An Aqueous Analog MAC Machine

Woo-Bin Jung, Han Sae Jung, Jun Wang, Henry Hinton, Maxime Fournier, Adrian Horgan, Xavier Godron, Robert Nicol, and Donhee Ham*

Using ions in aqueous milieu for signal processing, like in biological circuits, may potentially lead to a bioinspired information processing platform. Studies, however, have focused on individual ionic diodes and transistors rather than circuits comprising many such devices. Here a $16 \times 16$ array of new ionic transistors is developed in an aqueous quinone solution. Each transistor features a concentric ring electrode pair with a disk electrode at the center. The electrochemistry of these electrodes in the solution provides the basis for the transistor operation. The ring pair electrochemically tunes the local electrolytic concentration to modulate the disk’s Faradaic reaction rate. Thus, the disk current as a Faradaic reaction to the disk voltage is gated by the ring pair. The $16 \times 16$ array of these transistors performs analog multiply–accumulate (MAC) operations, a computing modality hotly pursued for low-power artificial neural networks. This exploits the transistor’s operating regime where the disk current is a multiplication of the disk voltage and a weight parameter tuned by the ring pair gating. Such disk currents from multiple transistors are summated in a global reference electrode to complete a MAC task. This ionic circuit demonstrating analog computing is a step toward sophisticated aqueous ionics.

1. Introduction

Ionic circuits in aqueous solutions seek to use ions as charge carriers for signal processing. Fluidic guides and ion-selective media have been arranged in aqueous electrolytic solutions to realize ionic diodes and ionic transistors as circuit building blocks.\[^{[1–6]}\] While ionic circuits are slow due to the low ionic mobility in aqueous solutions, they may ultimately bring an advantage of using diverse ionic species with differing physicochemical properties that can enrich the contents of the information processed. As biological circuits—particularly, neuronal networks in the brain—also process information by using various ionic species in an aqueous milieu, ionic circuits may one day be developed into what imitates biological circuits. Yet the field has so far focused mostly on individual ionic devices rather than more complex circuits comprising many such devices.\[^{[7–10]}\]

Here, we report an aqueous ionic circuit, an array of $16 \times 16 = 256$ ionic transistors, realized on the surface of, and operated by, a complementary metal–oxide–semiconductor (CMOS) electronic chip. Advances are two-fold: innovation of a new ionic transistor as a building block and organization of 256 such ionic transistors into a functional circuit. First, our ionic transistor electrochemically operates in a bare aqueous solution of quinones with no fluidic guides, channel materials, and ion-selective media. It has a center disk electrode surrounded by a concentric pair of ring electrodes (Figure 1a). The ring pair—with two rings driven by opposite sign currents with an identical tunable magnitude $I_g$ (gating current)—electrochemically tunes the concentrations of quinones and hydrogen ions ($H^+$) only in the local aqueous volume above, a technique we recently developed.\[^{[11]}\] Since the rates of electrochemical reactions of the center disk are influenced by the electrolytic concentration in the local aqueous volume surrounding the disk, the ionic current $i_{out}$ (output current) of the disk as an electrochemical reaction to a voltage $V_{in}$ (input voltage) applied to the disk depends on not only $V_{in}$ but also on $I_g$ that tunes the local electrolytic concentration. Hence the ring pair $I_g$ gates the disk current $i_{out}$ for a given disk voltage $V_{in}$. This is the essence of our ionic transistor.

Second, these electrochemically gated ionic transistors are arranged into the $16 \times 16$ array. This scalability is facilitated by the CMOS electronics that can integrate a large number of electrodes. We demonstrate the utility of this array-scale ionic circuit by performing physical, or analog, multiply–accumulate (MAC) operations. Analog MAC operations based on physical phenomena—which contrast digital MAC operations based on many digital logic gates in conjunction with Boolean algebra—are being actively pursued in hopes of lowering power consumption in the artificial neural network (ANN), which involves a large number of MAC operations to obtain dot products between synaptic weight vectors and input data vectors.\[^{[12–14]}\] In particular, a crossbar array of tunable solid-state resistors—notably memristors—has been used for analog MAC operations: each cross-point

W.-B. Jung, H. S. Jung, J. Wang, H. Hinton, D. Ham
John A. Paulson School of Engineering and Applied Sciences
Harvard University
Cambridge, MA 02138, USA
E-mail: donhee@seas.harvard.edu

M. Fournier, A. Horgan, X. Godron
DNA Script
Le Kremlin-Bicêtre 94270, France

R. Nicol
Broad Institute of MIT and Harvard
415 Main Street, Cambridge, MA 02142, 16, USA

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/adma.202205096.

DOI: 10.1002/adma.202205096
conductance acts as a synaptic weight, and input voltages fed to the rows of the array are multiplied by the weights via Ohm’s law, and the resulting currents are accumulated in each column by Kirchhoff’s law.[15–24] Therefore, each column current is a physically produced dot product between the input data vector and the synaptic weight vector of the column. Our ionic transistor array also performs analog MAC computation in the aqueous quinone solution based on physical phenomena (Figure 1b). In each ionic transistor where the disk current $I_{\text{out}}$ for an applied disk voltage $V_{\text{in}}$ is gated by $I_g$, we can find a region of $V_{\text{in}}$ where $I_{\text{out}} = W \times V_{\text{in}}$, with the proportionality constant—or weight—$W$, can be tuned by $I_g$. I.e., in this region, the ionic transistor performs a physical multiplication between the weight and the input voltage. Such $I_{\text{out}}$’s from multiple ionic transistors are then accumulated in the global reference electrode (Figure S1, Supporting Information) via Kirchhoff’s law, thus completing a MAC task. This demonstration of the functional ionic circuit capable of analog computing is a step toward more sophisticated aqueous ionics.

Figure 1. Ionic transistor array. a) Schematic illustration of our ionic transistor, consisting of a concentric electrode ring pair and a center disk electrode. b) Concept of analog MAC operation using a column (shown horizontally, given the space limitation) of ionic transistors. c) $16 \times 16 = 256$ electrochemical pixel array post-fabricated on a CMOS chip. This is the same hardware as in ref. [11], but here we newly create the ionic transistor function in each pixel, and perform analog MAC computing using the resulting ionic transistor array. d–f) The ring pair in each pixel electrochemically modifies local electrolytic concentrations in the aqueous quinone solution,[10] which in this work we harness as a gating for the pixel’s ionic transistor. As a review of these electrolytic concentration localizations first introduced in ref. [11], we show simulation and experiments in (d) through (f). d) COMSOL Multiphysics simulated spatial profiles of $[Q]$ and pH, representing $[H^+]$, with a single ring pair activated with $I_g = 59$ nA, where the left two figures are 3D profiles, while the right two figures are their cross sections. e) Left and middle: the OCP measured across the array when a single ring pair is activated with $I_g = 8.7$ nA and $I_g = 69.3$ nA; Right: OCP measured as a function of time at the center of a pixel, whose ring pair is activated with various $I_g$ values. f) Local electrolytic concentration control can be experimentally achieved simultaneously at any set of pixels we choose to activate.
2. Results and Discussion

2.1. An Array of Concentric Ring Electrode Pairs (Review)

We recently developed an array of concentric anode–cathode ring pairs on a CMOS chip to electrochemically localize electrolytic concentrations. We first review this system here, as our ionic transistor array is built from it (while this review is self-contained, see ref. [11] for full details). The CMOS chip features \(64 \times 64 = 4096\) aluminum (Al) pad electrodes on its surface. Each pad is connected to its own CMOS electronics that can configure to a galvanostat to inject a current, a potentiostat to apply a voltage, or an open-circuit potential (OCP) sensor to measure a voltage with no current. The chip surface is converted via post-fabrication to an array of \(16 \times 16 = 256\) electrochemical cells, or pixels, with each pixel containing a post-fabricated concentric platinum (Pt) ring pair (Figure 1c). Either Pt ring of the pair contacts 4 Al pads, with the 4 corresponding CMOS circuits being co-operated as one effective circuit. On the remaining Al pads unconnected to any of the Pt ring pairs, which are located at pixel centers and in between pixels, Pt disk electrodes are post fabricated (Figure 1c).

The concentric ring pair in each pixel can electrochemically modify local electrolytic concentrations, in conjunction with a quinone redox couple, \(2,5\text{-dimethyl-1,4-hydroquinone (H}_2\text{Q})\) and \(2,5\text{-dimethyl-1,4-benzoquinone (Q)}\), dissolved in the aqueous solution (Figure S2, Supporting Information).[11] To this end, both rings operate in the galvanostat mode, with the inner anodic ring and the outer cathodic ring injecting positive and negative currents, \(I_g\) and \(\pm I_g\) (\(I_g\) is tunable), respectively. Then oxidation, \(H_2Q \rightarrow Q + 2H^+ + 2e^-\), occurs at the inner ring, while reduction, \(Q + 2e^- \rightarrow Q^2-\), occurs at the outer ring.\[11,27\]

The oxidation at the inner ring increases \([Q]\) and \([H^+]\)—concentrations of Q and \(H^+\)—near the inner ring. These locally increased Q and \(H^+\) start spreading both inward (toward the pixel center) and outward. But the outward diffusion is soon blocked by the reduction at the outer ring. Specifically, \(H^+\) is consumed near the outer ring by \(Q^2-\) the reduction generates, and Q is converted to \(Q^2-\) near the outer ring by the reduction. Overall, \([Q]\) and \([H^+]\) end up being increased only in the local volume above around the inner ring and the pixel center, without being able to escape. This locally confined increase of \([Q]\) and \([H^+]\) can be tuned by adjusting \(I_g\); the larger the \(I_g\), the more the local intensification of \([Q]\) and \([H^+]\). Simulated spatial profiles of \([Q]\) and pH, representing \([H^+]\), with a single ring pair activated with \(I_g = 59\) nA show this local increase and confinement of \([Q]\) and \([H^+]\) (Figure 1d).

By using all disk electrodes and all non-activated ring pairs in OCP sensing mode, we can monitor pH to experimentally confirm the localization of \([H^+]\) in ring pairs activated with \(I_g\) (a larger OCP corresponds to a lower pH or a higher \([H^+]\)). Figure 1e shows array-wide measured maps of OCP when a single ring pair is activated with \(I_g = 8.7\) nA and \(I_g = 69.3\) nA. In either case, the OCP at the center of the activated ring pair is increased, indicating an \([H^+]\) increase in the local volume above the pixel center. This local OCP increase is more significant for \(I_g = 69.3\) nA than \(I_g = 8.7\) nA, showing that the locally confined \([H^+]\) increases with \(I_g\). This local concentration control, which is stable throughout the ring pair activation (Figure 1e, right; Figure S3a, Supporting Information), can be achieved simultaneously at any set of pixels we choose to activate (Figure 1f; Figure S3b, Supporting Information).

2.2. Electrochemically Gated Ionic Transistor

Each pixel consisting of a ring pair and a center disk can be turned into an ionic transistor. For this, we apply \(I_g\) in the ring pair to modify the local electrolytic concentration, while applying a voltage \(V_{in}\) to the center disk to evoke a disk current \(I_{out}\) as an electrochemical reaction, with the disk engaged in potentiostat mode. Since the disk is surrounded by the local volume with the modified electrolytic concentration, which influences the electrochemical reaction rate of the disk, \(I_{out}\) depends on not only \(V_{in}\) but also \(I_g\). Thus \(I_g\) gates \(I_{out}\) for a given \(V_{in}\). In this ionic transistor, since \(I_{out}\) is collected by a global reference electrode, the center disk and global reference may be viewed as a source and a drain (or vice versa, depending on definition).

We operate our ionic transistor with the aqueous quinone solution, for which \(I_{out}\) increases with \(I_g\) at any given \(V_{in}\). This specific gating behavior arises in connection with two electrochemical reactions possible at the center disk with \(V_{in} < 0\), the \(H^+\)-assisted reduction \((Q + 2H^+ + 2e^- \rightarrow H_2Q)\) and the normal reduction \((Q + 2e^- \rightarrow Q^2-)\): for small enough \(|V_{in}|\), only the former occurs; for large enough \(|V_{in}|\), both take place (Figure S4, Supporting Information).[27–40] Since the local \([Q]\) and \([H^+]\) increase with \(I_g\), i.e., since a larger \(I_g\) supplies more reactants of either reaction to accelerate it, a larger \(I_g\) leads to a larger \(I_{out}\) for a given \(V_{in} < 0\).

The cyclic voltammetry (CV) measurement of a center disk electrode configured in potentiostat mode for various \(I_g\) values shows a glimpse of this gating tendency of \(I_{out}\) increasing with \(I_g\) (Figure 2a). For the CV, \(V_{in}\) of the disk electrode is linearly ramped down and up between 0 and \(-0.4\) V with a scan rate of 200 mV s\(^{-1}\) (i.e., \(V_{in}\) vs time has a triangular waveform). The figure clearly shows that as \(I_g\) increases, the CV curve moves upward, or \(I_{out}\) is lifted overall across the \(V_{in}\) scan range.

The increase of \(I_{out}\) with \(I_g\) for any given fixed \(V_{in} < 0\)—as opposed to the time-varying triangular \(V_{in}\) sweep of the CV—is more definitively seen in the three sets of measurements shown in Figure 2b, left. In each set of measurements, we repeat the application of a \(V_{in}\) square pulse of a fixed amplitude for 1 s to the center disk and the measurement of the resulting steady-state disk current \(I_{out}\) settled to a (nearly) constant amplitude within that 1 s, for various \(I_g\) values. As seen in Figure 2b, left, for a given \(V_{in}\) pulse amplitude, as \(I_g\) grows, the steady-state \(I_{out}\) also increases. This measured relationship between \(I_g\) and \(I_{out}\) for each given \(V_{in}\) amplitude from Figure 2b, left is explicitly plotted in Figure 2b, right.

These measurements using square pulses of \(V_{in}\) are actually performed with more finely spaced \(V_{in}\) amplitudes between \(-0.36\) and \(-0.01\) V. From these, we can also obtain plots of the steady-state \(I_{out}\) versus the \(V_{in}\) square pulse amplitude, for various \(I_g\) values. Figure 2c shows such plots of \(I_{out}\) versus \(V_{in}\) for \(I_g\) of 0, 26.0, and 52.0 nA. As seen, the overall \(I_{out}\) versus \(V_{in}\) relationship for a given \(I_g\) is nonlinear (the shape is similar to the CV curves of Figure 2a, but not identical, due to the...
Figure 2. Characterization of an ionic transistor. a) CV ($V_\text{in}$, scanned between 0 and −0.4 V at 200 mV s$^{-1}$) at the center disk electrode for various $I_g$ values. b) We repeat the application of a $V_\text{in}$ pulse of a given amplitude for 1 s to the center disk and the measurement of the resulting disk current $I_{\text{out}}$ settled to a (nearly) constant amplitude, for various $I_g$ values ($I_g$ is increased from 0 to 69.3 nA with an approximate step of 8.7 nA). For each given $V_\text{in}$, pulse amplitude (−0.02, −0.08, or −0.2 V), as $I_g$ grows, the $I_{\text{out}}$ amplitude also increases. c) Measured $I_{\text{out}}$ versus $V_\text{in}$ for $I_g = 0, 26.0, \text{and } 52.0 \text{ nA}$. d) Zoom-in of Region A (0.01 V ≤ $|V_\text{in}|$ ≤ 0.1 V) of (c). e) Zoom-in of Region B (0.16 V ≤ $|V_\text{in}|$ ≤ 0.24 V) of (c).

In $V_\text{in}$ used: in the CV curves (Figure 2a) where $V_\text{in}$ is varied with time, $I_{\text{out}}$ includes not only Faradaic current but also capacitive current; in the pulsed measurement (Figure 2c), the steady-state $I_{\text{out}}$ for each constant $V_\text{in}$ amplitude includes only Faradaic current. At the same time, in Figure 2c, we can identify piecewise linear regions—Regions A (0.01 V ≤ $|V_\text{in}|$ ≤ 0.1 V) and B (0.16 V ≤ $|V_\text{in}|$ ≤ 0.24 V)—in the $I_{\text{out}}$ versus $V_\text{in}$ curves.

Figure 2d is a zoom-in of Region A of Figure 2c. Here the linear dependence of $I_{\text{out}}$ on $V_\text{in}$ may be modeled as $I_{\text{out}} = W \times V_\text{in}$, with the slope $W$, or “weight”, tunable by $I_g$ (the $W$ value for each $I_g$ extracted via linear fitting, is annotated in Figure 2d). Importantly, in Region A, $W$ increases with $I_g$, and hence for a given $V_\text{in}$, $I_{\text{out}}$ increases with $I_g$, which is consistent with our earlier discussion. Since $I_{\text{out}}$ is the multiplication of a weight $W$ and an input data $V_\text{in}$ and the weight $W$ is tunable by $I_g$, we exploit this Region A for the MAC operation later. We note that Region A is where the H$^+\text{-assisted reduction (Q + 2H}^+ + 2e^- \rightarrow \text{H}_2\text{Q})$ occurs at the center disk (see the Supplementary Note in the Supporting Information).

Figure 2e zooms in Region B of Figure 2c. This is where both the normal reduction (Q + 2e$^- \rightarrow Q^2$) and the H$^+$-assisted reduction (Q + 2H$^+ + 2e^- \rightarrow \text{H}_2\text{Q})$ occur at the center disk. In Region B, $I_{\text{out}}$ versus $V_\text{in}$ for a given $I_g$ can be again put on a straight line, but in contrast to Region A, the $I_{\text{out}}$ versus $V_\text{in}$ relationship is expressed as $I_{\text{out}} = W \times (V_\text{in} - T)$, where the slope $W$ is nearly constant, being independent of $I_g$, and the intercept $T$ is reduced with $I_g$. Here again, for a given $V_\text{in}$, $I_{\text{out}}$ increases with $I_g$, which is consistent with our earlier discussion. But importantly, the increase of $I_{\text{out}}$ with $I_g$ is not because $W$ increases with $I_g$, but because $T$ decreases with $I_g$ (for the mechanism, see Supplementary Note). This region cannot be utilized for the multiplication for the MAC operation, not only because $I_{\text{out}}$ cannot be expressed as a pure product of the weight $W$ and the input data $V_\text{in}$, but also because the non-tunable $W$ cannot accept different weight values. On the other hand, as Region B overall features a higher $W$ value than Region A, Region B would be more useful for signal amplification application, if $W$ is viewed as a transconductance of the ionic transistor. Furthermore, Region B could be utilized for analog subtraction operation with $T$ being tunable with $I_g$.

2.3. Simultaneous Operation of Ionic Transistors in a Single Column

Multiple ionic transistors in the array have to operate simultaneously to perform a signal processing task, for example, the analog MAC operation to be presented shortly. Therefore, the local environment set by $I_g$ for any given transistor should only gate that transistor, while not interfering with the operation of
any other transistors. Figure 3a shows three example measurements that together confirm no such interference (see also Figures S5 and S6 in the Supporting Information for additional experiments to verify no interference). In any of these three measurements, we choose the same four pixels in a specific column (pixel 4, 9, 12, and 16), where we apply four identical $V_{in}$ pulses (with amplitude $-0.08\ \text{V}$) and measure the resulting four $I_{out}$'s as a total sum ($\sum I_{out}$) that is collected in the global reference electrode. In any of the three cases, we always set $I_g = 0$ for the four chosen pixels. In Case 1, $I_g = 0$ for all twelve peripheral pixels as well. In Case 2, $I_g = 26.0\ \text{nA}$ for six peripheral pixels (pixel 1, 2, 11, 13, 14 and 15) while $I_g = 0$ for the remaining six peripheral pixels. In Case 3, $I_g = 26.0\ \text{nA}$ for all twelve peripheral pixels (pixel 1–3, 5–8, 10, 11, and 13–15). The measured $\sum I_{out}$'s for the three cases are nearly identical, confirming that the localized environment for a given transistor is effective only for that transistor, and hardly interferes with other transistors. b) MAC results in a single column with various assignments of $\{I_g, V_{in,1}\}$, $\{I_g, V_{in,2}\}$, $\{I_g, V_{in,1}\}$, and $\{I_g, V_{in,2}\}$ to the 16 transistors therein. Specifically, the number of transistors with $\{I_{g,1}, V_{in,1}\}$ and $\{I_{g,2}, V_{in,1}\}$ are fixed at 2 and 3. The number $N$ of transistors with $\{I_{g,2}, V_{in,2}\}$ is varied between 0 and 11, leaving $(11 - N)$ transistors with $\{I_{g,1}, V_{in,2}\}$. For a given $N$, there are a large number of different options for lining up the abovementioned $\{I_g, V_{in}\}$ assignments to the 16 transistors along the column, and we choose minimum 5 options, to obtain minimum 5 MAC results, for a given $N$. The MAC results versus $N$ are plotted on the right of the figure, along with the theoretical prediction.
pixels (pixel 1–3, 5–8, 10, 11 and 13–15). As seen in Figure 3a, right, the measured $\sum I_{\text{out}}'$s for the three cases are nearly identical (1.60, 1.61, and 1.59 nA), demonstrating that the localized environment for a given transistor is effective only for that transistor, and is localized enough not to affect other transistors.

2.4. Analog MAC Operation

We now use the array of ionic transistors to perform analog MAC operations, in particular, the dot product operation between an input data vector and a synaptic weight vector of a column, which is the most prevalent computation in the artificial neural network. Each ionic transistor in our system is randomly accessible. However, due to the particular CMOS chip we designed originally for neuronal recording,[31,32] at a given time, only three different $V_{\text{in}}$ value choices (with one being 0 V and each of the other two being widely tunable) and only two different $I_g$ value choices (with one being 0 A and the other being widely tunable) can be made available for the 256 ionic transistors in the array (this limitation has nothing to do with the fundamental concept of the ionic transistor array, but pertains only to the nature of the particular CMOS chip design, and should be readily lifted with a new dedicated CMOS chip design). Given this limitation, we binarize $I_g$ and $V_{\text{in}}$ values: for $I_g$, we choose between $I_g = I_{g,1} = 0$ and $I_g = I_{g,2} = 26.0$ nA; for $V_{\text{in}}$, we choose between $V_{\text{in}} = V_{\text{in},1} = -0.02$ V and $V_{\text{in}} = V_{\text{in},2} = -0.08$ V, to perform the analog MAC operation in the context of binary neural network (BNN).[31] Note that the $V_{\text{in}}$ value choices are in Region A, the region suitable for the multiplication operation. Importantly, in Region A, $I_g = I_{g,1} = 0$ and $I_g = I_{g,2} = 26.0$ nA translate to $W = W_1 = 4.9$ nA V$^{-1}$ and $W = W_2 = 17.4$ nA V$^{-1}$ (Figure 2d). Therefore, each transistor will be assigned to one of the four possible combinations of $\{I_g, V_{\text{in}}\}$, i.e., $\{I_{g,1}, V_{\text{in},1}\}$, $\{I_{g,2}, V_{\text{in},2}\}$, $\{I_{g,1}, V_{\text{in},1}\}$, and $\{I_{g,2}, V_{\text{in},2}\}$, to produce one of the four possible multiplication results as its $I_{\text{out}}$: $W_1 V_{\text{in},1}$, $W_1 V_{\text{in},2}$, $W_2 V_{\text{in},1}$, and $W_2 V_{\text{in},2}$. Such $I_{\text{out}}$s from 16 ionic transistors in a column are then added in the global reference electrode, which is the MAC result. This is the dot product between an input voltage data vector consisting of 16 $V_{\text{in}}$'s applied to the 16 center disks along the column, and a weight vector consisting of 16 $W$'s resulting from 16 $I_g$'s applied to the 16 ring pairs along the column.

Figure 3b shows many MAC computations using a column comprising 16 ionic transistors. For any single MAC computation here, we assign 2 randomly chosen transistors to $\{I_{g,1}, V_{\text{in},1}\}$, 3 random transistors to $\{I_{g,2}, V_{\text{in},1}\}$. N random transistors to $\{I_{g,N}, V_{\text{in},2}\}$ ($0 \leq N \leq 11$), and (11 – N) remaining transistors to $\{I_{g,2}, V_{\text{in},2}\}$. The theoretical MAC result is then $2W_1 V_{\text{in},1} + 3W_2 V_{\text{in},1} + NW_2 V_{\text{in},2} + (11 - N)W_1 V_{\text{in},2}$ = $2W_1 V_{\text{in},1} + 3W_2 V_{\text{in},1} + 11W_1 V_{\text{in},2} + N(W_2 V_{\text{in},2} - W_1 V_{\text{in},2})$. Note that for a given $N$, which produces the same theoretical MAC result as expressed above, there are a large number of different options for lining up the abovementioned $\{I_{g,1}, V_{\text{in}}\}$ assignments to the 16 transistors along the column, and we choose minimum 5 options (to obtain minimum 5 MAC results) for a given $N$. Figure 3b, right, shows the measured MAC results versus $N$, along with the theoretical straight line expressed above. The minimum 5 MAC results for each $N$, which must be theoretically identical, have variations in the measurements, as shown with the error bars. Overall, the measurements closely follow the theoretical straight line, confirming the ability of the ionic transistor column to perform analog MAC operation.

Our 16 × 16 ionic transistor array is then capable of producing a total of 16 MAC results from the 16 columns. In fact, a 1 × 16 output vector made up of these 16 MAC results is the product between the 1 × 16 input voltage data vector and the 16 × 16 synaptic weight matrix of the transistor array. To demonstrate the MAC computations using all 16 columns, we assign $I_g$ and $V_{\text{in}}$ values to the 256 ionic transistors of the array as shown in Figure 4a. Since we use only one global reference electrode, to resolve each column's MAC result, we perform the analog MAC operation column by column in a sequential manner (Figure 4b). Figure 4c shows the measured MAC results for all 16 columns in comparison to their corresponding theoretical calculations using $W_1$ and $W_2$ weight values for $I_{g,1}$ and $I_{g,2}$. The error remains below ±2.5%. 

Figure 4. MAC operations with the 16 × 16 ionic transistor array. a) $I_g$ and $V_{\text{in}}$ values assigned to the 256 ionic transistors of the array to demonstrate the MAC computations using all 16 columns. b) To resolve each column's MAC result given only one reference electrode, we perform the analog MAC operation column by column in a sequential manner. The figure shows the array-wide measured OCP map during each column's MAC operation. c) MAC operation results (blue dots) and theoretically calculated results (hollow red dots) for all 16 columns. The error (red dots) is below ±2.5%. 

2205096 (6 of 8) © 2022 Wiley-VCH GmbH
3. Conclusion

While solid-state electronics have provided the main stage for the efforts to develop bioinspired and biomimicking information processing systems,[12,13,31,32] aqueous ionics may add an additional dimension for bioinspired/biomimicking engineering due to their similarity to biological circuits in terms of using ions for signal processing in aqueous solutions. What is particularly attractive is the availability of a large number of ionic species with differing physical and chemical properties, which can be exploited to enrich the contents of the signal processing, just like in biological circuits. The field of aqueous ionics, however, has so far been mainly focused on the studies of individual ionic diodes and transistors, whereas efforts to develop more complex ionic circuits using many such devices together have been scarce. The development of the new type of fully electrochemical ionic transistor and the aqueous ionic circuit organizing such ionic transistors in the $16 \times 16$ array to perform analog MAC computing—a computing technique actively pursued to develop a low-power ANN accelerator—in this work thus suggests the potential for more sophisticated aqueous ionics.

The work also reveals technical limitations to overcome. For example, due to the lack of independent ionic routing pathways (which would draw parallel with metallic routings in electronic circuits) and local reference electrodes for different columns of the array, the 16 column currents cannot be separately resolved, if the MAC operations of the columns are performed in parallel, and thus the column MAC operations had to be performed column by column in a sequential manner, lowering the computing throughput. Therefore, creating separate ionic routing pathways via electrochemical means—for example, by providing long electrochemical guiding walls created by long electrode geometries (analogous to concentric ring pair geometries creating a localized electrochemical wall) together with distributed local reference electrodes—would be an example of meaningful next pursuits.

At the device level, our ionic transistor in its present form features a response time of 26 ms (Figure 2), has a size scale of 10's $\mu$m, and consumes a power as large as $\approx 11 \text{ nW}$ for the gating. All of these are inferior to solid-state electronic transistors—i.e., metal-oxide-semiconductor field-effect transistors—in integrated circuits. While ionic transistors with low ionic mobility in aqueous solutions can never be as fast as electronic transistors, the response time of 26 ms, which reflects the centimeter-scale distance between any center disk and the global reference electrode, can be substantially reduced by using a proximate local reference electrode. Also, our simulation (Figure S7, Supporting Information) shows that the ring pair structure scaled down to below 1 $\mu$m can still achieve the electrolytic concentration localization, while reducing the power consumption for the gating by nearly three orders of magnitude (the possibility for further downscaling toward what is comparable to the electronic transistor size pertains to the fundamental question of how electrochemistry scales at the deep sub-micrometer scale). In summary, our ionic transistor can be made faster, smaller, and more power efficient, albeit not to the level of the electronic transistor. In fact, the pursuit of ionics is not to compete with, but to complement, electronics, with unique possible features such as the use of diverse ionic charge carriers.

4. Experimental Section

Post-Fabrication and Packaging of the $16 \times 16$ Ionic Transistor Arrays: CMOS integrated circuit was outsourced to United Microelectronics Corporation for fabrication in 0.18 $\mu$m technology. Chip had an array of $64 \times 64 = 4096$ Al pads ($10.5 \times 10.5 \text{ $\mu$m square shape}$) with a pitch of 20 $\mu$m and a passivation layer on the top surfaces. An array of $16 \times 16 = 256$ ionic transistors were post-fabricated onto this pre-defined electrode pad array. First, photolithography was conducted to fabricate a desired pattern for disk electrodes, anodic rings, and cathodic rings. Second, the foundry passivation (silicon dioxide and silicon nitride layers) layers were etched to expose the Al pads via reactive-ion etching. Finally, a thick metal layer (20 nm titanium and 200 nm Pt) was deposited via sputtering. Ultimately, only Pt was exposed to the electrolytic solution, while Al was covered by either passivation or Pt. After the post-fabrication, the chip was glued onto a chip carrier and Al-wire-bonded with a designed interposer printed circuit board. To be encapsulated, poly(dimethylsiloxane) was used to fill between inner and outer rings (Figure S1, Supporting Information).

Array-Wide Generation of the Electrolytic Concentration Localization: The electrochemical setup for ionic transistor (Figure S1, Supporting Information) was filled with an aqueous quinone solution, consisting of $20 \times 10^{-3}$ M 2,5-dimethyl-1,4-hydroquinone (Alfa Chemistry, Ronkonkoma, NY, USA), $5 \times 10^{-3}$ M 2,5-dimethyl-1,4-benzoquinone (Sigma–Aldrich, Atlanta, GA, USA), $1 \times 10^{-3}$ M sodium chloride (aq), $1.9 \times 10^{-2}$ M sodium nitrate (aq), and 8.3% v/v dimethyl sulfoxide. Given this setup, a set of stimulation currents was applied to the anodic and cathodic rings at the selected transistors for array-wide electrolytic concentration localization, and the voltage stimulation was applied to the center disk electrode while peripheral disk electrodes between pixels monitored the spatiotemporal ion profile concurrently.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This research was based upon work supported in part by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via IARPA-19081900002. The views and conclusions contained herein were those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of ODNI, IARPA, or the U.S. Government. The publisher acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this article, or allow others to do so, for United States Government purposes only. Post-fabrication and characterization were performed, in part, at the Center for Nanoscale Systems at Harvard University.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflict of Interest
Keywords
analog computing, artificial neural networks, bioinspired engineering, electrochemical transistors, electrochemistry, ionic transistors, neuromorphic engineering

Received: June 5, 2022
Revised: August 1, 2022
Published online: September 9, 2022