Supplementary Materials for On modelling diversity in electrical cellular response: data-driven approach

Ablaikhan Akhazhanov*,† and Chi On Chui*,†,‡

†*Department of Electrical and Computer Engineering, University of California, Los Angeles*

‡*Department of Bioengineering, University of California, Los Angeles*

E-mail: gkcalat@ucla.edu; chui@ee.ucla.edu

To whom correspondence may be addressed Ablaikhan Akhazhanov or Chi On Chui.

This PDF file includes Materials and methods, Figs. S1 to S4, Tab. S1

Materials and Methods

Modelling electrical cellular response

We collected previously published experimental data and grouped them into five categories: adherent animal cells, non-adherent animal cells, yeast cells, bacteria, and other (tissues and bio-membranes). To visualize the entire dataset, we plot it in Figure 1 in the manuscript, which shows that electrical properties tend to form small clusters and sometimes have a significant overlap between different cell types. A larger dataset is necessary for further investigation with more sophisticated models, as they simply overfit the dataset presented in this work. Note, that a portion of the works did not include conductance of the membranes and a dielectric constant of the nucleoplasm (shown with the corresponding zero values). We exclude them from the modelling of the corresponding properties.

Correlation analysis. Although correlation between cytoplasm and culture medium is frequently reported in the literature, we did not find a strong evidence of consistent linear relationship between the variables. Moreover, no strong relationship between conductivity and permittivity (as well as conductance and capacitance) within any cell component was found. We, therefore, proceed with univariate analysis of each dielectric property of individual cell components.

Univariate analysis. Each electrical property is modelled as a random variable. Given clustered structure of the data shown in Figure 1 in the manuscript and knowing that biological diversity in cellular electrical response is driven by complex unobservable processes inside the cell, parametric modelling is a natural choice. Moreover, parametric models have physical interpretations, generalize better with small training datasets, and provide sufficient power when mixed together. A larger dataset is necessary for further investigation with more sophisticated models, as they simply overfit the dataset presented in this work. In our analysis, we choose candidate models as the closest to the observations on Cullen and Frey graph [1]

constructed using non-parametric bootstrap with 2000 iterations. To mitigate inherent constraints on functional forms of the parametric approach, we employ a comprehensive list of distributions including log-normal, normal, gamma, exponential, Weibull, t-distribution, inverse Gaussian (Wald), inverse normal, and inverse log-normal. Although more sophisticated statistical models were expected to produce a superior fit of biological diversity, they tend to overfit the data and result in poor generalization measured by cross-validation.

For each candidate distribution, we fit a mixture model by applying Expectation-Maximization (EM) algorithm.[2] At E-step, we estimate weights of the mixture components by maximizing the weighted log-likelihood for given model parameters. At M-step, we maximize the weighted log-likelihood to get new parameter estimates. To choose a proper number of mixture components without overfitting, we use 5-fold cross validation and pick the one with the largest out-of-sample log-likelihood. As an alternative to cross-validation, we use the gap statistics[3] of complete-linkage hierarchical clustering. Interestingly, both methods always suggested the same number of mixture components. To ensure global convergence, we run each optimization 50 times with different weights initialization and take the best performing models. In each run, we initialize the mixture weights with either K-Means, hierarchical clustering, or random uniform (continuous and discrete) assignment.

When the data is insufficient for EM algorithm to converge in 1000 iterations, we use hard cluster assignment of each data point. Similarly to the previous procedure, we use K-Means, hierarchical clustering and random uniform (discrete) cluster assignment. We, then, fit a parametric model with Maximum Likelihood Estimation (MLE) for each cluster.

To reject the models we conduct non-parametric Kolmogorov-Smirnov (KS) test, and calculate the corresponding p-values (confidence level of 0.1) of the simulated bootstrapping of the KS-statistics as was suggested in literature.[4] To choose the best performing model, we carefully analyze the quantile-quantile (Q-Q) plots and compare Bayesian information criterion (BIC).

Finite-Element Method FEM subdivides a large problem into smaller subdomains that are called finite elements. The systems of equations that model the finite elements are then combined into the entire model. FEM gives accurate representation of complex geometries, allows dissimilar material properties, and captures local effects. We employ COMSOL Multiphysics's AC/DC module in frequency domain to solve current conservation equation based on Ohm's law. Built-in discretization engine takes a randomly generated cell shape, places it onto a predefined array of microelectrodes, and builds an appropriate mesh (see Figure S2 (F)). To numerically solve the corresponding partial differential equations (PDEs), we use flexible generalized minimal residual method (FGMRES), an iterative procedure that gives an efficient trade-off between computational cost and solution quality. Our approach is easily scaled up to thousands of concurrent simulations and provides high quality results.

Experimental study

Microelectrode array. We designed a microelectrode array (MEA) of eight golden electrodes with pitch of 11 μm (see Figure S4). MEA was fabricated at Nanoelectronics Research Facility at UCLA. A layer of 20 nm of Ti and 100 nm of Au was vacuum-evaporated on fused silica wafer (100 mm in diameter) and the electrodes were patterned with the conventional lift-off process. On top of the metal, we deposited a passivation layer of 500 nm of silicon dioxide with plasma-enhanced chemical vapor deposition (PECVD). To pattern the passivation layer, we employed deep reactive ion etching. We made circular openings of 4 μm in diameter at the center of the MEA and square contact pads of 5 mm at the edges. The wafer was diced into 25-by-25 mm square dies and each of them was cleaned in ultrasonic bath of acetone with subsequent deionized water wash, nitrogen blow and thermal dehydration. On top of the MEAs we designed 500 microliter (μL) wells of polydimethylsiloxane (PDMS) to allow long-term cell culturing. In addition, MEAs were connected to the intermediate printed-circuit board (PCB) with conductive silver epoxy adhesive (MG Chemicals). The intermediate PCB was connected to switching PCB that contained analog multiplexers and

interfaces for vector network analyzer (VNA, HP 8753ES) and a computer. The computer was controlling stimulation power of the VNA, switching between the electrodes, and fetching data via General-Purpose Interface Bus (GPIB, Agilent 82357B).

HeLa cell line. Human cervical carcinoma (HeLa AC-free, Sigma-Aldrich) cells were grown at 37 °C in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 100 $units/mL$ penicillin, and 100 $\mu g/mL$ streptomycin (PS). Cells were regularly passaged to maintain exponential growth. Twenty-four hours before staining and electrical measurements, we trypsinized HeLa cells, diluted them $1:5$ with fresh DMEM medium without antibiotics, centrifuged, and suspended into fresh DMEM with 10% FBS and 1% PS. Then, we transferred 10 μL of the cell suspension ($2 \cdot 10^5$ $cells/mL$ concentration) to the MEA device with the attached PDMS well (500 μL). To enable the long-term incubation we added 390 μL of fresh DMEM medium with 10% FBS and 1% PS and left the cells in the incubator for 24 hours at 37 °C and 5% CO_2. The cells appeared healthy and attached thereafter. During the incubation, the entire assembly of MEA, PDMS well, and the intermediate PCB was enclosed in a small vented Petri dish. After the incubation, the well was gently desiccated and washed with DMEM medium. To stain the cell nuclei, we added 20 μL of 2 $\mu g/ml$ Hoechst 33342 (Invitrogen) solution and incubated the cells for 25 minutes at 37 °C. We then gently desiccated the well and washed it with DMEM medium. To stain the cell membrane, 20 μl of 5 $\mu g/ml$ of WGA Alexa Fluor®488 conjugate (Invitrogen) solution were added, and cells were incubated for another 10 minutes at 37 °C. After the last incubation, the well was gently desiccated, washed, and filled with a fresh DMEM medium with 10% FBS and 1% PS. We imaged the cells with fluorescent microscopy before and after the electrical measurements and observed no visible changes in structure, shape, and location caused by the electrical measurements. A single human cervical carcinoma (HeLa) cell at the middle of MEA, stained with green membrane and blue nucleus, is depicted in Figure S4.

Table S1: Computational complexity associated with levels of detail of cell geometry

Metric	**Realistic**	**Proposed**	**Naive**
Average degrees of freedom (entire model)	1,119,844	237,525	118,062
Mesh size of the cell, elements	113,402	15,880	5,297
Time to solve $[min, max]$, sec.	$[570, 635]$	$[114, 127]$	$[54, 60]$

procedure MAIN
  initialize $\Delta p, k_c, \lambda, c_0, z, \epsilon, \theta, \Delta A, \Delta V, R_{cell}, a_n, b_n, c_n$
  $A_m \leftarrow 4\pi R_{cell}^2$, $V_{cyto} \leftarrow \frac{4}{3}\pi(R_{cell}^3 - a_n b_n c_n)$
  $\rho_0 \leftarrow max\{x : x^2 - \frac{c_n}{\tan\theta}x - \frac{c_n^2}{3} - \frac{4(R_{cell}^3 - a_n b_n c_n)(1+\Delta V)}{3c_n} = 0\}$
  $\phi \leftarrow [-\pi, \pi)$
  stochastically initialize protrusions ρ^* at angles ϕ^*
  for all z_i **do**
    for all ϕ **do**
      $\rho(\phi) \leftarrow \left\{x : \sqrt{\frac{\rho_0^2}{\sin^2\theta} - x^2} - \frac{\rho_0}{\tan\theta} + \Pi\left(\frac{x}{2a_n}\right) \cdot \left(c_n\sqrt{1 - \frac{x^2}{a_n^2}}\right) - z_i = 0\right\}$
    end for
    for all ϕ^* **do**
      $\rho(\phi) \leftarrow \rho^*\sqrt{\frac{\rho_0^2}{\sin^2\theta} - \left(z_i + \frac{\rho_0}{\tan\theta}\right)^2}$
    end for
    $\rho_{OPT}(z_i) \leftarrow \text{OPT}(\rho(\phi), \phi^*)$
  end for
  $V \leftarrow \sum_j \Delta z_j \sum_i \frac{\rho_{ij}^2 \Delta\phi_i}{2}$
  $A \leftarrow \pi\rho_0^2 + \sum_j \Delta z_j \sum_i \rho_{ij}\Delta\phi_i$
  if $\frac{|A - A_m|}{A_m} > \Delta A$ **then**
    $c_a \leftarrow \left(\frac{A_m + \Delta A}{A}\right)^{1/2}$
    $\rho_{OPT} \leftarrow \rho_{OPT} \cdot c_a$
  end if
  if $\frac{|V - V_{cyt}|}{V_{cyt}} > \Delta V$ **then**
    $c_v \leftarrow \left(\frac{V_{cyt} + \Delta V}{V}\right)^{1/3}$
    $\rho_{OPT} \leftarrow \rho_{OPT} \cdot c_v$
  end if
  $surf \leftarrow$ transform ρ_{OPT} into a surface with cubic spline interpolation
  $obj \leftarrow$ convert $surf$ into a solid **return** obj
end procedure
procedure $\text{OPT}(\rho(\phi), \phi^*)$
  for all ϕ **do**
    $H(\phi) \leftarrow \frac{1}{\rho(\phi)}$
    $K(\phi) \leftarrow \frac{1}{\rho(\phi)}$
  end for
  $err \leftarrow \infty$
  while $err > \epsilon$ **do**
    for all $\phi_j \in \phi \setminus \phi^*$ **do**
      $\Delta\rho \leftarrow \left|1 - \frac{mean(\rho(\phi_{j-1\ldots j+1}))}{\rho(\phi_j)}\right| + c_0\mathcal{N}(0,1)$
      if $\left|\frac{\Delta\rho - \rho(\phi_j)}{\rho(\phi_j)}\right| > \gamma$ **then**
        $\Delta\rho \leftarrow \gamma$
      end if
      $\rho(\phi_j) \leftarrow \rho(\phi_j) \cdot (1 + \Delta\rho)$
    end for
    for all ϕ **do**
      $H(\phi) \leftarrow \frac{1}{\rho(\phi)}$
      $K(\phi) \leftarrow \frac{1}{\rho(\phi)}$
    end for
    $err \leftarrow \left|\Delta p - 2\lambda H + k_c(2H + c_0)(2H^2 - c_0 H - 2K) + 2k_c\nabla^2 H\right|$
  end whilereturn $\rho(\phi)$
end procedure

Figure S1: Pseudocode for cell geometry generation

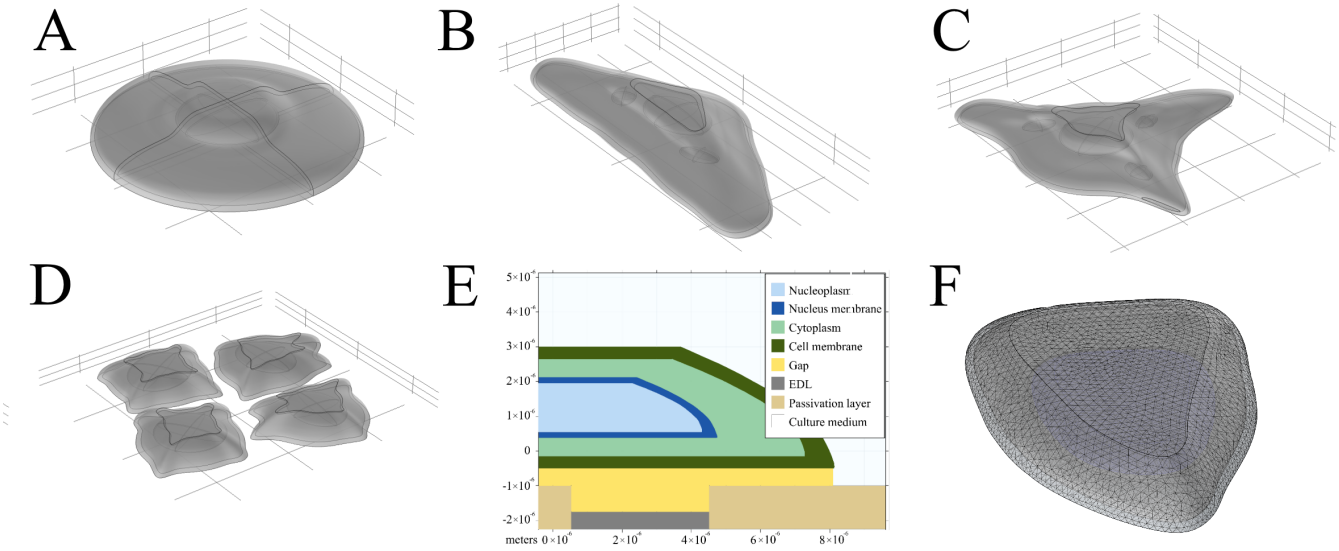

Figure S2: Examples of the cell geometries: (A) isotropic spreading, (B-C) anisotropic spreading, (D) assembly of cells, (E) cell structure schematic, (F) finite-element mesh

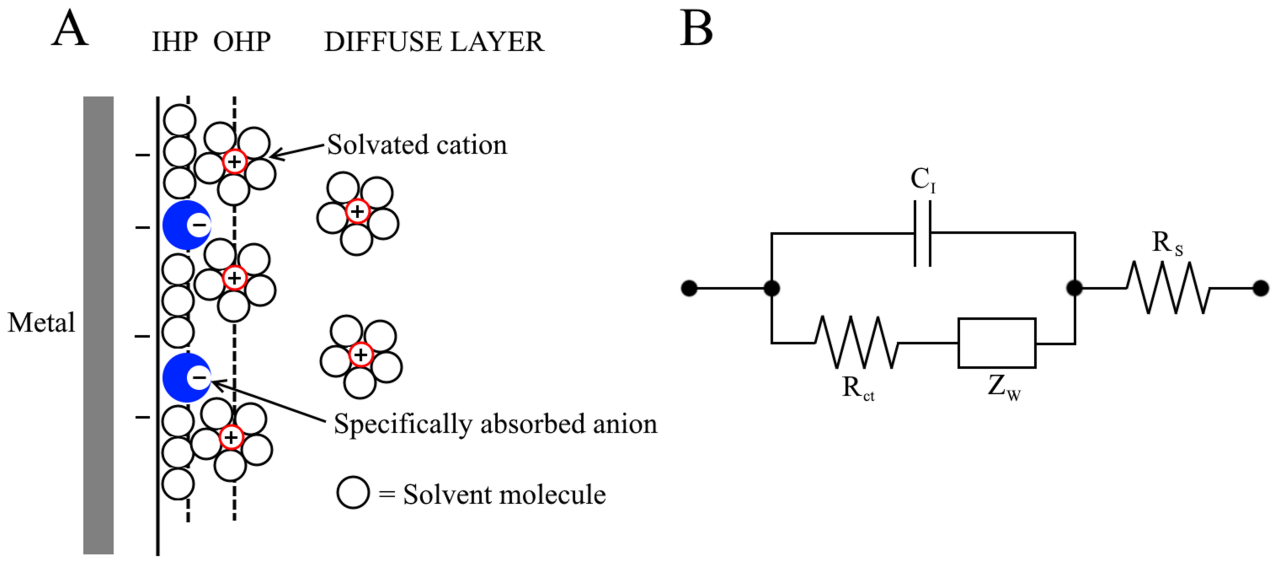

Figure S3: Electrical Double Layer (EDL): (A) physical structure and (B) its lumped-element model

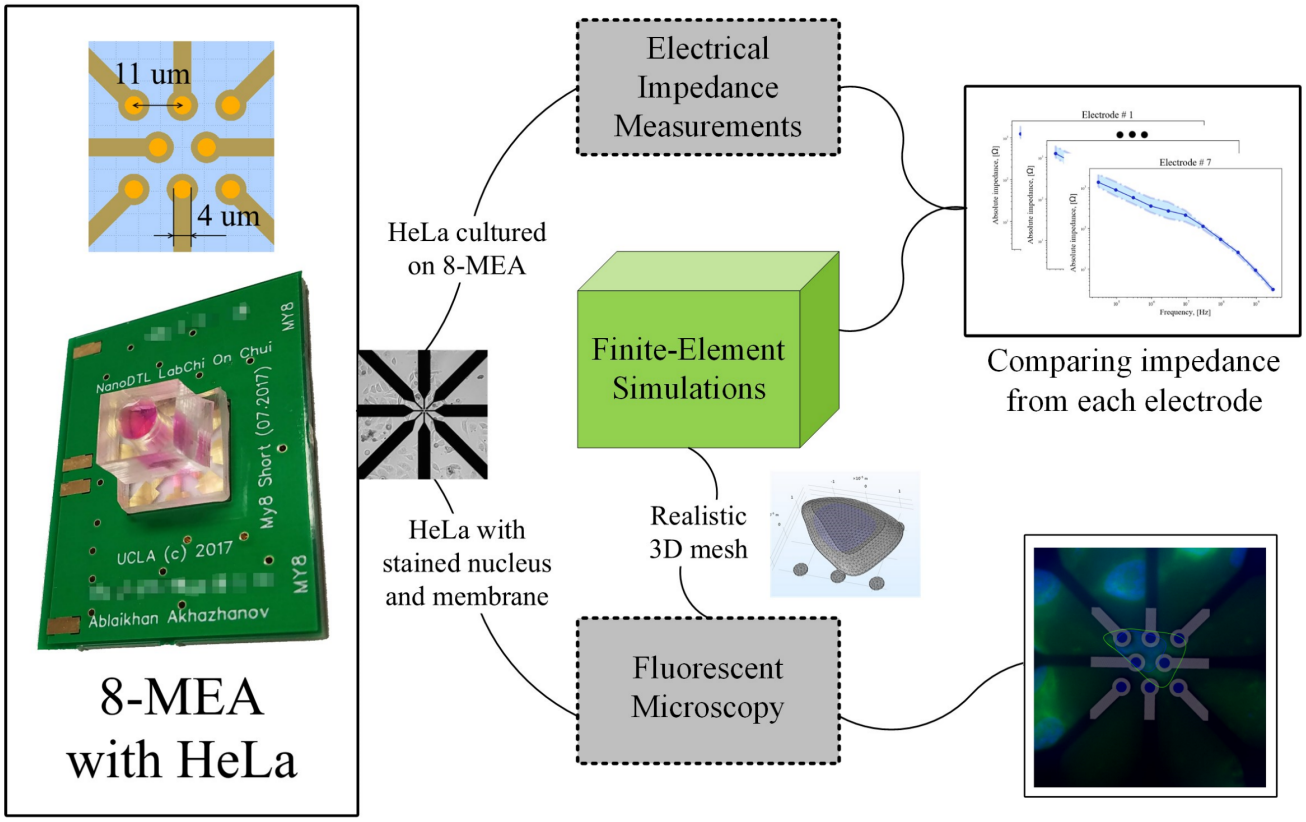

Figure S4: Experimental verification of the model with EIS measurements and FEM simulations of a HeLa cell on 8-electrode array

References

(1) Cullen, A. C.; Frey, H. C. Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs; Springer US, 1999; [Online; accessed 2018-09-26].

(2) Benaglia, T.; Chauveau, D.; Hunter, D. R.; Young, D. **2009**, 32, 1–29.

(3) Tibshirani, R.; Walther, G.; Hastie, T. Journal of the Royal Statistical Society: Series B (Statistical Methodology) **2001**, 63, 411–423.

(4) Clauset, A.; Shalizi, C. R.; Newman, M. E. J. SIAM Review **2009**, 51, 661–703, arXiv: 0706.1062.