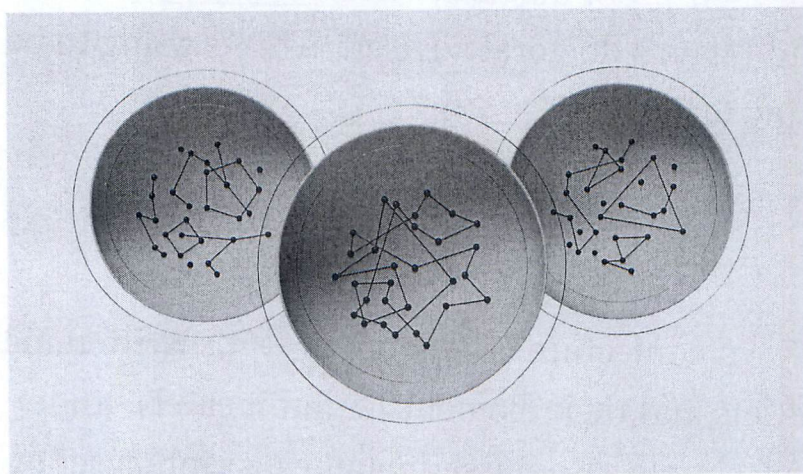


Elegant Six-Page Proof Reveals the Emergence of Random Structure

By Jordana Cepelewicz

April 25, 2022

Two young mathematicians have astonished their colleagues with a full proof of the Kahn-Kalai conjecture — a sweeping statement about how structure emerges in random sets and graphs.



Will a random graph result in a triangle (right), a Hamiltonian cycle (center), or some other property of interest?

Olena Shmahalo for Quanta Magazine

When the mathematicians Jeff Kahn and Gil Kalai first posed their “expectation threshold” conjecture in 2006, they didn’t believe it themselves. Their claim — a broad assertion about mathematical objects called random graphs — seemed too strong, too all-encompassing, too bold to possibly be true. It felt more like wishful thinking than a reflection of mathematical truth. Even so, no one could prove it false, and it quickly became one of the most important open problems in the field.

Now, more than 15 years later, a pair of young mathematicians at Stanford University have done what Kahn and Kalai thought borderline impossible:

In a surprisingly short preprint posted online just a few weeks ago, Jinyoung Park and Huy Tuan Pham have provided a complete proof of the conjecture.

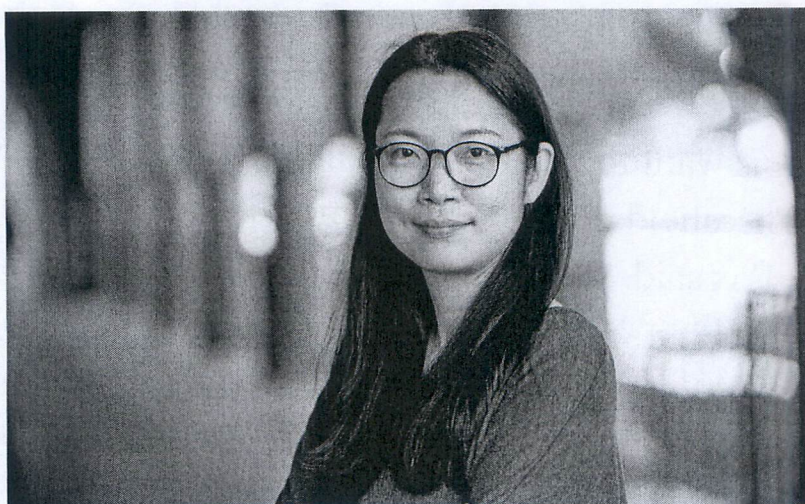
“It’s strikingly simple and ingenious,” said Kalai. “It’s stunning. It’s wonderful.”

The result automatically proves hundreds of more specific statements that would each be very difficult to prove on their own — and it has even deeper consequences for our understanding of random graphs and mathematical sets more broadly.

“I would call their proof magical,” said Jacob Fox, a mathematician at Stanford and Pham’s doctoral adviser. “This is going to be a major part of the field going forward.”

Freezing a Graph

The Kahn-Kalai conjecture is very broad — it’s written in the abstract language of sets and their elements — but it can be understood by considering a simple case. First, imagine a graph: a set of points, or vertices, connected by lines, or edges. To make a random graph, take a biased coin — one that lands on heads with a probability of 1%, or 30%, or any other percentage between zero and 100 — and flip it once for a given pair of vertices. If the coin lands on heads, connect those vertices with an edge; if the coin lands on tails, don’t. Repeat this process for every possible pair of vertices.



Jinyoung Park, a mathematician at Stanford University, “could feel the beauty and the strength of this conjecture,” she said. “But I never imagined that I would prove it.”

Rod Searcey

Mathematicians want to know when such a graph is likely to have some sort of interesting structure. Perhaps it will contain a triangle. Or maybe it will have a Hamiltonian cycle, a chain of edges that passes through every vertex exactly once. It’s possible to think about any property, so long as it is “increasing” — that is, if adding more edges to a graph that already contains the property will not destroy the property.

If the probability of the coin turning up heads is low, edges will be rare, and properties like Hamiltonian cycles are not likely to arise. But if you dial up the probability, something strange happens. Each property has what’s called a threshold: a probability at which the structure emerges, often very abruptly.

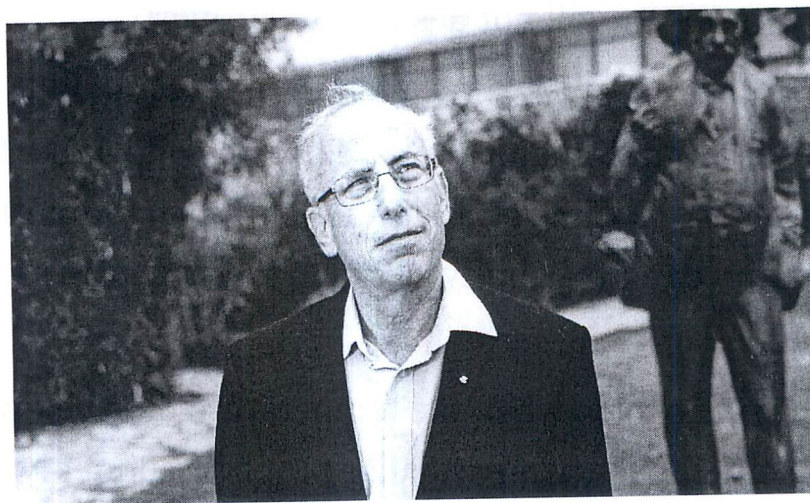
Just as ice crystals form when the temperature dips below zero degrees Celsius, the emergence of a particular property suddenly becomes extremely likely as more edges get added to the graph. When edges are added to a random graph of N vertices with a probability of less than $\log(N)/N$, for instance, the graph is unlikely to contain a Hamiltonian cycle.

But when that probability is adjusted to be just a hair greater than $\log(N)/N$, a Hamiltonian cycle becomes extremely likely.

Mathematicians want to determine such thresholds for various properties of interest. “Thresholds are maybe the most basic thing you’d try to understand,” Fox said. “I look at a random object; does it have the property that I’m interested in?” Yet while the threshold has been calculated for Hamiltonian cycles and some other specific structures, in most cases it remains very difficult to determine a precise threshold, or even a good estimate of one.

So mathematicians often rely on an easier computation, one that provides a minimum possible value, or lower bound, for the threshold. This “expectation threshold” is calculated by essentially taking a weighted average. “The nice thing about this expectation threshold is it’s very easy to calculate,” said David Conlon, a mathematician at the California Institute of Technology. “Generally speaking, you can calculate this expectation threshold in like two lines for almost anything.”

But averages can be misleading. For Hamiltonian cycles, for example, the expectation threshold is $1/N$, which is lower than the true value of $\log(N)/N$ by a factor of $\log(N)$.



Gil Kalai at the Hebrew University of Jerusalem.

Daniel Vaaknin for Quanta Magazine

In 2006, Kahn and Kalai posited that this was actually the worst-case scenario. Their eponymous conjecture states that the gap between the expectation threshold and the true threshold will never be greater than a logarithmic factor. The conjecture, according to Conlon, “essentially takes what is the central question in random graphs and gives a general answer for it.”

But that’s just a simple case. The conjecture pertains to far more than random graphs. If true, it holds for random sequences of numbers, for generalizations of graphs called hypergraphs, and for even broader types of systems. That’s because Kahn and Kalai wrote their statement in terms of abstract sets. Random graphs constitute one specific case — a random graph can be thought of as a random subset of the set of all possible edges — but there are many other objects that fall within the conjecture’s purview. “Weirdly, when you’re dealing with graphs, proving it in that context would be very hard,” Conlon said. “But somehow, jumping to this abstract setting reveals the navel of the thing.”

It was this generality that made the statement seem so unbelievable. “It was a very brave conjecture,” said Shachar Lovett, a theoretical computer scientist at the University of California, San Diego. For one thing, it would instantaneously streamline a huge effort in combinatorics — trying to calculate thresholds for different properties. “Questions where seemingly the proofs needed to be very long and complicated suddenly just disappear,” said Alan Frieze, a mathematician at Carnegie Mellon University. “The proofs became just trivial applications of this [conjecture].”

That so many seemingly unrelated problems could be settled by such a broad conjecture felt like a stretch to many mathematicians. “It seemed completely crazy, to be honest,” said Conlon. After devising their conjecture, Kahn and Kalai didn’t try to prove it. They worked to find a counterexample. There were so many settings they could explore, they figured they were bound to stumble across one eventually.

But as it turned out, “the story evolved in a very different way” than they’d expected, Kalai said.

The Sunflower Path

The methods that would eventually lead to the new proof of the Kahn-Kalai conjecture began with a breakthrough on a seemingly unrelated problem. In many ways, the story starts with the sunflower conjecture, a question posed by the mathematicians Paul Erdős and Richard Rado in 1960. The sunflower conjecture considers whether collections of sets can be constructed in ways that resemble the petals of a sunflower.

In 2019, Lovett was part of a team that came very close to a full solution of the sunflower problem. At the time, the work seemed completely separate from the Kahn-Kalai conjecture, which involves considerations of probability. “I didn’t see any connection with our conjecture,” said Kalai. Neither did Lovett, who said that “we weren’t aware of these [other] questions. We cared about sunflowers.”

But Kahn, Park (who was Kahn’s doctoral student at the time) and their colleagues ended up linking the two when they set out to prove a more relaxed version of the Kahn-Kalai conjecture a few months later. (Their proof was published in the *Annals of Mathematics* last year.) This weaker version, formulated by the French mathematician Michel Talagrand, replaced the Kahn-Kalai expectation threshold with a “fractional” expectation threshold — essentially a different way of taking a weighted average. The revised definition “gives you more wiggle room to work with things,” Lovett said.

Kahn and Park’s team realized they could export techniques from the 2019 sunflower result, fine-tune them, and apply them to the Talagrand conjecture. “This is certainly what got us started,” said Kahn.

The mathematicians took an iterative approach to the problem. They set out to show that if they picked a random set — say, a random graph — it would contain a structure such as a Hamiltonian cycle. But instead of

picking that random set all at once, they chose it in pieces, a process akin to the way Lovett and his colleagues approached the sunflower conjecture. “We perform some sort of random process iteratively,” Park said. “We choose some edges step by step” until they contain an entire Hamiltonian cycle.

To do this, the team turned to a notion of randomness called spread. If Hamiltonian cycles are “spread out” nicely, that means that not too many cycles contain the same edge or subset of edges. “Somehow, the collection of sets is spread out in space,” Pham said. “It’s not very clustered or concentrated on any part.” If the cycles are well distributed in this way, it guarantees that the process of random piece-by-piece containment — even if it fails for many Hamiltonian cycles — will succeed in capturing at least one.

The approach was possible only because of a critical equivalence: Spread could be quantified in a way that related directly to the fractional expectation threshold. Because of this, the mathematicians could rewrite the Talagrand conjecture in terms of spread.

Intriguingly, the proof of this weaker conjecture was enough to settle a torrent of threshold-related problems. “Every consequence that we know of for the full conjecture is also a consequence of the weak conjecture,” Kahn said. In fact, to him, Kalai and others, this suggested that the two conjectures might be more or less the same — that the values of the fractional and original expectation thresholds were basically equivalent. If someone could prove that equivalence, they’d prove the Kahn-Kalai conjecture. “I always thought the only way to prove our conjecture was to prove this,” Kahn said.

But that’s not what happened. While other mathematicians tried to follow this road map toward a full proof of the Kahn-Kalai conjecture, Park and Pham found a new approach altogether. “Jinyoung and Huy found this incredibly direct, incredibly short argument that just punches straight through all that,” Conlon said. “Which is extraordinary. I wasn’t expecting that at all.”

Kahn agreed. "This is one of the nice things that happens in mathematics," he said. "Things that people thought were hopeless turn out not only to be not hopeless, but not even hard."

A Surprise Approach

At first, neither Park nor Pham had any intention of tackling the original conjecture. Since first learning about the problem as a graduate student, Park "could feel the beauty and the strength of this conjecture," she said. "But I never imagined that I would prove it."

"It was not in our mind at all," Pham added.

Rather, they were working on another conjecture posed by Talagrand when they were "struck with a revelation," Pham said. They realized that "the picture that we have here, the ideas that we have, it somehow seemed like it's more powerful than it looked." Those ideas, they thought, might just be

powerful enough to take them all the way through a proof of the Kahn-Kalai problem.



Huy Tuan Pham.

Courtesy of Huy Tuan Pham

Over the course of a single sleepless night in March, they figured out how to make the proof work.

Unlike the fractional expectation threshold, the normal expectation threshold has no relationship with spread. Spread “gives you a starting point. And if you go to the original, non-fractional conjecture, that starting point just disappears,” Kahn said. “So it looked very tough.”

“So what do you do?” Pham said. “In this case, we switch our perspective.”

In particular, they thought about the problem in terms of a mathematical object called a cover. A cover is a collection of sets, where every object with a certain property contains one of those sets. For instance, one possible cover of all Hamiltonian cycles is the collection of all edges. Every Hamiltonian cycle will contain one of those edges.

Park and Pham rewrote the Kahn-Kalai conjecture in a way that let them make use of covers. The original conjecture puts constraints on what the probability of a weighted coin landing on heads should be in order to guarantee that a random graph or set contains some property. In

particular, it says that the probability has to be at least the expectation threshold for the property multiplied by a logarithmic factor. Park and Pham turned this around: If such a property is not likely to emerge, then the probability assigned to the weighted coin is lower than the expectation threshold multiplied by a logarithmic factor.

That's where covers come in: When a small cover can be constructed for a subset of structures (like a collection of Hamiltonian cycles), it means that the subset's contribution to the expectation threshold is small. (Remember that the expectation threshold is calculated by taking a kind of weighted average over all possible structures of a given type.) So what Park and Pham now needed to show was that if a random set is unlikely to contain one target structure, there must exist a small cover for all such target structures. The bulk of their proof was dedicated to constructing that small cover.

They did this by using a similar piece-by-piece sampling process to the one used in the previous results, while also introducing what Fox called a "very clever counting argument." One week after their sleepless night in March, they posted their elegant six-page paper online.

"Their proof is super simple. They take the basic idea we developed and [the ideas from] these other papers and add a twist to it," Lovett said. "And with this new twist, everything somehow becomes much, much easier."

Frieze agreed. "I cannot explain it, but amazingly it's true," he said.

Just like the fractional result, the Kahn-Kalai conjecture, now proved true, automatically implies a cornucopia of related conjectures. But more than that, "this is a powerful proof technique [that] will probably lead to new things," said Noga Alon, a mathematician at Princeton University. "They had to do it in the right way."

Park and Pham have now started to apply their method to other problems. They're particularly interested in getting a more precise understanding of the gap between the expectation threshold and the real threshold. By proving the Kahn-Kalai conjecture, they've shown that this gap is at most a

logarithmic factor — but sometimes the gap is smaller, or even nonexistent. At the moment, there's no broader mechanism for classifying when each of these scenarios might be true; mathematicians have to work it out case by case. Now, "we think that with this efficient technique we have, we can hopefully be much more precise in pinning down these thresholds," Pham said.

And their proof could have other consequences as well. "The Kahn-Kalai conjecture is not at all the end of the story," Park said.