Visualization of Incrementally Learned Projection Trajectories for Longitudinal Data

Tamasha Malepathirana[1], Damith Senanayake[1], Vini Gautam[2], Martin Engel[3,4], Rachelle Balez[3,4], Michael Lovelace[5,6], Gayathri Sundaram[5,6], Benjamin Heng[7], Sharron Chow[7], Chris Marquis[8], Gilles Guillemin[9], Bruce Brew[5,6,10], Chennupati Jagadish, Lezanne Ooi[3,4*] and Saman Halgamuge[1*]

[1]School of Electrical, Mechanical and Infrastructure Engineering, University of Melbourne, Melbourne, 3010, VIC, Australia.
[2]School of Chemical and Biomedical Engineering, University of Melbourne, Melbourne, 3010, VIC, Australia.
[3] Illawarra Health and Medical Research Institute, Northfields Avenue, Keiraville, 2522, NSW, Australia.
[4]School of Chemistry and Molecular Bioscience and Molecular Horizons, University of Wollongong, 2522, NSW, Australia.
[5]Applied Neurosciences Program, Peter Duncan Neurosciences Research Unit, St. Vincent's Centre for Applied Medical Research, 405 Liverpool St., Darlinghurst, Sydney, 2010, NSW, Australia.
[6] St Vincent's Clinical School, Faculty of Medicine and Health, UNSW Sydney, Sydney, 2010, NSW, Australia.
[7]Macquarie Medical School, Macquarie University Centre for Motor Neuron Disease Research, Sydney, 2109, NSW, Australia.
[8]School of Biotechnology and Biomolecular Sciences, UNSW Sydney, Sydney, 2052, NSW, Australia.
[9]Macquarie Medicine School, Faculty of Medicine, Health and Human Sciences, Macquarie University, Sydney, 2109, NSW, Australia.
[10]Department of Neurology and Immunology, St. Vincent's Hospital, Sydney, 2010, NSW, Australia.
[11]Research School of Physics, Australian National University, Canberra, 0200, ACT, Australia.

*Corresponding author(s). E-mail(s): lezanne@uow.edu.au; saman@unimelb.edu.au ;

Abstract

Human-induced pluripotent stem cell derived brain organoids provide models to study the human brain and its diseases. The neuronal activities of these organoids are recorded over time by Multi-electrode Arrays, calling for approaches to analyze their longitudinal variations to study the organoids' electrophysiological maturation. We introduce a pipeline (Il-Vis) that incrementally learns and visualizes a progression trajectory representing the organoids' electrophysiological properties. Unlike a static model that runs at the end of the experiment, Il-Vis generates a progression trajectory thus far at each sampling timepoint. Results on simulated data, for which the true progression trajectories are known, verified Il-Vis's ability to capture and visualize the trajectories accurately and relative to each other. Results on brain organoid data revealed useful insights about organoids' electrophysiological maturation stages and response patterns when exposed to Quinolinic Acid (and its blocking antibody), a metabolite contributing to many neuroinflammatory diseases including Alzheimer's.

Keywords: Multi-Electrode Arrays, Cortical Organoids, Alzheimer's, Machine Learning, Quinolinic Acid, Kynurenine Pathway

1 Introduction

Brain organoids derived from human-induced pluripotent stem cells (hiPSCs) facilitate the study of the human brain and the pathogenesis of brain diseases[1]. Hence, organoids are used to study the effects of potential drugs in many neurological diseases, including Alzheimer's disease(AD)[2].

Although existing genomic and transcriptomic analysis support the study of the cell types and developmental states of brain organoids[3, 4], the study of functional connectivity and network dynamics characterized by the electrophysiological properties is needed to realize the full potential of brain organoids[5]. Therefore, Multi-electrode Arrays (MEA) are of particular interest as they provide a high temporal resolution ($\simeq$25kHz) platform to study the electrophysiological properties of an organoid at a network scale[6–9]. MEAs contain multiple electrodes (9-4096) that record the extracellular activities of the surrounding neurons that vary with time, frequency, and amplitude thereby producing high-dimensional datasets. MEA data have been conventionally analyzed by manually extracting features about activity, such as spike rates, burst rates, etc.[10]. These hand-picked features albeit interpretable, may oversimplify and not adequately capture the complex dynamics of neuronal interactions in an organoid network. Consequently, a major challenge has been to develop new and innovative methods that extract complex dynamics from MEA data.

In this work, we present Il-Vis (Incrementally Learned Visualization Pipeline), a machine learning pipeline that visualizes the longitudinal progression of the electrophysiological properties of an organoid. We hypothesized that the progression of the electrophysiological properties of an organoid is a trajectory embedded in the high-dimensional MEA data. Our goal was to visualize this trajectory in a 2D space while keeping the complex relationships between electrodes and organoids (when modeling multiple organoids) intact to study and compare the organoids' electrophysiological progression. Hereafter, we refer to the original high-dimensional trajectory as T^H and the visualized 2D trajectory as T^L. A good visualization method should be able to project T^H to T^L as accurately as possible, implicitly capturing the time axis as shown in Fig. 1C.

Dimensionality reduction techniques attempt to map high-dimensional data (100-10k dimensions in this study) to a visually interpretable low-dimensional representation (2D/3D) while preserving the high-dimensional topological properties (e.g. connectedness and compactness). Static dimensionality reduction in which we present all data at once to the model can be easily applied in experiments where data is generated at only one sampling timepoint[11]. However, long-term experiments that run for several weeks or months generate data incrementally (progressively) at multiple discrete timepoints, challenging the direct application of static dimensionality reduction. Therefore, analysis is often conducted at a certain timepoint by training a static model using all data up to that timepoint. Expanding a static model in

this way has inherent limitations. First, integrating new data requires retraining a new model combining old and new data which is not desirable with large datasets. Second, even if we retrain a static model when new data arrives, the progression of visualizations obtained with each repetitive training may appear unsmooth and abrupt making it difficult to connect the visualizations. Therefore, Il-Vis combines incremental learning (the concept of training a model with currently available data and later integrating new data as they become available) with dimensionality reduction, eliminating the need for model retraining. Il-Vis learns a data increment (MEA data belonging to a sampling timepoint of an organoid) at-a-time (Fig. 1), generating visualizations of ongoing experiments with the existing data, and generating improved visualizations when new data of the same experiment is available. By tracing the resulting low-dimensional representations corresponding to each timepoint chronologically (Fig. 1C), we can then identify a trajectory of gradual changes followed by the electrophysiological properties of an organoid over the experimented period. Furthermore, this reduces the gap between the biological experiment and the data analysis by processing the data in the same order as they are recorded.

Nonlinear dimensionality reduction techniques such as t-Distributed Stochastic Nonlinear Embedding (t-SNE)[12] and Uniform Manifold Approximation and Projection (UMAP)[13] are designed as static dimensionality reduction techniques, often failing to insert new data points (an increment) to an already learned model without distorting the structure of the existing representation[14]. This makes it difficult to relate visualizations generated at different timepoints of the same experiment, particularly in longitudinal studies (Supplementary Fig.1). Recently, a new nonlinear dimensionality reduction method, Self-Organizing Nebulous Growths (SONG)[14] was proposed that showed superior performance in incremental dimensionality reduction by preserving the topology of previous increments when adapting visualizations to new data. Therefore, we used SONG as the dimensionality reduction technique in our proposed incremental pipeline.

First, we validated Il-Vis on a simulated dataset of high-dimensional nonlinear trajectories. We observed that Il-Vis projects the original trajectories onto progressive 2D visualizations while also capturing their trends and the relative similarities/dissimilarities. Il-Vis also showed robustness to varying degrees of noise in data and gaps between the increments. Then we applied Il-Vis to an experimental cortical organoid dataset derived from two AD and two healthy control cell lines. The generated visualizations were then used to obtain insights about organoids' influenced maturation when exposed to "Quinolinic Acid", a metabolite known to play a role in many neuroinflammatory diseases including AD[15] and, its blocking antibody: "monoclonal Quinolinic Acid Antibody". The organoids exposed to the antibody alone showed a distinct behavior from organoids in other treatments, regardless of the cell line. More importantly, organoids treated with both the Quinolinic Acid and its antibody together showed similar progression trajectories to the Untreated organoids, rather than

to the organoids treated with Quinolinic Acid or its antibody alone, suggesting that the antibody can normalize the phenotype.

2 Results

2.1 Il-Vis-Incrementally Learned Visualization framework

Il-Vis (Fig.1) incrementally learns a trajectory representing the progression of electrophysiological properties of an organoid over n sampling timepoints and visualizes it in 2D.

First, the signals from each sampling time point were segmented into 4-second segments ($t^{i,j}$; j^{th} timepoint of the ith sampling timepoint) where each segment was pre-processed using Fast Fourier Transformation (FFT)[16]. The resulting frequency components were represented as dimensions/features of the corresponding 4-second time segment. These were transformed to 2D using the incrementally trained dimensionality reduction model (Fig.1B) where SONG was used as the dimensionality reduction technique. The collection of pre-processed signal segments from the ith sampling timepoint was the ith data increment (I_i), represented as a table in Fig.1.

The model was trained for n sessions, one for each sampling timepoint. At session i, I_i was added to the existing data and all data including previous increments (D_i where $D_i = \{I_1 + ... + I_i\}$) were used to train the existing model parameters (θ^{i-1}) to obtain updated parameters (θ^i) producing a 2D visualization thus far (V_i). V_i visualizes all the increments from I_1 to I_i (Fig. 1C). The medians of the individual increments calculated using the representative 2D coordinates were then traced in their chronological order to obtain T^L that visualizes the progression of the organoid's electrophysiological properties over time. If n is sufficiently high, we would expect the consecutive and continuous data increments from an organoid to appear seamlessly connected, forming a continuous visualization. In absence of such a continuous stream of data, a good visualization method should provide a well-connected visualization even using discrete increments.

The pipeline can visualize data from multiple organoids i.e. we can jointly model multiple organoids' data in each sampling timepoint (see Figs.5A-D explained later). When a single organoid is modeled, the new data increments are visualized relative to its previous increments and the direction of the trajectory cannot be compared with another single organoid modeled independently. However, when two or more organoids are modeled together, the relative direction of a trajectory indicates whether the corresponding organoid's progression is similar/dissimilar to or moving towards/away from other organoids.

In this section, we report two sets of results using 1) a set of simulated datasets and 2) an experimental dataset of MEA recordings obtained from cortical organoids.

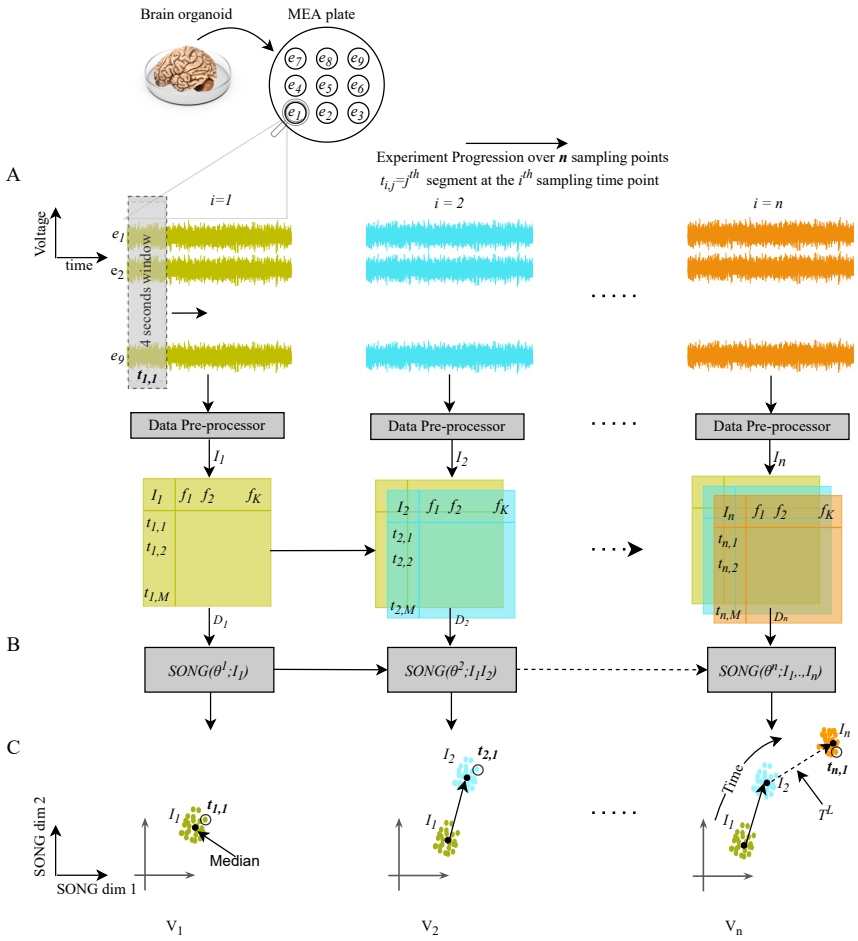

Fig. 1 Schematic representation of the proposed pipeline using (A) MEA signals from multiple "sampling timepoints" ("sessions" in the model's context) of the experiment, (B) incrementally learned dimension reduction model, and (C) 2D partial visualizations generated at each session (i) to cover the progression of electrophysiological properties of the modeled organoid up to the ith sampling timepoint. (A) The activity of a brain organoid is recorded by 9 electrodes (e_1 - e_9) of an MEA plate in n sampling timepoints. These recordings are segmented into 4-seconds segments and each segment; $t_{i,j}$ (j^{th} signal segment of the ith sampling timepoint) is pre-processed using FFT (Methods 4.6). The resulting frequency components for each segment are represented as the segment's features ($f_1, ..., f_k$). I_i is a collection of M pre-processed segments recorded at the ith sampling point ($I_i = \{t_{i,j}$ for $j\epsilon\{1,..,M\}\}$). The n increments corresponding to the n timepoints, are shown in different colors that are consistent throughout the diagram. (B) The dimensionality reduction model is trained for n sessions, each session corresponding to a sampling timepoint in the biological experiment. At session i, the model parameters θ^i are learned by updating the parameters from the previous session (θ^{i-1}) using $D_i = \{I_1 + ... + I_i\}$ (i.e. all the increments thus far), and a partial visualization V_i is generated. (C) Each data point in a partial visualization corresponds to a compact representation of a 4-second segment. The directed arrows map the medians of the increments learned so far in the chronological order to identify the trajectory T^L followed by the organoid's electrophysiological properties up to that point in time. Note that none of the axes corresponds to time but the time axis is implicit in the trajectory identified.

2.2 Simulated datasets

We validated Il-Vis using a series of simulated high-dimensional trajectories $\{T_r^H$ where $r\epsilon\{1, ..., 4\}\}$. The simulated trajectories were designed to be nonlinear and monotonically increasing (Methods 4.2.1), to simulate the monotonic progression of electrophysiological properties of an organoid, e.g. the increasing firing rate of an organoid with its increasing maturity[17] (before reaching full maturity in ideal conditions). Our goal was to validate that Il-Vis can capture these features and relative similarities between multiple simulated trajectories while producing progressive visualizations across the sessions. This validation may indicate that the deviations from such trends perhaps due to factors such as developmental pathology, aging of the organoid, or administered therapeutic stimuli can potentially be identified.

2.2.1 Il-Vis generated progressive visualizations capturing the underlying trends

The generated visualizations were progressive (Figs.2 and 3), i.e., in each experiment, the visualizations preserved the structure across incremental sessions, allowing to gain insights from incomplete or ongoing experiments. Theoretically, the experiment could be extended to any number of future increments and obtain continuing/incremental visualizations. The monotonic trend of the simulated trajectories is captured in the visualizations by visualizing each new increment I_i increasingly distant from its previous increments $I_1, .., I_{i-1}$. (Fig.2, see Supplementary Fig.1 for pairwise euclidean and geodesic distance heatmaps between the increments in T_r^H and T_r^L). The visualizations obtained when using the dimensionality reduction techniques PCA (Principle Component Analysis)[18] and UMAP are shown in Supplementary Fig.2, where PCA failed to preserve the nonlinearity of T_1^H and UMAP failed to generate progressive visualizations over the sessions. We further verified the manifold preservation quality of the SONG embedding (Supplementary Fig.3) and the method's capability in catering for an even higher dimensional space (up to 10000) (Supplementary Fig.4).

2.2.2 Il-Vis was robust to noise and gaps in the data

To replicate random noise arising from biological and non-biological variations in real data, random Gaussian noise (20%) was added to each increment in T_1^H (Methods 4.2.2). T_1^L still preserved T_1^H's trend and nonlinearity (Figs.2B-E). However, when the degree of noise was increased from 0-20% the smoothness of T_1^L decreased (Supplementary Fig.5).

Although the processes underlying the biological experiments are continuous, the recorded features are discrete and the sampling timepoints may have considerable gaps. To replicate this, we introduced minor, medium, and major gaps between the increments in T_1^H (Methods 4.2.2) (Figs.2C-E). T_1^L captured the monotonic trend of T_1^H despite such gaps. When major gaps were introduced (where $\delta = 99$), five different clusters were visualized each corresponding

to an increment (Fig.2E). This agreed with our expectation because Il-Vis had very limited information about the continuity. However, a monotonic and a nonlinear path along which the corresponding T_1^H follows could be identified from T_1^L, suggesting that Il-Vis is robust with datasets with considerable gaps and noise.

2.2.3 Il-Vis captured the relative similarity/dissimilarity between the trajectories

Observing the progression of an organoid's electrophysiological properties relative to another organoid's is of interest for neurodevelopmental and pharmacological studies. For instance, the difference between the maturity of organoids exposed and not exposed to a particular treatment would assist in understanding the treatment effectiveness. To obtain such insights, Il-Vis should capture and visualize the relative similarity and dissimilarity between the high-dimensional trajectories as they progress. For this validation, we jointly modeled and plotted two independent randomly generated nonlinear trajectories (T_1^H, and T_2^H) designed to originate at the same position while diverging at higher t values (Fig. 3A). As the new increments were added, T_1^L and T_2^L diverged from each other, exhibiting the increased difference towards the later increments.

We further modeled the same principal trajectories (T_1^H and T_2^H) together with two secondary trajectories T_3^H and T_4^H each of which was created to be closer to T_1^H and T_2^H respectively, based on a predefined similarity factor (α)(Methods 4.2.1). In Fig.3B, when α=0.8, by visualizing T_3^L closer to T_1^L than to T_2^L, and T_4^L closer to T_2^L than to T_1^L (Supplementary Table.1), Il-Vis captured each secondary trajectory's similarity to its principal trajectory and the dissimilarity to the other. Furthermore, in similar trajectory pairs, for instance, T_1^L and T_3^L, the trajectories became more separated in the later increments compared to the earlier increments, therefore capturing the subtle but important dissimilarities. In Fig. 3C, when α was increased from 0.6 to 0.9, T_3^L and T_4^L gradually visualized closer to T_1^L and T_2^L respectively. This together with previous experiments validated Il-Vis's capability in preserving the relative relationships of the high-dimensional trajectories in the obtained 2D visualizations.

2.3 Cortical organoids dataset

After validating Il-Vis on simulated data, it was then applied to a cortical organoid dataset. MEA data has been previously used to assess alterations in neuronal excitability, e.g. mediated by neuroinflammation[19]. Quinolinic Acid (QUIN) is an N-methyl-D-aspartate (NMDA) agonist generally synthesized and released during neuroinflammation. QUIN's over-activation of NMDA receptors leads to neuronal excitotoxicity, thus, QUIN has been implicated in a range of diseases in which neurodegeneration is a key component, including dementia, multiple sclerosis, and stroke[20–23]. QUIN antibodies

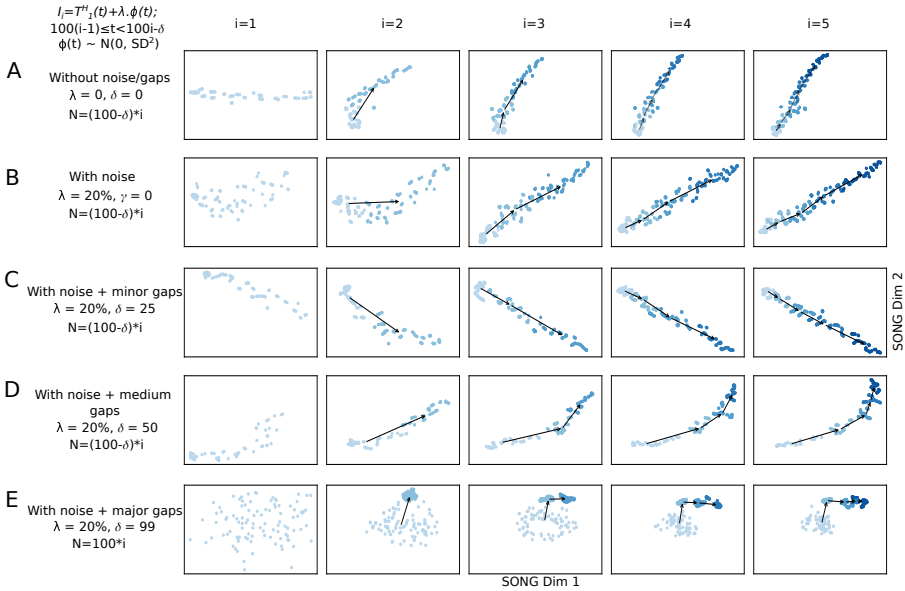

Fig. 2 2D visualizations obtained by the pipeline showed the method's capability in generating progressive visualizations capturing the monotonic and nonlinear trend of a simulated 100-dimensional nonlinear trajectory (T_1^H) (A) without noise/gaps, (B) with noise, (C) with noise and minor sampling gaps in between increments, (D) with noise and medium sampling gaps between increments, and (E) with noise and major sampling gaps between increments. T_1^H was defined on a pseudo-temporal parameter t (Equation 1). $\phi(t)$ is noise sampled from a Gaussian distribution with mean 0 and standard deviation SD (Methods 4.2.2). λ is the percentage of noise added: 0% for (A) and 20% for (B)-(E). i^{th} data increment (I_i) is a collection of data points generated using the equation T_1^H $+\lambda.\phi(t)$ where $100(i-1) \leq t \leq 100i-\delta$. The interval of t used in each experiment was restricted based on the sampling gap size (δ). The number of data points added in each new increment was determined based on δ and was 100, 100, 75, 50 and 100 for the experiments A-E respectively (see Methods 4.2.2). For each experiment, five data increments were formed (n=5). Each row contains the five visualizations generated at the five sessions where a new data increment is introduced to the pipeline at each session. At session i, all the increments thus far (i.e $D_i = \{I_1 + ... + I_i\}$) are visualized. In visualizations with multiple increments, different increments are displayed in varying intensities of the same color. The stronger the intensity; the newer the increment. By tracing the medians of the subsequent increments, T_1^L which is a 2D representation of the original trajectory (T_1^H) is identified. Each consecutive visualization has maintained the structure across the previous visualization allowing to relate between the visualizations. Furthermore, the progressiveness of the monotonic trajectory is captured by placing each new increment further away from the previous increments (Supplementary Fig.1). The pipeline shows robustness to noise (B) and minor to major gaps (C)-(E)

.

that prevent the binding of QUIN to NMDA receptors or its uptake, provide neuroprotective effects[24]. Therefore, reducing extracellular QUIN levels has been proposed to have therapeutic potential in a range of diseases, including AD. Previously[24], we showed that as little as 30 minutes of exposure to the same monoclonal Quinolinic Acid Antibody (QUIN Ab) used in this study could protect oligodendrocyte cell line monolayers from cell death induced by

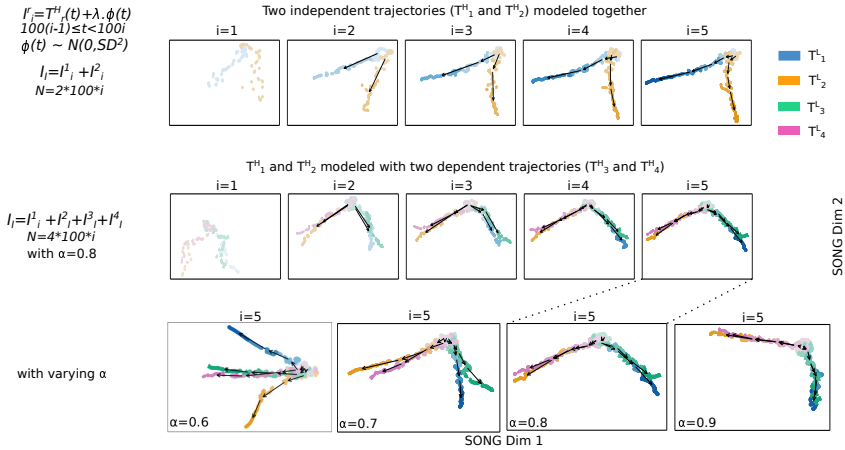

Fig. 3 2D visualizations obtained by the pipeline showed method's capability in capturing relative similarity/dissimilarity between (A) two trajectories; T^H_1 and T^H_2 and (B) four trajectories; T^H_1 and T^H_2 together with T^H_3 and T^H_4. T^H_1, T^H_2, T^H_3, and T^H_4 are 100-dimensional nonlinear trajectories. T^H_1 and T^H_2 are independent (principle trajecories) whereas T^H_3 and T^H_4 are dependent (secondary trajectories) on T^H_1 and T^H_2. Specifically, T^H_3 and T^H_4 were created as compositions of T^H_1 and T^H_2 where a similarity factor (α) was used to determine how closer a secondary trajectory is to a principal trajectory (see Equation 2). For instance, when $\alpha = 1$, T^H_3 and T^H_4 are equivalent to T^H_1 and T^H_2 respectively. $\phi(t)$ is noise sampled from a Gaussian distribution with mean 0 and standard deviation SD (Methods 4.2.2). λ is the percentage of noise added which is 0% for this experiment. I^r_i is the ith data increment of the trajectory T^H_r generated using the equation $T^H_r + \lambda.\phi(t)$ where $100(i-1) \leq t \leq 100i$. In each experiment, I_i is the collection of ith data increments of all trajectories used in that experiment (i.e. $I_i = I^1_i + .. + I^r_i$). Five increments of data points were formed (n=5). 100 data points were added in each increment per trajectory since no sampling gaps were introduced in between the increments. Thus, N corresponds to the total number of data points used in each visualization. Each experiment contains five visualizations generated at the five sessions where a new data increment is introduced to the pipeline at each session. At session i, all the increments thus far (i.e $D_i = \{I_1 + ... + I_i\}$) are used for training. In visualizations with multiple increments, different increments are displayed in varying intensities of the same color. The stronger the intensity; the newer the increment. By tracing the medians of the subsequent increments T^L_r which is a 2D representation of the original trajectory (T^H_r) was identified. (A) T^H_1 and T^H_2 originate at the same position and diverge from each other as the increments are added which is consistent with the definition of the trajectories (B) T^L_3 is visualized closer to T^L_1 than to T^L_2, and T^L_4 is visualized closer to T^L_2 than to T^L_1 capturing the relative similarity between the principle and secondary trajectories (C) Visualizations generated at the 5th session in four independent experiments with four different α values. The α was varied from 0.6 to 0.9 to observe the movement of T^L_3 and T^L_4 in reference to T^L_1 and T^L_2. T^L_3 and T^L_4 are visualized closer to T^L_1 and T^L_2 respectively as the α is increased. Altogether, visualizations show the pipeline's capability in capturing the relative similarity and dissimilarity between the trajectories.

QUIN. QUIN excitotoxicity and metabolism are very much associated with the aging process[25, 26], thus the long-term effect of QUIN Ab on QUIN should be studied. Therefore, we analyzed both the acute and chronic effects of exogenous QUIN Ab on both endogenous and exogenous QUIN using Il-Vis.

Our dataset contained MEA data from cortical organoids derived using four cell lines (biological replicates showing biological variation): two AD patients

(AD1 and AD2) and two healthy controls (CN1 and CN2). Four organoids were derived from each cell line exposed to one of the following four different treatments 1) Untreated, 2) QUIN added, 3) QUIN Ab added, and 4) both QUIN and QUIN Ab added. These 16 organoids from 4 cell lines were used to test the ability of QUIN Ab to sequester, and thus reduce the deleterious effects of QUIN. As expected, QUIN was also produced by the organoids endogenously and released into the media, where it was quantified at 16-30 nM (Supplementary, Fig.S6). Exogenous QUIN was added to the organoids in 2) at 100 nM (i.e. 3-5 times endogenous levels but lower than the circulating levels in blood[27]), to test the effect of increased QUIN on organoid activity. QUIN Ab was added alone in 3) to test its capability to sequester the endogenous levels of QUIN, and together with exogenous QUIN in 4) to test the QUIN Ab's capability to sequester the higher levels of QUIN. QUIN Ab was added at 10 μg/mL to reduce extracellular QUIN concentrations and prevent QUIN uptake and binding to NMDA receptors. Untreated organoids (1) were used as the controls for 2), 3), and 4), and to understand the normal maturation profile reflected in the plots.

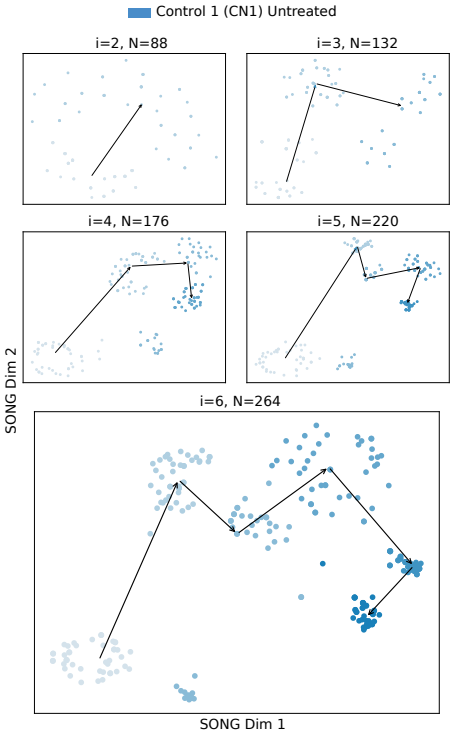

Fig. 4 2D visualizations obtained by the pipeline for the Untreated CN1 organoid when modeled independently. I_1 to I_6 are the six increments corresponding to the six different timepoints at which the recordings were taken (Table 3). At session i, all the increments thus far (i.e $D_i = \{I_1 + ... + I_i\}$) are used for training the model. N corresponds to the number of data points used in each visualization where each point corresponds to a compact representation of a 4-second recording obtained from the 9 electrodes in the respective MEA plate. The radius of a blue-filled circle is proportional to the number of high-dimensional data points mapped to the same x,y coordinate in the 2D representation. The 2D visualizations generated from the 2nd session ($i = 2$) onwards are shown in the sub-figures. The color intensity represents the organoid's maturity. At earlier increments, the color is lighter while in later increments the color is darker. The medians of the subsequent increments are connected to identify a trajectory that the organoid's electrophysiological properties have followed.

2.3.1 Il-Vis identified the progression of the electrophysiological properties of an organoid

The low-dimensional trajectories (T^L_{CN1}, T^L_{CN2}, T^L_{AD1}, T^L_{AD2}) corresponding to the progression of the electrophysiological properties of the Untreated CN1(T^H_{CN1}), CN2(T^H_{CN2}), AD1(T^H_{AD1}) and AD2(T^H_{AD2}) organoids were visualized independently using Il-Vis over the experimental period (Fig.4 and Supplementary Fig.7-9). The use of MEA signal segments instead of extracted spike characteristics (Supplementary Fig.10) allowed an in-depth representation of the change in an organoid's electrophysiological state over time. For instance, in the CN1 organoid (T^L_{CN1}), the 2nd, 3rd, and 4th increments were visualized as if they were moving away from the 1st (Fig.4). However, the 5th and 6th increments reversed the direction back and towards the 1st indicating that their states were becoming more similar to the initial state. As shown by the visualizations, we were able to visualize the independent low-dimensional trajectories that progressed through an implicit time axis.

2.3.2 Il-Vis visualized multiple organoids in the same context

We jointly visualized the four organoids (treated in different media) derived from the same cell line to visualize their relative electrophysiological progression. The final visualizations obtained for each cell line three months after the drugs being added were plotted in Fig.5. In these visualizations, the Untreated organoid trajectories that we saw in Fig.4, were adjusted automatically by the method relative to other organoids of the same cell line. When the increments of two trajectories are placed closer to each other as the new increments are added, it indicates that the electrophysiological states of the corresponding organoids are progressing similarly with their maturation. For instance, the corresponding increments in the Untreated and the QUIN+QUIN Ab treated trajectories in AD2 (Fig.5B) are placed closer to each other as new increments are added. To quantify this similarity over time, we used an Euclidean distance-based metric (Methods4.6.2).

2.3.3 The QUIN Ab treated organoid trajectory deviated from the other three conditions in all four cell lines

In each cell line, the visualized trajectory representing the organoid exposed to QUIN Ab, originated at a considerable distance from the trajectories corresponding to the organoids in the other three conditions (Untreated, QUIN, QUIN Ab+QUIN) as shown in Fig.5. This could be also confirmed by the quantitative distance measures given in Table1, wherein for all four cell lines, the QUIN Ab treated trajectory was the farthest to the Untreated organoid trajectory. Furthermore, the QUIN Ab treated organoid trajectory progressed in a different direction to the QUIN treated trajectory. For instance, in AD1 (Fig.5A), the QUIN treated trajectory progressed from E to W while the QUIN Ab treated trajectory progressed from SW to NE. Also, in AD2 (Fig. 5B), the QUIN treated trajectory progressed from NE to SW whereas the

QUIN Ab treated trajectory progressed from SW to NE in parallel but in the opposite direction. This suggests that QUIN Ab has a unique effect on electrophysiological activity compared to the other treatments.

2.3.4 Evidence for QUIN Ab sequestering QUIN

The distances from the QUIN Ab treated trajectory to the other three conditions were measured considering its unique behavior. The QUIN treated trajectory was visualized the farthest away from the QUIN Ab treated trajectory in cell lines AD1, AD2, and CN1 (Fig. 5A-C and Table 2). Furthermore, the distance ordering between the trajectories was consistent across these three cell lines (from Table 2) allowing us to derive the following ordering based on the trajectory similarity. QUIN Ab treated → Untreated → QUIN+QUIN Ab treated → QUIN treated meaning Untreated is the closest and QUIN treated is the farthest to/from QUIN Ab treated trajectory. By attributing this closeness to the level of the effective concentration of QUIN bathing the organoids, we could then conjecture that the order of effective concentration of QUIN in organoids to be QUIN Ab treated <Untreated <QUIN+QUIN Ab treated <QUIN treated in the cell lines discussed. QUIN+QUIN Ab being placed in between QUIN and QUIN Ab confirms QUIN Ab's capability to sequester the high levels of QUIN. Furthermore, the ordering indicates QUIN+QUIN Ab to have a stronger effect of QUIN than the Untreated organoid, suggesting that the used concentration of exogenous QUIN Ab is not sufficient to fully neutralize the effect of the exogenous QUIN added to the three cell lines.

For the remaining cell line (CN2), the relationship between the Untreated, QUIN treated, and QUIN+QUIN Ab treated trajectories were consistent with the other three cell lines, where QUIN treated trajectory was visualized closer to the QUIN+QUIN Ab treated trajectory than to Untreated trajectory by the pipeline. However, further investigation is needed to confirm the reason for QUIN Ab treated trajectory being visualized closer to QUIN treated trajectory than to the other two treatments and in particular to the Untreated. One possible explanation for this would be that the endogenous QUIN in the Untreated CN2 is much higher after nine months than the other three cell lines (Supplementary, Fig.6). In addition to the visualizations, the spiking activity of the MEA data was also plotted (Supplementary, Fig.10). However, with such an analysis a qualitative comparison of different treatments in organoids was challenging as it only showed an independent quantitative (spike frequency) variation over time.

3 Discussion

We proposed Il-Vis, a machine learning-based pipeline that can project and visualize high-dimensional trajectories embedded in longitudinal MEA data to a low-dimensional space, representative of organoids' electrophysiological properties, for the first time in literature. Il-Vis allows four key advantages

Table 1 Distance from the low-dimensional trajectory representing the Untreated organoid to low-dimensional trajectories representing the treated organoids

Treatment	AD1	AD2	CN1	CN2
QUIN Ab	**7.82**	**5.44**	**5.91**	**9.61**
QUIN	7.4	3.35	3.81	7.75
QUIN+QUIN Ab	5.93	2.05	3.98	3.03

Distance (Methods 4.6.2) from the Untreated organoid trajectory to the trajectories corresponding to the other three organoids of the same cell line after the final training session.

Table 2 Distance from the low-dimensional trajectory representing the QUIN Ab organoid to the low-dimensional trajectories representing the organoids in other media

Treatment	AD1	AD2	CN1	CN2
Untreated	7.82	5.44	5.91	**9.61**
QUIN	**11.10**	**6.06**	**7.99**	7.49
QUIN+QUIN Ab	9.45	5.62	7.47	8.23

Distance (Methods 4.6.2) from the QUIN Ab organoid trajectory to the trajectories corresponding to the other three organoids of the same cell line after the final training session.

that are important in the fields of neurodevelopmental, pharmacological, and disease modeling studies.

First, by being less reliant on human intervention for feature extraction, we overcome several weaknesses in conventional MEA analysis that manually extracts features such as spike rates, burst rates, etc.[28, 29]. Such methods require analyst expertise in choosing parameters and may introduce expert biases influencing all subsequent analysis tasks as discussed in[10]. Second, Il-Vis can model the currently available data in an ongoing experiment and incrementally improve the model as the new data arrives. This eliminates the requirement that the data used in the analysis should provide a complete picture of the underlying processes before beginning the modeling. Third, Il-Vis generates intuitive and interpretable 2D representations of progressions and deviations in the electrophysiological activity capitulating the complex biology of an organoid over time. Fourth, Il-Vis captures the relative progression of electrophysiological properties of multiple organoids thereby identifying the effect of different treatments applied to multiple organoids relative to each other. Il-Vis makes the third and fourth advantages possible by preserving the heterogeneity of responses at any given timepoint and in between different timepoints, reflecting anything from small to large variations in electrophysiological properties. We demonstrated these advantages in our results section by using simulated and experimental datasets. These features along with the results showing the pipeline's robustness to data with gaps and noise as frequently observed in experimental data provide increased confidence in applying Il-Vis to an unexplored dataset.

Our application of Il-Vis with the cortical organoids allowed us to identify the effect of QUIN and QUIN Ab on the electrophysiological activity of organoids which was not demonstrated in the spiking activity. The broader implications of the findings are that low-level QUIN may induce effects on

cellular health, neurodegeneration, and electrophysiological alterations. Furthermore, the significance of age-related changes in the kynurenine pathway in the brain, including QUIN accumulation, is starting to be appreciated in recent human studies of healthy and neurological disease patients[30, 31]. Therefore, we believe our findings supporting the evidence of QUIN Ab's capability in sequestering higher levels of QUIN are of greater importance to the field. More studies are needed to longitudinally assess the levels of QUIN, potential correlation with known neurodegeneration biomarkers, and the effect of potential interventions to reduce QUIN levels. Il-Vis can play a key role in modeling data from such future experiments.

Il-Vis could be used in many different scenarios including the three we have mentioned below. First, Il-Vis could be used to compare the effect of many other treatments e.g. the relative comparison of two metabolites with each other and also to an untreated organoid. Second, Il-Vis could assist organoid-based precision medicine to identify a particular treatment's effect on a certain patient's cell line. Third, the distance between the individual increments of a trajectory corresponding to an organoid subjected to a particular treatment could be used to determine how variable the responses are with respect to an organoid's age/treatment time. However, such an experiment needs to be carefully constructed with the above purpose in mind which should eventually generate data in regular and even more frequent time intervals. Furthermore, we note that the comparison of Alzheimer's vs control cell lines is not in the scope of this study because such a comparison requires more replicates to represent each disease/control population.

For studies using MEAs with higher spatial resolution, a 2D convolutional neural network-based feature extractor could be used before using the dimensionality reduction module (Fig. 1B), which would aid in capturing the grid information of the electrodes, thus extending Il-Vis. In addition, due to the practical limitations that may arise when storing/using all the training examples thus far, our pipeline could be directly used with a rehearsal-based strategy[32] that only stores/uses a few informative samples from the previous increments.

The following are suggested to obtain more insightful trajectories in biological experiments. First, MEA recordings from multiple timepoints before adding treatments would allow Il-Vis to calibrate and learn the variations among the organoids used. Second, smaller gaps between the increments would allow Il-Vis to learn a smooth/continuous trajectory for each organoid throughout the experiment. Third, organoids treated with varying concentrations of the same treatment would provide a detailed analysis where the relationship between trajectories could be determined with high confidence. Overall, we present a new pipeline applying the SONG algorithm to address the key issues arising from the generation of high-dimensional, longitudinal MEA datasets that fail to be adequately resolved by existing techniques.

4 Methods

4.1 Il-Vis

Code for the incremental learning framework is available at https://github.com/TamashaM/IncrementalFramework

4.1.1 Incremental learned dimensionality reduction model

The pseudo code for the incremental learning algorithm is as below.

Algorithm 1 Incremental Learning Algorithm

```
reduction_model = SONG(specs);
days = [day_1, day_2, .., day_n]
for day in days do
    if data == None then
        data = day.preprocess()
    else
        data = data.append(day.preprocess())
    end if
    Y = reduction_model.fit_transform(data)
end for
```

4.1.2 SONG

The functionality of SONG[14] algorithm can be summarised in three primary iterative steps. 1) Vector Quantization, 2) Self-Organization and 3) Dimensionality Reduction. In Vector Quantization, the dataset is partitioned based on the high density regions in the input distribution, where each partition is represented by a high-dimensional coding vector. In the Self-Organization step, these coding vectors are connected based on their proximity allowing SONG to approximate the input topology effectively parametrizing the input space. An edge between two coding vectors is assigned a higher weight if a dense region of input points lies in between the vicinity of the connected coding vectors, and a lower weight if otherwise. The resulting edge connections create a weighted graph that could be interpreted as a stochastic neighborhood distribution. This graph is then projected on to a low-dimensional space in the Dimensionality Reduction step by minimizing the cross entropy between the stochastic neighborhood graph in the input space, and the low-dimensional representations.

4.1.3 Median tracing

The medians of the visualized increments were calculated using the 2D coordinates of representative data points. The calculated medians were traced to

identify a low-dimensional trajectory. The median was used instead of the mean to remove the effect of possible outliers.

4.2 Simulated Dataset

4.2.1 Trajectory formation

For the simulation experiments, we mathematically formulated random nonlinear trajectories. To generate a D-dimensional nonlinear trajectory, we first generated a D-dimensional linear line L that passes through point p and in direction d where t is the pseudo temporal parameter describing a particular point on L. The vector form of L is shown in Equation 1. The intercepts and the coefficients are randomly generated using the in-built python function "random". L was then transformed to a nonlinear trajectory T using a nonlinear transformation. For simplicity we used the square transformation for randomly selected dimensions.

$$\begin{gathered} L =< d_1 t + p_1, d_2 t + p_2, ..., d_D t + p_D > \\ where\ \overrightarrow{d} =< d_1, d_2..d_D > and\ p =< p_1, p_2, ..., p_D > \\ T =< (d_1 t + p_1)^{a_1}, (d_2 t + p_2)^{a_2}, ..., (d_D t + p_D)^{a_D} > \\ where\ \{a_1, a_2, ..., a_D\} \epsilon \{1, 2\} \end{gathered} \tag{1}$$

T_1^H and T_2^H were created from Equation.1 using two random initializations. T_3^H and T_4^H were defined as compositions of T_1^H and T_2^H (Equation 2). The similarity factor (α) determines how closer the secondary trajectory (T_3^H or T_4^H) is to its principal trajectory (T_1^H or T_2^H).

$$\begin{gathered} if\ T_1^H =< (A_1^1 t + B_1^1)^{a_1^1}, (A_2^1 t + B_2^1)^{a_2^1}, .., (A_D^1 t + B_D^1)^{a_D^1} > and \\ T_2^H =< (A_1^2 t + B_1^2)^{a_1^2}, (A_2^2 t + B_2^2)^{a_2^2}, .., (A_D^2 t + B_D^2)^{a_D^2} > \end{gathered}$$

$$\begin{gathered} T_3^H = \alpha T_1^H + (1 - \alpha) T_2^H \\ T_3^H =< \alpha(A_1^1 t + B_1^1)^{a_1^1} + (1 - \alpha)(A_1^2 t + B_1^2)^{a_1^2}, \alpha(A_2^1 t + B_2^1)^{a_2^1} + (1 - \alpha)(A_2^2 t + B_2^2)^{a_2^2}, .., \\ \alpha(A_D^1 t + B_D^1)^{a_D^1} + (1 - \alpha)(A_D^2 t + B_D^2)^{a_D^2} > \end{gathered}$$

$$\begin{gathered} T_4^H = (1 - \alpha) T_1^H + \alpha T_2^H \\ T_4^H =< (1 - \alpha)(A_1^1 t + B_1^1)^{a_1^1} + \alpha(A_1^2 t + B_1^2)^{a_1^2}, (1 - \alpha)(A_2^1 t + B_2^1)^{a_2^1} + \alpha(A_2^2 t + B_2^2)^{a_2^2}, .., \\ (1 - \alpha)(A_D^1 t + B_D^1)^{a_D^1} + \alpha(A_D^2 t + B_D^2)^{a_D^2} > \\ where\ \{a_1^1, .., a_D^1\} \bigcup \{a_1^2, .., a_D^2\} \epsilon \{1, 2\} \end{gathered} \tag{2}$$

$$
\begin{gathered}
I_i = T_r^H(t) + \lambda.\phi_i(t) \\
100(i-1) \leq t \leq 100i - \delta \\
\phi_i(t) \sim N(0, SD^2)
\end{gathered} \tag{3}
$$

4.2.2 Increment formation

For the simulation experiments, the ith data increment (I_i) was formed based on the Equation.3 where λ and $\phi(t)$ are the percentage and the parent distribution of the noise respectively. The parent distribution is a Gaussian distribution with mean 0 and standard deviation SD where SD is the standard deviation of each dimension in a data increment.

The data points in an increment were generated by plugging in t values (in a pre-defined t range) to the trajectory equation. The range was constrained on the size of the sampling gaps added (δ) in the experiment and the increment number (i). For a sampling gap of size δ between consecutive increments the t range of the ith increment was $100(i-1) \leq t < 100i - \delta$. For each integer t value in this range, one data point was created forming $100 - \delta$ data points per an increment. For the experiment with major gaps (when $\delta = 99$), since this process would only result in one data point per increment, 100 data points were created around this base points (obtained from the single t value) by adding random noise.

4.3 Experimental Cortical Organoid Dataset

The dataset contains extracellular readings recorded from cortical organoids derived from 4 different patient cell lines: 2 healthy controls and 2 Alzheimer's patients. Both patients carry the Presenilin 1 mutation, the most common cause of Familial AD. The organoids from the same cell line were plated on a single MEA plate, thus four plates for the four cell lines. The plating was done after 5 months in culture. Four organoids from the same cell line were matured in four different media 1) Untreated medium 2) Quinolinic acid (QUIN) medium, Quinolinic acid Antibody (QUIN Ab) medium, QUIN+QUIN Ab medium. The concentrations for exogenous QUIN and QUIN Ab used in the media were 100 nM and 10 μg/mL respectively. The 100 nM QUIN concentration was chosen based upon our previous research of chronic QUIN toxicity and subcellular changes including dendritic varicosities and microtubule alterations in neurons using a 350 nM dose[33, 34] and also other studies which showed neuronal damage after exposure to only 100 nM QUIN for several weeks[35]. We chose the lower of these based on the longer timeframe of our study (three months after the treatment), giving us the ability to assess the effect of physiologically elevated levels of QUIN while likely retaining a substantial proportion of neurons alive, and thus being able to examine the effects on neuronal network electrophysiological firing parameters. The use of the antibody concentration (10 μg/mL) was based on our previous study of the antibody's ability to rescue oligodendrocyte cell lines from QUIN toxicity[24].

QUIN was also produced by the organoids endogenously which was quantified by Gas chromatography–mass spectrometry (GC-MS) in the range 16-30 nM (Supplementary Fig.6).

4.4 Human brain organoid production

Experimental protocols were approved by the University of Wollongong Human Research Ethics Committee (HE13/299). The iPSC lines used were characterised in [36, 37]. Human brain organoids[38] produced from iPSCs based on the protocol by Salick et al.[39] were cultured for up to 9 months. Briefly, iPSC colonies were passaged using PBS-EDTA as near single cell suspension, plated onto matrigel (Corning #354277) coated 60 mm cell culture dishes (5000 cells per dish) and maintained in TeSR-E8 media (StemCell Technologies, #5990) for five days in hypoxic conditions (3% O_2, 5% CO_2, 37°C). Five day old iPSC colonies were collected via PBS-EDTA (Thermo Fisher Scientific #14190-144, #15575-038), dissociated into a single cell suspension, and seeded at 1000 cells per well into U-bottom 96 well plates (Corning #7007) in iPS-Brew (Miltenyi Biotec, #130-107-086) containing 10 μM ROCK inhibitor (Y-27632, Calbiochem #688000), followed by brief centrifugation (2 min at 400 G) to focus the cell suspensions. Partial media changes were performed every 48h. Embryoid body media was used for days 2 to 11 (DMEM/F12 (Thermo Fisher Scientific #12500-096) supplemented with 1% NEAA (Thermo Fisher Scientific #11140-050), 1% Glutamax (Thermo Fisher Scientific #35050-061), 1% B27 (Thermo Fisher Scientific #17504-001), 1% N2 (Thermo Fisher Scientific #17502-001)) and contained 10 μM SB431542 (Stemcell Technologies #72234), 0.1 μM LDN193189 (MedChemExpress #HY-12071) and 1 μM IWP-2 (MedChemExpress #HY-13912) until day 8, and 20 ng/μl FGF2 (Stemcell Technologies #130-104-923) and EGF (Miltenyi Biotec #130097750) until day 11. Neural induction was triggered by phasing in neural media (Brainphys (Stemcell Technologies #05790) containing 1% NEAA, 1% Glutamax, 2% B27, 1% N2 until day 24) with partial media changes every 48h until day 24. Spheroids were then transferred to 24 well plates to accommodate the larger media volume needs, maintained in normoxic conditions (20% O_2, 5% CO_2, 37°C), and matured through the addition of 5 ng/ml BDNF (Miltenyi Biotec #130-096-286) and 10 ng/ml GDNF (Miltenyi Biotec #130-098-449) from day 26, and 10 ng/ml NRG1 (Peprotech #AF-100-03) and 20 ng/ml NT3 (Miltenyi Biotec #130-093-973) from day 34, after which partial media changes were performed every 3 to 4 days until the end of the experiment.

4.5 Anti-Quinolinic Acid antibody production and purification

The monoclonal QUIN Ab (clone 7AR, mouse anti-human, isotype IgG1) was developed by Millipore, and subsequently patented by B.J.B. and G.J.G (PCT/IB2013/055902 and WO/2015/008111). Briefly, hybridoma clones producing the antibody were grown in Hybridoma-SFM serum-free medium and

scaled up to a volume of 500 mL. After low-speed centrifugation (4,000×g at 4°C for 10 min) and sterile filtration (with a 0.22 μm filter), the antibodies were purified using a column packed with Protein G Sepharose 4 fast flow beads (GE Healthcare). Elution was carried out with 0.1 M glycine (pH 2.7), and the eluate was neutralized with 1 M Tris-HCl (pH 8.5), after which buffer exchange to PBS at pH 7.4 was carried out. Purified antibodies were quantified using spectrophotometer by measuring the absorbance at 280 nm. Affinity of interactions between biotinylated antibody and purified protein was measured by BLI (BLItz, ForteBio), at room temperature.

4.6 MEA recordings and Pre-processing

Each organoid was placed in a single well in a multi-well MEA plate, where each well consisted of a 3 x 3 grid of titanium nitride electrodes. Each electrode had a diameter of 30 μm with an inter-electrode distance of 200 μm. MEA recordings of each well were taken to record the electrophysiological activity of the organoid in that well at varying time intervals. The treatments were added after 6 months (see Table 3) and the activity was recorded after a day, a week, a month and three months. These time intervals were selected to measure both acute (one day) and chronic effects (one week, one month and three months) of the treatments. At each timepoint (represented by a row in Table 3), the media was changed and the recordings were taken before the media change. Filled cells in Table 3 represent the MEA recording availability for each cell line and their corresponding increment number for the incrementally learned algorithm. The recordings were recorded at a sampling frequency of 25 kHz using the Multi Channel Systems data acquisition system.

MEA recordings were passed through a third-order high pass Butterworth filter at a frequency of 300 Hz to remove the low-frequency artifact[10]. Then, each recording was segmented into 4-second non-overlapping windows. The segment length of 4 seconds was empirically determined by experimenting with varying segment lengths. MEA signals are often distorted with noise and disturbance which are well not understood in the time domain as it is the case in the frequency domain. Therefore, we used Fast Fourier Transformation (FFT) to represent the signal segments in the frequency domain[40]. The FFT is an efficient implementation of the discrete Fourier Transform (DFT). Once the FFT of the MEA signal segment is obtained, we compute the amplitude spectrum as it shows the distribution of constitutive components with discrete frequencies. Since the signal is real-valued, the obtained Fourier Transformation is conjugate symmetric. Therefore, we only use the first half of the spectrum while leaving out the second half. Lastly, we downsampled the segment to reduce the number of features in each segment to 1000.

4.6.1 Increment formation

The data increments were formed by dividing the dataset based on the sampling timepoints at which the MEA recordings were taken. The number of

Table 3 MEA signal recording availability for each cell line

Organoid age	Treatment time (t)	CN1	CN2	AD1	AD2
5 months	-	I_1	-	I_1	
6 months (Drugs added)	t=0	I_2	I_1	I_2	I_1
6 months & 1 Day	t=1 day	I_3	I_2	I_3	-
6 months & 1 Week	t=1 week	I_4	I_3	I_4	I_2
6 months & 1 Month	t=1 month	I_5	I_4	I_5	I_3
6 months & 3 Months	t=3 months	I_6	I_5	I_6	I_4

Four organoids from each cell line were plated on a multi-well MEA plate. The treatments were added after 6 months and the activity was recorded after a day, a week, a month and three months. Filled cells in the table represent the MEA recording availability for each cell line and their corresponding increment number for the incrementally learned algorithm.

increments in each cell line was different based on their recording availability as shown in Table 3.

4.6.2 Euclidean Distance Metric

An euclidean distance based metric was used to quantify the similarity/dissimilarity between any two trajectory pair. The metric calculates the average euclidean distance between the medians of the corresponding increments in the considered trajectories. Equation.4 shows the equation to calculate the distance between two 2D trajectories A and B, each with n increments. a_i^x and a_i^y correspond to the medians of the dimensions x and y in the ith increment of trajectory A. b_i^x and b_i^y correspond to the medians of the dimensions x and y in the ith increment of trajectory B.

$$D(A,B) = \frac{1}{n}\sum_{i=1}^{n}[(a_i^x - b_i^x)^2 + (a_i^y - b_i^y)^2]^{1/2} \tag{4}$$

4.7 Spike detection

MEA data were also analyzed using Neuroexplorer v5. The reference channel was subtracted, a band pass filter (Butterworth) from 300 to 1000 Hz was applied and spikes were detected using the threshold crossing algorithm. The threshold was calculated as -4.0*Median Sigma (MS), whereby MS is equal to the median of the signal absolute values (MA)/0.6745.

4.8 Code Availability

All code related to simulation trajectory formation, MEA data pre-processing, and proposed incremental pipeline could be found in the folowwing repository https://github.com/TamashaM/IncrementalFramework.

4.9 Data Availability

Experimental data supporting the findings of this paper are available from L.O. upon request.

Supplementary information. The supplementary material related to this manuscript can be found in the additional document attached in the submission.

Acknowledgments. The authors would like to acknowledge Dementia Australia Research Foundation and Yulgilbar Alzheimer's Research Program for funding. D.A.S. and S.K.H. acknowledge Australian Research Council grant DP210101135. L.O. is supported by a National Health and Medical Research Council (NHMRC) of Australia Boosting Dementia Research Leadership Fellowship (APP1135720). V.G. acknowledges Australian Research Council Discovery Early Career Researcher Award DE180100775. M.D.L, G.S. and B.J.B acknowledge funding of the Peter Duncan Neurosciences Research Unit at St. Vincent's Centre for Applied Medical Research. M.D.L., G.S. and B.J.B acknowledge the contribution of Prof. Daniel Christ and Dr. Peter Schofield toward expression and purification of the IgG1 antibody.

The authors would like to thank Bishakhdatta Gayen, Savas Peter, Maneesha Perera, Rashindrie Perera, Hai-Hang Wu, Nisal Ranasinghe and Deshani Geethika for proofreading.

Note. Please note the article has not yet been peer-reviewed.

References

[1] Di Lullo, E., Kriegstein, A.R.: The use of brain organoids to investigate neural development and disease. Nature Publishing Group (2017). https://doi.org/10.1038/nrn.2017.107

[2] Ooi, L., Dottori, M., Cook, A.L., Engel, M., Gautam, V., Grubman, A., Hernández, D., King, A.E., Maksour, S., Targa Dias Anastacio, H., Balez, R., Pébay, A., Pouton, C., Valenzuela, M., White, A., Williamson, R.: If Human Brain Organoids Are the Answer to Understanding Dementia, What Are the Questions? Neuroscientist **26**(5-6), 438–454 (2020). https://doi.org/10.1177/1073858420912404

[3] Quadrato, G., Nguyen, T., Macosko, E.Z., Sherwood, J.L., Yang, S.M., Berger, D.R., Maria, N., Scholvin, J., Goldman, M., Kinney, J.P., Boyden, E.S., Lichtman, J.W., Williams, Z.M., McCarroll, S.A., Arlotta, P.: Cell diversity and network dynamics in photosensitive human brain organoids. Nature **545**(7652), 48–53 (2017). https://doi.org/10.1038/nature22047

[4] Fischer, J., Heide, M., Huttner, W.B.: Genetic Modification of Brain Organoids. Frontiers in Cellular Neuroscience **13**, 558 (2019). https://doi.org/10.3389/fncel.2019.00558

[5] Passaro, A.P., Stice, S.L.: Electrophysiological Analysis of Brain Organoids: Current Approaches and Advancements. Frontiers in Neuroscience **14**(January), 1–13 (2021). https://doi.org/10.3389/fnins.2020.

622137

[6] Schröter, M., Girr, M., Boos, J., Renner, M., Gazorpak, M., Gong, W., Bartram, J., Müller, J., Hierlemann, A.: Mapping neuronal network dynamics in developing cerebral organoids. Frontiers in Cellular Neuroscience **12** (2018). https://doi.org/10.3389/conf.fncel.2018.38.00066

[7] Giandomenico, S.L., Mierau, S.B., Gibbons, G.M., Wenger, L.M.D., Masullo, L., Sit, T., Sutcliffe, M., Boulanger, J., Tripodi, M., Derivery, E., Paulsen, O., Lakatos, A., Lancaster, M.A.: Cerebral organoids at the air–liquid interface generate diverse nerve tracts with functional output. Nature Neuroscience **22**(4), 669–679 (2019). https://doi.org/10.1038/s41593-019-0350-2

[8] Monzel, A.S., Smits, L.M., Hemmer, K., Hachi, S., Moreno, E.L., van Wuellen, T., Jarazo, J., Walter, J., Brüggemann, I., Boussaad, I., Berger, E., Fleming, R.M.T., Bolognin, S., Schwamborn, J.C.: Derivation of Human Midbrain-Specific Organoids from Neuroepithelial Stem Cells. Stem Cell Reports **8**(5), 1144–1154 (2017). https://doi.org/10.1016/j.stemcr.2017.03.010

[9] Kawada, J., Kaneda, S., Kirihara, T., Maroof, A., Levi, T., Eggan, K., Fujii, T., Ikeuchi, Y.: Generation of a Motor Nerve Organoid with Human Stem Cell-Derived Neurons. Stem Cell Reports **9**(5), 1441–1449 (2017). https://doi.org/10.1016/j.stemcr.2017.09.021

[10] Dulini, G., Mendis, C.: Detection of Phenotypic Network Signatures in Cultured Neuronal Networks. PhD thesis, University of Melbourne (2017). https://minerva-access.unimelb.edu.au/handle/11343/208769

[11] Cunningham, J.P., Yu, B.M.: Dimensionality reduction for large-scale neural recordings. Nature Neuroscience **17**(11), 1500–1509 (2014). https://doi.org/10.1038/nn.3776

[12] Van Der Maaten, L., Hinton, G.: Visualizing Data using t-SNE. Journal of Machine Learning Research **9**, 2579–2605 (2008)

[13] McInnes, L., Healy, J., Melville, J.: UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction (2018)

[14] Senanayake, D.A., Wang, W., Naik, S.H., Halgamuge, S.: Self-Organizing Nebulous Growths for Robust and Incremental Data Visualization. IEEE Transactions on Neural Networks and Learning Systems, 1–15 (2020). https://doi.org/10.1109/tnnls.2020.3023941

[15] Guillemin, G.J., Williams, K.R., Smith, D.G., Smythe, G.A., Croitoru-Lamoury, J., Brew, B.J.: Quinolinic acid in the pathogenesis of alzheimer's

disease. Advances in Experimental Medicine and Biology **527**(61), 167–176 (2003).

[16] Nussbaumer, H.J.: Fast Fourier Transform and Convolution Algorithms **2** (1981). https://doi.org/10.1007/978-3-662-00551-4

[17] Trujillo, C.A., Gao, R., Negraes, P.D., Gu, J., Buchanan, J., Preissl, S., Wang, A., Wu, W., Haddad, G.G., Chaim, I.A., Domissy, A., Vandenberghe, M., Devor, A., Yeo, G.W., Voytek, B., Muotri, A.R.: Complex Oscillatory Waves Emerging from Cortical Organoids Model Early Human Brain Network Development. Cell Stem Cell **25**(4), 558–569 (2019). https://doi.org/10.1016/j.stem.2019.08.002

[18] Wold, S., Esbensen, K., Geladi, P.: Principal component analysis. Chemometrics and Intelligent Laboratory Systems **2**(1-3), 37–52 (1987). https://doi.org/10.1016/0169-7439(87)80084-9

[19] Gullo, F., Amadeo, A., Donvito, G., Lecchi, M., Costa, B., Constanti, A., Wanke, E.: Atypical "seizure-like" activity in cortical reverberating networks in vitro can be caused by LPS-induced inflammation: A multi-electrode array study from a hundred neurons. Frontiers in Cellular Neuroscience **8**(November), 1–18 (2014). https://doi.org/10.3389/fncel.2014.00361

[20] Colpo, G.D., Venna, V.R., McCullough, L.D., Teixeira, A.L.: Systematic review on the involvement of the kynurenine pathway in stroke: Pre-clinical and Clinical Evidence. Frontiers in Neurology **10**(JUL), 778 (2019). https://doi.org/10.3389/fneur.2019.00778

[21] Lovelace, M.D., Varney, B., Sundaram, G., Franco, N.F., Ng, M.L., Pai, S., Lim, C.K., Guillemin, G.J., Brew, B.J.: Current evidence for a role of the kynurenine pathway of tryptophan metabolism in multiple sclerosis. Frontiers in Immunology **7**(AUG) (2016). https://doi.org/10.3389/fimmu.2016.00246

[22] Sundaram, G., Sundaram, G., Lim, C.K., Brew, B.J., Brew, B.J., Brew, B.J., Guillemin, G.J.: Kynurenine pathway modulation reverses the experimental autoimmune encephalomyelitis mouse disease progression. Journal of Neuroinflammation **17**(1) (2020). https://doi.org/10.1186/s12974-020-01844-y

[23] Lovelace, M.D., Varney, B., Sundaram, G., Lennon, M.J., Lim, C.K., Jacobs, K., Guillemin, G.J., Brew, B.J.: Recent evidence for an expanded role of the kynurenine pathway of tryptophan metabolism in neurological diseases. Neuropharmacology **112**(Pt B), 373–388 (2017). https://doi.org/10.1016/j.neuropharm.2016.03.024

[24] Sundaram, G., Brew, B.J., Jones, S.P., Adams, S., Lim, C.K., Guillemin, G.J.: Quinolinic acid toxicity on oligodendroglial cells: relevance for multiple sclerosis and therapeutic strategies. Journal of Neuroinflammation 2014 11:1 **11**(1), 1–11 (2014). https://doi.org/10.1186/S12974-014-0204-5

[25] Moroni, F., Lombardi, G., Moneti, G., Aldinio, C.: The excitotoxin quinolinic acid is present in the brain of several mammals and its cortical content increases during the aging process. Neuroscience Letters **47**(1), 51–55 (1984). https://doi.org/10.1016/0304-3940(84)90385-9

[26] Lugo-Huitrón, R., Ugalde Muñiz, P., Pineda, B., Pedraza-Chaverrí, J., Ríos, C., Pérez-De La Cruz, V.: Quinolinic acid: An endogenous neurotoxin with multiple targets. Oxidative Medicine and Cellular Longevity **2013** (2013). https://doi.org/10.1155/2013/104024

[27] Lim, C.K., Bilgin, A., Lovejoy, D.B., Tan, V., Bustamante, S., Taylor, B.V., Bessede, A., Brew, B.J., Guillemin, G.J.: Kynurenine pathway metabolomics predicts and provides mechanistic insight into multiple sclerosis progression. Scientific Reports 2017 7:1 **7**(1), 1–9 (2017). https://doi.org/10.1038/srep41473

[28] Wagenaar, D.A., Pine, J., Potter, S.M.: An extremely rich repertoire of bursting patterns during the development of cortical cultures. BMC Neuroscience **7**(1), 11 (2006). https://doi.org/10.1186/1471-2202-7-11

[29] Johnstone, A.F.M., Gross, G.W., Weiss, D.G., Schroeder, O.H.U., Gramowski, A., Shafer, T.J.: Microelectrode arrays: A physiologically based neurotoxicity testing platform for the 21st century. NeuroToxicology **31**(4), 331–350 (2010). https://doi.org/10.1016/j.neuro.2010.04.001

[30] Sorgdrager, F.J.H., Vermeiren, Y., Van Faassen, M., van der Ley, C., Nollen, E.A.A., Kema, I.P., De Deyn, P.P.: Age- and disease-specific changes of the kynurenine pathway in Parkinson's and Alzheimer's disease. Journal of Neurochemistry **151**(5), 656–668 (2019). https://doi.org/10.1111/jnc.14843

[31] Sas, K., Szabó, E., Vécsei, L.: Mitochondria, oxidative stress and the kynurenine system, with a focus on ageing and neuroprotection. Molecules (2018). https://doi.org/10.3390/molecules23010191. https://pubmed.ncbi.nlm.nih.gov/29342113/

[32] Delange, M., Aljundi, R., Masana, M., Parisot, S., Jia, X., Leonardis, A., Slabaugh, G., Tuytelaars, T.: A continual learning survey: Defying forgetting in classification tasks. IEEE Transactions on Pattern Analysis and Machine Intelligence **8828**(c), 1–20 (2021). https://doi.org/10.1109/

TPAMI.2021.3057446

[33] Kerr, S.J., Armati, P.J., Guillemin, G.J., Brew, B.J.: Chronic exposure of human neurons to quinolinic acid results in neuronal changes consistent with AIDS dementia complex. AIDS **12**(4), 355–363 (1998). https://doi.org/10.1097/00002030-199804000-00003

[34] Kerr, S.J., Armati, P.J., Brew, B.J.: Neurocytotoxity of quinolinic acid in human brain cultures. Journal of neurovirology **1**(5-6), 375–380 (1995). https://doi.org/10.3109/13550289509111027

[35] Whetsell, W.O., Schwarcz, R.: Prolonged exposure to submicromolar concentrations of quinolinic acid causes excitotoxic damage in organotypic cultures of rat corticostriatal system. Neuroscience Letters **97**(3), 271–275 (1989). https://doi.org/10.1016/0304-3940(89)90609-5

[36] Engel, M., Balez, R., Muñoz, S.S., Cabral-da-Silva, M.C., Stevens, C.H., Bax, M., Do-Ha, D., Sidhu, K., Sachdev, P., Ooi, L.: Viral-free generation and characterization of a human induced pluripotent stem cell line from dermal fibroblasts. Stem Cell Research **32**, 135–138 (2018). https://doi.org/10.1016/J.SCR.2018.09.014

[37] SS, M., R, B., ME, C.C.-d.-S., T, B., M, E., M, B., D, D.-H., CH, S., M, G., A, B., L, O.: Generation and characterization of human induced pluripotent stem cell lines from a familial Alzheimer's disease PSEN1 A246E patient and a non-demented family member bearing wild-type PSEN1. Stem cell research **31**, 227–230 (2018). https://doi.org/10.1016/J.SCR.2018.08.006

[38] Ooi, L., Dottori, M., Cook, A.L., Engel, M., Gautam, V., Grubman, A., Hernández, D., King, A.E., Maksour, S., Targa Dias Anastacio, H., Balez, R., Pébay, A., Pouton, C., Valenzuela, M., White, A., Williamson, R.: If Human Brain Organoids Are the Answer to Understanding Dementia, What Are the Questions? Neuroscientist **26**(5-6), 438–454 (2020). https://doi.org/10.1177/1073858420912404

[39] Salick, M.R., Wells, M.F., Eggan, K., Kaykas, A.: Modelling Zika Virus Infection of the Developing Human Brain In Vitro Using Stem Cell Derived Cerebral Organoids. JoVE (Journal of Visualized Experiments) **2017**(127), 56404 (2017). https://doi.org/10.3791/56404

[40] Nussbaumer, H.J.: The Fast Fourier Transform, pp. 80–111. Springer (1981).

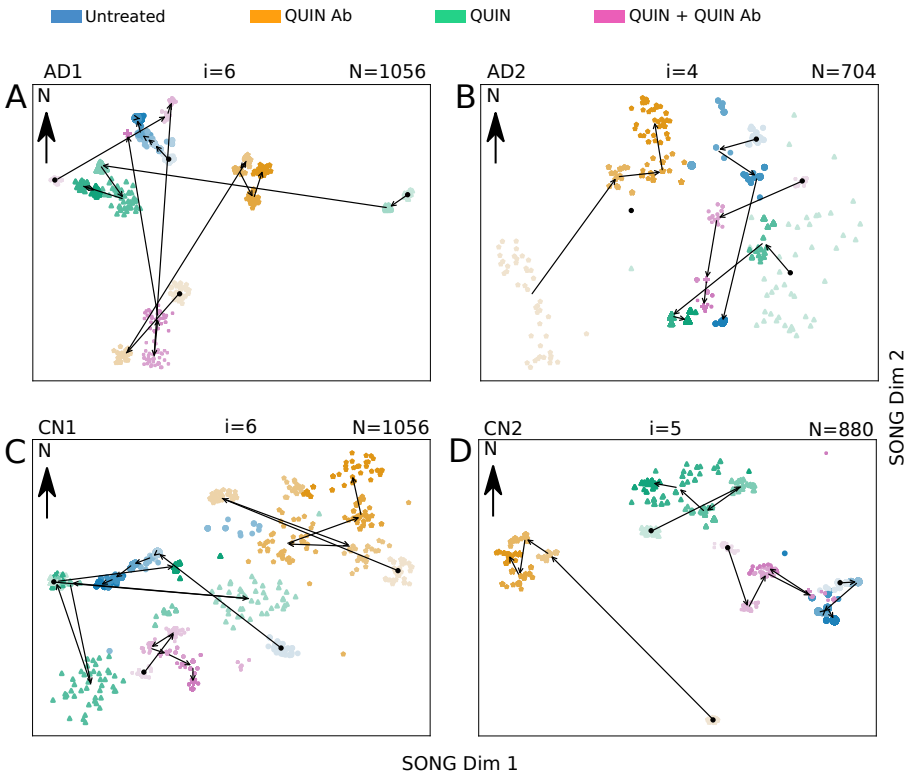

Fig. 5 2D visualizations obtained by the pipeline for each cell line at their final session when the organoids in four treatments (J={Untreated, QUIN, QUIN Ab, QUIN+QUIN Ab}) were modeled together. The data point representation is the same as Fig. 4. Table 3 shows the labels for increments in each cell line with the corresponding organoids' age. The visualizations show a unique behavior in AD/CN organoids treated with QUIN Ab alone, compared to Untreated, QUIN treated or QUIN+QUIN Ab treated organoids. (A) QUIN Ab treated organoid originates at a considerable distance from the other three organoids and also moves in a direction different to others. In contrast, both Untreated and QUIN+QUIN Ab treated organoids initiate and converge in nearby locations to each other. Although the organoid treated with QUIN is initiated at a considerable distance, it converges to a point closer to the Untreated and QUIN+QUIN Ab treated organoids. (B) QUIN Ab treated organoid originates from the bottom left corner of the diagram at a considerable distance from the other three organoids which are originating at the right side of the figure. QUIN Ab treated organoid propagates in the opposite direction (SW to NE and then towards N) than the other three organoids and converges in a distant location. The other three organoids converge closer to each other. (C) QUIN Ab treated organoid is originating at a considerable distance from the other three organoids and moving to a different direction (top right). In contrast, the other three organoids are located more towards the bottom left corner of the figure. (D) QUIN Ab treated organoid is originating at a considerable distance from all the other three organoids and moving towards the top left of the figure. In contrast, the other three organoids are located more towards the top right corner of the figure.