

# Electro-active Polymer Hydrogels Exhibit Emergent Learning When Embodied in a Simulated Game-Environment

Vincent Strong<sup>a</sup>, William Holderbaum<sup>a</sup>, and Yoshikatsu Hayashi<sup>a,1</sup>

<sup>a</sup>Department of Biomedical Sciences and Biomedical Engineering, School of Biological Sciences, University of Reading, Reading, Berkshire, RG6 7BE, UK

<sup>1</sup> Lead Contact

To whom correspondence should be addressed. E-mail: y.hayashi@reading.ac.uk

## Summary

The goal of artificial neural networks is to utilise the functions of biological brains to develop computational algorithms. However, these purely artificial implementations cannot achieve the emergent learning found in Biological Neural Networks (BNNs). Alternative computing mediums, that integrate biological neurons with computer hardware, have shown the emergent learning found in BNNs. By applying current theories in BNN learning, such as the Free Energy Principle (FEP) and active inference, to a different medium whose behaviour more directly governed by FEP, can emergent learning be achieved with alternative mediums? Electro-Active Polymer (EAP) hydrogels were embedded in the simulated game-world of Pong via custom Multi-Electrode Arrays and feedback between motor commands and stimulation. Through performance analysis within the game environment emergent learning was demonstrated, driven by ionic memory-like behaviours of the hydrogels. These observations enforced the theories of FEP within learning and its ability to allow learning in other mediums.

Electro-Active Polymers | Emergent Learning | Free Energy Principle

## Introduction

Emergent computing is a field inspired by the computation that takes place throughout nature and is described as highly complex processes arising from the cooperation of many simple processes,<sup>1</sup> or when behaviour of a system does not depend on its individual parts but on their relationships to one another.<sup>2</sup> Artificial neural networks were developed based on the emergent computing behaviour of biological brains. Both are large complex systems composed of simple machines whose combined interactions lead to complex computation or thought. The expansion of these theories led to reservoir computing and the application of physical non-linear systems, called reservoirs, as part of the computation process. By using reservoirs found in nature, reservoir computing techniques allowed the emergent computation of nature to be utilized in practical applications. For example, with water ripples for image analysis,<sup>3,4</sup> mycelium mold as maze solvers,<sup>5</sup> and chemical reactors as logic gates.<sup>6,7</sup> These systems solve calculation tasks and often provide unexpected solutions. However, they do not exhibit the emergent learning behaviour observed with biological brains. Without learning behaviour a reservoir system's capability to improve is limited. Continuous tasks that require behaviour based on task history are too complex to yield any useful result. With the correct combination of task

and computational medium, can the emergent learning behaviour similar to that observed in biological neural networks (BNNs) be achieved thus, allowing more complex and nuanced tasks to be pursued through reservoir computing techniques.

To harness the computational power of BNNs, for calculation problems, the field of Artificial Neural Networks (ANNs) was developed by applying the learning capabilities of biological structures. ANNs have made great strides in solving problems via machine learning that were previously thought to be impossible via computers. However, these solutions are limited by hardware implementation<sup>8,9</sup> as approximations of BNNs, ANNs are not capable of the learning found in their biological inspiration.

To better understand how the interactions of biological neurons allow for such complex useful behaviour, for improved ANN implementations, theories of morphological and embodied cognition were employed<sup>10,11</sup> both exploring the integration of computing behaviour with the physical 'hardware'.<sup>12</sup> Embodied cognition theorises that a body's environmental interactions constitute or contribute to cognition, meaning that learning behaviour results from the method information acquisition, thus resulting from its physical structure.<sup>13</sup> Both theories contribute to the idea that to achieve learning behaviour of BNNs in ANNs, hardware capable of the complex interactions present in BNNs is needed. The concept of ANN implementation outside of digital hardware lead to the development of reservoir computing.

Reservoir computing derives from Recurrent Neural Network (RNN) frameworks. The dynamics of a fixed non-linear system, called a reservoir, are used as part of a neural network, mapping input and output signals to higher dimensional space.<sup>14</sup> Reservoirs are typically physical systems exhibiting complex behaviour used to encode data, embodying computation not possible within a typical ANN structure. Reservoir computer structures consist of three main layers, excitation layer, reservoir, and readout layer. The excitation layer encodes network input into reservoir stimulation. The readout layer interprets reservoir responses into usable data using weights to allow performance tuning and optimisation. The reservoir can be any kind of medium that encodes temporal problems into higher dimensions creating recurrent connections in data.<sup>15</sup>

A promising application of reservoir computing to achieve biological learning capabilities, is through biological neurons themselves.<sup>16</sup> Integration of grown neuron cultures and electronic interfaces produce computational devices<sup>17</sup> referred to

as organic or 'wetware' computers. By stimulating BNNs with encoded task information, neurons will form pathways in response. This leads to learning behaviour in computation via Neural plasticity, the nervous system's ability to change activity in response to intrinsic or extrinsic stimuli by reorganizing its structure, functions, or connections.<sup>18</sup> Wetware computers are usually implemented through control of network plasticity, achieved via application of patterned or random stimulation.<sup>19</sup> By controlling plasticity, the rate of network reorganisation is controlled. A feedback loop is implemented where the network is 'punished' by increasing plasticity or 'rewarded' by reducing plasticity.<sup>20,19</sup> Implementations of this technology include simulated flight control<sup>21</sup> and the arcade game 'Pong'.<sup>22</sup>

In the paper "In vitro neurons learn and exhibit sentience when embodied in a simulated game-world"<sup>22</sup> a grown culture of neurons learned to play the game Pong in a virtual environment. The game environment is encoded into BNN stimulation with a region designated for sensing, used as motor commands to drive the paddle position. Using rally lengths as the performance metric with a closed loop that manages neural plasticity, the BNN (referred to as DishBrain) improved in ability while playing the game. The method of controlled plasticity, and BNN learning, is derived from theories of emergent intelligent behaviour in biological brains via the Free Energy Principle (FEP).<sup>23,24,25</sup>

FEP suggests that any self-organizing system that is at equilibrium with its environment, must minimize its free energy.<sup>23,26,27</sup> This is a formulation of how adaptive systems resist a natural tendency to disorder.<sup>27,28</sup> FEP interprets learning as an agent bounding it's maximum surprise or entropy, limiting free energy by entering states that alter sensed information.<sup>29,27,30,25</sup> This creates a feedback loop where internalised environmental information is continually improved by actions altering the environment, causing ordered interaction between the agent's internal states. This developed ordering of internal states is known as self-organisation, feedback corrects deviations from an ordered configuration.<sup>31,32</sup>

FEP in BNNs presents through self-organisation of synaptic pathways<sup>33,34</sup> via active inference. In active inference, an internal generative sudo model predicts inputs representing the environment,<sup>23,26,35</sup> continually updating the internal model to match external events. As sensed inputs, different from those expected via the model, are received, free energy increases. The system seeks to reduce free energy by the most direct means. This presents as the self-organisation of neural connections.<sup>36,37</sup> As inputs influence the structural change, the new structure is representative of the inputs, and by inference is representative of the environment the inputs came from. Restructuring presents behaviourally as either, action to reduce difference between the internal model and sensed environment by making the environment match the model or, by altering the internal model to better reflect the environment. Under this theory, BNNs hold 'beliefs' about the state of the environment, and learning behaviour emerges by minimising internal free energy through either, updating these beliefs, or taking action to change the environment to match these beliefs.<sup>26,38</sup> The aforementioned paper<sup>22</sup> showed these theories to hold merit, as the application of BNNs to the Pong game effectively allowed improved performance over time.<sup>22</sup> However, this learning is fundamentally achieved via FEP,<sup>23,35</sup> and BNNs are not the only system that present

FEP based behaviour.

Ionic Electro-Active Polymer (EAP) hydrogels are an active matter material<sup>39</sup> showing promise in soft robotics as actuators.<sup>40</sup> Recent studies have shown the potential for these materials behaviour to be used in computation<sup>41</sup> given an appropriate application and framework. As an ionic EAP, this material changes shape in response to electric field stimulation, ions acting as active agents. The dynamics of these active particles have shown memory like behaviour,<sup>42</sup> explainable through FEP. Electric fields cause an increase in free energy, the ions seek to minimise this free energy through migration.<sup>43</sup> Therefore, if the EAP hydrogel stimulation was representative of an environment, would the redistribution of ions represent an internal model of that environment, similar to the structure of BNNs within the brain? Furthermore, if both ionic EAP hydrogels and BNNs exhibit behaviour via the FEP, then is it possible that EAP hydrogels can exhibit the same learning capabilities if applied to an established BNN task such as Pong?

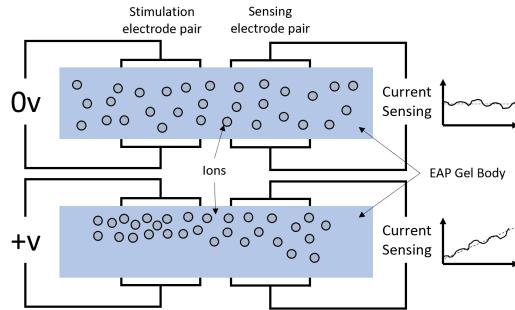
In summary, there are many computation techniques that are rapidly developing to achieve the kind of emergent learning behaviour found in BNNs. Theories in embodied computation have led to developments in ANNs, allowing physical mediums with complex behaviours to be used as computational resources.<sup>14</sup> Through reservoir computing, BNNs were integrated into computational tasks using fundamental theories of FEP.<sup>22</sup> However, BNNs are not the only medium whose behaviour is a result of FEP. Ionic EAP hydrogels are also governed by FEP, ions migrating under the influence of electric fields to reduce system free energy.<sup>43</sup> If similar emergent learning can be achieved with this medium, this confirms the core principles of free energy as the driving force of learning and allows for the potential for other FEP based systems to be used.

This study aims to show how EAP hydrogels can exhibit emergent learning behaviour, similar to that found in BNNs, when embodied in a simulated game-environment. First, by measuring ion concentrations through conductivity of the EAP hydrogel after periods of electric stimulation, the memory mechanics of the ionic EAP hydrogel were studied. Secondly, the EAP hydrogel was embodied in the simulated game environment of Pong through a custom Multi-Electrode Array (MEA). The game Pong was used as an established task for this study via current work in wetware computers.<sup>22</sup> The game environment was encoded into stimulations provided to the hydrogel, recorded ion concentrations were used as motor commands. Multiple game 'runs' were conducted with this setup and the game performance over time recorded and analysed.

## Nomenclature

$\bar{c}$	Mean value of $c$ , $\bar{c} \in \mathbb{R}^+$	181
$\epsilon$	Absolute permittivity of material, $\epsilon \in \mathbb{R}^{>0}$	182
$\mu_U$	Mean of $U$ , $\mu_U \in \mathbb{R}^{>0}$	183
$\Omega$	Number of microstates, $\Omega \in \mathbb{Z}^+$	184
$\Phi$	Normal distribution	185
$\sigma^2$	Variance, $\sigma^2 \in \mathbb{R}^{>0}$	186
$\sigma_U$	Standard deviation of $U$ , $\sigma_U \in \mathbb{R}^{>0}$	187
$A$	Surface area of electrode, $A \in \mathbb{R}^+$	188
$c_i$	Number of occurrences of the $i^{th}$ element of $O$ , $c_i \in \mathbb{Z}^+$	189
$C_{cont}$	Continuity adjustment, $C_{continuity} = -0.5$	190
$C_{ties}$	Adjustment due to tied ranks, $C_{ties} \in \mathbb{R}^{>0}$	191

192	$D$	Diffusion constant, $D \in \mathbb{R}^+$
193	$d$	Distance between opposing electrode plates, $d \in \mathbb{R}^+$
194	$e$	Electron charge, $e = 1.602 \times 10^{-19} C$
195	$F$	Faraday constant, $F = 9.649 \times 10^4 C \cdot mol^{-1}$
196	$K$	Set of unique ranks with ties, $K \in \mathbb{R}$
197	$k$	Rank, $k \in K$
198	$k_B$	Boltzmann constant, $k_B = 1.381 \times 10^{-23} m^2 \cdot kg \cdot s^{-2} \cdot K^{-1}$
199	$k_e$	Coulomb constant, $k_e = 8.988 \times 10^9 N \cdot m^2 \cdot C^{-2}$
200	$m_1$	Concentration at ion, $m_1 \in \mathbb{R}^{>0}$
201	$m_2$	Concentration between electrodes, $m_2 \in \mathbb{R}^{>0}$
203	$n$	Size of dataset, $n \in \mathbb{Z}^{>0}$
204	$O$	Set of elements in $Q$ without repetitions, $O \in Q$
205	$o$	Number of samples in $O$ , $o \in \mathbb{Z}^+$
206	$p$	P value, $p \in \mathbb{R}^{>0}$
207	$Q$	Sample used in distribution, $\{Q \in \mathbb{Z}   0 \leq Q_i \leq 1000\}$
208	$q$	Number of samples in $Q$ , $q \in \mathbb{Z}^+$
209	$R$	Gas constant, $R = 8.315 J \cdot mol^{-1} \cdot K^{-1}$
210	$r$	Distance of ion from electrode pair, $r \in \mathbb{R}^+$
211	$R_s$	Lowest ranked sum value, $R_s \in \mathbb{R}^{>0}$
212	$S$	Entropy, $S \in \mathbb{R}^{>0}$
213	$T$	Absolute Temperature, $T \in \mathbb{R}$
214	$t_k$	Number of ties for the $k^{th}$ rank, $t_k \in \mathbb{Z}^{>0}$
215	$U$	U statistic, $U \in \mathbb{R}^{>0}$
216	$u$	Free energy, $u \in \mathbb{R}^{>0}$
217	$v$	Voltage of electrodes, $v \in \mathbb{R}$
218	$v_1$	Ion Voltage potential at starting location, $v_1 \in \mathbb{R}$
219	$v_2$	Ion Voltage potential between electrodes, $v_2 \in \mathbb{R}$
220	$x$	Distance over which ion motion is measured, $x \in \mathbb{R}^+$
221	$Z$	Charge of ion, $Z \in \mathbb{R}$
222	$z$	Z score, $z \in \mathbb{R}$



**Fig. 1.** Diagrams show how ion motion is influenced by external stimulation via an electric field and recorded via current draw. Under stimulation ions will gather between the electrode pair from elsewhere in the EAP gel. As the concentration of ions in this location increases the conductivity also increases, the increase in conductivity can be measured as an increase in current draw when a small voltage is applied in parallel close to this location via a secondary electrode pair.

recording the ion concentrations before and after consecutive stimulations, the "remembered" state caused by the stimulation can be observed. Ion concentration can be measured through the conductivity of the gel. As ions collect in a location and ionic concentration increases, the local conductivity also increases.<sup>49</sup> By applying a small voltage to said location, the current draw on the power supply is directly proportional to the conductivity and thus directly proportional to the ion concentration. By utilising two adjacent pairs of electrodes, stimulation can be applied while conductivity is simultaneously measured. This setup is illustrated in figure 1.

To better understand hydrogel EAP system behavior, as shown in figure 1, the basic mechanisms can be described through Frick's first law of particle diffusion<sup>50</sup> characterising diffusive flux  $J$  over a free energy gradient  $\delta u$ , shown in equation 1. There are three main components to the free energy  $u$  of ions within the hydrogel system, shown in equation 2:

- Energy due to the electric field imposed by the electrode pair  $u_e$ .
- Energy imposed by the changing chemical concentration gradient within the hydrogel as ions migrate  $u_c$ .
- Energy due to the electrical potential gradient of ions within the hydrogel as ions migrate  $u_p$ .

$u_e$  and  $u_c + u_p$  (electrochemical gradient) oppose each other. As the electric field attracts ions to a location the local concentration and potential increases, causing an increase in free energy opposing that of the electric field.

$$J = \frac{D}{RT} \frac{\delta u}{\delta x} \quad [1]$$

Where:

$$\delta u = u_e + u_c + u_p \quad [2]$$

$$u_e = k_e \frac{v e A e}{dr} \quad [3]$$

$$u_c = RT \ln \left( \frac{m_2}{m_1} \right) \quad [4]$$

$$u_p = ZF(v_1 - v_2) \quad [5]$$

## Memory Mechanics Through Ion Migration and Conductivity Measurement

For a medium to be used in computation it must exhibit behaviours representative of properties key to computational system, the most fundamental of this is memory. Polyacrylamide hydrogel is an ionic EAP that displays mechanics that can be interpreted as a form of computational memory.

Polyacrylamide hydrogel is an active matter EAP,<sup>43</sup> hydrogen ions act as active agents within the hydrogel influenced by the polymer network and each other. Stimulation by an electric field causes an increase in free energy within the hydrogel, the ions migrate<sup>44</sup> to minimise the free energy of the system.<sup>43</sup> As the ions move, they drag water molecules causing changes in water distribution and localized deformation of the gel, driven by an equilibrium between osmotic pressure<sup>45</sup> and rubber elasticity<sup>46</sup> in the polymer network. Given a constant migration of ions into a location the localised rate of swelling gradually decreases, further decreasing free energy. This creates a hysteresis effect<sup>47</sup> which leads to complex dynamics in the gel's reactions to stimulation.

The hydrogen ions take relatively little time to migrate under stimulation by an electric field, but take considerably longer to diffuse to a homogeneous distribution under no stimulation. The difference in time scale allows previous stimulations to affect future stimulations, as the ion distributions persist between stimulations leading to a form of memory.<sup>48, 42</sup>

This experiment is designed to demonstrate this memory mechanic and detail the underlining methodology that is used in the subsequent experiment presented in this study. By

279  $u_e$  can be represented via Coulomb's law and electric potential energy, shown in equation 3. Ion charge is equal to electron charge, the charge imposed by the electrode pair is derived from the capacitance of the electrodes surrounding the hydrogel.  $u_c$  can be derived from the Gibbs free energy representation of entropy as the ratio between concentration increases,<sup>51</sup> shown in equation 4.  $u_p$  is represented via the Gibbs free energy derivation of cell potential as the combined charge of ions becomes inhomogeneous,<sup>51</sup> shown in equation 5.

280 Due to Coulomb's law, as the ions approach the charged electrode pair the  $u$  increases resulting in a larger flux  $J$ .  
281 However,  $u_c + u_p$  also increases opposing the electric field.  
282 These two opposing forces change at very different rates, the difference in magnitude being a key factor of the hysteresis component. As the ions migrate under the influence of the electric field, eventually the  $u$  is minimised as the opposing factors balance. This interpretation is, however, a simplification.  
283 In actuality there are many additional free energy components that contribute to the complexity of the hydrogel's EAP behavior such as terms associated with the deformation the gel's polymer structure as it swells due to local ion and water concentrations. These additional terms also influence factors of other terms such as diffusion coefficients and permittivity of the hydrogel material, resulting in extremely complex behavior as the free energy is minimised.

305 **Experimental Design.** Using this setup, a series of stimulations  
306 were applied to a polyacrylamide hydrogel sample to demonstrate  
307 the memory mechanism. Polyacrylamide is a relatively  
308 simple EAP hydrogel to synthesise,<sup>52</sup> allowing for it to be  
309 produced in large batches and moulded in many different ways  
310 for the required experiments. The gels were synthesised using  
311 the methodology described in the supplementary information  
312 section "Gel Synthesis Procedure". The gel was left in a sodium  
313 chloride solution (0.08%) to allow the ion concentrations to  
314 stabilize and provide consistent starting water concentration  
315 within the hydrogel. The gel was then placed between two  
316 electrode pairs spaced with a 7mm gap between them. A con-  
317 stant sensing voltage of 2 V was applied to one electrode pair,  
318 using a current sensor to record the response, while a series of  
319 20 V stimulations were applied to the second electrode pair.  
320 Both voltages were applied with the same polarity.

321 Temperature and water content within the hydrogel can  
322 affect free energy equilibrium.<sup>43</sup> Temperature alters the elastic  
323 properties and rate of ionic motion.<sup>52</sup> As the osmotic pressure  
324 difference between the polymer networks and ionic solution  
325 drives the swelling, changes in the hydrogel's water content  
326 change the degree to which it can swell, as well as altering  
327 the mechanical properties of the material.<sup>53</sup> Because of this,  
328 temperature and water content must be controlled. To mit-  
329 iate this as much as possible, the water was taken from a  
330 temperature controlled source (approximately 22C) to ensure  
331 the solution was a consistent temperature and provide the  
332 same starting temperature conditions. Furthermore, the gels  
333 were stored in water tight containers and left in the ionic  
334 solution for the same period of time to ensure the starting  
335 water content of the hydrogel was consistent.

336 **Results.** The results of this demonstration can be seen in figure  
337 2. The recorded current has a default value of approximately  
338 0.8 mA as the EAP hydrogel is inherently conductive shown

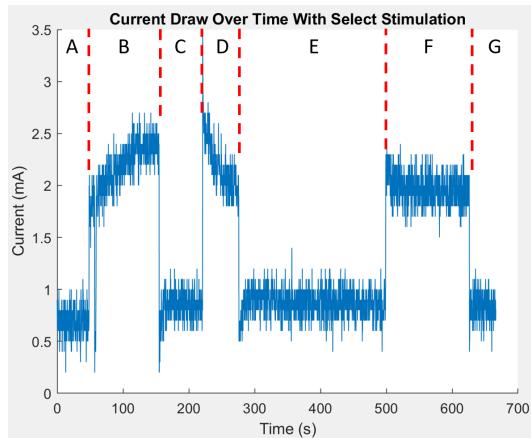


Fig. 2. This graph shows a sample of recorded current draw from 2 adjacent parallel electrodes as described in 1. The data has been segmented and labelled **A** through **G** for reference. The sensing voltage (2 V) was applied continually through the entire demonstration, stimulation (20 V) is applied in sections **B**, **D**, and **F**. The recorded current has a default value of approximately 0.8 mA as the EAP gel is inherently conductive shown in segment **A**. With application of an electric field there is an immediate rise to 2 mA as the field is picked up by the sensing electrode pair shown at the beginning of segment **B**.

339 in segment A. With application of an electric field there is  
340 an immediate rise to 2 mA as the field is picked up by the  
341 sensing electrode pair shown at the beginning of segment B.  
342 The stimulation also causes an increase in free energy. As a  
343 result, the ions move toward the electrode pair to minimise free  
344 energy. The current draw rises, as ions collect and conductivity  
345 rises, to 2.4 mA shown in segment B. Once the stimulation is  
346 removed the current drops back to default values as shown in  
347 segment C. Once the stimulation is reapplied, the current draw  
348 immediately rises back to the point where segment B ended as  
349 the ions still maintain their positions, as shown at the start of  
350 segment D. This observed behaviour exhibits a memory like  
351 mechanism where a value is "saved" to the EAP hydrogel by  
352 the redistribution of ions within the gel in response to the  
353 increase in free energy. This behaviour can then be seen again  
354 at the beginning of segment F where it continues from where  
355 segment D finished.

356 **Discussion.** From this demonstration, it is evident that  
357 changes in the ion distribution and polymer network can be  
358 recorded as current draw through the application of small volt-  
359 ages. This is evidenced by the change in current over time as  
360 stimulation is applied altering the ion distribution and polymer  
361 network. The demonstration also highlights how the increase  
362 in free energy through stimulation is resolved by the hydrogel  
363 through ion migration, and the last state of the EAP hydrogel  
364 "remembered" if stimulation is removed. With reapplication of  
365 the stimulation the changing current continues from the last  
366 state recorded during stimulation. From these observations  
367 it can be established that a form of memory is present in  
368 the EAP hydrogel, and the response due to the hydrogel's  
369 memory can be measured in parallel with the application of  
370 stimuli. To fully demonstrate this memory mechanic and its  
371 value within computation, a suitable activity is required, along  
372 with a closed loop control structure and hardware to interface  
373 with the hydrogel.

## 374 EAP Hydrogel Embodied in a Simulated Game-world

375 The behaviours of the EAP hydrogel can be characterized  
376 through free energy minimisation.<sup>43</sup> There are many systems  
377 whose mechanics can be represented through free energy min-  
378 imisation, in computation one of the most significant is that  
379 of learning within BNNs.<sup>23</sup> Furthermore, in the most basic  
380 sense, biological brains establish memory through the arrange-  
381 ment of neurons<sup>54</sup> much like the EAP hydrogel does through  
382 ion distribution. For the demonstrated memory mechanics  
383 to exhibit as learning, the EAP hydrogel would need to be  
384 embodied within a closed loop, as is present in BNN systems.  
385 To induce learning, the hydrogel must be able to influence  
386 actions within an environment. The change in environment as  
387 a result of those actions must feedback to the hydrogel leading  
388 to changes in actions and learning behaviour. To construct  
389 this closed loop and quantify learning a suitable activity is  
390 required.

391 The paper "In Vitro Neurons Learn and Exhibit Sentience  
392 When Embodied in a Simulated Game-World"<sup>22</sup> used the game  
393 Pong as the quantifying activity when integrating a BNN with  
394 computer architecture. The game provided a virtual environ-  
395 ment for the BNN to inhabit and learn from. The method  
396 used to interface the biological neurons with the simulated  
397 environment, fits neatly with the reservoir computing frame-  
398 work.<sup>14</sup> In the paper<sup>22</sup> the position of the ball was presented  
399 to the neurons through localized stimulation, aligning with  
400 the function of the excitation layer in a reservoir computer.  
401 The neuron responses were converted to motor commands to  
402 control the paddle, aligning with the function of the readout  
403 layer. Given the commonalities in the underlying mechanics,  
404 in relation to FEP and memory, the performance of compu-  
405 tational learning within the EAP hydrogel can be assessed  
406 using the same Pong game-based activity, in a simulated game  
407 environment, utilising the reservoir computing framework.

408 **Experimental Design.** An MEA was used to interface the EAP  
409 hydrogel with a computer system containing the Pong game  
410 environment. The array of electrodes were divided into two  
411 regions, used to provide stimulus as input to the gel via the  
412 application of electric fields at select locations, and record  
413 output via the electric current at select locations. This divi-  
414 sion of regions is similar to how MEAs are currently used in  
415 applications.<sup>22,55</sup> By dividing areas into specific tasks, inter-  
416 ference from different tasks happening in the same location is  
417 minimised. The layout of these regions can be seen in figure 3  
418 section A.

419 The stimulation region of the gel acts as the stimulation  
420 layer of the reservoir framework, consisting of a 2 by 3 array.  
421 To translate the ball's position into stimulations the game  
422 environment was divided into a 2 by 3 grid, as shown in figure  
423 4 section B. The location of the ball in the game environment is  
424 sent from the computer to the MEA, as the ball passes through  
425 those regions the corresponding electrode pair is activated to  
426 stimulate that area. The regions as they are labelled in black  
427 in figure 4 section B are how they will be referenced throughout  
428 this paper.

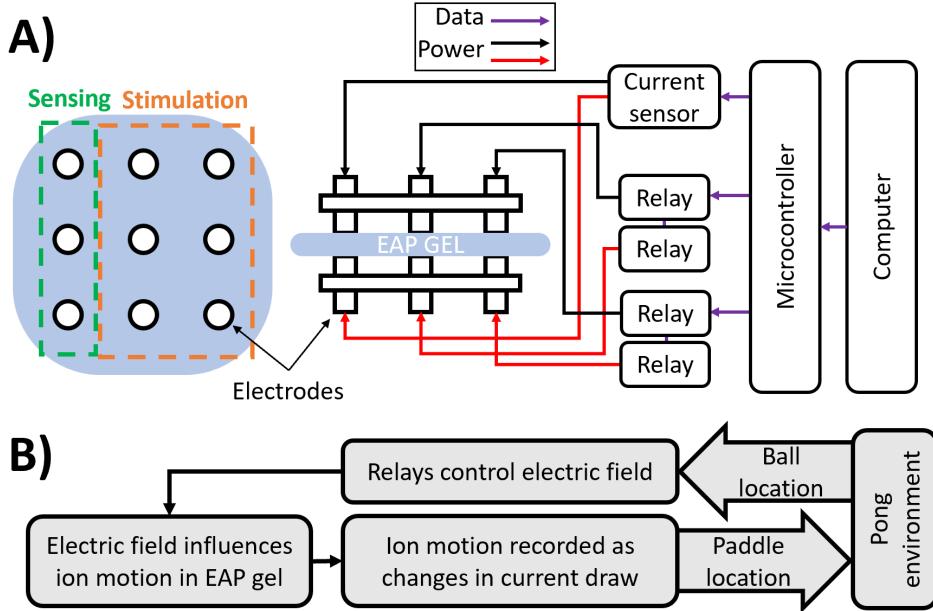
429 The sensing region acts as the readout layer of the reservoir  
430 framework. Current sensors were used to measure the electric  
431 current draw at the sensing electrodes, utilising the small  
432 sensing voltage (2v) defined in the previous experiment. These  
433 current values are recorded on screen, shown in figure 4 section

434 A recorded in mA, and interpreted into motor commands for  
435 the paddle. Recorded electrical current values are normalised  
436 using the maximum and minimum current values recorded  
437 from the hydrogel during the previous experimentation with  
438 the apparatus, found to be +3mA and -0.75mA respectively. A  
439 trend is then calculated using these normalized values against  
440 their locations represented in the Pong game environment.  
441 The trend is generated by fitting a 2nd degree polynomial to  
442 the 3 points. The peak of this trend is the predicted point  
443 of highest current, and highest ion concentration, in the gel  
444 and is where the paddle is placed. This is illustrated in the  
445 example in figure 4 section C.

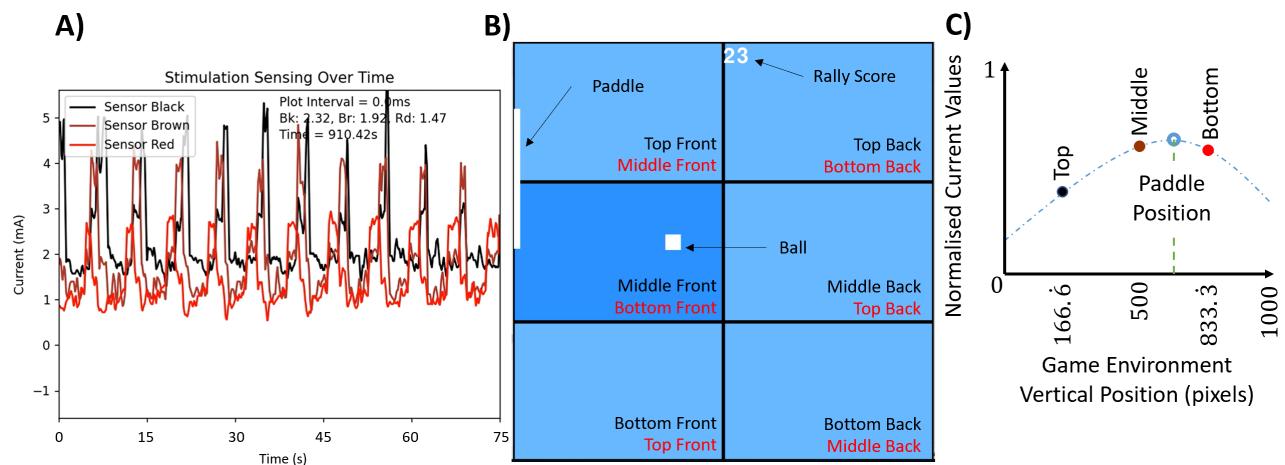
446 Through this layout of hardware and software, a closed  
447 loop was established where the computer outputs the game's  
448 environment, represented as ball location, to the MEA applied  
449 to the gel as shaped electric fields. In response, the ion con-  
450 centrations are measured by the MEA as electric current and  
451 provided to the computer and the Pong game to move the  
452 paddle. This closed loop is outlined in figure 3 section B. A  
453 more detailed description of the behaviour and implementation  
454 of the Pong game, as shown in figure 4, can be found in the  
455 supplementary information section "Pong Game Implemen-  
456 tation and Behaviour". Similarly, a more detailed description  
457 of the hardware implementation, as shown in figure 3 section  
458 A, can be found in the supplementary information section  
459 "MEA Hardware Implementation and Details". A circuitry  
460 schematic and image of the physical hardware can found in  
461 supplementary information figures S1 and S2. The full setup  
462 of the computer game environment and MEA apparatus can  
463 be seen in the supplemental information figure S3. The full  
464 experimental procedure for each run of the Pong game with  
465 the EAP hydrogel, can be found in the supplementary informa-  
466 tion section "Pong Game Experimental Procedure". In total  
467 21 separate EAP hydrogel runs were carried out, collecting  
468 3500 seconds of 'game play' for each run before the hydrogel  
469 degraded beyond the point of continuing.

470 From these games several metric were recorded to be used  
471 in performance analysis, electric current measurements, stim-  
472 ulation positions, and score. The score, shown in figure 4 section  
473 B, increases whenever the paddle successfully hits the ball and  
474 resets to 0 whenever the ball misses the paddle hitting the wall.  
475 All data from these separate runs was combined into a single  
476 dataset to view the overall averaged behaviour of the hydrogel  
477 and assess the performance as well as the repeatability.

478 A set of experiments were also carried out to create a  
479 baseline to compare performance against. This experiment  
480 explores the null hypothesis that the hydrogel's performance  
481 increase was a result of accurate game environment information.  
482 To explore this, the game environment's regional correlation  
483 to the MEA's stimulation electrodes were rearranged. This  
484 meant the gel received incorrect information about the ball's  
485 position. The rearrangement of game environment regions  
486 can be seen in figure 4 section B, indicated by the red region  
487 labels. For the null hypothesis to be rejected, the hydrogel  
488 would need to exhibit no learning. For this to be true, the  
489 hydrogel would need to perform worse than if the paddle  
490 was moving randomly. As the paddle is 1/3 the height of the  
491 simulated environment, the rejection would require a hit/miss  
492 ratio below 33%. This experiment was run 10 times to assess  
493 the performance and repeatability. Once ran and analysed, the  
494 results were deemed stable enough to not require additional



**Fig. 3.** System layout of closed loop communication between the computer containing the Pong environment and the EAP hydrogel. **A)** The layout of communication between components in the system and separation of the electrodes into regions. The electrodes are divided into driving, which provide stimulation, and sensing, which measure current draw. The driving electrodes are driven by relays that direct the electric field (20 V). The sensing electrodes provide a small voltage (2 V) and the current draw is measured by a current sensor. Both relays and sensors are coordinated by a microcontroller, which is in turn directed by the computer and Pong game environment. A more detailed circuitry schematic can also be found in the supplementary information figure S1. Additionally, an image of the physical hardware can be seen in the supplementary information figure S2. **B)** A flow chart of the information path from the Pong environment on the computer to the hydrogel containing the ions.



**Fig. 4.** The software representation of the Pong game within the computer. **A)** The current sensor readings as they are received by the computer. The black, brown, and red lines represent the top, middle, and bottom sensors respectively. The interval between received data is also measured and averages at 0.5 ms, labelled as "Plot Interval". The averaged current across the graph for each sensor is measured for the purposes of establishing default values, labelled as "Bk", "Br", and "Rd". The total recording time is also displayed labelled as "Time". These sensor values move paddle in the Pong game environment through conversion to a position based on sensor location, as visualised in figure 4 section C. This can be seen on the computer screen along with the apparatus in the supplemental information figure S3. **B)** The simulated Pong game environment, which is separated into 6 regions as described in figure 3 section B. When the ball is in a region it will darken and the corresponding electrode pair will stimulate the EAP hydrogel. The score is displayed at the top and resets when the ball hits the left wall behind the paddle. The regions are also labelled in black here for reference within this paper. The paddle is set to 1/3 the environment height. The rearrangement of regions used for the baseline comparison dataset, where 'vision' is impaired, is indicated by the red region labels. This can be seen on the computer screen along with the apparatus in the supplemental information figure S3. **C)** An example of the method used to place the paddle based off the current readings as measured in figure 4 section A, shown here as the same representative colours and labelled with their region locations. The current readings are normalised between 0 and 1 and plotted based on their region's centre vertical location. A 2nd degree polynomial is applied, as indicated by the blue dashed line. The point where the polynomial is at its peak is the maximum predicted current, as indicated by the blue circle, and this is where the paddle is placed.

repetitions as all data recorded satisfied the rejection of the null hypothesis.

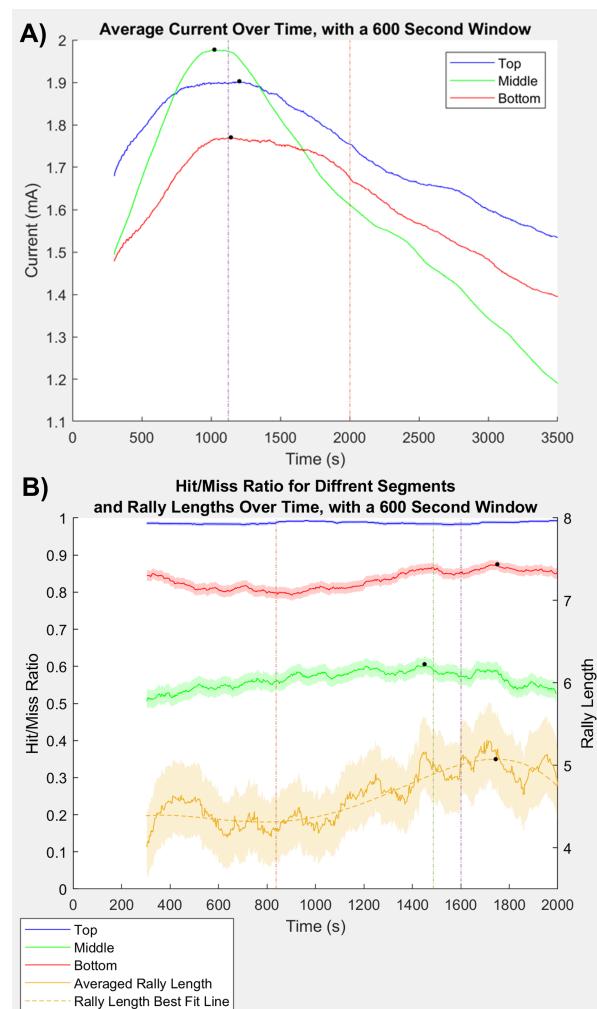
## Results.

**Memory Function of the EAP Hydrogels Via Electrical Current Measurement.** To show the memory mechanics in the combined MEA and Pong experimental structure, and follow the progression of the learning behaviour, first the measured current draw was plotted for each sensing electrode pair. This allowed an initial view of minimised energy in the form of electrical current, and established the time range in which learning occurs. Figure 5 section A represents the combined measured electric current data points of all 21 game runs, with data points sorted against time, averaged using a 600 second moving average window. The current initially rises as ions move towards the electrodes. As the ions collect their rate of migration decreases due to the increasing concentration gradient. Eventually, the ions stop moving and the electric current reaches its maximum, shown in figure 5 section A by the vertical purple dashed line. Then the hydrogel breaks down as the current falls, eventually leading to breakdown of the polymer structure as a result of electrolysis<sup>56</sup> and impairing connection between electrodes, as shown in the supplementary information figure S7. This breakdown marks a cutoff point, after which hydrogel is too degraded for results to be representative of performance, marked in figure 5 section A by the orange vertical dashed line at 2000 seconds. Thus, the graphs used to analyse the hydrogel's performance were limited to between 0 and 2000 seconds.

**Performance of the EAP Hydrogel Within The Pong Game.** To assess the performance of the hydrogel within the Pong game, the hit to miss ratio was calculated using the combined datapoints of the 21 game runs sorted by timestamp, shown in figure 5 section B as the left axis. 21 experiments were sufficient for reliable results shown via standard deviation analysis, described in the supplementary information section "Accuracy of Experiment Repetitions". Due to the boundary conditions of the paddle, the top and bottom regions of the wall were hit more often by the ball than the middle (see supplementary information figure S4). As hit rate is different between regions, learning rate may be different between regions. For this reason, the hit/miss data was segmented based on the region in which the hit/miss occurred. Figure 5 section B, shows the segmented paddle performance learning curve for the top, middle, and bottom game regions.

For the middle region, the paddle had an initial ball hit rate of 50%, but over the course of the game this rose to maximum of 60%, giving a rise of 10% over 1450 seconds. The middle region shows the greatest improvement in performance.

The top and bottom regions are subject to the boundary conditions leading to a higher initial hit/miss ratio, as seen in figure 5 section B, as the trend via the measured current in those regions is almost linear, seen in figure 5 section A. In the top region the paddle rarely missed the ball. The top sensor tended to read higher currents than the other regions, this is likely due to some bias in the MEA rig construction. This was however consistent through all experiments, including the baseline experiments, so did not impede the game performance analysis. For the bottom region the paddle initially hit the ball approximately four times more often than it missed, hitting



**Fig. 5. A)** The electrical current at each sensor over time for the combined unimpaired dataset of runs, smoothed using a moving average window of 600 seconds. Due to the averaging, data starts at 300 seconds. The point of maximum current draw is marked with a black dot. 1203, 1023, and 1142 seconds for top, middle, and bottom respectively. The average point of maximum current draw is marked by the purple vertical dashed line at 1122 seconds. The orange vertical dashed line marks the earliest point the EAP gels broke down. Standard error was calculated for each trend from the windowed samples and found to be at maximum, 0.0169, 0.0053, and 0.0116 for the top, middle, and bottom trends respectively. **B)** *Left Axis:* The hit to miss ratio over time for the combined unimpaired feedback loop dataset, separated based on the region the hit/miss occurred in as described by figure 4 section B. Smoothed using a moving average window of 600 seconds, this window size provided the clearest view, additional windows sizes can be seen in the supplementary information figure S5. The standard error for each windowed sample is shown by the shaded area. The point of maximum performance is marked on the "Middle" and "Bottom" lines by a black dot, 1450 and 1750 seconds respectively. The average between these times is marked by the purple vertical dashed line (1600 seconds). From the graph, the middle region saw a performance increase of 0.1, from 0.50 to 0.60. The bottom region decreased to 0.79 before increasing to 0.87, from its lowest point this is a performance increase of 0.08. *Right Axis:* The rally length (maximum score before a miss) against time for the combined unimpaired feedback loop dataset. Smoothed using a moving average window of 600 seconds. The standard error for each windowed sample is shown by the shaded area. A 4th degree polynomial best fit line was applied via the MATLAB Polyfit function.<sup>57</sup> This was the minimum order polynomial that fit the trend presented by the data. The maximum point of performance is marked by the black dot at 1744 seconds. This marks an increase of 0.8, from 4.3 to 5.1. This data, prior to the application of the moving average, was used to perform a one tailed t-test. From 0 seconds to the vertical orange dashed line (838 seconds), indicates the dataset used to represent before learning. From the vertical green dashed line (1487 seconds) to the cutoff of 2000 seconds, indicates the dataset used to represent after learning. The distributions of these datasets can be seen in the supplemental information figure S8.

554 79% of the time. As the bottom region is also a boundary condition  
555 this is expected. However, over time this performance  
556 did increase up to a maximum of 87% hit rate, showing an  
557 improvement of 8% over 1750 seconds.

558 To better show the overall performance increase through  
559 the course of the game, a graph was generated of the rally  
560 lengths against time. The rally length is a measure of how  
561 many times the ball was hit before it was missed, shown as the  
562 "rally score" in figure 4 section B. The score reached before  
563 it resets to zero when missed is the rally length. Each time  
564 the ball was missed the rally length achieved was recorded  
565 against the time, and plotted using an averaging window of  
566 600 seconds as with the previous plots. This plot is shown in  
567 figure 5 section B as the right axis, and shows an increase in  
568 rally length over the course of the game. This graph clearly  
569 shows the ability of the gel to perform within the Pong game  
570 before and after the learning period. With this separation  
571 of states, before and after learning, the significance of the  
572 increase in the hydrogel controlled paddle's performance, via  
573 the rally lengths, can be analysed.

574 The significance of the change in performance can be analysed  
575 through a statistical test (using null hypothesis  $\alpha = 0.05$ )  
576 between the rally length before and after the performance  
577 increase or "learning". The before learning sample is a subsample  
578 from 0 to 838 seconds in the rally length dataset, shown  
579 via the vertical orange dashed line in figure 5 section B. The  
580 after learning sample is a subsample from 1487 to 2000 seconds  
581 in the rally length dataset, shown via the vertical green dashed  
582 line in figure 5 section B. Details on how these ranges  
583 were selected can be found in the supplementary information  
584 section "Rally Length Subsampling for Learned and Unlearned  
585 States".

586 Although the pre and post learned values were recorded  
587 from the same gel samples, many data points were recorded  
588 per sample. Therefore, the variables representing before and  
589 after learning can be considered independent, as there is no  
590 one-to-one relationship between data points in the two states.  
591 Because of this, the statistical significance can be assessed via  
592 the p value of sample comparison, using critical limit theorem  
593 and the Mann–Whitney U test. Details on this calculation can  
594 be found in the supplementary information section "Statistical  
595 Significance and Mann–Whitney U Test Calculations".

596 These calculations give a  $p = 0.00041$ , placing the probability  
597 below 0.05% and rejecting the null hypothesis. This validates  
598 the improvement before and after learning as statistically  
599 significant and reinforces the observed increase in hydrogel  
600 paddle control ability in the Pong game environment. Showing  
601 that the improved hit rate, shown in figure 5 section B,  
602 was not regions exclusive, but contributed to improved  
603 overall game performance.

604 It can also be observed that the peak in rally length aligns  
605 with the average peak in hit rate at approximately 1600 seconds,  
606 observed in figure 5 section B, with a slight delay of around  
607 150 seconds. This delay makes sense as the effect of the  
608 improved hit rate would take some time to translate to an  
609 improvement in rally length, due to the additional layer  
610 of abstraction caused by the time taken for a full rally to  
611 be completed. Further to this, the average peak in hit rate  
612 observed in figure 5 section B, aligns with the peak current  
613 measurement observed in figure 5 section A, with another  
614 delay of around 600 seconds.

615 This increase in current represents the ion distribution of  
616 the hydrogel being influenced by the received environmental  
617 data, filling the memory similarly to how the memory component  
618 was observed in the first section of this paper as current  
619 measurement. As with the delay between hit rate and rally  
620 length, it would take some time for the base level current  
621 measurements to translate to improved hitting ability, due to  
622 the layers of abstraction caused by the mechanics of the game  
623 and MEA interface.

624 **Game World Performance in Relation to the Free Energy Principal.**  
625 The results of figure 5 sections A and B, show the learning  
626 capability through the improvement in game performance, via  
627 the EAPs control of the Pong paddle. However, the way in  
628 which the hydrogel dynamics, through free energy minimisation  
629 via ion migration, result in emergent learning is obscured.  
630 In other words, how does the change in free energy, as a result  
631 of polymer dynamics (electrostatic interactions between  
632 polymers and ions, entropy of polymer networks, ion migration,  
633 etc) and FEP, alter the paddle's behaviour with respect  
634 to ball's position result in increased game performance and  
635 emergent learning?

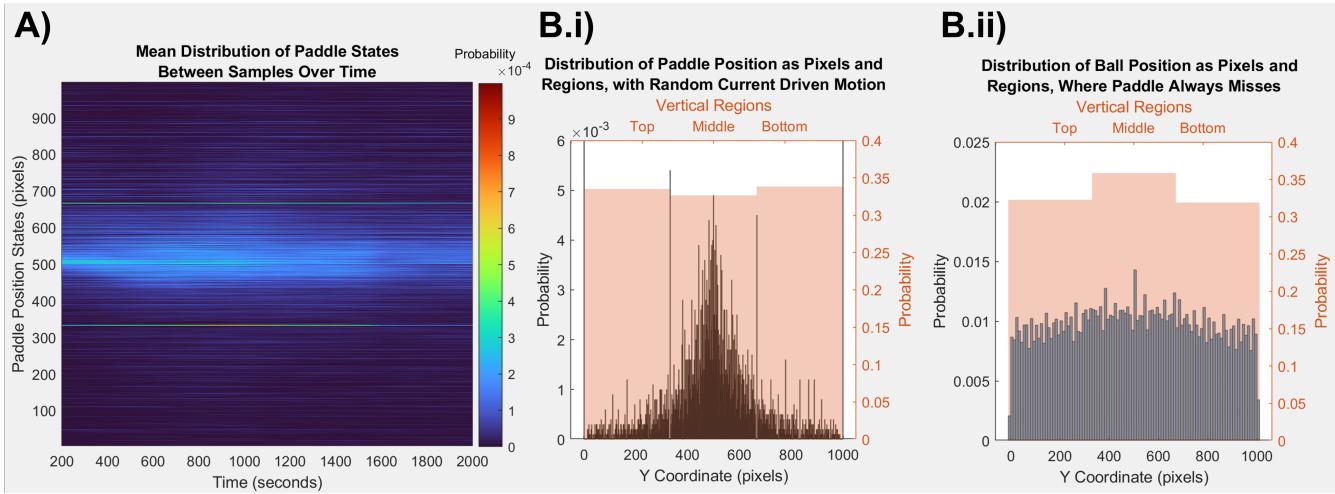
636 The molecular picture of the free energy is extremely complex,  
637 as the already complex EAP system (charged polymer  
638 networks interacting with ions and water molecules)<sup>43</sup> is, in  
639 this study, open, and coupled to the environmental dynamics  
640 of the Pong game through the electrostatic forces externally  
641 applied. However, recent theories of FEP can be a leading  
642 principle, since FEP incorporates direct coupling, or feedback  
643 loops, within the embodied environment and emergent learning  
644 process.

645 FEP is based on the assumption of the 'Bayesian brain'  
646 as an "inference generator", predicting environmental dynamics.<sup>29</sup>  
647 In FEP systems pursue actions of least surprise to minimise  
648 the difference between predictions based on their  
649 sudio internal model of the world and their perception of the  
650 world as input. This difference is quantified by variational  
651 free energy and is minimised through iterative feedback loops,  
652 through actively acting in the embodied environment to make  
653 it closer to the expected and predicted model of the environment.<sup>27,25</sup>

654 To summarise FEP in context of the EAP hydrogel, minimisation  
655 results as the equalisation between three main sources  
656 of free energy, shown in equation 2. These are; energy on  
657 ions due to the electric field, energy due to the ion and water  
658 chemical gradient, and energy due to the ion electrical gradient.  
659 The electric field application causes a gradient of increased  
660 free energy, shifting the point of equilibrium, so ions migrate  
661 to minimise this.

662 The game environment closed loop can be summarised as;  
663 measured electric current encoded into paddle motor commands,  
664 and the ball's position fed back via electric stimulation  
665 to the hydrogel. As the ions migrate to minimise free  
666 energy, under the electric field, the redistribution of ions affects  
667 the game environment through paddle motion since the paddle's  
668 motion alters ball's trajectory through collision. This  
669 implementation of closed loop free energy minimisation via  
670 environmental information, can be considered the framework  
671 of FEP. The results discussed indicate that through feedback,  
672 free energy minimisation, and environmental exploration FEP  
673 can lead to a form of emerged learning.

674 To develop the chain of events that culminate in improved  
675



**Fig. 6. A)** The mean distribution of paddle states for all 21 experimental samples, as a heatmap. Distributions were sampled every 20s starting at  $t=100$  using a window of 200s, providing enough datapoints for accurate representation with minimal sample overlap. The color of each point represents the state probability. States of 0 and 1000 pixels have been omitted, the probability of these states was much higher than others (0.0129, and 0.0177 respectively) due to the boundary conditions, so would obscure the data by scale biasing. An additional set of plots can be seen in the supplemental information figure S9, showing individual distributions for the timestamps 200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, and 2000. **B)** Distributions of positional states from paddle and ball simulations. Each plot shows the distribution divided into bins of each state, signified by the blue bars, and divided by regions shown by the orange transparent overlay bars. **B.i)** The simulated distribution of paddle states generated, using the algorithm described in figure 4 section C, with randomised current values. 10000 samples were simulated to show the distribution. To allow the central peak to be visible the left y-axis has been limited to 0.006, however the states of 0 and 1000 are higher than the central peak due to the boundary conditions. The states 0 and 1000 have probabilities of 0.2754 and 0.2796 respectively. **B.ii)** The simulated distribution of vertical ball states using the maximum hit/miss ratio for each region. The ball was simulated with the same game physics, starting with a random direction and resetting to the centre on a miss. When the ball hits the paddle wall the decision of if it will bounce follows the ratio at maximum performance, 0.98, 0.58, and 0.86 for the top, middle, and bottom regions respectively. As the ball moves the full path of each ball simulation was recorded resulting in 600000 samples, with the travel of each ball simulation before reset lasting on average 245 samples.

666 performance, the kinematics of the paddle were analysed over  
667 the course of the game through analysis of the changing distribution  
668 of paddle positions over time. This can also be approached through analysing the paddle's motion in response  
669 to the ball's location. This analysis is detailed in the supplementary information section "Paddle Motion and Standard  
670 Deviation Analysis".

671 Both the paddle and ball move in the y-axis, each vertical  
672 position the paddle or ball exists in can be thought of as a state  
673 that the paddle or ball inhabit. By applying a window to the  
674 data the change in positions can be shown as changing state  
675 distributions through the course of the game and analysed.

676 Figure 6 section A, shows the mean of the paddle positional  
677 state probability distributions as a function of time. Individual  
678 distributions can be seen in the supplemental information figure  
679 S9, showing probability distributions for timestamps 200,  
680 400, 600, 800, 1000, 1200, 1400, 1600, 1800, and 2000 using a  
681 larger window for direct comparisons of sample distributions.  
682 With these figures there is a clear change in the distribution  
683 of paddle positions as the game progresses. Initially starting  
684 with a large central peak that becomes less pronounced as the  
685 gel reaches its maximum performance point at 1800 seconds,  
686 as found from figure 5 section B, where the distribution be-  
687 comes more uniform. From this it can be deduced that the  
688 distribution of paddle positions changes through gameplay. In  
689 line with the principles of FEP this distribution of states can  
690 be thought of as a representation of the sudo internal model,  
691 as the paddle predicts the position of the ball. Similarly the  
692 ball's state distribution represents the actual environment. In  
693 figure 6 section A the states of  $y=0$  and  $y=1000$  were omitted  
694 as their larger values obscured the rest of the results. These  
695 larger values are a result of the boundary conditions. To fully

696 explore the learning mechanics, the impact of these boundary  
697 conditions must be investigated.

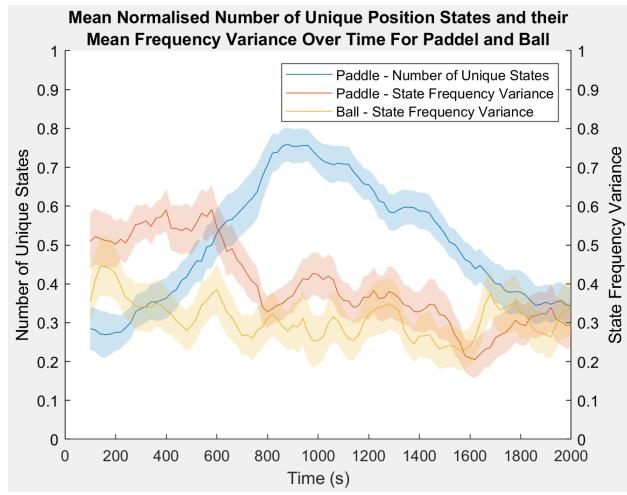
698 The boundary conditions occur as the paddle cannot pass  
699 beyond the edge of the game environment. If the measured  
700 current trend used to place the paddle, as described in figure  
701 4 section C, has a positive gradient toward either edge the  
702 paddle will be placed at the edge. This means that there are  
703 more combinations of current that lead to the paddle being  
704 at the edges than in-between. The boundary conditions can  
705 be visualised by simulating the paddle with random current  
706 inputs, using the algorithm described in figure 4 section C,  
707 and recording the paddle's positional states. A distribution  
708 of these simulated results is shown in figure 6 section B.i,  
709 this plot shows that with random current the paddle has a  
710 significant tendency toward either end with a peak distribution  
711 in the centre. The shape of distribution at the centre can also  
712 be explained as a result of the boundary conditions and is  
713 analogous with a depletion force. The boundaries in the  
714 paddle's motion act as two point forces with high attraction  
715 to the paddle; the forces interact at the mid-point between  
716 them creating a peak.<sup>58,59</sup> With greater resolution, the effect  
717 of the boundary conditions could be reduced. However, that  
718 is beyond the scope of this study. The distribution shown in  
719 figure 6 section B.ii is similar in shape to the initial distribution  
720 at  $t = 100$  in figure 6 section A, as the paddle's initial motion is  
721 almost entirely influenced by the boundary conditions having  
722 yet to gather enough information to diverge.

723 Although the boundary conditions explain the initial state  
724 distribution of the paddle, to understand the change in be-  
725 haviour the distribution at maximum performance ( $t=1800$ )  
726 must be assessed. Through the principles of FEP, if the pad-  
727 ple's state distribution represents the sudo internal model as  
728

predictions about the environment, and the ball's state distribution represents the actual environment, the system would seek to minimise the difference between these distributions. Indeed in an ideal Pong game the paddle position matches that of the ball's vertical position at all times, meaning that the ball and paddle should have identical distributions of vertical positional states. By comparing the paddle's state distribution against that of the ball, the paddle's change in behaviour can be assessed against what would be the ideal behaviour and compared to the principles of FEP. This allows analysis of what attributes of the paddle's distribution change to match that of the ball leading to improved performance. The ball's state distribution can be simulated using the Pong game environment. However, the ball resets on missing the paddle which alters the ball's distribution. This means that as performance improves, the ball's behaviour changes as it is hit more and follows different trajectories. This is evident when comparing two simulations of the ball's positional distribution, where the ball is never hit by the paddle and where it is always hit by the paddle, shown in supplementary figure S12. To accurately compare the paddle's distribution to that of the ball, the ball's distribution must represent the behaviour for the ball at the point of maximum performance. At this point the hit/miss ratios are 0.98, 0.58, and 0.86 for the top, middle, and bottom regions respectively, shown in figure 5 section B. When the ball hits the paddle wall, the ball is bounced according to the region's hit/miss ratio, this distribution can be seen in figure 6 section B.ii. From this distribution the ball's vertical motion over the game environment is more uniform than the paddle's initial distribution with fewer unique states (represented as the number of histogram bins), indicative of the predictable ball trajectories. From figure 6 section A, the paddle's distribution appears to become more uniform as the central peak becomes less pronounced. Additionally, the number of unique states reduces, also moving closer to the ball's distribution. These observations, however, need to be quantified via metrics representing the behavioural change of paddle and ball motion, as a result of the closed loop interaction them.

There are two main metrics used to analyse the attributes of the paddle and ball's motion distribution, both representative of entropy within the system aligning with theories of FEP and Bayesian inference.<sup>23, 26, 35, 60, 61</sup>

- Number of unique states: The number of unique states/positions the paddle/ball is recorded in within the sample window used to generate the distribution (also represented as the number of histogram bins). Defined in equation 7 as  $o$ , where  $Q$  is the sample used to make the distribution,  $O$  is the set of elements in  $Q$  without repetitions, and  $|O|$  represents the cardinality of  $O$ . When applied to the ball, due to the resolution of the ball's motion as applied to the gel, the number of unique states of the ball will remain constant, limited by the 6 regions the ball inhabits, so not providing usable information. Entropy  $S$  is a function of unique states  $u$  as the paddle position is determined by conductivity of ions and their distribution, in statistical thermodynamics entropy is defined as  $S = k_B \ln \Omega$ .<sup>62</sup>  $k_B$  is the Boltzmann constant, and  $\Omega$  is the number of microstates analogous with unique states  $u$ .
- Variance in frequency of unique states: The variance in quantity of occurrence of each unique state, used as a



**Fig. 7.** The mean of the normalised number of unique paddle states and the mean of the normalised variance between unique paddle state frequencies for both the paddle and ball, over the course of the game. Distributions were generated from each gel sample using a window size of 200 seconds, as in the plots in figure 6 Section A, sampled every 20 seconds from  $t=100$  to  $t=2000$ . The number of unique paddle states and state frequency variance were then normalised and the mean between gel samples found. The standard error between gel samples is show by the shaded area.

measure for uniformity in the distribution. Defined in equation set 6 as  $\sigma^2$ , where  $c_i$  is the number of occurrences of the  $i^{th}$  element of  $O$ ,  $\bar{c}_i$  is the mean value of  $c$ ,  $Q_j$  is the  $j^{th}$  element of  $Q$ , and  $q$  is the number of elements in  $Q$ . Shown in figure 6 section B.i, the boundary states of 0 and 1000 greatly overshadow the central peak distribution. As the edge conditions are results of the interfacing algorithm and only represent a fraction of the paddle's motion, they can be ignored for the purposes of calculating variance in the paddle. This allows the uniformity in the central peak to be more accurately compared to that of the ball which is not subject to boundary conditions, as evidenced by figure 6 section B.ii. Variance is linked to entropy through information theory. Although not directly proportional, entropy and variance are both measures of expected values and share trend directionality.<sup>63</sup>

$$\sigma^2 = \frac{\sum (c_i - \bar{c})^2}{o - 1} \quad [6]$$

$$o = |O| \quad [7]$$

Where:

$$c_i = \sum_{j=1}^q f(Q_j, O_i)$$

$$f(Q_j, O_i) = \begin{cases} 1 & Q_j = O_i \\ 0 & Q_j \neq O_i \end{cases}$$

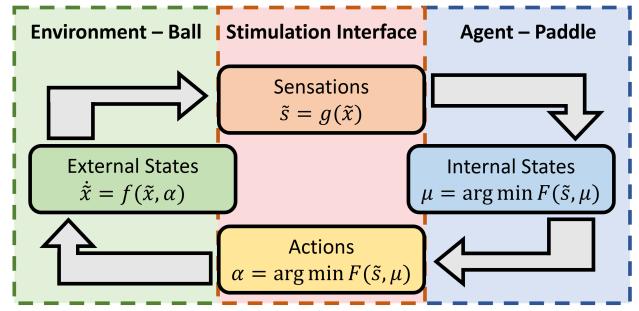
Due to the dynamic nature of the hydrogels, each sample will have slight variations in ion distribution, polymer density, and surface texture, this could lead to differences in recorded metrics. However, the learning behaviour is still driven by the same mechanics regardless of starting inconsistencies and, as a result, the trend shape of these metrics will be consistent between samples. To minimise the affect of gel synthesis inconsistencies, the metrics are normalised via feature scaling

801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816

817  
818  
819  
820  
821  
822  
823  
824

825 to better highlight the shape of the trends. The plot of these  
 826 trends is shown in figure 7. The lines represent the mean  
 827 between all samples with error bars representing the standard  
 828 error between all samples used.

- 829 1. At t=200: The number of unique paddle states is low while  
 830 the variance in both paddle and ball is high, showing few  
 831 states with large variations in frequency. This matches the  
 832 simulated distribution from figure 6 section B.i, expected  
 833 as the initial distribution will be almost entirely driven  
 834 by the boundary conditions, as no information has been  
 835 gathered about the game environment and ions remain  
 836 in a homogeneous distribution.
- 837 2. At t=600: The number of states increases as the variance  
 838 decreases in both paddle and ball. This happens as more  
 839 information about the environment is gathered and coincides  
 840 with the rise in current seen in figure 5 section A.  
 841 With the application of an electric field the ions start to  
 842 migrate within the hydrogel to represent the stimulation.  
 843 As ion mobility increases the number of unique states  
 844 increases and the paddle explores more of its working  
 845 area. Simultaneously, as more of the working area is ex-  
 846 plored, paddle motion becomes more evenly distributed  
 847 and so the variance decreases. As the ball is hit more  
 848 often its variance decreases as it moves through more of  
 849 the game environment before being reset on a miss. This  
 850 is evident when comparing variance of simulations where  
 851 the paddle always hits and always misses the ball shown  
 852 in supplementary figure S12.
- 853 3. At t=1000: The number of states reaches its maximum  
 854 and starts to decrease, this coincides with the behaviour  
 855 of the current in figure 5 section A and can be explained  
 856 through ion mobility. Initially the ions have total mobility  
 857 as the gel is yet to change in structure, as the ions mobility  
 858 increases hysteresis takes effect within the hydrogel  
 859 opposing ion migration. Eventually ion mobility reaches  
 860 its maximum causing the maximum peak in current seen  
 861 in figure 5 section A. After this point the hysteresis effect  
 862 continues and ion mobility decreases causing the number  
 863 of unique states to decrease, as ions start to settle into  
 864 their final positions.
- 865 4. At t=1800: Both the number of states and variance in  
 866 the paddle have reached their minimum, this coincides  
 867 with the point of maximum performance shown in figure  
 868 5 section B. The normalised variance in paddle and ball  
 869 now match as the paddle and ball's behaviour have a  
 870 stronger coupling. As the memory continues to saturate,  
 871 the ion distribution comes to better represent the ball's  
 872 behaviour. This leads to reduced paddle state variance  
 873 and reduced number of unique paddle states, causing  
 874 the behaviour of the paddle to become more like that  
 875 of the ball. This causes the ball to be missed less often,  
 876 resulting in a more homogeneous distribution of ball states  
 877 and less variance, as seen in supplementary figure S12  
 878 section B. The saturation of memory is caused by the  
 879 hysteresis and so is linked to the breakdown of the polymer  
 880 structure. Eventually, the same mechanic that allows the  
 881 gel to retain memory causes the polymer structure to  
 882 break down to a point that it inhibits conductivity and  
 883 performance reduces.



**Fig. 8.** Diagram of information flow and dependency between states of the free energy minimisation task derived from FEP equations.<sup>23,64,65</sup> External states  $\tilde{x}$  are of the environment in which the agent acts, synonymous with the probability distribution of the ball. Sensations  $\tilde{s}$  are the interaction by which the agent gathers information, synonymous with the regional stimulation. Internal states  $\mu$  are of the agent's internal model, synonymous with the probability distribution of the paddle. Actions  $\alpha$  are the interactions between agent and environment, synonymous with measured current interpretation.

The trends shown in figure 7, and the way in which they coincide with features in other figures, adds another step to the chain of events leading to the increased performance. This highlights interesting forms of emergent behavior occurring as the ions migrate to reduce free energy.

Baseline results were collected, as described in the experimental description, to further show that the hydrogel's improved performance was a direct result of accurate environmental information received. The results plot is shown the supplemental information figure S13. From these results the no region achieving a hit ratio beyond 30% and rally lengths averaged well below 1 point. The hit/miss was below 33% so the null hypothesis is rejected as the performance was worse than if the paddle were moving randomly. A more detailed analysis of this result can be found in the supplementary information section "Hit/Miss Baseline Comparison".

Even with this incorrect environmental information, the hydrogel is still being provided stimulation. This means the hydrogel is still learning as the ions move to a distribution representative of the information provided. Unfortunately, as the information provided is wrong, what is learnt by the hydrogel is also wrong. This observation helps to reinforce that the hydrogel's increase in performance is directly related to it being presented with an accurate representation of the virtual environment in which it is acting.

**Discussion.** From the results observed, there is a clear increase in performance of the hydrogel to play the Pong game, as shown in figure 5 section B, when given information that is representative of the virtual game environment. The development of the increased performance can also be directly linked to the underlying mechanics of the EAP hydrogel through ion migration and memory mechanics due to hysteresis. Through this linking of mechanics a logical similarity can be established between the properties of EAP hydrogels and theories of learning via FEP.

In FEP learning, external states of the environment influence internal states of the agent, in this case the EAP hydrogel, through a looped exchange of information called active inference. Active inference analogises an internal generative model used to predict inputs that represent the external world,<sup>23,26,35</sup> based on theories of Bayesian inference.<sup>60,61</sup>

This feedback of information flow is illustrated in figure 8 with equations showing informational exchange. Environmental states  $\tilde{s}$  are described by the constraints of the game physics defining the motion of the ball.  $\tilde{s}$  translate to sensations  $\tilde{s}$  via the agent's interface. In this case the stimulation interface translates the ball's pixel positions into regional stimulations. As such,  $\tilde{s}$  is synonymous with the probability distribution of ball regions represented in figure 6 section B.ii, supplemental figure S12, and used for state frequency variance of the ball in figure 7.  $\tilde{s}$  translate to internal states  $\mu$  in the agent as a minimisation of the agent's internal free energy  $F$ . The EAP hydrogel's internal structure is representative of these internal states, synonymous with the sudo internal model analogy of FEP learning and updated as sensations of information are received.  $\mu$  causes actions  $\alpha$  via the minimisation of  $F$  and as such  $\mu$  is synonymous with the probability distribution of paddle states represented in figure 6 section B.i, and used for number of unique states and state frequency variance of the paddle in figure 7.

It is clear from 7 the minimisation of the unique number of states and state frequency variance in both paddle and ball coincide with improved performance, and through figure 8 their part in the feedback of information can be seen as  $\tilde{s}$  and  $\mu$ . This link can be further highlighted through formal definitions of free energy  $F$  bound on Bayesian model evidence.  $F$  is defined in equation 8,<sup>64,65</sup> where  $D_{kl}$  is the Kullback–Leibler divergence,  $Q(\mu)$  is the probability density of internal states  $\mu$ , and the joint probability distribution of  $P(\tilde{s}, \mu)$  is such that  $P(\tilde{s}, \mu) = P(\tilde{s}|\mu)P(\mu)$ .

$$F(\tilde{s}, \mu) = \underbrace{D_{kl}[Q(\mu)||P(\mu|\tilde{s})]}_{\text{relative entropy}} - \underbrace{\ln P(\tilde{s})}_{\text{log of evidence}} \quad [8]$$

$$= \underbrace{D_{kl}[Q(\mu)||P(\mu)]}_{\text{complexity}} - \underbrace{E_Q[\ln P(\tilde{s}|\mu)]}_{\text{accuracy}} \quad [9]$$

Relative entropy cannot be less than zero, so that free energy is minimised when, in the Bayesian model, the approximate posterior becomes the true posterior and the free energy becomes the negative log of evidence for the internal sudo generative model.<sup>66</sup> Through minimisation of free energy, the difference between the  $\mu$  and  $\tilde{s}$  is minimised, as  $\tilde{x}$  is a component of  $\tilde{s}$  this results in environmental learning. The minimisation of relative entropy can be seen in figure 7 as the state frequency variance of both the paddle and ball become closer as performance increases. Log of evidence is representative of the hydrogel memory, as information is collected the memory becomes saturated leading to less impact on the overall system. Via the definition of equation 8, minimising free energy is equivalent to maximising model evidence, and equivalent to minimising the complexity of accurate explanations for sensations,<sup>65</sup> this is shown in the equivalent equation 9, where  $E_Q$  is the expected outcome respect to  $Q$ . Minimisation of free energy equates to minimisation of complexity, observed in the simplification of distributions in figure 6 section A, and maximisation of accuracy, observed in the improved performance of figure 5 section B.

The experimental results, presented in supplemental information figure S13, also show that when an incorrect representation of the virtual environment is provided to the hydrogel, the performance is severely inhibited. This behaviour can be

explained through the memory mechanics, as a result of ion interactions and migration. The hydrogel seeks to represent the information given via ion migration. If the information provided to the gel does not accurately represent the environment in which the gel's outputs are being applied, the ions will redistribute to represent this incorrect information. Thus the gel's sudo internal model of the environment, via FEP, is incorrectly updated, inhibiting active inference, and so the gel does not perform the task accurately.

## Conclusion

This study applied the theories of learning in BNNs through FEP to ionic EAP hydrogels, a medium whose behaviour is also driven by the theories of FEP. Firstly, by measuring ion concentrations through conductivity of the EAP hydrogel after periods of stimulation, the memory mechanics of the ionic EAP hydrogel were highlighted to provide a basis for further experimentation.

Secondly, utilising techniques in reservoir computing, the ionic EAP hydrogel is embodied in the simulated game world of Pong, through the use of a custom MEA. The game environment was encoded into stimulations provided to the hydrogel and recorded ion concentrations were used as motor commands within the game world. Through analysis of the behaviour of the EAP hydrogel within this simulated environment, improved performance was observed through the course of the game. Additionally, through analysis of how performance improvement manifested through behavioural changes, the mechanisms of the EAP hydrogels were linked back to FEP.

By investigating the chain of events leading from ion migration to game performance, emergent learning via theories of FEP present in the application of ionic EAP hydrogels was clarified. This was achieved by recording information from the various layers of abstraction, from electric field stimulation to rally lengths, and analysing the connections between these layers. The Pong game acts as a problem to be solved. When this problem is provided to the hydrogel as an electric field, it creates a gradient of increased free energy. Due to the memory mechanics and through FEP, the hydrogel updates a sudo internal model of the environment via ion migration, effectively learning how to play the game. The ions within the hydrogel seek to minimise the increased free energy caused by the electric stimulation and redistribute, the redistribution of ions then becomes the solution to the problem. The way in which the solution is interpreted allows it to be applied to the problem, in this case through conversion of localised ion concentration to paddle motion. As ions migrate to represent the game environment, presented as a series of changing electric fields, the memory in the hydrogel is filled resulting in an increase in current draw. This then causes a change in paddle motion to better intercept the ball, which in turn improves the hit/miss ratio in the regions, which in turn increases the rally lengths.

The results of this study demonstrate the importance of the part that FEP plays in theories of learning. BNNs themselves are only subject to FEP and active inference on a macro level, when the behaviour of a complete and embodied neural network is assessed. Furthermore, in reality BNNs are driven by many different interconnected electrical and chemical systems that rely on more than just free energy minimisation. Alternatively, EAP hydrogels are far more compatible with

pure free energy approximations, and offer a system whose behaviour can be explained via FEP at a fundamental level. The fact that a medium, whose fundamental mechanics relies on FEP, is able to exhibit emergent learning further implies that FEP may be more integral to emergent learning than previously thought.

**Future Work.** This study demonstrated that a form of emergent learning is possible in mediums other than BNNs, when the behaviour of the medium in question is also subject to the principles of free energy. The application demonstrated, however, does not achieve the same resolution of ability that can be achieved with BNNs. As future work, higher resolution MEAs can be tested to investigate the limits of this form of reservoir computing. Similarly, additional applications can be attempted to investigate how the learning behavior is influenced by task. Although FEP is found in BNNs, it only occurs in neurons as a network not individually. Comparatively, EAPs are subject to free energy on a more integrally at the physical chemistry level resulting in greater alignment with FEP theories. FEP is also found in many other forms of active matter and natural systems. Through exploration of these alternatives, a medium that is more capable, and just as accessible, could be found. All these avenues of research, however, fall outside the goals of this study but describe an interesting path for future work and application of this technology.

**Acknowledgments.** This project is fully funded by Process Vision Ltd (<https://www.processvision.com/>).

**Author Contributions.** Conceptualization, V.S. and Y.H.; Methodology, V.S. and Y.H.; Software, V.S.; Formal Analysis, V.S.; Investigation, V.S.; Writing – Original Draft, V.S., Y.H., and W.H.; Writing – Review & Editing, V.S., Y.H., and W.H.; Supervision, Y.H. and W.H.

**Declaration of Interests.** The authors declare no competing interests.

## References

- 1 H J Ruskin, R Walshe, (2006). Emergent computing-introduction to the special theme. *ERCIM News* **64**, 24–25.
- 2 S Bondar, JC Hsu, A Pfouga, J Stjepandić, (2017). Agile digitale transformation of enterprise architecture models in engineering collaboration. *Procedia Manuf.* **11**, 1343–1350.
- 3 C Fernando, S Sojakka, (2003). Pattern recognition in a bucket in *European conference on artificial life*. (Springer), pp. 588–597.
- 4 G Van der Sande, D Brunner, MC Soriano, (2017). Advances in photonic reservoir computing. *Nanophotonics* **6**, 561–576.
- 5 A Tero, et al., (2010). Rules for biologically inspired adaptive network design. *Science* **327**, 439–442.
- 6 M Dueñas-Diez, J Pérez-Mercader, (2019). How chemistry computes: language recognition by non-biochemical chemical automata. from finite automata to turing machines. *Iscience* **19**, 514–526.
- 7 A Wang, et al., (2016). Configurable nor gate arrays from belousov-zhabotinsky micro-droplets. *The Eur. Phys. J. Special Top.* **225**, 211–227.
- 8 J Misra, I Saha, (2010). Artificial neural networks in hardware: A survey of two decades of progress. *Neurocomputing* **74**, 239–255.
- 9 J Végh, (2019). How amdahl's law limits the performance of large artificial neural networks. *Brain informatics* **6**, 1–11.
- 10 K Mainzer, (2009). From embodied mind to embodied robotics: Humanities and system theoretical aspects. *J. Physiol.* **103**, 296–304.
- 11 A Cangelosi, J Bongard, MH Fischer, S Nolfi, (2015). Embodied intelligence in *Springer Handbook of Computational Intelligence*. (Springer), pp. 697–714.
- 12 VC Müller, M Hoffmann, (2017). What Is Morphological Computation? On How the Body Contributes to Cognition and Control. *Artif. Life* **23**, 1–24.
- 13 L Shapiro, S Spaulding, (2021). Embodied Cognition in *The Stanford Encyclopedia of Philosophy*, ed. EN Zalta. (Metaphysics Research Lab, Stanford University), Winter 2021 edition.
- 14 G Tanaka, et al., (2019). Recent advances in physical reservoir computing: A review. *Neural Networks* **115**, 100–123.
- 15 B Schrauwen, D Verstraeten, J Van Campenhout, (2007). An overview of reservoir computing: theory, applications and implementations in *Proceedings of the 15th european symposium on artificial neural networks*. p. 471–482 2007. pp. 471–482.
- 16 K Heiney, OH Ramstad, I Sandvig, A Sandvig, S Nichele, (2019). Assessment and manipulation of the computational capacity of in vitro neuronal networks through criticality in neuronal avalanches in *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*. (IEEE), pp. 247–254.
- 17 P Aaser, et al., (2017). Towards making a cyborg: A closed-loop reservoir-neuro system. *ECAL 2017, Fourteenth Eur. Conf. on Artif. Life* **14**, 430–437.
- 18 P Mateos-Aparicio, A Rodríguez-Moreno, (2019). The impact of studying brain plasticity. *Front. cellular neuroscience* **13**, 66.
- 19 ZC Chao, DJ Bakkum, SM Potter, (2008). Shaping embodied neural networks for adaptive goal-directed behavior. *PLoS computational biology* **4**, e1000042.
- 20 DJ Bakkum, ZC Chao, SM Potter, (2008). Spatio-temporal electrical stimuli shape behavior of an embodied cortical network in a goal-directed learning task. *J. neural engineering* **5**, 310.
- 21 TB DeMarse, KP Dockendorf, (2005). Adaptive flight control with living neuronal networks on microelectrode arrays in *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005*. (IEEE), Vol. 3, pp. 1548–1551.
- 22 BJ Kagan, et al., (2021). In vitro neurons learn and exhibit sentience when embodied in a simulated game-world.
- 23 K Friston, (2010). The free-energy principle: a unified brain theory? *Nat. reviews neuroscience* **11**, 127–138.
- 24 D Demekas, T Parr, KJ Friston, (2020). An investigation of the free energy principle for emotion recognition. *Front. Comput. Neurosci.* **14**, 30.
- 25 K Friston, (2009). The free-energy principle: a rough guide to the brain? *Trends cognitive sciences* **13**, 293–301.
- 26 T Parr, KJ Friston, (2019). Generalised free energy and active inference. *Biol. cybernetics* **113**, 495–513.
- 27 K Friston, J Kilner, L Harrison, (2006). A free energy principle for the brain. *J. physiology-Paris* **100**, 70–87.
- 28 WR Ashby, (1947). Principles of the self-organizing dynamic system. *The J. general psychology* **37**, 125–128.
- 29 K Friston, (2012). The history of the future of the bayesian brain. *NeuroImage* **62**, 1230–1233.
- 30 P Schwartenbeck, et al., (2015). Evidence for surprise minimization over value maximization in choice behavior. *Sci. reports* **5**, 1–14.
- 31 G Nicolis, I Prigogine, (1977). *Self-organization in Nonequilibrium Systems: From Dissipative Structures to Order Through Fluctuations*. (John Wiley, New York).
- 32 J Bruneberg, E Rietveld, (2014). Self-organization, free energy minimization, and optimal grip on a field of affordances. *Front. human neuroscience* **8**, 599.
- 33 JS Kelso, (1995). *Dynamic patterns: The self-organization of brain and behavior*. (MIT press).
- 34 V Pasquale, P Massobrio, L Bologna, M Chiappalone, S Martinoia, (2008). Self-organization and neuronal avalanches in networks of dissociated cortical neurons. *Neuroscience* **153**, 1354–1369.
- 35 K Friston, et al., (2013). The anatomy of choice: active inference and agency. *Front. human neuroscience* **7**, 598.
- 36 T Isomura, K Kotani, Y Jimbo, (2015). Cultured cortical neurons can perform blind source separation according to the free-energy principle. *PLoS computational biology* **11**, e1004643.
- 37 T Isomura, K Friston, (2018). In vitro neural networks minimise variational free energy. *Sci. reports* **8**, 1–14.
- 38 T Parr, KJ Friston, (2018). The discrete and continuous brain: from decisions to movement—and back again. *Neural computation* **30**, 2319–2347.
- 39 M Marchetti, et al., (2013). Hydrodynamics of soft active matter. *REVIEWS OF MODERN PHYSICS Rev Mod Phys* **85**, 1143.
- 40 M Otake, Y Kagami, M Inaba, H Inoue, (2002). Motion design of a starfish-shaped gel robot made of electro-active polymer gel. *Robotics Auton. Syst.* **40**, 185–191.
- 41 V Strong, W Holderbaum, Y Hayashi, (2022). Electroactive polymer gels as probabilistic reservoir automata for computation. *Iscience* **25**, 105558.
- 42 R Kürsten, V Sushkov, T Ihle, (2017). Giant kovacs-like memory effect for active particles. *Phys. Rev Lett* **119**, 188001.
- 43 T Tanaka, I Nishio, ST Sun, S Ueno-Nishio, (1982). Collapse of gels in an electric field. *Science* **218**, 467–469.
- 44 P Atkins, J De Paula, (2006). Atkins' physical chemistry, volume 8. ed.
- 45 VV Yashin, S Suzuki, R Yoshida, AC Balazs, (2012). Controlling the dynamic behavior of heterogeneous self-oscillating gels. *J. Mater. Chem.* **22**, 13625–13636.
- 46 O Kuksenok, VV Yashin, AC Balazs, (2007). Mechanically induced chemical oscillations and motion in responsive gels. *Soft Matter* **3**, 1138–1144.
- 47 D De Tommasi, G Puglisi, G Zurlo, (2014). Hysteresis in electroactive polymers. *Eur. J. Mech. A, Solid State Mech.* **48**, 16–22.
- 48 M Bassil, J Davenas, ME Tahchi, (2008). Electrochemical properties and actuation mechanisms of polyacrylamide hydrogel for artificial muscle application. *Sensors Actuators B* **134**, 496–501.
- 49 JR Gray, (2004). Conductivity analyzers and their application. *Environ. instrumentation analysis handbook* **1**, 491–510.
- 50 P Atkins, J De Paula, (2011). *Physical chemistry for the life sciences*. (Oxford University Press, USA).
- 51 P Atkins, PW Atkins, J De Paula, (2014). *Atkins' physical chemistry*. (Oxford university press).
- 52 T Tanaka, et al., (1980). Phase transitions in ionic gels. *Phys. Rev. Lett.* **45**, 1636.
- 53 M Oyen, (2014). Mechanical characterisation of hydrogel materials. *Int. Mater. Rev.* **59**, 44–59.
- 54 R Bourchouladze, (2002). *Memories are made of this: How memory works in humans and animals*. (Columbia University Press).
- 55 G Schiavone, et al., (2020). Guidelines to study and develop soft electrode systems for neural stimulation. *Neuron* **108**, 238–258.
- 56 K Jia, X Li, Y Wang, (2021). Electrochemical breakdown in hydrogel iontronic devices. *Soft Matter* **17**, 834–839.
- 57 Mathworks, *Polyfit*. <https://www.mathworks.com/help/matlab/ref/polyfit.html#bue6sxq-1-S>, Accessed 01/10/2021.
- 58 Y Mao, M Cates, H Lekkerkerker, (1995). Depletion force in colloidal systems. *Phys. A: Stat. Mech. Appl.* **221**, 111–125.

- 1187 59 *Mech. its Appl.* **222**, 10–24.  
1188 59 HN Lekkerkerker, R Tuinier, (2011). Depletion interaction in *Colloids and the depletion interaction*.  
1189 (Springer), pp. 57–108.  
1190 60 RM Neal, (2012). *Bayesian learning for neural networks*. (Springer Science & Business Media)  
1191 Vol. 118.  
1192 61 DC Knill, A Pouget, (2004). The bayesian brain: the role of uncertainty in neural coding and  
1193 computation. *TRENDS Neurosci.* **27**, 712–719.  
1194 62 DV Schroeder, (2000). *An introduction to thermal physic*. (San Francisco, CA: Addison Wesley)  
1195 Vol. 57.  
1196 63 N Ebrahimi, E Maasoumi, ES Soofi, (1999). Ordering univariate distributions by entropy and  
1197 variance. *J. Econom.* **90**, 317–336.  
1198 64 K Friston, G Buzsáki, (2016). The functional anatomy of time: what and when in the brain. *Trends  
1199 cognitive sciences* **20**, 500–511.  
1200 65 K Friston, T FitzGerald, F Rigoli, P Schwartenbeck, G Pezzulo, (2017). Active inference: a  
1201 process theory. *Neural computation* **29**, 1–49.  
1202 66 MJ Beal, (2003). *Variational algorithms for approximate Bayesian inference*. (University of London,  
1203 University College London (United Kingdom)).