue respectively), the most strong colors shows the detected event around 0 seconds. **d** The Q-transform of the event shows a rapidly increasing frequency bump in the power spectra right before the merger event (grey). The grey dashed lines show the lower (50 Hz) and upper (300 Hz) cutoff frequencies of the bandpass filter, which was applied on the time series as a preprocessing step before anomaly detection. **e** Filtered strain data at 0.1 second neighborhood around the event. LOF and TOF detected the merger event with different sensitivity, LOF detected more points of the event, while TOF found the period which has the highest power in the power spectra. ($E_{\mathrm{TOF}} = 6$, $\tau_{\mathrm{TOF}} = 1.953$ ms, $k_{\mathrm{TOF}} = 12$, $M = 146.484$ ms, $w = 7$; $E_{\mathrm{LOF}} = 11$, $\tau_{\mathrm{LOF}} = 1.953$ ms, $k_{\mathrm{LOF}} = 100$, threshold$= 0.5\%$)

is a normal background behavior which generates the majority of the measurements and outliers form only a small minority.

Detection performance comparison of TOF and LOF on different simulated datasets highlighted the conceptual difference between the traditional outliers and the unique events as well. LOF detected classic outliers based on the drop in the local density, therefore it detected well those anomalies which generated low density sets in the state space. The short segment of tent map within a logistic map background and the higher frequency beating of tachycardia within the background of the normal heart rhythm generated

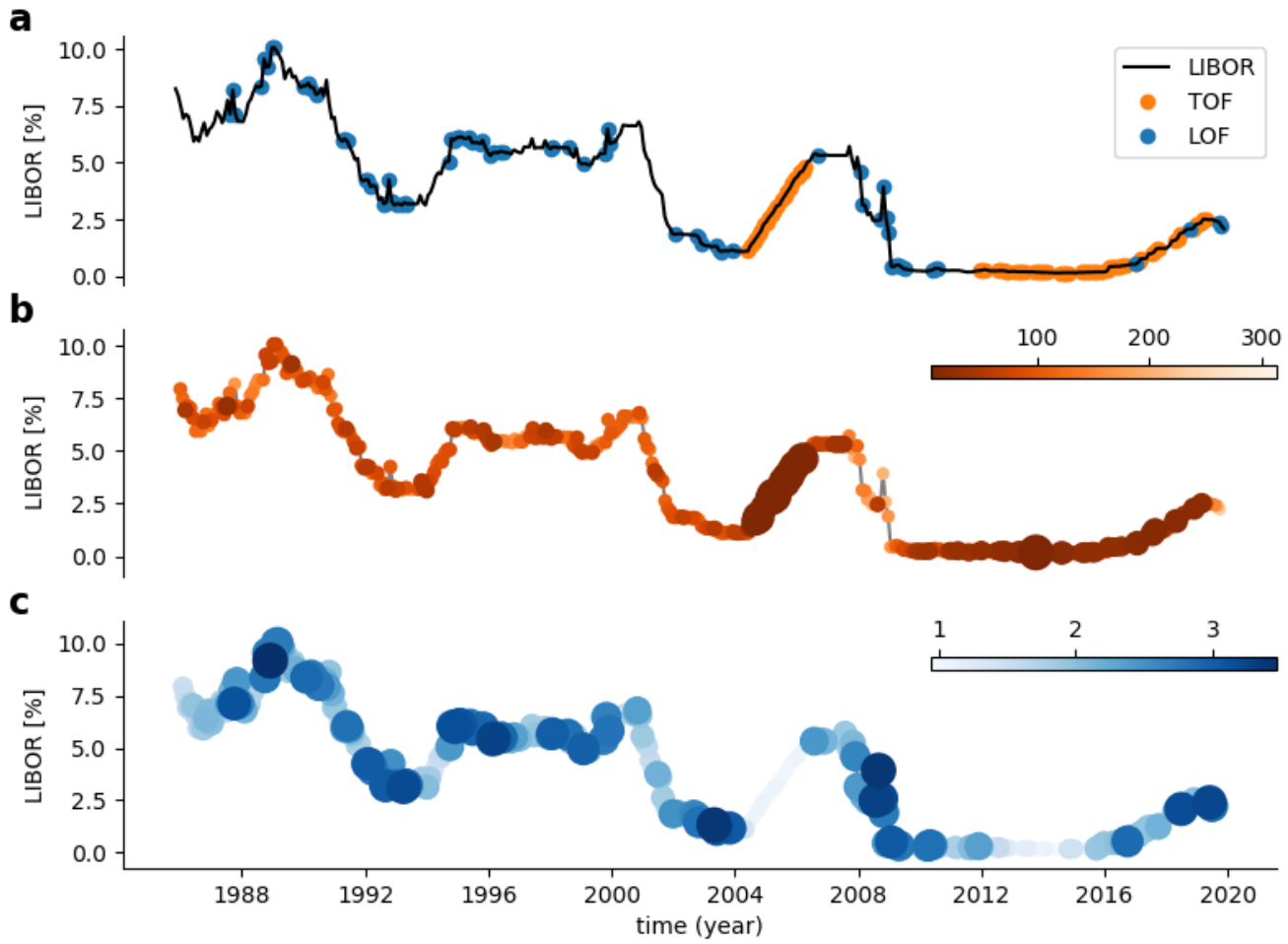


Figure 7. Analysis of LIBOR dataset. The detections were run on the temporal derivative of the LIBOR time series. **a** time series with detections. **b** TOF score values. **c** LOF score values. TOF detected two rising periods: the first between 2005 and 2007 and a second, started in 2012 and lasts until now. While both periods exhibit unique dynamics, they differ from each other as well.

such low density sets. However, depending on the parameters, linear segment anomalies can form higher, lower or equal density sets compared to the background. In our simulations the linear anomaly formed similar density sets than the logistic background and much higher than the random walk background, which made them invisible for the LOF. As our simulations showed, TOF with the same parameter settings was able to find both higher and lower density anomalies, based on the sole property that they were unique events. As a striking difference to the outlier concept, in case of the detrended random walk dataset with linear anomaly, the anomalous points not only formed a higher density set, but it was located right in the center of the normal data distribution. In our tests, TOF showed good precision and recall on simulated dynamical and stochastic anomaly datasets, as well as on simulated ECG time series. The algorithm has very low false detection rate, but not all the points of an outlier were found or not all the points of the event were unique. As an example, QRS waves of ECG simulations appear not to be different from the normal ones, therefore the algorithms did not find them.

Also it was shown in the simulations with multiple events: TOF detects only unique events or unicorns. However when two outliers were so close to each other that the elapsed time was in the same order of

magnitude as the threshold event length, then TOF identified both events as outliers, since the two events formed one longer one. On the contrary, when the elapsed time between outlier segments is much bigger than the threshold event length, TOF detected none of the events. In this latter case, ROC AUC values were much smaller than by-chance meaning, that TOF score during the events were higher than average score over the whole time period.

On the polysomnographic dataset, the anomaly was known from the parallel respiratory signal, but there was no evident change on the ECG signal. It was shown that TOF and LOF both can detect parts of the apnoe event from an ECG measurement. TOF found the actual stalling period of respiration, but LOF detected the preceding irregular breaths. While ECG analysis mostly concentrates onto the temporal relations of the identified wave components, here we apply the detection methods to the continuous ECG data. Curiously, TOF marked mainly the T waves of the heart cycle as anomalous points. T-waves are signs of the ventricular repolarization and known to be largely variable, thus they are often omitted from the ECG analysis. This example showed, that they can carry relevant information as well.

On the gravitational wave dataset both TOF and LOF was able to detect the merger GW150914 event, however TOF needed lower embedding dimension ($E = 6$) and neighbor number ($k = 11$) than LOF ($E = 11, k = 100$). Clearly, those specific, model-based detectors used originally to recognize gravitational waves are more sensitive to the actual waveforms of merger black holes or neutron stars, than the model-free methods we used[10]. However, model-free methods can have role in finding signs of events with unpredicted waveforms such as gravitational waves of supernovas.

TOF detected the two rising periods on temporal derivative of USD LIBOR dataset: one preceding the 2008 crisis, and an other from 2012 till nowdays, while LOF showed no specific detection pattern. Both detected periods showed unique dynamics that differs from each other as well. The period between 2005-2007 can be considered unique in many ways: not only because of the upswing of the global market but it become infamous about the LIBOR scandal, where investigations revealed that several banks colluded in manipulation and rigging of LIBOR rates[35]. Note, that this was not the only case, when LIBOR was manipulated: During the economic breakdown in 2008 the Barclys Bank submitted artificially low rates to show healthier appearance[36–38]. As a consequence of these scandals, significant reorganization took place in controlling LIBOR calculation, starting from 2012.

To sum it up, gravitational waves of the merger black-holes on the filtered dataset formed a traditional outlier which was well detectable both by the TOF and LOF, while LIBOR exhibited longer periods of unique events only detectable by the TOF. Hypopnoe generated a mixed event on ECG, where the period of irregular breathing formed outliers and was detectable by LOF, while during the failed respiration, the apnoe generated a unique event on the ECG only detectable by the TOF.

Comparing TOF and LOF proved that temporal scoring has advantageous properties and adds a new aspect to anomaly detection. One advantage of TOF can be experienced when it comes to threshold selection. Since TOF score has time dimension, an actual threshold value means the maximal expected length of the to be found event. Also, on the reverse of the medal the neighborhood size k parameter sets the minimal event length. Because of these properties, domain knowledge about possible event lengths renders threshold selection to a simple task. An other advantage of TOF is from the computational point of view: the method performs optimally on small embedding dimensions and neighborhood sizes, which makes computations faster and less memory hungry.

Time indices of k nearest neighbors have been previously utilized differently in nonlinear time series analysis to diagnose nonstationary time series[23,24,39], measure intrinsic dimensionality of system's attractors[25–27], monitor changes in dynamics[28] and even for fault detection[29]. Rieke et al. utilized very resembling statistics to TOF: the average absolute temporal distances of k nearest neighbors from the points. However they analyzed the distribution of temporal distances to determine nonsationarity and did

not interpret the resulting distance scores locally. Gao & Hu and Martinez-Rego et al. used recurrence times (T_1 and T_2) to monitor dynamical changes in time series locally, but these statistics are not specialized for detecting extremely rare unique events. TOF utilizes the temporal distance of k nearest neighbors at each point, thus provides a locally interpretable outlier score. This score takes small values when the system visits an undiscovered territory of state-space for a short time period, therefore it is suitable to detect unicorns.

Future directions to develop TOF would be to form a model which is able to represent uncertainty over detections by creating temporal outlier probabilities just like Local Outlier Probabilities[40] created from LOF. Moreover, an interesting possibility would be to make TOF applicable also on different classes of data, for example on point processes like spike-trains, network traffic time-stamps or earthquake dates.

Methods

TOF Analysis workflow

1. Preprocessing and applicability check:

   This step varies from case to case, and depends on the data itself or on the goals of analysis. Usually it is advisable to make the data stationary. For example in the case of oscillatory signals, the signal must contain many periods even from the lowest frequency components. If this latter condition does not hold, then Fourier filtering can be applied to get rid of the low frequency components of the signal.

2. Time delay embedding:

   We embed the scalar time series into an E dimensional space with even time delays (τ, Eq. (1), SFig. 3 a). The embedding parameters can be set with prior knowledge of the dynamics or by other optimization methods. Such optimization methods are the first minimum or zerocrossing of the autocorrelation function (for delay selection), the false nearest neighbor method[41,42] or we applied the differential entropy based embedding optimizer[43].

3. kNN Neighbor search:

   We search for k-Neighborhoods around each datapoint in statespace and save the distance and temporal index of neighbors.

4. Compute TOF score:

   We compute the TOF score according to equation (2).

5. Apply a threshold θ on TOF score to detect unicorns (SFig. 3 c):

   The threshold can be established by prior knowledge, by clustering techniques or supervised learning. Because embedding causes uncertainty in the timing of the points and the threshold also has an event length meaning, we can extend the detections around the points to a maximum of half the event length or minimum by an half event length corresponding to the threshold k.

We implemented these steps in the python programming language (python 3). Time delay embedding and embedding parameter optimization was carried out by custom python scripts. We used the scikit-learn package[44] to calculate LOF was computed, and the neighbor search was established by the kd-tree algorithm of the scipy package[45].

Local Outlier Factor

The Local Outlier Factor[11] compares local density around a point (X) with the density around its neighbors (Eq. 9).

$$\mathrm{LOF}_k(X) = \frac{1}{|N_k(X)|} \sum_{o \in N_{\mathrm{k}(X)}} \frac{\mathrm{lrd}_k(o)}{\mathrm{lrd}_k(X)} \tag{9}$$

Where $|N_k(X)|$ is the cardinality of the k-distance neighborhood of X, lrd_k is the local reaching density for k-neighborhood (see Breunig et al.[11] for details, SFig. 3).

Generation of simulated datasets

Simulated logistic map and stochastical datasets

We simulated 4 systems: logistic map with linear tent map outlier segment, logistic map with linear outlier segment, simulated ECG data with tachycardia outlier segment and random walk with linear outlier segment. The first three datasets stem from deterministic dynamics, whereas the last simulated dataset has stochastic nature.

We generated 100 time series from each type, the length and the position of outlier segments were determined randomly in each case.

Logistic map with tent-map anomaly

100 instances of logistic map data-series were simulated ($N = 2000$) with one randomly (uniform) inserted outlier period in each dataset. The length of outlier periods was randomly chosen with length between $(2-200)$. The basic dynamics in normal conditions were governed by the update rule:

$$x_{t+1} = r x_t (1 - x_t) \tag{10}$$

where $r = 3.9$. The equation was changed during anomaly periods:

$$x_{t+1} = 1.59 - 2.15 \times |x_t - 0.7| - 0.9 \times x_t \tag{11}$$

where $a = \pm 0.001$. To make sure that the time series were bounded into the $I = [0,1]$ interval we changed the sign of a if it was required: for a start $a > 0$ and changes its sign to negative when $x_t >= 1$ condition met, thus bounced back into I.

Also to test the robustness against noise, noisy datasets were generated with Gaussian observational noise levels ranging from $\sigma = 0.005$ to 0.1 in 0.005 steps.

Logistic map with linear anomaly

The background generation process exhibited the logistic dynamics (Eq. 10) while the anomaly can be described by linear time dependence:

$$x_{t+1} = a * x_t + x_t \tag{12}$$

Here we used $a = \pm 0.001$, where the sign of the slope is positive by default and changes when the border of the $(0,1)$ domain is reached ensuring reflective boundary condition.

Random walk data with linear anomaly

We simulated 100 instances of multiplicative random walks with 2-200 timestep long linear outlier-insets. The generation procedure was as follows:

1. Generate w_i random numbers from a normal distribution with $\mu = 0.001$ and $sigma = 0.01$
2. Transform w_i to get the multiplicative random walk data as follows: $x_i = \prod_{j=1}^{i}(1+w_j)$
3. Draw the length (L) and position of outlier-section from discrete uniform distributions between $2-200$ and $1-(N-L)$ respectively.
4. Use linear interpolation between the section-endpoint values.

Simulated ECG datasets with tachyarrhythmic segments

We generated artificial ECG data series according to the model of Ryzhii and Ryzhii[31]. The pacemakers of the heart: the sinoatrial node (SA), the atroventricluar node (AV) and the His-Purkinje system (HP) are simulated by van der Pol equations:

$$SN \begin{cases} \dot{x}_1 &= y_1 \\ \dot{y}_1 &= -a_1 y_1 (x_1 - u_{11})(x_1 - u_{12}) - f_1 x_1 (x_1 + d_1)(x_1 + e_1) \end{cases} \tag{13}$$

$$AV \begin{cases} \dot{x}_2 &= y_2 \\ \dot{y}_2 &= -a_2 y_2 (x_2 - u_{21})(x_2 - u_{22}) - f_2 x_2 (x_2 + d_2)(x_2 + e_2) + K_{SA-AV}(y_1^{\tau_{SA-AV}} - y_2) \end{cases} \tag{14}$$

$$HP \begin{cases} \dot{x}_3 &= y_3 \\ \dot{y}_3 &= -a_3 y_3 (x_3 - u_{31})(x_1 - u_{32}) - f_3 x_3 (x_3 + d_3)(x_3 + e_3) + K_{AV-HP}(y_2^{\tau_{AV-HP}} - y_3) \end{cases} \tag{15}$$

where the parameters were set according to Ryzhii[31]: $a_1 = 40$, $a_2 = a_3 = 50$, $u_{11} = u_{21} = u_{31} = 0.83$, $u_{12} = u_{22} = u_{32} = -0.83$, $f_1 = 22$, $f_2 = 8.4$, $f_3 = 1.5$, $d_1 = d_2 = d_3 = 3$, $e_1 = 3.5$, $e_2 = 5$, $e_3 = 12$ and $K_{\mathrm{SA-AV}} = K_{\mathrm{AV-HP}} = f_1$.

The following FitzHugh-Nagumo equations describe the atrial and ventricular muscle depolarization and repolarization responses to pacemaker activity:

$$\text{P wave} \begin{cases} \dot{z}_1 = k_1(-c_1 z_1 (z_1 - w_{11})(z_1 - w_{12}) - b_1 v_1 - d_1 v_1 z_1 + I_{\mathrm{AT_{De}}}) \\ \dot{v}_1 = k_1 h_1 (z_1 - g_1 v_1) \end{cases} \tag{16}$$

$$\text{Ta wave} \begin{cases} \dot{z}_2 = k_2(-c_2 z_2 (z_2 - w_{21})(z_2 - w_{22}) - b_2 v_2 - d_2 v_2 z_2 + I_{\mathrm{AT_{Re}}}) \\ \dot{v}_2 = k_2 h_2 (z_2 - g_2 v_2) \end{cases} \tag{17}$$

$$\text{QRS} \begin{cases} \dot{z}_3 = k_3(-c_3 z_3 (z_3 - w_{31})(z_2 - w_{32}) - b_3 v_3 - d_3 v_3 z_3 + I_{\mathrm{VN_{De}}}) \\ \dot{v}_3 = k_3 h_3 (z_3 - g_3 v_3) \end{cases} \tag{18}$$

$$\text{T wave}\begin{cases}\dot{z}_4 = k_4(-c_4 z_4(z_4 - w_{41})(z_4 - w_{42}) - b_4 v_4 - d_4 v_4 z_4 + I_{\mathrm{VN_{Re}}}) \\ \dot{v}_4 = k_4 h_4(z_4 - g_4 v_4)\end{cases} \tag{19}$$

where $k_1 = 2\times10^3$, $k_2 = 4\times10^2$, $k_3 = 10^4$, $k_4 = 2\times10^3$, $c_1 = c_2 = 0.26$, $c_3 = 0.12$, $c_4 = 0.1$ $b_1 = b_2 = b_4 = 0$, $b_3 = 0.015$, $d_1 = d_2 = 0.4$, $d_3 = 0.09$, $d_4 = 0.1$, $h1 = h2 = 0.004$, $h_3 = h_4 = 0.008$, $g_1 = g_2 = g_3 = g_4 = 1$, $w11 = 0.13$, $w_{12} == w_{22} = 1$, $w_{21} = 0.19$, $w_{31} = 0.12$, $w_{32} = 0.11$, $w_{41} = 0.22$, $w_{42} = 0.8$.

The input-currents (I_i) are caused by pacemaker centra.

$$I_{\mathrm{AT_{De}}} = \begin{cases}0 & \text{for} \quad y_1 \le 0 \\ K_{\mathrm{AT_{De}}} y_1 & \text{for} \quad y_1 > 0\end{cases} \tag{20}$$

$$I_{\mathrm{AT_{Re}}} = \begin{cases}-K_{\mathrm{AT_{Re}}} y_1 & \text{for} \quad y_1 \le 0 \\ 0 & \text{for} \quad y_1 > 0\end{cases} \tag{21}$$

$$I_{\mathrm{VN_{De}}} = \begin{cases}0 & \text{for} \quad y_3 \le 0 \\ K_{\mathrm{VN_{De}}} y_3 & \text{for} \quad y_3 > 0\end{cases} \tag{22}$$

$$I_{\mathrm{VN_{Re}}} = \begin{cases}-K_{\mathrm{VN_{Re}}} y_3 & \text{for} \quad y_3 \le 0 \\ 0 & \text{for} \quad y_3 > 0\end{cases} \tag{23}$$

where $K_{\mathrm{AT_{De}}} = 4\times10^{-5}$, $K_{\mathrm{AT_{Re}}} = 4\times10^{-5}$, $K_{\mathrm{VN_{De}}} = 9\times10^{-5}$ and $K_{\mathrm{VN_{Re}}} = 6\times10^{-5}$.

The net ECG signal is given by the weighted sum of muscle depolarization and repolarization responses:

$$ECG = z_0 + z_1 - z_2 + z_3 + z_4 \tag{24}$$

where $z_0 = 0.2$ is a constant offset.

We simulated 100 instances of $t = 100$ seconds long ECG data with base rate parameter chosen from a Gaussian distribution ($f_1 \sim \mathcal{N}(\mu = 22, \sigma = 3)$). We randomly inserted $2-20$ seconds long fast heart-beat segments by adjusting the rate parameter ($f_1 \sim \mathcal{N}(\mu = 82, \sigma = 3)$). The simulations were carried out by the ddeint python package, with simulation time-step $\Delta t = 0.001$ from random initial condition and warmup time of 2 seconds. Also, a $10\times$ rolling-mean downsampling were applied on the data series before analysis.

Generating non-unique anomalies dataset

To show the selectiveness of TOF for the detection of unicorns, we simulated logistic map data with two tent-map outlier segments. The governing equations were the same as in the previous section, but instead of one, we placed randomly two not overlapping outlier segments into the time series during data generation, ($N = 2000, L = 20-200$).

Analysis steps on simulations

We applied optional preprocessing, embedded the time series, then applied TOF and LOF with specific threshold value and neighborhood size on all simulated datasets.

For logistic map with tent map anomaly datasets we applied no preprocessing. Since the generated dynamics is known and simple, we could chose embedding parameters manually, without any sophisticated selection criteria. The dynamics is approximately one dimensional ($d = 1$), therefore $E = 3$ is sufficient to embed the dynamics according to Takens theorem ($2d + 1 = 3$). Also for this discrete time dynamics $\tau = 1$ is a good choice of embedding delay. The expected anomaly length was calculated from the simulations and the threshold was set accordingly, more specifically $M = 110$ time steps (Eq. 25).

$$M = \frac{L_{\max} + L_{\min}}{2} = \frac{200 + 20}{2} = 110 \tag{25}$$

Where L is the possible length of inserted anomalies. Since this event length is the 5.5 % of the whole data series' length, we set LOF threshold to detect the top 5.5 % of points based on LOF score. The neighborhood size was scanned between 1 and 100 points and ROC AUC were computed on the range for TOF and LOF. Also the F_1, precision and recall score were calculated, where $k_{\mathrm{TOF}=4}$ neighbors and the best neighborhood size based on ROC AUC were selected for TOF and LOF respectively. We used this setting for k_{TOF} as an uninformed rule of thumb, since $E + 1$ points define a simplex in embedding-space determining unequivocally the position of the neighborhood's centre. For LOF we used $k_{\mathrm{LOF}} = 28$. This parameter setting scheme was used for all simulated datasets.

For the logistic map with linear anomaly segment datasets we set the parameters as in the tent map anomaly's case, the only difference is that the neighborhood size for LOF was set to $k_{\mathrm{LOF}} = 28$ as the maximal place of ROC AUC.

For the simulated ECG data we applied a tenfold downsampling, thus sampling period became $\Delta t = 0.01$ s. The dynamics seems approxiamtely 2 dimensional, so we set $E = 3$, which may be enough to reconstruct the dynamics, also $\tau = 0.01$ s was set as embedding delay. The threshold was set to $M = 11$ and 11%s as the expectation of tachycardia event length. The neighborhood size was scanned for ROC AUC values, and $k = 4$ and $k = 99$ was used to compute F_1 score, precision and recall for TOF and LOF respectively.

We applied a logarithmic difference for the multiplicative random walk with linear anomaly dataset as preprocessing step in order to get rid of unstationarity in the time series (Eq. 26).

$$y_t = \log(x_t) - \log(x_{t-1}) \tag{26}$$

where x is the original time series, log is the natural logarithm and y is the preprocessed time series. Embedding parameters was set to $E = 3$ and $\tau = 1$ arbitrarily, since the basic dynamics is stochastic. We set the event threshold parameter to $M = 110$ and 5.5% for tof and LOF respectively, also, we used $k_{\mathrm{TOF}} = 4$ and $k_{\mathrm{LOF}} = 1$ as the later was the maximal of a ROC AUC performance scanning.

On the logistic map with two tentmap anomalies dataset we did not apply any preprocessing, we set embedding parameters to $E = 3$ and $\tau = 1$. Also the neighborhood size was set to $k_{\mathrm{TOF}} = 4$ and $k_{\mathrm{LOF}} = 28$ for TOF and LOF respectively. We calculated the ROC AUC values for each simulated instance and plotted these values in the function of inter event interval (Fig. 4).

Model Evaluation metrics

We used precision, recall, F_1 score and ROC-AUC to evaluate the detection-performance on the simulated datasets.

The precision metrics characterizes that how precise the detections were:

$$\text{precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}} \tag{27}$$

The recall evaluates that what fraction of the would be detected points were actually detected:

$$\text{recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}} \tag{28}$$

As a mixture of precision and recall F_1 score provides a single scalar to rate model performance:

$$F_1 = 2\,\frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{29}$$

As an alternative evaluation metrics we applied the area under Receiver Operating Characteristic curve[46]. The ROC curve consists of point-pairs of True Positive Rate (recall) and False Positive Rate parametrized by a threshold (α, Eq. 30).

$$ROC(\alpha) := (\text{precision}(\alpha), \text{FPR}(\alpha)) \tag{30}$$

where $\text{FPR} = 1 - \text{recall}$ is the false positive rate and $\alpha \in [-\infty, \infty]$.

We computed the median and median absolute deviance from the 100 simulations on each simulated datasets (Fig. 3).

Analysis of real-world data

Polysomnography dataset

We analysed a part of the first recording of the MIT-BIH polysomnographic database[32] on Physionet[33]. The ECG data was sampled at 250 Hz. At the 300. s of the recording we selected a 160 s long segment to be analysed. The embedding parameters were set by a manual procedure to $E_{\text{TOF}} = 3$ and $\tau = 0.02$ s. The embedding delay was set according to the first zero-crossing of the autocorrelation function, embedding dimension was determined by an iterative embedding process, where the intrinsic dimensionality[47] of the dataset was measured for various embedding dimensions (SFig 5). The embedding dimension where the intrinsic dimensionality started to saturate were selected. For LOF, the embedding dimension was set higher ($E_{\text{LOF}} = 7$), because the results became more informative about the apnoe event. The neihgborhood size was set according to simulation results, we used a smaller neighborhood for TOF ($k = 11$) and a large neighborhood for LOF ($k = 200$). Moreover we set the event length to $M = 5$ s for TOF, which would correspond to 3.125% for LOF, which turned out to be a too loose condition. Therefore we used the more conservative 0.5% threshold for LOF to get more informative results.

Gravitational wave dataset

We analysed the 4096 Hz sampling rate strain data of the LIGO Hanford (H1) detector around the GW150914 merger event. The analysed 12 s recording starts 10 s before the event. We investigated the q transform spectogram of the time series around the event at 5×10^{-4} s time resolution by using the gwpy python package (https://doi.org/10.5281/zenodo.3598469). Based on the spectogram we applied 50-300 Hz bandpass filtering on the time series as a preprocessing step. Embedding parameters were selected manually (SFig. 6), by choosing the first minima of the autocorrelation function for the embedding delay ($\tau = 8$ sampling periods ≈ 1.95 ms) and then we selected the embedding dimension according to a manual procedure. Successive embedding of the time series into higher and higher dimensional space showed,

that the intrinsic dimensionality of the dataset starts to deviate from the embedding dimension at $E = 6$. Thus, we set this latter value as embedding dimension for TOF. For LOF a higher embedding dimension ($E = 11$) led to informative results. We set the neighborhood sizes based on our experiences with the simulated datasets: smaller vale were set for TOF ($k = 12$) and larger for LOF ($k = 100$). The event length was set to $M = 146$ ms for TOF as the visible length of the chirp on the spectogram and 0.5 % for LOF. Also, a $w = 7$ widening window were applied on the TOF detections.

LIBOR dataset

The monthly LIBOR dataset was analysed to identify interesting periods. As a preprocessing step, the first difference was applied for detrending purposes.

Optimal Embedding parameters were selected according to the minima of the relative entropy ($E = 3$, $\tau = 1$ month, SFig. 7-8). The neighborhood size was set manually to $k_{\mathrm{TOF}} = 5$ and $k_{\mathrm{TOF}} = 30$ for TOF and LOF respectively. Also, the event length was $M = 30$ for TOF and the threshold was set to 18.86 % for LOF. Also, a widening window $w = 3$ was applied on the TOF detections.

References

1. Chandola, V., Banerjee, A. & Kumar, V. Anomaly detection: A survey. *ACM Computing Surveys* **41**, 1–58 (2009). URL http://portal.acm.org/citation.cfm?doid=1541880.1541882. DOI 10.1145/1541880.1541882. arXiv:1011.1669v3.

2. Taleb, N. N. *The Black Swan: The Impact of the Highly Improbable* (2007).

3. Sornette, D. Dragon-Kings, Black Swans and the Prediction of Crises. *International Journal of Terraspace Science and Engineering* **2**, 1–18 (2009). URL https://arxiv.org/abs/0907.4290.

4. Hodge, V. J. & Austin, J. A Survey of Outlier Detection Methodologies. *Artificial Intelligence Review* **22**, 85–126 (2004). URL https://doi.org/10.1007/s10462-004-4304-y. DOI 10.1007/s10462-004-4304-y.

5. Pimentel, M. A. F., Clifton, D. A., Clifton, L. & Tarassenko, L. A review of novelty detection. *Signal Processing* **99**, 215–249 (2014). URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-84893296219{&}doi=10.1016{%}2Fj.sigpro.2013.12.026{&}partnerID=40{&}md5=0d3380f6f44195f22684bd2e83d06cd8. DOI 10.1016/j.sigpro.2013.12.026.

6. Chalapathy, R. & Chawla, S. Deep Learning for Anomaly Detection: A Survey 1–50 (2019). URL http://arxiv.org/abs/1901.03407. 1901.03407.

7. A survey of deep learning-based network anomaly detection. *Cluster Computing* (2019). DOI 10.1007/s10586-017-1117-8.

8. Braei, M. & Wagner, S. Anomaly Detection in Univariate Time-series: A Survey on the State-of-the-Art (2020). URL http://arxiv.org/abs/2004.00433. 2004.00433.

9. Using machine learning to predict extreme events in complex systems. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 52–59 (2020). DOI 10.1073/pnas.1917285117.

10. Abbott, B. P. *et al.* Observation of Gravitational Waves from a Binary Black Hole Merger. *Physical Review Letters* **116**, 061102 (2016). URL https://link.aps.org/doi/10.1103/PhysRevLett.116.061102. DOI 10.1103/PhysRevLett.116.061102.

11. LOF: Identifying density-based local outliers. *SIGMOD Record (ACM Special Interest Group on Management of Data)* (2000). DOI 10.1145/335191.335388.

12. Oehmcke, S., Zielinski, O. & Kramer, O. Event detection in marine time series data. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **9324**, 279–286 (2015).

13. Packard, N. H., Crutchfield, J. P., Farmer, J. D. & Shaw, R. S. Geometry from a Time Series. *Physical Review Letters* **45**, 712–716 (1980). URL https://link.aps.org/doi/10.1103/PhysRevLett.45.712. DOI 10.1103/PhysRevLett.45.712.

14. Takens, F. Detecting strange attractors in turbulence. *Dynamical Systems and Turbulence, Warwick 1980* **898**, 366–381 (1981). URL http://link.springer.com/10.1007/BFb0091924. DOI 10.1007/BFb0091924. arXiv:1011.1669v3.

15. Hegger, R., Kantz, H., Matassini, L. & Schreiber, T. Coping with nonstationarity by overembedding. *Physical Review Letters* **84**, 4092–4095 (2000). DOI 10.1103/PhysRevLett.84.4092.

16. Ye, H. *et al.* Equation-free mechanistic ecosystem forecasting using empirical dynamic modeling. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E1569–E1576 (2015). DOI 10.1073/pnas.1417063112.

17. Schreiber, T. & Kaplan, D. T. Nonlinear noise reduction for electrocardiograms. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **6**, 87–92 (1996). URL http://aip.scitation.org/doi/10.1063/1.166148. DOI 10.1063/1.166148.

18. Hamilton, F., Berry, T. & Sauer, T. Ensemble Kalman filtering without a model. *Physical Review X* (2016). DOI 10.1103/PhysRevX.6.011021.

19. Sugihara, G. *et al.* Detecting causality in complex ecosystems. *Science (New York, N.Y.)* **338**, 496–500 (2012). DOI 10.1126/science.1227079.

20. Benkő, Z. *et al.* Causal relationship between local field potential and intrinsic optical signal in epileptiform activity in vitro. *Scientific Reports* **9**, 1–12 (2019). DOI 10.1038/s41598-019-41554-x.

21. Selmeczy, G. B. *et al.* Old sins have long shadows: climate change weakens efficiency of trophic coupling of phyto- and zooplankton in a deep oligo-mesotrophic lowland lake (Stechlin, Germany)—a causality analysis. *Hydrobiologia* (2019). DOI 10.1007/s10750-018-3793-7.

22. Benkő, Z. *et al.* Complete Inference of Causal Relations between Dynamical Systems 1–43 (2018). 1808.10806.

23. Kennel, M. B. Statistical test for dynamical nonstationarity in observed time-series data. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* **56**, 316–321 (1997). DOI 10.1103/PhysRevE.56.316. 9512005.

24. Rieke, C. *et al.* Measuring Nonstationarity by Analyzing the Loss of Recurrence in Dynamical Systems. *Physical Review Letters* **88**, 4 (2002). DOI 10.1103/PhysRevLett.88.244102.

25. Gao, J. B. Recurrence time statistics for chaotic systems and their applications. *Physical Review Letters* **83**, 3178–3181 (1999). DOI 10.1103/PhysRevLett.83.3178.

26. Carletti, T. & Galatolo, S. Numerical estimates of local dimension by waiting time and quantitative recurrence. *Physica A: Statistical Mechanics and its Applications* **364**, 120–128 (2006). DOI 10.1016/j.physa.2005.10.003.

27. MARWAN, N., CARMENROMANO, M., THIEL, M. & KURTHS, J. Recurrence plots for the analysis of complex systems. *Physics Reports* **438**, 237–329 (2007). URL https://linkinghub.elsevier.com/retrieve/pii/S0370157306004066. DOI 10.1016/j.physrep.2006.11.001.

28. Gao, J. & Hu, J. Fast monitoring of epileptic seizures using recurrence time statistics of electroencephalography. *Frontiers in Computational Neuroscience* **7**, 1–8 (2013). DOI 10.3389/fncom.2013.00122.

29. Martínez-Rego, D., Fontenla-Romero, O., Alonso-Betanzos, A. & Principe, J. C. Fault detection via recurrence time statistics and one-class classification. *Pattern Recognition Letters* **84**, 8–14 (2016). DOI 10.1016/j.patrec.2016.07.019.

30. May, R. M. DOI 10.1038/261459a0.

31. Ryzhii, E. & Ryzhii, M. A heterogeneous coupled oscillator model for simulation of ECG signals. *Computer Methods and Programs in Biomedicine* **117**, 40–49 (2014). URL http://dx.doi.org/10.1016/j.cmpb.2014.04.009.

32. Ichimaru, Y. & Moody, G. B. Development of the polysomnographic database on CD-ROM. *Psychiatry and Clinical Neurosciences* (1999). DOI 10.1046/j.1440-1819.1999.00527.x.

33. Goldberger, A. L. *et al.* PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* (2000).

34. Abbott, R. *et al.* Open data from the first and second observing runs of Advanced LIGO and Advanced Virgo (2019). URL http://arxiv.org/abs/1912.11716. 1912.11716.

35. Department of Justice of The United States. Barclays bank PLC admits misconduct related to submissions for the london interbank offered rate and the euro interbank offered rate and agrees to pay $160 million penalty. https://www.justice.gov/opa/pr/barclays-bank-plc-admits-misconduct-related-submissions-london-interbank-offered-rate-and (2012).

36. Snider, C. & Youle, T. Diagnosing the libor: strategic manipulation member portfolio positions. *Working paper- faculty.washington.edu* (2009).

37. Snider, C. & Youle, T. Does the libor reflect banks' borrowing costs? *Social Science Research Network: SSRN.1569603* (2010).

38. Snider, C. & Youle, T. The fix is in: Detecting portfolio driven manipulation of the libor. *Social Science Research Network: SSRN.2189015* (2012).

39. Rieke, C., Andrzejak, R. G., Mormann, F. & Lehnertz, K. Improved statistical test for nonstationarity using recurrence time statistics. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* **69**, 9 (2004). DOI 10.1103/PhysRevE.69.046111.

40. Kriegel, H. P., Kröger, P., Schubert, E. & Zimek, A. LoOP: Local outlier probabilities. In *International Conference on Information and Knowledge Management, Proceedings* (2009).

41. Rhodes, C. & Morari, M. The false nearest neighbors algorithm: An overview. *Computers & Chemical Engineering* (1997). DOI 10.1016/S0098-1354(97)87657-0.

42. Krakovská, A., Mezeiová, K. & Budáčová, H. Use of False Nearest Neighbours for Selecting Variables and Embedding Parameters for State Space Reconstruction. *Journal of Complex Systems* **2015**, 1–12 (2015). DOI 10.1155/2015/932750.

43. Gautama, T., Mandic, D. P. & Van Hulle, M. M. A differential entropy based method for determining the optimal embedding parameters of a signal. In *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 6, 29–32 (2003).

44. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011).

45. Virtanen, P. *et al.* SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* **17**, 261–272 (2020). DOI https://doi.org/10.1038/s41592-019-0686-2.

46. Bradley, A. P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition* **30**, 1145–1159 (1997). DOI 10.1016/S0031-3203(96)00142-2.

47. Massoud Farahmand, A., Szepesvári, C. & Audibert, J.-Y. Manifold-adaptive dimension estimation. In *Proceedings of the 24th international conference on Machine learning - ICML '07*, 265–272 (ACM Press, New York, New York, USA, 2007).

Acknowledgements

We are grateful to prof. Róbert Gábor Kiss MD, PhD for his helpful comments on ECG data sets.

This research has made use of data, software and/or web tools obtained from the Gravitational Wave Open Science Center (https://www.gw-openscience.org), a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. LIGO is funded by the U.S. National Science Foundation. Virgo is funded by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale della Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by Polish and Hungarian institutes.

This research was supported by grants from the Hungarian National Research, Development and Innovation Fund NKFIH K 113147 and Human Brain Project associative grant CANON, under grant number NN 118902, and the Hungarian National Brain Research Program KTIA NAP 2017-1.2.1-NKP-2017-00002.

Author contributions statement

Z.S. designed the methods, Z.B., T.B. and Z.S. conceived the analysis, run the simulations and the analysis, wrote and reviewed the manuscript.

Additional information

Competing interests statement The authors of the paper state that they do not have any competing interest.

Appendix

Mean and variance for $q = 1$

The mean and the variance of TOF can be computed for uncorrelated noise in the continuous-time limit, where the typical properties of the metrics can be introduced. The expectation of the first neighbor is easy to compute (Eq. 31), if we take the probability density function ($p(\tau)$) as uniform, this is the assumption of white noise. Also, the pdf is independent of the rank of the neighbor (k), thus the mean is same for all neighborhood sizes. By the previous assumptions the mean is simply a quadratic expression:

$$\langle TOF_{q=1} \rangle \;\; = \int_0^T |t-\tau| \; p(\tau) \, \mathrm{d}\tau \; = \; \tfrac{1}{T} \int_0^T |t-\tau| \, \mathrm{d}\tau \; = \; \tfrac{t^2}{T} - t + \tfrac{T}{2} \tag{31}$$

with the method of moments, we calculate the variance for $k = 1$:

$$\langle TOF^2_{q=1} \rangle \;\; = \int_0^T (t-\tau)^2 \; p(\tau) \, \mathrm{d}\tau \; = \; \tfrac{1}{T} \int_0^T (t-\tau)^2 \, \mathrm{d}\tau \; = \; t^2 - tT + \tfrac{T^2}{3} \tag{32}$$

$$\sigma^2_{q=1} \;\; = \langle TOF^2_{q=1} \rangle - \langle TOF_{q=1} \rangle^2 \; = \; -\tfrac{t^4}{T^2} + \tfrac{2t^3}{T} - t^2 + \tfrac{T^2}{12} \tag{33}$$

if we have k neighbors, then the variance is reduced by a $1/k$ factor:

$$\sigma^2_{q=1,k} \;\; = \langle TOF^2_{q=1} \rangle - \langle TOF_{q=1} \rangle^2 \; = \; \tfrac{1}{k} \left(-\tfrac{t^4}{T^2} + \tfrac{2t^3}{T} - t^2 + \tfrac{T^2}{12} \right) \tag{34}$$

To test whether these theoretical arguments fits to data, we simulated random noise time series ($n = 100, T = 1000$) and computed mean TOF score and standard deviation (SFig. 1). We found, that theoretical formulas described perfectly the behaviour of TOF.

Mean and variance for $q = 2$

The exact statistics is hard to calculate, when the value of the q exponent is not equals to one, here we compute a vague approximation for $q = 2$. By computing the mean and variance for squared TOF, and taking the squareroot of these values can give a feeling about the properties of $\mathrm{TOF}_{q=2}$ respectively.

$$\left\langle \mathrm{TOF}^2_{\mathrm{noise},q=2} \right\rangle = \int_0^T (t-\tau)^2 \; p(\tau) \, \mathrm{d}\tau \; = \; \tfrac{1}{T} \int_0^T (t-\tau)^2 \, \mathrm{d}\tau \; = \; t^2 - tT + \tfrac{T^2}{3} \tag{35}$$

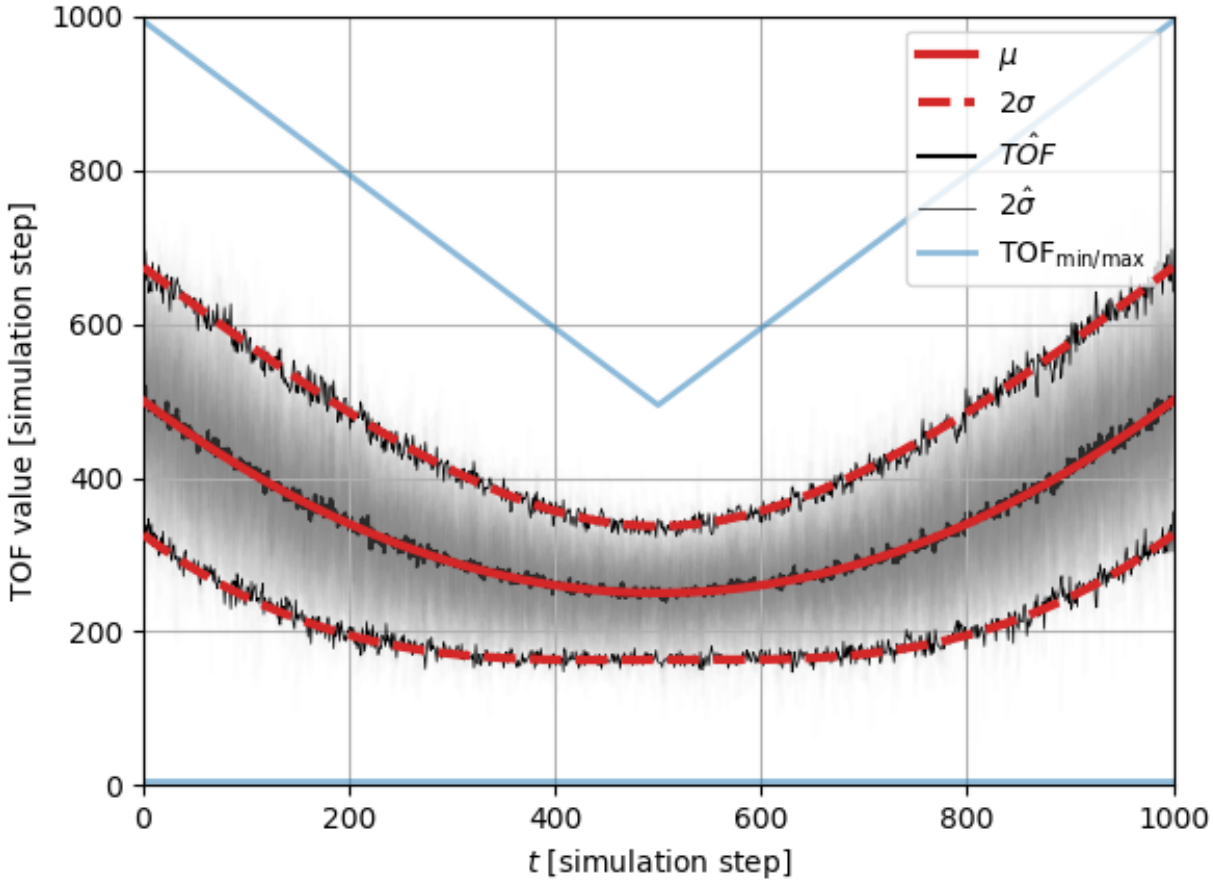
the second moment is as follows:

$$\left\langle \mathrm{TOF}^4_{\mathrm{noise},q=2} \right\rangle = \int_0^T (t-\tau)^4 \; p(\tau) \, \mathrm{d}\tau \; = \; \tfrac{1}{T} \int_0^T (t-\tau)^4 \, \mathrm{d}\tau \; = \; \tfrac{t^5 + (T-t)^5}{5T} \tag{36}$$

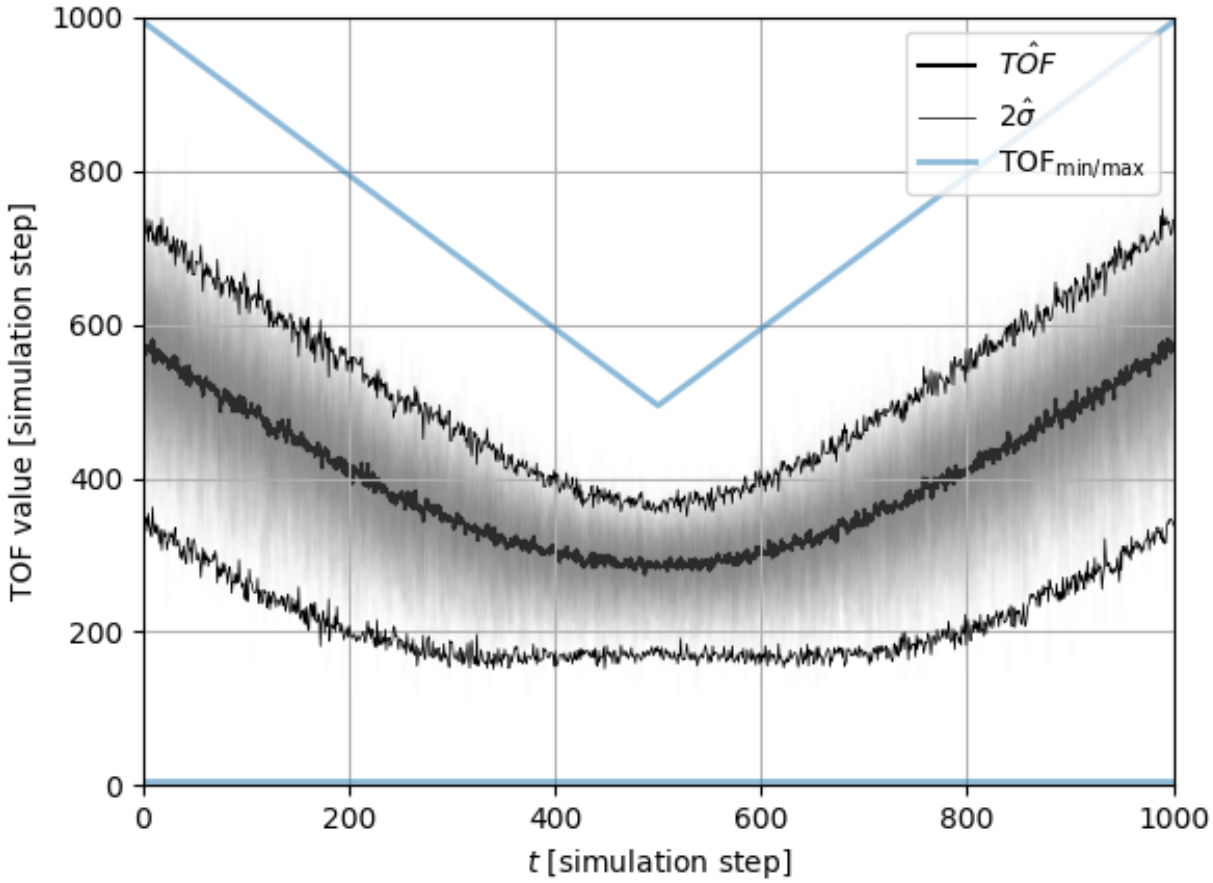
Thus using the method of moments we can get the variance of the $TOF^2_{q=2}$:

$$\mathrm{Var}\left(\mathrm{TOF}^2_{\mathrm{noise},q=2} \right) \; = \; \tfrac{t^5 + (T-t)^5}{5T} - \left(t^2 - tT + \tfrac{T^2}{3} \right)^2 \tag{37}$$

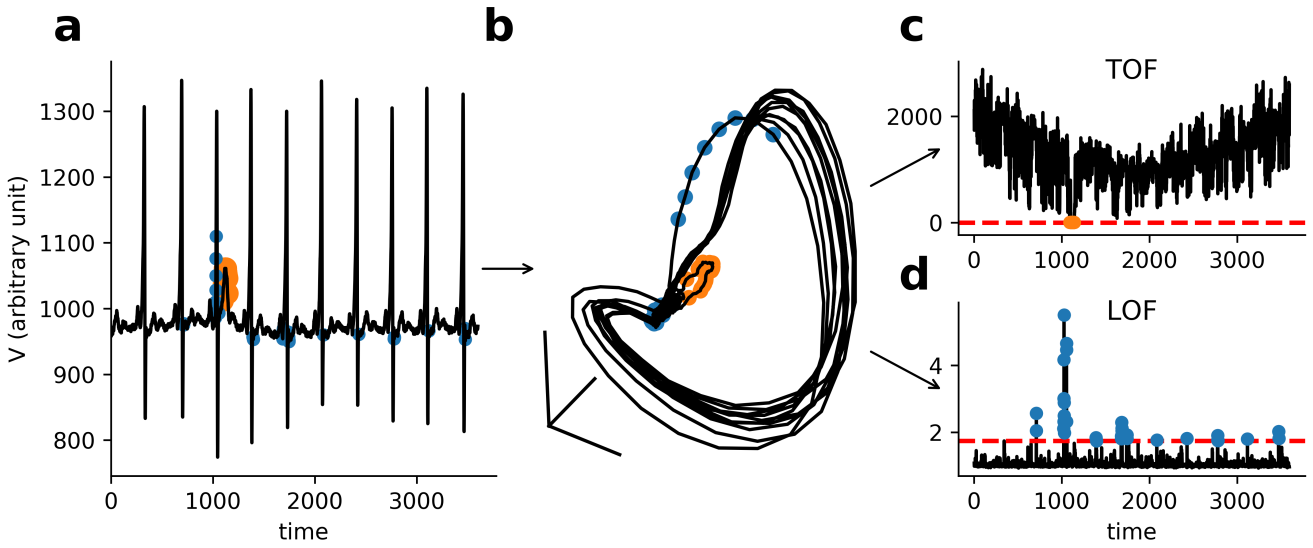
Additional Figures



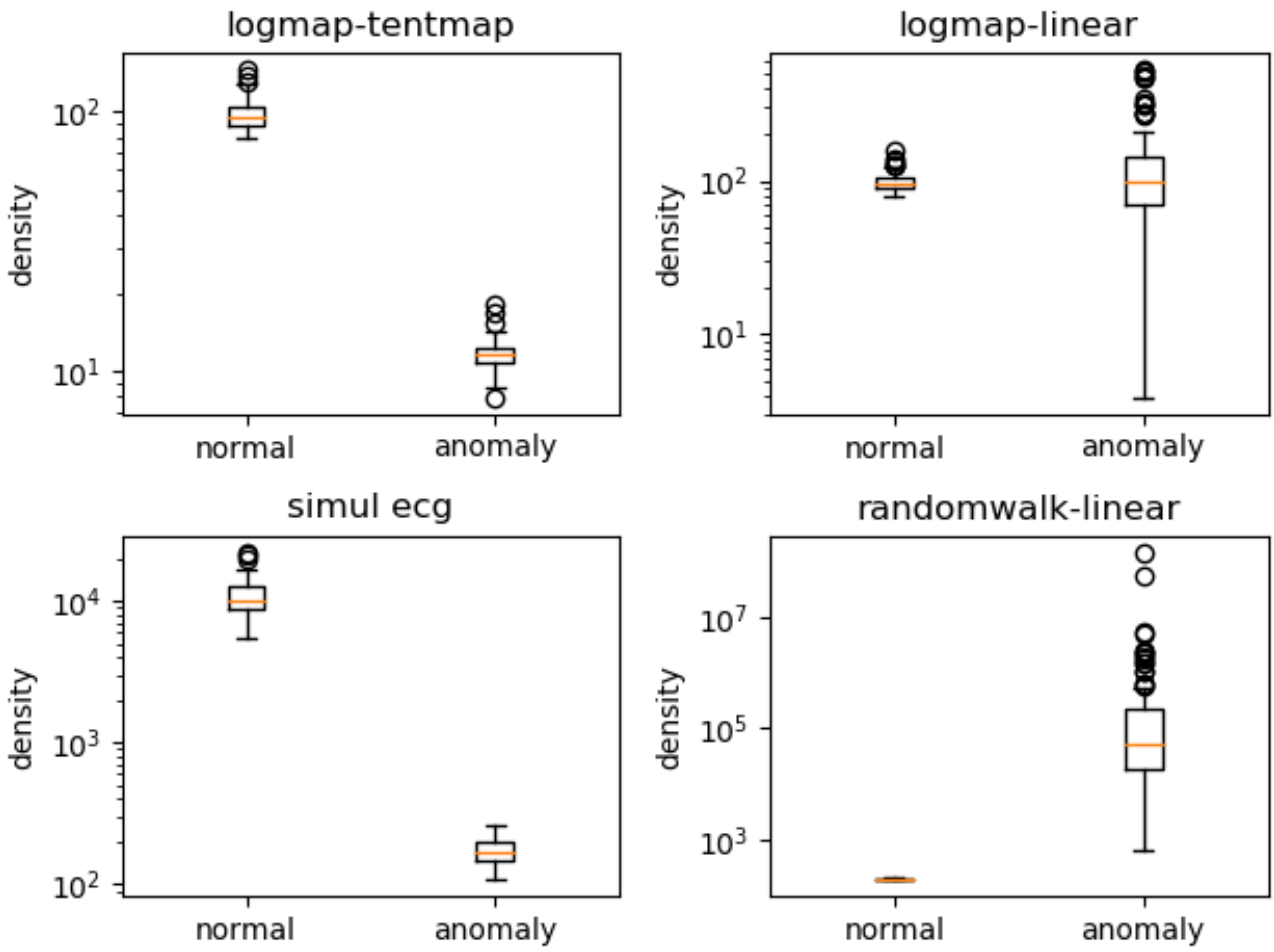
Supplementary Fig. 1. Properties of TOF for white noise data: theory and simulations. The expectation of TOF is computed as a function of temporal position in the time series ($q = 1$, thick red line), also the standard deviation was calculated (dashed red line). The average (thick black line) and standard deviation (thin black line) of $n = 100$ instances (grey shading). The minimal and maximal possible TOF vales are also charted (blue lines).



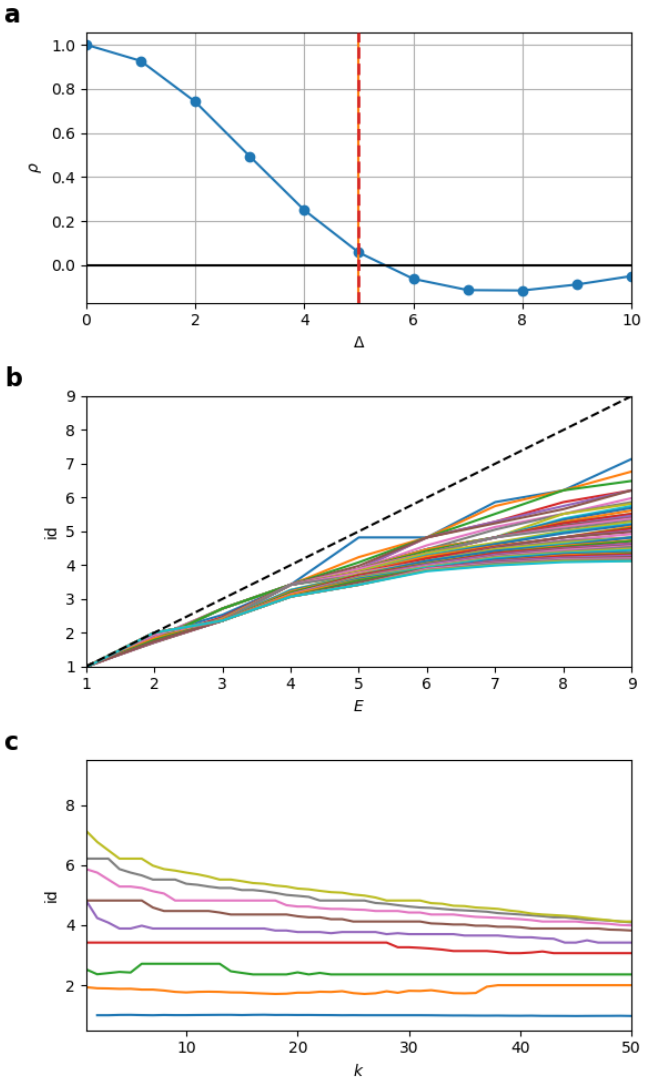
Supplementary Fig. 2. Properties of TOF for white noise data 2: simulations The baseline of TOF with $q = 2$. The average (thick black line) and standard deviation (thin black line) of $n = 100$ instances (grey shading).



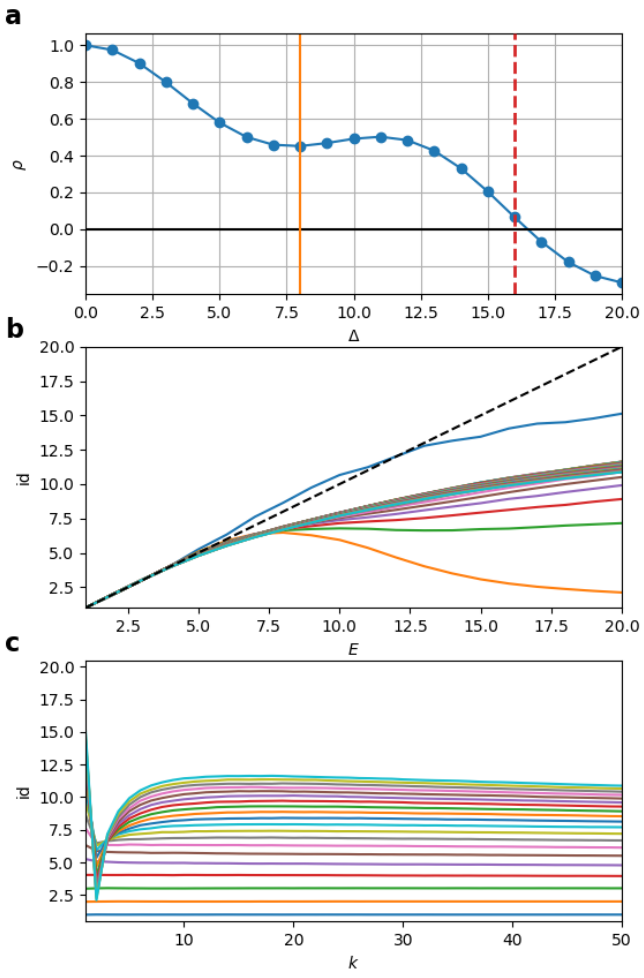
Supplementary Fig. 3. The worklow for TOF and LOF analysis for time series. (a) We start with a time series generated by a dynamical system, orange and blue marks TOF and LOF detections respectively. (b) As a next step of our analysis we apply time delay embedding, then kNN search in the reconstructed statespace. (c) We calculate TOF and LOF scores and apply thresholds on the outlier scores to detect anomalies.



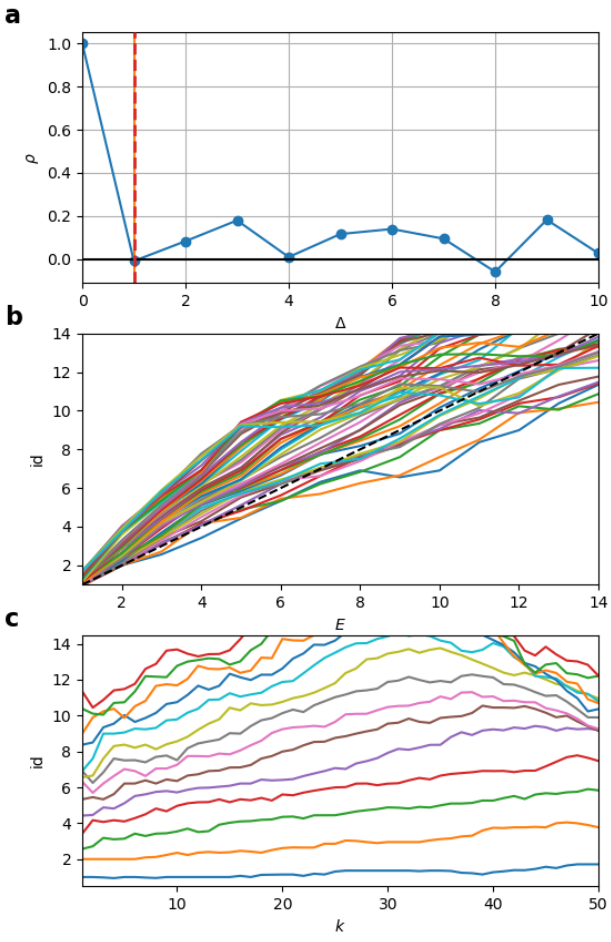
Supplementary Fig. 4. Point density in normal and anomlous data. The density around each data point is computed by the inverse distance from the k-th neighbor ($E = 3$, $\tau = 1$, $k = 20$).



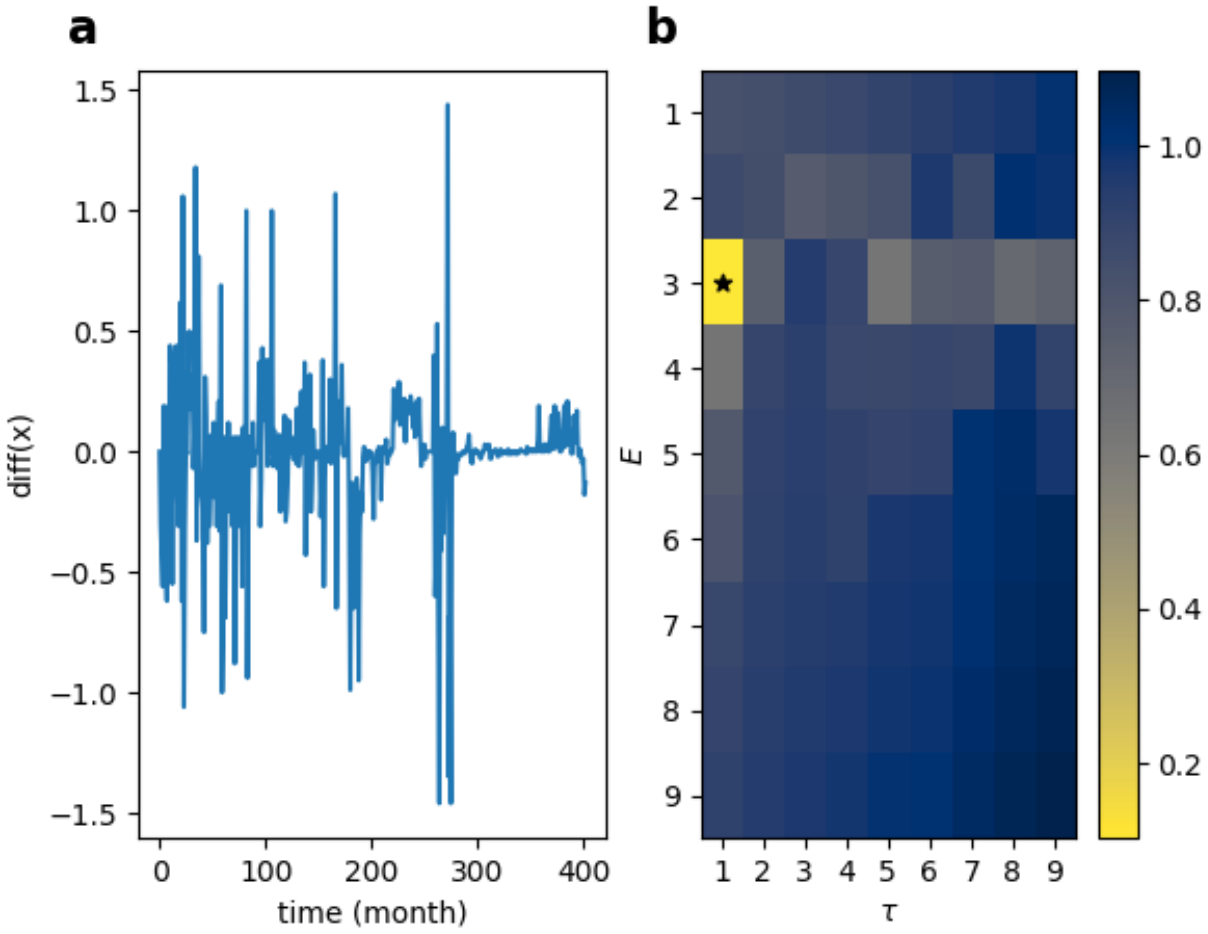
Supplementary Fig. 5. Embedding parameter selection for the polysomnography data with differential entropy. The minima of the entropy landscape marks the optimal embeddingparameters ($E = 3, \tau = 4$ timestep).



Supplementary Fig. 6. Embedding parameter selection for the gravitational wave data. (top) Embedding delay was selected ($\tau = 8$ sampling period) according to the first minima of the autocorrelation function. The first zeropoint was between 16 and 17 sampling periods. (middle) The intrinsic dimensionality is measured in the function of embedding dimension (E) for various neighborhood sizes. The dimesnion estimates start to deviate from the diagonal at $E = 5$. (bottom) Intrinsic dimensionality in the function of neighborhood size (k) for various embedding dimensions.



Supplementary Fig. 7. Autocorrelation and intrinsic dimension measurement of the preprocessed LIBOR time series.



Supplementary Fig. 8. Preprocessing and embedding parameter selection for the LIBOR time series with differential entropy. (left) The simple difference of the original time series were taken to detrend the data (right) The minima of the entropy landscape marks the optimal embeddingparameters ($E = 3, \tau = 1$ timestep).