Quantitative Analysis and Statistical Mechanics of Granular Pack Structures

Joseph Francis Patrick Jordan

A thesis submitted to

Imperial College London

presented for the degree of

Doctor of Philosophy

Department of Earth Science and Engineering

December 2014

Declaration of Originality

I declare that the work presented in this thesis is original and my own, except where the work of others is explicitly referenced.

Copyright Notice

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Abstract

Disordered granular packs present many challenges over regular structures in description, analysis and modelling. Within the granular statistical mechanics, which aims to compute bulk properties from details of the local structure, the quadron description was defined over a decade ago. This quantitative characterisation of disordered systems enumerates the structural degrees of freedom, by dividing the system into tessellating volume elements, quadrons, and assigning each a quantitative structure tensor computed from its shape. Presented in this work is a theoretical derivation of an equipartition principle of the volume, which is the analogue of the well known $\frac{3}{2}k_BT$ in thermal physics. This computation is included in full, for two-dimensional systems, using the quadron description. Following this, quadron tensors and statistics in three dimensional disordered structures are computed and analysed for the first time. A new computer program to achieve this was developed, requiring a new class of solution for cell/pore identification, a custom 3D rendering engine, and several extensions to the quadron description as originally defined. The program was successful in analysing structures from simulations, experimental colloidal suspensions and granular systems of a few thousand grains. By a subsystem-based parallelisation method, much larger data-sets were successfully analysed, including an X-ray tomography experiment with 60,000 non-spherical grains. The quadron statistics of a triaxial shear experiment were computed at multiple stages, and the chirality, a pseudo-vector computed from the structure tensor which measures the deviation from orthogonality of the volume element boundary, develops an anisotropy during the formation of a shear band. While the software performed very well on dense granular packs, its performance in less dense systems or regions, like a shear band, was both slower and less reliable. This was traced to the increased ambiguity in

cell structure as the density decreases, and is something that future analysis could improve.

Contents

1 Introduction **17**

2 Literature Review **22**

2.1 Mathematical Methods: Graphs 22

2.1.1 Graphs and their Representations 22

2.1.2 Special Cases 25

2.1.3 Graph Structure 26

2.1.4 Analysis 29

2.2 Computational Methods 38

2.2.1 Preliminary Computational Definitions 38

2.2.2 Enumerating All Cycles 40

2.2.3 Maximum Independent Set 41

2.2.4 Contact Detection 41

2.2.5 Voronoï Tessellation 42

2.2.6 Convex Hulls 43

2.2.7 Segmentation 43

2.3 A Granular Statistical Mechanics 49

2.3.1 Historical Work on Granular Materials 49

2.3.2 Theory of Powders 52

2.3.3 Quantifying the Degrees of Freedom 54

2.4 Experimental Techniques in Granular Physics 57

2.4.1 Two Dimensional Experiments 57
2.4.2 Three Dimensional Experiments 58
2.4.3 X-ray Computed Tomography 58
2.4.4 Colloidal Optics 62
2.4.5 Tests of the Statistical Mechanical Theory 63

3 Two-Dimensional Methods 72

3.1 Two Dimensional Quadrons 72
3.1.1 Quadron Definition 73
3.2 Equipartition of Volume in Two Dimensional Isostatic Granular Media 79
3.2.1 Derivation of a Granular Equipartition Relation 80

4 Three-Dimensional Methods 87

4.1 Preliminary Mathematical Definitions 88
4.1.1 The Granular Graph 88
4.1.2 The Contact Point Graph 89
4.2 Three Dimensional Quadrons 90
4.2.1 Ideal Quadrons in Three Dimensions 92
4.2.2 Geometrical Constraints 100
4.3 Searching for Cells 106
4.3.1 Existing Techniques 106
4.3.2 Development of the Cycle-Centric Algorithm 108
4.3.3 Overview of the Algorithm as a Whole 118

5 Results and Discussion 122

5.1 Physics Results from using the Cycle-Centric Algorithm 122
5.1.1 Distribution of Quadron Volumes 123
5.1.2 Classification of Cell Type 125

5.1.3 Conditional Quadron Volumes by Cell Type 126
5.1.4 The Collapse of the Conditional Volume Distributions . . . 129
5.1.5 Quadron Chirality 130
5.1.6 Sensitivity of Statistics to Physical Properties 133
5.1.7 Quality of Outcomes 140

6 Conclusion **144**

Appendices **158**

A Definitions **159**

B Known Integrals **161**

B.1 Gaussian Integrals 161
B.2 Dawson Functions 163
B.3 An Exact form for Definite Integrals 163

C Comprehensive Results **166**

C.1 Distributions of Quadron Volumes 166
C.2 Distributions of Quadron Chiralities 202
C.3 Zoo of Cell Types 209

D Copyrighted Works Permissions **226**

D.1 Permission from Dr. Blumenfeld 226
D.2 Permission from Dr. Andó 228
D.3 Request for permission to Facebook 229

List of Tables

2.1 The duality relations for graph embeddings in two and three dimensions. 35

2.2 A summary of the equivalent equations between statistical mechanics and Edwards' theory of powders, from Edwards et al[1] 53

5.1 Data Sources for the statistical results presented. 123

5.2 Proportion of the system volume successfully solved by the new software. 141

A.1 Mathematical definitions 159

C.1 Cell Zoo - the different cell types detected in the 3SR Caicos initial condition and the number of quadrons detected in such cells. . . . 209

List of Figures

2.1 The Facebook Graph 23

2.2 Various embeddings of graph G. 24

2.3 The complete graphs with three, four and five vertices. 25

2.4 Examples of bipartite complete graphs. 26

2.5 Various subgraphs of K_5 27

2.6 A path, P. 28

2.7 A cycle, C. 28

2.8 Example of a cycle basis. 31

2.9 Drawing K_4 to emphasise its planarity. 32

2.10 K_4 and its dual, K_4. 34

2.11 Construction of the Voronoï tessellation, and the Delaunay triangulation as its dual, in two dimensions. 37

2.12 An example of a convex hull. 43

2.13 The watershed algorithm. 46

2.14 Glantz and Hilpert's Voronoï complex 48

2.15 A schematic of a classic blurred X-ray tomography experiment. . . 59

2.16 The experimental layout of the Caicos Ooid triaxial strain experiment. 61

2.17 9th X-ray tomograph image of the Caicos Ooid triaxial strain experiment 61

2.18 Grain displacement and rotation in the Caicos Ooid triaxial strain experiment 62
2.19 Total and conditional quadron volume distributions in two dimensions 69
2.20 Different geometrical arrangements of cells with $e = 6$. 70
2.21 Collapse of the conditional volume distributions in two dimensions 71

3.1 Two dimensional grain polygon. 73
3.2 Two dimensional cell polygon. 74
3.3 Two dimensional quadron construction. 77
3.4 Two Dimensional Chirality 79

4.1 A three dimensional cell. 91
4.2 The quadron construction in three dimensions. 98
4.3 Definitions for *throats* and *throat-facets*. 101
4.4 Two dimensional centroidal normality. 103
4.5 Intersection of $\vec{\alpha}$ with a grain facet. 105
4.6 A phantom facet. 113
4.7 Intersection of two $\vec{\alpha}$s with grain facets adjoining a phantom facet. 115
4.8 Data flow within the novel cycle-centric algorithm. 119
4.9 Run-time performance of the program on different sized systems. . 120

5.1 The quadron volume probability distribution for the Caicos[2] initial condition with a $k - \Gamma$ PDF fit. 124
5.2 Conditional quadron volume frequency distributions. 128
5.3 Normalised per-cell volume probability densities collapse. 129
5.4 Isotropy of oil droplet suspension[3] quadron chiralities. 131
5.5 The difference between the chirality distributions initial and sheared Caicos systems[2]. 133
5.6 PDF (v^q) in and out of the shear band. 134

5.7 PDF $(v^q|t)$ for common cell types plotted for all the available datasets. 138
5.8 Particle size distributions. 140
5.9 Heatmap of quadron volumes projected onto the xz plane, for system COEA01-09-. 142

B.1 A definite integral, illustrated as the area under a curve between two limits. 164
B.2 A definite integral as the average function value between the limits of integration. 165

C.1 Total quadron volume frequency distribution $f(v^q)$, plotted with the seven most frequent cells' conditional frequency distributions $f(v^q|t)$, for system ABEA02-01-. 167
C.2 Seven most frequent cells' conditional probability distributions PDF $(v^q|t)$, for system ABEA02-01-. 168
C.3 Heatmap of quadron volumes projected onto the xy plane, for system ABEA02-01-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 169
C.4 Heatmap of quadron volumes projected onto the xz plane, for system ABEA02-01-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 170
C.5 Heatmap of quadron volumes projected onto the yz plane, for system ABEA02-01-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 171

C.6 Total quadron volume frequency distribution $f(v^q)$, plotted with the seven most frequent cells' conditional frequency distributions $f(v^q|t)$, for system COEA01-01-. 172

C.7 Seven most frequent cells' conditional probability distributions PDF $(v^q|t)$, for system COEA01-01-. 173

C.8 Heatmap of quadron volumes projected onto the xy plane, for system COEA01-01-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 174

C.9 Heatmap of quadron volumes projected onto the xz plane, for system COEA01-01-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 175

C.10 Heatmap of quadron volumes projected onto the yz plane, for system COEA01-01-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 176

C.11 Total quadron volume frequency distribution $f(v^q)$, plotted with the seven most frequent cells' conditional frequency distributions $f(v^q|t)$, for system COEA01-09-. 177

C.12 Seven most frequent cells' conditional probability distributions PDF $(v^q|t)$, for system COEA01-09-. 178

C.13 Heatmap of quadron volumes projected onto the xy plane, for system COEA01-09-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 179

C.14 Heatmap of quadron volumes projected onto the xz plane, for system COEA01-09-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 180

C.15 Heatmap of quadron volumes projected onto the yz plane, for system COEA01-09-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 181

C.16 Total quadron volume frequency distribution $f(v^q)$, plotted with the seven most frequent cells' conditional frequency distributions $f(v^q|t)$, for system large0. 182

C.17 Seven most frequent cells' conditional probability distributions PDF $(v^q|t)$, for system large0. 183

C.18 Heatmap of quadron volumes projected onto the xy plane, for system large0. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 184

C.19 Heatmap of quadron volumes projected onto the xz plane, for system large0. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 185

C.20 Heatmap of quadron volumes projected onto the yz plane, for system large0. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 186

C.21 Total quadron volume frequency distribution $f(v^q)$, plotted with the seven most frequent cells' conditional frequency distributions $f(v^q|t)$, for system t0. 187

C.22 Seven most frequent cells' conditional probability distributions PDF $(v^q|t)$, for system t0. 188

C.23 Heatmap of quadron volumes projected onto the xy plane, for system t0. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 189

C.24 Heatmap of quadron volumes projected onto the xz plane, for system t0. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 190

C.25 Heatmap of quadron volumes projected onto the yz plane, for system t0. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 191

C.26 Total quadron volume frequency distribution $f\,(v^q)$, plotted with the seven most frequent cells' conditional frequency distributions $f\,(v^q|t)$, for system t5. 192

C.27 Seven most frequent cells' conditional probability distributions PDF $(v^q|t)$, for system t5. 193

C.28 Heatmap of quadron volumes projected onto the xy plane, for system t5. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 194

C.29 Heatmap of quadron volumes projected onto the xz plane, for system t5. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 195

C.30 Heatmap of quadron volumes projected onto the yz plane, for system t5. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 196

C.31 Total quadron volume frequency distribution $f\left(v^q\right)$, plotted with the seven most frequent cells' conditional frequency distributions $f\left(v^q|t\right)$, for system t10. 197

C.32 Seven most frequent cells' conditional probability distributions PDF $\left(v^q|t\right)$, for system t10. 198

C.33 Heatmap of quadron volumes projected onto the xy plane, for system t10. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 199

C.34 Heatmap of quadron volumes projected onto the xz plane, for system t10. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 200

C.35 Heatmap of quadron volumes projected onto the yz plane, for system t10. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. 201

C.36 Normalised quadron chirality components – where the pseudo-vector has been given length 1 and each component plotted as a probability distribution, for system ABEA02-01-. 202

C.37 Normalised quadron chirality components – where the pseudo-vector has been given length 1 and each component plotted as a probability distribution, for system COEA01-01-. 203

C.38 Normalised quadron chirality components – where the pseudo-vector has been given length 1 and each component plotted as a probability distribution, for system COEA01-09-. 204
C.39 Normalised quadron chirality components – where the pseudo-vector has been given length 1 and each component plotted as a probability distribution, for system large0. 205
C.40 Normalised quadron chirality components – where the pseudo-vector has been given length 1 and each component plotted as a probability distribution, for system t0. 206
C.41 Normalised quadron chirality components – where the pseudo-vector has been given length 1 and each component plotted as a probability distribution, for system t5. 207
C.42 Normalised quadron chirality components – where the pseudo-vector has been given length 1 and each component plotted as a probability distribution, for system t10. 208

Chapter 1

Introduction

Granular materials make up a large part of the human environment, and yet they aren't completely understood. When a building is constructed on a granular material, as most of them are, the workers can be reasonably sure that the grains won't rearrange under them and swallow their work whole. However, if granular packs are stressed hard enough, they yield in short bursts which are the result of the universe solving thousands of tiny equations for us and finding a more stable arrangement to support the external force. Then, if the container is turned on its head, the grains will flow out in a manner not dissimilar to a liquid. While many granular materials can be modelled using approximations, we lack an empirically verified fundamental theory of granular behaviour – an aim which much of the work both before in this field and in this thesis contributes to.

Granular materials can be represented analytically using graphs — mathematical objects which are made up of vertices connected by edges (like the grains connected by contacts.) This work makes extensive use of terms from graph theory, which are defined in a preliminary section to the Literature Review.

This work is a computational advance in the description of the microstructure of

granular materials, that is the details of how the particles are arranged, how they came to be there, and how that affects their behaviour. In order to set the scene for communicating this work, the next section of the Literature Review contains a summary of the relevant computer science — technical definitions and a detailed description of some algorithms.

The third section of the Literature Review can then focus on the relevant engineering and physics research into the behaviour and description of granular matter. It begins with a historical overview of attempts to describe and predict behaviour in granular systems, notably the work in the fields of soil mechanics and soil structure.

Many ground-breaking experiments which advanced the field of study in their time are covered, including: Reynolds who in 1885 showed that granular materials counter-intuitively dilate upon compression[4]; Smith who in 1929 first measured precisely the number of contacts between grains in a three dimensional system[5]; Murayama and Matsuoka who in 1970 were the first to use birefringence to observe the stress state of a two dimensional granular system directly[6]; Nagel et al who in 1997 in the famous Chicago experiment measured quantitatively and formally the irreversible granular compaction which occurs following tapping[7], a phenomenon known to anyone who has ever refilled a sugar bowl; Richard et al in 2003 who were the first to apply X-ray tomography to a granular system and recover the positions of individual particles[8]; and Andó et al in 2011 who were the first to re-identify the same particles before and after shear in such a tomograph[2].

This section also contains an overview of theories of granular behaviour: Starting with Coulomb's 1776 essay which formed the basis for elasticity theory[9], which was erroneously applied to soils for many years; Terzaghi's 1925 Magnum Opus "Erdbaumechanik"[10], which condensed many decades of experience as an engineer building dams into the basis of modern soil mechanical theory; Rowe and

Horne in the nineteen sixties, who were forming elaborate theories from properties like sliding angles and fitting them to data using empirically determined constants[11, 12]; and Edwards' ambitious theory of powders from 1989[1].

Edwards' theory[1] aims to re-purpose the powerful Statistical Mechanics, a theory used to turn microscopic details about gas molecules into heat capacities, to take us from grain shape and roughness to density or yield behaviour. With no relevant thermal fluctuations this is not a simple task – the lack of ergodicity (the symmetry between the ensembles of the system at different times, and a set of different systems) and the static Hamiltonian mean the theory must be redefined in new terms. The star of the show in the modified theory is the Entropy, which enumerates all the states that a granular arrangement could be in. The analogue of the Hamiltonian is the Volume Function, which makes the Chicago experiment[7] an extremely important empirical verification – the discovery of an irreversible loss of the quantity playing the role of energy!

Finally, the relevant literature is concluded with contributions from Blumenfeld which enumerate the degrees of freedom in granular systems by defining the quadron volume elements[13, 14]. These latch on to the complex structures found in disordered granular media and define their configuration in two and three dimensions, yielding a tensor quantity as a function of position in the system which describes the local environment.

The first section of original work in this thesis is a piece of algebra relating to the two dimensional quadron coordinates to macroscopic properties of the system. Starting with the Edwards theory statistical mechanical description, namely the volume function as a Hamiltonian-analogue and the usual statistical mechanics apparatus of a weighted sum normalised by a partition function, and plugging into it the enumerated degrees of freedom defined by the quadron coordinates, expectation values of two dimensional isostatic granular arrangements were cal-

culated. One of these expectation values, the volume expectation, simplifies to a simple function of the *compactivity* (the granular temperature-analogue) and the number of degrees of freedom – the granular equipartition of volume. This result was published during the project[15], but the full algebraic derivation is presented in chapter 3.

Then, in chapter 4 we come to the primary advance in the state of the art — a method to compute the quadron coordinates in three dimensions for the first time. In two dimensions, the problem is simple because the granular graph is planar (which means, among other things, that it has a dual.) In three dimensions, finding a suitable dual representation involves computing a pore-space decomposition into discrete cells which satisfies the strict geometrical constraints of the quadron coordinates. Two existing methods for decomposing pore spaces were implemented, including Beucher and Lantuéjoul's watershed algorithm[16] and Glantz and Hilpert's Tight Dual Method[17, 18], but neither of these produced cells which satisfy the constraints.

Thus, a new cycle-centric technique was developed from scratch, taking on several hard problems in computational graph theory, and implementing a geometrical constraint validation scheme which can handle the many edge cases present in granular systems from experiments and simulations. A detailed description of the development of this software and its performance and applicability form the last part of chapter 4.

The novel outcomes from applying this new computational technique to granular systems from other researchers is presented in the Results chapter. Of note are: the quadron volume distributions, which are an analogous result to published work on two dimensional granular systems; the classification of cells, which is similar to the classification of bubbles in foams; the maps of quadron volume by position; and the quadron chirality, which is a measure of anisotropy and directionality in

the structure.

This work gives us a brand new perspective on granular systems, which have confounded scientists for hundreds of years - the author is very pleased to present it to you.

Chapter 2

Literature Review

2.1 Mathematical Methods: Graphs

2.1.1 Graphs and their Representations

Granular systems are naturally described by mathematical objects called graphs. The Graph Theory branch of mathematics is large and diverse, but what is needed to understand this work is summarised in this short section, which uses definitions from Bollobás[19], and Bondy and Murty[20]

2.1.1.1 Graph

A *graph* is a network made up of two types of objects: *vertices*, which are point-like; and *edges*, which are line-like connections between the *vertices*. An example of a graph is the Facebook graph, where people are the vertices and friendships are the edges. An example of a graph in granular systems is the contact network, where each particle is a vertex, and vertices are joined by edges, corresponding to inter-granular contacts. Both these graphs can be seen to have a geometri-

cal form, where vertices are placed in space according to their properties: with Facebook, this might mean placing each vertex according to its listed home town; with granular particles the centre of mass of particles might be used. Edges are then added as lines between the relevant vertex positions; in Fig 2.1, we see the Facebook graph as rendered by Butler[21], with edges rendered as lines following geodesics around the world. Such representations, introduced formally shortly, are called *embeddings*, and are distinct objects from the graphs themselves.



Figure 2.1 – Paul Butler's embedding of the Facebook graph from December 2010[21], with vertices being global position, and edges drawn as geodesics projected onto the map. This image is copyright Facebook 2010 (see section D for usage request.)

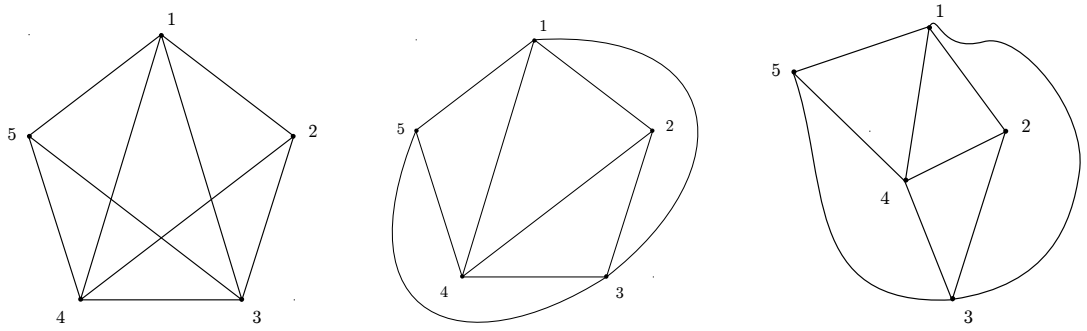
A formal definition of this geometry-independent object follows:

A *graph* G is an ordered pair, (V, E), of *sets* (unordered, non-repeating collections) of vertices $V = V(G) = \{v_0, v_1, ...\}$, and edges $E = E(G) = \{(v_i, v_j), (v_i, v_k), ...\}$ (that is, each edge is an unordered pair of elements from V.) Vertices that are connected or joined by an edge are said to be adjacent to and neighbouring one another, and the number of edges entering or leaving a vertex is called its *degree*.

2.1.1.2 Embeddings of Graphs

A graph can be drawn, with or without any reference to vertices' properties, by assigning each vertex v_i a position, and then connecting adjacent vertices with lines or curves to represent edges. Such diagrams are referred to as *embeddings*, and a given graph can be represented by many embeddings, with different choices of vertex positions and edge curves.

For some examples of embeddings, a graph $G = (V, E)$ where $V = \{1, 2, 3, 4, 5\}$ and $E = \{(1, 2), (1, 3), (1, 4), (1, 5), (2, 3), (2, 4), (3, 4), (3, 5), (4, 5)\}$) is rendered as three different embeddings, G_1, G_2, and G_3, in Fig 2.2. Note that embeddings technically need not only be two dimensional, but for practical purposes the ones illustrated here are.



(a) G_1 is an embedding where the vertices are the positions of a regular pentagon, and the edges are straight lines.

(b) G_2 is an embedding where the vertices are as in G_1, but where some of the edges have been drawn as curves to avoid intersecting others.

(c) G_3 is an embedding where the vertices are moved and several edges are curved unnecessarily - this is in order to emphasise that embeddings are *arbitrary*.

Figure 2.2 – Various embeddings of graph G.

2.1.1.2.1 Manifolds A *manifold* is a non-Euclidean space, which resembles Euclidean space locally. An example of a two dimensional manifold is the surface of the sphere or torus; although there are some two dimensional manifolds which

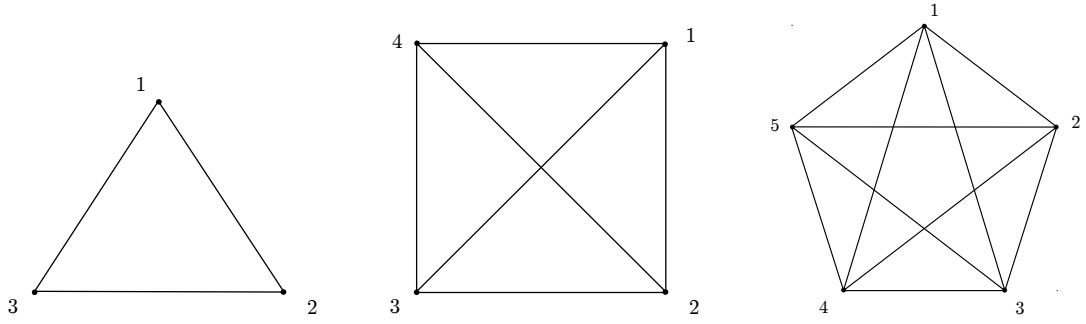
cannot be represented in three dimensions without intersecting themselves. Manifolds are often rendered using local *charts*, which are Euclidean. We might find it useful to consider graph embeddings on manifolds, for example a granular experiment on the outer surface of a cylinder, or a simulation with periodic boundary conditions.

2.1.2 Special Cases

A few famous graphs, which will be referenced later in the work, will be discussed here before proceeding.

2.1.2.1 Complete Graphs

The complete graphs are identified by a single integer – the number of vertices. The *complete* part of the name refers to the fact that every possible edge is present. So, K_3 has three edges (forming a triangle), K_4 has six edges (in 3D granular systems, K_4 often emerges where four grains all touch in a tetrahedral shape), and K_5 has ten. This relationship is encapsulated in the form $|E_{K_i}| = \frac{i(i-1)}{2}$, since each node i has $i-1$ edges, and each is counted twice. Complete graphs three to five are illustrated in Fig 2.3.



(a) Complete Graph with three vertices, K_3

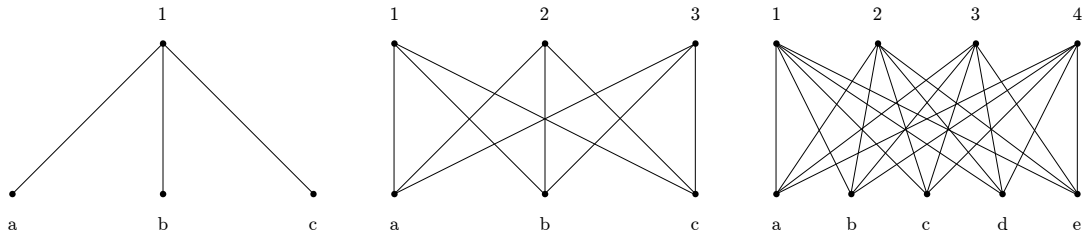
(b) Complete Graph with four vertices, K_4

(c) Complete Graph with five vertices, K_5

Figure 2.3 – The complete graphs with three, four and five vertices.

2.1.2.2 Bipartite Graphs

Bipartite graphs are where the set of vertices V can be split into two subsets such that no edges exist between vertices in the same subset, only between vertices in different subsets. This is most easily illustrated using the bipartite complete graphs, which are identified by two integers - one for the size of the first subset and one for the second. Obviously, here complete means that every vertex in the first set is connected to every vertex in the second, and (therefore) vice versa. Some of these graphs are presented in Fig 2.4.



(a) Complete Bipartite Graph with one and three vertices, $K_{1,3}$

(b) Complete Bipartite Graph with three and three vertices, $K_{3,3}$

(c) Complete Bipartite Graph with four and five vertices, $K_{4,5}$

Figure 2.4 – Examples of bipartite complete graphs.

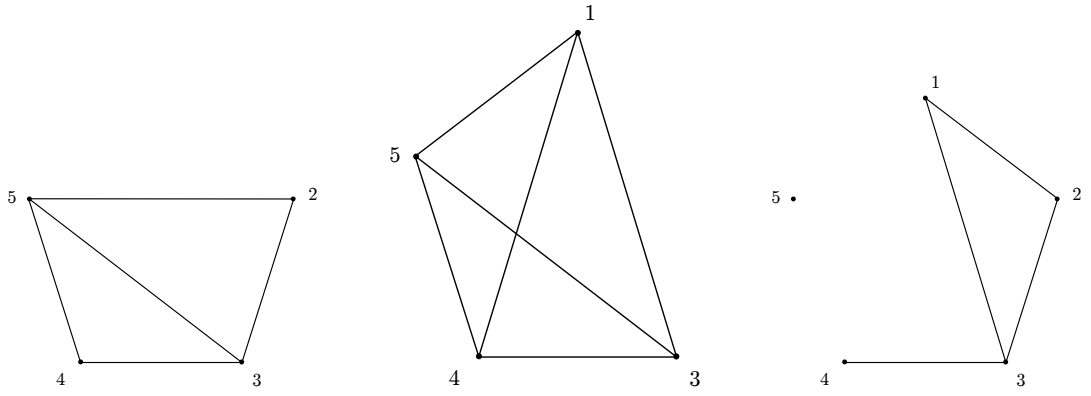
2.1.3 Graph Structure

2.1.3.1 Isomorphism

The reader may have noticed that technically the vertices in graphs are arbitrary sets – it would be possible to relabel any of the graphs in the preceding sections with different integers, or indeed any other object that is valid in a set, but retain the edge structure. What such disparate graphs have in common is *isomorphism*, that is if there exists a mapping between two graphs G and H such that all nodes and edges are preserved, the two graphs are isomorphic, $G \cong H$. Synonymously, the graphs share the same *structure*.

2.1.3.2 Subgraphs

Borrowing from the notion of a *subset* (a set made up of less than all of the members of some other set), a *subgraph* $G' = (V', E')$ can be defined of a graph $G = (V, E)$, where $V' \subseteq V$ and $E' \subset E$. A subgraph is called *induced* if it can be obtained by simply deleting vertices and the edges that point to them, and *spanning* if $V' \equiv V$. Arbitrary, induced and spanning subgraphs of K_5 are illustrated in Fig 2.5.



(a) An arbitrary subgraph of K_5, with vertex 1 and all its edges removed, and edge $(4, 2)$ removed.

(b) A subgraph of K_5 *induced* by removing vertex 2. What remains is isomorphic to K_4, another complete graph (as with all induced subgraphs of the complete graphs.)

(c) A subgraph of K_5 containing all its nodes, but with many of its edges removed. Note that node 5 has no edges, but is still part of the graph.

Figure 2.5 – Various subgraphs of K_5

2.1.3.3 Subdivisions

Adding new vertices of degree 2 on an existing edge (breaking an edge in half, and adding a new vertex at the break) does not change the topology of a graph – such graphs are said to be *subdivisions* of their parents.

2.1.3.4 Paths

A *path* is a subgraph of some G, P, which is a simple chain; that is if $V(P) = \{1, 3, 7, 19\}$ then $E(P) = \{(1, 3), (3, 7), (7, 19)\}$. Note that a path cannot contain a vertex more than once, but that the parent graph may contain additional edges between the elements of $V(P)$, or equivalently P need not be an induced subgraph. The example path P is illustrated in Fig 2.6.



Figure 2.6 – A path, P.

2.1.3.5 Cycles

A *cycle* is a subgraph of some G, C, which is a simple loop. Equivalently, this is a path whose endpoints are connected, e.g. using the example from the Paths section, the two end vertices 1 and 19 would share a new edge (this would only be valid if this edge was present in the original G - this cycle is illustrated in Fig 2.7.

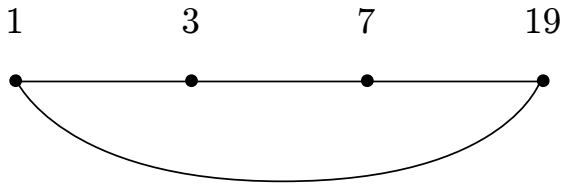


Figure 2.7 – A cycle, C.

A cycle is said to be *chorded* if any of its vertices are adjacent in the parent graph but not in the subgraph - i.e. if there is an edge which forms a "chord" (as used in geometry) across the cycle. Equivalently, a cycle is chorded if it is not an induced subgraph of its parent.

Both the vertices of a cycle $V(C)$ and the edges of a cycle $E(C)$ uniquely identify a cycle in a particular graph.

2.1.3.6 Trees

If there are no cycles in a graph, then it is called a *tree*.

2.1.3.7 Connectedness

If there is no path between some pair of vertices, then the graph is *disconnected*. In this eventuality, each separated part of the graph is a *connected component*. If the graph consists of multiple components, each of which is a tree, it is sometimes called a *forest*.

2.1.4 Analysis

2.1.4.1 Spanning Trees and Cycle Bases

A spanning subgraph (section 2.1.3.2) of a graph which is a tree (section 2.1.3.6) is called a *spanning tree*. Only connected graphs (section 2.1.3.7) can have spanning trees – the equivalent for a disconnected graph would be a *spanning forest*.

A spanning tree can be obtained by repeatedly removing any edge of a connected graph which is a member of a cycle – eventually there will be no cycles left and the subgraph will be a tree, and since the procedure only involves removing edges and not vertices it will be a spanning tree. Obviously there are many different choices to make during this algorithm (since each cycle has more than one edge, by definition) so the choice of spanning tree is arbitrary.

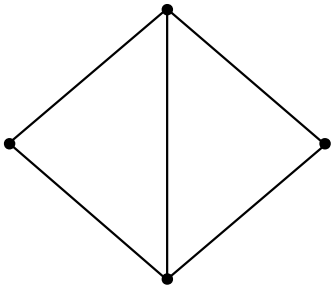
Once a spanning tree T of a graph G has been chosen, the associated *cycle basis* of G can be computed as follows.

Each edge e_c that is in G but not T (the set $E_C = E(G) \setminus E(T)$) can be added back into T to complete exactly one cycle, since it is a tree. Such a cycle is called

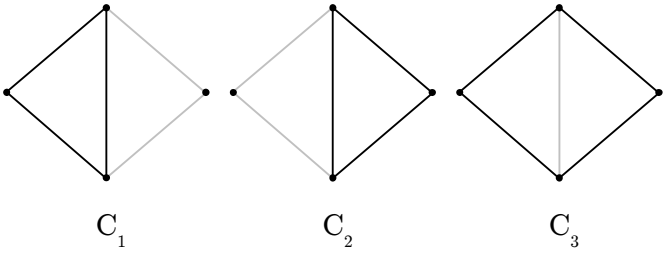
a *fundamental cycle* of G with respect to the spanning tree T and the edge e_c.

There is one such fundamental cycle for each edge e_c for a particular choice of spanning tree, and the set of fundamental cycles $B(G) = \{C_1, C_2...\}$ is called a cycle basis, which analogous to a vector basis can be used to construct any cycle in G by cycle addition and subtraction. A simple cycle basis and examples of addition and subtraction of cycles is given in Fig 2.8.

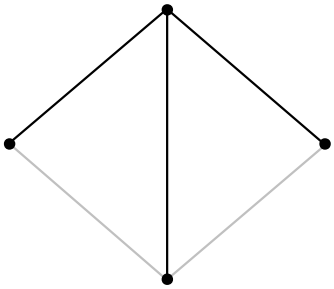
If the vertices of a graph are points, and the edges represent relative vectors between those points, then the spanning tree's edges are a choice of independent vectors in the edge set. The fundamental cycles, in such a graph, yield the linear combinations needed to represent the dependent vectors in terms of the independent ones.



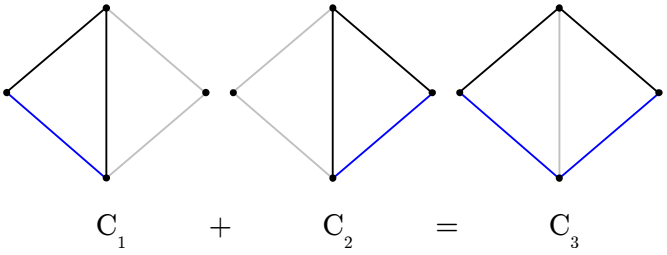
(a) A graph G.



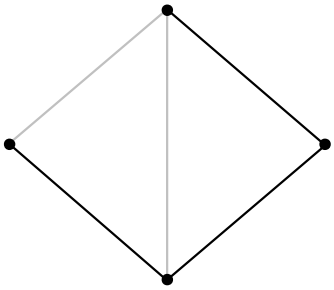
(b) The three cycles in G, C_1, C_2 and C_3.



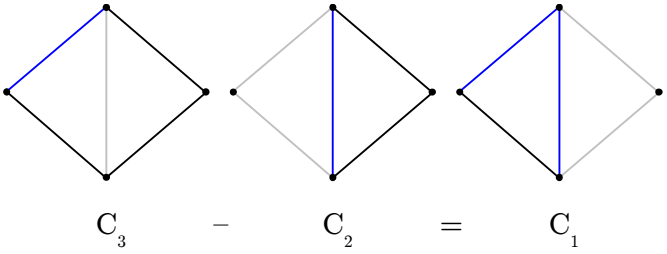
(c) A spanning tree of G, T_1.



(d) The fundamental cycles of G with respect to T_1: C_1 and C_2, and their sum C_3.



(e) A spanning tree of G, T_2.

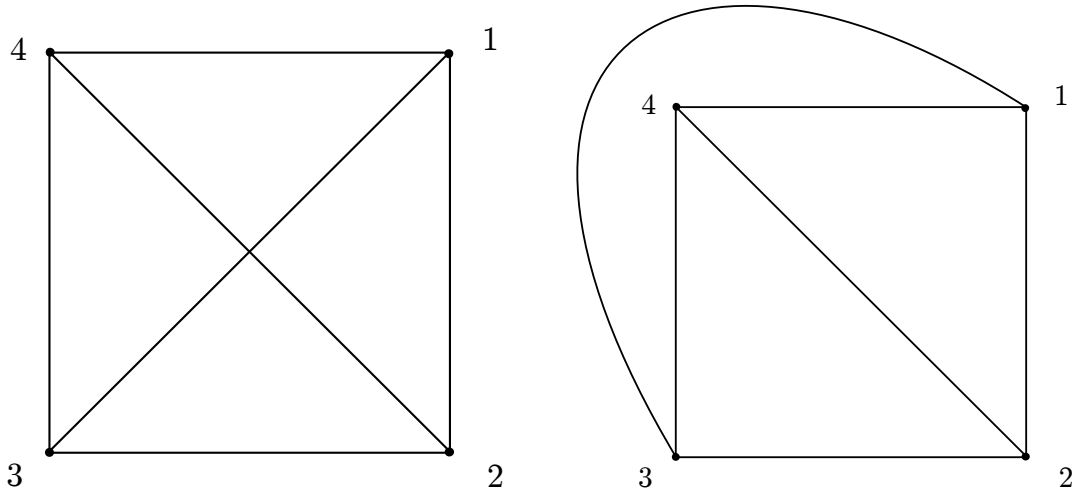


(f) The fundamental cycles of G with respect to T_2: C_2 and C_3, and their difference C_1.

Figure 2.8 – Two alternate cycle bases for an example graph, and explicit computations of the remaining cycle as a linear combination of the fundamental cycles.

2.1.4.2 Planarity

All graphs are either planar or not, regardless of whether their vertices have geometrical information. The simplest way to discuss planarity is to consider the embedding of a graph in the plane; if a graph is *planar*, then it can be embedded such that edge lines only meet at vertices, and never cross in other circumstances. It can be shown that K_4 is planar by re-drawing the $(2, 4)$ edge as demonstrated in Fig 2.9. From this drawing of K_4, it should be obvious that K_5 is *non-planar* – if not, consider where on the diagram vertex 5 could be placed such that it can be joined to all the other vertices without intersecting an edge.



(a) The usual way of rendering the complete graph with four vertices, K_4.

(b) The complete graph with four vertices, K_4, rendered so as to emphasise its planarity.

Figure 2.9 – Drawing K_4 to emphasise its planarity.

Planar embeddings like those in Fig 2.9b have well-defined *faces*, that is polygons in the plane made up of exactly one cycle of vertices and edges. Faces are unique to a particular embedding, and a planar graph may have many valid embeddings and thus many possible sets of faces. Note that the area "outside" the embedding is also a face, and thus a tree always has one face.

Trivially, all graphs which arise from two dimensional physical systems are planar[1], but note that some three dimensional systems also give rise to planar graphs. For example a graph which is "planar on the surface of a sphere" is also strictly planar: simply make one of the faces the boundary in a flat plane. Thus, planar graphs (especially as subgraphs) can be found in unexpected places, including inside three dimensional complex structures, like granular packs. Physical systems posses inherent geometrical information which usually defines a single embedding and a single set of faces.

Formally, according to Kuratowski's Theorem[20], a graph is planar if and only if it does not have any subgraph which is a subdivision of K_5 or $K_{3,3}$ (which are the simplest non-planar graphs.)

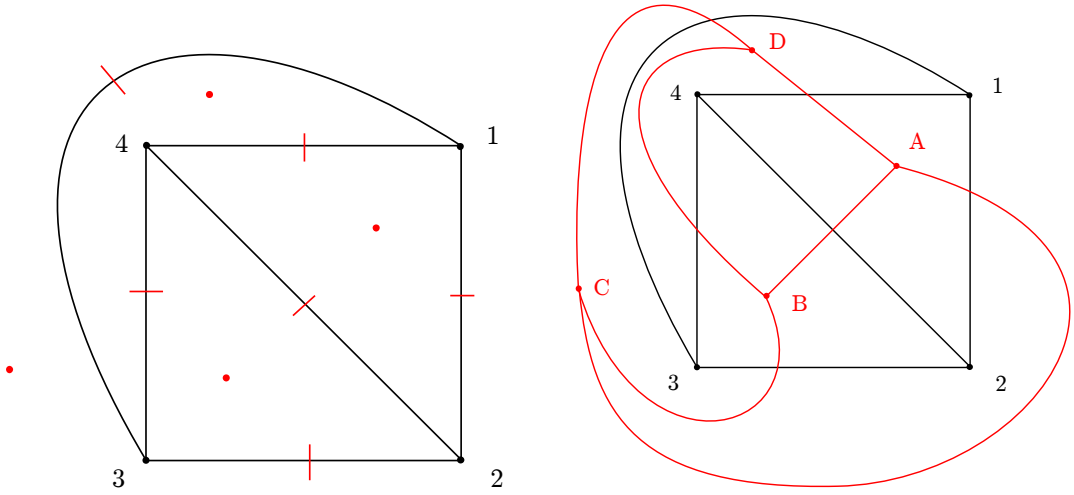
2.1.4.3 Duality

A duality relation between two graphs is a non-isomorphic reversible map between them. For example, the well-discussed dual relationship between planar graphs embedded on the same manifold[20], maps the vertices of one graph to the faces of the other, with a one-to-one correspondence between edges. The construction of such a dual, starting with the "rotated" edges and progressing to the vertices that correspond to the input graph's faces, is illustrated in Fig 2.10.

Referring to planar duality as in $d = 2$, duality in arbitrary dimension can be defined by recalling the Voronoï and Delaunay tessellations.

2.1.4.3.1 Voronoï and Delaunay Tessellations The Voronoï and Delaunay graphs are individually defined from a set of distinct points in d-dimensional Euclidean space, but are also one another's dual for a particular set of points, and

[1] I exclude some two dimensional simulations from this statement, particularly those with periodic boundary conditions in both x and y: such systems are planar on a non-trivial manifold.



(a) A planar embedding of K_4 is here annotated in preparation for drawing the dual; note that each face, and the boundary, have had a vertex added, and each edge has had the start of a dual edge drawn. Dual elements are in red.

(b) The dual of K_4 is rendered in red. Note that the dual rendered is actually isomorphic to K_4, and that since this graph cannot be rendered any other way in the plane, it is *always* true that K_4 is isomorphic to its own dual.

Figure 2.10 – K_4 and its dual, K_4.

so one can be inferred from the definition of the other (since both are defined in terms of a particular embedding).

The Voronoï graph is defined by building a region of d dimension around each input point, corresponding to the space that is closer to that point than any of the others. In two dimensions, these regions are the faces of a planar graph G_V embedded in the original space, while in higher dimensions they form cells, hyper-cells and so on. In any dimension, the regions are made up of "straight-edged" sides, that is straight lines in two dimensions (where the regions form polygons,) and flat faces in 3D (where the regions form polyhedra), and so on in higher dimensionality. All such shapes ultimately break down to vertices and edges, where the vertices V_V are local maxima in the distance function from the nearest input point. Normally, the degree of the vertices is $d + 1$, although symmetry can increase it[2]. An example of the Voronoï tessellation and arising graph in two

[2]For example in two dimensions consider the four corners of a square as the input points.

dimensions is rendered in Fig 2.11b.

The Delaunay graph G_D takes the input points as its vertices V_D (recall that for planar graphs vertices correspond to the *faces* of their dual). The edges are constructed by taking the d-sphere for the relevant Euclidean space (a circle in two dimensions and a sphere in three), and arranging it on the points such that it is the largest it can be without containing any of the points (i.e. its centre is the position of a Voronoï vertex.) This will position $d + 1$ vertices on its surface, or more if there are symmetry relations. Then, a surface mesh of the vertices is constructed to add the edges E_V (this constitutes triangles in two dimensions, or tetrahedra in three, assuming no symmetry relations.) These edges correspond exactly to $d - 1$-sized entities in the Voronoï tessellation bounding exactly two regions – edges between faces in two dimensions and faces between cells in three. The Delaunay tessellation in two dimensions is rendered in Fig 2.11c on top of its Voronoï dual graph.

Thus, the duality of the two graphs is established, even for the higher dimensional (non-planar) embeddings. The duality mapping for two and three dimensions is illustrated in Table 2.1b.

Table 2.1 – The duality relations for graph embeddings in two and three dimensions.

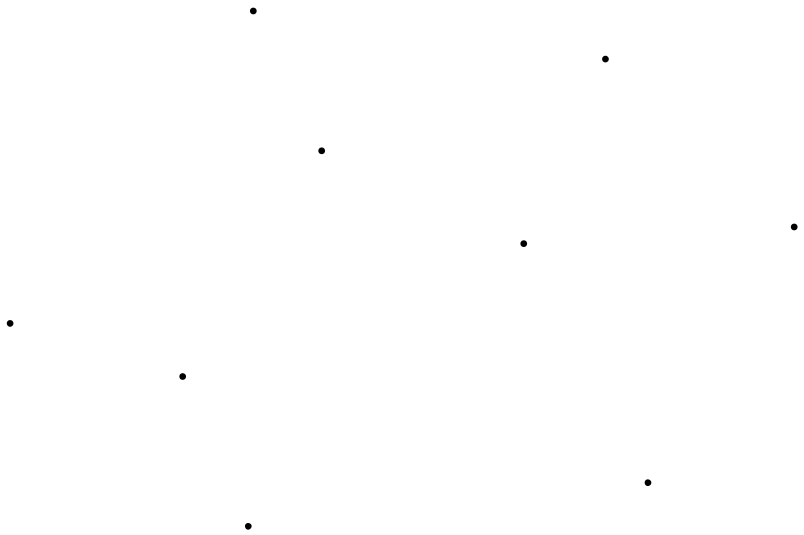
Graph Embedding	**Dual**
Vertices	Faces
Edges	Edges
Faces	Vertices

(a) Two dimensional duality

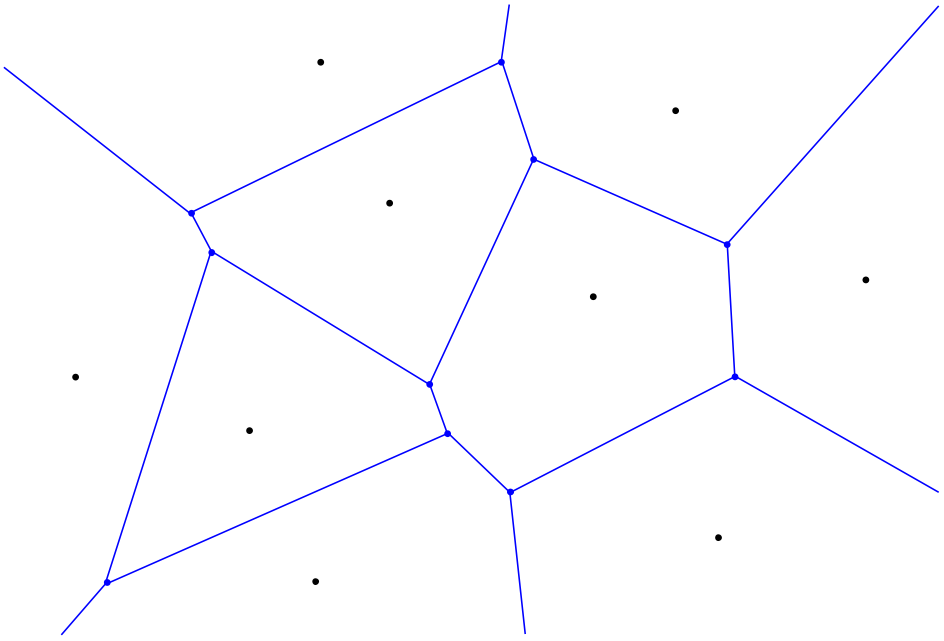
Graph Embedding	**Dual**
Vertices	Cells
Edges	Faces
Faces	Edges
Cells	Vertices

(b) Three dimensional duality

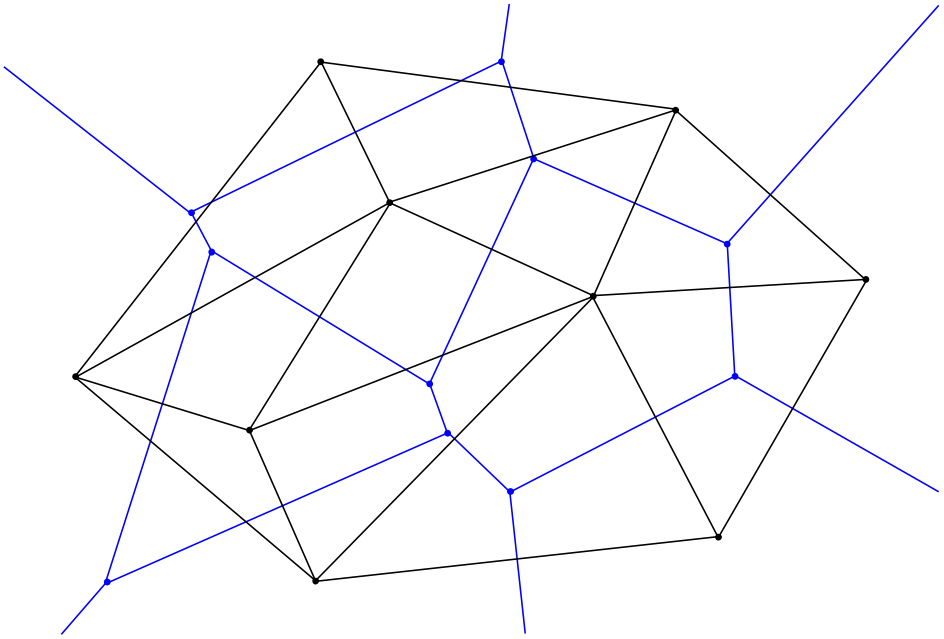
There will be a local maximum of the distance function in the centre of the square, which is equidistant from all $4 = d+2$ input points. Similar problems can arise in higher dimensions and on other manifolds too.



(a) To construct the Voronoï and Delaunay graphs, we require an input of a set of points, here in two dimensions.



(b) Each input point is surrounded by a polygonal face delimiting the space nearer that input point than any other. The boundary of these faces form the Voronoï graph (blue). The edges are segments of the lines bisecting the input points.



(c) Finally, we add the Delaunay edges (black). Note that there is a one-to-one correspondence between Voronoï vertices (blue dots) and Delaunay faces (black triangles), although each triangle does not contain exactly one vertex: rather, each blue vertex is the centre of the circle whose boundary lies on all three of the triangle's vertices.

Figure 2.11 – Construction of the Voronoï tessellation, and the Delaunay triangulation as its dual, in two dimensions.

2.2 Computational Methods

2.2.1 Preliminary Computational Definitions

A few terms which are common in computer science for specifying computational methods (and their difficulty) will be defined before any such methods are discussed. This section also uses definitions from Bondy and Murty[20].

Algorithm – An algorithm is any procedure for performing a mathematical computation. For example the method taught in schools for dividing numbers with many digits, sometimes called *long-division*, is an algorithm – its input, or arguments, are the numerator and the denominator, its output is the result of the division, and the procedure for juggling individual digits is the algorithm itself. In this work, the algorithms that are discussed mostly take more complicated mathematical objects, often graphs or sets, and compute results which, like the division, vary according to the specific input, for example identifying all the cycles in a particular graph.

Computational Complexity – The computational complexity of an algorithm is the number of instructions it must run to produce the output. This is almost always a function of the size of the input, for example the number of bits required to represent the graph in question or the number of elements in the set, and is sometimes written as such to an approximation in "big O notation" (the O is for *order*, as in the order of magnitude.) Examples of complexity are *linear* or $O(n)$, *quadratic* or $O(n^2)$, more generally *polynomial* or $O(n^m)$, or in the worst cases *exponential* or $O(x^n)$. Problems are often analogous to other problems, where a mapping exists between the inputs and outputs, and by discovering such analogies

the complexity of the harder of two problems can be deduced to be that of the easier plus the mappings.

Problem Class $\mathcal{P}$ – The class of problems known to be solvable in *polynomial* time is labelled $\mathcal{P}$. Examples of *polynomial* time problems include checking whether an element is contained in a set (*linear* in the number of elements in the set), or comparing each member of a set with every other member (*quadratic* in the number of elements). Determining whether a problem is in the class $\mathcal{P}$ is a matter of devising an algorithm which solves the problem in polynomial time, thus determining that a problem is not in $\mathcal{P}$ is sometimes difficult.

Problem Classes $\mathcal{NP}$ and co-$\mathcal{NP}$ – Decision problems (where the answer is either *yes* or *no*) may exist in certain problem classes as follows: if the answer can be computed in polynomial time by some algorithm, then the problem straightforwardly belongs to class $\mathcal{P}$. If not, but there is some verification algorithm (a certificate) which can check the validity of a particular solution in polynomial time, and thus resolve the problem with a *yes* if the solution is valid, then the problem is *non-deterministically polynomial*, and belongs to the class $\mathcal{NP}$. Similarly, an analogous certificate which resolves the main problem as *no* if a counter-example is found, places the problem in the class co-$\mathcal{NP}$.

Problem Class $\mathcal{NP}$-complete – Some problems in the class $\mathcal{NP}$ (and co-$\mathcal{NP}$) are, in spite of a known polynomial test for a particular solution, not solvable in general in any reasonable time for large examples. The $\mathcal{NP}$-complete class of problems are (loosely) defined as the most difficult problems in the $\mathcal{NP}$ class, in that they require a complete search of the possible solutions (until one is found which completes the certificate.) An example of such a problem is the satisfiability of Boolean formulae – that is, given a Boolean formula with Boolean variables and

operators (like *and*, *or* and *not*), whether the formula ever be satisfied (evaluated to true.) In order to test whether a particular formula can ever be satisfied, after simplifying the expression using Boolean algebra, every combination of the remaining Boolean variables must be evaluated – a polynomial time operation – until the expression evaluates to true, making the complexity $O(2^n)$ for n variables. $\mathcal{NP}$-complete problems are defined to be at least this complicated.

Problem Class $\mathcal{NP}$-hard – So far, most the problem classes defined have only contained decision problems – problems with a *yes* or *no* answer – there are many combinatorial problems that require enumerating all the examples of something, or finding the best of something, which by definition do not belong to these classes whatever their complexity. The problem class $\mathcal{NP}$-hard contains problems which are at least as hard as $\mathcal{NP}$-complete problems, but which do not yield decision answers. An example of an $\mathcal{NP}$-hard problem is the famous Travelling Salesman Problem (TSP), which is finding the shortest route that visits every city on a map (this is analogous to finding the lowest edge-weight-sum path that visits every node in a graph.) It can be shown[20, p. 188] that to solve the TSP one must also solve a Hamilton Cycle – an $\mathcal{NP}$-complete decision problem – which is enough to show that the TSP is $\mathcal{NP}$-hard.

2.2.2 Enumerating All Cycles

Given a graph G, determining whether some ordered list of elements is a cycle in G is a linear problem – verifying that each adjacent pair in the list, including the start and end elements, is also adjacent in G. Hence, a naive algorithm to compute all cycles in a graph must perform a linear test on every path with at least two edges to see if the terminating elements are adjacent – enumerating all such paths is $O(2^{|V(G)|})$, implying that the problem is $\mathcal{NP}$-hard.

2.2.3 Maximum Independent Set

An Independent Set of a graph G is the vertex set of an induced subgraph G_s with no edges – that is, $V(G_s)$ cannot contain any two elements which are adjacent in G. A *maximal* independent set is one which is not a subset of another independent set. Such a set can be discovered in polynomial time by a *greedy* algorithm[22]: choose a starting vertex; eliminate all adjacent vertices, choose a remaining vertex and repeat until there are no more vertices available. Such an algorithm is of linear complexity in the number of nodes.

The *maximum* independent set is the largest of the maximal sets, and finding it involves enumerating all of the maximal sets to find the largest. This problem is $\mathcal{NP}$-hard.

An analogous problem is one where each vertex has some *weight*, and the objective is to maximise the sum of the weights rather than the number of vertices – the procedure being the same, this alternative problem is also $\mathcal{NP}$-hard.

2.2.4 Contact Detection

The problem of detecting colliding objects in geometrical space, whatever their shape, can be reduced initially to the overlap of spheres (where each sphere, relating to one object, is the smallest shell which contains the whole object,) and such a problem is trivially polynomial – $O(n^2)$ to compare each possible pair of objects. Advanced algorithms for detecting intersection of non-trivial surfaces can then be performed on pairs which are near each other (whose shells overlap), which is much more efficient.

In fact, this $O(n^2)$ algorithm can be improved upon by binning the objects' positions and then operating on each bin only once (using the maximum shell diameter

as the bin size,) as detailed by Munjiza[23]. This algorithm can be implemented in arbitrary dimension, and reduces the complexity class of the contact detection problem to $O(n)$, making it practical to compute for even large granular systems where required.

2.2.5 Voronoï Tessellation

The problem of computing the Voronoï tessellation of a set of points in Euclidean space is a geometry problem defined in section 2.1.4.3.1. The computational complexity of computing it in two dimensions is $O(n \log n)$, by a divide and conquer algorithm (subdividing the space into recursive half-spaces until each contains two points, finding the bisecting line and then combining the adjacent half-spaces by a path of new edges.)

In three dimensions the complexity varies considerably depending on the result, as Golin and Na [24] show – theoretically the tessellation can contain $O(n^2)$ vertices[3], requiring that any algorithm will be at least quadratic.

In cases where the points are uniformly distributed in a convex shape (the authors give cubes and spheres as examples) the number of neighbours is roughly constant per point away from the boundary, and so the number of edges and faces might be as low as $O(n)$. Golin and Na[24] also show that when the points are taken from the surface of a convex shape, the resulting tessellation graph also has $O(n)$ faces, edges and nodes. Such Voronoï tessellations may therefore be computed in less than quadratic time.

[3]The reference uses *complexity* to mean the number of objects (vertices, edges and faces) in a graph, and is not referring to the computational complexity.

2.2.6 Convex Hulls

A convex hull is the smallest convex region containing a set of input points. For a set of discrete input points in three dimensions this region will be a polyhedron, with some or all of the input points as polyhedral vertices.

Convex hulls for such discrete sets of points are the dual of the Voronoï tessellation of the points as shown by Brown [25], and thus this problem shares its computational complexity with the other – in particular with the sub-problem of points taken from a surface, which has less than quadratic complexity, if all the points are on a sensible convex surface.

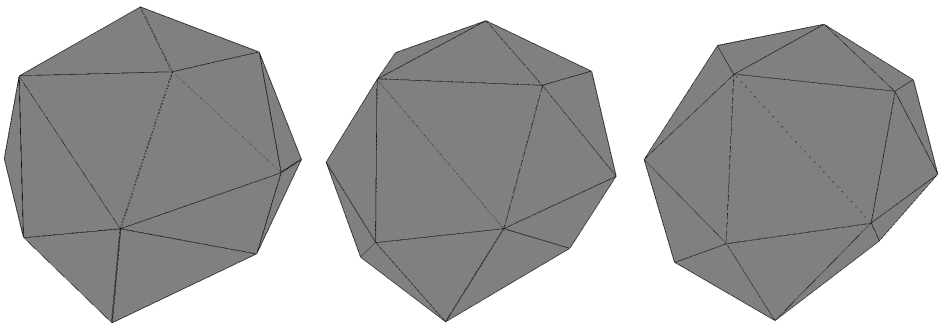


Figure 2.12 – Three different angles of the same convex hull, as if the viewer is rolling it horizontally on the page.

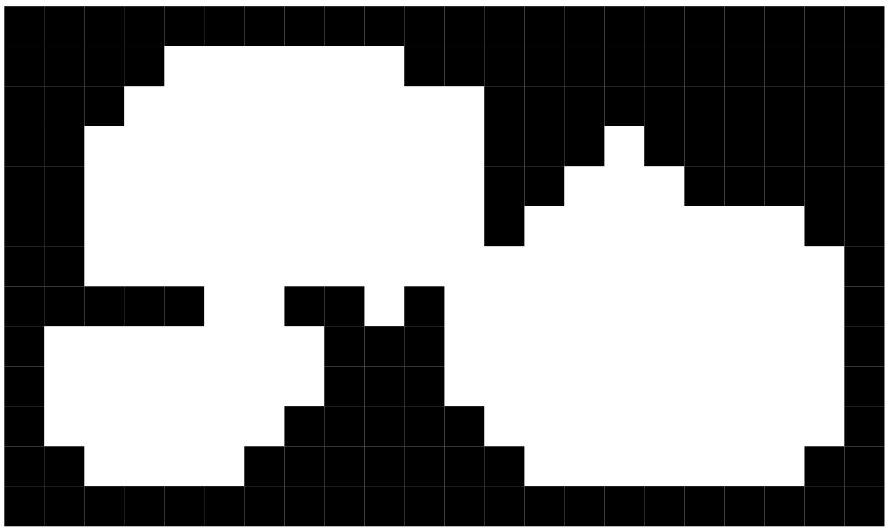
2.2.7 Segmentation

Segmentation is the problem of taking a continuous region and dividing it into segments. In the study of porous media, this is to achieve an equivalent network that can be used to approximate flow through the space between the solid phase or grains – where the gap narrows and constricts flow the feature is better represented as a narrow channel between the pores, where it widens out again. In granular matter the technique is also used on the solid phase itself to distinguish granular particles from each other – it is normally most effective when the grains are convex.

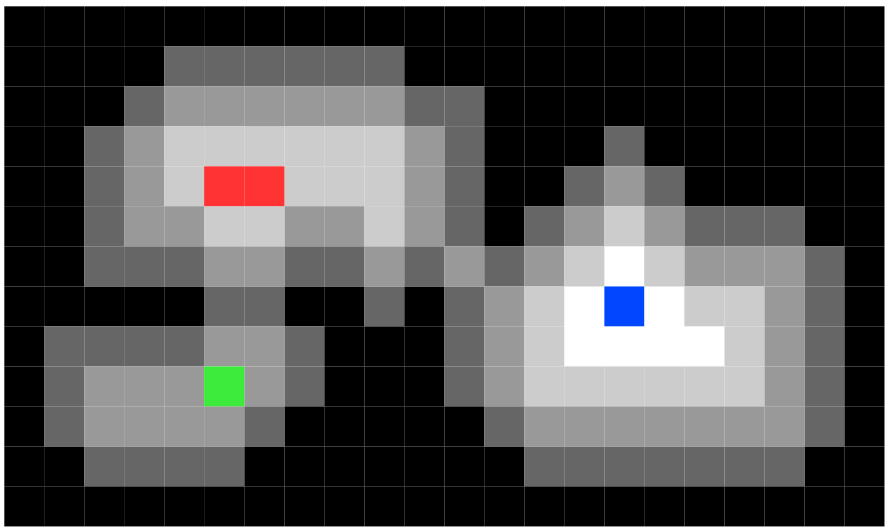
Some segmentation algorithms work on pixel images, and others work with surface meshes – either of which can form the input data about the barrier between the phase of interest and the remainder. Pixel images are the usual form for experimental data (in two and three dimensions), so they are ubiquitous and often unavoidable, in spite of their $O(l^3)$ RAM requirements, and the subsequent increase in instructions to work with. Surface meshes can be much more efficient (to represent and to iterate over), but unless your data is in an "object oriented" form to start with (a particle list rather than a photograph) you must convert your data (and worry about the quality of the resulting mesh.)

2.2.7.1 Watershed Algorithm

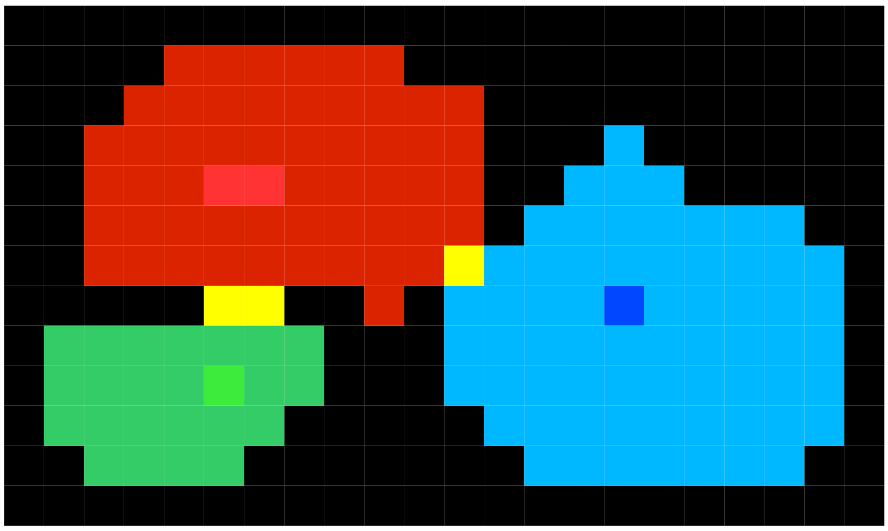
The Watershed algorithm was first proposed by Beucher and Lantuéjoul[16], and is a method for finding objects in a pixel image (including in three dimensional pixels, *voxels*.) The algorithm consists of a cellular automaton which (conceptually) strips layers of pixels away from the phase in question. As this process progresses the widest parts of the phase are separated from one another, because the narrower channels between them have been completely stripped away. There is then a consolidating phase where each pixel is labelled according to the pore it belongs to by looking at the gradients (pass during which each pixel was eliminated).



(a) The input to the watershed algorithm is a pixel image, where for example the solid phase could be masked away, leaving the pore space to be operated on by the cellular automaton.



(b) In the second phase, "flooding", the distance function is computed (here illustrated in shades of grey) away from the solid phase (black) walls. Note that at the end of this procedure it is possible to identify maxima which distinguish the objects to be identified, which are presented in bright colours here.



(c) The final phase of the algorithm is computing the steepest uphill gradient from each pore space pixel, conceptually each cell becomes a tree with the original maximum as the root node. Some pixels (here marked in yellow) may have two equally steep paths which lead to different maxima – these can simply be marked as the interfaces between the cells.

Figure 2.13 – The three stages of the watershed algorithm for segmenting pore spaces, illustrated in two dimensions.

The complexity of the watershed algorithm, in terms of the number of pixels n, is estimated as follows. Each pixel must be considered once in each pass, so each pass is $O(n)$. The number of passes required varies according to the image, but in a worst-case it will be $O(n^{\frac{1}{d}})$, in d dimensions, because that is how long it will take to progress from the edge of a square or cube image to the centre, one layer at a time. Thus, the overall complexity of the watershed procedure is $O(n^{1+\frac{1}{d}})$, assuming that the consolidating phase is $O(p)$, the number of pores $p \ll n$. Note that while it may initially look like 3D watershed is faster than 2D watershed, remember that three dimensional images have more pixels, so for the same system length and resolution, $n_{d=3} = O(n_{d=2}^{\frac{3}{2}})$.

2.2.7.2 Tight Dual Method

Watershed is very useful when the data is provided as a pixel image, but what is the best approach for 3D data-sets with a surface mesh? One approach might be to "render" the mesh into voxels at a particular resolution, but Glantz and Hilpert developed the Tight Dual Method[17, 18] as a means of generating higher quality equivalent networks where mesh data is available. The method is also a more efficient[4] way to produce an equivalent network, when the mesh is known.

The Tight Dual Method operates on a Voronoï complex – some of the elements (vertices, edges and faces) from a Voronoï tessellation – which is constructed as follows. First, a Voronoï tessellation is computed over the mesh vertices. Note that, since each mesh vertex is by definition on the boundary between phases, its corresponding cell in the Voronoï tessellation will be partly in each phase. The complex is the result of removing from the Voronoï tessellation all vertices, edges

[4]The Watershed complexity was stated in terms of the number of voxels, which is $O(l^3)$ (cubic in the length of the system.) In addition to applying to the time of computation, this scaling also applies to the RAM usage to represent the image – a mesh system is *considerably* more efficient in RAM, even at high resolution with many vertices per grain.

and faces which are wholly or partly in the solid phase (which will remove some part of each cell, leaving them all open.) The resulting Voronoï complex shares its topology with the pore space, and the remainder of the method is a set of rules for merging the Voronoï vertices, edges and faces into distinct sets that form pores.

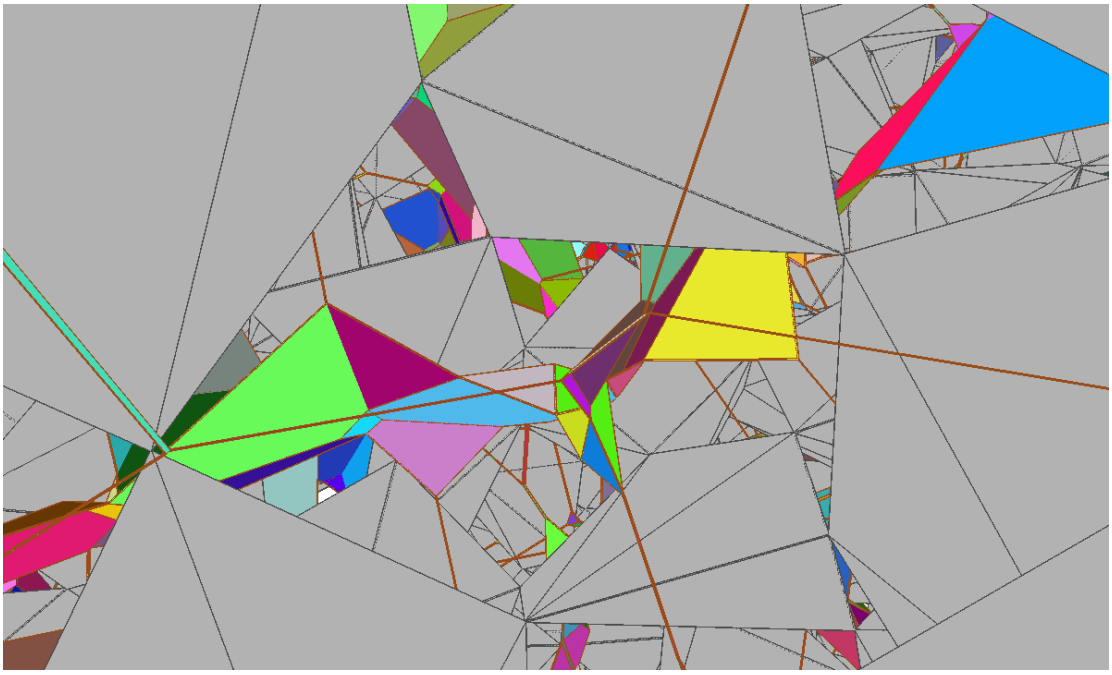


Figure 2.14 – Glantz and Hilpert's Voronoï complex in many colours, for a set of granular convex polyhedra in grey.

Because the pores consist of sets of Voronoï vertices, their dual representation as Delaunay tetrahedra can be trivially deduced – the union of the tetrahedra in a pore gives the geometry of each pore in terms of the original mesh (which is very useful for subsequent flow calculations.)

The complexity of the Tight Dual Method is the same as that of the Voronoï tessellation - the method computes such a tessellation and then performs linear operations on its result.

2.3 A Granular Statistical Mechanics

2.3.1 Historical Work on Granular Materials

2.3.1.1 Soil Mechanics

In an essay in 1776 Coulomb[9] introduced what would become elasticity theory, which was applied to granular materials by engineers until the middle of the twentieth century. It is a hugely successful theory, predicting the yield behaviour of many materials.

Something fundamentally different about granular materials was first described by Reynolds[4] who observed that it changes its volume during the yield process – something which cannot be accounted for by elasticity theory which operates at constant volume. Reynolds performed experiments to quantify these observations, where grains were confined in a canvas bag and flooded with water. The bag was sealed excepting a tube to an apparatus for measuring fluid flow in and out. It was observed that when the bag was compressed fluid flowed *in* – the re-arrangements of the particles had caused an increase in pore volume. Reynolds named this phenomenon *dilatency*.

The history of soil mechanics as a discipline, that is the study of the stress-strain relation of soils and similar materials, cannot be written without crediting Terzaghi[10], who was a great engineer who solved many problems relating to building structures which must interact with the materials which make up the earth. Terzaghi's Principle is the separation of the stress σ from the *effective stress* σ', which removes the effect of any fluid contained in the pores and looks only at the stresses between grains.

In 1962, Rowe[11] pulled together quantitative experimental work from the pre-

ceding decades[26, 27, 28] and consolidated an analysis of existing theories to account for the collected observations. Summarising Rowe's quasi-static work, Horne[12] writes about granular systems under deformation that instantaneously only a few grains are actually moving relative to their neighbours – that large groups of particles move relative to one another (although the sets of particles in the groups change frequently.) It is also pointed out that Rowe's equations only apply in the absence of rolling or grain rotation. Horne goes on to hypothesise that grain rotation should be minimised following from the mechanics of particle interactions.

In the second part of his work "The behaviour of an assembly of rotund, rigid cohesionless particles" Horne[29] introduces a characteristic of a random arrangement of grains, the *mean projected solid path*. They argue that anisotropy in a granular system can be expressed in terms of the mpsp along each axis, m_i, and that the changes to the granular system's response to applied stresses during yield are related to changes in the anisotropy, characterised by ratios $\frac{m_i}{m_j}$. They also compute explicitly the m_i for close packed structures, and show agreement with Rowe's measurements[11] of regular systems.

At a similar time, Litwiniszyn[30, 31] developed a theory of granular motion undergoing a random walk. Later other studies with the aim of deducing the stress-strain relation for granular materials from analysis of the possible mechanics of the individual grains, were continued by others[32, 33, 34], producing several ad-hoc theories with empirically determined parameters. Some of this work built on concepts from Horne, notably the mean projected solid path in Oda[34].

An interesting development in Soil Mechanics occurred when the innovation of photo-elastic beads was conceived. Murayama and Matsuoka[6, 35] used this to investigate Murayama's earlier theories[36] in two dimensional systems by shearing systems of cylinders. More complex experiments using photo-elastic beads were

later used by Majmudar and Behringer[37].

2.3.1.2 Structure and Microstructure

In 1929, Smith et al[5] reported an experiment to measure the numbers of contacts in granular materials. This was an ingenious experimental method, which used residual acid between small lead spheres ("shot") held by capillary pressure to mark contacts with white lead oxide that were then counted by hand, in a painstaking method similar to that employed by Matske later on foams [38]. Smith reported on six different compaction levels, the more dense being achieved by compression with a piston, and noted the increasing proportion of crystalline (close packed) regions in this monodisperse (single radius) system, signified by grains with 12 contacts. This was the first measurement of a granular *microstructure*, the observation of the precise configuration of the grains.

Another fruitful source of granular invistigation was the observation and qualitative description of rocks in the field. In 1966 Lafeber[39] took many such investigations and presented a quantitative mechanism for characterising soil microstructures, in particular compared the frequency of contacts between particles of different phases to expected values in random arrangements. They showed that a common theory of the time, that the hard particles in the soil formed a skeleton structure that would stand without the softer components, was unfounded. Correlations of the orientation of grains and pores were also observed.

The soil mechanics theories of granular assemblies, where reasoning about the mechanisms of deformation from local structure and degrees of freedom are extrapolated to macro-behaviours, were not the only path of theoretical investigation at this time. In 1965, Mogami[40] characterises a static structure of grains by dividing up the void space into discrete elements and assigning a "void ratio" to

each - a ratio of the volume of the surrounding grains to the volume of void space in this element. They proceed to derive a form for the *probability of a particular state*, upon which to perform statistics to make predictions. Elsewhere, particularly in the soil mechanics literature, void ratio was used to mean the average packing over the whole system, in contrast to Mogami's definition.

2.3.2 Theory of Powders

Towards the end of the Nineteen Seventies, some interest was growing in using the Entropy of granular arrangements to describe their mechanics – specifically in work by Jowitt and Munro[41] and Brown[42]. It is this idea which formed the basis for the Edwards Theory, introduced in 1989[1], which this work builds on.

Statistical Mechanics is a theory of thermodynamics developed by Ludwig Boltzmann which uses information about the micro scale of a system (normally atoms and molecules) and derives from it expressions which describe all kinds of macro scale properties: the heat capacity of gasses; the ideal gas law which relates the temperature, pressure and quantity of a gas to the volume it occupies; the entropy or disorder present in a system; and many more.

The mechanism by which the expectation values of such properties are computed is by means of a weighted sum normalised by a partition function, the numerical method for which is summarised in section B.1.

Edwards et al[1, 43] introduced in 1989 a theoretical framework for analysis of granular materials and their behaviours, which drew from thermal statistical mechanics in order to define a similar set of macroscopic properties. There were several important differences to overcome in doing so – granular systems are unlike thermal systems which explore the different microstates available to them over time, a property called ergodicity, in that they sit on the bench (or beach)

doing nothing until some energy from outside the system is added. Edwards proposed that the ensemble of states be all those in which a set of grains could find themselves given the same initial protocol (pouring into the sugar bowl, for example). The entropy in the analogy was thus defined as the log of the number of microstates as in thermal systems, albeit in this new ensemble. Other quantities in the theory still needed definition – in the absence of significant thermal fluctuations, what of the temperature and energy? Table 2.2 shows the definitions that were chosen.

Statistical mechanics	Powders
E	V
$\mathcal{H}$	$\mathcal{W}$
S	S
k_B (dimensions of energy)	λ (dimensions of volume)
$\int \prod \mathrm{d}r_i \prod \mathrm{d}p_i$	$\int \prod \mathrm{d}r_i \, Q$
$\exp \frac{-S}{k_B} \delta(E - \mathcal{H})$	$\exp \frac{-S}{\lambda} \delta(V - \mathcal{W})$
$T = \frac{\partial E}{\partial S}$	$X = \frac{\partial V}{\partial S}$
$\exp\left(\frac{-F}{k_B T}\right) = \int \exp\left(\frac{-\mathcal{H}}{k_B T}\right)$	$\exp\left(\frac{-Y}{\lambda X}\right) = \int \exp\left(\frac{-\mathcal{W}}{\lambda X}\right)$
$F = E - TS$	$Y = V - XS$
$E = F - T\frac{\partial F}{\partial T}$	$V = Y - X\frac{\partial Y}{\partial X}$

Table 2.2 – A summary of the equivalent equations between statistical mechanics and Edwards' theory of powders, from Edwards et al[1]

Of particular importance in this analogy are the replacement of energy E with volume V, and their respective functional forms the Hamiltonian $\mathcal{H}$ with the volume function $\mathcal{W}$. It is important that Energy is replaced with another extensive

property, that is a property which is additive as the system size is increased, and that the Hamiltonian is also replaced with a function which is sensitive to the degrees of freedom of the system. If we see the volume of the system as a sum of the volume elements associated with each particle (or perhaps some other partitioning), with each such volume element depending on the degree(s) of freedom associated with each particle, then it can be seen that volume works as such a function.

Another important parameter in Edwards' analogy is the temperature T which is replaced by a quantity with the units as $\mathcal{W}$ called the compactivity, X.

2.3.2.1 The Stress Ensemble

In addition to the degrees of freedom associated with the physical positions and configurational arrangements of the particles, there are further degrees of freedom in the forces between particles which can vary independently of this configuration, to a small degree. In *isostatic* systems (where each grain has exactly enough contacts to be in mechanical equilibrium and no more) the forces are all linear combinations of the forces on grains from the boundaries, which form the degrees of freedom in the stress ensemble[44, 14, 45]. In systems with higher numbers of contact points, there will be additional degrees of freedom from some independent set of interior forces as well.

2.3.3 Quantifying the Degrees of Freedom

In order to make use of Edwards' proposed[1] statistical mechanical theory, the functional form of $\mathcal{W}$, the volume function, must be derived in terms of the degrees of freedom in a granular system.

2.3.3.1 In general

In arbitrary dimension, the granular degrees of freedom can be defined in terms of the positions of grains. To do so, it would be necessary to quantify the position of a grain *and* its shape, which for simple systems like sphere assemblies is easy, but for many physical systems is non-trivial. The shape is required because the stress state of the assembly must be computable, which necessitates determining the position of contact points between grains (the positions of the forces acting on the grains.)

It is far simpler, then, to directly define the positions and identities of the contact points themselves (where identities means the labels of the two grains in contact,) and the force applied to each grain at that contact point, as the degrees of freedom of a granular system. Note that area or multiple-point contacts between grains is equivalent to a single force, according to Blumenfeld[46], and thus as a single contact point, from the perspective of the overall stress state and thus the degrees of freedom.

Finally, this definition can be simplified further, since by Newton's laws of force and torque balance not all the forces are independent (this is obvious if the granular graph is considered – notice how many cycles it has.) These dependencies are eliminated if only a subset of the forces are included. In isostatic systems, this subset is only the forces on the boundary.

If working in Edwards' statistical mechanics[1, 43], the final degrees of freedom in the partition function must be tessellating volume elements. Such sets of tessellating elements are defined in two and three dimensions, such that their definition relies only on the structural degrees of freedom (the contact point positions.) The stress ensemble, consisting of the loop forces in the system[13], is a (interdependent[15]) second term in the partition function, which is defined later.

2.3.3.2 Two Dimensional Volume Elements

In two dimensions, the granular graph is planar (or planar on a manifold, for artificial systems with periodic boundary conditions), and in its natural embedding has a dual made up of the discrete void spaces between grains, which are named cells. Blumenfeld and Ball[13] exploit this simplicity, defining a set of tessellating volume elements with a two-to-one correspondence with the contact positions (which were used as part of its boundary). The details of this tessellation are explained in section 3.1. The volume elements, which are quadrilaterals, are named quadrons.

Blumenfeld and co-workers use a generalised tensor to represent the state of the grain[13], and the quadron[14], which is used to compute the volume and other quantities like the chirality (in two dimensions, a scalar measuring the skewness of the quadron.)

Quadron co-ordinates are simple to compute for any granular assembly in two dimensions for which one has the contact positions, and some research has been done into their statistical properties in simulated and experimental systems, described in section 2.4.5.4.

2.3.3.3 Three Dimensional Volume Elements

In three dimensions, the granular graph is no longer planar, and therefore computing its dual is no longer straightforward, as was hinted in table 2.1b and illustrated in the segmentation methods in section 2.2.7. Under the assumption that a valid dual exists, and thus that cells are a well-defined concept in granular materials, Blumenfeld and Edwards[14] provide an extension of the two dimensional quadrilaterals into three dimensions. The 3D volume elements are octohedral, and no longer share such a simple relation with the number of contacts, depending on the

co-ordination number and the cell geometries.

Just as in two dimensions, the quadron has an associated tensor, from which its volume and chirality can be computed. The chirality in three dimensions has not been defined in the literature before. A full description of the volume element boundaries, the tensor and the chirality are given in section 4.2.

The statistics and properties of three dimensional volume elements have not been analysed for granular media because of the difficulty in defining the cells. This is not true for foams, where three dimensional quadrons have been analysed in section 4.3 of Hihinashvili[47].

2.4 Experimental Techniques in Granular Physics

Experimental measurement of granular systems varies from conventional photography to X-ray based three dimensional image deduction. In the following sections, some of the techniques used will be described. Several of the data-sets mentioned will be analysed in more detail later in section 5.1.

2.4.1 Two Dimensional Experiments

Many experiments in two dimensions, like those of Murayama and Matsuoka[6, 35], Behringer[48, 37], Daniels[49], Combe[50] and coworkers, photograph the grains from outside the plane of the experiment. This technique allows for the collection of the data about individual particles, positions and contacts, and even stresses.

The details of such experiments are often not as simple as this sounds, for example in Combe et al [50] the end of each cylinder is speckled so that both position and rotation can be computed precisely from the resulting images.

Similarly Murayama and Matsuoka[6, 35] and Behringer et al [48] use grains which are *photoelastic*, that is grains whose birefringences vary with stress, photographed between crossed polarisers in order to measure the highly anisotropic stress fields in constrained granular matter.

In three dimensions, there is no "out of the plane" to photograph from, and so more advanced imaging techniques are used to obtain the details of the structure than those from two dimensional work.

2.4.2 Three Dimensional Experiments

2.4.3 X-ray Computed Tomography

Tomography, the technique of taking a slice image from the middle of a three dimensional structure, is widely used in medicine and the sciences.

Early Non-computerised Tomography

The principle of tomography is to use the fact that a three dimensional structure looks different from different angles. The original X-ray technique was designed to work by blurring structures out of the chosen focal plane, by sweeping the detector and film about some small θ close to the perpendicular to the focal plane, and thus producing a clear image only of the pertinent structure. A diagram of a simple geometry which illustrates this principle is included in figure 2.15.

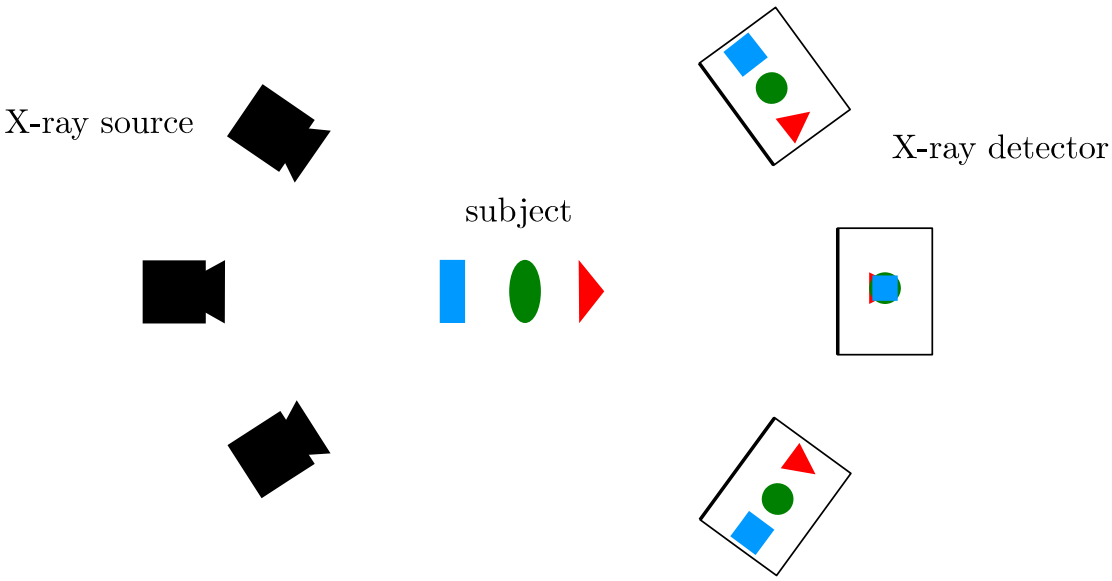


Figure 2.15 – A schematic of an experiment to compute a classic blurred X-ray tomography image. Note that the three different X-ray sources and detectors signify the entire apparatus rotating, and the three images would be blurred so that only the green circle would be in focus. Obviously the colours of the shapes are for illustrative purposes only.

Computerised Tomography

In modern tomography, an image is not formed on an x-ray film as in the original form. Instead, a series of static X-rays are taken (usually using a digital detector) from different angles, like the conventional tomographic sweep discretised. By looking for features that do not move between images, using a reconstruction algorithm in software, a clearer image can be obtained without the superimposed blurs of the other features.

By taking images fully around one half rotation of a sample, rather than around some particular orientation close to a desired focal plane, a fully three dimensional image can be obtained. In principle, this works by using the adjacent few images either side of a particular orientation as the sweep to get a focused image for that plane, and then repeating for each scan plane – correlating such focused images yields a three dimensional image.

2.4.3.1 Granular X-ray CT experiments

Software specialisation for the analysis of granular materials with X-ray CT has been developed. Particular innovations include pore size distribution statistics by Richard et al[8], and contact point identification by Aste et al[51].

The work presented in later chapters benefits greatly from the data provided by Andó et al[2], obtained from a later three dimensional reconstructed X-ray CT experiment. The objective of their work was to see the behaviours of individual grains during the formation of a shear band in a triaxial strain experiment, illustrated in figure 2.16, in which an initially dense sample is stressed *in situ* and scanned during deformation without unloading. Following the collection of the three dimensional images, individual grains were identified using a segmentation procedure (see section 2.2.7,) and correlated between different states – that is, particular grains were identified in multiple different scans and their trajectories between states (translation and rotation) were computed, in a process named ID-Track (similar to a process developed independently by Smit[52].)

The experimental results of Andó et al are very striking – looking at figure 2.18 the shear band can be clearly observed as a signal in grain rotation and displacement. Note that, for clarity, the figure is rendering the properties of grains which intersect a vertical plane with the steepest shear band slope – the data itself is fully three dimensional.

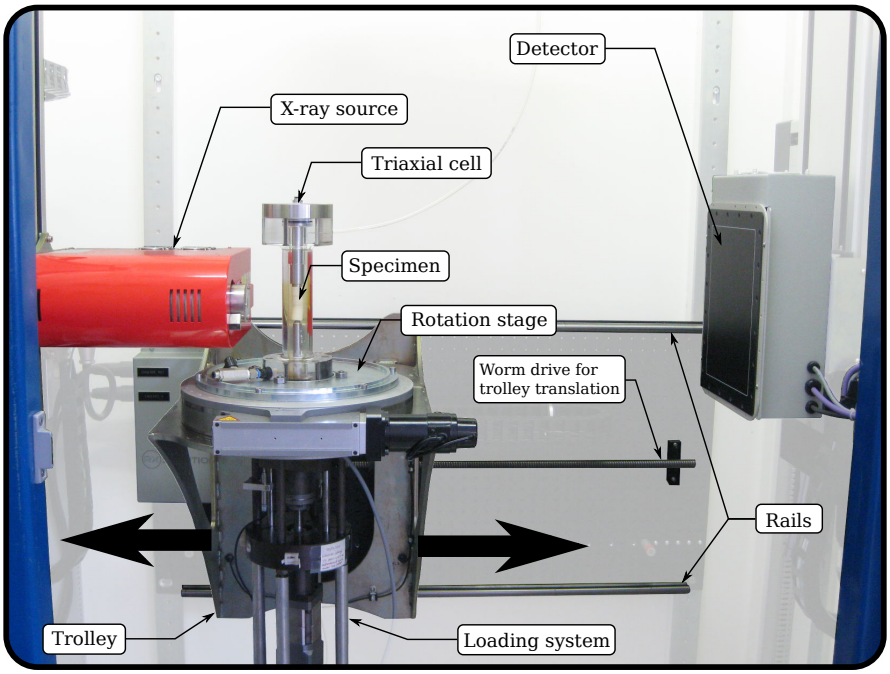


Figure 2.16 – The experimental layout of the Caicos Ooid triaxial strain experiment, supplied by Andó et al (used with permission, see section D)

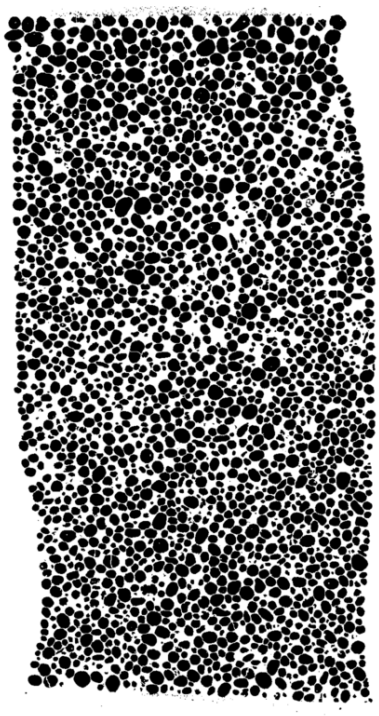


Figure 2.17 – 9th X-ray tomograph image of the Caicos Ooid triaxial strain experiment, supplied by Andó et al (used with permission, see section D). This is a plot of the particle positions, as a reference for figure 2.18 where the same grains have been coloured according to their position and rotation.

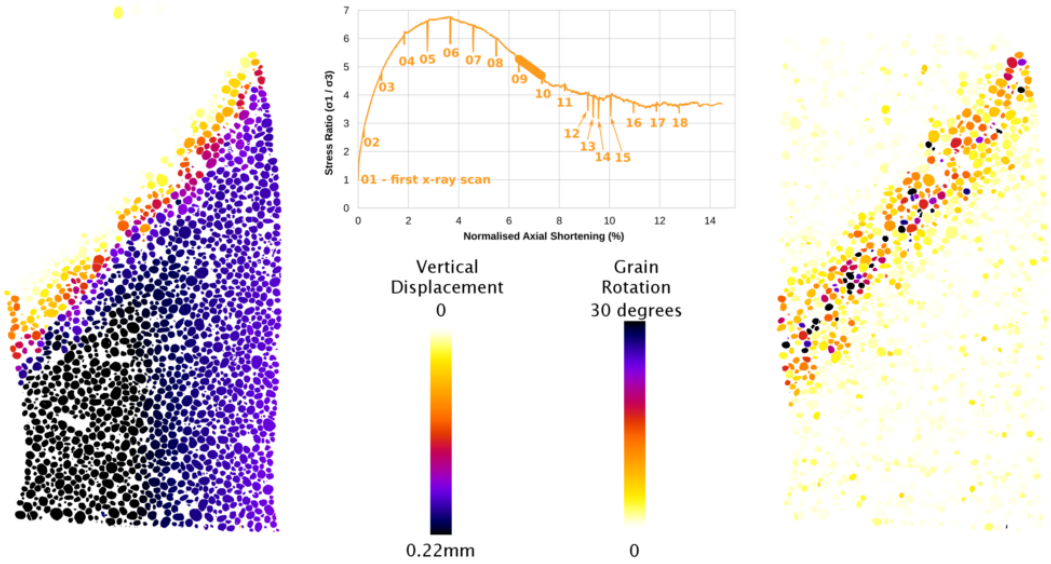


Figure 2.18 – Grain displacement and rotation in the 9th X-ray tomograph image of the Caicos Ooid triaxial strain experiment, supplied by Andó et al (used with permission, see section D). The slice on the left shows the displacement of grains since their starting positions, and on the right are the grain rotations (about any axis.)

The study by Andó et al[2] comes from the tradition of soil mechanics, and is the ultimate extension of the work of Rowe and Horne[11, 12] in that they are tracking the grain displacement and rotation during deformation – the literature from the 1960s was full of debates about how the particles were moving, and now it can be directly observed! The objective of this work is broader in nature, seeking to characterise static structures in general, although the raw data from this experiment is an extremely good test of the tools developed (in both size and complexity.)

2.4.4 Colloidal Optics

Colloidal systems, namely those of small particles suspended in a liquid, can also have similar structures to granular systems, with similar origins of repulsive forces between particles. Brujić et al[3] have developed a technique for measuring the configuration of water droplets in oil, including precisely measuring the contact

positions.

The experiment uses silicone oil droplets tagged with Nile red dye, suspended in a 1:1.05 water to glycerol volume ratio solution (this ratio is chosen so that the two substances have matched refractive indices,) with a sodium dodecyl sulphate surfactant to prevent attraction between droplets. Optically, the sample is excited at a 488nm wavelength, and then two emission wavelengths were detected in a confocal microscope – one at 505-550nm, which revealed the position of the oil droplets, and another at 585nm, which illuminated the interface between droplet and solution. The superposition of the images reveals the contacts as the brightest objects, the droplets as the next brightest and the solution as the darkest. By this technique very accurate system coordinates for particles and contacts were obtained.

The colloid was spun in a centrifuge at different artificial gravities to obtain different packing fractions and arrangements.

2.4.5 Tests of the Statistical Mechanical Theory

2.4.5.1 Irreversibility

In the historic Chicago experiment (Nagel et al[53]), the irreversible nature of granular compaction was demonstrated formally. They measured the density of a shaken granular system as a function of the shaking amplitude, by simply measuring the height of the granular pack in its cylindrical container. They demonstrated an initial irreversible settling during increasing amplitude, followed by reversible behaviour where the density was reproducible during both increasing and decreasing amplitudes, remaining at a higher density than the initial condition.

This result is important to the wider field because it came after the suggestions

of Edwards et al [1], proposing an entropy-based statistical theory. Finding an experimentally reproducible irreversible transition in granular density, although not explicitly predicted by Edwards, provided supporting evidence that the entropy of the number of arrangements of the particles may indeed account for the macroscopic state of the static assemblies produced.

That such a (tomographically) simple experiment yielded such an important result is encouraging, although this work seeks to leverage the micro details of the granular structure to try to solve the mystery that irreversible compaction presents.

2.4.5.2 Computing Entropy by Counting States

D Frenkel and co-workers[54, 55, 56] have developed a way of testing whether Edwards' proposed entropy is valid in practise using computer simulations.

Their method is to enumerate some possible configurations of jammed grains, and determine the probability of the jammed states obtained. For small systems ($N < 20$ grains) it was possible to enumerate all states[54], but for larger systems this quickly becomes impractical and the statistics of the remaining states had to be estimated.

A granular system's configuration can be described by a point in a high-dimensional configuration space (for N static spheres, with no rotational coordinates, this space will contain dN independent coordinates in d dimensional physical space.) If a point in this configuration space is chosen, and a physics engine is used to bring the system into mechanical equilibrium, then one jammed state has been discovered. In order to obtain the probability of this state, D Frenkel et al measure the configuration-space volume (the possible randomly chosen starting configurations) which yield the same jammed state, and divide it by the total volume of configuration space. This can be thought of as a potential energy landscape,

where the system slides down the slopes until it reaches a minimum or basin, in equilibrium – the technique measures the size of that basin. Obviously this technique is only valid for soft spheres, where overlap is allowed, but the jammed states themselves are valid since they have zero or very small overlap.

For the small systems[54], it was computationally feasible to sample all of configuration space, and thus it was deduced that every jammed arrangement was discovered, and the number of times a particular state was obtained yielded the volume of its associated energy landscape basin, after starting the system off at every point on a high resolution lattice in configuration space.

For larger systems, random regions of configuration space were sampled, and once a particular jammed state was obtained, the system used a Monte-Carlo sampling of the surrounding energy space to calculate the basin volume. This method has the drawback that larger basins are more likely to be sampled, and thus smaller basins are less well represented – this biased distribution must be unbiased, but D Frenkel et al[55] report that not enough statistics from such small basins were obtained to draw robust conclusions about the total number of states, or to derive such an unbiased distribution.

In their later study[56] this problem is sidestepped by assuming that the functional form of the unbiased distribution is similar to the complete distribution obtained for smaller systems. This allows a re-defined Edwards entropy (where the probability of each state is weighted rather than equal) to be supported by the work, and critically for this entropy to be extensive, providing evidence supporting a granular statistical mechanics.

2.4.5.3 Voronoï and Delaunay Volumes

Aste et al[57] have shown that, under some assumptions, a k-Γ distribution should model Voronoï tessellations of granular systems. They report that these predictions are supported in a variety of experimental conditions[58, 59, 60]. Their theoretical analysis begins with the statistical mechanical partition function for indistinguishable particles with quantised energies (with the volume being substituted for the energy as proposed by Edwards):

$$Z_{quantum} = \frac{\left[\sum_n \exp\left(-\frac{v - v_{min}}{X}\right)\right]^N}{N!}$$

Where the volume is represented, via a variable substitution, as the volume above some v_{min}, N is the number of *elementary* volume cells, and the sum over n is over all the possible (quantised) volume states. Aste then substitutes the quantum partition function for a "state counting" system, in which the Boltzmann factor is replaced with the width of the continuum space available to the elementary cell volume microstate:

$$Z_{continuum} = Z = \frac{[V_T - N v_{min}]^N}{N!}$$

where V_T is the total volume of the system. This equation counts all of the available states, so the states left in the system can easily be obtained if one cell has a volume v:

$$Z(v) = \frac{[V_T - N v_{min} - (v - v_{min})]^{N-1}}{(N-1)!}$$

By setting $Z(v)/Z$ the probability of a cell with volume v can be found:

$$p(v) = \frac{N}{V_T - N v_{min}} \left(1 - \frac{v - v_{min}}{V_T - N v_{min}}\right)$$

Taking the limit that $N \to \infty$, where V_T/N remains finite, the exponential form is recovered:

$$p(v) = \frac{1}{\chi} \exp\left[-\left(v - v_{min}\right)/\chi\right]$$

with $\chi = \frac{V_T}{N} - v_{min}$, an intensive temperature-analogue. They then argue that any particular choice of *agglomerate* cells may be a sum of the original *elementary* ones, specifically that an agglomerate Voronoï cell V be equivalent to k elementary cells v. Thus, they substitute variables $V = kv$ to obtain $V_{min} = kv_{min}$, and $\langle V \rangle = k \langle v \rangle = kV_T/N$. The probability density function of the volumes of such agglomerate cells, from pure statistical mechanics[61], is:

$$\mathrm{PDF}(V, k) = \frac{k^k}{(k-1)!} \frac{(V - V_{min})^{k-1}}{(\langle V \rangle - V_{min})^k} \exp\left(-k \frac{V - V_{min}}{\langle V \rangle - V_{min}}\right)$$

from which they deduce:

$$\chi = \frac{\langle V \rangle - V_{min}}{k}$$

and discover empirically[58, 59, 60] that $k \approx 12$.

This work is important in the context of granular statistical mechanics because it aims to simplify the quantitative measurement of granular statistics, using an accessible technique (Voronoï partition of the volume). However, there is one criticism of this statistical method.

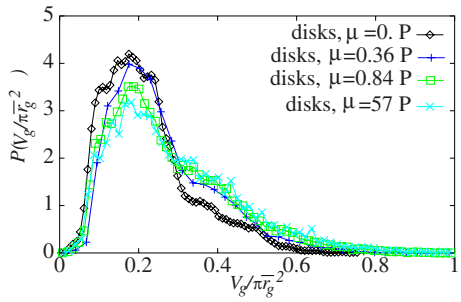
In order to argue that some tessellation's elements represent k-sums of fundamen-

tal elements, it must be assumed that each element contains a whole number, or close to a whole number, of fundamental elements. It follows that the boundaries of the tessellation should coincide with some of the fundamental element boundaries – which means that, far from the fundamental volume elements being unknown, they should be packed neatly inside whichever agglomerate elements are chosen (and which are distributed in k-Γ probability density functions as confirmed by experiment). Since the authors report both Voronoï and Delaunay tessellations fitting into such a distribution, it is either implied that both tessellations fit neatly onto the boundaries of these fundamental cells (which should be enough constraint to deduce them, or something close to them, analytically) or that there is some other origin for k-Γ distributions in Voronoï and Delaunay tessellation cell volumes' probability density functions.

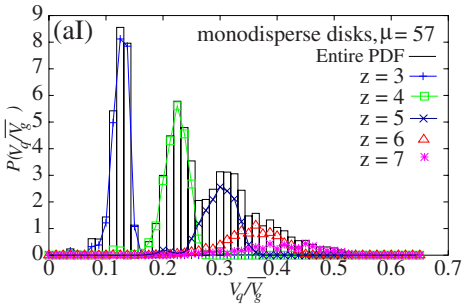
2.4.5.4 Two Dimensional Quadron Volume Distributions

The probability density function for quadron volumes, PDF (v^q) has been measured by G Frenkel et al[62] in two dimensions, where they find shoulders for systems of monodisperse discs as illustrated in figure 2.19a. These they studied by looking at *conditional* quadron volume distributions, PDF $(v^q|e)$, as follows. Taking all the quadrons from a particular cell order e, that is the number of grains in a two dimensional granular cell, and plotting those individual PDFs yields a family of unimodal curves. These peaks correspond to the shoulders in PDF (v^q), as illustrated in figure 2.19b.

The broadening of the peaks with increasing e was accounted for by Toikka[63] and later Hihinashvili[47], who showed that as the cell order gets larger the configurations have more scope to vary the aspect ratio, and therefore the average quadron volume. Some example configurations which illustrate this fact are given in figure 2.20.



(a) Figure 4 from [62], which shows the quadron volume distribution PDF (v^q) from a series of two dimensional granular packs, with different friction coefficients. Note the shoulders above the main peak visible in all the distributions.



(b) Figure 5aI from [62] showing one total volume distribution and its corresponding conditional distributions PDF $(v^q|e)$, which are unimodal curves whose peaks correspond to the shoulders and/or peaks in the total distribution.

Figure 2.19 – Figures taken from Frenkel et al[62], used with permission (see section D.)

Because of mechanical stability, Matsushima and Blumenfeld have argued[64, 65, 66] that cells with higher e are more likely to have aspect ratios closer to 1 (are likely to be more rounded). This should lead to a bias in the PDF $(v^q|e)$ towards higher volumes, compared to the predictions from the geometrical arguments in Toikka[63] and Hihinashvili[47]. They (Matsushima and Blumenfeld) report that this effect is small but measurable.

Some very recent work by Matsushima and Blumenfeld[64, 65, 66] shows that this family of curves can collapse under a suitable normalisation onto the same curve. In particular, they took PDF $(v^q|e)$ from simulated systems of discs generated with different friction coefficients μ and showed that normalised curves collapse independent of μ separately for each value of e, as illustrated in figure 2.21.

All this work suggests that in two dimensional granular systems the conditional probability distributions PDF $(v^q|e)$ are almost entirely accounted for by geometry. However, in the total distribution PDF (v^q) these distributions can be com-

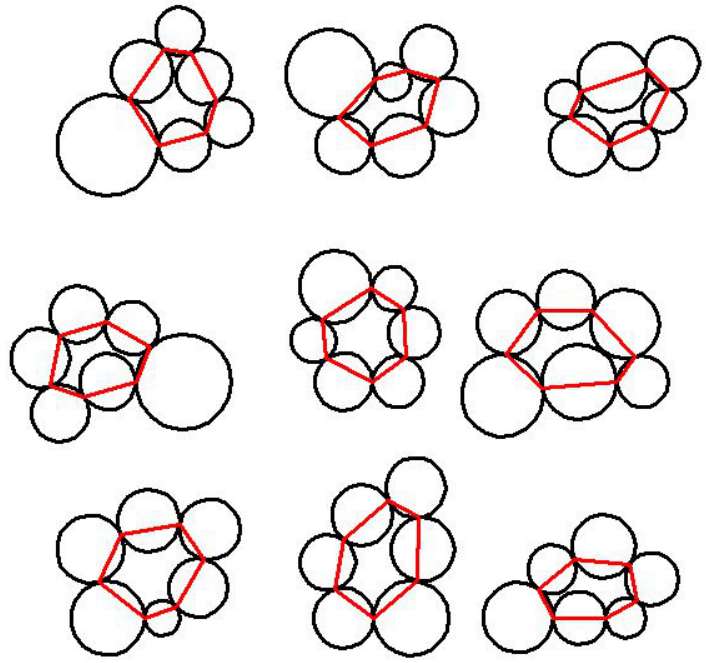
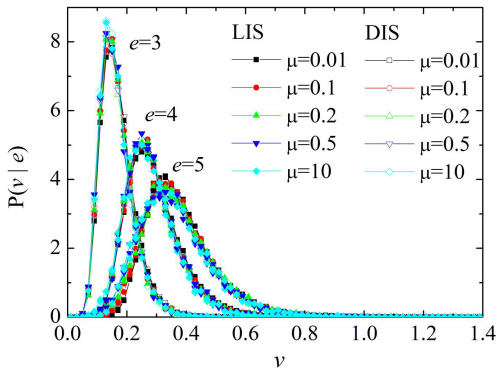
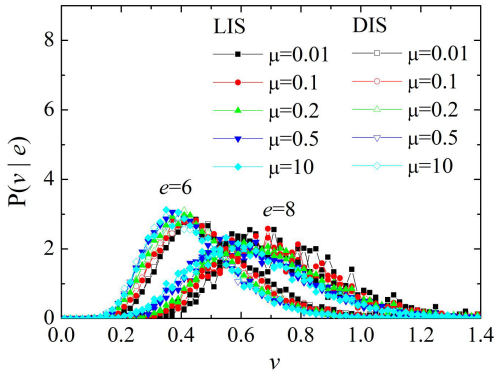


Figure 2.20 – Different geometrical arrangements of cells with $e = 6$, figure 20 taken from [64] (used with permission, see section D.)

bined differently depending on the friction and preparation protocol, having both different weights and different scaling as a function of these factors.



(a) Figure 15 from [64] showing how PDF $(v^q|e)$ for $e = 3, 4, 5$ each collapse onto a curve independent of friction.



(b) Figure 16 from [64] showing how PDF $(v^q|e)$ for $e = 6, 8$ each collapse onto a curve independent of friction.

Figure 2.21 – Figures taken from Matsushima and Blumenfeld[64], used with permission (see section D.)

Chapter 3

Two-Dimensional Methods

One further disadvantage of Edwards' formalism for understanding granular media, beyond the lack of ergodicity mentioned in section 2.3.2, is the lack of a compactometer – a device for directly measuring the granular temperature-analogue, *compactivity*. Such difficulties can be overcome using the tools of statistical mechanics.

In this section, the two dimensional quadron volume elements will be explicitly defined, and then put to use in a straightforward, if technical, application of the method of a weighted sum of probabilities to compute the expectation value of the volume function in a two dimensional isostatic granular pack. The resulting equation can be used to compute the compactivity from other more easily measurable quantities.

3.1 Two Dimensional Quadrons

This section is the author's own description, including some original illustrations, of the definitions of Ball and Blumenfeld[13] concerning the form of the volume

function in two dimensions.

Centroids

Below, some positions are defined purely as averages of sets of contact point positions – such average positions are referred to as *centroids*, $\vec{x}_c$. Usually a centroid is given with respect to a particular object, for example a *grain-centroid* $\vec{g}_c$ would be the average of the positions of the contact points on the surface of a particular grain. Other centroids are defined where similar local objects bounded by contact points arise.

3.1.1 Quadron Definition

Taking the contact points as the vertices of a graph, C, edges are added around the perimeter of each grain as illustrated in Fig 3.1, forming a *grain polygon*.

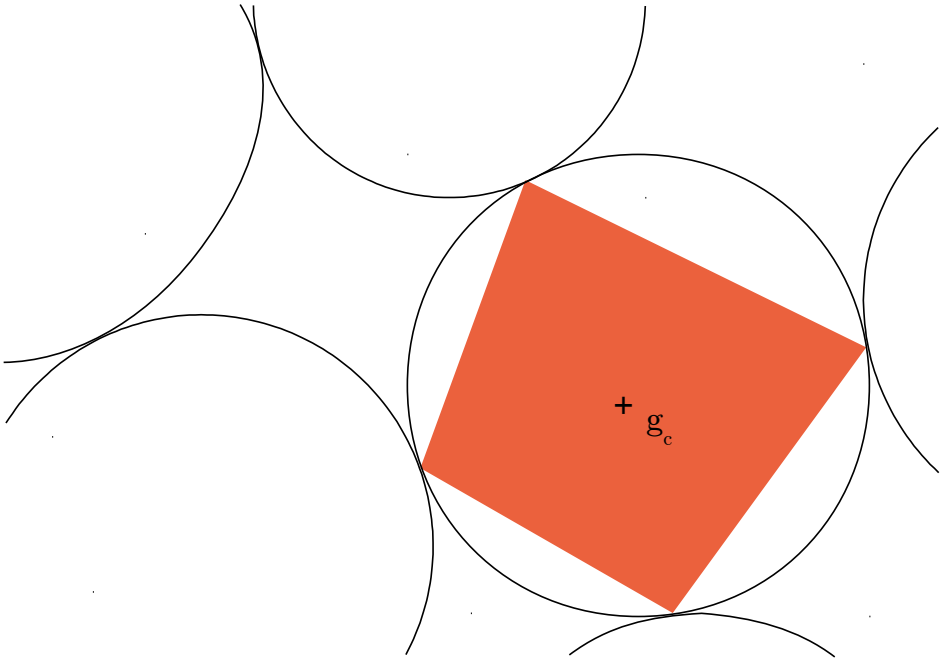


Figure 3.1 – Two dimensional grain polygon.

These grain polygons form faces in this embedding of C. Looking at a wider system

where this procedure has been applied to each grain, we notice that a second set of faces also arises, where cycles of grains (and thus cycles of contact points in C) form a further set of polygons in the spaces between the grain polygons, as illustrated in Fig 3.2. Such emergent polygons are referred to as *cell polygons*. Note that *cell* is used in two dimensions rather than *face*, as is traditional in graph theory when in two dimensions, for consistent terminology and labelling between two and three dimensions for this description.

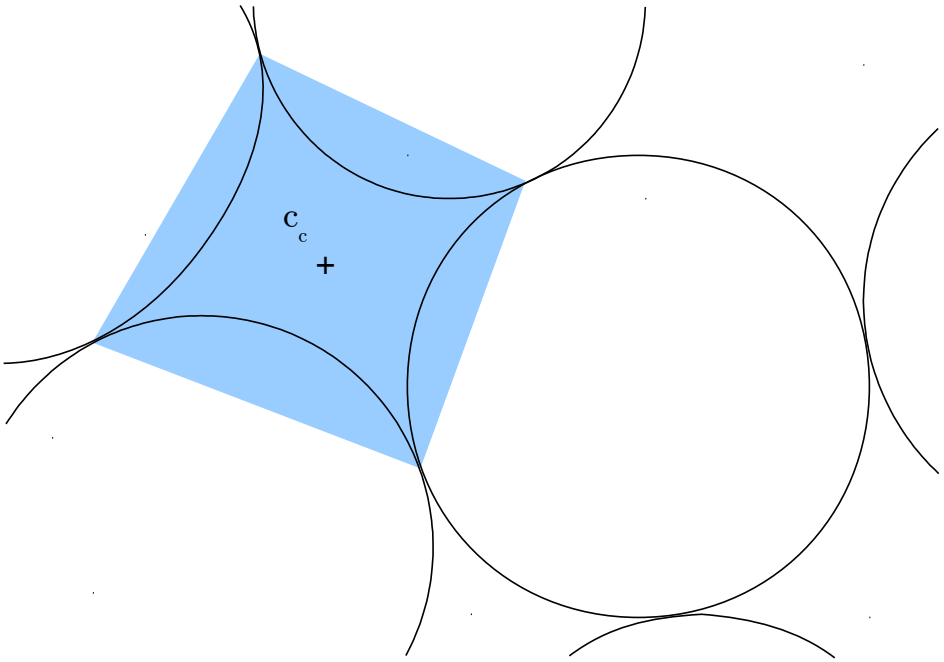


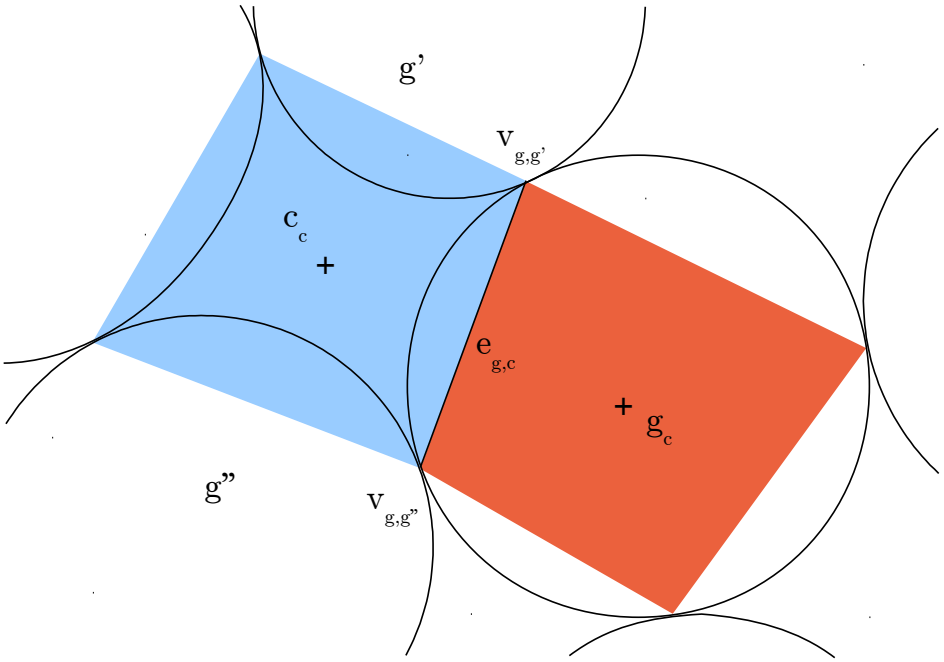
Figure 3.2 – Two dimensional cell polygon.

These faces (the grain polygons and the cell polygons) are further subdivided into tessellating quadrilaterals, *quadrons* for short, whose number is identical to the number of structural degrees of freedom, as follows.

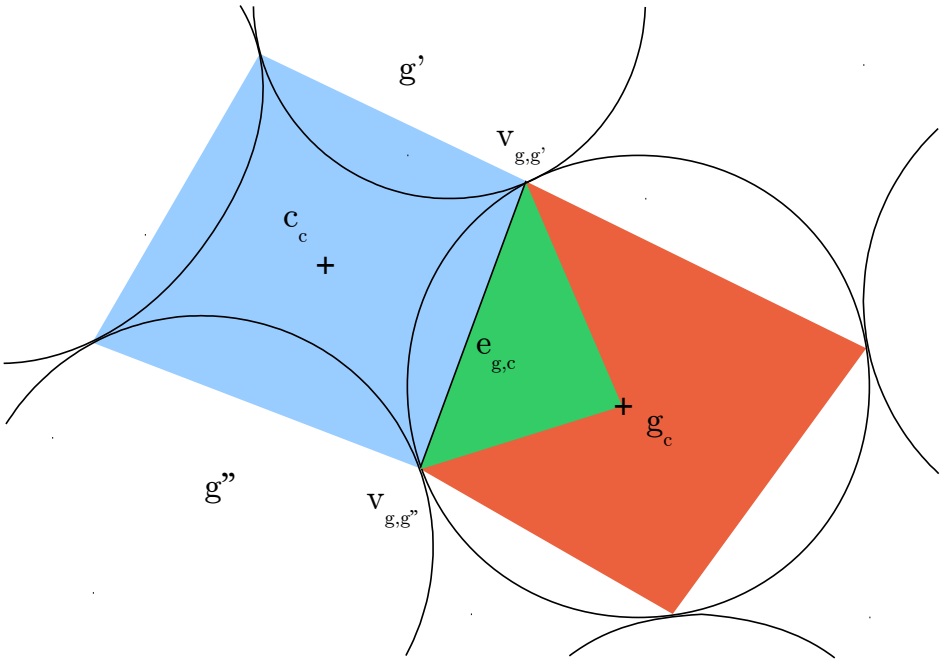
Take an edge e from $E(C)$, and notice that this edge is in exactly two faces, one grain polygon and one cell polygon (see Fig 3.3a.) Thus, the edge can be uniquely indexed as $e_{g,c}$. To construct a tessellating quadrilateral, two triangles are combined with this $e_{g,c}$ as their common edge. The first triangle extends into the grain, and consists of the two contact points at each end of $e_{g,c}$, $v_{g,g'}$ and $v_{g,g''}$,

and the grain centroid g_c (see Fig 3.3b.) The second triangle extends into the cell, and consists of the same two contact points, $v_{g,g'}$ and $v_{g,g''}$, and the cell centroid c_c (see Fig 3.3c.)

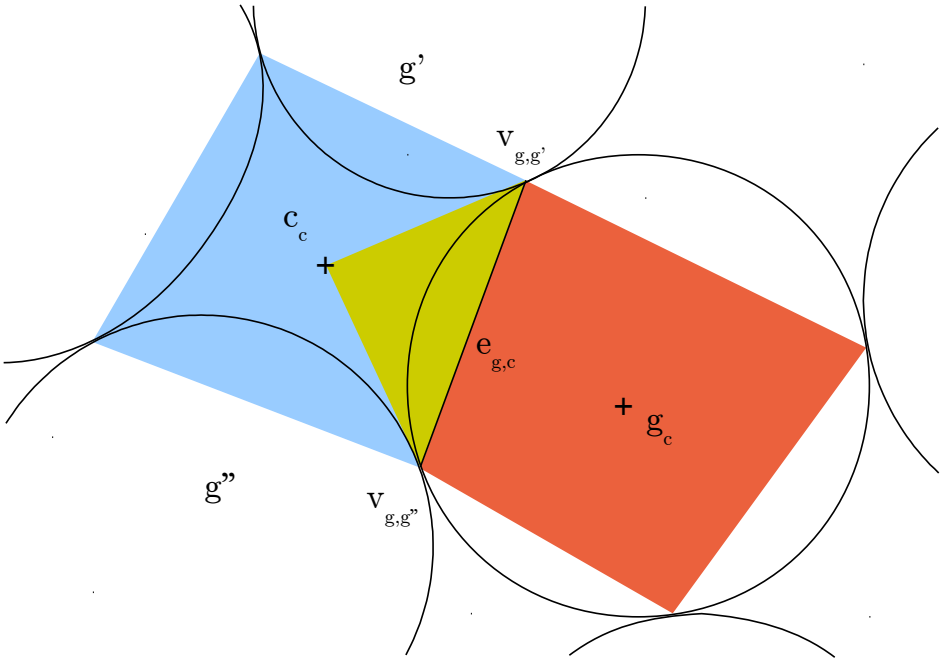
The union of these two triangles, a quadrilateral the set of which over the system tessellate the volume, is named a quadron. Two vectors are defined[13] to make the quadron an even more useful tool for analysis; $\vec{r}$, which is the edge $e_{g,c}$ with a direction defined by a global convention, and $\vec{R} = \vec{c_c} - \vec{g_c}$ (see Fig 3.3d.)



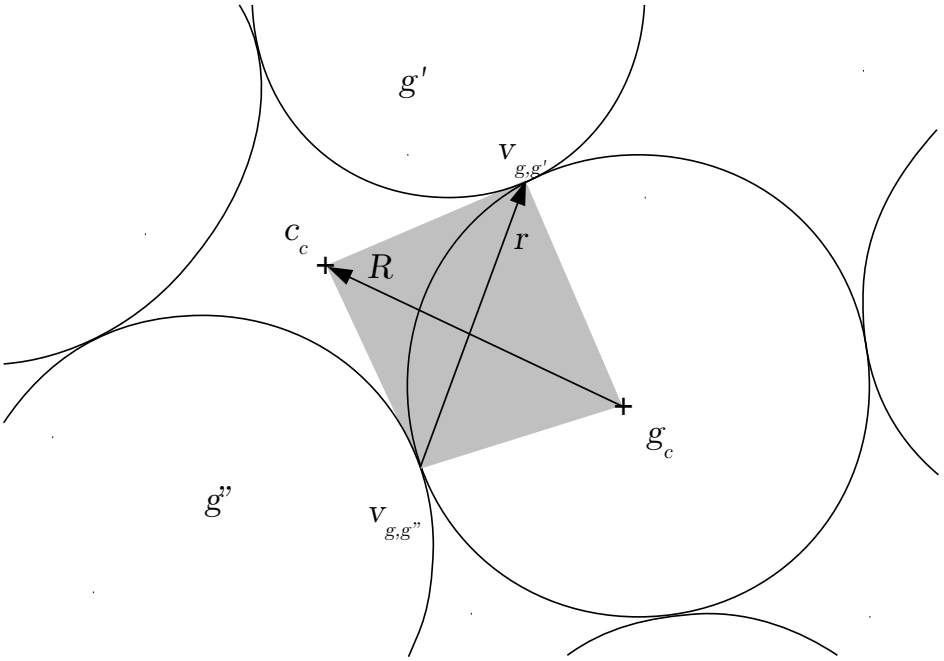
(a) The shared edge (from contact point graph C) unique to a particular cell and grain.



(b) Define a grain triangle using the shared edge and the grain's centroid.



(c) Define a cell triangle using the shared edge and the cell's centroid.



(d) The union of these two triangles form a quadrilateral tessellating volume, quadron, which can be encapsulated using the two labelled diagonal vectors $\vec{r}$ and $\vec{R}$.

Figure 3.3 – Two dimensional quadron construction.

The structure tensor Note also that $\vec{r}$ and $\vec{R}$ are both defined using only contact point positions. This is trivial to see for $\vec{r}$, since it is the difference between two contact point positions, but for $\vec{R}$ recall that the centroids are themselves averages of (cycles of) contact point positions, and thus $\vec{R}$ is also an, admittedly more complex, linear combination of contact point positions[14].

A locally unique structure tensor can now be defined for each quadron[67] (index q), which encapsulates all the shape and orientation information:

$$\hat{C}_{ij}^{q} = r_i^q R_j^q \tag{3.1}$$

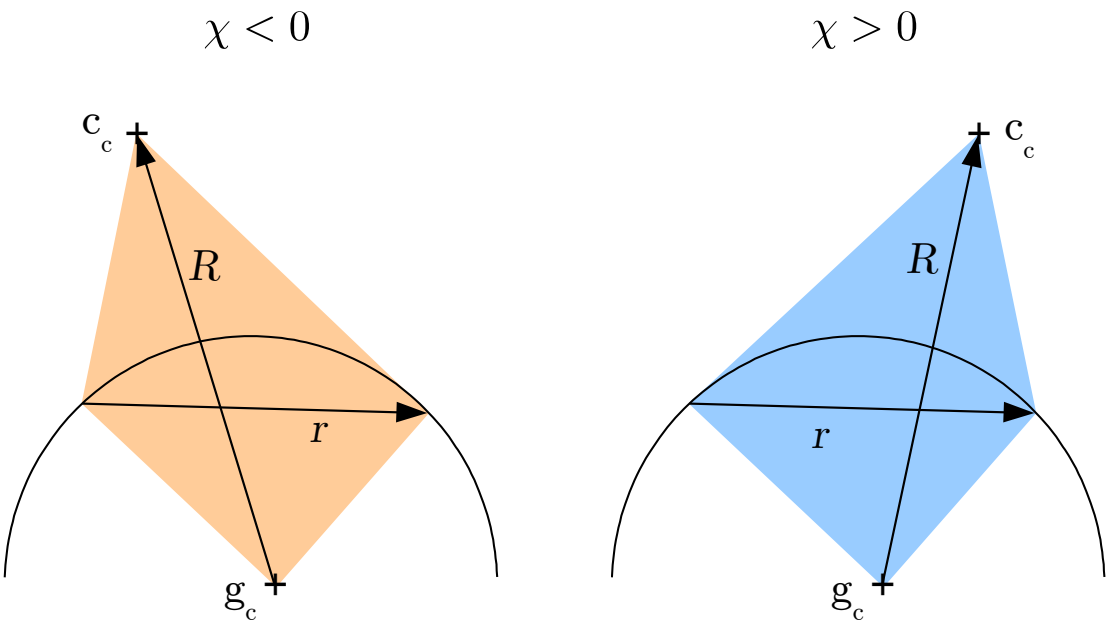
An example of extracting useful information from this tensor is computing the volume using:

$$v^q = \frac{1}{2}\left|\vec{r}^{\,q} \wedge \vec{R}^q\right| = \frac{1}{2}\,\mathrm{Tr}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \hat{C}_{ij}^{q}\right) \tag{3.2}$$

The structure tensor can also reveal a quantity called the chirality, χ, which is a measure of which way (in the quadron's own frame) it is twisted. It can be expressed in terms of the symmetric part of $\hat{C}_{jk}$ as follows:

$$\chi = \vec{r}^{\,q} \cdot \vec{R}^q = \mathrm{Tr}\,\frac{1}{2}\left(\hat{C}_{jk} + \hat{C}_{jk}^{T}\right) \tag{3.3}$$

This definition matches that given by Ball and Blumenfeld[13]. The chirality can be easily interpreted using figure 3.4.



(a) A negative chirality quadron in two dimensions.

(b) A positive chirality quadron in two dimensions.

Figure 3.4 – Values of chirality in two dimensions illustrated.

3.2 Equipartition of Volume in Two Dimensional Isostatic Granular Media

Boltzmann's statistical mechanics, as was illustrated in section 2.3.2, allows one to compute expectation values of quantities related to the degrees of freedom of the system by performing a sum over microstates weighted by their likelihood. Such techniques are also applicable to Edwards' granular statistical mechanics, with the volume function substituted for the Hamiltonian, as was defined in section 2.3.2.

In two dimensions, where the volume function $\mathcal{W}$ is in units of length squared, the granular partition function is Gaussian. In order to evaluate such integrals, we refer to the standard results in section B.1, and to the parameterised form in section B.3.

3.2.1 Derivation of a Granular Equipartition Relation

3.2.1.1 Simplifying the Combined Partition Function

Expectation value expressions from the combined force and volume ensemble partition function can now be derived, as presented by Blumenfeld et al[15]. This section refers only to problems in two dimensions, except where explicitly stated otherwise.

It is argued that the volume elements which make up the sum in the volume function are the quadron volumes as defined in section 3.1. An individual quadron volume can be expressed as a function of its two vectors, and thus the volume function represented as a sum over these individual volume elements:

$$\mathcal{W} = \sum_q \frac{1}{2} \left| \vec{r}^{\,q} \wedge \vec{R}^q \right| \tag{3.4}$$

Since the definition of $\vec{R}$ is the difference between averages of contact positions, which can be expressed as linear combinations of other $\vec{r}$s, a sum over such quadron volumes can be written as a quadratic form where $\vec{x}$ is all the $\vec{r}$s in one vector, and D is a matrix which contains all the cross products and linear combinations:

$$\mathcal{W} = \frac{1}{2} \vec{x} D \vec{x} \tag{3.5}$$

All the $\vec{r}$s, however, do not need to be considered since by the definition of the contact graph C, they form loops around cells and grains. This means that the $\vec{r}$s are not all independent, and therefore the expression $\vec{x}$ over counts the degrees of freedom. An independent subset of them is defined, containing exactly one $\vec{r}$

pointing to each contact position[1], and placed in a new combined vector $\vec{\rho}$ which has length $N\bar{z}$. Under this change of basis, we rewrite D as a new matrix A, which contains all the same sums and cross products in terms of the $\vec{r}$s that were included in $\vec{\rho}$. We also include in A a factor of $1/X$, the inverse compactivity.

$$\frac{\mathcal{W}}{X} = \frac{1}{2}\vec{\rho}A\vec{\rho} \tag{3.6}$$

Thus, having computed the form of the structural degrees of freedom, the volume partition function can be presented:

$$Z_v = \int \exp\left(-\frac{1}{2}\vec{\rho}\,A\,\vec{\rho}\right)\left(\mathrm{d}^{N\bar{z}}\vec{\rho}\right) \tag{3.7}$$

However, this expression does not account for *all* degrees of freedom, only the structural degrees of freedom. Next, the stress ensemble must be taken into account.

Considering an isostatic granular system, that is for two dimensions where the average coordination number $\bar{z} = 3$, Newton's equations can fully determine the forces between all particles given only the forces at the boundary of the system. It is this class of system that is considered in the following analysis.

The force moment function (as alluded to in section 2.3.2.1) $\mathcal{F}_{\alpha\beta}$ is an additive Hamiltonian-equivalent for encapsulating the stress states. Summing over all quadrons, and where the loop force $\vec{f}$ is as defined in [13] the relevant expression for the force moment is:

$$\mathcal{F}_{\alpha\beta} = \sum_q \vec{f}^q_\alpha \vec{r}^q_\beta \tag{3.8}$$

[1]this independent subset is a spanning tree of C, as mentioned in section 2.1.4.1.

Recalling that the loop forces are linear combinations of the inter-granular forces, which themselves can be expressed in terms of the $M = O(\sqrt{N})$ perimeter boundary forces (which are placed sequentially, as before, in one large vector $\vec{g}$), the expression for the force moment divided by its tensorial temperature analogue $X_{\alpha\beta}$, the angoricity, can be re-written as follows:

$$\sum_{\alpha,\beta} \frac{\mathcal{F}_{\alpha\beta}}{X_{\alpha\beta}} = \vec{g} C^T E^T B \vec{\rho} \tag{3.9}$$

Where B is the inverse angoricity components, E transforms from inter-granular forces to loop forces, and C from loop forces to boundary forces. Using shorthand $Q = B^T E C$, the two partition functions can be combined:

$$Z_{combined} = \iint \exp\left(-\frac{1}{2}\vec{\rho}\, A\, \vec{\rho} - \vec{g}\, Q^T\, \vec{\rho}\right) \left(\mathrm{d}^{N_{\tilde{z}}}\vec{\rho}\right) \left(\mathrm{d}^{2M}\vec{g}\right) \tag{3.10}$$

This integral can be solved analytically in spite of the mixture of linear and quadratic terms in the exponent by completing the square. Changing basis $\vec{\tilde{\rho}} = \vec{\rho} + A^{-1} Q \vec{g}$, the exponent becomes:

$$-\frac{1}{2}\vec{\rho}\, A\, \vec{\rho} - \vec{g}\, Q^T\, \vec{\rho} = -\frac{1}{2}\vec{\tilde{\rho}}\, A\, \vec{\tilde{\rho}} + \frac{1}{2}\vec{g}\, Q^T A^{-1} Q\, \vec{g} \tag{3.11}$$

now, relabelling $P = Q^T A^{-1} Q$, and noting that since the integral is being performed in $\vec{\rho}$ over an unbounded interval we can shift it by a constant factor $A^{-1}Q\vec{g}$ without altering the result, we obtain:

$$Z_{combined} = \iint \exp\left(-\frac{1}{2}\vec{\tilde{\rho}}\, A\, \vec{\tilde{\rho}} + \frac{1}{2}\vec{g}\, P\, \vec{g}\right) \left(\mathrm{d}^{N_{\tilde{z}}}\vec{\tilde{\rho}}\right) \left(\mathrm{d}^{2M}\vec{g}\right) \tag{3.12}$$

This expression makes it appear as if the equipartition result[15] is within reach

– two orthogonal quadratic terms in our partition function are surely soluble analytically using standard Gaussian results (see section B.1). However, note the sign of the $\vec{g}\,P\,\vec{g}$ expression is positive, not negative, leading to a divergence in the unbounded integrals. Since these expressions will be both on the top and bottom of the expectation value equation, it is helpful to write them in terms of Dawson functions (see section B.2,) which may then cancel some of the time. It is also necessary to set the limits of the boundary force integral to some finite $g_{\max}$, in order to evaluate the partition function:

$$Z_{combined} = \sqrt{\frac{(2\pi)^{N\bar{z}}\,2^{2M}}{|A||P|}} \prod_{i=0}^{2M} \mathrm{Daw}\left(\sqrt{\frac{p_i}{2}}\,g_{\max}\right) \tag{3.13}$$

where p_i is the ith eigenvalue of P. Because of the definition of Daw () in equation B.7, this value may diverge. Placing a limit on the size of the boundary forces does not prevent any valid microstates from being enumerated – the force ensemble only sums over the possible boundary forces which keep the macroscopic stress the same, and which do not cause re-arrangements and change the value of A.

3.2.1.2 Obtaining the equipartition relation

Having identified a means to simplify the combined partition function, it is now possible to compute an expectation value, the volume expectation. The volume function $W(\vec{\rho})$ has already been used in the partition function in the form:

$$\frac{\mathcal{W}(\vec{\rho})}{X_0} = \frac{1}{2}\vec{\rho}\,A\,\vec{\rho} \tag{3.14}$$

and thus the expectation value calculation can simply be written as:

$$\langle V \rangle = \frac{\iint \frac{X_0}{2} \vec{\rho}\, A\, \vec{\rho} \exp\left(-\frac{1}{2}\vec{\rho}\, A\, \vec{\rho} - \vec{g}\, Q^T\, \vec{\rho}\right) \left(\mathrm{d}^{N\bar{z}}\vec{\rho}\right) \left(\mathrm{d}^{2M}\vec{g}\right)}{\iint \exp\left(-\frac{1}{2}\vec{\rho}\, A\, \vec{\rho} - \vec{g}\, Q^T\, \vec{\rho}\right) \left(\mathrm{d}^{N\bar{z}}\vec{\rho}\right) \left(\mathrm{d}^{2M}\vec{g}\right)} \tag{3.15}$$

Performing the coordinate transformation from Equation 3.11, we obtain:

$\langle V \rangle =$

$$\frac{\frac{X_0}{2} \iint \left(\vec{\tilde{\rho}} - A^{-1}Q\vec{g}\right)^T A \left(\vec{\tilde{\rho}} - A^{-1}Q\vec{g}\right) \exp\left(-\frac{1}{2}\vec{\tilde{\rho}}\, A\, \vec{\tilde{\rho}} + \frac{1}{2}\vec{g}\, P\, \vec{g}\right) \left(\mathrm{d}^{N\bar{z}}\vec{\tilde{\rho}}\right) \left(\mathrm{d}^{2M}\vec{g}\right)}{\iint \exp\left(-\frac{1}{2}\vec{\tilde{\rho}}\, A\, \vec{\tilde{\rho}} + \frac{1}{2}\vec{g}\, P\, \vec{g}\right) \left(\mathrm{d}^{N\bar{z}}\vec{\tilde{\rho}}\right) \left(\mathrm{d}^{2M}\vec{g}\right)} \tag{3.16}$$

Expanding the integrand, cancelling linear terms because they are odd functions, and separating out the two integrals (since the exponential multiplier for the other variables simply cancels for each term in the sum,) leaves:

$$\begin{aligned} \langle V \rangle = & \frac{\frac{X_0}{2} \int \vec{\tilde{\rho}}\, A\, \vec{\tilde{\rho}} \exp\left(-\frac{1}{2}\vec{\tilde{\rho}}\, A\, \vec{\tilde{\rho}}\right) \left(\mathrm{d}^{N\bar{z}}\vec{\tilde{\rho}}\right)}{\int \exp\left(-\frac{1}{2}\vec{\tilde{\rho}}\, A\, \vec{\tilde{\rho}}\right) \left(\mathrm{d}^{N\bar{z}}\vec{\tilde{\rho}}\right)} + \\ & \frac{\frac{X_0}{2} \int \vec{g}\, P\, \vec{g} \exp\left(+\frac{1}{2}\vec{g}\, P\, \vec{g}\right) \left(\mathrm{d}^{2M}\vec{g}\right)}{\int \exp\left(+\frac{1}{2}\vec{g}\, P\, \vec{g}\right) \left(\mathrm{d}^{2M}\vec{g}\right)} \end{aligned} \tag{3.17}$$

Here one term is a synonym of an identity presented earlier in equation B.4 and another can be expressed in terms of the Dawson functions. The former, the fraction in $\vec{\tilde{\rho}}$, simply evaluates to:

$$\frac{X_0 \bar{z} N}{2}$$

but the latter expression, let it be called $\mathscr{F}$, results in:

$$\frac{2}{X_0}\mathscr{F} = \sum_i^{2M} \frac{\int_{-g_{\max}}^{g_{\max}} p_i \vec{g}_i^2 \exp\left(+\frac{p_i}{2}\vec{g}_i^2\right) \mathrm{d}\vec{g}_i}{\int_{-g_{\max}}^{g_{\max}} \exp\left(+\frac{p_i}{2}\vec{g}_i^2\right) \mathrm{d}\vec{g}_i}$$

This term could be written in terms of the Dawson functions, as defined in section B.2, but since this is an expression with multiple such integrals, it may be simpler to use the exact form for definite integrals introduced in section B.3:

$$\begin{aligned} \frac{2}{X_0}\mathscr{F} &= \sum_i^{2M} \frac{2g_{\max} p_i \left(\vec{g}_i\right)_0^2 \exp\left(+\frac{p_i}{2}\left(\vec{g}_i\right)_0^2\right)}{2g_{\max}\exp\left(+\frac{p_i}{2}\left(\vec{g}_i\right)_{0'}^2\right)} \\ &= \sum_i^{2M} p_i \left(\vec{g}_i\right)_0^2 \exp\left[\frac{p_i}{2}\left(\left(\vec{g}_i\right)_0^2 - \left(\vec{g}_i\right)_{0'}^2\right)\right] \end{aligned} \tag{3.18}$$

Where we have defined $\left(\vec{g}_i\right)_0$ as the value which gives the average of the quadratic Dawson function (the numerator of $\mathscr{F}$) for the particular p_i eigenvalue, and similarly $\left(\vec{g}_i\right)_{0'}$ for the non-quadratic Dawson function in the denominator. Which of these two values is larger depends on p_i and $g_{\max}$, since although the quadratic Dawson increases faster, it is smaller for $p_i g_i^2 < 1$.

The sum in equation 3.18 is a constant c_f of order 1, a new granular equipartition relation is obtained, indicating that the expected volume contains contributions from each degree of freedom, regardless of its type:

$$\langle V \rangle = \left(\bar{z}N + 2Mc_f\right)\frac{X_0}{2} \tag{3.19}$$

3.2.1.3 Notes on the Interpretation of the new Equipartition Relation

It is important to point out several things about this new relation, in particular its 2D specificity, and the fact that this method of describing the forces is only valid in isostatic systems, over-connected systems have additional degrees of freedom which are unaccounted for here. That said, the relation can be used to investigate behaviour of the compactivity which may well hold for more general systems, including those in higher dimensions.

For simulated and carefully-measured experimental systems this relation functions as a compactometer, since it defines the compactivity in terms of the observable quantities in the average volume, the number of contacts and the number of boundary forces. We can further see that Edwards' original definition of the compactivity with a zero and infinite limit cannot hold for granular systems, because the limits of the compactivity are proportional to finite expected volumes of quadrons, and as such both limits must be finite and dependant on the particle size. Finally, we can see that the compactivity is only comparable on systems with the same or similar sized grains — we can change the compactivity of a system without any rearrangement simply by scaling the coordinates (and thus the volume.) It is relatively simple to normalise, for example by the mean quadron volume, if systems of such different scales are to be compared statistically.

Chapter 4

Three-Dimensional Methods

In two dimensions any complex system can be represented as a planar graph embedding, which by definition has a dual. Another way of describing this is to say that there is only one way to define a dual network of pores or cells in 2D granular systems, since each dual cell is completely separated from the others.

In contrast, real-world three dimensional granular materials form permeable beaches and cemented ones make up porous rock like sandstone. The property of permeability necessarily means that in order to construct the quadron tessellation one must divide the pore space up artificially into cells – a step-change in difficulty of the problem from two to three dimensions. This forms the most pressing problem in three dimensional granular media — the computation of a suitable non-planar dual, as was introduced in section 2.3.3.3. Indeed, no prior work has ever presented an analysis of the quadron statistics in three dimensions for grains.

In this section, the three dimensional quadron description will be precisely defined, as by Bumenfeld and Edwards[14], using original diagrams devised by the author, and including additional information which has not been explicitly defined in the literature (see section 4.2.2.6. Such a description applies to both foams and cellu-

lar structures as well, where the problem of finding a dual does not arise because of the well-defined divisions between bubbles or cells. Two existing approaches, introduced in the Literature Review in section 2.2.7, for computing duals in undivided three dimensional systems were implemented and their limitations will be discussed.

A new software procedure was developed to solve this problem, and in its simplest form is immediately applicable to foams (that is, it can recover the bubbles of a system after they have been stripped.) Further modifications were necessary to ensure the software's reliability in granular matter, and as part of this work edge cases and exceptions to the original three dimensional quadron description were discovered, and the description modified to admit such cases. This additional work is presented as a narrative, demonstrating each problem observable in granular systems and the solution devised for it (and the modifications required to the software procedure.)

Finally, the overall procedure is summarised, and its remaining limitations are discussed.

4.1 Preliminary Mathematical Definitions

In order to define and compare these methods for computing a 3D dual, a few terms must be defined.

4.1.1 The Granular Graph

Although grains in a mechanically stable granular configuration may have some definition of a position, e.g. the position of the centre of mass, we do not use this information to define our first graph. Instead, we use only the identities of

distinct grains, and which grains they have a contact-point with (their adjacency.) Graph G can be defined as

$$G \equiv \begin{cases} \text{vertices} & = \text{grains} \\ \text{edges} & = \text{contacts between grains} \end{cases}$$

This graph will be referred to as G or *the granular graph* interchangeably in this chapter. This graph does not have a particular embedding, although it may be helpful to consider an embedding using the centroids of the grains as the vertex positions, and interpolating the edges as straight lines between these points.

4.1.2 The Contact Point Graph

The structural degrees of freedom can be defined[67] by the positions and identities of the contact points between grains. Following Blumenfeld and Edwards[14] further, the full 3D quadron tessellation is constructed, with a polyhedron computed for each grain, where the vertices on the surface are the contact points of that particular grain. Such a polyhedron can be defined by computing the convex hull of the contact point positions, via the 3D Delaunay tetrahedralisation's external facets[1]. Having computed such a hull for each grain, a Graph embedding C can be defined as

$$C \equiv \begin{cases} \text{vertices} & = \text{contact points (including positions)} \\ \text{edges} & = \text{convex hull edges from all grains} \\ \text{faces} & = \text{convex hull facets from all grains} \end{cases}$$

[1] Usually, a convex hull is defined as a *closed set* of points, which includes all the points inside this polyhedron, i.e. the union of the Delaunay tetrahedra. It is used in this work to mean the boundary of that set, subdivided into facets.

This is exactly the arrangement of grain polyhedra defined by Blumenfeld[14]. This embedded graph will be referred to as C and *the contact point graph* interchangeably in this chapter. It provides an unambiguous definition for the boundary between the solid and pore phases.

4.2 Three Dimensional Quadrons

Recalling the three dimensional duality relationships (Table 2.1b) it is clear that in three dimensions the definition of a tessellating volume element dependent only on the structural degrees of freedom is more complicated. Consider initially graph G, embedded in $\mathbb{R}^3$ such that $V(G)$ are placed at their corresponding grains' positions and $E(G)$ are straight line segments between those positions (even though, in general, contacts may exist off this line). It is to *this* graph that faces and cells which define a three-dimensional dual must be added. The rest of this region of $\mathbb{R}^3$ not accounted for by these lines and points can be divided up by "cutting" the space along unpenetrated cycles of edges. If such cycles are considered *faces* of G, and enough such faces are chosen, then a region of $\mathbb{R}^3$, $\mathbf{S}$ will be enclosed by such faces at a boundary, and any remaining internal faces will further subdivide $\mathbf{S}$ into *cells* $\{c_i\}$.

Having obtained a graph G with an appropriate structure for a three dimensional duality relationship, using some procedure to choose a set of faces to cut along, the equivalent structures required in C can be deduced as follows. Since each grain surface (convex hull) is enclosed, the only remaining problem is the contiguous pore space, which has already been divided into cells of G, $\{c_i\}$. Taking the faces of G, we can find equivalent cycles in C, which must never run along a convex hull for more than one edge (any such cycle in G would not make a sane face.) These cycles in C cut the pore space into cells corresponding to $\{c_i\}$, and they

are referred to in this theory as *throats*, which is a term borrowed from the study of flow in porous media. An example of a cell in C is illustrated in figure 4.1.

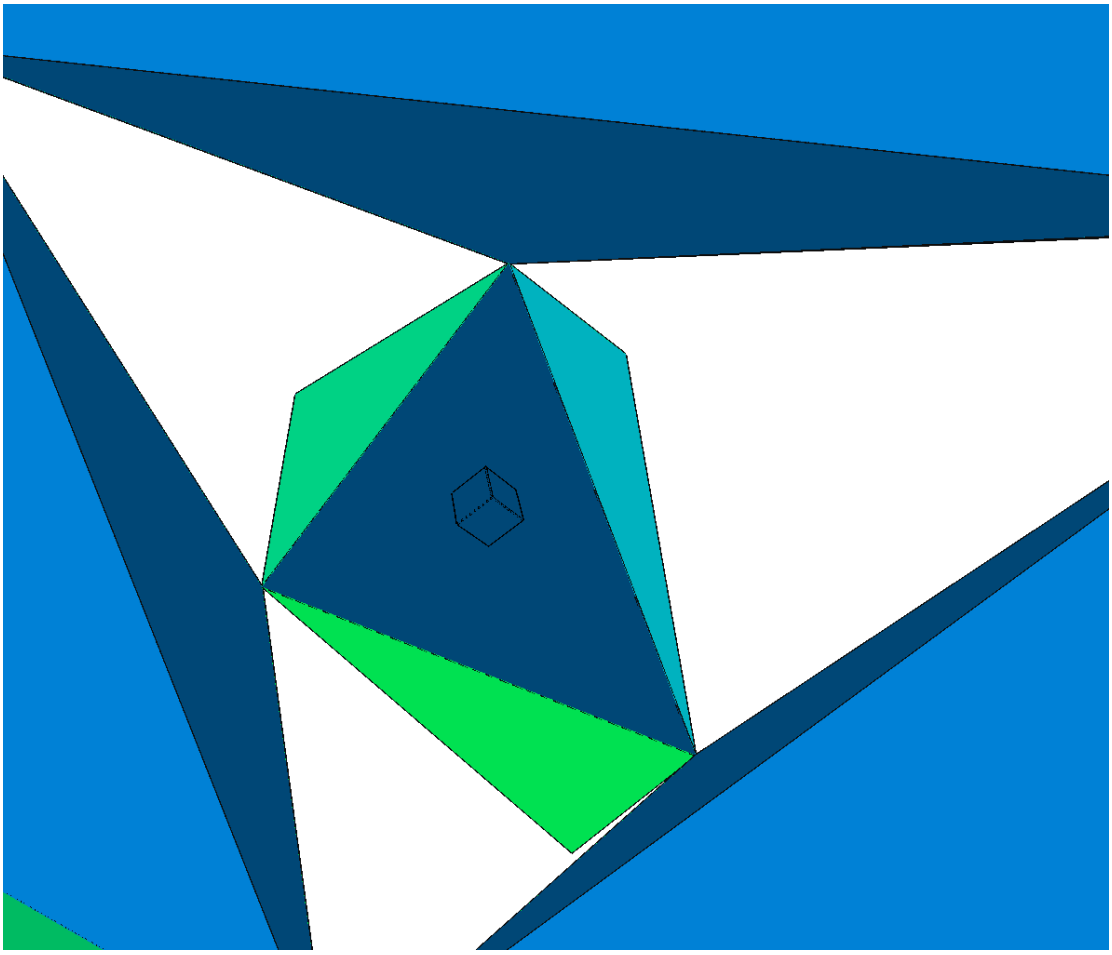


Figure 4.1 – A cell in a granular system made up of grain facets (dark blue), and with the cell centroid rendered as a cube in the same colour. Grain polyhedra facets pointing to other cells can be seen in other colours.

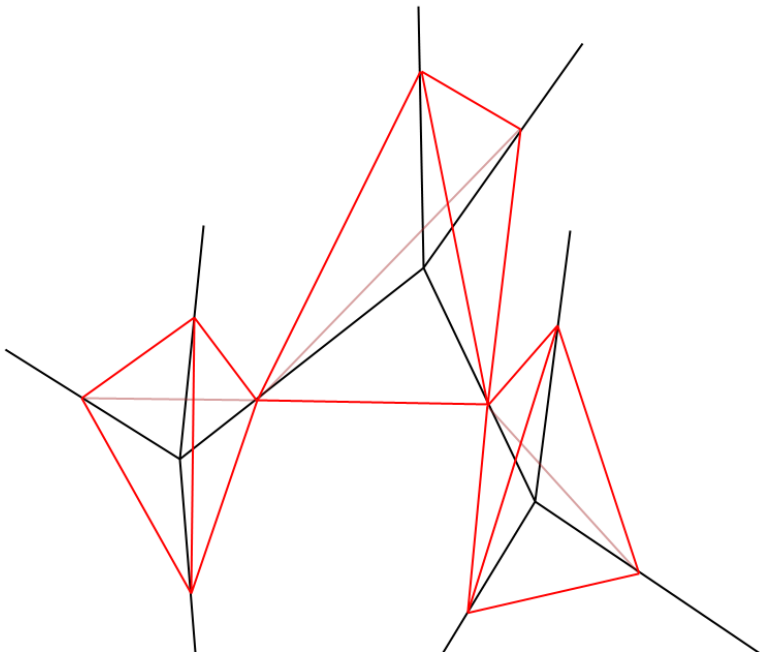
4.2.0.1 Three Dimensional Visualisations

In order to solve the computational problems in finding a non-planar dual which meets all the constraints which will be enumerated shortly, it quickly became necessary to visualise particular three dimensional structures. To this end, an original three dimensional viewer was implemented as part of this work using Python[68] and OpenGL[69]. Many of the figures in this work are screen captures from this viewer, including the illustration of a granular cell in figure 4.1.

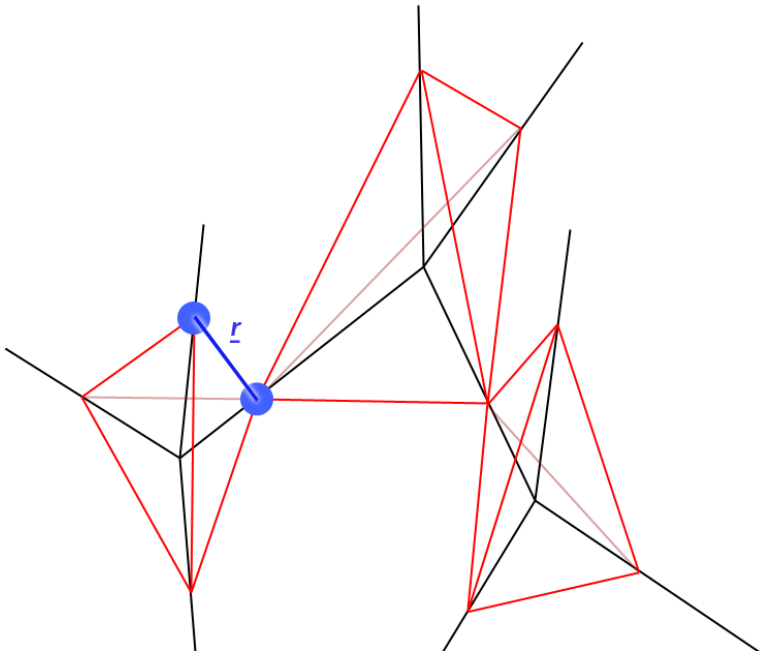
4.2.1 Ideal Quadrons in Three Dimensions

All of the necessary pieces of information required to define three dimensional quadrons are now present: the grain has a *facet*, f_i of its convex hull pointing to a particular *cell*, c_i. Each such facet is bounded by *edges* $E(C)$, which can now be identified as the $\vec{r}$ of three dimensions, and these vectors form the edges of a *throat*, t_i, which forms a non-planar polygon dividing c_i from the adjacent c_j. Using centroids from each of these entities, the three dimensional $\vec{R}$ can be identified as the vector from the grain to the cell centroid, that is $\vec{R} = \vec{c_c} - \vec{g_c}$, and a new vector $\vec{\xi}$, defined as the vector from the facet centroid to the throat centroid, that is $\vec{\xi} = \vec{t_c} - \vec{f_c}$.

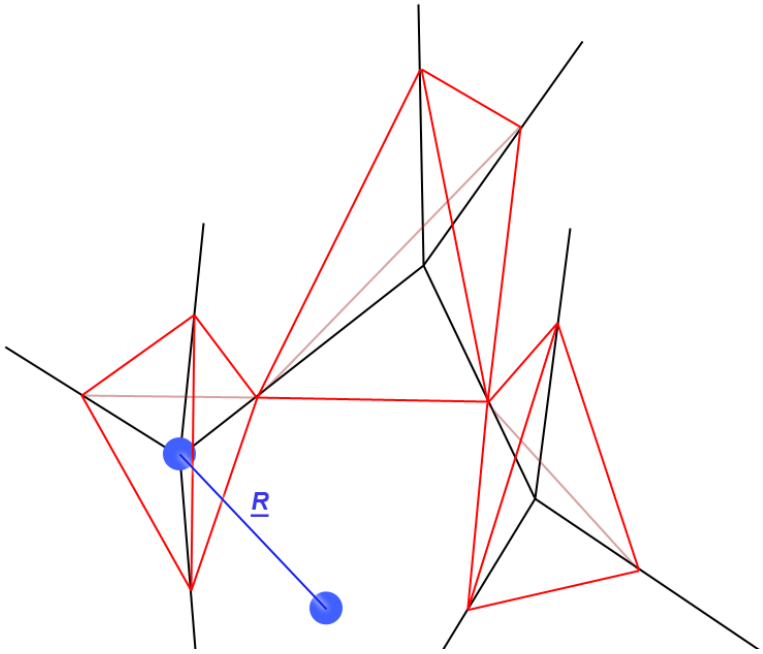
These new vectors are formed into an octohedral volume element as shown in Fig 4.2k, which is named the three dimensional quadron. Exactly as in two dimensions, all the quantities are defined as linear combinations of the structural degrees of freedom, the contact points.



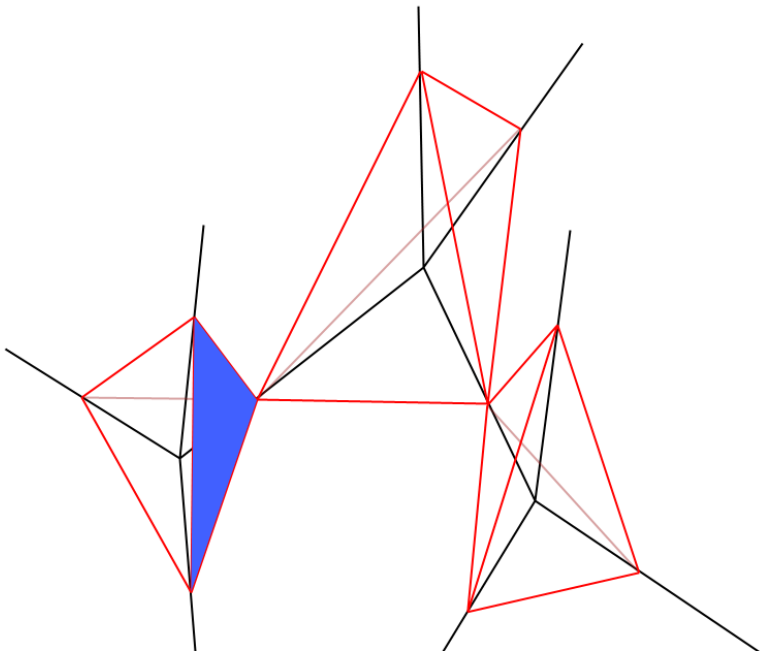
(a) The construction begins on a set of convex hulls, in this case three tetrahedral elements as might be found in a foam for simplicity of the drawing – the same concepts work equally well in grains with higher contact numbers.



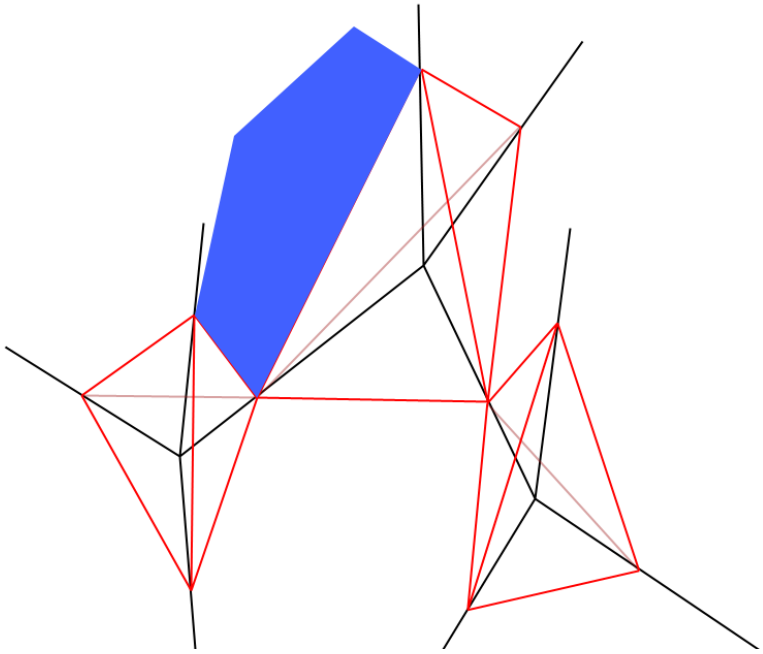
(b) Firstly, $\vec{r}$ is defined as in two dimensions as an edge running along the surface of a grain.



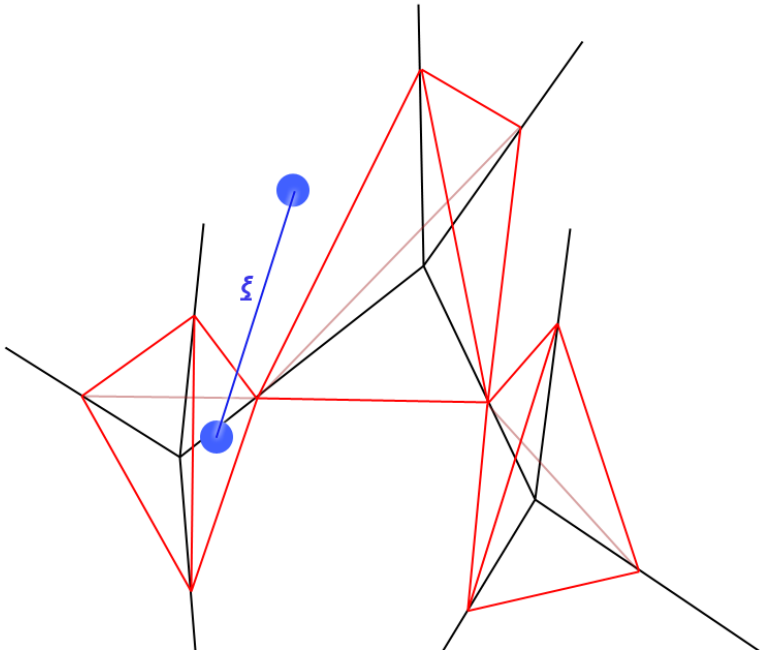
(c) Additionally, $\vec{R}$ is defined, as in two dimensions, as the difference between the grain and cell centroids.



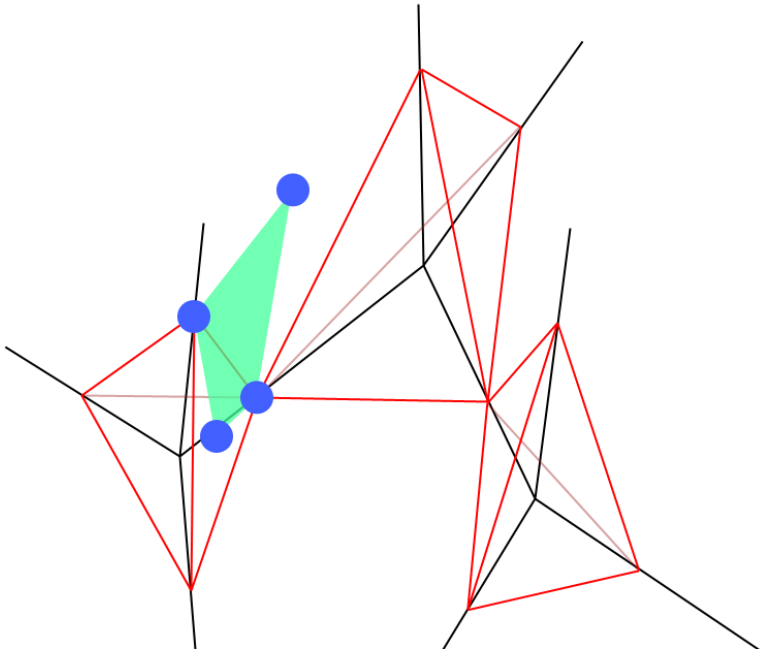
(d) To construct the final vector the facet adjoining $\vec{r}$ and pointing to this cell is highlighted...



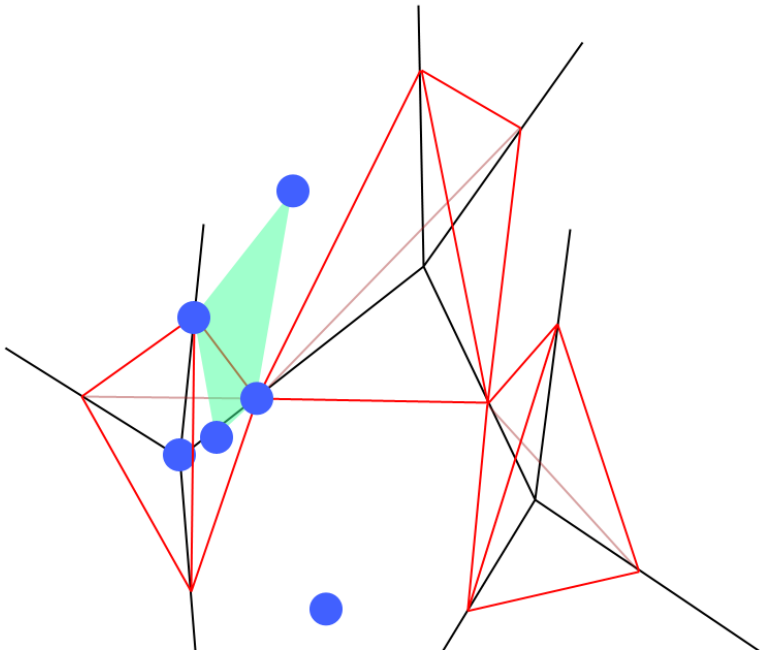
(e) ...and the throat which cuts the pore space along $\vec{r}$ is also highlighted.



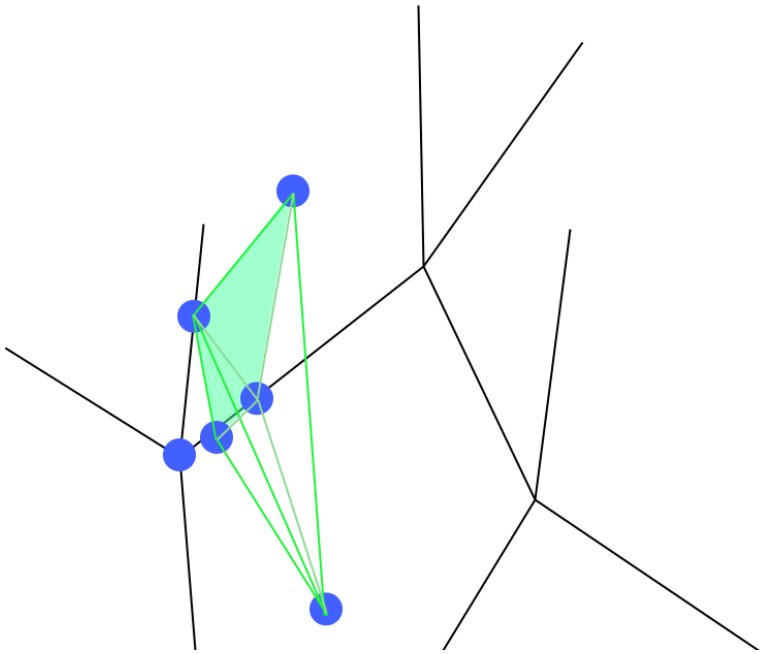
(f) Thus, $\vec{\xi}$ is defined as the difference between the centroids of these two structures.



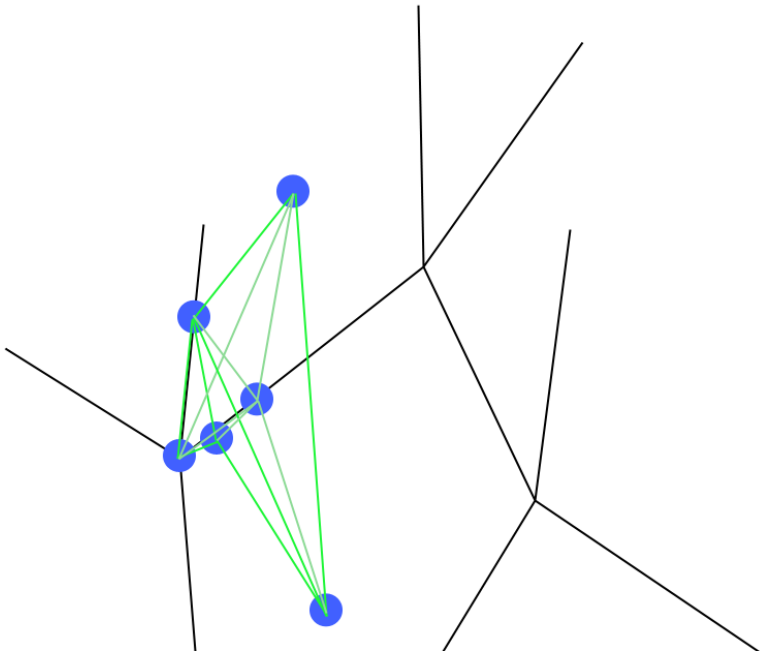
(g) Constructing the volume element itself – a non-planar quadrilateral is imagined where the vectors $\vec{r}$ and $\vec{\xi}$ cross over in the middle.



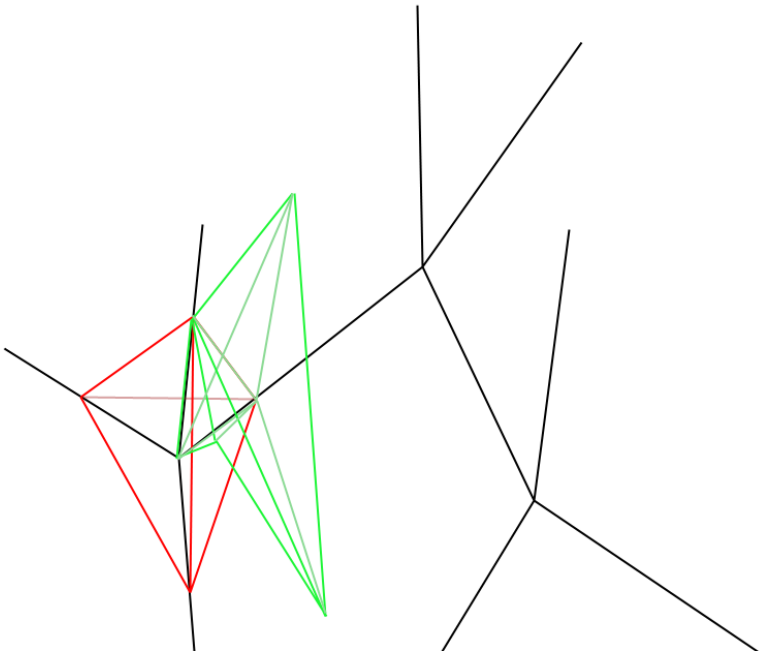
(h) Recalling the start and end of $\vec{R}$ as the centroid of the grain and cell...



(i) ...two non-flat based pyramids are constructed, the first extending from the non-flat quadrilateral out to the cell centroid...



(j) ...and the second extending back into the grain centroid.



(k) Thus an octahedral volume element is defined, where the set of such elements tessellates the structure.

Figure 4.2 – The quadron construction in three dimensions.

Analogous to two dimensions, the 3D structure tensor is defined as follows:

$$\hat{\mathbf{C}}^q = \left(\vec{r}^{\,q} \wedge \vec{\xi}^q\right) \otimes \vec{R}^q \tag{4.1}$$

where $\wedge$ denotes the cross product, and $\otimes$ the outer product. The volume of the quadron can be computed, analogous to two dimensions, as follows:

$$v^q = \frac{1}{2}\left(\vec{r}^{\,q} \wedge \vec{\xi}^q\right) \cdot \vec{R}^q = \frac{1}{2}\,\mathrm{Tr}\left(\hat{C}^q_{ij}\right) \tag{4.2}$$

Again, the chirality can be computed from the structure tensor. In three dimensions, the cross product becomes a pseudo-vector, rather than a z-component, as illustrated in equation 4.3.

$$\chi_i = -\left(\vec{r}^{\,q} \wedge \vec{\xi}^q\right) \wedge \vec{R}^q = \epsilon_{ijk}\hat{C}_{jk} \tag{4.3}$$

Unlike in two dimensions an "ideal" quadron (e.g. like that from a crystalline structure) does not have zero chirality, because $\vec{\xi}$ and $\vec{R}$ are only orthogonal in irregular cases — however the chiralities of the quadrons *on the same facet* in a crystalline structure do sum to zero. Since it is usual for the chirality to be summed per-grain, it measures a fluctuation away from an ideal close packing other than density.

There are n such quadrons around each facet on a grain, where usually $n = 3$ because the facets are triangular.

4.2.1.1 Quadron Position

In two dimensions the quadron position can easily be given either as the centroid of the four points that bound the quadrilateral or as the intersection point between

$\vec{r}$ and $\vec{R}$, as both of these points will (for any valid quadron) be inside the volume element.

In three dimensions, this problem requires more care, as the fourth tetrahedron (see section 4.2.2.6) means that many assumptions about which points are inside a quadron may not hold. Where such a position is necessary for the presentation of results – for example to correlate statistics spatially – it is defined as the midpoint of $\vec{\xi}$. This point, just as in two dimensions, is always inside the quadron as long as its geometry is valid according to the constraints listed in section 4.2.2.

4.2.2 Geometrical Constraints

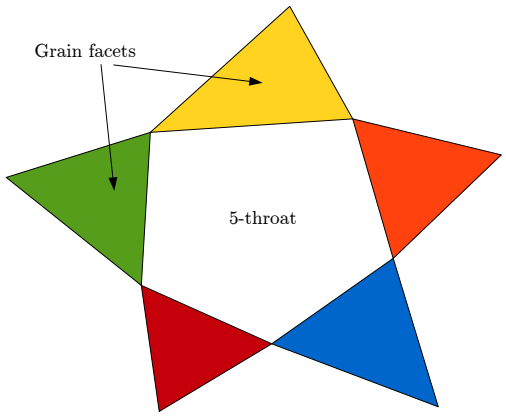
In order for quadrons to be tessellating elements, they must meet a few geometrical conditions. If they fail to meet these conditions, they become inside out or self-intersecting, which causes overlaps and makes the structure tensor function multi-valued at some points in $\mathbb{R}^3$. This requirement applies some specific geometrical constraints on quadrons, or rather on the structures that they are pinned to. Any procedure for computing three dimensional quadrons for granular materials must produce grain hulls and cell boundaries which meet these constraints.

4.2.2.1 Grain and Throat Facets

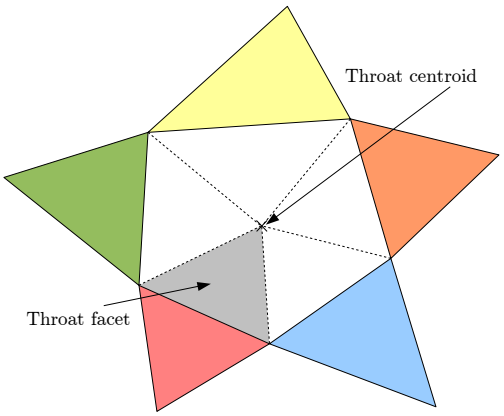
The convex hull facets on the grains are referred to as *grain facets*, and are also the faces of C. Each triangular facet is split about the facet centroid into three *grain sub-facets*.

A *throat*, t_{ij}, is a cut through the pore space such that all of the points on one side of the throat are contained in one cell, c_i, and all the points on the other side are contained in a different cell, c_j. It should be implicit in this definition that every cell must be a region bounded entirely by throat- and grain-facets. In

the quadron description, a throat must be an n-cycle in G, that is every throat must be made up of a ring of n grains, and must also be an n-cycle in C, that is every throat must run along convex hull edges through n contact points. This means that each throat is a non-planar n-polygon in the pore space, its boundary made up of contact points and convex hull edges on grain surfaces, as illustrated in Fig 4.3a.



(a) The definition of a throat, as a cycle of grains or contact points along the convex hull edges. Note that this diagram has been projected into the plane of the page.



(b) The definition of a throat facet, as a triangle made up of a throat centroid and a pair of contact points that share a convex hull edge.

Figure 4.3 – Definitions for *throats* and *throat-facets*.

Given these definitions, we define a throat centroid to be the average of its contact point positions, and a throat facet to be a triangle made up of the throat centroid and one of its convex hull edges (there will be n such facets for a given throat,) as is illustrated in Fig 4.3b.

4.2.2.2 Convexity

A polyhedron is convex if and only if all internal angles between faces are less than π. A convex hull is always convex because it is the solution to a Delaunay tessellation; any two concave faces in three dimensions would form a tetrahedron

which would be included in the tessellation, and thus they would no longer be external faces of the hull.

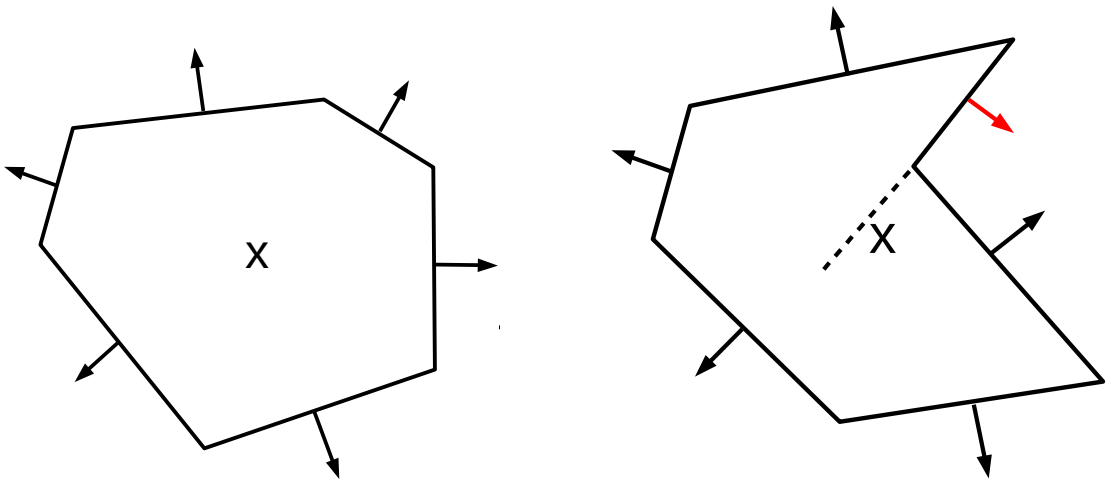
4.2.2.3 Centroidal-normality

A polyhedron p has a centroid $\vec{p}_c$, defined as the average position of its vertices. A polyhedron has facets f_i, which are triangular segments of planes which, near the surface of the triangle, divide the set of points inside and outside the polyhedron. The *facet normal*, $\hat{\vec{n}}_f$, is defined as the normal to the plane of the facet which points to the side of the plane which, in the locality of the triangle, is outside the polyhedron.

A polyhedron is *centro-normal* if and only if all its facet normals, when placed on the facet, point away from its centroid. That is if $(\vec{f}_i)_c$ is the facet centroid (a point on the facet) and if $\vec{p}_c$ is the polyhedron's centroid, the condition can be written as

$$\left(\vec{p}_c - (\vec{f}_i)_c\right) \cdot \hat{\vec{n}}_f < 0, \ \forall f_i \in p$$

There are several polyhedrons which are required to be centro-normal for the quadron tessellation to yield non-overlapping quadrons. A two-dimensional analogy for centroidal-normality is illustrated in Figure 4.4



(a) A centro-normal polygon, where each external normal faces away from the centroid (marked).

(b) A non-centro-normal polygon, where one of the normals to the external surface points *towards* the centroid.

Figure 4.4 – Two dimensional centroidal normality.

It is a condition on all two dimensional systems that grains and cells be centro-normal – a condition always met by grains, by definition of C in two dimensions, but not always by some esoteric cells which occur in frictional systems without body forces.

4.2.2.4 Grain Polyhedra

The condition for valid, non-overlapping quadrons is that grain polyhedra must be centro-normal, a condition which is always true for convex hulls.

4.2.2.5 Cell Polyhedra

We also define a cell polyhedron around each cell, where the union of all the cell and grain polyhedra is the total volume occupied by the system. Thus, one part of the boundary of cell polyhedra are the grain facets pointing to that particular cell, and the rest of the surface are the throat facets for each throat connecting this cell with another one.

The condition for valid, non-overlapping quadrons is that cell polyhedra, including grain polyhedron facets and throat facets, must be centro-normal.

4.2.2.6 The Fourth Tetrahedron

A three dimensional quadron is made up of four tetrahedra – two from the grain sub-facet out to the grain and cell centroids, and two similarly from the throat facet. The two tetrahedra towards the cell are already accounted for in our centro-normal criteria from above, since the cell polyhedron consisted of both the grain and throat facets, but the grain polyhedron only contains the grain facets, and not the throat ones. This is not an oversight – consider the throat facet, which is shared between two quadrons one for each of the cells the throat divides, and precisely which fourth point completes this fourth tetrahedron for the quadron with each cell. It is precisely the same fourth point – the grain centroid – which completes the tetrahedron in both cases, implying that both quadrons overlap on this same tetrahedron.

There is no overlap inherent in the definition of a quadron, the tetrahedron is assigned to one of the quadrons, and defined to be a negative space removed from the other (taking chunks out of both the throat-cell and grain-grain tetrahedra). This definition is made according to where the line segment between throat centroid $\vec{t_c}$ and grain centroid $\vec{g_c}$ intersects the grain polyhedron (this line segment is called $\vec{\alpha}$, for clarity). This line, $\vec{\alpha}$, must intersect one or other of the sub-facets pointing to the cells divided by t, see the projection in figure 4.5, and we take the negative space out of the side that it intersects. This fits the original "pyramid" definition because the tetrahedron emerges from the back of the throat facet in this case, meaning the non-flat 4-pyramid's edges up from the base are not all convex (octohedral quadrons were not convex before this problem arose.)

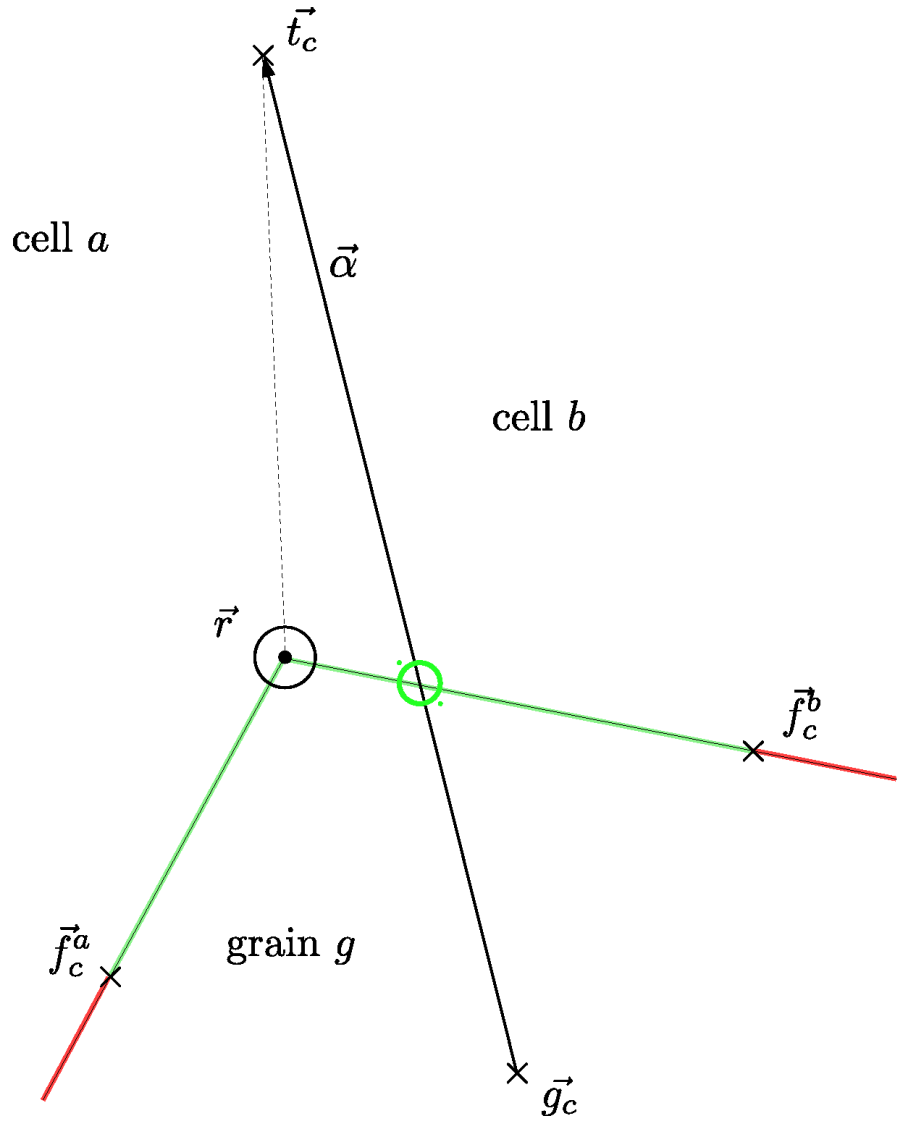


Figure 4.5 – Two facets, f^a and f^b, pointing to cells a and b respectively and sharing an edge $\vec{r}$, which is used by a throat t. Places where $\vec{\alpha}$ may intersect the facets relative to their centroids are labelled in green (and similarly prohibited line segments in red), in this projection with $\vec{r}$ perpendicular to the plane.

This condition can lead to an invalid state, another overlap, although not between the two quadrons which share that throat facet. If the line $\vec{\alpha}$ intersects the grain polyhedron at some place other than the sub-facet of either of the two quadrons unique to this grain and throat, then both quadrons would overlap with the other quadron whose sub-facet was intersected, corresponding to the red line segments illustrated in figure 4.5

Thus, the validity condition on this fourth tetrahedron is that the intersection point between the line segment between $\vec{t_c}$ and $\vec{g_c}$ and the grain polyhedron must lie on one of the sub-facets on either side of the throat edge.

4.3 Searching for Cells

4.3.1 Existing Techniques

The segmentation algorithms referred to in the Literature Review in section 2.2.7, were both tested for solving the problem of determining the cells in three dimensions. Each has a drawback which leads to an insufficient solution to this problem if the cells are to be compatible with the geometrical constraints in section **??**.

4.3.1.1 Watershed

Beucher and Lantuéjoul's watershed algorithm[16] as defined does not remember details about which voxel belongs to which grain, or indeed facet, in the phase not being considered. It can be (and was) re-implemented in order to capture this extra information, but the results were extremely disappointing. Each facet should ideally be adjacent to pore phase voxels belonging to one object or pore detected by watershed, or at least the overwhelming majority of the adjacent

voxels should belong to one such object. Instead it was observed that the facet voxels were adjacent to many such objects, often with as many as five different cells having the same order of magnitude of claim on the facet. It was concluded that such a technique required significant redesign to yield a clear answer as to the identities of the cells – before any of the geometrical constraints listed above could even be tested – and as such the watershed approach was retired in this problem.

4.3.1.2 Tight Dual Model

Glantz and Hilpert's tight dual model[17, 18] (TDM), working conceptually with the mesh rather than some other representation, had no problems assigning cells to facets – once the algorithm had been implemented it was easy to render the choices it had made (using the custom three dimensional visualiser mentioned in section 4.2.0.1) and perform a preliminary examination of the solution with reference to the requirements of the quadron description.

TDM operates entirely in the pore-space topology, and while its output does uniquely map each facet to a cell, it doesn't necessarily ensure that each cell is bounded by a surface of facets and valid throats. This is because it divides up the pore space by slicing the Voronoï complex in arbitrary ways – for example by having three cuts to the complex meet in the middle of a Voronoï face (see figure 2.14), causing three throats to adjoin one another in the middle of the pore space, which is disallowed in the quadron description (they must only run along convex hull surfaces.)

If a slice through the Voronoï complex splits up or stops, the dual structure (the throat) must jump across the pore space to reach the next contact point. This makes it an invalid throat for the purpose of the quadron description – although

it can still be used in equivalent networks for solving flow behaviour, which is why it is useful in its original context.

4.3.2 Development of the Cycle-Centric Algorithm

The following cycle-centric algorithm was implemented in Python 2.7[68], using several existing open source libraries, namely NetworkX[70] for graph representations, and Voro++[71] and CGAL[72] for geometrical tessellations. The Cython[73] language was of great help interfacing with C, C++ and FORTRAN routines. Granular systems from simulations and experiments were used during development of the code, as well as a foam system for a sanity check, and the particular systems investigated are introduced formally in section 5.1.

4.3.2.1 Tetrahedral Complex Structures

Graphs which admit "sane" three dimensional embeddings (those with edges to nearby elements only) are sparse and contain many small (spatially local) cycles. They retain these properties independent of their original physical embedding (if any), and can be analysed as pure graphs.

One method for computing a set of throats might be to search for cycles in the embedding-independent representation of the granular graph G. Enumerating all cycles is an $\mathcal{NP}$-hard problem, as discussed in section 2.2.2, but the sparsity and correlations of the vertices in graphs from physical systems can be used to narrow the search.

Searching for short, chordless cycles (up to some length threshold, l) yields a set of cycles $\{a_i\}$. An algorithm which executes this task performantly is, for example, the one presented by Dias et al[74], which begins the $\mathcal{NP}$-hard search at each

unique triadof adjacent vertices, and limits the search to the subgraph within l edges of the middle vertex.

Initially the chordless cycle search in [74] was implemented in python from the pseudo-code, and the remainder of the necessary logic – converting cycles in G to cycles in C, and deducing distinct cells $\{c_i\}$ and their facet sets from C from the set of cycles – was added by looking at the geometry of the grain facets relative to the throat facets (see figure 4.3b), "colouring" facets on each side of a throat, and then collecting facets together in sets based on common colours from different throats. Part of this process involves verifying that the detected cycles in G do also run simply along convex hull edges in C – this can filter the original set of cycles $\{a_i\}$ down to a subset of *valid* throats $\{v_i\}$.

In a foam, like one generated using Surface Evolver[75, 47], each vertex is of degree 4, and all short chordless cycles are bubble walls (analogous to throats in the granular graph). One such foam was loaded into the first draft of the novel software, having first been stripped of all information about bubbles and walls so that it resembled a tetrahedral granular system, like the one shown in the original quadron description definition in section 4.2.1.

The bubbles for this foam were immediately recoverable simply by choosing an adequate value for l – a value which was easy to determine following the work of Matske who showed that the largest order bubble walls observed contained 9 vertices – which was $l = 10$. For foams, $\{a_i\} \equiv \{v_i\}$ because of the simple convex hulls.

This quick, positive result showed that if anything could determine a valid set of quadrons for a three dimensional granular system, it would be a cycle-searching algorithm.

4.3.2.2 High-Degree Vertices

In foam systems, all vertices are degree $|g_i| = 4$. This means the convex hulls are simple tetrahedra, and every facet on the same grain (foam vertex) points to a unique cell (bubble). Another way of stating this condition is that every contact point has three convex hull edges on each grain, and that *all* such edges correspond exactly to a single short cycle each. It is not necessary to check for centroidal-normality of the cell polyhedra, since this must be satisfied according to the geometrical construction of the tetrahedral facets in foams. In granular systems vertex degree is much higher, and this has three consequences immediately for the efficacy of the algorithm.

Firstly, it is no longer obvious that the geometrical constraints on cell polyhedra would be met easily. As such, a stage of geometrical validation was added to the computed cells at this stage, which verifies the constraints outlined in section 4.2.2 are satisfied for each cell in turn. This validation involves many floating point manipulations, and is $O(|G|)$ with a large coefficient (and therefore relatively slow.)

Secondly, many detected cycles may overlap with one another on convex hull edges $e_i \in E(C)$, which means they cannot all be applied simultaneously. Thus, the procedure was modified to choose a non-clashing subset of $\{v_i\}$, $\{d_i\}$. Finding an optimal subset (one which satisfies the geometrical constraints) is not practical since the validation step (including both determining which facets belong to which cells, and the geometrical constraint checking) must be computed for every non-clashing subset, a problem analogous to the $\mathcal{NP}$-hard Maximal Independent Set introduced in section 2.2.3. This problem was overcome by using a greedy algorithm as proposed by Edwards[22] (ranking throats according to their suitability, choosing the best throat, and eliminating all clashing throats) – a choice

which is robust to increases in l if the throat length is used as the primary ranking criterion. Additional criteria added to distinguish between cycles of the same order were the average convexity of the hulls (the larger the external angle the better), and the deviation from a co-planar set of points (the flatter the better.)

Thirdly, the degree of each contact in C may not be equal for both of its grains, and so it is no longer obvious that every convex hull edge is used in an applied throat. This implies that two adjacent facets may point to the same cell, an edge-case not accounted for in the original quadron description. When two adjacent facets on a grain both point to the same cell, they are merged into one *multi-facet*, which has a single centroid but more than three contact points which are not all in the same plane. In this case, each quadron uses two adjacent contact points in the cycle and the facet centroid as before, but the quadrons from the multi-facet are not all in the same facet plane, since they are divided into *sub-facets* each with its own normal direction. As long as no contact point is completely surrounded by facets pointing to the same cell, it is possible to allow multi-facets with more than 2 of C's faces, provided the contact points around the circumference of the union of the hull facets form a cycle. This modifies the geometrical constraints slightly, since now the grain surface may longer be convex – the new constraint is, just as with the cell polyhedron, the grain polyhedron after substituting in the (non-planar) sub-facets of a multi-facet must remain centro-normal (see section 4.2.2.3). This must also be true from the point of view of the cell polyhedron, which also is re-shaped in the same way.

After these modifications were made to the software, it was capable of analysing dense, spherical granular systems including simple DEM simulation data and the most dense colloidal suspension data-set from Brujić[3], t0. In dense systems like these, throats above order-6 were very rare, and so it was computationally efficient to reduce l to 6.

4.3.2.3 Less Dense Systems

All particles in a granular system to be analysed in the quadron description must be at rest in mechanical equilibrium. Normally, this requires grains in three dimensions to have at least four contact points, which gives a tetrahedral (non-flat) grain polyhedron. However a grain under body forces, like gravity, may have only three contact points and still be "held down" in equilibrium by the body force. In this case, the grain polyhedron degenerates to a triangle, and is flat. This was observed in less-dense systems.

This means that some of the vertices in the quadrons' octahedrons are coincident, namely both facet centroids and the grain centroid, and so the octahedron degenerates to a hexahedron (the apex of the grain's non-flat square-based pyramid has merged into one of its base's points, leaving behind two faces instead of four.)

Another problem, which arises in dense but especially frequently in less dense granular packs, is what should occur if a cell, after the greedy algorithm has been applied, fails the geometrical validation. The three dimensional viewer was especially important in eliminating several bugs in the program at the stage where this problem was being considered, but even after they were all fixed some edge-cases would arise to which the author's eye could see no obvious solution – the set of cells $\{c_i\}$ implied by the throats $\{d_i\}$ contains in general some subset, $\{b_i\}$, which do not satisfy the geometrical constraints and are termed *invalid*.

Reluctant to revisit the greedy algorithm, the compromise was to admit a small number of clashing throats from $\{s_i\} = \{v_i\} \setminus \{d_i\}$[2], where they rendered cells geometrically valid. Of course, if two throats run along the same convex hull edge, the definition of quadrons must be revisited (and the geometrical validation updated once more.) In this case, we define a *phantom-facet* to be a one dimen-

[2]The set of throats $\{s_i\}$ which remain after the greedy algorithm chooses $\{d_i\}$ throats from $\{v_i\}$, the set of valid ones.

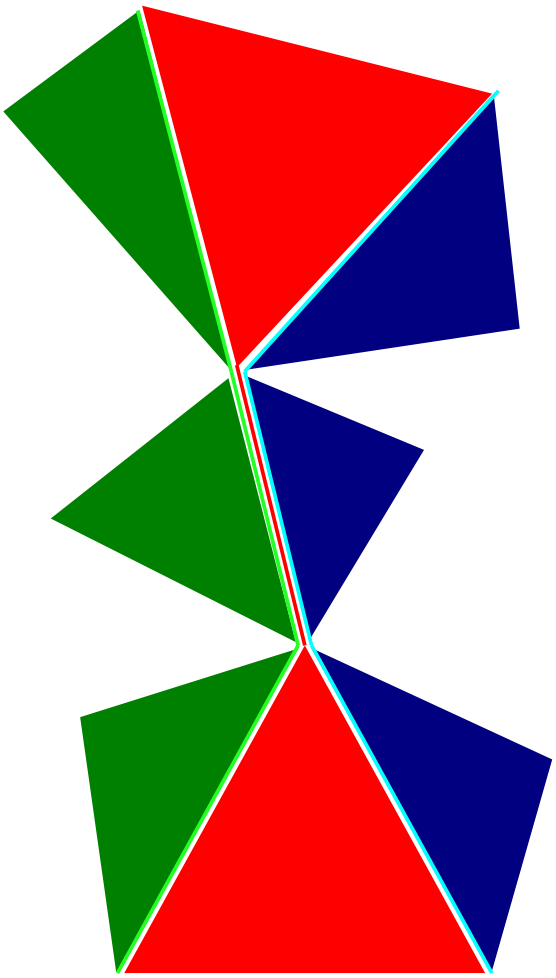


Figure 4.6 – A phantom facet (red) in between a blue and a green cell. The two throats' paths have been rendered indicating how they both flow along the same convex hull edge, which is the one which is a phantom facet.

sional facet running along the conflicted convex hull edge – conceptually it is "in between" the two triangular facets either side, see figure 4.6.

Quadrons built on a phantom facet are degenerate in that two tetrahedra of the four, the two which reference the facet centroid (one towards the grain and the other towards the cell centroid), collapse into a triangle because the facet centroid is collinear with the two contact points (since it is straightforwardly their midpoint.) This has no consequence on any other part of the quadron, except to constrain the quadrons on either side on the same grain a little more than before. The quadron's volume is accounted for in a single tetrahedron, with

a triangle sticking out of one of the edges. As with planar grains, the structure tensor definition remains unchanged.

4.3.2.3.1 Fourth Tetrahedra in Phantom Facets

For a phantom facet, there are now two vectors running from $\vec{g}_c$ to different throat centroids, $\vec{t_{1_c}}$ and $\vec{t_{2_c}}$, yielding line segments $\vec{\alpha}_1$ and $\vec{\alpha}_2$. In order for neither phantom facet quadron to overlap the regular quadron from the opposite side, the condition above must be true for each throat, and furthermore each line segment $\vec{\alpha}_i$ must intersect a different sub-facet. This condition is illustrated in figure 4.7.

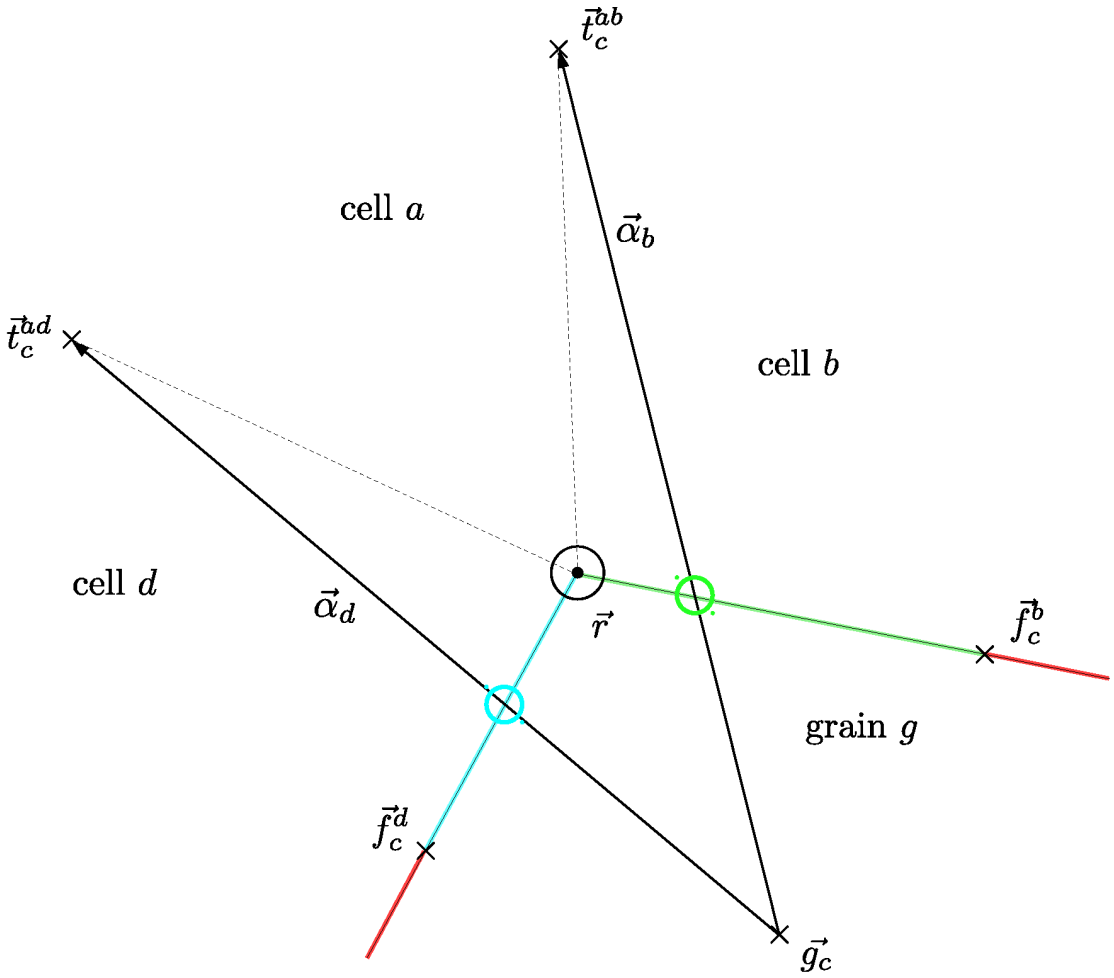


Figure 4.7 – Two facets, f^a and f^d, pointing to cells a and d respectively, projected into the same plane as figure 4.5. When a phantom facet causes two throats to share one $\vec{r}$, each corresponding $\vec{\alpha}^i$ is restricted further to its own side (indicated in lime green and cyan in this plot.) This ensures that the phantom facet's quadrons will not overlap with either.

Specifically, the method for introducing the minimum number of phantom facets was to add a routine which operated recursively on invalid cells, splitting them using the throats $\{s_i\}$ rejected by the greedy algorithm until they became valid, or the routine ran out of throats to test. Sometimes, in this worst case scenario, the software must simply give in and leave invalid cells present. The final sets of valid throats and cells are then named $\{t_i\}$ and $\{c_i\}$ respectively.

The pathological case of this invalid cell problem is where a large throat near the edge of the system allows the boundary layer to "leak" into the rest of the system. The software was always careful to handle the boundary layer separately from the rest of the structure (initially is was simply discarded) – once this pathological case was observed in less-dense systems a special non-recursive version of the cell-splitting routine was developed to avoid excessive duplication.

Systems that required this extra analysis, including some highly frictional DEM simulation data and the less dense colloidal systems from Brujić[3], t5 and t10, not only required the higher $l = 10$, but also spent much more time in the cell splitting routine than any other part of the algorithm. It was clear that for less dense systems this would be the dominating computational limit.

4.3.2.4 Non-Spherical Particles

The next granular system to be tested was the glass bead pack from 3SR, ABEA01-01-, which consists of notionally spherical particles, but contained some anomalous data – mostly where the spheres had broken up. Workarounds were computed for such particles, so that their presence did not cause errors to propagate around the system and allowed the program to continue, while ensuring they were marked as excluded for the purposes of statistics. Additionally the Caicos ooids[2] were also used, and further tested the assumptions that had been made about particles and

their convex hulls.

Several changes were made to the program in relation to these problems, in particular the temporary re-computation of contact points entirely contained within a convex hull, while marking both grains involved as unsuitable for statistics. Also, the method used to compute the convex hull, which was initially a routine which computed the Delaunay tessellation on the surface of a sphere, was replaced. First, with a routine which computed the three dimensional Delaunay tessellation for the entire system (the same as was required for the purposes of the Glantz-Hilpert Voronoï Complex described in section 2.2.7.2) which proved impractical as a grain's hull could be influenced by a contact between two other grains that happened to be near its surface. Finally, a routine which computed the convex hull as the union of the Delaunay tetrahedra of each grain independently was used.

4.3.2.5 Large Systems

The program was impractically inefficient in RAM requirements when it came to very large systems, like those of the 3SR experiments with many tens of thousands of grains. The main culprit for this was the NetworkX[70] library, which uses native python types and lots of redundancy to represent graph objects. Large systems like the Caicos data-sets could occupy several gigabytes of RAM before any analysis had begun, and since the main development machine was furnished with only 8GB of RAM (which should be plenty for any such analysis with an optimised program) this soon became a problem as run-time died as the system began to page/swap (use the hard disk as memory.)

One solution would have been to re-implement a NetworkX compatible interface to use instead that was more efficient, and with less time pressure this would be

a very fruitful enhancement to the project, but a simpler solution presented itself which also allowed parallel processing – running the program on a small subset of a large system.

This compromise meant that the performance problems went away, but the results were inconsistent – a cell in one subsystem might overlap with a cell in an adjacent subsystem. In order to obtain valid statistics from such subsystems, they were determined as follows.

A *slice* was defined to be a cubic region inside a system, where any grains inside that region were copied into the slice subsystem. Slices were chosen to overlap, so that any individual grain could be in up to 8 slices at once (2x2x2 cubes each offset by half their side), and so that the majority of grains in the bulk of a large system would be in exactly 8 slices. When reconstructing the statistics after the analysis, the most successful[3] slice for each grain was identified, and then all of the quadrons for that grain were imported from that slice. This meant that quadrons in the final statistics would often overlap where slices met, but that a good sample of quadrons from diverse cell types all over a large pack were achievable without a complete re-implementation.

By trial and error, a slice size of a cubic volume which could be expected to contain 700 grains was found to give the best balance of performance and output statistics.

4.3.3 Overview of the Algorithm as a Whole

The detailed algorithm for choosing a valid set of cells (with sane constraints on RAM usage and CPU time) that has been developed is summarised in figure 4.8.

[3]where *successful* means here the slice where the grain had the most quadrons from valid cells.

Mathematical terms from the text have been used to make the figure as concise as possible.

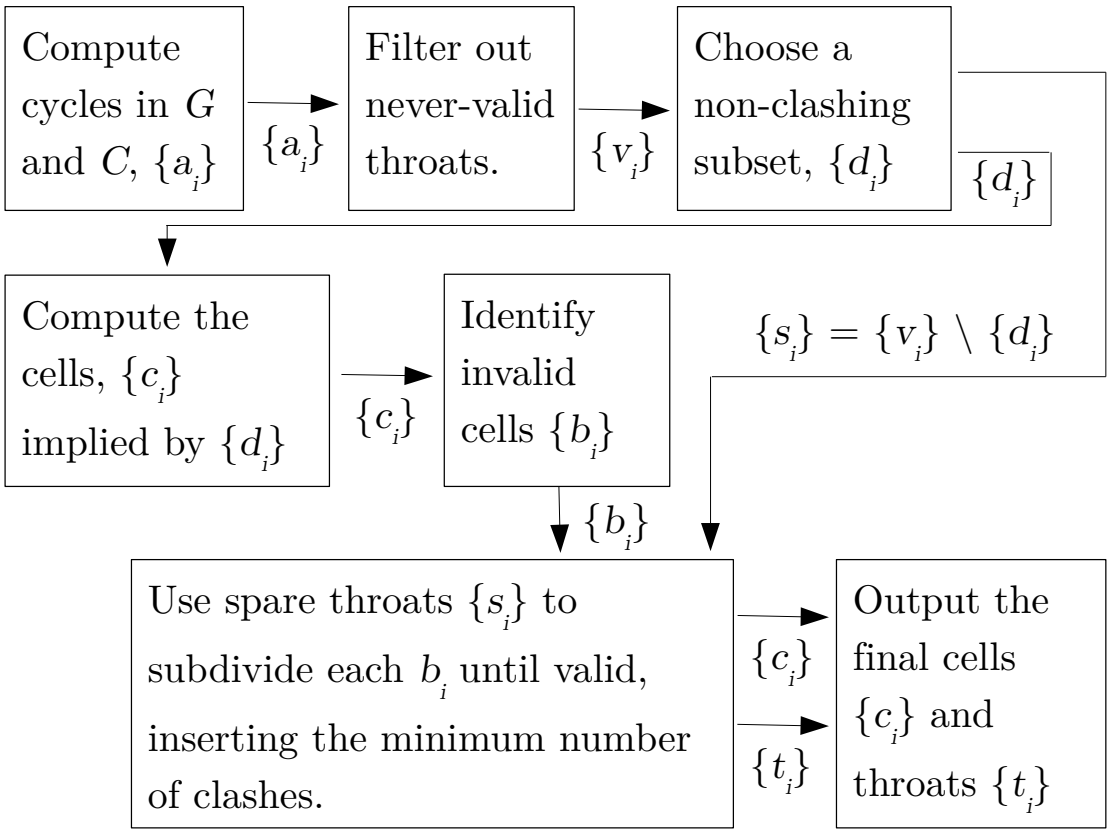


Figure 4.8 – Data flow within the novel cycle-centric algorithm.

4.3.3.1 Limitations of the final algorithm

The complexity class of the program is extremely difficult to determine, because it depends more strongly on the density than the overall size of the system. A plot of the number of quadrons produced (which roughly equates to system size) versus the time taken is provided in figure 4.9 to illustrate the problem. The systems in the plot are some slices from the Caicos initial and sheared systems[2], and the three oil droplet densities[3].

Note that there is a linear trend along the bottom of the plot, which corresponds to simple dense packs, and then there is a very wide range of run-times for both

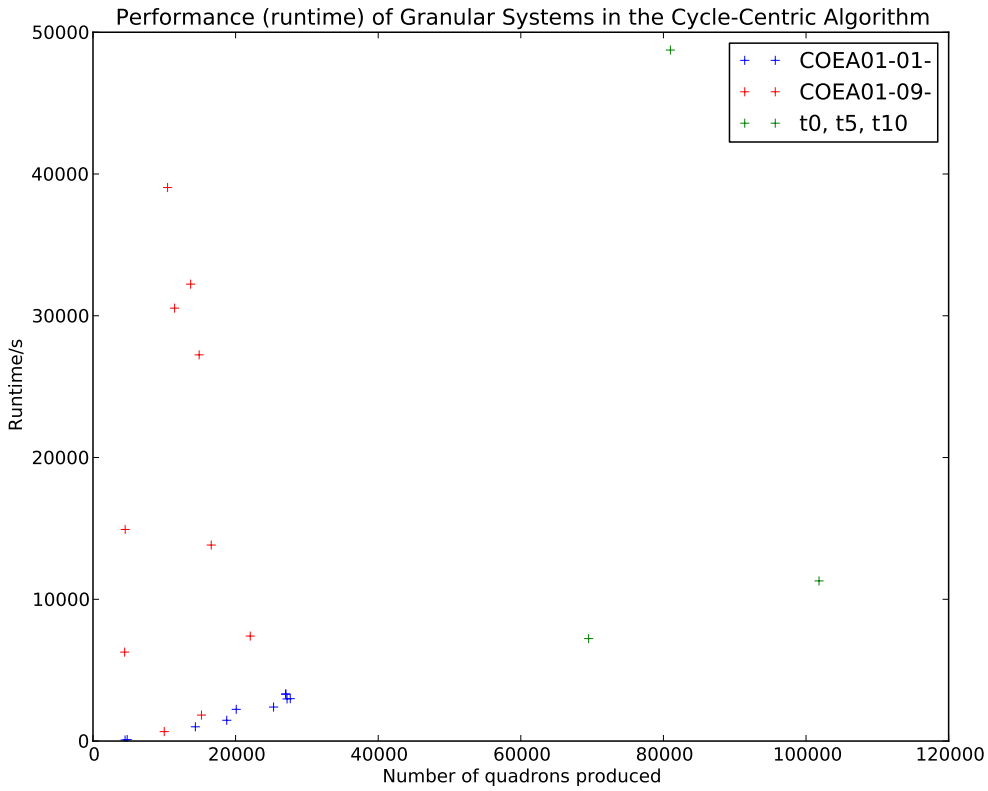


Figure 4.9 – Run-time performance of the program on different sized systems. The Caicos Ooid[2] data are for individual slices, not the whole systems.

small and large systems with lower densities – in particular note that the sheared Caicos system's slices in red are at best following the linear trend of the initial condition, but at worst are orders of magnitude slower. This may be a signal that this is in general an $\mathcal{NP}$-hard problem, where the dense pack is a special case which is easier to solve via a greedy algorithm. It is certainly true that the certificate to validate a choice of throats is polynomial, if a high exponent and with a large constant.

There are two reasons for the poor performance of the cell splitting procedure. Firstly the logic itself, which performs an exhaustive search of the spare throats $\{s_i\}$, filtered to work on the subset contained within the cell to be split. This is unavoidable even when on a second recursion the same throat is considered, since

the cell centroid changes at each stack level mean *all* the geometrical constraints must be re-evaluated. Secondly in order to safely verify the geometrical constraints the (often large) data structures are copied in RAM, which in a RAM-heavy program can be a large cause of inefficiency. The memory problem is more general, in fact, in that the program uses many GB of RAM to hold the data structures necessary for large systems (this was why the slicing technique was originally devised.)

The geometrical constraint verification is currently implemented as a set of Python functions acting on a pure-python vector class, which is doing much unnecessary type-checking and bounds-checking at run-time.

The quadron description is slightly more yielding to non-convex grains than the implemented software – the anomalous glass chips in the glass bead pack from 3SR (data-set AOEA02-01-) were excluded from the statistics as a workaround, even while they were influencing the behaviour of the experiment and would almost certainly have valid quadron structure.

Additionally, some parts of the algorithm (notably the contact detection routines) are not robust at all to wide polydispersity – at present they will run very slowly rather than using large amounts of RAM. This could be optimised by providing a specialised case of the C routines for widely polydisperse samples.

Chapter 5

Results and Discussion

Summary results are presented in this chapter with commentary – the exhaustive plots for every system are present in appendix C, along with a zoo of cell types.

5.1 Physics Results from using the Cycle-Centric Algorithm

Once the implementation of the new algorithm had been finalised, the definitive version was re-run on all available granular data-sets (details and labels for which are given in table 5.1). The most notable results, and the new descriptions required to classify them, are explained in this section.

Label	Particles	Quadrons	Slices	Description
COEA01-01-	56,265	1,641,400	751	Caicos Ooid experiment[2] initial condition (1/17).
COEA01-09-	56,395	861,862	665	Caicos Ooid experiment[2] sheared system (9/17).
ABEA01-01-	65,911	1,765,730	859	Glass bead pack from 3SR[76], initial condition.
large0	9,383	240,963	124	Sedimentation simulation[77].
t10	2,880	69,508	1	Oil droplet suspension experiment[3], lowest centrifuge speed.
t5	3,044	80,978	1	Oil droplet suspension experiment[3], middle centrifuge speed.
t0	3,327	101,824	1	Oil droplet suspension experiment[3], highest centrifuge speed.

Table 5.1 – Data Sources for the statistical results presented.

5.1.1 Distribution of Quadron Volumes

In micro-studies of granular media where tessellating volume elements are computed, volume distributions of the elements, PDF (v^q) are often presented[57, 58, 59, 62, 60, 64, 47, 65, 66]. In order to make sense in a global context, tessellating element volumes should be normalised so that systems at different scales can be compared like-for-like. The normalisation used in this work, except where overridden explicitly, is to define the average volume of a granular particle to be 1.

Volume probability distributions are presented here unmodified in figure 5.1, and exhaustively for all available systems in appendix section C.1. The distribution depicted is that for the initial condition of the Caicos Ooid triaxial compression

test from 3SR Grenoble, by Vianccino and Ando et al [2].

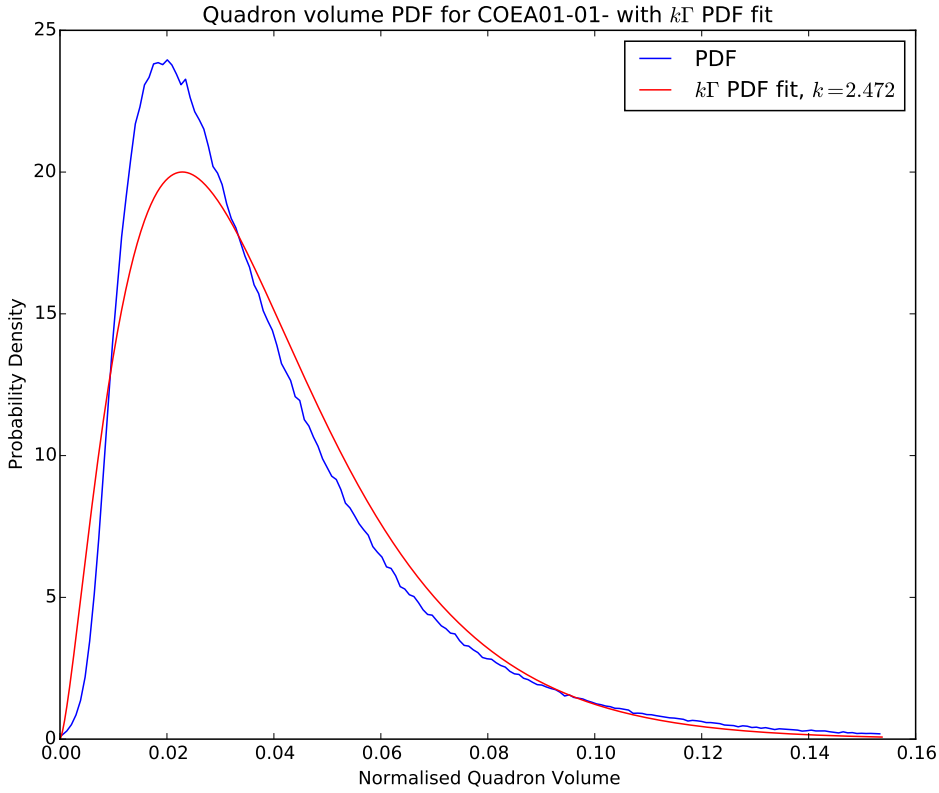


Figure 5.1 – The quadron volume probability distribution for the Caicos[2] initial condition with a $k - \Gamma$ PDF fit.

It should be noted that this distribution, and the analogous distributions in the appendix, are emphatically not $k - \Gamma$ functions, in contrast to the prediction in [57], although they do have a single maximum with no shoulders unlike the corresponding quadron distributions in two dimensions, examples of which can be found in the literature review in section 2.4.5.4. In such two dimensional systems the conditional distributions, PDF $(v^q|e)$, suggest the origins of the multi-modal curves. To conduct an analogous analysis in three dimensions, cell types must be classified.

5.1.2 Classification of Cell Type

In two dimensions, cells can be classified by a single parameter since they constitute a cycle of grains – the number of grains in the cycle is adequate. In three dimensions cells are planar graphs and can be imagined embedded on the surface of a sphere, the question of how to classify the different types is less obvious.

Interestingly, in three dimensional complex structures there exists a type of system, already extensively studied, where the cells are trivial to determine, indeed where they define the physics of the system's evolution – in foams. Kraynik[75] has used a software tool called Surface Evolver to reproduce and further investigate the pioneering work of Matzke[38], where the foam "cells", more simply called bubbles, are classified as follows.

A foam bubble is made up of walls where two bubbles meet and are separated by a two dimensional face, edges where three bubbles meet and the faces between each pair intersect, and vertices where four bubbles meet and four edges, one from each set of three that can be taken from the four, intersect. The quadron description is also valid for foams, where the vertices are analogous to grains, edges between them to contact points (the point is taken as the midpoint of the two vertices,) the faces to throats, and the bubbles to cells. The foam bubble resembles a subgraph of G in a granular system induced according to the grains in a single cell, were it to be embedded with grains at some position (their centre of mass, or their centroid.)

The foam bubble has been classified[75] as a frequency list of the number of vertices in each face. For example: a bubble with one quadrilateral face, ten pentagonal faces and two hexagonal faces would be labelled $(1, 10, 2)$[1]. Kraynik proposed

[1]Kraynik[75] uses notation where this cell would be written 1-10-2, but this notation in parentheses is used to mimic the Python "tuple" data structure.

calling this cell the Matzke cell, since it is the most common type in the soap film foams he originally studied[38]. The frequency list in foams always begins with the number of quadrilateral faces, since triangles have not been observed in foams, and nearly always has only three indices, because heptagonal faces are extremely rare.

In granular systems it proposed here to use the same numbering with two modifications necessitated by the types to be described. Firstly, the smallest throat size is three grains, not four, so the indices start with the frequency of three-throats (even in cells with no 3-throats, this index will be included with value 0). Secondly, granular cells vary widely in the upper limit of the throat orders – some cells only have 3- and 4-throats, others may contain 10-throats – so the number of indices in any one type will vary according to the type itself. The last index included will always be the last non-zero one, so for the unique case of the tetrahedral K_4 cell with four 3-throats the type will be written as $t = (4,)$.

The exhaustive list of all cell types discovered in the Caicos initial condition are presented in a table, in appendix section C.3.

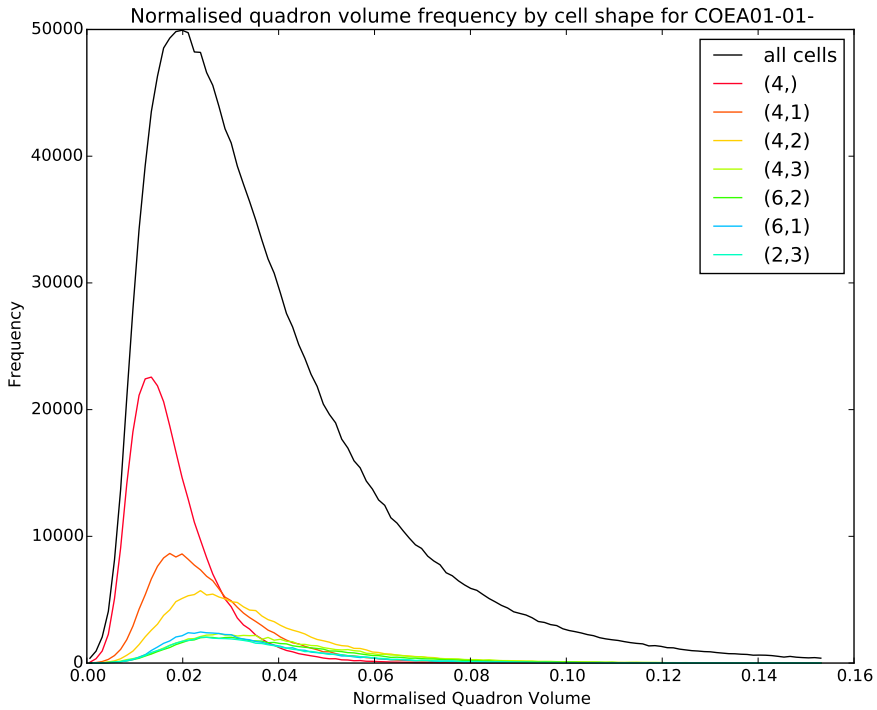
5.1.3 Conditional Quadron Volumes by Cell Type

Having defined a means of classifying three dimensional cells, the quadron volume histogram can be separated out into distributions of quadrons from the same cell type, PDF $(v^q|t)$. Initially, this is presented in a straight frequency plot, not normalised into a probability distribution, with the total frequency included for reference. Only the 7 most frequent cell types are included for each system so that the long tail of infrequent cells doesn't make the plot too crowded. This plot is included as figure 5.2, for the oil droplet system from Brujić[3], and the Caicos Ooid initial condition[2].

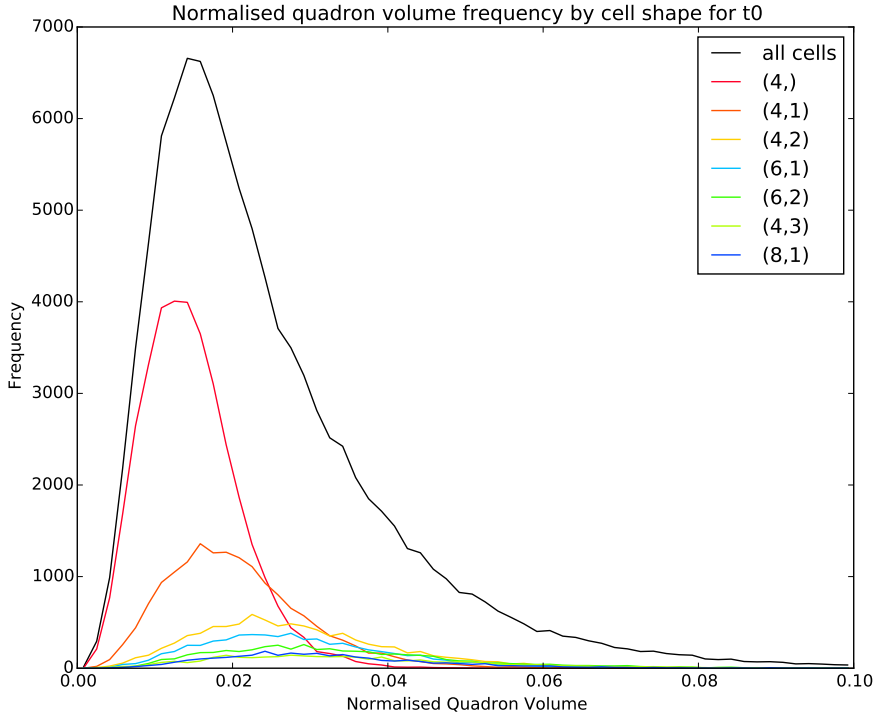
The two distributions in figure 5.2 appear similar, but it is important to notice that the cell types (the seven most frequent from each experiment) are different. The frequency of different cell types, as in two dimensions, is likely a function of different packing protocols, and in this case the differing interaction physics of non-deformable sand grains in figure 5.2a and deformable oil droplets in figure 5.2b.

The analogous result in two dimensions, in figure 2.19b looks similar in the sense of an envelope function composing a sum of other distributions, but the final form is radically different because a two dimensional system has so many fewer cell types, which make up correspondingly larger proportions of the total volume, and thus their individual peaks are distinguishable in the total PDF (v^q).

The diversification of cell types in three dimensions is interesting, and difficult to plot in the same way as in two dimensions. The long tail of cell types is presented in tabular form where it is more easily understood – the cell types in table C.1 in the appendix are taken from the Caicos initial condition and ordered by the number of quadrons from cells of that type, which is also presented. The final form of the quadron volume probability distributions in three dimensions are single-peaked and shoulder-less because of this long tail.



(a) Quadron volume frequency for the 7 most frequent cells in the Caicos[2] initial condition, plus the total frequency for reference.



(b) Quadron volume frequency for the 7 most frequent cells in the most dense oil droplet[3] sample, plus the total frequency for reference.

Figure 5.2 – Cell shapes make up much smaller proportions of the total volume distribution in three dimensions, resulting in a PDF (v^q) with no shoulders.

5.1.4 The Collapse of the Conditional Volume Distributions

A further normalisation is now introduced, to compare the probability density functions of quadron volumes between cells. The conditional quadron volume distribution is plotted for each cell type, after normalisation by its mean quadron volume. This normalisation is illustrated in figure 5.3 for the Caicos experiment's initial condition.

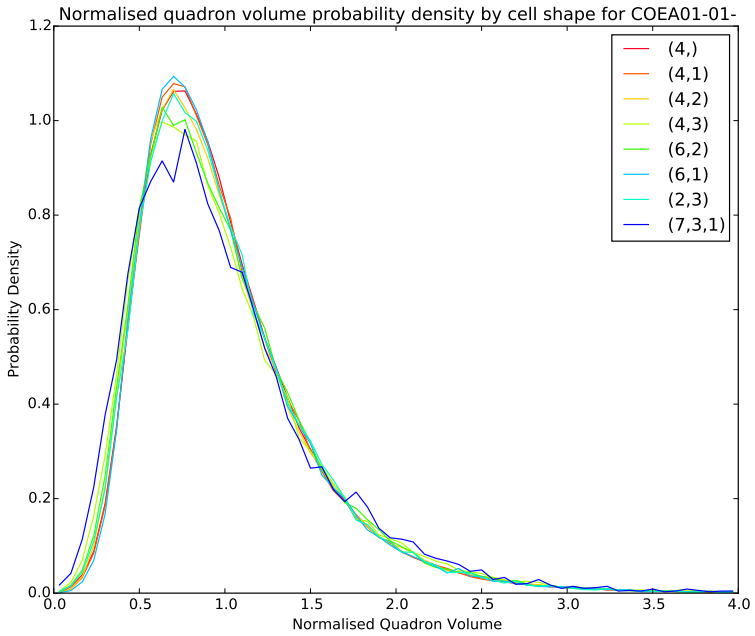


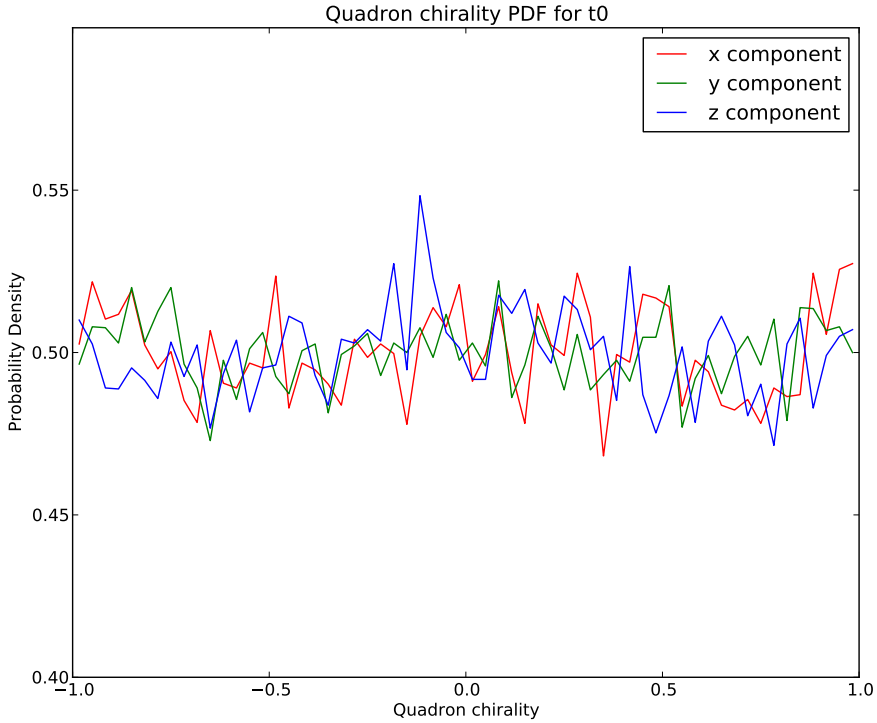
Figure 5.3 – Normalised per-cell volume probability densities for the 7 most frequent cell shapes in the Caicos initial condition, and the 26^{th} most frequent.

Note that included are both the 7 most frequent cell types, which have the best statistics, and one more example, $(7, 3, 1)$. This larger cell type suggests the subtle trend that as the cell size increases, so the normalised distribution shifts, becoming wider. This is quite difficult to see in these systems, because as the cells get larger they become considerably less frequent, and thus have worse statistics – the conclusion is at best tentative. In two dimensions, the this effect has been

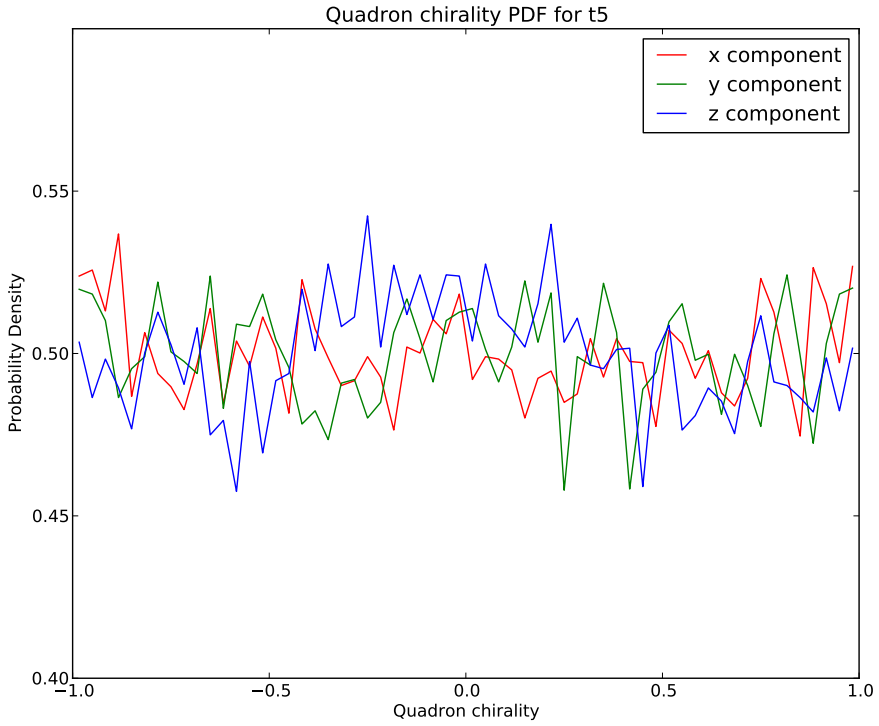
observed[62], and accounted for[63, 47] by a geometrical argument summarised in section 2.4.5.4.

5.1.5 Quadron Chirality

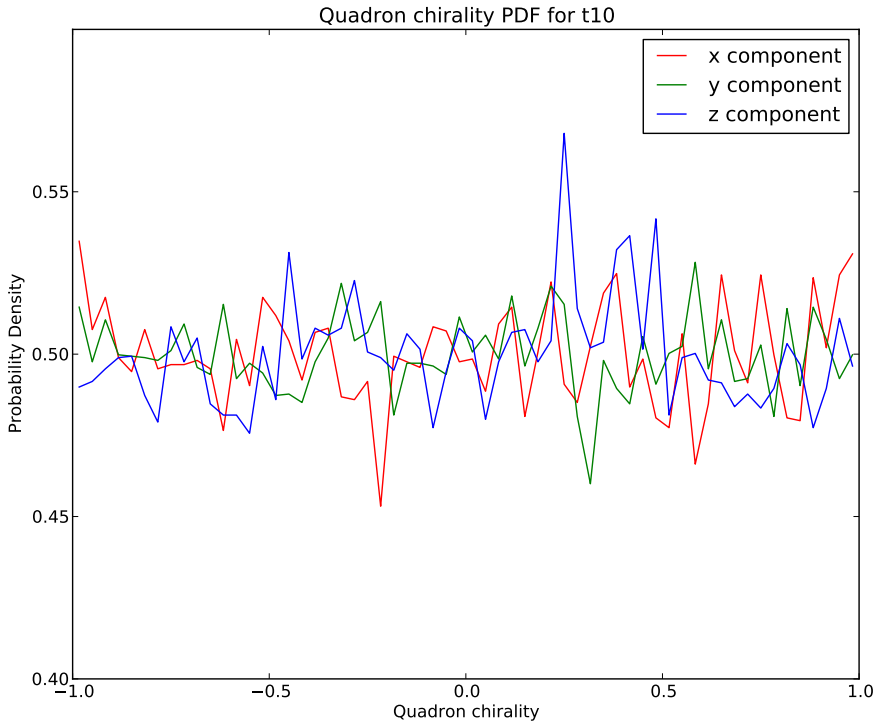
In the three different densities of oil droplet systems, the quadron chirality is isotropic, as can be observed in figure 5.4, where all three components of the normalised chirality pseudo-vector are roughly flat probability functions with some moderate noise.



(a) Isotropy of the quadron chirality in the densest oil droplet suspension[3].



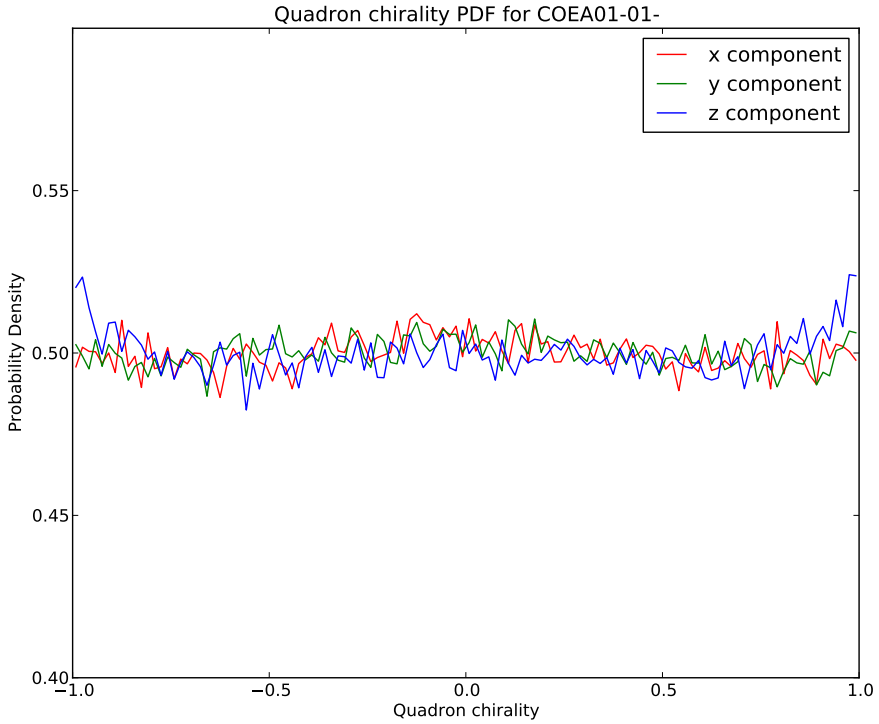
(b) Isotropy of the quadron chirality in the middle-density oil droplet suspension[3].



(c) Isotropy of the quadron chirality in the least dense oil droplet suspension[3].

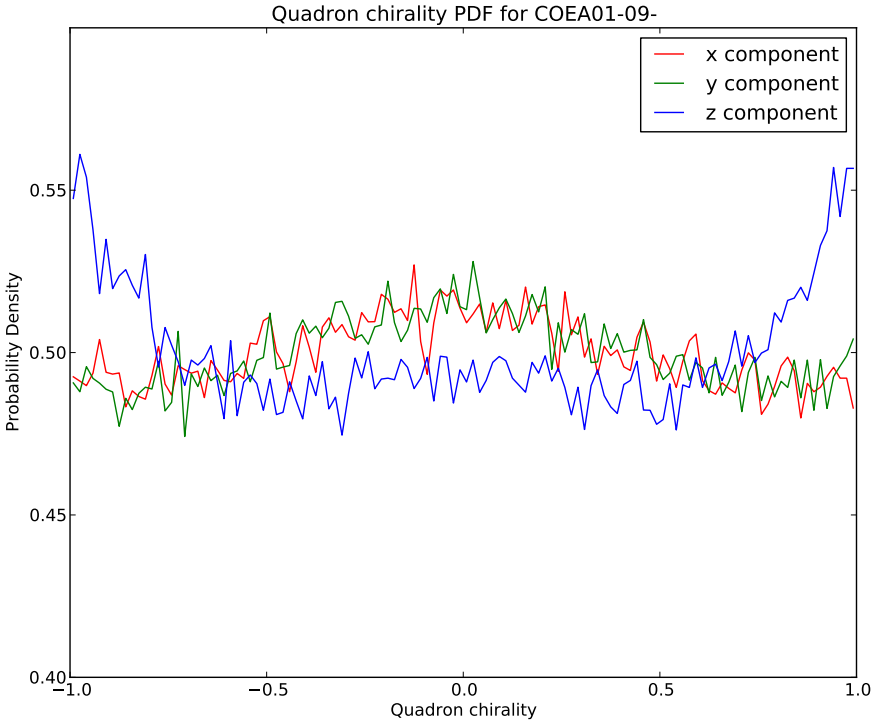
Figure 5.4 – Isotropy of oil droplet suspension[3] quadron chiralities.

In contrast, the quadron chirality in the Caicos systems shows some anisotropic bias towards vertical vectors. The effect is so small as to be discountable in the initial condition, but in the sheared system a distinctive signal emerges of the z component of the pseudo-vector preferring ± 1, and the other components compensating by preferring lower absolute values, as illustrated in figure 5.5



(a) Isotropy of the quadron chirality in the Caicos system initial condition[2].

Do note the scale – this is a plot of probability between limits 0.4 and 0.6.



(b) Anisotropy of the quadron chirality in the sheared Caicos system[2].

Figure 5.5 – The difference between the chirality distributions initial and sheared Caicos systems[2].

5.1.6 Sensitivity of Statistics to Physical Properties

5.1.6.1 Caicos Initial State versus Shear Band

The quadron chirality discussed previously is the first significant result demonstrating the sensitivity of the statistics of quadrons to physical properties – recall that the triaxial strain experiment performed by Andò et al at 3SR[2] applied vertical stress (plus a confining pressure) and it is this vertical direction that displays a bias in the quadron chirality pseudo-vectors.

In addition to this, one further significant result can be obtained from the same data-set, again by comparing the initial and sheared systems. In the original study[2], Andò et al observed the changes in grain position between scans, which were after each yield – a shear band was quite obvious in terms of the grain

rotation adjacent to a particular plane, and grain displacements were large and similarly oriented on one side of the plane and much smaller on the other. These observations are apparent in figure 2.18, taken from that study.

However, the quadron volume probability distribution PDF (v^q) depicted in figure 5.6 suggests that the system does not have two relatively stable regions connected by a chaotic shear band, as is clearly visible in the plots by individual grain behaviour – indeed, the bulk and shear band quadron volume distributions are quite similar, and both different to the initial condition. This suggests that, although the largest rearrangements take place in the shear band, the structure of the whole pack is altered significantly by the triaxial strain process.

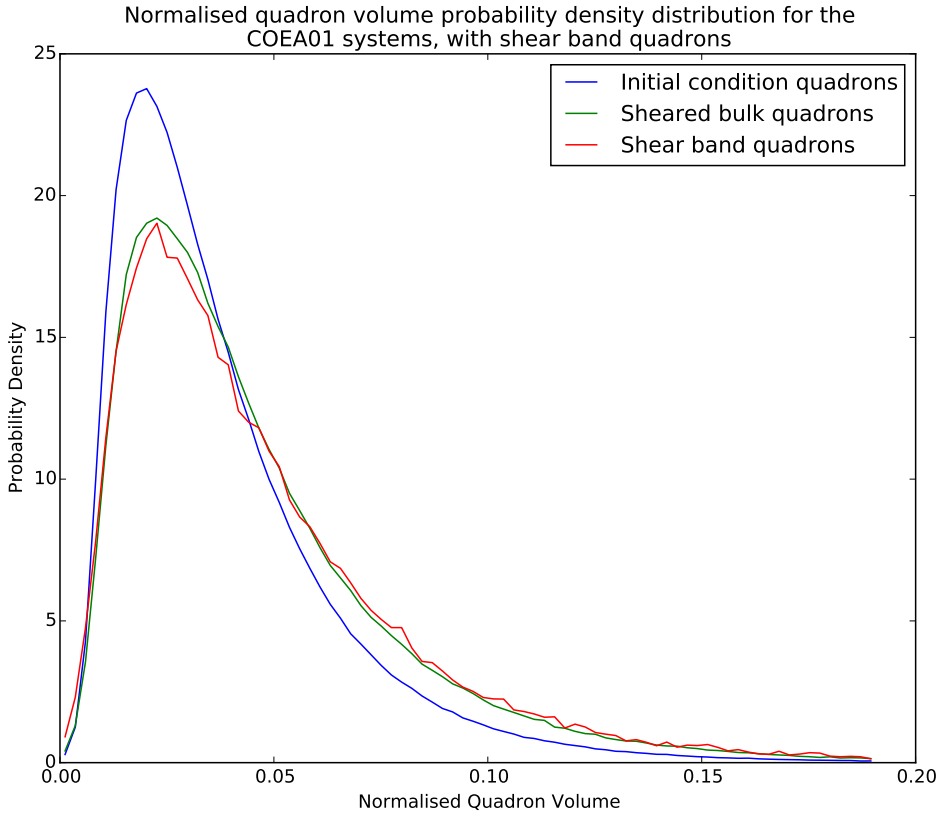
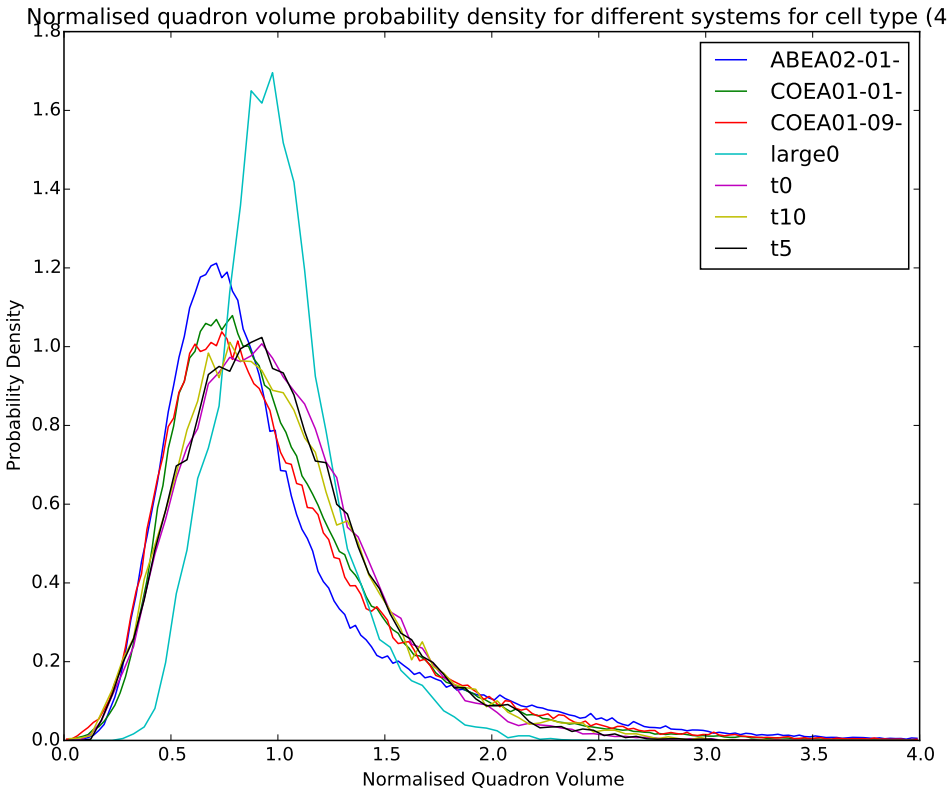


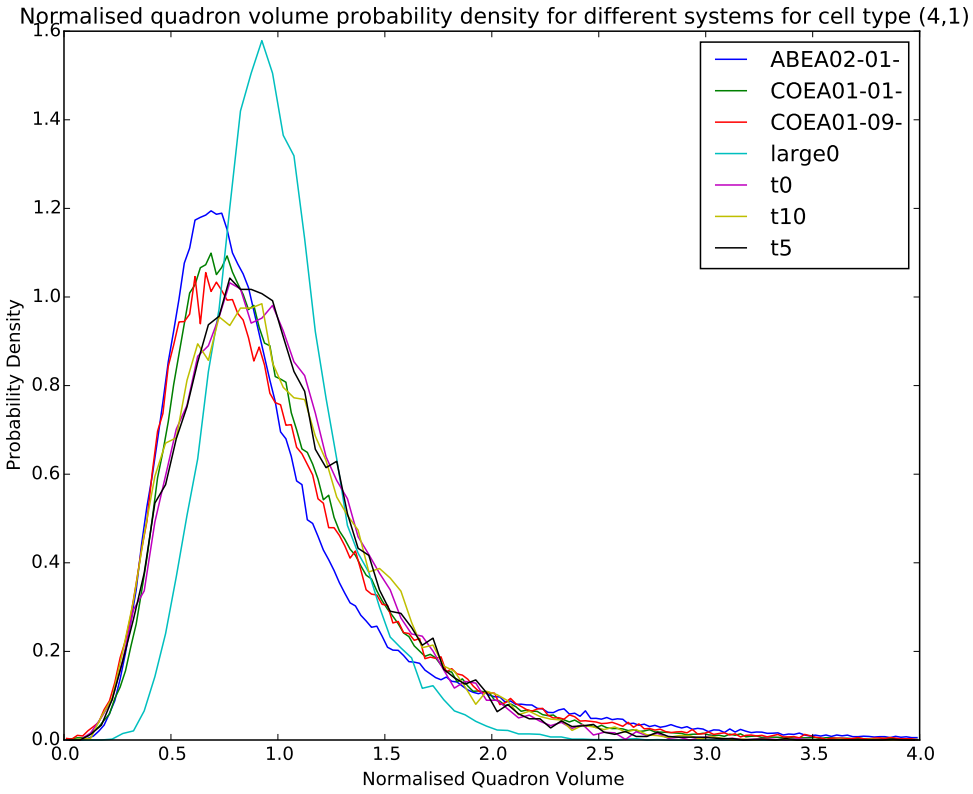
Figure 5.6 – PDF (v^q) for the Caicos experiment[2] initial condition, and the sheared system (at scan 9) separated by quadrons from the shear band and those in the bulk.

5.1.6.2 Comparison of Diverse Systems

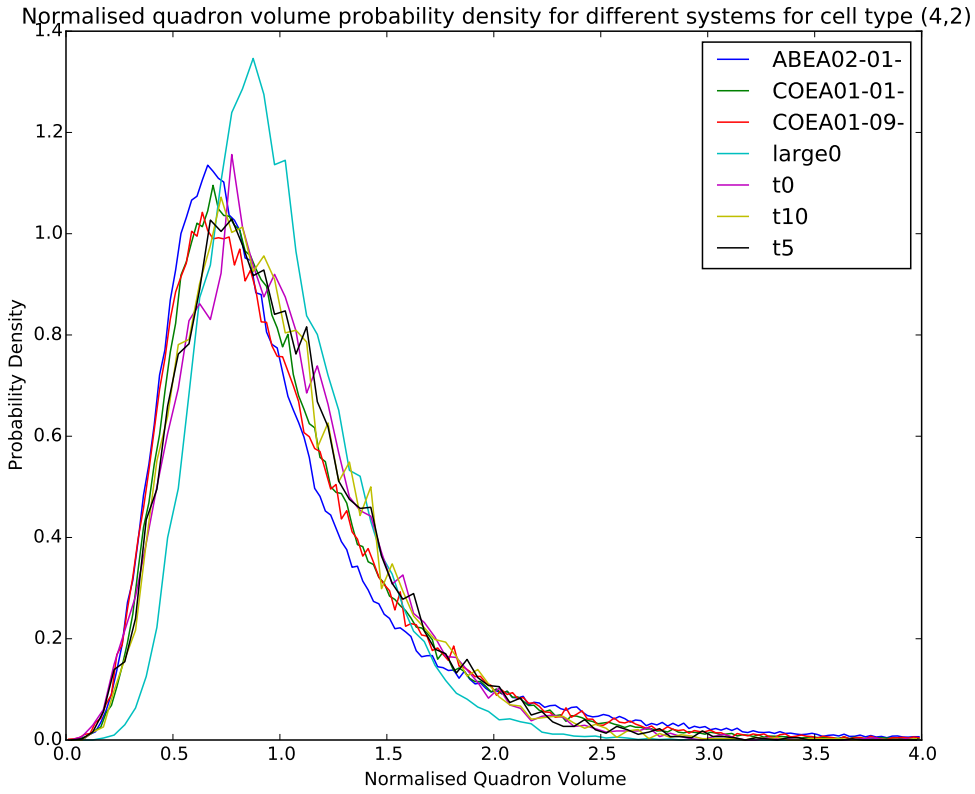
One significant result in the literature review, that of the collapse of the PDF $(v^q|e)$ conditional distributions for different friction coefficients[64, 65, 66] described in section 2.4.5.4, has a three dimensional analogue replacing cell order e with cell type t. The available data-sets do not include a set of experiments which *only* change friction coefficient, sadly, but the curves of different cell types can be plotted for all the available data-sets and any similarities observed. Plots for the three most common cell types, and one less common type, are given in figure 5.7.



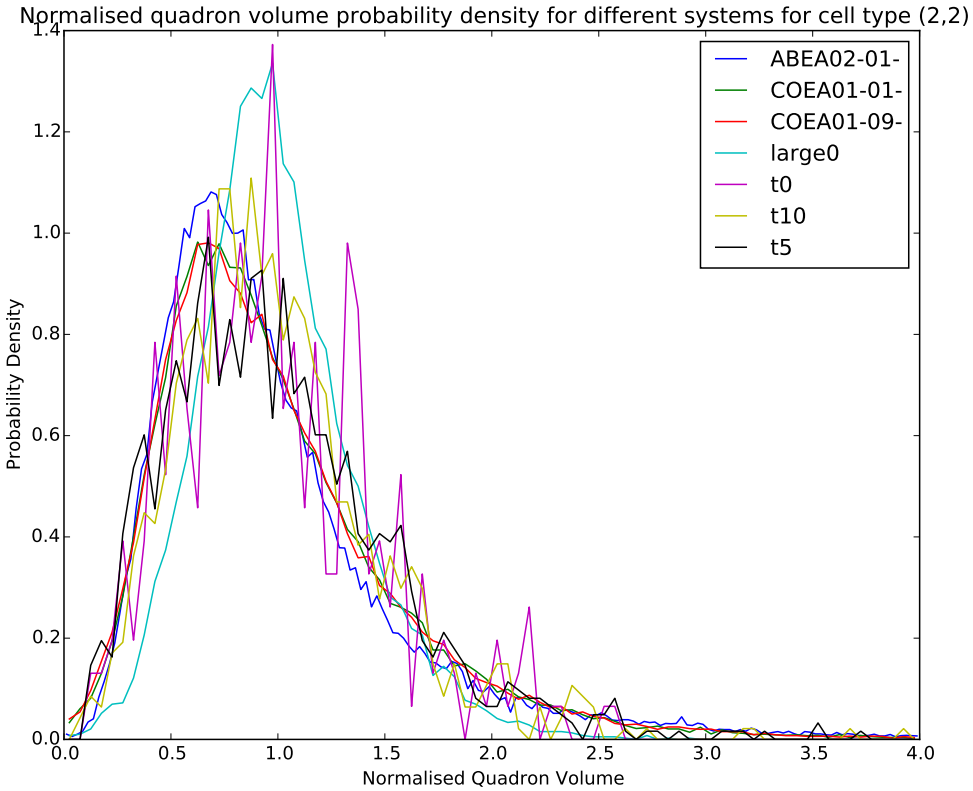
(a) PDF $(v^q|t)$ for $t = (4,)$ in all available data-sets, normalised so that $\bar{v^q} = 1$.



(b) PDF $(v^q|t)$ for $t = (4, 1)$ in all available data-sets, normalised so that $\bar{v^q} = 1$.



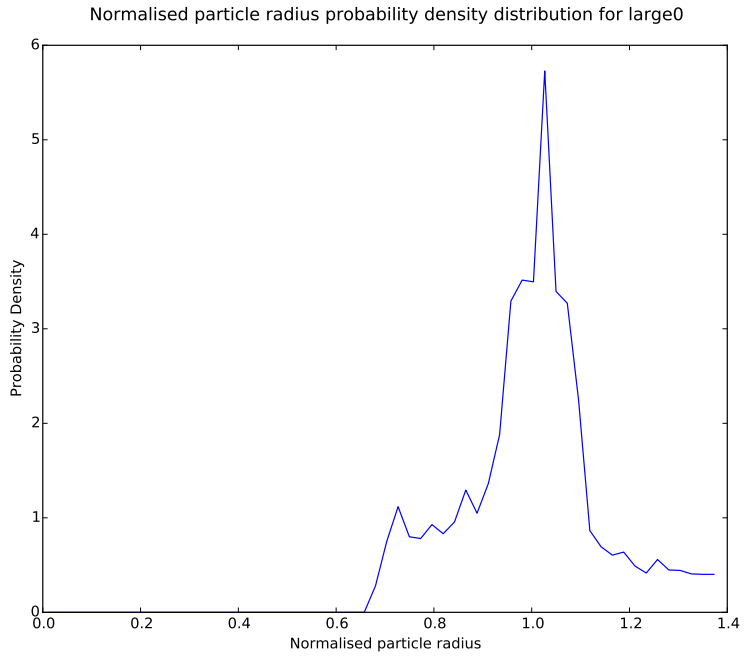
(c) PDF $(v^q|t)$ for $t = (4, 2)$ in all available data-sets, normalised so that $\bar{v^q} = 1$.



(d) PDF $(v^q|t)$ for $t = (2, 2)$ in all available data-sets, normalised so that $\bar{v^q} = 1$. Note the problem with cell types that are not one of the most frequent – it is very difficult to get good enough statistics to determine

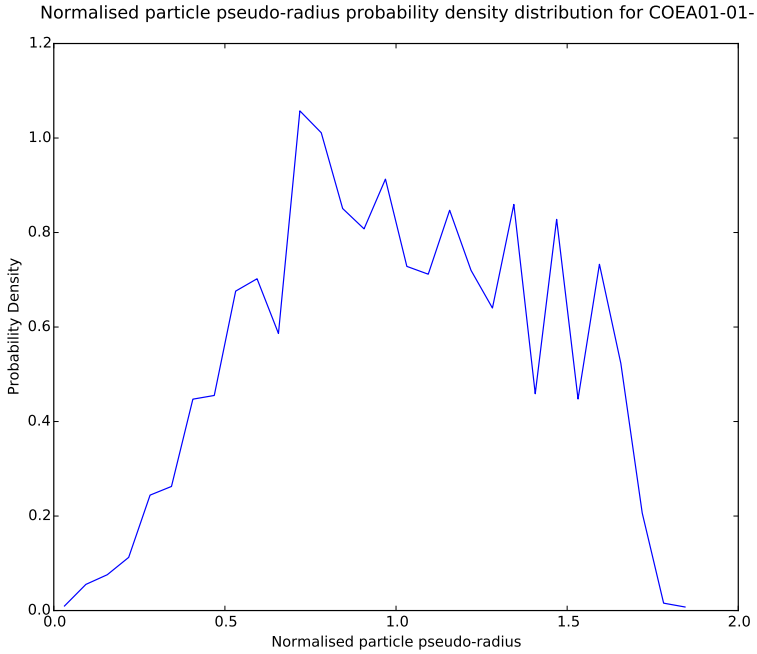
Figure 5.7 – PDF $(v^q|t)$ for common cell types plotted for all the available data-sets.

These curves do not collapse as in two dimensions, although these system differ significantly from the original work[64, 65, 66] and so they should not necessarily be expected to. In particular, the particles in these systems have different size distributions, which mean that the geometrical possibilities for cell geometries described in section 2.4.5.4 is different between them. For example, figure 5.8 shows the particle radii (or pseudo-radii[2]) for the sedimentary simulation[77], which is modelled after Fontainebleau sandstone, and the Caicos initial condition[2].



(a) Particle radius distribution in the large0 sedimentary simulation[77], modelled after Fontainebleau sandstone.

[2]for the non-spherical Caicos ooids, the pseudo-radii is the average distance from centroid to contact point



(b) Particle pseudo-radius distribution in the Caicos initial condition[2].

Figure 5.8 – Particle size distributions, which have been normalised so that the average radius is 1.

Further lending credence to the possibility that size distribution accounts in large part for the shapes of the conditional distributions, it should be observed in figure 5.7 that the two Caicos samples, which share the same particle set, collapse nicely for all the plots shown (they are depicted in red and green in each plot.)

5.1.7 Quality of Outcomes

It is an important caveat to the results presented above that the tool is not completely reliable when solving granular systems. Table 5.2 shows the *coverage* (the fraction of the system volume covered by valid quadrons), and the best system only reaches 90%. Note that the "total volume" used as the denominator in these fractions takes a convex hull of the whole system, which is perhaps a small overestimate of the volume, so a value in the high nineties is probably the maximum

attainable without using solid walls and defining boundary quadrons. Additionally presented in table 5.2 is the average coverage per slice for systems which were decomposed and reassembled. It is possible for quadrons from different slices to overlap in the final reassembled system, but this is as likely as gaps between the quadrons, so it is argued that this does not apply a systematic bias to the results.

Label	Particles	Quadrons	Slices	Coverage	Mean Slice Coverage
COEA01-01-	56,265	1,641,400	751	77%	62%
COEA01-09-	56,395	861,862	665	46%	36%
ABEA01-01-	65,911	1,765,730	859	78%	62%
large0	9,383	240,963	124	91%	67%
t10	2,880	69,508	1	75%	N/A
t5	3,044	80,978	1	80%	N/A
t0	3,327	101,824	1	78%	N/A

Table 5.2 – Proportion of the system volume successfully solved by the new software.

It is obvious that the sheared system COEA01-09-, with the large dilated shear band, was very difficult to solve. To further illustrate where the difficulties lay in this, refer to figure 5.9, which is a plot of average quadron volume by position. The volumes are averaged in a pixel-shaped prism projected in the direction of the plot – where the diagram is blank or white is where no quadron positions were found within that prism. Most diagrams of systems drawn like this (see section C.1) are solid colour up to the edge of the system boundary, which is clearly not the case for figure 5.9, which has a sprinkling of values in the shear band, and has a large square hole in the centre of the system: a symptom of the shear band slices being completely overwhelmed by the large pores.

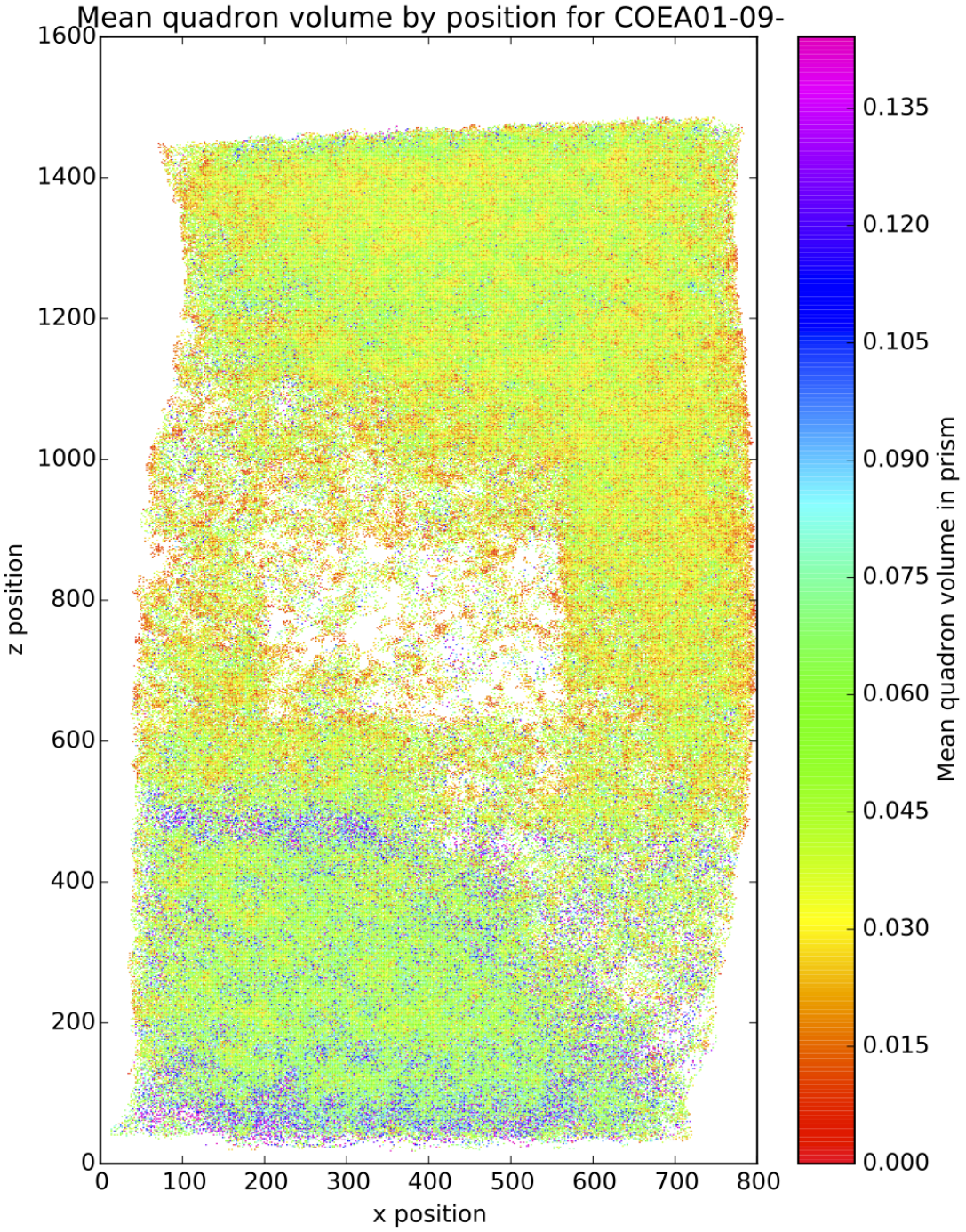


Figure 5.9 – Heatmap of quadron volumes projected onto the xz plane, for system COEA01-09-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it. Note that this projection is perpendicular to the shear band, which is visible (as gaps in the data, rather than as differently coloured points) and that the plot contains a large artefact of the slicing (the square region in the middle with no data values.)

Several improvements were suggested for the project in section 4.3.3.1, which would improve the possibility of overcoming obstacles like those presented by the Caicos triaxial shear experiment shear band.

Chapter 6

Conclusion

In this work the author has presented a two dimensional analytical solution which reveals an equipartition principle in granular systems – equiparition of volume among the degrees of freedom – and an original three dimensional algorithm which can compute the quadron coordinates for a three dimensional granular system for the first time.

Equiparition in Two Dimensions

The equipartition result:

$$\langle V \rangle = (\bar{z}N + 2Mc_f)\,\frac{X_0}{2}$$

is important for three reasons. Firstly, it demonstrates the power of the statistical mechanical approach for generating quantitative relations from first principals, and provides a template for future work towards an equation of state. Secondly, the result itself holds a mirror to thermal statistical mechanics mimicing the famous

$$\langle E \rangle = \frac{3}{2} k_B T$$

equipartition of energy. Finally, it provides an analytical expression for the compactivity in terms of quantities – the number of internal and boundary contacts and the total volume of the system – which are easily computable or measureable and thus acts as a means to evaluate the compactivity for two dimensional experimental systems.

It also reveals several factors about the compactivity which were previously unknown: it scales with grain size, so compactivity values are only compatible for similar particles without normalisation; it cannot be zero or infinite as the original theory of powders hypothesised; and its approximate value is the volume per contact.

The equipartition result is however quite limited in its applicability, being only valid for two dimensional systems, and being correct under the assumption of isostaticity (which corresponds to a restriction on the system density.)

Cycle-Centric Algorithm

The quadron statistics of three dimensional granular packs, which have never before been computed, have been presented in this work because of the successful conception and implementation of a new algorithm based on a cycle search.

The software as it stands solves medium-sized systems of a few thousand grains, which are also relatively densely packed, in a linear amount of time without excessive RAM requirements. It can also reliably solve large densly packed systems via a parallelisation method involving breaking them into smaller pieces and reassembling them (although the re-assembly stage can require larger amounts of

RAM as it must hold the whole system in memory at once.)

On less dense systems, like the sheared Caicos system, it solves as much of the system as it can, and returns results which can be plotted without including any of the invalid regions. The runtime of the software suffers significantly from the combinatorial problems inherant in finding a suitable partitioning of an invalid cell, and the memory copies required to make draft changes to the cell structure.

Quadron Solutions Analysis

It has been shown that the statistics from the quadron description, the volumes and the chiralities, are sensitive to the system history and the complex set of factors which determine granular behaviour, as was hoped. New results presented include:

- The cell classification is of similar complexity to three dimensional foams — in the larger systems hundreds of cell types were observed (see appendix section C.3).
- The quadron volume distributions break down by cell type to form single-peaked curves very similar to the two dimensional equivalent, but for the vastly increased number of cell types.
- The spatial distributions of quadron volumes reveal density variation, and in spite of the gaps in the data do detect the shear band.
- Quadron chirality distributions show that quadrons are in general homogeneous in skewness, except in the sheared Caicos system where a small bias towards the direction of principal stress was shown.

Going Further

The tools developed in this project are immidiately applicable to a wide range of problems in granular media: as a tool for the experimentalist, to allow them a rich statistical view on their data-sets; as a tool for a theorist working with simulation data and testing aspects of the statistical mechanics numerically, the possibilities for use are very wide.

Additionally, the statistics presented in this work are only a fraction of what can be done with the structure tensor: researchers could invesitgate distributions of throat sizes, orders and orientations; the volumes of cell elements and shapes, and cross-correlations between any of these and the other quantities already explored including spatial distribution.

Open Challenges

The algorithm has been implemented in Python, which is a good language to prototype in, and is sometimes performant enough to keep using beyond that stage. However, the finished solution to this problem doesn't seem to be a good fit for it.

One advantage of a language like Python – which has good support for compiled extensions in lower level languages – is that you can identify problem areas in the code and optimise only those, rather than committing to a wholesale reimplementation. In this way, Python can act as a high level shell for lower level data structures and routines.

The three problem areas with the existing implementation are: RAM usage, particularly for large systems; the speed of the geometrical constraint validation; and the limitation on the grain geometry, requiring convexity where the quadron de-

scription can handle slightly non-convex structures. Improvements could therefore be made to this work to increase applicability and utility as follows:

- Commonly used data structures, like the NetworkX Graph class, should be re-implemented using a more compact memory representation. The default structure has been optimised for run-time speed, using Python dictionaries with $O(1)$ look-up time, at the expense of RAM footprint. This upgrade would allow the existing routines to run on larger data-sets without further modification.

- The geometrical constraint verification could be consolidated into a single, optimised routine. This would have an immediate impact on the run-time for less dense systems, and make the software much more practical for such cases to be explored fully, in all their combinatorial complexity.

- Extending the routines which compute the initial grain surfaces from a cloud of contact points – which deal very harshly with any non-convexity – to more sympathetically map a surface structure to use as a starting point would extend the possible inputs to many highly aspherical granular systems. No doubt, though, such work would also include making the geometrical constraint verification more robust to diverse and unexpected structures.

Bibliography

[1] S. F. Edwards and R. B. S. Oakeshott. Theory of powders. *Physica A*, 157(3):1080–1090, 1989.

[2] Edward Andò, Stephen A. Hall, Gioacchino Viggiani, Jacques Desrues, and Pierre Bésuelle. Grain-scale experimental investigation of localised deformation in sand: a discrete particle tracking approach. *Acta Geotechnica*, 7(1):1–13, 2012.

[3] Jasna Brujić, C. Song, Ping Wang, C. Briscoe, G. Marty, and H. A. Makse. Measuring the coordination number and entropy of a 3d jammed emulsion packing by confocal microscopy. *Physical Review Letters*, 98(24):248001, June 2007.

[4] Osborne Reynolds. Lvii. on the dilatancy of media composed of rigid particles in contact. with experimental illustrations. *Philosophical Magazine Series 5*, 20(127):469–481, 1885.

[5] W. O. Smith, Paul D. Foote, and P. F. Busang. Packing of homogeneous spheres. *Phys. Rev.*, 34:1271–1274, Nov 1929.

[6] S Murayama and H Matsuoka. A microscopic consideration of the shearing behavior of granular materials using two-dimensional models. *Disaster Prevention Res Inst Bul/Jap/*, 19:60, 1970.

[7] E. R. Nowak, J. B. Knight, M. L. Povinelli, H. M. Jaeger, and S. R. Nagel. Reversibility and irreversibility in the packing of vibrated granular material. *Powder Technology*, 94(1):79–83, 1997.

[8] P. Richard, P. Philippe, F. Barbe, S. Bourlès, X. Thibault, and D. Bideau. Analysis by x-ray microtomography of a granular packing undergoing compaction. *Physical Review E*, 68(2):020301, August 2003.

[9] C. A. Coulomb. *Essai sur une application des règles de maximis & minimis à quelques problèmes de statique, relatifs à l'architecture.* De l'Imprimerie Royale, 1776. Title translated to English: Test on an application of the maximum and minimum rules and some static problems related to architecture.

[10] Karl Terzaghi. *Erdbaumechanik (Soil Mechanics).* Leipzig/Vienna, 1925.

[11] P. W. Rowe. The stress-dilatancy relation for static equilibrium of an assembly of particles in contact. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 269(1339):500–527, 1962.

[12] M. R. Horne. The behaviour of an assembly of rotund, rigid, cohesionless particles. i. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 286(1404):62–78, 1965.

[13] Robin C. Ball and Raphael Blumenfeld. Stress field in granular systems: Loop forces and potential formulation. *Physical Review Letters*, 88(11):115505, Feb 2002.

[14] R. Blumenfeld and S. F. Edwards. Geometric partition functions of cellular systems: Explicit calculation of the entropy in two and three dimensions. *The European Physical Journal E: Soft Matter and Biological Physics*, 19:23–30, 2006. 10.1140/epje/e2006-00014-7.

[15] Raphael Blumenfeld, Joe F. Jordan, and Sam F. Edwards. Interdependence of the volume and stress ensembles and equipartition in statistical mechanics of granular systems. *Phys. Rev. Lett.*, 109:238001, Dec 2012.

[16] Serge Beucher and Christian Lantuéjoul. Use of watersheds in contour detection. In *International Workshop on Image Processing: Real-time Edge and Motion Detection/Estimation, Rennes, France.*, September 1979.

[17] Roland Glantz and Markus Hilpert. Dual models of pore spaces. *Advances in Water Resources*, 30(2):227 – 248, 2007.

[18] Roland Glantz and Markus Hilpert. Tight dual models of pore spaces. *Advances in Water Resources*, 31(5):787 – 806, 2008.

[19] Béla Bollobás. *Modern Graph Theory*, volume 184 of *Graduate Texts in Mathematics*. Springer, 1998.

[20] U. S. R. Murty J. A. Bondy. *Graph Theory*. Graduate Texts in Mathematics. Springer, 2008.

[21] Paul Butler. Visualizing friendships. December 2010.

[22] Jack Edmonds. Matroids and the greedy algorithm. *Mathematical Programming*, 1(1):127–136, 1971.

[23] Antonio A. Munjiza. *The Combined Finite-Discrete Element Method*. Wiley, 2004.

[24] Mordecai J. Golin and Hyeon-Suk Na. On the average complexity of 3d-voronoi diagrams of random points on convex polytopes. *Computational Geometry*, 25(3):197 – 231, 2003.

[25] Kevin Q. Brown. Voronoi diagrams from convex hulls. *Information Processing Letters*, 9(5):223 – 228, 1979.

[26] Donald W. Taylor. *Fundamentals of Soil Mechanics*. John Wiley & Sons-New York Inc., Chapman & Hall-London, 1948.

[27] A. W. Skempton; A. W. Bishop. The measurement of the shear strength of soils. *Gotechnique*, 2:90–108, December 1950.

[28] P. L. Newland and B. H. Allely. Volume changes in drained taixial tests on granular materials. *Gotechnique*, 7:17–34(17), 1957.

[29] M. R. Horne. The behaviour of an assembly of rotund, rigid, cohesionless particles. ii. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 286(1404):79–97, 1965.

[30] J Litwiniszyn. The model of a random walk of particles adopted to researches on problem of mechanics of loose media. *Bulletin lAcademie Polonaise Sciences*, 11:61–70, 1963.

[31] Jerzy Litwiniszyn. An application of the random walk argument to the mechanics of granular media. In Julien Kravtchenko and PierreMarcel Sirieys, editors, *Rheology and Soil Mechanics*, International Union of Theoretical and Applied Mechanics, pages 82–89. Springer Berlin Heidelberg, 1966.

[32] J.K. Wilkins. A theory for the shear strength of rockfill. *Rock mechanics*, 2(4):205–222, 1970.

[33] Masanobu Oda. The mechanism of fabric changes during compressional deformation of sand. *Soils and foundations*, 12(2):1–18, 1972.

[34] Masanobu Oda. A mechanical and statistical model of granular material. *Soils and Foundations*, 14(1):13–27, 1974.

[35] S Murayama and H Matsuoka. A microscopic study on shearing mechanism of soils. In *Proceedings of the 8th International Conference on Soil Mechanics*

and Foundation Engineering. Rotterdam: Balkema AA, volume 1, pages 293–298, 1973.

[36] Sakuro Murayama. A theoretical consideration on a behaviour of sand. In J. Kravtchenko and P.M. Sirieys, editors, *Rheology and Soil Mechanics*, pages 146–159. Springer Berlin Heidelberg, 1966.

[37] Trushant S Majmudar and Robert P Behringer. Contact force measurements and stress-induced anisotropy in granular materials. *Nature*, 435(7045):1079–1082, 2005.

[38] Edwin B. Matzke and Jane Nestler. Volume-shape relationships in variant foams. a further study of the role of surface forces in three-dimensional cell shape determination. *American Journal of Botany*, 33(2):pp. 130–144, 1946.

[39] D. Lafeber. Soil structural concepts. *Engineering Geology*, 1(4):261 – 290, 1966.

[40] Takeo Mogami. A statistical approach to the mechanics of granular materials. *Soil and Foundations*, 5(2):26–36, mar 1965.

[41] PW Jowitt and J Munro. The influence of void distribution and entropy on the engineering properties of granular media. In *Proc. 2nd Int. Conf. on Applications of Statistics and Probability in Soil and Structural Engineering, Aachen*, pages 365–385, 1975.

[42] C. B. Brown. The use of maximum entropy in the characterization of granular media. In S. C. Cowin and M. Satake, editors,
textitProc. US-Japan Seminar on Continuum Mechanics and Statistical Approaches in the Mechanics of Granular Media, pages 98–108, 1978.

[43] Anita Mehta and S.F. Edwards. Statistical mechanics of powder mixtures. *Physica A*, 157(3):1091–1100, 1989.

[44] S.F. Edwards. The full canonical ensemble of a granular system. *Physica A: Statistical Mechanics and its Applications*, 353(0):114 – 118, 2005.

[45] Silke Henkes, Corey S. O'Hern, and Bulbul Chakraborty. Entropy and temperature of a static granular assembly: An ab initio approach. *Physical Review Letters*, 99(3):038002, Jul 20 2007.

[46] Raphael Blumenfeld. Stress transmission and isostatic states of non-rigid particulate systems. *IMA Volume in Mathematics and its Applications*, 141:235–246, 2005.

[47] Rebecca Hihinashvili. *Morphological Characterisation Of Porous Materials For Fuel Cell Technology*. PhD thesis, Imperial College London, 2013.

[48] B. Miller, C. O'Hern, and R.P. Behringer. Stress fluctuations for continuously sheared granular materials. *Physical Review Letters*, 77(15):3110–13, 1996.

[49] James G. Puckett, Frédéric Lechenault, and Karen E. Daniels. Local origins of volume fraction fluctuations in dense granular materials. *Physical Review E*, 83:041301, Apr 2011.

[50] V. Richefeu, G. Combe, and G. Viggiani. An experimental assessment of displacement fluctuations in a 2d granular material subjected to shear. *Gotechnique Letters*, 2:113–118(5), 2012.

[51] M Saadatfar, A Kabla, TJ Senden, and T Aste. The geometry and the number of contacts of monodisperse sphere packs using x-ray tomography. In *Powders and Grains 2005-Proceedings of the 5th International Conference on Micromechanics of Granular Media*, volume 1, pages 33–36, 2005.

[52] Rutger CA Smit. Discrete 3d digital image correlation (dic) using particle reconnaissance. *Delft University of Technology*, 2010.

[53] Edmund R. Nowak, James B. Knight, Eli Ben-Naim, Heinrich M. Jaeger, and Sidney R. Nagel. Density fluctuations in vibrated granular materials. *Physical Review E*, 57(2):1971–1982, Feb 1998.

[54] Ning Xu, Daan Frenkel, and Andrea J. Liu. Direct determination of the size of basins of attraction of jammed solids. *Physical Review Letters*, 106:245502, Jun 2011.

[55] Daan Frenkel, Daniel Asenjo, and Fabien Paillusson. The other entropy. *Molecular Physics*, 111(22-23):3641–3650, 2013.

[56] Daniel Asenjo, Fabien Paillusson, and Daan Frenkel. Numerical calculation of granular entropy. *Phys. Rev. Lett.*, 112:098002, Mar 2014.

[57] T. Aste, T. Di Matteo, M. Saadatfar, T. J. Senden, M. Schröter, and H. L. Swinney. An invariant distribution in static granular media. *Europhysics Letters*, 792:24003, July 2007.

[58] T. Aste and T. Di Matteo. Emergence of gamma distributions in granular materials and packing models. *Physical Review E*, 77(2):021309, February 2008.

[59] T. Aste and T. Di Matteo. Structural transitions in granular packs: statistical mechanics and statistical geometry investigations. *The European Physical Journal B*, 64:511–517, 2008. 10.1140/epjb/e2008-00224-8.

[60] T. Aste, G. W. Delaney, and T. Di Matteo. kgamma distributions in granular packs. *AIP Conference Proceedings*, 1227(1):157–166, 2010.

[61] R. V. Hogg and A. T. Craig. *Introduction to Mathematical Statistics*. Macmillan, New York, 1978.

[62] Gad Frenkel, Raphael Blumenfeld, Zdenek Grof, and Peter R. King. Structural characterization and statistical properties of two-dimensional granular

systems. *Physical Review E*, 77(4):041304, Apr 2008.

[63] L Toikka. Private communication. Reference 8 on [78], 2009.

[64] Takashi Matsushima and Raphael Blumenfeld. Fundamental structural characteristics of planar granular assemblies: scaling away friction and initial state. *ArXiv e-prints*, 2012.

[65] Takashi Matsushima and Raphael Blumenfeld. Microstructural characteristics of planar granular solids. *AIP Conference Proceedings*, 1542(1):325–328, 2013.

[66] Takashi Matsushima and Raphael Blumenfeld. Universal structural characteristics of planar granular packs. *Phys. Rev. Lett.*, 112:098003, Mar 2014.

[67] Raphael Blumenfeld and Sam F. Edwards. Granular entropy: Explicit calculations for planar assemblies. *Physical Review Letters*, 90:114303, Mar 2003.

[68] Guido van Rossem. The python programming language, version 2.7. `http://www.python.org`.

[69] Later the Khronos Group. Silicon Graphics Computer Systems. The opengl graphics programming interface, 1992.

[70] Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th Python in Science Conference (SciPy2008)*, pages 11–15, Pasadena, CA USA, August 2008.

[71] Chris H. Rycroft. Voro++: A three-dimensional voronoi cell library in c++. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 19(4):–, 2009.

[72] Cgal, Computational Geometry Algorithms Library. http://www.cgal.org.

[73] S. Behnel, R. Bradshaw, C. Citro, L. Dalcin, D.S. Seljebotn, and K. Smith. Cython: The best of both worlds. *Computing in Science Engineering*, 13(2):31 –39, 2011.

[74] Elisângela Silva Dias, Diane Castonguay, Humberto J. Longo, and Walid Abdala Rfaei Jradi. Efficient enumeration of all chordless cycles in graphs. *Algorithmica (submitted)*, abs/1309.1051, 2013.

[75] Andrew M. Kraynik, Douglas A. Reinelt, and Frank van Swol. Structure of random monodisperse foam. *Phys. Rev. E*, 67:031403, Mar 2003.

[76] Edward Andò. Private communication.

[77] Stig Bakke. Private communication. Communication between Joe F Jordan and Stig Bakke at Lithicon.

[78] Raphael Blumenfeld. Structural evolution of granular systems: Theory. *ArXiv e-prints*, 1412.6933, 2014.

[79] H G Dawson. On the numberical value of $\int_0^h e^{x^2} dx$. *Proceedings of the London Mathematical Society*, s1-29:519–522, 1898.

Appendices

Appendix A

Definitions

Here are summarised all of the mathematical notations used in this work.

Table A.1 – Mathematical definitions

Notation	Meaning
$\vec{x}$	A vector.
$\hat{\vec{x}}$	A unit vector.
$\vec{x}_i$	A component of a vector.
$\lvert\vec{x}\rvert$	A length of a vector.
$\mathbb{R}^n$	n-dimensional Euclidean space.
$\mathbf{S}$	A region of a space.
$\vec{a_c}$	A centroid.
G	A graph.
$V(G)$ or $E(G)$	G is a graph, and V and E are the vertex and edge functions.
C_{ij} or C	A rank-2 tensor, C short-form where explicitly defined in text.
$\lvert C\rvert$	The determinant of tensor C.

Continued on next page

Table A.1 – continued from previous page

Notation	Meaning
$\mathrm{Tr}\, C$	The trace of tensor C.
$\hat{C}_{ij}$	*the structure tensor.*
$\{s_i\}$	A set.
s_i	A member of a set.
PDF ()	A probability density function.
$\vec{x} \cdot \vec{y}$	A dot product of two vectors.
$\vec{x} \wedge \vec{y}$	A cross product of two vectors.
$\vec{x} \otimes \vec{y}$	An outer product of two vectors.
$\mathcal{H}$	A Hamiltonian function.
$\mathcal{W}$	A volume function.

Appendix B

Known Integrals

B.1 Gaussian Integrals

Starting with the simplest Gaussian curve, with amplitude a, standard deviation $\frac{1}{\sqrt{2b}}$ and mean 0:

$$\int_{-\infty}^{\infty} a \exp\left(-bx^2\right) \mathrm{d}x = a\sqrt{\frac{\pi}{b}} \tag{B.1}$$

And then the same, where a factor of x^2 has multiplied the integrand:

$$\int_{-\infty}^{\infty} ax^2 \exp\left(-bx^2\right) \mathrm{d}x = \frac{a\sqrt{\pi}}{2\sqrt{b^3}} \tag{B.2}$$

Next the quadratic form $\vec{x}A\vec{x}$ is considered, where A is a diagonalisable matrix:

$$\int\!\!\int_{-\infty}^{\infty} \exp\left(-\vec{x}A\vec{x}\right) \mathrm{d}^N\vec{x} = \prod_i^N \sqrt{\frac{\pi}{a_i}} \tag{B.3}$$

where a_i is the ith eigenvalue of A in the diagonalisation chosen (transformations

don't change the eigenvalues, only their ordering.) This expression is basis independent, as the bottom half of the fraction is simply $\sqrt{\det A}$. Note that the constant b from the one dimensional case has been absorbed into A, and multiplies every eigenvalue. This result gives the value of a quadratic partition function for N degrees of freedom.

Using the forms from equations B.2 and B.3, the following expectation value can be deduced:

$$\langle \vec{x} A \vec{x} \rangle = \frac{\iint_{-\infty}^{\infty} \vec{x} A \vec{x} \, \exp\left(-\vec{x} A \vec{x}\right) \mathrm{d}^N \vec{x}}{\iint_{-\infty}^{\infty} \exp\left(-\vec{x} A \vec{x}\right) \mathrm{d}^N \vec{x}} = \sum_i^N \frac{a_i \sqrt{\frac{\pi}{a_i^3}}}{\sqrt{\frac{\pi}{a_i}}} = N \tag{B.4}$$

Note that in the first step, the large and unwieldy product/determinant from equation B.3 (running over a different dummy index j) cancels directly for all $j \neq i$.

Finally, a non-trivial expectation value, $\langle \vec{x} B \vec{x} \rangle$, is demonstrated:

$$\langle \vec{x} B \vec{x} \rangle = \frac{\iint_{-\infty}^{\infty} \vec{x} B \vec{x} \, \exp\left(-\vec{x} A \vec{x}\right) \mathrm{d}^N \vec{x}}{\iint_{-\infty}^{\infty} \exp\left(-\vec{x} A \vec{x}\right) \mathrm{d}^N \vec{x}} = \sum_i^N \frac{B'_{ii} \sqrt{\frac{\pi}{a_i^3}}}{\sqrt{\frac{\pi}{a_i}}} = \sum_i^N \frac{B'_{ii}}{a_i} \tag{B.5}$$

Where B' has been defined as $B' = Y^T B Y$, and where Y is the orthogonal matrix used to diagonalise A. The final form is much less elegant than in equation B.3 – this is because the diagonalisations of the matrices do not coincide as before (hence the guts of the diagonalisation chosen, Y, being visible in the result.)

B.2 Dawson Functions

The Dawson function was originally defined[79] as:

$$\mathrm{Daw}_0(\alpha) = \int_0^{\alpha} \exp\left(x^2\right) \mathrm{d}x \tag{B.6}$$

In this work it is useful to consider the same function over different limits, so Daw () is defined as:

$$\mathrm{Daw}(\alpha) = \int_{-\alpha}^{\alpha} \exp\left(x^2\right) \mathrm{d}x \tag{B.7}$$

Note that if there is some constant factor β in the exponent, the dummy variable can be re-scaled to give:

$$\int_{-\alpha}^{\alpha} \exp\left(\beta x^2\right) \mathrm{d}x = \frac{1}{\sqrt{\beta}} \mathrm{Daw}\left(\sqrt{\beta}\alpha\right) \tag{B.8}$$

In the case that the integral which needs to be evaluated is the Dawson integral with an additional factor of x^2, then by parts this can be simplified to:

$$\int_{-\alpha}^{\alpha} \beta x^2 \exp\left(\beta x^2\right) \mathrm{d}x = \alpha \exp\left(\beta\alpha^2\right) - \frac{1}{2\sqrt{\beta}} \mathrm{Daw}\left(\sqrt{\beta}\alpha\right) \tag{B.9}$$

B.3 An Exact form for Definite Integrals

A definite integral, that is an integral with finite limits, is the area under the curve in a plot of the integrand as a function of the variable of integration, see figure B.1.

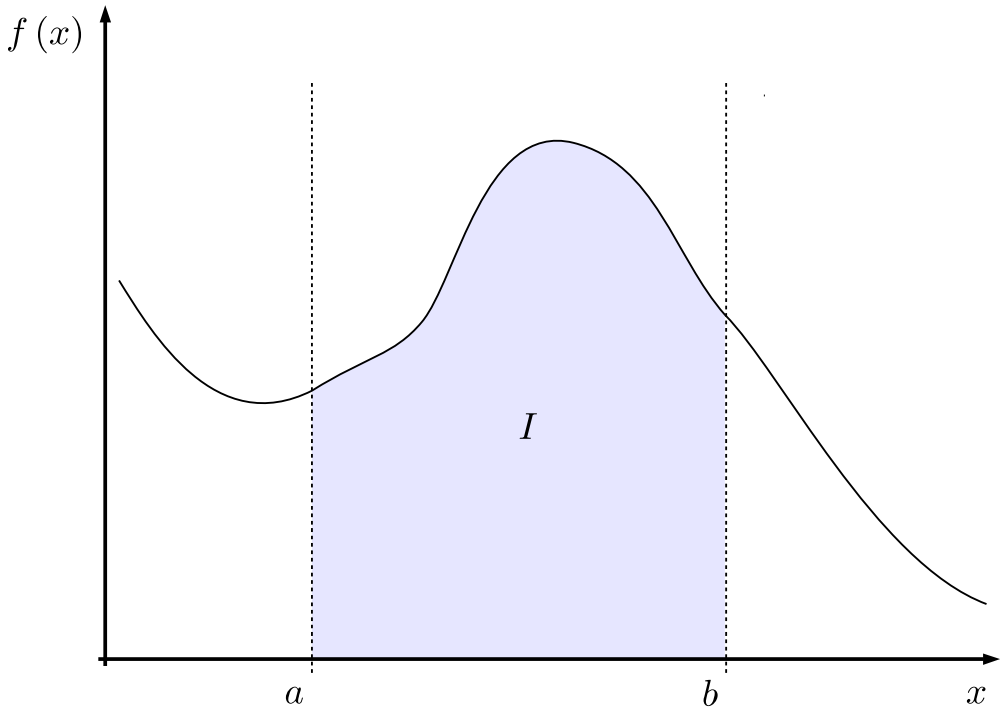


Figure B.1 – A definite integral, illustrated as the area under a curve between two limits.

The same value (area) can always be represented as a rectangle, with the same width as the integration limits and some height which represents the average value of the integrand between those limits, see figure B.2.

This allows any integral to be written as the difference of its limits multiplied by the function average, which can be written as the function taking some argument between the limits of integration for all continuous functions:

$$\int_a^b f(x)\ \mathrm{d}x = (b-a) f(x_0)\ ,\ a < x_0 < b \tag{B.10}$$

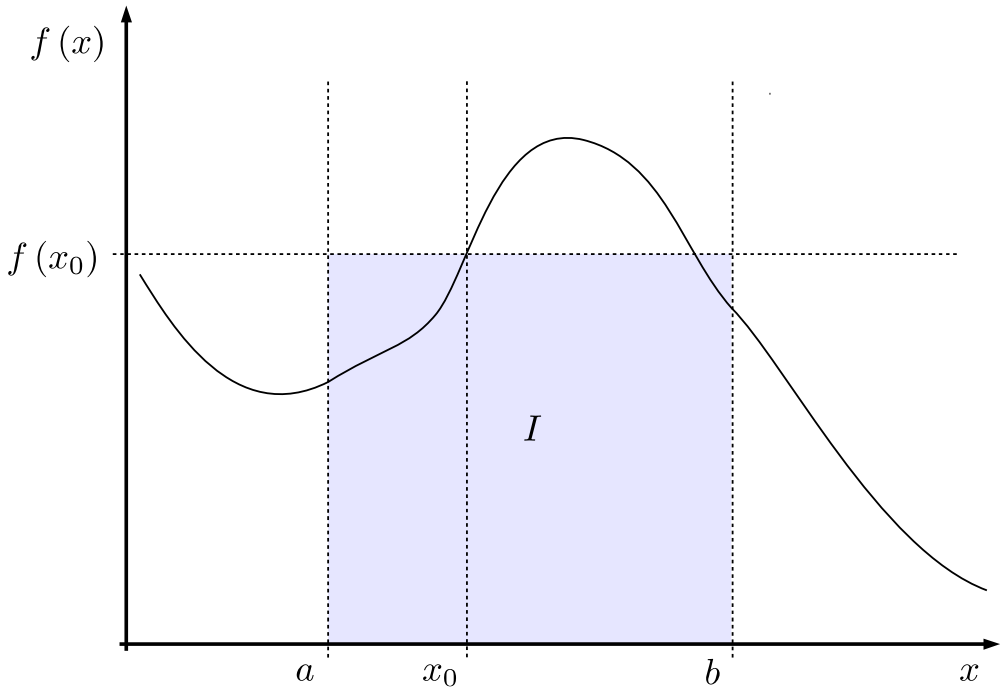


Figure B.2 – A definite integral as the average function value between the limits of integration.

Appendix C

Comprehensive Results

This appendix contains the distributions computed for every available data-set.

C.1 Distributions of Quadron Volumes

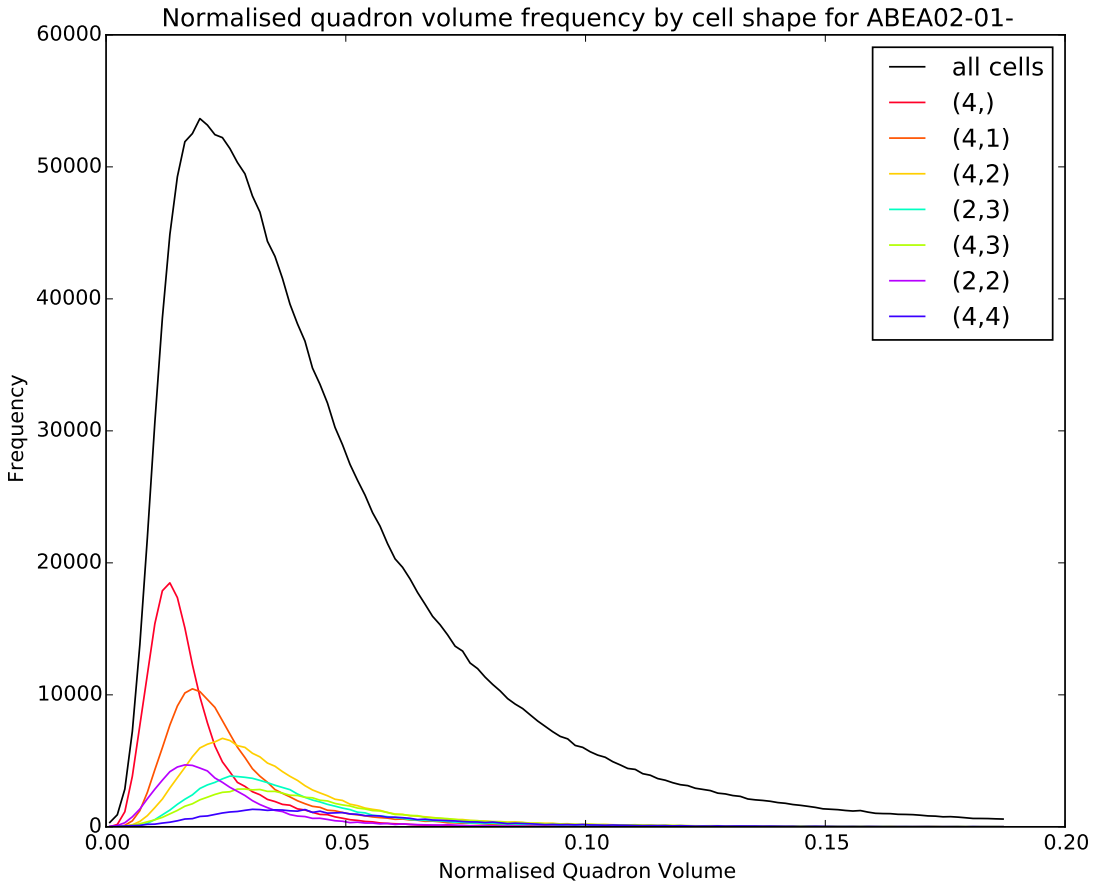


Figure C.1 – Total quadron volume frequency distribution $f\left(v^q\right)$, plotted with the seven most frequent cells' conditional frequency distributions $f\left(v^q|t\right)$, for system ABEA02-01-.

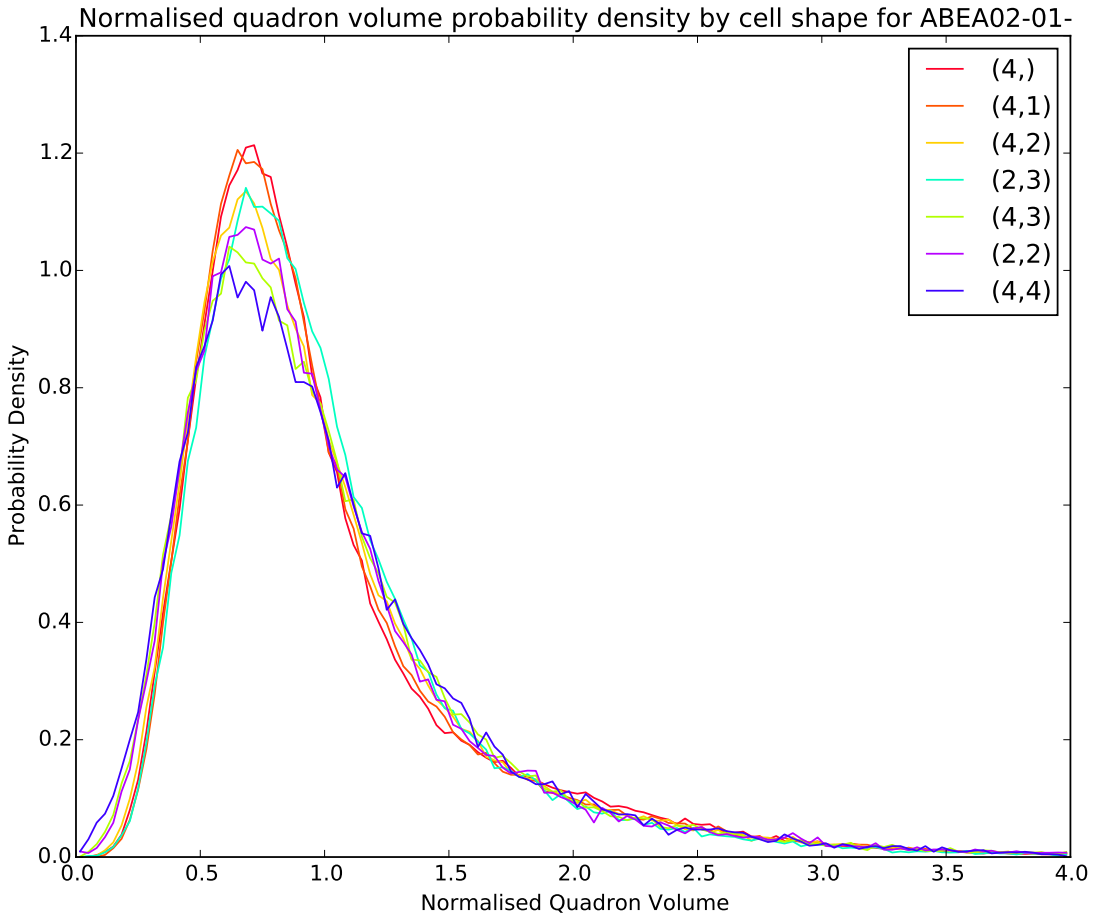


Figure C.2 – Seven most frequent cells' conditional probability distributions PDF $(v^q|t)$, for system ABEA02-01-.

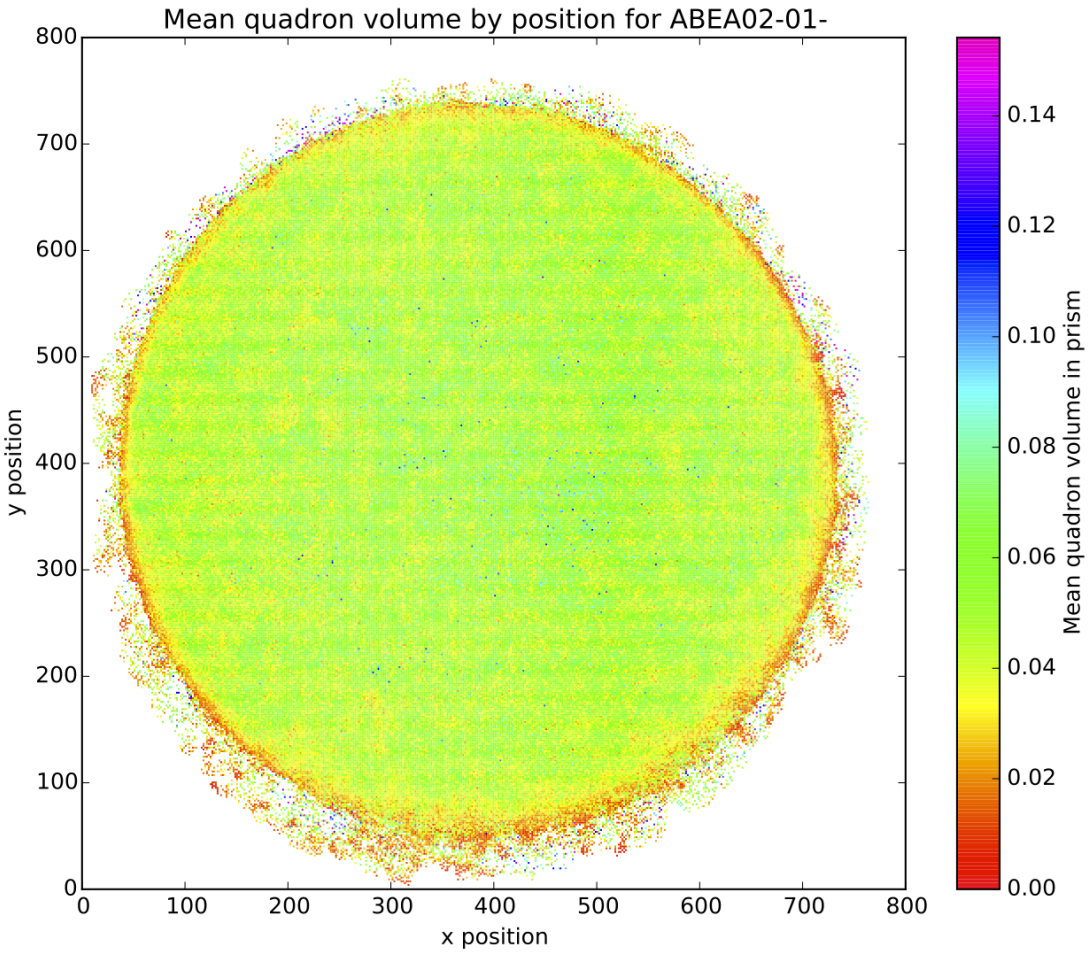


Figure C.3 – Heatmap of quadron volumes projected onto the xy plane, for system ABEA02-01-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

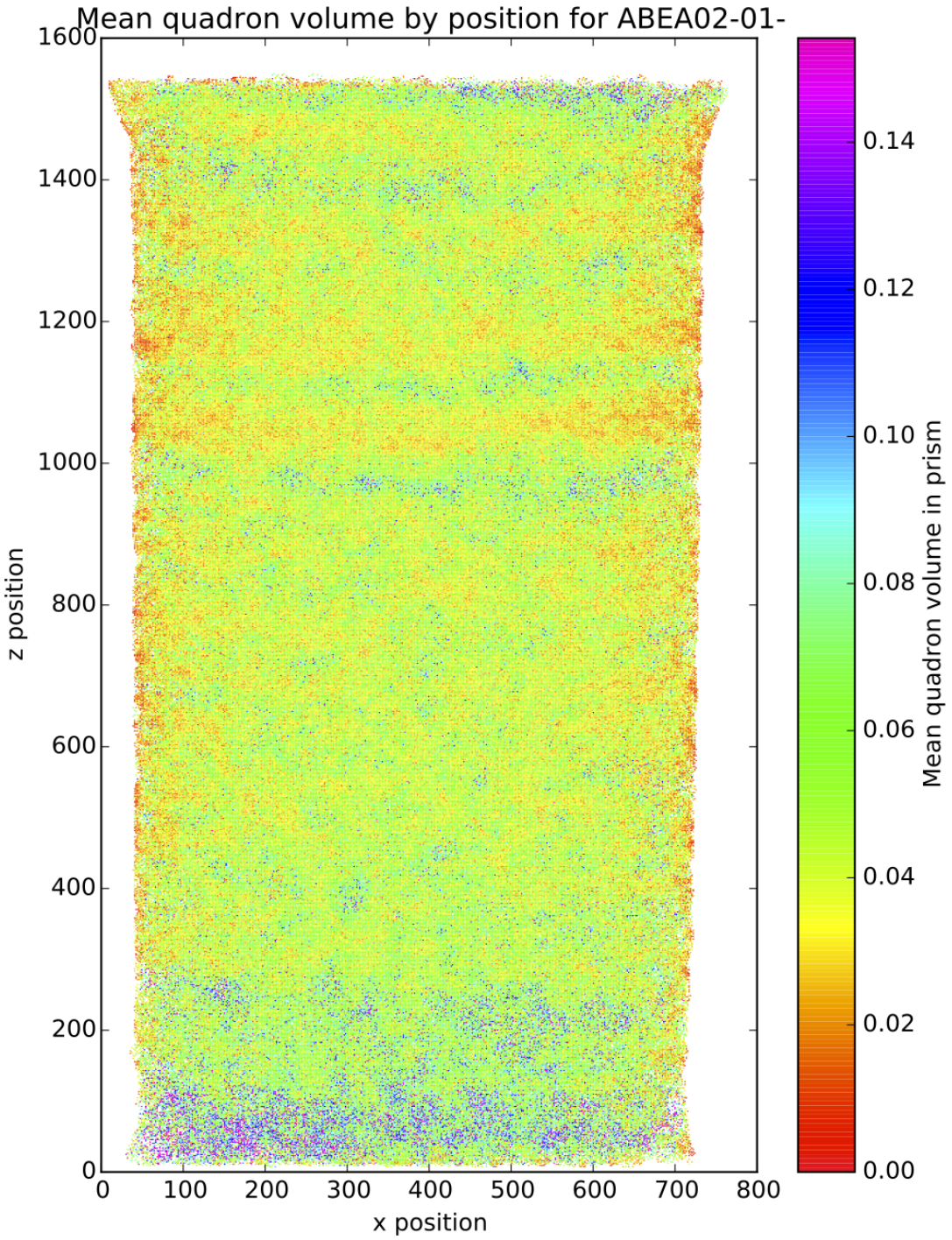


Figure C.4 – Heatmap of quadron volumes projected onto the xz plane, for system ABEA02-01-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

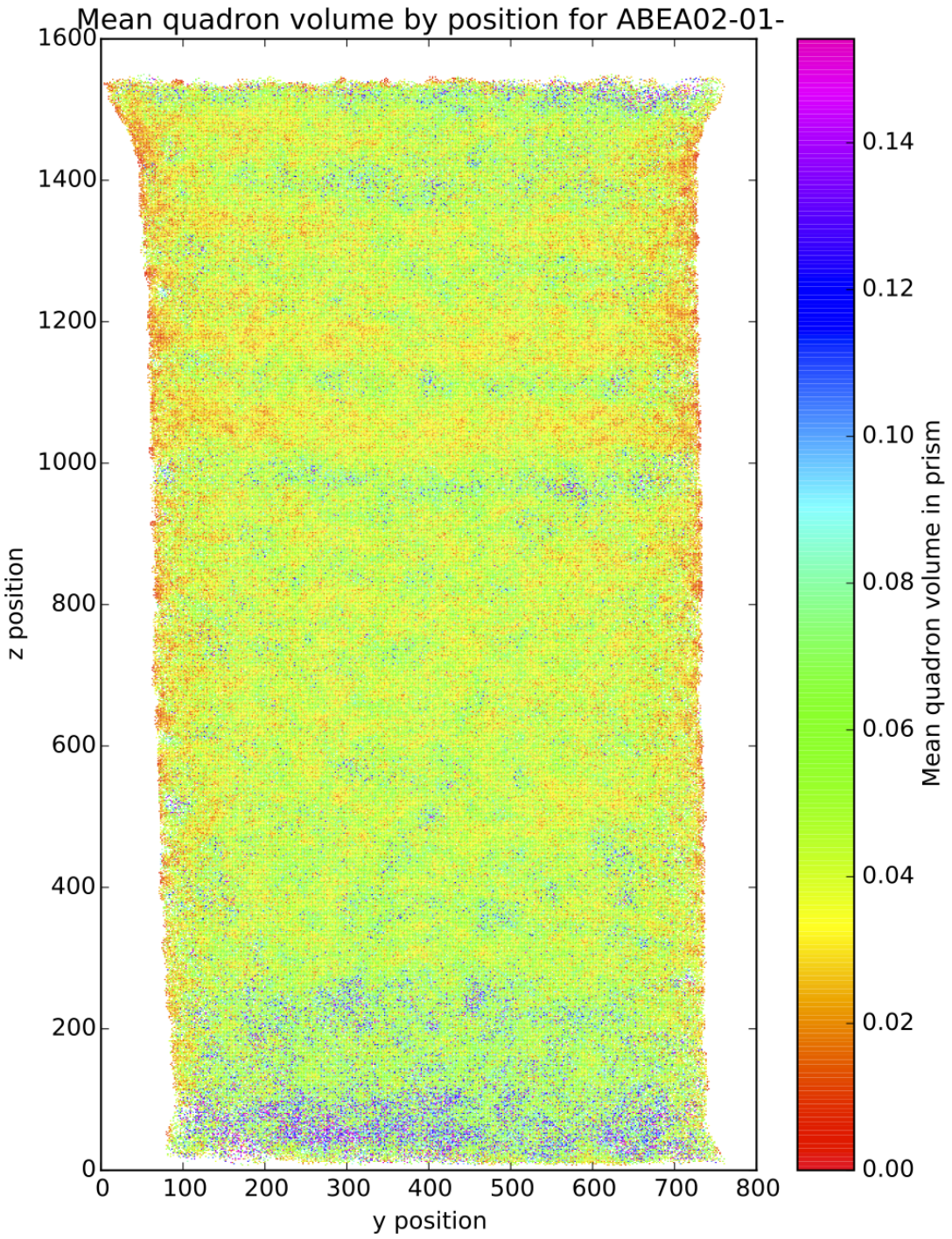


Figure C.5 – Heatmap of quadron volumes projected onto the yz plane, for system ABEA02-01-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

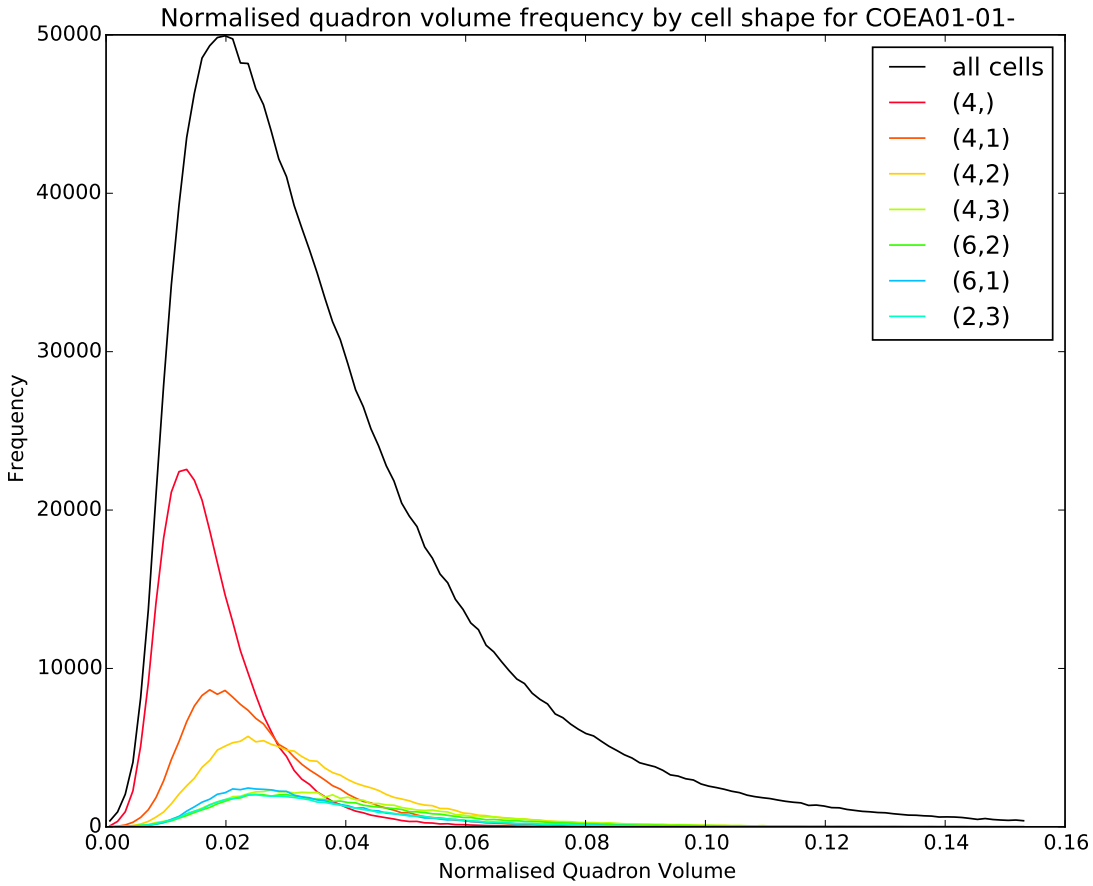


Figure C.6 – Total quadron volume frequency distribution $f\left(v^q\right)$, plotted with the seven most frequent cells' conditional frequency distributions $f\left(v^q|t\right)$, for system COEA01-01-.

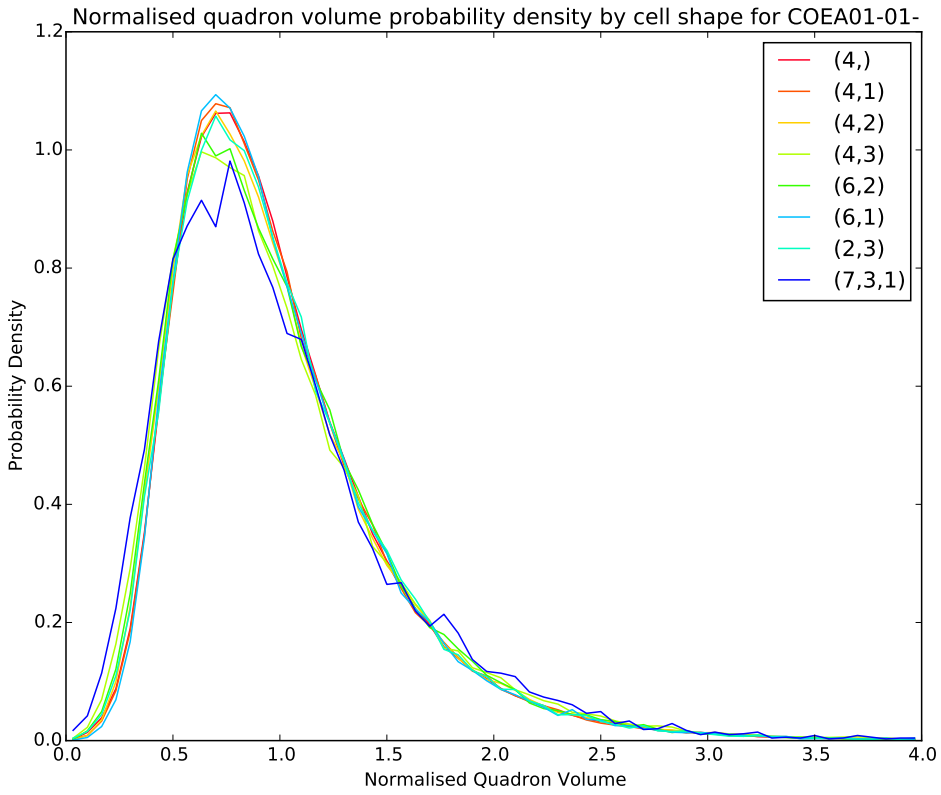


Figure C.7 – Seven most frequent cells' conditional probability distributions PDF $(v^q|t)$, for system COEA01-01-.

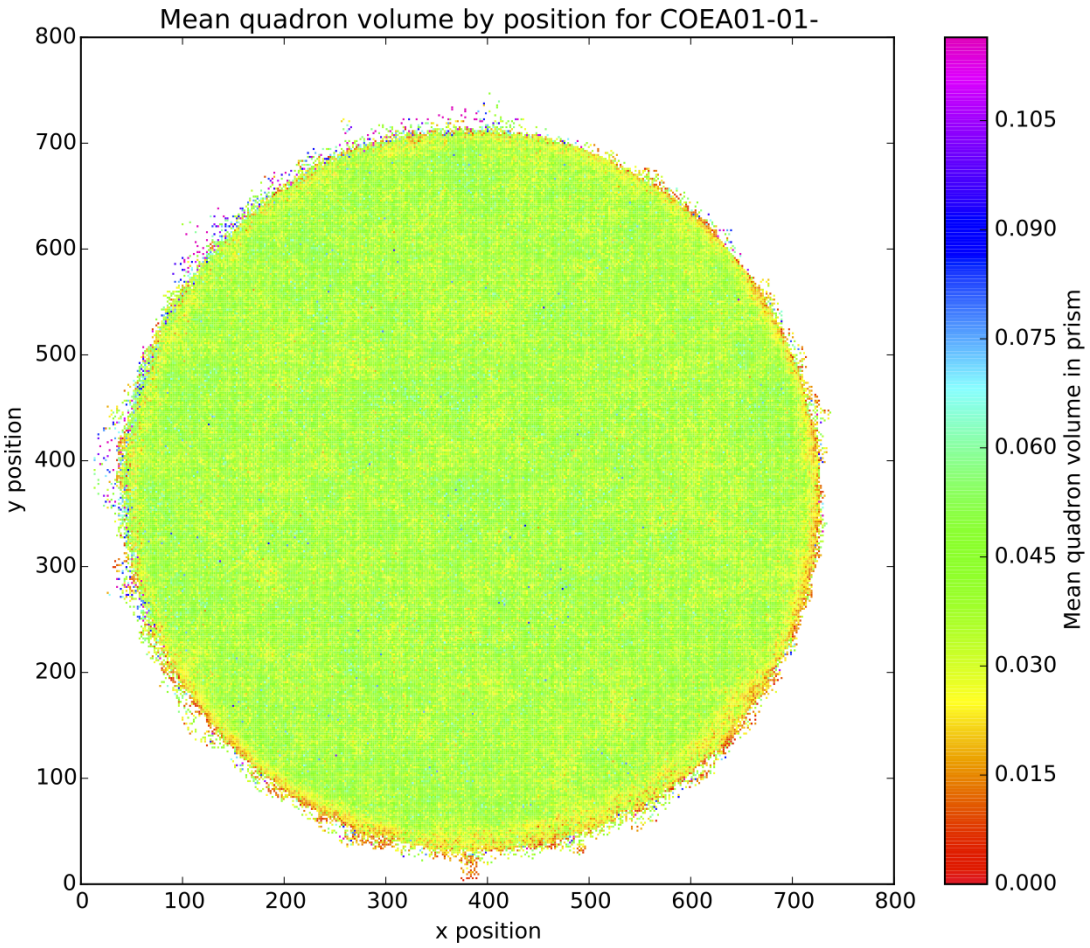


Figure C.8 – Heatmap of quadron volumes projected onto the xy plane, for system COEA01-01-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

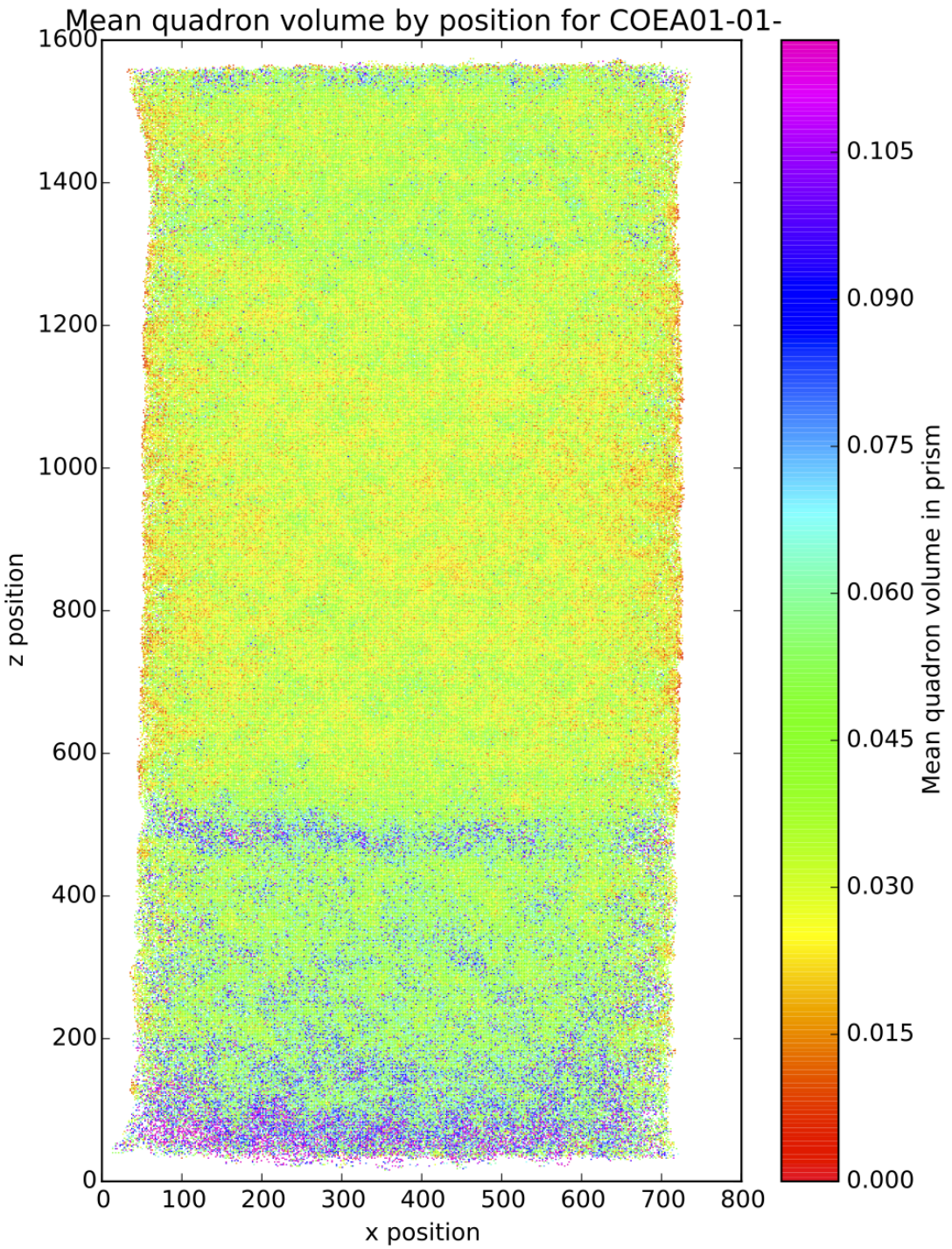


Figure C.9 – Heatmap of quadron volumes projected onto the xz plane, for system COEA01-01-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

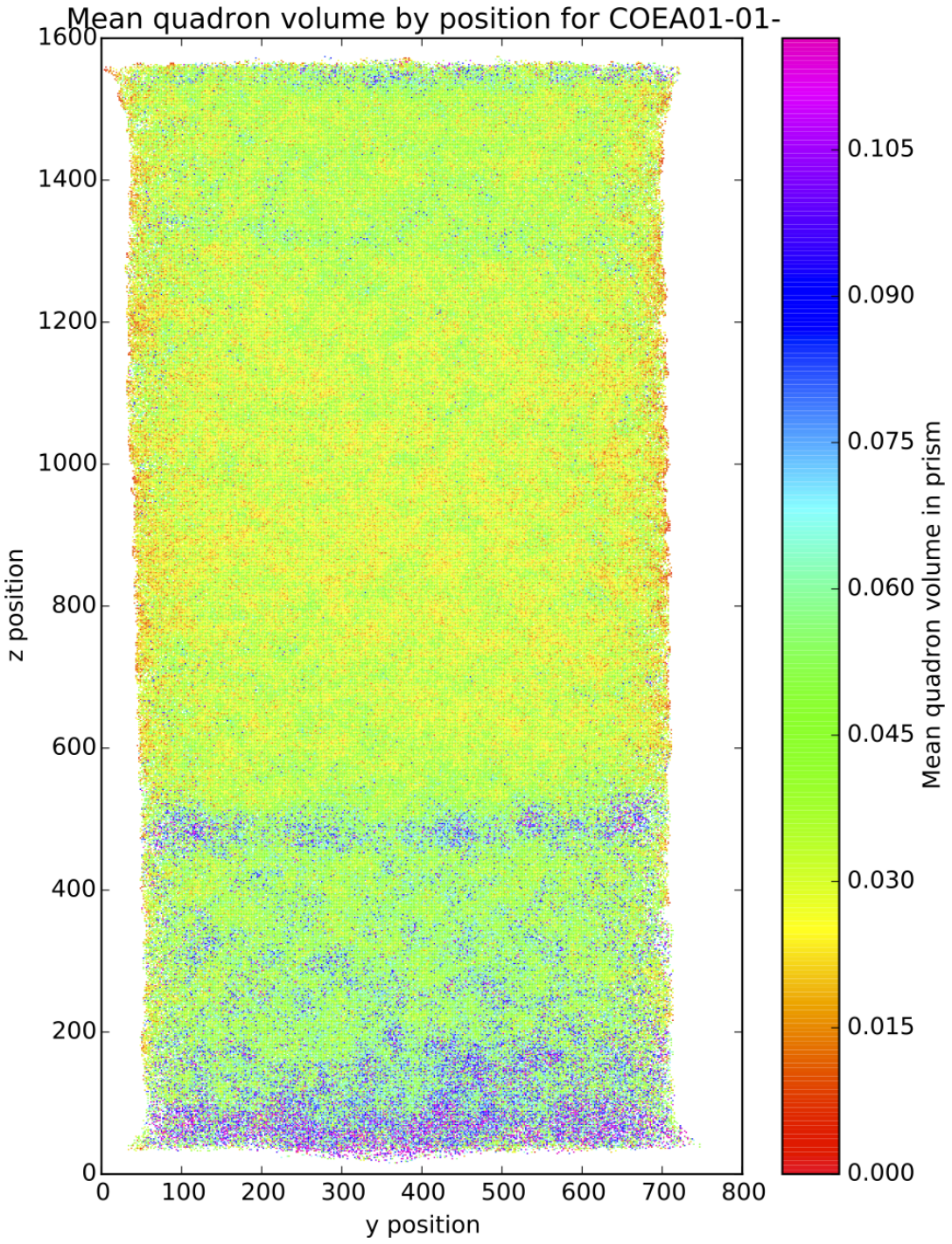


Figure C.10 – Heatmap of quadron volumes projected onto the yz plane, for system COEA01-01-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

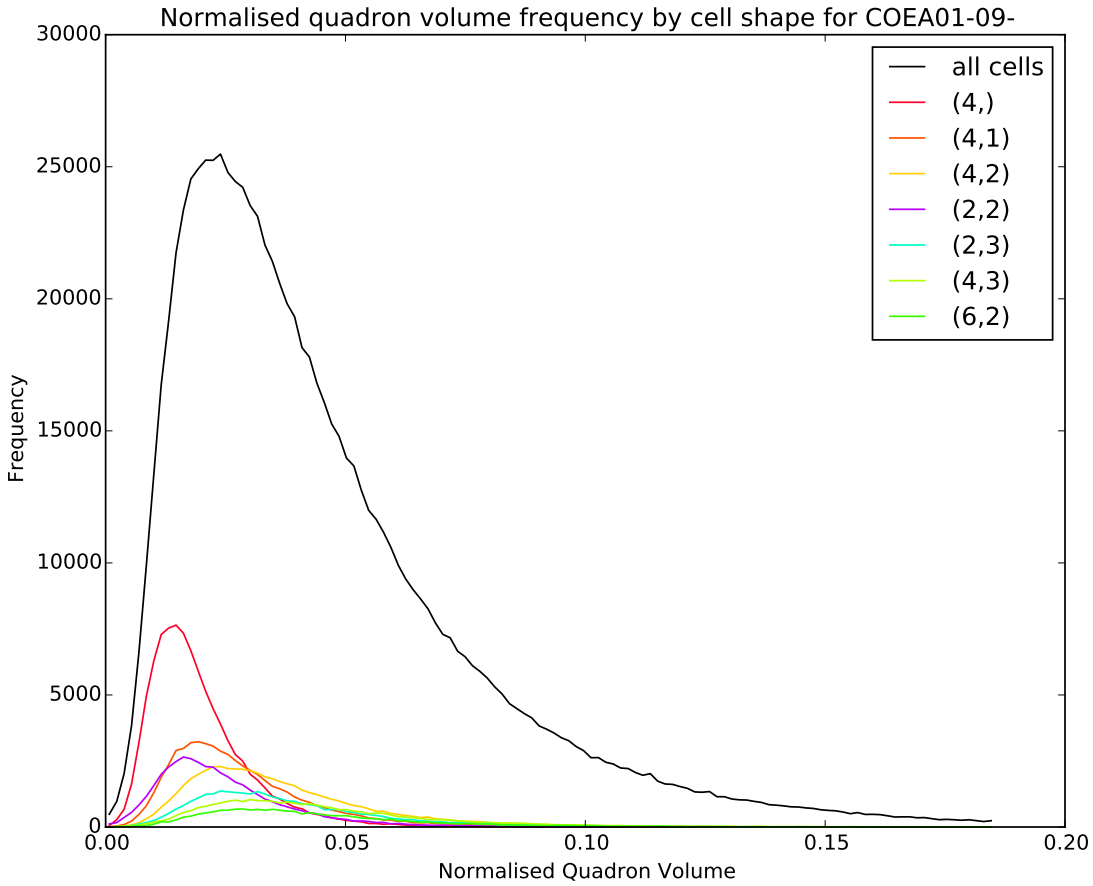


Figure C.11 – Total quadron volume frequency distribution $f\left(v^q\right)$, plotted with the seven most frequent cells' conditional frequency distributions $f\left(v^q|t\right)$, for system COEA01-09-.

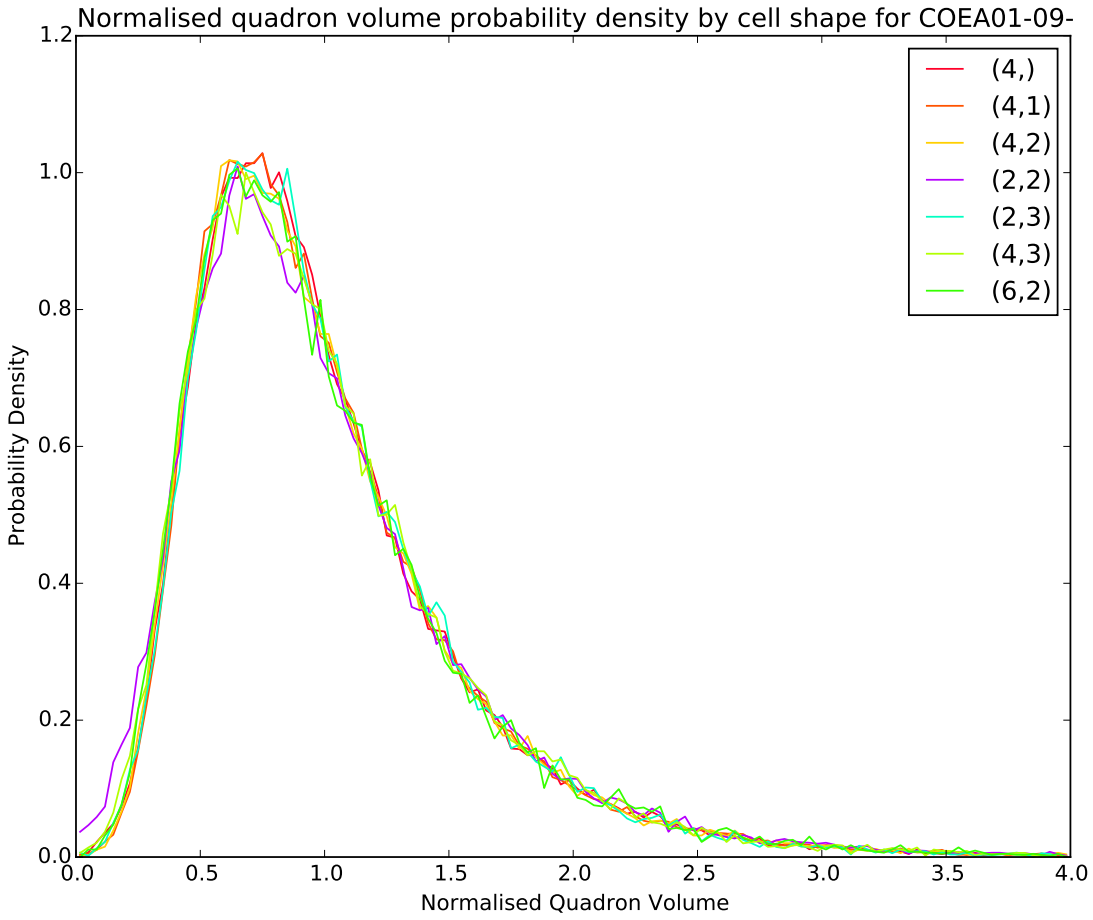


Figure C.12 – Seven most frequent cells' conditional probability distributions PDF $(v^q|t)$, for system COEA01-09-.

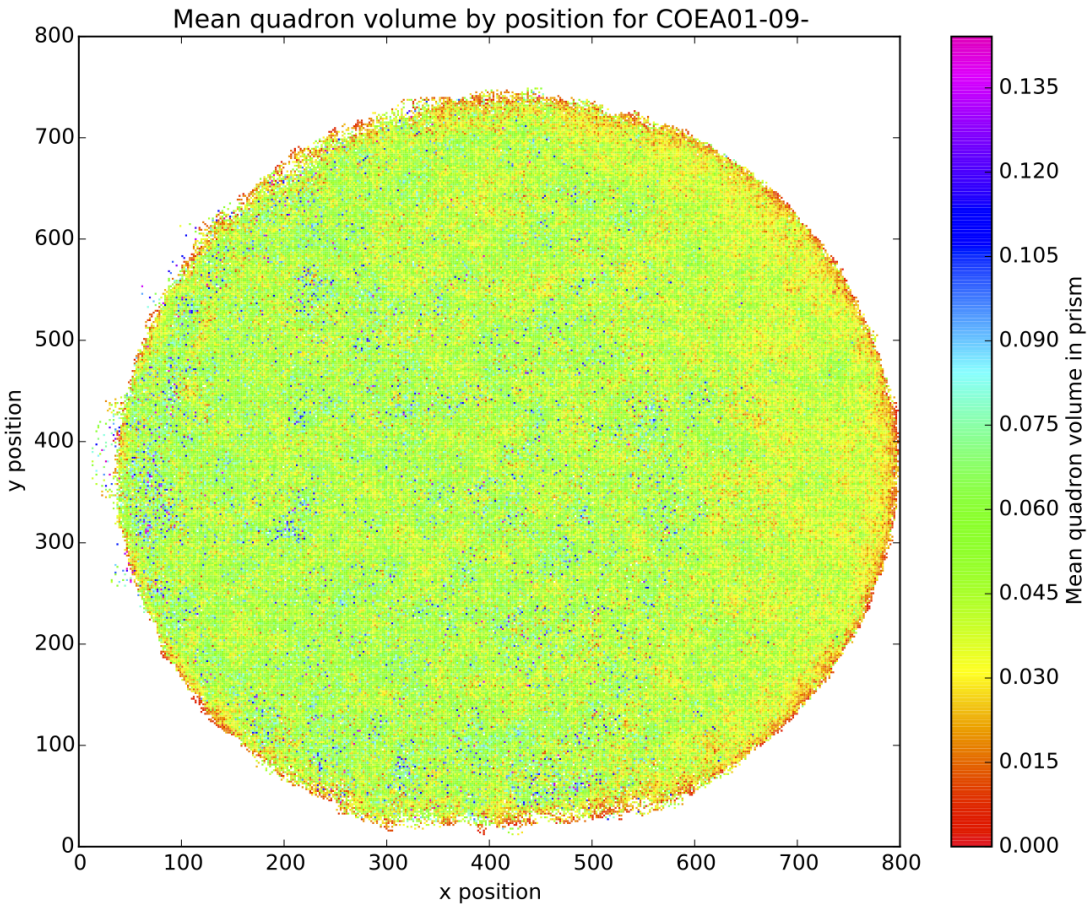


Figure C.13 – Heatmap of quadron volumes projected onto the xy plane, for system COEA01-09-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

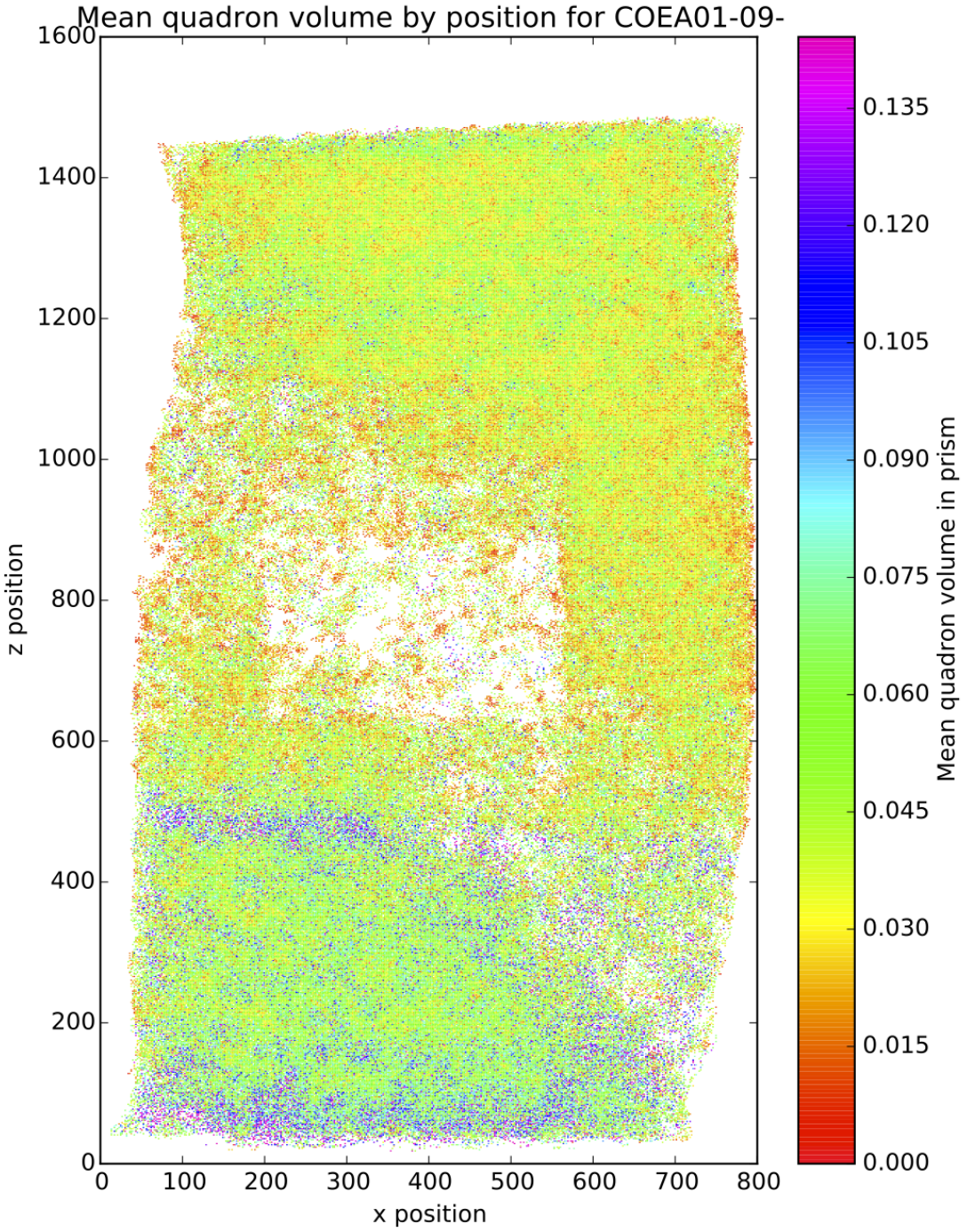


Figure C.14 – Heatmap of quadron volumes projected onto the xz plane, for system COEA01-09-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

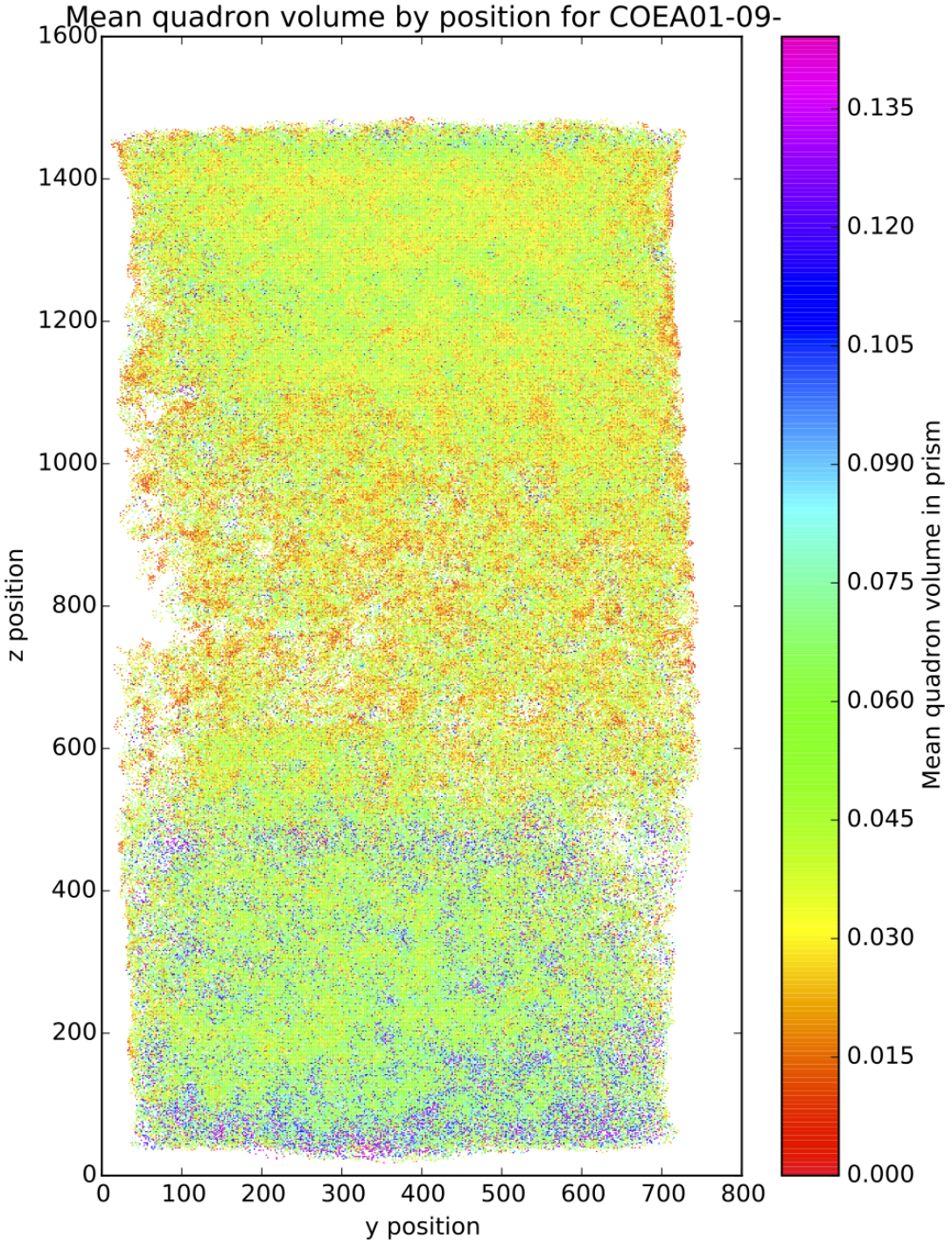


Figure C.15 – Heatmap of quadron volumes projected onto the yz plane, for system COEA01-09-. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

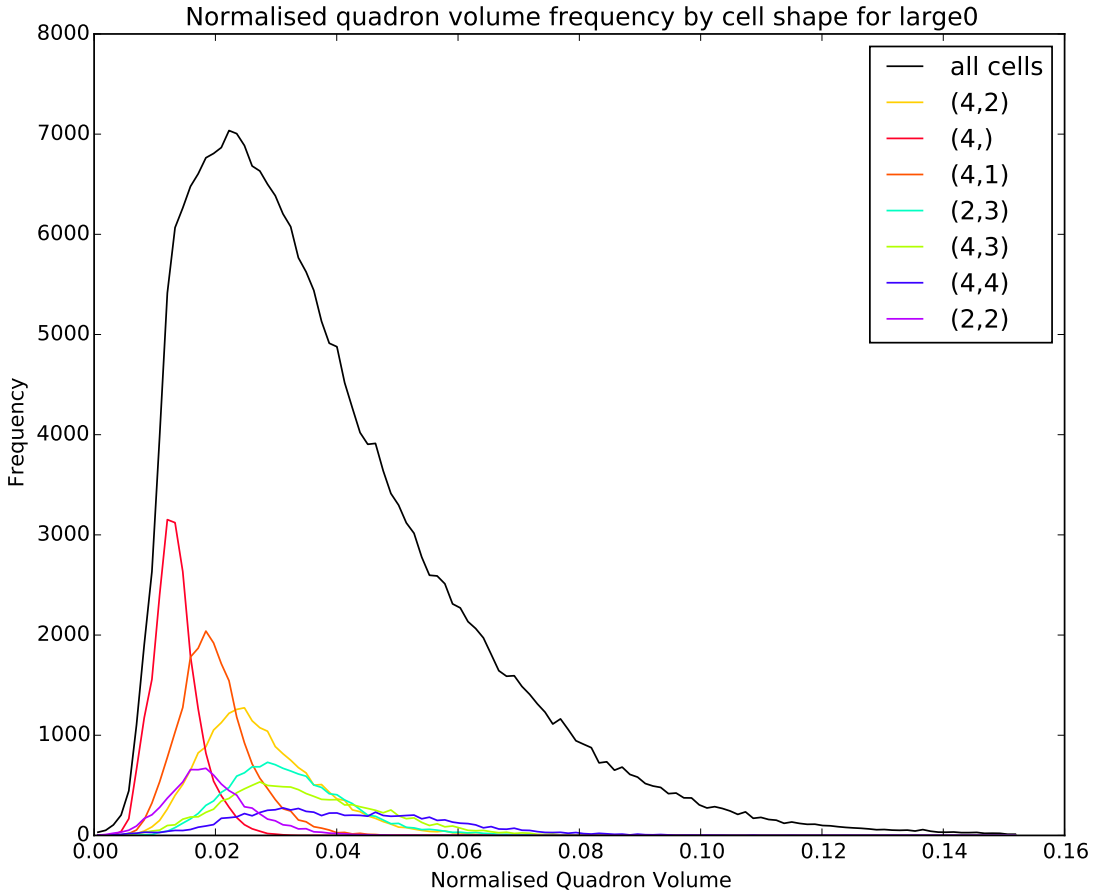


Figure C.16 – Total quadron volume frequency distribution $f\left(v^q\right)$, plotted with the seven most frequent cells' conditional frequency distributions $f\left(v^q|t\right)$, for system large0.

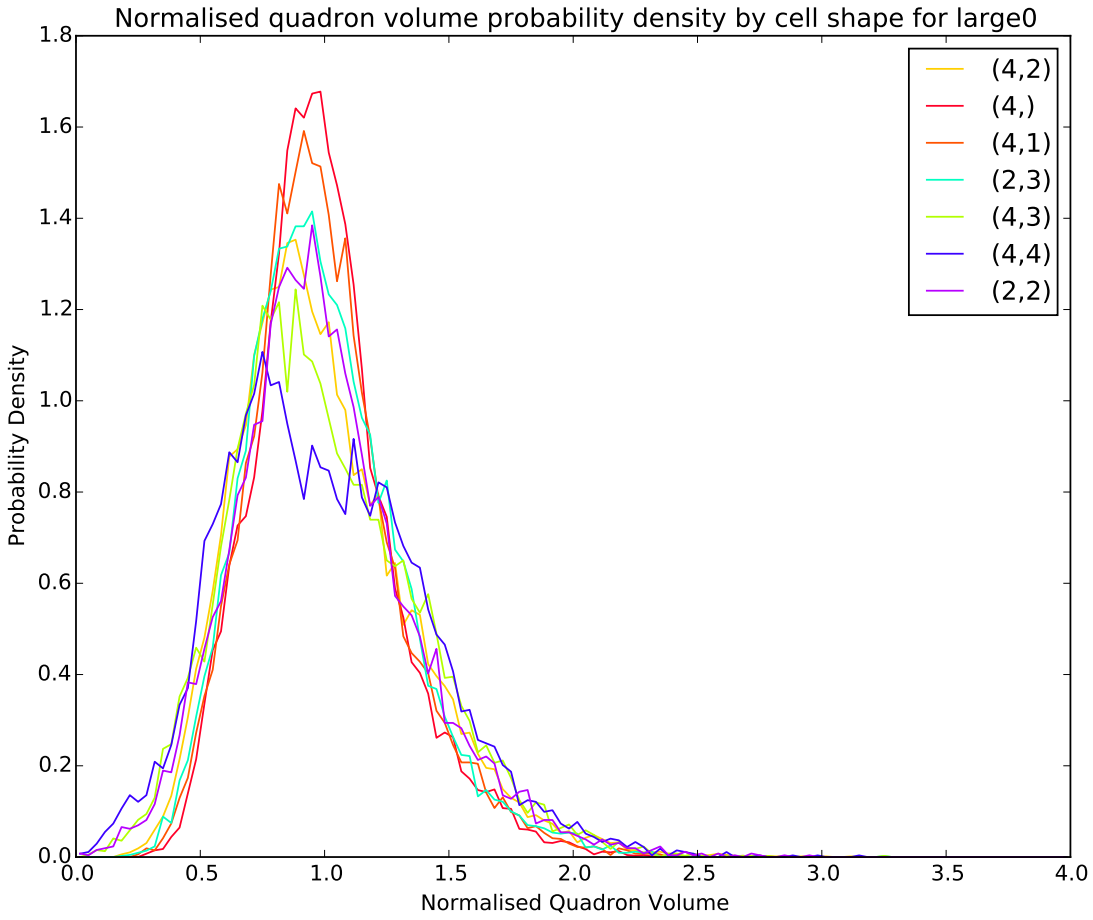


Figure C.17 – Seven most frequent cells' conditional probability distributions PDF $(v^q|t)$, for system large0.

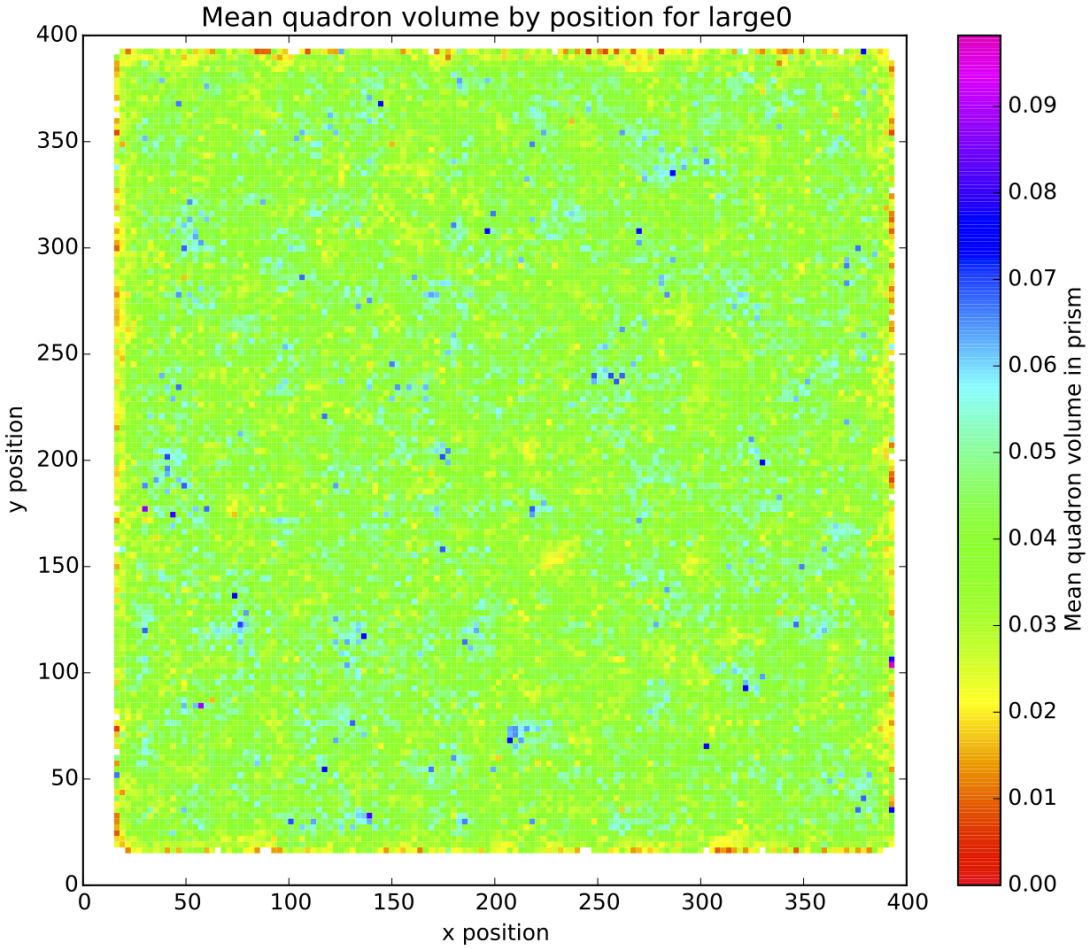


Figure C.18 – Heatmap of quadron volumes projected onto the xy plane, for system large0. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

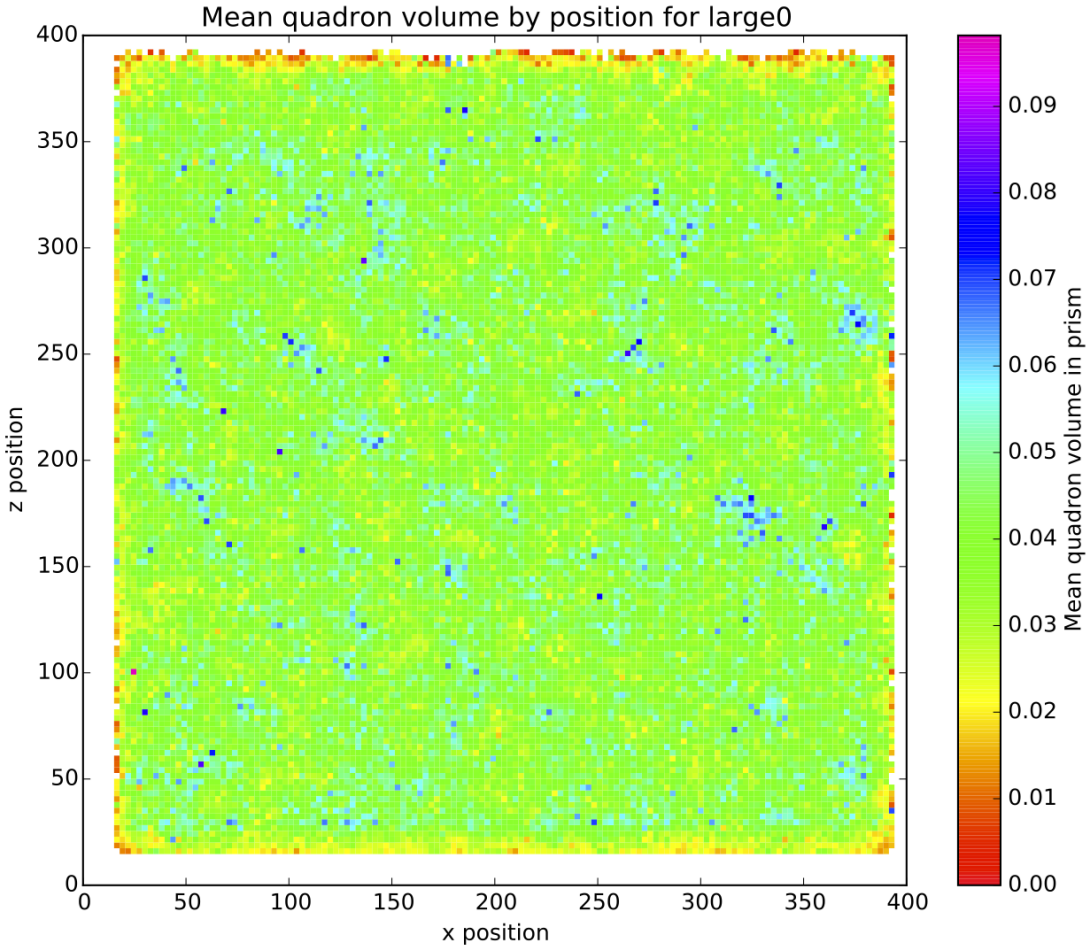


Figure C.19 – Heatmap of quadron volumes projected onto the xz plane, for system large0. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

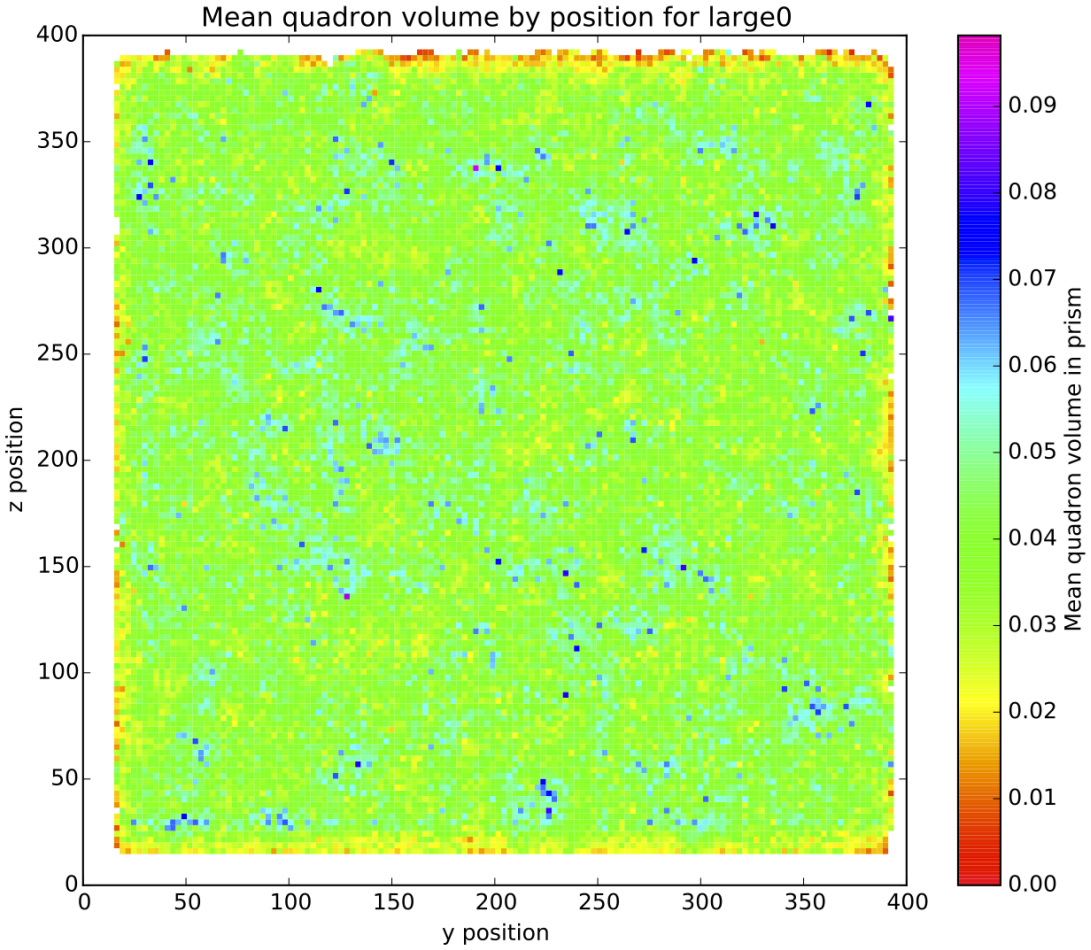


Figure C.20 – Heatmap of quadron volumes projected onto the yz plane, for system large0. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

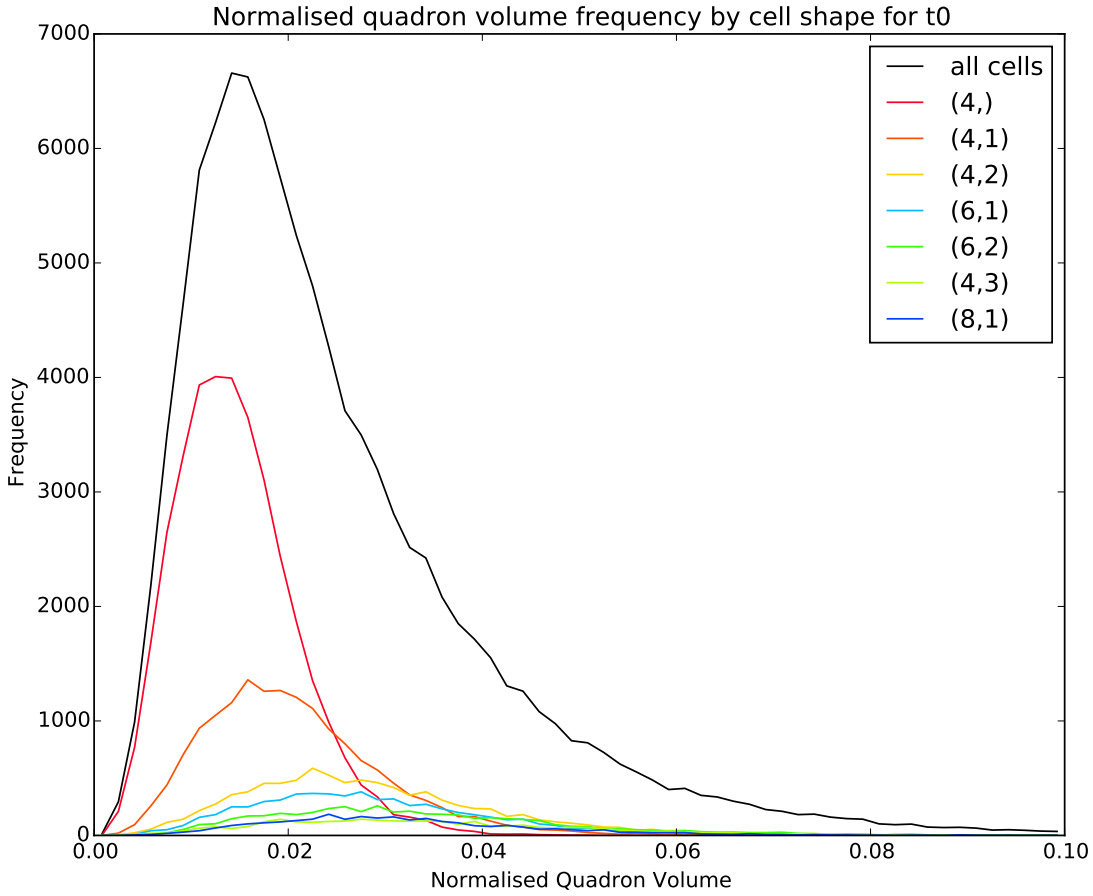


Figure C.21 – Total quadron volume frequency distribution $f\left(v^q\right)$, plotted with the seven most frequent cells' conditional frequency distributions $f\left(v^q|t\right)$, for system t0.

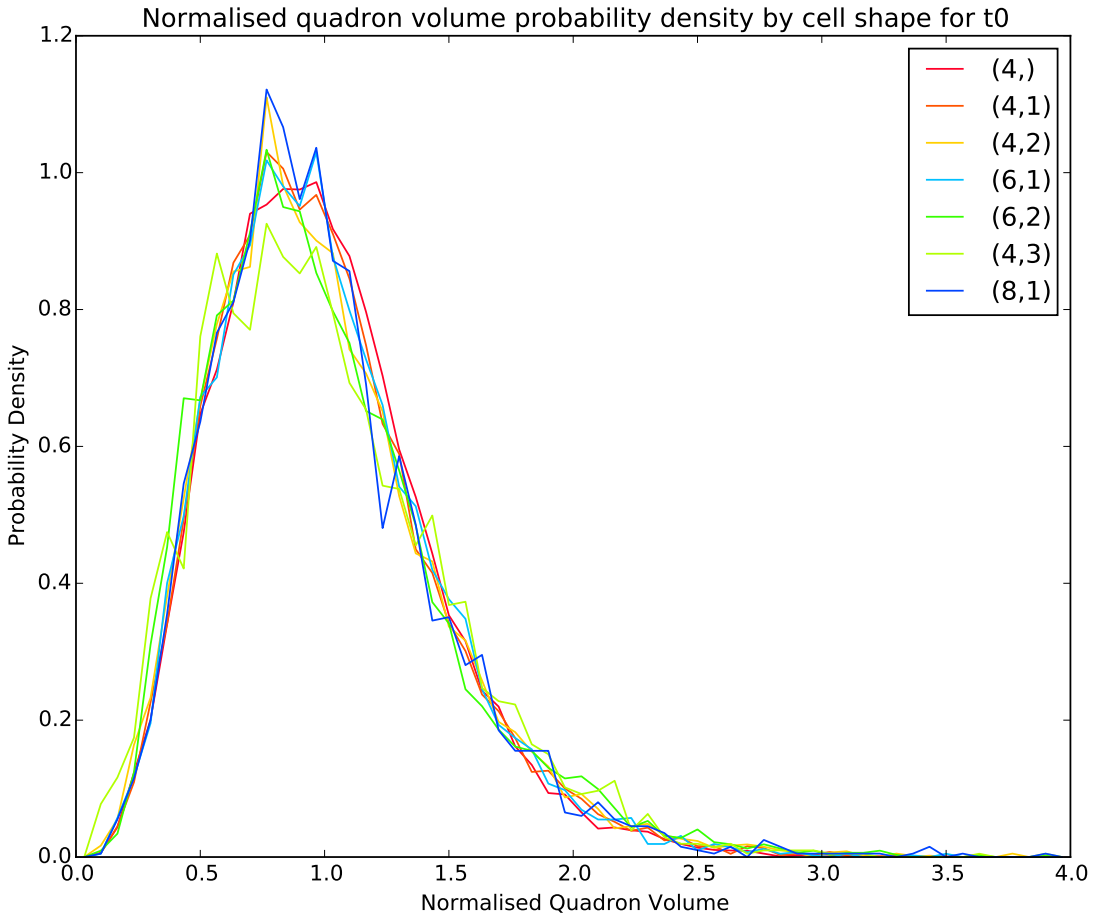


Figure C.22 – Seven most frequent cells' conditional probability distributions PDF $(v^q|t)$, for system t0.

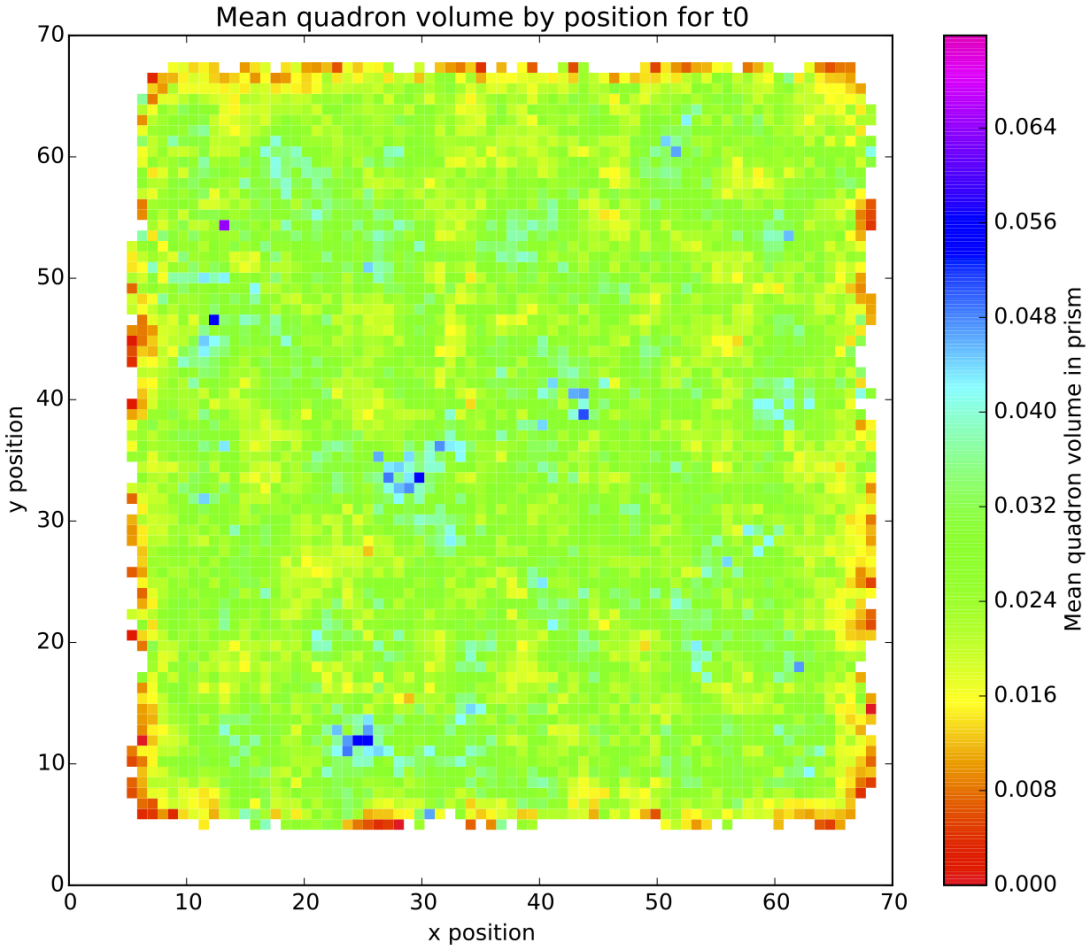


Figure C.23 – Heatmap of quadron volumes projected onto the xy plane, for system t0. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

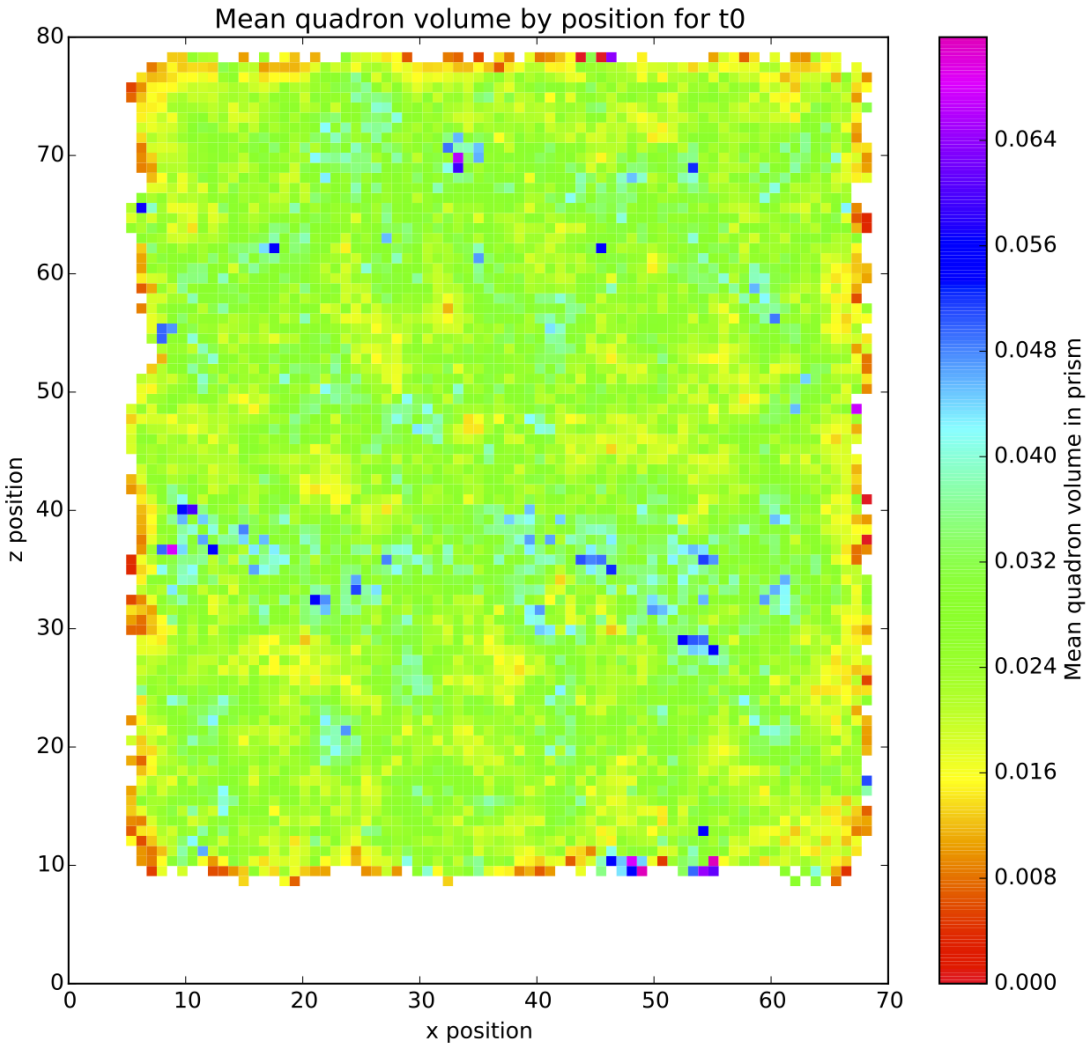


Figure C.24 – Heatmap of quadron volumes projected onto the xz plane, for system t0. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

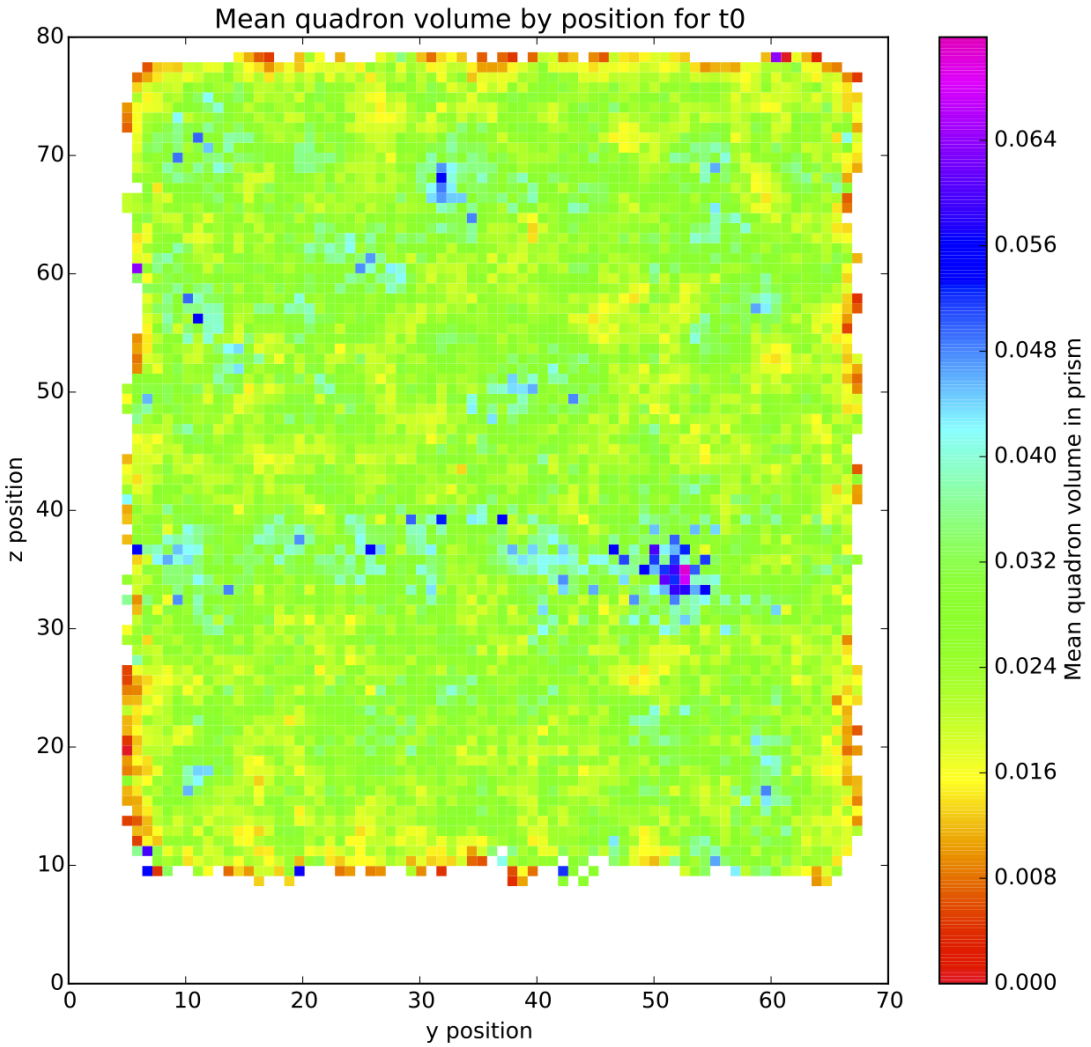


Figure C.25 – Heatmap of quadron volumes projected onto the yz plane, for system t0. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

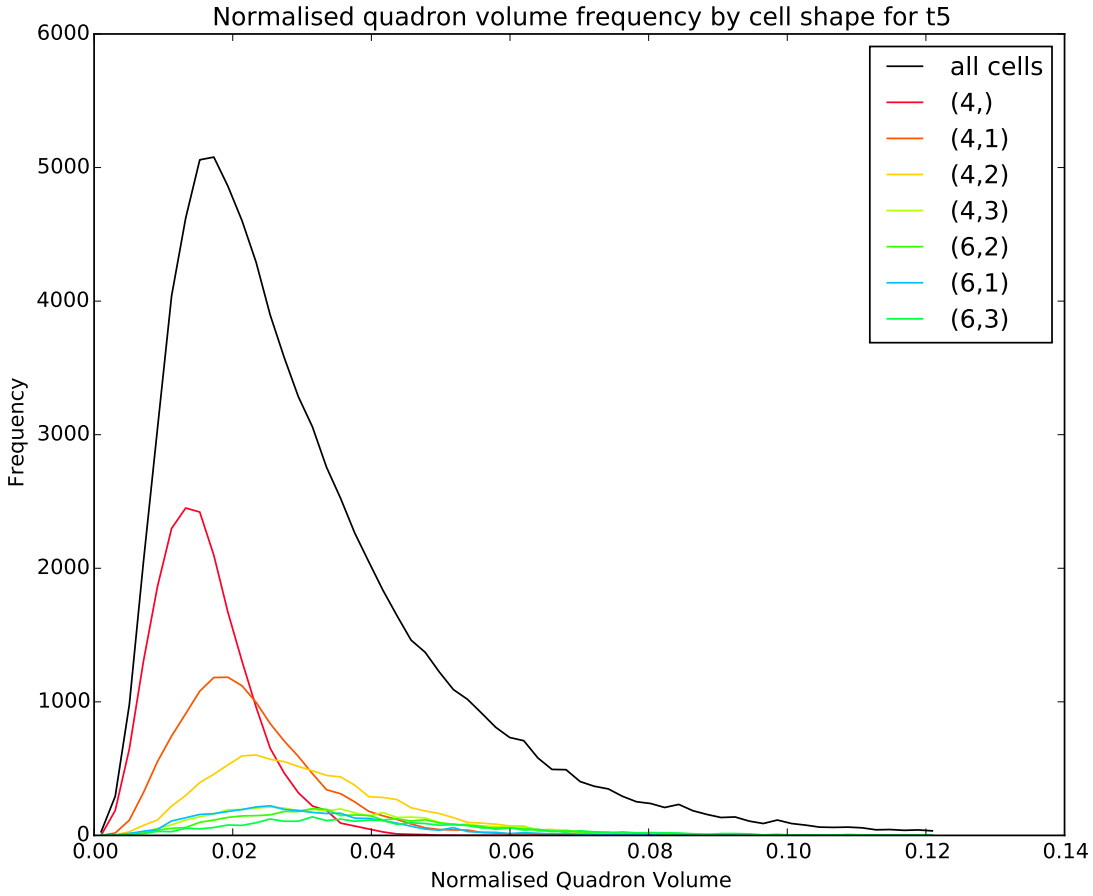


Figure C.26 – Total quadron volume frequency distribution $f\left(v^q\right)$, plotted with the seven most frequent cells' conditional frequency distributions $f\left(v^q|t\right)$, for system t5.

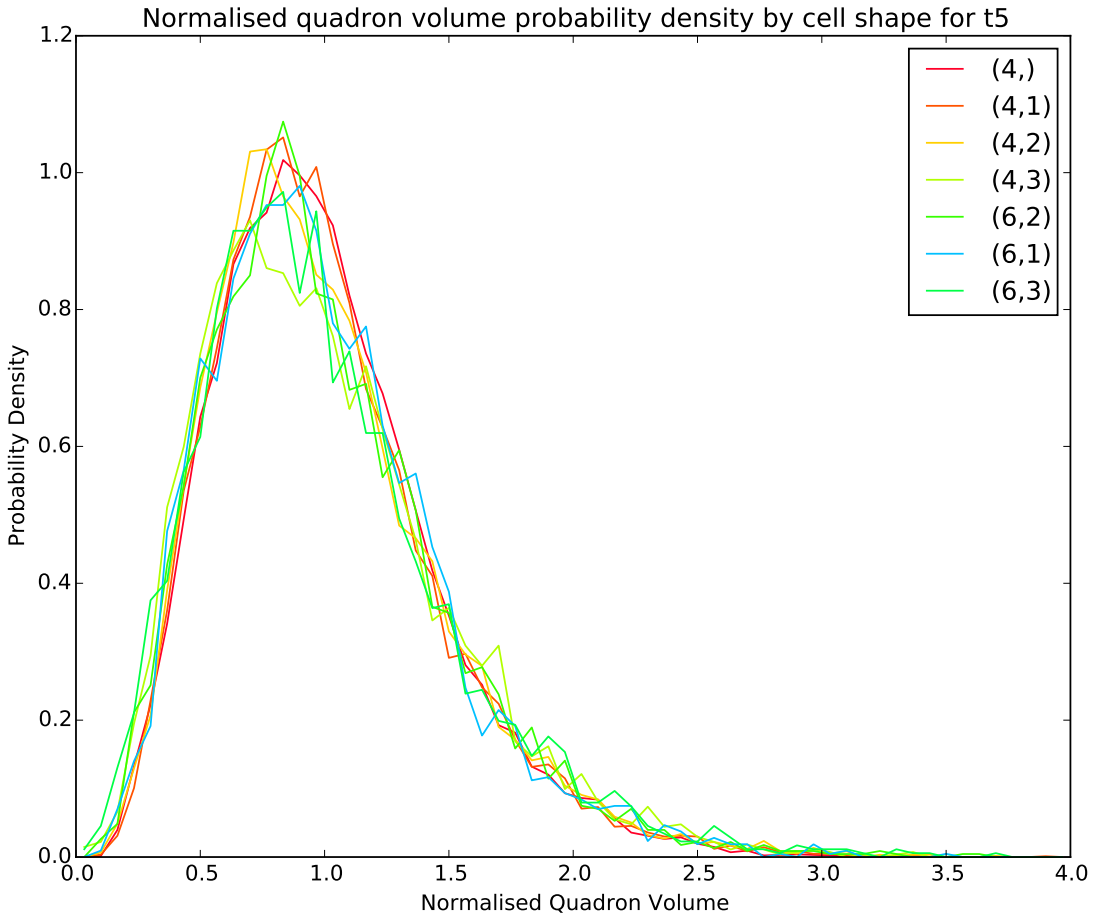


Figure C.27 – Seven most frequent cells' conditional probability distributions PDF $(v^q|t)$, for system t5.

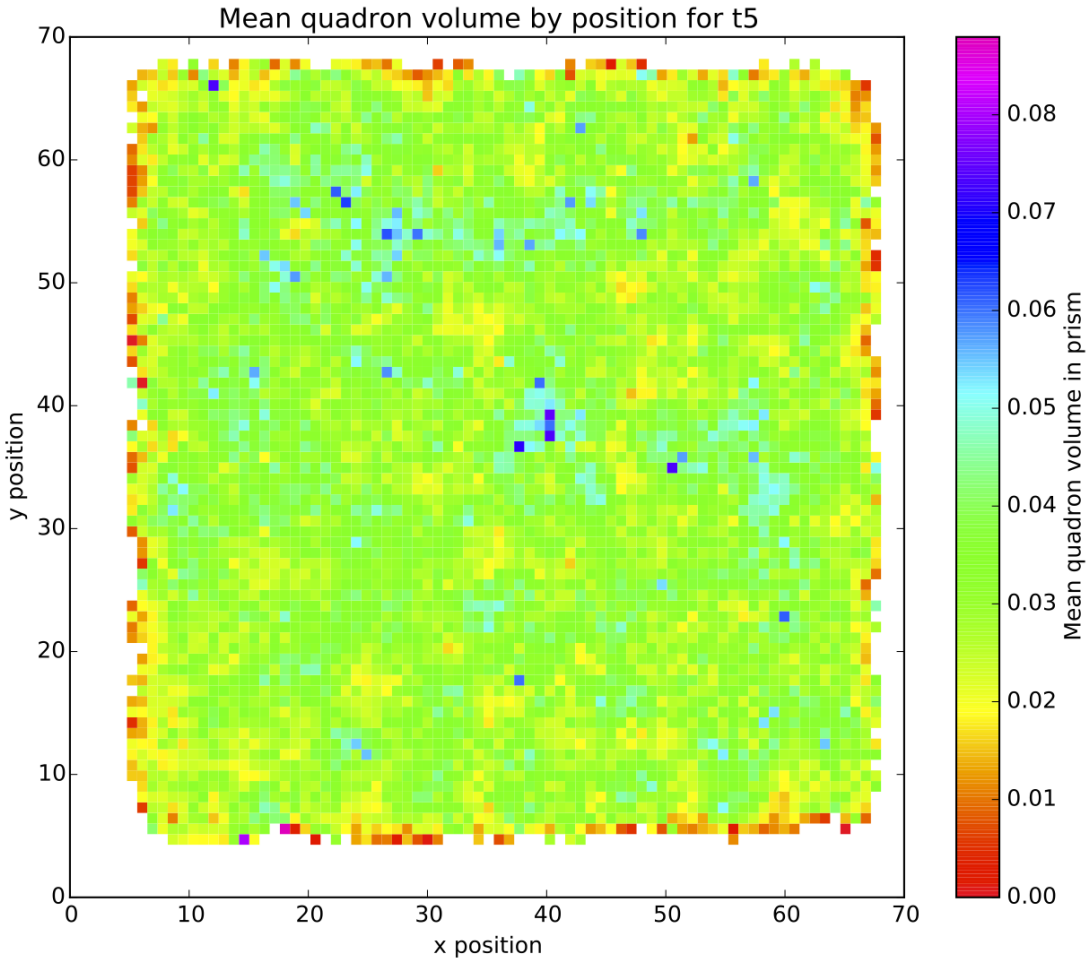


Figure C.28 – Heatmap of quadron volumes projected onto the xy plane, for system t5. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

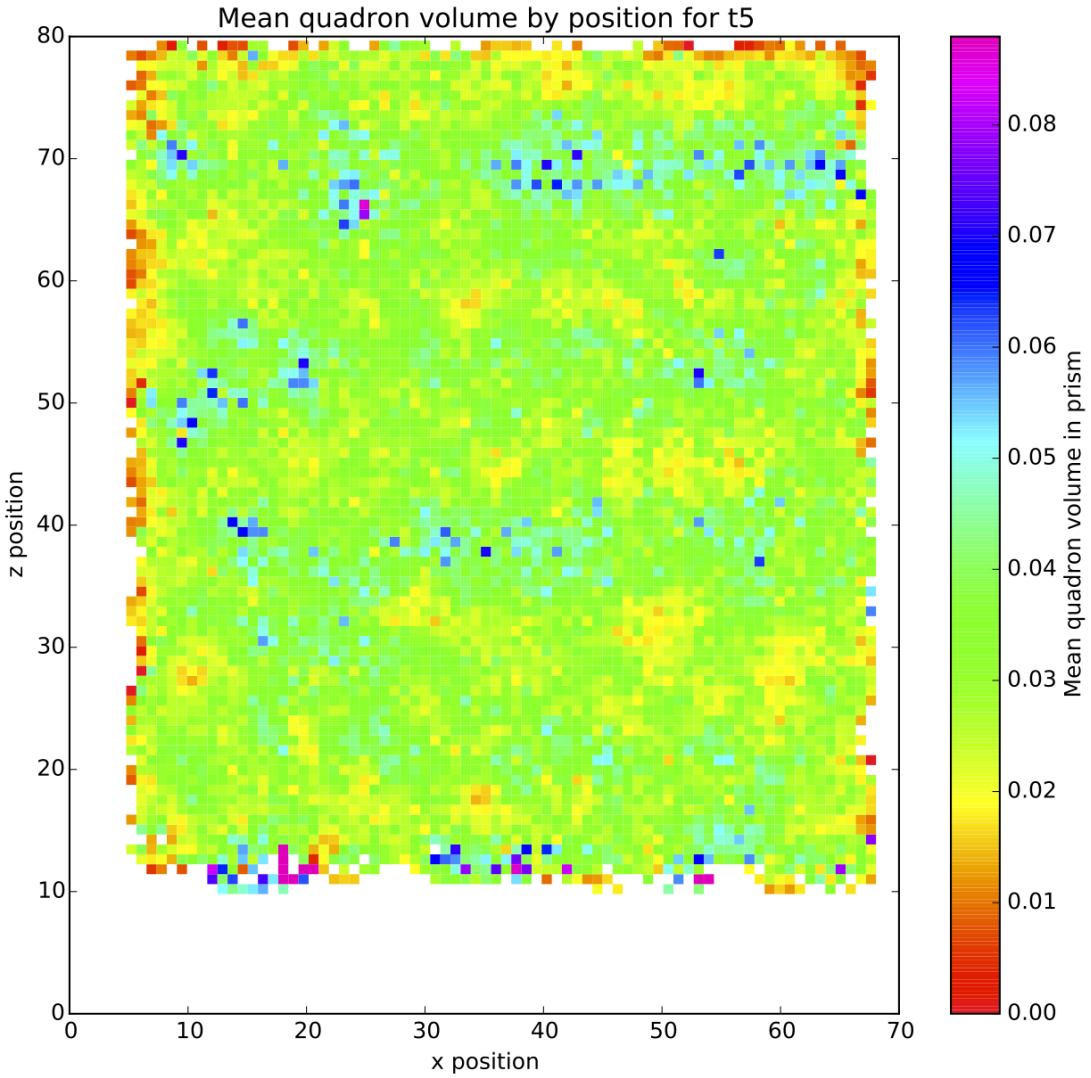


Figure C.29 – Heatmap of quadron volumes projected onto the xz plane, for system t5. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

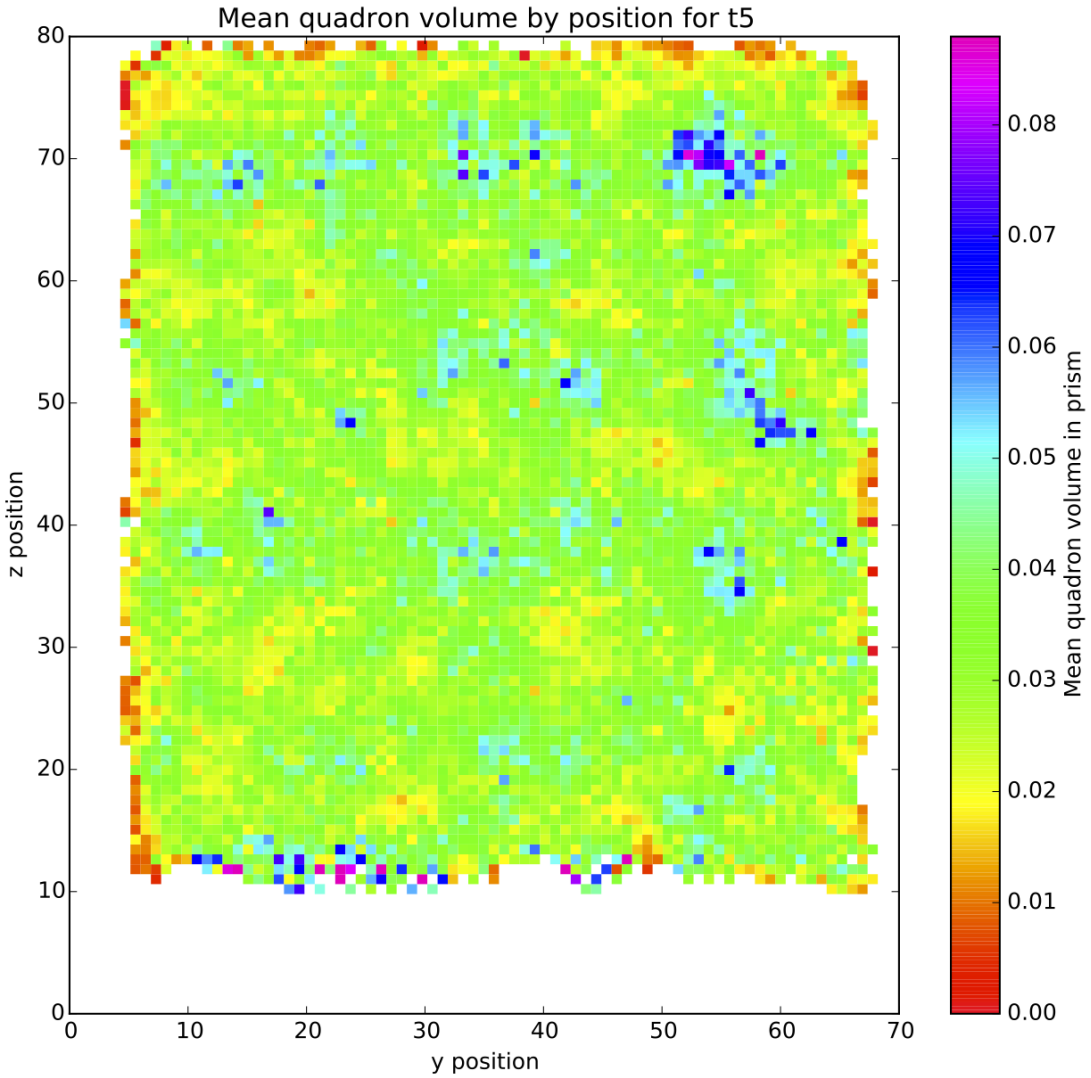


Figure C.30 – Heatmap of quadron volumes projected onto the yz plane, for system t5. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

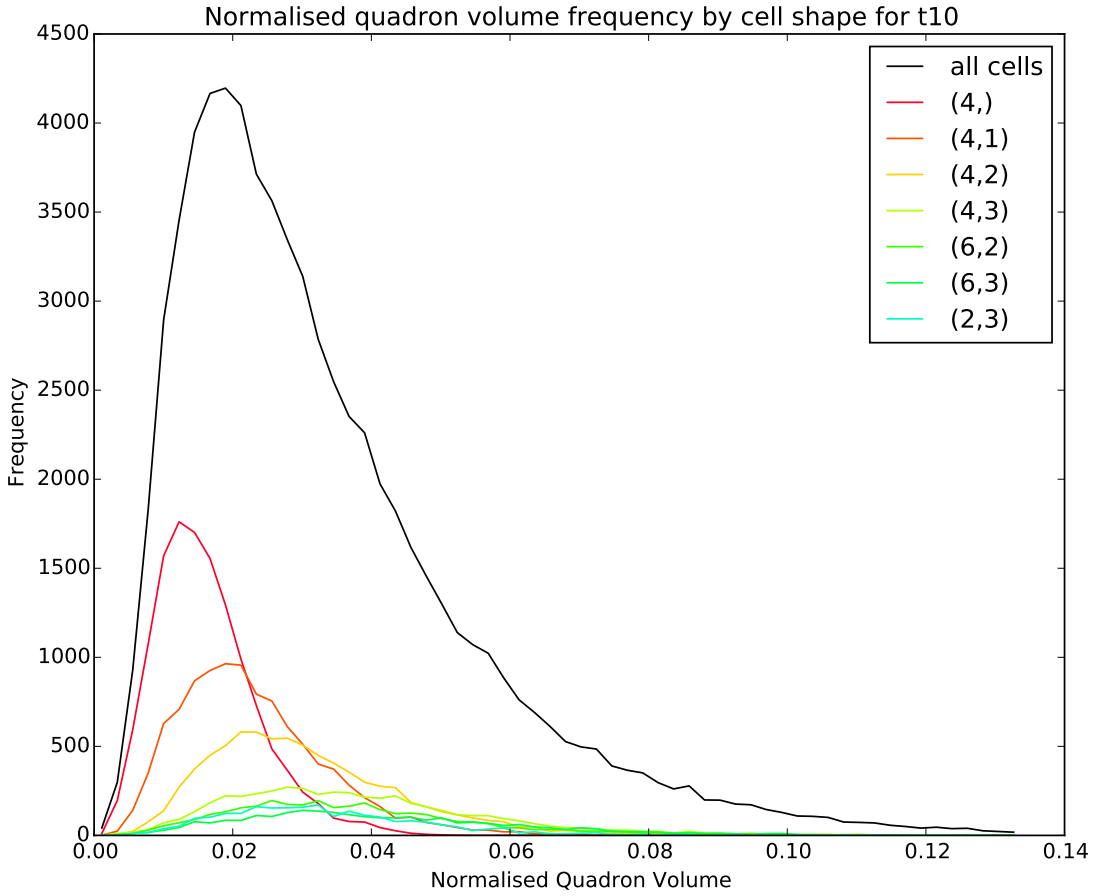


Figure C.31 – Total quadron volume frequency distribution $f\left(v^q\right)$, plotted with the seven most frequent cells' conditional frequency distributions $f\left(v^q|t\right)$, for system t10.

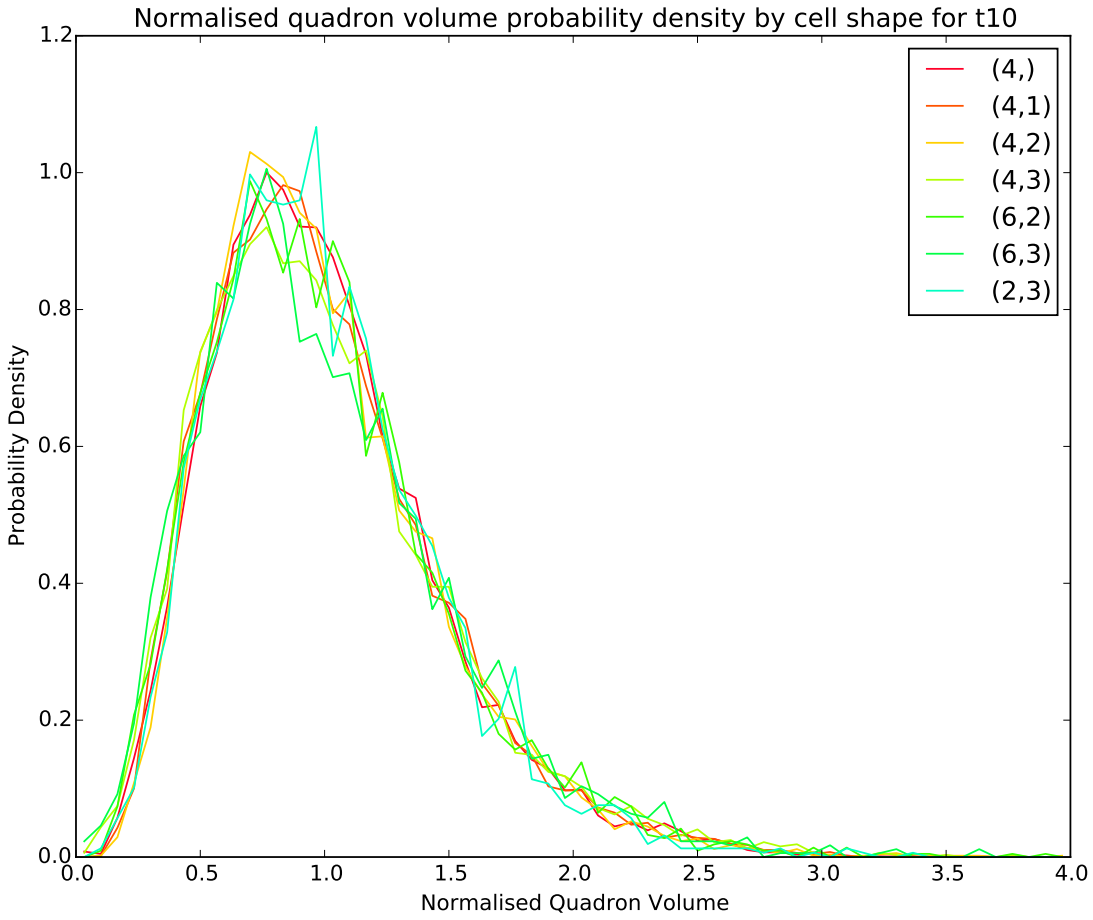


Figure C.32 – Seven most frequent cells' conditional probability distributions PDF $(v^q|t)$, for system t10.

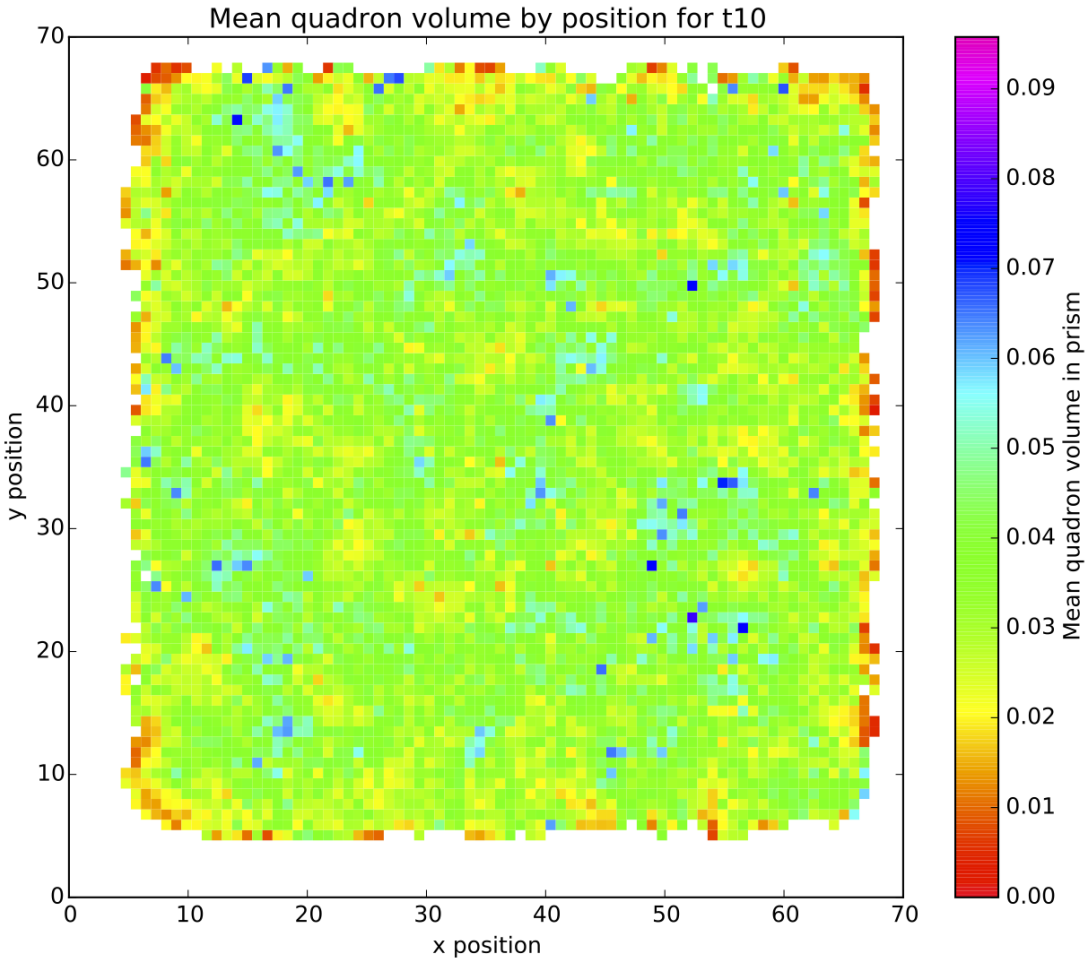


Figure C.33 – Heatmap of quadron volumes projected onto the xy plane, for system t10. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

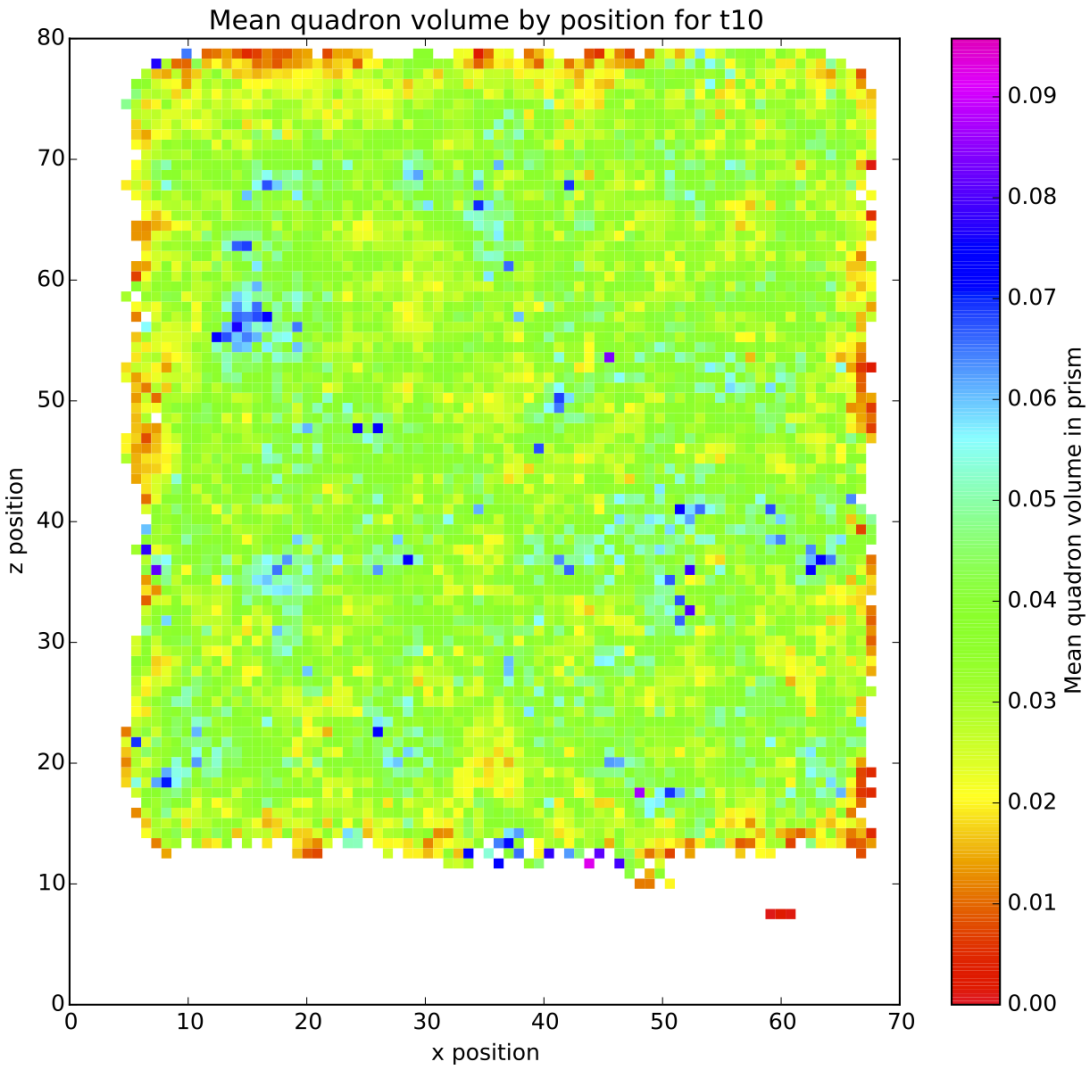


Figure C.34 – Heatmap of quadron volumes projected onto the xz plane, for system t10. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

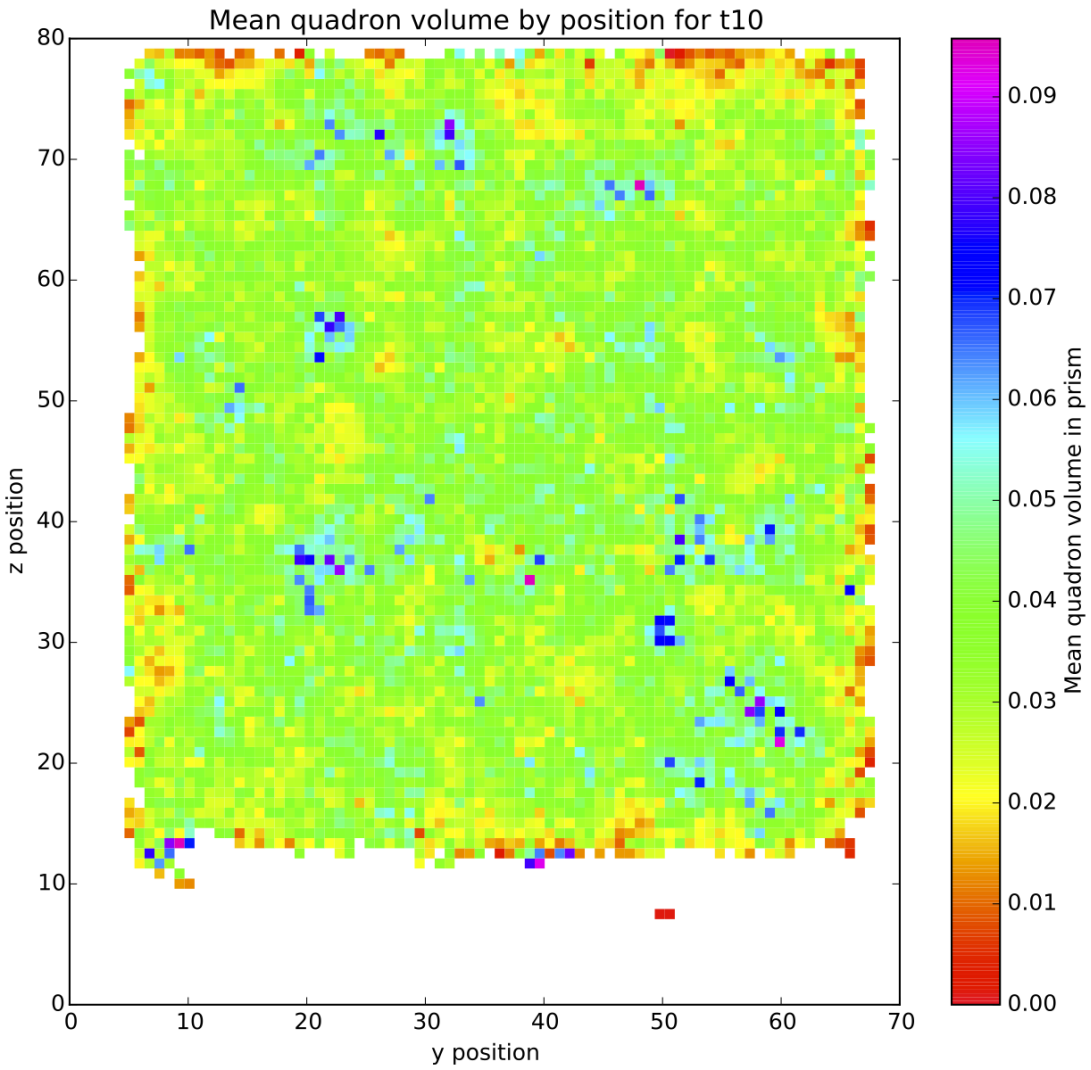


Figure C.35 – Heatmap of quadron volumes projected onto the yz plane, for system t10. Each pixel in the heatmap is coloured by the average volume of quadrons whose positions are projected into the plane within it.

C.2 Distributions of Quadron Chiralities

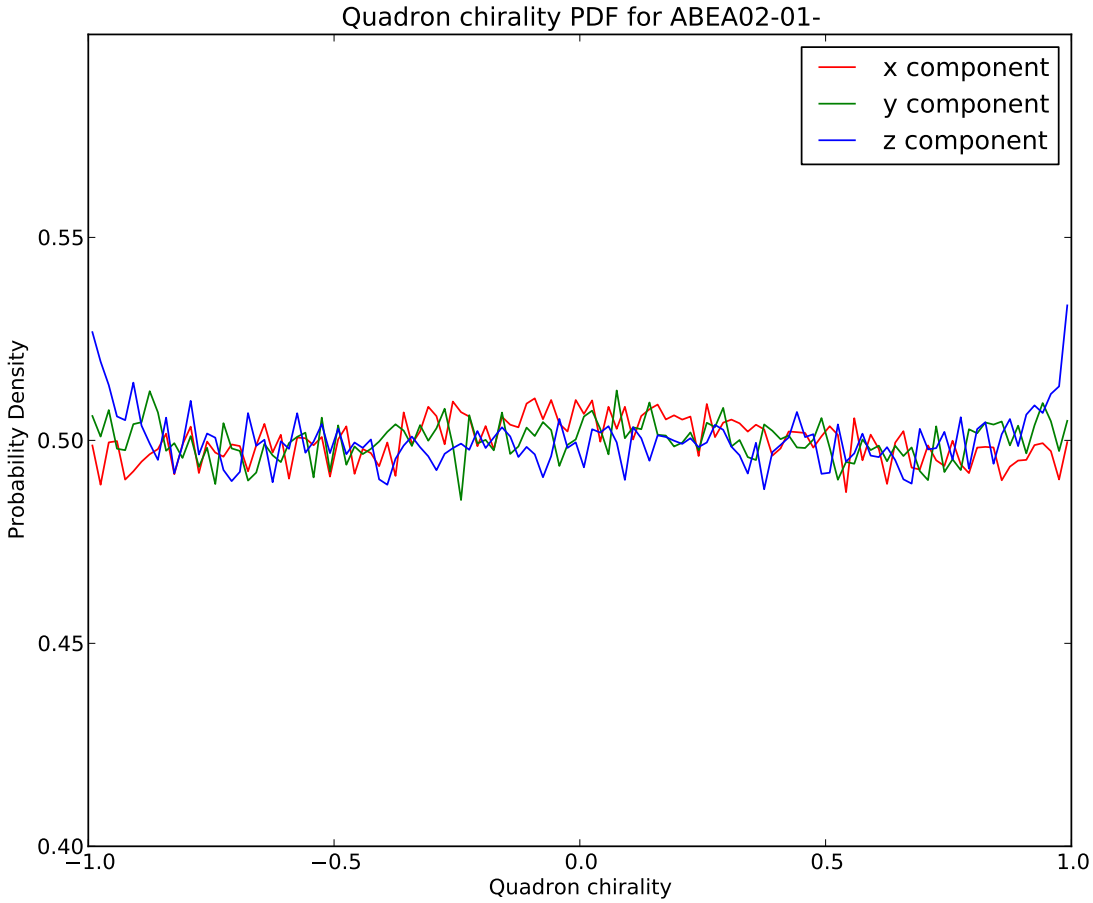


Figure C.36 – Normalised quadron chirality components – where the pseudo-vector has been given length 1 and each component plotted as a probability distribution, for system ABEA02-01-.

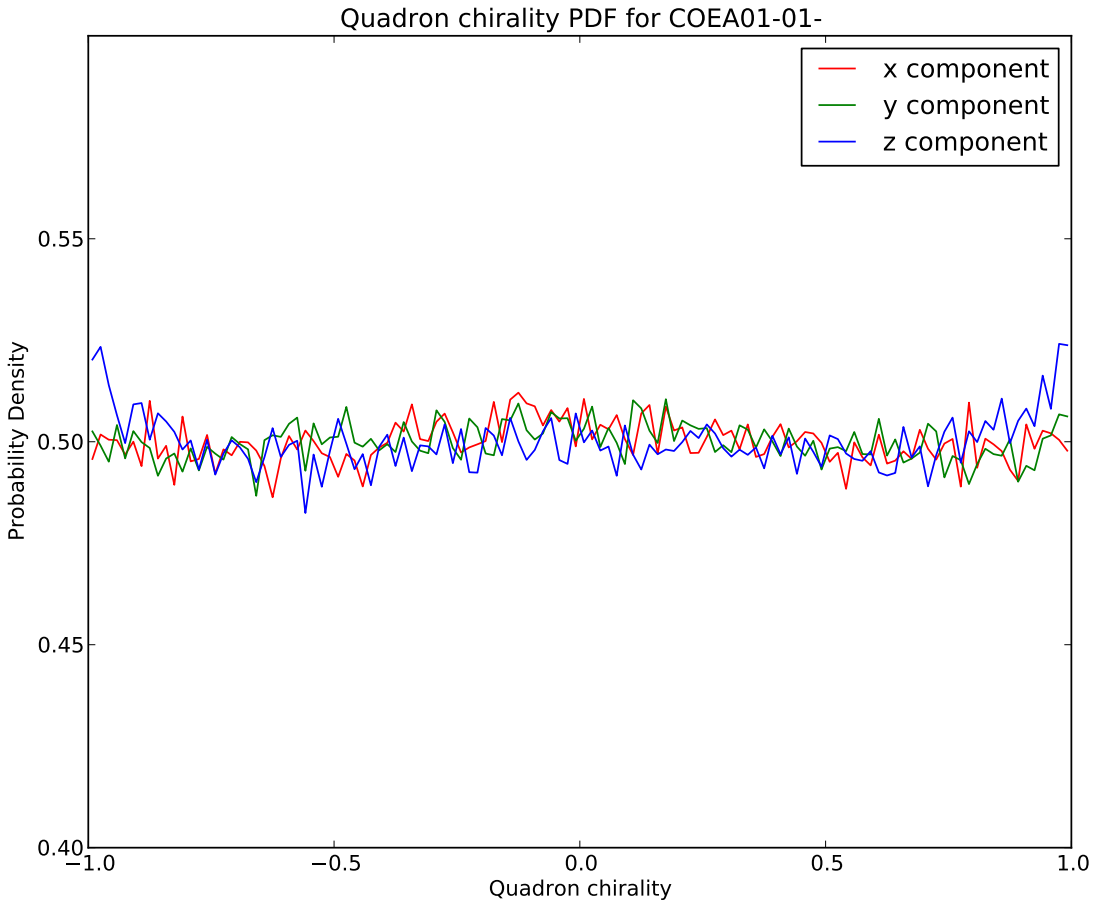


Figure C.37 – Normalised quadron chirality components – where the pseudo-vector has been given length 1 and each component plotted as a probability distribution, for system COEA01-01-.

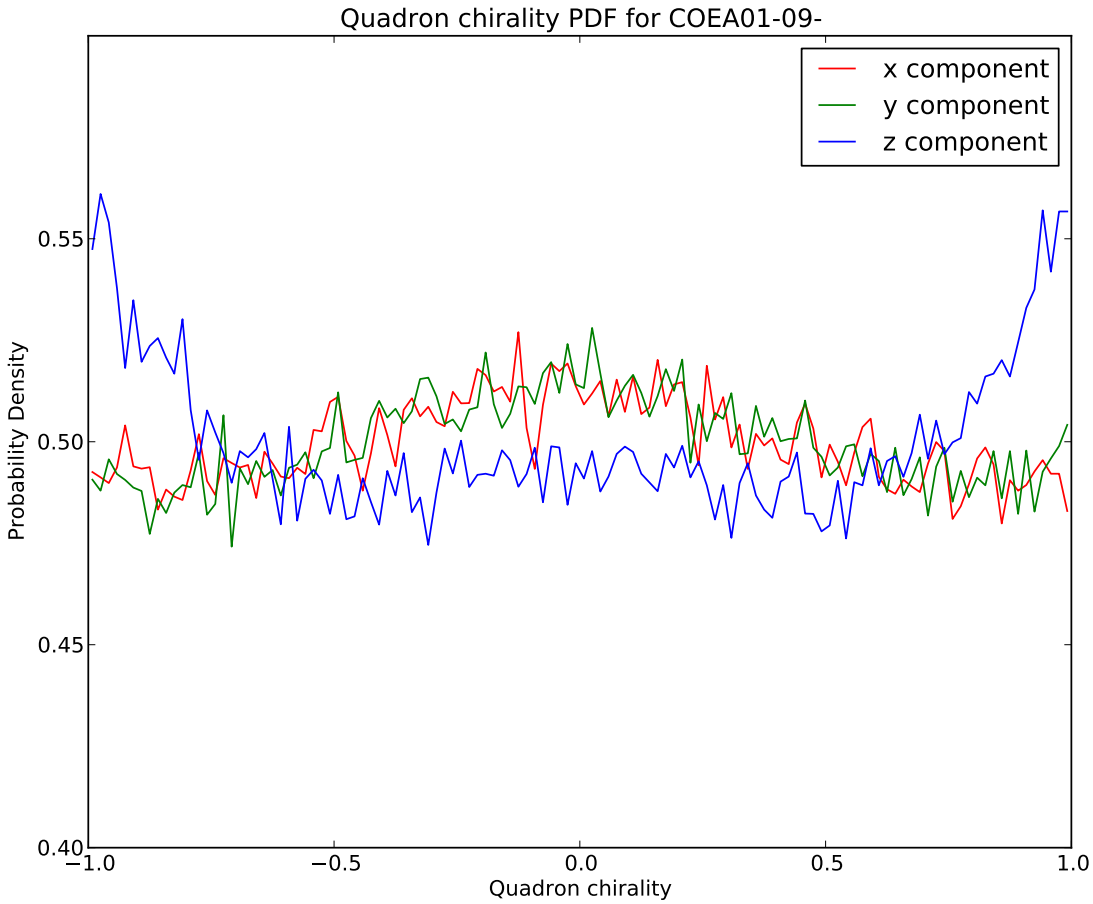


Figure C.38 – Normalised quadron chirality components – where the pseudo-vector has been given length 1 and each component plotted as a probability distribution, for system COEA01-09-.

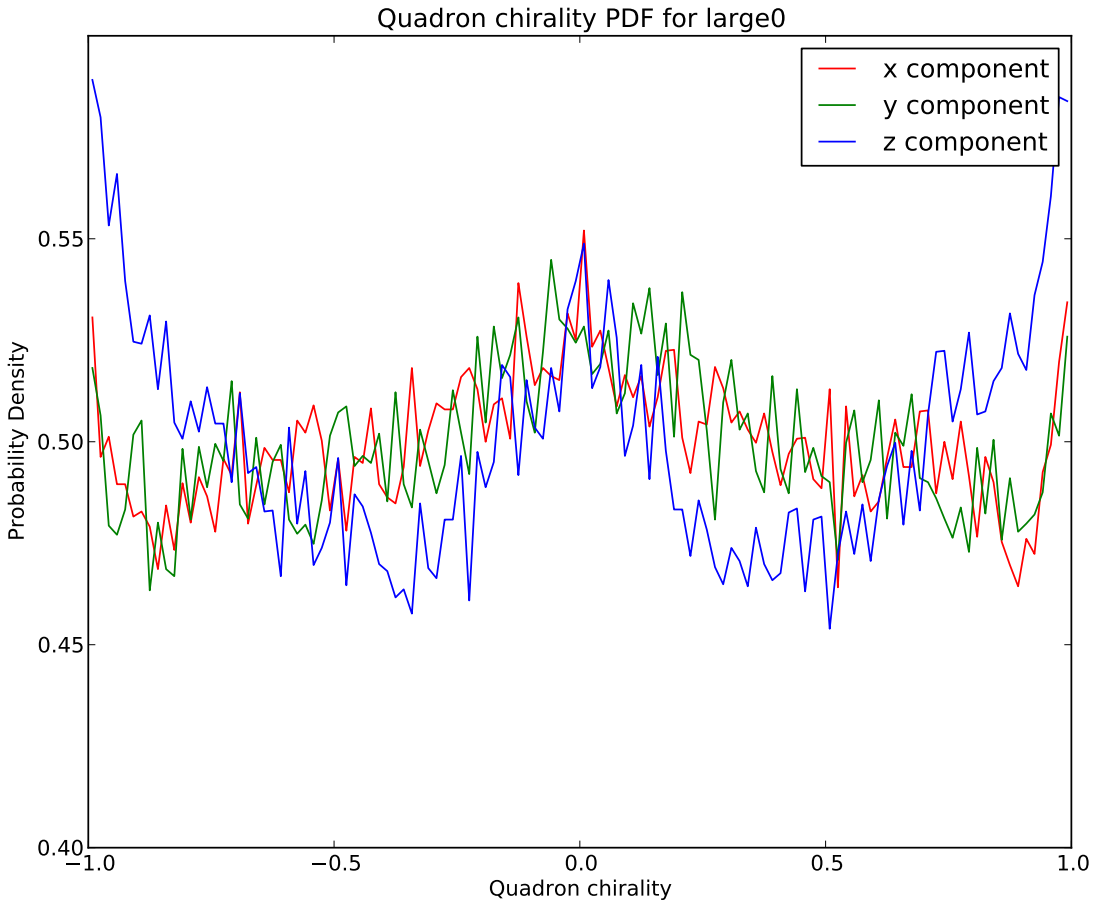


Figure C.39 – Normalised quadron chirality components – where the pseudo-vector has been given length 1 and each component plotted as a probability distribution, for system large0.

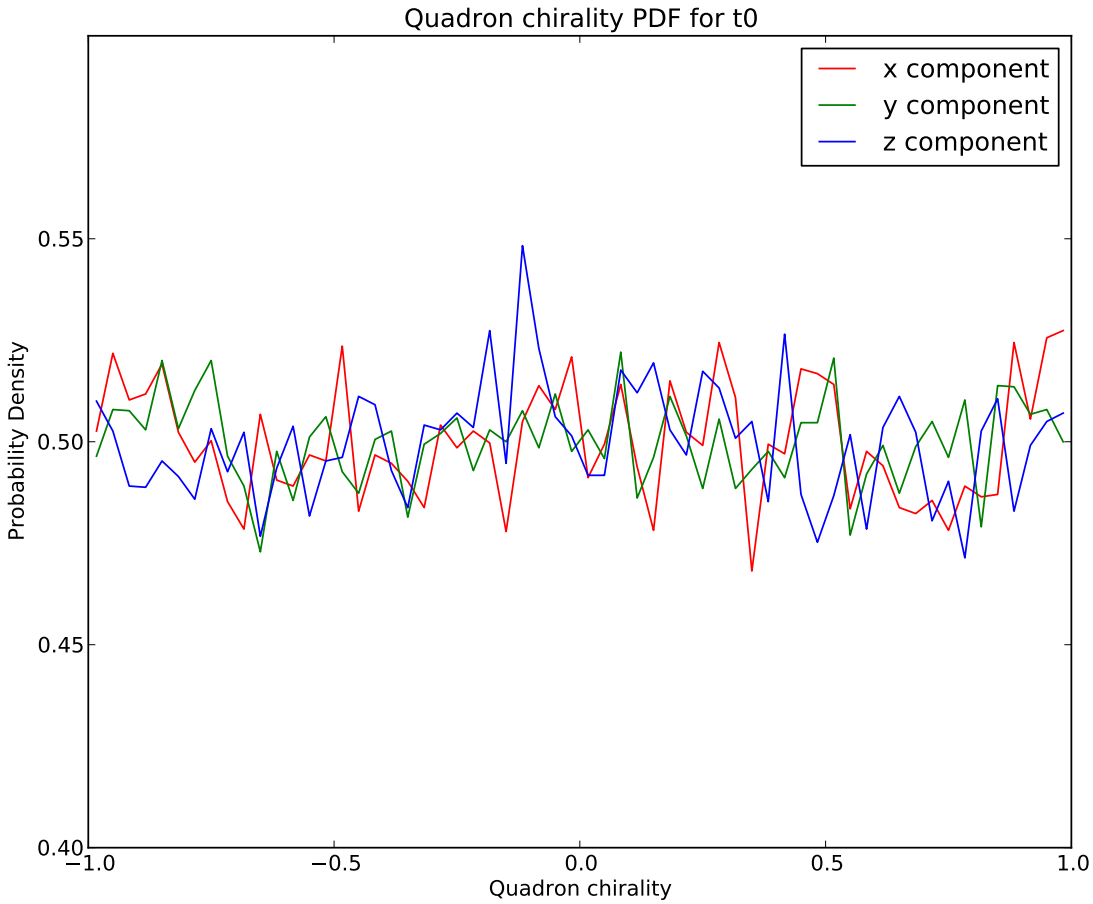


Figure C.40 – Normalised quadron chirality components – where the pseudo-vector has been given length 1 and each component plotted as a probability distribution, for system t0.

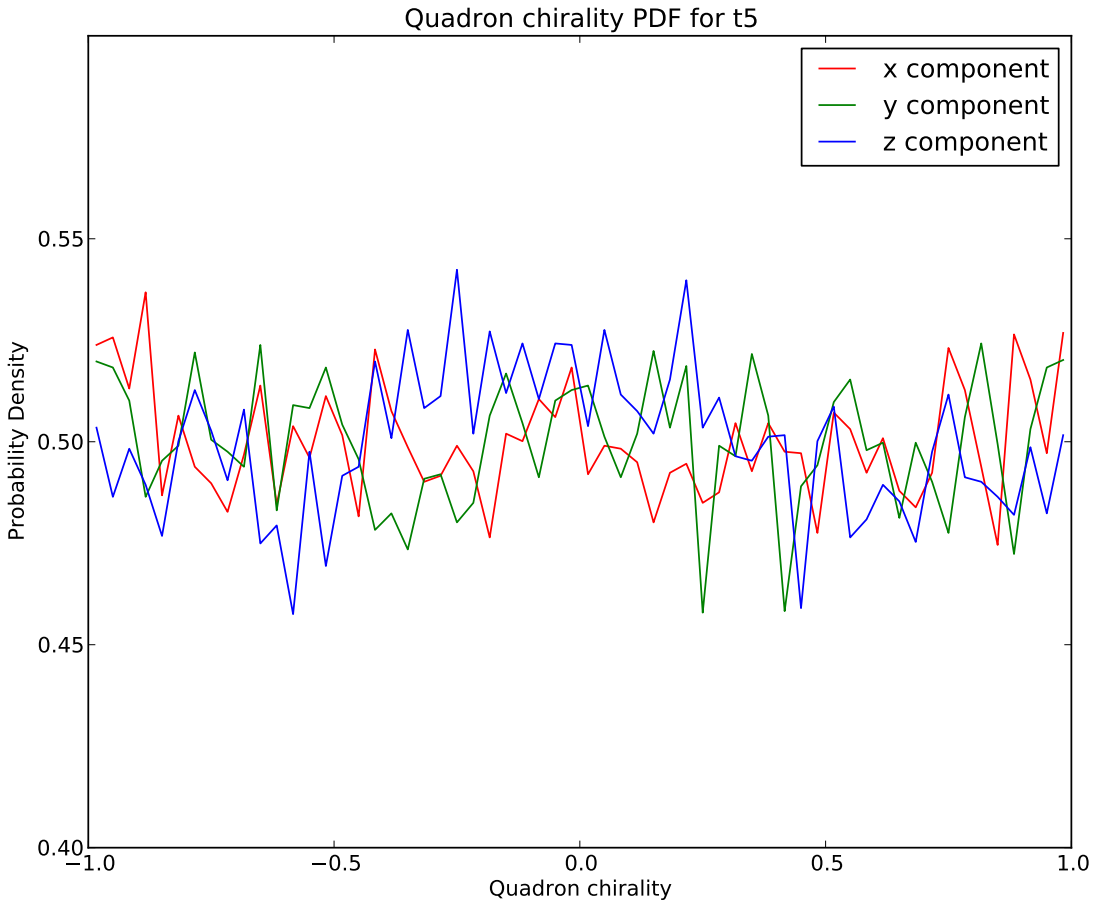


Figure C.41 – Normalised quadron chirality components – where the pseudo-vector has been given length 1 and each component plotted as a probability distribution, for system t5.

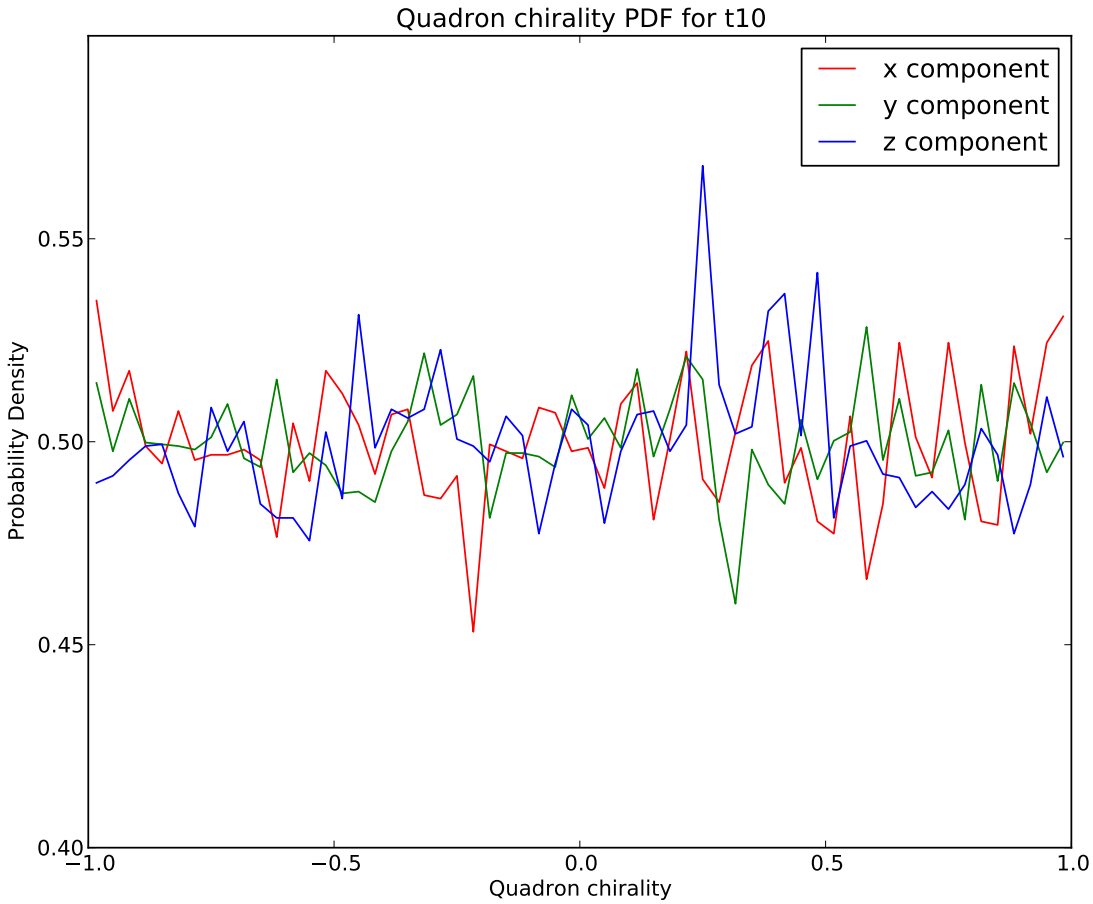


Figure C.42 – Normalised quadron chirality components – where the pseudo-vector has been given length 1 and each component plotted as a probability distribution, for system t10.

C.3 Zoo of Cell Types

Table C.1 – Cell Zoo - the different cell types detected in the 3SR Caicos initial condition and the number of quadrons detected in such cells.

Cell Type	**Frequency**	**Cell Type**	**Frequency**
$(4,)$	298,413	$(4,1)$	160,220
$(4,2)$	139,390	$(4,3)$	72,914
$(6,2)$	61,754	$(6,1)$	59,228
$(2,3)$	53,640	$(6,3)$	50,300
$(2,2)$	46,539	$(4,4)$	34,496
$(6,4)$	30,149	$(8,3)$	27,946
$(8,2)$	25,400	$(2,4)$	21,941
$(8,4)$	21,037	$(8,1)$	18,973
$(5,2,1)$	16,743	$(6,5)$	16,549
$(5,1,1)$	15,936	$(4,5)$	14,944
$(5,3,1)$	13,217	$(8,5)$	13,050
$(10,3)$	12,735	$(10,4)$	11,577
$(7,2,1)$	11,089	$(3,2,1)$	10,767
$(10,2)$	10,724	$(7,3,1)$	10,390
$(3,3,1)$	9,785	$(8,)$	9,672
$(6,6)$	8,623	$(2,5)$	8,174
$(10,5)$	7,994	$(7,1,1)$	7,858
$(0,3)$	7,676	$(5,4,1)$	7,483
$(7,4,1)$	7,404	$(9,3,1)$	6,735
$(8,6)$	6,134	$(9,2,1)$	5,652
$(4,6)$	5,555	$(10,1)$	5,496

Continued on next page

Table C.1 – continued from previous page

Cell Type	Frequency	Cell Type	Frequency
$(12, 3)$	5,425	$(10, 6)$	5,351
$(3, 4, 1)$	5,312	$(12, 4)$	5,246
$(7, 5, 1)$	5,141	$(12, 5)$	5,026
$(5, 5, 1)$	5,014	$(9, 4, 1)$	4,868
$(12, 2)$	4,493	$(1, 2, 1)$	3,870
$(9, 5, 1)$	3,837	$(6, 7)$	3,758
$(3, 1, 1)$	3,668	$(8, 7)$	3,503
$(11, 3, 1)$	3,500	$(9, 1, 1)$	3,333
$(11, 4, 1)$	3,207	$(7, 0, 1)$	3,166
$(10, 7)$	3,063	$(3, 5, 1)$	2,708
$(14, 4)$	2,544	$(12, 6)$	2,340
$(7, 6, 1)$	2,317	$(11, 5, 1)$	2,315
$(6, 2, 2)$	2,299	$(11, 2, 1)$	2,288
$(4, 2, 2)$	2,284	$(5, 0, 1)$	2,251
$(10,)$	2,232	$(2, 6)$	2,228
$(4, 1, 2)$	2,197	$(5, 6, 1)$	2,177
$(0, 4)$	2,164	$(6, 3, 2)$	2,117
$(1, 3, 1)$	2,006	$(6, 1, 2)$	1,971
$(4, 7)$	1,966	$(2, 2, 2)$	1,898
$(9, 6, 1)$	1,878	$(4, 3, 2)$	1,851
$(14, 3)$	1,806	$(14, 5)$	1,791
$(13, 4, 1)$	1,680	$(8, 3, 2)$	1,625
$(13, 3, 1)$	1,622	$(6, 2, 0, 1)$	1,621
$(8, 8)$	1,606	$(12, 7)$	1,570

Continued on next page

Table C.1 – continued from previous page

Cell Type	Frequency	Cell Type	Frequency
$(2, 1, 2)$	1,560	$(8, 2, 2)$	1,539
$(12, 1)$	1,491	$(1, 4, 1)$	1,416
$(4, 4, 2)$	1,416	$(10, 3, 2)$	1,409
$(4, 0, 2)$	1,353	$(0, 5)$	1,325
$(14, 2)$	1,281	$(11, 7, 1)$	1,281
$(11, 6, 1)$	1,262	$(6, 4, 2)$	1,219
$(14, 6)$	1,216	$(6, 8)$	1,200
$(11, 1, 1)$	1,193	$(13, 5, 1)$	1,110
$(10, 4, 2)$	1,110	$(9, 7, 1)$	1,090
$(8, 1, 2)$	1,075	$(8, 2, 0, 1)$	1,075
$(8, 4, 2)$	1,042	$(4, 2, 0, 1)$	1,037
$(13, 2, 1)$	1,030	$(3, 6, 1)$	1,011
$(6, 4, 0, 1)$	1,007	$(2, 3, 2)$	982
$(7, 7, 1)$	970	$(13, 6, 1)$	965
$(6,)$	943	$(9, 0, 1)$	940
$(14, 7)$	924	$(10, 8)$	914
$(4, 3, 0, 1)$	883	$(4, 8)$	882
$(8, 5, 2)$	873	$(8, 9)$	873
$(6, 1, 0, 1)$	856	$(6, 3, 0, 1)$	839
$(14, 8)$	828	$(4, 4, 0, 1)$	821
$(6, 5, 2)$	813	$(6, 0, 2)$	804
$(16, 4)$	765	$(13, 1, 1)$	753
$(15, 5, 1)$	744	$(10, 5, 2)$	734
$(8, 3, 0, 1)$	712	$(5, 7, 1)$	708

Continued on next page

Table C.1 – continued from previous page

Cell Type	Frequency	Cell Type	Frequency
$(0, 1, 2)$	706	$(4, 5, 2)$	701
$(10, 2, 0, 1)$	695	$(14, 1)$	684
$(3, 1, 1, 1)$	656	$(10, 2, 2)$	645
$(2, 7)$	639	$(1, 5, 1)$	632
$(10, 6, 2)$	618	$(5, 2, 1, 1)$	607
$(7, 9, 1)$	584	$(9, 8, 1)$	584
$(7, 2, 1, 1)$	580	$(5, 3, 1, 1)$	574
$(3, 2, 1, 1)$	557	$(8, 1, 0, 1)$	544
$(16, 2)$	542	$(2, 4, 2)$	539
$(12, 4, 2)$	537	$(0, 6)$	534
$(13, 8, 1)$	524	$(2, 0, 2)$	515
$(12, 8)$	515	$(16, 5)$	507
$(5, 8, 1)$	502	$(0, 2, 2)$	502
$(6, 5, 0, 1)$	492	$(18, 4)$	486
$(17, 5, 1)$	479	$(8, 6, 2)$	475
$(10, 1, 2)$	453	$(6, 6, 2)$	445
$(16, 6)$	443	$(5, 3, 3)$	428
$(2, 3, 0, 1)$	425	$(7, 3, 1, 1)$	415
$(5, 1, 1, 1)$	413	$(15, 6, 1)$	403
$(8, 8, 2)$	393	$(8, 4, 0, 1)$	392
$(10, 4, 0, 1)$	377	$(12, 9)$	375
$(7, 8, 1)$	374	$(10, 9)$	374
$(3, 1, 3)$	359	$(6, 9)$	358
$(16, 7)$	349	$(12,)$	346

Continued on next page

Table C.1 – continued from previous page

Cell Type	Frequency	Cell Type	Frequency
$(10, 1, 0, 1)$	343	$(10, 3, 0, 1)$	342
$(18, 6)$	338	$(12, 6, 2)$	334
$(15, 4, 1)$	333	$(8, 0, 2)$	327
$(9, 9, 1)$	321	$(5, 1, 3)$	321
$(3, 7, 1)$	318	$(3, 3, 1, 1)$	316
$(11, 0, 1)$	311	$(19, 3, 1)$	308
$(4, 5, 0, 1)$	305	$(9, 2, 3)$	305
$(2, 2, 0, 1)$	296	$(12, 3, 2)$	296
$(14, 4, 2)$	294	$(11, 4, 3)$	288
$(10, 10)$	287	$(7, 1, 1, 1)$	286
$(12, 5, 2)$	278	$(11, 8, 1)$	276
$(9, 4, 3)$	275	$(9, 3, 3)$	261
$(15, 2, 1)$	260	$(7, 5, 1, 1)$	257
$(2, 5, 2)$	254	$(4, 6, 2)$	250
$(5, 3, 0, 0, 1)$	250	$(18, 5)$	242
$(14, 5, 2)$	241	$(3, 0, 3)$	235
$(1, 6, 1)$	232	$(15, 7, 1)$	227
$(8, 5, 0, 1)$	226	$(5, 0, 1, 1)$	224
$(6, 6, 0, 1)$	222	$(1, 2, 1, 1)$	217
$(4, 1, 0, 1)$	217	$(19, 5, 1)$	209
$(11, 3, 1, 1)$	209	$(12, 4, 0, 1)$	209
$(6, 7, 2)$	207	$(5, 4, 3)$	207
$(5, 4, 0, 0, 1)$	205	$(15, 3, 1)$	204
$(4, 7, 2)$	204	$(3, 2, 3)$	203

Continued on next page

Table C.1 – continued from previous page

Cell Type	Frequency	Cell Type	Frequency
$(1, 3, 1, 1)$	202	$(17, 8, 1)$	200
$(9, 3, 0, 0, 1)$	198	$(8, 3, 4)$	198
$(7, 1, 3)$	198	$(11, 5, 3)$	196
$(12, 2, 2)$	193	$(16, 5, 2)$	192
$(0, 3, 2)$	190	$(7, 3, 3)$	190
$(8, 10)$	189	$(3, 3, 3)$	186
$(17, 4, 1)$	185	$(8, 0, 0, 1)$	184
$(11, 2, 1, 1)$	181	$(9, 5, 1, 1)$	180
$(7, 3, 0, 0, 1)$	180	$(3, 8, 1)$	177
$(5, 2, 0, 0, 1)$	176	$(5, 4, 1, 1)$	174
$(9, 1, 3)$	174	$(4, 9)$	173
$(14, 10)$	170	$(10, 8, 2)$	170
$(8, 7, 2)$	170	$(19, 7, 1)$	170
$(1, 1, 1, 1)$	169	$(10, 7, 2)$	169
$(13, 0, 1)$	169	$(13, 7, 1)$	167
$(16, 3)$	162	$(3, 4, 1, 1)$	161
$(5, 6, 1, 1)$	161	$(9, 2, 1, 1)$	160
$(4, 1, 2, 1)$	157	$(11, 2, 3)$	156
$(2, 8)$	156	$(2, 2, 2, 1)$	155
$(7, 2, 0, 0, 1)$	153	$(2, 1, 2, 1)$	152
$(17, 2, 1)$	151	$(3, 4, 3)$	149
$(10, 5, 0, 1)$	147	$(15, 6, 1, 1)$	144
$(14, 6, 0, 1)$	144	$(1, 2, 3)$	142
$(20, 8)$	142	$(12, 5, 0, 1)$	139

Continued on next page

Table C.1 – continued from previous page

Cell Type	Frequency	Cell Type	Frequency
$(7, 5, 3)$	138	$(7, 4, 3)$	138
$(6, 0, 0, 1)$	138	$(13, 4, 1, 1)$	137
$(5, 0, 3)$	136	$(4, 1, 1, 0, 1)$	136
$(1, 1, 3)$	135	$(12, 2, 0, 1)$	134
$(16, 2, 2)$	132	$(7, 4, 0, 0, 1)$	131
$(16, 4, 2)$	131	$(8, 11)$	131
$(16, 8)$	131	$(6, 2, 0, 2)$	129
$(2, 0, 2, 1)$	129	$(5, 2, 3)$	126
$(14, 2, 2)$	126	$(2, 4, 0, 1)$	125
$(6, 1, 1, 0, 1)$	125	$(19, 2, 1)$	123
$(20, 7)$	123	$(4, 2, 1, 0, 1)$	121
$(7, 2, 3)$	121	$(3, 0, 1, 1)$	121
$(6, 1, 2, 1)$	120	$(9, 4, 1, 1)$	119
$(8, 2, 4)$	118	$(7, 0, 3)$	116
$(6, 4, 2, 1)$	116	$(18, 1)$	116
$(11, 3, 3)$	113	$(7, 4, 1, 1)$	112
$(12, 1, 2)$	111	$(6, 3, 2, 1)$	111
$(12, 3, 0, 1)$	110	$(24, 7, 2)$	110
$(0, 3, 0, 1)$	110	$(9, 3, 1, 1)$	110
$(9, 10, 1)$	109	$(16, 6, 2)$	108
$(5, 9, 1)$	107	$(16, 1)$	104
$(7, 0, 1, 1)$	104	$(14, 3, 0, 1)$	103
$(1, 4, 1, 1)$	103	$(9, 1, 1, 1)$	103
$(13, 5, 3)$	102	$(9, 6, 1, 1)$	101

Continued on next page

Table C.1 – continued from previous page

Cell Type	Frequency	Cell Type	Frequency
(1, 7, 1)	101	(5, 5, 0, 0, 1)	98
(4, 2, 2, 1)	98	(2, 7, 2)	98
(11, 5, 1, 1)	97	(5, 6, 0, 0, 1)	97
(6, 2, 4)	92	(3, 6, 1, 1)	92
(15, 8, 1)	90	(21, 7, 1)	89
(6, 2, 2, 1)	89	(12, 3, 2, 1)	88
(3, 1, 2, 0, 1)	88	(15, 1, 1)	86
(20, 12)	85	(19, 6, 1, 1)	85
(8, 6, 0, 1)	85	(11, 7, 1, 1)	85
(18, 7)	84	(14,)	84
(16, 9)	84	(8, 2, 0, 0, 0, 1)	83
(11, 3, 0, 0, 1)	83	(12, 1, 0, 1)	82
(11, 11, 1)	82	(4, 4, 1, 0, 1)	81
(2, 9)	80	(14, 8, 0, 1)	80
(8, 2, 1, 0, 1)	80	(6, 10)	79
(18, 6, 0, 1)	78	(20, 2, 2)	78
(12, 9, 0, 1)	78	(4, 3, 1, 0, 1)	78
(14, 14)	78	(9, 9, 3)	78
(11, 9, 1)	78	(13, 4, 3)	76
(20, 4)	76	(12, 10)	76
(1, 0, 3, 1)	76	(5, 5, 1, 1)	76
(1, 0, 3)	76	(14, 9, 2)	76
(13, 10, 1, 1)	75	(6, 6, 4)	75
(17, 10, 1)	75	(10, 4, 2, 1)	74

Continued on next page

Table C.1 – continued from previous page

Cell Type	Frequency	Cell Type	Frequency
$(6, 7, 0, 1)$	74	$(12, 7, 2)$	73
$(16, 2, 2, 1)$	72	$(4, 6, 0, 1)$	71
$(10, 11)$	70	$(2, 10)$	70
$(13, 3, 1, 1)$	70	$(11, 0, 3)$	69
$(14, 5, 0, 1)$	69	$(1, 8, 1)$	69
$(12, 4, 3, 0, 1)$	68	$(10, 0, 0, 1)$	68
$(6, 3, 1, 0, 1)$	68	$(7, 4, 2, 0, 1)$	67
$(17, 5, 0, 0, 1)$	66	$(18, 3)$	66
$(2, 5, 0, 1)$	66	$(17, 3, 3)$	65
$(3, 4, 0, 0, 1)$	65	$(10, 2, 1, 0, 1)$	65
$(10, 3, 1, 0, 1)$	65	$(22, 8)$	65
$(13, 6, 1, 2)$	65	$(2, 3, 2, 1)$	65
$(14, 3, 2)$	64	$(4, 0, 2, 1)$	64
$(11, 1, 3)$	64	$(8, 2, 2, 1)$	64
$(20, 5)$	63	$(8, 4, 1, 0, 1)$	63
$(12, 6, 0, 1)$	62	$(7, 11, 1)$	62
$(13, 2, 3)$	62	$(16, 3, 0, 1)$	62
$(6, 2, 1, 0, 1)$	62	$(4, 3, 2, 1)$	62
$(12, 3, 1, 0, 1)$	62	$(2, 3, 0, 2)$	61
$(3, 5, 0, 0, 1)$	60	$(10, 2, 0, 0, 0, 1)$	60
$(3, 3, 0, 0, 1)$	60	$(7, 6, 3)$	60
$(15, 1, 1, 1)$	60	$(4, 10, 2)$	59
$(20, 6)$	59	$(7, 1, 0, 0, 1)$	59
$(13, 6, 1, 1)$	59	$(7, 6, 2, 0, 0, 0, 1)$	59

Continued on next page

Table C.1 – continued from previous page

Cell Type	Frequency	Cell Type	Frequency
$(18, 4, 2)$	58	$(6, 1, 4)$	58
$(8, 5, 2, 1)$	57	$(4, 3, 0, 0, 0, 1)$	56
$(3, 1, 1, 2)$	56	$(17, 1, 1)$	56
$(20, 3)$	56	$(14, 1, 2)$	56
$(10, 4, 1, 0, 1)$	55	$(16, 3, 2)$	55
$(17, 6, 2, 0, 1)$	55	$(0, 4, 0, 1)$	55
$(18, 8)$	55	$(5, 6, 3)$	54
$(7, 3, 2, 0, 0, 0, 1)$	54	$(6, 9, 2)$	54
$(9, 2, 3, 2)$	54	$(13, 2, 0, 0, 1)$	54
$(9, 11, 3)$	53	$(4, 2, 4)$	53
$(13, 1, 1, 1)$	53	$(5, 1, 2, 0, 1)$	53
$(11, 1, 3, 1)$	53	$(16, 8, 0, 1)$	52
$(11, 1, 0, 0, 1)$	52	$(8, 3, 2, 1)$	52
$(11, 10, 1)$	52	$(6, 3, 4, 1)$	52
$(14, 2, 0, 1)$	52	$(5, 2, 2, 0, 1)$	52
$(9, 1, 3, 1)$	52	$(10, 6, 0, 1)$	51
$(9, 2, 0, 0, 1)$	51	$(12, 2, 4)$	51
$(10, 4, 0, 0, 0, 1)$	51	$(7, 2, 4, 0, 0, 0, 1)$	51
$(15, 0, 1)$	50	$(18, 2)$	50
$(1, 5, 1, 1)$	50	$(6, 3, 0, 2)$	50
$(7, 2, 3, 1)$	50	$(4, 8, 0, 1)$	50
$(9, 1, 1, 2)$	50	$(5, 4, 5)$	49
$(3, 5, 1, 1)$	49	$(3, 9, 1)$	49
$(12, 1, 1, 0, 1)$	48	$(4, 7, 0, 1)$	48

Continued on next page

Table C.1 – continued from previous page

Cell Type	Frequency	Cell Type	Frequency
$(4, 1, 0, 2)$	48	$(19, 1, 1)$	47
$(7, 1, 5, 1)$	47	$(8, 1, 0, 2)$	47
$(4, 8, 2)$	47	$(9, 2, 1, 0, 0, 1)$	46
$(6, 0, 2, 1)$	46	$(0, 8)$	46
$(8, 7, 2, 1)$	46	$(7, 6, 3, 1)$	46
$(6, 4, 4)$	46	$(11, 6, 1, 1)$	45
$(17, 6, 1)$	45	$(8, 2, 0, 2)$	44
$(3, 5, 3)$	44	$(4, 0, 0, 2)$	44
$(6, 4, 1, 0, 1)$	44	$(10, 2, 2, 1)$	44
$(1, 9, 1)$	44	$(6, 2, 1, 1, 1)$	44
$(6, 12)$	44	$(13, 3, 0, 0, 1)$	44
$(9, 0, 3, 1)$	43	$(4, 3, 4)$	43
$(21, 4, 1)$	43	$(6, 2, 4, 1)$	43
$(18, 0, 2)$	43	$(4, 5, 4)$	43
$(5, 2, 4, 0, 1)$	43	$(6, 4, 0, 0, 0, 1)$	42
$(7, 0, 3, 1)$	42	$(3, 3, 3, 1)$	42
$(4, 2, 0, 2)$	42	$(9, 0, 3)$	42
$(1, 4, 0, 0, 1)$	41	$(14, 9)$	41
$(7, 2, 1, 0, 0, 1)$	41	$(0, 0, 2, 1)$	41
$(9, 0, 1, 1)$	41	$(7, 5, 0, 0, 1)$	40
$(5, 1, 3, 1)$	40	$(7, 7, 1, 1)$	39
$(7, 2, 0, 1, 1)$	39	$(9, 7, 3)$	39
$(8, 3, 0, 0, 0, 1)$	39	$(2, 1, 0, 2)$	39
$(5, 1, 0, 1, 1)$	38	$(2, 4, 2, 1)$	38

Continued on next page

Table C.1 – continued from previous page

Cell Type	Frequency	Cell Type	Frequency
$(7, 6, 1, 1)$	38	$(0, 1, 0, 2)$	38
$(6, 7, 4)$	37	$(4, 3, 1, 0, 0, 0, 1)$	37
$(6, 13)$	37	$(2, 2, 0, 2)$	37
$(2, 4, 1, 0, 1)$	36	$(12, 0, 2)$	36
$(6, 8, 2)$	36	$(0, 5, 2)$	36
$(4, 1, 4)$	36	$(4, 5, 1, 0, 1)$	36
$(3, 1, 5)$	35	$(1, 3, 2, 0, 1)$	35
$(9, 1, 0, 0, 1)$	35	$(3, 1, 0, 1, 1)$	35
$(5, 10, 1)$	34	$(6, 0, 1, 0, 1)$	34
$(9, 0, 0, 0, 1)$	34	$(3, 6, 0, 0, 1)$	34
$(9, 2, 5)$	33	$(6, 9, 0, 1)$	33
$(4, 2, 4, 1)$	33	$(16, 1, 0, 1)$	32
$(1, 2, 2, 0, 1)$	31	$(22, 9)$	31
$(13, 2, 1, 1)$	31	$(10, 6, 1, 0, 1)$	30
$(9, 0, 2, 0, 1)$	30	$(4, 0, 4)$	30
$(1, 5, 0, 0, 1)$	30	$(1, 1, 2, 0, 1)$	30
$(8, 8, 0, 1)$	29	$(6, 3, 4)$	29
$(0, 7)$	29	$(4, 4, 0, 0, 0, 1)$	29
$(2, 1, 4)$	28	$(5, 4, 2, 0, 1)$	28
$(6, 1, 2, 2)$	28	$(0, 2, 2, 1)$	28
$(10, 6, 2, 1)$	28	$(2, 2, 1, 0, 1)$	28
$(9, 3, 0, 0, 0, 0, 1)$	27	$(0, 4, 2)$	27
$(7, 4, 2, 1, 1)$	26	$(2, 0, 4)$	26
$(17, 0, 1, 1)$	26	$(5, 0, 3, 1)$	26

Continued on next page

Table C.1 – continued from previous page

Cell Type	Frequency	Cell Type	Frequency
$(5, 4, 1, 0, 0, 1)$	26	$(5, 2, 1, 0, 0, 1)$	25
$(7, 6, 0, 0, 1)$	25	$(8, 3, 1, 0, 1)$	25
$(8, 4, 2, 1)$	25	$(8, 2, 1, 0, 0, 0, 1)$	24
$(4, 5, 0, 2)$	24	$(6, 5, 0, 0, 0, 1)$	24
$(5, 5, 3)$	23	$(3, 3, 5)$	23
$(0, 1, 2, 1)$	23	$(4, 2, 0, 0, 0, 1)$	23
$(3, 2, 0, 0, 1)$	23	$(4, 5, 2, 1)$	23
$(11, 1, 1, 1)$	22	$(10, 6, 0, 0, 0, 1)$	22
$(12, 9, 2)$	22	$(0, 5, 1, 0, 1)$	22
$(0, 2, 4)$	22	$(5, 2, 0, 1, 1)$	22
$(10, 2, 2, 0, 0, 0, 0, 1)$	22	$(7, 10, 1)$	22
$(3, 0, 0, 1, 1)$	22	$(2, 2, 4, 1)$	21
$(6, 2, 1, 0, 0, 0, 1)$	21	$(2, 1, 0, 0, 2)$	21
$(8, 5, 0, 0, 0, 1)$	21	$(5, 7, 1, 1)$	21
$(7, 3, 1, 0, 0, 1)$	21	$(11, 3, 1, 2)$	21
$(4, 6, 0, 2)$	20	$(12, 1, 0, 2)$	20
$(6, 3, 0, 0, 0, 1)$	20	$(9, 0, 2, 1, 1)$	20
$(4, 1, 2, 0, 0, 1)$	20	$(6, 0, 0, 2)$	20
$(14, 0, 2)$	19	$(12, 0, 0, 1)$	19
$(5, 3, 2, 0, 1)$	19	$(6, 3, 1, 1, 1)$	19
$(8, 0, 2, 1)$	19	$(4, 6, 1, 0, 1)$	18
$(7, 7, 0, 0, 1)$	18	$(2, 2, 4)$	18
$(13, 2, 2, 0, 0, 0, 1)$	18	$(5, 3, 3, 1)$	18
$(4, 10)$	18	$(3, 2, 1, 0, 0, 1)$	18

Continued on next page

Table C.1 – continued from previous page

Cell Type	Frequency	Cell Type	Frequency
$(5, 3, 1, 0, 0, 1)$	18	$(5, 1, 1, 0, 0, 1)$	17
$(10, 3, 0, 0, 0, 1)$	17	$(0, 6, 2)$	17
$(6, 5, 2, 1)$	17	$(14, 7, 0, 1)$	17
$(7, 0, 1, 0, 0, 1)$	17	$(9, 5, 0, 0, 0, 0, 1)$	17
$(12, 4, 1, 0, 1)$	17	$(10, 9, 2)$	17
$(4, 3, 0, 2)$	17	$(9, 4, 0, 0, 1)$	17
$(10, 5, 0, 0, 0, 1)$	16	$(3, 6, 3)$	16
$(8, 4, 4)$	16	$(1, 3, 3)$	16
$(1, 1, 3, 1)$	16	$(11, 6, 3)$	16
$(0, 2, 0, 0, 2)$	16	$(11, 7, 1, 2)$	16
$(8, 4, 0, 0, 0, 1)$	16	$(9, 6, 3)$	16
$(7, 2, 2, 0, 1)$	16	$(3, 0, 2, 0, 1)$	15
$(6, 0, 4)$	15	$(2, 6, 2, 1)$	15
$(8, 7, 0, 0, 0, 1)$	15	$(1, 3, 0, 0, 1)$	15
$(13, 8, 0, 0, 1)$	15	$(2, 3, 1, 0, 1)$	15
$(13, 2, 0, 0, 0, 0, 1)$	15	$(1, 0, 3, 0, 0, 1)$	14
$(11, 4, 1, 1)$	14	$(2, 3, 4)$	14
$(0, 0, 1, 1, 1)$	14	$(0, 0, 1, 0, 1, 1)$	14
$(11, 1, 1, 2)$	14	$(1, 2, 3, 1)$	14
$(12, 10, 2)$	13	$(0, 3, 1, 0, 1)$	13
$(2, 7, 0, 1)$	12	$(8, 7, 0, 1)$	12
$(9, 1, 1, 1, 0, 1)$	12	$(6, 1, 0, 1, 0, 1)$	12
$(1, 2, 1, 2)$	11	$(9, 2, 3, 1)$	11
$(0, 3, 2, 1)$	11	$(7, 3, 3, 0, 0, 0, 0, 1)$	11

Continued on next page

Table C.1 – continued from previous page

Cell Type	Frequency	Cell Type	Frequency
$(11, 3, 2, 0, 0, 0, 1)$	11	$(2, 6, 2)$	11
$(2, 6, 1, 0, 1)$	11	$(5, 3, 0, 0, 1, 1)$	11
$(0, 2, 0, 2)$	11	$(11, 4, 0, 1, 1)$	11
$(10, 7, 0, 0, 0, 1)$	11	$(3, 0, 1, 2)$	11
$(7, 2, 1, 2)$	11	$(3, 5, 2, 0, 1)$	11
$(7, 3, 2, 0, 1)$	11	$(7, 3, 0, 1, 1)$	11
$(6, 4, 1, 0, 0, 0, 1)$	10	$(2, 2, 1, 1, 1)$	10
$(9, 2, 1, 1, 0, 1)$	10	$(3, 3, 0, 1, 1)$	10
$(2, 1, 1, 0, 1)$	10	$(11, 2, 0, 0, 1)$	10
$(3, 2, 3, 1)$	10	$(7, 0, 1, 2)$	10
$(4, 0, 1, 1, 1)$	10	$(6, 5, 4)$	9
$(6, 6, 2, 1)$	9	$(0, 5, 0, 1)$	9
$(3, 2, 0, 1, 1)$	9	$(12, 2, 2, 1)$	9
$(3, 2, 1, 2)$	8	$(8, 0, 1, 0, 1)$	8
$(0, 1, 2, 2)$	8	$(10, 6, 0, 0, 0, 0, 0, 1)$	8
$(8, 1, 1, 0, 1)$	8	$(4, 2, 0, 1, 0, 1)$	8
$(9, 5, 0, 0, 1)$	8	$(9, 5, 1, 0, 0, 1)$	8
$(6, 1, 0, 2)$	8	$(10, 3, 0, 0, 0, 0, 0, 1)$	8
$(14, 3, 0, 0, 0, 1)$	8	$(10, 2, 1, 0, 0, 0, 1)$	8
$(6, 2, 0, 1, 0, 0, 0, 1)$	8	$(7, 2, 0, 1, 0, 0, 1)$	8
$(4, 4, 4)$	7	$(9, 3, 1, 0, 0, 1)$	7
$(3, 6, 1, 0, 0, 1)$	7	$(6, 1, 0, 0, 2)$	7
$(9, 5, 1, 0, 0, 0, 0, 1)$	7	$(3, 5, 1, 0, 0, 1)$	7
$(4, 2, 1, 1, 1)$	7	$(2, 3, 1, 1, 0, 0, 1)$	7

Continued on next page

Table C.1 – continued from previous page

Cell Type	Frequency	Cell Type	Frequency
$(9, 2, 2, 0, 0, 0, 1)$	7	$(18, 6, 2)$	7
$(7, 4, 1, 0, 0, 0, 0, 1)$	7	$(10, 1, 1, 0, 1)$	7
$(3, 3, 1, 0, 0, 1)$	7	$(9, 3, 3, 1)$	7
$(9, 0, 0, 0, 1, 1)$	6	$(2, 5, 0, 0, 0, 1)$	6
$(8, 3, 6)$	6	$(4, 4, 0, 2)$	6
$(0, 2, 2, 0, 0, 1)$	6	$(7, 3, 1, 2)$	6
$(13, 0, 1, 1)$	6	$(0, 5, 0, 0, 0, 1)$	6
$(1, 2, 0, 1, 1)$	6	$(5, 0, 2, 0, 1)$	6
$(7, 2, 2, 0, 0, 0, 1)$	6	$(6, 2, 2, 0, 0, 1)$	6
$(7, 8, 0, 0, 1)$	6	$(9, 0, 2, 0, 0, 0, 1)$	5
$(12, 5, 0, 0, 0, 1)$	5	$(5, 1, 1, 1, 0, 1)$	5
$(11, 4, 1, 0, 0, 1)$	5	$(4, 2, 1, 0, 0, 0, 1)$	5
$(8, 3, 0, 0, 0, 0, 0, 1)$	5	$(4, 0, 1, 0, 1)$	5
$(11, 1, 2, 0, 0, 0, 1)$	5	$(15, 1, 0, 0, 0, 0, 1)$	4
$(13, 0, 2, 0, 1)$	4	$(9, 2, 1, 0, 0, 0, 0, 1)$	4
$(7, 1, 2, 0, 0, 0, 1)$	4	$(6, 5, 2, 0, 0, 1)$	4
$(2, 5, 2, 1)$	4	$(6, 0, 0, 3)$	4
$(8, 4, 0, 0, 0, 0, 0, 1)$	4	$(3, 4, 0, 0, 0, 0, 1)$	4
$(1, 4, 3)$	4	$(5, 5, 0, 0, 0, 0, 1)$	4
$(10, 1, 0, 1, 0, 1)$	4	$(3, 4, 2, 0, 1)$	4
$(11, 1, 1, 0, 0, 0, 0, 1)$	4	$(8, 1, 2, 0, 0, 1)$	4
$(9, 0, 0, 1, 1)$	4	$(4, 11)$	4
$(10, 3, 4)$	4	$(12, 0, 1, 0, 1)$	3
$(6, 0, 2, 0, 0, 1)$	3	$(5, 2, 1, 2)$	3

Continued on next page

Table C.1 – continued from previous page

Cell Type	**Frequency**	**Cell Type**	**Frequency**
$(4, 6, 0, 0, 0, 1)$	3	$(2, 1, 2, 0, 0, 1)$	3
$(2, 6, 0, 1)$	3	$(3, 7, 0, 0, 1)$	3
$(5, 4, 0, 0, 0, 0, 1)$	3	$(3, 1, 3, 1)$	3
$(2, 3, 0, 0, 1, 0, 1)$	3	$(0, 2, 0, 3)$	3
$(4, 5, 0, 0, 0, 1)$	3	$(7, 0, 2, 0, 1)$	3
$(4, 2, 0, 1, 0, 0, 0, 1)$	3	$(3, 2, 0, 0, 1, 1)$	3
$(4, 0, 2, 0, 0, 1)$	3	$(4, 1, 0, 1, 0, 1)$	3
$(1, 0, 1, 2)$	2	$(2, 5, 1, 0, 1)$	2

Appendix D

Copyrighted Works Permissions

Emails in this section presented as the headers, less information that may be considered sensitive.

D.1 Permission from Dr. Blumenfeld

Outgoing email:

Subject: Permission for use of Figures from your publications in my Thesis

From: Joe Jordan <j.jordan10@imperial.ac.uk>

To: Rafi Blumenfeld <r.blumenfeld@imperial.ac.uk>

Dear Dr. Blumenfeld,

I am writing to request the use of the following figures, which are your and your co-authors' copyright, from the following publications:

Figures 4 and 5aI from:

Gad Frenkel, Raphael Blumenfeld, Zdenek Grof, and Peter R. King. Structural characterization and statistical properties of two-dimensional granular systems. Physical Review E, 77(4):041304, Apr 2008.

Figures 15, 16 and 20 from:

Takashi Matsushima and Raphael Blumenfeld. Fundamental structural characteristics of planar granular assemblies: scaling away friction and initial state. ArXiv e-prints, 2012.

These will be presented in a thesis for the degree of Doctor of Philosophy at Imperial College, London, which builds on your original work in these publications.

I look forward to hearing from you,

Joe Jordan

Response:

From: "Blumenfeld, Rafi" <r.blumenfeld@imperial.ac.uk>
To: "Jordan, Joe" <j.jordan10@imperial.ac.uk>
Subject: Re: Permission for use of Figures from your publications in my Thesis
Thread-Topic: Permission for use of Figures from your publications in my Thesis

You have my permission.

Rafi

D.2 Permission from Dr. Andó

Outgoing email:

Subject: Permission to use figures in my Thesis
From: Joe Jordan <j.jordan10@imperial.ac.uk>
To: "Edward Ando'" <edward.ando@3sr-grenoble.fr>

Dear Dr. Ando,

I am writing to request the use of the following figures, which are your and your co-authors' copyright, from the following communications:

Displacement and Rotation of particles in the COEA01 experiment from:
Private communication between Joe Jordan and Edward Ando on the 4th of May 2014.

These will be presented in a thesis for the degree of Doctor of Philosophy at Imperial College, London, which builds on your original work in the publication:
Edward And=C3=B2, Stephen A. Hall, Gioacchino Viggiani, Jacques Desrues, = and

Pierre B=C2=B4 esuelle. Grain-scale experimental investigation of localised deformation in sand: a discrete particle tracking approach. Acta Geotechnica, 7(1):1=E2=80=9313, 2012.

I look forward to hearing from you,

Joe Jordan

Response:

From: Edward Ando <edward.ando@3sr-grenoble.fr>
To: "Jordan, Joe" <j.jordan10@imperial.ac.uk>
Subject: Re: Permission to use figures in my Thesis
Thread-Topic: Permission to use figures in my Thesis

yes of course you can use it!

Eddy

D.3 Request for permission to Facebook

Originally the author of the figure in question, Paul Butler, was contacted. He replied to inform the author of this work that the copyright to the figure remained with his former employer, Facebook, and that the appropriate way to send such a request would be to their press office.

Thus, the following message was sent, and at time of publication no reply has yet been received.

Subject: Permission to use the "Visualizing Friendships" image in an academic work
From: Joe Jordan <j.jordan10@imperial.ac.uk>
To: press@fb.com

Hello,

I am writing to request permission to use the following diagram, which is your copyright, from the following source:

"Visualising Friendships" Facebook graph projection from:

https://www.facebook.com/notes/facebook-engineering/visualizing-friendships/469716
published
in December 2010.

This will be presented in a thesis for the degree of Doctor of Philosophy at Imperial College, London, which uses your diagram to aid explaining the concept of a graph. The thesis concerns the structure of granular materials, which can also be represented as graphs. The thesis will eventually be published in the public domain under a Creative Commons Attribution Non-commercial No Derivatives license, and your copyrighted figure will be flagged as your property within the text.

I first contacted Paul Butler, the former employee who created the diagram,

and he referred me to you as the copyright holders.

I look forward to hearing from you,

Joe Jordan