



Cite this: *Energy Environ. Sci.*,  
2022, 15, 2874

## Machine learning-assisted design of flow fields for redox flow batteries†

Shuaibin Wan, <sup>a</sup> Haoran Jiang, <sup>be</sup> Zixiao Guo, <sup>a</sup> Changxiang He, <sup>a</sup> Xiongwei Liang, <sup>a</sup> Ned Djilali\*<sup>cd</sup> and Tianshou Zhao <sup>\*a</sup>

Flow fields are a crucial component of redox flow batteries (RFBs). Conventional flow fields, designed by trial-and-error approaches and limited human intuition, are difficult to optimize, thus limiting the performance of RFBs. Here, we develop an end-to-end approach to the design of flow fields by combining machine learning and experimental methods. A library of 11 564 flow fields is generated using a custom-made path generation algorithm, in which flow fields are elegantly encoded by two-dimensional binary images. To accelerate the discovery process, we train convolutional neural networks with low test errors for predicting the uniformity factor and pressure drop of flow fields (0.59% and 1.37%, respectively). Through a collaborative screening process, eight promising candidates are successfully identified. Experimental validation shows that the battery with the flow fields designed with this approach yields higher electrolyte utilization and exhibits about a 22% increase in limiting current density and up to 11% improvement in energy efficiency compared to the conventional serpentine flow field. Furthermore, statistical analysis suggests that the promising candidates have a saved channel length of  $1490 \pm 100$  and a torque integral of  $20.1 \pm 1.8$ , revealing the quantitative design rules of flow fields for the first time.

Received 14th October 2021,  
Accepted 26th May 2022

DOI: 10.1039/d1ee03224k

rsc.li/ees

### Broader context

The redox flow battery is one of the most promising energy storage technologies, benefiting from its decoupled energy and power, flexible scalability, fast response, and high safety. The high capital cost remains a major barrier to the widespread application of the redox flow battery. One effective strategy for cost reduction is to develop batteries with higher power density, which can be realized by the design of flow fields. Conventional flow fields, designed by trial-and-error approaches and limited human intuition, are difficult to optimize. With the exponential growth of computing resources and improvements in simulation tools, machine learning-assisted screening has provided a promising way to solve large-scale design and optimization problems across various disciplines. When it comes to the design of flow fields, there remain two major challenges: one is how to generate a custom library containing numerous flow field designs, and the other is how to screen promising candidates from the search library. In this study, we successfully address these two challenges and develop an end-to-end approach to the design of flow fields. Furthermore, we reveal the quantitative design rules of flow fields for redox flow batteries for the first time.

## 1 Introduction

Developing large-scale energy storage systems is an effective strategy to mitigate power fluctuations of electric grids with a high proportion of renewable energy sources (e.g., solar and wind).<sup>1–3</sup> As one of the most promising energy storage technologies, the redox flow battery (RFB) is attracting increasing attention due to its decoupled energy and power, flexible scalability, fast response, and high safety.<sup>4–6</sup> Among various RFB systems, the all-vanadium redox flow battery (VRFB) is the most widely studied and commercialized as it relies on vanadium ions as both positive and negative electroactive species, which significantly extends the cycle life of the battery by alleviating cross-contamination.<sup>7–9</sup> However, the high capital cost poses a

<sup>a</sup> Department of Mechanical and Aerospace Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR, China.  
E-mail: metzhao@ust.hk

<sup>b</sup> Key Laboratory of Efficient Utilization of Low and Medium Grade Energy (MOE), Tianjin University, Tianjin, China

<sup>c</sup> Department of Mechanical Engineering, University of Victoria, Victoria, BC, Canada. E-mail: ndjilali@uvic.ca

<sup>d</sup> Institute for Integrated Energy Systems, University of Victoria, Victoria, BC, Canada

<sup>e</sup> Institute for Advanced Study, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR, China

† Electronic supplementary information (ESI) available: Fig. S1–S24, Tables S1–S6, and Notes S1–S3, and the Supplementary Video. See DOI: <https://doi.org/10.1039/d1ee03224k>

‡ These authors contributed equally to this work.

major barrier to the widespread application of the VRFB.<sup>10–12</sup> An effective strategy for cost reduction is to develop VRFBs with higher power density.<sup>13–15</sup> Over the past decades, tremendous progress has been made in increasing reaction sites and improving reaction kinetics of electrodes by surface modification, such as heteroatom doping, catalyst deposition, and surface etching, which are aimed at reducing the activation loss of the battery.<sup>16–18</sup> Another important way to enhance the power density of VRFBs is to decrease the concentration loss by improving electrolyte distribution within the electrode, which can be realized by the design of flow fields.<sup>19–21</sup>

Historically, the earlier VRFBs have been constructed using the flow-through configuration, where electrolytes enter a long channel packed with a thick porous electrode.<sup>22</sup> This configuration, however, leads to an uneven distribution of electroactive species and high ohmic resistance, especially when scaling up the battery stack.<sup>23</sup> The application of flow fields in the VRFB has interested researchers since 2012. Experimentally, Aaron *et al.*<sup>24</sup> reported a VRFB with a “zero-gap” serpentine flow field. Their battery yielded a peak power density of 557 mW cm<sup>−2</sup>, much higher than the previous ones with the flow-through configuration. Subsequent studies from different groups also demonstrated the superior performance of the VRFB with serpentine flow fields to that without flow fields.<sup>25–27</sup> The dramatic improvement in the battery performance brought by flow fields stems from the two aspects: first, the battery with flow fields enables the use of thinner electrodes as compared with the flow-through configuration, which decreases the ohmic resistance of the system;<sup>28</sup> second, well-designed flow fields allow electrolytes to distribute uniformly over the electrode, which increases the local concentration of reactants and thus reduces the concentration loss.<sup>29</sup>

Generally, an ideal flow field for RFBs should maximize the uniformity of the reactant distribution over the electrode with a relatively low pressure drop.<sup>30–32</sup> To pursue this goal, some studies have been conducted to develop novel flow fields,<sup>33–36</sup> in which the spatial arrangement of flow channels is re-designed by combining expertise and intuition. One main idea is to enhance the under-rib convection, which is caused by the pressure difference between adjacent flow channels. The under-rib convection can facilitate the penetration of electrolytes into the porous electrode, which improves the transport of electroactive species within the electrode and thus alleviates the concentration loss.<sup>37</sup> When it comes to the design of flow fields, however, the limits of human intuition severely restrict the discovery of the optimal flow field, and traditional trial-and-error experimental and computational studies do not ensure optimal designs. Therefore, a systematic and pragmatic approach to the design of flow fields is highly desirable for developing high-performance RFBs. In this respect, a possible solution is topology optimization, which can optimize the layout of materials within a prescribed design domain while satisfying the required constraints.<sup>38–40</sup> For example, Yaji *et al.*<sup>41</sup> reported a two-dimensional (2-D) topology optimization method to design a flow field with the maximized generation ratio of vanadium ions. They then extended the method to a three-dimensional (3-D) level.<sup>42</sup> Their simulation results suggested that

the VRFB with the optimized flow field delivered better performance than those with conventional flow fields. However, the objective functions used in the previous studies only consider maximizing the reaction rate but ignore minimizing the pressure drop, hence not reflecting the real problem behind the design of flow fields for RFBs. Unfortunately, even though successful in mechanical applications,<sup>43</sup> multi-objective topology optimization remains challenging for solving complex systems such as RFBs where multiple physicochemical mechanisms are coupled. In addition, the density approach used in the previous studies suffers from numerical instability (mesh dependency and local minima in particular<sup>44</sup>), which may result in large computational errors. More robust and effective methods to design flow fields for RFBs are required.

Machine learning-assisted screening, on the other hand, provides a promising avenue. With the exponential growth of computing resources and improvements in simulation tools and machine learning techniques, this powerful approach has led to the discovery of novel materials and molecules across diverse applications<sup>45–49</sup> and has promoted advances in fluid mechanics.<sup>69–72</sup> Compared with traditional high-throughput screening, it benefits from acceleration in data collection and is thus particularly cost-effective for solving complex problems. However, using machine learning-assisted screening to design flow fields faces two challenges. The first is how to generate a custom library containing numerous flow field designs. Mathematically, the problem of generating flow fields can be formulated as the search for a Hamiltonian path in an undirected graph with a rectangular grid of cells, which has been proven to be NP-complete.<sup>50</sup> Similar problems have been investigated in the research area of labyrinth and maze generation.<sup>51–53</sup> Due to inherent flaws and high computational cost, algorithms reported in previous studies are difficult to traverse the solution set; ergodicity is important for screening-based design approaches. Therefore, a new algorithm that ensures the diversity of the generation result is needed. The second challenge is how to screen promising candidates from the search library. Though powerful and robust, machine learning-assisted screening remains restricted in terms of the inverse design with the goal of optimizing properties or performance objectives.<sup>54</sup> To overcome this difficulty, some studies have combined machine learning with optimization algorithms (*e.g.*, genetic algorithm) to tackle material design problems related to polymers,<sup>55</sup> battery electrodes,<sup>56</sup> and optical glass.<sup>57</sup> Such methods, however, are infeasible when the solution set of the problem to be addressed is not continuous. In this situation, an effective strategy is to use explicit criteria (*e.g.*, superior to a specific target) to trawl promising regions of the solution space and then leverage human experts to identify the most promising candidates.<sup>58</sup>

Here, we present a framework for the machine learning-assisted design of flow fields for RFBs. As illustrated in Fig. 1, we first develop an end-to-end approach to designing flow fields, encompassing library generation, multi-physics simulation, machine learning, and experimental validation, which will be presented in Sections 3.1–3.5 in order. After exploring

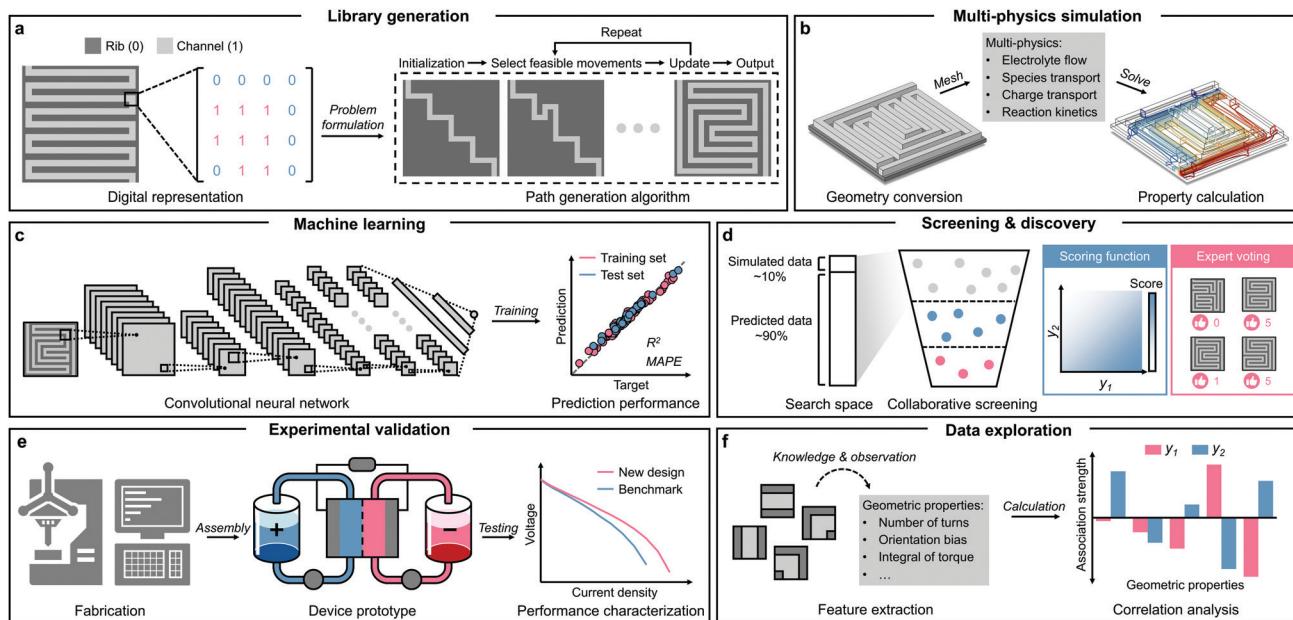


Fig. 1 Framework of the machine learning-assisted design of flow fields for redox flow batteries. In this framework, an end-to-end approach to designing flow fields is first developed, which incorporates library generation (a), multi-physics simulation (b), machine learning (c), screening and discovery (d), and experimental validation (e). After that, quantitative design rules of flow fields are revealed through data exploration (f).

11 564 flow fields, the data-driven search process identifies eight promising candidates. Three of these are fabricated and tested, and experimental results show that the battery with the newly designed flow fields exhibits about a 22% increase in limiting current density and up to 11% improvement in energy efficiency compared to the conventional serpentine flow field. In Section 3.6, a statistical analysis of geometric properties shows that the promising candidates have the saved channel length of  $1490 \pm 100$  and the torque integral of  $20.1 \pm 1.8$ , revealing the design rules of flow fields on a quantitative level for the first time.

## 2 Experimental

### 2.1 Path generation algorithm

A path generation algorithm was developed to generate a library containing as many as 11 564 flow field designs, represented by 2-D binary images. This algorithm starts with an undirected graph with a rectangular grid of cells and continuously extends a path between the inlet and the outlet until a Hamiltonian path is found. Specifically, the initialization of the generation process is shown as follows: (I) given the inputs, the number of rows and columns of the grid,  $n_1$  and  $n_2$ , and the cell size,  $\delta$ , are calculated. (II) An undirected graph  $G$  with an  $n_1 \times n_2$  rectangular grid of cells is generated (Fig. 2(a1)). For each cell, a vector is used to record if the surrounding walls exist and if the cell is already visited. (III) An initial path  $G^0$  (e.g., diagonal), which connects the inlet with the outlet, is generated (Fig. 2(a2)). If two adjacent cells are connected, the wall between them will be removed, and meanwhile, the edge will be added accordingly. After the initialization, the program enters the main loop.

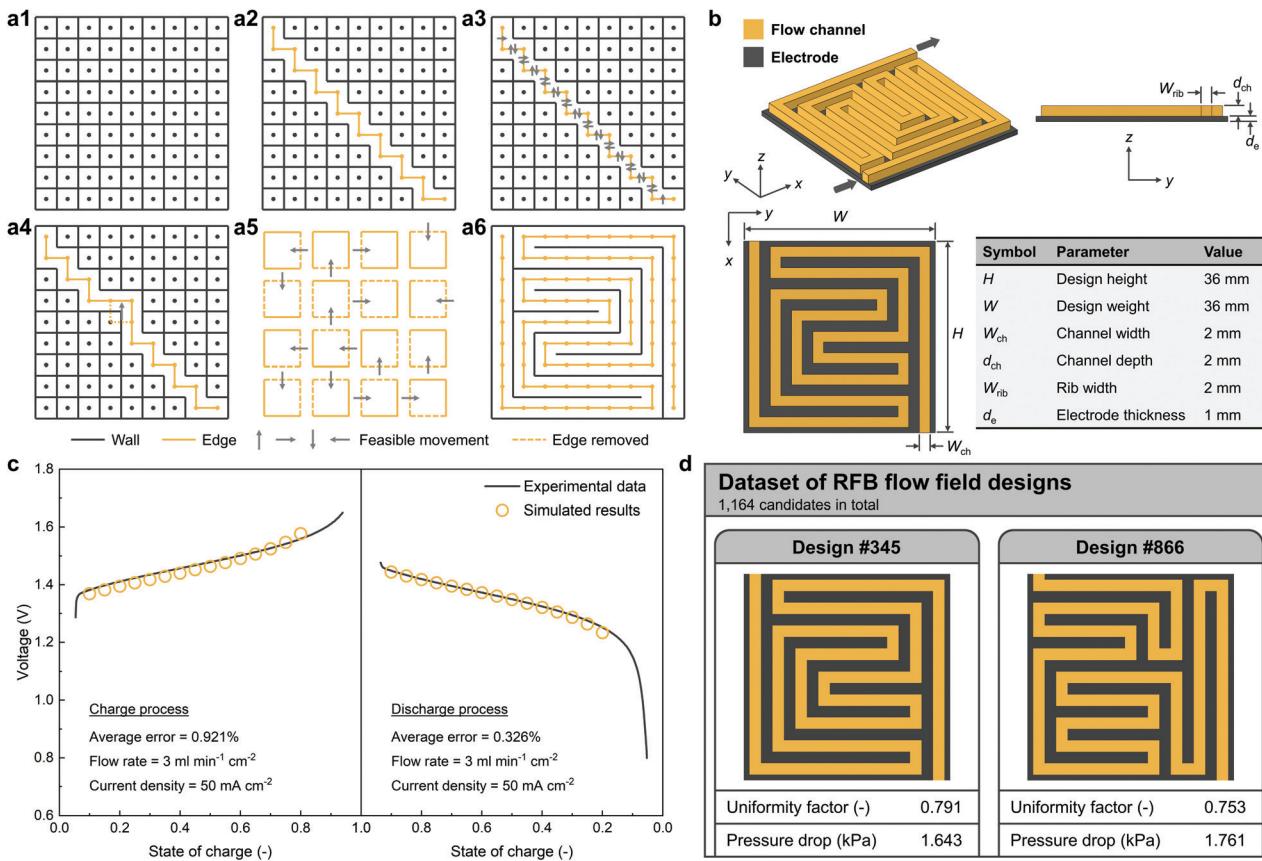
Assuming  $k - 1$  iterations have been completed, the evolution of the current path  $G^k$  includes the following steps: (IV) all the edges of the current path,  $e_1, \dots, e_p$ , are identified, and then the possible moving directions for each edge,  $\Gamma$ , are determined (Fig. 2(a3)). For a horizontal edge,  $\Gamma = [01]$  (up) or  $\Gamma = [0\bar{1}]$  (down); for a vertical edge,  $\Gamma = [\bar{1}0]$  (left) or  $\Gamma = [10]$  (right). Note that a moving direction will be considered infeasible if it makes the edge exceed the entire grid or encounter the existing path. (V) A candidate path  $G^k|_{e_i, \Gamma}$  is obtained by moving  $e_i$  of  $G^k$  along  $\Gamma$  (Fig. 2(a4)). Then the energy of the candidate path,  $E(G^k|_{e_i, \Gamma})$ , is computed. Here  $E$  is a hypothetical energy operator, which is given by:

$$E(G^k|_{e_i, \Gamma}) = \frac{N_{\text{turn}}}{(L_{\text{path}})^2} \quad (1)$$

where  $N_{\text{turn}}$  is the number of turns of the path, and  $L_{\text{path}}$  is the length of the path. (VI) The roulette wheel selection method is used to select one candidate path based on the softmax-normalized probability  $P(e_i, \Gamma)$ , which is given by:

$$P(e_i, \Gamma) = \frac{\exp(E(G^k|_{e_i, \Gamma}))}{\sum_{i, \Gamma} \exp(E(G^k|_{e_i, \Gamma}))} \quad (2)$$

Combining the energy operator and the roulette wheel selection method can generate more promising solutions while maintaining a certain degree of the diversity of solutions, which is particularly important for building a dataset for screening-based design approaches. (VII) The graph is updated, i.e.,  $G^k \leftarrow G^{k+1}$ . For one iteration, there are 16 types of possible updates in total (Fig. 2(a5)). The program will repeat steps (IV)–(VII) until all the cells have been visited or the maximum number of



**Fig. 2** Library generation and multi-physics simulation. (a) Schematic of the key steps of the path generation algorithm. (b) 3-D representation of the computational domain for the multi-physics model. The inlet and outlet are depicted by the gray arrows. All the geometric sizes are listed in the inserted table. (c) Validation of the multi-physics model by comparing the simulated charge–discharge curve against the experimental measurement (at an area-specific flow rate of  $3 \text{ ml min}^{-1} \text{ cm}^{-2}$  and a current density of  $50 \text{ mA cm}^{-2}$ ). The average error is quantified by the mean absolute percentage error. (d) Illustration of two training examples stored in the dataset. Each training example has one input variable (flow field image) and two output variables (uniformity factor and pressure drop).

iterations is reached. Finally, the graph  $G$  is filled with pixel values (*i.e.*, channel – 1 and rib – 0) to create a binary image  $M$  with a height of  $n_1 \times \delta$  and a width of  $n_2 \times \delta$  (Fig. 2(a6)). The pseudocode of the algorithm is presented in Algorithm 1. The path generation algorithm was realized using MATLAB. The library generation took about 13 days on an 8 GB-RAM computer with 2 CPUs (Intel® Core™ i5-8400, 2.8 GHz).

## 2.2 3-D multi-physics simulation

A 3-D multi-physics model was used to calculate the uniformity factor and pressure drop of 1164 flow fields in the library. The model simulated physicochemical processes occurring in the VRFB system, such as electrolyte flow, species and charge transport, and electrochemical reactions. For convenience, the following assumptions were made:<sup>8,59</sup> (I) All the physical processes were steady-state. (II) The electrolyte flow was incompressible. (III) All the material properties were isotropic and homogeneous. (IV) Side reactions (*e.g.*, hydrogen evolution) were not considered. (V) The dilute solution approximation was used. (VI) The second-step dissociation of sulfuric acid was neglected. The computational domain for the model consists of a flow channel and a porous electrode (Fig. 2(b)). The main

governing equations for the model are presented as follows, and a detailed description is given in Note S1 (ESI†).

The conservation of mass in the flow channel and the porous electrode is given by the continuity equation:

$$\nabla \cdot \bar{u} = 0 \quad (3)$$

where  $\bar{u}$  is the velocity field of the electrolyte flow.

The conservation of momentum in the flow channel and the porous electrode are described by the Navier–Stokes equation and the Brinkman equation, respectively:

$$\rho (\bar{u} \cdot \nabla) \bar{u} = -\nabla P + \mu \nabla^2 \bar{u} \quad (4)$$

$$\frac{\rho}{\varepsilon^2} (\bar{u} \cdot \nabla) \bar{u} = -\nabla P + \frac{\mu}{\varepsilon} \nabla^2 \bar{u} - \frac{\mu}{\kappa} \bar{u} \quad (5)$$

where  $\rho$  is the electrolyte density,  $P$  is the pressure field,  $\mu$  is the electrolyte viscosity,  $\varepsilon$  is the electrode porosity, and  $\kappa$  is the hydraulic permeability of the porous electrode.

The transport of electroactive species in the electrolyte can be described by the Nernst–Planck equation:

$$\nabla \cdot \left( -D_i^{\text{eff}} \nabla c_i - \frac{z_i c_i D_i^{\text{eff}}}{RT} F \nabla \phi_i + c_i \bar{u} \right) = -S_i \quad (6)$$

where the subscript  $i \in \{V^{2+}, V^{3+}, VO^{2+}, VO_2^+, H^+\}$ ,  $c_i$  is the molar concentration of the species  $i$ ,  $z_i$  is the charge number,  $S_i$  is the source term,  $\phi_i$  is the electric potential in the electrolyte,  $R$  is the universal gas constant,  $T$  is the temperature,  $F$  is the Faraday constant, and  $D_i^{\text{eff}}$  is the effective diffusion coefficient.

The source term of the species concentration is directly correlated with the local current density, which can be described by the Butler–Volmer equation:

$$i_{\text{loc,neg}} = i_{\text{neg}}^0 \left[ \exp\left(\frac{\alpha_{\text{a,neg}} F \eta_{\text{neg}}}{RT}\right) - \exp\left(-\frac{\alpha_{\text{c,neg}} F \eta_{\text{neg}}}{RT}\right) \right] \quad (7)$$

$$i_{\text{loc,pos}} = i_{\text{pos}}^0 \left[ \exp\left(\frac{\alpha_{\text{a,pos}} F \eta_{\text{pos}}}{RT}\right) - \exp\left(-\frac{\alpha_{\text{c,pos}} F \eta_{\text{pos}}}{RT}\right) \right] \quad (8)$$

where subscripts neg and pos denote the negative and positive sides respectively,  $i^0$  denotes the exchange current density,  $\alpha_a$  and  $\alpha_c$  represent the anodic and cathodic charge transfer coefficients, respectively, and  $\eta$  represents the overpotential.

The charge transport in the electrode and the electrolyte is respectively described by:

$$\nabla \cdot \bar{i}_s = -\sigma_s^{\text{eff}} \nabla^2 \phi_s = i_{\text{loc}} \quad (9)$$

$$\nabla \cdot \bar{i}_l = -\sigma_l^{\text{eff}} \nabla^2 \phi_l = -i_{\text{loc}} \quad (10)$$

where  $\bar{i}_s$  and  $\bar{i}_l$  are the local current density in the electrode and the electrolyte, respectively, and  $\sigma_s^{\text{eff}}$  and  $\sigma_l^{\text{eff}}$  are the effective electronic conductivity of the electrode and the electrolyte, respectively.

Based on the simulation results, the uniformity factor and pressure drop were calculated. As illustrated in Fig. S1a (ESI†), the uniformity factor, which quantifies the uniformity of the distribution of electroactive species within the porous electrode, was calculated by:

$$y_1 = 1 - \frac{1}{c_{i,\text{avg}}} \sqrt{\frac{1}{V_e} \iiint (c_i - c_{i,\text{avg}})^2 dV_e} \quad (11)$$

where  $y_1$  is the uniformity factor,  $c_i$  is the molar concentration of the species  $i$  within the electrode,  $c_{i,\text{avg}}$  is the average value of  $c_i$ , and  $V_e$  is the volume of the electrode. For convenience, this work only considered the uniformity factor of  $V^{2+}$  in the negative electrode. The pressure drop, on the other hand, was computed by the pressure difference between the inlet and outlet of the flow field (Fig. S1b, ESI†), as given by:

$$y_2 = \frac{1}{S_{\text{in}}} \iint P dS_{\text{in}} - \frac{1}{S_{\text{out}}} \iint P dS_{\text{out}} \quad (12)$$

where  $y_2$  is the pressure drop, and  $S_{\text{in}}$  and  $S_{\text{out}}$  denote the cross-sectional area of the inlet and the outlet, respectively (here  $S_{\text{in}} = S_{\text{out}}$ ).

The parameters of material properties and electrochemistry are listed in Table S1 (ESI†), and the parameters related to

boundary conditions are listed in Table S2 (ESI†). The 3-D multi-physics simulation was performed on the software COMSOL Multiphysics®. Fig. S2 (ESI†) shows the mesh sensitivity study of the finite element model, indicating that the mesh used in this study is sufficiently fine with relative errors of lower than 0.5%. The simulation took about a total of 36 days on a 64 GB-RAM workstation with 48 CPUs (Intel® Xeon® E5-2680, 2.5 GHz).

## Algorithm 1 path generation algorithm

**Input:** design height  $H$  (mm); design width  $W$  (mm); channel width  $W_{\text{ch}}$  (mm); open ratio  $r_o$  (–); pixel resolution  $r_p$  (mm)

**Output:** 2-D binary image  $M$

- 1: Calculate the following variables: the number of rows of the grid  $n_1 = r_o H / W_{\text{ch}}$ ; the number of columns of the grid  $n_2 = r_o W / W_{\text{ch}}$ ; cell size  $\delta = W_{\text{ch}} / r_o r_p$
- 2: Create an undirected graph  $G$  with an  $n_1 \times n_2$  rectangular grid
- 3: Generate an initial path:  $G \leftarrow G^0$
- 4: **repeat**
- 5:   Identify all the edges in the current path  $e_1, \dots, e_p$
- 6:   **for**  $i = 1$  to  $p$  **do**
- 7:     Identify the feasible moving direction  $\Gamma$  of the edge  $e_i$
- 8:     Calculate the energy of the movement:  $E(G^k|_{e_i, \Gamma}) \leftarrow N_{\text{turn}} / (L_{\text{path}})^2$
- 9:   **end for**
- 10:   Compute the probability of a movement to be selected:  $P(e_i, \Gamma) \leftarrow \exp(E(G^k|_{e_i, \Gamma})) / \sum_{i, \Gamma} \exp(E(G^k|_{e_i, \Gamma}))$
- 11:   Use the roulette wheel selection method to select a movement
- 12:   Implement the selected movement to update the path:  $G^k \leftarrow G^{k+1}$
- 13: **until** all the cells have been visited or the maximum number of iterations is reached
- 14: Fill the obtained graph  $G$  given the cell size  $\delta$  to obtain a binary image  $M$
- 15: **return** image  $M$

### 2.3 Convolutional neural networks

Convolution neural networks (CNNs) were trained to capture the functional mapping of the input variable (flow field image) to the output variable (either uniformity factor or pressure drop). As shown in Fig. S3 (ESI†), the CNNs for predicting  $y_1$  and  $y_2$  share the same architecture, which consists of 19 layers: one input layer, four convolutional layers, four batch normalization layers, four ReLU (rectified linear unit) layers, two max-pooling layers, two fully connected layers, one dropout layer, and one output layer. Table S3 (ESI†) summarizes the hyper-parameters used in these layers, and the operation of the layers is detailed in Note S2 (ESI†).

The original dataset was split into two subsets: a training set and a test set, which occupy 75% and 25% of the dataset, respectively (the two subsets contain 873 and 291 examples, respectively). As shown in Fig. S4 (ESI†), the statistical distributions of the output variables of the training and test sets are consistent, suggesting that there is no apparent bias during data splitting. The training set was used to teach the CNN model to capture the functional relationship between the input and output variables; the test set was used to evaluate the performance of the trained model. During the training process, the data were fed to the CNN in a mini-batch manner. The loss function was optimized using the stochastic gradient descent with momentum (SGDM) algorithm, in which the momentum was set as 0.9, and the weight for L2-regularization was 0.001. Initially, the learning rate was set as 0.001; after 20 epochs, it would decrease by 50%. Note that normalization was applied to the training data to improve the stability and convergence of the training process. We recorded the *R*-squared  $R^2$  and the mean absolute percentage error (MAPE) for the test set to evaluate the prediction performance of the CNNs. These metrics are computed by:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (13)$$

$$\text{MAPE} = \frac{100\%}{N} \sum_{i=1}^N \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (14)$$

where  $\hat{y}_i$  is the predicted output variable, and  $y_i$  is the actual output variable for the  $i$ -th sample in the test set, and  $\bar{y}$  is the average value of  $y_i$ , and  $N$  is the number of samples. The CNN described above was developed using the deep learning toolbox from MATLAB.

#### 2.4 Experimental test

The tested VRFB system comprises a single cell, two external electrolyte tanks, pumps, and other accessories, and the configuration of the single cell is illustrated in Fig. S5 (ESI†). HTS01, HTS04, and HTS02 flow fields and the two benchmarks (SFF and IFF) were fabricated by machining channels on the graphite plates. The ion exchange membrane was Dupont™ Nafion® 212. Both the negative and positive electrodes were 4 layers of Sigracet® 39 AA carbon paper (the nominal thickness for each layer is about 280 µm), which had been thermally treated in a muffle furnace in ambient air at 400 °C for 10 hours. During the assembly of the battery, the carbon paper electrodes were compressed to about 1000 µm.

The setup for the pressure drop test is illustrated in Fig. S6 (ESI†). The water was used as the working medium and was circulated at various volumetric flow rates of 35–75 ml min<sup>-1</sup> by a 2-channel peristaltic pump (LongerPump, BT100-1L). A pressure transmitter (Suzhou Xuansheng Meter Technology Co., Ltd, PCM950A) was used to measure the pressure difference between the inlet and outlet of the flow field plate.

The electrochemical test was carried out on a battery test instrument (Neware Technology Ltd). When testing the charge-discharge performance, the cut-off voltage of charge and discharge processes was 1.65 V and 0.80 V, respectively. Then the polarization curves were obtained starting from around 90% state of charge; the battery was discharged with increasing the current density and lasted 10 seconds at each test step until the voltage reached zero. The volume of both the negative and positive electrolyte solutions was 20 mL. The negative electrolyte solution contained 1.0 M V<sup>3+</sup> and 3.0 M H<sub>2</sub>SO<sub>4</sub>; the positive electrolyte solution contained 1.0 M VO<sup>2+</sup> and 3.0 M H<sub>2</sub>SO<sub>4</sub>. Both electrolytes were circulated by a 2-channel peristaltic pump (LongerPump, BT100-1L). In addition, nitrogen gas was used to exhaust the air in the battery setup to avoid undesirable side reactions.

#### 2.5 Geometric properties of flow fields

To analyze the design rules of flow fields, we defined five types of geometric properties to describe the morphological features of flow channels. To facilitate the definition of the geometric properties, we first introduced some basic concepts. As mentioned in the path generation algorithm, the flow field is represented by a rectangular grid of cells, and these cells fall into two categories:<sup>60</sup> “straight” and “turn” (Fig. S7, ESI†). The turn cells divide the flow channel into several sections; these sections are henceforth called straight channels. Obviously, the number of turns ( $P_1$ ) is one of the important geometric properties. Furthermore, by recording the order of positions of the turn cells, a flow channel can be shown as a vector plot (Fig. 6(a)), where the vectors (colored in gold) represent the magnitude and direction of the respective straight channels. Thus the average length and standard deviated length of straight channels are important features. However, the average length of straight channels is a redundant variable with respect to  $P_1$  since they are completely inversely proportional. In this regard, the standard deviated length of straight channels ( $P_2$ ) is regarded as another geometric property to be studied, as calculated by:

$$P_2 = \sqrt{\frac{1}{P_1} \sum_{i=1}^{P_1+1} \left( |\vec{s}_i| - \langle |\vec{s}_i| \rangle \right)^2} \quad (15)$$

where  $|\vec{s}_i|$  denotes the length of the straight channel  $\vec{s}_i$ , and  $\langle \cdot \rangle$  denotes the average value of the variable. In addition, a geometric property called orientation bias ( $P_3$ ) is defined to describe the statistical distribution of the direction of straight channels (Fig. S8, ESI†), as given by:

$$P_3 = \frac{|(L_h - \lambda_h) - (L_v - \lambda_v)|}{L - \lambda_h - \lambda_v} \quad (16)$$

where  $L_h$  and  $L_v$  denote the total length of the horizontal and vertical straight channels, respectively, and  $\lambda_h$  and  $\lambda_v$  denote the minimal length of the horizontal and vertical straight channels that are needed to connect the inlet with the outlet, respectively (here  $\lambda_h = 8$  and  $\lambda_v = 9$ ), and  $L$  represents the total length of the

flow channel (here  $L = 81$ ). With this definition, the value of  $P_3$  varies from 0 to 1, where the magnitude indicates how significant the difference between the length of horizontal and vertical straight channels is.

In addition to the three basic properties mentioned above, we proposed two high-level geometric properties. One is called the length of saved flow channels ( $P_4$ ), which is specifically defined as the total length of flow channels saved by the way that electrolytes move to the neighboring flow channels by under-rib convection (Fig. S9, ESI†), as given by:

$$P_4 = \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n (d_{i,j;\leftarrow} + d_{i,j;\uparrow} + d_{i,j;\rightarrow} + d_{i,j;\downarrow}) \quad (17)$$

where  $d_{i,j;\leftarrow}$ ,  $d_{i,j;\uparrow}$ ,  $d_{i,j;\rightarrow}$ , and  $d_{i,j;\downarrow}$  are the saved channel length when electrolytes at the cell  $[i,j]$  moves to the left, upward, right, and downward neighbors, respectively, and the weighting factor 1/2 is used to deduct repeated calculation. Here  $P_4$  is proposed to roughly describe the strength of under-rib convection, as under-rib convection is essentially caused by the pressure difference between adjacent flow channels, which is largely dependent on their distance along the channel. The other high-level geometric property to be studied is called the torque integral ( $P_5$ ). Here the torque is defined as the cross product of two neighboring straight channels (*i.e.*,  $\vec{T}_i = \vec{s}_i \times \vec{s}_{i+1}$ ). The torque for all the turns can be plotted as a function of the relative channel length (the ratio of the length of the passed flow channel to the total length of the flow channel). With linear interpolation,  $P_5$  is calculated by the area under the torque distribution curve (Fig. S10, ESI†), as given by:

$$P_5 = \frac{1}{n} \sum_{i=1}^n T\left(\frac{i}{n}\right) \quad (18)$$

where  $T(x)$  denotes the interpolated torque distribution function, and  $n$  denotes the number of intervals (here  $n = 162$ ). In this study,  $P_5$  is proposed to integrate the three basic properties to include the statistical features of the number, magnitude, and orientation of the straight channels into a single property. It is noted that, among these geometric properties,  $P_1$  and  $P_3$  are dimensionless, and  $P_2$  and  $P_4$  are in unit lattice length, and  $P_5$  is in square of unit lattice length. The geometric properties described above were calculated using MATLAB.

## 3 Results and discussion

### 3.1 Library generation

The library generation process is crucial because it delimits the possible outcomes from the beginning, and successful candidates can be selected only from the initial search library.<sup>61</sup> One important attribute when generating a search library is to select a suitable descriptor for the input variable to be studied. In this work, a good digital representation of a flow field is provided by a 2-D binary image where the flow channel and the rib are distinguished by different pixel values (*e.g.*, channel – 1 and rib – 0). Using 2-D binary images to represent flow fields is visually

intuitive and, more importantly, enables the use of powerful deep learning models in subsequent data learning.

To generate a custom library of flow field designs, we developed a path generation algorithm (see Section 2.1). This algorithm, inspired by ink diffusion, starts with an initial path in a rectangular grid graph and then gradually extends it until a Hamiltonian path has been found in the grid (Fig. 2(a)). In particular, the evolution of the path is driven by a hypothetical energy operator and a roulette wheel selection method, which add randomness to the generation process and thus ensure the diversity of the obtained solutions. The pseudocode of the algorithm is presented in Algorithm 1 (Section 2.1), and a video recording of the generation process of a flow field can be found in the Video (ESI†). To generate flow fields that meet various dimensional needs, the inputs to our algorithm include design height  $H$ , design width  $W$ , channel width  $W_{ch}$ , open ratio  $r_o$ , and pixel resolution  $r_p$ . The output of the algorithm is a 2-D binary image  $M$ , which represents a flow field design.

In this study, both  $H$  and  $W$  were set as 36 mm, which is in the range of typical laboratory-scale RFBs.<sup>62</sup> Given this design size,  $W_{ch}$  was set as 2 mm. In addition,  $r_o$ , which represents the ratio of the area of the channel region to the total area of the flow field, was set as 0.5, which is a typical value for flow fields in RFBs. Furthermore,  $r_p$  was set as 1 mm. With these inputs, the number of rows and columns of the grid ( $n_1$  and  $n_2$ ) were calculated as 9, and the cell size ( $\delta$ ) was 4, and hence  $M$  has a dimension of  $36 \times 36$ . Besides, following the design convention of flow fields in RFBs, the top-left and bottom-right corners were labeled as the inlet and outlet, respectively. By using the developed algorithm, we generated a total of 11 564 flow field designs. Note that it is theoretically possible that the path generation algorithm could traverse the entire solution set of the problem, but to avoid excessively high computational cost of the exhaustive search, we terminated the program when the distributions of the first two principal components of the collected subset were converged (Fig. S11, ESI†).

### 3.2 Multi-physics simulation

The uniformity factor and the pressure drop are widely used criteria for evaluating the performance of flow fields for RFBs.<sup>63–65</sup> Here, the uniformity factor and pressure drop of the flow fields were calculated using a 3-D multi-physics model (see Section 2.2 and Note S1 of ESI†), which simulates the coupled physico-chemical phenomena occurring in the VRFB, including electrolyte flow, species transport, charge transport, and electrochemical reactions.

As depicted in Fig. 2(b), the computational domain for the model consists of a flow channel and a porous electrode. As shown in Fig. S12 (ESI†), the geometry of the flow channel is obtained by extruding the “channel” subgraph of the 2-D binary image by a height of 2 mm (*i.e.*, the channel depth is 2 mm), and the geometry of the porous electrode is represented by a  $36 \times 36 \times 1$  mm<sup>3</sup> cuboid (*i.e.*, the electrode thickness is 1 mm). To validate the reliability of the multi-physics model built in this study, Fig. 2(c) compares the simulated charge-discharge curve against the experimental measurement (at an area-specific

flow rate of  $3 \text{ ml min}^{-1} \text{ cm}^{-2}$  and a current density of  $50 \text{ mA cm}^{-2}$ ). Results show that the simulation results agree very well with the experimental data with an average error of lower than 1%, indicating that the calculation results of the model can be regarded as “ground truth” when training machine learning models.

With the multi-physics model, the uniformity factor and pressure drop of 1164 flow fields randomly selected from the library were calculated at an area-specific flow rate of  $3 \text{ ml min}^{-1} \text{ cm}^{-2}$  and a discharge current density of  $50 \text{ mA cm}^{-2}$ . Fig. S13 (ESI†) compares the statistical distribution of the first two principal components of these 1164 flow field images and all the 11 564 samples, indicating that the simulated dataset is sufficiently representative of the search library. After that, the 1164 flow fields (*i.e.*, input variable) and the corresponding uniformity factor and pressure drop (*i.e.*, output variables) were compiled into a dataset, as illustrated in Fig. 2(d). To facilitate the presentation of the subsequent analysis and results, we will use  $y_1$  and  $y_2$  to represent the uniformity factor and pressure drop, respectively.

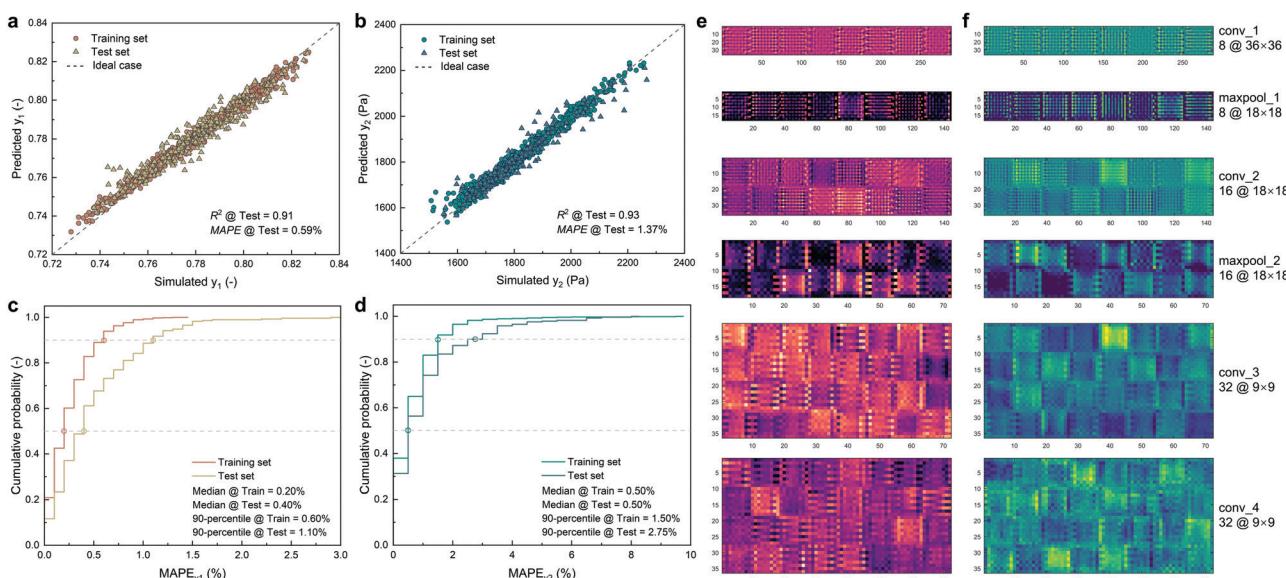
### 3.3 Machine learning

In this study, the objective of machine learning was to learn a mapping from the input variable to the output variable. The well-trained machine learning models were then used to make rapid and accurate predictions for the cases unseen in the training process (also known as the regression task of supervised learning), which accelerated the discovery of promising candidates considerably from a vast solution space.

Generally, the predictive power of machine learning is significantly affected by the selection of mathematical models

and model hyper-parameters. Among machine learning regression models, we used the convolutional neural network (CNN) because the convolution operation makes it more suitable for analyzing spatial data (*e.g.*, images). The effect of hyper-parameters (learning rate and mini-batch size) on the prediction performance of the CNN is presented in Fig. S14 (ESI†). It is found that the model loss for  $y_1$  achieves the lowest value when the learning rate is 0.001 and the mini-batch size is 8, and the preferable hyper-parameters for  $y_2$  are 0.001 and 16, respectively. In addition, Fig. S15 (ESI†) shows the effect of the amount of training data on the prediction performance. Results show that, with increasing training data, the model loss for either  $y_1$  or  $y_2$  decreases rapidly and gradually keeps stable, proving that the volume of data collected in this study is sufficient to train CNNs with the optimal performance.

After training the CNN, we evaluated its performance by recording the  $R$ -squared ( $R^2$ ) and the mean absolute percentage error (MAPE) for the test set.  $R^2$  and MAPE are widely used metrics in training machine learning models.<sup>66–68</sup> Fig. 3(a) and (b) compare the CNN-predicted  $y_1$  and  $y_2$  against the corresponding ground truth, respectively. Results corroborate that the uniformity factor ( $R^2 = 0.91$  and MAPE = 0.59%) and pressure drop ( $R^2 = 0.93$  and MAPE = 1.37%) of a flow field can be accurately predicted by simply using the 2-D binary image as the input. Fig. 3(c) and (d) display the cumulative probability plot of the MAPE for predicting  $y_1$  and  $y_2$ , respectively (the corresponding histogram plots are given in Fig. S16 of ESI†). These results indicate that for both  $y_1$  and  $y_2$ , the probability distribution of the prediction error on the training



**Fig. 3** Machine learning. (a) Comparison between the CNN-predicted  $y_1$  and the ground truth. (b) Comparison between the CNN-predicted  $y_2$  and the ground truth. The black dashed line in (a) and (b) represents the situation where the CNN reaches the accuracy of 100%. (c) Cumulative probability plots of the prediction error for  $y_1$  on the training and test sets. (d) Cumulative probability plots of the prediction error for  $y_2$  on the training and test sets. The gray dashed lines in (c) and (d) highlight the median and 90-percentile of the prediction error. (e) Feature maps of the CNN trained for predicting  $y_1$ . (f) Feature maps of the CNN trained for predicting  $y_2$ . In (e) and (f), “conv” and “maxpool” stand for convolutional and max-pooling layers, respectively, and the number behind the terms denotes the order of the layer.

and test sets is broadly similar, proving that there is no obvious overfitting during the model training.

Furthermore, to analyze the working mechanism of the well-trained CNNs, we visualized the networks by plotting the feature maps of the convolutional and max-pooling layers, as shown in Fig. 3(e) and (f). Here a serpentine flow field was taken as the input for demonstration purposes. Results suggest that the filters in the initial layers (*e.g.*, conv\_1 and conv\_2) act as primitive edge detectors responding to the outline of the flow channel in the input image. However, in deep layers (*e.g.*, conv\_3 and conv\_4) the filters extracted by the filters become more visually less interpretable, primarily because the filters respond to the features that are not present in the input image. In addition, the CNNs were visualized using activation maximization (Fig. S17, ESI†), which synthesizes an input image pattern by maximizing the activation of a specific neuron. Results show that most of the filters in the initial layers behave like Gabor filters that are sensitive to information about edges, shape, and orientation, in agreement with the observation on the feature maps. Interestingly, other filters in the initial layers (*e.g.*, conv\_1–4 in Fig. S17a, ESI†) seem to respond to the central region of the input image, indicating that the flow field in the central region is probably the key factor in determining the uniformity factor and pressure drop. It is also found that from the initial layers to the deep layers, the complexity and variation of the visualized image patterns are increased, suggesting that invariant features are gradually learned by the neurons. To some extent, the results of CNN visualization also reveal that it is less rewarding to directly map the input variable to the output variables when exploring the design rules of flow fields.

### 3.4 Screening and discovery

The well-trained CNNs were applied to the 10 400 unexplored flow fields in the library to predict their uniformity factor and pressure drop. In contrast with the time-consuming multi-physics simulation, this process was completed in less than 1 second, highlighting the tremendous potential of machine learning in accelerating high-throughput screening. To identify promising candidates from the pool of 11 564 flow field designs, we performed a two-round screening to gradually reduce the search space by over three orders of magnitude, as depicted in Fig. 4(a).

The first-round screening was based on a heuristic rule that selects flow fields with a higher uniformity factor ( $y_1$ ) and a lower pressure drop ( $y_2$ ). As such, we introduced a scoring function that linearly combines these two objectives, as described by:

$$S_{\text{hr}} = w y_1^* - (1 - w) y_2^* \quad (19)$$

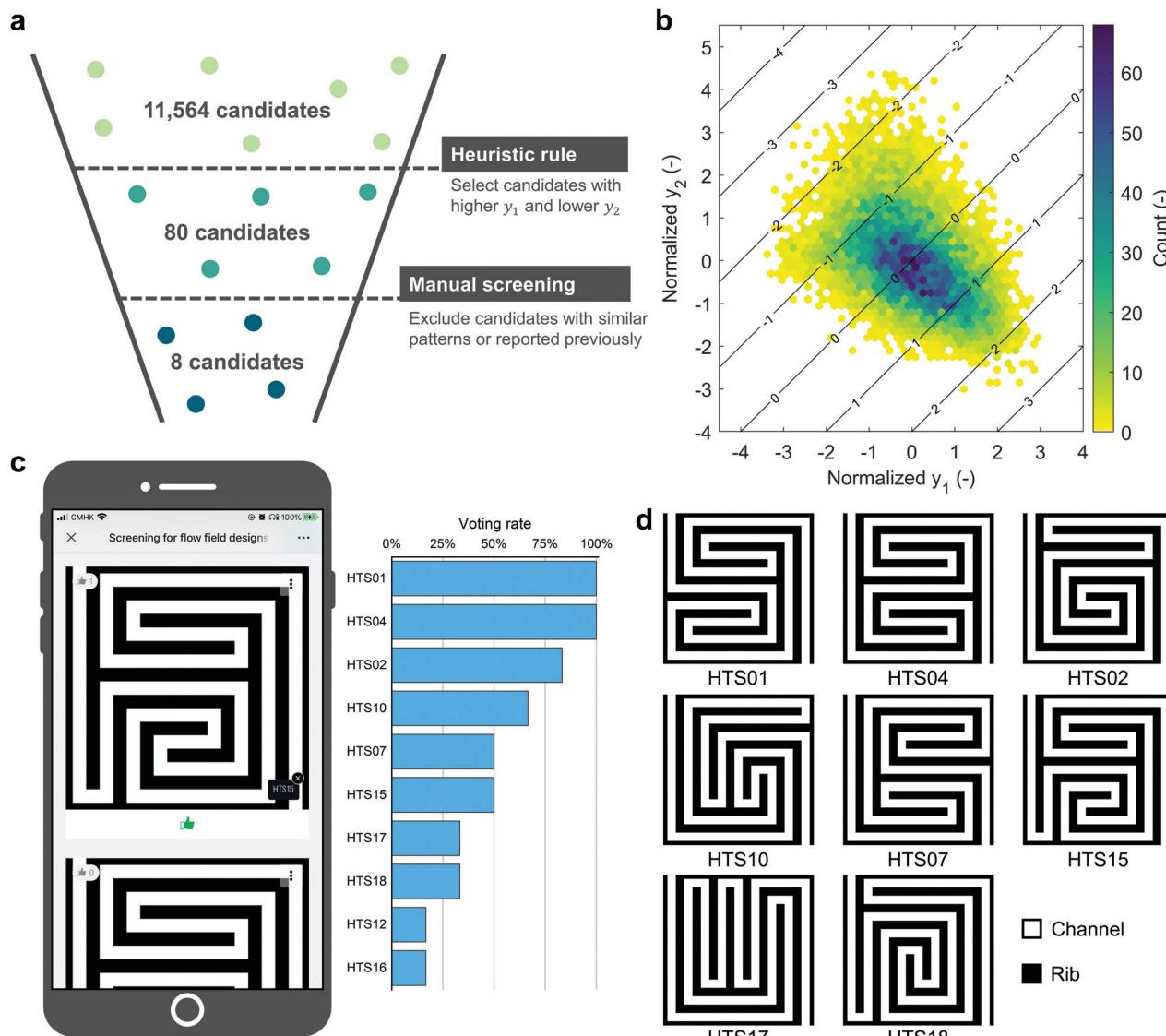
where  $S_{\text{hr}}$  is the score,  $w$  is the weighting factor ranging from 0 to 1, and  $y_1^*$  and  $y_2^*$  are the normalized output variables (normalization was used to eliminate variations in range of variables with different physical units). Note that although  $w$  has a certain influence on the screening results, it fails to determine the exact value because it remains unclear which of

the uniformity factor or the pressure drop contributes more to the design of flow fields. In this study,  $w$  was simply set as 0.5, which, though somewhat arbitrary, is adequate for obtaining satisfactory results (Fig. S18, ESI†). As shown in Fig. 4(b), the output variables of the 11 564 candidates exhibit a triangle-like distribution, where the Pareto frontiers indicate an apparent fundamental limit to the balance between the uniformity factor and pressure drop of flow fields for RFBs. In addition, the contour plot of the calculated score clearly shows that the ones at the lower right corner have the higher scores and thus are of the greatest importance. From the distribution map, 80 candidates with the highest scores were screened, and they were renamed according to their rankings (*e.g.*, the no. 1 candidate was renamed as “HTS01”). The full list of the 80 candidates is available *via* this link.<sup>73</sup>

After the first-round screening, the original search space ( $\sim 10^4$ ) was dramatically reduced to a human-tractable range ( $\sim 10^2$ ). Then human experts are leveraged to further review the pre-screened candidates through a web voting tool, including data visualization and sorting interfaces (Fig. 4(c)). A group of 6 experts rates the 80 flow fields on a positive/negative scale. For demonstration purposes, Fig. S19 (ESI†) presents the top 40 candidates selected from the first round. An interesting finding is that some of these candidates have a similar pattern, as highlighted in the same color. For example, HTS01 is quite similar to HTS03, HTS05, HTS06, and HTS11, despite slight differences in a partial area (Fig. S20, ESI†). In a high-dimensional solution space, these similar ones can be regarded as falling into the neighborhood of a local optimum (Fig. S21, ESI†). To ensure the diversity of our optimal solutions, however, repeating similar patterns should be avoided. From similar candidates, experts are therefore asked only to select the one that has earned the highest score in the first-round screening (a higher score means a higher  $y_1$  and a lower  $y_2$ ). For example, from HTS01, HTS03, HTS05, HTS06, and HTS11, only HTS01 will be selected (the candidates are named according to their rankings in the first round). In addition, to ensure the novelty of optimal solutions, the experts should exclude flow field designs reported previously. It is worth mentioning that in the research area of material design, many studies (as listed in Table S4 of ESI†) rely on human experts to refine screening results and discover promising candidates, because screening criteria are sometimes challenging to compute explicitly. Finally, eight candidates with the highest number of votes were selected as the optimal solutions to the design problem of flow fields, the binary images of which are displayed in Fig. 4(d).

### 3.5 Experimental validation

To validate the effectiveness of the design approach presented in this study, 3 representative flow fields with the highest number of votes (*i.e.*, HTS01, HTS04, and HTS02 in order) were fabricated and then tested experimentally in real redox flow batteries (see Section 2.4). In addition, the HTS18 flow field with the lower number of votes was also studied, as shown in Fig. S22 (ESI†). Here a serpentine flow field (SFF) and an interdigitated flow field (IFF) were fabricated as well and



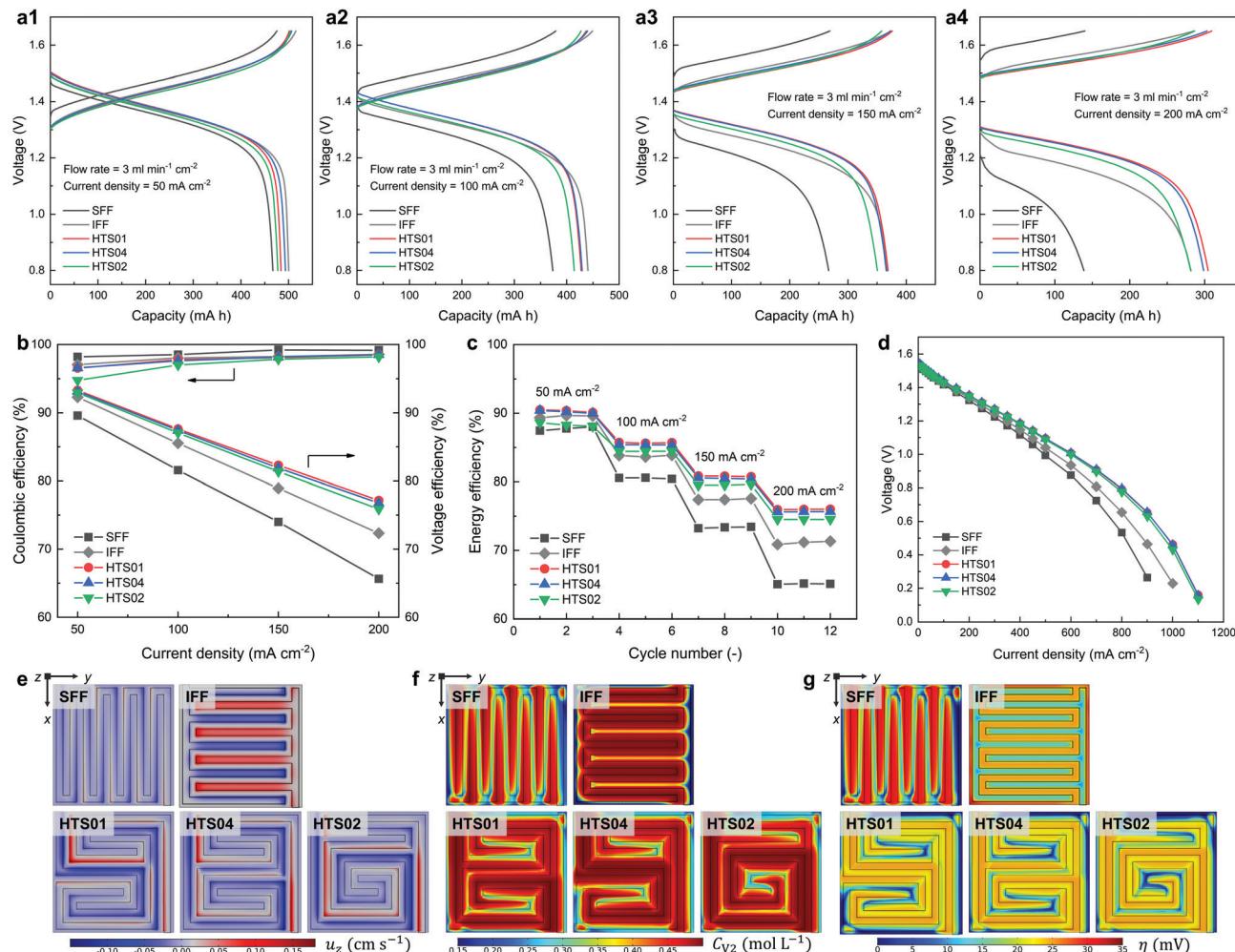
**Fig. 4** Screening and discovery. (a) Diagram of the two-round screening process: the search space decreases by over three orders of magnitude as the screening progresses. (b) Scatter plot for  $y_1$  with  $y_2$  of the 11 564 candidates. The contour plot illustrates the distribution of the calculated scores. (c) Screenshot of the interactive web voting tool (left), and the bar plot showing the voting results (right). (d) Images of the 8 promising flow fields selected from the screening process. The channel and rib regions are colored in white and black, respectively.

considered as benchmarks for evaluating the newly designed flow fields.

Fig. S23 (ESI†) compares the pressure drop of electrolyte flow in the RFB with different flow fields (at inlet flow rates from 35 to 75  $\text{mL min}^{-1}$ ). It is found that all the three new flow fields yield comparable pressure drop with the IFF while exhibiting broadly lower pressure drop as compared with the SFF. Particularly, with increasing the flow rate, the difference between the SFF and the other four flow fields becomes increasingly significant, primarily because the pressure loss of the flow channel is positively associated with the local electrolyte velocity.

Fig. 5(a) shows the charge–discharge curves of the VRFB with different flow fields (at current densities of 50, 100, 150, and 200  $\text{mA cm}^{-2}$  and a flow rate of 3  $\text{ml min}^{-1} \text{cm}^{-2}$ ). Results suggest

that the battery with the new flow fields has a comparable electrolyte utilization than that with the IFF while having lower charge and discharge overpotentials and a higher capacity than that with the SFF. At the current density of 100  $\text{mA cm}^{-2}$ , for example, the battery with the HTS04 flow field yields an electrolyte utilization of 78.3%, 10.8% higher than that with the SFF, indicating the important role of designing flow fields in reducing the use of precious electroactive species and the capital cost of VRFBs. The coulombic, voltage, and energy efficiencies of the VRFB with different flow fields are summarized in Fig. 5(b) and (c). Results show that the coulombic efficiency of the battery with all the five flow fields is higher than 95% (Fig. 5(b)), demonstrating the reduced side reaction and low crossover rate of our battery setup. Note that the coulombic efficiency of the battery with the SFF is slightly



**Fig. 5** Experimental validation. (a) Charge–discharge curves of the battery with different flow fields (SFF, IFF, HTS01, HTS04, and HTS02) at an area-specific flow rate of  $3 \text{ ml min}^{-1} \text{ cm}^{-2}$  and current densities of  $50 \text{ mA cm}^{-2}$  (a1),  $100 \text{ mA cm}^{-2}$  (a2),  $150 \text{ mA cm}^{-2}$  (a3), and  $200 \text{ mA cm}^{-2}$  (a4). (b) Coulombic efficiency and voltage efficiency of the battery with different flow fields as a function of current density. (c) Energy efficiency of the battery with different flow fields as a function of charge–discharge cycle numbers. (d) Polarization curve of the battery with different flow fields at an area-specific flow rate of  $3 \text{ ml min}^{-1} \text{ cm}^{-2}$ . (e)–(g) Simulated distribution of the z-component electrolyte velocity ( $u_z$ ) (e),  $V^{2+}$  concentration ( $C_{V2}$ ) (f), and overpotential ( $\eta$ ) (g) across an xy-plane (near the channel/electrode interface) inside the porous electrode.

higher than that of the battery with the other four flow fields due to its decreased charge–discharge duration. Moreover, consistent with the charge–discharge curves, the battery with the new flow fields delivers a higher voltage efficiency than that with the two benchmarks (Fig. 5(b)). Additionally, the battery with all the flow fields yields a stable energy efficiency as the cycle number increases and the current density varies (Fig. 5(c)). Since the energy efficiency is the product of the coulombic efficiency and the voltage efficiency, the battery with the new flow fields shows a higher energy efficiency as well. For example, at the current density of 50, 100, 150, and 200 mA cm<sup>-2</sup>, the battery with the HTS01 flow field exhibits an energy efficiency of 90.1%, 85.7%, 80.7%, and 76.0%, which is 2.7%, 5.3%, 7.3%, and 10.9% higher than that with the SFF and 0.5%, 1.85%, 3.18%, 4.69% higher than that with the IFF, respectively. Fig. 5(d) compares the polarization curves of the VRFB with different flow fields (at a flow rate of  $3 \text{ mL min}^{-1} \text{ cm}^{-2}$ ), suggesting that the

battery with the new flow fields yields lower polarization losses than that with the SFF and IFF, in agreement with the results of the charge–discharge test. Furthermore, it is shown that using the new flow fields brings about up to 22% increase in limiting current density, from 900 to 1100 mA cm<sup>-2</sup>, indicating a considerable improvement in the transport of active species within porous electrodes.

To further explain the above experimental results, the multi-physics simulation described in Section 3.2 was performed. Essentially, the superior performance of the battery with the new flow fields can be attributed to improvement in mass transfer and the resulting reduction in concentration polarization losses. As shown in Fig. 5(e), the new flow fields yield higher z-component electrolyte velocity ( $u_z$ ) within the electrode, indicating the stronger under-rib convection. For example, the maximal  $u_z$  across the HTS01 flow field reaches  $0.17 \text{ cm s}^{-1}$ , higher than the SFF ( $0.04 \text{ cm s}^{-1}$ ) and the IFF ( $0.16 \text{ cm s}^{-1}$ ).

As a result, using the new flow fields can distribute electroactive species more uniformly throughout the electrode (Fig. 5(f)), which is corroborated by the higher uniformity factor of the new flow fields (0.821–0.827) than the SFF (0.731) and the IFF (0.818). Since the improved species transport increases the local concentration of reactants within the electrode, the concentration loss as well as the total overpotential can be significantly decreased (Fig. 5(g)). For example, the battery with the HTS01 flow field suffers from an overpotential of 25.5 mV, which is 20% lower than that with the SFF (31.9 mV) and slightly lower than the IFF (26.8 mV).

### 3.6 Data exploration

The superior performance of the new flow fields has been validated experimentally, but the underlying characteristics shared by the promising candidates and that account for their success remain unclear. Uncovering these characteristics is

critical for guiding the design of flow fields. In this regard, direct mapping of the flow field image to the output variables has been found of limited values and fails to provide much physically meaningful information, as discussed in Section 3.3. Instead, we defined five types of geometric properties that represent the morphological features of flow channels, including the number of turns  $P_1$ , the standard deviation of straight channels  $P_2$ , the orientation bias  $P_3$ , the length of saved flow channels  $P_4$ , and the torque integral  $P_5$  (see Section 2.5). Based on their computational complexity, these geometric properties can be divided into basic ( $P_1$ ,  $P_2$  and  $P_3$ ) and high-level properties ( $P_4$  and  $P_5$ ), as shown in Fig. 6(a).

Fig. 6(b)–(f) display the scatter plot of the geometric properties with the output variables for 1164 flow fields in the original dataset, where three representative data points are presented as well to facilitate the explanation for the effect of the corresponding geometric property on the morphology of the flow channel.

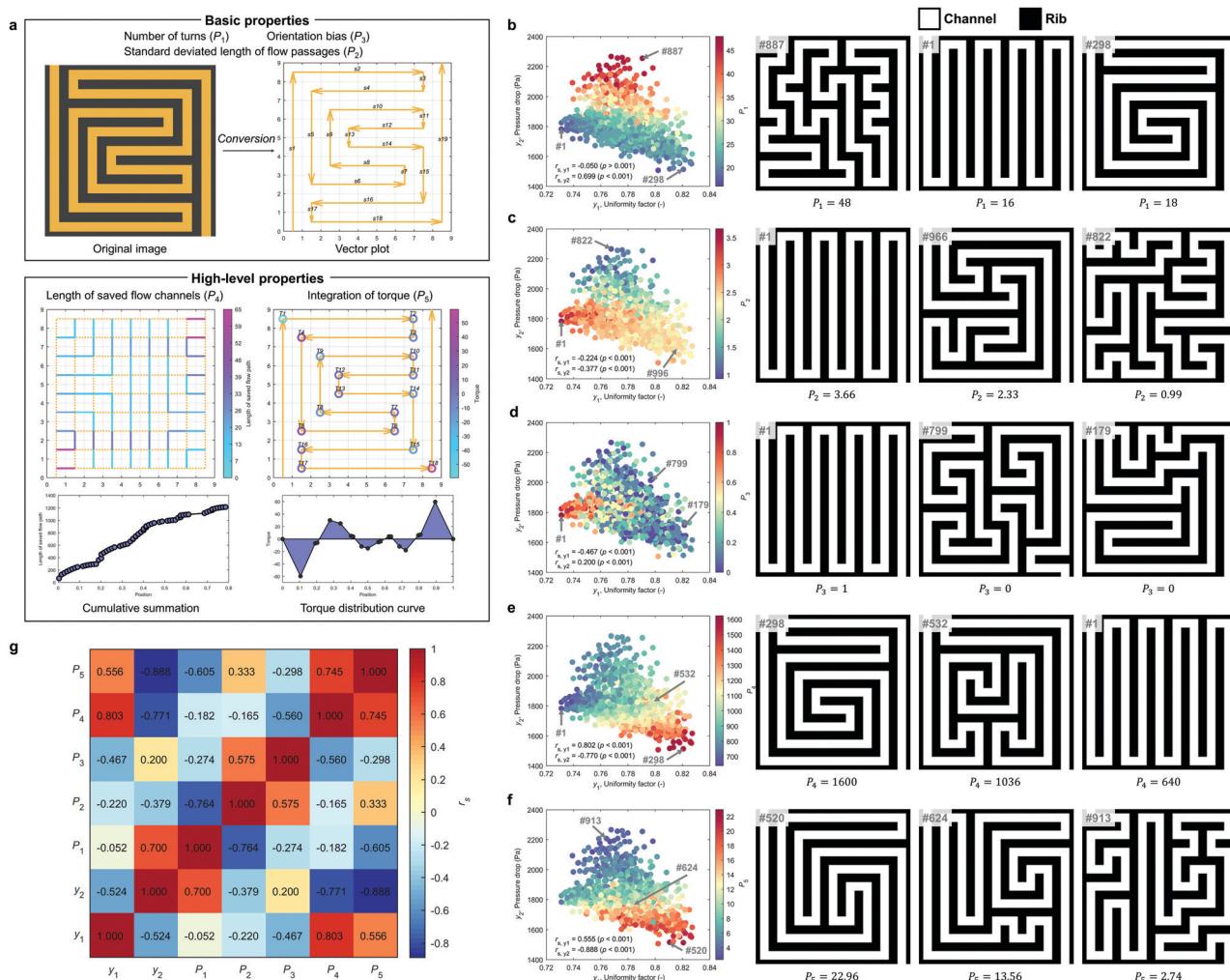


Fig. 6 Data exploration. (a) Three basic geometric properties: number of turns ( $P_1$ ), standard deviated length of straight channels ( $P_2$ ), and orientation bias ( $P_3$ ), and two high-level geometric properties: length of saved flow channels ( $P_4$ ), and torque integral ( $P_5$ ). (b)–(f) Scatter plot of  $y_1$  and  $y_2$  showing the distribution of the geometric property  $P_1$  (b),  $P_2$  (c),  $P_3$  (d),  $P_4$  (e), and  $P_5$  (f) in the original dataset of 1164 samples. From the scatter plot, three representative flow fields (channel and rib regions are colored in white and black, respectively) are selected to demonstrate the effect of the corresponding geometric property on the morphology of the flow channel. (g) Correlation matrix of the Spearman's rank correlation coefficient ( $r_s$ ) between every two variables.

**Table 1** Comparison of the geometric properties of the 8 promising candidates identified in this study with those of the 1164 flow fields in the original dataset. In the table, SD and IQR stand for standard deviation and interquartile range, respectively, and  $\bar{L}$  denotes the unit lattice length

	Promising candidates			Original dataset		
	Mean $\pm$ SD	Range	IQR	Mean $\pm$ SD	Range	IQR
$P_1 (-)$	$18 \pm 1$	16–20	18–18	$26 \pm 7$	16–48	22–30
$P_2 (\bar{L})$	$2.45 \pm 0.16$	2.28–2.79	2.36–2.51	$2.34 \pm 0.42$	0.93–3.66	2.17–2.58
$P_3 (-)$	$0.23 \pm 0.18$	0.00–0.63	0.13–0.25	$0.28 \pm 0.22$	0.00–1.00	0.13–0.38
$P_4 (\bar{L})$	$1490 \pm 100$	1360–1620	1400–1550	$1020 \pm 200$	640–1620	872–1140
$P_5 (\bar{L}^2)$	$20.1 \pm 1.8$	17.4–23.0	19.2–20.7	$11.2 \pm 4.6$	2.4–23.0	7.3–15.2

Moreover, Spearman's rank correlation coefficient  $r_s$  and  $p$ -value were calculated to quantify the association strength and statistical dependence between the observed variables, respectively (see Note S3 of ESI†). From the results, we can infer that  $P_4$  and  $P_5$  have the greatest impact on  $y_1$  ( $r_s = 0.802$ ) and  $y_2$  ( $r_s = 0.888$ ), respectively, suggesting that these geometric properties can provide a good measure of the intricate morphological features of flow channels. The results suggest that an ideal flow field should have low  $P_3$ , high  $P_4$ , and high  $P_5$  to maximize  $y_1$ , while low  $P_1$ , moderate  $P_2$ , high  $P_4$ , and high  $P_5$  to minimize  $y_2$ , indicating the qualitative design rules of flow fields. Furthermore, in terms of morphology, a good flow field design should be achieved by satisfying the following three features. First, the flow channel should have fewer turns (Fig. 6(b), (c) and (f)), which is beneficial for decreasing the pressure drop of electrolyte flow. Second, the straight sections of the channel should show a less aligned arrangement (Fig. 6(d)) to prevent preferential electrolyte distribution along a specific direction and to promote a relatively uniform reactant distribution. Third, the channel should have a rotating pattern (Fig. 6(e) and (f)), which increases the channel length between adjacent channel regions and thus contributes to stronger under-rib convection. Fig. 6(g) displays the correlation matrix of  $r_s$  among the geometric properties and the output variables (the correlation matrix of  $p$  is given in Table S5 of ESI†), indicating that some of the geometric properties are also strongly associated with each other.  $P_1$  and  $P_2$ , for example, are strongly negatively correlated with an association strength of  $-0.764$ . This is because increasing the number of turns decreases the average length of straight channels, thus flattening the variance of the length of straight channels.

When using the defined geometric properties as criteria for assessing flow fields, we can investigate the design rules of flow fields on a quantitative level. Table 1 compares the geometric properties of the eight promising candidates identified in this study with those of the 1164 ones in the original dataset. The table summarizes the overall pattern of the data by listing the mean, standard deviation (SD), range, and interquartile range (IQR). Results suggest that the eight promising candidates have lower  $P_1$  ( $18 \pm 1$ ), more moderate  $P_2$  ( $2.45 \pm 0.16$ ), lower  $P_3$  ( $0.23 \pm 0.18$ ), higher  $P_4$  ( $1490 \pm 100$ ), and higher  $P_5$  ( $20.1 \pm 1.8$ ), compared with the original dataset. Thus, taking the eight promising candidates as the optimal solutions to the design problem of flow fields and analyzing their geometric properties have allowed us to reveal the quantitative design rules of flow

fields for RFBs. To the authors' best knowledge, such design rules have not been reported previously. To demonstrate the effectiveness of the design rules, Table S6 (ESI†) compares the geometric properties of the SFF with the design rules. It is found that the SFF has inferior properties:  $P_2$  (3.66),  $P_3$  (1.00),  $P_4$  (640), and  $P_5$  (4.45), indicating that the SFF is highly likely to exhibit both a lower uniformity factor and a higher pressure drop, in agreement with the experimental results.

## 4 Conclusions

In this study, we have developed an end-to-end approach to the design of flow fields for RFBs. A search library of 11 564 flow field designs has been generated by combining a custom-made path generation algorithm, hundreds and thousands of multi-physics simulations, and well-trained CNN regression models. Through a collaborative screening process, eight promising candidates highly different from currently known flow fields for RFBs have been successfully identified. Experimental results have shown that the battery with the newly designed flow fields exhibits around a 22% increase in limiting current density and up to 11% improvement in energy efficiency compared to the conventional SFF. Furthermore, to explore the design rules of flow fields, five types of geometric properties have been proposed to describe the morphological features of flow channels. They are the number of turns ( $P_1$ ), the standard deviated length of straight channels ( $P_2$ ), the orientation bias ( $P_3$ ), the length of saved flow channels ( $P_4$ ), and the torque integral ( $P_5$ ). The identification of the geometric properties shared among the eight lead candidates has revealed the quantitative design rules of flow fields: low  $P_1$  ( $18 \pm 1$ ), moderate  $P_2$  ( $2.45 \pm 0.16$ ), low  $P_3$  ( $0.23 \pm 0.18$ ), high  $P_4$  ( $1490 \pm 100$ ), and high  $P_5$  ( $20.1 \pm 1.8$ ). The discovery of the design rules can contribute to a deeper understanding of the problem of designing flow fields for RFBs. Furthermore, since the geometric properties are easy to compute, the design rules are believed useful for future research and development.

It must be stressed that the methodology developed in this study has broad generality in terms of three aspects. First, this study considers RFB flow fields with a single channel but can readily be extended to flow fields with two or more channels after modifying the path generation algorithm, which will be the focus of our future work. Second, although this study is demonstrated with a laboratory-scale RFB, our design approach

will transfer well to scaled-up systems if an appropriate scaling-up method is included in the path generation algorithm (an example is given in Fig. S24 of ESI†). Third, in addition to the application in designing flow fields for RFBs, the methodology is applicable to the investigation and optimization of flow fields in other devices, such as flow fields of fuel cells and cooling plates of lithium-ion battery stacks, as long as reliable simulation tools, accumulated design intuition, and convenient device fabrication are available.

## Author contributions

T. Z., S. W. and H. J. conceived the project; S. W. and C. H. developed the path generation algorithm; S. W. and Z. G. performed the multi-physics simulation; S. W. and X. L. trained the CNN model; H. J. and S. W. conducted the experimental validation. S. W., H. J., N. D. and T. Z. prepared the manuscript. All authors discussed the results and commented on the manuscript.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

The work described in this paper was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. T23-601/17-R) and Shenzhen-Hong Kong-Macao Science and Technology Program (Category C, No. SGDX2020110309460000).

## References

- B. Dunn, H. Kamath and J. Tarascon, *Science*, 2011, **334**, 928–935.
- W. Wang, Q. Luo, B. Li, X. Wei, L. Li and Z. Yang, *Adv. Funct. Mater.*, 2013, **23**, 970–986.
- G. L. Soloveichik, *Chem. Rev.*, 2015, **115**, 11533–11558.
- M. Skyllas-Kazacos and M. Kazacos, *J. Power Sources*, 2011, **196**, 8822–8827.
- C. Ding, H. Zhang, X. Li, T. Liu and F. Xing, *J. Phys. Chem. Lett.*, 2013, **4**, 1281–1294.
- H. R. Jiang, Y. K. Zeng, M. C. Wu, W. Shyy and T. S. Zhao, *Appl. Energy*, 2019, **240**, 226–235.
- W. Lu, X. Li and H. Zhang, *Phys. Chem. Chem. Phys.*, 2018, **20**, 23–35.
- Z. Cheng, K. M. Tenny, A. Pizzolato, A. Forner-Cuenca, V. Verda, Y. Chiang, F. R. Brushett and R. Behrou, *Appl. Energy*, 2020, **279**, 115530.
- S. Tsushima and T. Suzuki, *J. Electrochem. Soc.*, 2020, **167**, 020553.
- A. Z. Weber, M. M. Mench, J. P. Meyers, P. N. Ross, J. T. Gostick and Q. Liu, *J. Appl. Electrochem.*, 2011, **41**, 1137.
- M. Skyllas-Kazacos, M. H. Chakrabarti, S. A. Hajimolana, F. S. Mjalli and M. Saleem, *J. Electrochem. Soc.*, 2011, **158**, R55.
- M. Zhang, M. Moore, J. S. Watson, T. A. Zawodzinski and R. M. Counce, *J. Electrochem. Soc.*, 2012, **159**, A1183.
- K. J. Kim, M. Park, Y. Kim, J. H. Kim, S. X. Dou and M. Skyllas-Kazacos, *J. Mater. Chem. A*, 2015, **3**, 16913–16933.
- H. R. Jiang, J. Sun, L. Wei, M. C. Wu, W. Shyy and T. S. Zhao, *Energy Storage Mater.*, 2020, **24**, 529–540.
- L. Wei, C. Xiong, H. R. Jiang, X. Z. Fan and T. S. Zhao, *Energy Storage Mater.*, 2020, **25**, 885–892.
- W. Zhang, J. Xi, Z. Li, H. Zhou, L. Liu, Z. Wu and X. Qiu, *Electrochim. Acta*, 2013, **89**, 429–435.
- H. R. Jiang, W. Shyy, M. C. Wu, L. Wei and T. S. Zhao, *J. Power Sources*, 2017, **365**, 34–42.
- L. Wei, T. S. Zhao, L. Zeng, Y. K. Zeng and H. R. Jiang, *J. Power Sources*, 2017, **341**, 318–326.
- M. R. Gerhardt, A. A. Wong and M. J. Aziz, *J. Electrochem. Soc.*, 2018, **165**, A2625.
- J. Houser, J. Clement, A. Pezeshki and M. M. Mench, *J. Power Sources*, 2016, **302**, 369–377.
- T. J. Latha and S. Jayanti, *J. Appl. Electrochem.*, 2014, **44**, 995–1006.
- M. MacDonald and R. M. Darling, *AIChE J.*, 2018, **64**, 3746–3755.
- J. Sun, M. Zheng, Z. Yang and Z. Yu, *Energy*, 2019, **173**, 637–646.
- D. S. Aaron, Q. Liu, Z. Tang, G. M. Grim, A. B. Papandrew, A. Turhan, T. A. Zawodzinski and M. M. Mench, *J. Power Sources*, 2012, **206**, 450–453.
- Q. Xu, T. S. Zhao and C. Zhang, *Electrochim. Acta*, 2014, **142**, 61–67.
- T. J. Latha and S. Jayanti, *J. Power Sources*, 2014, **248**, 140–146.
- S. Maurya, P. T. Nguyen, Y. S. Kim, Q. Kang and R. Mukundan, *J. Power Sources*, 2018, **404**, 20–27.
- Q. H. Liu, G. M. Grim, A. B. Papandrew, A. Turhan, T. A. Zawodzinski and M. M. Mench, *J. Electrochem. Soc.*, 2012, **159**, A1246.
- X. Ke, J. I.-D. Alexander, J. M. Prahl and R. F. Savinell, *J. Power Sources*, 2014, **270**, 646–657.
- B. W. Zhang, Y. Lei, B. F. Bai and T. S. Zhao, *Int. J. Heat Mass Transfer*, 2019, **135**, 460–469.
- E. Knudsen, P. Albertus, K. T. Cho, A. Z. Weber and A. Kojic, *J. Power Sources*, 2015, **299**, 617–628.
- S. Kumar and S. Jayanti, *J. Power Sources*, 2016, **307**, 782–787.
- S. Tsushima, S. Sasaki and S. Hirai, ECS Meeting Abstracts, 2013.
- R. M. Darling and M. L. Perry, *J. Electrochem. Soc.*, 2014, **161**, A1381.
- C. R. Dennison, E. Agar, B. Akuzum and E. C. Kumbur, *J. Electrochem. Soc.*, 2015, **163**, A5163.
- M. Lu, Y. Deng, W. Yang, M. Ye, Y. Jiao and Q. Xu, *Electrochim. Acta*, 2020, **361**, 137089.
- M. Messaggi, P. Canzi, R. Mereu, A. Baricci, F. Inzoli, A. Casalegno and M. Zago, *Appl. Energy*, 2018, **228**, 1057–1070.
- H. A. Eschenauer and N. Olhoff, *Appl. Mech. Rev.*, 2001, **54**, 331–390.

- 39 O. Sigmund, *Struct. Multidiscipl. Optim.*, 2007, **33**, 401–424.
- 40 G. I. Rozvany, *Struct. Multidiscipl. Optim.*, 2009, **37**, 217–237.
- 41 K. Yaji, S. Yamasaki, S. Tsushima, T. Suzuki and K. Fujita, *Struct. Multidiscipl. Optim.*, 2018, **57**, 535–546.
- 42 C. Chen, K. Yaji, S. Yamasaki, S. Tsushima and K. Fujita, *J. Energy Storage*, 2019, **26**, 100990.
- 43 J. Lim, C. You and I. Dayyani, *Mater. Des.*, 2020, **190**, 108552.
- 44 O. Sigmund and J. Petersson., *Struct. Opt.*, 1998, **16**, 68–75.
- 45 R. Gómez-Bombarelli, J. Aguilera-Iparraguirre, T. D. Hirzel, D. Duvenaud, D. Maclaurin, M. A. Blood-Forsythe, H. S. Chae, M. Einzinger, D. Ha and T. Wu, *Nat. Mater.*, 2016, **15**, 1120.
- 46 L. Ward, A. Agrawal, A. Choudhary and C. Wolverton, *npj Comput. Mater.*, 2016, **2**, 16028.
- 47 S. Lu, Q. Zhou, Y. Ouyang, Y. Guo, Q. Li and J. Wang, *Nat. Commun.*, 2018, **9**, 3405.
- 48 R. Yuan, Z. Liu, P. V. Balachandran, D. Xue, Y. Zhou, X. Ding, J. Sun, D. Xue and T. Lookman, *Adv. Mater.*, 2018, **30**, 1702884.
- 49 A. Mansouri Tehrani, A. O. Oliynyk, M. Parry, Z. Rizvi, S. Couper, F. Lin, L. Miyagi, T. D. Sparks and J. Brigo, *J. Am. Chem. Soc.*, 2018, **140**, 9844–9853.
- 50 A. Itai, C. H. Papadimitriou and J. L. Szwarcfiter, *SIAM J. Comput.*, 1982, **11**, 676–686.
- 51 H. Pedersen and K. Singh, Proceedings of the 4th International Symposium on Non-photorealistic Animation and Rendering, 2006.
- 52 Y. Nagata, A. Imamiya and N. Ono, *Comput. Oper. Res.*, 2020, **115**, 104860.
- 53 Y. Okamoto and R. Uehara, CCCG, 2009.
- 54 P. C. Jennings, S. Lysgaard, J. S. Hummelshøj, T. Vegge and T. Bligaard, *npj Comput. Mater.*, 2019, **5**, 1–6.
- 55 C. Kim, R. Batra, L. Chen, H. Tran and R. Ramprasad, *Comput. Mater. Sci.*, 2021, **186**, 110067.
- 56 S. Wan, X. Liang, H. Jiang, J. Sun, N. Djilali and T. Zhao, *Appl. Energy*, 2021, **298**, 117177.
- 57 D. R. Cassar, G. G. Santos and E. D. Zanotto, *Ceram. Int.*, 2021, **47**, 10555–10564.
- 58 G. H. Gu, J. Noh, I. Kim and Y. Jung, *J. Mater. Chem. A*, 2019, **7**, 17096–17117.
- 59 C. Yin, Y. Gao, S. Guo and H. Tang, *Energy*, 2014, **74**, 886–895.
- 60 P. H. Kim, J. Grove, S. Wurster and R. Crawfis, Proceedings of the 14th International Conference on the Foundations of Digital Games, 2019.
- 61 E. O. Pyzer-Knapp, C. Suh, R. Gómez-Bombarelli, J. Aguilera-Iparraguirre and A. Aspuru-Guzik, *Annu. Rev. Mater. Res.*, 2015, **45**, 195–216.
- 62 X. Ke, J. M. Prahl, J. I.-D. Alexander, J. S. Wainright, T. A. Zawodzinski and R. F. Savinell, *Chem. Soc. Rev.*, 2018, **47**, 8721–8743.
- 63 Q. Wang, Z. G. Qu, Z. Y. Jiang and W. W. Yang, *Appl. Energy*, 2018, **220**, 106–116.
- 64 Q. Xu, T. S. Zhao and P. K. Leung, *Appl. Energy*, 2013, **105**, 47–56.
- 65 M. Lu, Y. Jiao, X. Tang, W. Yang, M. Ye and Q. Xu, *J. Energy Storage*, 2021, **35**, 102284.
- 66 T. Li, F. Xing, T. Liu, J. Sun, D. Shi, H. Zhang and X. Li, *Energy Environ. Sci.*, 2020, **13**, 4353–4361.
- 67 J. W. Barnett, C. R. Bilchak, Y. Wang, B. C. Benicewicz, L. A. Murdock, T. Bereau and S. K. Kumar, *Sci. Adv.*, 2020, **6**, eaaz4301.
- 68 K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring and D. Fragedakis, *Nat. Energy*, 2019, **4**, 383.
- 69 S. Kamrava, M. Sahimi and P. Tahmasebi, *npj Comput. Mater.*, 2021, **7**, 1–9.
- 70 J. E. Santos, Y. Yin, H. Jo, W. Pan, Q. Kang, H. S. Viswanathan, M. Prodanović, M. J. Pyrcz and N. Lubbers, *Transp. Porous Media*, 2021, 1–32.
- 71 J. Willard, X. Jia, S. Xu, M. Steinbach and V. Kumar, 2020, arXiv preprint arXiv:2003.04919, vol. 1, pp. 1–34.
- 72 M. P. Brenner and P. Koumoutsakos, *Phys. Rev. Fluids*, 2021, **6**, 070001.
- 73 <https://github.com/HarryBinary/Flow-Field-Design.git>.