

OPEN

# 3D-Printing and Machine Learning Control of Soft Ionic Polymer-Metal Composite Actuators

James D. Carrico<sup>1</sup>, Tucker Hermans<sup>2</sup>, Kwang J. Kim<sup>3</sup> & Kam K. Leang<sup>4\*</sup>

This paper presents a new manufacturing and control paradigm for developing soft ionic polymer-metal composite (IPMC) actuators for soft robotics applications. First, an additive manufacturing method that exploits the fused-filament (3D printing) process is described to overcome challenges with existing methods of creating custom-shaped IPMC actuators. By working with ionomeric precursor material, the 3D-printing process enables the creation of 3D monolithic IPMC devices where ultimately integrated sensors and actuators can be achieved. Second, Bayesian optimization is used as a learning-based control approach to help mitigate complex time-varying dynamic effects in 3D-printed actuators. This approach overcomes the challenges with existing methods where complex models or continuous sensor feedback are needed. The manufacturing and control paradigm is applied to create and control the behavior of example actuators, and subsequently the actuator components are combined to create an example modular reconfigurable IPMC soft crawling robot to demonstrate feasibility. Two hypotheses related to the effectiveness of the machine-learning process are tested. Results show enhancement of actuator performance through machine learning, and the proof-of-concepts can be leveraged for continued advancement of more complex IPMC devices. Emerging challenges are also highlighted.

Ionic polymer metal composites (IPMC)s are electroactive-polymer soft actuators with application in biomedical devices and soft robotics. IPMCs consist of an ion-exchange membrane (also referred to as a polyelectrolyte) and electrode layers on opposing sides of the ion-exchange membrane. When hydrated and a voltage is applied across its electrodes, an IPMC deforms. Conversely, a hydrated IPMC device when deformed produces a measurable voltage and functions as a sensor<sup>1,2</sup>. The advantages of IPMCs include low activation voltage (<3 V), flexibility, and that they operate in aqueous environments<sup>1</sup>. Additionally, they are functional down to sub-micron scales<sup>3</sup>. Because of these advantages, IPMCs are attractive for applications such as active catheters<sup>4</sup>, manipulators<sup>3,5-7</sup>, grippers<sup>8</sup>, microfluidic valves and pumps<sup>9-11</sup>, and propulsion and sensing mechanisms in mobile robots<sup>12-20</sup>. However, the main challenges of working with IPMCs are the lack of novel manufacturing methods for creating complex monolithic actuator and sensor designs and limited approaches for effective actuator motion control that do not rely on complex modeling and analysis. Therefore, the main contribution of this paper is a paradigm whereby custom-shaped IPMC actuators can be manufactured as monolithic devices through 3D printing and effective motion control can be achieved through machine learning that avoids the need for modeling the complex behaviors. Furthermore, the machine learning approach is also adaptable to time-varying degradation of the material through repeated use.

The polymer matrix of the ion exchange material in IPMCs contains hydrophilic networks that form due to ionic clustering. Consequently, the hydrated ion exchange material conducts charge via the soluble counterions that neutralize the material. When a voltage is applied across the electrodes of an IPMC, charge redistributes affecting the hydrophilic regions of the polymer, causing change in volume which leads to the material deforming (e.g., macroscopic bending in cantilever-type actuators)<sup>1,2</sup>. Conversely, deforming the hydrated composite produces a voltage and thus the material acts as a sensor. IPMCs are most commonly fabricated as composites of a Nafion membrane with platinum metal electrodes, but a variety of other materials can be used to manufacture

<sup>1</sup>University of Mary, School of Engineering, Bismarck, ND, 58504, USA. <sup>2</sup>University of Utah, School of Computing, Utah Learning Lab for Manipulation Autonomy, University of Utah Robotics Center, Salt Lake City, UT, 84112, USA. <sup>3</sup>University of Nevada, Las Vegas, Department of Mechanical Engineering, Active Materials and Smart Living (AMSL) Laboratory, Las Vegas, NV, 89154, USA. <sup>4</sup>University of Utah, Department of Mechanical Engineering, Design Automation Robotics and Control (DARC) Lab, University of Utah Robotics Center, Salt Lake City, 84112, USA. \*email: [kam.k.leang@utah.edu](mailto:kam.k.leang@utah.edu)

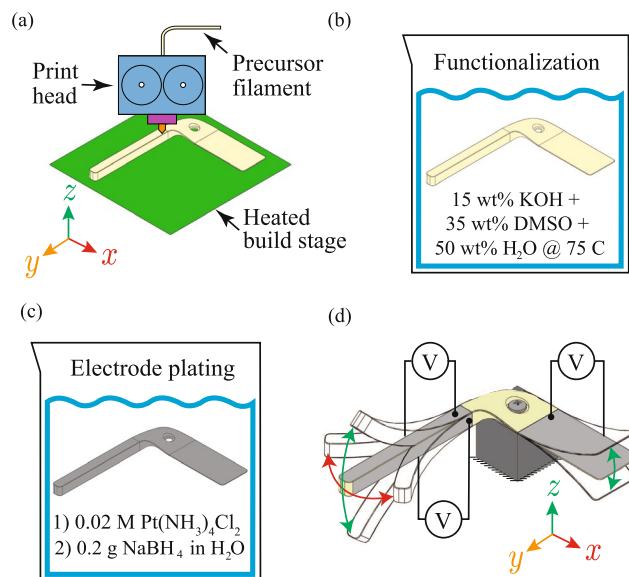
IPMCs with tuned performance characteristics or reduced cost. For example, IPMCs have been successfully manufactured from block ionomers and blends using polystyrene sulfonic acid (PSSA)<sup>21,22</sup>. Other membrane materials include blends of polyvinylidene fluoride (PVDF), hexafluoropropylene (HFP), polyethylene oxide (PEO), and nitrile butadiene rubber (NBR)<sup>23</sup>. Researchers have selected membrane materials to have favorable water uptake and ion exchange capacities as well as high ion conductivity<sup>22</sup>. Researchers have also used additives with Nafion to improve the membrane quality. For example, these approaches include blending amphiphatic organic molecules with Nafion to improve the membrane's ion conductivity and adding hygroscopic alumina layers to reduce water loss<sup>24</sup>. Notably, an actuator was recently fabricated with a 300 fold improved ion conductivity and a millisecond response time through the use of a single-ion-conducting polymer as the membrane material<sup>25</sup>. Researchers have employed carbon nanotubes and graphene paper as the electrode layer to create highly flexible actuators that are durable and have low solvent evaporation<sup>26</sup>. Other conductive materials that have been used in the electrode layers include polypyrrole (PPy) or Poly(3,4-ethylenedioxythiophene) (PEDOT)<sup>23</sup>.

Fabricating IPMCs typically involves shaping and plating commercially-available sheets and tubular structures of Nafion, Flemion, or Aquivion material. Unfortunately, this is a laborious and unreliable means of fabricating IPMC-based actuators and devices, and thus significantly hinders the use of IPMCs in practical applications. To overcome this challenge, dispersions of ionomeric material or molten precursor to ionomeric material are used in a mold<sup>6,27</sup>. However, these approaches are restrictive in that new designs require new molds, which are time consuming and costly to create. In comparison to these methods, free-form-based techniques where structures are created using a layer-by-layer manufacturing process has been proposed<sup>28,29</sup>, where layers of Nafion are dispensed into silicone casts and the solvent in the dispersion is allowed to evaporate. This process creates both the electrode layer and the ionomeric substrate. Unfortunately, a plasticizer is required to minimize brittleness which increases drying time. The IPMCs that have been created had lower blocking force compared to traditionally-made IPMCs<sup>28</sup>. Alternatively, it has been shown that fused deposition modeling (3D printing) can be used to deposit a precursor to an ionomeric material as a monolithic body which can subsequently be functionalized and plated, creating IPMCs of any desired shape<sup>20,30</sup>. However, with increasingly complex actuator designs, IPMCs become increasingly difficult to control for practical applications.

There are a number of challenges to controlling IPMCs. Specifically, two identical IPMC-based actuators, fabricated through the same process will exhibit appreciably different behaviors<sup>31</sup>. Also, the behavior of a single IPMC-based actuator will vary over time<sup>31</sup> due to solvent evaporation, aging, and degradation from repeated use. IPMCs also exhibit nonlinearity and other complex dynamic behavior such as back-relaxation and higher order resonances<sup>32</sup>. Additionally, sophisticated, custom-shaped, multi-input-multi-output (MIMO), IPMC-based systems (as may be fabricated through 3D printing) will exhibit coupled nonlinear behavior that may be difficult to model and control. To address these challenges, prior works have developed sophisticated control-oriented dynamics models and advanced feedback-control methods to compensate for time-varying and complex dynamic behavior. However, existing models are limited in their applicability to custom-shaped IPMCs. Moreover, unresolved challenges to integrating IPMCs with sensors inhibits the use of feedback control<sup>33,34</sup>. A possible solution to these challenges is the use of feedforward learning-based control methods such as Bayesian optimization. These methods are especially applicable to devices that operate repetitively, where the processes of iteration can be exploited to evaluate and adjust control inputs that optimizes a relevant performance metric. Recent work on Bayesian optimization has shown it to succeed in spite of complex, idiosyncratic, and time-varying behavior<sup>35,36</sup>. These features, together with the ability of Bayesian optimization to incorporate prior knowledge to speed up convergence, make it an attractive candidate for use on more complex 3D-printed IPMC actuators and devices.

Currently, there are other closely-related soft actuators with similar challenges to IPMCs. For instance, hydrogels are a stimuli responsive material that have the potential for use in a variety of biomedical and robotic applications<sup>37</sup>. When these gels are realized as polyelectrolytes and immersed in an electrolytic environment, they can be controlled by electric fields<sup>38</sup>. The electric field causes the counter-ions to migrate, leading to a concentration gradient that causes bending of the actuator<sup>38</sup>. In order to fabricate custom-shaped hydrogel actuators, a variety of 3D-printing techniques have been developed including extrusion-based, sterolithographic, and jetting-based methods<sup>37</sup>. Control of hydrogels is generally through thermal or hygroscopic stimuli, while polyelectrolyte gels are also stimulated by electric fields<sup>38</sup>. Lacking electrodes, polyelectrolyte gels are not individually controllable and must be immersed in an electrolytic environment in order to respond to electric stimuli. The response time of hydrogels is also significantly slower than other ionic electroactive polymers<sup>37</sup>. Another closely-related material to IPMCs are conjugated polymer actuators. Conjugated polymers, such as polypyrrole (PPy) and polyaniline (PANI), which have alternating single and double bonds, are organic semiconductors. When a sufficiently-large voltage is applied to a conjugated polymer, electrons are removed. If the material is in contact with an electrolyte, the material will then change size in proportion to its change in oxidation level<sup>39</sup>. Conjugated polymer actuators are generally layered bending actuators like IPMCs, involving a composite of a conjugated polymer with at least one noble metal electrode layer<sup>40</sup>. Addition of polyelectrolyte gels or electrolyte infused membranes to such composites was developed to allow the actuator to operate outside of an electrolytic environment<sup>41</sup>. Well established micro-machining techniques have been developed for fabricating conjugated polymer actuators<sup>40</sup>, but not a 3D printing process. 3D printing has been developed for some conjugated polymers, but not for the composite necessary for fabricating an actuator<sup>42</sup>. Like IPMCs, the behavior of polyelectrolyte gels and conjugated polymer actuators is complex and time varying, and there are significant challenges to integrating sensors<sup>43-46</sup>. The machine learning approach proposed here is equally relevant to these materials to help improve performance.

To address the challenges to manufacturing and control of IPMCs and related materials, this paper considers 3D printing for manufacturing and machine learning to control example IPMC actuator components, which together can be assembled to create a modular reconfigurable soft crawling robot platform to demonstrate feasibility. The development of the platform is presented to illustrate the proof-of-concept of 3D-printing of IPMC actuators and application of machine learning for effective motion control. Since the use of Bayesian optimization



**Figure 1.** Fused-filament 3D printing process for IPMCs: (a) printing of soft structure using precursor of ionomeric material, (b) functionalization of printed precursor structure, (c) plating process, and (d) segmenting and wiring of electrodes for multi-DOF actuation and/or sensing.

on IPMCs is new, this paper tests two hypotheses: (1) that Bayesian optimization will lead to convergence in fewer trials than a finite-difference policy gradient method making it more suitable for controlling IPMCs; and (2) that prior knowledge from a dynamics model (especially a known achievable target value) will lead to convergence in fewer number of trials than simply optimizing from a uniform prior distribution.

To investigate these hypotheses, first, the linear dynamics of the 3D-printed soft crawling platform are modeled and then validated through simulation and in physical experiments. The dynamics model of the robot is then used to contrast the performance of Bayesian optimization and finite-difference policy gradient method in simulation, where convergence is achieved in fewer trials using Bayesian optimization<sup>35,47,48</sup>. Since these simulations show the superiority of Bayesian optimization, the process is then applied to the real IPMC-based system. Experiments are presented to compare the performance of Bayesian optimization from a uniform prior with the performance of Bayesian optimization using a prior distribution obtained from optimizing in simulation on the dynamics model.

The contribution of this work is 3D printing of IPMCs and machine-learning control of the 3D-printed IPMCs and related actuators. The 3D-printing process, design and modeling of the actuator components, and Bayesian optimization as a learning-based control method are discussed in detail. Additionally, this paper compares the effectiveness of a finite-difference policy gradient method and Bayesian optimization as control methods for IPMCs. The results of the comparison highlight the advantage of Bayesian optimization's ability to encode prior knowledge in the form of known reachable target values and simulation results from a linear dynamics model.

## Methods

**3D Printing of IPMCs.** The 3D-printing process used to fabricate custom-shaped IPMCs involves four major steps as illustrated in Fig. 1. First, the fused-filament fabrication process requires manufacturing filament using Nafion's (or Aquivion's<sup>49</sup>) sulfonyl fluoride precursor. Next, a custom-designed 3D printer is required to process the precursor material for 3D printing. Since the precursor is not electroactive immediately after being 3D printed, an *in situ* functionalization process is then employed to convert the precursor material to fully ionomeric Nafion. Subsequently, the ion-exchange properties of the ionomeric material is leveraged in an electroless plating process to create platinum electrodes on the ionomeric material's surface.

**Filament extrusion and 3d printing of ionomeric precursor.** Since commercially-available ionomeric materials, such as Nafion (typically, in the sulfonic acid form  $-\text{SO}_3^-\text{H}^+$ ), are not melt processable<sup>50,51</sup>, the sulfonyl fluoride precursor of Nafion ( $-\text{SO}_2\text{F}$ ), sometimes referred to as XR resin, is employed. Nafion precursor filament is created using a filament extruder<sup>52</sup>. For example, XR resin is fed into a hopper system, where a drive mechanism, such as an auger or a piston, advances the precursor material for extrusion through a nozzle to form the filament. Cooling and a drawing mechanism may also be used to maintain dimensional consistency. Typical extrusion parameters include extrusion temperature range between 220 to 300 °C, and extrusion and drawing rates between 10 to 125 mm/s, depending on desired filament diameters.

There are several challenges to 3D printing the Nafion precursor material including Nafion's high glass-transition temperature, poor adhesion, material compressibility, and the tendency for the material to buckle in the drive gears of conventional 3D printers. These issues make it challenging to control the rate at which the material is extruded from the nozzle. Commercial 3D print heads cannot effectively handle the Nafion precursor

material without modifications, thus a custom-designed system is often required. Such custom print heads can be designed and integrated into commercial motion control platforms. Additionally, the build stage must be able to heat to approximately 200 °C to enhance adhesion characteristics during printing. A constrained filament path is often needed to prevent filament buckling. A thermal barrier, which confines the transition and melt zones to the heater block and heated nozzle, prevents the filament from progressively softening as it approaches the heater block. The components of the heater block can reach temperatures exceeding 300 °C to melt the filament material. The melted filament collects in the nozzle and is extruded out onto a high temperature print bed to ensure that the printed material adheres<sup>52</sup>. More recently, it was determined that commercially-available print heads designed for soft polymer materials can be adapted to print Nafion precursor by incorporating a heater capable of reaching approximately 300 °C. Figure 2(a,b,d) show examples of 3D-printed Nafion precursor devices.

**Functionalization process.** After printing, the component has to be made electroactive through a functionalization process. This involves converting the non-ionic Nafion precursor into fully ionic Nafion, which has sulfonate terminal groups that give Nafion its characteristic ion conducting ability. This process is done by base hydrolysis using a solution of 15 wt% KOH, 35 wt% DMSO, and 50 wt% deionized water at 75 °C as prescribed by the manufacturer<sup>53</sup>. In this process, Fluorine is exchanged nucleophilically with a hydroxyl group, and the hydrogen ion is then exchanged electrophilically for a potassium ion. A hydration sphere then forms around the potassium ion which swells the material and permits the ingress of the functionalizing solution. At 75 °C, the process proceeds at a rate of approximately 1.3 μm per minute<sup>54</sup>. The printed components are left in the functionalization solution for additional time (approximately 4 hours) after the estimated completion of the hydrolysis process, to facilitate the development of ionic clusters<sup>54</sup>. The material is then rinsed in two successive baths of deionized water at 75 °C for 30 minutes each time. To confirm complete hydrolysis (*i.e.*, functionalization), the material can be stained using methylene blue, which only dyes the “activated” material<sup>54</sup>.

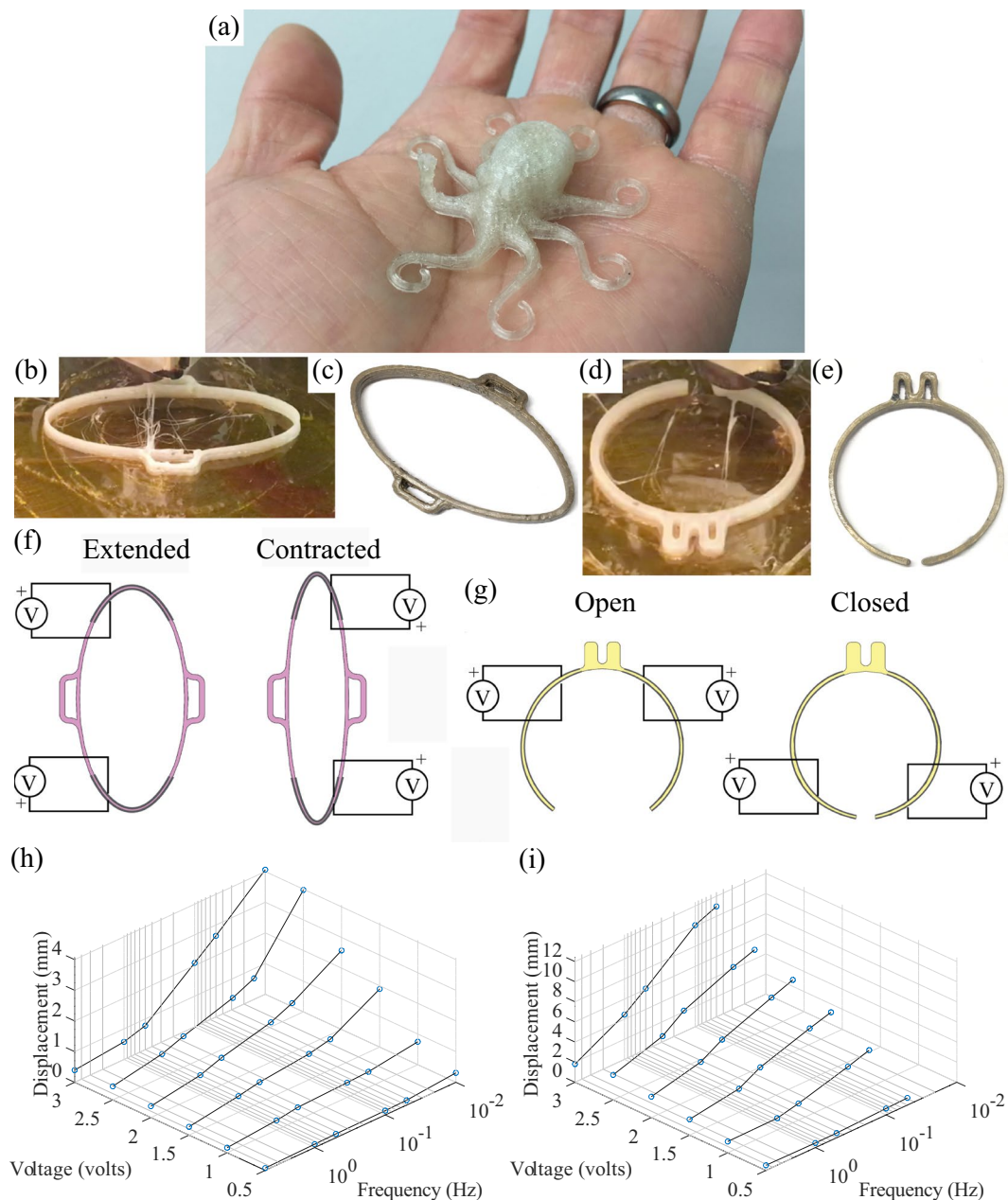
**Electrode plating process.** The ion-exchange properties of the now ionic material can be leveraged in an electroless plating process. Following functionalization, the material is converted to its acid form using 20% nitric acid (sulfuric acid can also be used). This exchanges the potassium ions neutralizing the material with hydrogen ions and is done as an intermediate step whenever attempting to change the neutralizing ionic species. The material is then rinsed again in deionized water, following which it is immersed in a 0.2 molar solution of tetraamineplatinum (II) chloridemonohydrate ( $[\text{Pt}(\text{NH}_3)_4] \cdot \text{Cl}_2 \times \text{H}_2\text{O}$ , Alfa Aesar) for a period of 2 to 4 hours. This deposits platinum ions in the surface layer of the Nafion. The platinum ions are then converted to elemental platinum by placing them in deionized water and repeatedly introducing a reducing agent such as sodium borohydride ( $\text{NaBH}_4$ , Aldrich) into the solution over the course of 3 hours. Repetition of this process two or three times will significantly improve the responsiveness of the resulting IPMC. Subsequently, either electrical or electroless plating methods can be used to further develop these electrodes and decrease their surface resistance. The results of plating are shown in Fig. 2(c) and (e) for the printed robot body and leg components, respectively.

**Fabrication and characterization of 3d-printed IPMC actuators.** The IPMC 3D-printing process was used to fabricate a variety of custom-shaped IPMC actuators as shown in Fig. 2. These include an octopus printed in Nafion precursor material, as shown in Fig. 2(a), as well as an IPMC-based linear actuator, as shown in Fig. 2(b,c), and an IPMC-based gripper, as shown in Fig. 2(d,e). The linear actuator and gripper were first 3D-printed using the fused-filament technique as described above. This consisted of first printing the desired geometry using the precursor filament material, then applying chemical activation, followed by plating and conditioning the components, and finally segmenting the plated electrodes on the surfaces to create independent electrode pairs. As shown in Fig. 2(b,d), the printed product has a yellowish discoloration. The material is soft and pliable and has a Teflon-like texture. After the subsequent chemical processes, the resulting ionic material beneath the plated surface is clear and though still distensible, is stiffer than the original printed product. As can be seen in Fig. 2(c,e), the plated linear actuator and gripper have a dull gray coloration after the two step plating process. This indicates a high conductivity across the plated surface. The linear actuator is elliptically-shaped with its interior and exterior electrodes segmented into four distinct regions, as shown in Fig. 2(f). The linear actuator extends with an applied voltage such that the exterior is the anode and the interior is the cathode. The linear actuator contracts with applied voltage such that the exterior is the cathode and the interior is the anode. The linear actuator has pockets for coupling with other devices such as the 3D-printed gripper. As shown in Fig. 2(g), the gripper is nominally a circular structure designed to fit around a tube. The gripper has two independent electrode pairs on the inside and outside of the individual arms, as shown in Fig. 2(h). When a positive voltage is applied such that the exterior electrodes are the anode side and the interior electrodes are the cathode side, the gripper opens, as shown in Fig. 3(d). When the voltage is applied to these electrode pairs in opposite polarity, the gripper closes, as shown in Fig. 2(h). Each gripper has a protrusion that extends off of it which engages in an interference fit with other devices, such as the 3D-printed linear actuators.

As shown in Fig. 2(h,i), periodic voltage signals were applied to the linear actuator and gripper over a range of amplitudes and frequencies to characterize the actuation performance. The linear actuator reaches its maximum displacement in response to 3 volt signals applied at 0.01 Hz, as can be seen in Fig. 2(h). The gripper reaches its maximum displacement in response to 3 volt signals applied at 0.05 Hz, as can be seen in Fig. 2(i). The actuation performance was deemed appropriate for developing a soft robotic testbed to demonstrate feasibility, as described below.

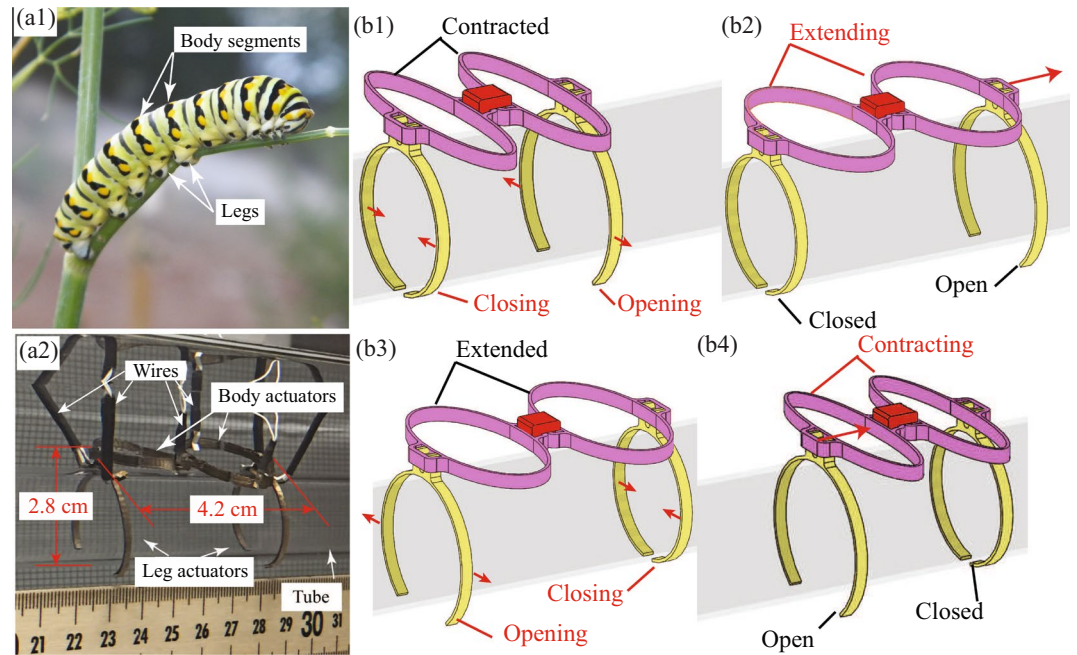
**Example modular reconfigurable soft crawling platform.** There has been prior work on the additive manufacturing of soft actuators for soft robotic platforms, but the approaches required assembly with actuators and sensors after fabricating the soft robot body<sup>55,56</sup>. Notably, researchers have very successfully constructed soft





**Figure 2.** Example 3D-printed Nafion precursor and functionalized and plated devices: **(a)** monolithic Nafion precursor octopus-shaped device prior to functionalization and electrode application; **(b,c)** linear actuator component for modular robotic device, printed, functionalized, and plated examples; **(d,e)** gripper mechanism for modular robotic device, printed, functionalized, and plated examples; **(f)** 3D-printed linear actuator extension and contraction as a function of applied voltage; **(g)** 3D-printed gripper actuator opening and closing as a function of applied voltage; **(h)** peak displacement response of linear actuator for a range of amplitudes and frequencies of the input signal, **(i)** peak displacement response of gripper for a range of amplitudes and frequencies of the input signal.

robots based on pneumatic and fluid elastomeric actuators<sup>57–62</sup>. Recently, multi-material additive manufacturing methods have even been successfully employed to create soft robotic devices<sup>63,64</sup>. These devices utilize pressure gradients provided by hosing, internal chemical reactions or electric motors for actuation. Consequently, the scale of the robot is limited by the size of the mechanisms required for actuation. Also, recent research has developed “4D printing” and origami-based robots, that are able to reconfigure into a functional shape after being fabricated (often as a flat sheet). These are very promising, but shape change is generally accomplished *via* heating or other nonelectric stimuli<sup>65–68</sup>. In general, the development of functional soft electroactive polymer based robotic platforms is limited. Thus, an important complimentary approach to these other techniques is the proposed 3D-printing process to create actuators for realizing soft robotic systems. As an example application to demonstrate proof-of-concept, the linear actuators and grippers shown in Fig. 2(b–e) are used to create a modular



**Figure 3.** Example 3D-printed IPMC soft crawling robot inspired by a caterpillar: **(a1)** biological caterpillar showing body and leg segments that were modeled in the design of the crawling robot; **(a2)** assembled robot with two legs and two body sections, with electrical connections; **(b1)** back leg actuator closing and front leg actuator opening while body is contracted; **(b2)** extension of body actuators while back leg actuator is closed and front leg actuator is open propelling the robot forward; **(b3)** back leg actuator opening and front leg actuator closing while body is extended; **(b4)** contraction of body actuators while back leg actuator is open and front leg actuator is closed drawing up the back of the robot.

reconfigurable soft crawling robot. The 3D-printed soft IPMC robot is inspired by a caterpillar or worm-like organism, as shown in Fig. 3(a1). The robot platform consists of modular grippers that act as its legs and linear actuators that act as its body segments. The robot is designed to “grip” onto a cylindrical tube and “inch” its way along the tube, as shown in Fig. 3(b1–b4). The modular components can be assembled into various configurations to create a robot of varying lengths. A leg-body segment of a crawling robot is assembled from a gripper and linear actuator. These segments can be chained together to create longer sections of the robot. The robot can be assembled into multiple configurations to test which results in faster locomotion and how the different configurations effect gait performance. An example of the locomotion of a simple leg-body-body-leg robot is illustrated in Fig. 3(b1–b4). In this configuration, the nominal gait begins with the linear actuators contracted while the rear gripper closes and front gripper opens, as shown in Fig. 3(b1). Then, the linear actuators extends propelling the front end of the robot forward, as shown in Fig. 3(b2). Next, the front gripper closes and the rear gripper opens, as shown in Fig. 3(b3). Finally, the linear actuators contract, as shown in Fig. 3(b4), bring the rear end of the robot towards the stationary front end. The process is repeated enabling the robot to move forward.

**Dynamics modeling and parameter estimation.** The actuation behavior of the 3D-printed IPMC actuators and gripper mechanisms described above are modeled and the model is used to predict the performance of the overall IPMC-based robotic testbed. Additionally, the model is exploited by the machine learning algorithm for gait optimization.

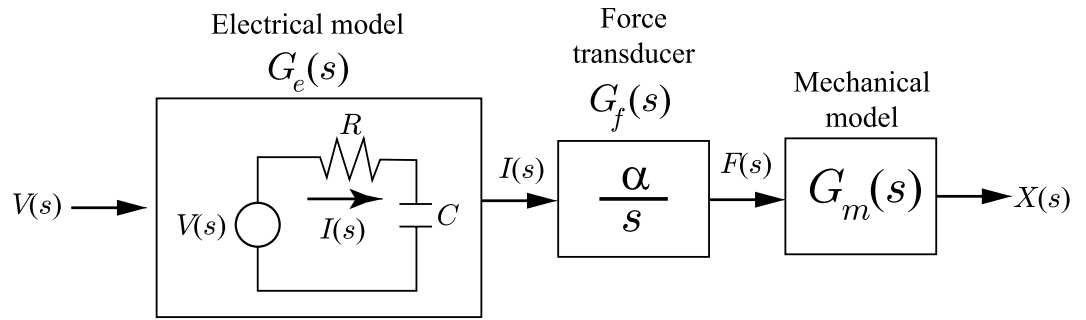
**Electromechanical model.** The electromechanical dynamics of an IPMC actuator consists of an electrical model, force transducer, and mechanical model as illustrated in Fig. 4. The electrical model is an equivalent circuit that relates the applied voltage  $V(s)$  (where ‘ $s$ ’ is the Laplace variable) to the current  $I(s)$ . The current  $I(s)$  then becomes the input to the force transducer  $G_f(s)$ , modeled by an integral term, where the output is the mechanical force  $F(s)$ . Finally, the force drives a lumped-parameter model of the structural dynamics of the actuator with output displacement  $X(s)$ .

The transfer function that relates the applied voltage  $V(s)$  to the output current  $I(s)$  is given by

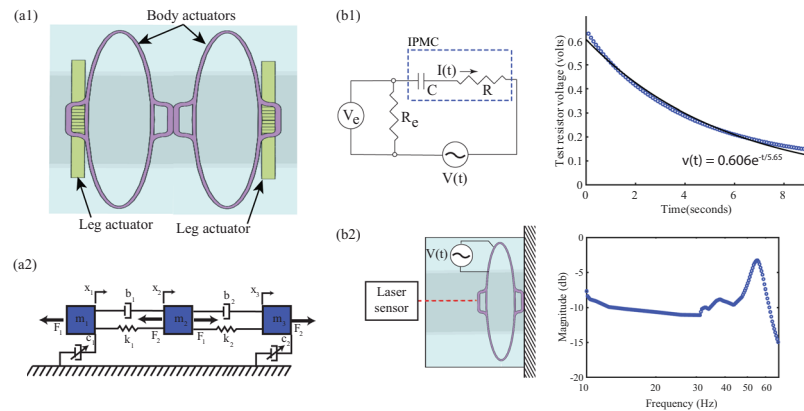
$$G_e(s) = \frac{I(s)}{V(s)} = \frac{\frac{1}{R}s}{s + \frac{1}{RC}}. \quad (1)$$

In this model,  $R$  is the effective electrical resistance and  $C$  is the effective capacitance of the actuator.

Assuming that the effective IPMC force  $F(s)$  is linearly related to the density of the transferred charges<sup>3</sup>, then



**Figure 4.** An IPMC model: The applied voltage  $V(s)$  is the input to an equivalent electrical circuit model, where current  $I(s)$  is the output. A force transducer model (integral term) relates current  $I(s)$  to force  $F(s)$ . Finally, the force is applied to a mechanical model where the output is the displacement of the IPMC mechanism.



**Figure 5.** Modeling and parameter estimation of crawling robot: **(a1)** two body actuators and two leg actuators configured as a crawling robot; **(a2)** model of crawling robot as a chain of lumped masses, springs and dampers where  $m_1, m_2,$  and  $m_3$  are effective masses,  $k_1$  and  $k_2$  are springs,  $b_1$  and  $b_2$  are dampers, and  $c_1$  and  $c_2$  are damping elements that approximate the viscous friction between the legs and the tube; **(b1)** equivalent RC-circuit model for the IPMC actuator and the measured step response of the body-segment behavior compared to the model  $V(t) = 0.6e^{-t/5.65}$  (R-squared value of 0.9911), where  $R_e = 220\Omega$ . **(b2)** schematic of the experiment for characterizing mechanical model parameters and the measured experimental frequency response of the actuator.

$$G_f(s) = \frac{F(s)}{I(s)} = \frac{\alpha}{s}, \tag{2}$$

where  $\alpha$  is a constant. It is noted that this model is only valid over operating frequencies where back-relaxation effect, which is characteristic of IPMCs, is not significant<sup>69</sup>.

The robot's structural dynamics,  $G_m(s)$ , is modeled by inertial, elastic, and damping elements, as shown in Fig. 5(a2). The variable damping elements,  $c_1, \dots, c_n$ , shown in Fig. 5(a2), model the frictional contact behavior between the legs and the tube. Balancing forces on the masses in Fig. 5(b2) yields the following relationships:

$$m_1\ddot{x}_1 = b_1(\dot{x}_2 - \dot{x}_1) - c_1\dot{x}_1 + k_1(x_2 - x_1) - F_1, \tag{3}$$

$$m_i\ddot{x}_i = b_{i-1}(\dot{x}_{i-1} - \dot{x}_i) + b_i(\dot{x}_{i+1} - \dot{x}_i) - c_i\dot{x}_i + k_{i-1}(x_{i-1} - x_i) + k_i(x_{i+1} - x_i) + F_{i-1} - F_i, \tag{4}$$

$$m_n\ddot{x}_n = b_{n-1}(\dot{x}_n - \dot{x}_{n-1}) - c_n\dot{x}_n + k_{n-1}(x_n - x_{n-1}) + F_{n-1}, \tag{5}$$

where  $x_1, \dots, x_n$  are the displacements of the effective masses  $m_1, \dots, m_n$  respectively;  $b_1, \dots, b_{n-1}$  and  $k_1, \dots, k_{n-1}$  are the constants for the damping and elastic elements, respectively; and  $F_1, \dots, F_{n-1}$  are the forces applied to each inertial element.

The equations of motion can be written in the following matrix form:

	R (Ohms)	C (mF)	$\tau$ (sec)	$\omega_c$ (rad-sec <sup>-1</sup> )	$m_1$ (g)	$m_2$ (g)	$b$ (Dyn-s-cm <sup>-1</sup> )	$c$ (Dyn-s-cm <sup>-1</sup> )	$k$ (N-cm <sup>-1</sup> )
Body-actuator one	140	16	2.2	0.45	1.4	1.0	59	56	1.30
Body-actuator two	260	24	6.2	0.16	1.2	0.7	63	57	0.65
Leg-actuator one	100	55	5.6	0.17	N/M	N/M	N/M	N/M	N/M
Leg-actuator two	170	56	9.5	0.11	N/M	N/M	N/M	N/M	N/M
Leg-actuator three	210	52	10.9	0.09	N/M	N/M	N/M	N/M	N/M

**Table 1.** Values for parameters of the electromechanical model, where “N/M” denotes not measured.

$$\begin{bmatrix} \dot{\vec{p}} \\ \dot{\vec{q}} \end{bmatrix} = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{0} \end{bmatrix} \begin{bmatrix} \vec{p} \\ \vec{q} \end{bmatrix} + \begin{bmatrix} \tilde{B}_1 \\ \tilde{0} \end{bmatrix} \vec{F}, \quad (6)$$

where  $\vec{p}$ ,  $\vec{q}$ , and  $\vec{F}$  are vectors of the momentums, displacements, and forces, respectively. Matrices  $\tilde{A}_{11}$ ,  $\tilde{A}_{12}$  and  $\tilde{A}_{21}$  have dimension  $n \times n$  and  $\tilde{B}_1$  has dimension  $n \times (n - 1)$ . The vectors  $\vec{p}$  and  $\vec{q}$  have dimension  $n \times 1$  and  $\vec{F}$  is dimension  $(n - 1) \times 1$ . For an example robot with two bodies and three legs,  $\tilde{A}_{11}$ ,  $\tilde{A}_{12}$ ,  $\tilde{A}_{21}$  and  $\tilde{B}_1$  are given by

$$\tilde{A}_{11} = \begin{bmatrix} -c_1 - b_1 & b_1 & 0 \\ m_1 & m_2 & \\ b_1 & -c_2 - b_1 - b_2 & b_2 \\ m_1 & m_2 & m_3 \\ 0 & b_2 & -c_3 - b_2 \\ & m_2 & m_3 \end{bmatrix},$$

$$\tilde{A}_{12} = \begin{bmatrix} -k_1 & k_1 & 0 \\ k_1 & -k_1 - k_2 & k_2 \\ 0 & k_2 & -k_2 \end{bmatrix},$$

$$\tilde{A}_{21} = \begin{bmatrix} \frac{1}{m_1} & 0 & 0 \\ 0 & \frac{1}{m_2} & 0 \\ 0 & 0 & \frac{1}{m_3} \end{bmatrix}, \text{ and } \tilde{B}_1 = \begin{bmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}.$$

The output of the model given by Eq. (6) can be compared to experimental results to validate the robot's gait performance. Furthermore, the model can be used to help design a controller to enhance the robot's gait.

*Identifying the model parameters.* From Fig. 5(b1), the transfer function for circuit model relating the applied voltage  $V(s)$  to the voltage  $V_e(s)$  across the resistor  $R_e$  is given by

$$\frac{V_e(s)}{V(s)} = \left( \frac{R_e}{R + R_e} \right) \left( \frac{s}{s + \frac{1}{\tau}} \right), \quad (7)$$

where  $\tau = (R + R_e)C$ . Then for a step input, the time response is

$$v_e(t) = \left( \frac{R_e}{R + R_e} \right) e^{-\frac{t}{\tau}}. \quad (8)$$

To find the model parameters, a step input is applied to the IPMC and the measured voltage across a series resistor with the IPMC is used to determine the parameters of the electrical model through curve-fitting. Table 1 lists the experimentally-determined parameters of the electrical model, where “n/m” denotes “not measured”.

For the mechanical model, the effective spring constant is determined by the ratio between an applied known load and the measured actuator deflection. The equivalent mass and damping constant are determined by analyzing the measured frequency response (obtained with an HP 35665 A digital signal analyzer) of the IPMC actuator, as shown in Fig. 5(b2). The boundary conditions for these experiments are shown in Fig. 5(b2), where one side of a body or joined body and leg are fixed and the displacement of the free end is measured by a laser displacement sensor (Keyence LK-2001). For body sections, experiments were conducted both with an attached gripper (leg) and without a gripper. More details about the experimental parameter identification process can be found in prior work<sup>20</sup>.



**Machine learning control via bayesian optimization.** Modeling, control, and higher level planning of IPMC-based actuators is complicated by both the flexibility of the material as well as the material's complex electrochemical behavior. The flexibility of the material can potentially be addressed in manners similar to that of other soft and elastomeric devices<sup>70</sup>. As with pneumatic or fluid elastomer actuators, curvature is controlled in IPMC-based systems. Kinematic models can thus be developed by modeling IPMC-based systems with piece-wise functions that describe segments' deformations from their original undeflected states<sup>3,71</sup>. Further, the mechanical behavior of IPMCs can be modeled similar to that of other polymers or elastomers<sup>72</sup>. But the response of IPMCs to electrical inputs suffers from non-repeatability, nonlinear effects, and there are still significant challenges to integrating sensors on IPMC actuators to measure state. These difficulties make it difficult to implement conventional feedback controllers<sup>32,33</sup>. Machine learning control (which encompasses a wide variety of techniques) is an attractive solution for control of IPMC-based devices because, like conventional adaptive control, the approach can adjust control policy parameters to address non-repeatability and complex non-linear effects. However, the fundamental difference between conventional adaptive control and machine learning control is that conventional adaptive control theoretically employs continuous-time feedback signals, whereas machine learning control is an iterative process, where improvements are done from one operating cycle to the next. Consequently, the requirements for sensor feedback can be minimal for some machine learning control techniques. For instance, use of direct policy search methods in highly repetitive tasks (like gait optimization for a walking robot or optimization of the stroke of a diaphragm pump) are excellent examples of machine learning being employed with minimal sensor feedback. In these cases, a gross performance metric is only sampled once at the end of an iteration (such as a period of attempted walking or a period of attempted pumping). This is particularly relevant to soft actuators because in practical applications it may often be possible to take periodic measurements of cumulative performance (like distance traveled or volume of fluid pumped) but not to continuously monitor a system output (like actuator displacement). This use of reinforcement learning reflects a larger trend in which control methods are being developed to suite the compromised accuracy, repeatability, and internal sensing capabilities of low-cost robots and devices<sup>73</sup>.

Direct policy search methods belong to the broader family of reinforcement learning which is extensively used in soft robotics. Reinforcement learning iteratively evaluates the reward associated with states, policies and/or actions in an attempt to optimize a control policy. Broadly speaking, there are three kinds of reinforcement learning methods: model-based methods, value-based model-free methods, and policy-based model-free methods (or direct policy search methods). Model-based methods attempt to learn a model of the system (generally in the form of a Markov decision process) and then solve for the optimal control policy based on cumulative reward<sup>73,74</sup>. Most recently, model-based methods employing neural networks as models have been deployed on a variety of platforms including a tensegrity-based mobile robot and a pneumatic manipulator<sup>75,76</sup>. Value-based model-free methods attempt to learn a function of the total possible cumulative reward, but do not employ a state transition distribution. Most recently, the temporal difference (TD) algorithm has been employed to control the position and stiffness of a fluid actuated soft robotic arm<sup>77,78</sup> and Q learning has been employed to learn an effective control policy for the dielectric elastomer actuators employed in an artificial cuttlefish robot<sup>79</sup>. In contrast to these methods, policy-based model-free methods do not employ either a model or a cumulative value function, but rather directly optimize control parameters in the policy space<sup>74</sup>. Direct policy search methods include policy gradient methods<sup>47</sup> and genetic algorithms<sup>80</sup>. The general advantages of direct policy search methods are that they are simpler and often are more computationally efficient.

Recently, Bayesian optimization has been successfully employed as a direct policy search method to learn a walking gait for a biped robot<sup>35,81</sup>. In comparison to gradient-based methods and methods employing genetic algorithms, Bayesian optimization converges in significantly fewer iterations and is theoretically guaranteed to find the global optimum<sup>82</sup>. Bayesian optimization assumes the objective function can be modeled as a Gaussian process (GP) and utilizes all previous evaluation points to select future ones. Therefore, it sacrifices computational efficiency to make better use of available data<sup>48</sup>. Previous works have emphasized Bayesian optimization's ability to function as a black box optimizer without requiring expert knowledge (such as in the form of a dynamics model)<sup>35,74</sup>. However, another notable feature of Bayesian optimization is precisely its ability to incorporate incomplete or imperfect expert knowledge to speed up convergence to the true global optimum. This can be done by encoding known reachable performance targets. It can also be done by incorporating a prior distribution based on simulation data that, while not capturing every aspect of a system dynamics, is known to be generally accurate<sup>74</sup>. In previous uses of Bayesian optimization since optimization was conducted without the benefit of a model to generate simulated results, the selection of the initial evaluation points were effectively random<sup>35</sup>. In one of the methods employed here, a dynamics model of the robot is exploited and Bayesian optimization is conducted first in simulation on the model. The Gaussian process trained in simulation is then used as the prior for Bayesian optimization on the real robot.

In deploying Bayesian optimization on the crawling robot, the goal is to determine the input amplitudes and relative phases of the inputs that are applied to the leg and body section such that the robot achieves robust locomotion. To do so, an effective crawling gait is learned using Bayesian optimization. In this approach, a Gaussian process (GP) models robot speed,  $f(\theta)$ , over a domain of gait parameters,  $\theta$  (the phase and amplitudes of inputs into the robot actuators). The GP is trained from the set of past evaluation points of the objective function, that constitutes the training data,  $\mathbb{T}$ . The GP is then mapped to a surface *via* an acquisition function. A new evaluation point is selected by optimizing over the resultant surface<sup>35</sup>. This technique is employed first on the robot in simulation and then on the real robot experimentally using the GP from optimizing in simulation as the prior distribution. The general Bayesian optimization algorithm is given in Alg. 1.

A Gaussian process is used in Bayesian optimization to define what is termed the "virtual objective function",  $\hat{f}(\theta)$ . The Gaussian process models the relation between parameters,  $\theta$ , and the experimental performance results

from testing those parameters,  $f(\theta)$ , as a distribution over functions<sup>35</sup>. The function  $\hat{f}(\theta)$  is defined by a mean function  $m(\theta)$  and a covariance function  $k(\theta)$ :

$$\hat{f}(\theta) \sim GP(m(\theta), k(\theta, \theta')).$$

The virtual objective function is used to predict probable performance results  $\mathbf{f}_*$  at selected parameter values  $\Theta_*$ . These future predictions are modeled as jointly Gaussian with prior data,  $\{\Theta, \mathbf{f}\}$ . This is expressed as follows:

---

**Algorithm 1.** Bayesian optimization algorithm.

---

```

1 while optimizing do
2   Train Gaussian Process model from  $\mathbb{T}$ 
3   Compute the response surface  $\hat{f}(\cdot)$ 
4   Compute the acquisition surface  $\alpha(\cdot)$ 
5   Find  $\theta^*$  that optimizes  $\alpha(\cdot)$ 
6   Evaluate  $f$  at  $\theta^*$ 
7   Add  $\{\theta^*, f(\theta^*)\}$  to  $\mathbb{T}$ 

```

---

$$\begin{bmatrix} \mathbf{f} \\ \mathbf{f}_* \end{bmatrix} \sim N \left( \begin{bmatrix} \mathbf{m}(\Theta) \\ \mathbf{m}(\Theta_*) \end{bmatrix}, \begin{bmatrix} K(\Theta, \Theta) & K(\Theta, \Theta_*) \\ K(\Theta_*, \Theta) & K(\Theta_*, \Theta_*) \end{bmatrix} \right),$$

where  $\mathbf{m}$  is a vector of mean values and  $K$  is the covariance matrix<sup>83</sup>. It follows that the posterior probability distribution of  $\mathbf{f}_*$  is given by the equation,

$$\begin{aligned} \mathbf{f}_* | \Theta_*, \Theta, \mathbf{f} &\sim N(\mathbf{m}(\Theta_*) + K(\Theta_*, \Theta)K(\Theta, \Theta)^{-1}(\mathbf{f} - \mathbf{m}(\Theta)), \\ &K(\Theta_*, \Theta_*) - K(\Theta_*, \Theta)K(\Theta, \Theta)^{-1}K(\Theta, \Theta_*)). \end{aligned} \quad (9)$$

In the case of a zero prior mean, Eq. (9) becomes

$$\begin{aligned} \mathbf{f}_* | \Theta_*, \Theta, \mathbf{f} &\sim N(K(\Theta_*, \Theta)K(\Theta, \Theta)^{-1}\mathbf{f}, \\ &K(\Theta_*, \Theta_*) - K(\Theta_*, \Theta)K(\Theta, \Theta)^{-1}K(\Theta, \Theta_*)). \end{aligned} \quad (10)$$

If a model of the objective function exists, Eq. (10) can still be used to predict the error between future predictions and predictions of the existing model. In that case error is defined as follows:  $e(\theta) = f(\theta) - f_e(\theta)$ , where  $e(\theta)$  is the error and  $f_e(\theta)$  is the prediction of the existing model at  $\theta$ . Equation (10) then becomes

$$\begin{aligned} \mathbf{e}_* | \Theta_*, \Theta, \mathbf{e} &\sim N(K(\Theta_*, \Theta)K(\Theta, \Theta)^{-1}\mathbf{e}, \\ &K(\Theta_*, \Theta_*) - K(\Theta_*, \Theta)K(\Theta, \Theta)^{-1}K(\Theta, \Theta_*)). \end{aligned}$$

Subsequently,  $f(\theta^*)$  can then be determined from  $f(\theta^*) = e(\theta^*) + f_e(\theta^*)$ , which is the summation of two normally distributed random variables. Use of the error requires modification to the Bayesian optimization, however. The Gaussian process is now trained on the error, but optimization is still conducted with respect to  $\hat{f}(\theta)$ . Therefore, computing the response surface involves combining the existing model,  $\hat{f}_e(\theta)$ , and the predicted error,  $\hat{e}(\theta)$ . Moreover, the actual error at an evaluation point must be computed from the difference between the actual performance and the prediction of the existing model. This modified Bayesian optimization algorithm is shown in Alg. 2.

Here, the covariance function used is the squared exponential kernel, which is defined as

$$k(\theta_p, \theta_q) = \sigma_f^2 \exp\left(-\frac{1}{2}(\theta_p - \theta_q)^T \lambda (\theta_p - \theta_q)\right) + \sigma_w^2 \delta_{pq} \quad (11)$$

where  $\sigma_f^2$  is the functional variance,  $\lambda$  is a diagonal matrix of characteristic lengths,  $\sigma_w^2$  is the noise variance, and  $\delta_{pq}$  is the Kronecker delta. The parameters  $\sigma_f^2$ ,  $\lambda$ , and  $\sigma_w^2$  are set by way of automatic relevance determination. The acquisition function employed to map the response surface to the acquisition surface is the probability of improvement (PI) which is defined as

$$\alpha(\theta^*) = \frac{T - \mu(\theta^*)}{\sigma(\theta^*)}, \quad (12)$$

where  $T$  is the target,  $\mu(\theta^*)$  is the predicted mean, and  $\sigma(\theta^*)$  is the predicted standard deviation of  $f(\theta^*)$ . The target, is generally defined as the best evaluated response on the real system; that is  $T = \max(\mathbf{f})$ , where,  $\max(\mathbf{f}) = \max(\mathbf{e} + \mathbf{f}_e)$ , when using a model to inform the prior. Optionally, a value that diminishes with successive evaluations may be added to this. Alternatively, the target may be set to a known reachable value. Both the option of adding a successively diminishing value and defining the target as a known reachable value causes the acquisition function to be more explorative in selecting earlier evaluation steps.

**Algorithm 2.** Modified algorithm.

---

```

1 while optimizing do
2   Train GP model from  $\mathbb{T}$ 
3   Compute response surface  $\hat{e}(\cdot)$  and  $\hat{f}_e(\cdot)$ 
4   Compute acquisition surface  $\alpha(\cdot)$ 
5   Find  $\theta^*$  that optimizes  $\alpha(\cdot)$ 
6   Evaluate  $f$  and  $f_e$  at  $\theta^*$ 
7   algorithm Add $\{\theta^*, e(\theta^*)\}$  to  $\mathbb{T}$ 

```

---

	Body 1	Body 2	Leg 1	Leg 2	Leg 3
leg-body-body-leg	0°	0°	270°	n/a	90°
leg-body-leg-body-leg	0°	30°	270°	0°	110°

**Table 2.** Optimized phases for Bayesian optimization in simulation.**Results and Discussion**

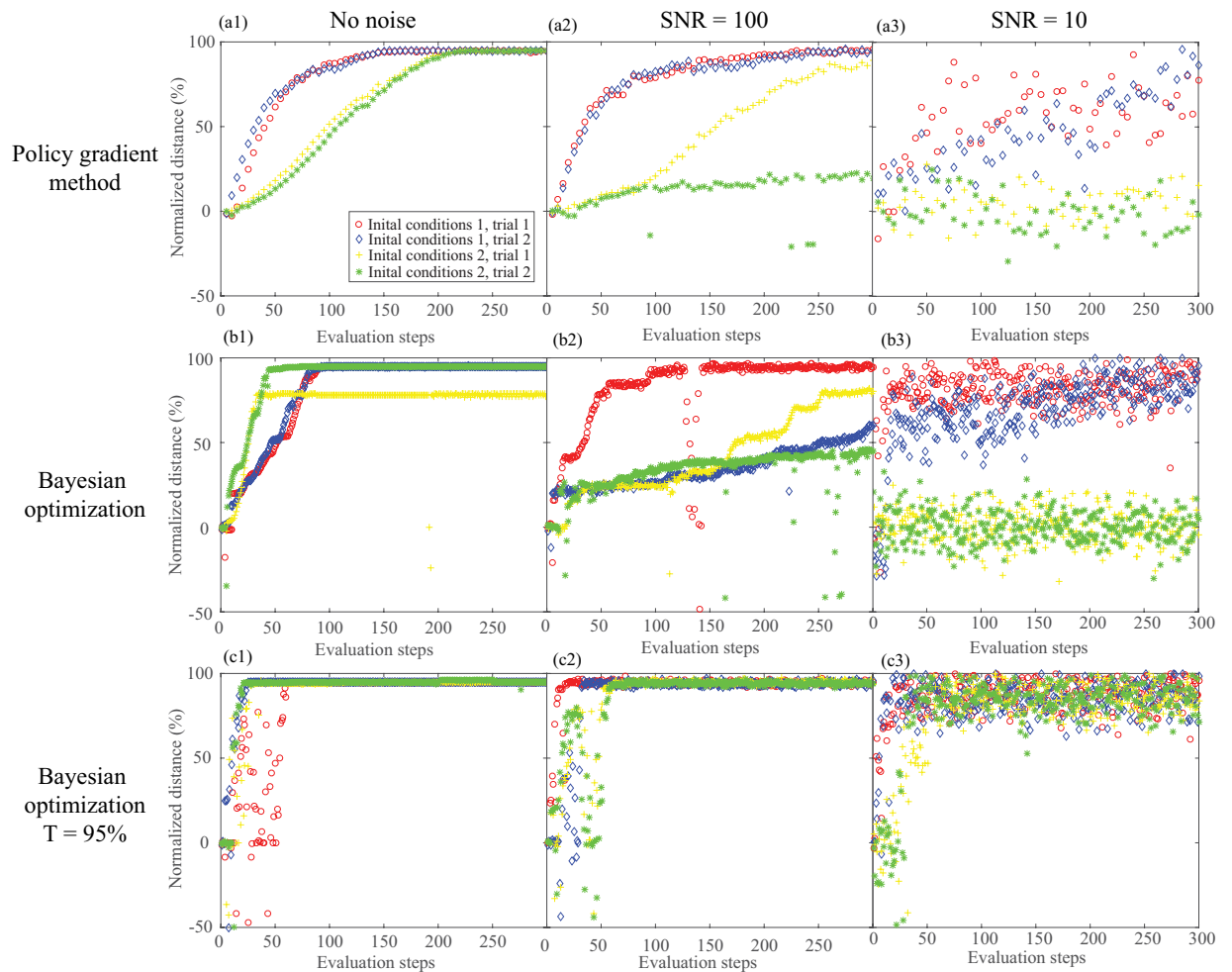
To validate the dynamics model of both the leg-body-body-leg and leg-body-leg-body-leg configurations of the crawling robot, test gaits were determined by first optimizing on the respective dynamics models in simulation. These gaits were then applied to the real robot and the performance compared to that of the simulations. These results show good agreement between the performance of the real crawling robot and its performance in simulation. The validated dynamics model for the leg-body-body-leg configuration was then used in simulation to compare the performance of Bayesian optimization to a finite difference policy gradient method to motivate the use of Bayesian optimization as a gait learning method. These simulations showed that Bayesian optimization outperformed a finite difference policy gradient method, supporting the hypothesis that Bayesian optimization will lead to convergence in fewer trials than a finite-difference policy search method making it more suitable for deployment on IPMC based robots in real-time. Bayesian optimization was then conducted on the real crawling robot, specifically to compare Bayesian optimization conducted from a uniform prior distribution and Bayesian optimization using a Gaussian process trained in simulation as the prior. Unfortunately, unexpected degradation of the IPMCs was observed with extended use. By the end of successive experiments the IPMC could not replicate its same level of performance upon application of the same crawling gait. This technically violates the assumption that successive evaluations would be independent and identically distributed (IID) and thus invalidates the use of a Gaussian process to model the objective function. In spite of this, both approaches to Bayesian optimization show improving performance with successive evaluations. This section will first discuss the simulation methods and results and then the experiment methods and results.

**Simulation results and discussion.** Optimization conducted to determine the test gaits for both the leg-body-body-leg and leg-body-leg-body-leg configurations was done by running successive simulations on the dynamics models in Simulink using 10-minute runs (simulation-time), using automatic solver selection with a variable step size. The Gaussian process used as the prior for subsequent optimization on the real robot was also obtained from the simulation on the dynamics model of the leg-body-body-leg configuration.

Comparison of Bayesian optimization and the policy gradient method were conducted by running successive simulations on the dynamics model of the leg-body-body-leg configuration of the crawling robot in Simulink using 15-second runs (simulation-time), using automatic solver selection with a variable step size. A random variable was added to the simulated distance traveled to add noise to the results. The decision variables used for optimization were the phases of the inputs to the individual body and leg components. Each optimization method was tested under three different noise levels (no noise, a signal to noise ratio of 1 and a signal to noise ratio of 10). Two initial conditions were tested under each noise level. For each initial condition, two trials were run to account for randomness in the simulations. Since, an advantage of Bayesian optimization is its ability to encode a known reachable target value, the performance of both Bayesian optimization where the target is defined as the maximum of the training data and Bayesian optimization where the target is defined as 95% the maximum possible distance are tested.

Table 2 gives the optimized phases of the inputs for both the the leg-body-body-leg and the leg-body-leg-body-leg configurations. In this table, the bodies and legs are enumerated moving forward along the direction of travel. As can be seen from Table 2, Bayesian optimization for the leg-body-body-leg configuration results in converged gait phases of 0°, 0°, 270° and 90° for body 1, body 2, leg 1 and leg 3 respectively and Bayesian optimization for the leg-body-leg-body-leg configuration results in converged gait phases of 0°, 30°, 270°, 0° and 110° for body 1, body 2, leg 1, leg 2, and leg 3 respectively. The effect of adding the middle leg to the robot is to cause the maximum possible speed to decrease. Notably the converged gait for the leg-body-leg-body-leg configuration only has body 2 and leg 3 slightly lagging their converged phases in the leg-body-body-leg configuration.

Figure 6 compares the use of finite-difference policy gradient method<sup>47</sup> and Bayesian optimization in simulation on the dynamics model of the leg-body-body-leg configuration of the crawling robot. Figure 6(a1–a3) shows the performance of the finite difference policy gradient method. Figure 6(b1–b3) shows the performance of Bayesian optimization using the probability of improvement acquisition function where the target is defined



**Figure 6.** Policy gradient method using finite-difference gradient estimation vs. Bayesian optimization using probability of improvement (PI) acquisition function, with multiple trials from different ICs and signal to noise ratios (SNR): **(a1)** policy gradient with no noise, **(a2)** policy gradient with SNR = 100, **(a3)** policy gradient with SNR = 10, **(b1)** Bayesian optimization with target equal to maximum of training data with no noise, **(b2)** Bayesian optimization with target equal to maximum of training data with SNR = 100, **(b3)** Bayesian optimization with target equal to maximum of training data with SNR = 10, **(c1)** Bayesian optimization with target equal to 95% maximum possible distance with no noise, **(c2)** Bayesian optimization with target equal to 95% maximum possible distance with SNR = 100, **(c3)** Bayesian optimization with target equal to 95% maximum possible distance with SNR = 10.

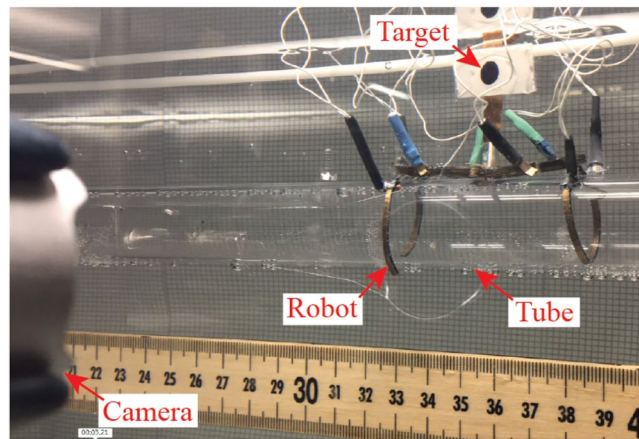
as the maximum of the training data. Figure 6(c1–c3) shows the performance of Bayesian optimization using the probability of improvement acquisition function where the target is defined as 95% the maximum possible distance of the robot. Bayesian optimization using both the maximum of the training data as the target and 95% the maximum possible distance as the target, perform as well or better than the policy gradient method. These results support the hypothesis that Bayesian optimization will be superior for use on the real crawling robot.

**Experiment results and discussion.** Experiments on the crawling robot were conducted using Simulink real-time and Ni-Daq control hardware to output control signals and LM675T power amplifiers were used to power the signals. For comparison of the simulated and experimental performance of the crawling robot in response to the test gaits determined by optimizing in simulation, video recordings were taken and the Kanade-Lucas-Tomasi (KLT) algorithm was used to track points on the robots at the locations of the rear leg ( $x_1$ ), central leg or central joint ( $x_2$ ), and front leg ( $x_3$ ), respectively, in the videos of the experimental trials.

To compare Bayesian optimization conducted starting from a uniform prior distribution to that of Bayesian optimization using Gaussian process trained in simulation as the prior, the leg-body-body-leg configuration was used and a webcam was added to track the robot's position for 60 second evaluations of a set of gait parameters, as shown in Fig. 7. The direction of each trial was selected to automatically center the robot (*i.e.*, if the robot was to the right of its origin at the beginning of a trial it would move left and vice-versa).

Figure 8 shows the experimental motion of the tracked locations for the leg-body-body-leg configuration and the leg-body-leg-body-leg configuration over 20 seconds compared to their simulated motion. Figure 8(a2,b2)





**Figure 7.** Experimental setup for optimization showing the robot, and the target and camera used to acquire the robot's position. Optimization and control is conducted using Matlab and Simulink Real-Time software package.

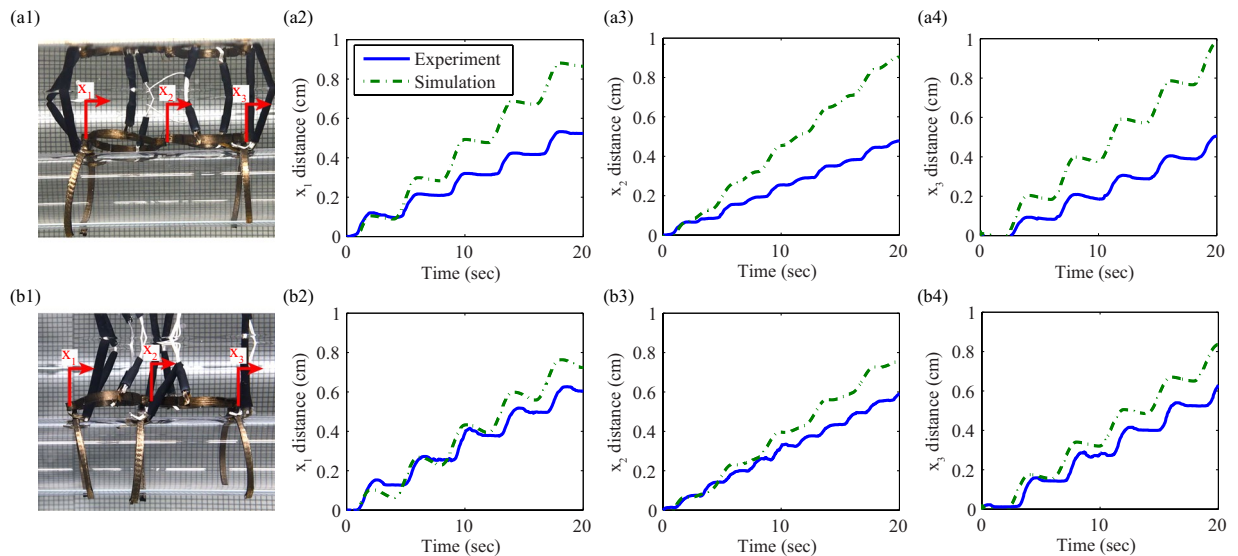
show the motion of the rear legs, Fig. 8(a3,b3) show the motion of the central joint, and Fig. 8(a4,b4) show the motion of the front legs. The real speed of the leg-body-body-leg configuration with input phases of  $0^\circ$ ,  $0^\circ$ ,  $270^\circ$  and  $90^\circ$  for body 1, body 2, leg 1 and leg 2 respectively is 2.15 cm/min as compared to the simulated speed of 2.9 cm/min. The real speed of the leg-body-leg-body-leg configuration with input phases of  $0^\circ$ ,  $30^\circ$ ,  $270^\circ$ ,  $0^\circ$  and  $110^\circ$  for body 1, body 2, leg 1, leg 2, and leg 3 respectively is 1.8 cm/min as compared to the simulated speed of 2.3 cm/min. Neglecting non-linear friction in the model is responsible for the discrepancies between the simulated and experimental performance. For instance, comparing the plots of the experimental and simulated motion in Fig. 8(b2), there are spans of time during which the velocity of the leg of the real robot is 0. By contrast, the velocity of the leg of the simulated robot is only momentarily zero. This difference is because the dynamics of crawling robot were characterized assuming a linear viscous friction effect. This neglects a nonlinear stiction effect by which there is a range of inputs for which the robot will not move, after coming to rest.

Figure 9 shows the results of Bayesian optimization on the crawling robot using a uniform prior distribution and of Bayesian optimization using a Gaussian process trained in simulation as the prior. Specifically, Fig. 9 shows the speed attained at each evaluation step for both optimization methods. Figure 9(a) shows these speeds in the leftwards direction and Fig. 9(b) shows these speeds in rightwards direction. Neither the optimization from a uniform prior nor the optimization using a GP trained in simulation as the prior show clear convergence for the leftwards direction. The optimization from a uniform prior reaches a maximum speed of 0.5 cm/min and optimization using the GP trained in simulation as the prior reaches a maximum speed of about 0.6 cm/min. Both optimization from a uniform prior and using the GP trained in simulation as the prior converge to a speed of 0.5 cm/min for optimization in the rightwards direction. The optimization from a uniform prior converges after 20 evaluations and the optimization using the GP trained in simulation as the prior converges after 15 evaluations. Table 3 shows the final optimized gait phases for both the leftwards and rightwards directions. For the leftwards direction the optimization does not seem to converge but the final input phases were  $0^\circ$ ,  $40^\circ$ ,  $100^\circ$  and  $250^\circ$  for body 1, body 2, leg 1 and leg 2 respectively. For the rightwards direction the converged input phases were  $0^\circ$ ,  $0^\circ$ ,  $240^\circ$  and  $100^\circ$  for body 1, body 2, leg 1 and leg 2 respectively. The lower speeds compared with those achieved using test gaits derived from optimizing in simulation could be due to the degradation of the IPMC performance with successive evaluations. This would also effect the optimization algorithm, possibly preventing it from converging to the optimum gait. As a consequence of the degradation of the IPMC actuators performance over time, the objective function being optimized over would in fact be time dependent. This would invalidate the assumption implicit in a Gaussian process, that successive evaluations of the same parameter set would be from identical distributions. This problem could possibly be ameliorated by using different policy representations and ensemble methods in the future.

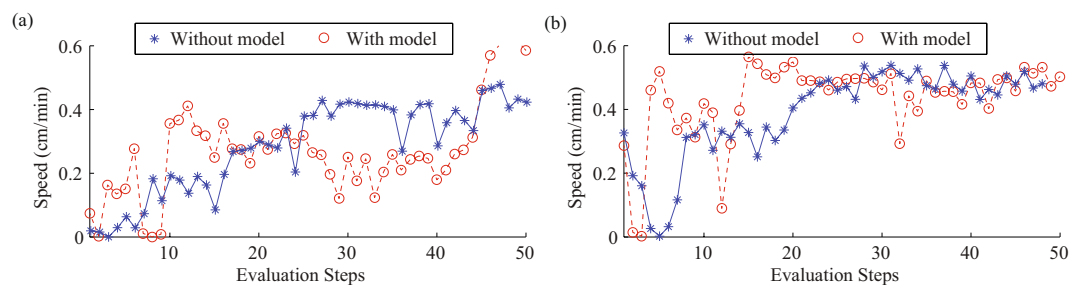
### Emerging Challenges

In the development of the 3D printing and learning-based control of IPMC actuators, several challenges have emerged that can be addressed by future research. One of the emerging challenges specific to 3D printing IPMCs is that increased printing resolution leads to increased failures in the printing process. This impedes the printing of thin-walled actuators that exhibit large displacements when actuated. This can potentially be ameliorated by a thorough study of the printing process variables. Additionally, alternative melt-processing-based 3D printing methods for Nafion precursor such as selective laser sintering (SLS) or use of alternative precursor materials, such as polystyrene in the FDM 3D printing process could also lead to increased resolution. In the use of Nafion precursor in an SLS 3D printing process, the critical inquiries would be into how the powder should be prepared and the required print settings for fabricating solid fused structures in Nafion precursor with high-resolution features, such as walls as thin as 0.1 mm or less. Alternative materials, such as Polystyrene, which are more easily printed than Nafion precursor through the FDM process, might also be considered. Materials such as polystyrene are more controllable in FDM 3D printing, because they are more rigid in their solid state, and have a lower melting





**Figure 8.** Experimental and simulated performance for optimized gaits for both the leg-body-body-leg and leg-body-leg-body-leg configurations showing the positions of the rear leg, the central joint and the front leg respectively as functions of time: (a1) leg-body-body-leg configuration of real robot, (a2) position of front leg of leg-body-body-leg configuration, (a3) position of central joint of leg-body-body-leg configuration, (a4) position of rear leg of leg-body-body-leg configuration, (b1) leg-body-leg-body-leg configuration of real robot, (b2) position of front leg of leg-body-leg-body-leg configuration, (b3) position of central leg of leg-body-leg-body-leg configuration and (b4) position of rear leg of leg-body-leg-body-leg configuration.



**Figure 9.** Speed for each evaluation step in Bayesian optimization on the real robot done both from a uniform prior (without a model) and using the GP trained in simulation as the prior (with a model): (a) in the leftwards direction (b) in the rightwards direction.

	Body 1	Body 2	Leg 1	Leg 2
Leftwards	0	40°	100°	250°
Rightwards	0	0°	240°	100°

**Table 3.** Optimized phases for Bayesian optimization on robot.

temperature. However, a critical research consideration in adapting alternative materials to the print then functionalize paradigm is that these materials require *in situ* functionalization processes. In the case of polystyrene, for instance, this material requires identification of an *in situ* partial sulfonation process to obtain the functional material, lightly-sulfonated polystyrene, which acts as an ion-exchange membrane.

Another challenge to the 3D printing of IPMCs is control of electrode development in the post printing process, such that discrete electrode regions are generated without need for segmentation. This is especially needed for the manufacturing of complex multi-input-multi-output IPMC devices and systems, which require multiple electrode pairs. A potential solution to this challenge is incorporating a secondary material in the printing process, specifically to control electrode development in the post-printing steps. Since electrode development using reduction processes, such as are used for developing platinum on Nafion membranes, can only deposit platinum ions in functional materials, secondary materials can prevent the development of electrodes in selected regions by masking the ion-exchange material. Alternatively, another approach might be to print electrode materials directly onto the precursor or functionalized material.

In the area of motion control, performance degradation, especially when operated above the electrolysis voltage of the hydrating solvent<sup>84</sup>, is a significant emerging challenge to the control of IPMCs. This performance degradation is not well understood and can lead to significant under performance if performance degradation is not planned for or avoided. Since performance degradation changes the dynamic behavior of IPMCs over time, it invalidates control policies trained on an IPMC's past performance. This is a challenge for conventional Bayesian optimization which treats each data point as drawn from an identical distribution. Moreover, since conventional control methods do not optimize the performance of a device over its functional lifetime, these methods could lead to accelerated degradation. To address this challenge, control techniques are needed that anticipate and plan for performance degradation. One possible approach is use of Gaussian processes to project IPMC performance over time, based on general trends for similar IPMCs and real-time degradation data of the IPMC being controlled.

## Conclusions

This paper described 3D printing for manufacturing and machine learning for control of example IPMC actuator components. A prototype modular reconfigurable soft crawling robot platform was described. The development of the platform was presented to illustrate the proof-of-concept of 3D-printing of IPMCs actuators and the application of machine learning for effective motion control. Furthermore, this paper tested two hypotheses: (1) that Bayesian optimization will lead to convergence in fewer trials than a finite-difference policy gradient method making it more suitable and practical for controlling IPMCs; and (2) that prior knowledge from a dynamics model (especially a known achievable target value) will lead to convergence in fewer number of trials than simply optimizing from a uniform prior distribution. To investigate these hypotheses, the dynamics of the crawling robot were modeled and the model was used to test various learning-based control methods in simulation. Bayesian optimization was also applied directly to the crawling robot to investigate the practical impact of the degradation on this motion control strategy. Simulation results indicate that Bayesian optimization employing prior knowledge in the form of a known achievable performance target does lead to much faster convergence than either Bayesian optimization without this knowledge or a conventional finite difference policy gradient method. Experiments directly on the crawling robot indicated that performance degradation is significant (reducing the crawling robot performance to less than a quarter of its initial achievable performance over the duration of optimization). This invalidates the performance level explicitly incorporated as the target in the acquisition function for Bayesian optimization. Technically, this would also invalidate the assumption implicit in the manner in which the GP was used, that successive evaluations of the same parameter set would be from identical distributions. None-the-less, Bayesian optimization still demonstrated continuous performance improvement with successive trials on the crawling robot. These results also established proof-of-concept for continued advancement of more complex IPMC devices.

Received: 28 March 2019; Accepted: 31 October 2019;

Published online: 25 November 2019

## References

- Shahinpoor, M. & Kim, K. J. Ionic polymer-metal composites: I. fundamentals. *Smart Materials and Structures* **10**, 819 (2001).
- Bhandari, B., Lee, G.-Y. & Ahn, S.-H. A review on IPMC material as actuators and sensors: fabrications, characteristics and applications. *International Journal of Precision Engineering and Manufacturing* **13**, 141–163 (2012).
- Moghadam, A. A. A., Kouzani, A., Torabi, K., Kaynak, A. & Shahinpoor, M. Development of a novel soft parallel robot equipped with polymeric artificial muscles. *Smart Materials and Structures* **24**, 035017 (2015).
- Fang, B.-K., Ju, M.-S. & Lin, C.-C. K. A new approach to develop ionic polymer-metal composites (IPMC) actuator: Fabrication and control for active catheter systems. *Sensors and Actuators A: Physical* **137**, 321–329 (2007).
- Kruusmaa, M., Hunt, A., Punning, A., Anton, M. & Aabloo, A. A linked manipulator with ion-polymer metal composite (IPMC) joints for soft-and micromanipulation. In Robotics and Automation, 2008. ICRA 2008. *IEEE International Conference on*, 3588–3593 (IEEE, 2008).
- McDaid, A., Aw, K., Haemmerle, E., Shahinpoor, M. & Xie, S. Adaptive tuning of a 2dof controller for robust cell manipulation using IPMC actuators. *Journal of Micromechanics and Microengineering* **21**, 125004 (2011).
- Wang, H., Chen, J., Lau, H. Y. & Ren, H. Motion planning based on learning from demonstration for multiple-segment flexible soft robots actuated by electroactive polymers. *IEEE Robotics and Automation Letters* **1**, 391–398 (2016).
- Jain, R. K., Majumder, S. & Dutta, A. Scara based peg-in-hole assembly using compliant IPMC micro gripper. *Robotics and Autonomous Systems* **61**, 297–311 (2013).
- Nguyen, T. T., Goo, N. S., Nguyen, V. K., Yoo, Y. & Park, S. Design, fabrication, and experimental characterization of a flap valve IPMC micropump with a flexibly supported diaphragm. *Sensors and Actuators A: Physical* **141**, 640–648 (2008).
- Ramírez-García, S. & Diamond, D. Biomimetic, low power pumps based on soft actuators. *Sensors and Actuators A: Physical* **135**, 229–235 (2007).
- Lee, S. & Kim, K. J. Design of IPMC actuator-driven valve-less micropump and its flow rate estimation at low reynolds numbers. *Smart Materials and Structures* **15**, 1103 (2006).
- Arena, P., Bonomo, C., Fortuna, L., Frasca, M. & Graziani, S. Design and control of an IPMC wormlike robot. *IEEE Trans. Systems, Man and Cybernetics, Part B (Cybernetics)* **36**, 1044–1052 (2006).
- Yamakita, M., Sera, A., Kamamichi, N., Asaka, K. & Luo, Z.-W. Integrated design of IPMC actuator/sensor. In *IEEE International Conference on Robotics and Automation*, 1834–1839.
- Aureli, M., Kopman, V. & Porfiri, M. Free-locomotion of underwater vehicles actuated by ionic polymer metal composites. *IEEE/ASME Transactions on Mechatronics* **15**, 603–614 (2010).
- Chen, Z. & Tan, X. Mems-based fabrication of multiple-degree-of-freedom ionic polymer-metal composite actuators. In *SPIE Smart Structures and Materials+ Nondestructive Evaluation and Health Monitoring*, 76420X–76420X (International Society for Optics and Photonics, 2010).
- Kim, K. J., Pugal, D. & Leang, K. K. A twistable ionic polymer-metal composite artificial muscle for marine applications. *Marine Technology Society Journal* **45**, 83–98 (2011).
- Hubbard, J. J. *et al.* Monolithic IPMC fins for propulsion and maneuvering in bio-inspired underwater robotics. *IEEE Journal of Oceanic Engineering* **39**, 540–551 (2014).

18. Shi, L., Guo, S. & Asaka, K. A bio-inspired underwater microrobot with compact structure and multifunctional locomotion. In *IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, 203–208 (IEEE, 2011).
19. Firouzeh, A., Ozmaeian, M., Alasty, A. & Irajizad, A. An IPMC-made deformable-ring-like robot. *Smart Materials and Structures* **21**, 065011 (2012).
20. Carrico, J. D., Kim, K. J. & Leang, K. K. 3D-printed ionic polymer-metal composite soft crawling robot. In *IEEE International Conference on Robotics and Automation (ICRA)*, 4313–4320 (2017).
21. Vargantwar, P. H., Roskov, K. E., Ghosh, T. K. & Spontak, R. J. Enhanced biomimetic performance of ionic polymer-metal composite actuators prepared with nanostructured block ionomers. *MACROMOLECULAR RAPID COMMUNICATIONS* **33**, 61–68 (2012).
22. Panwar, V., Ko, S. Y., Park, J.-O. & Park, S. Enhanced and fast actuation of fullerene/pvdf/pvp/psaa based ionic polymer metal composite actuators. *Sensors and Actuators B: Chemical* **183**, 504–517 (2013).
23. Punning, A. *et al.* Ionic electroactive polymer artificial muscles in space applications. *Scientific Reports* **4**, 6913 (2014).
24. Lee, J.-W., Yoo, Y.-T. & Lee, J. Y. Ionic polymer-metal composite actuators based on triple-layered polyelectrolytes composed of individually functionalized layers. *ACS APPLIED MATERIALS & INTERFACES* **6**, 1266–1271 (2014).
25. Kim, O., Kim, H., Choi, U. H. & Park, M. J. One-volt-driven superfast polymer actuators based on single-ion conductors. *Nature Communications* **7**, 13576 (2016).
26. Kim, J., Jeon, J.-H., Kim, H.-J., Lim, H. & Oh, I.-K. Durable and water-floatable ionic polymer actuator with hydrophobic and asymmetrically laser-scribed reduced graphene oxide paper electrodes. *ACS Nano* **8**, 2986–2997 (2014).
27. Kim, S. J., Pugal, D., Wong, J., Kim, K. J. & Yim, W. A bio-inspired multi degree of freedom actuator based on a novel cylindrical ionic polymer-metal composite material. *Robotics and Autonomous Systems* **62**, 53–60 (2014).
28. Malone, E. & Lipson, H. Freeform fabrication of ionomeric polymer-metal composite actuators. *Rapid Prototyping Journal* **12**, 244–253 (2006).
29. Malone, E. & Lipson, H. Multi-material freeform fabrication of active systems. In *Proceedings 9th Biennial Conference on Engineering Systems Design and Analysis*, 345–353 (ASME, 2008).
30. Carrico, J. D., Traeden, N. W., Aureli, M. & Leang, K. K. Fused filament 3D printing of ionic polymer-metal composites (IPMCs). *Smart Materials and Structures* **24**, 125021 (11 pages) (2015).
31. Kang, S., Shin, J., Kim, S. J., Kim, H. J. & Kim, Y. H. Robust control of ionic polymer-metal composites. *Smart Materials and Structures* **16**, 2457 (2007).
32. Shan, Y. & Leang, K. K. Frequency-weighted feedforward control for dynamic compensation in ionic polymer-metal composite actuators. *Smart Materials and Structures* **18**, 125016 (2009).
33. Kruusamäe, K., Punning, A., Aabloo, A. & Asaka, K. Self-sensing ionic polymer actuators: a review. In *Actuators*, vol. 4, 17–38 (Multidisciplinary Digital Publishing Institute, 2015).
34. Chen, Z., Kwon, K.-Y. & Tan, X. Integrated IPMC/PVDF sensory actuator and its validation in feedback control. *Sensors and Actuators A: Physical* **144**, 231–241 (2008).
35. Calandra, R., Seyfarth, A., Peters, J. & Deisenroth, M. P. An experimental comparison of bayesian optimization for bipedal locomotion. In *IEEE International Conference on Robotics and Automation (ICRA)*, 1951–1958 (IEEE, 2014).
36. Calandra, R., Gopalan, N., Seyfarth, A., Peters, J. & Deisenroth, M. P. Bayesian gait optimization for bipedal locomotion. In *International Conference on Learning and Intelligent Optimization*, 274–290 (Springer, 2014).
37. Zolfagharian, A., Kouzani, A. Z., Khoo, S. Y., Gibson, I. & Kaynak, A. 3d printed hydrogel soft actuators. In *IEEE Region 10 Conference (TENCON)*, 2272–2277 (IEEE, 2016).
38. Hines, L., Petersen, K., Lum, G. Z. & Sitti, M. Soft actuators for small-scale robotics. *Advanced Materials* (2017).
39. Smela, E. Conjugated polymer actuators for biomedical applications. *Advanced Materials* **15**, 481–494 (2003).
40. Jager, E. W., Smela, E. & Inganäs, O. Microfabricating conjugated polymer actuators. *Science* **290**, 1540–1545 (2000).
41. Temmer, R. *et al.* In search of better electroactive polymer actuator materials: Ppy versus pedot versus pedot-ppy composites. *Smart Materials and Structures* **22**, 104006 (2013).
42. Weng, B., Shepherd, R. L., Crowley, K., Killard, A. & Wallace, G. G. Printing conducting polymers. *Analyst* **135**, 2779–2789 (2010).
43. Wu, Y., Alici, G., Madden, J. D., Spinks, G. M. & Wallace, G. G. Soft mechanical sensors through reverse actuation in polypyrrole. *Advanced Functional Materials* **17**, 3216–3222 (2007).
44. John, S. W., Alici, G. & Cook, C. D. Towards the position control of conducting polymer trilayer bending actuators with integrated feedback sensor. In *IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, 65–70 (IEEE, 2009).
45. John, S. W., Alici, G. & Cook, C. D. Inversion-based feedforward control of polypyrrole trilayer bender actuators. *IEEE/ASME Transactions on Mechatronics* **15**, 149–156 (2010).
46. Zolfagharian, A., Kouzani, A. Z., Maheepala, M., Khoo, S. Y. & Kaynak, A. Bending control of a 3d printed polyelectrolyte soft actuator with uncertain model. *Sensors and Actuators A: Physical* **288**, 134–143 (2019).
47. Kohl, N. & Stone, P. Policy gradient reinforcement learning for fast quadrupedal locomotion. In *IEEE International Conference on Robotics and Automation*, vol. 3, 2619–2624 (IEEE, 2004).
48. Snoek, J., Larochelle, H. & Adams, R. P. Practical bayesian optimization of machine learning algorithms. In *Advances in Neural Information Processing Systems*, 2951–2959 (2012).
49. Trabia, S., Olsen, Z., Hwang, T. & Kim, K. J. Producing intricate IPMC shapes by means of painting and printing. In *SPIE Smart Structures and Materials+Nondestructive Evaluation and Health Monitoring* (2017).
50. Moore, R. B., Cable, K. M. & Croley, T. L. Barriers to flow in semicrystalline ionomers. a procedure for preparing melt-processed perfluorosulfonate ionomer films and membranes. *Journal of Membrane Science* **75**, 7–14 (1992).
51. Mauritz, K. A. & Moore, R. B. State of understanding of Nafion. *Chemical Reviews* **104**, 4535–4586 (2004).
52. Carrico, J. D. & Leang, K. K. Fused filament 3d printing of ionic polymer-metal composites for soft robotics. In *SPIE Smart Structures and Materials+ Nondestructive Evaluation and Health Monitoring*, 101630I–101630I (International Society for Optics and Photonics, 2017).
53. DuPont Fluoroproducts. Chemical treatment of Nafion PFSA resins R-1100 and R-1000 (2002).
54. Elliott, J. A. *et al.* Hydrolysis of the Nafion precursor studied by x-ray scattering and *in-situ* atomic force microscopy. *e-Polymers* **1**, 210–220 (2001).
55. Seok, S. *et al.* Meshworm: a peristaltic soft robot with antagonistic nickel titanium coil actuators. *IEEE/ASME Transactions on Mechatronics* **18**, 1485–1497 (2013).
56. Umedachi, T., Vikas, V. & Trimmer, B. A. Highly deformable 3-D printed soft robot generating inching and crawling locomotions with variable friction legs. In *Proceedings Intelligent Robots and Systems (IROS)*, 4590–4595 (IEEE, 2013).
57. Tolley, M. T. *et al.* A resilient, untethered soft robot. *Soft Robotics* **1**, 213–223 (2014).
58. Onal, C. D. & Rus, D. Autonomous undulatory serpentine locomotion utilizing body dynamics of a fluidic soft robot. *Bioinspiration & Biomimetics* **8**, 026003 (2013).
59. Marchese, A. D., Onal, C. D. & Rus, D. Autonomous soft robotic fish capable of escape maneuvers using fluidic elastomer actuators. *Soft Robotics* **1**, 75–87 (2014).
60. Marchese, A. D., Katschmann, R. K. & Rus, D. A recipe for soft fluidic elastomer robots. *Soft Robotics* **2**, 7–25 (2015).
61. Rus, D. & Tolley, M. T. Design, fabrication and control of soft robots. *Nature* **521**, 467–475 (2015).
62. Paez, L., Agarwal, G. & Paik, J. Design and analysis of a soft pneumatic actuator with origami shell reinforcement. *Soft Robotics* **3**, 109–119 (2016).

63. MacCurdy, R., Katzschmann, R., Kim, Y. & Rus, D. Printable hydraulics: a method for fabricating robots by 3d co-printing solids and liquids. In *IEEE International Conference on Robotics and Automation (ICRA)*, 3878–3885 (IEEE, 2016).
64. Wehner, M. *et al.* An integrated design and fabrication strategy for entirely soft, autonomous robots. *Nature* **536**, 451–455 (2016).
65. Mu, Q. *et al.* Thermal cure effects on electromechanical properties of conductive wires by direct ink write for 4d printing and soft machines. *Smart Materials and Structures* **26**, 045008 (2017).
66. Mao, Y. *et al.* 3d printed reversible shape changing components with stimuli responsive materials. *Scientific Reports* **6** (2016).
67. Gladman, A. S., Matsumoto, E. A., Nuzzo, R. G., Mahadevan, L. & Lewis, J. A. Biomimetic 4d printing. *Nature Materials* (2016).
68. Bakarich, S. E. *et al.* 4d printing with mechanically robust, thermally actuating hydrogels. *Macromolecular Rapid Communications* **36**, 1211–1217 (2015).
69. Fleming, M. J., Kim, K. J. & Leang, K. K. Mitigating IPMC back relaxation through feedforward and feedback control of patterned electrodes. *Smart Materials and Structures* **21**, 085002 (12 pages) (2012).
70. Marchese, A. D. & Rus, D. Design, kinematics, and control of a soft spatial fluidic elastomer manipulator. *International Journal of Robotics Research* **35**, 840–869 (2016).
71. Marchese, A. D., Komorowski, K., Onal, C. D. & Rus, D. Design and control of a soft and continuously deformable 2d robotic manipulation system. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2189–2196 (IEEE, 2014).
72. Marchese, A. D., Tedrake, R. & Rus, D. Dynamics and trajectory optimization for a soft spatial fluidic elastomer manipulator. *International Journal of Robotics Research* **35**, 1000–1019 (2016).
73. Polydoros, A. S. & Nalpanitidis, L. Survey of model-based reinforcement learning: Applications on robotics. *Journal of Intelligent & Robotic Systems* **86**, 153–173 (2017).
74. Pautrat, R., Chatzilygeroudis, K. & Mouret, J.-B. Bayesian optimization with automatic prior selection for data-efficient direct policy search. In *IEEE International Conference on Robotics and Automation (ICRA)*, 7571–7578 (IEEE, 2018).
75. Zhang, M. *et al.* Deep reinforcement learning for tensegrity robot locomotion. In *IEEE International Conference on Robotics and Automation (ICRA)*, 634–641 (IEEE, 2017).
76. Thuruthel, T. G., Falotico, E., Renda, F. & Laschi, C. Model-based reinforcement learning for closed-loop dynamic control of soft robotic manipulators. *IEEE Transactions on Robotics* **35**, 124–134 (2018).
77. Ansari, Y., Manti, M., Falotico, E., Cianchetti, M. & Laschi, C. Multiobjective optimization for stiffness and position control in a soft robot arm module. *IEEE Robotics and Automation Letters* **3**, 108–115 (2017).
78. Ansari, Y. *et al.* Towards the development of a soft manipulator as an assistive robot for personal care of elderly people. *International Journal of Advanced Robotic Systems* **14**, 1729881416687132 (2017).
79. Yang, T. *et al.* A soft artificial muscle driven robot with reinforcement learning. *SCIENTIFIC REPORTS* **8**, 14518 (2018).
80. Chernova, S. & Veloso, M. An evolutionary approach to gait learning for four-legged robots. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 3, 2562–2567 (IEEE, 2004).
81. Tesch, M., Schneider, J. & Choset, H. Using response surfaces and expected improvement to optimize snake robot gait parameters. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1069–1074 (IEEE, 2011).
82. Calandra, R., Seyfarth, A., Peters, J. & Deisenroth, M. P. Bayesian optimization for learning gaits under uncertainty. *Annals of Mathematics and Artificial Intelligence* **76**, 5–23 (2016).
83. Rasmussen, C. E. Gaussian processes in machine learning. In *Advanced Lectures on Machine Learning*, 63–71 (Springer, 2004).
84. Bar-Cohen, Y. *et al.* Challenges to the application of IPMC as actuators of planetary mechanisms. In *Smart Structures and Materials 2000: Electroactive Polymer Actuators and Devices (EAPAD)*, vol. 3987, 140–147 (International Society for Optics and Photonics, 2000).

## Acknowledgements

Authors acknowledge financial support from the National Science Foundation (NSF), Partnerships for International Research and Education (PIRE) Program, Grant No. 1545857. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF. Authors also thank Prof. Il-Kwon Oh of KAIST in South Korea for fruitful discussions on IPMC development and Justin Ngo from the DARC Lab for creating the 3D-printed example in Fig. 2.

## Author contributions

J.C. and K.L. conceived the original ideas, with further development from T.H. and K.K. J.C. performed the experiments and collected data. All authors interpreted the results. J.C. and K.L. the wrote the manuscript with contributions from T.H. and K.K. All authors discussed the results and conclusions drawn.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to K.K.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019