MODELLING THE SHEAR MODULUS OF CHEESE

Natakala Dakshesh

CID: 01869590
SUPERVISOR: WILLIAM PROUD
ASSESSOR: YASMIN ANDREW
5900 Words

Contents

Declaration of work taken: I handled the modelling aspect, whilst my project partner focused on experimental methods, so little was shared, apart from the starting model.

Abstract

The first principles colloidal gel model of cheese protein to calculate the shear modulus was confirmed for small volume fractions of protein by fitting, with $A = 2.4 \times 10^5 \pm 1 \times 10^5$ and $b = 2.7 \pm 0.7$, with a 74% chance of falling within the predicted b range of 2.3-3.2. For large volume fractions, powers on a modified Lennard Jones potential have tentatively been found, but cannot be confirmed without more data, and until errors in temperature and continuity across the volume fraction range have been solved. At high volume fractions, colloidal gels act like uniform spheres following a potential. This was simulated, which confirmed that this monodisperse sphere model has incorrect volume fraction dependence, and again confirmed estimates for model parameters were correct. Large variations limit its power to predict the potential it does not solve temperature issues.

1 Introduction

Foods are complex materials, with non-uniform structures that can be modified by changing the composition, adding new ingredients, or by using new processing methods. This gives it new structural properties, improving the taste or making manufacturing easier and cheaper – important for a major global industry. However, this is difficult since properties are linked to composition and manufacturing processes through unknown, multivariate relationships [1] - when developing new foods, the only way to know the properties is to create and test them, which is expensive, especially if testing mass production processes [2] Therefore this is worth modelling to speed up this process. Since foods are so diverse, models must be built up for each type of food from their underlying physics, focusing on the specific factors that are important to each food type [3].

Cheese is the perfect application for multivariate modelling. The core of any cheese is its composition, which cheesemakers control using manufacturing processes. The key elements are the volume fractions of protein and fat, pH, salt, minerals and bacteria [4] – in that order– so any model must start here. Cheesemaker's have multiple processes [5] to choose from making models very useful to save time e.g. from maturation times which can take months [6]. It's also ideal since its utility is very sensitive to its key structural properties, such as bulk and shear moduli, melting points and density. Cheese is primarily used as an ingredient, so valued for its structural properties [4, 7] - characteristics like meltability, shreddability, and stretchability are very important [8] and easier to model than taste. Scientists even invented a new type of mozzarella to melt better on pizza [4]. Shear modulus, the ratio of stress to shear strain, is a key way to quantify these factors, so is the focus of this project [9-12].

Cheese modelling has traditionally focused on empirical models, where independent variables are changed, the dependent variable is measured, and then simple mathematical models are fitted to try and parameterise the relationship. These have many limitations. For example Yang et al. [13] measured the shear modulus of cheese for different fat contents and fit 16 different possible models linking shear modulus to fat content, but found multiple models that worked with conflicting physics. Furthermore, they only work for one variable, in the narrow conditions tested – a general problem with empirical models in food physics [2, 3], or any complex material. Finally, experimental data can be very slow to gather –

microbial counts, which are important in some cheese, take a day to measure, making it difficult to get enough data [2].

The solution is first principles microscopic modelling. This is powerful since it only requires the microscopic structure and interparticle interactions to calculate macroscopic properties [3, 14]. Fat and protein are the most important variables, but the properties of fat are well known [15] and can be combined with protein akin to Yang – protein is the focus of this project. Cheese is a viscoelastic material – it exhibits some solid and some liquid characteristics – so standard shear modulus and viscosity models do not work [4]. Instead it is closer to a polymer [12]. Qualitative microscopic models, treating the protein as a bespoke polymer have existed for decades [16], but no prebuilt models exist. However, only in 2019 Graeme Gillies [12] made a quantitative model, estimating the interparticle pair potential as a modified Lennard-Jones potential, and using approaches from soft matter physics to approximate the structure as a colloid (large particles in a solvent), allowing him to get an equation linking shear modulus to the volume fraction of protein. Using colloidal gels as models for food is common [17], but this was the first time it was applied to cheese.

This project aimed to build on Gillies, firstly by comparing the model more rigorously to real world data to validate it, and if possible, extend the model to more factors that affect the structure and shear modulus of cheese, such as fat and pH. I did this by analysing the model by Gillies, and then running computer simulations to understand how to improve it. This is a multiscale, multivariate modelling problem, with applications across food and materials physics, as colloid gels (and at high volume fractions, glasses) have many uses, and the general approach here applies to any modelling problem.

2 Methods

2.1 Qualitative Protein Model

To make a microscopic model, we need the interparticle potential and the microscopic structure of protein in cheese, and to do that, we must understand how cheese is formed. Cheese is made from milk, which has 3 key components: water (the solvent), fat globules, and casein micelles.

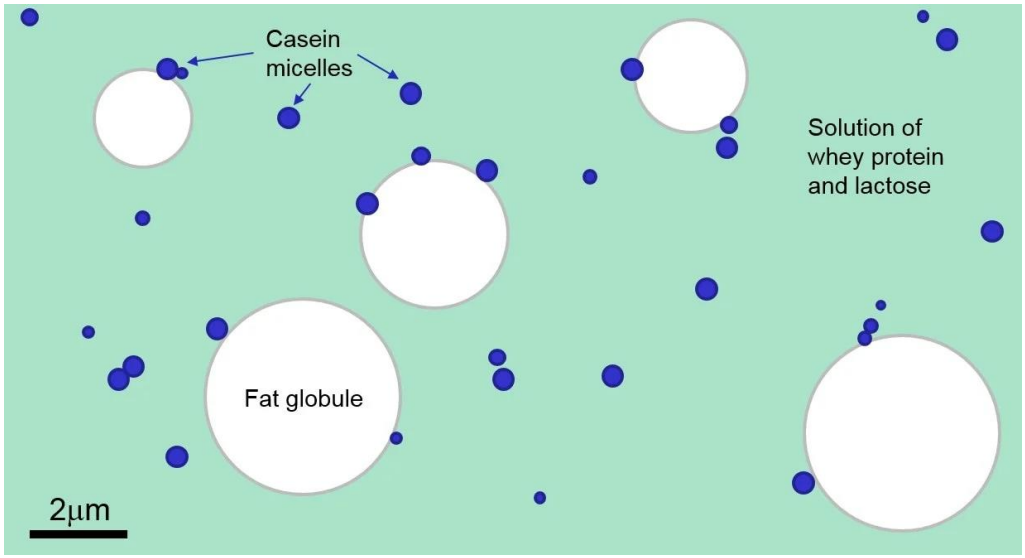

Figure 1: Structure of milk [18]

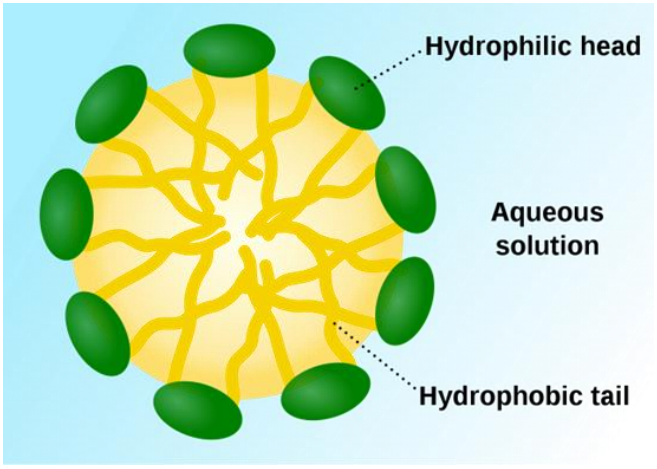

Figure 2**:** Diagram of casein micelle [19]. Some parts of casein are hydrophobic, so gather in the middle of the ball, whereas other parts are hydrophilic, so stay on the outside.

Caseins are proteins, and micelles are a ball shaped supra-structure of particles, shown in figure 2. When cheese is made, milk is dehydrated so most of the water leaves, forcing the micelles closer together. Then the outer layer of the micelles is stripped away, letting the internal structures of the micelles merge into a continuous protein matrix [16]. Since the fat particles are on a higher length scale to the protein [13], they don't directly affect the properties of the protein matrix so can be incorporated later. Other variables e.g. pH can be added later as perturbations to the potential or structural changes.

We need the structure of the continuous protein matrix. Horne's dual binding model [16], shown in figure 3, hypothesises interlinking chains of casein proteins and 'core-shell calcium phosphate nanoclusters' (CCP). CCP are clusters of $\sim 10^2$ calcium phosphate molecules [20].

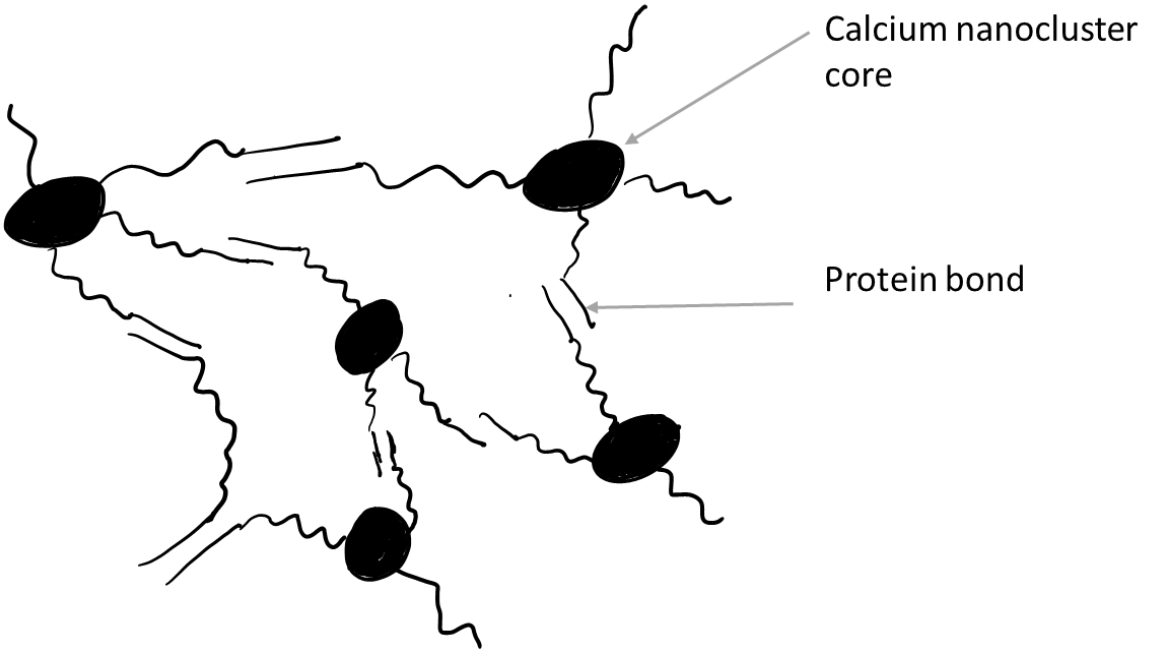

Figure 3: Structure of protein matrix. One end of a casein protein bonds to a calcium nanocluster, whilst the other end bonds to a protein – the chain length is limited to 1. This is randomised.

This type of structure isn't modelled by current polymer theories. Therefore, Gillies simplified the structure and modelled it as a colloidal gel. By modelling each calcium nanocluster and its bonded caseins as one soft colloid, as shown in figure 4, he could use colloidal physics to calculate bulk structural properties. Its key dimensions are R_{HS}, the hard sphere radius, below which colloid deformation requires significant work, and c, the number of caseins per colloid. This allows us to find the relationship between the volume fraction of protein and the shear modulus.

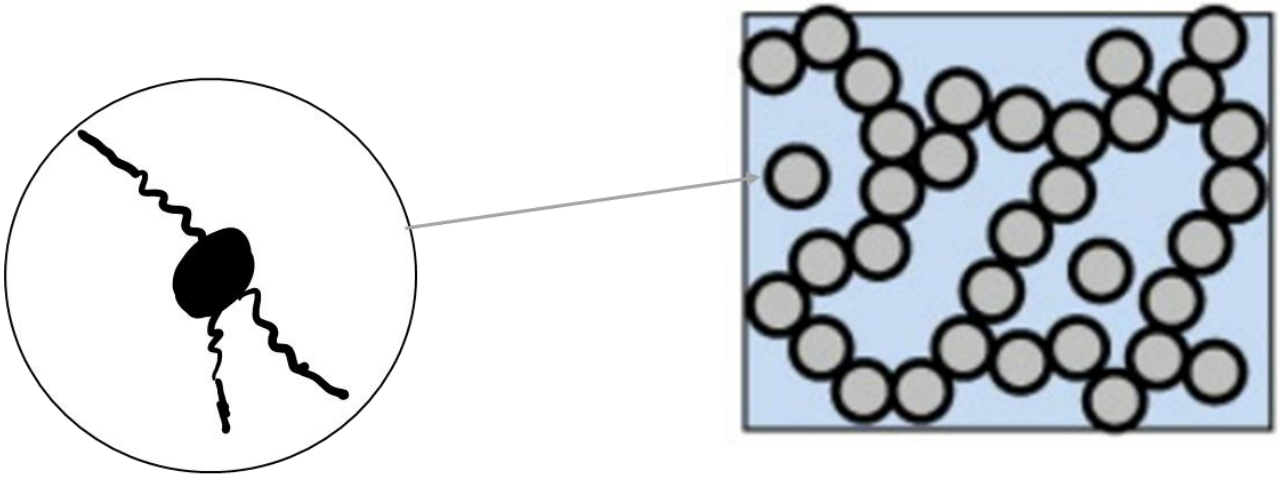

Figure 4: (Right) Colloidal gel model. (Left) how Horne model is approximated as a colloidal gel by Gillies.

2.2 Quantitative Shear Modulus and Protein Model

From soft matter physics, Gillies found general equations for the shear modulus (G) of a colloid gel, where φ is the volume fraction of protein:

$$G = \begin{cases} A\varphi^b, & \varphi \leq \varphi_{CP} \\ Nk_BT + \frac{\varphi_{cp} j}{5\pi r^{\star 2}}\left(4\frac{dU(r^\star)}{dr} + r^\star \frac{d^2U(r^\star)}{dr^2}\right), & \varphi \geq \varphi_{CP} \end{cases} \qquad (1, 2)$$

U is the inter-colloidal pair potential, given by (3). Each term is explained in table I, and estimated values are given. If $\varphi \leq \varphi_{CP}$, where φ_{CP} is the close packing ratio, the average distance between colloids is great enough that only the largest aggregation of colloids determines the shear modulus, which in turn depends on the volume fraction itself. If $\varphi \geq \varphi_{CP}$, the colloid gel tends to a collection of monodisperse spheres [21] - a set of mechanics - as interparticle interactions play a greater role. (2) holds if the potential is a linear sum of power laws. Gillies approximated this using the Mie Potential, a modified Lennard Jones potential, as in (3), guessing n = 6 and m = 3 by comparison with star polymers, but allowing for flexibility.

$$U(d) = \frac{n}{n-m}\left(\frac{n}{m}\right)^{\frac{m}{n-m}} \epsilon\left(\left(\frac{\sigma}{d}\right)^n - \left(\frac{\sigma}{d}\right)^m\right) \qquad (3)$$

b can be estimated from first principles:

$$b = \frac{1+\beta}{d_f - 3} \qquad [22], (4)$$

Where d_f is the gel fractal dimension, which ranges from 1.68 to 2.03 for all colloidal gels [23], and $\beta \approx 3.1 \pm 0.1$, the bond dimension [22]. d_f will depend on the current conditions on cheese, so further simulations or calculations can improve this.

To validate the model, it was firstly compared against experimental data. Gillies did some comparisons, the fits (figure 5) are poor, especially for (2).

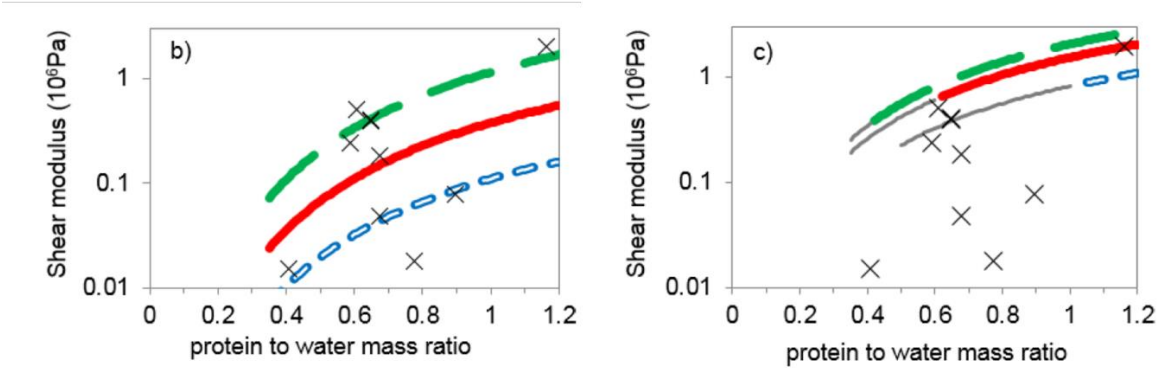

Figure 5: Comparisons to the model from Gillies [12]. (1) (left) is reasonable but (2) (right) is poor. Use protein to moisture ratio instead of volume fraction.

TABLE I

MODEL PARAMETERS AND ESTIMATES

Symbol	NAME	Value	Explanation
R_{HS}	Hard Sphere Radius	$5 \pm 0.5\ nm$	Reasonable assumption for room temperature. Full range is 1 – 10nm, bounded by total size of colloid.
A	N/A	Unknown	For volume fractions below close packing, there is a proportionality relationship between volume fraction and G.
b	N/A	2.3 – 3.2	Power in proportionality relationship, range of colloidal gels.
c	Caseins per colloid		
N	Number density of stress bearing bonds	$\frac{c\varphi}{\frac{4}{3}\pi R_{HS}^3}$	Where *c* is number of caseins per colloid. First calculate the number density of colloids, then multiply by caseins which determine bonds
φ_{CP}	Close Packing Fraction	~0.6	Above this, the colloids are no longer isolated from neighbours and may be deformed. Material dependent.
J	Nearest neighbours	~11	From random packing of polydisperse spheres.
r	Interparticle distance	N/A	Required for pair potential
R_{eq}	Equilibrium Radius	$2R_{HS} = \int_0^{2R_{eq}} 1 - e^{-\frac{U(r)}{kT}}\ dr$ Found to be 3.8nm.	For hard spheres. The probability of r being below $2R_{HS} \sim 0$. (limited deformation). Limit to $2R_{eq}$, since beyond that interaction potential is very small. [26] If applied to standard hard spheres, $U = \infty$, $r \leq R_{HS}$, else $U = 0$, so would return $2R_{HS}$ – we're comparing with our potential
$r^\star$	Expected interparticle distance	$\begin{cases} 2R_{eq}, & \varphi \leq \varphi_{CP} \\ 2R_{eq}\left(\frac{\varphi_{cp}}{\varphi}\right)^{\frac{1}{3}}, & \varphi \geq \varphi_{CP} \end{cases}$ **AND** $r^\star \geq 2R_{Calcium\ Core}$	If particles are closely packed, find particles cannot get close enough to reach minimum energy. **Note:** centre-centre separation must be greater than incompressible calcium core diameter, 6nm
ϵ	Minimum Energy	$2 - 4k_bT$	By comparison to polymer gels
n, m	N/A	6, 3	Repulsive and attractive power law in Mie potential. By comparison to star polymers.
σ	Zero potential	$2R_{eq}\left(\frac{m}{n}\right)^{\frac{1}{n-m}}$, ~6nm	Distance where U = 0. Linked to R_{eq}, therefore R_{HS}

We need more to check its validity. To do this, I found experimental data for shear and Young's Moduli online. Due to limited data, I assumed constant temperature T = 298K, so only φ was changed, and ignored other factors such as pH and fat content. Temperature only varied by 5K, which at most affects shear modulus by 15% for soft cheeses which melt at lower temperatures [24]. Typically cheese composition is given in percentage by mass of fat, moisture, solid content and protein. To convert into a volume fraction, we use:

$$\varphi = N_{Colloid}\frac{4}{3}\pi R_{HS}^3 \qquad (5)$$

Where $N_{Colloid}$ is the number density of colloids:

$$N_{Colloid} = \frac{protein\ percentage}{cm_c} \qquad (6)$$

Where c is caseins per colloid and m_c = 25,000 atomic mass units is the casein mass [12]. The rest gives the volume of 1 colloid. Sometimes protein wasn't given, so was calculated from other components. For soft cheeses, most fat was in liquid phase, whilst protein was solid, so protein was the solid content. For hard cheeses, protein could only be calculated if the solidified fat content was given. For some cheeses, the shear modulus could not be found, but Young's Modulus (E) could.

$$E = 2G(1 + \upsilon) \qquad (7)$$

and for cheese $\upsilon = 0.5$ is reasonable [25], so G is a factor of 3 smaller than E.

Most model parameters are given in Table I, but R_{eq} was more complex to calculate. To solve for R_{eq}, a 1D root finding problem was set up as in (8) and solved numerically. The integral was also evaluated numerically.

$$f(R_{eq}) = 2R_{HS} - \int_0^{2R_{eq}} 1 - e^{-\frac{U(r)}{kT}}\, dr = 0 \quad [26]\ (8)$$

Note, $R_{eq} \geq Calcium\ core\ radius$ = 3 nm [20].

To validate (1) I fitted to the experimental data. For (2) fitting is unfeasible due to the number of variables. Instead estimated values were taken from Table I, the equation is plotted, and then a sensitivity analysis is performed to see which variables are most important, by calculating numerical derivatives calculated.

3 Modelling Results

3.1 Low Volume Fraction

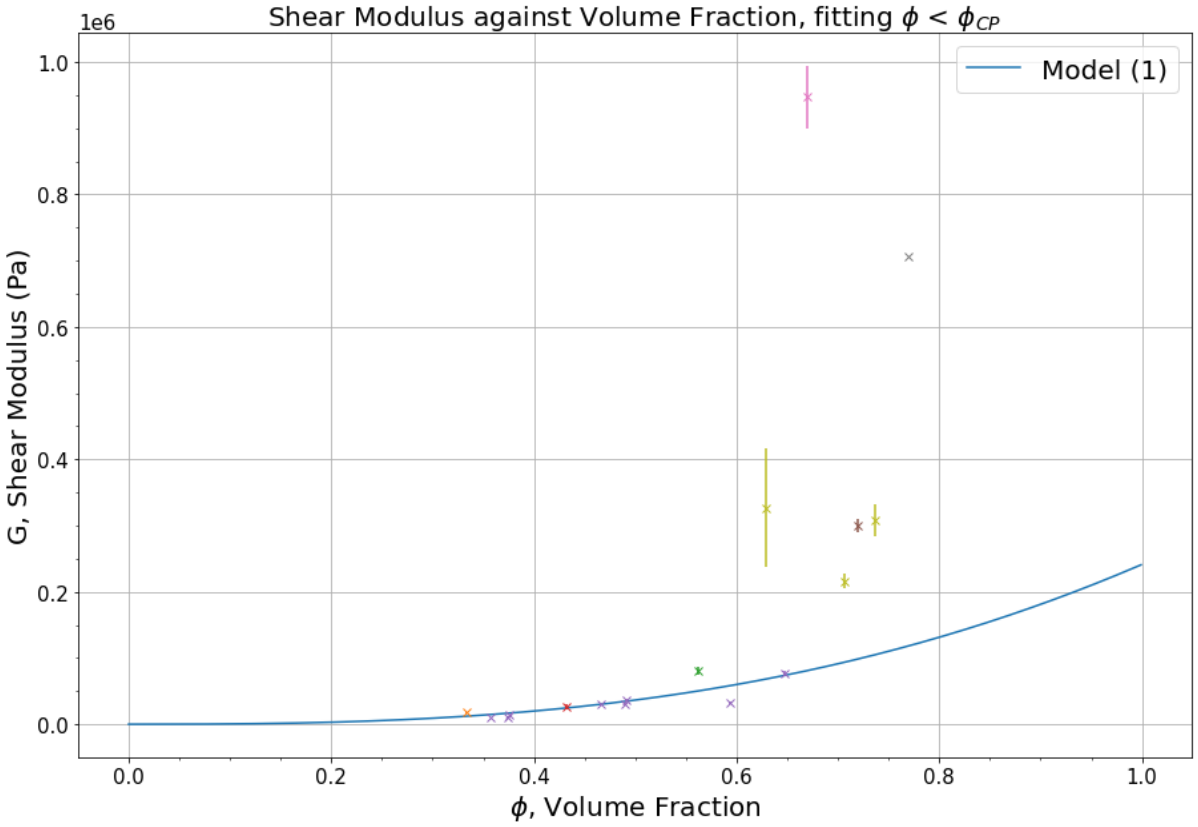

Figure 6: (1) is fitted to experimental data for $\varphi \leq \varphi_{CP} = 0.6$. Larger φ are ignored in the fit, shown for comparison. Data from [27-34]

The fit values for (1) are found to be $A = 2.4 \times 10^5 \pm 1 \times 10^5$, and $b = 2.7 \pm 0.7$. The large uncertainty on b means we cannot fully accept this model, as there is a p-value of 0.24 that b lies below 2.3, the estimated parameter. However, the model isn't tuned for cheese, instead general colloidal gels, so the range of b may be inaccurate, as β is not known. Therefore, this confirms that colloidal gel relationships hold for cheese. It also supports the concept of high-volume fraction cheeses falling into a different set of physics to low volume fraction cheeses, and that $\varphi_{CP} \sim 0.6$. If all data is included, $b = 4.7 \pm 1.4$, with a p-value of only 0.16 for the hypothesis that the fitted b falls into the modelled b, meaning we can reject this as a universal model with 16% significance. Worst of all, only fitting (1) for points $\varphi \geq 0.6$ causes convergence to fail. To reduce uncertainties, more data points are needed, with cheeses of similar pH and fat levels, and temperatures should be identical – although all data taken had T = 293-298K to limit this.

3.2 High Volume Fractions

This part was more difficult. In [12], (2) is written incorrectly, with $\frac{1}{r^{\star}}$ instead of $\frac{1}{r^{\star 2}}$, by comparison with [21]. This version suggested that the shear modulus was constant with all variables except temperature, for which it increased – which is clearly wrong. My original conclusion was either the equation or the potential were very wrong, so I which inspired me to run a simulation.

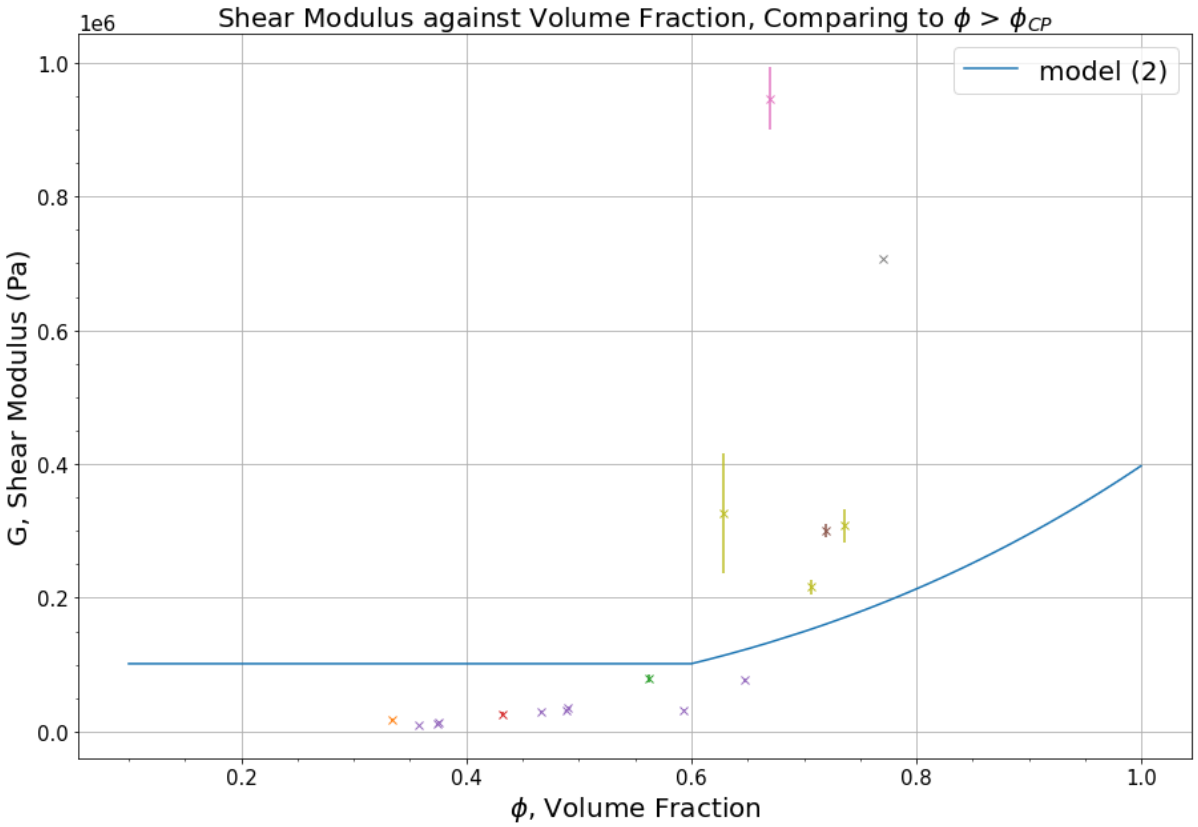

Figure 7: Plot of (2) with estimated parameters from table I. Underestimates most points $\varphi \geq \varphi_{CP}$. Data from [27 – 34].

When the equation is fixed, the predictions for high volume fraction from (2) are still poor, but significantly better (and are better than Gillies results' in figure 5) And it's notable the model has the correct order of magnitude for G, despite all terms coming from first principles. As expected, when $\varphi \leq \varphi_{CP}$ the modulus is constant as φ leaves the equation. However, when $\varphi \geq \varphi_{CP}$ (2) consistently underestimates G – even with large variations in data. To understand which parameters are worth changing, I performed a sensitivity analysis as shown in Table II, using estimates for values from Table I. Note $\varphi = 0.7$, since there are multiple points are in this area, and T = 298K, room temperature.

I didn't consider N for two reasons. Firstly, variation in $N_{Colloid}$ from (5) affects φ calculations. The key term in (5) is c, as uncertainties on the protein percentage are around 1%. Meanwhile φ changes by $\sim$0.2 if c changes by 2 (c must be even, and is constrained by measurements of protein mass, so this is the greatest change [12]). It's more intuitive to study φ. Secondly, N affects the first term in (2), NkT. This term is inconsequential. Numerically, it's only $O(10^4)$, and physically, we expect modulus to decrease as temperature increased,

since inter-colloidal bonds are easier to break. Hence, we can ignore it, and N from our analysis.

TABLE II
SENSITIVITY ANALYSIS

Variable	DIFFERENTIAL (CHANGE PER ONE SI UNIT)	Order of Magnitude 0Change	Expected Variation	Comments
R_{HS}	-2.74×10^{14}	$O(10^{-9})$	-2.74×10^{5}	Assume φ is constant, as it's the independent variable.
φ	6.05×10^{5}	~ 0.25	9.10×10^{4}	Magnitude calculated from how variation in $R_{HS} = 5 \pm 0.5\, nm$ and c affects calculated φ.
φ_{CP}	1.01×10^{5}	~ 0.1	1.01×10^{5}	Too small, since this is the most it can be changed.
J	1.17×10^{4}	1	1.17×10^{4}	Too small
ε	1.04×10^{25}	$O(k_b T)$	$138T \sim 1 \times 10^{4}$	Too small. Also unphysical, as temperature effects are much greater and negative
n	5.20×10^{4}	O(1)	5.20×10^{4}	Seems useful
m	1.10×10^{5}	O(1)	1.10×10^{5}	Seems useful

The sensitivity analysis offers some options to tune the model. Whilst R_{HS} would make G a reasonable estimate of the current data points at $\varphi = 0.7$, R_{HS} is linked to φ by (4). This means decreasing R_{HS} by 0.5nm would increase G but shift all points down by 0.15, making the model still underestimate our values, or pushing them into $\varphi \leq \varphi_{CP}$, negatively affecting the fit to (1). This does validate the choice for $R_{HS} = 5nm$, and shows R_{HS} is not suitable to change.

However, R_{HS} does explain the large variations in the data, through φ. Other reasons include T (discussed earlier) and fat content. The other main reason for variation is fat content. [13] found the modulus would half with no fat for hard cheeses. Fat content varies by around 10% (with a mean of 30%), introducing variation - and this theory is supported by the greater variation at higher volume fractions, since that is where solid fat reinforces the protein matrix, increasing the shear modulus.

n and m are more promising and were expected to be tuned by Gillies. However, increasing n or m individually causes the whole function to rise, shown by figures 8&9.

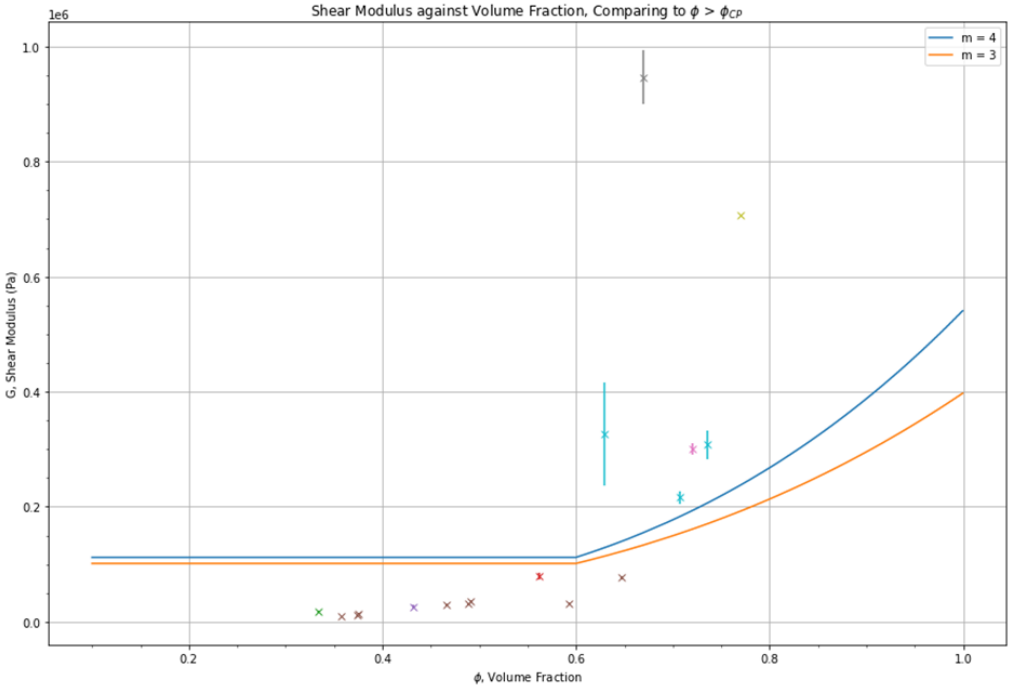

Figure 8: Comparison of shear modulus with m = 3, 4

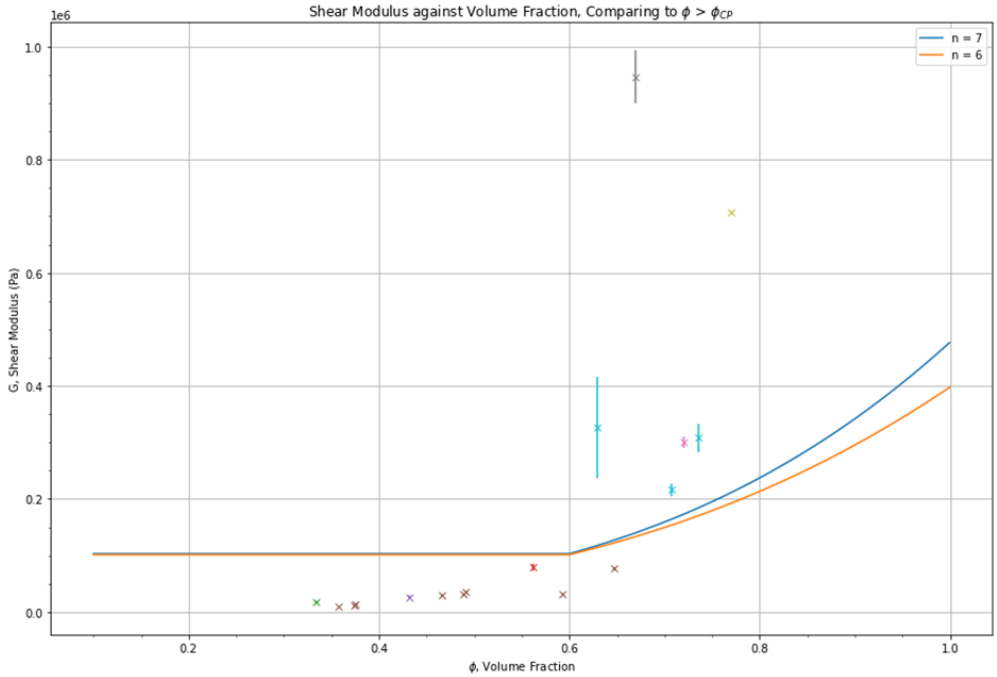

Figure 9: Comparison of shear modulus with n = 6, 7. Difference is smaller, as expected.

This increases the discontinuity at $\varphi = 0.6$ with the previous model. We need to find the best fit of m and n to all the data, but this is difficult since they must be integers so ordinary fitting methods fail [35]**.** Instead, it's best to manually compare the model output to the mean value

of the data points of $\varphi \leq 0.6$. I found n = 10 and m = 9 to give values of G closest to the data. The relationship between m, n and G is visualised in figure 10**.**

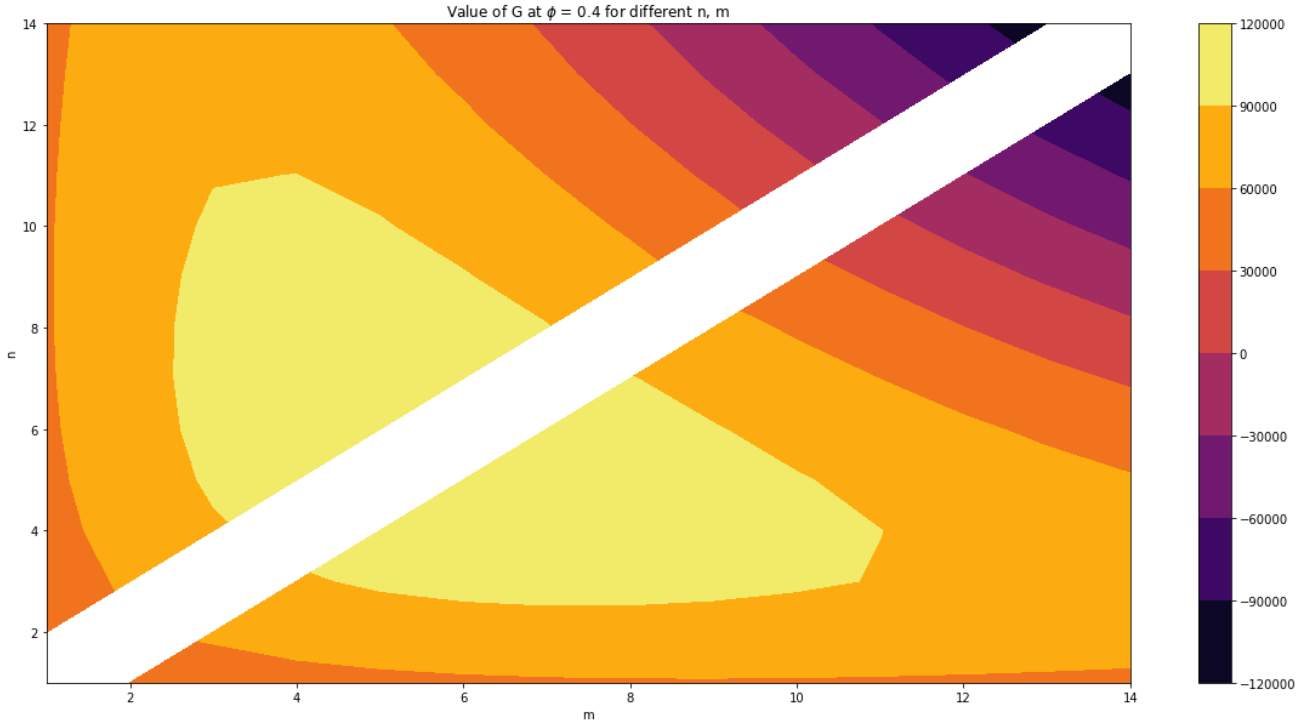

Figure 10: Contour plot of shear modulus, n and m. n cannot be m, so central diagonal is blank. Only valid for integer values. Shows both n and m must be increased to reduce the modulus to ~30,000 Pa, the average value of points $\varphi \leq \varphi_{CP}$

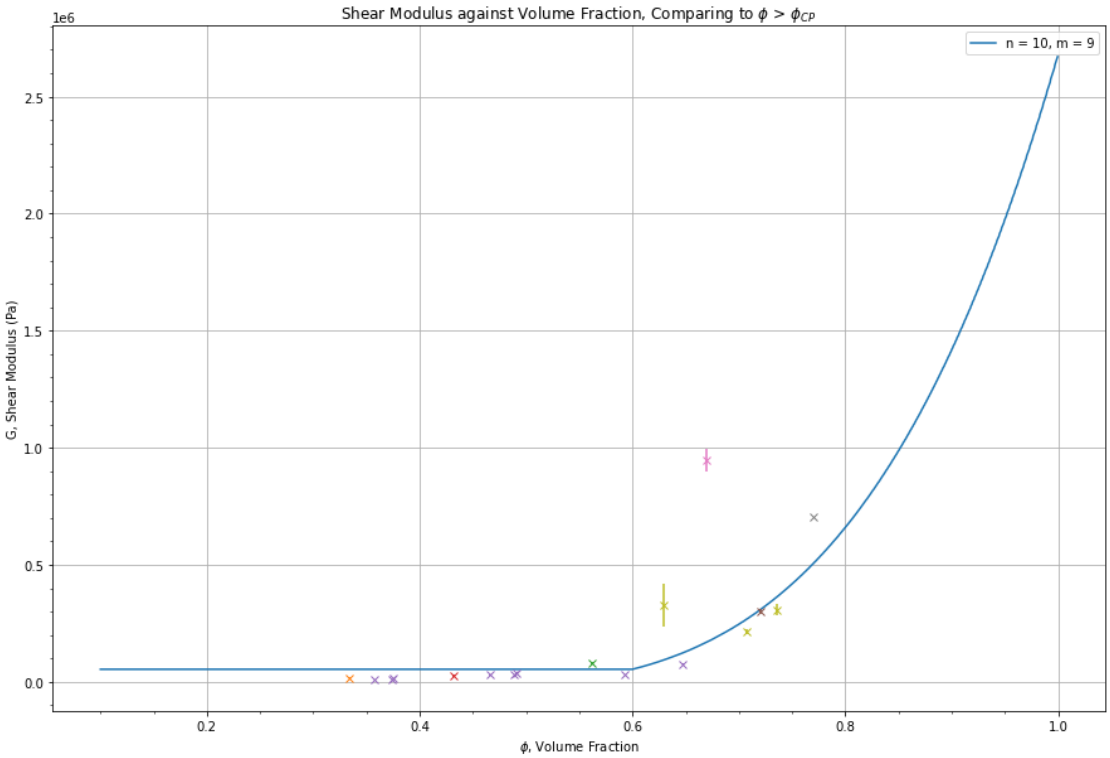

Figure 11: Model for n =10, m = 9. Good fit.

Figure 11 shows the results. It shows a much better fit to the data, and the very large G values for large φ can be ignored, as they are unphysical for cheese.

3.3 Issues and Conclusions

Still, this approach is unsatisfactory since there is no physical explanation for these powers. As cheese data is hugely variable – which is why we want to use a first principles approach, based on physical reasoning. Furthermore, there are some strange behaviours. Firstly, the temperature is unphysical. The temperature dependence is in ε and NkT (as it cancels out of (6)), but both suggest small linear increases in G with T, which is completely unphysical – real cheeses weaken temperature [24]**.** This is true for most materials – metals, polymers, colloidal gels etc. suggesting there is a big problem with the model. This suggests the monodisperse spheres approximation for large volume fractions isn't perfect – this temperature dependence is true for all monodisperse sphere solutions following Lennard Jones style potentials [36].

It has been observed that as temperatures increase, 'clusters form' [37] in colloidal gels. Large clusters break up the uniformity that (2) assumes – derived from equations to describe monodisperse spheres – and could fall into a different shear modulus regime. This suggests a transition scheme is needed to better describe the transition region between (1) and (2), although this would depend on the cluster sizes and therefore the temperature difference. Differential effective medium theories, which have been used to describe the impact of fat in cheese [13] can also be applied multiple times for multiple cluster sizes, akin to the 'dirty binder approximation' used in explosives manufacturing [38]**.** Alternatively, the temperature dependence of the potential could be wrong, but this is unlikely since it comes from basic thermodynamics.

Secondly is the lack of φ dependence for $\varphi \leq 0.6$. It's simply untrue that φ dependence disappears, as again (2) comes from a general formula describing monodisperse spheres. This suggests an issue with $r^\star$ (therefore (2) in general), as that contains the φ dependence. Perhaps at lower volume fractions, $r^\star$ is the average distance between colloids.

These two issues make it worth questioning (2), $r^\star$ and the potential.

4 Modelling Conclusions

Overall, I have shown that (1) is a good fit to data for $\varphi \leq \varphi_{CP}$, showing that modelling the protein matrix in cheese as a colloidal gel is reasonable. $A = 2.4 \times 10^5 \pm 1 \times 10^5$ and $b = 2.7 \pm 0.7$, with a 74% chance of falling within the predicted b range of 2.3-3.2**.** However, this fails at high volume fractions and requires more structural simulations or data to solve improve. (2), the proposed model for high volume fractions can fit the data, for n = 10, m = 9. This is unsatisfactory, as there is no physical reasoning to support this, and the data is variable – although it is reasonable to say these are the ballpark values. Issues with temperature and φ dependence also occur: temperature could be solved by considering how the structure evolves with temperature; it could also be solved by better understanding the

potential. The φ dependence problem is trickier and altogether, suggests a more fundamental problem with the model or potential.

Therefore, I simulated monodisperse spheres following the Mie potential, to understand if the issue was with the model or the potential.

5 Computational Method

5.1 Choosing a Method

The first question was what to simulate – the potential or the shear modulus directly. Ideally we would simulate the potential, since simulating the shear modulus following (3) is a multiscale modelling problem (linking microscopic scales to macroscopic properties), introducing difficult approximations [2]. However, we cannot model the true potential. First principles quantum mechanics simulations, only using the structures are unfeasible due to the size of the calcium core (~250 atoms) and the solvent – casein interactions, so simulation takes too long. Semi-empirical and theory based models aren't usable, as they haven't been solved for colloidal gels [39]**.**

Simulating the shear modulus is the only option. I simulated a monodisperse particle system – which colloidal gels tend to - following the Mie potential (4), letting us study if the problem is with (2) or the Mie Potential. We must choose between Monte Carlo simulations, which propose movements which are then accepted or rejected according to a probability distribution, or molecular dynamics simulations which integrate Newton's equations of motion to get particle trajectories. Monte Carlo – where particles steps are proposed following a probability distribution - simulations are better for larger systems and macroscopic properties [39, 40]. Even so, for the shear modulus of colloidal systems, molecular dynamics is traditionally chosen [41, 42]. This is because with periodic boundary conditions [42], we can simulate materials of infinite size, removing the size advantage of Monte Carlo, and we don't have to calculate how applying strain affects the probability. And since we are looking at static, not dynamic shear modulus, timescales aren't an issue for the instantaneous elastic stress response to strain. This is possible in python, which I used due to my familiarity existing guides.

I simulated Young's Modulus instead of shear modulus, as it's simpler and faster to apply a constant strain across the material instead of the variable strain of shear modulus, and they are linked using Poisson's ratio. There are 3 stages to simulation: initialisation, equilibration, and sampling [40].

The final choice is choosing an ensemble of what is held constant. I chose NVE, or constant particles, volume and energy, as this is computationally the simplest.

5.2 Description of Method

5.2.1 Initialisation

To simply the simulation, reduced units were used. $Length = \sigma$ and $Energy = \varepsilon$ these cancel in (4), and:

$$Time = \sigma\sqrt{\frac{m}{\varepsilon}} \quad (9)$$

$$Pressure\ (Modulus) = \frac{\varepsilon}{\sigma^3} \quad (10)$$

Where m is the mass of one colloid. The rescaled time in (9) is key, since timesteps (Δt) should be less than 1 in natural units [43].

Firstly, a cube size V = L^3 is made, and random distribution of initial particles is generated using the Poisson disc algorithm. The number of initial particles and the volume are set by the volume fraction of the colloids, where the volume of each colloid is determined using R_{HS}. The particle number was held roughly constant (at 59 +/-4, the variation is useful to test how sensitive the simulation is to variations in particle number), so if the volume fraction decreased, the box size had to increase. The Poisson disc algorithm was used to randomly generate initial positions of particles, whilst ensuring particles are a set distance away from each other. The calcium cores weren't allowed to interact. The code for this algorithm was taken from [44].

Periodic boundary conditions were applied here. If particles went outside the cell, they were shifted as in figure 12. This is reasonable if treating this cube as an identical unit cell in a larger material.

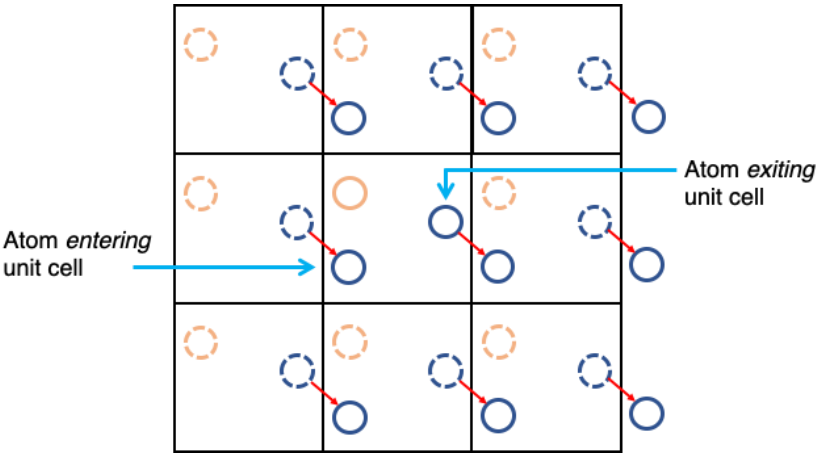

Figure 12: Diagram of periodic boundary conditions. The blue atom leaves cell so is shifted to the other side [45].

However, this means we need to account for the effects of neighbouring unit cells. This is simulated using the minimum image convention [45]:

1. 26 'image cells', identical to the original are made and placed around the original as in figure 13.

2. Forces are calculated for all particles, with particles in the original cell influenced by the images.
3. Particles in the original cell step forward in time – outer cell particles aren't influenced by neighbouring cells.
4. Repeat for each timestep

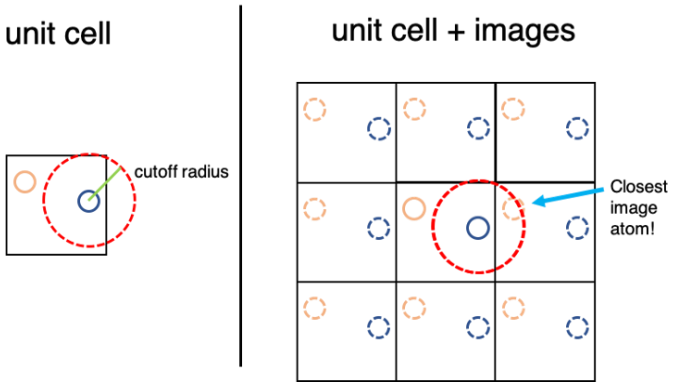

Figure 13: Diagram of unit cell and images. Demonstration of cutoff radius [45]

This only works if the cell size is $\sim 2\sigma$, so particles cannot interact with their images – in my simulations, the minimum box size was 5, so this wasn't an issue.

5.2.2 Equilibration

The particles are then free to move from their original positions, following Newton's 2nd Law (11) until they reach equilibrium.

$$m\ddot{r} = -\frac{dU}{dr} \qquad (11)$$

To step forward particles in time following (11), two methods were tested: the Euler algorithm and the velocity Verlet algorithm.

Euler:

1. Update position vector using current velocity: $\boldsymbol{r}(t + \Delta t) = \boldsymbol{r}(t) + \Delta t \boldsymbol{v}(t)$
2. Update velocity vector using calculated force: $\boldsymbol{v}(t + \Delta t) = \boldsymbol{v}(\boldsymbol{t}) + \Delta t \boldsymbol{F}(t)$. This is used to update the position for the next time step. Mass is absorbed into time, so cancels from (10)

Velocity Verlet:

1. $\boldsymbol{v}\left(t + \frac{1}{2}\Delta t\right) = \boldsymbol{v}(t) + \frac{1}{2}\boldsymbol{F}(t)\Delta t$
2. $\boldsymbol{r}(t + \Delta t) = \boldsymbol{r}(t) + \boldsymbol{v}\left(t + \frac{1}{2}\Delta t\right)\Delta t$
3. Update forces using position update.
4. $\boldsymbol{v}(t + \Delta t) = \boldsymbol{v}\left(t + \frac{1}{2}\Delta t\right) + \frac{1}{2}\boldsymbol{F}(t + \Delta t)\Delta t$

The velocity Verlet estimates the velocity at half a time step using Euler, which is used to update the force midway through the step, resulting in an improved velocity estimation. Both algorithms were investigated [46].

I stopped the algorithms when equilibrium was reached. I tracked stress until percentage changes in stress fell below 1×10^{-6}, which by trial and error, was a reasonable stopping value. I chose stress since it is required for young's modulus, so controlling its fluctuations is key. Macroscopic stress was calculated by summing the interparticle stress tensor across each particle:

$$\tau^{ab}(\mathbf{r}) = -\left\langle \sum_{i=1}^{N} \left(m v_i^a v_i^b + \frac{1}{2} \sum_{j(\neq i)=1}^{N} f_{ij}^a \mathcal{F}_{ij}^b(\boldsymbol{r}_{ij}) \right) \right\rangle \qquad \text{[47] (12)}$$

Where i and j are particles, and a, b are dimensions xyz. This is an average in time and space, so I made it a rolling average of the last 3 steps. We don't need to spatial average since as this is achieved by periodic boundary conditions [48]**.** It creates a 3x3 stress matrix. Since strain is applied uniformly in one direction, we're only interested in the diagonals so the sum of the diagonals is the stress scalar.

5.3.3 Sampling

Strain is characterised by strain rate $\dot{\gamma} = \frac{v(t)}{L}$, where the material is stretched by velocity v in a direction – I chose the x direction. This means the new update equation is:

$$\boldsymbol{v}(t + \Delta t) \rightarrow \boldsymbol{v}(t + \Delta t) + \dot{\gamma} x(t) \qquad (13)$$

where strain is applied in the x direction. $\dot{\gamma}$ is multiplied by x, not L so the particles move further apart, causing stress [41, 42]. The number of steps where strain was applied was chosen to get enough of the stress-strain curve and depended on the step size.

5.3 Implementation

To speed up the code as much as possible, I used of array vectorization and Einstein sums (for 1) instead of for loops to calculate forces, distances and stress between particles. This made calculating the Force vectors 8 times faster for 59 particles**.** To calculate r from the xyz distance vectors, square roots were slow when working with large arrays. Instead I used the function scipy.spatial.distance.pdist, which doubles the speed (from 0.002 to 0.001) for 59 particles.

The main choice was between Euler and velocity Verlet. Euler converges well (figure 14) but has stability issues – if the timestep is too big, it broke down and started oscillating around the solution (figure 15).

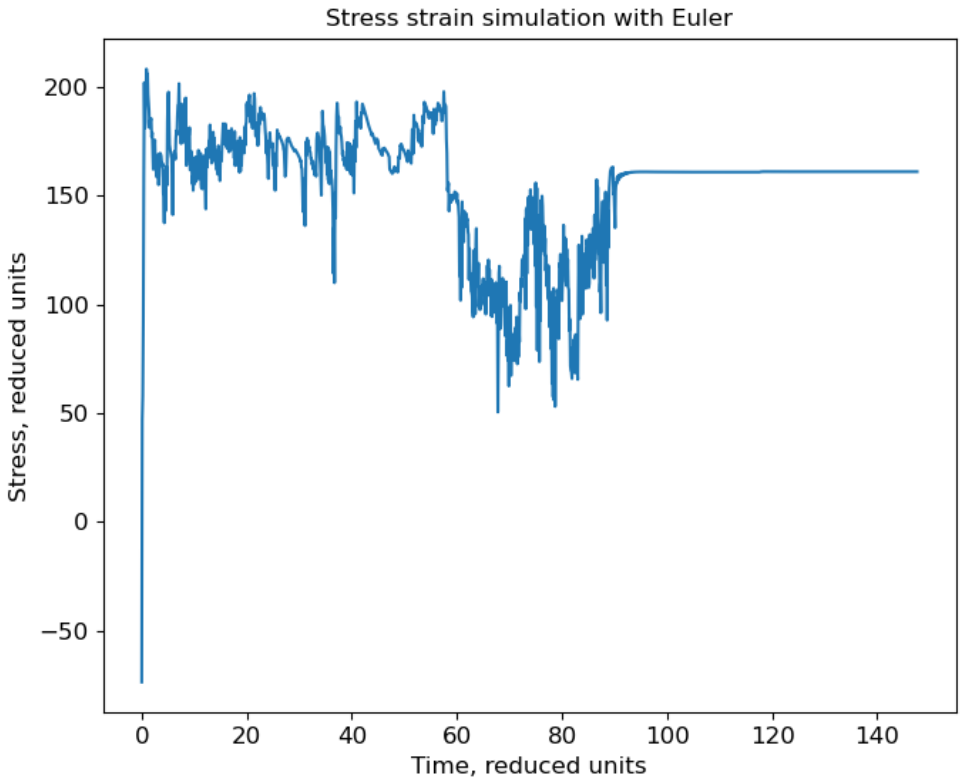

Figure 14: Stress against time for Euler step. Reaches stable equilibrium

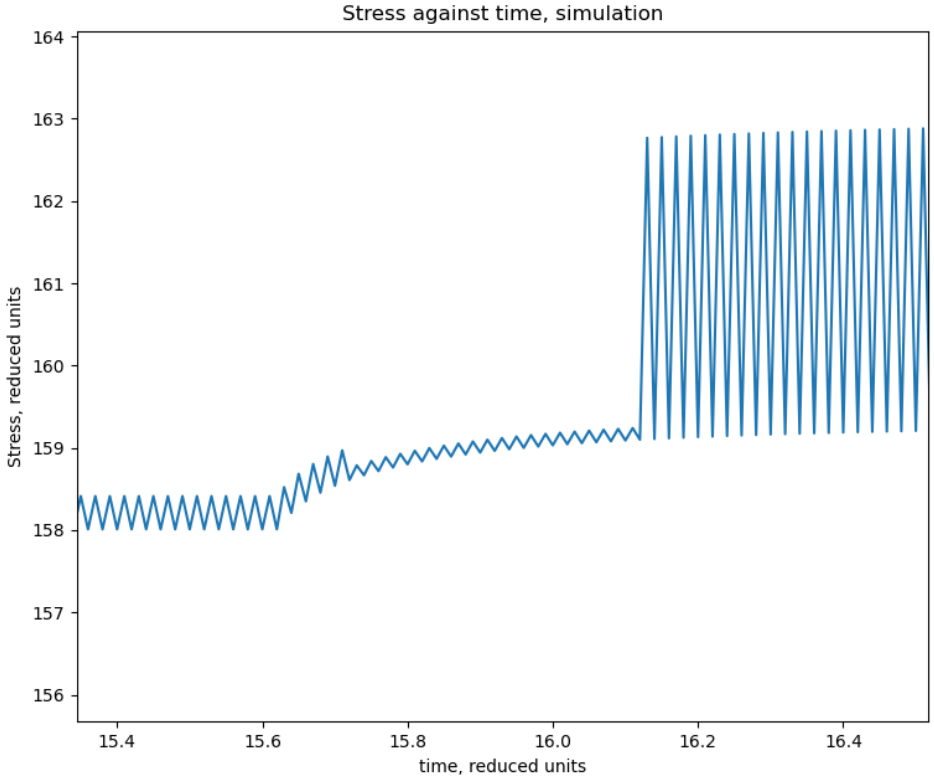

Figure 15: Breakdown of Euler step for larger timesteps.

It also doesn't follow the NVE ensemble – total energy increases with time (figure 16). Meanwhile, velocity Verlet was much better at conserving energy, (figure 17), and worked with greater timesteps, converging even with $\Delta t = 1$, whereas Euler needed at least 0.2.

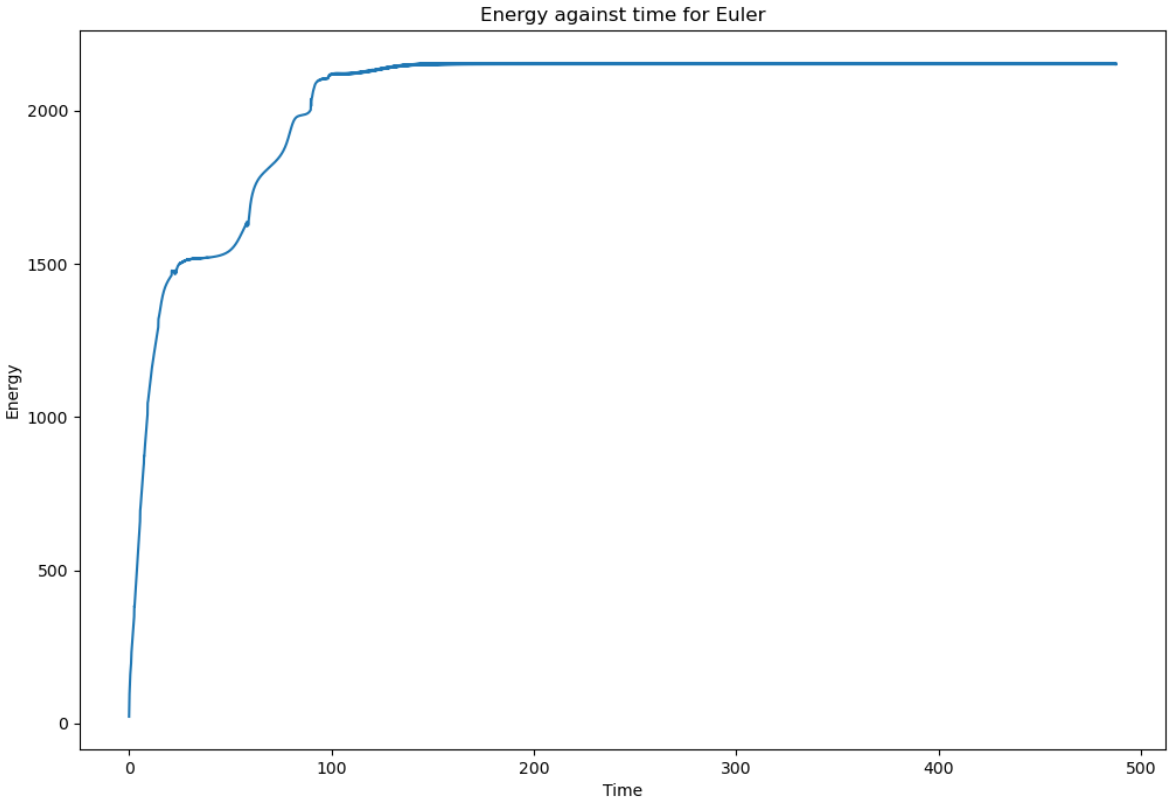

Figure 16: Energy against time. Not constant but stabilises so eventually follows NVE.

However, I used Euler. Figures 17 and 18 also show Verlet is much more unstable, and its equilibrium is much noisy than Euler's equilibrium, making stress-strain curves unusable. Furthermore, its speed advantage isn't used, since the large timesteps cause particles to leave travel outside the 2L wide image box, which is too far so cannot be used. This places an upper limit on the on Δt of 0.5. It also takes much longer to reach equilibrium for the same Δt – 5000 steps vs ~1000. Finally, the lack of energy conservation for Euler is only a problem if the system cannot reach stable equilibrium - but ours can, so its usable. See figure 19 for an equilibrium breakdown example

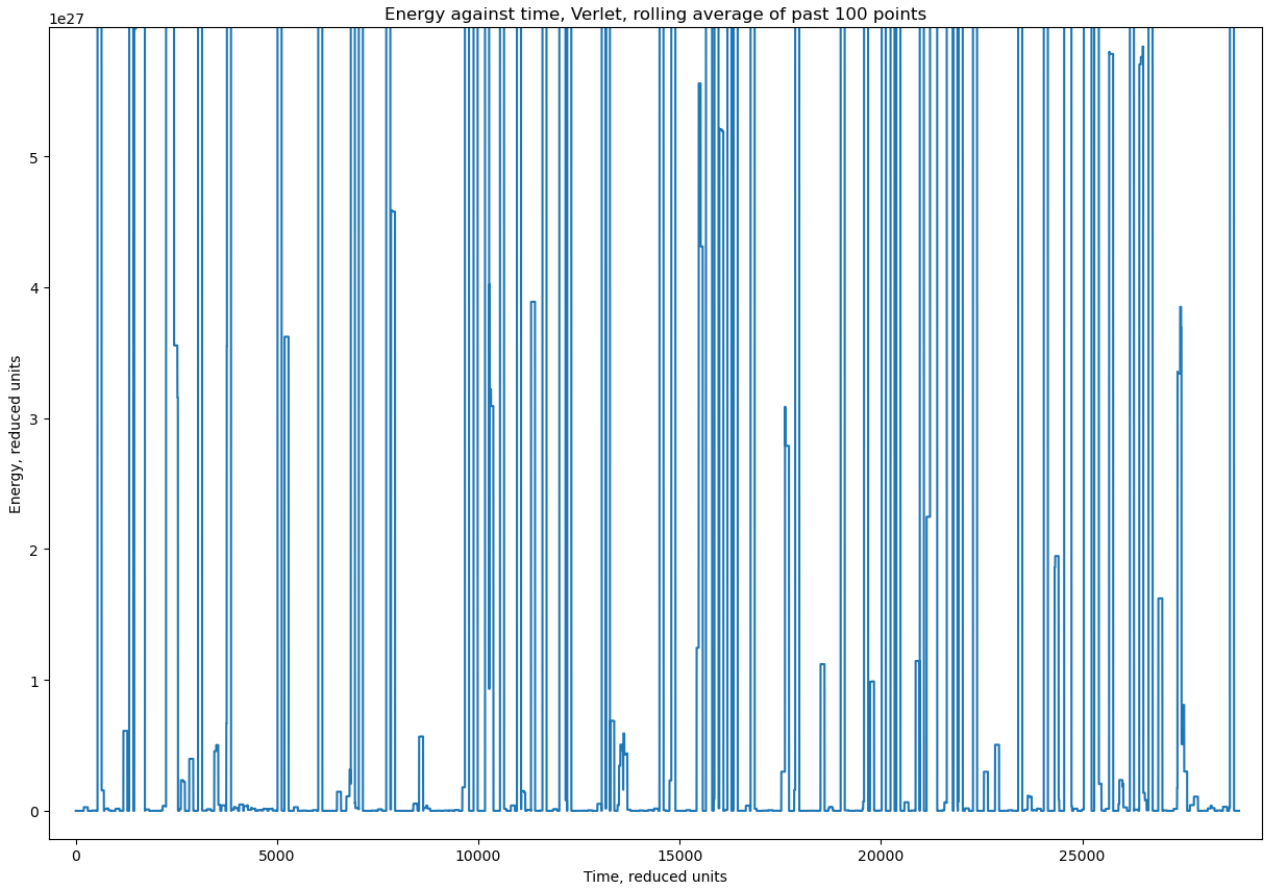

Figure 17: Verlet energy against time. Energy value doesn't rise, but is very unstable, even when a rolling average is applied.

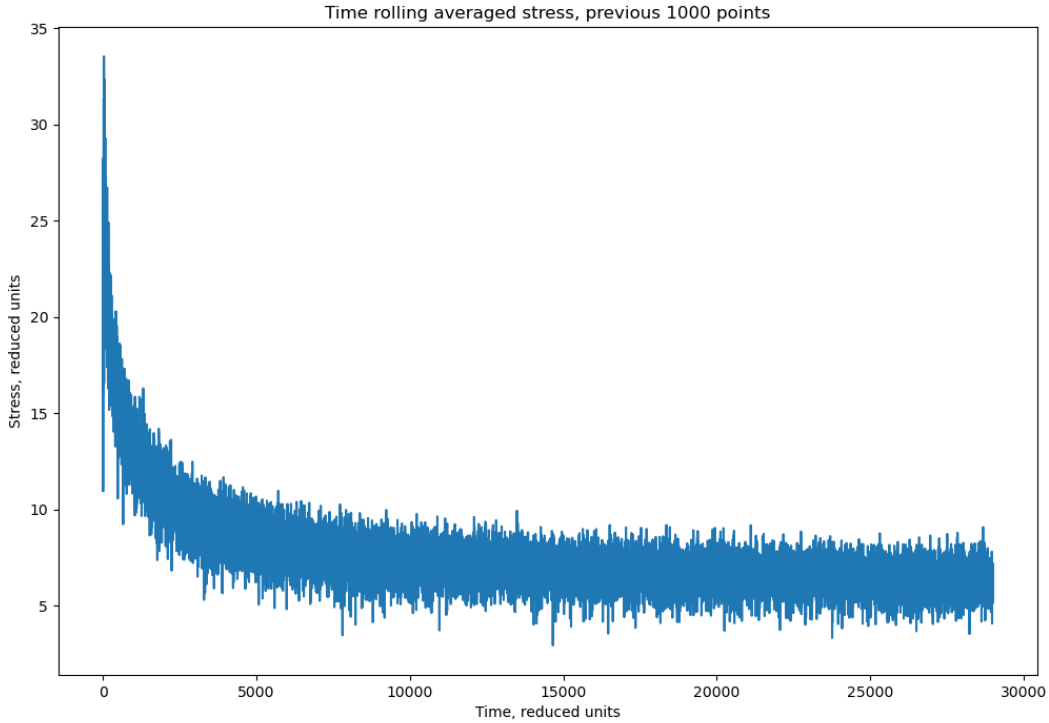

Figure 18: Verlet stress against time. Reaches equilibrium, but very unstable and slow.

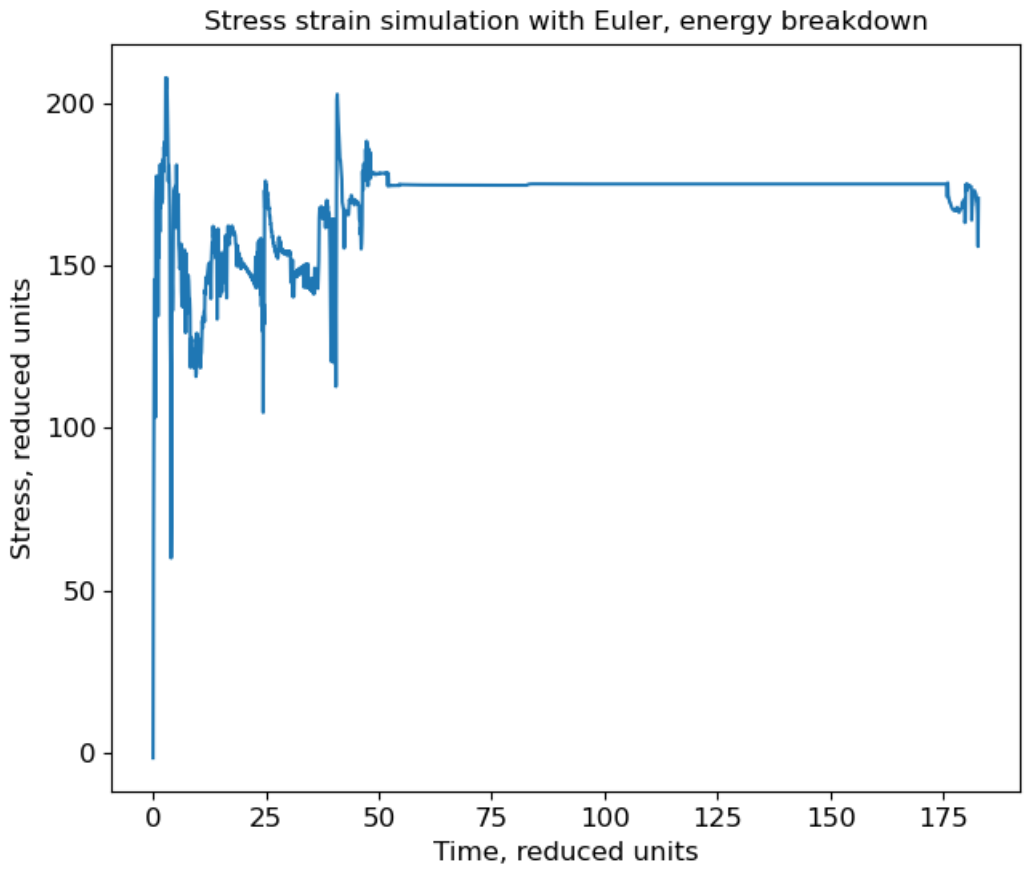

Figure 19: Euler stress against time. Breakdown after long equilibrium, due to lack of energy preservation.

$\Delta t = 0.1$ was chosen as the time step, balancing speed and stability. Large arrays caused memory limitations – the most particles I couldn't simulate above 800 particles as my laptop ran out of memory to store the interparticle force vectors, and anything above 100 was very slow. However this wasn't an issue – the minimum image convention means we're simulating 27 times the specified particles, so small numbers still work well. 59 particles balanced speed – each run took 25 minutes – with stability, as the artifacts caused that occur with small numbers of particles (figure 20) aren't seen.

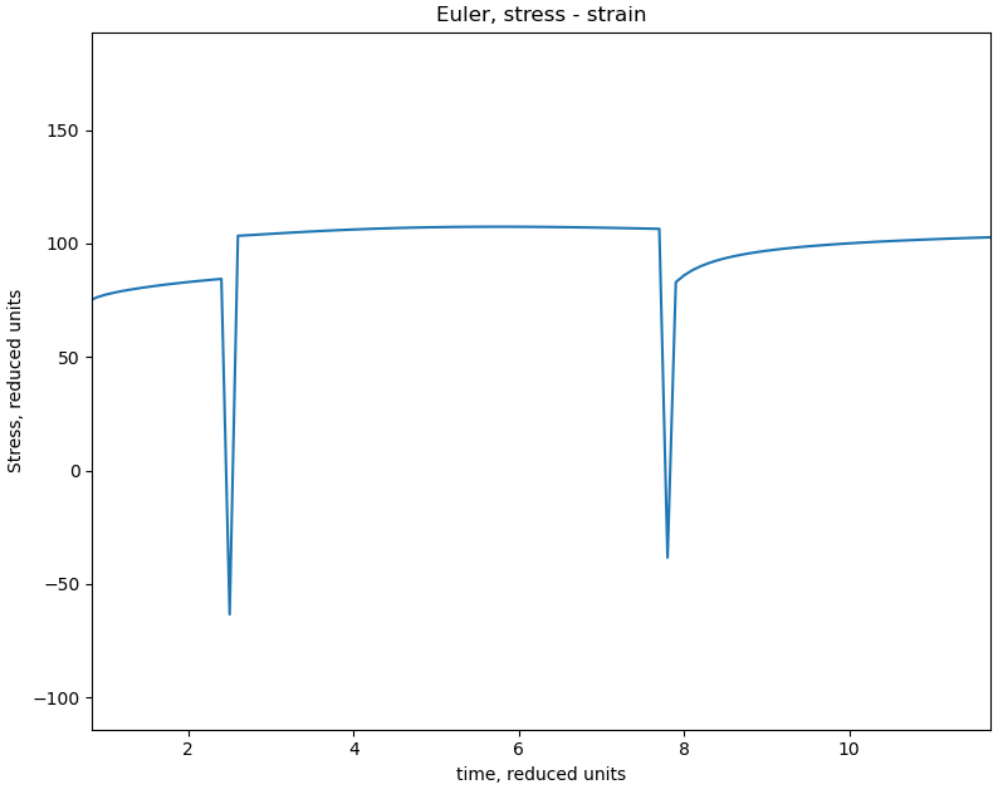

Figure 20: Simulation with only 8 particles. Equilibrium broke by sudden spikes as particles leave cell.

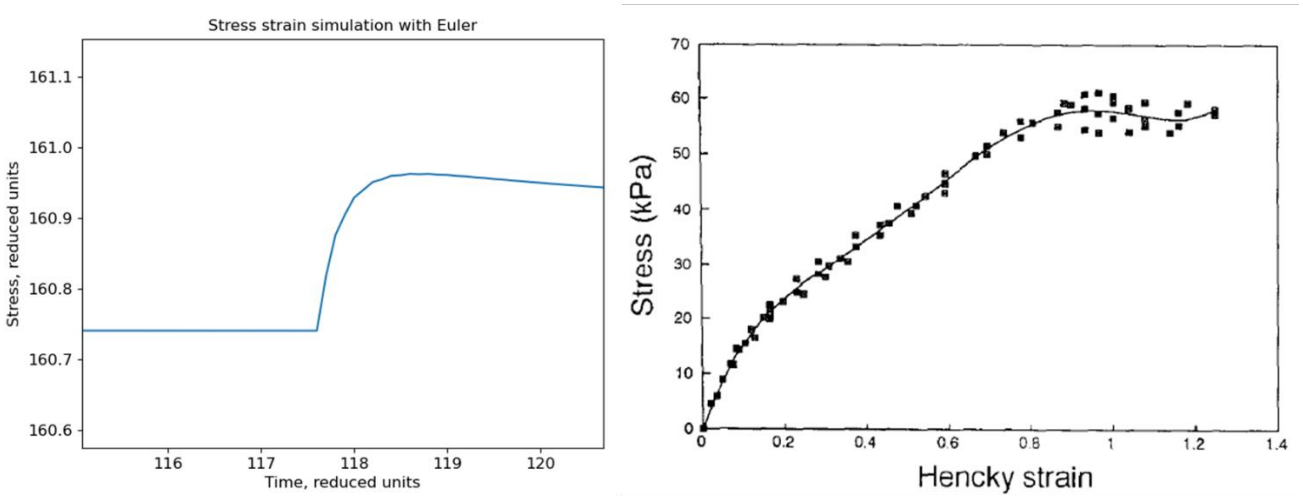

Figure 21: (Left) our simulation stress – strain curve. (Right) stress – strain curve for cheddar at a constant strain rate. Both stabilise.

When equilibration was reached, a strain was applied affecting stress. The shape of the stress-strain curve was correct, showing that the simulation was working as in figure 21.

We want to calculation the elastic, quasistatic modulus without deformation – this means we took the initial slope as the modulus. This is difficult in figure 21, so to get more points without sacrificing speed, $\Delta t = 0.01$ after equilibration. $\dot{\gamma} = 0.01$, a value which gave a large stress/strain curve without causing breakdowns.

6 Computational Results

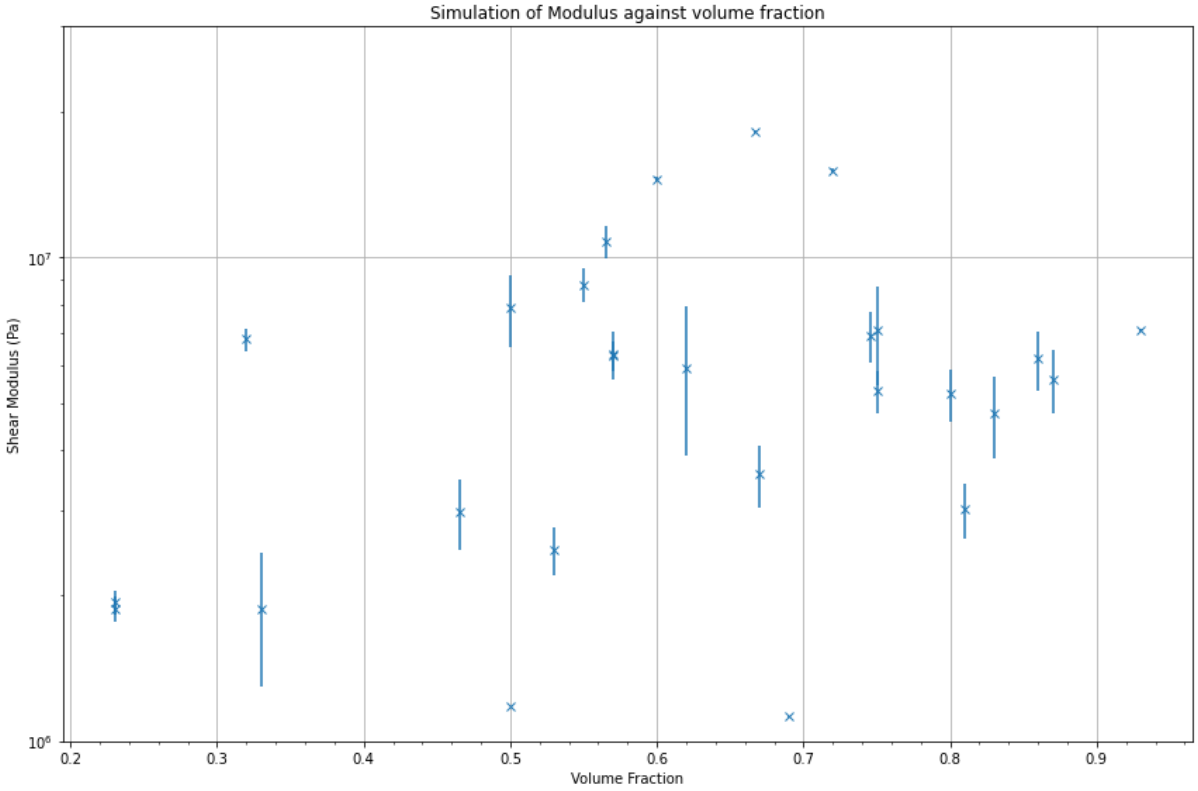

Figure 22: Volume fraction against shear modulus results.

Figure 22 shows the results of the simulation, which have been converted to shear modulus with Poisson's ratio. This is a monodisperse particle simulation – not a pure colloidal gel – so we'd only expect the results to match up with the data or (2) at high volume fractions. Even then, it is poor. The values are much greater - 6×10^6 instead of $\sim 1 \times 10^6$, although this could be fixed by altering σ and ε in (10). If $\varphi \geq 0.8$, we can see a linear trend, which could suggest close packing is higher than predicted, and our data sample is too small. There is also a general increase in modulus with higher volume fractions as expected.

However, this ignores the large variations. Errors in the model come from multiple places. Changes in the particle number and the Euler step have been discussed and are unlikely. However, the failure of the superior Verlet algorithm to stabilise suggests deeper problems – and Euler gets stuck in different, stable, but not equilibrium states, which is a problem with molecular dynamics [40]. Verlet does reach equilibrium, it's just very unstable, perhaps due to particles leaving the cell which was common for Verlet, causing issues such as in figure 20, which could be fixed with more particles.

Another issue could be with the ensemble used, NVE. When strain is applied, the energy of the system is not constant. However, colloid modulus simulations work in NVE [42, 43], so it's unclear why this would be an issue – and the Euler step has shown it adapts.

Even so, the fact we can see variation for $\varphi \leq \varphi_{CP}$ – it's generally smaller - confirms that there is an issue with the φ of (2), and the temperature scaling (through (9)) is unchanged,

showing the problem isn't simply the potential, but the model itself (although the temperature effects could change if an NVT ensemble is implemented). Perhaps for high volume fractions, more complicated polymer – colloid models would be ideal, as the proteins get pushed together, potentially making the soft colloid model unsuitable [49].

7 Computational and Final Conclusions

Building on the modelling conclusions from section 4, which accepted Gillies colloidal gel model and the quantitative relationship for $\varphi \leq \varphi_{CP}$, but not for $\varphi \geq \varphi_{CP}$, although a likely potential, with n = 10 and m = 9, was found – different to Gilles predictions. The $\varphi \leq \varphi_{CP}$ model can be improved with better predictions for b, which can be simulated [41]. The monodisperse sphere simulation following the Mie Potential with n = 6, m = 3 cannot confirm the form of the potential or fix temperature dependence. However, it can confirm φ dependence is missing for small φ from the high-volume fraction model, which approximates colloidal gels as monodisperse spheres.

The colloidal gel model is a promising first principles approach to model cheese. Future research should focus on correctly understanding temperature and volume fraction dependence by considering structural changes away from the monodisperse sphere approximation. Simulations can be improved by switching to the NVT equilibrium, giving guidance for temperature dependence, and they should be tested with more particles and the velocity Verlet algorithm to see if variation can be reduced. If the monodisperse sphere approximation is correct at high volume fractions, these simulations combined with experiment can confirm the form of the potential. Once this foundation is formed, other variables, such as fat, can be added, and manufacturing processes modelled.

Acknowledgements

I wish to thank my project partner, my supervisor William Proud, other students taking projects with him, and Graeme Gillies who kindly helped me understand his model.

References

[1] Ho QT, Carmeliet J, Datta AK, Defraeye T, Delele MA, Herremans E, Opara L, Ramon H, Tijskens E, Van Der Sman R, Van Liedekerke P. Multiscale modeling in food engineering. Journal of food Engineering. 2013 Feb 1;114(3):279-91. Available at: https://doi.org/10.1016/j.jfoodeng.2012.08.019

[2] Datta AK. Status of physics-based models in the design of food products, processes, and equipment. Comprehensive Reviews in Food Science and Food Safety. 2008 Jan;7(1):121-9. Available at: https://doi.org/10.1111/j.1541-4337.2007.00030.x

[3] Perrot N, Trelea IC, Baudrit C, Trystram G, Bourgine P. Modelling and analysis of complex food systems: state of the art and new trends. Trends in Food Science & Technology. 2011 Jun 1;22(6):304-14. Available at: https://doi.org/10.1016/j.tifs.2011.03.008

[4] Everett DW, Auty MA. Cheese structure and current methods of analysis. International Dairy Journal. 2008 Jul 1;18(7):759-73. Available at: https://doi.org/10.1016/j.idairyj.2008.03.012

[5] Bintsis T, Papademas P. An overview of the cheesemaking process. Global cheesemaking technology: Cheese quality and characteristics. 2017 Nov 29:120-56. Available at: https://doi.org/10.1002/9781119046165.ch0f

[6] Al-Otaibi MM, Wilbey RA. Effect of temperature and salt on the maturation of white-salted cheese. International journal of dairy technology. 2004 Feb;57(1):57-63. Available at: https://doi.org/10.1111/j.1471-0307.2004.00123.x

[7] Ah J, Tagalpallewar GP. Functional properties of Mozzarella cheese for its end use application. Journal of food science and technology. 2017 Nov;54(12):3766-78. Available at: https://doi.org/10.1007/s13197-017-2886-z

[8] Guinee TP. Protein in cheese and cheese products: Structure-function relationships. InAdvanced dairy chemistry 2016 (pp. 347-415). Springer, New York, NY.

[9] Gunasekaran S, Ak MM. Cheese rheology and texture. CRC press; 2002 Dec 23.

[10] Lucey JA, Johnson ME, Horne DS. Invited review: perspectives on the basis of the rheology and texture properties of cheese. Journal of Dairy Science. 2003 Sep 1;86(9):2725-43. Available at: https://doi.org/10.3168/jds.S0022-0302(03)73869-7

[11] Prentice JH, Langley KR, Marshall RJ. Cheese rheology. InCheese: chemistry, physics and microbiology 1993 (pp. 303-340). Springer, Boston, MA.

[12] Gillies G. Predictions of the shear modulus of cheese, a soft matter approach. Applied Rheology. 2019 Jan 1;29(1):58-68. Available at: https://doi.org/10.1515/arh-2019-0006

[13] Yang XI, Rogers NR, Berry TK, Foegeding EA. Modeling the rheological properties of cheddar cheese with different fat contents at various temperatures. Journal of Texture Studies. 2011 Oct;42(5):331-48. Available at: https://doi.org/10.1111/j.1745-4603.2011.00283.x

[14] Porter D. Group interaction modelling of polymer properties. CRC Press; 1995 Feb 8.

[15] Timmen H, Patton S. Milk fat globules: fatty acid composition, size and in vivo regulation of fat liquidity. Lipids. 1988 Jul;23(7):685-9. Available at: https://doi.org/10.1007/BF02535669

[16] Horne DS. Casein interactions: casting light on the black boxes, the structure in dairy products. International Dairy Journal. 1998 Mar 1;8(3):171-7. Available at: https://doi.org/10.1016/S0958-6946(98)00040-5

[17] Genovese DB, Lozano JE, Rao MA. The rheology of colloidal and noncolloidal food dispersions. Journal of Food Science. 2007 Mar;72(2):R11-20. Available at: https://doi.org/10.1111/j.1750-3841.2006.00253.x

[18] Goff, H. Douglas, et al. *Dairy Science and Technology EBook*, University of Guelph, Available at: https://books.lib.uoguelph.ca/dairyscienceandtechnologyebook/.

[19] Manu, Super. "Micelles." *Wikimedia Commons,* Available at: https://commons.wikimedia.org/wiki/File:Micelle_scheme-en.svg.

[20] Choi J, Horne DS, Lucey JA. Determination of molecular weight of a purified fraction of colloidal calcium phosphate derived from the casein micelles of bovine milk. Journal of dairy science. 2011 Jul 1;94(7):3250-61. Available at: https://doi.org/10.3168/jds.2010-3762

[21] Evans ID, Lips A. Concentration dependence of the linear elastic behaviour of model microgel dispersions. Journal of the Chemical Society, Faraday Transactions. 1990 Jan 1;86(20):3413-7. Available at: https://doi.org/10.1039/FT9908603413

[22] Krall AH, Weitz DA. Internal dynamics and elasticity of fractal colloidal gels. Physical review letters. 1998 Jan 26;80(4):778. Available at: https://doi.org/10.1103/PhysRevLett.80.778

[23] Asnaghi D, Carpineti M, Giglio M, Sozzi M. Coagulation kinetics and aggregate morphology in the intermediate regimes between diffusion-limited and reaction-limited cluster aggregation. Physical Review A. 1992 Jan 1;45(2):1018. Available at: https://doi.org/10.1103/PhysRevA.45.1018

[24] Joshi NS, Muthukumarappan K, Dave RI. Viscoelastic properties of part skim Mozzarella cheese: Effect of calcium, storage, and test temperature. International Journal of Food Properties. 2004 Dec 31;7(2):239-52. Available at: https://doi.org/10.1081/JFP-120026060

[25] Kondyli E, Pappa EC, Vlachou AM. Effect of package type on the composition and volatile compounds of Feta cheese. Small Ruminant Research. 2012 Nov 1;108(1-3):95-101. Available at: https://doi.org/10.1016/j.smallrumres.2012.06.014

[26] Weeks JD, Chandler D, Andersen HC. Role of repulsive forces in determining the equilibrium structure of simple liquids. The Journal of chemical physics. 1971 Jun 15;54(12):5237-47. Available at: https://doi.org/10.1063/1.1674820

[27] Van Hekken DL, Tunick MH, Park YW. Effect of frozen storage on the proteolytic and rheological properties of soft caprine milk cheese. Journal of Dairy Science. 2005 Jun 1;88(6):1966-72. Available at: https://doi.org/10.3168/jds.S0022-0302(05)72872-1

[28] Wium H, Qvist KB. Prediction of sensory texture of feta cheese made from ultrafiltered milk by uniaxial compression and shear testing. Journal of texture studies. 1998 May;29(2):215-32. https://doi.org/10.1111/j.1745-4603.1998.tb00165.x

[29] Dairy UK. Code of practice on compositional standards for UK named ... - dairy UK [Internet]. Cheese compositional standards UK. Dairy UK; 2018 [cited 2023Jan17].

Available from: https://www.dairyuk.org/wp-content/uploads/2018/12/FINAL-Compositional-Standards-Cheese.pdf

[30] Tunick MH. Activation energy measurements in rheological analysis of cheese. International dairy journal. 2010 Oct 1;20(10):680-5. Available at: https://doi.org/10.1016/j.idairyj.2010.03.010

[31] Luyten H, Van Vliet T, Walstra P. Characterization of the consistency of Gouda cheese: Rheological properties. Nederlands melk en Zuiveltijdschrift. 1991;45(1):33-53.

[32] Ak MM, Gunasekaran S. Stress-strain curve analysis of Cheddar cheese under uniaxial compression. Journal of Food Science. 1992 Sep;57(5):1078-81. Available at: https://doi.org/10.1111/j.1365-2621.1992.tb11268.x

[33] Johnston DE, Darcy PC. The effects of high pressure treatment on immature Mozzarella cheese. Milchwissenschaft. 2000;55(11):617-20.

[34] Gunasekaran S., Mehmet Ak M., Cheese Rheology and Texture. (CRC press 2002).

[35] Toth P. Optimization engineering techniques for the exact solution of NP-hard combinatorial optimization problems. European journal of operational research. 2000 Sep 1;125(2):222-38. Available at: https://doi.org/10.1016/S0377-2217(99)00453-1

[36] Zwanzig R, Mountain RD. High-frequency elastic moduli of simple fluids. The Journal of Chemical Physics. 1965 Dec 15;43(12):4464-71. Available at: https://doi.org/10.1063/1.1696718

[37] Rueb CJ, Zukoski CF. Viscoelastic properties of colloidal gels. Journal of Rheology. 1997 Mar;41(2):207. Available at: https://doi.org/10.1122/1.550812

[38] Xue L, Borodin O, Smith GD, Nairn J. Micromechanics simulations of the viscoelastic properties of PBX-9501 by material point method. InThe 2006 Annual Meeting 2006.

[39] Porter D. Group interaction modelling of polymer properties. CRC Press; 1995 Feb 8.

[40] Gartner III TE, Jayaraman A. Modeling and simulations of polymers: a roadmap. Macromolecules. 2019 Jan 22;52(3):755-86.

[41] WHITTLE BM, Dickinson E. Brownian dynamics simulation of gelation in soft sphere systems with irreversible bond formation. Molecular physics. 1997 Apr 1;90(5):739-58. Available at: https://doi.org/10.1080/002689797172101

[42] Dickinson E., Patino, R J.M, Food emulsions and foams: interfaces, interactions and stability. Royal Society of Chemistry; 1999. p344

[43] Richards M., Dancy P., Basics of Numerical Programming. 2022, unpublished. p.18

[44] IHautal, Poisson disc, (2015), GitHub repository, https://github.com/IHautaI/poisson-disc

[45] Mcelfresh C. Molecular dynamics: Periodic boundary conditions [Internet]. Medium. Python in Plain English; 2021 [cited 2023Jan17]. Available from:

https://python.plainenglish.io/molecular-dynamics-periodic-boundary-conditions-21f957bbb294

[46] Verlet integration [Internet]. Verlet Integration · Arcane Algorithm Archive. [cited 2023Jan17]. Available from: https://www.algorithm-archive.org/contents/verlet_integration/verlet_integration.html

[47] Morante S, Rossi GC, Testa M. The stress tensor of a molecular system: An exercise in statistical mechanics. The Journal of chemical physics. 2006 Jul 21;125(3):034101. Available at: https://doi.org/10.1063/1.2214719

[48] Smith ER. Calculating the pressure in simulations using periodic boundary conditions. Journal of Statistical Physics. 1994 Oct;77(1):449-72. Available at: https://doi.org/10.1007/BF02186852

[49] Gibson, Jonathan B. et al. Simulation of Colloid-Polymer Systems using Dissipative Particle Dynamics. *Molecular Simulation* 23 1999;23:1-41.