

# Locality sensitive hashing via mechanical behavior

Emma Lejeune<sup>\*</sup>, Peerasait Prachaseree

Department of Mechanical Engineering, Boston University, Boston, MA, United States of America

## ARTICLE INFO

### Article history:

Received 8 March 2023

Received in revised form 14 May 2023

Accepted 18 June 2023

Available online 22 June 2023

Dataset link: <https://github.com/elejeune11/mechHS>

### Keywords:

Physical computing  
Morphological computing  
Programmable matter  
Mechanical hashing

## ABSTRACT

From healing wounds to maintaining homeostasis in cyclically loaded tissue, living systems have a phenomenal ability to sense, store, and respond to mechanical stimuli. Broadly speaking, there is significant interest in designing engineered systems to recapitulate this incredible functionality. In engineered systems, we have seen significant recent computationally driven advances in sensing and control. And, there has been a growing interest – inspired in part by the incredible distributed and emergent functionality observed in the natural world – in exploring the ability of engineered systems to perform computation through mechanisms that are fundamentally driven by physical laws. In this work, we focus on a small segment of this broad and evolving field: *locality sensitive hashing via mechanical behavior*. Specifically, we will address the question: can mechanical information (i.e., loads) be transformed by mechanical systems (i.e., converted into sensor readouts) such that the mechanical system meets the requirements for a locality sensitive hash function? Overall, we not only find that mechanical systems are able to perform this function, but also that different mechanical systems vary widely in their efficacy at this task. Looking forward, we view this work as a starting point for significant future investigation into the design and optimization of mechanical systems for conveying mechanical information for downstream computing.

© 2023 Elsevier Ltd. All rights reserved.

## 1. Introduction

From the cells embedded in our skin deciding if they should activate to heal a wound [1,2], to robotic systems dexterously manipulating delicate objects [3–5], the ability to effectively transmit and interpret mechanical signals can lead to incredible functionality [6–8]. In natural systems, this ability leads to complex emergent behavior such as the maintenance of homeostasis in mechanically loaded tissue [9,10]. And, in engineered systems, we can design for responsiveness by controlling the transmission of mechanical signals through material selection and structural form [11–13].

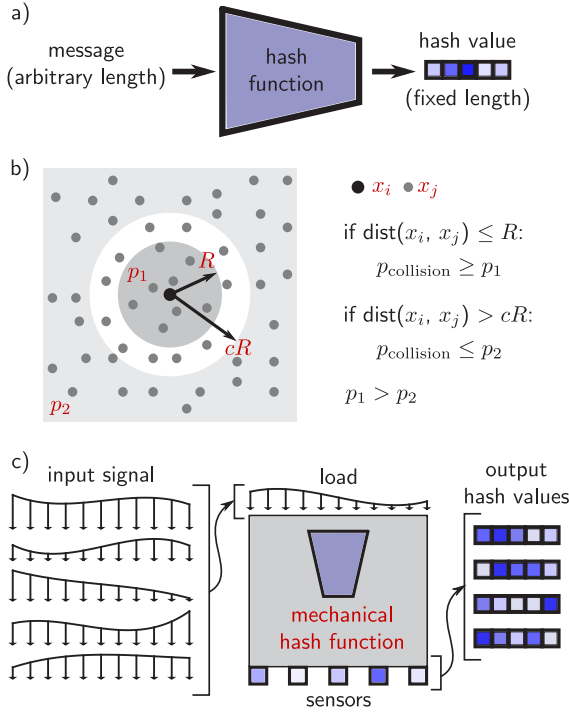
Transmission and interpretation of mechanical signals is especially relevant to growing interests in “morphological computing” [14], “physical learning” [15], and “programmable matter” [16,17]. Broadly speaking, these are all paradigms where a physical system is either programmed, or used to perform some form of “computation”. For example, researchers have experimentally realized physical logic gates [18–20], as well as responsive mechanisms that trigger functional behavior when activated [21–23]. And, within the scope of dynamical systems,

researchers have used physical bodies to perform “reservoir computing” where a higher dimensional computational space is created by multiple non-linear responses to an input signal [24,25], and cryptographic hashing where researchers have shown that chaotic hydrodynamics can be used to store and manipulate information in a fluid system [26]. In this paper, we will focus on a small segment of this broad and emerging field: *locality sensitive hashing via mechanical behavior*. Here, our goal is to explore this specific type of computation in the context of mechanical systems.

Hashing, the process of converting arbitrarily sized inputs to outputs of a fixed size, is schematically illustrated in Fig. 1a [27]. In most popular applications of hashing (e.g., storing sensitive information), it is desirable to minimize collisions (i.e., the occurrence of different inputs mapping to the same output) and obfuscate the relationship between inputs and outputs. However, there has been a growing interest in alternative types of hashing algorithms – specifically hashing algorithms for applications such as similarity search, see Fig. 1b [28]. In these algorithms, the goal is to compress input data while preserving essential aspects of the relationship between input data points. In Section 2.1, we lay out the mathematical definition for “locality sensitive hashing” [29,30]. In this paper, we will focus on the concept schematically illustrated in Fig. 1c. Can mechanical information (i.e., loads) be transformed by mechanical systems (i.e., converted into sensor readouts) such that the mechanical system meets the

<sup>\*</sup> Corresponding author.

E-mail addresses: [elejeune@bu.edu](mailto:elejeune@bu.edu) (E. Lejeune), [pprachas@bu.edu](mailto:pprachas@bu.edu) (P. Prachaseree).



**Fig. 1.** (a) Schematic illustration of a generally defined hash function; (b) Schematic illustration of the requirements for locality sensitive hashing where  $p_{\text{collision}}$  is the probability of a hash collision that is larger for inputs that are closer together; (c) Schematic illustration of a mechanical system performing locality sensitive hashing.

requirements for a locality sensitive hash function? In exploring this specific type of computing in mechanical systems, our goal is to lay a solid ideological foundation for future applications of physical computing where mechanical systems are tailored to act as a “physical computing layer” that transforms mechanical information to enable downstream responsiveness and control.

The remainder of the paper is organized as follows. In Section 2, we further define locality sensitive hashing, elaborate on the concept of mechanical systems as locality sensitive hash functions, and define an example problem to explore the performance of different mechanical systems for locality sensitive hashing. Then, in Section 3, we show the results of our investigation of our example problem, and conclude in Section 4. Overall, our goal is threefold: (1) to introduce the concept of locality sensitive hashing in the context of mechanical systems, (2) to provide a straightforward “proof of concept” that mechanical systems can be used to perform locality sensitive hashing, and (3) to lay the foundation for future investigations on optimizing mechanical behavior to perform hashing for similarity search.

## 2. Methods

We will begin in Section 2.1 by defining Locality Sensitive Hashing (LSH), then in Section 2.2 we will demonstrate how the concept behind LSH can be applied to mechanical systems as a “proof of concept”. Finally, in Section 2.3 we define an example problem that will set up the main investigation presented in this paper.

### 2.1. Locality sensitive hashing

In simple terms, a “hash function” is a function that maps input data of an arbitrary size to a fixed size output, referred to as a

“hash value” [31]. This is schematically illustrated in Fig. 1a. Hash functions have broad societal applications ranging from storing passwords, to checking if files match, to enabling data structures (e.g., dictionaries in Python) [32]. For these applications, hash functions are designed to minimize “hash collisions”. To minimize “hash collisions”, hash functions typically convert similar yet different inputs to drastically different hash values [33]. Therefore, for typical hash algorithms, it would not make sense to perform downstream applications that rely on the distance between hash values. However, there has been recent interest in an alternative type of hash algorithm referred to as “locality sensitive hashing” where the goal is to create hash functions that *encourage* collisions between similar inputs [28]. For these Locality Sensitive Hash (LSH) approaches, similar inputs should lead to similar or identical hash values. To date, these techniques have primarily been used for dimensionality reduction prior to nearest neighbor search [34]. More formally, we can introduce LSH through the following definition [28].

First, we describe our input data as points in a  $N$  dimensional metric space  $\mathcal{M}$  with distance function  $d$ . Here we will choose  $d$  as the  $L^\infty$  norm<sup>1</sup> –  $\|\mathbf{x}\|^\infty = \max\{|\mathbf{x}_1|, |\mathbf{x}_2|, \dots, |\mathbf{x}_N|\}$  where  $\mathbf{x}$  is the difference between two points in  $\mathcal{M}$ . We define a family of  $h$  hash functions as  $\mathcal{F}$  where for any two points  $q_i$  and  $q_j$  in  $\mathcal{M}$  and any hash function  $h$  chosen uniformly at random from  $\mathcal{F}$  the following conditions hold:

$$\text{if } d(q_i, q_j) \leq R : \Pr[h(q_i) = h(q_j)] \geq p_1 \quad (1)$$

$$\text{if } d(q_i, q_j) > cR : \Pr[h(q_i) = h(q_j)] \leq p_2$$

where threshold  $R > 0$ , approximation factor  $c > 1$ , and  $\Pr[\cdot]$  computes probabilities  $p_1$  and  $p_2$  with  $0 \leq p_1, p_2 \leq 1$ . In this work, we will store hash values  $h(q_i)$  with the `numpy.float64` data type, thus  $\Pr[h(q_i) = h(q_j)] \approx 0$ . In physical implementations of these systems, the precision of  $h(q_i)$  will depend on the choice of sensor. Therefore, in establishing our mechanical analogue to a locality sensitive hashing algorithm, we re-write Eq. (1) in terms of a positive value  $S$  as:

$$\text{if } d(q_i, q_j) \leq R : \Pr[d(h(q_i), h(q_j)) < S] \geq p_1 \quad (2)$$

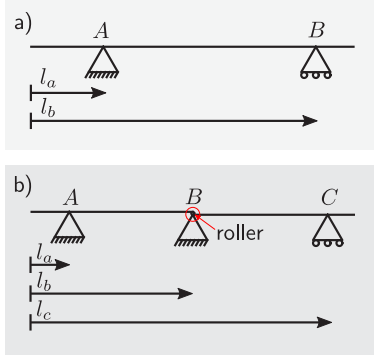
$$\text{if } d(q_i, q_j) > cR : \Pr[d(h(q_i), h(q_j)) < S] \leq p_2$$

which is elaborated on in Appendix A.1. With this definition, a family  $\mathcal{F}$  is referred to as a “Locality Sensitive Hash” family, or alternatively as  $(R, cR, p_1, p_2)$ -sensitive if  $p_1 > p_2$ . In simple and functional terms, illustrated schematically in Fig. 1b, a family of hash functions will exhibit LSH behavior if the probability of a hash collision is higher for points that are closer together in the input space.

### 2.2. Introduction to mechanical systems as locality sensitive hash functions

In this paper, we will explore the idea of using mechanical behavior to perform locality sensitive hashing where  $\mathcal{F}$  will define a class of mechanical systems. In Fig. 1c, we schematically illustrate our approach to defining this problem. Specifically, we will consider a vertical distributed load  $w(x)$  applied on the surface of a mechanical system. The mechanical system will be drawn from a family  $\mathcal{F}$ , where the mechanical behavior of the system will lead to multiple force sensor readouts at discrete locations, treated as hash values.

<sup>1</sup> Alternative choices of  $\|\cdot\|$  would also be acceptable, here we choose  $\|\cdot\|^\infty$  to simplify future calculations, see Appendices A.1 and A.2. We note briefly that our GitHub page hosts the code necessary to re-implement the numerical portions of our study with  $\|\cdot\|^2$ , which leads to very similar results and identical conclusions to what we find using  $\|\cdot\|^\infty$ .



**Fig. 2.** (a) Schematic illustration of  $\mathcal{F}_{ss}$ , a family of simply supported beams with two supports (A and B); (b) Schematic illustration of  $\mathcal{F}_{ss-c3}$ , a family of simply supported composite beams with three supports (A, B, and C).

For this setup, we define the input continuous distributed load  $w(x)$  as an evenly spaced  $N \times 1$  dimensional vector (i.e.,  $w(x)$  is a continuous interpolation of  $N$  points) and the process of “hashing” entails converting this  $N$  dimensional vector into  $ns$  (number of sensors) force sensor readouts. In future work the distributed load could be conceptualized as either multi-dimensional, i.e.,  $w(x, y)$ , or displacement driven, and the sensor readouts could capture alternative forms of behavior (e.g., strain).

Here we will define our first family of mechanical hash functions  $\mathcal{F}_{ss}$  as a family of simply supported beams, illustrated in Fig. 2a, where supports A and B are randomly placed at positions  $l_a$  and  $l_b$  such that they are separated by a minimum distance  $mL$  where  $0 < m < 1$  and  $L$  is the length of the beam. Here,  $w(x)$  will be the input distributed load, and force at each of the  $ns$  sensors is the hash value output. Following Eq. (2), we will treat two hash values as a collision if the readouts at all sensors are within a tolerance  $S$ . We can then assess if the conditions defined in Eqs. (1)–(2) hold for  $\mathcal{F}_{ss}$ . In Appendix A.1, we expand on this in detail and explicitly define  $R$ ,  $cR$ ,  $p_1$ , and  $p_2$  for  $\mathcal{F}_{ss}$ . However, for  $\mathcal{F}_{ss}$ , it only requires a simple thought experiment to demonstrate that  $\mathcal{F}_{ss}$  is not  $(R, cR, p_1, p_2)$ -sensitive.

In brief, we can demonstrate that  $\mathcal{F}_{ss}$  is not  $(R, cR, p_1, p_2)$ -sensitive by showing that we can choose two points in our inputs space ( $w_1, w_2$ ), that are arbitrarily far apart ( $d(w_1, w_2) > cR$  for an arbitrarily large  $cR$ ) yet still lead to a hash collision for all possible mechanical hash functions in  $\mathcal{F}_{ss}$  ( $p_2 = 1$ ). In the context of mechanics, because all distributed loads with the same resultant force and centroid lead to the same reaction forces, simply supported beams do not meet the definition for locality sensitive mechanical hash functions.

However, if we consider even a slightly more complicated family of mechanical systems, simply supported composite beams with 3 supports, referred to as  $\mathcal{F}_{ss-c3}$  and illustrated in Fig. 2b, the situation changes. For  $\mathcal{F}_{ss-c3}$ , we consider supports A, B, and C that are randomly placed at positions  $l_a$ ,  $l_b$ , and  $l_c$  such that they are each separated by a minimum distance  $mL$  where  $0 < m < 0.5$  and segments AB and BC are connected through a roller support. Because  $l_a$ ,  $l_b$ , and  $l_c$  change for each hash function in  $\mathcal{F}_{ss-c3}$ , two far apart distributed loads will collide with  $p_2 < 1$ , thus if we define  $R$  as small enough such that the readout at each sensor will be within tolerance  $S$  and thus  $p_1 = 1$ , we can show that  $\mathcal{F}_{ss-c3}$  is  $(R, cR, p_1, p_2)$ -sensitive. An explicit computation of  $R$ ,  $c$ ,  $p_1$ , and  $p_2$  for  $\mathcal{F}_{ss-c3}$  is expanded on in Appendix A.2. Overall, this simple demonstration is a proof of concept for mechanical systems as locality sensitive hash functions. Though this framework is straightforward, it is important to define as this work will lay the foundation for addressing more complex problems where we can consider broader definitions of mechanical hash

functions, inputs, outputs, and system structure. In Section 3, we will examine the functional performance of these simple beams alongside more complicated mechanical systems that may lead to more desirable functional LSH behavior. In addition, the straightforward LSH approach defined here will serve as a baseline for novel strategies to optimize mechanical systems for downstream signal processing.

### 2.3. Example problem definition

Beyond satisfying the criteria defined in Eq. (1), we are interested in assessing the functional utility of mechanical systems for performing locality sensitive hashing. To this end, we now define an example problem consisting of example loading, defined in Section 2.3.1, example mechanical systems, defined in Section 2.3.2, and evaluation metrics, defined in Section 2.3.3. In brief, we will investigate how different mechanical systems transform mechanical signals in the context of LSH.

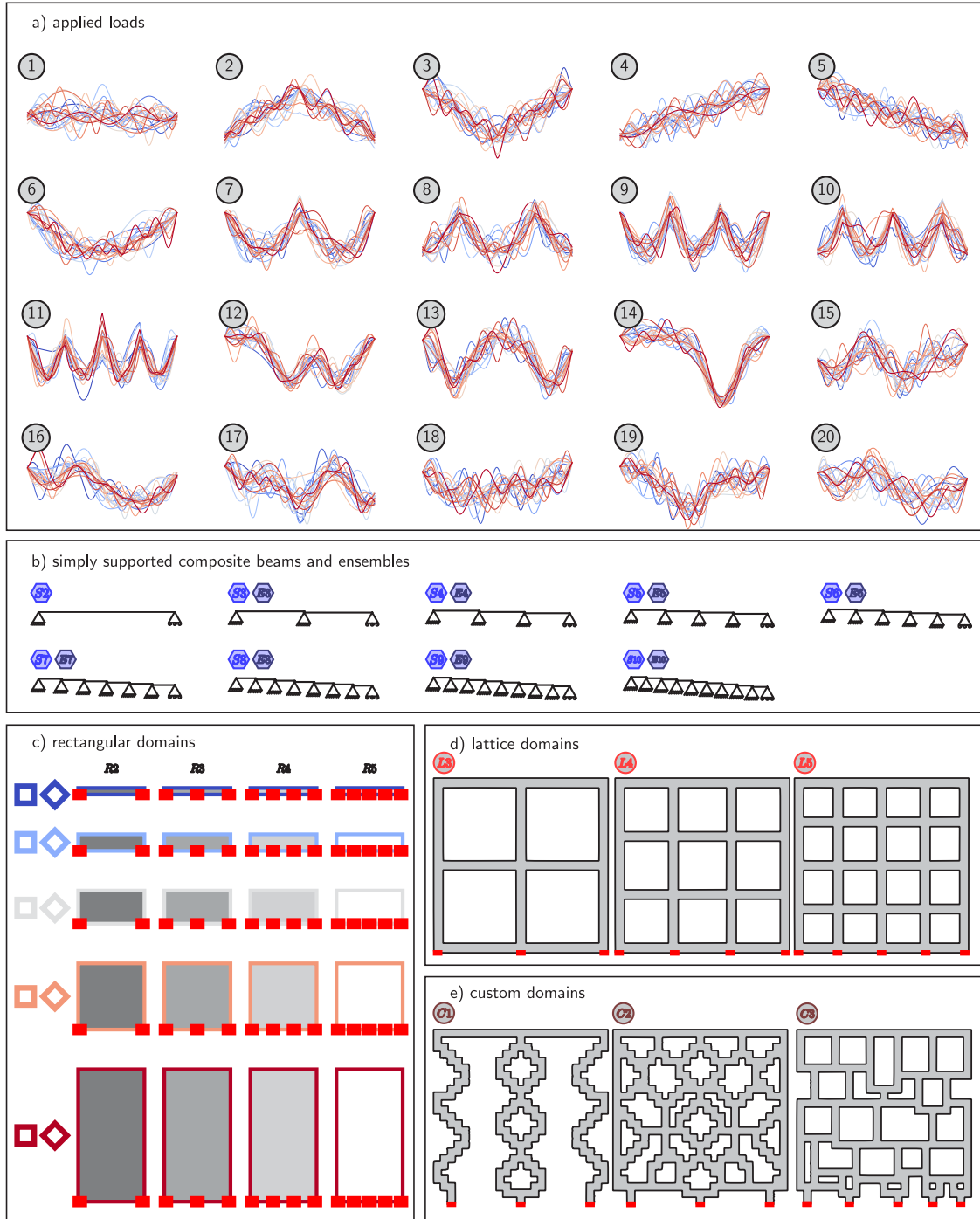
#### 2.3.1. Example loads

In Fig. 3a, we show our 200 randomly generated applied loads across 20 different categorical classes. In brief, we consider: constant loads, linear loads, piecewise linear loads, absolute value sinusoidal loads, and kernel density estimate loads based on randomly generated point densities. Note that, as illustrated in Fig. 1c, we apply all loads pointing downwards. Building on the framework laid out in Section 2.2, these loads are chosen to: (1) pose a risk of unwanted hash collisions across categories (i.e., identical centroids and resultant forces), while (2) being qualitatively different. For each of the 20 categorical classes, we generate 20 examples with different selections of correlated Perlin noise [35,36]. Specifically, we add Perlin noise with a randomly selected seed and octave (random integer with range [2, 10]) – this is the source of the variation across each category in the curves illustrated in Fig. 3a. Additional details for describing all categories of load are given in Appendix B.2, and the link to the code for re-generating these loads including selected random seeds is given in Section 5.

#### 2.3.2. Example mechanical systems

In Fig. 3b–e, we show the mechanical systems investigated in this study. For all examples, we set the top surface length  $L = 10$  (length units). In brief, we consider the following classes of mechanical systems:

- Simply supported and simply supported composite beams defined in Section 2.2. We will consider composite beams with up to 10 supports, and we will report the performance of both single beam instances and hard voting based ensemble behavior for a total of 17 different scenarios (see Fig. 3b). For the beam ensembles, each of the 100 composite beams will have different randomly generated support locations. The final performance of each ensemble will then be represented as a single value that combines information from all 100 randomly generated composite beams. An illustration of a representative beam ensemble is included in Appendix B.1.
- Homogeneous rectangular domains with variable depth, number of sensors, and fixity. We will consider rectangular domains with 2, 3, 4, and 5 force sensors, depths 1.0, 2.5, 5.0, 10.0, and 20.0, and both with and without bottom fixity. For reference, a rectangular domain with  $depth = 1.0$  will have dimension  $1.0 \times 10.0$ , and a rectangular domain with  $depth = 10.0$  will be a square. The combination of 4 sensor options, 5 depths, and 2 bottom fixities leads to a total of  $4 \times 5 \times 2 = 40$  different rectangular domains (see Fig. 3c where all potential sensor placements and depths are illustrated).



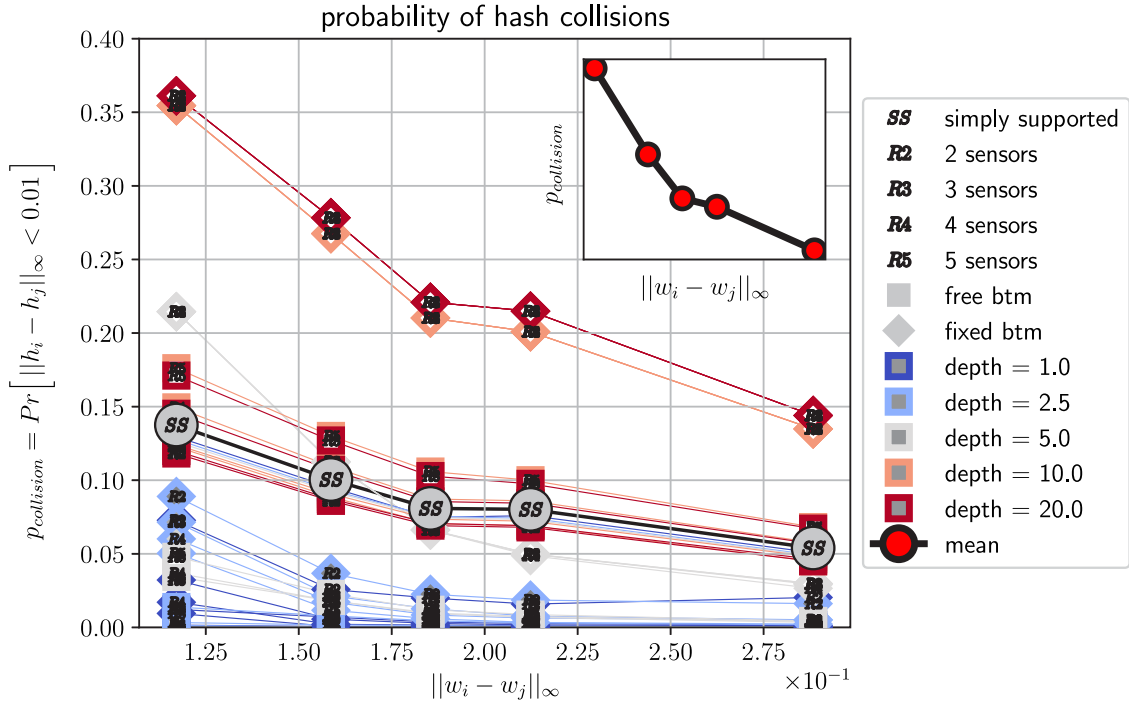
**Fig. 3.** (a) Illustration of the 20 classes of applied loads; (b) Schematic illustration of the simply supported and ensemble beam mechanical systems (note an example ensemble is explicitly illustrated in Fig. 8); (c) Schematic illustration of the rectangular domains with different depths ( $depth = [1.0, 2.5, 5.0, 10.0, 20.0]$ ) and number of sensors ( $ns = [2, 3, 4, 5]$ ). For all rectangular domains, we simulate the bottom fixed only at sensors ( $\square$ ) and the whole bottom fixed ( $\diamond$ ); (d) Schematic illustration of the lattice domains ( $L3, L4, L5$ ); (e) Schematic illustration of the custom domains ( $C1, C2, C3$ ). Note that all lattice and custom domains have  $depth = 10.0$ . In b–e, the markers next to each mechanical system match the markers used in Figs. 4–5.

- Lattice domains with 3, 4, and 5 sensors and corresponding  $2 \times 2$ ,  $3 \times 3$ , and  $4 \times 4$  window grids (see Fig. 3d). All lattice domains have  $depth = 10.0$ .
- Custom domains with three different geometries and variable sensor numbers  $ns$  (see Fig. 3e). All custom domains have  $depth = 10.0$ .

For each mechanical system, we run 400 simulations corresponding to each of the applied loads illustrated in Fig. 3a and report the  $y$  direction force at each sensor location. In the context of LSH,

the applied loads are the input, and these forces are the hash function output. In all cases, sensor locations are fixed in both the  $x$  and  $y$  direction, but only the  $y$  direction force is used in subsequent analysis. Simply supported beams are simulated in a Python [37] script that we link to in Section 5. All finite element simulations are conducted using open source finite element software FEniCS [38,39] and built in FEniCS mesh generation software mshr. To avoid numerical artifacts, we simulate all domains as Neo-Hookean materials with  $\nu = 0.3$ , and a fine mesh (mshr mesh parameter set to 200) of quadratic triangular elements.





**Fig. 4.** Plot of  $p_{\text{collision}}$  of hash values (defined as the probability that  $\|h_i - h_j\|_\infty < 0.01$ ) vs. distance between input loads  $\|w_i - w_j\|_\infty$  for all rectangular domains. For context, we also include the curve for a simply supported beam with two supports, and the curve for the mean of all rectangular samples (see inset, where  $p_{\text{collision}} = [0.14, 0.099, 0.075, 0.070, 0.047]$ ).

The link to the code for re-generating all simulation results is given in Section 5. In addition, our code contains a tutorial for designing and simulating user defined architected domains to make it straightforward to expand on the initial results presented in this paper.

### 2.3.3. Evaluation metrics

Based on the suite of 400 loads defined in Section 2.3.1 and the 63 mechanical systems defined in Section 2.3.2 (see Fig. 3), we will have over 25,000 simulation results to analyze. To draw conclusions from these results, we will examine the relationship between three different quantities of interest: (1) the probability of a hash collision with respect to the  $L^\infty$  distance between loads, (2) the Spearman's Rank correlation coefficient  $\rho$  between the  $L^\infty$  distance between loads and the  $L^\infty$  distance between hash values [40,41], and (3) the classification accuracy based on a single nearest neighbor (note that there are 20 categories of loads, with 20 examples in each category, illustrated in Fig. 3a) [42].

To begin, we will report the probability of hash collision  $p_{\text{collision}}$  as function of the  $L^\infty$  distance between input loads and hash values for all 40 rectangular domains and compare to a baseline simply supported beam from  $\mathcal{F}_{\text{ss}}$ . To approximate  $p_{\text{collision}}$  vs.  $L^\infty$  distance for our selection of example loads, we divide all load pairs into five equally sized bins based on the distance between input loads. Then, we compute  $p_{\text{collision}}$  as the fraction of hash values in each bin with  $L^\infty < 0.01$  (input loads are normalized to sum up to 1). In Fig. 4, we plot  $p_{\text{collision}}$  with respect to the average input distance value associated with each equally sized bin. Thus, for each device we will have a curve that shows binned  $p_{\text{collision}}$  vs.  $L^\infty$  distance between loads. In Section 3, we visualize this behavior in Fig. 4 as one (limited) approach to observing LSH behavior.

Beyond  $p_{\text{collision}}$  vs.  $L^\infty$  distance, we will also compute Spearman's  $\rho$  designed to capture the rank correlation between the  $L^\infty$  distance between loads and the  $L^\infty$  distance between hash values. For all 79,800 load pairs ( $400 \times 399/2 = 79,800$  pairs) we rank

order the distances of both the loads and the hash values and then compute  $\rho$  as:

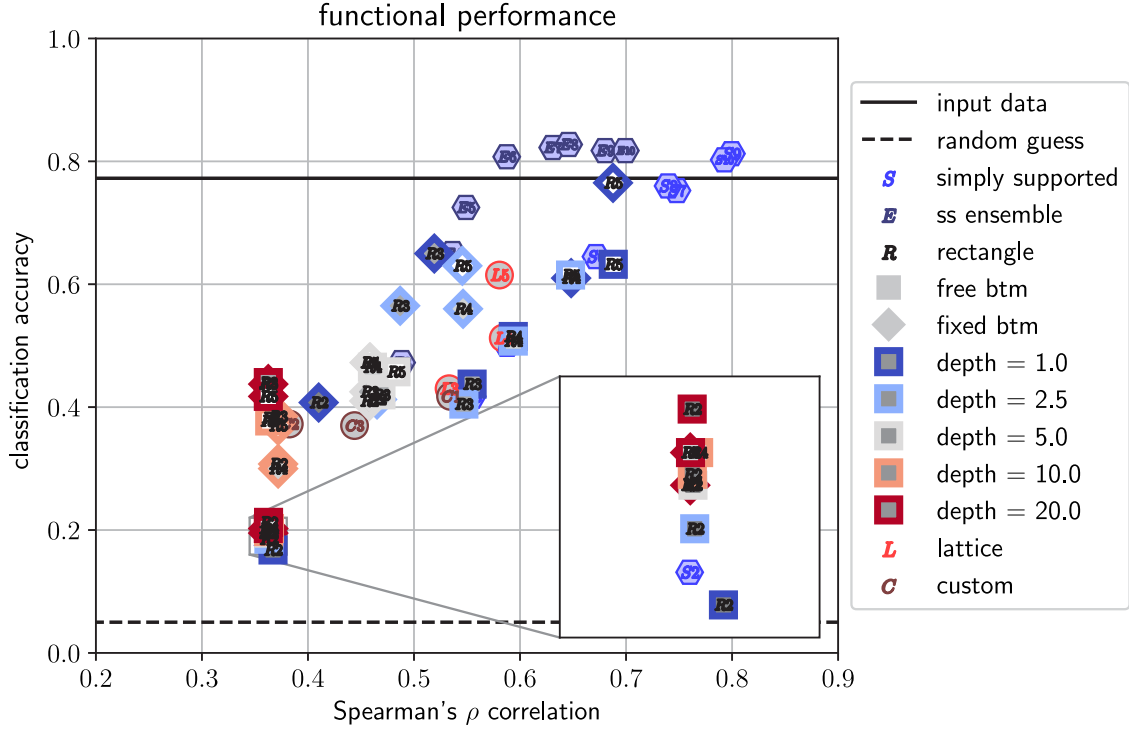
$$\rho = 1 - \frac{6 \sum_{i=1}^n r_i^2}{n(n^2 - 1)} \quad (3)$$

where  $r_i$  is the difference between the ranks of the load and hash values distances and  $n$  is the number of load pairs. For perfect monotonic correlation  $\rho = 1$  and for no correlation  $\rho = 0$ . We perform this operation with the Python function `scipy.stats.spearmanr()` [41]. Though visualizing  $p_{\text{collision}}$  vs.  $L^\infty$  is a more intuitive match to the LSH definitions in Eqs. (1)–(2), the quantitative comparison of preserving input relationships is perhaps more interpretable via Spearman's  $\rho$ . In Section 3, we visualize Spearman's  $\rho$  in each device both with respect to classification accuracy (Fig. 5), and mechanical system properties (Fig. 6, see also Fig. 12).

The third evaluation quantity is chosen as a set up for future applications in “learning to hash”. In future “learning to hash” applications, the hashing efficacy of the mechanical system will be measured via the performance on a functional task. Here we choose classification accuracy as an example functional task. In brief, each of the 400 loads belongs to one of 20 applied load classes (20 classes, 20 loads per class). This is illustrated in Fig. 3a. Here, we compute classification accuracy using a simple nearest neighbor algorithm where we predict the class of a given load based on its hash value. For each of the 400 loads, we implement a new k-nearest neighbor classifier based on the 399 other loads and then see what class is predicted for the held out load with  $k = 1$  [43]. Classification accuracy is then defined based on the ability of this algorithm to predict the class of the held out load:

$$\text{accuracy} = \frac{\text{correct predictions}}{\text{total predictions}} \quad (4)$$

where the “correct” prediction is which of the 20 applied load classes (see Fig. 3a) the input load belongs to. We implement this algorithm with scikit-learn [42] and report the average prediction



**Fig. 5.** Plot of classification accuracy vs. Spearman's  $\rho$  for all mechanical systems explored in this study. Note the inset plot which contains S2 (simply supported beam with two supports) and similarly performing mechanical systems. Fig. 12 also presents Spearman's  $\rho$  and classification accuracy plotted with respect to number of sensors and domain depth.

accuracy across all 400 individual hold out cases. Because there are 20 identically sized labeled classes according to the problem definition, the baseline prediction accuracy that represents random guessing is 0.05. In Section 3, classification accuracy is reported in Fig. 5. As a brief note, the link to the code for computing all of these quantities of interest is given in Section 5.

### 3. Results and discussion

In this Section, we will summarize the results of the investigation detailed in Section 2.3. On one hand, the results of this study are largely intuitive – applied loads influence mechanical response, and different applied loads lead to different mechanical response. On the other hand, this initial investigation is an important step because it lays the groundwork for significant future investigation in designing mechanical systems that outperform the baseline proof of concept results shown here. Multiple future directions are explicitly stated in Section 4.

#### 3.1. Probability of hash collision decreases with increasing distance between input loads

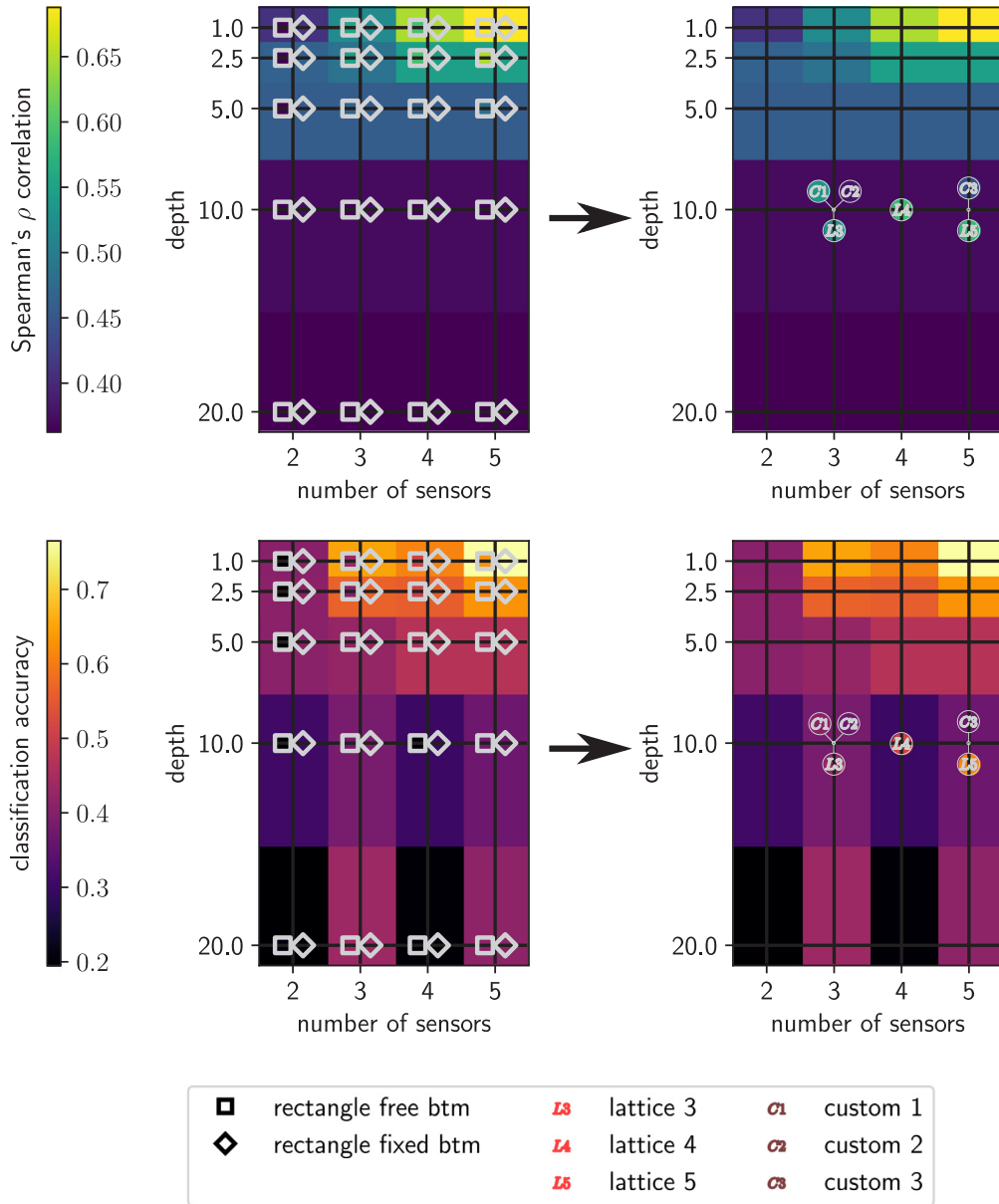
The first major result of this investigation is shown in Fig. 4 where we visualize the probability of a hash collision  $p_{\text{collision}}$  vs. the  $L^\infty$  distance between the normalized input loads (sampled with  $N = 1000$ ) for all rectangular domains. Following Section 2.3.2 and Section 2.3, we define a collision as the circumstance where every component of the hash value is within distance  $S = 0.01$  (with all input loads normalized to have the same total resultant force). For example,  $[0.5, 0.25, 0.25]$  and  $[0.495, 0.2525, 0.2525]$  would “collide”. The critical outcome shown in this plot is that  $p_{\text{collision}}$  decreases as the  $L^\infty$  distance increases, which is desirable behavior for locality sensitive hashing and hashing for similarity search in general. We note briefly

that the inset plot of Fig. 4 shows the mean  $p_{\text{collision}}$  curve for all rectangular domains where this decrease is readily visible. However, in Fig. 4, we also plot a baseline  $p_{\text{collision}}$  curve for a simply supported beam with two supports which shows that observing a decrease in  $p_{\text{collision}}$  vs. the  $L^\infty$  input distance for this selection of input loads (see Fig. 3a) is not sufficient to claim LSH behavior. Therefore, we turn to the other metrics defined in Section 2.3.3, Spearman's  $\rho$  and classification accuracy, to add needed context to our investigation.

#### 3.2. Spearman's $\rho$ and classification accuracy vary across mechanical systems

In Fig. 5, we explore two key components of the functional behavior of our mechanical hashing systems, Spearman's  $\rho$  and classification accuracy, and visualize their relationship. From a functional perspective, Spearman's  $\rho$  captures a picture of the potentially non-linear distance preservation between the input loads and the hash values. And, from a functional perspective, assessing classification accuracy will set up a toy problem and baseline functional performance for future work in optimizing mechanical domains to perform hashing functions. Specifically, we anticipate future work in designing application specific mechanical systems that target specific functionality (e.g., high classification accuracy) where the results shown here will serve as baselines at these tasks.

Overall, we observe that accuracy ranges from  $\approx 0.17$ , the accuracy of a simply supported beam with two supports, to  $\approx 0.8$ , the accuracy for simply supported ensembles with  $> 5$  sensors (note that force sensors are all located at the supports). For reference,  $\text{accuracy} = 0.05$  corresponds to random guessing, and  $\text{accuracy} = 0.77$  corresponds to the prediction accuracy that is obtained by performing classification with the input loads directly rather than with the hash values. And, overall,



**Fig. 6.** Visualization of Spearman's  $\rho$  (upper) and classification accuracy (lower) with respect to domain depth and number of sensors. In the left column, results from individual rectangular domains are indicated by the  $\square$  and  $\diamond$  markers, and the background shading is based on the rectangular domain with a fixed bottom. In the right column, results from the lattice and custom domains are superimposed on the same background shading, thus comparing the lattice and custom domains to the rectangular domain baseline. Note that the fill color of all markers is dictated by Spearman's  $\rho$  (upper) and classification accuracy (lower). Fig. 12 also presents Spearman's  $\rho$  and classification accuracy plotted with respect to number of sensors and domain depth. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Spearman's  $\rho$  and classification accuracy appear to be correlated, which is consistent with LSH behavior for scenarios where the hash function is not specifically learned to perform a non-linear transformation on the distribution of inputs. In addition, it is worth mentioning that the inset plot in Fig. 5 contains multiple “poorly performing” domains with both low Spearman's  $\rho$  and low classification accuracy. Notably, the domains in this region are rectangular domains that are deep and/or have only two sensors. And, as demonstrated by the inset plot, these domains all perform similarly to the simply supported beam with two supports.

For each of the mechanical systems introduced in Section 2.3.2, key observations are as follows:

- For the simply supported composite beams and beam ensembles, increasing  $ns$  leads to both higher Spearman's  $\rho$

and higher accuracy. And, the hard voting based ensemble predictions tend to outperform the single mechanical system results. Notably, the transition between  $ns = 2$  and  $ns = 3$  leads to the largest incremental improvement in performance for the simply supported beams (accuracy = 0.175 to accuracy = 0.42), which is also consistent with our introduction to the concept of LSH in Section 2.2.

- For the rectangular domains, accuracy and Spearman's  $\rho$  vary widely. As stated previously, some designs perform poorly – similar to the simply supported beam with  $ns=2$  – whereas other designs exceed the simply supported composite beam performance for the equivalent number of sensors. In general, increasing  $ns$  and decreasing depth correspond to increases in  $\rho$  and improvements in accuracy. From a mechanics perspective, this is a logical result as increasing  $ns$  will provide more information about mechanical behavior

while increasing *depth* will lead to diminishing differentiation between applied tractions with the same resultant force and centroid following Saint-Venant's principle [44].

- In comparison to the rectangular domains, the lattice domains led to consistent performance improvements ( $accuracy_{L3} = 0.43$ ,  $accuracy_{L4} = 0.51$ ,  $accuracy_{L5} = 0.61$ ).
- In comparison to the rectangular domains and the lattice domains, the custom domains selected offered little performance improvement ( $accuracy_{C1} = 0.42$ ,  $accuracy_{C2} = 0.37$ ,  $accuracy_{C3} = 0.37$ ).

Further visualizations are shown in Fig. 12 where Spearman's  $\rho$  and classification accuracy are plotted with respect to the number of sensors and domain depth. And, in Table 1, we also list Spearman's  $\rho$  and classification accuracy for each domain directly. Finally, it is worth re-emphasizing that even when classification is performed on the input signals directly, we only achieve  $accuracy = 0.77$ . This is because we defined an example problem with loads that can be difficult to disaggregate in the presence of noise. In Appendix C Fig. 6, we provide the confusion matrix for the unaltered input signals to highlight which loads are leading to overlapping predictions. This quantitative outcome is consistent with the qualitative comparison that can be made by examining the plots in Fig. 3a.

### 3.3. Architected domains change, and can be used to enhance, task specific LSH performance

In Fig. 6, we re-organize the rectangular, lattice, and custom domain data shown in Fig. 5 to better visualize the influence of architected domains on Spearman's  $\rho$  and classification accuracy with respect to both domain depth and number of sensors. Critically, this figure demonstrates that even for a fixed number of sensors, there is a large variation in domain performance. And, by comparing Spearman's  $\rho$  and classification accuracy of the lattice domains to the rectangular domains, we can see that architecting these domains can help overcome the drop off in performance with respect to domain depth. Finally, the spread in performance between the lattice and custom domains also indicates that there is potential richness to this problem where selecting domains that will perform well is non-trivial. These observations taken together point towards a strong future opportunity to engineer systems that, for a given suite of possible loads and allowable number of sensors, are specifically designed to maximize Spearman's  $\rho$ , classification accuracy, and/or achieve an alternative type of engineered relationship between mechanical inputs and sensor readouts. Though the notion that architected domains can alter force transmission is a straightforward result, the framework that we have established here will directly enable future work in the design and optimization of architected domains to perform desirable application specific signal transformations.

## 4. Conclusion

In this paper, we began with a brief introduction to the concept of hashing and hashing for similarity search. Then, we define locality sensitive hashing and lay the foundation for considering mechanical systems as locality sensitive hash functions. From both our analytical and computational investigations, we find that mechanical systems can exhibit the properties required for locality sensitive hashing, and we find that by tuning mechanical inputs (e.g., boundary conditions, domain architecture) we can change the functional efficacy of mechanical systems for this task. Based on our observations, and the very general scope of "mechanical systems", we anticipate that there will be a significantly broader potential range of behavior than what we captured in the systems selected for this study. Overall, the main contributions

**Table 1**

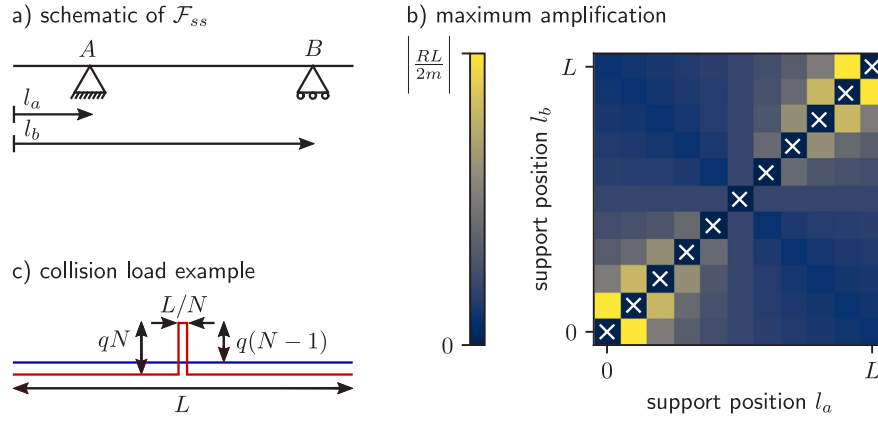
Summary of results, supporting information for Fig. 5.

System type	Number of sensors $n_s$	Domain depth $d$	Spearman's $\rho$	Classification accuracy
simply supported	2	n/a	0.36	0.17
simply supported	3	n/a	0.55	0.42
simply supported	4	n/a	0.59	0.5
simply supported	5	n/a	0.59	0.5
simply supported	6	n/a	0.67	0.65
simply supported	7	n/a	0.75	0.75
simply supported	8	n/a	0.74	0.76
simply supported	9	n/a	0.8	0.81
simply supported	10	n/a	0.79	0.8
ss ensemble	3	n/a	0.49	0.47
ss ensemble	4	n/a	0.54	0.65
ss ensemble	5	n/a	0.55	0.72
ss ensemble	6	n/a	0.59	0.81
ss ensemble	7	n/a	0.63	0.82
ss ensemble	8	n/a	0.65	0.83
ss ensemble	9	n/a	0.68	0.82
ss ensemble	10	n/a	0.7	0.82
rect fixed btm	2	1	0.41	0.41
rect fixed btm	3	1	0.52	0.65
rect fixed btm	4	1	0.65	0.61
rect fixed btm	5	1	0.69	0.77
rect fixed btm	2	2.5	0.47	0.41
rect fixed btm	3	2.5	0.49	0.56
rect fixed btm	4	2.5	0.55	0.56
rect fixed btm	5	2.5	0.55	0.63
rect fixed btm	2	5	0.46	0.41
rect fixed btm	3	5	0.46	0.42
rect fixed btm	4	5	0.46	0.47
rect fixed btm	5	5	0.46	0.47
rect fixed btm	2	10	0.37	0.31
rect fixed btm	3	10	0.37	0.39
rect fixed btm	4	10	0.37	0.3
rect fixed btm	5	10	0.37	0.37
rect fixed btm	2	20	0.36	0.2
rect fixed btm	3	20	0.36	0.44
rect fixed btm	4	20	0.36	0.2
rect fixed btm	5	20	0.36	0.42
rect	2	1	0.37	0.17
rect	3	1	0.55	0.44
rect	4	1	0.59	0.52
rect	5	1	0.69	0.63
rect	2	2.5	0.36	0.18
rect	3	2.5	0.55	0.41
rect	4	2.5	0.59	0.51
rect	5	2.5	0.65	0.61
rect	2	5	0.36	0.2
rect	3	5	0.47	0.42
rect	4	5	0.46	0.47
rect	5	5	0.48	0.46
rect	2	10	0.36	0.2
rect	3	10	0.36	0.38
rect	4	10	0.36	0.2
rect	5	10	0.36	0.38
rect	2	20	0.36	0.21
rect	3	20	0.36	0.44
rect	4	20	0.36	0.2
rect	5	20	0.36	0.42
lattice	3	10	0.53	0.43
lattice	4	10	0.58	0.51
lattice	5	10	0.58	0.61
custom 1	3	10	0.53	0.42
custom 2	3	10	0.38	0.37
custom 3	5	10	0.44	0.37

of this work are to: (1) introduce the concept of locality sensitive hashing via mechanical behavior, (2) define a numerical approach to readily assessing metrics that indicate functional locality sensitive hashing behavior, and (3) establish a baseline performance for future comparison where mechanical systems are optimized to perform hashing for similarity search related tasks.

Following this thread, is worth highlighting that we view this investigation as a starting point for significant further study of





**Fig. 7.** (a) Schematic illustration of  $\mathcal{F}_{ss}$ ; (b) Visualization of the amplification of the  $\pm R1$  term defined in Eq. (6) for different support positions  $l_a$  and  $l_b$  (here  $m = 0.1$ , the  $\times$  markers indicate choices where  $|l_a - l_b| < mL$ , and we discretize 11 evenly spaced potential support positions  $[0, L]$ ); (c) Schematic illustration of an example of two loads that will always collide for  $\mathcal{F}_{ss}$ .

hashing performed by mechanical systems. Looking forward, we anticipate four key endeavors that will build on this work. First, future work is required to demonstrate that mechanical systems besides simply supported and simply supported composite beams do or do not meet the *formal* definition of locality sensitive hash functions. We anticipate that this work will be conducted by beginning with our straightforward to compute *proxies* for locality sensitive behavior, and then showing formally that a given mechanical system is able to meet the definition laid out in Eq. (2) by anticipating extreme load pairs following the procedure in Appendix A.1. Second, an important next step is the physical realization of mechanical systems for locality sensitive hashing. Constraints imposed by the need for ready constructability and current sensing capabilities will pose a challenge [45], and it will be important to determine if our findings remain consistent in the equivalent experimentally realized systems. Third, it will be interesting to explore the efficacy of a *learning to hash* approach in mechanical systems where the input distribution is known and the hash function (i.e., the mechanical system) is designed specifically to perform a desired task [46,47]. The framework and metrics defined here will allow us to construct an optimization problem where both system mechanical behavior and sensor placement can be jointly tailored to serving a specific function. In future learning to hash applications, the mechanical systems explored in this work can serve as a baseline for comparison, where optimized systems should lead to better performance of a desired task. Because this is a challenging optimization problem, we anticipate that there will be a need to implement efficient modeling and optimization strategies [48–51]. Finally, we anticipate that the structural form of architected materials that are highly effective at hashing for similarity search may exist in nature, and identifying relevant motifs may help us better understand force transmission in biological cells and tissue. Though this initial study is quite straightforward, we anticipate that it will directly enable a highly novel approach to physical computing.

## 5. Additional information

The data and code to reproduce and build on the results in this paper are provided on our GitHub page (<https://github.com/elejeune11/mechHS>).

## Declaration of competing interest

The author declares that there are no conflicts of interest.

## Data availability

The data and code to reproduce and build on the results in this paper are provided on our GitHub page (<https://github.com/elejeune11/mechHS>).

## Acknowledgments

This work was made possible with funding through the Boston University David R. Dalton Career Development Professorship, the Hariri Institute Junior Faculty Fellowship, the Haythornthwaite Foundation Research Initiation Grant, and the National Science Foundation, United States Grant CMMI-2127864. This support is gratefully acknowledge.

## Appendix A. Simply supported composite beams as locality sensitive hash functions

Here we provide supporting information for the work presented in Section 2.1. As stated in the main text, we are exploring the problem schematically illustrated in Fig. 1c. Following the definition in Eq. (1), our goal is to determine if a given family of mechanical hash functions  $\mathcal{F}$  is  $(R, cR, p_1, p_2)$ -sensitive.

### A.1. Simply supported beams

As our first exploration, we define a family of mechanical hash functions  $\mathcal{F}_{ss}$  as a family of simply supported beam reaction forces. This family, illustrated in Fig. 7a, is defined by the randomly generated placement of reaction supports, A and B placed at  $l_a$  and  $l_b$  respectively. These supports can have any location as long as they are separated by distance  $mL$ , where  $0 < m < 1$  and  $L$  is the length of the beam. To begin, we will first establish  $R$ , the threshold distance, and  $p_1$ , the probability of a collision for two loads within the threshold distance. For simplicity, we choose to define  $R$  for  $p_1 = 1$ . To do this, we need to define the threshold distance between two loads  $w(x)_1$  and  $w(x)_2$  that will always hash to the same value, defined as  $h[w_1(x)] = h[w_2(x)]$  and alternatively written as  $h_1 = h_2$ . To explicitly define distance  $R$ , we need to define the size of our hash buckets  $S$ . Because our outputs will come in the form of continuous numerical values, we will conceptualize our hash buckets as discrete bins that break up this continuous space into bins of size  $S$ . To mitigate the influence of the placement of the bin boundaries, we will simply consider any two numbers within distance  $S$  as identical. Therefore, for a simply supported beam, two loads  $w_1(x)$  and

$w_2(x)$  will experience a hash collision when both reaction forces collide, defined as:

$$|A_{y1} - A_{y2}| < S \quad \text{and} \quad |B_{y1} - B_{y2}| < S. \quad (5)$$

For  $\mathcal{F}_{ss}$ , we can define  $w_1$  and  $w_2$  in terms of  $R$  as:

$$w_2(x) = w_1(x) \pm R \mathbf{1} \quad (6)$$

where  $\mathbf{1}$  is a vector with the same length  $N$  as  $w_1(x)$  and  $w_2(x)$  such that  $R$  is the  $L^\infty$  norm of the distance between  $w_1(x)$  and  $w_2(x)$ . For  $\mathcal{F}_{ss}$ , as illustrated in Fig. 7b, the most extreme amplification of the  $\pm R \mathbf{1}$  term will occur when  $l_a = 0$  and  $l_b = mL$  (alternatively  $l_a = L - mL$ ,  $l_b = L$ ). For  $l_a = 0$  and  $l_b = mL$ , we perform a simple equilibrium calculation to compute:

$$\begin{aligned} A_{y1} &= \int_0^L w_1(x) dx - \frac{1}{mL} \int_0^L w_1(x) x dx \\ B_{y1} &= \frac{1}{mL} \int_0^L w_1(x) x dx \\ A_{y2} &= A_{y1} \pm \left(1 - \frac{1}{2m}\right) RL \\ B_{y2} &= B_{y1} \pm \left(\frac{1}{2m}\right) RL \end{aligned} \quad (7)$$

which allows us to compute:

$$\begin{aligned} |A_{y1} - A_{y2}| &= \left| \left(1 - \frac{1}{2m}\right) RL \right| \\ |B_{y1} - B_{y2}| &= \left| \left(\frac{1}{2m}\right) RL \right| \end{aligned} \quad (8)$$

where:

$$\begin{aligned} S &= \max \left( \left| \left(1 - \frac{1}{2m}\right) RL \right|, \left| \left(\frac{1}{2m}\right) RL \right| \right) \\ &= \left| \left(\frac{1}{2m}\right) RL \right| \end{aligned} \quad (9)$$

for  $0 < m < 1$  which allows us to compute:

$$R = 2mS/L. \quad (10)$$

Therefore, for  $d(p, q) < 2mS/L$ ,  $p_1 = 1$ .

Following our identification of  $R$  for  $p_1 = 1$ , we need to see if it is possible to specify  $c$  such that  $p_2 < p_1$ . Here is where we encounter the fundamental limitation of  $\mathcal{F}_{ss}$  as a  $(R, cR, p_1, p_2)$ -sensitive hash function.

Of course, for  $\mathcal{F}_{ss}$ , we can define two loads  $w_1(x)$  and  $w_2(x)$  that are arbitrarily far apart (i.e., no upper limit on  $c$ ) that will always lead to a hash collision. Specifically, we just need to choose two different loads with the same resultant force and centroid. For example, illustrated in Fig. 7c, we can define one load as:

$$w_1(x) = q \mathbf{1} \quad (11)$$

where  $q$  is a constant and  $\mathbf{1}$  is length  $N$  vector of ones, and another load as a central spike, written as:

$$w_2(x) = \begin{cases} 0, & x < \frac{L(N-1)}{2N} \\ qN, & \frac{L(N-1)}{2N} \leq x \leq \frac{L(N+1)}{2N} \\ 0, & \frac{L(N+1)}{2N} < x \end{cases} \quad (12)$$

and illustrated in Fig. 7c. For this choice of  $w_1(x)$  and  $w_2(x)$ , we can compute  $\|w_1(x) - w_2(x)\|_\infty$  as:

$$\|w_1(x) - w_2(x)\|_\infty = q(N - 1) \quad (13)$$

which can become arbitrarily large.<sup>2</sup> Because  $p_2 = 1$  for any value of  $c$ , we can formally say that  $\mathcal{F}_{ss}$  is not  $(R, cR, p_1, p_2)$ -sensitive.

<sup>2</sup>  $\|w_1(x) - w_2(x)\|_\infty = q(N - 1)$  if  $N$  is odd,  $\|w_1(x) - w_2(x)\|_\infty = q(N/2 - 1)$  if  $N$  is even.

## A.2. Simply supported composite beams with $> 2$ supports

If we instead consider a slightly more complicated family of mechanical systems, simply supported composite beams with 3 supports, referred to as  $\mathcal{F}_{ss-c3}$  and illustrated in Fig. 8a, this picture changes. Here, we consider a composite beam where segments  $AB$  and  $BC$  are connected via a roller support with supports  $A$ ,  $B$ , and  $C$  located at  $l_a$ ,  $l_b$ , and  $l_c$  respectively. In this case, a hash collision between two loads  $w_1(x)$  and  $w_2(x)$  requires all three support reactions to collide, written as:

$$|A_{y1} - A_{y2}| < S \quad \text{and} \quad |B_{y1} - B_{y2}| < S \quad |C_{y1} - C_{y2}| < S. \quad (14)$$

We can follow the same logic to compute  $R$  as the prior  $\mathcal{F}_{ss}$  example. Specifically, for  $\mathcal{F}_{ss-c3}$  we identify the most extreme amplification of the  $\pm R \mathbf{1}$  term introduced in Eq. (6) for  $l_a = 0$ ,  $l_b = mL$ , and  $l_c = 2mL$ , where  $m$  controls the minimum allowable distance between supports (see Fig. 8b for a visualization of this amplification). Again, we perform a simple equilibrium calculation to compute the reaction supports  $A_y$ ,  $B_y$ ,  $C_y$  for  $w_1(x)$  and  $w_2(x)$  as:

$$\begin{aligned} A_{y1} &= \frac{mL \int_0^{mL} w_1(x) dx - \int_0^{mL} w_1(x) x dx}{mL} \\ B_{y1} &= \int_0^L w_1(x) dx - \frac{mL \int_0^{mL} w_1(x) dx - \int_0^{mL} w_1(x) x dx}{mL} \\ &\quad - \frac{\int_{mL}^L w_1(x) x dx - mL \int_{mL}^L w_1(x) dx}{mL} \\ C_{y1} &= \frac{\int_{mL}^L w_1(x) x dx - mL \int_{mL}^L w_1(x) dx}{mL} \\ A_{y2} &= A_{y1} \pm \frac{RmL}{2} \\ B_{y2} &= B_{y1} \pm \left( RL - \frac{RmL}{2} - \frac{R(L - mL)^2}{2mL} \right) \\ C_{y2} &= C_{y1} \pm \frac{(L - mL)^2 R}{2mL} \end{aligned} \quad (15)$$

which allows us to compute:

$$\begin{aligned} |A_{y1} - A_{y2}| &= \left| \frac{RmL}{2} \right| \\ |B_{y1} - B_{y2}| &= \left| RL - \frac{RmL}{2} - \frac{R(L - mL)^2}{2mL} \right| \\ |C_{y1} - C_{y2}| &= \left| \frac{(L - mL)^2 R}{2mL} \right| \end{aligned} \quad (16)$$

which can be manipulated, following the example in Eq. (9), as:

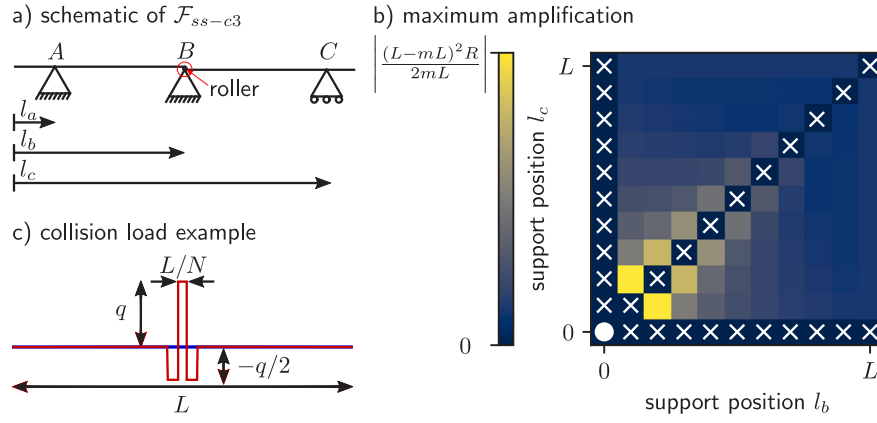
$$S = \max \left( \left| \frac{RmL}{2} \right|, \left| RL - \frac{RmL}{2} - \frac{R(L - mL)^2}{2mL} \right|, \left| \frac{(L - mL)^2 R}{2mL} \right| \right) \quad (17)$$

$$S = \begin{cases} \left| \frac{(L - mL)^2 R}{2mL} \right|, & 0 < m < 1 - \frac{\sqrt{12}}{6} \\ \left| RL - \frac{RmL}{2} - \frac{R(L - mL)^2}{2mL} \right|, & 1 - \frac{\sqrt{12}}{6} \leq m \leq \frac{1}{2} \end{cases} \quad (18)$$

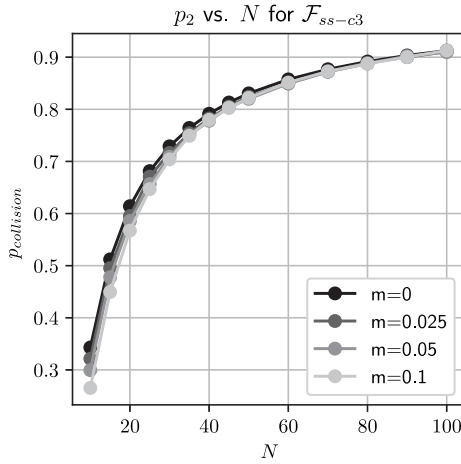
which allows us to determine:

$$R = \frac{2SmL}{(L - mL)^2} \quad (19)$$

where  $0 < m \leq \frac{1}{3}$ . As a brief note, we define  $0 < m \leq \frac{1}{3}$  for the general case of 3 supports in order to accommodate all supports within total length  $L$  while simultaneously allowing realizations that place supports at any location throughout the domain. Eq. (19) will hold for  $0 < m < 1 - \frac{\sqrt{12}}{6}$ .



**Fig. 8.** (a) Schematic illustration of  $\mathcal{F}_{ss-c3}$ ; (b) Visualization of the amplification of the  $\pm R1$  term defined in Eq. (6) for different support positions  $l_b$  and  $l_c$  (here  $m = 0.1$ , the  $\bullet$  marker indicates the selection of  $l_a = 0$ , the  $\times$  markers indicate choices where  $|l_a - l_b| < mL$  or  $|l_a - l_c| < mL$  or  $|l_b - l_c| < mL$ , and we discretize 11 evenly spaced potential support positions  $[0, L]$ ); (c) Schematic illustration of a pair of loads that we believe will have the highest probability of a hash collision for mechanical systems in  $\mathcal{F}_{ss-c3}$ .



**Fig. 9.** Simulated probability of a collision ( $p_{\text{collision}}$  with respect to the discretization of the input load ( $N$ ) for different values of  $m$ . For  $m = 0$  the curve plotted matches Eq. (21).

After we have shown that we can compute  $R$  for  $p_1 = 1$ , we then need to show that for some value of  $c$  greater than 1,  $p_1$  will be greater than  $p_2$ . To do this, we conceptualize a worst case example of  $w_1(x)$  and  $w_2(x)$  where substantially different loads will lead to hash collisions by considering the two loads illustrated in Fig. 8c. Here,  $w_1(x) = 0$  and we can define  $w_2(x)$  as a piecewise function:

$$w_2(x) = \begin{cases} 0, & x < \frac{tL}{N}, \\ -cR/2, & \frac{tL}{N} \leq x < \frac{(t+1)L}{N}, \\ cR, & \frac{(t+1)L}{N} \leq x < \frac{(t+2)L}{N}, \\ -cR/2, & \frac{(t+2)L}{N} \leq x < \frac{(t+3)L}{N}, \\ 0, & \frac{(t+3)L}{N} \leq x \end{cases} \quad (20)$$

where  $t$  is an integer, and  $L/N$  with  $N \geq 3$  represents the discretization of the load, and  $c > 1$  following the definition introduced with Eq. (1). In this case, a hash collision will occur when either: (1)  $c$  is small enough that  $w_2(x)$  will always lead to a change in support force  $< S$  regardless of where the support positions are located, or (2) when the support positions defined by distances  $l_a$ ,  $l_b$ , and  $l_c$  (see Fig. 7b-i) all fall outside the range  $[tL/N, (t+3)L/N]$ . To satisfy the conditions for locality sensitive hashing laid out in Eqs. (1)–(2), we need to show that arbitrarily

far apart functions (i.e., large  $c$  and thus large distance between functions  $d(w_1, w_2) = cR$ ) experience a hash collision with  $p_2 < p_1$ . In other words, we need to determine if there is an upper bound on  $p_2$  as values of  $c$  become arbitrarily large. To do this, we consider scenario (2) and compute  $p_2$  as the probability that distances  $l_a$ ,  $l_b$ , and  $l_c$  all fall outside the range  $[tL/N, (t+3)L/N]$  as a function of our discretization  $N$  and our minimum distance between support  $m$ . In Fig. 8c, we plot  $p_2$  vs.  $N$  for multiple values of  $m$ . Note that for  $m \rightarrow 0$ , we can readily compute:

$$p_2 = \left[ \frac{N-3}{N} \right]^3 \quad (21)$$

as the probability that all three supports will be placed outside of the  $tL/N \leq x \leq (t+3)L/N$  zone if their locations are randomly generated. From Fig. 9, we also see that higher values of  $N$  lead to higher values of  $p_2$ . However, it is clear that even for this worst case of comparison points with large  $c$ ,  $p_2 < p_1$  and thus the system is  $(R, cR, p_1, p_2)$ -sensitive. Notably, this simplest case example paves the way for the investigation of more complex mechanical systems as  $(R, cR, p_1, p_2)$ -sensitive.

## Appendix B. Example problem additional details

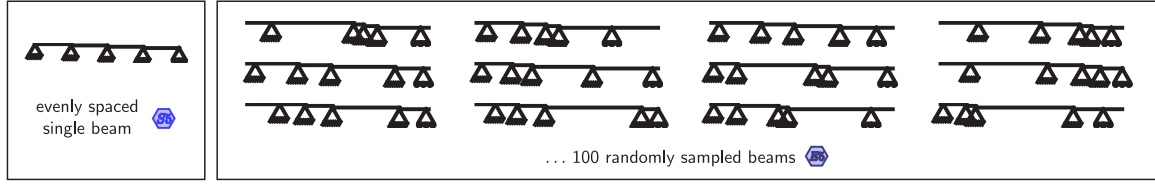
In Section 2.3, we define the example problem that will lead to the results proposed in Section 3. Here we provide additional information to ensure that our problem definition is clear.

### B.1. Simply supported ensemble

In Fig. 3b-d, we illustrate the mechanical systems that we explore as hash functions. Here, in Fig. 10, we explicitly illustrate what we mean by an “ensemble” of simply supported beams. Namely, each ensemble contains 100 simply supported beams with randomly generated support locations. Each one of these devices may lead to a different load class prediction, and the final prediction of the ensemble is the hard voting based outcome of combining all 100 of these predictions. In hard voting, the class labels with the highest frequency from the ensemble predictions becomes the final prediction.

### B.2. Applied load categories

The applied loads introduced in Section 2.3.1 are described in more detail as follows:



**Fig. 10.** Explicit illustration of the difference between simply supported “S” and simply supported ensemble “E” mechanical systems for the case with 5 supports.

- Class 1, constant load ( $c = 1/L$ ):

$$w_1(x) = -c \quad (22)$$

- Class 2, piecewise linear load ( $c = 2/L$ ):

$$w_2(x) = \begin{cases} -c + \frac{2c}{L}x, & 0 \leq x < \frac{L}{2}, \\ (\frac{L}{2} - x)\frac{2c}{L}, & \frac{L}{2} \leq x \leq L \end{cases} \quad (23)$$

- Class 3, piecewise linear load ( $c = 2/L$ ):

$$w_3(x) = \begin{cases} -\frac{2c}{L}x, & 0 \leq x < \frac{L}{2}, \\ \frac{2c}{L}x - 2c, & \frac{L}{2} \leq x \leq L \end{cases} \quad (24)$$

- Class 4, linear load ( $c = 2/L$ ):

$$w_4(x) = \frac{c}{L}x - c \quad (25)$$

- Class 5, linear load ( $c = 2/L$ ):

$$w_5(x) = -\frac{c}{L}x \quad (26)$$

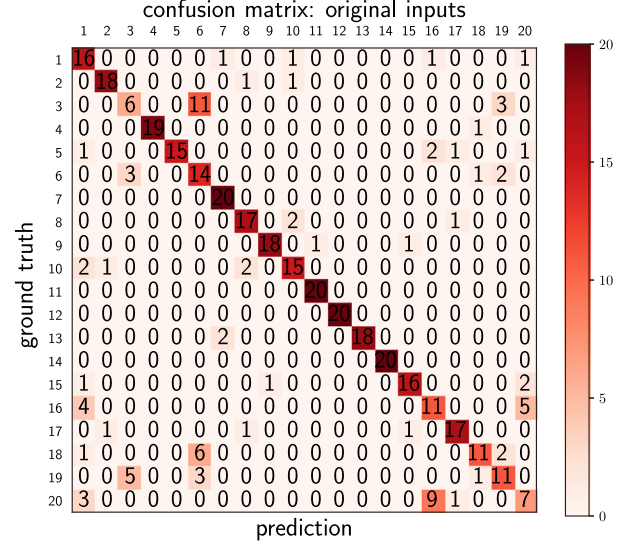
- Class 6, sine wave with wave number  $k = 0.5$  and offset  $\varphi = 0$ :

$$w_6(x) = -\frac{\pi}{2L} \left| \sin\left(\frac{k}{2\pi L}x - 2\pi\varphi\right) \right| \quad (27)$$

- Class 7, sine wave with wave number  $k = 1.0$  and offset  $\varphi = 0$ , see Eq. (27).
- Class 8, sine wave with wave number  $k = 1.0$  and offset  $\varphi = 0.25$ , see Eq. (27).
- Class 9, sine wave with wave number  $k = 1.5$  and offset  $\varphi = 0$ , see Eq. (27).
- Class 10, sine wave with wave number  $k = 1.5$  and offset  $\varphi = 0.25$ , see Eq. (27).
- Class 11, sine wave with wave number  $k = 2.0$  and offset  $\varphi = 0$ , see Eq. (27).
- Class 12, negative kernel density estimate (kde) based on  $n = 2$  points  $p_i$  with uniform random location on the  $x$ -axis. The negative kde for a Gaussian kernel  $\rho_K(x)$  with bandwidth  $h = 0.1L$  is written as:

$$w_{12}(x) = -\frac{1}{nh} \sum_{i=1}^n \exp\left(\frac{-(x - p_i)^2}{2h^2}\right) \quad (28)$$

- Class 13, negative kernel density estimate (kde) based on  $n = 2$  points  $p_i$  with uniform random location on the  $x$ -axis, see Eq. (28).
- Class 14, negative kernel density estimate (kde) based on  $n = 2$  points  $p_i$  with uniform random location on the  $x$ -axis, see Eq. (28).
- Class 15, negative kernel density estimate (kde) based on  $n = 5$  points  $p_i$  with uniform random location on the  $x$ -axis, see Eq. (28).
- Class 16, negative kernel density estimate (kde) based on  $n = 5$  points  $p_i$  with uniform random location on the  $x$ -axis, see Eq. (28).



**Fig. 11.** Confusion matrix based on the original input loads. Note that this confusion matrix corresponds to a classification accuracy of 0.77 across 20 classes.

- Class 17, negative kernel density estimate (kde) based on  $n = 5$  points  $p_i$  with uniform random location on the  $x$ -axis, see Eq. (28).
- Class 18, negative kernel density estimate (kde) based on  $n = 25$  points  $p_i$  with uniform random location on the  $x$ -axis, see Eq. (28).
- Class 19, negative kernel density estimate (kde) based on  $n = 25$  points  $p_i$  with uniform random location on the  $x$ -axis, see Eq. (28).
- Class 20, negative kernel density estimate (kde) based on  $n = 25$  points  $p_i$  with uniform random location on the  $x$ -axis, see Eq. (28).

Each load is set up so that the area under the curve is equal to 1. There are 20 examples for each of the 20 classes of loads. The 20 examples are all differentiated from each other through the addition of Perlin noise with randomly selected initial seed and integer octave in range [2 – 10]. Details for accessing the code to exactly reproduce these loads including the randomly generated Perlin noise are given in Section 5.

### Appendix C. Results additional details

This Appendix contains additional supporting results to supplement the information presented in Section 3. In Section 3, we compare the classification accuracy of our mechanical systems to both random guessing  $accuracy = 0.05$  and direct analysis of the original input data  $accuracy = 0.77$ . Here, in Fig. 11, we show the confusion matrix for nearest neighbor load classification based on the original input data. The purpose of showing this graphic is to demonstrate that we have chosen a challenging suite of applied loads that are non-trivial to distinguish. Next,



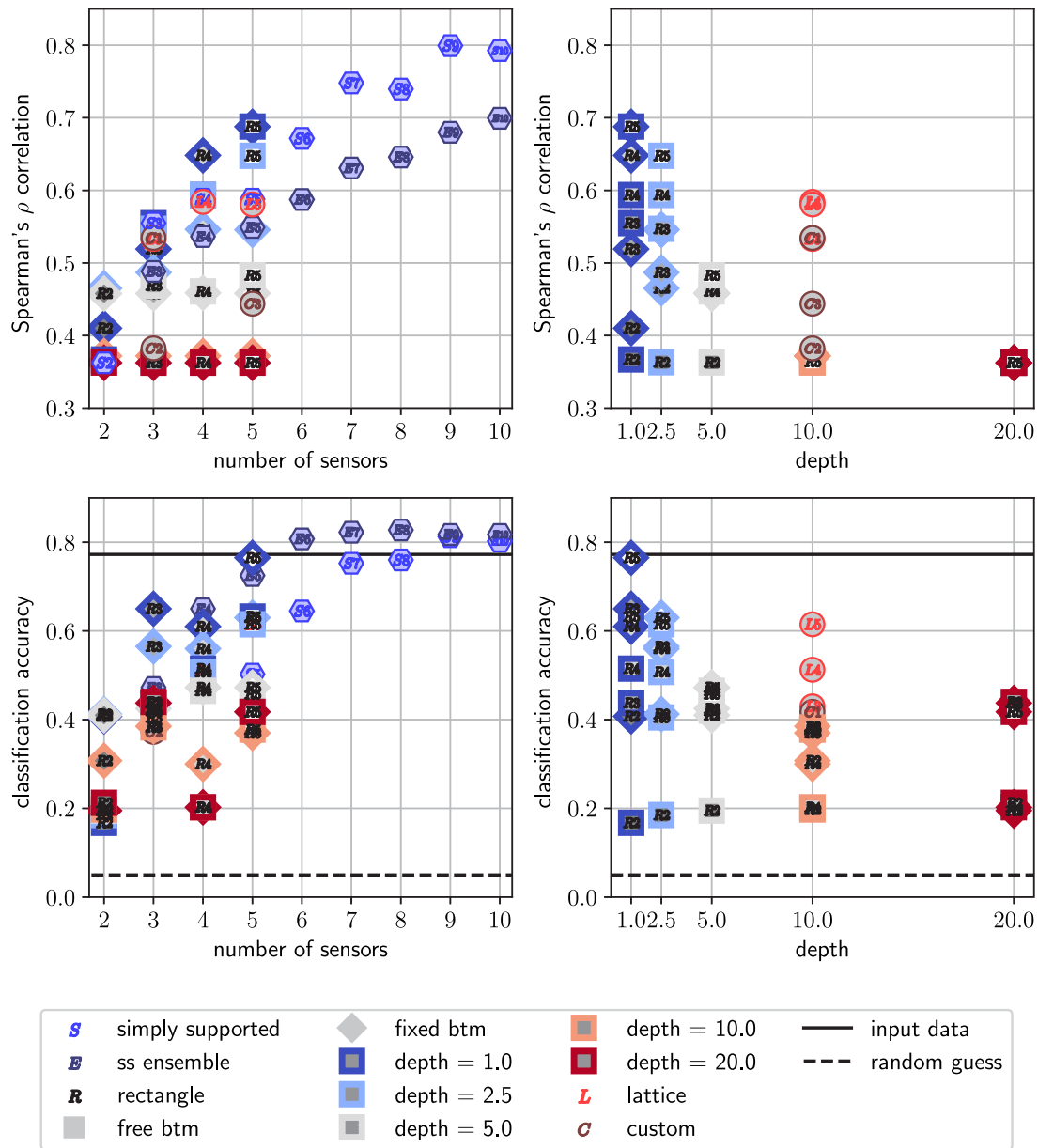


Fig. 12. As a supplement to Fig. 5, we provide Spearman's  $\rho$  and classification accuracy plotted with respect to both number of sensors and domain depth.

for completeness and as an alternative view of the relationship between Spearman's  $\rho$ , classification accuracy, and mechanical domain properties, we provide Fig. 12 as a supplement to Figs. 5 and 6. Finally, we provide Table 1 which contains Spearman's  $\rho$  and classification accuracy for every system investigated in this study. These data are directly visualized in Figs. 5, 6, and 12.

## References

- [1] Shoshana L Das, Prasenjit Bose, Emma Lejeune, Daniel H Reich, Christopher Chen, Jeroen Eyckmans, Extracellular matrix alignment directs provisional matrix assembly and three dimensional fibrous tissue closure, *Tissue Eng. A* 27 (23–24) (2021) 1447–1457.
- [2] Vivek D. Sree, Adrian B. Tepole, Computational systems mechanobiology of growth and remodeling: Integration of tissue mechanics and cell regulatory network dynamics, *Curr. Opin. Biomed. Eng.* 15 (2020) 75–80.
- [3] Tomas Amadeo, Daniel Van Lewen, Taylor Janke, Tommaso Ranzani, Anand Devaiah, Urvashi Upadhyay, Sheila Russo, Soft robotic deployable origami actuators for neurosurgical brain retraction, *Front. Robot. AI* 8 (2022) 437.
- [4] Arincheyan Gerald, Max McCandless, Avani Sheth, Hiroyuki Aihara, Sheila Russo, A soft sensor for bleeding detection in colonoscopies, *Adv. Intell. Syst.* 4 (4) (2022) 2100254.
- [5] Yi Yang, Katherine Vella, Douglas P. Holmes, Grasping with kirigami shells, *Science Robotics* 6 (54) (2021) eabd6426.
- [6] Lillian Chin, Jeffrey Lipton, Michelle C Yuen, Rebecca Kramer-Bottiglio, Daniela Rus, Automated recycling separation enabled by soft robotic material classification, in: 2019 2nd IEEE International Conference on Soft Robotics (RoboSoft), IEEE, 2019, pp. 322–329.
- [7] Ryan L Truby, Robert K Katzschnmann, Jennifer A Lewis, Daniela Rus, Soft robotic fingers with embedded ionogel sensors and discrete actuation modes for somatosensitive manipulation, in: 2019 2nd IEEE International Conference on Soft Robotics (RoboSoft), IEEE, 2019, pp. 322–329.
- [8] Andrew Spielberg, Alexander Amini, Lillian Chin, Wojciech Matusik, Daniela Rus, Co-learning of task and sensor placement for soft robotics, *IEEE Robot. Autom. Lett.* 6 (2) (2021) 1208–1215.
- [9] William D Meador, Mrudang Mathur, Gabriella P Sugerman, Marcin Malinowski, Tomasz Jazwiec, Xinmei Wang, Carla MR Lacerda, Tomasz A Timek, Manuel K Rausch, The tricuspid valve also maladapt as shown in sheep with biventricular heart failure, *Elife* 9 (2020) e63855.
- [10] Tianhong Han, Taeksang Lee, Joanna Ledwon, Elbert Vaca, Sergey Turin, Aaron Kearney, Arun K Gosain, Adrian B Tepole, Bayesian calibration of a

- computational model of tissue expansion based on a porcine animal model, *Acta Biomater.* 137 (2022) 136–146.
- [11] Luke E Osborn, Andrei Dragomir, Joseph L Betthausen, Christopher L Hunt, Harrison H Nguyen, Rahul R Kaliki, Nitish V Thakor, Prosthesis with neuromorphic multilayered e-dermis perceives touch and pain, *Science Robotics* 3 (19) (2018) eaat3818.
  - [12] Ryan L. Truby, Designing soft robots as robotic materials, *Accounts Mater. Res.* 2 (10) (2021) 854–857.
  - [13] John T. Klein, Eduard G. Karpov, Spectral entropy and strain energy trends in composite mechanical metamaterials, *Extreme Mech. Lett.* 45 (2021) 101289.
  - [14] Rudolf M Fuchsli, Andrej Dzyakanchuk, Dandolo Flumini, Helmut Hauser, Kenneth J Hunt, Rolf H Luchsinger, Benedikt Reller, Stephan Scheidegger, Richard Walker, Morphological computation and morphological control: steps toward a formal theory and applications, *Artif. Life* 19 (1) (2013) 9–34.
  - [15] Menachem Stern, Arvind Murugan, Learning without neurons in physical systems, 2022, arXiv preprint arXiv:2206.05831.
  - [16] Elliot Hawkes, B An, Nadia M Benbernou, H Tanaka, Sangbae Kim, Erik D Demaine, D Rus, Robert J Wood, Programmable matter by folding, *Proc. Natl. Acad. Sci.* 107 (28) (2010) 12441–12445.
  - [17] Tian Chen, Mark Pauly, Pedro M. Reis, A reprogrammable mechanical metamaterial with stable memory, *Nature* 589 (7842) (2021) 386–390.
  - [18] Yuanping Song, Robert M Panas, Samira Chizari, Lucas A Shaw, Julie A Jackson, Jonathan B Hopkins, Andrew J Pascall, Additively manufacturable micro-mechanical logic gates, *Nature Commun.* 10 (1) (2019) 882.
  - [19] Charles El Helou, Philip R Buskohl, Christopher E Tabor, Ryan L Harne, Digital logic gates in soft, conductive mechanical metamaterials, *Nature Commun.* 12 (1) (2021) 1633.
  - [20] Zhiqiang Meng, Weitong Chen, Tie Mei, Yuchen Lai, Yixiao Li, CQ Chen, Bistability-based foldable origami mechanical logic gates, *Extreme Mech. Lett.* 43 (2021) 101180.
  - [21] Tian Chen, Osama R. Bilal, Kristina Shea, Chiara Daraio, Harnessing bistability for directional propulsion of soft, untethered robots, *Proc. Natl. Acad. Sci.* 115 (22) (2018) 5698–5702.
  - [22] Yi Zhu, Mayur Birla, Kenn R. Oldham, Evgueni T. Filipov, Elastically and plastically foldable electrothermal micro-origami for controllable and rapid shape morphing, *Adv. Funct. Mater.* 30 (40) (2020) 2003741.
  - [23] Zhengxuan Wei, Ruobing Bai, Temperature-modulated photomechanical actuation of photoactive liquid crystal elastomers, *Extreme Mech. Lett.* 51 (2022) 101614.
  - [24] Helmut Hauser, Physical reservoir computing in robotics, *Reserv. Comput.: Theory Phys. Implement. Appl.* (2021) 169–190.
  - [25] Kohei Nakajima, Helmut Hauser, Tao Li, Rolf Pfeifer, Exploiting the dynamics of soft materials for machine learning, *Soft Robot.* 5 (3) (2018) 339–347.
  - [26] William Gilpin, Cryptographic hashing using chaotic hydrodynamics, *Proc. Natl. Acad. Sci.* 115 (19) (2018) 4869–4874.
  - [27] Johannes Buchmann, Introduction to Cryptography, Vol. 335, Springer, 2004.
  - [28] Jingdong Wang, Heng Tao Shen, Jingquan Song, Jianqiu Ji, Hashing for similarity search: A survey, 2014, arXiv preprint arXiv:1408.2927.
  - [29] Omid Jafari, Preeti Maurya, Parth Nagarkar, Khandker Mushfiqul Islam, Chidambaram Crushev, A survey on locality sensitive hashing algorithms and their applications, 2021, arXiv preprint arXiv:2102.08942.
  - [30] Loïc Paulevé, Hervé Jégou, Laurent Amsaleg, Locality sensitive hashing: A comparison of hash function types and querying mechanisms, *Pattern Recognit. Lett.* 31 (11) (2010) 1348–1358.
  - [31] Lianhua Chi, Xingquan Zhu, Hashing techniques: A survey and taxonomy, *ACM Comput. Surv.* 50 (1) (2017) 1–36.
  - [32] Guido Van Rossum, Fred L. Drake, Python 3 Reference Manual, CreateSpace, Scotts Valley, CA, ISBN: 1441412697, 2009.
  - [33] Ron Rivest, The MD5 message-digest algorithm, April 1992, 1992.
  - [34] Malcolm Slaney, Michael Casey, Locality-sensitive hashing for finding nearest neighbors [lecture notes], *IEEE Signal Process. Mag.* 25 (2) (2008) 128–131.
  - [35] Ken Perlin, An image synthesizer, *ACM Siggraph Comput. Graph.* 19 (3) (1985) 287–296.
  - [36] Perlin-noise python package, <https://pypi.org/project/perlin-noise/>.
  - [37] Charles R. Harris, K. Jarrod Millman, Stéfan J van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, Travis E. Oliphant, Array programming with NumPy, *Nature* 585 (2020) 357–362, <http://dx.doi.org/10.1038/s41586-020-2649-2>.
  - [38] Anders Logg, Kent-Andre Mardal, Garth Wells, Automated Solution of Differential Equations by the Finite Element Method: The FEniCS Book, Vol. 84, Springer Science & Business Media, 2012.
  - [39] Martin Alnæs s, Jan Blechta, Johan Hake, August Johansson, Benjamin Kehlet, Anders Logg, Chris Richardson, Johannes Ring, Marie E Rognes, Garth N Wells, The FEniCS project version 1.5, *Arch. Numer. Softw.* 3 (100) (2015).
  - [40] Jerome L. Myers, Arnold D. Well, Robert F. Lorch, Research Design and Statistical Analysis, Routledge, 2013.
  - [41] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R.J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E.A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, SciPy 1.0 Contributors, SciPy 1.0: Fundamental algorithms for scientific computing in Python, *Nature Methods* 17 (2020) 261–272, <http://dx.doi.org/10.1038/s41592-019-0686-2>.
  - [42] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al., Scikit-learn: Machine learning in Python, *J. Mach. Learn. Res.* 12 (Oct) (2011) 2825–2830.
  - [43] Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, An Introduction to Statistical Learning, Vol. 112, Springer, 2013.
  - [44] Allan F. Bower, Applied Mechanics of Solids, CRC Press, 2009.
  - [45] Javier Tapia, Espen Knoop, Mojmir Mutný, Miguel A Otaduy, Moritz Bächer, Makesense: Automated sensor design for proprioceptive soft robots, *Soft Robot.* 7 (3) (2020) 332–345.
  - [46] Jingdong Wang, Ting Zhang, Nicu Sebe, Heng Tao Shen, et al., A survey on learning to hash, *IEEE Trans. Pattern Anal. Mach. Intell.* 40 (4) (2017) 769–790.
  - [47] Jun Wang, Wei Liu, Sanjiv Kumar, Shih-Fu Chang, Learning to hash for indexing big data—A survey, *Proc. IEEE* 104 (1) (2015) 34–57.
  - [48] Peerasait Prachaseree, Emma Lejeune, Learning mechanically driven emergent behavior with message passing neural networks, *Comput. Struct.* 270 (2022) 106825.
  - [49] Saeed Mohammadzadeh, Emma Lejeune, Predicting mechanically driven full-field quantities of interest with deep learning-based metamaterials, *Extreme Mech. Lett.* 50 (2022) 101566.
  - [50] Miguel A. Bessa, Piotr Glowacki, Michael Houlder, Bayesian machine learning in metamaterial design: Fragile becomes supercompressible, *Adv. Mater.* 31 (48) (2019) 1904845.
  - [51] Fernando V Senhora, Heng Chi, Yuyu Zhang, Lucia Mirabella, Tsz Ling Elaine Tang, Glaucio H Paulino, Machine learning for topology optimization: Physics-based learning through an independent training strategy, *Comput. Methods Appl. Mech. Engrg.* 398 (2022) 115116.