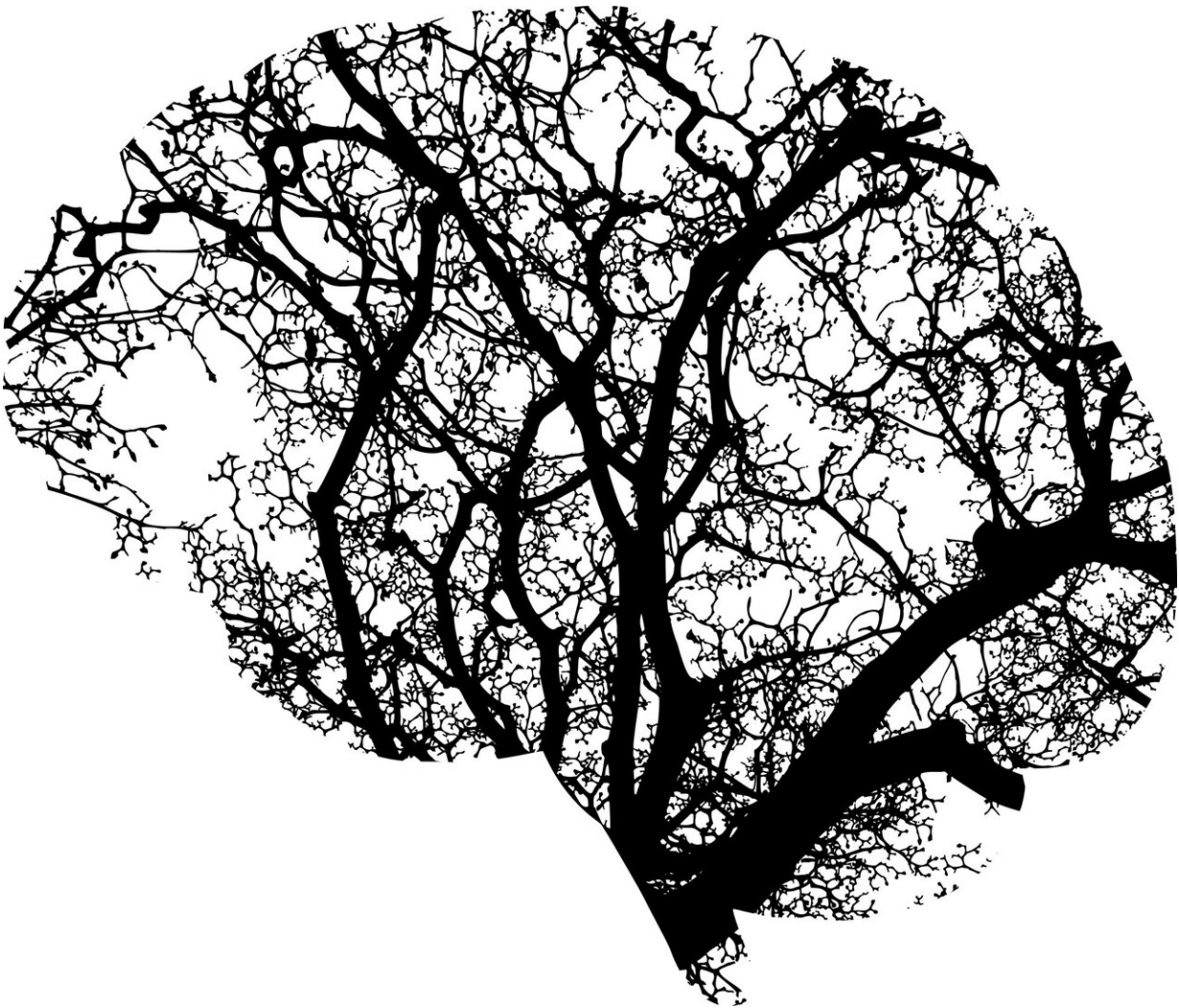# How AI is improving simulations with smarter sampling techniques

October 2 2024, by Rachel Gordon

Credit: CC0 Public Domain

Imagine you're tasked with sending a team of football players onto a field to assess the condition of the grass (a likely task for them, of course). If you pick their positions randomly, they might cluster together in some areas while completely neglecting others. But if you give them a strategy, like spreading out uniformly across the field, you might get a far more accurate picture of the grass condition.

Now, imagine needing to spread out not just in two dimensions, but across tens or even hundreds. That's the challenge MIT CSAIL researchers are getting ahead of. They've developed an AI-driven approach to "low-discrepancy sampling," a method that improves simulation accuracy by distributing data points more uniformly across space.

A key novelty lies in using Graph Neural Networks (GNNs), which allow points to "communicate" and self-optimize for better uniformity. Their approach marks a pivotal enhancement for simulations in fields like robotics, finance, and computational science, particularly in handling complex, multi-dimensional problems critical for accurate simulations and numerical computations.

"In many problems, the more uniformly you can spread out points, the more accurately you can simulate complex systems," says T. Konstantin Rusch, lead author of the new paper and MIT CSAIL postdoctoral associate. "We've developed a method called Message-Passing Monte Carlo (MPMC) to generate uniformly spaced points, using geometric deep learning techniques.

"This further allows us to generate points that emphasize dimensions which are particularly important for a problem at hand, a property that is highly important in many applications. The model's underlying Graph Neural Networks lets the points 'talk' with each other, achieving far better uniformity than previous methods."

Their work will be [published](#) in the September issue of the *Proceedings of the National Academy of Sciences*.

Take me to Monte Carlo

The idea of Monte Carlo methods is to learn about a system by simulating it with random sampling. Sampling is the selection of a subset of a population to estimate characteristics of the whole population. Historically, it's already been used in the 18th century, when mathematician Pierre-Simon Laplace employed it to estimate the population of France without having to count each individual.

Low-discrepancy sequences, which are sequences with low discrepancy, i.e., high uniformity, such as Sobol, Halton, and Niederreiter have long been the gold standard for quasi-random sampling, which exchanges random sampling with low-discrepancy sampling. They are widely used in fields like [computer graphics](#) and computational finance, for everything from pricing options to risk assessment, where uniformly filling spaces with points can lead to more accurate results.

The MPMC framework suggested by the team transforms random samples into points with high uniformity. This is done by processing the random samples with a GNN that minimizes a specific discrepancy measure.

One big challenge of using AI for generating highly uniform points is that the usual way to measure point uniformity is very slow to compute and hard to work with. To solve this, the team switched to a quicker and more flexible uniformity measure called L2-discrepancy.

For high-dimensional problems, where this method isn't enough on its own, they use a novel technique that focuses on important lower-dimensional projections of the points. This way, they can create point

sets that are better suited for specific applications.

The implications extend far beyond academia, the team says. In computational finance, for example, simulations rely heavily on the quality of the sampling points.

"With these types of methods, random points are often inefficient, but our GNN-generated low-discrepancy points lead to higher precision," says Rusch. "For instance, we considered a classical problem from computational finance in 32 dimensions, where our MPMC points beat previous state-of-the-art quasi-[random sampling](#) methods by a factor of 4 to 24."

Robots in Monte Carlo

In robotics, path and motion planning often rely on sampling-based algorithms, which guide robots through real-time decision-making processes. The improved uniformity of MPMC could lead to more efficient robotic navigation and real-time adaptations for things like autonomous driving or drone technology.

"In fact, in a recent preprint, we demonstrated that our MPMC points achieve a fourfold improvement over previous low-discrepancy methods when applied to real-world robotics motion planning problems," says Rusch.

"Traditional low-discrepancy sequences were a major advancement in their time, but the world has become more complex, and the problems we're solving now often exist in 10, 20, or even 100-dimensional spaces," says Daniela Rus, CSAIL director and professor of electrical engineering and computer science (EECS).

"We needed something smarter, something that adapts as the

dimensionality grows. GNNs are a paradigm shift in how we generate low-discrepancy point sets. Unlike traditional methods, where points are generated independently, GNNs allow points to 'chat' with one another so the network learns to place points in a way that reduces clustering and gaps—common issues with typical approaches."

More information: T. Konstantin Rusch et al, Message-Passing Monte Carlo: Generating low-discrepancy point sets via graph neural networks, *Proceedings of the National Academy of Sciences* (2024). [DOI: 10.1073/pnas.2409913121](https://doi.org/10.1073/pnas.2409913121)

Provided by Massachusetts Institute of Technology