Surveying the Energy Landscapes of Multistable Elastic Structures

SAMUEL JOSEPH AVIS

AVIS, SAMUEL JOSEPH (2023) Surveying the Energy Landscapes of Multistable Elastic Structures. Doctoral thesis, Durham University.

University of Durham

Surveying the Energy Landscapes of Multistable Elastic Structures

Author:

Mr. Samuel Joseph Avis

Supervisor:

Prof. Halim Kusumaatmaja

A thesis presented for the degree of

Doctor of Philosophy

Department of Physics
Stockton Road
Durham
DH1 3LE

31st March 2023

Surveying the Energy Landscapes of Multistable Elastic Structures

Samuel Joseph Avis

Abstract

Energy landscapes analysis is a versatile approach to study multistable systems by identifying the network of stable states and reconfiguration pathways. Thus far, it has primarily been used in microscale systems, such as studying chemical reaction rates and to characterise the behaviour of how protein fold. Here, however, we aim to utilise energy landscape techniques to study multistable elastic structures, in particular, complex 3D structures that have been buckled from 2D patterns, which are of interest for applications such as flexible electronics and microelectromechanical systems.

To this end we have developed new energy landscape methods and software that are well suited to continuous, macroscale systems with many degrees of freedom. The first is the binary image transition state search method (BITSS), which offers greater efficiency for large scale systems compared to traditional transition state search methods, and it is well suited to complex, non-linear pathways. Next, a new software library is introduced that contains a variety of energy landscape methods and potentials which are parallelised to study large-scale continuous systems. This library can be flexibly used for any chosen application, and has been designed to be easily extensible for new methods and potentials.

Furthermore, we exploit energy landscape analysis to tailor the stable states and reconfiguration paths of various reconfigurable buckled mesostructures. We establish stability phase diagrams and identify the corresponding available reconfiguration pathways by varying essential structural parameters. Furthermore, we identify how the introduction of creases affects the multistability of the structures, finding that a small number can increase the number of distinct states, but more creases can lead to a loss of multistability. Taken together, these results and methodology can be used to influence the design of new structures for a variety of different applications.

Statement of Authorship

The work presented in this thesis has been developed under the supervision of Prof. Halim Kusumaatmaja of the Department of Physics at Durham University. All text and figures are the work of the author, unless otherwise stated. No part of this thesis has been presented for any other degree or qualification.

Statement of Copyright

Publications

The following published works form the majority of chapters 5, 7 and 8:

- S. J. Avis, J. R. Panter, and H. Kusumaatmaja, “A robust and memory-efficient transition state search method for complex energy landscapes”, *J. Chem. Phys.*, 157, 124107 (2022).
- Y. Li*, S. J. Avis*, J. Chen, G. Wu, T. Zhang, H. Kusumaatmaja, and X. Wang, “Reconfiguration of multistable 3D ferromagnetic mesostructures guided by energy landscape surveys”, *Extreme Mech. Lett.*, 48, 101428 (2021).
- Y. Li*, S. J. Avis*, T. Zhang, H. Kusumaatmaja, and X. Wang, “Tailoring the multistability of origami-inspired, buckled magnetic structures via compression and creasing”, *Mater. Horiz.*, 8, 3324 (2021).

In the latter two, the experimental work was carried out by Y. Li, while the simulations and energy landscape analysis are my own work.

---

[*] These authors contributed equally to this work.

Acknowledgements

First and foremost, I would like to thank my supervisor, Prof. Halim Kusumaatmaja, for his unwavering support and guidance over the last several years which has made this work possible. I am also deeply grateful to Dr. Jack Panter for his help in getting me started with energy landscapes and providing a basis for my research. I would also like to thank Prof. Teng Zhang, Junbo Chen, and Guangchao Wan for their insightful discussions and introducing me to methods for studying elastic materials, and also Prof. Xueju Wang, Yi Li, and Zizheng Wang, whose expertise and collaboration in experimental work helped advance my research. And lastly, thank you to my parents and my partner, Alison, for their constant support and encouragement, without which I wouldn't have made it this far.

Contents

Chapter 1

Introduction

Energy landscapes have been widely utilised in physics and chemistry to understand the various stable configurations of molecular and particulate systems, and to identify the possible reconfiguration pathways between them. These concepts have been crucial for understanding topics such as the folding behaviour of proteins[1] and the effect of catalysis upon chemical reactions[2], as well as characterising the statistical behaviours of systems such as glasses and amorphous solids[3].

Meanwhile, a separate field of study is that of multistable elastic structures, which are systems that have multiple stable states that can be reconfigured using external stimuli, modifying their functionality or structural arrangement. There are a variety of structures such as mechanical metamaterials[4], which can exhibit unusual mechanical properties, origami[5], which involves folding flat sheets along predefined creases to create complex 3D shapes, and buckled mesostructures[6], which are thin films that buckle into 3D geometries under compression. Such multistable structures have found increasing applications in fields like robotics[7,8], flexible electronics[6,9], and deployable structures[10,11].

The purpose of this thesis is to combine these two fields and to apply energy landscape techniques to continuous, macroscopic applications in general. To achieve this aim, new methods and software are developed to address computational challenges posed by these applications and systems with very large numbers of degrees of freedom, which can often become infeasible using existing approaches. This thesis then explores the energy landscapes of some reconfigurable buckled mesostructures, in particular a table-like structure and a ribbon with origami-inspired creases. The effect of changes to their geometry, material properties, and external forces are studied, as well as how these alter the stability of states and their reconfigurable behaviour. Through this we hope to gain new insights into the fundamental principles of reconfigurable elastic structures and to develop new approaches to design structures with desired functionality.

1.1 Structure of this thesis

Firstly, in chapter 2, the fundamental concepts of energy landscapes will be introduced, as well as an overview of some of the diverse applications that they have been used to study. This is followed, in chapter 3, by an introduction to several different types of reconfigurable elastic structures, as well as some of the recent and ongoing research into these topics. After this, there is a summary of some of the existing studies that have applied various energy landscape techniques to investigate elastic structures. Chapter 4 will then explore a variety of computational methods used to investigate energy landscapes and their particular advantages and disadvantages will be discussed.

The main original contributions of this thesis are then divided into four chapters. The first, chapter 5, introduces a new method for finding transition states, the binary image transition state search (BITSS). This method is evaluated by performing comparisons with other methods for locating transition states, and some of its unique and attractive features are demonstrated, such as the ability to find transition states even in discontinuous landscapes.

Next, chapter 6 outlines the implementation of a software library, ‘ELLib’, that incorporates a variety of energy landscape methods. The primary goal of this is to create a cohesive and adaptable framework that can be applied to a range of complex applications. The library has also been designed from the ground up with parallel computing in mind in order to study large-scale systems. To demonstrate this, an example is presented that involves optimising the geometric parameters of an elastic structure in order to enhance the stability of its two distinct states.

In chapter 7, these ideas are used in the analysis of 3D buckled mesostructures, leading to the creation of stability phase diagrams that are dependent upon the structural parameters. In addition to these, the chapter provides detailed information about the various reconfiguration pathways that are accessible to the structures. The results gained from this may be used to assist in the design of structures with desired behaviour and to inform the remote manipulation between states.

Building upon this, chapter 8 explores the energy landscapes of origami-inspired structures, focusing on the simple arrangement of a single ribbon with multiple creases, which can be used as building block for more complex structures. The chapter examines the impact of compressive strain and the number of creases on the different stable states that are accessible. The findings provide valuable insight into how creases can introduce further multistability to buckled mesostructures.

Finally, chapter 9 brings together the work in this thesis to provide some overarching conclusions, and outlines some possible avenues for future work related to the energy landscapes of elastic structures.

Chapter 2

Energy Landscapes

This chapter aims to introduce key concepts in the theory of energy landscapes and how they can be utilised. These include definitions of salient critical points in the energy landscapes of interest in this thesis. Then, a summary is provided for some of the various applications to which they have been used.

2.1 Features of energy landscapes

An energy landscape describes how the energy of a system changes throughout its configuration space. These are often very high-dimensional landscapes, as the total number of degrees of freedom are typically very large, particularly when modelling continuous macroscale systems.

The primary points of interest are the stable states. These correspond to local minima in the energy landscape, as any small perturbation would raise the energy and result in a driving force restoring it back to the minimum. Mathematically, the energy in the vicinity of a minimum at $\boldsymbol{x}_0$ can be expressed as,

$$E(\boldsymbol{x}) = E(\boldsymbol{x}_0) + \sum_i \frac{a_i}{2} \left.\frac{\partial^2 E}{\partial x_i^2}\right|_{\boldsymbol{x}_0} (x_i - x_{0,i})^2, \tag{2.1}$$

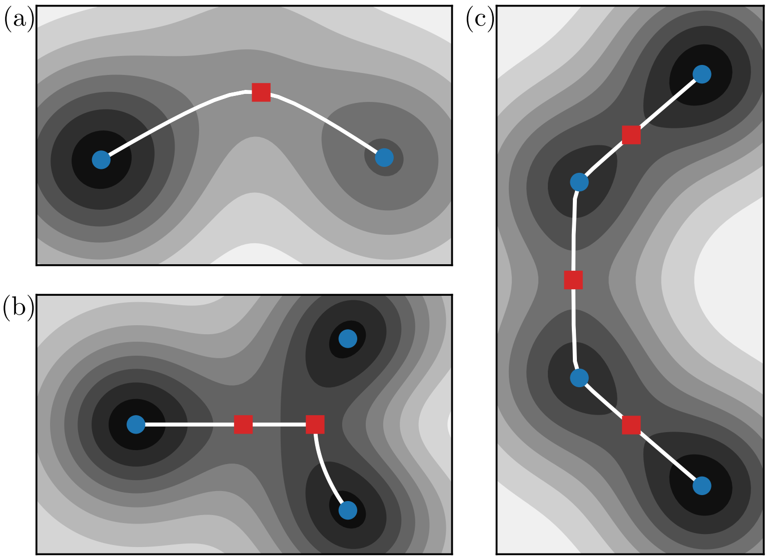


Figure 2.1: Schematics showing 2D energy landscapes with the features of interest. Blue circles represent the minima, red squares show the transition states, and white paths show the MEPs. (a) A simple landscape with two minima connected by a single MEP. (b) A landscape where an MEP should terminate at a transition state. However, running downhill from the left-most transition state goes past this and continues to a minimum. (c) A multi-step pathway with two intermediate minima. The full MEP between the top and bottom minima is a combination of the three individual MEPs.

where i represents each of the principal directions of curvature, and all of the curvatures, a_i, are positive. Using energy minimisation methods to identify these stable states is central to the exploration of energy landscapes. As an example, fig. 2.1(a) shows a 2D landscape containing two minima, as well as highlighting other features of interest.

Each minimum also possesses a basin of attraction, which is the surrounding region of points that will converge to the minimum upon following the gradient downhill. Taking the analogy of a mountainous landscape, these basins correspond to valleys for which the dividing lines are ridges and watersheds. The energy depth and the extent of a basin can be used to quantify the stability of the state; the former providing the energy required by a interaction to change the state, while the latter provides the required perturbation size.

Transition states (TS) are first-order saddle points on the energy landscape. These are stationary points where the energy is defined by eq. (2.1) with a single negative curvature, $a_1 < 0$, while all others are positive or zero, $a_{i>1} \geq 0$. These correspond to minima on the ridges separating basins of attraction, so they provide the lowest possible energy required to swap between two stable states. This energy barrier is incredibly useful for quantifying the stability of states and the probability of a transition between them[12]. The energy barrier is also an important quantity to estimate transition times in microscale systems, such as chemical reactions, using transition state theory[13]. Each transition state also uniquely defines a possible reconfiguration pathway, known as the minimum energy pathway (MEP).

The MEP, $\boldsymbol{x}(s)$, connects the transition state to its adjacent minima by following the path of steepest descent from either side of the saddle point,

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}s} = \hat{\tau} = -\frac{\boldsymbol{\nabla} E(\boldsymbol{x})}{|\boldsymbol{\nabla} E(\boldsymbol{x})|}. \tag{2.2}$$

Here the pathway parameter s is the integrated path length, so the derivative is the unit tangent vector of the pathway. Therefore, if a transition state has already been found it is a simple process to obtain the MEP. An equivalent statement, that can provide a convergence criterion for pathway search algorithms, is that the local gradient has no component perpendicular to the direction of the pathway at all points on its length,

$$\boldsymbol{\nabla}^{\perp} E(\boldsymbol{x}) = \boldsymbol{\nabla} E(\boldsymbol{x}) - (\hat{\boldsymbol{\tau}} \cdot \boldsymbol{\nabla} E(\boldsymbol{x}))\hat{\boldsymbol{\tau}} = 0. \tag{2.3}$$

Sometimes the precise MEP can end at a transition state rather than a minimum if the steepest-descent pathway reaches a bifurcation point[14]. A separate MEP then runs downhill from this to the minima (or to other transition states). However, in practice, small deviations from the steepest pathway will cause a breaking of symmetry in both experiment and simulations, and it will proceed all the way to one of the minima. An example of this is shown in fig. 2.1(b). MEPs can also be obtained between states that are separated by intermediate minima. In this case, the MEP will consist of the combination of the individual MEPs between each adjacent pair of minima, each passing over a separate transition state (fig. 2.1(c)). Hence, it is possible to build up a network of pathways between adjacent minima to identify all possible reconfiguration pathways of a system. A single pair of minima can also have multiple pathways between them, even if they are adjacent to one another. These different pathways can often be difficult to locate and there can be a large number of them in a high-dimensional landscape. However, the most important are those with the lowest energy and most direct, which are the most easily identifiable.

The MEPs that are located are time independent, providing no direct information about the dynamics along the pathway. On one hand, this allows energy landscape methods to be more efficient than performing the corresponding dynamical simulations, such as molecular dynamics or finite element analysis, particularly in the case of rare event transitions with barriers larger than $k_\mathrm{B}T$[15]. On the other hand, the true pathway taken by the system will depend upon the force driving the transition, and inertial effects can cause it to deviate from the steepest descent pathway. Therefore, the MEP instead provides an ideal pathway that is independent of the force driving the reconfiguration, and provides the average and most probable path that will be taken if driven by random forces or thermal noise[16].

2.2 Visualising energy landscapes

It can be difficult to visualise the high-dimensional landscapes, but one simple way of doing so is using disconnectivity graphs[14]. These have a vertical axis showing the energies of all the minima and the transition states with the minimum required energies that connect them. Such visualisations are useful because they clearly separate the energy landscape into a hierarchy of groups of states according to the ease of reconfiguring between them. Additionally, the system can be characterised according to the shape of the graph[17]. If the energy barriers are tall the system will be trapped it in a single state, and it is referred to as a glassy landscape. In contrast, small barriers allow the system to easily explore the landscape. This can be seen by contrasting fig. 2.2(a) and fig. 2.2(b). Furthermore, the energy landscape may be funnel-shaped, in which case it can quickly converge to the global minimum (fig. 2.2(b)), while for a flat landscape (fig. 2.2(c)) the system is not biased toward any particular state.

2.3 Free energy landscapes

An important concept for characterising the behaviour of a system is that of free energy landscapes. Until now the discussion has focused mainly upon potential energy landscapes which differ from free energy landscapes in a couple of key ways. Firstly, while potential energy landscapes are high-dimensional and depend upon all coordinates of the system, free energy landscapes often involve averaging over the coordinates to produce a lower-dimensional system that represents the overall behaviour[14]. Also, the free energy includes information about the thermodynamic nature of the system, because separate minima in the potential energy can correspond to a single minimum in the free energy if it can freely swap between them. Therefore, the free energy landscape changes depending upon the temperature.

Fig. 2.2(d) shows an example of this for a particle cluster, where the potential energy landscape has been reduced to a two-dimensional free energy landscape. These two order parameters are the potential energy, V, and a face-centered cubic packing order parameter, Q_4. Although, in this example there are no free energy minima with high Q_4. At a high temperature there is just a single minimum for high potential energy states, which correspond to a liquid phase. As the temperature lowers another minimum appears with low potential energy, representing the crystallised structure. At this point the liquid and crystal states are both stable. Finally, at the lowest temperatures the liquid minimum disappears and the cluster will only be found in its highly ordered configuration.

This distinction between free and potential energy landscapes are important for

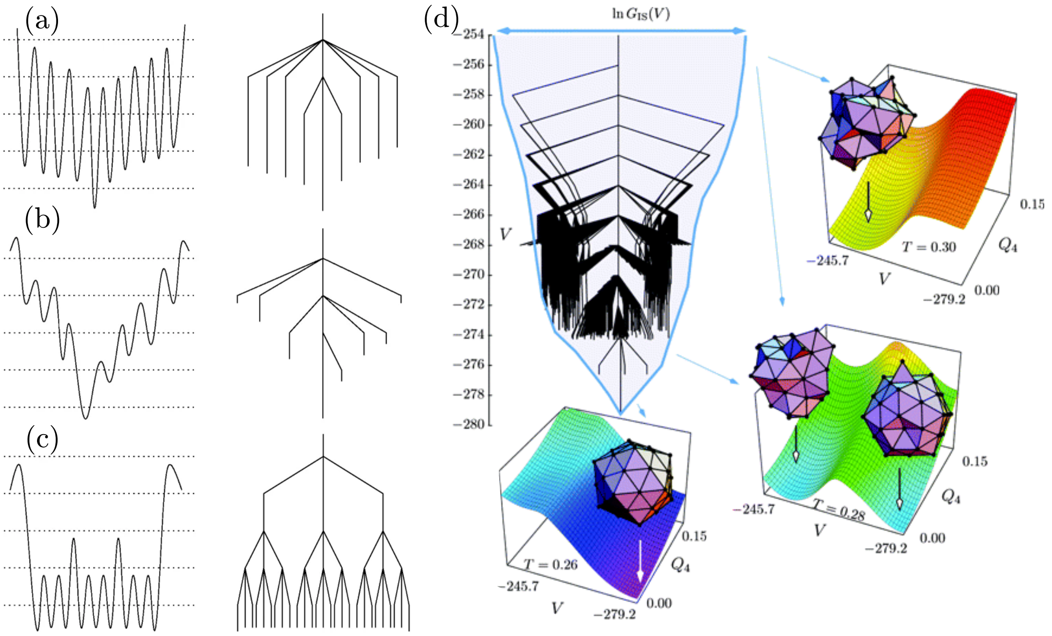


Figure 2.2: (a-c) Disconnectivity graphs for three one-dimensional potentials, each exhibiting a different behaviour (see text for discussion). (d) Disconnectivity graph for a 55 particle cluster, with the envelope size representing the number of states at, or below, that energy. Three 2D free energy landscapes show the states of the system at different temperatures. Representative configurations are shown in each of the minima. Adapted with permission from[18].

microscale and particulate systems where thermal noise is sufficient for the system to spontaneously explore different states. However, for macroscale systems, which this thesis primarily focusses upon, we are already dealing with effective free energy models that correspond to a thermodynamic average of microstates. Although there may be some random perturbing effects, such as vibrations, we generally do not consider any further averaging of the states. Therefore, for brevity, we will simply term our study as an exploration of energy landscapes.

2.4 Applications of energy landscapes

Here, we will summarise known applications of energy landscape studies. To illustrate their wide applicability, we will discuss phenomena across several orders of magnitude, from less than 1nm to larger than 1m.

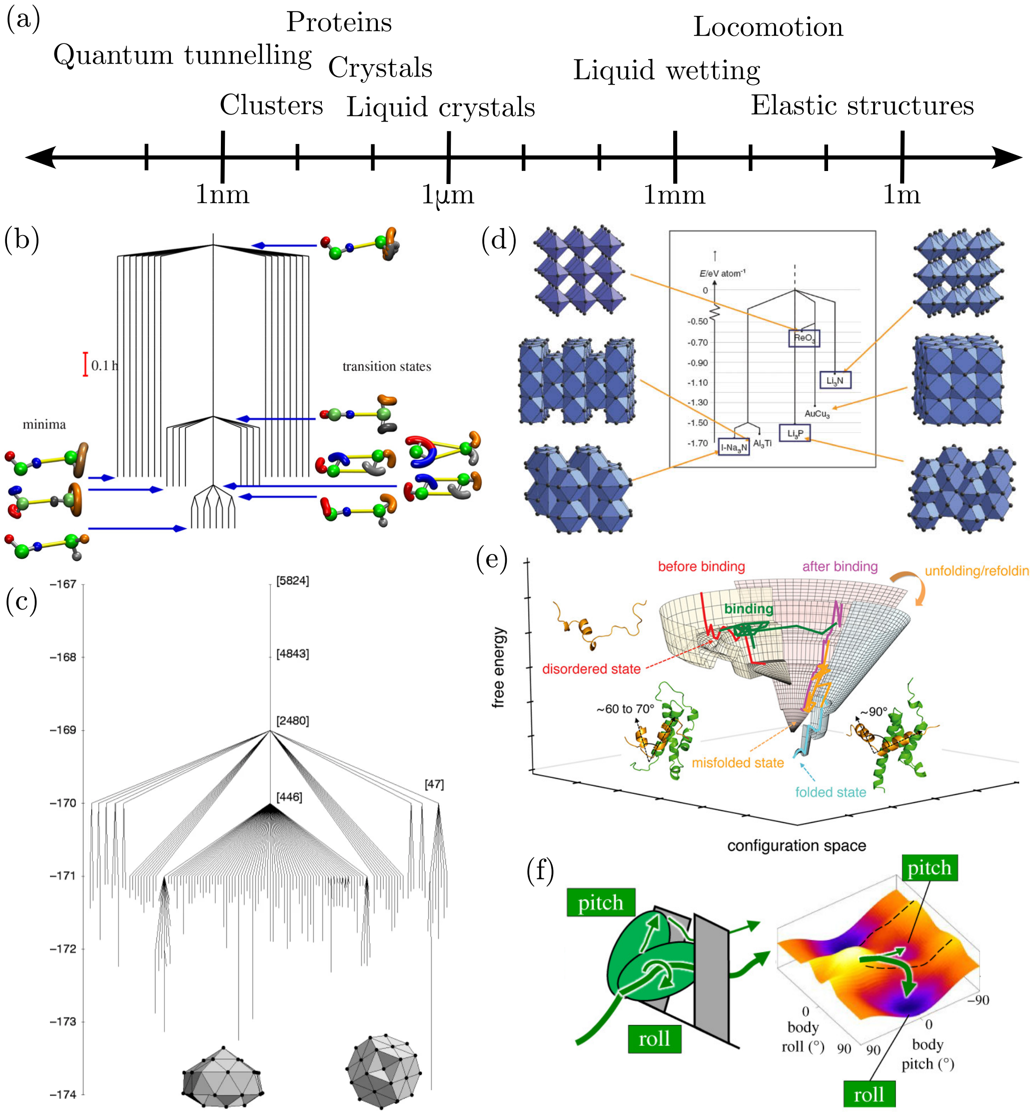


Figure 2.3: (a) Approximate length scales of various applications of energy landscapes techniques. (b) Energy landscape of a pair of hydrogen molecules where the transition states show instanton tunnelling pathways. Adapted with permission from[19]. (c) Energy landscape showing the candidates for the crystal structure of Na_3N. Adapted with permission from[20]. (d) Disconnectivity graph for a Lennard-Jones 38 particle cluster showing a large funnel for the native icosahedral structure, and a smaller funnel for the truncated octahedron structure. Adapted with permission from[17]. (e) Energy landscape of an intrinsically disordered protein, pKID, that folds upon binding with another molecule. Adapted with permission from[21]. (f) Diagram showing the two modes of locomotion to pass a barrier and the associate energy landscape. This shows the locomotor transition from the pitching to rolling motion. Adapted with permission from[22].

To date, energy landscapes exploration techniques see their most prominent applications in particulate and molecular systems. At the smallest scales, these are used to study the transition pathways of chemical reactions, where the stable states represent the reactants and products, as well as any reaction intermediates. Such applications are the source of the term ‘transition state’ which has particular importance for finding reaction rates according to transition state theory.

To obtain accurate chemical reaction rates it is often necessary to consider the effects of quantum tunnelling, which can cause transitions to occur that are classically disallowed. It is possible to use information about energy barriers and MEP to provide tunnelling corrections to the classical rates from transition state theory[23,24]. However, quantum tunnelling depends upon the width of a barrier, not just its height, so shorter paths than the MEP may be more important for determining accurate tunnelling rates. One approach is to use instantons, which are based upon the path integral formulation of quantum mechanics. These provide a temperature-dependent pathway that ‘cuts the corner’ of the MEP, providing a more direct path at the expense of a higher barrier[25,26]. By finding the transition states of these instantons the effective energy barriers of the transition can be determined, providing more accurate tunnelling rates[19].

Another topic important to many chemical reactions is that of catalysis, which has been used with energy landscapes for both biological and industrial processes[2,27–29]. This involves a catalyst introducing intermediate reaction steps in order to reduce the maximum energy barrier and increase the reaction rate[2]. These additional reaction steps, as well as multiple possible competing pathways[28], make it necessary to study the entire energy landscape in order to gain an accurate understanding of the effect that a catalyst will have.

Moving to slightly larger scales beyond chemical reactions, energy landscapes can be used to investigate the formation and rearrangement of complex structures, of which protein folding is perhaps the most pertinent[1,30]. Proteins are made up of a very long chain of amino acids and, since the number of stable states tends to grow exponentially with system size, can fold into 3d structures with an incredibly large number of configurations. In fact, assuming the protein performs a random walk through these structures it may take considerably longer than the age of the universe to find a given state. And yet, in nature many proteins spontaneously fold to a single state at short timescales. This is Levinthal's paradox[31], and can be resolved using energy landscapes[1]. The landscapes of these proteins are found to be highly funnel-shaped and so the protein does not perform a random search, but is instead driven quickly towards the global minimum configuration. Many proteins can also be found in multiple states, or are intrinsically disordered proteins that do not fold into any particular states[21]. Such proteins can be identified by the shape of their landscapes, according to the number of funnels, or if the landscape is flat, respectively. The misfolding of proteins is also important to study[32], owing to the role aggregates of misfolded proteins play in diseases such as Alzheimer's and Parkinson's, or the infectious misfolding that causes prion diseases. In this case the transition pathways can provide insight into cause of the misfolding, as well as the rate at which it occurs.

Particle clusters, which have widespread applications from catalysis to nanotechnology, are another well studied subject in the context of energy landscapes. Here the landscapes are often characterised using disconnectivity graphs in order to understand the structural arrangements formed by a given system, as well as properties such as relaxation timescales[17,33] As an example, the Lennard-Jones 38 particle cluster is found to have a double funnel landscape producing either a truncated octahedral or icosahedral structure[34]. The truncated octahedron is the global minimum configuration and so might be expected to be favoured. However, the icosahedral funnel is much wider with more states, explaining why this state is generally favoured. Related are crystalline solids, which differ due to their lack of a surface, but use energy landscape analysis similarly. For example, free energy minima have been surveyed to predict the stable structures for a given chemical system[20,35,36], and energy barriers have been used to study phase transitions[37], nucleation[38], and defects[36,39].

In contrast to crystal structures, which tend to have simple funnel landscapes, amorphous solids, glasses, and granular materials all have very complex energy landscapes with high barriers that cause them to get stuck in local minima when cooled below a critical temperature. Due to the huge complexity in these structures the statistical properties of the energy landscapes are investigated to gain insight into their behaviour[3,40]. Notably, the glassy landscapes tend to contain fractal basins that result in a secondary relaxation, even if large scale rearrangements are blocked[41]. This causes aging of the materials as the structure switches to progressively deeper basins in the landscape[3]. Shear strain has been found to affect this aging process by moving into shallower or deeper minima depending on size of strain[42,43]. Studies have also looked at glassy landscapes for a variety of different topics, such as active matter[44] and soft glassy materials[45].

Another topic involving the collective arrangement of many particles is liquid crystals[46–48]. Here, the transitions between different orientational arrangements are of particular interest for their applications in liquid crystal displays. Much like the other microscale systems described so far, these can be studied using particulate models[46]. In this case, each particle is given an orientation and anisotropic interactions are used. On the other hand, because liquid crystal systems involve a large number of particles, continuum phase-field models can be used to represent the smooth variation of the particle orientations[47,48]. Liquid crystals can therefore be considered somewhere between the microscale, particulate systems, and the macroscale which are often represented by continuum models.

Progressing to larger scale systems, liquid wetting on surfaces has also been studied using energy landscapes[49,50]. Similarly to the liquid crystals, these often use a continuum phase-field model that uses a smoothly varying order parameter to represent different fluids. Wetting is important for both biological systems and engineering applications, and one topic relevant to both is structured surfaces. These can introduce superhydrophobicity by suspending droplets so they can be easily shed, keeping the surface dry. Using energy landscapes these structures have been investigated to identify how this suspended droplet state may fail, and how this can be avoided[49,50].

Even highly non-equilibrium systems can benefit from the use of energy landscapes, for example, transition pathways have been used to describe locomotor transitions in animals[22,51]. This represents different modes of motion (climbing, walking, rolling, etc.) as minima in a free-energy landscape. This landscape is modified by interaction with the environment, which changes the energy barriers associated with locomotor transitions, and can cause them to occur spontaneously. This is seen, for example, in the motion of a cockroach either climbing or rolling past grass-like barriers[51].

Among some of these largest scale applications using energy landscapes (but by no means the upper limit) are elastic structures. Here the primary interest lies in the structural changes between different states, so pathway analysis will be important. In the following chapter, the different types of elastic structures will be discussed as well as the existing work that has been performed using the theory of energy landscapes.

Chapter 3

Elastic Structures

Elastic behaviour is ubiquitous in both nature and engineering applications, and in many cases, they exhibit multistability and can reversibly change between different configurations. Such systems have wide-ranging potential applications, such as deployable structures[10,11,52], micro-electromechanical systems[53,54], robotics[7,8,55–57], and energy absorption[4,58]. These are often complex structures, but the multistable behaviour is usually derived from simple elements, such as buckled beams or surfaces, and elastic joints or creases. In this chapter various types of multistable elastic structures will be outlined, along with some of the recent and ongoing research on this topic.

3.1 Buckling

Buckling is a sudden change of shape that occurs when a structure is subject to compressive strain. Generally, this is an undesired effect that is associated with structural failure and can have severe consequences, particularly if it occurs in buildings or other infrastructure. Therefore, fields such as structural engineering utilise a range of techniques to prevent buckling and ensure the stability of structures. However, recently there has been significant interest in exploiting buckling transitions to develop flexible devices and structures that can reversibly transform from an undeformed state to a buckled state, or between two different buckled states. Some key examples will be discussed here.

The most basic, and perhaps most versatile, buckled system is that of a column or thin beam under compression. This is possible to be monostable or bistable, depending upon the strain applied and angles at which the end points are held. Previous research has extensively surveyed the stability of the states under different sets of parameters[59]. This can therefore be applied to understand the behaviour of more complex structures constructed from simple beam elements.

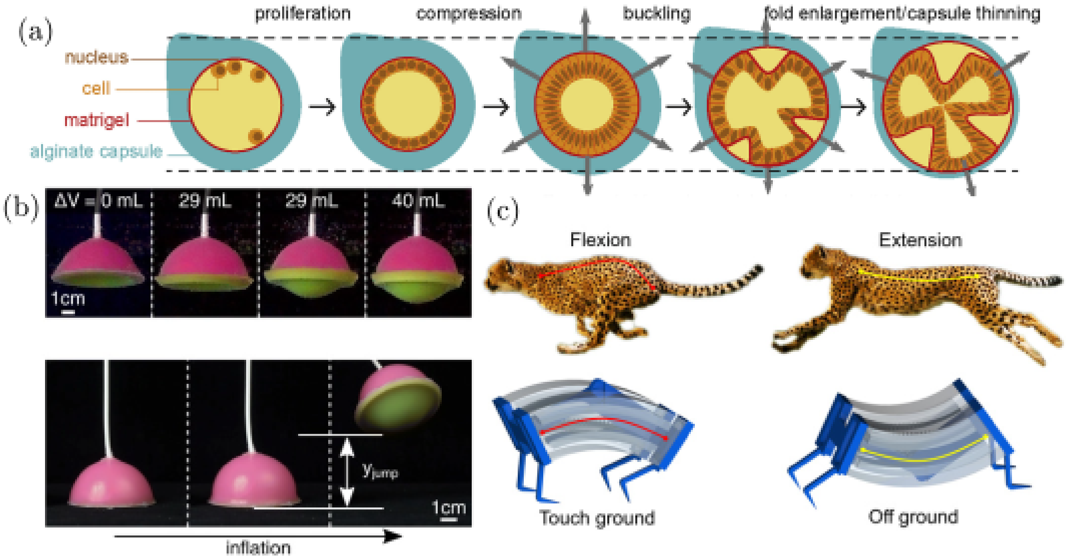


Figure 3.1: (a) Schematic showing how the growth of a membrane of cells can lead to buckling and the formation of folds. Adapted with permission from[60]. (b) A soft jumper that uses the snap-through buckling of a spherical shell upon inflation. Adapted with permission from[7]. (c) A cheetah-inspired soft actuator for robotic movement. A central bistable linkage is used in combination with with inflatable channels to change states. Adapted with permission from[57].

While global buckling modes are typically observed in thin beams, surfaces and thin shells can exhibit local buckle through the formation of dimples on the surface. This is generally undesired in engineering applications, such as containers or rockets, and is important to understand how it occurs in order to prevent and control structural failure. The size of disturbances that are required in order to buckle both cylindrical and spherical shells under strain have been investigated by applying probing forces[61,62]. In addition, various works have investigated the effect of imperfections, which can significantly reduce resistance to buckling[62,63]. These findings highlight the precise manufacturing that is required to ensure that buckling cannot occur.

The buckling of surfaces and membranes also plays an important role in various biological functions. One case is in the formation of folded structures of cell layers, which are essential in many anatomical structures such as glands or crypts. Buckling is a possible mechanism for the formation of these structures. For example, it can be caused by the growth of cells when constrained by a surface[60,64], fig. 3.1(a). Many of these biological membranes, such as lipid bilayers, are a liquid surface that can redistribute to reduce local stress, but these too can buckle. An example of this is the budding of cell membranes, where the binding of curved structures to the membrane can result in stress that causes the surface to buckle[65]. Furthermore, it is possible to use the onset of buckling to measure the mechanical properties of cells[66].

An area where buckling can be highly beneficial is in the development of soft actuators for robotics. In contrast to traditional robotic motion, soft actuators can provide simpler systems that are lighter, less prone to failure, and safer around humans. Using buckling instabilities it is possible to amplify actuation of these soft structures, causing large, sudden displacements which can be used for propulsion. For example, bistable beams have been used to propel swimming robots[56], and the snap-through of spherical shells has been used for inflatable jumpers[7], shown in fig. 3.1(b). Buckling can further be used to emulate biological behaviour such as swimming fish[8], and the galloping of cheetahs[57], fig. 3.1(c). Finally, bistable structures have been used in pneumatic soft grippers that can be pressurised to pick up objects and continue to hold them even when the pressure is released[67,68].

3.2 Buckled mesostructures

The class of elastic structures that will primarily be investigated in this thesis are known as 3D buckled mesostructures. These are produced as a flat pattern which buckles under a compressive strain to form a three-dimensional structure. Typically, this is carried out by binding the structure at various points to a pre-stretched substrate, which is then released, compressing the structure. This provides a simple method for fabricating complex structures, and can be used to create structures down to the microscale. These have been investigated for applications such as flexible electronics[6], microelectromechanical systems[6,69], biomedical devices[70,71], and microrobotics[71].

Many works have investigated the resultant three-dimensional structures for various different precursor patterns. Simple patterns may consist of a single ribbon or a connected grid of ribbons with periodic bonding sites, in which case the ribbons will buckle into a wavy structure, section 3.2(a). Or if the ribbons are curved, they can twist into helical coil shape when buckled[72]. By connecting these ribbons together in complex networks, it is possible to create multilevel, buckled designs. Generally, these structures will buckle up fully. However, if the adhesive force to the substrate is strong ,such as for very small scale structures, it may remain attached to the substrate in various locations[73].

It is possible for some mesostructures to buckle into multiple different stable configurations, such as those seen in section 3.2(b). However, these structures are limited and it may be difficult to obtain more than a single state in practice. This multistability can be enhanced by incorporating origami-like creases into the structure where it can more easily bend[6]. This greatly increases the variety of structures with multiple stable stables and allows greater control over which states are stable. This can be utilised to create switchable devices, such as an antenna that can transmit in one configuration, but is shielded in another, as depicted in section 3.2(c).

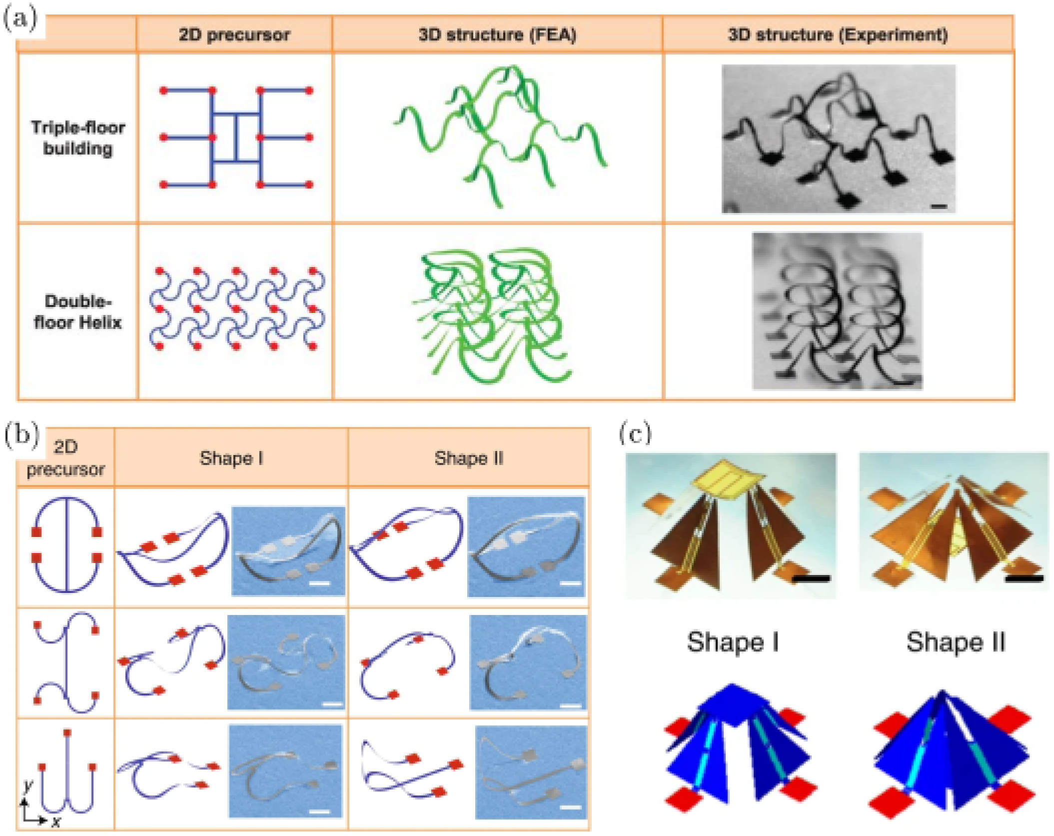


Figure 3.2: (a) Buckled mesostructures with patterns that produce multilevel structures. Adapted with permission from[72]. (b) Multistable mesostructures that can each take two different buckled states. Adapted with permission from[6]. (c) A multistable mesostructure with a radio-frequency antenna on top. In shape 1 the antenna is free to transmit, but in state 2 it is blocked. Adapted with permission from[74].

A number of different approaches have been suggested in order to obtain these different states. Fu et al.[6] discovered a method of producing different shapes by controlling the release of the underlying substrate over time. By independently releasing the substrate in different directions they were able to obtain structures with three separate states. In this approach, the available states are limited to those that are spontaneous under buckling. Park et al.[75] were able to control the shape of the structure over time by dissolving a constricting film. However, this approach is irreversible and the possible shapes are limited. Furthermore, they suggested a method for keeping the mesostructures buckled after removing from the substrate, by interlocking parts of the structure to keep it in place. A different approach has been considered by Hao et al.[76]. Instead of using a stretched substrate, they used structures made from hydrogels which swell, causing the structure to buckle. They were able to control the buckling to obtain various stable configurations by preswelling parts of the structure.

Despite these methods for creating multiple stable states, reconfigurability after the structure is created has largely remained unaddressed. An approach addressing this problem will motivate later work in chapters 7 and 8.

3.3 Mechanical metamaterials

Although we focus primarily upon buckled mesostructures in this thesis, it is worth considering the related, and popular topic of mechanical metamaterials. These are materials that have properties defined by their structure and not the properties of the constituent material. They can show a variety of interesting and unusual properties such as negative Poisson ratio (auxetic materials)[4,77], where the material expands (or contracts) in all directions simultaneously, or negative compressiblity[78,79], where the material contracts under a tensile force. There are many different types of mechanical metamaterials, but each of which is typically made up of a repeating pattern of simple elements, such as beams or hinges[77]. If these individual elements are multistable by undergoing a buckling transition, then the metamaterial itself will be multistable.

Some metamaterials can have large numbers of elements that buckle at once under compression or tension[4,80]. An example is the 3D elastomeric structure shown in section 3.3(a), where the array of holes can collectively buckle, resulting in a highly non-linear compressive behaviour with a negative Poisson ratio[4,81–83]. Similar structures with buckling beams have also been proposed for energy absorption and storage[4,58,80,84]. These have advantages over traditional materials used for collision protection because they can easily unbuckled and reused. The structures can also exhibit hysteresis loops under cyclic loading and unloading, and so can be used for dissipating energy and damping vibration[4]. The order of buckling in these structures is unpredictable because each cell is identical and will depend upon manufacturing imperfections, but this can be controlled by varying the thickness of the bistable beams[80,84].

Energy absorbing metamaterials are generally constructed in arrays, with rows of cells that buckle together, but not all at once. However, for some metamaterials, the transition of one cell can destabilise neighbouring sections and lead to a cascade across the entire structure. These are attractive for realising complex, self-deploying structures[11]. The change of state occurs through transition waves that travel at a constant speed through the structure, behaving similarly to phase transitions observed in crystalline solids[85]. The transition waves can also be controlled and redirected by introducing defects into the periodic structure, as seen in section 3.3(b).

In addition to exhibiting interesting behaviour under compression, metamaterials can also be used to create multifunctional structures that morph into a variety of different configurations[11,86–89]. One approach consists of tessellating prismatic structures with flexible hinges, section 3.3(c), which are able to change into a large variety of different shapes thanks to their 3D structure[88,89]. These can be used to create structures that modify their shape and behaviour based upon changing environments, and be used to perform multiple tasks. These structures are able to remain in the different configurations without an applied force because of multiple stable states resulting from the buckling of the rigid faces[89]. A similar idea is seen in[86]. Here, the authors design a metamaterial that consists of an elastic spherical sphere that can undergo transformation from a large sphere with a relatively open surface to a small sphere with a dense surface. Such a device could offer reversible encapsulation functionality in order to transport small quantities of substances.

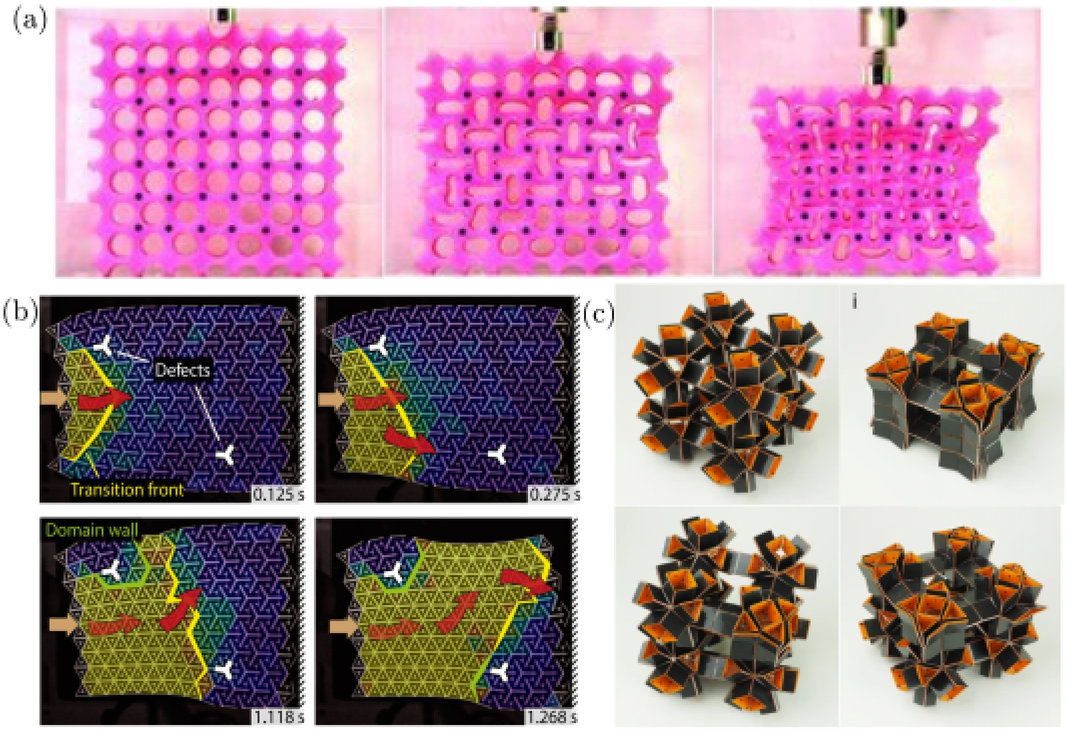


Figure 3.3: (a) An elastic metamaterial with holes that buckle under compression. Adapted with permission from[90]. (b) A transition wave through an auxetic metamaterial that is redirected by lattice defects. Adapted with permission from[85]. (c) A 3D prismatic metamaterial that is able to reconfigure into a variety of stable states. Adapted with permission from[89].

3.4 Origami

Origami structures are a specific type of mechanical metamaterial that have garnered considerable attention in particular. They consist of a network of rigid flat surfaces joined by creases that are free to bend, which enables simple two-dimensional surfaces to form complex three-dimensional structures. Origami is used for both artistic purposes and engineering applications, largely for its ability to produce strong, but light structures and materials[91]. One particularly useful class of origami structures are those that can fold from a flat surface, known as flat foldable, and are therefore easy to produce. The simple reconfigurability of flat foldable structures also makes them suitable for deployable structures, such as solar panels for satellites[92,93].

By considering elastic folds with an optimal angle and allowing some bending of the rigid faces, many origami structures are found to be multistable. Indeed, one of the most basic components of origami (vertices of four creases) is multistable[94] and that these can be tiled to create large reconfigurable structures. One such tiled structure is the Miura-ori pattern, fig. 3.4(b), which is well known for its ability to unfold from a compact structure to a flat surface in a single motion as a result of a negative Poisson ratio[92,93]. For this usage, the structure is elastically monostable and therefore requires actuation to remain in a deployed state. Bistability has been observed, however, through the inversion of singular cells of the Miura-ori pattern to create "pop-through defects". These can be introduced to programmatically modify the stiffness of the structure, while remaining completely reversible[74].

Beyond simple sheets, origami can be used to produce three-dimensional structures. For example, by stacking origami sheets, cellular metamaterials can be produced that are auxetic or can be used for impact energy absorption[95,96]. The behaviour is then easily tunable by modifying the underlying two-dimensional fold pattern. Alternatively, the origami sheets can be curved and turned into tubes and cylinders[10,52]. These tubes are then able to easily extend and collapse simply by deforming the end, providing another useful approach for realising deployable structures.

Kirigami is a variation of origami whereby cuts are introduced into the surface in addition to folding. This allows for structures with greater flexibility and additional stable states, with each end of a cut able to take one of two different states[97]. The introduction of cuts also allows the structure to deform and buckle under compression[98], which has been exploited to create a variety of simple actuators that rotate or lift[99].

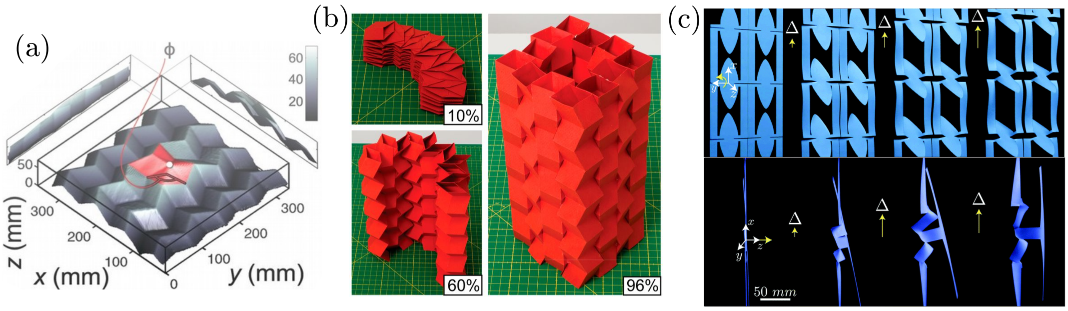


Figure 3.4: (a) A Miura-ori pattern with a pop-through defect (red) that alters the stiffness of the structure. Adapted with permission from [74]. (b) A combination of 12 origami tubes that deploy to produce an enclosed structure. Adapted with permission from [10]. (c) Two kirigami actuators that rotate and lift, respectively, when extended. Adapted with permission from [99].

3.5 Energy landscapes applied to elastic structures

Recently there has been increasing interest in using energy landscape methods to investigate multistable elastic structures [52,61,62,83,89,100,101]. Since these structures are primarily macroscopic, thermal effects are insignificant and are external forces are needed to reconfigure them instead. Correspondingly, there is no need to investigate their statistical behaviour. Instead, energy landscape analysis can be used to design structures, and to inform how best to reconfigure them.

Perhaps the most common use of energy landscapes in the field is to determine the various states that a given structure can take for given structural parameters. For example, this has been used in the study of some reconfigurable mechanical metamaterials. Iniguez-Rabago et al. have looked into several shape-changing 3D prismatic structures [89]. By varying just two of the fold angles they were able to produce a simplified landscape, shown in section 3.5(a), displaying the stable regions of each of the structure's various states. A similar survey was also carried out for the elastomeric metamaterial with cylindrical holes that buckles under compression [83]. To understand how it behaves when compressed, all of the possible states were identified at increasing levels of strain, section 3.5(b). Furthermore, by introducing slight imperfections to the structure they were able to identify which of the states are obtained upon experimentally deforming the system.

Energy landscapes have also been utilised in the study of origami structures. In particular, these have looked at the reconfiguration pathways, which are often complex and non-linear due to the folding through large angles. Silverberg et al.[100] investigated the square twist structure which undergoes a bistable transition from an unfolded to folded state, during which the faces are required to bend. They were able to identify how the rigidity and shape of the structure affected the transition pathway and its barrier, and thereby how they could be tuned for specific applications. Another use of the pathway is demonstrated for an origami cylinder[52], which is easily deployable as a result of the small energy barrier. In this case, the MEP causes the cylinder to first extend beyond its deployed state before returning. Therefore, the structure must take a different pathway when the deployed structure is placed under compression, ultimately resulting in a much higher energy barrier and stiffness. However, in each of these cases the pathways have been identified by applying forces to the structure, rather than locating the MEP.

Recently, MEP methods have also enabled the identification of completely rigid folding pathways in origami structures, even when crease bending energy is disregarded, such that there is no energy barrier between the unfolded and folded states. By employing a pathway search algorithm that was fixed at the two end points with an energy that penalised any deviations from rigid folding, Zhou et al. were able to locate the rigid folding pathway for the Miura-ori pattern[102]. A similar method could also be applied to other reconfigurable structures that are not necessarily multistable, as long as the two end states to be known.

In addition to providing information about desired reconfigurabilty, pathway analysis can be used to identify the failure modes of a system. For example, the buckling of spherical and cylindrical shells under axial compression is hard to predict and depends strongly upon the size of any imperfections[62]. However, by studying the energy barrier it is possible to identify how these imperfections affect the critical pressure before buckling occurs[61,62]. Further work has also been undertaken to identify how the shape of the cylinders affects the structure of the energy landscape[101]. Panter et al. identified that short cylinders had simple funnelled landscapes, and long cylinders had complex landscapes with many more minima. Furthermore, using the information about the initial buckling transition state, they were able to modify the distribution of the structure's thickness in order to maximise the energy barrier, and to control the energy landscape, reducing the number of states, section 3.5(d).

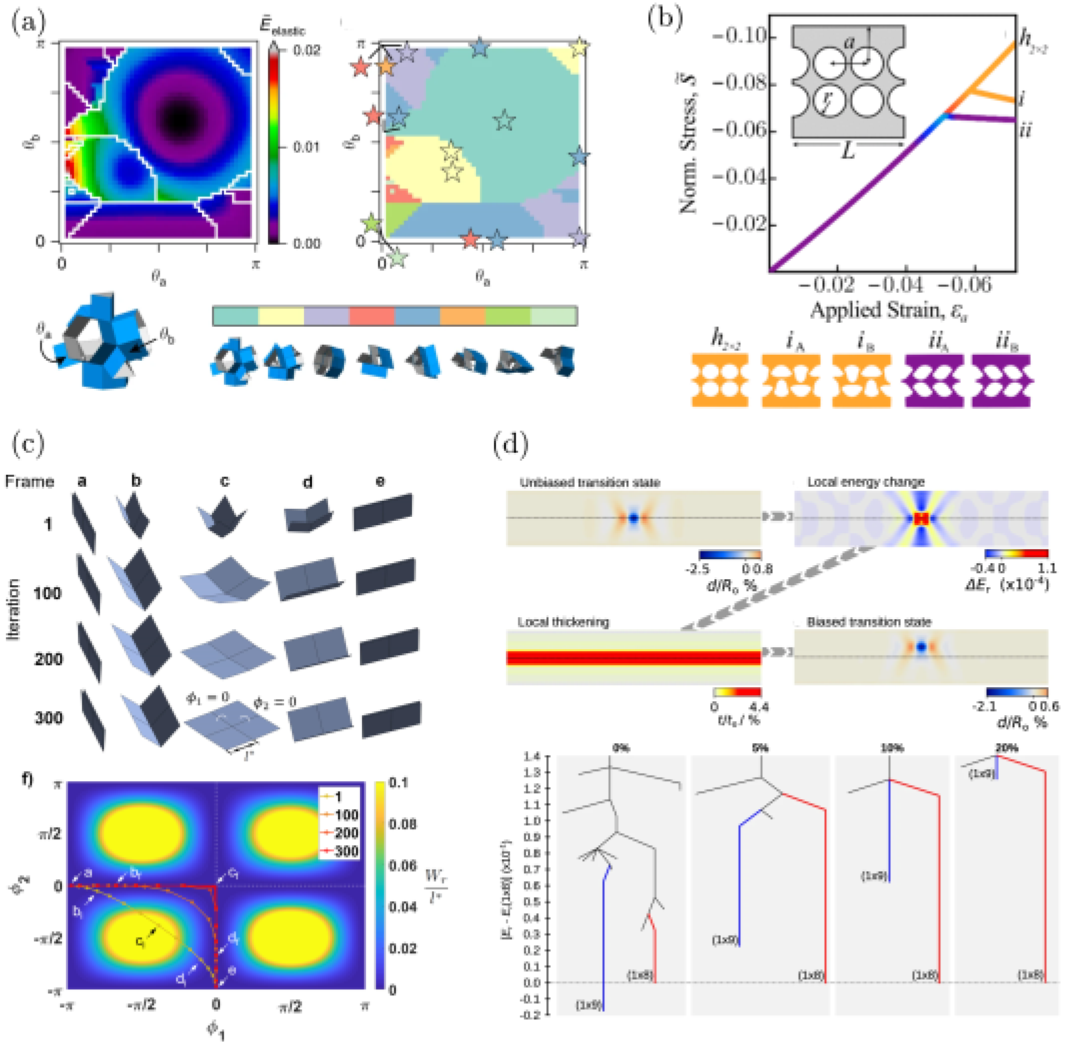


Figure 3.5: Energy landscape methods applied to various elastic structures. (a) The landscape of a prismatic structure when varying two of the angles. The elastic energies of each configuration are shown on the left, and the basins of attraction for each state are on the right. Adapted with permission from [89]. (b) The stable states of a 2×2 holar metamaterial at different values of the compressive strain. Adapted with permission from [83]. (c) Identifying the rigid folding pathway of a Miura-ori vertex using an NEB pathway. The top shows snapshots along the pathway as it is optimised. At the bottom the 2D energy landscape is shown. The pathway is refined towards the rigid folding pathway where the energy is zero. Adapted with permission from [102]. (d) Using information about the transition state of the buckling of a cylinder to alter the buckling behaviour. The top shows how the thickness is increased where the bending energy of the transition state is locally higher, resulting in an altered transition state. The bottom shows that, as the thickness is increased, the landscape changes and a low energy state is suppressed. Adapted with permission from [101].

Chapter 4

Energy Landscape Methods

In this chapter we will explore some of the existing methods for energy landscape exploration. Then we will discuss the advantages and disadvantages for each of the methods, including their use for elastic structures and other continuous systems. It is not feasible to discuss all such methods, so we will restrict our focus to some of the most commonly used approaches.

Firstly, we will look at energy minimisation methods, which are central to the exploration of energy landscapes. They are used to locate local minima in the landscapes, and are utilised in almost all energy landscape methods to find saddle points and minimum energy pathways. So, efficient minimisation algorithms are crucial.

After this we will look at methods for locating transition states (TS). These can use a single state to find a nearby TS, so-called single-ended methods, or can use two states that converge towards a TS from either side, known as double-ended methods. Finally, we will look at some methods for locating minimum energy pathways using a chain of states. These methods are also often referred to as double-ended transition state methods, while the bracketting methods are often overlooked. However, here we will not be refer to them as such, and disambiguate between the two groups according to whether a method is primarily locating the pathway or the TS.

4.1 Energy minimisation

4.1.1 Gradient descent

Perhaps the simplest minimisation algorithm is gradient descent. In this method a point on the landscape is iteratively refined by taking steps in the direction opposite to the gradient. This causes the point to follow the path of steepest descent towards the local minimum. Explicitly, the coordinates $\boldsymbol{x}_i$ at iteration i are updated according to,

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i - \alpha_i \boldsymbol{\nabla} E(\boldsymbol{x_i}). \tag{4.1}$$

The step size, α_i, may be set to a single value, however this would slow down convergence, and may cause the method to diverge if it is too large. A more common method is to use a line search (described below) to perform a one-dimensional optimisation in the direction of the step.

Line search

Given a search direction $\boldsymbol{p}$ (for gradient descent $\boldsymbol{p} = -\boldsymbol{\nabla} E(\boldsymbol{x})$) the aim of the line search is to find the step size, α, that minimises the function $E(\boldsymbol{x} + \alpha\boldsymbol{p})$. An exact solution could be found, but this is likely inefficient. Instead, it is best to find a value that is good enough and then proceed with the next iteration of the method. Two common ways of determining this are using the Wolfe conditions [103,104] or a backtracking line search [105].

The Wolfe conditions are that the objective function must reduce sufficiently (the Armijo rule), and that the gradient must also reduce sufficiently (the curvature rule):

$$E(\boldsymbol{x} + \alpha\boldsymbol{p}) \leq E(\boldsymbol{x}) + c_1 \alpha \boldsymbol{p} \cdot \boldsymbol{\nabla} E(\boldsymbol{x}), \qquad \text{(Armijo rule)} \tag{4.2}$$

$$-\boldsymbol{p} \cdot \boldsymbol{\nabla} E(\boldsymbol{x} + \alpha\boldsymbol{p}) \leq -c_2 \boldsymbol{p} \cdot \boldsymbol{\nabla} E(\boldsymbol{x}). \qquad \text{(Curvature rule)} \tag{4.3}$$

The parameters c_1 and c_2 must be chosen to satisfy $0 < c_1 < c_2 < 1$. Each of these rules places a bound on the step size, with the Armijo rule placing a maximum bound, and the curvature rule enacting a minimum bound.

The backtracking line search is a little different. The Armijo rule is still used to ensure that the step size is not too large, however the curvature rule is not. Instead, the method first begins with a large step size and incrementally reduces it until it satisfies the Armijo rule.

4.1.2 Newton's method

Newton's method is a second order minimisation method that uses the Hessian to provide a faster rate of convergence than first order methods such as gradient descent. Each iteration of this method effectively uses the gradient and curvature to fit a paraboloid to the energy landscape surface, and then proceed to the critical point. As a result, this method may also find maxima or saddle points instead of local minima.

To obtain the step, $\boldsymbol{s}_i$, to take from a position $\boldsymbol{x}_i$ at iteration i, we consider the second-order Taylor expansion,

$$E(\boldsymbol{x}_i + \boldsymbol{s}_i) = E(\boldsymbol{x}_i) + \boldsymbol{s}_i^{\mathrm{T}} \boldsymbol{\nabla} E(\boldsymbol{x}_i) + \frac{1}{2} \boldsymbol{s}_i^{\mathrm{T}} \mathrm{H}(\boldsymbol{x}_i) \boldsymbol{s}_i, \tag{4.4}$$

where H is Hessian matrix. The step should take it to the critical point, where the gradient is zero,

$$\boldsymbol{\nabla} E(\boldsymbol{x}_i + \boldsymbol{s}_i) = \boldsymbol{\nabla} E(\boldsymbol{x}_i) + \mathrm{H}(\boldsymbol{x}_i) \boldsymbol{s}_i = 0. \tag{4.5}$$

So the step that should be taken at each iteration is

$$\boldsymbol{s}_i = -\mathrm{H}^{-1}(\boldsymbol{x}_i)\ \boldsymbol{\nabla} E(\boldsymbol{x}_i). \tag{4.6}$$

Once the step has been found it is also possible to perform a line search along that direction. This can be used to ensure that the energy always reduces and a minimum is found rather than a saddle point or maximum.

4.1.3 L-BFGS

The limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm[106,107] addresses two difficulties associated with Newton's method. The first is the calculation of the Hessian, which may be complicated or not known for many systems. Furthermore, for an n-dimensional system it has a computational complexity of $O(n^2)$, so it can be computationally costly for many energy landscapes which have a very large number of dimensions. The second is the inversion of the Hessian in eq. (4.6), which is even more costly, scaling as $O(n^3)$. Since L-BFGS is a quasi-Newtonian method which approximates eq. (4.6) rather than calculating the Hessian directly, it does not suffer from these same problems.

It is based upon the BFGS algorithm[108], which continually updates an approximation of the inverse Hessian matrix using the changes to the position and gradient at each iteration. This prevents the need to know the Hessian, or perform the matrix inversion. However, the whole matrix H^{-1} must still be calculated, with a complexity of $O(n^2)$.

L-BFGS addresses this by only storing several vectors representing changes in the positions and gradients of the past m steps ($\boldsymbol{s}_j$ and $\boldsymbol{y}_j$ for $j = i - m, \ldots, i - 1$, respectively), and the approximation of eq. (4.6) is calculated directly. This bypasses the need to calculate the full inverse Hessian matrix, and results in a much lower complexity of $O(nm)$. To do this, two sets of vectors (q_j and z_j for $j = i - m, \ldots, i$) are calculated iteratively based upon the recursive method used by BFGS. The first vectors are given by

$$\boldsymbol{q}_i = \boldsymbol{\nabla} E(\boldsymbol{x}_i), \tag{4.7}$$

$$\boldsymbol{q}_j = \boldsymbol{q}_{j+1} - \left(\frac{\boldsymbol{s}_j \cdot \boldsymbol{q}_{j+1}}{\boldsymbol{s}_j \cdot \boldsymbol{y}_j} \right) \boldsymbol{y}_j \quad \text{for} \quad j = i-1, \ldots, i-m. \tag{4.8}$$

Then the second set of vectors are calculated as

$$\boldsymbol{z}_{i-m} = - \left(\frac{\boldsymbol{s}_{i-1} \cdot \boldsymbol{y}_{i-1}}{\boldsymbol{y}_{i-1}^2} \right) \boldsymbol{q}_{i-m}, \tag{4.9}$$

$$\boldsymbol{z}_{j+1} = \boldsymbol{z}_j - \left(\frac{\boldsymbol{s}_j \cdot \boldsymbol{q}_{j+1}}{\boldsymbol{s}_j \cdot \boldsymbol{y}_j} - \frac{\boldsymbol{z}_j \cdot \boldsymbol{y}_j}{\boldsymbol{s}_j \cdot \boldsymbol{y}_j} \right) \boldsymbol{s}_j \quad \text{for} \quad j = i-m, \ldots, i-1. \tag{4.10}$$

The step to take is then given by $\boldsymbol{s}_i = \boldsymbol{z}_i$. It is also generally desirable to carry out a line search in this direction to ensure that the step is acceptable and is going downhill. This is then repeated until convergence is reached.

4.1.4 FIRE

Another useful method is the fast inertial relaxation engine (FIRE)[109]. Unlike the other methods this uses an inertial minimiser inspired by molecular dynamics. This involves a velocity, $\boldsymbol{v}$, that quantifies the rate of change of the system coordinates, which is driven by the force $\boldsymbol{F} = -\boldsymbol{\nabla} E(\boldsymbol{x})$. Additional modifications are also made to the velocity in order to drive it more quickly in the downhill direction, or stop if it is moving uphill,

$$\boldsymbol{v} \to \begin{cases} \boldsymbol{v} + \alpha |\boldsymbol{v}| (\hat{\boldsymbol{F}} - \hat{\boldsymbol{v}}) & \text{if } \boldsymbol{v} \cdot \boldsymbol{F} > 0 \\ 0 & \text{if } \boldsymbol{v} \cdot \boldsymbol{F} \leq 0 \end{cases} \tag{4.11}$$

where α parametrises the downhill velocity modification and Δt is the timestep. These values are also dynamically changed to accelerate the minimisation if it is successfully proceeding downhill: If $\boldsymbol{v} \cdot \boldsymbol{F} > 0$ for the past five or more iterations, Δt is increased by a factor of 1.1 until it reaches a defined maximum value, and α is reduced by a factor of 0.99. Otherwise, if $\boldsymbol{v} \cdot \boldsymbol{F} \leq 0$, α is reset to its starting value of 0.1, and the timestep is halved. In summary the method follows the following algorithm:

1. Initialise coordinates $\boldsymbol{x}$, and velocity $\boldsymbol{v} = \mathbf{0}$.
2. Calculate the force $\boldsymbol{F} = -\boldsymbol{\nabla} E(\boldsymbol{x})$.
3. Update the velocity, $\boldsymbol{v} \to \boldsymbol{v} + \boldsymbol{F}\Delta t/m$.
4. Modify the velocity according to eq. (4.11).
5. Update α and Δt depending upon the value of $\boldsymbol{v} \cdot \boldsymbol{F}$.
6. Update the position, $\boldsymbol{x} \to \boldsymbol{x} + \boldsymbol{v}\Delta t$.
7. Check for convergence, otherwise return to 2.

4.1.5 Simulated Annealing

For some applications, when the gradient is undefined or difficult to calculate it may be necessary to use a gradient-free minimisation algorithm. This is most typically the case for discrete problems. Simulated annealing[110] is one such algorithm that is based upon the thermodynamic annealing of metals. This uses random displacements and chooses whether to accept them depending upon the change in energy. Unlike the methods above, simulated annealing is primarily a global minimiser because it can jump into the basins of other minima and so is not simply restricted to finding the local minimum closest to its initial position.

For each iteration the random perturbation is chosen such that the new state is not too far from the previous. This may be performed by adding small perturbations to each degree of freedom separately, or ensuring that the total displacement is below a certain value. The probability of accepting this depends upon the energies of the proposed state, E', and the current state, E, according to a Metropolis acceptance criteria[111],

$$P(E', E, T) = \begin{cases} 1, & \text{if } E' \leq E \\ \exp(-(E' - E)/T), & \text{if } E' > E \end{cases} \tag{4.12}$$

The temperature, T, is reduced over the course of the simulation from a high initial temperature, T_0, to zero. This allows the state to initially explore the entire configuration space, but it is steadily constrained to a single minimum. Usually the rate of cooling should be sufficiently slow so that the state is remains at thermodynamic equilibrium, otherwise it is highly likely to get stuck in a local minimum rather than finding the global minimum. This rate will depend upon the specific application, and it is hard to predict a reasonable value a priori. The precise cooling schedule can use a variety of different functions, such as a linear reduction in temperature or an exponential decrease.

Simulated annealing may also be used to emulate local minimisation by setting $T = 0$, ensuring that the energy can only decrease. However, the method is not guaranteed to find the closest local minimum, because the random perturbations may still cause it to jump into the basin of another minimum. The probability of this can be reduced by setting a small limit for the maximum perturbation, but it cannot be avoided entirely.

4.1.6 Basin hopping

Basin hopping is a related method for finding global minima[112]. Like simulated annealing, it employs Monte Carlo sampling and evaluates moves according to the criteria given in eq. (4.12). However, there are two main differences between the methods. Firstly and most significantly, the energy of each trial state is computed by performing a local minimisation. This ensures that all states within a basin of attraction have the same energy for the acceptance test, effectively eliminating the barriers between adjacent basins. The coordinates of the state may be left unchanged by the minimisation process, or they may be updated to the position of the local minimum. The latter approach, given by Li and Scheraga[113], is generally seen to be the more effective strategy[14].

The second notable difference is that the temperature is held constant throughout the simulation. Instead of depending upon a changing temperature to adequately explore the energy landscape, the step size is adjusted dynamically to ensure that the average acceptance ratio aligns with a chosen value.

The basin hopping algorithm will tend towards deeper basins until it locates the global minimum. However, the finite temperature also allows it to move to higher basins so that it can pass between multiple funnels in the landscape. This allows it to successfully locate global minima, even in landscapes with multiple deep funnels. To do this, the temperature must be approximately the size of the barriers separating funnels. So if the funnels are particularly deep or the temperature is too low the method may potentially get stuck.

4.1.7 Genetic algorithm

Genetic algorithm is another method for performing global minimisation[114] that is based upon genetic evolution. In contrast to the other methods that have been described, genetic algorithm does not optimise using a single state, but instead an ensemble of states. Each iteration of the algorithm selects the fittest individuals (with the lowest energies) in the population which are used to produce a new generation. These children are produced through a combination of crossover the parents genes, and mutation.

The selection of the parents generally occurs stochastically where the fitter individuals are more likely to be selected. There are a number of such methods, but a simple approach that is often used is the roulette wheel selection. For this, the probability that each individual is picked is proportional to their fitness (or inversely proportional to their energy),

$$p(\boldsymbol{x}_i) = \frac{1}{E(\boldsymbol{x}_i)} \Bigg/ \sum_j \frac{1}{E(\boldsymbol{x}_j)}. \tag{4.13}$$

An approach known as elitism may also be used, where a small number of the most fit individuals are selected for the next generation without undergoing any changes. This ensures that the best solutions are never lost and the global minimum estimate can only ever improve during the simulation.

For each child in the new generation the parents are randomly selected. Typically, two parents are chosen, but more parents can be used. Then, the most simple approach is to use uniform crossover, where each coordinate is randomly selected from each of the parents. Alternatively, the parents' coordinates may be split at a random point and swapped, resulting in a block of coordinates from one parent and the remaining coordinates from the other. This is known as one-point crossover, and can be generalised by splitting at a larger number of points.

The selection of the most fit individuals reduces the genetic diversity of the overall population and can lead to premature convergence, so mutation is used to keep the diversity high. For discrete applications the genetic information of an individual can be expressed as a string of binary values. Mutation is then performed by randomly flipping certain bits at a given probability, analogous to real genetic mutation. However, for continuous systems mutation is usually carried out by altering each coordinate by a random amount chosen from a normal distribution.

For continuous applications it also can be helpful to consider hybrid methods[115]. These can be used to perform a gradient-based local minimisation on each individual when calculating the fitness. This is also known as a Lamarckian genetic algorithm because of the ability for individuals to pass on what they have 'learned'. Performing the local search is beneficial because it can locate local minima much faster than using genetic algorithm alone. Also, it may be useful for studying energy landscapes where the minima sit in deep wells, in which case the energy in the neighbouring region may be unrelated to the energy of the minima.

However, a local search may cause issues due to many duplicates, which can result in a lack of genetic diversity and stagnation of the population before reaching the global minimum. One approach to resolve this is to remove individuals that are too close to another using a duplicate predation operator[116,117]. This ensures that the population remains diverse, so it can converge more quickly to the global solution, and also finds more local minima. However, duplicate predation may result in too few parents if there are not many minima on the landscape, in which case more individuals would need to be initialised, or the method restarted.

4.2 Single-ended transition state methods

One approach to locate saddle points is by evolving a single state such that it ascends up valleys in the landscape, effectively maximising one degree of freedom while minimising others. Such methods are known as walker methods or single-ended transition state search methods. We will discuss the climbing image, eigenvector-following, and dimer methods, although other methods also exist, such as the climbing string method[118] or the Lanczos iterative method[119].

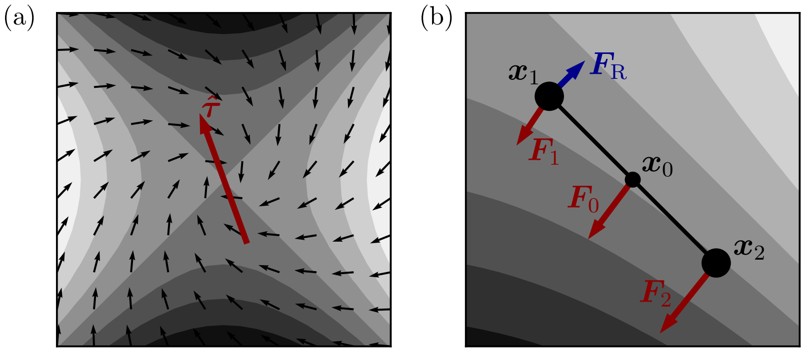


Figure 4.1: (a) The effect of $\hat{\boldsymbol{\tau}}$ on the climbing image force in the vicinity of a saddle point (white is high energy, black is low). Here $\hat{\boldsymbol{\tau}}$ is pointing 20° away from the optimal vertical direction, resulting in an inward spiralling force. The arrows show only the direction of the force. (b) The forces used to rotate a dimer, including the downhill forces on each state, $\boldsymbol{F}_i$, and the rotational force applied to state 1 about the midpoint, $\boldsymbol{F}_\mathrm{R}$.

4.2.1 Climbing image

The climbing image method[16,120] is perhaps the most simple single-ended method. It involves minimising the energy of a state, but the gradient is inverted in a chosen direction, $\hat{\boldsymbol{\tau}}$, resulting in a driving force of the form,

$$\boldsymbol{F} = -\boldsymbol{\nabla} E + 2(\boldsymbol{\nabla} E \cdot \hat{\boldsymbol{\tau}})\hat{\boldsymbol{\tau}}. \tag{4.14}$$

An example of this force about a saddle point is shown in fig. 4.1(a).

Because the gradient has been modified, this can no longer be considered a minimisation problem. Therefore, optimisation methods that ensure that the energy decreases, such as line search, can not be used here. This is also the case for most of the other energy landscape methods.

The climbing image is able to reach the TS so long as it starts sufficiently close to it and $\hat{\boldsymbol{\tau}}$ is approximately the direction of negative curvature at the TS. If not, this method is likely to diverge. Therefore, this method is often used in conjunction with other methods, especially pathway methods, in order to refine an initial estimate for the transition state, but it cannot usually locate a TS from a minimum.

4.2.2 Eigenvector-following

Unlike the climbing image method which requires the maximisation direction to be provided, the eigenvector-following method[121,122] instead calculates it by finding the eigenvector with the lowest eigenvalue of the Hessian matrix, H. Also, because this direction is updated at each step, the state does not need to be initialised close to the TS. The original method by Cerjan and Miller[121] required calculating the inverse Hessian matrix, with later methods finding ways around the costly matrix inversion[123]. Munro and Wales then introduced the hybrid eigenvector-following method that did not require the Hessian matrix to be calculated at all, making it useful for large systems[122].

The method used by hybrid eigenvector-following to find the smallest eigenvalue is to first consider the following function for a vector $\boldsymbol{y}$,

$$\lambda(\boldsymbol{y}) = \frac{\boldsymbol{y}^{\mathrm{T}}\mathrm{H}\boldsymbol{y}}{\boldsymbol{y}^2}. \tag{4.15}$$

When this function is minimised with respect to $\boldsymbol{y}$ it will give the value of the smallest eigenvalue, λ_{min}, and $\boldsymbol{y}$ will be in the direction of the corresponding eigenvalue, $\hat{\boldsymbol{e}}_{\mathrm{min}}$. To do this without knowledge of the Hessian, $\lambda(\boldsymbol{y})$ is approximated using the function,

$$\lambda(\boldsymbol{y}) \approx \frac{E(\boldsymbol{x}_0 + \delta\boldsymbol{y}) + E(\boldsymbol{x}_0 - \delta\boldsymbol{y}) - 2E(\boldsymbol{x}_0)}{(\delta\boldsymbol{y})^2}, \tag{4.16}$$

where $\boldsymbol{x}_0$ is the position where we are evaluating the Hessian, and δ is a very small value.

Once the minimum eigenvalue and eigenvector have been found the state is stepped uphill along the eigenvector with a step size of

$$h = \frac{2F}{|\lambda_{\mathrm{min}}|\,(1 + \sqrt{1 + 4F^2/\lambda_{\mathrm{min}}^2})}, \tag{4.17}$$

with

$$F = \boldsymbol{\nabla} E \cdot \hat{\boldsymbol{e}}_{\mathrm{min}}. \tag{4.18}$$

Then the energy is minimised along the hyperplane orthogonal to $\hat{\boldsymbol{e}}_{\mathrm{min}}$. This entire process is repeated, first finding the eigenvector then stepping uphill and minimising, until it converges to the saddle point.

4.2.3 Dimer method

The dimer method[124–126] is a variant of the eigenvector-following method, that identifies the smallest eigenvector of the Hessian using a pair of close-lying states. It does this by using the gradient on each state to generate a torque that turns the dimer to minimise the curvature.

By considering a close pair of states $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ with midpoint $\boldsymbol{x}_0$, each experiencing a force $\boldsymbol{F}_i = \boldsymbol{\nabla} E(\boldsymbol{x}_i)$ (as demonstrated in fig. 4.1(b)), the rotational force applied to state 1 is found to be

$$\begin{aligned} \boldsymbol{F}_{\mathrm{R}} &= (\boldsymbol{F}_1 - \boldsymbol{F}_2) - [(\boldsymbol{F}_1 - \boldsymbol{F}_2) \cdot \hat{\boldsymbol{\tau}}]\hat{\boldsymbol{\tau}} \\ &= 2(\boldsymbol{F}_1 - \boldsymbol{F}_0) - 2[(\boldsymbol{F}_1 - \boldsymbol{F}_0) \cdot \hat{\boldsymbol{\tau}}]\hat{\boldsymbol{\tau}}, \end{aligned} \tag{4.19}$$

where $\boldsymbol{\tau} = \boldsymbol{x}_1 - \boldsymbol{x}_0$. This has been re-expressed using the force at the midpoint, because this is needed for the translation step, so one fewer gradient calculation is necessary. This rotation force can be used to iteratively reorient the dimer until it reaches convergence.

Once this has occurred the dimer proceeds similarly to the climbing image method, maximising in the direction of the gradient while minimising in other directions using eq. (4.14). However, if the dimer is located in a convex region where the curvature along the dimer is positive (such as close to minima) a different force is instead used to ensure that it does not get trapped:

$$\boldsymbol{F} = (\boldsymbol{\nabla} E \cdot \hat{\boldsymbol{\tau}})\hat{\boldsymbol{\tau}}. \tag{4.20}$$

4.3 Double-ended transition state methods

If both endpoints of a transition are known then single-ended methods are inadequate because they are not guaranteed to find the correct TS. Instead, double-ended methods that converge towards the saddle from either side can be used. Unlike pathway methods these do not typically require an initial estimate for the pathway, simplifying the process for applications with complex non-linear pathways.

There are a number of such methods, including the method introduced by Dewar, Healy, and Stewart (DHS)[127], the step and slide method[128], the ridge-walking method[129], and the double-ended surface walking method[130]. Here we will focus on the first two.

4.3.1 DHS method

The DHS method[127] involves individually minimising each of the two states under a constant separation from the other state. This separation is then steadily reduced until the method converges to the transition state.

Each iteration of the method involves first identifying which of the two states has the lowest energy. This state is typically further from the transition state, so it is chosen to be minimised. It is then brought towards the other, typically by about 5%, and minimised under a fixed separation constraint. After one or more iterations, this state gains a higher energy than the other and the minimisation switches to the other state. The behaviour of this method is demonstrated in fig. 4.2.

4.3.2 Step-and-slide method

The step and slide method[128] is similar to the DHS method, although, the roles are switched. Instead of fixing the separation and minimising the energy, the separation is minimised at constant energy. This is iteratively performed using a two-step process. First, an energy is first chosen and the two states are 'stepped' directly towards one another until they reach the isosurface for this energy. Then, the states 'slide' along this isosurface until the separation is minimised. If the chosen energy is below that of the transition state then the two states will approach the TS. Whereas, if it is too high, the states will meet and should be reverted back to their previous positions. In this way, step and slide is able to provide an energy bound for the transition state.

In order to slide along the isosurface a bespoke minimisation procedure is required. It must also be implemented in such a way that one state cannot accidentally jump over the saddle point, otherwise the method will fail. For example, this may occur if the state is moved tangentially to the isosurface when close to the saddle. Therefore, the approach used in ref. 128 is to instead move one state downhill slightly according to the gradient and then step back up to the isosurface. This is alternatively performed on each state until the separation has converged.

In order to choose the target energy for each iteration E_{trial}, the method uses maximum and minimum bounds for the transition state energy. The minimum energy is given by the energies of the two states $E_{1,2}$, while the maximum energy, E_{max}, is updated any time the states combine after the slide. Then the target energy can be chosen according to a binary search, $E_{\text{trial}} = (E_{1,2} + E_{\text{max}})/2$. To initialise the method the maximum bound can be set by finding the maximum energy of a linear interpolation. However, if the result is very large, it may be better to increase the energy by a set amount until an upper bound is found that way.

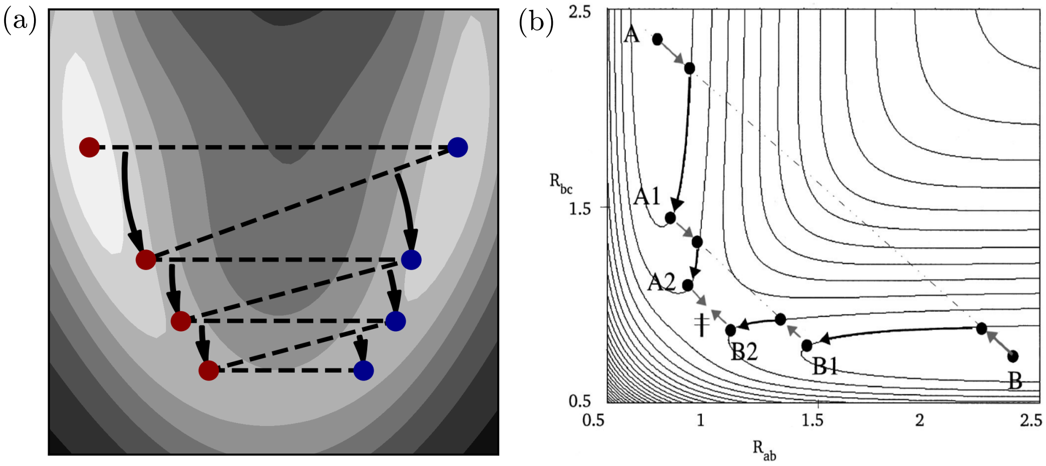


Figure 4.2: (a) A diagram demonstrating the behaviour of the DHS method. (b) An illustration of two iterations of the step and slide method. Adapted with permission from [128].

4.4 Pathway methods

Once transition states have been identified, the full MEPs can be found by performing a downhill minimisation from either side of the saddle to the minima and tracking the trajectory. To accurately follow the steepest descent, the step size of this minimisation should not be too large, and ideally a gradient descent method should be used. However, other minimisation methods, such as L-BFGS, can give an approximation of the pathway, particularly for large and complex systems where gradient descent may be slow.

Alternatively, there are a number of commonly used methods that can directly find MEPs between two end points. Each of these approaches utilise a chain of states $\{\boldsymbol{x}_1, ..., \boldsymbol{x}_N\}$, where each state contains a full set of system coordinates. This chain can be initialised as a linear interpolation between the two end states, but for some systems more complex pathways may be required[129,131,132]. The energy of the chain is then minimised while keeping the states equally separated until it converges to the MEP. Convergence is satisfied when the gradient orthogonal to the local pathway tangent vector goes to zero for every state in the chain,

$$\boldsymbol{\nabla}^{\perp} E(\boldsymbol{x}_i) = \boldsymbol{\nabla} E(\boldsymbol{x}_i) - (\hat{\boldsymbol{\tau}} \cdot \boldsymbol{\nabla} E(\boldsymbol{x}_i))\hat{\boldsymbol{\tau}}, \tag{4.21}$$

where $\hat{\boldsymbol{\tau}}_{\boldsymbol{i}}$ is the tangent vector to the pathway at image $\boldsymbol{x}_i$ along the chain.

These methods are also often used in order to find the approximate locations of transition states, which occupy local maxima along the MEP. While these methods generally do not precisely find these transition states, they can be used in conjunction with a single-ended TS search method for this purpose.

A number of pathway search methods have been proposed, although here we will focus on two commonly used approaches: the nudged elastic band (NEB) and string methods, as well as some of their variants.

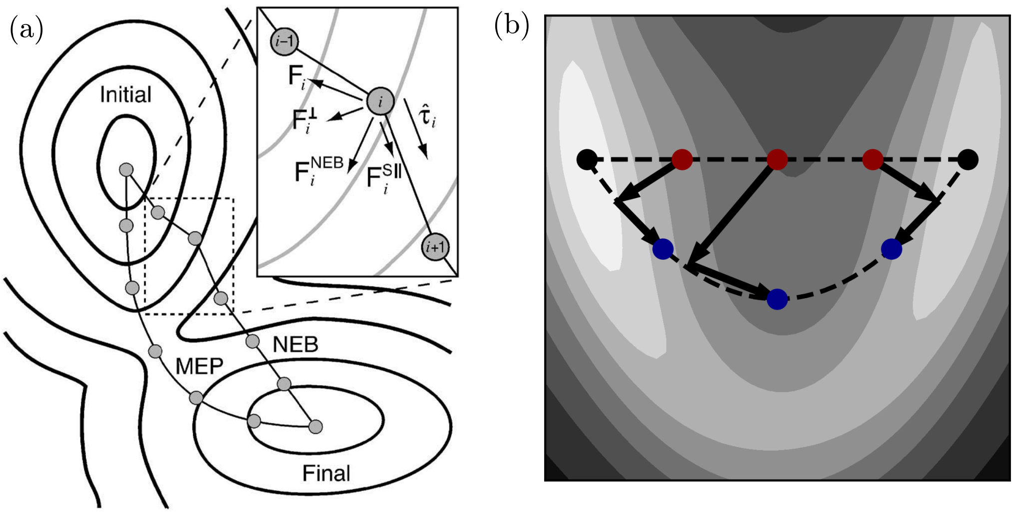


Figure 4.3: (a) A diagram showing how NEB converges to the MEP for a 2D landscape. The inset shows a schematic of the forces on an state in the chain. Adapted with permission from[133]. (b) A schematic illustrating how the string method evolves. The arrows show the partial minimisation followed by a reparametrisation along the new interpolation.

4.4.1 Nudged elastic band method

The core feature of NEB methods[134,135] are that the states are kept separated along the chain using elastic spring interactions. Therefore, the aim of the method can be considered to minimise the following total energy of the string,

$$E_{\text{NEB}} = \sum_{i=1}^{N} \left[E(\boldsymbol{x}_i) + \frac{1}{2} k \sum_{j=i+1}^{N} |\boldsymbol{x}_i - \boldsymbol{x}_j|^2 \right], \tag{4.22}$$

where k is a parameter for the spring constant between the states. However, a couple of modifications ('nudges') to the gradient of eq. (4.22) are made to reduce unwanted interference between the two contributions. Firstly, the gradient of the underlying potential is projected perpendicular to the tangent vector, $\hat{\boldsymbol{\tau}}_i$. This prevents the images from sliding down the pathway and ensures that the potential does not affect the parametrisation of the path, only its course. The second modification aims to reduce the corner cutting around bends in the pathway by using only the component of the spring force parallel to the tangent vector. Therefore the total gradient on state i in the chain is

$$\boldsymbol{G}_{\mathrm{NEB},i} = \boldsymbol{\nabla}^{\perp} E(\boldsymbol{x}_i) + \boldsymbol{G}^{\|}_{\mathrm{S},i}, \tag{4.23}$$

where

$$\boldsymbol{G}^{\|}_{\mathrm{S},i} = (\hat{\boldsymbol{\tau}}_i \cdot \boldsymbol{G}_{\mathrm{S},i})\hat{\boldsymbol{\tau}}_i \tag{4.24}$$

$$\boldsymbol{G}_{\mathrm{S},i} = k(\boldsymbol{x}_{i+1} + \boldsymbol{x}_{i-1} - 2\boldsymbol{x}_i). \tag{4.25}$$

The tangent vector at each point can be computed in a few ways. The simplest approach is to use central difference between the points and normalise the result,

$$\hat{\boldsymbol{\tau}}_{\boldsymbol{i}} = \frac{\boldsymbol{x}_{i+1} - \boldsymbol{x}_{i-1}}{|\boldsymbol{x}_{i+1} - \boldsymbol{x}_{i-1}|}. \tag{4.26}$$

However, an approach that provides better results in areas of high curvature is to bisect the tangents given by forward and backward difference[134],

$$\boldsymbol{\tau}_i = \frac{\boldsymbol{x}_{i+1} - \boldsymbol{x}_i}{|\boldsymbol{x}_{i+1} - \boldsymbol{x}_i|} + \frac{\boldsymbol{x}_i - \boldsymbol{x}_{i-1}}{|\boldsymbol{x}_i - \boldsymbol{x}_{i-1}|} \quad \text{and} \quad \hat{\boldsymbol{\tau}}_i = \frac{\boldsymbol{\tau}_i}{|\boldsymbol{\tau}_i|}. \tag{4.27}$$

A variant of NEB known as the doubly-nudged elastic band method (DNEB) brings further improvements to the method[135], and has since seen widespread use. This improves stability by introducing an additional term to the NEB gradient, eq. (4.23), such that the perpendicular component of the spring force is not completely removed. This correction term is

$$\boldsymbol{G}^{*}_{\mathrm{S}} = \boldsymbol{G}^{\perp}_{\mathrm{S}} - (\boldsymbol{G}^{\perp}_{\mathrm{S}} \cdot \boldsymbol{\nabla}^{\perp} E)\boldsymbol{\nabla}^{\perp} E. \tag{4.28}$$

The states along the chain in NEB and DNEB are unlikely to provide a particularly accurate approximation to the transition state, especially if there is significant corner cutting due to the spring interaction. Therefore they are often used in hybrid methods with single-ended transition search methods, such as climbing image[120], or eigenvector-following[135]. Once the chain has converged these can be run from the highest energy state, or, can be applied to the highest state at the same time as the elastic band is minimised. The latter requires some extra care because it is possible for the climbing state to diverge if it starts too far from the transition state.

4.4.2 String method

The string method[16,136] shares many similarities with NEB, however it differs essentially by the process that keeps the states separated along the chain. Instead of using a spring force, the states are reparameterised along the pathway occasionally throughout the minimisation.

In the originally proposed method[136] the states in the chain are driven downhill according to the gradient of the potential orthogonal to the direction of the chain. Explicitly, for state i along the chain this driving force would be,

$$\boldsymbol{F}_i = -\boldsymbol{\nabla} E_i + (\hat{\boldsymbol{\tau}}_i \cdot \boldsymbol{\nabla} E_i)\hat{\boldsymbol{\tau}}_i. \tag{4.29}$$

In a later modification to the method[16] this driving force was simplified to remove the orthogonal projection because the reparametrisation step makes this unnecessary, leaving $\boldsymbol{F}_i = -\boldsymbol{\nabla} E_i$. The tangent vector can be simply computed according to the difference between the neighbouring states, $\hat{\boldsymbol{\tau}}_i = (\boldsymbol{x}_{i+1} + \boldsymbol{x}_{i-1})/|\boldsymbol{x}_{i+1} + \boldsymbol{x}_{i-1}|$.

After a certain number of minimisation iterations have passed the states must be reparametrised along the chain. A suitable choice for this depends upon the system and the minimisation algorithm. This reparametrisation may be performed using either a linear, or a cubic-spline interpolation along the set of points. Then the new points can be equally separated along the interpolation. The number of these new points does not necessarily need to be same as the previous step, which can make it easy to refine the pathway over time. Additionally, the points do not need to be evenly distributed, and could instead be clustered close to the transition state by weighting the length of the interpolated path according to the local value of the potential[16].

Similar to NEB it is possible to use a hybrid method with climbing image[16] or eigenvector-following in order to accurately find the transition state. In this case, the state with the highest energy using the chosen method and should be left out of the reparametrisation step. Another climbing image / string method hybrid, called the climbing string method, has also been proposed[118]. However, unlike the other methods, this is a single-ended saddle point search method. This uses the string of states to guarantee that it has not left the basin of attraction for a minimum rather than for locating a MEP.

4.5 Methods comparison

Minimisation

Minimisation methods that use the Hessian, like Newton's method, generally provide faster convergence compared to other methods. However, the calculation of the Hessian can become the limiting factor if the number of degrees of freedom is large, which it generally is in the continuous systems that we are interested in. Since L-BFGS is quasi-Newtonian, it benefits from this faster convergence without being bogged down

Method	Local	Global	Gradient	Hessian	Ensemble
Gradient descent	✓	✗	✓	✗	✗
L-BFGS	✓	✗	✓	✗	✗
FIRE	✓	✗	✓	✗	✗
Newton's method	✓	✗	✓	✓	✗
Simulated annealing	✓	✓	✗	✗	✗
Basin hopping	✗	✓	✓	✗	✗
Genetic algorithm	✗	✓	✗	✗	✓

Table 4.1: A comparison of the different minimisation methods, indicating which can be used as local or global minimisers, which require the gradient or the Hessian, and which use an ensemble of states.

by the calculation of the Hessian, therefore it is ideal for our purpose.

FIRE is also a good option with convergence speeds approaching that of L-BFGS[109], so it is a good alternative for cases where L-BFGS has difficulty. It is also more amenable to parallelisation along the degrees of freedom, because there are much fewer vector magnitudes and dot products calculated per iteration compared to L-BFGS. On the other hand, FIRE has a free parameter, the maximum time step, which can require some tuning for each system and makes the method a little more difficult to use.

As for simulated annealing, it is primarily a global minimiser, but it can also be used to find local minima. Compared to L-BFGS and FIRE it converges much more slowly, and the random perturbations make it possible to jump into the basin of another minima. Therefore, where possible other methods should be used, but it is particularly useful if the gradient cannot be calculated or is undefined.

However, when it comes to global optimisation or locating multiple minima, basin hopping and the genetic algorithm tend to win out. For genetic algorithm, the ensemble of states provides a much larger accessible search region, resulting in better performance for most problems[137,138]. Furthermore, it is well suited to parallelisation because each state can be run separately, with a minimal number of interactions in each generation. As for basin hopping, it is particularly attractive because of its simplicity, the only variable parameter being the temperature. This makes it more straightforward to apply to new systems in comparison to simulated annealing or genetic algorithm, which depend heavily upon the choice of cooling schedule and genetic operators, respectively, as well as their related parameters.

Transition states and pathways

Single-ended search methods are generally efficient at locating saddle points quickly. This makes them particularly well suited for surveying large numbers of transition

Method	Number of states	TS from endpoints	Can find multiple TS	No interpolation
Single-ended TS	1	✗	✗	✓
Double-ended TS	2	✓	✗	✓
Chain-of-states	≥ 3	✓*	✓*	✗

Table 4.2: A comparison of the three types of TS and pathway methods, including the number of states that are optimised, if the method can find one or multiple TS between two chosen endpoints, and if an interpolation is required. An asterisk indicates a hybrid method is required.

states or refining estimates, such as those obtained from chain-of-states methods. However, if a specific pathway between two known end states is of interest, single-ended methods are often inadequate and double-ended methods and chain-of-states methods should be used instead.

The simplicity of the climbing image method makes it excellent for refining transition state estimates, but it cannot usually be started far from the transition state, such as from minima. The requirement for a tangent vector estimate also restricts its use. The eigenvector-following and dimer methods solve these issues and can be used in most situations, but their success depends strongly upon the number of iterations allowed when calculating the uphill direction and the corresponding convergence criteria. This can also take some time to adequately converge for large numbers of degrees of freedom. If there are zero eigenmodes of the Hessian, such as global translation and rotation, then the dimer and eigenvector-following methods can also struggle and special consideration is required.

For finding the transition states and pathways between two end points, the choice between using double-ended transition state methods or chain-of-state methods will depend upon the situation. Chain-of-states methods will be preferable if full pathways are desired, or precise transition states are not required. Likewise, if there are many intermediate minima and transition states, which is often the case for particulate systems, chain-of-states methods can identify the approximate locations for many of these at once. Chain-of-states methods are also more reliable, whereas double-ended methods can often fail to locate a transition state. As a result, chain-of-states methods are much more widely used than other methods. However, double-ended bracketting methods can be particularly efficient for finding transition states for pathways with no (or very few) intermediate minima. This is especially true if the pathways are highly non-linear on the landscape, where chain-of-states methods require a large number of images to accurately approximate the pathway. It can also make it difficult to choose an appropriate initial pathway estimate, which is necessary for chain-of-states methods. These advantages make double-ended methods attractive for elastic structures and other continuous systems, which have comparatively few minima, yet complex, non-linear pathways.

Comparing the double-ended TS methods, we find that each has different advantages and disadvantages. The DHS method is particularly simple to implement and use, and it can be good for obtaining the TS within a certain distance error. However, because only one state is minimised at a time, the distance must be reduced quite slowly for it not to pass over the ridge. As for the step and slide method, it is useful for getting a bound on the energy of the transition state. However, the minimisation along a constant energy surface requires a more complicated implementation that is not as efficient as standard minimisation algorithms. Also, it can be hard to know how much to initially increment the energy, if it is too large or too small the method may take a lot longer to converge.

Of the pathway methods, both DNEB and string methods have a comparable effectiveness. Each can be advantageous for different systems and sets of parameters, but it is difficult to know which will be better a priori[139].

Chapter 5

Binary Image Transition State Search

5.1 Introduction

Transition states are central to the description of reconfiguration mechanisms for systems in chemistry, condensed matter physics, and engineering. Historically, many computational methods for locating transition states have grown from an atomistic or particulate perspective. These have proven to be important tools for understanding, for example, protein folding[1,30], biological and industrial catalysis[2,27,29], quantum tunnelling[19,25], crystallisation[38], and cluster formation[17,33].

More recently, it is increasingly being recognised that transition states are useful in mesoscale or macroscale systems. Here, the minimum energy barriers provide important lower bounds to the energy input required for transitions to occur. This has been used to understand failure in structural engineering applications[62,101], for the development of super liquid-repellent surfaces[50,140,141], and investigating locomotion through complex terrain for robotics[51]. Moreover, it is becoming desirable to tailor elastic deformation transitions to enable technologies such as advanced deployable structures[10,52], mechanical sensors and actuators[77,142–144], and energy absorbers[58,84] to name but a few.

As discussed in sections 4.2 to 4.4, transition state search methods generally fall into two categories, single- and double-ended methods. Single-ended methods are initialised at a single state and attempt to climb to a nearby saddle point. Examples include eigenvector following[121], the dimer method[125,126,145], and climbing image methods[16,118]. Double-ended methods can be further subdivided into two groups. The first are bracketing methods, which involve two states converging to the transition state from either side. These include the Dewar-Healy-Stewart (DHS) algorithm[127], ridge method[129], the step and slide method[128], and the double-ended surface walking method[130]. The second group utilise chain-of-states methods for finding pathways, such as string method[16,136] and doubly-nudged elastic band (DNEB)[135]. In order to locate transition states, these are combined with single-ended methods, resulting in a hybrid approach. However, chain-of-states methods require an appropriate initial interpolation, which can sometimes be challenging to obtain[132].

Unfortunately, a large range of landscapes prove challenging or impossible to explore via these methods. One such problem arises from the push towards larger and more complex systems[146–148], resulting in the need to develop algorithms that are more computationally and memory efficient, and can incorporate optimisation strategies such as on-the-fly adaptive remeshing and coarse-graining. These typically involve changing the resolution or discretisation of the systems to focus the computational time on important regions, such as using a higher resolution mesh in regions of high stress in finite element simulations[149]. However, chain-of-states methods involve a coupling between the configurations of each state, and so there is an issue if they have different discretisations and numbers of degrees of freedom. Meanwhile, single-ended methods can be inefficient because they are not well suited for identifying specific pathways and can spend a large amount of time searching for undesired transition states. Another major challenge in studying complex energy landscapes relates to the presence of locally flat or discontinuous regions, such as when considering patchy[150–152] and hard-body[38,153] interactions in atomistic simulations, systems of polymer chains[12], or collision constraints for macroscopic objects[154]. Flat zero-modes in the landscape pose issues for single-ended search methods and current bracketing methods that rely only upon local information. Specialist treatment can sometimes be used such as in the case of global rotation and translation[155], but they are thwarted by local zero-modes. Finally, current methods cannot typically be applied in the case of discontinuous potentials, or if the gradient is prohibitively expensive to compute, because continuous, differentiable optimisation functions are required.

In this chapter we introduce a new double-ended bracketing method, the Binary-Image Transition State Search (BITSS). Using a range of different applications, we demonstrate that it successfully addresses each of the above challenges. In addition, we show that BITSS is superior compared to existing bracketing methods, allowing us to access the transition states when other methods fail.

5.2 BITSS method

The method begins by first initialising the states, $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, in the basins of attraction of different local minima, such as the two blue spots in the 2d potential in fig. 5.2(a). These can be set to the minima, but this is not a necessary requirement. The energies of these two states are then minimised, while constraining their separation. This is iteratively reduced to zero, such that, at iteration i, their separation is

$$d_i = (1 - f)d_{i-1}, \tag{5.1}$$

with d_0 taking the value of the separation between the two initial states. A reduction factor of $f = 0.5$ is successful for most applications, but this can be made smaller to ensure that the states do not slide off the ridge between the two basins of attraction. This and other parameters in this chapter are listed in table 5.4. Different metrics may be used to compute this distance, although we will simply use the Euclidean distance,

$$d(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sqrt{\sum_i (x_{1,i} - x_{2,i})^2}. \tag{5.2}$$

To further ensure that neither state is pulled over the ridge, a secondary constraint enforces equal energies for the two states. Using this strategy, the two states will meet at the lowest point on the ridge, the transition state.

The two constraints are implemented using energy penalty terms, which result in driving forces on the two states if the constraints are not met, such as in fig. 5.2(b). Including these energy penalty terms gives the total BITSS energy for the pair of states,

$$E_{\text{BITSS}}(\boldsymbol{x}_1, \boldsymbol{x}_2) = E_1 + E_2 + \kappa_e \left(E_1 - E_2\right)^2 + \kappa_d \left(d(\boldsymbol{x}_1, \boldsymbol{x}_2) - d_i\right)^2, \tag{5.3}$$

where E_1 and E_2 are the single-state energies, and κ_e and κ_d parametrise the strengths of the energy and distance constraints.

The L-BFGS algorithm is chosen to minimise this energy, owing to its fast convergence and low memory requirement for large numbers of degrees of freedom[107]. However, any other minimisation method can be used in practice.

5.2.1 Choosing the constraint coefficients

To ensure that the transition state is located successfully, the constraint strengths κ_d and κ_e are updated as the algorithm proceeds using information from the system. These are initially calculated at the start of each minimisation, and regularly recalculated throughout (once per 100 iterations is used here). These are set such that the driving forces due to the constraints and single-state energies are of similar size, in order to prevent the constraints from dominating the underlying potential or causing large jumps that make a state pass over the ridge.

To obtain the expression for the energy coefficient, κ_e, we first assume that the separation is fixed, so the distance term can be ignored. The coefficient κ_e must be high enough to prevent one state from being pulled over the ridge, for which the greatest risk occurs when the gradient on one state is much greater than the other, e.g. $|\boldsymbol{\nabla} E_2| \gg |\boldsymbol{\nabla} E_1|$. In this case the total gradient is approximated by

$$\boldsymbol{\nabla} E_\mathrm{BITSS} = \left[1 - 2\kappa_\mathrm{e}(E_1 - E_2)\right] \boldsymbol{\nabla} E_2. \tag{5.4}$$

Therefore, when not at a transition state or a minimum in the landscape ($|\boldsymbol{\nabla} E_2| \neq 0$) convergence will occur when the term in the square brackets is zero, resulting in $E_1 - E_2 = 1/2\kappa_\mathrm{e}$. This energy difference should be less than current energy barrier, so we can substitute it with E_B/α, where α is a constant greater than one, and E_B is an estimation for the current energy barrier, evaluated using a linear interpolation between the two states. This leaves us with an expression for κ_e.

$$\kappa_e = \frac{\alpha}{2E_\mathrm{B}}. \tag{5.5}$$

The distance coefficient is determined by assuming that the energies are equal and thus the energy constraint can be ignored. In this case, convergence will occur when $\boldsymbol{\nabla} E_1 + \boldsymbol{\nabla} E_2 = -2\kappa_\mathrm{d}(d - d_i)\boldsymbol{\nabla} d$. It is then possible to find the value of κ_d for which the magnitude of each side of this equation is equal for a desired relative error in the distance, $\beta = (d - d_i)/d_i$,

$$\kappa_\mathrm{d} = \frac{\sqrt{|\boldsymbol{\nabla} E_1|^2 + |\boldsymbol{\nabla} E_2|^2}}{2\sqrt{2}\beta d_i}. \tag{5.6}$$

Here, it has been used that the gradient of the distance with respect to a single state has a magnitude of 1, so the total magnitude for the pair of states is $|\boldsymbol{\nabla} d| =$

$\sqrt{1^2+1^2} = \sqrt{2}$. To ensure that the coefficient is not too small if the gradient is close to zero, such as when the states are initialised at the minima, a lower bound is set by replacing $|\boldsymbol{\nabla} E_1|$ and $|\boldsymbol{\nabla} E_2|$ with $2E_\mathrm{B}/d_i$. This results in the following overall equation for κ_d.

$$\kappa_d = \max\left(\frac{\sqrt{|\boldsymbol{\nabla} E_1|^2 + |\boldsymbol{\nabla} E_2|^2}}{2\sqrt{2}\beta d_i}\, ,\ \frac{E_\mathrm{B}}{\beta d_i^2}\right), \tag{5.7}$$

In practice, when numerically minimising, the states will jump about slightly which can result in large gradients perpendicular to the optimal movement direction. To reduce this effect, the gradients used in eq. (5.7) are projected in the direction of the separation between the two states:

$$|\boldsymbol{\nabla} E_n| \approx \frac{|(\boldsymbol{x}_1 - \boldsymbol{x}_2)\cdot \boldsymbol{\nabla} E_n|}{|\boldsymbol{x}_1 - \boldsymbol{x}_2|}. \tag{5.8}$$

In order to choose suitable values for the constant parameters α and β we have tested different parameter choices using the seven-particle cluster and cylindrical buckling examples, described in section 5.3. These are used with a fractional separation decreases of $f = 0.5$ and $f = 0.4$, respectively. Fig. 5.1 shows which choices lead to convergence and the speed at which this occurs.

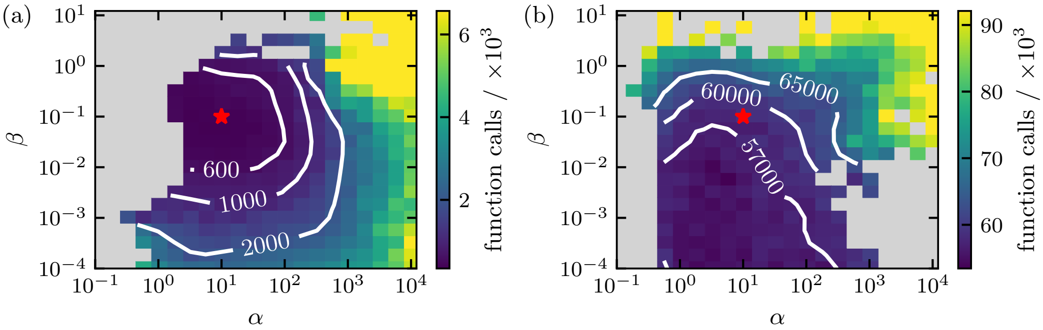


Figure 5.1: Speed of convergence of the BITSS method under different choices of parameters for (a) the seven-particle cluster, and (b) cylindrical buckling. The speed is given by the number of evaluations of the gradient until the two states are separated by less than a thousandth of the initial separation. The combinations that do not converge to the transition state are shown in grey. The chosen parameters are marked by a red star.

The parameters $\alpha = 10$ and $\beta = 0.1$ are chosen for converging quickly, while remaining far from the regions of non-convergence in both test cases. Hence, this choice is likely to still succeed even if the boundaries of the non-convergence region shifts when using different systems. However, these parameters may also be adapted to a specific system if desired.

5.2.2 Method summary

In summary, the BITSS method involves iteratively performing the following three steps:

1. Reduce the constrained separation, according to eq. (5.1).
2. Minimise the potential of the pair of states, eq. (5.3).
3. Recompute the constraint coefficients, κ_e and κ_d, at regular intervals using eqs. (5.5) and (5.7).

This process is completed once a suitable convergence criterion is reached. This can either be based upon the separation between the states, the size of gradient at the midpoint between them, or the change in the position of the midpoint.

Using this approach, the typical trajectories of the states are demonstrated for a simple 2D potential in fig. 5.2(a). Initially, the lower energy state jumps up to satisfy the equal energy constraint and then moves to minimise the separation without increasing its energy. Then, the two states converge directly towards one another, before being deflected towards the saddle in the ridge. Consequently, if there are multiple possible pathways between two states, BITSS will be biased towards identifying those that are more direct or with lower energy. Furthermore, the final two states are positioned either side of the transition state in the direction of the negative curvature eigenvector, $\hat{\boldsymbol{\tau}}$ (fig. 5.2(c)). So, BITSS automatically identifies the 'reactive mode' and associated eigenvalue in addition to the transition state. Once the transition state has been identified, it is possible to find the full minimum energy pathway by tracing the trajectory of downhill minimisations from the two final states, which are either side of the saddle.

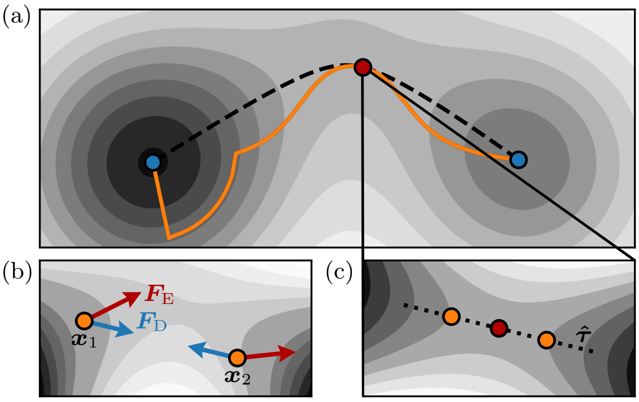


Figure 5.2: Schematics of the BITSS method on a simple 2D potential with two minima. (a) The orange line shows the trajectories of the two states from the minima (blue) to the transition states (red) under the BITSS method. The minimum energy pathway is shown by the dashed line. (b) A snapshot of the BITSS minimisation showing the driving forces on each state due to the energy constraint, $\boldsymbol{F}_\mathrm{E}$, and distance constraint, $\boldsymbol{F}_\mathrm{D}$, with $E_1 < E_2$ and $d(\boldsymbol{x}_1, \boldsymbol{x}_2) > d_i$. (c) The final configuration of the BITSS method showing the two states in orange, the transition state in red, and the negative curvature eigenvector, $\hat{\boldsymbol{\tau}}$.

In the event that there are intermediate stable states, there will be a chain of multiple transition states between the two minima. In this case, the equal-energy constraint will not prevent the states from passing over the lower energy transition states, so BITSS should converge to the transition state with the highest energy. This enables the identification of the overall energy barrier, providing estimates about the overall ease of the transition, or the rate for chemical processes. However, as demonstrated later in section 5.4.5, if multiple transition states have very similar energies then a smaller distance reduction factor, f, may be necessary to ensure that it does indeed converge to the highest transition state. Furthermore, if all of transition states or the full pathway are desired, BITSS can be continually repeated from one of the minima downhill from the located transition state and one of the initial minima until the initial minima are piece-wise connected by a full pathway.

5.2.3 Changes for undefined gradients

In order to use BITSS in situations where the gradients are unknown a couple of alterations to the method must be made. First, the calculation of κ_d in eq. (5.7) must be adapted to avoid the use of gradients. This can be done by simply removing the first term and just using the second term in the equation. Secondly, L-BFGS can no longer be used because it requires knowledge of the gradients. We must instead use a minimiser that does not require a differentiable optimisation function, for which we use simulated annealing[110]. This has a chance of randomly jumping one state over the dividing barrier, but we can reduce this probability by limiting the initial temperature and maximum random displacement. We typically employ $T_0 = E_\mathrm{B}/10$, and $d_\mathrm{max} = d(\boldsymbol{x}_1, \boldsymbol{x}_2)/100$.

5.3 Details of the test systems

To test the BITSS method several diverse systems are used which exhibit a broad range of energy landscapes. These are described below, and further details of the potentials can be found in section 6.3, including the full functions for their energy and gradient.

Firstly, the 2D potentials in figs. 5.2 and 5.4 use a sum of Gaussian potentials, eq. (5.9), to create a landscape with two minima and a curved pathway between them.

$$E(x, y) = \sum_i a_i \exp\left(-\frac{(x - b_{x,i})^2}{c_{x,i}} - \frac{(y - b_{y,i})^2}{c_{y,i}} \right) \tag{5.9}$$

The parameters for each figure are provided in tables 5.1 and 5.2.

The next system is a two-dimensional cluster of seven particles, with 14 degrees of freedom representing the particle coordinates. A Lennard-Jones potential is used for the interaction between each pair of particles. This is frequently used as a test

a	b_x	b_y	c_x	c_y
-3	-1.4	0	1	1
-2	1.4	0	1	1
-1	0.07	1	1	1

Table 5.1: Parameters for the Gaussians to produce the potential used in fig. 5.2.

a	b_x	b_y	c_x	c_y
-1	0	0	10	10
1	0	0	1	1
5	2	0	1	0.1
-1	1	1	0.1	0.1
-1	1	-1	0.1	0.1
0.01	0	0	1	1
0.5	-2	0	1	1

Table 5.2: Parameters for the Gaussians to produce the potential used in fig. 5.4.

system for studying transition rates[15,156]. The cluster has a global minimum when the particles are arranged in a hexagon, and it undergoes a transition to a second minimum when a pair of the outer particles slide past one of the other particles, as shown in fig. 5.3(a).

The third test case is the buckling of a thin cylindrical shell under axial compression. For this, the ends of the cylinder are compressed towards one another by 0.14% and then fixed in place. The characteristic transition in fig. 5.3(b) shows the formation of a stable dimple from an initially unbuckled cylinder. This transition is essential to capture and predict mechanical failure under strain[61,101]. This is modelled using a 2D triangular mesh to represent the surface of the cylinder, with 35 400 degrees of freedom for the node coordinates in three-dimensional space. The elastic energy is calculated using a bar-and-hinge model. This treats all the bonds in the mesh as elastic springs and connects all adjacent pairs of triangles with elastic hinges in order to calculate the stretching and bending energies, respectively.

The final system is the droplet on a striped surface, with 40 000 degrees of freedom representing the local fluid compositions at each site of the discretised domain. The regions on the surface are hydrophilic and hydrophobic, with contact angles of $\theta_a =$

60° and $\theta_b = 110°$, respectively. In this example, the transition involves the movement of the droplet from being centred upon one of the hydrophilic regions, to straddling across two of them, as shown in fig. 5.3(c). Droplet transitions on patterned surfaces such as this are vital to understand as powerful bio-inspired liquid manipulation strategies[157,158]. The system is modelled on a 200×200 2D grid (and 400×400 in section 5.4.3) with a phase-field model[50]. This uses an order parameter to represent the phase of the fluid (liquid or gas) at each point on the grid, and it varies smoothly, resulting in a diffuse interface between the phases. This interface width is set to a value of 2.5 lattice units. The total volume of the droplet is also constrained to ensure that it does not fully evaporate, using an energy penalty for any variation from the target volume. Because the order parameter only varies when it is at the interface of the droplet, the transitions follow highly non-linear pathways across the model's configuration space. Therefore, it is a useful test for how BITSS behaves for complex pathways.

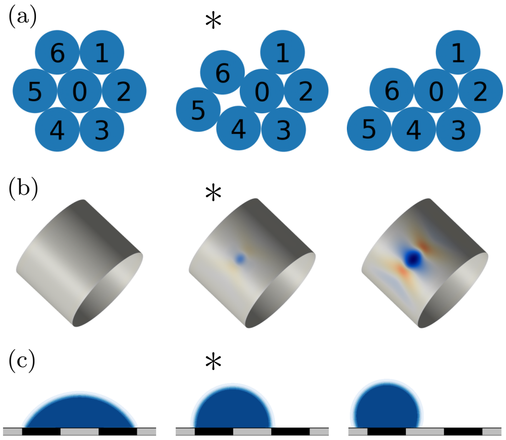


Figure 5.3: Three of the transitions used to test the BITSS method. These are: (a) a Lennard-Jones seven-particle cluster, (b) cylindrical shell buckling, and (c) wetting of a chemically-striped surface. The configurations shown correspond to the two minimum energy states and the transition state, marked by an asterisk.

5.4 Results and Discussion

5.4.1 Comparison with other bracketing methods

The BITSS potential in eq. (5.3) and iterative steps above offer key advantages over existing bracketing methods that also use two states to locate the transition state. For instance, in the ridge method[129], the two images are initially chosen to bracket the largest energy point on an interpolated path between the two endpoints. However, this is not guaranteed to be on the ridge containing the transition state, and specialist methods are required to avoid high-energy local maxima, or when the initial path contains multiple candidate maxima. In another example, the double ended surface walking method[130] requires Gaussian bias potentials to be added at each iteration to force two dimers to climb uphill in the landscape. For high numbers of degrees of freedom and many iterations, this becomes very computationally expensive.

The two methods most similar to BITSS are the DHS[127] and step and slide[128] methods. In the step and slide method, the separation between two images is minimised while their energy is fixed (iteratively increasing the energy up to the transition state). Conversely in the DHS method, the energy of an image is minimised while the image separation is fixed (iteratively decreasing the separation and changing the frozen image up to the transition state). To illustrate how BITSS is superior compared to these methods, we consider the hooked 2d potential in fig. 5.4. For this potential the energies of the images ascend higher than that of the transition state and consequently both of these methods fail to converge to the saddle point regardless of the parameters that are used. The step and slide method fails in this situation because it always expects that the energy of the two states is below the saddle point if they have not converged, so it has no means of descending down the ridge. For DHS, the images reach a certain point at which one state can pass over the ridge by minimising its energy. At this point DHS will fail even if the distance is reduced very slowly. In contrast, the BITSS method is successful for this potential. This is because the combination of distance and energy constraints allow BITSS to approach a transition state from both below and above (by sliding down a ridge).

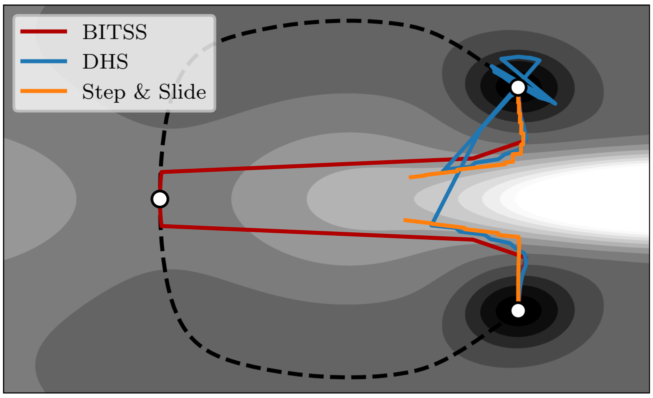


Figure 5.4: Trajectories of the bracketing methods on a hooked potential with a single saddle point. The minimum energy pathway is shown by the dashed line.

Furthermore, using both an energy and distance constraint with BITSS provides improved efficiency over these methods which each use just one of the constraints. In the case of DHS, fixing one state in place and optimising the other means that the amount that the separation is reduced must be much smaller than BITSS to ensure that it does not pass over the ridge. Meanwhile, in step and slide, it is difficult to obtain a reasonable energy increment when the two states are far from the transition state, leading to a larger than necessary number of iterations. Also, restricting the minimisation to a constant energy surface can result in a considerably more complex method, as the states must be constantly projected back onto this surface.

5.4.2 Comparison with chain-of-states methods

In contrast to BITSS, chain-of-states methods do not typically find transition states directly. Instead, they are designed to find the full pathway (or an approximation thereof), and a secondary method can then be used to refine to the transition state. As we will demonstrate, this strategy is successful for simple, linear pathways, but faces two key challenges when the pathway is highly non-linear. Firstly, for such complex pathways, a large number of states are required to sufficiently approximate the minimum energy pathway. The second is that choosing a suitable initial interpolation can be problematic to achieve. BITSS can be advantageous in both these regards, as only two states are evolved, regardless of the pathway complexity, and no initial interpolation is required.

Here we compare the speeds of convergence of BITSS to two widely used approaches for finding transition states that employ chain-of-states methods: climbing image nudged elastic band (CINEB)[120], and DNEB with hybrid eigenvector following (DNEB-HEVF)[121]. Three diverse systems are used for this comparison, the seven-particle cluster, cylindrical buckling, and wetting of a droplet on a striped surface. For the latter, the high non-linearity of the pathway causes the chain-of-states methods to fail to find the correct pathway if they use a simple linear interpolation. Instead, the position of a semi-circular droplet is interpolated between the two final positions.

These methods use NEB and DNEB which are described in section 4.4.1. Put simply, they involve minimising the total energy of a chain of states that are connected by elastic springs to keep them equally spaced along the transition pathway. The optimal values for these spring constants are system dependent. They are chosen such that they keep the states equidistant without overwhelming the gradients arising from the potential energy landscapes under consideration. We employ 10^{-1} for the Lennard-Jones particle cluster, 10^{-2} for the cylindrical buckling example, and 10^{-6} for the striped wetting system. We also fix the two end-points at the minima, so the number of states that are minimised is two fewer than the number of states in the chain.

CINEB modifies NEB by altering the behaviour of the state with the highest energy. The direction of minimisation on this state is inverted along the pathway direction, effectively converting the saddle point into a local minimum. Alternatively, DNEB-HEVF involves first minimising the chain of states until a convergence criterion is met, and then performing hybrid eigenvector following from the highest energy

System	BITSS	BITSS-HEVF	DNEB-HEVF			CINEB		
			3	5	20	3	5	20
LJ-7	148	144	135	138	361	30	153	1692
Buckling	12 866	14 446	73 694	7315	12 094	—	8352	77 904
Wetting	12 100	17 262	17 840	17 721	48 286	—	—	—

Table 5.3: Number of potential gradient calculations required to reach the transition state for the three comparison examples. The climbing image nudged elastic band (CINEB) and DNEB with hybrid eigenvector following (DNEB-HEVF) methods have been run for different numbers of images. Convergence is determined to be when the root-mean-square of the gradient at the estimate for the transition state is less than 10^{-4}. The fields left blank indicate that the method has not converged to the correct transition state.

state, moving uphill along the smallest eigenvector of the Hessian until it reaches the transition state. For completeness, we also combine hybrid eigenvector following with BITSS and include the results in the convergence comparison.

When using hybrid eigenvector following, we need to set out criteria to determine when the double-ended method has sufficiently converged, at which point the eigenvector following method begins. For BITSS, the criterion is when the average of the two states changes by less than a tenth of the reduction in the separation given by eq. (5.1) during a BITSS step. For DNEB, the root-mean-square of the total energy gradient of the chain of states is used with a convergence criterion of 10^{-3} for the particle cluster and cylindrical buckling systems, and 10^{-5} for the striped wetting. We employ the hybrid eigenvector following method implemented in the program OPTIM[159].

It is also possible to use the string method with a climbing image[16] or eigenvector following[160]; although, in this case, the results are expected to be similar to the nudged elastic band methods. We note that our aim in this section is to observe how the BITSS method behaves for different systems, rather than providing a comprehensive comparison of the currently available methods, which has been performed in other works[133,139].

The results for the three systems are shown in table 5.3. First, we note that for all three systems, using hybrid eigenvector following does not significantly improve the speed of BITSS. Indeed, for buckling and especially wetting, HEVF is detrimental to performance. Next, it is interesting to compare each method's performance between simple and complex pathways. In contrast to the wetting transition's highly non-linear pathway, the pathways of the LJ-7 rearrangement and the buckling system can be simply tracked following a gradual variation in the order parameters. For LJ-7, this is the translation of atoms 5 and 6, and for buckling, this is the radial displacement of the centre of the dimple[101]. For these simpler pathways, BITSS is generally slower, but for the complex pathway, BITSS is faster. Moreover, we see that for the wetting example, CINEB does not converge to the transition state because the estimated tangent vector is highly inaccurate due to the non-linearity of the pathway.

For situations where memory is limited, it is important to minimise the number of images used. However, efficiently finding the TS is challenging for both CINEB and DNEB-HEVF if too few images are used, as observed for the cylindrical buckling with three images. BITSS, on the other hand, converges using only two images.

5.4.3 Adaptive discretisation

Adaptive remeshing and coarse-graining are widely used techniques that we can utilise to further increase the efficiency of BITSS. These techniques cause issues for most existing double-ended methods because the coupled states may end up with different degrees of freedom. However, in BITSS the only direct coupling is in the distance measure, $d(\boldsymbol{x}_1, \boldsymbol{x}_2)$, which is relatively easy to adapt. Here we will demonstrate the feasibility of adaptive remeshing by considering two issues separately: a changing discretisaton, and different discretisations on the two states.

To test the discretisation changing we use the cylindrical buckling example and increase the resolution each time the separation is halved. In total the number of triangles around the cylinder is increased from 40 to 100, corresponding to an increase from 1760 to 11 000 degrees of freedom. To determine the three-dimensional coordi-

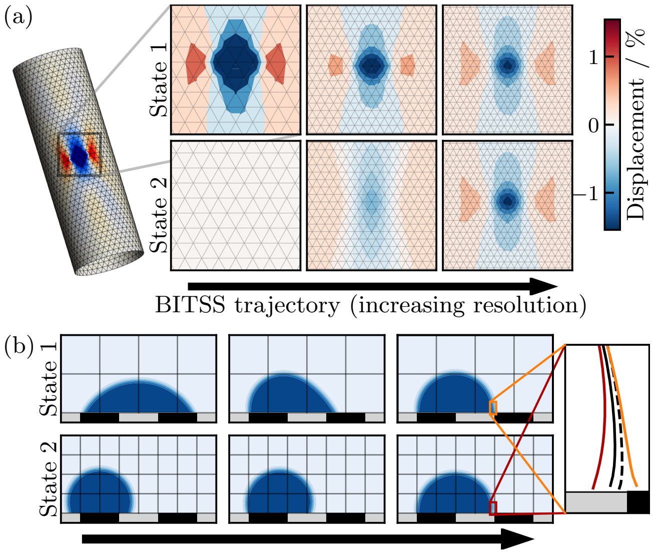


Figure 5.5: Demonstration of BITSS addressing the challenges associated with adaptive remeshing. (a) Snapshots of the BITSS method for the buckling of a cylinder with a changing mesh. The radial displacement relative to the unbuckled cylinder is shown, as well as the underlying triangular mesh. (b) Snapshots for the striped wetting example with different resolutions for the two states. Each grid cell denotes 50 × 50 lattice nodes. The zoomed axis shows the difference in the fluid interface between the two final states, as well as the approximated transition state (solid black line). This is compared to the transition state found using a high resolution (dashed line).

nate of each new grid point, a linear interpolation from the previous coordinates is performed based upon the positions of the unbuckled meshes. The results for this in fig. 5.5(a) demonstrate that BITSS is indeed able to handle the discretisation and number of degrees of freedom changing as the method runs. Therefore, BITSS is able to converge to the transition state so long as the remeshing is not so significant as to shift a state into the basin of attraction of the other minimum.

In the second test, shown in fig. 5.5(b), we demonstrate the use of different meshes for the two states in the striped wetting example. In this case, the distance measure is adapted by mapping the higher resolution state onto the other mesh and then computing the Euclidean distance. However, for some applications a simpler measure may be sufficient, such as the difference between average values of the system. Using this approach, BITSS is able to closely approach the transition state. The precision of this convergence is now limited by the transition state energy differing slightly on each grid, but this effect will be reduced when using a higher resolution or an

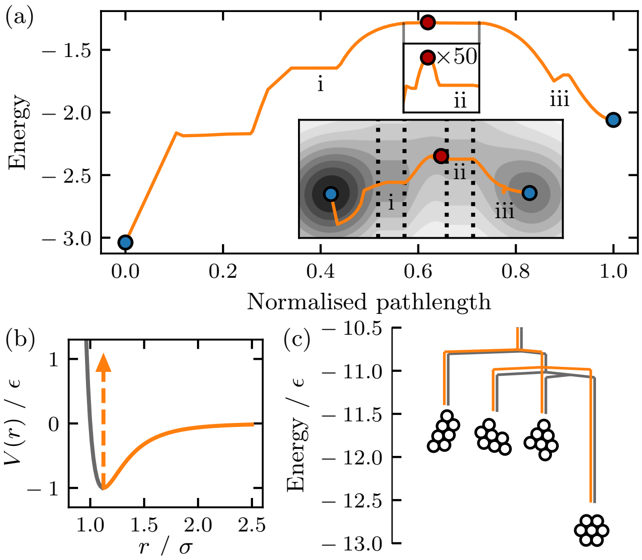


Figure 5.6: Demonstration of BITSS applied to flat and discontinuous potentials. (a) Energy profile of the BITSS pathway on a 2D potential with flat regions. Blue and red dots denote the minima and transition state, respectively. Points of interest are labelled by i–iii (see text). Top inset: A zoomed in view around the transition state. Bottom inset: The pathway taken, with the edges of the flat regions marked by dashed lines. (b) The discontinuous hard-core pair potential used in the seven-particle cluster (orange). The standard Lennard-Jones potential is also shown in grey. (c) Disconnectivity graphs of the energy landscapes for the seven-particle cluster with the two potentials. The two graphs are offset for visibility.

adaptive method where the grids themselves converge.

5.4.4 Complex landscapes

The final challenges we will address are those related to complex landscapes that prove challenging for previous algorithms. The first is the presence of flat regions in the landscape. Fig. 5.6(a) shows BITSS applied to a 2D landscape with two such regions (i & ii) that are flat in the x-direction. We see BITSS is able to successfully converge past these flat regions, even with one very close to the transition state (ii). In these regions there are no driving forces due to the potential and the energy constraint, which use purely local information about the gradient. However, the distance constraint continues to pull the states together, preventing them from getting stuck. When only a single state has a zero-gradient mode then the other is likely to slide down the potential slightly (iii), but the two states still remain either side of the dividing ridge and so the result is unaffected.

An additional consideration is the case where the potential energy surface is flat at the top of the pathway. There are two possibilities here, one is that the potential is flat in a direction perpendicular to the tangent of the pathway, such that the ridge is level. In this case BITSS is unaffected and it will be able to converge to some point along the ridge. An example of this is the free global rotation and translation of the Lennard-Jones cluster in fig. 5.3(a). The other possibility is that the flat mode is in the direction of the pathway. In this case there is no single transition state along the pathway, but instead a region. BITSS would be ill suited in this situation because the equal-energy constraint would not prevent the images from passing over the saddle and falling to a minimum.

Finally, we investigate the application of BITSS to systems with undefined gradients, such as when the landscape is discontinuous. To account for this, the equations for the coefficients must be adapted to not depend upon the gradients, and a gradient-free minimiser (simulated annealing) is used. These changes are detailed in the methods section. This has been tested using a 7-particle cluster with a hard-core Lennard-Jones pair-potential, shown in fig. 5.6(b), which results in a discontinuous landscape. Using the gradient-free approach, BITSS is able to successfully find the transition states, allowing us to plot the disconnectivity graph of the system, shown in fig. 5.6(c). Compared with the results for the standard Lennard-Jones cluster, the energies of the minima are largely unchanged, but the energies of the transition states are found to be slightly higher. This indicates that the particles in the Lennard-Jones cluster cut the corner slightly as they transition, whereas this is not possible using the discontinuous potential, resulting in higher energies. Despite this gradient-free method being feasible, it is worth noting that a gradient-based approach is significantly more efficient, and so should be preferred if possible.

5.4.5 Multiple transition states

Here we test how the BITSS method performs when there are multiple transition states in the pathway between the two starting minima. For this we use a 2D potential, shown in fig. 5.7(a), with a pathway that follows two connected 135° circular arcs. The energy is given by the squared distance from this path, plus two Gaussian barriers resulting in transition states A and B, with energies E_{A} and E_{B}. We then vary E_{B} between 0 and E_{A}, and the size of the distance reduction factor, f, to see if BITSS successfully converges to the higher transition state A. For each pair of parameters we perform 5 runs with slight variations in the starting positions, with the results in fig. 5.7(b) showing the points at which all 5 runs converge to the higher TS, A.

We see that if the difference between the two barriers is sufficiently large ($\gtrsim 10\%$) then BITSS always converges to the higher transition state. However, as the difference decreases it starts to sometimes converge to the lower transition state if the separation is decreased quickly. Therefore, to ensure that BITSS always converges to the higher energy transition state it may be necessary to restrict the separation step size.

Although the equal-energy constraint should cause the states to converge to A, this is not always the case because the discrete steps in the minimisation can cause the left image to jump over A before the other image passes B. If the minimisation is fast and takes large steps then the chance for this to occur is increased. Therefore, systems with complex, high-dimensional landscapes can probably successfully locate the highest transition state for larger values of f and smaller height differences than simpler systems such as this 2D example.

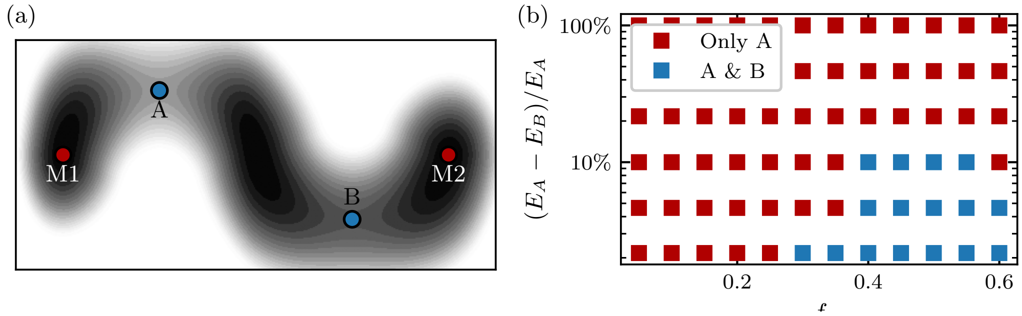


Figure 5.7: (a) The potential used in the multiple TS test. M1 and M2 denote the two starting minima, and A and B are the two transition states. The energy of the barrier to B, E_{B}, is varied between 0 and E_{A}. (b) The results for the parameters under which BITSS reliably converges to A, and the points at which it sometimes converges to B. The x-axis is the BITSS distance reduction factor, f, and the y-axis is the relative difference between the two barrier heights.

5.5 Conclusion

Overall, we have developed the binary image transition state search (BITSS) algorithm for the efficient location of transition states in traditionally challenging landscapes. This has distinct advantages for complex pathways owing to the lack of a required initial pathway estimate, as well as the identification of the transition state that provides the overall energy barrier in multi-step pathways. From the speed analysis, we find that the combination of chain-of-states methods with single-ended transition state search methods provides good performance for near-linear pathways, such as for the Lennard-Jones cluster and cylindrical buckling. However, for highly complex and non-linear pathways, as exhibited by the striped wetting example, BITSS is superior. Indeed, the demonstrated speed and memory-efficiency will be key as we move towards studying larger and more complex systems using BITSS.

A second source of efficiency in the BITSS method comes from the ability to adaptively change the degrees of freedom as the algorithm proceeds. We demonstrated how transition states could be found by both increasing the resolution upon convergence, and coupling systems with different discretisations. The ease of coupling two copies of a system and adaptive remeshing, now leads to the possibility of incorporating BITSS into existing open-source optimisation methods, such as surface evolver[161] or finite element methods[162], to provide important energy barrier functionality.

We also showed how BITSS can be used to survey discontinuous energy landscapes, demonstrated for a system of attractive hard-core particles. This opens up possibilities for studying a broad range of systems previously out of reach of conventional landscape methods, but where transition information is valuable. These include systems with very short range interactions, such as in colloidal clusters, or hard contact forces, such as in the folding of elastic materials, or locomotion and environmental interaction in robotics.

The distance metric between the two BITSS images is interesting to analyse further. One question that emerges is whether transition states can be located by coupling two images through a small number of collective properties, rather than the total distance between all degrees of freedom in the system. A second question concerns landscapes with multiple competing pathways between states. In such cases, it may be possible to access transition states different from the most direct one by using a biased distance metric. A further investigation that is now open to pursue is when discontinuities in the landscape occur at 'stationary points' (now properly referred to as critical points). In this case, a transition state can no longer be defined by its Hessian eigenvalues, but instead is more broadly defined as a region of locally minimal energy that separates two basins of attraction to minima. Overall, it will be interesting to explore how BITSS enables access to even more challenging landscapes, and those not yet amenable to traditional landscape exploration techniques.

Topic	Symbol	Parameter description	Value
BITSS	α	Energy constraint strength	10
	β	Distance constraint error	0.1
	f	Distance reduction factor	0.5
	—	Coefficient recalculation regularity	100 iterations
Buckling	—	Cylinder radius	50
	—	Cylinder length	80
	—	Mesh resolution	1.57
	—	Compressive strain	0.14%
	k^{S}	Stretching rigidity	100
	k^{B}	Bending rigidity	1
Wetting	—	x grid size	200
	—	y grid size	200
	θ_a	Hydrophilic contact angle	60°
	θ_b	Hydrophobic contact angle	110°
	α_{12}	Interfacial width	1.77
	γ_{12}	Surface tension	0.943
	V_0	Droplet volume	4375
	k_{V}	Volume constraint coefficient	10^4
NEB	k	NEB spring constant (LJ-7)	0.1
	k	NEB spring constant (Buckling)	0.01
	k	NEB spring constant (Wetting)	10^{-6}

Table 5.4: Values of the parameters that are used throughout this chapter, unless otherwise specified. Some of the parameters for the buckling and wetting examples are described in the following chapter in section 6.3.

Chapter 6

Energy Landscape Software Library

In this chapter, we discuss the aims and structure of the energy landscape library 'ELLib', which has been developed in order to study large and versatile systems. We detail the various potentials and methods that have been included in the library and how they have parallelised. Finally, there is a demonstration of the library being applied to a structural optimisation problem.

The software library and the program used for the structural optimisation are both accessible in the supplementary materials, detailed in appendix A. Alternatively, the most up-to-date version of 'ELLib' is available on GitHub at the link: `https://github.com/sjavis/ellib`. This chapter provides details about how it can be used and will be further developed into comprehensive documentation in the near future.

6.1 Software structure

The energy landscape software is contained within an object-oriented C++ library that can be imported and run inside a user script or as a part of another program. It has been designed in a modular manner so that different methods and applications can be easily interchanged and new methods can be implemented without difficulty.

At the core of almost all energy landscape methods is some form of energy minimisation, often with modifications, therefore ELLib contains a sub-library specifically for the minimisation methods. Outside of this are the landscape exploration methods which make use of the minimisation, such as pathway and transition state search. However, there is not always a clear division between the minimisation methods and the landscape methods. For example, the genetic algorithm is a minimisation method but is not contained within the sub-library. This distinction is made according to whether the method involves the minimisation of just a single state or multiple states, and whether further minimisation steps are involved within the method.

The minimisation sub-library is split into several core components. Firstly, there is a class called `Potential`. This provides the interface for calculating the energy and gradient for a given set of coordinates, and includes any potential specific parameters. A `Potential` object can be created using functions for the energy and gradient, or using a predefined subclass. The `State` class is used to create systems consisting of a potential and a set of coordinates. These `State` objects are what the landscape and minimisation methods operate upon, and can be used to easily obtain the coordinates, energy, or gradient. Each state can be parallelised onto multiple processors. This uses a `Communicator` object that abstracts away the parallelisation using MPI, the details for which are given in section 6.2 below. The final key component is the `Minimiser` class. This provides the framework for the methods to minimise a single state.

A typical user script would likely use the energy minimisation methods in the following way:

- Read in or initialise any required data.
- Create a `Potential` object using a predefined potential or make their own.
- Modify any potential parameters.
- Create a `State` object using the potential and some initial system coordinates.
- Create a `Minimiser` object, choosing the specific minimiser to use.
- Modify the minimisation parameters.
- Minimise the state.
- Final data analysis / output data.

A typical energy landscape user script may look like:

- Set up data
- Create a `Potential` object and modify any parameters.
- Optionally, create a `Minimiser` object.
- Initialise a landscape method with the potential and minimiser and choose parameters.
- Run the landscape method.
- Final data analysis / output data.

Users may also combine multiple methods into a single script or use them as part of a larger program.

6.2 Parallelisation

A primary aim of the software is to be able to utilise computational clusters for high performance computing in order to speed up large-scale systems. Therefore, the code is designed to be parallelised using MPI (message passing interface), which distributes the memory across the computational processors and uses messages to pass necessary data among them. This has been implemented in such a way as to abstract away the details of the parallelisation. This enables users who are either developing new methods and potentials or primarily interested in applications to make use of the parallelisation with minimal effort or knowledge of MPI.

The parallelisation is performed on each `State` object. By default each state is split among all of the processors, but specific processors may also be assigned to a given state when it is created. The state's coordinates are split into blocks, one for each processor, which are then stored in a vector locally. Any coordinates that the processor requires for computing the energy and gradient of its block are then stored at the end of this vector (the halo coordinates). The updating of these halo regions and all other communication between the processors is handled by an internal `Communicator` object. `State` offers functions to either calculate the energy and gradient contributions specific to each processor's block (`blockEnergy` and `blockGradient`), or the total energy and gradient (`energy` and `gradient`) which uses MPI communication in the background. Likewise, the `blockCoords` functions can be used to read and assign the local coordinate vector, and `coords` can be used to read and assign the whole system coordinates.

In order for a specific potential to make use of the software's parallel framework it must be defined in such a way that the halo coordinates can be determined automatically. To do this, the calculation of the energy and gradient are split into separate elements that each rely only upon a small number of degrees of freedom. These elements also include a label for the type and any parameters specific to that element. The potential must then have an `elementEnergyGradient` function defined that will compute the contributions to the energy and gradient for each individual element. For some potentials it may not be possible to split it neatly into separate elements. If this is the case, then there is a backup `blockEnergyGradient` potential that can be defined. The state's `Communicator` object is passed to this function, which can be used to manually communicate between the processors.

The `Communicator` object itself is initialised automatically when a State is created. This will first split the system coordinates into the blocks for each processor. Currently there is no attempt to optimise this, it requires that the user define the coordinates in a logical order so that the halo regions are kept small. There is potential scope to improve this in the future so that the order of coordinates is changed to reduce the halo regions. Next, each element is assigned to a processor based upon which block contains the most of its degrees of freedom. Any of this element's coordinates that are contained on other blocks are therefore in the halo region. Once this is done, new MPI data types are created that will send the required data to update the halos of the other processors. In order to use this, a `communicate` function is defined which will update the halo of whatever vector is passed to it, typically, the coordinates or gradient vector. The `Communicator` also has several other useful functions, such as `assignBlock` and `assignProc`, which take a vector with the total number of degrees of freedom and return a vector of the local block, or the block and the halo, respectively. Furthermore, there are additional functions included, such as `sum` and `dotProduct`, that first perform operations on vectors locally before summing the result across the processors used in this Communicator object.

6.3 Potentials

An underlying system potential is necessary in order to study an energy landscape. In the software library, these can be created using the `Potential` class or a subclass. This potential can be defined by the user by providing `energy` and `gradient` functions, or a single function to compute both, `energyGradient`. If instead, it is to be run in parallel then the `elementEnergyGradient` and `blockEnergyGradient` functions should be used, as described in the previous section. If the potential also needs to set up parameters, then there is an `init` function, which is called when the potential is used to create `State` object.

The user may also make use one of several predefined potentials: `Lj2d`, `Lj3d`, `BarAndHinge`, `PFWetting`. Each of these potentials is described below.

6.3.1 Lennard-Jones

A Lennard-Jones particulate model has been implemented for both 2D and 3D, which can be accessed using `Lj2d` and `Lj3d`. For a set of particles with positions $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ the model gives an energy of

$$E(\boldsymbol{x}_i) = \sum_{i,j \neq i} 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right] \quad \text{where} \quad r_{ij} = \sqrt{\sum_k (x_{i,k} - x_{j,k})^2}. \tag{6.1}$$

The parameters ϵ and σ are the inter-particle potential well depth and the particle size, respectively. These can be set in the code using the `epsilon` and `sigma` parameters, each beginning with a value of one. The gradient of the energy with respect to each degree of freedom is then given by,

$$\frac{\partial E}{\partial x_{i,k}} = 24\epsilon \sum_{j \neq i} (x_{i,k} - x_{j,k}) \left(\frac{\sigma^6}{r_{ij}^8} - \frac{2\sigma^{12}}{r_{ij}^{14}} \right). \tag{6.2}$$

At present this potential cannot be run in parallel, as Lennard-Jones systems are not the focus of this PhD work. In future parallelisation could be implemented in order to study large clusters with short range interactions. However, this would require the particles to be redistributed among the processors according to their location, which is beyond the current limits of the software.

6.3.2 Bar and Hinge Triangulated Surface Model

The software contains a bar-and-hinge model to simulate thin elastic surfaces (`BarAndHinge`). This represents the surface as a triangular mesh, within which, each bond is an elastic spring, and each pair of adjacent triangles is connected by an elastic hinge. The total elastic energy is, therefore,

$$E = \sum_i \frac{k_i^{\mathrm{S}}}{2} (r_i - r_{0i})^2 + \sum_j \frac{k_j^{\mathrm{B}}}{2} (\theta_j - \theta_{0j})^2, \tag{6.3}$$

where the first sum gives the stretching energy due to each bond i, with length r_i, equilibrium length r_{0i}, and stretching rigidity k_i^{S}. The second sum provides the bending energy, where θ_j is the dihedral angle, θ_{0j} is the equilibrium angle, and k_j^{B} is the bending rigidity of hinge j. In the software k^{S} and k^{B} values can be set using the `setRigidity` function. It is also possible to set the equilibrium lengths and angles, but otherwise they will be calculated from the initial configuration of the mesh. The triangulation itself is applied using `setTriangulation`. This takes a list of the three node indices for each triangle. These must all be listed in the same order, either clockwise or anticlockwise, to ensure that the surface is properly oriented and so that the dihedral angles can be evaluated across the full 2π range.

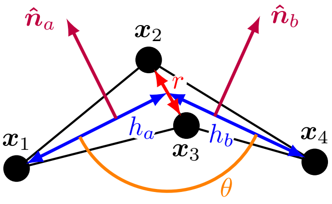


Figure 6.1: Schematic of the bar and hinge model showing the relevant parameters for a single hinge element. h_i and $\hat{\boldsymbol{n}}_i$ denote the height and unit normal of each triangle respectively.

The expressions for the gradient of the energy can be obtained by considering the stretching and bending energies of an individual bond and hinge. For simplicity, the indices for the bond and hinge are ignored. Using the variables shown in the schematic in fig. 6.1, the gradient of the stretching energy of the bond between $\boldsymbol{x}_2$ and $\boldsymbol{x}_3$ is given by

$$\frac{\partial E^{\mathrm{S}}}{\partial \boldsymbol{x}_2} = -\frac{\partial E^{\mathrm{S}}}{\partial \boldsymbol{x}_3} = k^{\mathrm{S}}(r - r_0)(\boldsymbol{x_2} - \boldsymbol{x_3}). \tag{6.4}$$

The gradients of the bending energy of the hinge are

$$\frac{\partial E^{\mathrm{B}}}{\partial \boldsymbol{x}_1} = k^{\mathrm{B}}(\theta - \theta_0)\frac{\hat{\boldsymbol{n}}_a}{h_a}, \tag{6.5}$$

$$\frac{\partial E^{\mathrm{B}}}{\partial \boldsymbol{x}_2} = -k^{\mathrm{B}}(\theta - \theta_0)\left[\frac{\hat{\boldsymbol{n}}_a}{h_a} + \frac{\hat{\boldsymbol{n}}_a + \hat{\boldsymbol{n}}_b}{r}\right], \tag{6.6}$$

$$\frac{\partial E^{\mathrm{B}}}{\partial \boldsymbol{x}_3} = -k^{\mathrm{B}}(\theta - \theta_0)\left[\frac{\hat{\boldsymbol{n}}_b}{h_b} + \frac{\hat{\boldsymbol{n}}_a + \hat{\boldsymbol{n}}_b}{r}\right], \tag{6.7}$$

$$\frac{\partial E^{\mathrm{B}}}{\partial \boldsymbol{x}_4} = k^{\mathrm{B}}(\theta - \theta_0)\frac{\hat{\boldsymbol{n}}_b}{h_b}. \tag{6.8}$$

These models have been widely applied to study origami structures[102,163,164], where the triangular mesh is made to match the geometry of the crease pattern. Alternatively, it can also be used to model continuous elastic surfaces by choosing the spring constants in such a way that they accurately represent the elastic moduli in the continuum limit. In which case, the spring coefficients are given by[165–167],

$$k_{\mathrm{S}} = Yt \cdot \frac{A_{\mathrm{sum}}}{r_0^2}, \tag{6.9}$$

$$k_{\mathrm{B}} = \frac{Yt^3}{12(1-\nu^2)} \cdot \frac{r_0^2}{A_{\mathrm{sum}}}, \tag{6.10}$$

where Y is the Young's modulus, t is the thickness of the surface, and ν is the Poisson's ratio. The value A_{sum} is the total area of the adjacent triangles at equilibrium, of which there will be one for bonds on the boundary, and two for the others. This method is used if the rigidities have not been directly set, and the `thickness` parameter is provided. This can take a different value for each node of the mesh in order to create a varying thickness. In this case an average is taken for each bond and hinge.

This model is particularly useful for its simplicity and speed, however it has some limitations on its accuracy compared to real structures. Firstly, because it is a 2D surface, the thickness is assumed to be much smaller than the dimensions along the surface itself. This is particularly important since the one-dimensional folding of adjacent triangles cannot accurately model the effect of twisting, but this will be negligible as long as the thickness is sufficiently small. This is because any twisting of a thin surface requires some in-plane stretching, the energy of which scales proportionally to the thickness, t, while twisting energy scales as t^3. Similarly, it cannot model the energy due to Gaussian curvature, the product of the two principal curvatures. However, the high resistance to stretching makes this negligible in most circumstances, and, if the boundary is fixed or the surface is closed, the total energy due to the Gaussian curvature is topologically constant according to the Gauss-Bonnet theorem[168], so it may be safely ignored. Finally, another limitation of the model is that the Poisson ratio under stretching cannot freely be tuned. Instead, it depends upon the geometry of the mesh. For example, if an equilateral triangular mesh is used then the Poisson ratio will be 1/3, regardless of the spring constants that are applied[165]. These limitations mean that the model is primarily of use to study elasticity problems where bending is the dominant effect.

An additional energy contribution is possible by incorporating an external force with the `force` parameter. This can be a uniform force to model gravity, or a magnetic field; or applied inhomogeneously across the structure such as to initialise random structures. It is also possible to include interactions with a substrate using `setWall`, which turns on a Lennard-Jones 3-9 potential, which is obtained by considering an interaction between a particle and an infinite wall of particles that interact by the standard Lennard-Jones potential, eq. (6.1).

$$E(h) = \epsilon \left(\frac{2}{15} \left(\frac{\sigma}{h - h_0} \right)^9 - \left(\frac{\sigma}{h - h_0} \right)^3 \right), \tag{6.11}$$

where the height above the wall, h, has been shifted by h_0 to ensure that the energy is minimal at $h = 0$. The Lennard-Jones parameters σ and ϵ are initially set to 10^{-12} and 10^{-5}, which tend to give results that correspond to levels of adhesion observed in experiments. But the precise adhesion energy, controlled by ϵ, will depend upon system size and so should be adjusted for different applications. The adhesion arises because the potential is attractive for $h > 0$, but if this is not desired, it can be switched of, in which case the potential is cut off at $h = 0$ and shifted so that $E(0) = 0$.

This potential uses the software's parallel framework by splitting the contributions to the energy into energy elements. The stretching energy uses one element for each pair of particles, and the bending energy has one element for the four particles in pair of triangles. If the force and substrate interactions are switched on, they each have one element per particle.

6.3.3 Phase Field Model

In order to model multicomponent fluid systems a phase field model has been implemented, `PFWetting`. This potential has been implemented with two variants: a binary fluid and an N-fluid version, which can be accessed by modifying the `nFluid` parameter. This method uses a square grid (resolution size δ) where each node is marked as either solid or a fluid using `setSolid`. If a fluid node is adjacent to a solid node then the solid surface passes through the center of the node, and the volume V and surface area A take different values from the bulk. Each fluid has a concentration, C_i, for each fluid i (for the N-fluid variant) or an order parameter, $\phi = C_1 - C_2$ for the binary model. These concentrations take bulk values of 0 or 1 depending on if the fluid is located at that position, and smoothly vary, resulting in diffuse fluid interfaces.

This potential has four or five contributions to the overall energy depending upon whether it is using the binary or N-fluid model. One sets the bulk fluid values, E^{B}, and another includes the interfacial energy of the fluids, E^{I}. If the N-fluid model is used, there is also a soft energy constraint to ensure that the total concentration is the same everywhere, E^{D}. Next, there is a fluid-solid interaction, E^{S}. Then, finally, there is either an energy penalty to set a constant volume across the system, E^{V}, or the energy due to the pressure of the system, E^{P}.

The E^{B}, E^{I}, and E^{D} terms each apply to the whole of the fluid domain. Therefore, they are split into single elements for each fluid node in the system so that the potential can be computed in parallel. The energies of these discrete elements are then given by,

$$E^{\mathrm{B}} = \begin{cases} \frac{\kappa}{16}(\phi+1)^2(\phi-1)^2V & \text{binary,} \\ \sum_{i=1}^{N} \frac{\kappa_i}{2} C_i^2 (C_i - 1)^2 V & N\text{-fluid,} \end{cases} \tag{6.12}$$

$$E^{\mathrm{I}} = \begin{cases} \frac{\kappa'}{4}(\boldsymbol{\nabla}\phi)^2V & \text{binary,} \\ \sum_{i=1}^{N} \frac{\kappa'_i}{2} (\boldsymbol{\nabla} C_i)^2 V & N\text{-fluid,} \end{cases} \tag{6.13}$$

$$E^{\mathrm{D}} = \begin{cases} 0 & \text{binary,} \\ k^{\mathrm{V}} \left(1 - \sum_{i=1}^{N} C_i\right)^2 V & N\text{-fluid.} \end{cases} \tag{6.14}$$

Equation (6.12) uses a double well potential to ensure that the energy is minimised when the fluid concentrations are 0 or 1. Equation (6.13) penalises any fluid interfaces where the gradient of the concentrations / order parameters are non-zero. Equation (6.14) ensures that the sum of the concentrations is equal to one by using a large constraint coefficient k^{V}. The parameters κ_i and κ'_i are related to the interfacial tensions, γ_{ij}, and widths, α_{ij}. This can be determined by finding the fluid concentration function that minimises the energy given by the continuous analogue of eqs. (6.12) and (6.13), giving the following relations,

$$\gamma_{ij} = \frac{\sqrt{\kappa_i + \kappa_j}\sqrt{\kappa'_i + \kappa'_j}}{6}, \tag{6.15}$$

$$\alpha_{ij} = \sqrt{\frac{\kappa'_i + \kappa'_j}{\kappa_i + \kappa_j}}. \tag{6.16}$$

They are assigned by the `setSurfaceTension` and `setInterfaceWidth` functions, and used to calculate κ_i and κ'_i. However, if four or more fluids are used, the full set of surface tensions cannot be matched because there are not enough parameters in the model. Alternative models that allow all surface tensions to be prescribed are available in the literature, such as[169].

Each of these energy elements includes the neighbouring nodes in order to calculate the gradients in eq. (6.13). Generally these use central difference, however forward difference is used if one side is a solid node. For example, the x-gradient of ϕ at a point where $x = x_0$ is

$$\frac{\partial \phi}{\partial x} = \begin{cases} \frac{\phi(x_0+\delta)-\phi(x_0-\delta)}{2\delta} & \text{for bulk fluid,} \\ \frac{\phi(x_0+\delta)-\phi(x_0)}{\delta} & \text{if solid at } x_0 - \delta, \\ \frac{\phi(x_0)-\phi(x_0-\delta)}{\delta} & \text{if solid at } x_0 + \delta. \end{cases} \tag{6.17}$$

The equations for the energy gradients at each element are therefore given by

$$\frac{\partial E^{\mathrm{B}}}{\partial \phi} = \frac{\kappa}{4}\phi(\phi^2 - 1)V \qquad \text{binary,} \tag{6.18}$$

$$\frac{\partial E^{\mathrm{B}}}{\partial C_i} = \sum_{i=1}^{N} \frac{\kappa_i}{2}(2C_i^3 - 3C_i^2 + C_i)V \qquad N\text{-fluid.} \tag{6.19}$$

The interaction between the fluid and solid is designed to apply partial wetting behaviour, such that the fluid interface meets the solid surface at a specific contact angle, θ. This contact angle can be defined by a function across the domain using `setContactAngle`. The interaction uses a cubic wetting potential, where the gradient is zero at the bulk fluid concentrations in order to avoid enrichment of the fluid concentration at the surface. This energy relies only upon the local fluid values at the surface, so the energy is split into one element per surface node with

$$E^{\mathrm{S}} = \frac{1}{\sqrt{2}}\cos\theta\left(\frac{\phi^3}{3} - \phi - \frac{2}{3}\right)A. \tag{6.20}$$

Finally the pressure and volume energies are given by,

$$E^{\mathrm{P}} = pV_{\mathrm{tot}}, \tag{6.21}$$

$$E^{\mathrm{V}} = k_{\mathrm{V}}(V_{\mathrm{tot}} - V_0)^2, \tag{6.22}$$

where p and V_0 are the chosen pressure and volume, and V_{tot} is the total volume of the fluid. Unlike the energy contributions, the calculation of V_{tot} relies upon the whole system, so it cannot be assigned to separate energy elements and requires special treatment using the `blockEnergyGradient` function. Therefore, the fluid volume is first calculated on each block and then added up across the processors, reducing the overhead required for MPI.

6.4 Implementation of minimisation methods

The `Minimiser` class contains the framework for optimising a single state to a (usually local) minimum in the energy landscape. Therefore, it does not encompass ensemble minimisation methods, such as Genetic Algorithm, which will be discussed in the next section. The methods that have been implemented are:

- Gradient descent (`GradDescent`)
- L-BFGS (`Lbfgs`)
- FIRE (`Fire`)
- Simulated annealing (`Anneal`)

These methods are themselves subclasses of `Minimiser` and implement functions that are called during a minimisation. The first is an `init` function that can be used if any set up is required at the start, including the instantiation of parameters. Then the main loop is run, during which the `iteration` function is called which contains the main part of the methods. At the end of each iteration a `checkConvergence` function detects if it has reached the minimum, at which point the method ends. A user defined function may also be passed into the minimisation that is called at the start of each iteration to modify system parameters. This provides flexibility to the user and could be used, for instance, to steadily perturb the system during the start of the minimisation, or to save the trajectory of the minimisation.

`Minimiser` also has parameters generic to most or all of its subclasses. Namely, the maximum number of iterations and which line search method should be used, if at all. Currently, backtracking line search is included, which ensures that the energy decreases sufficiently to satisfy the Armijo rule, eq. (4.2), with $c_1 = 0.5$, or optionally, that the energy never increases more than a chosen amount.

6.4.1 Gradient descent

The gradient descent implementation has one additional parameter `alpha`, which is the step size factor, α_i, in eq. (4.1). This has a default value of 0.1, although the actual step size will be determined using a line search. So, unless it is turned off, the value of `alpha` is not hugely important. In order to check for convergence, the root mean square of the gradient, $|\boldsymbol{\nabla} E| / \sqrt{N}$, where N is the number of degrees of freedom, is compared to the `convergence` parameter that is chosen for the `State` object. This same criterion is also used for both L-BFGS and FIRE.

The gradient descent algorithm is also very amenable to parallelisation by distributing the degrees of freedom between processors. This is because communication between the processors are only required at three points: to communicate the gradient to the halo nodes, computing the magnitude of the gradient for the root mean square, and calculating the energy for the line search. The latter may require multiple communications, depending upon the number of iterations, but if no line search is used then just the two communications are necessary.

6.4.2 L-BFGS

As for the implementation of L-BFGS, there are two additional parameters. The first is the number of steps to store to calculate the search direction, m, which has a default value of five. Secondly, `maxStep` can be used to restrict the step size in each iteration, such that, if the step is greater than this, it is rescaled to the maximum value. This can be useful, for example, to ensure that the minimisation is smooth when the trajectory is used to approximate an MEP. If this value is not manually set, then no restriction is made.

Currently, if L-BFGS is run in parallel, then only the calculations of the gradient and line search are split among the processors, while the calculation of the search direction is instead performed in serial. This is because $2m + 2$ dot products are required when calculating eqs. (4.7) to (4.10), and each would require a separate communication. Therefore, the search direction is calculated in serial by collecting the whole of the gradient on each processor in an attempt to reduce the number of MPI communications.

6.4.3 FIRE

The `Fire` object has a parameter `dtMax`, which is the maximum timestep that is allowed when updating the velocity and the position. Ideally, an appropriate value should be assigned by the user. However, if none are provided, then a reasonable value will be estimated using the energy gradient at the start of the initialisation,

$$\Delta t_{\max} = \frac{1}{10\sqrt{|\boldsymbol{\nabla} E|}}. \tag{6.23}$$

This will result in a step size that is independent of the scale of the energy, so, if the gradients are very high it will not cause very large step sizes. The step size would instead be on the order of 0.1, which may not be accurate for particularly large or small system sizes. In the unusual circumstance that the gradient is initially zero, `dtMax` is simply set to one. Since FIRE is an inertial optimisation method that uses a continually updated velocity, the line search is off by default. However, it can be switched on, in which case the importance of the timestep is reduced.

In contrast to L-BFGS, FIRE does not require so many communications between processor nodes, and so is efficient for parallelisation. Aside from communicating to update the halo nodes, the only other communication that is required is to calculate the magnitudes of the gradient and the velocity, and their dot product. Therefore the method never requires collecting all degrees of freedom on a single processor.

6.4.4 Simulated Annealing

To use simulated annealing an `Anneal` object can be created using two parameters: `displacement` and `tempInit`. Suitable values for these cannot be easily guessed, so the user is required to provide them. The former is used in calculating the random trial state for each iteration. For each degree of freedom a perturbation is picked from a uniform distribution where the maximum value is given by `displacement`. The second parameter `tempInit` describes the initial temperature, which should be high enough to allow the system to adequately explore the energy landscape. The precise function describing how the temperature decreases over time may be set by the user using the `coolingSchedule` parameter, but the default function uses linear multiplicative cooling[170], such that, at iteration i, the temperature is

$$T(i) = \frac{T_0}{1 + ki}, \tag{6.24}$$

where T_0 is the initial temperature, and k is the cooling rate, which can be adjusted via the parameter `coolingRate` which has a default value of 1.

Because simulated annealing is not a gradient-based minimiser and it is attempting to find the global minimum rather than local minima, convergence should not be determined using the root mean square of the gradient. Instead, the state is determined to be converged once the trial state has been rejected more than a certain number of iterations in a row, which is given by the `maxRejections` parameter. This is initially set to zero, in which case the method is allowed to continue until it reaches the maximum number of iterations.

Simulated annealing is performed almost completely in parallel. The only communications required are to sum up the total energy of the state, to communicate the random perturbations to the halo coordinates, and to share the same random test value so that all processors agree on whether to accept or reject the step.

6.5 Implementation of energy landscape methods

The main ELLib library contains the more general energy landscape methods that typically make use of the minimisation algorithms in order to find transition states, pathways, and in the search for minima. Because they cover a variety of methods, they do not follow any rigid structure, unlike the minimisation methods. Each method is implemented in a class which can be used by creating an object for the method, modifying any parameters, and then calling a `run` function. This function outputs the result for the method, whether that be a global minimum configuration, transition state, or pathway. Each class may also have functions for obtaining other information from the method.

The currently implemented methods are the genetic algorithm for finding global minima, the BITSS method for locating transition states, and NEB to find pathways. In the future we hope to expand this to include other methods, such as basin hopping, eigenvector-following, and the string method.

6.5.1 Genetic Algorithm

The genetic algorithm is implemented in the `GenAlg` class. This takes a `Potential` object that is be used to generate a population of states. The size of this population is given by `popSize`, and the total number of generations, `maxIter`. Each of these are initially set to 100.

There are two possibilities for initialising the population. Firstly, a function may be provided by the user to individually generate states (assiged using `stateGen`). Alternatively, `bounds` can be set to the upper and lower bounds for each degree of freedom. The coordinates for the states are then randomly generated between these bounds. If neither of these are set then there is no way of determining suitable coordinates for the states, so the method will not run.

In order to create a new generation, the parents can be selected using one of two methods, either roulette selection (the default), or simply choosing the most fit. This is set by the `selectionMethod` parameter, with the number of parents chosen as fraction of the population according to `selectionRate`. Elitist selection is also possible by setting `numElites` to be the number of the best states that should be passed directly to the next generation. The remaining children are then created using uniform crossover between pairs of parents. Each of the child's coordinates may then be mutated with a probability given by `mutationRate`. This mutation is chosen from a uniform distribution with a standard deviation chosen by the `perturbation` parameter. If this is not set, but the bounds are, then the standard deviation is 1% of the total allowed range.

The genetic algorithm implementation also has a couple of other useful features. Firstly, it can also be run as a Lamarckian hybrid approach by passing a `Minimiser` object to use for the local search. This allows it to more quickly identify and explore local minima in the landscape. Secondly, a generic user-defined function can be called in each generation to access and modify the population. This can be used, for example, to output the fittest state from each generation.

The standard behaviour is to continue until the full number of generations has been reached, but there are two possible methods for stopping early if it is deemed to have converged. The first is a user-defined energy limit, such that if a state is found with a lower energy it will stop. This could be useful if the precise global minimum is not needed and all that is desired is a low energy solution. The second method halts the method if it has gone a certain number of generations without seeing any improvement in the fittest individual. These two methods and their criteria can be applied using the `setConvergence` function.

The method by which the genetic algorithm is parallelised depends on the number of processors relative to the size of the population. If the population size is greater than or equal to the number of processors then each state is run on just a single processor. This is particularly efficient because the processors only need to communicate when setting each new generation, while the energy of each state is computed in serial. On the other hand, if there are more processors, then each state is distributed among two or more processors. This may be useful if the potential is particularly complex or a hybrid method is being used. This parallelisation is most efficient when the population size is a multiple / factor of the number of processors. Otherwise, some processors will be assigned a larger share of the work, and the others will spend time idle.

6.5.2 BITSS

The `Bitss` class provides the implementation of the BITSS method, and is initialised using two states and a minimiser. These two states provide the initial coordinates for the method, so they are typically selected to be the minima. The `Bitss` class creates a new `State` object and an associated `BitssPotential`. These wrap around the two individual states in order to calculate the full energy of the pair of states, including the constraints.

The most significant parameter for BITSS is the fraction to reduce the distance for each iteration. This is given by `distStep` with a value of 0.5, but this may be too rapid for some applications, and should be changed accordingly. Typically, the method should run until convergence or failure, but `maxIter` can be set to halt the method after a large number of iterations, with an initial value of 100. If `distStep` is reduced significantly then this will likely need to be increased. To calculate the separation between the states the Euclidean distance is often used, but the user is free to use any metric. This can be done by passing functions for the distance and its gradient to the `distFunc` parameter.

The constraint coefficients are recalculated throughout each minimisation according to the `coefIter` parameter, which is initially every 100 iterations. During this recalculation the energy barrier is estimated using a linear interpolation, but for some applications this may be many orders of magnitude too large when the states are far from TS. To prevent this from causing an issue, the `maxBarrier` parameter may be set to a reasonable value. It is also possible to alter the α and β parameters from eqs. (5.5) and (5.7), using `alpha` and `beta`, but they can usually be left at their recommended values of 10 and 0.1, respectively. To prevent issues associated with recalculating the coefficients, it is recommended to switch off the line search for the minimisation. Furthermore, if no minimiser is provided then it defaults to L-BFGS with the line search automatically switched off.

The convergence to the TS can be based on one of five methods. Two of these are based upon the separation constraint reaching a chosen size, whether that be the actual distance, or the distance relative to the initial separation. The next criteria are based upon the estimate of the TS. The third criterion relates to the root mean square of the gradient at the estimate, while the fourth uses the change of the estimate from one iteration to the next. Finally, the uncertainty of the TS energy can be used. This is calculated as the difference between the energy of the TS estimate, E_{TS}, and the pair of states, E_i, relative to its height above the initial state energies, $E_{i,0}$,

$$\epsilon_E = \frac{E_{\mathrm{TS}} - \langle E_1, E_2 \rangle}{E_{\mathrm{TS}} - \langle E_{1,0}, E_{2,0} \rangle}, \tag{6.25}$$

where $\langle \cdots \rangle$ denotes the average. Each of these methods can be set using the `setConvergenceMethod` function, which is initially set to a relative distance criteria of 1%.

The coordinates of the TS are output by the `run`, but other information can also be extracted. The final two states can be obtained using the `getPair` and `getPairCoords` functions, which return two `State` objects or vectors, respectively. Meanwhile, the `getTS` and `getTSCoords` functions give the state and coordinates of the TS estimate midway between the pair. `Bitss` can also report if it believes it has not successfully found the TS if the `getFailed` function is called. This is based upon any of three conditions: if the final state did not properly converge, if $\langle E_1, E_2 \rangle < \max(E_{1,0}, E_{2,0})$, or if $E_{\text{TS}} < \langle E_1, E_2 \rangle$.

The BITSS method can be run in parallel to distribute the calculation of the energies and gradients of the individual states, but the energy contributions of the constraints are computed in serial. This is because the total energy needs to be known for the energy constraint, and the full set of coordinates are needed to calculate the separation between the two states. However, the effect of this is relatively minor because the calculation of the underlying potential tends to be much more costly.

6.5.3 Nudged elastic band (NEB)

The `Neb` class encompasses several variations of the NEB method, including a hybrid climbing image method and DNEB. The latter of which may also be accessed using the `Dneb` wrapper class. There are two ways that the method may be initialised, either using a set of coordinates for the initial chain, or by passing the end points and the number of states (including the end points), in which case the chain is set by a linear interpolation.

Similar to BITSS, the NEB implementation uses a single `State` object with `NebPotential` which calculates the total energy of the elastic band. This delegates the computation of the single state energies to a set of `State` objects for each image along the chain. This main `State` is then minimised to find the pathway. Because NEB and DNEB modify the gradient of the energy, the minimisation should, again, not use a line search, and the default L-BFGS minimiser has it switched off by default.

The elastic spring constant, k in eq. (4.22), can be set using the `kSpring` parameter, which is initially set to one. This should not be too large or too small because it may cause the potential or equal spacing constraint to be washed out and the method may take longer to converge. The tangent vectors used to project the components of the forces are obtained using the bisection method, described by eq. (4.27).

To utilise a hybrid climbing image method the `setHybrid` function may be called. This is also used to set the `hybridIter` variable, which activates the climbing image at a chosen iteration, allowing the the chain to first approximate the pathway. After this, at each iteration, the state with the highest energy is selected and its driving force is replaced by the climbing image force, eq. (4.14).

The `run` function performs the initialisation and optimisation of the chain and outputs the coordinates along the pathway. Subsequent outputs can be acquired using `getChain`, `getCoords`, and `getEnergies`. These methods provide the final list of `State` objects, the list of coordinates, and the list of energies, respectively.

The method allows the chain of states to be distributed among the available processors so that the calculations of the single state energies and gradient are performed in parallel. These are either allocated with multiple states per processor or multiple processors per state, using the same approach as the genetic algorithm. The results from the individual states are then broadcast to all processors so that the total energy and gradient can be computed serially. Although the full energy and gradient could possibly be calculated in parallel, this would significantly complicate the method, and much like BITSS, the main computational cost is usually associated with the single state energy calculations.

6.6 Demonstration: structural optimisation

In this section the versatility of the software library will be demonstrated by optimising an elastic structure to maximise the energy barrier between two states, thereby identifying the structure for which these states are at their most stable. The structure that will be investigated is a table-shaped buckled mesostructure that will be studied in further detail in the next chapter. Fig. 6.2(a) shows a schematic of this structure with the four parameters that will be optimised: the leg width, the thickness, and the outer and inner radii of the centre. At the end of legs are bonding sites that are compressed by a 20% strain and then fixed in place. The structure will not restricted by a substrate, so may buckle down as well as up. The two stable states, between which the barrier will be computed, are shown in fig. 6.2(b).

The parameters are bounded during the optimisation to ensure that they remain realistic. The thickness, for example, would likely continue to rise if left unrestricted. Therefore, a maximum is placed to ensure that it does not become unreasonably large and respects the model's requirement that the thickness remains relatively small compared to the other dimensions of the structure. The parameters are given relative to the total length of the structure, ignoring binding sites, which is given a size of 1. It also depends upon the resolution of the mesh, which is made as close as possible to an equilateral triangular mesh with side lengths of 0.05. These parameter bounds are listed in table 6.1. The other fixed parameters are listed in table 6.2.

Parameter	Minimum	Maximum
w	0.0866^a	0.25
t	0	0.05
r_{out}	0	1/3
r_{in}	0	r_{out}

[a] This is two times the optimal triangular height.

Table 6.1: The bounds on each of the four structural parameters.

To perform this structural optimisation it is necessary to combine a variety of energy landscape methods. First, energy minimisation is required to identify the two stable states for each set of parameters. L-BFGS will be used for this. Then, the transition state between the states will be identified using BITSS. At this point the energies of the TS, E_{TS}, and minima, $E_{1/2}$, are known, so the energy barriers in each direction can be evaluated and added together to give the 'total energy barrier',

$$E_{\text{B}} = 2E_{\text{TS}} - E_1 - E_2. \tag{6.26}$$

To maximise this, the four structural parameters will be optimised using the genetic algorithm. A script is written that carries out this procedure using ELLib, with its structure described here.

In order to be used with the genetic algorithm a `Potential` subclass is created, whose energy is the inverse of the barrier given in eq. (6.26). Therefore, minimising this energy will maximise the barrier. The potential also takes a blank function for the gradient because the genetic algorithm does not require it.

The energy function takes four coordinates between 0 and 1, which it converts into the parameters using the bounds in table 6.1. An external program is then called to generate a mesh based upon the parameters, using a Delaunay triangulation[171] to make the triangles as close to equilateral as possible. Using this mesh, the thickness, and the `BarAndHinge` potential, an initial `State` object is created. Two copies of this state are then minimised using `Lbfgs`, each with a function that is called for the first 500 iterations of the minimisation. This function has two roles. First, it slowly moves the binding sites in until they reach the final position. Second, it applies forces to the structure to obtain the desired configuration. After 500 iterations these are turned off to allow the structure to relax. Once the states have converged, they are checked to ensure they converged correctly. If not, the energy barrier is determined to be zero, and the function ends. Otherwise, the two minimum states are passed to `Bitss` in order to find the TS. At this point the barrier is evaluated using eq. (6.26), and the function returning the inverse of this value.

Using this potential, the script creates a `GenAlg` object and sets the following parameters: `popSize` $= 48$, `maxIter` $= 20$, `bounds` $= [0, 1]$, `perturbation` $= 0.3$, `mutationRate` $= 0.25$, `selectionRate` $= 0.4$. Additionally, it uses elitism to ensure the best structure is not lost, and it uses a function to save the population in each generation. Finally, the genetic algorithm is run until the full 20 generations are complete.

The results in fig. 6.2 show how the optimal structure evolves over time. The leg width and thickness very quickly converge to their lower and upper bounds, respectively, and the inner radius also becomes small early on. However, the outer radius takes more time to evolve, beginning with a large value, and only changing to a small

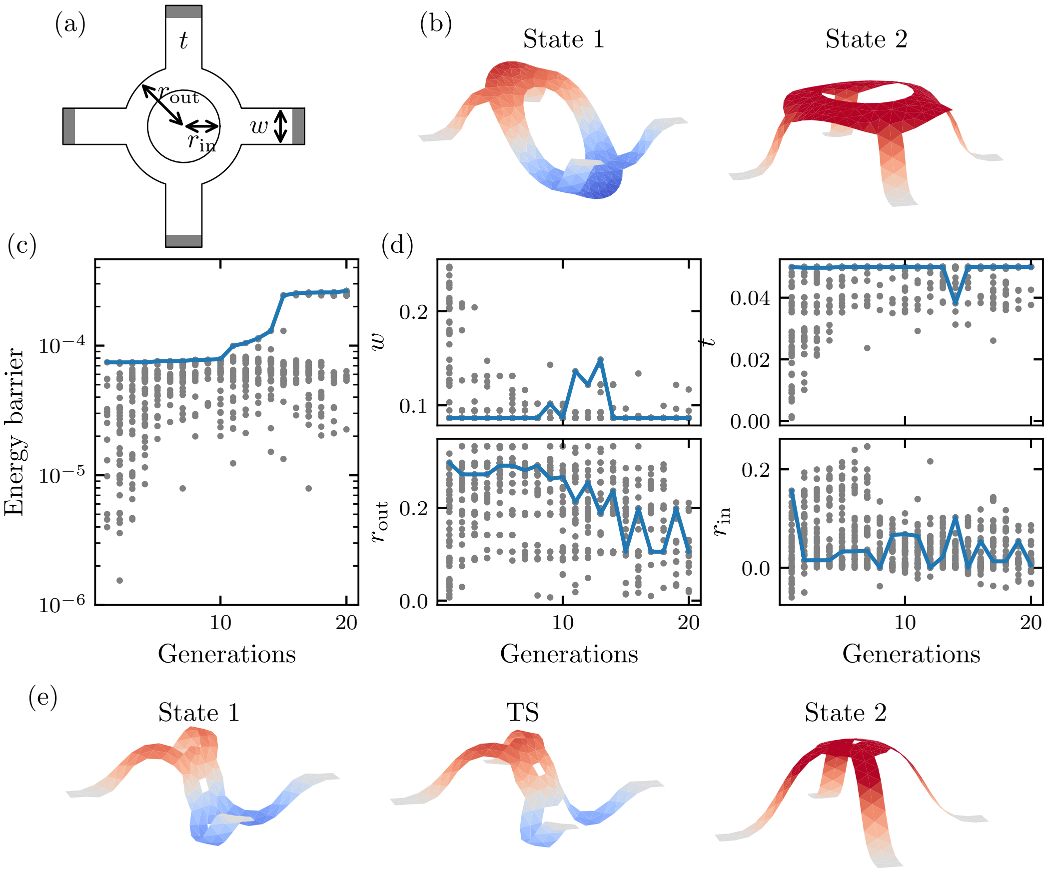


Figure 6.2: (a) Schematic of the structure showing the four parameters: leg width, w, thickness, t, outer radius, r_{out}, and inner radius, r_{in}. The fixed bonding sites are shown in grey. (b) The two minima for the structure with the highest barrier in the first generation. The colour represents the height above the horizontal plane, positive height in red, negative in blue. (c) The energy barrier for each individual in the population. The blue line follows the states with the highest barrier in each generation. (d) The values for each of the four parameters for each individual throughout the optimisation process. The parameters are scaled such that the total length of the structure, excluding bonding sites, is 1. (e) The minima and transition state for the most optimal state after 20 generations.

radius around generation 15, at which point the maximum barrier correspondingly increases. After 20 generations the optimal structure is: $w = 0.0866$, $t = 0.05$, $r_{\mathrm{out}} = 0.1145$, and $r_{\mathrm{in}} = 0.0137$.

These results demonstrate that a highly complex optimisation study is feasible using ELLib, and more generally, illustrates the flexibility of the library.

Topic	Symbol	Parameter description	Value
Structure	—	Compressive strain	20%
	—	Total structure length	1
	—	Mesh resolution	0.05
	Y	Young's modulus	1.4
	ν	Poisson ratio	0.3
BITSS	α	Energy constraint strength	10
	β	Distance constraint error	0.1
	f	Distance reduction factor	0.15
	—	Coefficient calculation regularity	100 iterations
	—	Max BITSS iterations	30
Genetic Algorithm	—	Population size	48
	—	Number of generations	20
	—	Mutation rate	0.25
	—	Parent selection rate	0.4
	—	Pertubation size	30%

Table 6.2: Values of the parameters that are used in this section.

Chapter 7

3D Buckled Mesostructures

This chapter is based upon the joint work: Y. Li, S. J. Avis, J. Chen, G. Wu, T. Zhang, H. Kusumaatmaja, and X. Wang, "Reconfiguration of multistable 3D ferromagnetic mesostructures guided by energy landscape surveys", *Extreme Mech. Lett.*, 48, 101428 (2021).

In particular, the experimental ferromagnetic structures are the work of Y. Li, while the simulations and energy landscape analysis are my own work.

7.1 Introduction

Reconfigurable three-dimensional structures that can actively change their geometries and thereby their functionalities upon external stimuli (like mechanical forces, magnetic fields, hydration, and temperature)[10,88,172–175] are promising for a diverse range of applications including deployable solar panels[176], metamaterials[177–179], phononics[180,181], biomedical devices[182,183], robotics[7,184,185], metasurfaces[186,187], and many others. The design of reconfigurable structures usually relies on structural instabilities[11,188,189], stimuli-responsive constituent materials (like swellable gels, shape memory polymers and magnetic materials)[174,190–193], or their heterogeneous combinations[194].

As discussed in section 3.2, structural instabilities of simple 2D patterns under compression have been exploited to produce 3D buckled mesostructures, and can be used to realise a diverse range of complex 3D functional architectures and electronics[72,195–198]. Through elastically deforming the assembly platform in different time sequences (i.e., loading-path control), the 3D structures have been shown to morph into different shapes[6,199]. In addition to mechanical deformation, reconfigurable structures can also be realised through active materials including ferromagnetic composites[173,200–202], hydrogels[174], shape memory polymers (SMPs)[175,196], liquid crystal elastomers (LCEs)[203], and electrochemically active materials[204], which can change their shape upon external stimuli.

In recent years, an emerging attractive strategy to realise robust reconfigurable structures is to encode multiple stable states by design that do not require persistent external stimuli[205–208]. Each of these stable states is a local minimum in a complex energy landscape, and therefore only a temporary external stimulus is needed to guide the structure from one state to another, after which it will remain in that state.

Despite intensive studies, a number of critical challenges remain, especially when asymmetric and complicated modes are involved in the architectural reconfiguration. The first challenge is associated with the presence of multiple local minimum configurations, which could affect the stability of the targeted stable state under perturbations (e.g., environmental noises) or trap the structure in an intermediate state during the process of reconfiguration. For example, it is recently recognised by researchers that hidden local minimum configurations can destroy the designed pathway of deploying origami structures[100,209]. Thus, it is crucial to provide means of manipulating the energy well depth of targeted configurations and eliminating unfavorable local minima in the design of reconfigurable structures.

The second challenge stems from the complexity of the transition paths from one local minimum state to another, especially when the two states are separated by other stable and transition states[61,83,101,210,211]. A recent work[101] shows that the transition from a uniformly deformed cylindrical shell to a 9-dimple buckled pattern needs to pass 7 local minimum configurations, which cannot be achieved by directly applying local probes to the shell. Complicated transition paths among local minima are also found in mechanical metamaterials, which further provides new design parameters to program the deformed configurations of the structures[83].

Finally, it is also very challenging to realise remote, locally controlled, and rapid reconfiguration of the deformed, highly nonlinear structures[6,212]. These challenges call for a new means of systematically surveying the energy landscapes of the multi-stable structures to probe and tailor the energy barrier height among different local minima.

In this chapter, we develop a set of strategies to address these challenges and to guide the design of ferromagnetic 3D mesostructures, which are developed by our collaborators. The use of ferromagnetic materials enables remote and rapid reconfiguration among different stable states via a portable magnet, eliminating the need for mechanical stages to apply controlled, global deformations to the substrate, as used in previous works[6,199]. Thus providing opportunities for remote, tetherless shape morphing, which is advantageous for many applications, especially those that require shape morphing in enclosed or delicate environments such as robotics and biomedical devices.

Theoretically, we employ a discrete shell model[166,213] and path-finding algorithms[16,101] to conduct fundamental studies on the complicated energy landscapes of the 3D ferromagnetic structures. We construct design phase diagrams showing how the available stable states sensitively depend on the essential material and geometrical parameters of the elastic structures. Furthermore, the minimum energy reconfiguration paths between the stable states and their energy barriers are identified using a combination of our developed BITSS method and the string method[16]. A salient advantage of our approach is that it allows the computation of reconfiguration paths following the minimum energy pathways (MEPs) systematically without the need to assume a predefined path for the structural transformation.

Our methodology is versatile and is applied to design and realise reconfigurable structures with a wide range of geometries, including ribbons and structures that resemble tables, baskets, flowers, boxes, single and double beams. In all cases, the theoretically predicted and experimentally realised shapes and reconfiguration paths are in excellent agreement. We also demonstrate reconfiguration of structures fabricated from non-active materials that are locally patterned with ferromagnetic films, which significantly expands the design space and material libraries for reconfigurable structures.

To further demonstrate the versatility of these reconfigurable 3D ferromagnetic structures, we highlight a number of applications that have been created by our experimental collaborators. This includes a reconfigurable light emitting system, which illustrates the capacity to integrate other functional components into the 3D structures. Finally, magnetically actuated particle delivery is shown, which utilises the different reconfigurational modes of a table structure.

7.2 Results

7.2.1 Phase diagram of stable states for a table structure

Fig. 7.1(a) schematically illustrates the assembly process of a 3D structure from ferromagnetic composite thin films via compressive buckling, and the architectural reconfiguration of the structure via magnetic actuation. The ferromagnetic composite is prepared by homogeneously embedding hard NdFeB (neodymium-iron-boron) microparticles into a soft elastomer, polydimethylsiloxane (PDMS). These are then patterned into 2D layouts using a CO2 laser, followed by their magnetization using impulse magnetic fields (about 2.7 T). They are then attached to a prestretched elastomeric substrate using a very thin layer of superglue to generate strong bonding at select location on the 2D patterns. Releasing the prestrain in the substrate leads to compressive forces on the 2D ferromagnetic pattern at the bonding sites and geometrically transforms the pattern into a 3D structure. We define the 2D layout to lie on the x-y plane so that its normal vector is in the z direction.

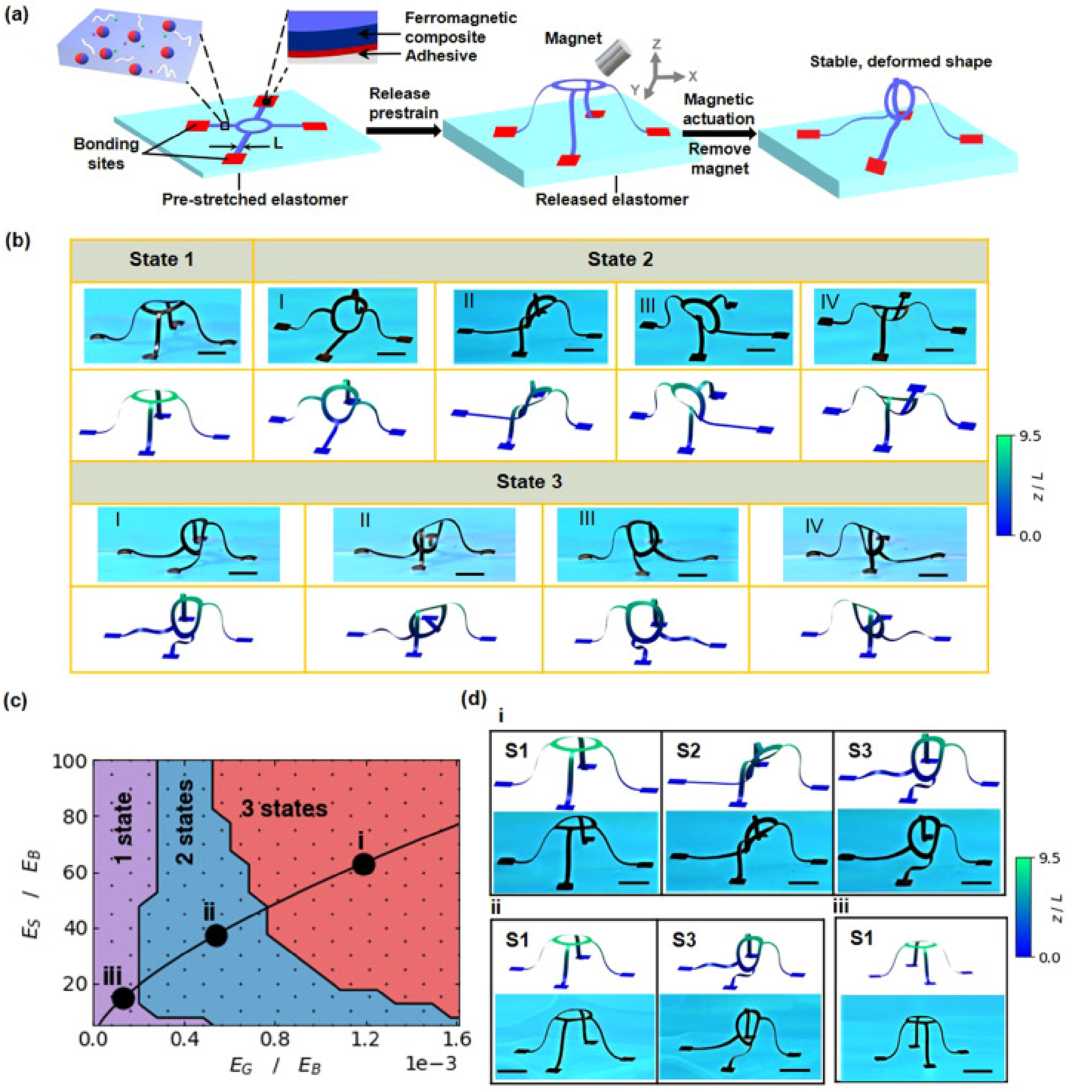


Figure 7.1: Multistability of magnetically reconfigurable 3D structures. (a) Schematic illustrating the assembly and reconfiguration of 3D ferromagnetic structures. (b) Experimental and computational results of multistable states in a 3D table structure. (c) Design phase diagram showing the number of distinct stable states based on the energy scales of the table structure. Representative examples in each regime with different values of L are marked. (d) Experimental and computational results of the distinct stable states for the three representative examples marked in (c). S1-S3 refer to state 1-3, respectively. Scale bars, 1 cm.

To efficiently explore the possible stable states existing in the table structure, we simulate the structure using a simple discrete shell model. The surface is discretised using a Delaunay triangulation[171], with a target mesh size of a tenth of the ribbon width, and then the energy is calculated using the bar-and-hinge model discussed in section 6.3.2. In these simulations, we assign the Young's modulus of the ferromagnetic film to be $Y = 1.4\,\mathrm{MPa}$[200] and use a thickness, $t = 150\,\mu\mathrm{m}$. The gravitational force is also included by setting the film density to its measured experimental value of $\rho = 2.59\,\mathrm{g\,cm^{-3}}$. These parameters, and others used in this chapter, are listed in table 8.3. Finally, we model the repulsive substrate interaction using a Lennard-Jones 3-9 potential with the attractive region removed. Consequently, the model does not take into consideration the effect of adhesion which may become significant at small scales.

The stable states are obtained by minimizing the structures energy using the L-BFGS algorithm[106,107], which is highly efficient for minimization involving a large number of degrees of freedom. The structures are buckled by steadily moving the binding sites from their positions in the flat precursor to their final positions over the first 10 000 iterations. While this buckling is taking place, the structure is split into a number of different regions and each is given a random force to enable the structure to take different configurations. This process is repeated several hundreds of times to identify the different local minima in the energy landscape.

Fig. 7.1(b) presents experimental and corresponding computational results of a 3D ferromagnetic table structure and its multiple stable states. The table structure is oriented in a way such that gravity is perpendicular to the substrate (in the negative z direction), which is the case for most of the structures in this study unless specified otherwise.

Three distinct states (state 1-3) with a total of nine configurations are discovered. Magnetically deforming the assembled table structure via a portable disk magnet, with one leg becoming flat while the other three legs remain buckled, rapidly reconfigures the table structure from state 1 into a distinct stable state 2, with four different orientational configurations (state 2: shape I-IV). Further magnetic actuation induces an additional leg of the table structure to become flat and results in the transition from state 2 into another stable state with four different orientations (state 3: shape I-IV). In addition, unlike previously reported reconfigurable 3D structures that rely on persistent external stimuli to maintain their deformed shapes[173,174,203,214], the deformed configurations of the table structure shown here can well maintain their shapes after the applied magnetic field is removed, as they are locally stable configurations. Computational modelling results of the configurations for the stable states and their different orientations via the discrete shell model are highly consistent with experiments, with the colour in the results denoting the displacement of the structure along the z direction.

Using the combined experimental and computational strategy, we further study the effect of essential geometry and material properties on the entire spectrum of existing stable states, yielding the design phase diagram shown in fig. 7.1(c). The diagram illustrates the total number of stable states according to two dimensionless parameters that are determined from the ratios between three relevant energy scales – gravitational (E_G), stretching (E_S), and bending (E_B). These energy scales can be expressed in terms of the material and geometric parameters of the 3D structure:

$$E_G = \rho g L^3 t, \qquad E_S = Y L^2 t, \qquad E_B = \frac{Y t^3}{1-\nu^2}, \tag{7.1}$$

where ρ is the material density, Y is the Young's modulus, ν is the Poisson ratio, t is the film thickness, and L represents the in-plane size of the structure (the width of table legs is used for this study). The cubic term of L in E_G comes from the fact that the volume of the structure scales with L^2 and the displacement in the direction of gravity is proportional to L.

Three regimes in terms of the number of existing stable states can be identified in fig. 7.1(c): (i) 3 distinct stable states, (ii) 2 distinct stable states, and (iii) 1 distinct stable state, which are controlled by two dimensionless variables $E_S/E_B = L^2(1-\nu^2)/t^2$ and $E_G/E_B = \rho g L^3(1-\nu^2)/Yt^2$. The variable E_S/E_B characterises the bendability of the structure, with a larger value indicating that it is easier to generate bending deformation. E_G/E_B describes the competition between the gravitational and bending energies. When E_S/E_B and E_G/E_B are relatively large, three distinct stable states as shown in fig. 7.1(b) exist in regime i, with states 2 and 3 having 4 different orientations. As the two dimensionless variables decrease, state 2 becomes unstable, resulting in two distinct stable states (states 1 and 3) for regime ii. When E_G/E_B is further reduced to an even smaller value, only the buckled-up state (state 1) is stable in regime iii.

To validate the phase diagram experimentally, we fabricate 3D ferromagnetic structures of the same film thickness (t) and material properties (ρ, Y, and ν) but of different sizes (L), and magnetically reconfigure them into their multistable states. Fig. 7.1(d) (i-iii) demonstrates the number and configurations of the stable states for structures with $L = 1.04\,\mathrm{mm}$, $0.8\,\mathrm{mm}$, and $0.5\,\mathrm{mm}$, respectively, which agree well with the prediction of the design phase diagram. Please note that the overall in-plane dimensions of the structures are scaled proportionally with L as it varies.

The diagram provides very important guidelines for achieving a well-controlled number of stable states or for removing undesired stable states in various applications by tuning the geometry and material properties. For instance, if state 2 is undesired, the structure can be designed with E_S/E_B and E_G/E_B falling within regime ii, where only states 1 and 3 are stable.

7.2.2 Reconfiguration pathways of a table structure

Identifying the optimal transition pathways between the stable states of a 3D structure is critical for well-controlled, energy-efficient reconfigurations. By using the table structure as an example, we further examine its energy landscape to identify the minimum energy pathways (MEPs) between metastable states, in which the maximum energy along the path occurs at a saddle point (the transition state). However, the non-linear reconfigurations of the structure can thwart standard pathway methods unless they are initialised close to the correct path, and the large number of degrees of freedom used to model the structure can make the problem computationally intractable. Instead, to locate the pathway efficiently, a combination of methods are used. First, the transition state is located using the BITSS method, which does not require an initial pathway guess. By minimising downhill from either side of the transition state with L-BFGS a pathway is generated, which is then refined using the string method (introduced in section 4.4.2).

Our approach allows automatic and systematic computation of reconfiguration paths following the MEPs without the need to assume the pathways along which the structure morphs. This is useful because the energy-efficient reconfiguration paths can be highly complex for 3D elastic structures, and it is not always possible to prescribe an expected pathway a priori, as typically adopted in previous works where energy barriers are calculated[6,61,95,215].

As shown in fig. 7.2(a), we find the transition pathways to progress from state 3 to state 1 of the table structure can be broadly grouped into three regions: (I) two paths, 1 and 2, exist (red region); (II) two paths, 1 and 2*, exist (orange region); (III) only path 1 exists (yellow region). The distinction between paths 2 and 2* is that the intermediate state 2 is no longer stable on the latter. For illustration, we construct the design diagram by varying the in-plane size of the structures (L), while keeping their material properties (ρ, Y, and ν) and film thickness ($t = 125\,\mu\text{m}$) constant. The marked red and blue points correspond to those in the 3-state and 2-state regimes identified in the design diagram in fig. 7.1(c), respectively.

For $L > 0.89\,\text{mm}$ (region I), two independent minimum energy pathways are found and illustrated in fig. 7.2(b) for a representative table structure with $L = 1.04\,\text{mm}$. The first (path 1) is a direct pathway with energy barrier ΔE_1, and the second (path

2), a path that passes through state 2 as an intermediary, with energy barriers ΔE_2 between states 3 and 2 and ΔE_3 between states 2 and 1.

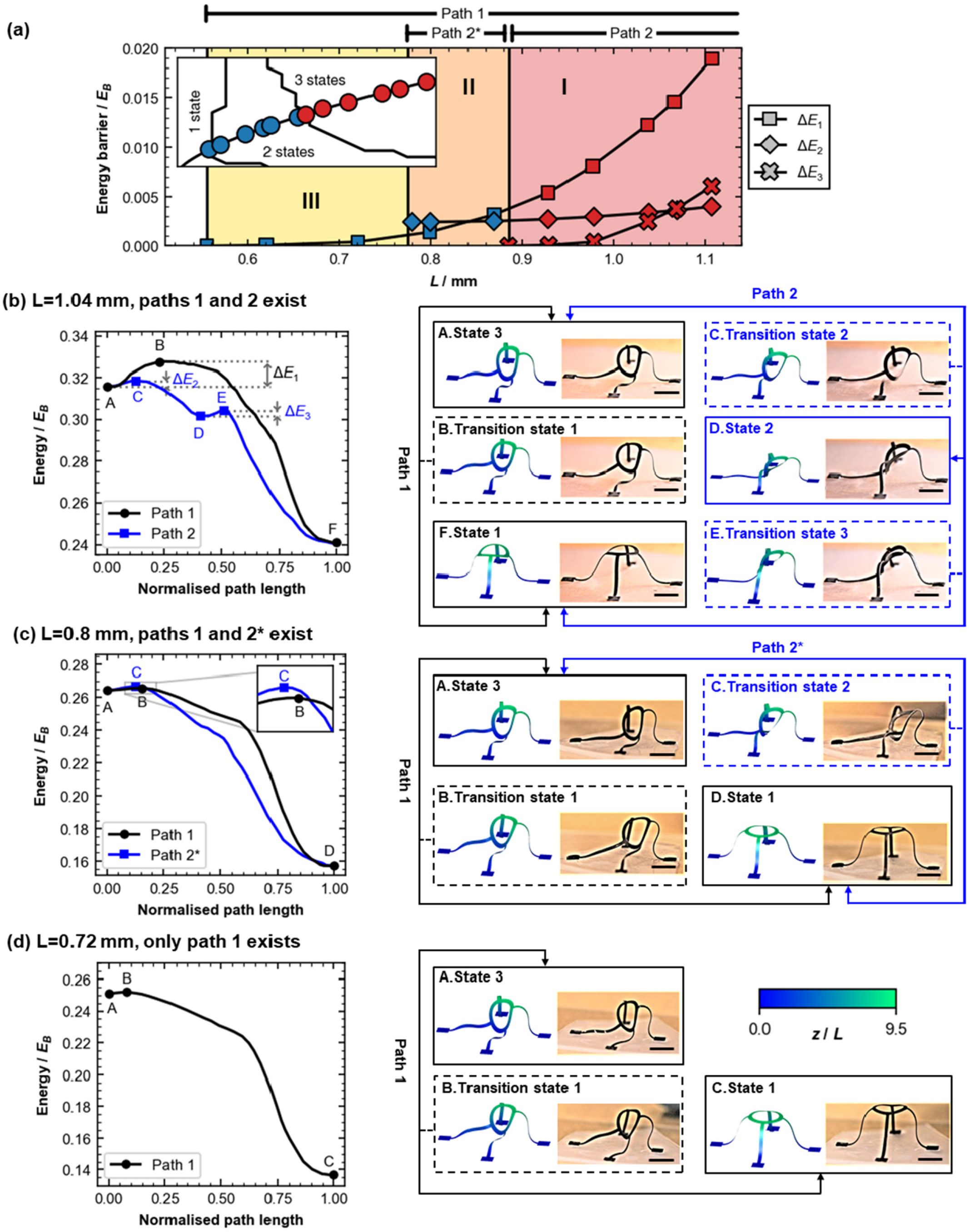


Figure 7.2: Energy barriers and reconfiguration pathways. (a) Variation of the energy barriers against L. Three regions (I-III) with different available pathways are highlighted. Inset: The locations within the design phase diagram. (b-d) Energy profiles and snapshots of stable and transition states for the pathways from state 3 to 1 with three values of L: (b) 1.04 mm, (c) 0.8 mm, and (d) 0.72 mm. The energies are normalised by the bending energy. The definitions of the three energy barriers (ΔE_1, ΔE_2, ΔE_3) are shown in (b). Scale bars, 1 cm.

Snapshots of configurations along the pathways are captured in simulations and magnetic reconfiguration experiments and are shown in fig. 7.2(b), with the complete structural transformation process and corresponding energy profiles simultaneously demonstrated in videos A.1 and A.2.

The energy barriers of the pathways demonstrate both the stability of the states and the ease of transitioning between them. For example, the energy barrier to progress from state 3 to state 1 is much smaller than that from the inverse direction, thereby making state 1 more stable. Comparing paths 1 and 2 also suggests that a smaller input energy is needed when reconfiguring the table structure along path 2 (passing two saddle points) than that along path 1. However, the finite energy barrier ΔE_3 also indicates that the transition can be potentially trapped in state 2 if insufficient external energy is provided for reconfiguration. In addition, reconfiguration along path 2 requires more efforts in manipulating the direction and the strength of external magnetic forces, compared to the direct pathway following path 1.

These complexities clearly show the importance of harnessing the energy landscape analysis for guiding the choice of transition paths based on specific applications and the nature of available external magnetic fields (or other external stimuli). In addition, it is also worth noting that the pathways observed in experiments, where a portable magnet is utilised to reconfigure the structure from state 3 to state 1, are highly consistent with those from simulations, which indicates the high fidelity and reliability of the approach used in this study. Moreover, the experimental reconfigurations can be completed within a few seconds (videos A.1 and A.2). Such fast and remotely controlled transitions are desired for numerous applications, such as in soft robotics[57,216] and multifunctional metasurfaces[186,187].

As we decrease L, all the energy barriers become smaller monotonically, as shown in fig. 7.2(a). At $L = 0.89\,\mathrm{mm}$, the boundary between the 3-state and the 2-state regimes in the phase diagram of fig. 7.1(c) is reached, where ΔE_3 reduces to zero, a signature that state 2 is no longer stable. For $0.78\,\mathrm{mm} < L < 0.89\,\mathrm{mm}$ (region

II in fig. 7.2(a)), there remains two distinct minimum energy pathways to transition from state 3 to state 1, even in the absence of a stable state 2. The typical energy profiles for this range of L values are illustrated in fig. 7.2(c) for $L = 0.80\,\mathrm{mm}$. Snapshots of the configurations from simulations and experiments for $L = 0.80\,\mathrm{mm}$ are further shown in fig. 7.2(c), with reconfiguration along the full pathways provided in videos A.3 and A.4. The shape evolution along the pathways remains qualitatively similar to that shown in fig. 7.2(b).

Interestingly, we also find that there is a crossover in the preferred energy-efficient pathway at $L = 0.84\,\mathrm{mm}$, where $\Delta E_1 = \Delta E_2$ in fig. 7.2(a). For a larger L ($0.84\,\mathrm{mm} < L < 0.89\,\mathrm{mm}$), path 2* is favorable because the corresponding energy barrier is smaller, while for a smaller L ($0.78\,\mathrm{mm} < L < 0.84\,\mathrm{mm}$), path 1 has a lower energy barrier and thereby is desired.

The preference for path 1 at smaller L values is further accentuated by the loss of path 2* as a minimum energy pathway for $0.56\,\mathrm{mm} < L < 0.78\,\mathrm{mm}$ (region III in fig. 7.2(a)). For this range of L values, we are unable to identify path 2* in simulations, as indicated by the lack of data points in fig. 7.2(a). Furthermore, attempts to experimentally reconfigure along this pathway fail, with the structure instead following path 1. Therefore, only path 1 remains. The typical energy profile for the transition along path 1 in this range of L values is shown in fig. 7.2(d) for $L = 0.72\,\mathrm{mm}$, with the corresponding snapshots of the reconfiguration and the full pathway given in fig. 7.2(d) and video A.5, respectively. State 3 itself becomes unstable when L is further reduced below 0.56 mm, where we enter the 1-state regime. At this point the energy barrier ΔE_1 becomes zero, as seen in fig. 7.2(a).

Taken together, these highly nonlinear behaviors observed for the transition pathways make the energy landscape analysis a necessary tool for an efficient and robust design of the multistable configurations and the transition paths among them, which has not been systematically studied before for continuous 3D mesostructures.

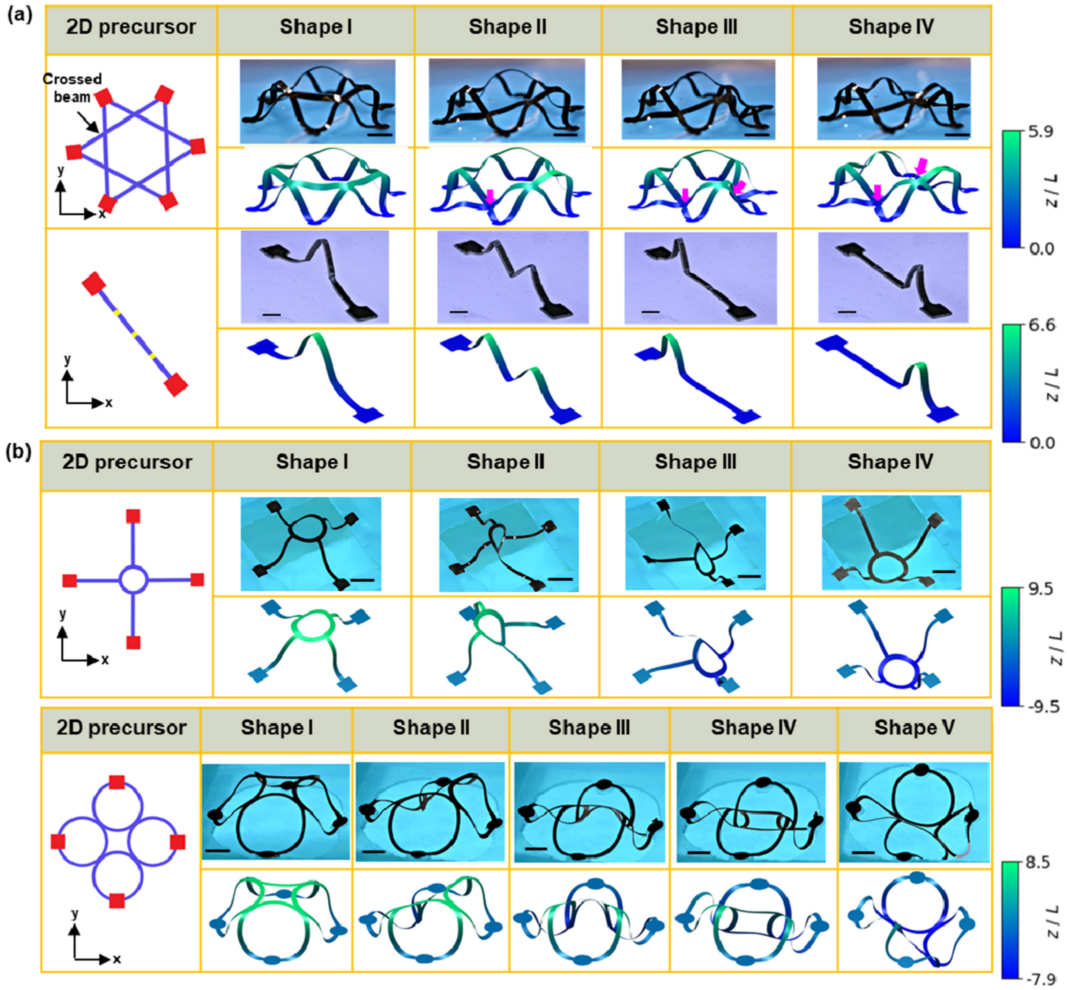


Figure 7.3: Diverse magnetically reconfigurable 3D structures. Optical images and computational modelling of multistable states for three types of structures: (a) 3D structures attached to a substrate. Pink arrows in the first structure indicate differences among the four shapes. Scale bars in the first and second structure are 1 cm and 2 mm, respectively. (b) 3D structures with the substrate underneath the structure removed. Scale bars, 1 cm.

7.2.3 3D structures with diverse geometries

The reconfiguration strategy and the energy landscape analysis described above are versatile and can be extended to different types of 3D structures in diverse geometries, as illustrated in fig. 7.3. Fig. 7.3(a) shows an assembled ferromagnetic basket structure (shape I; film thickness: 125 µm) and its multiple stable states (shapes II-IV) that can be achieved via magnetic control. The pink arrows indicate differences among the four shapes. In addition, by introducing creases to selective locations of a ribbon structure (crease thickness: 70 µm; non-crease thickness: 125 µm), we create a ferromagnetic origami structure, which is further tuned to display multiple stable configurations by using magnetic forces.

The elastomer substrate used for 3D assembly in figs. 7.1 and 7.2 provides essential support for the assembled 3D structures, but it also poses some limitations for the reconfiguration in the out of plane direction. To allow more freedom for spatial reconfiguration, the substrate underneath the 3D table structure is removed, while the substrate adjacent to the bonding sites is maintained to support the 3D structure. Such 3D structure on a hollow substrate allows bending up/down deformations across the plane of the substrate. Fig. 7.3(b) demonstrates a selection from a large variety of reconfigurable 3D structures on hollow substrates that have been investigated, including basket, single- and double-beam structures.

It is very interesting to note that the multistable configurations on hollow substates are sensitive to the orientation of the structures. Take the table structure for example, when placed horizontally (i.e., the direction of gravity is perpendicular to the substrate), the structure only displays two stable states, buckled up and down. However, when the table structure is placed vertically with the direction of gravity in the x direction (fig. 7.3(b)), in addition to the buckled up (Shape I) and down (Shape IV) states, it can deform into an asymmetric, twisted shape (Shape II), similar to state 3 of the table structure with an intact substrate. This deformed shape can be further magnetically deformed into shape III, with the four bonding sites attached to the substrate and the rest of the structure self-supported in the air.

We also carry out energy landscape analysis for this structure, the details of which are presented in fig. 7.4. Similar to the case with an intact substrate, we can distinguish several regimes with different numbers of stable states, depending on the three relevant energy scales: gravitational, stretching, and bending energies.

In addition, we quantitatively compare the dimensions of three representative 3D structures from experiments and modelling results and find negligible difference between them (table 7.1), which further highlights the consistency between modelling and experimental results.

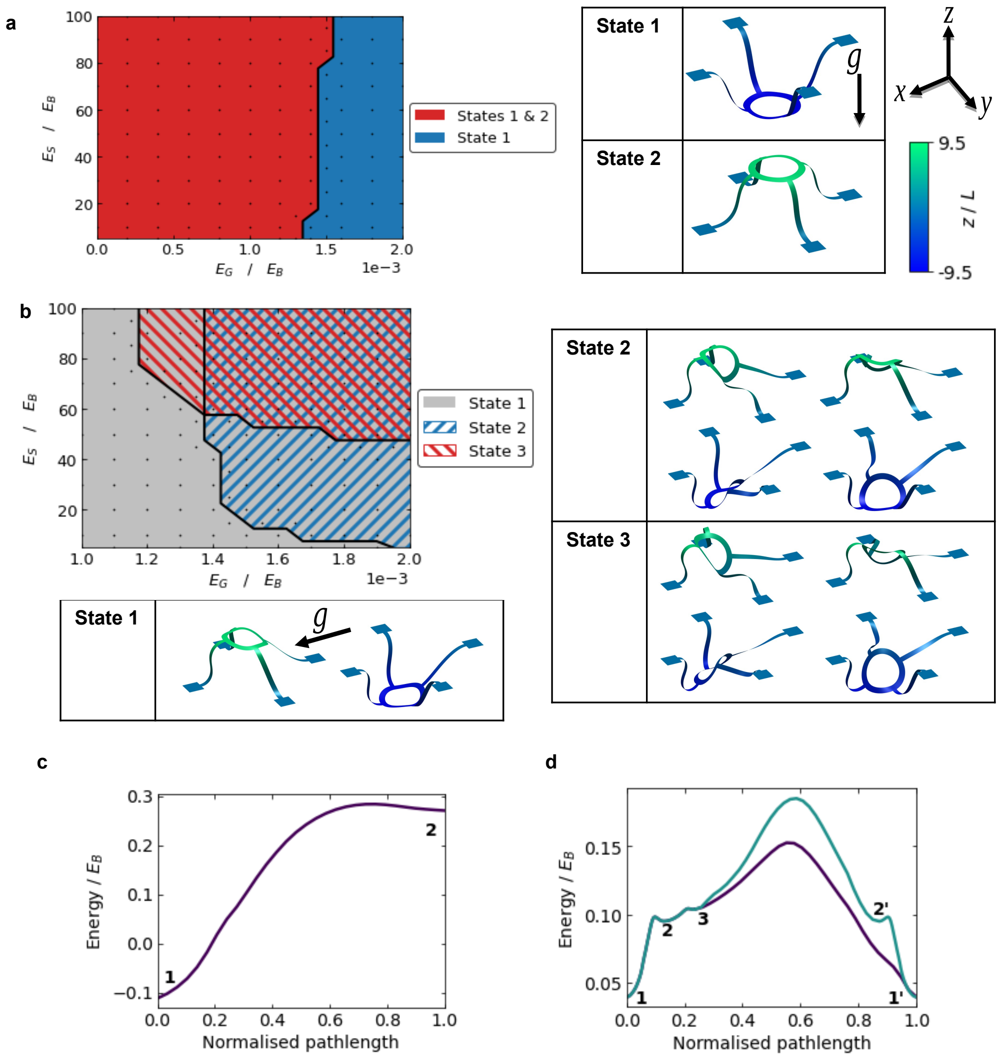


Figure 7.4: (a-b) Phase diagrams of table structures on hollow substrates, with the direction of gravity in (a) the z-direction, (b) the x-direction. In panel (b), state I has two equivalent configurations, while both states II and III have four equivalent configurations. (c) Energy profile of the transition pathway from the buckled-down (state 1) to buckled-up (state 2) configurations when gravity is in the z-direction. (d) Energy profile of the transition pathway from the buckled-down (state 1) to buckled-up (state 1') configurations when gravity is in the x-direction.

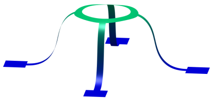
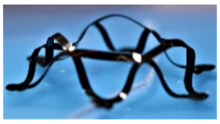
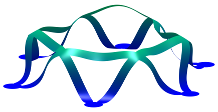
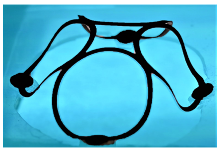
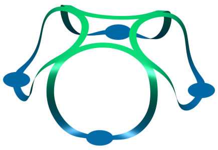
Experiment			Simulation		
State	Width (mm)	Height (mm)	State	Width (mm)	Height (mm)
	32.2	9.2		33.5	9.0
	37.6	10.7		38.0	10.9
	31.5	10.1		31.9	9.9

Table 7.1: The dimension comparison of experimental 3D structures and simulated 3D structures.

7.2.4 Hybrid ferromagnetic structures

In addition to tuning the strength and the direction of external magnetic forces for reconfiguration, an alternative and potentially more versatile means to realise magnetically controlled reconfiguration is through manipulating the distribution of magnetic materials within 3D hybrid structures. Fig. 7.5(a) schematically illustrates the fabrication process of such structures by locally integrating laser-cut ferromagnetic composite film patches (125 µm thick) onto a thin layer of inactive materials (PDMS, 180 µm thick) and compressively buckling the 2D hybrid pattern into a 3D structure.

Following this strategy, reconfigurable 3D hybrid structures that resemble tables and flowers are fabricated and magnetically tuned to display five stable configurations (Shape I-V) using a portable magnet. The deformed shapes can well maintain their configuration after the magnetic force is removed, the capability of which is the same as structures composed of a single layer of ferromagnetic composites.

In addition, 3D hybrid structures can potentially allow more local magnetic control by varying the direction and the strength of the residual magnetic field in each ferromagnetic patch through magnetization in different directions and mixing different concentrations of magnetic particles in the composite film, respectively. This could be combined with information from the energy landscape analysis in order to design the patches so the deformation in a magnetic field most accurately follows the MEP. However, this analysis and the exploration of other composition materials for reconfigurable 3D hybrid structures are beyond the scope of the current work.

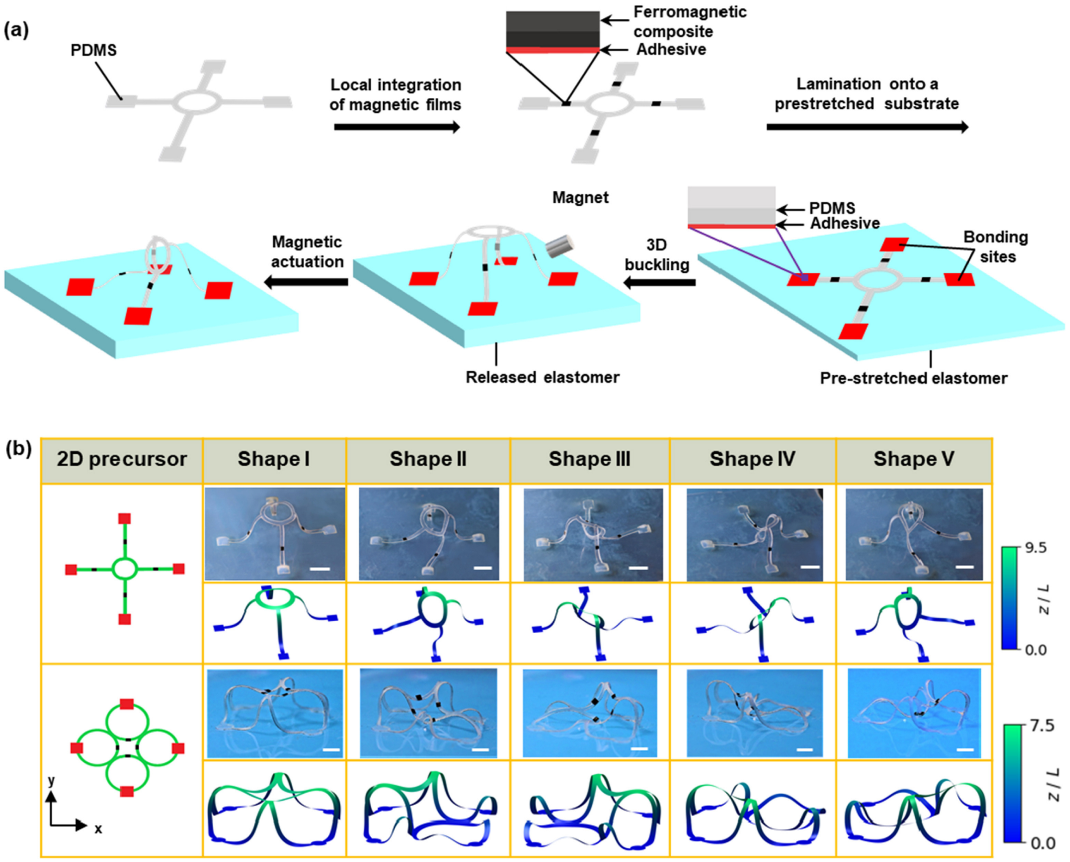


Figure 7.5: Locally tunable, heterogeneous 3D structures. (a) Schematic illustration of the fabrication process for locally controlled heterogeneous 3D ferromagnetic structures. (b) Optical images and computational modelling of multistable states for 3D structures composed of locally patterned ferromagnetic films on a thin layer of PDMS. Scale bars, 1 cm.

7.2.5 Applications of reconfigurable 3D structures

Using the techniques described above, our experimental collaborators have designed two proof-of-principle applications that demonstrate the potential offered by reconfigurable ferromagnetic mesostructures.

Firstly, functional materials and components may be integrated onto the 2D precursor, providing access to tunable 3D functional devices. This is demonstrated using a reconfigurable light emitting system in fig. 7.6(a). Here, two light emitting diodes (LEDs) are integrated onto designated locations of a double-beam structure with copper thin films connected to the LED serving as contacts for LED activation. The cross-links in the structure can be individually buckled using magnetic actuation, deforming the structure from shape I to II (or III). This enables a firm contact between the LED on the structure with the copper contacts on the substrate, connecting to an external power supply and activating the LED. The multistability of the structure then enables the LED to maintain its “on” status after the magnetic force is removed. When the two beams are simultaneously deformed to form stable configuration IV using a magnet, both LEDs are turned on. This simple example suggests a broad range of possibilities in other types of electronic and optoelectronic devices.

Based upon our investigations of the 3D table structure, a device is created for the delivery of particles in a well-controlled manner. As shown in fig. 7.6(b), a thin layer of pure PDMS film is placed in a concave shape in the centre of the table, in order to support a silica particle. This film that supports the particle is insensitive to the applied magnetic field when the table structure is magnetically actuated, thereby exerting minimal influence on the reconfiguration of the whole structure. The table structure itself is designed to be in the 2 state region of the phase diagram, fig. 7.1(c), such that it exhibits one upright state and a second state in four different orientations (modes I-IV). To trigger the delivery of the particle, a portable magnet is used to manipulate the structure into one of these different orientations, leading to well-controlled particle delivery in a chosen direction. Therefore, such flexible, remotely reconfigurable structures could have a lot of potential applications in controlled delivery.

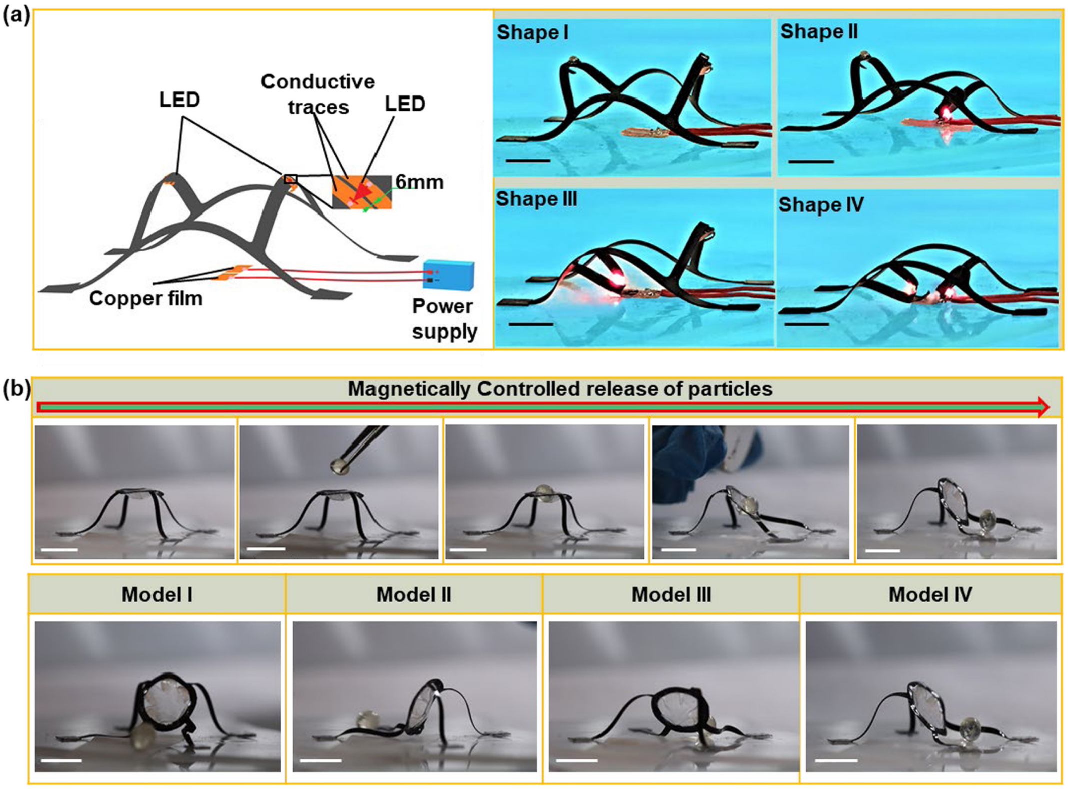


Figure 7.6: Applications of reconfigurable structures in on-demand 3D light emitting systems and controlled particle delivery. (a) Schematic illustration and optical images of a 3D light emitting system that can be reconfigured among its multistable states to switch the LEDs on and off. (b) Controlled release of particles in four different directions (Modes I-IV) enabled by four possible stable states of the table structure. Scale bars, 1 cm.

7.3 Discussion

To conclude, we have tightly integrated experiments and modelling to survey the energy landscapes of highly nonlinear 3D ferromagnetic structures for the development of reconfigurable systems. Our systematic computational analysis predicts a design phase diagram for a targeted number of stable states by varying essential material properties and geometries. It also allows us to systematically identify highly complex transition paths among the distinct stable states that minimise the energy barriers, without the need to prescribe an expected pathway for structural transformations a priori. These predictions are thoroughly validated by our experiments of creating multistable 3D ferromagnetic structures and magnetically reconfiguring the structures among their multistable states following different pathways rapidly, in a remote, on-demand manner.

Using this strategy, a plethora of symmetric and asymmetric 3D mesostructures of diverse geometries are realised and reconfigured among their multistable states, including those that resemble ribbons, tables, baskets, flowers, boxes, single and double beams. These versatile reconfigurable structures have a range of possible applications, illustrated by the integration of electronic components to create reconfigurable light-emitting functional devices, and a magnetically actuated table structure for particle delivery.

We believe our fundamental studies on the reconfigurational behaviours of 3D structures establish the foundations for a robust and reproducible modelling approach that may be applied more broadly to the design and development of diverse types of reconfigurable structures and systems. This may include origami, kirigami and stimuli-responsive structures beyond magnetic actuation, where the exploration of energy landscapes could aid in understanding the reconfiguration and the tunability of multistable states.

It is also interesting to explore the integration of functional materials with locally patterned magnetic films of different magnetization directions[173,216], and stiffness manipulation[217] to expand the design space of possible hybrid structures and to potentially enable multi-stimuli-responsive structures and soft machines. We believe these studies will provide important insights for the design of reconfigurable structures and functional systems for wide-ranging applications, such as in deployable solar panels, phononics, morphable architected materials, and soft robotics.

Topic	Symbol	Parameter description	Value
Structure	L	Arm width	1.04 / 0.8 / 0.5 mm
	—	Arm length	$12.5L$
	r_{out}	Outer ring radius	$5L$
	r_{in}	Inner ring radius	$3.75L$
	—	Mesh resolution	$0.1L$
	t	Thickness	150 µm
	ρ	Density	$2.59\,\mathrm{g\,cm^{-3}}$
	Y	Young's modulus	1.4 MPa
	ν	Poisson ratio	0.3
	—	Compressive strain	0.3
L-BFGS	m	Iteration memory	10
	—	Initialisation iterations	10 000
	—	Initialisation force density	$10\,\mathrm{N\,m^{-2}}$
	—	Gradient RMS convergence	10^{-10}
Pathways	α	BITSS energy constraint strength	10
	β	BITSS distance constraint error	0.1
	f	BITSS distance reduction factor	0.5
	—	BITSS coefficient regularity	100 iterations
	—	String reparametrisation regularity	300 iterations
	—	String number of images	20

Table 7.2: Values of the parameters that are used throughout this chapter, unless otherwise specified.

Chapter 8

Origami-inspired Mesostructures

This chapter is based upon the joint work: Y. Li, S. J. Avis, T. Zhang, H. Kusumaatmaja, and X. Wang, "Tailoring the multistability of origami-inspired, buckled magnetic structures via compression and creasing", *Mater. Horiz.*, 8, 3324 (2021).

In particular, the experimental ferromagnetic structures are the work of Y. Li, while the simulations and energy landscape analysis are my own work.

8.1 Introduction

Origami, the ancient art of folding two-dimensional (2D) thin sheets along predefined creases to create three-dimensional (3D) objects[69,218,219], has inspired the design of many engineering structures for a wide range of applications, including deployable systems[176,220,221], self-folding machines[53], reconfigurable metamaterials[74,206,222], and DNA origami[223]. For those applications, a key design feature of the structures is their ability to have multiple stable states as well as the tailoring of those states for tunability and adaptability. Existing works so far have primarily focused on bistable systems for rigid origami patterns (like the Miura folding[92] and its derivatives[10]) and deformable origami (like the twisted square pattern[100]). For example, Sadeghi and Li realised rapid and reversible folding by harnessing the asymmetric bistability of designed origami structures[224]. Liu et al. demonstrated that a folded hypar origami, obtained by folding a piece of paper along concentric squares and their diagonals to arrive at a seemingly smooth saddle shape, exhibits bistability between two symmetric configurations[193]. More recently, Melancon et al. realised pressure-deployable origami structures characterised by two stable configurations — one compact and one expanded – at the meter scale[220]. Furthermore, Fang et al. showed that the potential energy landscapes of stacked Miura-ori and the Kresling-ori structures, and therefore their stability profiles and constitutive force–displacement relations, can be effectively tuned by embedded magnets[225].

In addition to multistability, reconfiguration among the different stable states of origami-inspired structures in a well-controlled manner is demanded in many engineering devices and structures. Theoretical and experimental studies have been performed on reconfiguration paths in origami-inspired structures. For example, Zhai et al. created an origami-inspired mechanical metamaterial that can be deployed and collapsed along different configuration paths[52]. Silverberg et al. showed that hidden degrees of freedom in square twist origami structures give rise to a critical transition from mono- to bi-stability[100]. Moreover, significant advances have been made to improve the tunability and adaptability of origami-inspired structures by incorporating stimuli-responsive materials, which provides opportunities for micro robotics, biomedical devices, and many others[192,226–229].

Previous studies have laid solid foundations for the design and development of origami-inspired structures with bistability and tunability. One of the remaining challenges in this field is to actively control the number of possible stable states and their reconfiguration paths. Here we investigate this by extending our studies of mesostructures from the previous chapter to incorporate origami-inspired creases, which can increase the prevalence of multistable states. There are innumerable structures that can be constructed from networks of creased ribbons, so we instead restrict our attention to the most simple structure, a single ribbon with multiple creases, to gain fundamental insights about the effect of creases on multistability. However, the same approaches could also be applied to study other structures.

The two key parameters that we control are compressive strain and creasing in the ribbon, which are selected because they are easy to tune and very effective to alter the multistable behaviors of the structure for real applications. We utilise a similar computational approach as the last chapter in order to construct phase diagrams showing how the available states sensitively depend on the compressive strain and the number of creases. Then we generate the the reconfiguration pathways to understand the stability of the various states and how they can interchange upon applying an external magnetic force or varying the applied strain.

Finally, guided by this analysis, our collaborators have produced diverse complex origami-inspired structures, including structure arrays that be individually tuned based on the multistability of the structural unit and a biomimetic insect.

8.2 Results and Discussion

8.2.1 Concept of origami-inspired, multistable magnetic structures

Fig. 8.1 schematically illustrates the experimental assembly process and reconfiguration of origami-inspired structures compressively buckled from ferromagnetic composites[73,195,196]. These are constructed using the approach described in chapter 7, however laser patterning is also used in order to create crease sections with a thickness that is 46.15% of the non-crease regions. Due to this reduced thickness, the crease regions act as flexible hinges that can bend into different states, affecting the potential multistable reconfigurations[185].

Under an external magnetic field generated by a manually manipulative magnet,

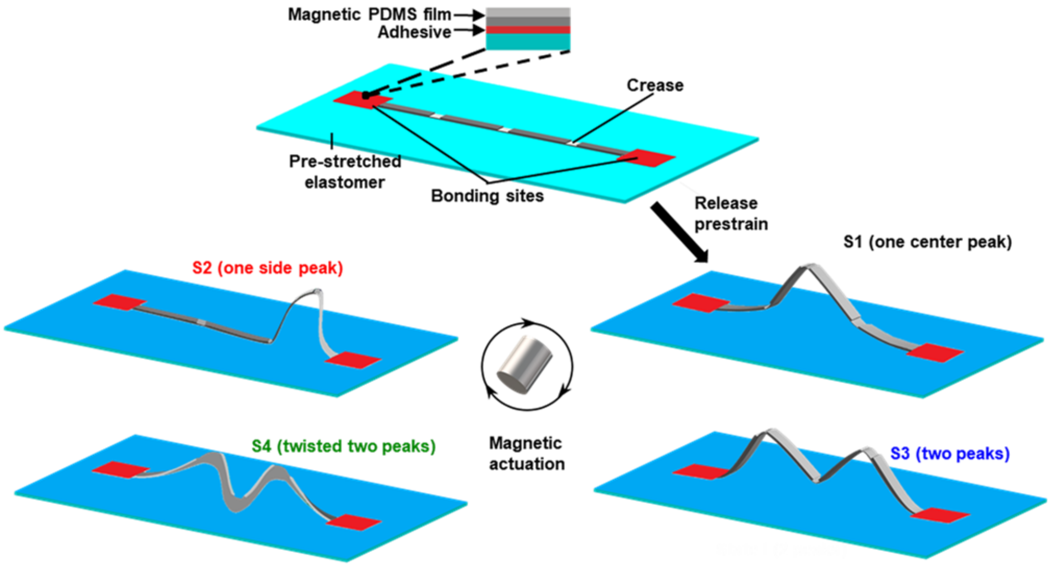


Figure 8.1: Schematic illustration of the assembly and magnetic reconfiguration schemes of multistable origami-inspired ribbon structures.

the assembled origami-inspired structure can be reconfigured into up to four distinct stable states: state 1 (S1, one center peak), state 2 (S2, one side peak), state 3 (S3, two peaks), and state 4 (S4, twisted two peaks). These are then able to remain stable after the removal of the applied magnetic field.

8.2.2 Phase diagram of stable states for origami-inspired structures

We first study the effect of two essential parameters, the number of creases and compressive strains, on the number and configurations of stable states by using a ribbon structure as an example. A phase diagram, indicating the stable states of the structure with up to seven creases and at strain levels of 0%-40%, is shown in fig. 8.2(a) and serves as a theoretical foundation for the multistability design. The phase diagram is constructed computationally by identifying the local minima in the energy landscape of the structure under various conditions of strain levels and crease numbers.

As in chapter 7, we have used the discrete shell model, this time with two thicknesses: 130 µm and 60 µm, combined with a random search algorithm for energy minimisation. At selected points (typically 5 different strain levels for each crease number) across the phase diagram, the structure is randomly initialised several hundreds of times and relaxed to the minima. Once the qualitatively distinct configurations are identified, we systematically vary the compressive strain to observe the full range over which each state is stable. It is worth noting that while it is difficult to guarantee that all possible stable states have been obtained, we are unable to find additional states when more points in the phase diagram are sampled. Hence, we consider the identified stable states to be representative in this study.

As shown in fig. 8.2(a), the number of creases (n_{crease}) has notable effects on the number and configurations of stable states. To better illustrate this point, we demonstrate the available states of structures as n_{crease} is increased from 0 to 7 under a constant compressive strain of 15% in fig. 8.2(b), with the computational and experimental results shown side by side. When the n_{crease} is zero, i.e., the continuous ribbon structure, only S1 (one center peak) exists at all strain levels (e.g., point i). An additional asymmetric state S2 becomes available when n_{crease} rises to two at a compressive strain of 15% (point ii). A further increase to three creases causes the structure to admit state S3 (two peaks) at a strain of 15% (point iii) and S4 (twisted two peaks) at a higher strain level. As n_{crease} increases above three, the number of stable states begins to decline, with S2 becoming unstable for structures with four creases (point iv). For structures with five creases and above, only S1 remains stable at a compressive strain of 15%, although S3 is still stable under lower strains for $n_{\text{crease}} = 5$ and 6. This can be expected because we are approaching the continuous case when the number of creases is large.

The available stable states also depend strongly on the magnitude of the applied compressive strain used in the assembly process of the structure. We illustrate this using a structure with three creases due to its rich phenomena.

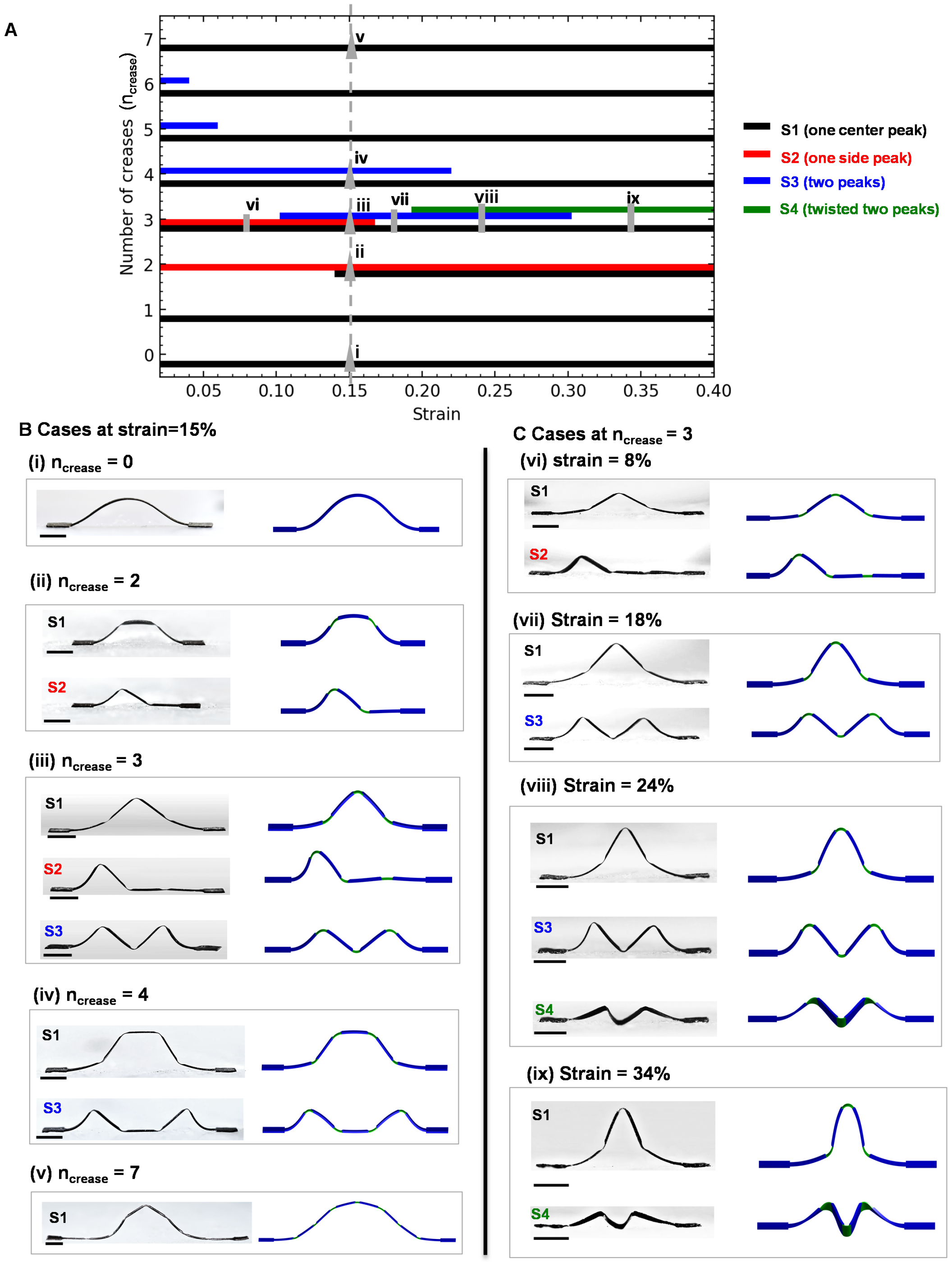


Figure 8.2: Multistability of origami-inspired ferromagnetic ribbon structures under varying creases and compressive strains. (A) Design phase diagram showing the effect of the crease number and compressive strain on the number and configurations of distinct stable states of the ribbon. (B) Experimental and computational results of multistable states under a constant strain of 15% and varying crease numbers, corresponding to points i-v labelled in fig. 8.2(a). Scale bars, 2 mm. and (C) Experimental and computational results of multistable states at a constant number of creases ($n_{\text{crease}} = 3$) and varying strain levels, corresponding to points vi-ix in fig. 8.2(a). Scale bars, 2 mm.

Based on the number and configurations of stable states, five regimes are identified: a) two distinct stable states (S1 and S2; 0-10% strain), b) three distinct stable states (S1, S2, and S3; 10-17% strain), c) two distinct stable states (S1 and S3; 17-19% strain), d) three distinct stable states (S1, S3, and S4; 19-30% strain), and e) two distinct stable states (S1 and S4; 30-40% strain). Fig. 8.2(c) shows experimental and computational results of distinct stable states at representative strain levels for the five regimes. Under a relatively low compressive strain of 8% (point vi), states S1 (one center peak) and S2 (one side peak) exist, which can be reversibly reconfigured into each other via manually controlling the magnetic force and direction. Increasing the strain to 15% (point iii) leads to the appearance of a third state, S3 with two peaks, which increases the number of stable states to three. However, when the strain is further increased to 18% (point vii), S2 with one side peak disappears. Further increasing the strain to 24% (point viii) causes the formation of an interesting twisted state (S4) and therefore increases the number of stable states to three (S1, S3 and S4). As the strain becomes even larger, S3 disappears and only S1 and S4 exist at the strain of 34% (point ix).

The experimental and computational results shown above agree reasonably well. The discrepancy in the configuration of some stable states (e.g., (v) S1 in fig. 8.2(b)) between experiments and modeling likely results from the adhesion and friction between the structure and the substrate in experiments, which is neglected in computational modeling. The effect of friction will be further discussed in section 8.2.3.

Furthermore, by continuously changing the strain levels from 0% to 40% using the 3-crease structure, we record the dynamic progression of the five regimes and the fast, remote magnetic reconfiguration among the stable states within each regime in video A.6, which is highly consistent with the predictions of the phase diagram in fig. 8.2(a). We also show that the states are stable under perturbations perpendicular to the vertical surface (along the height direction) of the ribbon, except in cases where S4 (twisted two peaks) is present.

The studies above provide important guidelines for tailoring the number and configurations of stable states in origami-inspired structures by tuning the number of creases and the assembly strain. For example, for applications like digital logics where multiple stable states are desired, structures with three creases may offer more design space and tunability, while structures with a lower or higher number of creases is preferred for applications where a single stable state is needed.

Furthermore, from the phase diagram, we observe that structures with two and three creases have multistable states under a wide range of strain (15%-40% strain for the case of two creases and 0%-40% strain for the case of three creases), which can be used for applications of multistable structures that demand a wide working strain range.

Effect of crease thickness and length

As a brief aside, we also study the effects of modifying the crease thickness ratio and length ratio (defined with respect to the length of each flat segment) on the multistability. For this, the 3-crease ribbon under a strain level of 14% is used as an example.

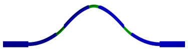
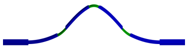
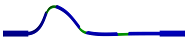
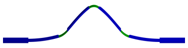
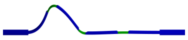
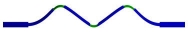
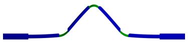
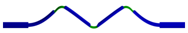
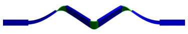
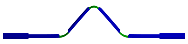
Thickness ratio	State 1	State 2	State 3	State 4
0.67		✗	✗	✗
0.56			✗	✗
0.46				✗
0.33		✗		
0.25		✗		✗

Table 8.1: Effect of the crease thickness ratio on the multistability of the ribbon structure (strain: 14%; length ratio: 0.32)

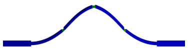
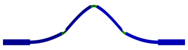
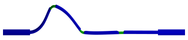
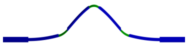
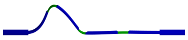
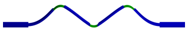
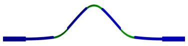
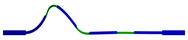
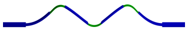
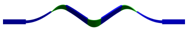
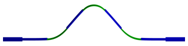
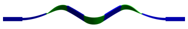
Length ratio	State 1	State 2	State 3	State 4
0.10		✗	✗	✗
0.18			✗	✗
0.32				✗
0.56				
1.00		✗		

Table 8.2: Effect of the crease length ratio on the multistability of the ribbon structure (strain: 14%; thickness ratio: 0.46)

The results in tables 8.1 and 8.2 show that both ratios can have a significant effect on the multistability of the ribbon structure. In particular, reducing the crease thickness and increasing its length has an overall similar effect on the stable states as increasing the compressive strain. The main difference is at the extreme when state 3 remains stable, rather than progressing to the twisted state. This is because reducing the thickness of the crease, or increasing its length, grants the creases more flexibility, which reduces the amount of bending that is required in the edge segments. However, it is likely that at larger compressive strains state 3 will become unstable.

8.2.3 Reconfiguration pathways for a 3-crease ribbon structure

In addition to locating the stable configurations, we also investigate available pathways that the structure can be reconfigured among its stable states, which is important for many applications. The pathways have been computed using the same method as chapter 7, first by finding the TS with BITSS, for which the separation of the two states is reduced by 30% each iteration, and then computing the pathways using a combination of downhill minimisation and the string method. In addition, once a transition state is found for a given strain, a continuation approach is then used with the BITSS method in order to identify the energy barriers across the whole

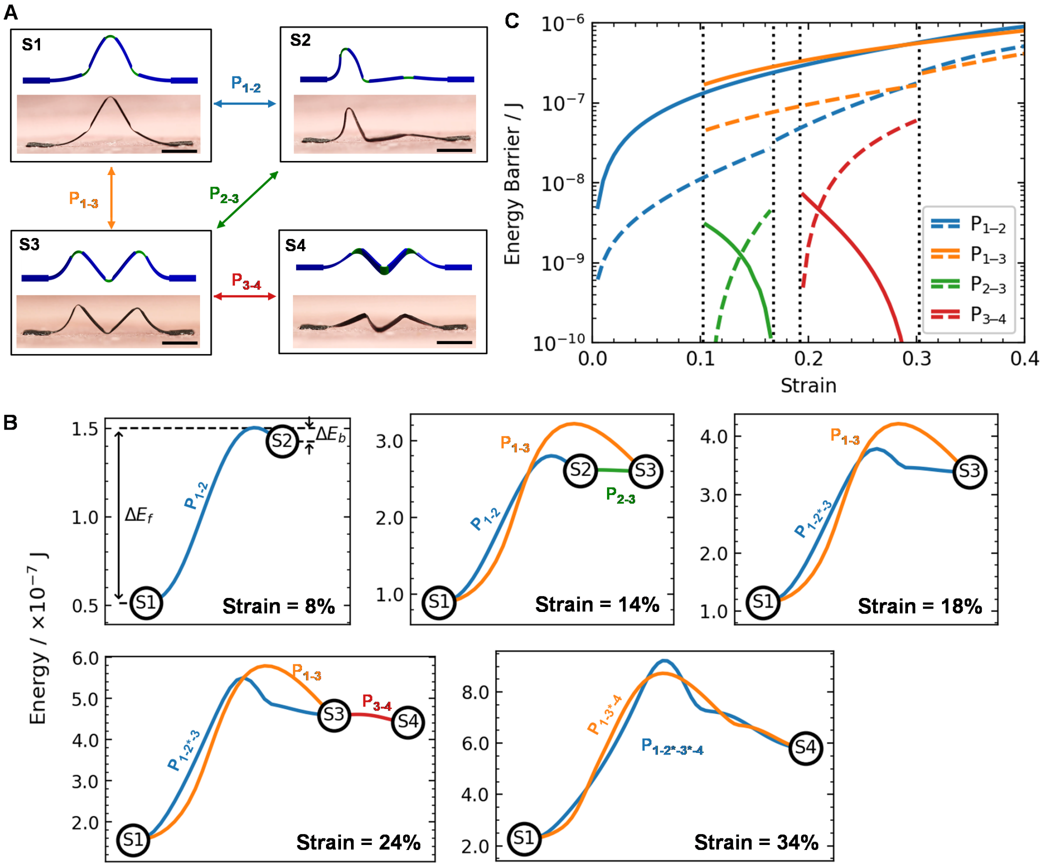


Figure 8.3: (A) Reconfiguration paths among the distinct stable states of origami-inspired ferromagnetic ribbon structures. Scale bars, 2 mm. (B) Energy profiles of pathways for the 3-crease ribbon structure at various strain levels. (C) Minimum energy barriers for pathways in fig. 8.2(b) as a function of the strain. The solid lines indicate the energy barriers from the lower numbered states to the higher numbered states and the dashed lines are the barriers from the higher numbered states to the lowered numbered states (illustrated by ΔE_f and ΔE_b in fig. 8.3(b), respectively) . The line colors correspond those in fig. 8.3(a). The blue and orange lines also encompass the extended pathways P$_{1\text{-}2^*\text{-}3}$, P$_{1\text{-}2^*\text{-}3^*\text{-}4}$, and P$_{1\text{-}3^*\text{-}4}$.

range of strains that the pathway exists.

Here we use the three-crease structure as a representative case for study because of its complexity in the configurations and the number of stable states, which results in a large number of reconfiguration paths. As illustrated in fig. 8.3(a), we identify four distinct pathways among the stable states of the structure: P$_{1\text{-}2}$ for the transition path between S1 and S2, P$_{1\text{-}3}$ between S1 and S3, P$_{2\text{-}3}$ between S2 and S3, and P$_{3\text{-}4}$ between S3 and S4. It is worth noting that in cases where one or more of those states are not stable, the corresponding individual paths may merge into one, as we will detail below.

Fig. 8.3(b) shows the available pathways and the associated energy profiles for a representative case in each of the five regimes for a 3-crease ribbon (identified in fig. 8.2(a)) as the strain is increased from 0% to 40%. In the first regime (0-10% strain), only states S1 and S2 exist, with a single pathway ($P_{1\text{-}2}$) to reconfigure between them, as shown in video A.7. In the second regime (10%-17% strain), state S3 appears, so two additional pathways, $P_{1\text{-}3}$ and $P_{2\text{-}3}$, become accessible (video A.8).

Our analysis also highlights that there are not only multiple stable states existing in the structure, but also multiple available paths for transforming the structure from one state to another. For example, we can observe that there are two main pathways that can be taken from state S1 to S3. The first path, initially following $P_{1\text{-}2}$ to reconfigure S1 to S2, subsequently involves the creation of an asymmetric peak at one side of the structure to form state S3 from S2 ($P_{2\text{-}3}$). Due to symmetry, there are two equivalent scenarios depending on which side the peak is created to form S3. The second path, $P_{1\text{-}3}$, directly reconfigures S1 to S3 by forming two edge peaks simultaneously while maintaining symmetry.

As the strain increases to the regime of 17%-19%, state S2 becomes unstable, so $P_{1\text{-}2}$ and $P_{2\text{-}3}$ merge into a single path $P_{1\text{-}2^{*}\text{-}3}$ (video A.9), where 2* denotes that state 2 is no longer stable. Above the strain of 19%, pathway $P_{3\text{-}4}$ appears between state S3 and twisted state S4, while the two pathways from S1 to S3 still exist (video A.10). Finally, at the strain of 30% and above, S3 is no longer stable leaving just two merged paths $P_{1\text{-}2^{*}\text{-}3^{*}\text{-}4}$ and $P_{1\text{-}3^{*}\text{-}4}$ between S1 and S4 (video A.11).

We further investigate the energy barriers of each pathway as a function of strain. For each pathway, there are two relevant energy barriers, as illustrated in fig. 8.3(b) using $P_{1\text{-}2}$ in the case of 8% strain as an example. We label ΔE_f for the energy required to transition from a lowered numbered state to a higher numbered state in each pathway (here, S1 to S2), while ΔE_b is the energy required for the opposite transition (here, S2 to S1). We summarise the values of the energy barriers for all the available pathways in fig. 8.3(c), where ΔE_f and ΔE_b for each pathway are represented by solid and dotted lines, respectively.

We can see that the energy barriers for pathways reconfiguring from S1 to other states including S2, S3 and S4 depending on the strain levels ($\mathrm{P}_{1\text{-}2}$, $\mathrm{P}_{1\text{-}2^*\text{-}3}$, $\mathrm{P}_{1\text{-}2^*\text{-}3^*\text{-}4}$, $\mathrm{P}_{1\text{-}3}$, $\mathrm{P}_{1\text{-}3^*\text{-}4}$; solid blue and orange lines) are much larger than their corresponding reverse barriers (dotted blue and orange lines, respectively). Such difference suggests that state S1 is more stable than the other states (S2, S3 and S4), which is consistent with the phase diagram in fig. 8.2(a), where S1 is almost always stable for cases under all crease numbers.

It is also seen that the energy barriers of paths $\mathrm{P}_{2\text{-}3}$ and $\mathrm{P}_{3\text{-}4}$ are significantly lower than those of pathways involving state S1. This suggests that reconfigurations among states S2, S3, and S4 are much easier than those between S1 and S2/S3/S4. Correspondingly, the required forces to disturb the reconfiguration paths or trap locally stable states are low. Hence, forces like friction between the structure and the substrate (assembly platform) in experiments may be sufficient to trap the structure along these paths or at states S2, S3 and S4. This may explain why experimentally the structure appears to get trapped in some states which are shown to be unstable in computational results (videos A.9 to A.11). For example, in video A.9, when the structure is reconfigured from S2 to S3, the 2nd edge peak is not fully formed due to possible friction between the structure and the substrate, causing the slight asymmetric configuration of S3.

From fig. 8.3(c), we can also see that the energy barriers for the majority of the reconfiguration paths increase monotonically with the strain, implying that it becomes more difficult to reconfigure the states at larger strains. Exceptions occur for reconfiguration paths $\mathrm{P}_{2\text{-}3}$ (solid green line, transition from S2 to S3) and $\mathrm{P}_{3\text{-}4}$ (solid red line, transition from S3 to S4), the energy barriers of which decrease with increasing strain levels. Here, the cases when the energy barriers tend to be zero correspond to the instances where S2 and S3 become unstable, i.e., the instability modes of these states. Overall, the energy barrier analysis reinforces the previous observation that, as the strain is varied from 0% to 40%, state S1 always exists while the availability of states S2, S3 and S4 strongly depends on the strain imposed.

8.2.4 Multi-ribbon structures

The multistable ribbon configurations studied above can serve as building blocks for complicated geometries and diverse types of origami-inspired structures. Several such structures are detailed in this section that have been created by our collaborators.

Fig. 8.4(a) show a 3 × 3 array of ribbons with three creases at different strain levels, which is consistent with the prediction of the phase diagram in fig. 8.2(a). Particularly, as predicted by the phase diagram, three distinct stable states exist at the strain levels of 14% (S1, S2, and S3) and 28% (S1, S3, and S4), respectively. In experiments, each individual ribbon unit is separately addressable and therefore the structure array can be magnetically tuned in a sequential manner to simultaneously display all the three stable states (S1 in row 1, S2 in row 2, and S3 in row 3 at a strain of 14%; S1 in column 1, S3 in column 2, and S4 in column 3 at a strain of 28%). Such capability of dynamically and reversibly tuning the patterns in a structure array represents an important option for applications including digital coding and smart switches[222], especially when integrated with other functional materials or elements.

In addition, the complexity of origami-inspired structures and the number of stable states can be increased by incorporating a larger number of interconnected structural units, like ribbons with two creases. Fig. 8.4(b) presents a structure consisting of two creased table structures connected with a creased ribbon, which can be magnetically reconfigured into six distinct stable states that are enabled by the multistability of the ribbon units. Furthermore, the multistable origami-inspired structures can be extended to diverse geometries. For example, fig. 8.4(c) demonstrates a structure that

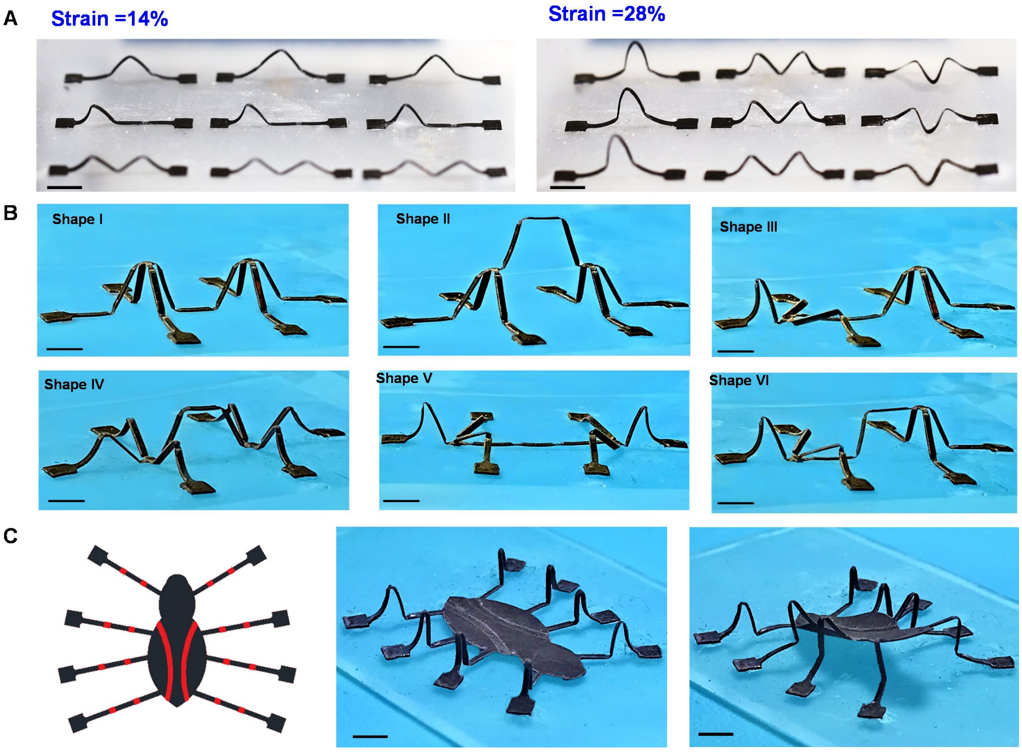


Figure 8.4: Diverse origami-inspired, multistable ferromagnetic structures that can be actuated by magnetic forces. (A) An array of 3-crease ribbon structures showing different patterns formed by the multistable states of the ribbon under the strain of 14% and 28%, respectively. Scale bars, 3 mm. (B) Multistable states and magnetic reconfiguration of a double-table structure composed of creased ribbon segments. (C) A multistable biomimetic "insect" in its flat and standing states, respectively. Scale bars, 2 mm.

mimics insect flexion, with various standing states that show biomimetic movements. Here the legs can be treated as two-crease ribbons with each switching between their two asymmetric cases (S2 in fig. 8.2). When the peak is beside the binding site, it corresponds to the insect lying flat. But if the peak is beside the body then it causes the insect to stand upright.

8.3 Conclusion

To sum up, it has been shown that the multistability and the associated transition paths of origami-inspired, compressively buckled ferromagnetic structures can be tailored by controlling the number of creases and assembly strain. Our constructed phase diagram of a representative creased ribbon structure from the energy landscape analysis provides important guidelines for the targeted number of stable states by varying the two control factors. In addition, transition pathways among the distinct stable states are computed and illustrate how the structure can be manipulated to be reconfigured along different pathways. These have been validated using experimental results of origami-inspired ferromagnetic structures, which show targeted multistable states following designed pathways, which highly agree with modelling.

The fundamental understanding of the multistability of creased ribbon structures provides important guidelines for the design and application of multistable origami-inspired systems, with demonstrated examples including an arrays of creased ribbon units and the various states of a biomimetic "insect" structure. Also, the energy landscape analysis offers an approach to study the multistable states and pathways, and can be tailored for specific applications.

These results highlight potential opportunities in the future to exploit multistable, origami-inspired ferromagnetic structures for intelligent and adaptive systems such as programmable logic array by integrating complex structure design and functional materials like stimuli-responsive polymers and electronics. It will also be interesting to explore the concept of actively manipulating the number of stable states and their reconfiguration paths in the design of other types of functional structures by using alternative control parameters.

Topic	Symbol	Parameter description	Value
Structure	t	Thickness	130 µm
	t_{crease}	Crease thickness	60 µm
	l	Segment length	2.5 mm
	l_{crease}	Crease length	0.8 mm
	w	Ribbon width	0.8 mm
	ρ	Density	$2.59\,\mathrm{g\,cm^{-3}}$
	Y	Young's modulus	1.4 MPa
	ν	Poisson ratio	0.3
L-BFGS	m	Iteration memory	5
	—	Initialisation iterations	30 000
	—	Initialisation force density	$10\,\mathrm{N\,m^{-2}}$
	—	Gradient RMS convergence	10^{-11}
Pathways	α	BITSS energy constraint strength	10
	β	BITSS distance constraint error	0.1
	f	BITSS distance reduction factor	0.2
	—	BITSS coefficient regularity	100 iterations
	—	String reparametrisation regularity	5000 iterations
	—	String number of images	60

Table 8.3: Values of the parameters that are used throughout this chapter, unless otherwise specified.

Chapter 9

Conclusions and Outlook

In this thesis, we set out to apply energy landscape techniques to study complex macroscale systems, in particular multistable elastic structures. However, we found that traditional techniques, such as chain-of-states pathway methods, are often inadequate for these applications. As a result, we have developed new techniques to study the energy landscapes of these large-scale systems.

In chapter 5, we introduced a new transition state search method, BITSS. Unlike traditional methods, BITSS does not require any initial estimate for the pathway, making it effective for locating complex pathways without any a priori knowledge of their route. We have shown that BITSS is compatible with adaptive remeshing techniques, which can offer simultaneous improvements to both efficiency and accuracy. Furthermore, BITSS has been successfully applied to various challenging energy landscapes, such as where it is locally flat, or even discontinuous. Therefore, this opens up a range of previously inaccessible problems to energy landscape analysis.

This and several other methods have then been incorporated into an energy landscape software library, 'ELLib', as detailed in chapter 6. The library is designed to be modular, flexible to use, and may be easily extended to incorporate new methods and potentials. Additionally, it is designed to utilise parallel methods for calculating potentials, minimisation, and exploring energy landscape for large system sizes.

Having developed this new methodology, we then applied it to investigate the behaviour of multistable buckled mesostructures. We demonstrated in chapter 7 that our approach can accurately locate stable states and reconfiguration pathways for a range of different structures, with results that match observations from experimentally manipulating the structures. Moreover, by investigating a table structure we were able to create phase diagrams that highlight how the various states and pathways depend upon the chosen system parameters. These findings can therefore be used to design a structure with specific functionality for a given application.

In chapter 8 we extended our investigations to explore the effect of origami-inspired creases on the multistability of simple ribbons, which can be used as a building block for many larger structures. Our results show that the number of creases plays a critical role in the number of available states, with three creases showing the greatest multistability, then quickly reducing to just a single stable state as the number of creases is increased further. Furthermore, we found that increasing the strain led to changes to the available states and pathways, with a group of states that could be easily interchanged while another was separated by a large barrier.

Taken together, these results provide new insight into how the multistability of buckled mesostructures is affected by various structural and material parameters. Our methodology and findings outline a process by which these, and other more general structures, may be tailored to obtain chosen states and reconfiguration pathways. The identified reconfiguration pathways can also be serve as targets for remote actuation, providing a critical step towards developing rapidly reconfigurable devices for applications in soft robotics and deployable structures, among other fields. This work has thus contributed to advancing the field of multistable elastic structures and has potential implications for a wide range of technological applications.

9.1 Future work

9.1.1 Locating multiple pathways

Many applications have multiple competing pathways between a pair of states. For example, the snap-through buckling of a clamped beam, fig. 9.1(a), can take one of two asymmetric pathways. Therefore, we may wish to use a method that enables us to locate these additional pathways.

However, locating multiple pathways between a single pair of minima is a difficult problem that has not been fully addressed by existing methods. One possibility is to use single-ended TS search methods from the regions around each of the minima, however this is not guaranteed to find all of the TS and those that are found may not be connected to the states of interest. Another approach is to make initial guesses using a chain-of-states method, but this requires some knowledge of the expected pathways and the initial interpolation may be difficult. Alternatives are possible, such as transition path sampling[230] or using molecular dynamics over long time-scales, but these can be very computationally costly.

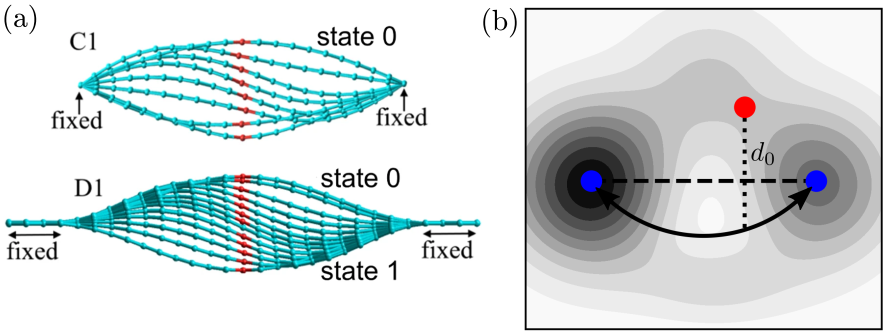


Figure 9.1: (a) The asymmetric snap-through buckling pathways for constrained thin sheets (specifically graphene membranes). The endpoints are allowed to rotate in the first, and they are fixed in the second. An equivalent, reflected pathway is possible in both. Adapted with permission from[231]. (b) Schematic illustrating a 2D landscape with two pathways, and a possible modified distance metric that would enable BITSS to locate the second TS. The pathway's radius of curvature is calculated to set a certain distance d_0 from the known TS (red dot).

One possibility would be to extend the BITSS method to locate multiple TS by modifying the distance metric. Taking inspiration from general relativity, in which mass can curve space-time and reduce the distance between two points, once BITSS has located a TS a "negative mass" could be applied, effectively expanding space around that point. This would result in the two states attempting to move away from the known TS as they converge, and hopefully converging to another. This could then be repeated to find multiple different transition states that separate the two minima. Such a method would bear a resemblance to the deflated Newton method[232], but for finding multiple TS instead of minima.

Of course attempting to calculate an accurate distance based upon a curved space-time would be exceedingly complicated, however it does provide a useful analogy. Instead, a simpler approach could be to consider a circular arc between the states that has a radius of curvature that points it away from the TS so that the images would converge a chosen distance away, as illustrated in fig. 9.1(b). By repeating BITSS and continually increasing this distance bias, it could be possible to locate all available pathways.

9.1.2 3D elastic metamaterials

In this thesis, we have investigated the behavior of two-dimensional elastic structures through the use of a discretised shell model that is particularly effective in analyzing the bending of thin films. However, this approach is insufficient for studying a variety of complex reconfigurable structures such as three-dimensional mechanical metamaterials, as discussed in section 3.3. As a result, a natural extension to our work would be to incorporate a three-dimensional elastic potential, such as one based on the finite element method, to better examine these systems.

Energy landscape analysis could be a promising tool for addressing a wide range of problems associated with mechanical metamaterials. For example, the effectiveness of collisional energy-absorbing structures[58,84] is heavily influenced by material and

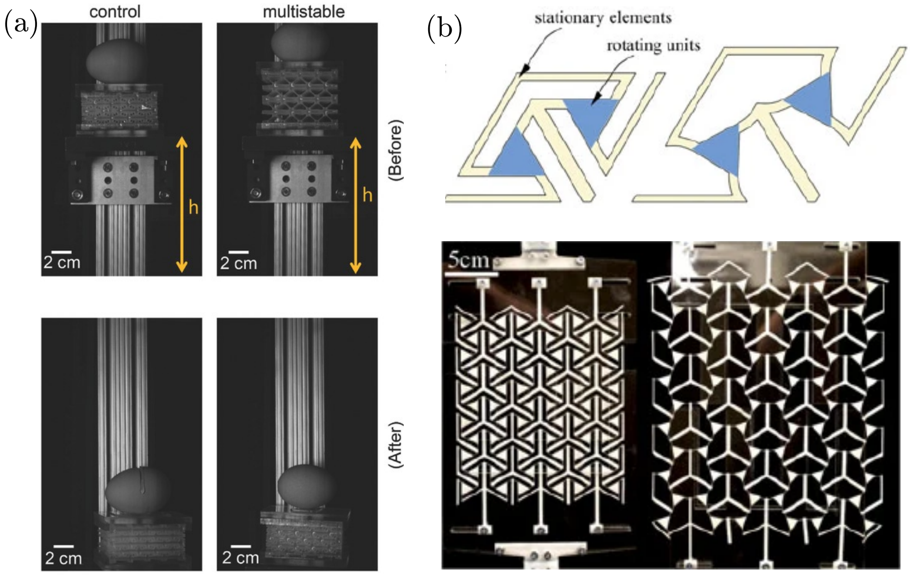


Figure 9.2: (a) Energy absorbing structures. Adapted with permission from[58]. (b) Bistable auxetic metamaterial that undergoes a transition to a deployed state. Adapted with permission from[215].

geometric factors, which modify the energy barriers of the various reconfiguration pathways. Therefore, by optimising the barriers as we demonstrated in section 6.6, we could gain valuable insights into how to maximize the efficiency of these structures.

A similar analysis could be applied to flexible metamaterials that undergo a large-scale, bistable transformation, which are of interest to develop deployable structures[11,215]. Here, an important consideration is their durability, i.e. how many times they can be reconfigured without breaking, which is found to be inversely correlated with the elastic strain energy during the transition[215]. In this case, investigating the minimum energy pathway could provide the structural parameters that minimise the energy barrier, thereby increasing durability, while allowing the structure to remain bistable.

9.1.3 Reconfigurable surface structures for adaptable wetting

An interesting application of the creased ribbon structures investigated in chapter 8 could be to create an array of adaptable surface structures that could be remotely

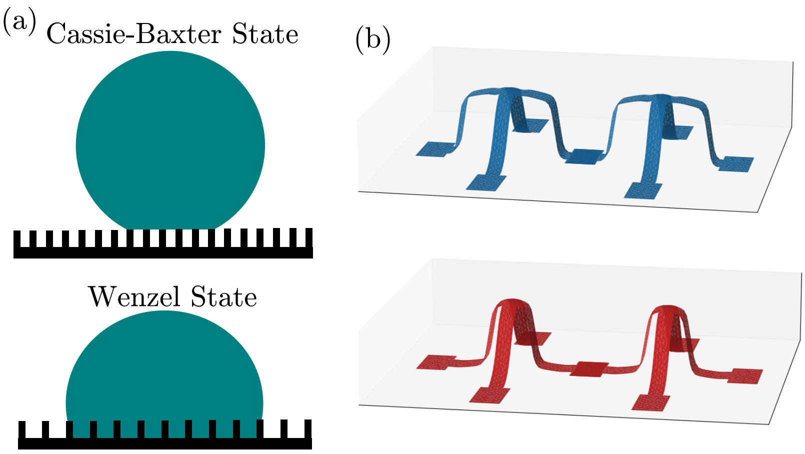


Figure 9.3: (a) Illustration of the Cassie–Baxter and Wenzel states of a droplet on a structured surface. (b) An origami-inspired mesostructure with two states that could be used to realise an adaptable wetting surface.

switched between two states in order to alter the wetting properties of a surface.

For example, it is known that an array of pillars on a surface can keep a droplet in a suspended Cassie–Baxter state where it can exhibit superhydrophobic behaviour. But this depends upon the geometry of the posts, and if they are further apart the droplet may instead transition to a Wenzel state, sinking into the structure and losing the superhydrophobic behaviour. These two states could potentially be switched from one to the other using the structure in fig. 9.3(b), which has one state where the ‘pillars’ are larger with less spacing between then, and another state with smaller pillars and larger spacing.

However, due to their more complicated design, flexible ribbon structures would restricted to larger sizes than a simple pillar geometry, making the Wenzel state more prone to failure. Energy landscape methods could be used to address this using a similar analysis as performed in[50]. This would provide information as to how the structure should be modified to maximise the energy barrier that prevents the droplet from penetrating the structure.

Appendix A

Supplementary Material

Supplementary material for this thesis can be found at this link: `https://doi.org/10.15128/r1n870zq87j`

This contains the ELLib software library, the script used to perform the structural optimisation in section 6.6, and supplementary videos for chapters 7 and 8, which are described below.

Video A.1: Reconfiguration path 1 for a table structure with $L = 1.04\,\text{mm}$.

Video A.2: Reconfiguration path 2 for a table structure with $L = 1.04\,\text{mm}$.

Video A.3: Reconfiguration path 1 for a table structure with $L = 0.8\,\text{mm}$.

Video A.4: Reconfiguration path 2* for a table structure with $L = 0.8\,\text{mm}$.

Video A.5: Reconfiguration path 1 for a table structure with $L = 0.72\,\text{mm}$.

Video A.6: Progression of the five regimes of the 3-crease ribbon when the strain is increased from 0% to 40%, with a magnetic reconfiguration among the stable states in each regime.

Video A.7: The single pathway ($\text{P}_{\text{1-2}}$) for a 3-crease ribbon at 8% strain.

Video A.8: The three pathways ($\text{P}_{\text{1-2}}$, $\text{P}_{\text{1-3}}$, and $\text{P}_{\text{2-3}}$) for a 3-crease ribbon at 14% strain.

Video A.9: The two pathways ($\text{P}_{\text{1-2*-3}}$ and $\text{P}_{\text{1-3}}$) for a 3-crease ribbon at 18% strain.

Video A.10: The three pathways ($\text{P}_{\text{1-2*-3}}$, $\text{P}_{\text{1-3}}$, and $\text{P}_{\text{3-4}}$) for a 3-crease ribbon at 24% strain.

Video A.11: The two pathways ($\text{P}_{\text{1-2*-4}}$ and $\text{P}_{\text{1-3*-4}}$) for a 3-crease ribbon at 34% strain.

Bibliography

[1] Joseph D. Bryngelson, José Nelson Onuchic, Nicholas D. Socci, and Peter G. Wolynes. Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins: Structure, Function, and Bioinformatics*, 21(3):167–195, 1995.

[2] S. Jordan Kerns, Roman V. Agafonov, Young Jin Cho, Francesco Pontiggia, Renee Otten, Dimitar V. Pachov, Steffen Kutter, Lien A. Phung, Padraig N. Murphy, Vu Thai, Tom Alber, Michael F. Hagan, and Dorothee Kern. The energy landscape of adenylate kinase during catalysis. *Nature Structural and Molecular Biology*, 22(2):124–131, 2015.

[3] Zamaan Raza, Björn Alling, and Igor A. Abrikosov. Computer simulations of glasses: The potential energy landscape. *Journal of Physics: Condensed Matter*, 27(29):293201, July 2015.

[4] Katia Bertoldi. Harnessing Instabilities to Design Tunable Architected Cellular Materials. *Annual Review of Materials Research*, 47:51, 2017.

[5] Edwin A. Peraza-Hernandez, Darren J. Hartl, Richard J. Malak Jr, and Dimitris C. Lagoudas. Origami-inspired active structures: A synthesis and review. *Smart Materials and Structures*, 23(9):094001, August 2014.

[6] Haoran Fu, Kewang Nan, Wubin Bai, Wen Huang, Ke Bai, Luyao Lu, Chaoqun Zhou, Yunpeng Liu, Fei Liu, Juntong Wang, Mengdi Han, Zheng Yan, Haiwen Luan, Yijie Zhang, Yutong Zhang, Jianing Zhao, Xu Cheng, Moyang Li, Jung Woo Lee, Yuan Liu, Daining Fang, Xiuling Li, Yonggang Huang, Yihui Zhang, and John A. Rogers. Morphable 3D mesostructures and microelectronic devices by multistable buckling mechanics. *Nature Materials*, 17(3):268–276, 2018.

[7] Benjamin Gorissen, David Melancon, Nikolaos Vasios, Mehdi Torbati, and Katia Bertoldi. Inflatable soft jumper inspired by shell snapping. *Science Robotics*, 5(42):eabb1967, May 2020.

[8] Todd M. Currier, Samuel Lheron, and Yahya Modarres-Sadeghi. A bio-inspired robotic fish utilizes the snap-through buckling of its spine to generate accelerations of more than 20g. *Bioinspiration & Biomimetics*, 15(5):055006, August 2020.

[9] E Hawkes, B An, N M Benbernou, H Tanaka, S Kim, E D Demaine, D Rus, and R J Wood. Programmable matter by folding. *Proceedings of the National Academy of Sciences of the United States of America*, 107(28):12441–12445, 2010.

[10] Evgueni T Filipov, Tomohiro Tachi, Glaucio H Paulino, and David A Weitz. Origami tubes assembled into stiff, yet reconfigurable structures and metamaterials. *Proceedings of the National Academy of Sciences of the United States of America*, 112(40):12321, 2015.

[11] Ahmad Zareei, Bolei Deng, and Katia Bertoldi. Harnessing transition waves to realize deployable structures. *Proceedings of the National Academy of Sciences of the United States of America*, 117(8):4015–4020, 2020.

[12] Harri Mökkönen, Tapio Ala-Nissila, and Hannes Jónsson. Efficient dynamical correction of the transition state theory rate estimate for a flat energy barrier. *The Journal of Chemical Physics*, 145(9):094901, September 2016.

[13] Donald G. Truhlar, Bruce C. Garrett, and Stephen J. Klippenstein. Current Status of Transition-State Theory. *The Journal of Physical Chemistry*, 100(31):12771–12800, January 1996.

[14] David Wales. *Energy Landscapes.* Cambridge University Press, Cambridge, 2004.

[15] Daniele Passerone and Michele Parrinello. Action-Derived Molecular Dynamics in the Study of Rare Events. *Physical Review Letters*, 87(10):108302, August 2001.

[16] Weinan E, Weiqing Ren, and Eric Vanden-Eijnden. Simplified and improved string method for computing the minimum energy paths in barrier-crossing events. *The Journal of Chemical Physics*, 126:164103, 2007.

[17] David J. Wales, Mark A. Miller, and Tiffany R. Walsh. Archetypal energy landscapes. *Nature*, 394(6695):758–760, 1998.

[18] David J. Wales and Tetyana V. Bogdan. Potential Energy and Free Energy Landscapes. *The Journal of Physical Chemistry B*, 110(42):20765–20776, October 2006.

[19] Christophe L. Vaillant, Stuart C. Althorpe, and David J. Wales. Path Integral Energy Landscapes for Water Clusters. *Journal of Chemical Theory and Computation*, 15(1):33–42, 2019.

[20] Martin Jansen. The energy landscape concept and its implications for synthesis planning. *Pure and Applied Chemistry*, 86(6):883–898, June 2014.

[21] Song-Ho Chong, Haeri Im, and Sihyun Ham. Explicit Characterization of the Free Energy Landscape of pKID–KIX Coupled Folding and Binding. *ACS Central Science*, 5(8):1342–1351, August 2019.

[22] Ratan Othayoth, Qihan Xuan, Yaqing Wang, and Chen Li. Locomotor transitions in the potential energy landscape-dominated regime. *Proceedings of the Royal Society B: Biological Sciences*, 288(1949):20202734, April 2021.

[23] Donald G. Truhlar and Aron Kuppermann. Exact tunneling calculations. *Journal of the American Chemical Society*, 93(8):1840–1851, April 1971.

[24] William H. Miller. Tunneling corrections to unimolecular rate constants, with application to formaldehyde. *Journal of the American Chemical Society*, 101(23):6810–6814, November 1979.

[25] Jeremy O. Richardson, Cristóbal Pérez, Simon Lobsiger, Adam A. Reid, Berhane Temelso, George C. Shields, Zbigniew Kisiel, David J. Wales, Brooks H. Pate, and Stuart C. Althorpe. Concerted hydrogen-bond breaking by quantum tunneling in the water hexamer prism. *Science*, 351(6279):1310, 2016.

[26] Jeremy O. Richardson. Perspective: Ring-polymer instanton theory. *The Journal of Chemical Physics*, 148(20):200901, May 2018.

[27] David D. Boehr, Dan McElheny, H. Jane Dyson, and Peter E. Wrightt. The dynamic energy landscape of dihydrofolate reductase catalysis. *Science*, 313(5793):1638–1642, 2006.

[28] Stephen J. Benkovic, Gordon G. Hammes, and Sharon Hammes-Schiffer. Free-energy landscape of enzyme catalysis. *Biochemistry*, 47(11):3317–3321, 2008.

[29] Chenxi Guo, Ziyun Wang, Dong Wang, Hai Feng Wang, and P. Hu. First-Principles Determination of CO Adsorption and Desorption on Pt(111) in the Free Energy Landscape. *Journal of Physical Chemistry C*, 122(37):21478–21483, 2018.

[30] José Nelson Onuchic, Zaida Luthey-Schulten, and Peter G. Wolynes. Theory of Protein Folding: The Energy Landscape Perspective. *Annual Review of Physical Chemistry*, 48(1):545–600, 1997.

[31] R Zwanzig, A Szabo, and B Bagchi. Levinthal's paradox. *Proceedings of the National Academy of Sciences of the United States of America*, 89(1):20–22, January 1992.

[32] Hao Yu, Derek R. Dee, Xia Liu, Angela M. Brigley, Iveta Sosova, and Michael T. Woodside. Protein misfolding occurs by slow diffusion across multiple barriers in a rough energy landscape. *Proceedings of the National Academy of Sciences*, 112(27):8308–8313, July 2015.

[33] David J. Wales. Decoding the energy landscape: Extracting structure, dynamics and thermodynamics. *Philosophical Transactions of the Royal Society A*, 370(1969):2877–2899, 2012.

[34] Jonathan P. K. Doye, Mark A. Miller, and David J. Wales. The double-funnel energy landscape of the 38-atom Lennard-Jones cluster. *The Journal of Chemical Physics*, 110(14):6896–6906, April 1999.

[35] Artem R. Oganov and Mario Valle. How to quantify energy landscapes of solids. *The Journal of Chemical Physics*, 130(10):104504, March 2009.

[36] Neil L. Allan, Sergio Conejeros, Judy N. Hart, and Chris E. Mohn. Energy landscapes of perfect and defective solids: From structure prediction to ion conduction. *Theoretical Chemistry Accounts*, 140(11):151, October 2021.

[37] Martin Jansen, Klaus Doll, and J. Christian Schön. Addressing chemical diversity by employing the energy landscape concept. *Acta Crystallographica Section A*, 66(5):518–534, 2010.

[38] David Richard and Thomas Speck. Crystallization of hard spheres revisited. I. Extracting kinetics and free energy landscape from forward flux sampling. *The Journal of Chemical Physics*, 148(12):124110, March 2018.

[39] Shigenobu Ogata, Ju Li, and Sidney Yip. Energy landscape of deformation twinning in bcc and fcc metals. *Physical Review B*, 71(22):224102, June 2005.

[40] Andreas Heuer. Exploring the potential energy landscape of glass-forming systems: From inherent structures via metabasins to macroscopic transport. *Journal of Physics: Condensed Matter*, 20(37):373101, August 2008.

[41] Patrick Charbonneau, Jorge Kurchan, Giorgio Parisi, Pierfrancesco Urbani, and Francesco Zamponi. Fractal free energy landscapes in structural glasses. *Nature Communications*, 5(1):3725, April 2014.

[42] Daniel J. Lacks and Mark J. Osborne. Energy Landscape Picture of Overaging and Rejuvenation in a Sheared Glass. *Physical Review Letters*, 93(25):255501, December 2004.

[43] Yuliang Jin and Hajime Yoshino. Exploring the complex free-energy landscape of the simplest glass by rheology. *Nature Communications*, 8(1):14935, April 2017.

[44] Peter K. Morse, Sudeshna Roy, Elisabeth Agoritsas, Ethan Stanifer, Eric I. Corwin, and M. Lisa Manning. A direct link between active matter and sheared granular systems. *Proceedings of the National Academy of Sciences*, 118(18):e2019909118, May 2021.

[45] Hyun Joo Hwang, Robert A. Riggleman, and John C. Crocker. Understanding soft glassy materials using an energy landscape approach. *Nature Materials*, 15(9):1031–1036, September 2016.

[46] Dwaipayan Chakrabarti and David J. Wales. Energy landscape of a model discotic liquid crystal. *Physical Review E*, 77(5):051709, May 2008.

[47] Halim Kusumaatmaja and Apala Majumdar. Free energy pathways of a multistable liquid crystal device. *Soft Matter*, 11(24):4809–4817, 2015.

[48] Halim Kusumaatmaja. Surveying the free energy landscapes of continuum models: Application to soft matter systems. *Journal of Chemical Physics*, 142:124112, 2015.

[49] Weiqing Ren. Wetting Transition on Patterned Surfaces: Transition States and Energy Barriers. *Langmuir*, 30(10):2879–2885, March 2014.

[50] J. R. Panter, Y. Gizaw, and H. Kusumaatmaja. Multifaceted design optimization for superomniphobic surfaces. *Science Advances*, 5(6):eaav7328, June 2019.

[51] Ratan Othayoth, G Thoms, and Chen Li. An energy landscape approach to locomotor transitions in complex 3D terrain. *Proceedings of the National Academy of Sciences of the United States of America*, 117(26):14987–14995, 2020.

[52] Zirui Zhai, Yong Wang, and Hanqing Jiang. Origami-inspired, on-demand deployable and collapsible mechanical metamaterials with tunable stiffness. *Proceedings of the National Academy of Sciences of the United States of America*, 115(9):2032–2037, February 2018.

[53] S. Felton, M. Tolley, E. Demaine, D. Rus, and R. Wood. A method for building self-folding machines. *Science*, 345(6197):644–646, August 2014.

[54] Hyunhyub Ko and Ali Javey. Smart Actuators and Adhesives for Reconfigurable Matter. *Accounts of Chemical Research*, 50(4):691, 2017.

[55] Sen W. Kwok, Stephen A. Morin, Bobak Mosadegh, Ju-Hee So, Robert F. Shepherd, Ramses V. Martinez, Barbara Smith, Felice C. Simeone, Adam A. Stokes, and George M. Whitesides. Magnetic Assembly of Soft Robots with Hard Components. *Advanced Functional Materials*, 24(15):2180–2187, April 2014.

[56] Tian Chen, Osama R Bilal, Kristina Shea, and Chiara Daraio. Harnessing bistability for directional propulsion of soft, untethered robots. *Proceedings of the National Academy of Sciences of the United States of America*, 115(22):5698, 2018.

[57] Yichao Tang, Yinding Chi, Jiefeng Sun, Tzu-Hao Huang, Omid H. Maghsoudi, Andrew Spence, Jianguo Zhao, Hao Su, and Jie Yin. Leveraging elastic instabilities for amplified performance: Spine-inspired high-speed and high-force soft robots. *Science Advances*, 6(19):eaaz6912, May 2020.

[58] Sicong Shan, Sung H. Kang, Jordan R. Raney, Pai Wang, Lichen Fang, Francisco Candido, Jennifer A. Lewis, and Katia Bertoldi. Multistable Architected Materials for Trapping Elastic Strain Energy. *Advanced Materials*, 27(29):4296–4301, 2015.

[59] Guangchao Wan, Yin Liu, Zhe Xu, Congran Jin, Lin Dong, Xiaomin Han, John X.J. Zhang, and Zi Chen. Tunable bistability of a clamped elastic beam. *Extreme Mechanics Letters*, 34:100603, 2020.

[60] Anastasiya Trushko, Ilaria Di Meglio, Aziza Merzouki, Carles Blanch-Mercader, Shada Abuhattum, Jochen Guck, Kevin Alessandri, Pierre Nassoy, Karsten Kruse, Bastien Chopard, and Aurélien Roux. Buckling of an Epithelium Growing under Spherical Confinement. *Developmental Cell*, 54(5):655–668.e6, September 2020.

[61] Emmanuel Virot, Tobias Kreilos, Tobias M. Schneider, and Shmuel M. Rubinstein. Stability Landscape of Shell Buckling. *Physical Review Letters*, 119(22):224101, November 2017.

[62] John W. Hutchinson and J. Michael T. Thompson. Imperfections and energy barriers in shell buckling. *International Journal of Solids and Structures*, 148–149:157–168, September 2018.

[63] M. Deml and W. Wunderlich. Direct evaluation of the ‘worst’ imperfection shape in shell buckling. *Computer Methods in Applied Mechanics and Engineering*, 149(1):201–222, October 1997.

[64] M. R. Nelson, D. Howard, O. E. Jensen, J. R. King, F. R. A. J. Rose, and S. L. Waters. Growth-induced buckling of an epithelial layer. *Biomechanics and Modeling in Mechanobiology*, 10(6):883–900, December 2011.

[65] Martin Lenz, Daniel J. G. Crow, and Jean-François Joanny. Membrane Buckling Induced by Curved Filaments. *Physical Review Letters*, 103(3):038101, July 2009.

[66] E. A. Evans. Bending elastic modulus of red blood cell membrane derived from buckling instability in micropipet aspiration tests. *Biophysical Journal*, 43(1):27–30, July 1983.

[67] Dian Yang, Bobak Mosadegh, Alar Ainla, Benjamin Lee, Fatemeh Khashai, Zhigang Suo, Katia Bertoldi, and George M. Whitesides. Buckling of Elastomeric Beams Enables Actuation of Soft Machines. *Advanced Materials*, 27(41):6323–6327, 2015.

[68] Jakob A. Faber, Janav P. Udani, Katherine S. Riley, André R. Studart, and Andres F. Arrieta. Dome-Patterned Metamaterial Sheets. *Advanced Science*, 7(22):2001955, 2020.

[69] Xin Ning, Xinge Yu, Heling Wang, Rujie Sun, R. E. Corman, Haibo Li, Chan Mi Lee, Yeguang Xue, Aditya Chempakasseril, Yao Yao, Ziqi Zhang, Haiwen Luan, Zizheng Wang, Wei Xia, Xue Feng, Randy H. Ewoldt, Yonggang Huang, Yihui Zhang, and John A. Rogers. Mechanically active materials in three-dimensional mesostructures. *Science Advances*, September 2018.

[70] Joselle M. McCracken, Sheng Xu, Adina Badea, Kyung In Jang, Zheng Yan, David J. Wetzel, Kewang Nan, Qing Lin, Mengdi Han, Mikayla A. Anderson, Jung Woo Lee, Zijun Wei, Matt Pharr, Renhan Wang, Jessica Su, Stanislav S. Rubakhin, Jonathan V. Sweedler, John A. Rogers, and Ralph G. Nuzzo. Deterministic Integration of Biological and Soft Materials onto 3D Microscale Cellular Frameworks. *Advanced Biosystems*, 1(9):1–16, 2017.

[71] Zheng Yan, Mengdi Han, Yan Shi, Adina Badea, Yiyuan Yang, Ashish Kulkarni, Erik Hanson, Mikhail E. Kandel, Xiewen Wen, Fan Zhang, Yiyue Luo, Qing Lin, Hang Zhang, Xiaogang Guo, Yuming Huang, Kewang Nan, Shuai Jia, Aaron W. Oraham, Molly B. Mevis, Jaeman Lim, Xuelin Guo, Mingye Gao, Woomi Ryu, Ki Jun Yu, Bruno G. Nicolau, Aaron Petronico, Stanislav S.

Rubakhin, Jun Lou, Pulickel M. Ajayan, Katsuyo Thornton, Gabriel Popescu, Daining Fang, Jonathan V. Sweedler, Paul V. Braun, Haixia Zhang, Ralph G. Nuzzo, Yonggang Huang, Yihui Zhang, and John A. Rogers. Three-dimensional mesostructures as high-temperature growth templates, electronic cellular scaffolds, and self-propelled microrobots. *Proceedings of the National Academy of Sciences of the United States of America*, 114(45):E9455–E9464, November 2017.

[72] Sheng Xu, Zheng Yan, Kyung-In Jang, Wen Huang, Haoran Fu, Jeonghyun Kim, Zijun Wei, Matthew Flavin, Joselle McCracken, Renhan Wang, Adina Badea, Yuhao Liu, Dongqing Xiao, Guoyan Zhou, Jungwoo Lee, Ha Uk Chung, Huanyu Cheng, Wen Ren, Anthony Banks, Xiuling Li, Ungyu Paik, Ralph G. Nuzzo, Yonggang Huang, Yihui Zhang, and John A. Rogers. Assembly of micro/nanomaterials into complex, three-dimensional architectures by compressive buckling. *Science*, 347(6218):154–159, January 2015.

[73] Yuan Liu, Xueju Wang, Yameng Xu, Zhaoguo Xue, Yi Zhang, Xin Ning, Xu Cheng, Yeguang Xue, Di Lu, Qihui Zhang, Fan Zhang, Jianxing Liu, Xiaogang Guo, Keh-Chih Hwang, Yonggang Huang, John A. Rogers, and Yihui Zhang. Harnessing the interface mechanics of hard films and soft substrates for 3D assembly by controlled buckling. *Proceedings of the National Academy of Sciences*, 116(31):15368–15377, July 2019.

[74] Jesse L. Silverberg, Arthur A. Evans, Lauren McLeod, Ryan C. Hayward, Thomas Hull, Christian D. Santangelo, and Itai Cohen. Using origami design principles to fold reprogrammable mechanical metamaterials. *Science*, 345(6197):647–650, August 2014.

[75] Yoonseok Park, Haiwen Luan, Kyeongha Kwon, Shiwei Zhao, Daniel Franklin, Heling Wang, Hangbo Zhao, Wubin Bai, Jong Uk Kim, Wei Lu, Jae-Hwan Kim, Yonggang Huang, Yihui Zhang, and John A. Rogers. Transformable, Freestanding 3D Mesostructures Based on Transient Materials and Mechanical Interlocking. *Advanced Functional Materials*, 29(40):1903181, October 2019.

[76] Xing Peng Hao, Zhao Xu, Chen Yu Li, Wei Hong, Qiang Zheng, and Zi Liang Wu. Kirigami-Design-Enabled Hydrogel Multimorphs with Application as a Multistate Switch. *Advanced Materials*, n/a(n/a):2000781, April 2020.

[77] Katia Bertoldi, Vincenzo Vitelli, Johan Christensen, and Martin Van Hecke. Flexible mechanical metamaterials. *Nature Reviews Materials*, 2:17066, 2017.

[78] Zachary G Nicolaou and Adilson E Motter. Mechanical metamaterials with negative compressibility transitions. *Nature Materials*, 11:608, 2012.

[79] R. S. Lakes, T. Lee, A. Bersie, and Y. C. Wang. Extreme damping in composite materials with negative-stiffness inclusions. *Nature*, 410:565, 2001.

[80] Kaikai Che, Chao Yuan, Jiangtao Wu, H Jerry Qi, and Julien Meaud. Three-Dimensional-Printed Multistable Mechanical Metamaterials With a Deterministic Deformation Sequence. *Journal of Applied Mechanics*, 84(1):011004, 2017.

[81] Katia Bertoldi, Pedro M. Reis, Stephen Willshaw, and Tom Mullin. Negative Poisson's Ratio Behavior Induced by an Elastic Instability. *Advanced Materials*, 22(3):361–366, 2010.

[82] Bastiaan Florijn, Corentin Coulais, and Martin van Hecke. Programmable Mechanical Metamaterials. *Physical Review Letters*, 113(17):175503, October 2014.

[83] Eder Medina, Patrick E Farrell, Katia Bertoldi, and Chris H Rycroft. Navigating the landscape of nonlinear mechanical metamaterials for advanced programmability. *Physical Review B*, 101:064101, 2020.

[84] Tark Raj Giri and Russell Mailen. Controlled snapping sequence and energy absorption in multistable mechanical metamaterial cylinders. *International Journal of Mechanical Sciences*, 204(May):106541, 2021.

[85] Lishuai Jin, Romik Khajehtourian, Jochen Mueller, Ahmad Rafsanjani, Vincent Tournat, Katia Bertoldi, and Dennis M. Kochmann. Guided transition waves in multistable mechanical metamaterials. *Proceedings of the National Academy of Sciences*, 117(5):2319–2325, February 2020.

[86] Jongmin Shim, Claude Perdigou, Elizabeth R. Chen, Katia Bertoldi, and Pedro M. Reis. Buckling-induced encapsulation of structured elastic shells under pressure. *Proceedings of the National Academy of Sciences*, 109(16):5978–5983, April 2012.

[87] Babak Haghpanah, Ladan Salari-Sharif, Peyman Pourrajab, Jonathan Hopkins, and Lorenzo Valdevit. Multistable Shape-Reconfigurable Architected Materials. *Advanced Materials*, 28(36):7915–7920, September 2016.

[88] Johannes T. B. Overvelde, James C. Weaver, Chuck Hoberman, and Katia Bertoldi. Rational design of reconfigurable prismatic architected materials. *Nature*, 541(7637):347–352, January 2017.

[89] Agustin Iniguez-Rabago, Yun Li, and Johannes T. B. Overvelde. Exploring multistability in prismatic metamaterials through local actuation. *Nature Communications*, 10:5577, 2019.

[90] J. T. B. Overvelde, S. Shan, and K. Bertoldi. Compaction Through Buckling in 2D Periodic, Soft and Porous Structures: Effect of Pore Shape. *Advanced Materials*, 24(17):2337–2342, 2012.

[91] Cheng Lv, Deepakshyam Krishnaraju, Goran Konjevod, Hongyu Yu, and Hanqing Jiang. Origami based Mechanical Metamaterials. *Scientific Reports*, 4(1):5979, August 2014.

[92] K Miura. Method of packaging and deployment of large membranes in space. Technical Report 618, The Institute of Space and Astronautical Science, 1985.

[93] Arthur A Evans, Jesse L Silverberg, and Christian D Santangelo. Lattice Mechanics of Origami Tessellations. *Physical Review E*, 92(1):13205, March 2015.

[94] Scott Waitukaitis, Rémi Menaut, Bryan Gin-ge Chen, and Martin van Hecke. Origami Multistability: From Single Vertices to Metasheets. *Physical Review Letters*, 114(5):055503, February 2015.

[95] Hongbin Fang, K. W. Wang, and Suyi Li. Asymmetric energy barrier and mechanical diode effect from folding multi-stable stacked-origami. *Extreme Mechanics Letters*, 17:7–15, November 2017.

[96] J. A. Harris and G. J. McShane. Impact response of metallic stacked origami cellular materials. *International Journal of Impact Engineering*, 147:103730, January 2021.

[97] Souhayl Sadik and Marcelo A. Dias. On local kirigami mechanics I: Isometric conical solutions. *Journal of the Mechanics and Physics of Solids*, 151(February):104370, 2021.

[98] Roberto Brighenti. Buckling of cracked thin-plates under tension or compression. *Thin-Walled Structures*, 43(2):209–224, 2005.

[99] Marcelo A. Dias, Michael P. McCarron, Daniel Rayneau-Kirkhope, Paul Z. Hanakata, David K. Campbell, Harold S. Park, and Douglas P. Holmes. Kirigami actuators. *Soft Matter*, 13(48):9087–9092, 2017.

[100] Jesse L. Silverberg, Jun-Hee Na, Arthur A. Evans, Bin Liu, Thomas C. Hull, Christian D. Santangelo, Robert J. Lang, Ryan C. Hayward, and Itai Cohen. Origami structures with a critical transition to bistability arising from hidden degrees of freedom. *Nature Materials*, 14(4):389–393, April 2015.

[101] J. R. Panter, J. Chen, T. Zhang, and H. Kusumaatmaja. Harnessing energy landscape exploration to control the buckling of cylindrical shells. *Communications Physics*, 2(1):151, 2019.

[102] Hao Zhou, Matthew Grasinger, Philip Buskohl, and Kaushik Bhattacharya. Low energy fold paths in multistable origami structures. *International Journal of Solids and Structures*, 265–266:112125, March 2023.

[103] Philip Wolfe. Convergence Conditions for Ascent Methods. *SIAM Review*, 11(2):226–235, 1969.

[104] Philip Wolfe. Convergence Conditions for Ascent Methods. II: Some Corrections. *SIAM Review*, 13(2):185–188, 1971.

[105] Larry Armijo. Minimization of functions having Lipschitz continuous first partial derivatives. *Pacific Journal of Mathematics*, 16(1):1–3, January 1966.

[106] Jorge Nocedal. Updating quasi-Newton matrices with limited storage. *Mathematics of Computation*, 35(151):773–782, 1980.

[107] Dong C. Liu and Jorge Nocedal. On the limited memory bfgs method for large scale optimization. *Mathematical Programming*, 45:503–528, 1989.

[108] R. (Roger) Fletcher. *Practical Methods of Optimization.* Chichester ; New York : Wiley, 1987.

[109] Erik Bitzek, Pekka Koskinen, Franz Gähler, Michael Moseler, and Peter Gumbsch. Structural relaxation made simple. *Physical Review Letters*, 97(17):170201, 2006.

[110] S Kirkpatrick, C D Gelatt, and M P Vecchi. Optimization By Simulated Annealing. *Science*, 220:671–680, 1983.

[111] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6):1087–1092, June 1953.

[112] David J. Wales and Jonathan P. K. Doye. Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms. *The Journal of Physical Chemistry A*, 101(28):5111–5116, July 1997.

[113] Z Li and H A Scheraga. Monte Carlo-minimization approach to the multiple-minima problem in protein folding. *Proceedings of the National Academy of Sciences*, 84(19):6611–6615, October 1987.

[114] Sourabh Katoch, Sumit Singh Chauhan, and Vijay Kumar. A review on genetic algorithm: Past, present, and future. *Multimedia Tools and Applications*, 80(5):8091–8126, February 2021.

[115] Tarek A El-Mihoub, Adrian A Hopgood, Lars Nolle, and Alan Battersby. Hybrid genetic algorithms: A review. *Engineering Letters*, 13(2):124–137, 2006.

[116] FR Manby, Roy L Johnston, and C Roberts. Predatory genetic algorithms. *MATCH (Communications in Mathematical and Computational Chemistry)*, 38:111–122, 1998.

[117] Graham A. Cox, Thomas V. Mortimer-Jones, Robert P. Taylor, and Roy L. Johnston. Development and optimisation of a novel genetic algorithm for studying model protein folding. *Theoretical Chemistry Accounts*, 112(3):163–178, July 2004.

[118] Weiqing Ren and Eric Vanden-Eijnden. A climbing string method for saddle point search. *Journal of Chemical Physics*, 138:134105, April 2013.

[119] Rachid Malek and Normand Mousseau. Dynamics of Lennard-Jones clusters: A characterization of the activation-relaxation technique. *Physical Review E*, 62(6):7723–7728, December 2000.

[120] Graeme Henkelman, Blas P. Uberuaga, and Hannes Jónsson. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *The Journal of Chemical Physics*, 113(22):9901–9904, December 2000.

[121] Charles J. Cerjan and William H. Miller. On finding transition states. *The Journal of Chemical Physics*, 75(6):2800–2806, 1981.

[122] Lindsey J. Munro and David J. Wales. Defect migration in crystalline silicon. *Physical Review B*, 59(6):3969–3980, February 1999.

[123] David J. Wales. Rearrangements of 55-atom Lennard-Jones and (C60)55 clusters. *The Journal of Chemical Physics*, 101(5):3750–3762, September 1994.

[124] Graeme Henkelman and Hannes Jónsson. A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. *The Journal of Chemical Physics*, 111(15):7010–7022, October 1999.

[125] Andreas Heyden, Alexis T. Bell, and Frerich J. Keil. Efficient methods for finding transition states in chemical reactions: Comparison of improved dimer method and partitioned rational function optimization method. *Journal of Chemical Physics*, 123:224101, 2005.

[126] Johannes Kästner and Paul Sherwood. Superlinearly converging dimer method for transition state search. *Journal of Chemical Physics*, 128:014106, 2008.

[127] Michael J. S. Dewar, Eamonn F. Healy, and James J. P. Stewart. Location of transition states in reaction mechanisms. *Journal of the Chemical Society, Faraday Transactions 2*, 80(3):227, 1984.

[128] Radu A. Miron and Kristen A. Fichthorn. The Step and Slide method for finding saddle points on multidimensional potential surfaces. *The Journal of Chemical Physics*, 115(19):8742–8747, November 2001.

[129] Irina V. Ionova and Emily A. Carter. Ridge method for finding saddle points on potential energy surfaces. *The Journal of Chemical Physics*, 98(8):6377–6386, April 1993.

[130] Xiao-Jie Zhang, Cheng Shang, and Zhi-Pan Liu. Double-Ended Surface Walking Method for Pathway Building and Transition State Location of Complex Reactions. *Journal of Chemical Theory and Computation*, 9(12):5745–5753, December 2013.

[131] Jhih-Wei Chu, Bernhardt L. Trout, and Bernard R. Brooks. A super-linear minimization scheme for the nudged elastic band method. *The Journal of Chemical Physics*, 119(24):12708–12717, December 2003.

[132] David J. Wales and Joanne M. Carr. Quasi-Continuous Interpolation Scheme for Pathways between Distant Configurations. *Journal of Chemical Theory and Computation*, 8(12):5020–5034, December 2012.

[133] Daniel Sheppard, Rye Terrell, and Graeme Henkelman. Optimization methods for finding minimum energy paths. *The Journal of Chemical Physics*, 128(13):134106, April 2008.

[134] Graeme Henkelman and Hannes Jónsson. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *Journal of Chemical Physics*, 113(22):9978–9985, 2000.

[135] Semen A. Trygubenko and David J. Wales. A doubly nudged elastic band method for finding transition states. *Journal of Chemical Physics*, 120(5):2082–2094, 2004.

[136] Weinan E, Weiqing Ren, and Eric Vanden-Eijnden. String method for the study of rare events. *Physical Review B*, 66(5):052301, August 2002.

[137] Randall S. Sexton, Robert E. Dorsey, and John D. Johnson. Optimization of neural networks: A comparative analysis of the genetic algorithm and simulated annealing. *European Journal of Operational Research*, 114(3):589–601, May 1999.

[138] Alexander Kerr and Kieran Mullen. A comparison of genetic algorithms and simulated annealing in maximizing the thermal conductance of harmonic lattices. *Computational Materials Science*, 157:31–36, February 2019.

[139] Elena F. Koslover and David J. Wales. Comparison of double-ended transition state search methods. *The Journal of Chemical Physics*, 127(13):134102, October 2007.

[140] Yanan Zhang and Weiqing Ren. Numerical study of the effects of surface topography and chemistry on the wetting transition using the string method. *The Journal of Chemical Physics*, 141(24):244705, 2014.

[141] Matteo Amabili, Alberto Giacomello, Simone Meloni, and Carlo Massimo Casciola. Collapse of superhydrophobicity on nanopillared surfaces. *Physical Review Fluids*, 2(3):034202, March 2017.

[142] Ryan L. Truby and Jennifer A. Lewis. Printing soft matter in three dimensions. *Nature*, 540(7633):371–378, 2016.

[143] Yinding Chi, Yanbin Li, Yao Zhao, Yaoye Hong, Yichao Tang, and Jie Yin. Bistable and Multistable Actuators for Soft Robots: Structures, Materials, and Functionalities. *Advanced Materials*, 34(19):2110384, May 2022.

[144] Silvia Bonfanti, Roberto Guerra, Francesc Font-Clos, Daniel Rayneau-Kirkhope, and Stefano Zapperi. Automatic design of mechanical metamaterial actuators. *Nature Communications*, 11:4162, 2020.

[145] Lei Zhang, Qiang Du, and Zhenzhen Zheng. Optimization-based Shrinking Dimer Method for Finding Transition States. *SIAM Journal on Scientific Computing*, 38(1):A528–A544, January 2016.

[146] Anne E. Trefethen and Jeyarajan Thiyagalingam. Energy-aware software: Challenges, opportunities and strategies. *Journal of Computational Science*, 4(6):444–449, 2013.

[147] John Shalf. The future of computing beyond Moore's Law. *Philosophical Transactions of the Royal Society A*, 378(2166):20190061, 2020.

[148] Francis Alexander, Ann Almgren, John Bell, Amitava Bhattacharjee, Jacqueline Chen, Phil Colella, David Daniel, Jack DeSlippe, Lori Diachin, Erik Draeger, Anshu Dubey, Thom Dunning, Thomas Evans, Ian Foster, Marianne Francois, Tim Germann, Mark Gordon, Salman Habib, Mahantesh Halappanavar, Steven Hamilton, William Hart, Zhenyu Huang, Aimee Hungerford, Daniel Kasen, Paul R.C. Kent, Tzanio Kolev, Douglas B. Kothe, Andreas

Kronfeld, Ye Luo, Paul Mackenzie, David McCallen, Bronson Messer, Sue Mniszewski, Chris Oehmen, Amedeo Perazzo, Danny Perez, David Richards, William J. Rider, Rob Rieben, Kenneth Roche, Andrew Siegel, Michael Sprague, Carl Steefel, Rick Stevens, Madhava Syamlal, Mark Taylor, John Turner, Jean Luc Vay, Artur F. Voter, Theresa L. Windus, and Katherine Yelick. Exascale applications: Skin in the game. *Philosophical Transactions of the Royal Society A*, 378(2166):20190056, 2020.

[149] Nam-Sua Lee and Klaus-Jürgen Bathe. Error indicators and adaptive remeshing in large deformation finite element analysis. *Finite Elements in Analysis and Design*, 16(2):99–139, May 1994.

[150] Angus McMullen, Miranda Holmes-Cerfon, Francesco Sciortino, Alexander Y. Grosberg, and Jasna Brujic. Freely Jointed Polymers Made of Droplets. *Physical Review Letters*, 121(13):138002, 2018.

[151] Valery Nguemaha and Huan Xiang Zhou. Liquid-Liquid Phase Separation of Patchy Particles Illuminates Diverse Effects of Regulatory Components on Protein Droplet Formation. *Scientific Reports*, 8(1):1–11, 2018.

[152] D. Chen, G. Zhang, and S. Torquato. Inverse Design of Colloidal Crystals via Optimized Patchy Interactions. *Journal of Physical Chemistry B*, 122(35):8462–8468, 2018.

[153] Mantu Santra, Rakesh S. Singh, and Biman Bagchi. Polymorph selection during crystallization of a model colloidal fluid with a free energy landscape containing a metastable solid. *Physical Review E*, 98(3):032606, September 2018.

[154] Peter Wriggers. *Computational Contact Mechanics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[155] Michael Page and James W. McIver. On evaluating the reaction path Hamiltonian. *The Journal of Chemical Physics*, 88(2):922–935, 1988.

[156] David J. Wales. Discrete path sampling. *Molecular Physics*, 100(2):3285–3305, October 2002.

[157] H. Kusumaatmaja, J. Léopoldès, A. Dupuis, and J. M. Yeomans. Drop dynamics on chemically patterned surfaces. *Europhysics Letters*, 73(5):740–746, March 2006.

[158] Philip S. Brown and Bharat Bhushan. Bioinspired materials for water supply and management: Water collection, water purification and separation of water from oil. *Philosophical Transactions of the Royal Society A*, 374(2073):20160135, August 2016.

[159] David J. Wales. OPTIM: A program for optimizing geometries and calculating reaction pathways.

[160] Paul Zimmerman. Reliable Transition State Searches Integrated with the Growing String Method. *Journal of Chemical Theory and Computation*, 9(7):3043–3050, July 2013.

[161] Kenneth A. Brakke. The Surface Evolver. *Experimental Mathematics*, 1(2):141–165, January 1992.

[162] Tzanio Kolev, Paul Fischer, Misun Min, Jack Dongarra, Jed Brown, Veselin Dobrev, Tim Warburton, Stanimire Tomov, Mark S. Shephard, Ahmad Abdelfattah, Valeria Barra, Natalie Beams, Jean-Sylvain Camier, Noel Chalmers, Yohann Dudouit, Ali Karakus, Ian Karlin, Stefan Kerkemeier, Yu-Hsiang Lan, David Medina, Elia Merzari, Aleksandr Obabko, Will Pazner, Thilina Rathnayake, Cameron W. Smith, Lukas Spies, Kasia Swirydowicz, Jeremy Thompson, Ananias Tomboulides, and Vladimir Tomov. Efficient exascale discretizations: High-order finite element methods. *The International Journal of High Performance Computing Applications*, 35(6):527–552, November 2021.

[163] K. Liu and G. H. Paulino. Nonlinear mechanics of non-rigid origami: An efficient computational approach. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 473(2206), 2017.

[164] A. Gillman, K. Fuchi, and P. R. Buskohl. Truss-based nonlinear mechanical analysis for origami structures exhibiting bifurcation and limit point instabilities. *International Journal of Solids and Structures*, 147:80–93, August 2018.

[165] Allen Van Gelder. Approximate Simulation of Elastic Membranes by Triangulated Spring Meshes. *Journal of Graphics Tools*, 3(2):21–41, 1998.

[166] H. S. Seung and David R. Nelson. Defects in flexible membranes with crystalline order. *Physical Review A*, 38(2):1005–1018, July 1988.

[167] Eitan Grinspun, Yotam Gingold, Jason Reisman, and Denis Zorin. Computing discrete shape operators on general meshes. *Computer Graphics Forum*, 25(3):547–556, 2006.

[168] Mara Denisse Rueda-Contreras, Andreu F. Gallen, J. Roberto Romero-Arias, Aurora Hernandez-Machado, and Rafael A. Barrio. On Gaussian curvature and membrane fission. *Scientific Reports*, 11(1):9562, May 2021.

[169] S. Dong. Multiphase flows of N immiscible incompressible fluids: A reduction-consistent and thermodynamically-consistent formulation and associated algorithm. *Journal of Computational Physics*, 361:1–49, May 2018.

[170] Kuan Yik Junn, Joe Henry Obit, and Rayner Alfred. Comparison of Simulated Annealing and Great Deluge Algorithms for University Course Timetabling Problems (UCTP). *Advanced Science Letters*, 23(11):11413–11417, November 2017.

[171] S. Rebay. Efficient Unstructured Mesh Generation by Means of Delaunay Triangulation and Bowyer-Watson Algorithm. *Journal of Computational Physics*, 106(1):125, 1993.

[172] Michael Wehner, Ryan L. Truby, Daniel J. Fitzgerald, Bobak Mosadegh, George M. Whitesides, Jennifer A. Lewis, and Robert J. Wood. An integrated design and fabrication strategy for entirely soft, autonomous robots. *Nature*, 536(7617):451–455, August 2016.

[173] Yoonho Kim, Hyunwoo Yuk, Ruike Zhao, Shawn A. Chester, and Xuanhe Zhao. Printing ferromagnetic domains for untethered fast-transforming soft materials. *Nature*, 558(7709):274–279, June 2018.

[174] A. Sydney Gladman, Elisabetta A. Matsumoto, Ralph G. Nuzzo, L. Mahadevan, and Jennifer A. Lewis. Biomimetic 4D printing. *Nature Materials*, 15(4):413–418, April 2016.

[175] Tao Xie. Tunable polymer multi-shape memory effect. *Nature*, 464(7286):267–270, March 2010.

[176] Dignesh Thesiya, A. R. Srinivas, and Piyush Shukla. A Novel Lateral Deployment Mechanism for Segmented Mirror/Solar Panel of Space Telescope. *Journal of Astronomical Instrumentation*, 04(03n04):1550006, December 2015.

[177] Yuan Hsing Fu, Ai Qun Liu, Wei Ming Zhu, Xu Ming Zhang, Din Ping Tsai, Jing Bo Zhang, Ting Mei, Ji Fang Tao, Hong Chen Guo, Xin Hai Zhang, Jing Hua Teng, Nikolay I. Zheludev, Guo Qiang Lo, and Dim Lee Kwong. A Micromachined Reconfigurable Metamaterial via Reconfiguration of Asymmetric Split-Ring Resonators. *Advanced Functional Materials*, 21(18):3589–3594, 2011.

[178] Wei Wang, Chenzhe Li, Hugo Rodrigue, Fengpei Yuan, Min Woo Han, Maenghyo Cho, and Sung Hoon Ahn. Kirigami/Origami-Based Soft Deployable Reflector for Optical Beam Steering. *Advanced Functional Materials*, 27(7), 2017.

[179] Kyung Hoon Lee, Kunhao Yu, Hasan Al Ba'ba'a, An Xin, Zhangzhengrong Feng, and Qiming Wang. Sharkskin-Inspired Magnetoactive Reconfigurable Acoustic Metamaterials. *Research*, 2020, February 2020.

[180] Sahab Babaee, Johannes T. B. Overvelde, Elizabeth R. Chen, Vincent Tournat, and Katia Bertoldi. Reconfigurable origami-inspired acoustic waveguides. *Science Advances*, 2(11):e1601019, November 2016.

[181] Bolei Deng, Pai Wang, Vincent Tournat, and Katia Bertoldi. Nonlinear transition waves in free-standing bistable chains. *Journal of the Mechanics and Physics of Solids*, 136:103661, March 2020.

[182] A. N. Damdam, N. Qaisar, and M. M. Hussain. Honeycomb-serpentine silicon platform for reconfigurable electronics. *Applied Physics Letters*, 115(11):112105, September 2019.

[183] Patrick L. Anderson, Arthur W. Mahoney, and Robert James Webster. Continuum Reconfigurable Parallel Robots for Surgery: Shape Sensing and State Estimation With Uncertainty. *IEEE Robotics and Automation Letters*, 2(3):1617–1624, July 2017.

[184] Nicolae Plitea, Dorin Lese, Doina Pisla, and Calin Vaida. Structural design and kinematics of a new parallel reconfigurable robot. *Robotics and Computer-Integrated Manufacturing*, 29(1):219–235, February 2013.

[185] Yichao Tang, Yanbin Li, Yaoye Hong, Shu Yang, and Jie Yin. Programmable active kirigami metasheets with more freedom of actuation. *Proceedings of the National Academy of Sciences*, 116(52):26407–26413, December 2019.

[186] Ke Liu, Tomohiro Tachi, and Glaucio H. Paulino. Invariant and smooth limit of discrete geometry folded from bistable origami leading to multistable metasurfaces. *Nature Communications*, 10(1):4238, September 2019.

[187] Qian Wang, Edward T. F. Rogers, Behrad Gholipour, Chih-Ming Wang, Guanghui Yuan, Jinghua Teng, and Nikolay I. Zheludev. Optically reconfigurable metasurfaces and photonic devices based on phase change materials. *Nature Photonics*, 10(1):60–65, January 2016.

[188] Zi Chen, Qiaohang Guo, Carmel Majidi, Wenzhe Chen, David J. Srolovitz, and Mikko P. Haataja. Nonlinear Geometric Effects in Mechanical Bistable Morphing Structures. *Physical Review Letters*, 109(11):114302, September 2012.

[189] Yijie Jiang, Lucia M. Korpas, and Jordan R. Raney. Bifurcation-based embodied logic and autonomous actuation. *Nature Communications*, 10(1):128, January 2019.

[190] Andreas Lendlein, Hongyan Jiang, Oliver Jünger, and Robert Langer. Light-induced shape-memory polymers. *Nature*, 434(7035):879–882, April 2005.

[191] Jungwook Kim, James A. Hanna, Myunghwan Byun, Christian D. Santangelo, and Ryan C. Hayward. Designing Responsive Buckled Surfaces by Halftone Gel Lithography. *Science*, 335(6073):1201–1205, March 2012.

[192] Larissa S. Novelino, Qiji Ze, Shuai Wu, Glaucio H. Paulino, and Ruike Zhao. Untethered control of functional origami microrobots with distributed actuation. *Proceedings of the National Academy of Sciences*, 117(39):24096–24101, September 2020.

[193] Jiayu Liu, Ozan Erol, Aishwarya Pantula, Wangqu Liu, Zhuoran Jiang, Kunihiko Kobayashi, Devina Chatterjee, Narutoshi Hibino, Lewis H. Romer, Sung Hoon Kang, Thao D. Nguyen, and David H. Gracias. Dual-Gel 4D Printing of Bioinspired Tubes. *ACS Applied Materials & Interfaces*, 11(8):8492–8498, February 2019.

[194] Zhen Ding, Chao Yuan, Xirui Peng, Tiejun Wang, H. Jerry Qi, and Martin L. Dunn. Direct 4D printing via active composite materials. *Science Advances*, 3(4):e1602890, April 2017.

[195] Yihui Zhang, Zheng Yan, Kewang Nan, Dongqing Xiao, Yuhao Liu, Haiwen Luan, Haoran Fu, Xizhu Wang, Qinglin Yang, Jiechen Wang, Wen Ren, Hongzhi Si, Fei Liu, Lihen Yang, Hejun Li, Juntong Wang, Xuelin Guo, Hongying Luo, Liang Wang, Yonggang Huang, and John A. Rogers. A mechanically driven form of Kirigami as a route to 3D mesostructures in micro/nanomembranes. *Proceedings of the National Academy of Sciences*, 112(38):11757–11764, September 2015.

[196] Xueju Wang, Xiaogang Guo, Jilong Ye, Ning Zheng, Punit Kohli, Dongwhi Choi, Yi Zhang, Zhaoqian Xie, Qihui Zhang, Haiwen Luan, Kewang Nan, Bong Hoon Kim, Yameng Xu, Xiwei Shan, Wubin Bai, Rujie Sun, Zizheng Wang, Hokyung Jang, Fan Zhang, Yinji Ma, Zheng Xu, Xue Feng, Tao Xie, Yonggang Huang, Yihui Zhang, and John A. Rogers. Freestanding 3D Mesostructures, Functional Devices, and Shape-Programmable Systems Based on Mechanically Induced Assembly with Shape Memory Polymers. *Advanced Materials*, 31(2):1805615, 2019.

[197] Kyung-In Jang, Kan Li, Ha Uk Chung, Sheng Xu, Han Na Jung, Yiyuan Yang, Jean Won Kwak, Han Hee Jung, Juwon Song, Ce Yang, Ao Wang, Zhuangjian Liu, Jong Yoon Lee, Bong Hoon Kim, Jae-Hwan Kim, Jungyup Lee, Yongjoon Yu, Bum Jun Kim, Hokyung Jang, Ki Jun Yu, Jeonghyun Kim, Jung Woo Lee, Jae-Woong Jeong, Young Min Song, Yonggang Huang, Yihui Zhang, and John A. Rogers. Self-assembled three dimensional network designs for soft electronics. *Nature Communications*, 8(1):15894, June 2017.

[198] Wonho Lee, Yuan Liu, Yongjun Lee, Bhupendra K. Sharma, Sachin M. Shinde, Seong Dae Kim, Kewang Nan, Zheng Yan, Mengdi Han, Yonggang Huang, Yihui Zhang, Jong-Hyun Ahn, and John A. Rogers. Two-dimensional materials in functional three-dimensional architectures with applications in photodetection and imaging. *Nature Communications*, 9(1):1417, April 2018.

[199] Wenbo Pang, Xu Cheng, Haojie Zhao, Xiaogang Guo, Ziyao Ji, Guorui Li, Yiming Liang, Zhaoguo Xue, Honglie Song, Fan Zhang, Zheng Xu, Lei Sang, Wen Huang, Tiefeng Li, and Yihui Zhang. Electro-mechanically controlled assembly of reconfigurable 3D mesostructures and electronic devices based on dielectric elastomer platforms. *National Science Review*, 7(2):342–354, February 2020.

[200] Yoonho Kim, German A. Parada, Shengduo Liu, and Xuanhe Zhao. Ferromagnetic soft continuum robots. *Science Robotics*, 4(33):eaax7329, August 2019.

[201] Wenqi Hu, Guo Zhan Lum, Massimo Mastrangeli, and Metin Sitti. Small-scale soft-bodied robot with multimodal locomotion. *Nature*, 554(7690):81–85, February 2018.

[202] Mila Boncheva, Stefan A. Andreev, L. Mahadevan, Adam Winkleman, David R. Reichman, Mara G. Prentiss, Sue Whitesides, and George M. Whitesides. Magnetic self-assembly of three-dimensional surfaces from planar sheets. *Proceedings of the National Academy of Sciences*, 102(11):3924–3929, March 2005.

[203] Timothy J. White and Dirk J. Broer. Programmable and adaptive mechanics with liquid crystal polymer networks and elastomers. *Nature Materials*, 14(11):1087–1098, November 2015.

[204] Xiaoxing Xia, Arman Afshar, Heng Yang, Carlos M. Portela, Dennis M. Kochmann, Claudio V. Di Leo, and Julia R. Greer. Electrochemically reconfigurable architected materials. *Nature*, 573(7773):205–213, September 2019.

[205] Yuanping Song, Robert M. Panas, Samira Chizari, Lucas A. Shaw, Julie A. Jackson, Jonathan B. Hopkins, and Andrew J. Pascall. Additively manufacturable micro-mechanical logic gates. *Nature Communications*, 10(1):882, February 2019.

[206] Hang Yang and Li Ma. Multi-stable mechanical metamaterials by elastic buckling instability. *Journal of Materials Science*, 54(4):3509–3526, February 2019.

[207] Matthew P. O'Donnell, Madeleine Towes, Rainer M. J. Groh, and Isaac V. Chenchiah. Exploring Adaptive Behavior of Non-linear Hexagonal Frameworks. *Frontiers in Materials*, 7, 2020.

[208] Yang Li and Sergio Pellegrino. A Theory for the Design of Multi-Stable Morphing Structures. *Journal of the Mechanics and Physics of Solids*, 136:103772, March 2020.

[209] Jun-Hee Na, Arthur A. Evans, Jinhye Bae, Maria C. Chiappelli, Christian D. Santangelo, Robert J. Lang, Thomas C. Hull, and Ryan C. Hayward. Programming Reversibly Self-Folding Origami with Micropatterned Photo-Crosslinkable Polymer Trilayers. *Advanced Materials*, 27(1):79–85, 2015.

[210] R. M. J. Groh and A. Pirrera. On the role of localizations in buckling of axially compressed cylinders. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 475(2224):20190006, April 2019.

[211] Joel Marthelot, Francisco López Jiménez, Anna Lee, John W. Hutchinson, and Pedro M. Reis. Buckling of a Pressurized Hemispherical Shell Subjected to a Probing Force. *Journal of Applied Mechanics*, 84(12), October 2017.

[212] Guoquan Luo, Haoran Fu, Xu Cheng, Ke Bai, Liping Shi, Xiaodong He, John A. Rogers, Yonggang Huang, and Yihui Zhang. Mechanics of bistable cross-shaped structures through loading-path controlled 3D assembly. *Journal of the Mechanics and Physics of Solids*, 129:261–277, August 2019.

[213] Eitan Grinspun, Anil N. Hirani, Mathieu Desbrun, and Peter Schröder. Discrete shells. In *Proceedings of the 2003 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pages 62–68, 2003.

[214] Ehsan Hajiesmaili and David R. Clarke. Reconfigurable shape-morphing dielectric elastomers using spatially varying electric fields. *Nature Communications*, 10(1):183, January 2019.

[215] Xiao Shang, Lu Liu, Ahmad Rafsanjani, and Damiano Pasini. Durable bistable auxetics made of rigid solids. *Journal of Materials Research*, 33(3):300–308, February 2018.

[216] Yunus Alapan, Alp C. Karacakol, Seyda N. Guzelhan, Irem Isik, and Metin Sitti. Reprogrammable shape morphing of magnetic soft machines. *Science Advances*, 6(38):eabc6414, September 2020.

[217] Zirui Zhai, Yong Wang, Ken Lin, Lingling Wu, and Hanqing Jiang. In situ stiffness manipulation using elegant curved origami. *Science Advances*, 6(47):eabe2000, November 2020.

[218] Yan Chen, Rui Peng, and Zhong You. Origami of thick panels. *Science*, 349(6246):396–400, July 2015.

[219] Levi H. Dudte, Etienne Vouga, Tomohiro Tachi, and L. Mahadevan. Programming curvature using origami tessellations. *Nature Materials*, 15(5):583–588, May 2016.

[220] David Melancon, Benjamin Gorissen, Carlos J. García-Mora, Chuck Hoberman, and Katia Bertoldi. Multistable inflatable origami structures at the metre scale. *Nature*, 592(7855):545–550, April 2021.

[221] Mingchao Liu, Lucie Domino, and Dominic Vella. Tapered elasticæ as a route for axisymmetric morphing structures. *Soft Matter*, 16(33):7739–7750, August 2020.

[222] Tian Chen, Mark Pauly, and Pedro M. Reis. A reprogrammable mechanical metamaterial with stable memory. *Nature*, 589(7842):386–390, 2021.

[223] Joschka Hellmeier, Rene Platzer, Alexandra S. Eklund, Thomas Schlichthaerle, Andreas Karner, Viktoria Motsch, Magdalena C. Schneider, Elke Kurz, Victor Bamieh, Mario Brameshuber, Johannes Preiner, Ralf Jungmann, Hannes Stockinger, Gerhard J. Schütz, Johannes B. Huppa, and Eva Sevcsik. DNA origami demonstrate the unique stimulatory power of single pMHCs as T cell antigens. *Proceedings of the National Academy of Sciences*, 118(4):e2016857118, January 2021.

[224] Sahand Sadeghi and Suyi Li. Dynamic folding of origami by exploiting asymmetric bi-stability. *Extreme Mechanics Letters*, 40:100958, October 2020.

[225] Hongbin Fang, Tse-Shao Chang, and K. W. Wang. Magneto-origami structures: Engineering multi-stability and dynamics via magnetic-elastic coupling. *Smart Materials and Structures*, 29(1):015026, December 2019.

[226] Mustapha Jamal, Sachin S. Kadam, Rui Xiao, Faraz Jivan, Tzia-Ming Onn, Rohan Fernandes, Thao D. Nguyen, and David H. Gracias. Bio-Origami Hydrogel Scaffolds Composed of Photocrosslinked PEG Bilayers. *Advanced Healthcare Materials*, 2(8):1142–1150, 2013.

[227] Shanshan Chen, Jianfeng Chen, Xiangdong Zhang, Zhi-Yuan Li, and Jiafang Li. Kirigami/origami: Unfolding the new regime of advanced 3D microfabrication/nanofabrication with “folding”. *Light: Science & Applications*, 9(1):75, April 2020.

[228] Weinan Xu, Zhao Qin, Chun-Teh Chen, Hye Rin Kwag, Qinli Ma, Anjishnu Sarkar, Markus J. Buehler, and David H. Gracias. Ultrathin thermoresponsive self-folding 3D graphene. *Science Advances*, 3(10):e1701084, October 2017.

[229] Lizhi Xu, Terry C. Shyu, and Nicholas A. Kotov. Origami and Kirigami Nanocomposites. *ACS Nano*, 11(8):7587–7599, August 2017.

[230] Peter G. Bolhuis, David Chandler, Christoph Dellago, and Phillip L. Geissler. TRANSITION PATH SAMPLING: Throwing Ropes Over Rough Mountain Passes, in the Dark. *Annual Review of Physical Chemistry*, 53(1):291–318, 2002.

[231] Ruslan D. Yamaletdinov, Oleg V. Ivakhnenko, Olga V. Sedelnikova, Sergey N. Shevchenko, and Yuriy V. Pershin. Snap-through transition of buckled graphene membranes for memcapacitor applications. *Scientific Reports*, 8(1):3566, February 2018.

[232] Marcus M. Noack and Simon W. Funke. Hybrid genetic deflated Newton method for global optimisation. *Journal of Computational and Applied Mathematics*, 325:97–112, December 2017.