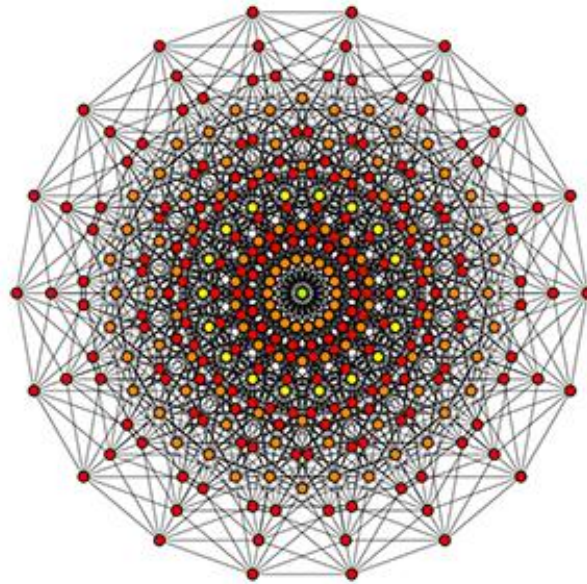


New method for making effective calculations in 'high-dimensional space'

October 13 2016



Projection of a 9-dimensional cube. High-dimensional spaces pose considerable problems when trying to make calculations and predictions – something that the new method devised by researchers aims to address. Credit: Tom Ruen via Wikimedia Commons

Researchers have developed a new method for making effective calculations in "high-dimensional space" – and proved its worth by using it to solve a 93-dimensional problem.

Researchers have developed a new technique for making calculations in

"high-dimensional space" – mathematical problems so wide-ranging in their scope, that they seem at first to be beyond the limits of human calculation.

In what sounds like the title of a rejected script for an Indiana Jones movie, the method improves on existing approaches to beat a well-known problem known as "The Curse Of Dimensionality". It was devised by a team of researchers at the University of Cambridge.

In rough terms, the "Curse", refers to the apparent impossibility of making calculations in situations where the number of variables, attributes, and possible outcomes is so large that it seems futile even to try to comprehend the problem in the first place.

A simple example is this: Imagine that you have a cup containing 100 grains of rice. You pick it up, shake it, and put it down again. The arrangement within the cup changes, but what are the chances of that arrangement occurring, relative to all other possibilities?

While most people would reasonably consider that problem not just impossible, but largely pointless, it illustrates the type of maths needed to make predictions about much bigger – and more meaningful – issues.

Those include, for example, trying to model the likely shape and impact of a decaying ecosystem, such as a developing area of deforestation, or the potential effect of different levels of demand on a power grid. More fundamentally, the same class of calculation would theoretically enable us to get to grips with the statistical probability of our own existence on Earth, or the chances that life might happen again, elsewhere in the Universe.

The new study was led by Stefano Martiniani, a Gates Scholar at St John's College, Cambridge, who carried out the work with colleagues in

the Department of Chemistry and at Microsoft Research in Cambridge.

"There is a very large class of problems that can be solved through the sort of approach that we have devised," Martiniani said. "It opens up a whole world of possibilities in the study of things like dynamical systems, chemical structure prediction, or [artificial neural networks](#)."

Most people understand "dimensions" to mean height, width, depth and time, but in Mathematics the term is also used flexibly to describe the number of parameters needed to specify a "state" for any given problem. The more complicated the problem in question, the greater the space you need to express the parameters. They therefore become "high-dimensional spaces".

Similarly, working out the likelihood of a particular outcome in a situation where all sorts of different variables apply – such as the grains of rice in a cup arranging themselves in a particular way – is a high-dimensional problem. Expressing and plotting the combined impact of the many parameters that might affect the outcome involves imagining a graph with multiple axes, as if working in numerous dimensions at once.

The method devised by Martiniani and colleagues, like other approaches, begins by characterising such challenges as an "energy landscape". The range of possible states in which a system such as the cup of rice may exist is envisaged as a landscape of mountains and valleys, in which the base of each valley is a stable state.

The set of initial conditions leading to this stable state is called a "basin of attraction". The fundamental theory is that, if the volume of each basin of attraction can be calculated, then this begins to provide some sort of indication of the probability of a given state's occurrence.

To do that, researchers build computer software which models high

dimensional systems, using the landscape analogy, and makes calculations within it.

The simplest model is a brute force approach, which essentially takes a reading, shakes the system up, takes another reading, and repeats the process – many millions of times – in an attempt to establish the probability of certain outcomes. A more sophisticated strategy recurrently starts in the same place and measures the average distance within the energy landscape in which the system finds the same basin of attraction, through which the user gradually develops an appreciation of its volume.

"In most cases you are like a blindfolded person, walking around drunk in the energy landscape," Martiniani said. "At any given moment, you only really know where you are and where you have just come from."

In the new study, however, the team applied a different approach to the same kind of problem. Borrowing a technique widely used in biomolecular simulations, called the Multistate Bennett Acceptance Ratio, they developed a method which systematically tests the limits of one particular basin of attraction. Rather than gauging its volume by just taking an average from random samples, it looks for the furthest and least likely limits.

The net result is a much more efficient sampling technique, which enables a much broader range of calculations in high-dimensional space.

To test this, the team modelled an imaginary 93-dimensional "system" made up of 32 soft spheres that could be packed together in multiple ways. They found that they were able to sample and quantify outcomes within that system that would only be found randomly one in every 10^{100} times. In other words, the chances of stumbling across those outcomes by chance would be one in ten duotrigintillion.

"In basic terms it goes where brute force sampling never will, because if you started to try, you would never finish," Martiniani added.

"Technically, the limits of the problems we can solve are now not those of the approach, but the computing power we need to simulate the underlying [energy landscape](#). When addressing these kinds of problems in high-dimensional space, this should now be the technique of choice."

More information: Stefano Martiniani et al. Structural analysis of high-dimensional basins of attraction, *Physical Review E* (2016). [DOI: 10.1103/PhysRevE.94.031301](https://doi.org/10.1103/PhysRevE.94.031301)

Provided by University of Cambridge

Citation: New method for making effective calculations in 'high-dimensional space' (2016, October 13) retrieved 21 April 2024 from <https://phys.org/news/2016-10-method-effective-high-dimensional-space.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.