Mathematical Landscapes and String Theory

Michael R. Douglas

[1]Simons Center / Stony Brook University

Microsoft, April 25, 2019

Abstract

Usually we think of datasets as derived from the physical world: images, waveforms, measurements, *etc.*, or the social world: text, networks of relationships, *etc.*.

But there are also "natural" or "platonic" datasets with *ab initio* definitions. The primary examples are lists of mathematical objects: finite groups, topological classes of manifolds, graphs with special properties, *etc.*. Such datasets also arise from fundamental physical theories, *e.g.* lists of solutions in quantum chemistry, and lists of vacua of superstring theory.

We refer to the structures underlying such data as "mathematical landscapes," both for their permanence and because of the important roles of energy and complexity measures in their study.

In this talk we introduce a few of these datasets and discuss how ML and other computational methods can be fruitfully applied to their study.

Mathematics is the classification and study of all possible patterns. (Prelude to Mathematics, W. W. Sawyer, 1955)

Mathematics is the art of giving the same name to different things. (Science and Method, Henri Poincaré, 1914)

Some patterns are more prevalent and more fruitful to notice than others. They are the ones which are easy to define precisely, which appear in many contexts, and for which we can develop a rich theoretical description.

One of the most ubiquitous mathematical concepts is the group. Much is known about groups, and there are developed systems for computational group theory such as GAP and Magma. Rederiving this knowledge in simpler and more accessible ways could be a valuable benchmark problem for the use of AI and ML in mathematics.

We then discuss classification of the possibilities for the extra dimensions of string theory

Mathematics is the classification and study of all possible patterns. (Prelude to Mathematics, W. W. Sawyer, 1955)

Mathematics is the art of giving the same name to different things. (Science and Method, Henri Poincaré, 1914)

Some patterns are more prevalent and more fruitful to notice than others. They are the ones which are easy to define precisely, which appear in many contexts, and for which we can develop a rich theoretical description.

One of the most ubiquitous mathematical concepts is the group. Much is known about groups, and there are developed systems for computational group theory such as GAP and Magma. Rederiving this knowledge in simpler and more accessible ways could be a valuable benchmark problem for the use of AI and ML in mathematics.

We then discuss classification of the possibilities for the extra dimensions of string theory

Mathematics is the classification and study of all possible patterns. (Prelude to Mathematics, W. W. Sawyer, 1955)

Mathematics is the art of giving the same name to different things. (Science and Method, Henri Poincaré, 1914)

Some patterns are more prevalent and more fruitful to notice than others. They are the ones which are easy to define precisely, which appear in many contexts, and for which we can develop a rich theoretical description.

One of the most ubiquitous mathematical concepts is the group. Much is known about groups, and there are developed systems for computational group theory such as GAP and Magma. Rederiving this knowledge in simpler and more accessible ways could be a valuable benchmark problem for the use of AI and ML in mathematics.

We then discuss classification of the possibilities for the extra dimensions of string theory

Mathematics is the classification and study of all possible patterns. (Prelude to Mathematics, W. W. Sawyer, 1955)

Mathematics is the art of giving the same name to different things. (Science and Method, Henri Poincaré, 1914)

Some patterns are more prevalent and more fruitful to notice than others. They are the ones which are easy to define precisely, which appear in many contexts, and for which we can develop a rich theoretical description.

One of the most ubiquitous mathematical concepts is the group. Much is known about groups, and there are developed systems for computational group theory such as GAP and Magma. Rederiving this knowledge in simpler and more accessible ways could be a valuable benchmark problem for the use of AI and ML in mathematics.

We then discuss classification of the possibilities for the extra dimensions of string theory

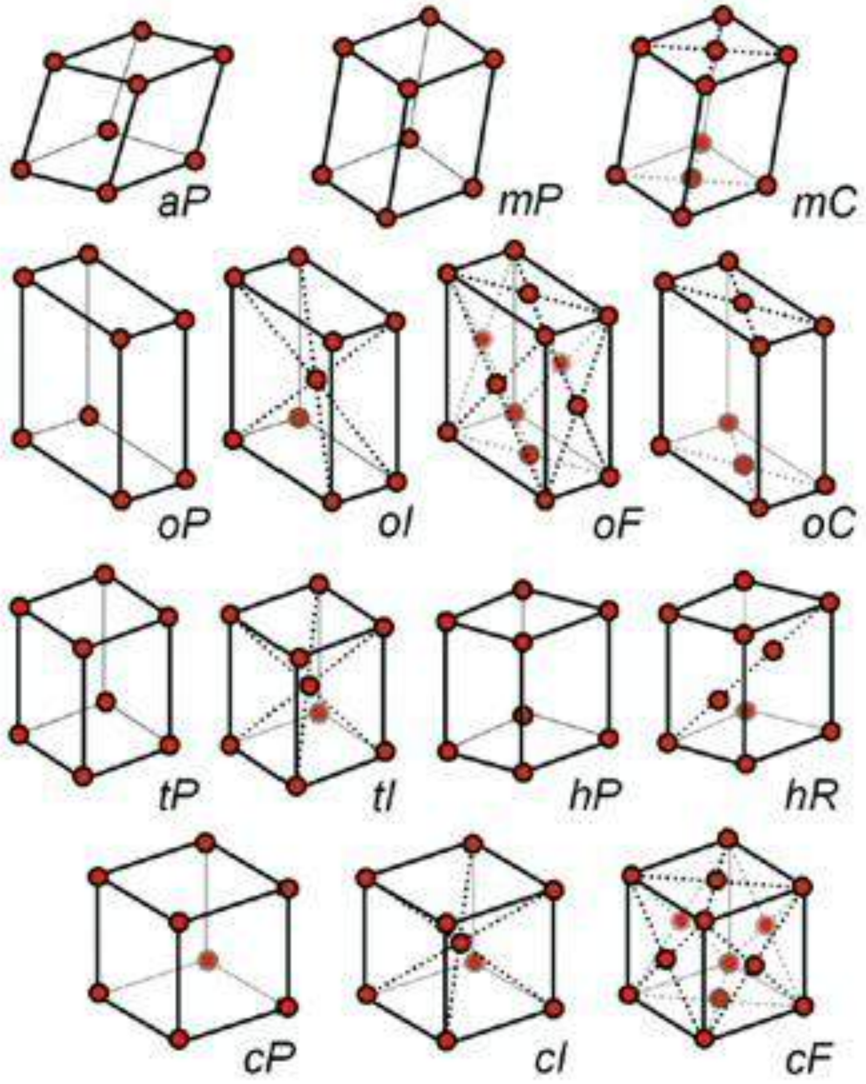
Groups

Symmetries of mathematical structures are naturally described by group actions.

Example: d-dimensional Euclidean space, and its isometry group $ISO(n)$ (rigid motions).

Example: tilings of a plane or a volume, crystal lattices: space groups $\equiv$ discrete subgroups of $ISO(n)$.

Example: algebraic equations: Galois groups, *e.g.* $\sqrt[p]{x} \to e^{2\pi i/p} \sqrt[p]{x}$.



The classification of groups and their actions has been a central mathematical question since the late 19th century. Two highlights of this project were the classification of the finite dimensional simple Lie groups, and the classification of the finite simple groups.

A group is simple if it has no nontrivial quotient groups (or normal subgroups). For example, a space group is not simple because the translations form a normal subgroup. All finite groups can be built up from simple groups.

A group can act on other objects in many ways – it can have many representations. For example, the symmetric group S_n acts on a set of n elements by permutation. This is a very special case of a linear representation: a homomorphism $G \rightarrow \mathrm{Mat}_n$.

The classification of groups and their actions has been a central mathematical question since the late 19th century. Two highlights of this project were the classification of the finite dimensional simple Lie groups, and the classification of the finite simple groups.

A group is simple if it has no nontrivial quotient groups (or normal subgroups). For example, a space group is not simple because the translations form a normal subgroup. All finite groups can be built up from simple groups.

A group can act on other objects in many ways – it can have many representations. For example, the symmetric group S_n acts on a set of n elements by permutation. This is a very special case of a linear representation: a homomorphism $G \to \mathrm{Mat}_n$.

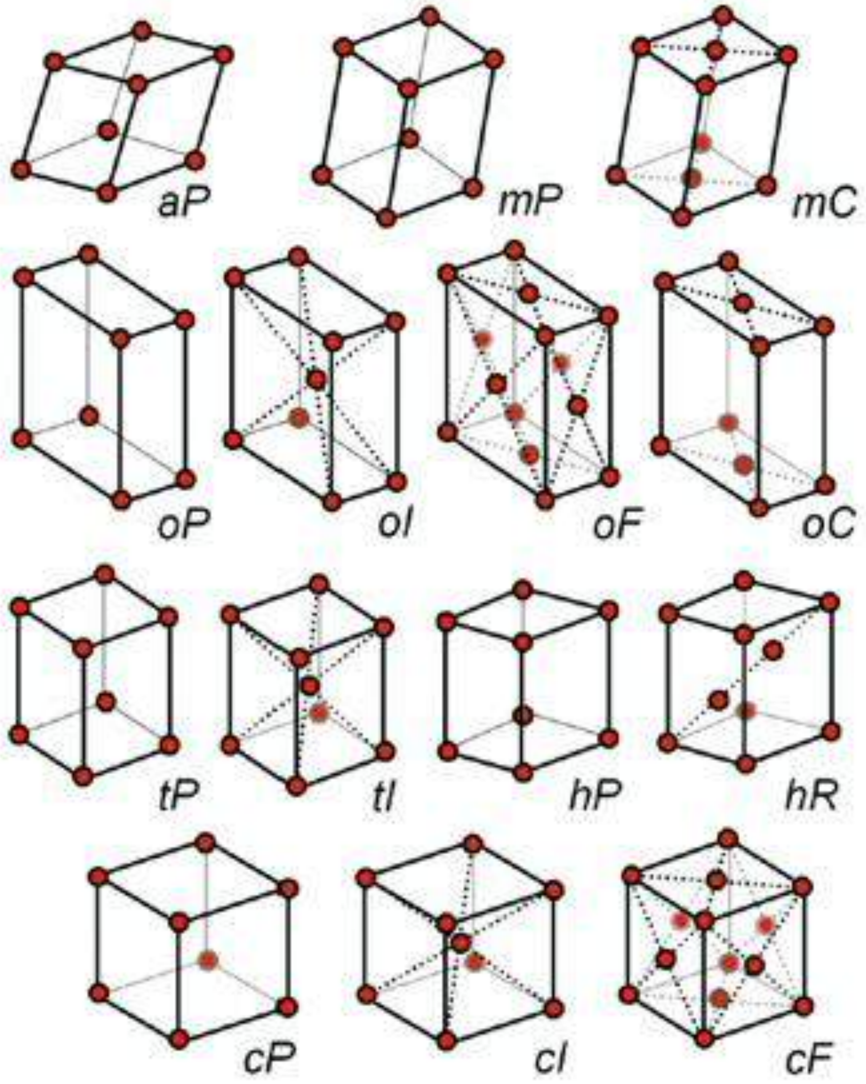
Groups

Symmetries of mathematical structures are naturally described by group actions.

Example: d-dimensional Euclidean space, and its isometry group $ISO(n)$ (rigid motions).

Example: tilings of a plane or a volume, crystal lattices: space groups $\equiv$ discrete subgroups of $ISO(n)$.

Example: algebraic equations: Galois groups, *e.g.* $\sqrt[p]{x} \to e^{2\pi i/p} \sqrt[p]{x}$.



The classification of groups and their actions has been a central mathematical question since the late 19th century. Two highlights of this project were the classification of the finite dimensional simple Lie groups, and the classification of the finite simple groups.

A group is simple if it has no nontrivial quotient groups (or normal subgroups). For example, a space group is not simple because the translations form a normal subgroup. All finite groups can be built up from simple groups.

A group can act on other objects in many ways – it can have many representations. For example, the symmetric group S_n acts on a set of n elements by permutation. This is a very special case of a linear representation: a homomorphism $G \to \mathrm{Mat}_n$.

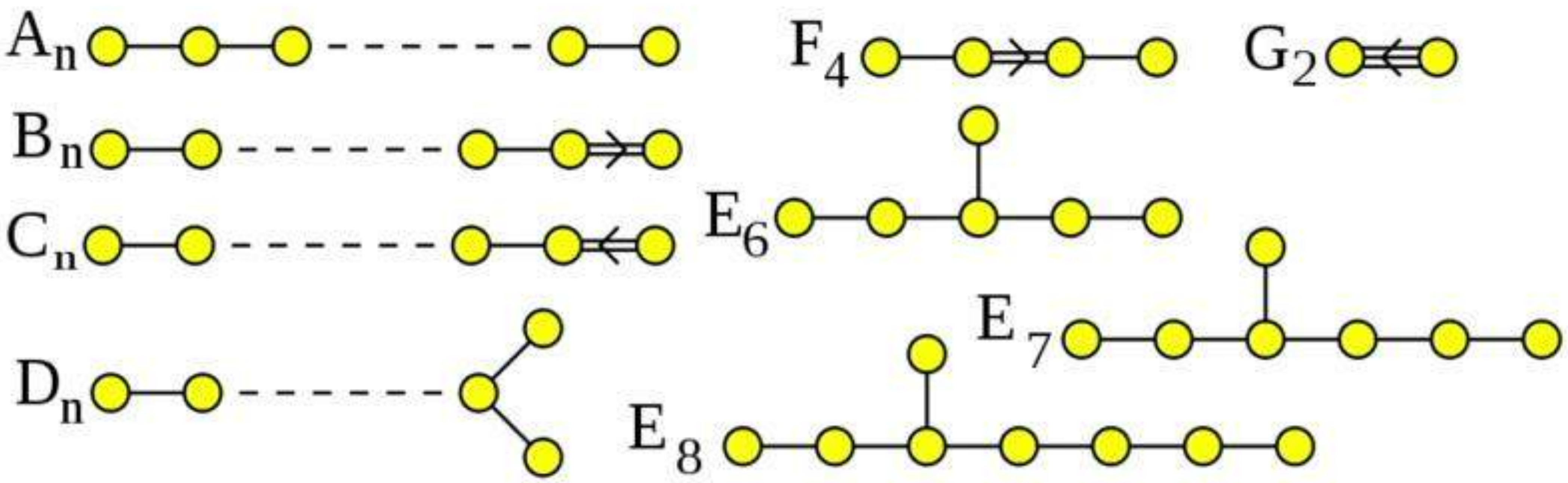
The classification of groups and their actions has been a central mathematical question since the late 19th century. Two highlights of this project were the classification of the finite dimensional simple Lie groups, and the classification of the finite simple groups.

A group is simple if it has no nontrivial quotient groups (or normal subgroups). For example, a space group is not simple because the translations form a normal subgroup. All finite groups can be built up from simple groups.

A group can act on other objects in many ways – it can have many representations. For example, the symmetric group S_n acts on a set of n elements by permutation. This is a very special case of a linear representation: a homomorphism $G \to \mathrm{Mat}_n$.

Semisimple Lie groups

These are manifolds (they have continuous parameters) and are the bread and butter of theoretical physics, describing symmetries of space-time, gauge groups in Yang-Mills theory, *etc.*. Their classification can be reduced to that of their infinitesimal symmetry generators (Lie algebras) and was done in the early 20th century. The results are conveniently described by Dynkin diagrams:



Classification of finite simple groups

This is much more intricate and was only completed in 2004. Even a simplified proof will be $>$ 5000 pages long. We quote from Wikipedia:

Theorem — Every finite simple group is isomorphic to one of the following groups:

- a member of one of three infinite classes of such, namely:
  - the cyclic groups of prime order,
  - the alternating groups of degree at least 5,
  - the groups of Lie type
- one of 26 groups called the "sporadic groups"
- the Tits group (which is sometimes considered a 27th sporadic group).

The largest of the sporadic groups is the Fischer-Greiss monster group with $\sim 8 \times 10^{53}$ elements. It has no simple description. The most conceptual is as the automorphism group of the “monster vertex algebra” constructed by Frenkel-Lepowsky-Meurman. This has a reasonably simple explanation in string theory terms, suggesting some important though still mysterious role for the monster in string theory.

Both of these sets (Lie groups, finite groups) are "mathematical landscapes" in the sense of this talk. They have precise definitions, and their classification and elaboration leads to large datasets. This is clear for finite groups, and is also the case for representations of Lie algebras and groups.

These are well developed theories, as is computational group theory. Is there any role for AI/ML ? As an example for a single group, Rubik's cube was recently solved by RL (McAleer *et al*, `cs.AI/1805.07470`).

This problem can be generalized. Given a set of group generators $\{g_i\}$, the Cayley graph of a group has a vertex for each element, and an edge between each pair of vertices related by $v_1 = g_i v_2$ for some i. So the problem is to learn how to find the identity from a random starting point. It might be interesting to see for which groups this can be solved by RL.

But a real "landscape" problem would explore the set of all groups. Let us come back to this.

Both of these sets (Lie groups, finite groups) are "mathematical landscapes" in the sense of this talk. They have precise definitions, and their classification and elaboration leads to large datasets. This is clear for finite groups, and is also the case for representations of Lie algebras and groups.

These are well developed theories, as is computational group theory. Is there any role for AI/ML ? As an example for a single group, Rubik's cube was recently solved by RL (McAleer *et al*, `cs.AI/1805.07470`).

This problem can be generalized. Given a set of group generators $\{g_i\}$, the Cayley graph of a group has a vertex for each element, and an edge between each pair of vertices related by $v_1 = g_i v_2$ for some i. So the problem is to learn how to find the identity from a random starting point. It might be interesting to see for which groups this can be solved by RL.

But a real "landscape" problem would explore the set of all groups. Let us come back to this.

Both of these sets (Lie groups, finite groups) are "mathematical landscapes" in the sense of this talk. They have precise definitions, and their classification and elaboration leads to large datasets. This is clear for finite groups, and is also the case for representations of Lie algebras and groups.

These are well developed theories, as is computational group theory. Is there any role for AI/ML ? As an example for a single group, Rubik's cube was recently solved by RL (McAleer *et al*, `cs.AI/1805.07470`).

This problem can be generalized. Given a set of group generators $\{g_i\}$, the Cayley graph of a group has a vertex for each element, and an edge between each pair of vertices related by $v_1 = g_i v_2$ for some i. So the problem is to learn how to find the identity from a random starting point. It might be interesting to see for which groups this can be solved by RL.

But a real "landscape" problem would explore the set of all groups. Let us come back to this.

Both of these sets (Lie groups, finite groups) are "mathematical landscapes" in the sense of this talk. They have precise definitions, and their classification and elaboration leads to large datasets. This is clear for finite groups, and is also the case for representations of Lie algebras and groups.

These are well developed theories, as is computational group theory. Is there any role for AI/ML ? As an example for a single group, Rubik's cube was recently solved by RL (McAleer *et al*, `cs.AI/1805.07470`).

This problem can be generalized. Given a set of group generators $\{g_i\}$, the Cayley graph of a group has a vertex for each element, and an edge between each pair of vertices related by $v_1 = g_i v_2$ for some i. So the problem is to learn how to find the identity from a random starting point. It might be interesting to see for which groups this can be solved by RL.

But a real "landscape" problem would explore the set of all groups. Let us come back to this.

Solutions of physical theories

In a fundamental physical theory, there is a precise mathematical definition of all possible physical states and histories. This could be in terms of an ODE (Newton's equations), a PDE (Maxwell's equations; the Schrödinger equation), a path integral (for quantum field theory), or even more general concepts (string/M theory).

Arguably, the most important general class of physical landscapes are the energy landscapes of quantum chemistry and material science. These are defined in terms of the Schrödinger equation,

$$E\psi = \left[-\frac{\hbar^2}{2m_e} \sum_i \frac{\partial^2}{\partial \vec{x}_i \cdot \partial \vec{x}_i} + \sum_{i<j} \frac{e^2}{|\vec{x}_i - \vec{x}_j|} - \sum_{i,a} \frac{Q_a e^2}{|\vec{x}_i - \vec{X}_a|} \right] \psi.$$

Typically one solves for the quantum state of the electrons as a function of the positions of the nuclei to get an effective potential , and then studies its metastable minima, transitions between them, *etc.*

Solutions of physical theories

In a fundamental physical theory, there is a precise mathematical definition of all possible physical states and histories. This could be in terms of an ODE (Newton's equations), a PDE (Maxwell's equations; the Schrödinger equation), a path integral (for quantum field theory), or even more general concepts (string/M theory).

Arguably, the most important general class of physical landscapes are the energy landscapes of quantum chemistry and material science. These are defined in terms of the Schrödinger equation,

$$E\psi = \left[-\frac{\hbar^2}{2m_e} \sum_i \frac{\partial^2}{\partial \vec{x}_i \cdot \partial \vec{x}_i} + \sum_{i<j} \frac{e^2}{|\vec{x}_i - \vec{x}_j|} - \sum_{i,a} \frac{Q_a e^2}{|\vec{x}_i - \vec{X}_a|} \right] \psi .$$

Typically one solves for the quantum state of the electrons as a function of the positions of the nuclei to get an effective potential , and then studies its metastable minima, transitions between them, *etc.*

Working out the structures of molecules and materials *ab initio* from quantum physics is very challenging. Fortunately we have a vast amount of data.
Many researchers are applying machine learning to these problems. For example, one can apply graph learning techniques to the structural formulas of molecules, and do supervised learning of energies, favored reaction paths, *etc.*.
Still, except for atoms and simple materials, quantum chemistry is generally too special and complicated to lead to interesting mathematics.

In the rest of this talk, we will describe some ingredients of the analogous *ab initio* calculation of the laws of our universe, from the fundamental laws of string/M theory. The prospects for doing this are arguably much better than for *ab initio* quantum chemistry. We already have constructions which at least qualitatively reproduce the "standard models" both of particle physics and of cosmology. Furthermore these can come out of relatively simple dynamics of quantum cosmology.

Working out the structures of molecules and materials *ab initio* from quantum physics is very challenging. Fortunately we have a vast amount of data.
Many researchers are applying machine learning to these problems.
For example, one can apply graph learning techniques to the structural formulas of molecules, and do supervised learning of energies, favored reaction paths, *etc.*.
Still, except for atoms and simple materials, quantum chemistry is generally too special and complicated to lead to interesting mathematics.

In the rest of this talk, we will describe some ingredients of the analogous *ab initio* calculation of the laws of our universe, from the fundamental laws of string/M theory. The prospects for doing this are arguably much better than for *ab initio* quantum chemistry. We already have constructions which at least qualitatively reproduce the "standard models" both of particle physics and of cosmology. Furthermore these can come out of relatively simple dynamics of quantum cosmology.

Solutions of physical theories

In a fundamental physical theory, there is a precise mathematical definition of all possible physical states and histories. This could be in terms of an ODE (Newton's equations), a PDE (Maxwell's equations; the Schrödinger equation), a path integral (for quantum field theory), or even more general concepts (string/M theory).

Arguably, the most important general class of physical landscapes are the energy landscapes of quantum chemistry and material science. These are defined in terms of the Schrödinger equation,

$$E\psi = \left[-\frac{\hbar^2}{2m_e} \sum_i \frac{\partial^2}{\partial \vec{x}_i \cdot \partial \vec{x}_i} + \sum_{i<j} \frac{e^2}{|\vec{x}_i - \vec{x}_j|} - \sum_{i,a} \frac{Q_a e^2}{|\vec{x}_i - \vec{X}_a|} \right] \psi.$$

Typically one solves for the quantum state of the electrons as a function of the positions of the nuclei to get an effective potential , and then studies its metastable minima, transitions between them, *etc.*

Working out the structures of molecules and materials *ab initio* from quantum physics is very challenging. Fortunately we have a vast amount of data.
Many researchers are applying machine learning to these problems.
For example, one can apply graph learning techniques to the structural formulas of molecules, and do supervised learning of energies, favored reaction paths, *etc.*.
Still, except for atoms and simple materials, quantum chemistry is generally too special and complicated to lead to interesting mathematics.

In the rest of this talk, we will describe some ingredients of the analogous *ab initio* calculation of the laws of our universe, from the fundamental laws of string/M theory. The prospects for doing this are arguably much better than for *ab initio* quantum chemistry. We already have constructions which at least qualitatively reproduce the “standard models” both of particle physics and of cosmology. Furthermore these can come out of relatively simple dynamics of quantum cosmology.

String theory

To make a long story short, string/M theory is by far our best candidate for a theory of quantum gravity unified with Yang-Mills theory, but is defined in ten and eleven space-time dimensions.

One can take six or seven of the dimensions to be a small compact manifold *M*, leading to an effective theory in four-dimensional space-time. Good arguments have been given that there are choices of *M* and associated extra data which lead to the Standard Model of particle physics, starting with Candelas *et al* in 1985. Much progress has been made on this picture, and while the claims have borne up, the number of possibilities for *M* and the extra data are combinatorially large. There are even physical arguments that this should be expected, based on the anthropic principle.

Developing a precise yet understandable picture of this set of possibilities is a key step towards understanding whether string theory can describe our universe and how to test this claim.

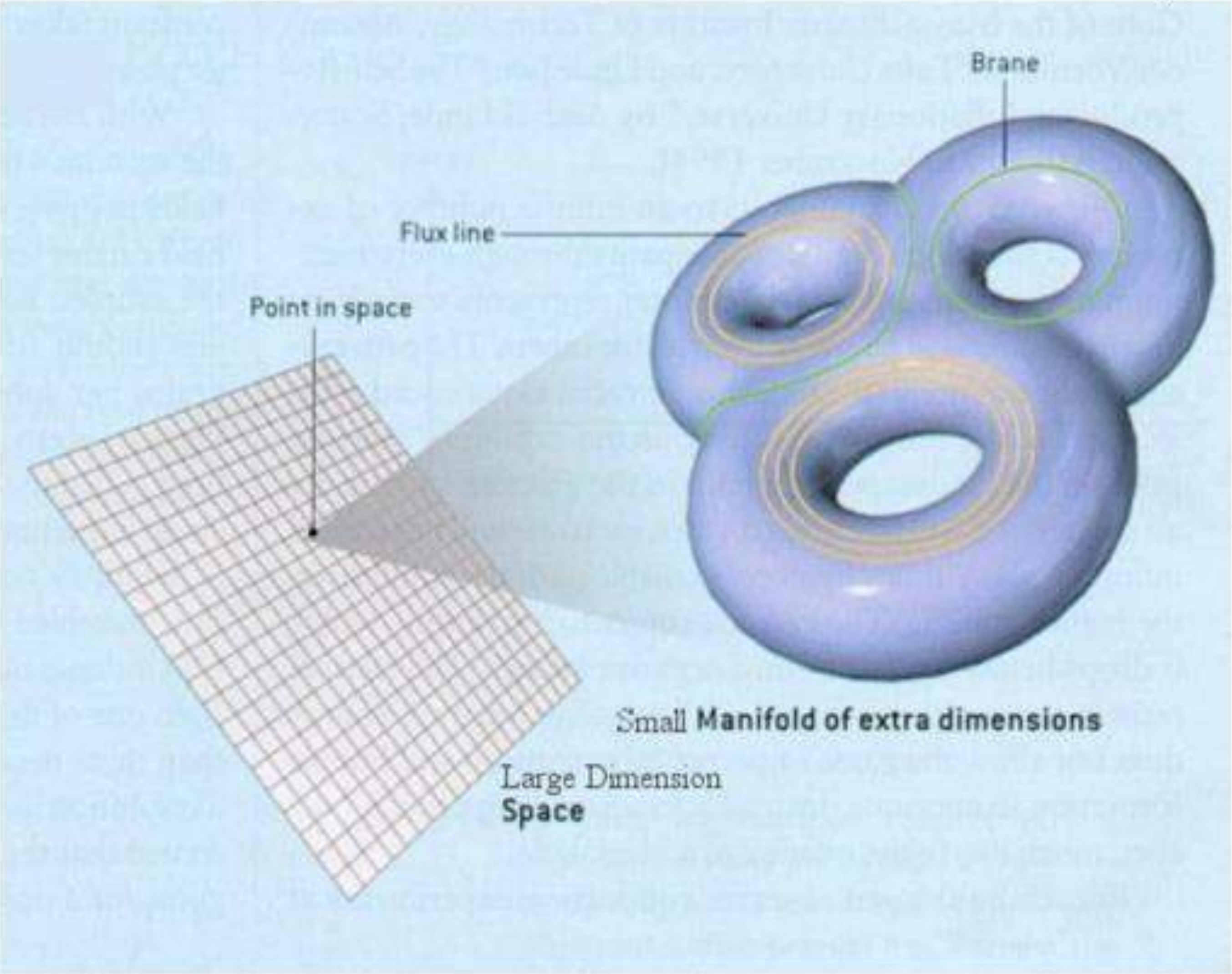
String theory

To make a long story short, string/M theory is by far our best candidate for a theory of quantum gravity unified with Yang-Mills theory, but is defined in ten and eleven space-time dimensions.

One can take six or seven of the dimensions to be a small compact manifold *M*, leading to an effective theory in four-dimensional space-time. Good arguments have been given that there are choices of *M* and associated extra data which lead to the Standard Model of particle physics, starting with Candelas *et al* in 1985. Much progress has been made on this picture, and while the claims have borne up, the number of possibilities for *M* and the extra data are combinatorially large. There are even physical arguments that this should be expected, based on the anthropic principle.

Developing a precise yet understandable picture of this set of possibilities is a key step towards understanding whether string theory can describe our universe and how to test this claim.

There are many possible choices for M – it might be a six-torus, a six-torus quotiented by a discrete group (an "orbifold"), a Calabi-Yau manifold, or many other possibilities. In addition there is additional data (branes, fluxes, *etc.*) to be chosen on M, call such a choice V (we will be a little bit more concrete below).



Each (M, V) can lead to different predictions for observed physics. We do not know *a priori* which of these choices describes our universe. Thus, we need to work out the predictions for many such choices.

Toroidal and orbifold compactifications

Almost all quasi-realistic string compactifications start out with a Ricci-flat manifold M, *i.e.* one which solves the vacuum Einstein equations. Otherwise, the curvature of M produces a huge contribution to the vacuum energy.

The simplest way to be Ricci-flat is to be flat. And there is a canonical flat manifold T^d in each dimension, the torus. It can be defined by starting from Euclidean space $\mathbb{R}^d$ and identifying $\vec{x} \cong \vec{x} + \vec{e}_i$ for d independent vectors $\vec{e}_a \in \mathbb{R}^d$, $1 \leq a \leq d$. This produces a very special and symmetric compactification, too much so to describe our universe.

Fortunately there is a simple variation on this, which is to quotient T^d by additional symmetries. For example one could identify $x \cong -x$ to get the orbifold $T^d/\mathbb{Z}_2$. This works for $d = 4$ but not in $d = 6$, as can be seen from holonomy. But one can use $T^6/\mathbb{Z}_2 \times \mathbb{Z}_2$ where each $\mathbb{Z}_2$ flips four signs.

Toroidal and orbifold compactifications

Almost all quasi-realistic string compactifications start out with a Ricci-flat manifold M, *i.e.* one which solves the vacuum Einstein equations. Otherwise, the curvature of M produces a huge contribution to the vacuum energy.

The simplest way to be Ricci-flat is to be flat. And there is a canonical flat manifold T^d in each dimension, the torus. It can be defined by starting from Euclidean space $\mathbb{R}^d$ and identifying $\vec{x} \cong \vec{x} + \vec{e}_i$ for d independent vectors $\vec{e}_a \in \mathbb{R}^d$, $1 \leq a \leq d$. This produces a very special and symmetric compactification, too much so to describe our universe.

Fortunately there is a simple variation on this, which is to quotient T^d by additional symmetries. For example one could identify $x \cong -x$ to get the orbifold $T^d/\mathbb{Z}_2$. This works for $d = 4$ but not in $d = 6$, as can be seen from holonomy. But one can use $T^6/\mathbb{Z}_2 \times \mathbb{Z}_2$ where each $\mathbb{Z}_2$ flips four signs.

Toroidal and orbifold compactifications

Almost all quasi-realistic string compactifications start out with a Ricci-flat manifold M, *i.e.* one which solves the vacuum Einstein equations. Otherwise, the curvature of M produces a huge contribution to the vacuum energy.

The simplest way to be Ricci-flat is to be flat. And there is a canonical flat manifold T^d in each dimension, the torus. It can be defined by starting from Euclidean space $\mathbb{R}^d$ and identifying $\vec{x} \cong \vec{x} + \vec{e}_i$ for d independent vectors $\vec{e}_a \in \mathbb{R}^d$, $1 \leq a \leq d$. This produces a very special and symmetric compactification, too much so to describe our universe.

Fortunately there is a simple variation on this, which is to quotient T^d by additional symmetries. For example one could identify $x \cong -x$ to get the orbifold $T^d/\mathbb{Z}_2$. This works for $d = 4$ but not in $d = 6$, as can be seen from holonomy. But one can use $T^6/\mathbb{Z}_2 \times \mathbb{Z}_2$ where each $\mathbb{Z}_2$ flips four signs.

Following up this idea, and again to make a long story short, one obtains the $T^6/\mathbb{Z}_2 \times \mathbb{Z}_2$ compactifications of type IIa superstring theory. This was worked out in a series of papers starting in the 90's and by now there is a list of data and constraints which uniquely specify such a compactification, and ways to compute the spectrum and couplings of the resulting four-dimensional theory.

Without going into all the details, here is the flavor of the problem. The bulk of the data is a list of vectors of six integers with multiplicities N_a,

$$\oplus_a N_a(n_a^1, m_a^1, n_a^2, m_a^2, n_a^3, m_a^3).$$

Each vector specifies a 3-dimensional hyperplane in T^6, into which we embed a D6-brane. These vectors must satisfy a fixed set of nonlinear constraints, some with undetermined parameters. They are linear in terms of cubic monomials $X^0 \sim n^1 n^2 n^3$, $X^1 \sim n^1 m^2 m^3$, *etc.*, so

$$T^I = \sum_a N_a X_a^I \,\forall I; \qquad \exists U^I \; s.t. \sum_I X_a^I U^I = 0 \;\forall a; \;\; \cdots .$$

The 4D Yang-Mills gauge group and matter is then (up to some $\mathbb{Z}_2$'s)

$$\otimes_a U(N_a); \qquad \oplus_{a,b} v_a \times v_b (N_a, \bar{N}_b)$$

where the cross product is $v_a \times v_b = \prod_{i=1,2,3} (n_a^i m_b^i - m_a^i n_b^i)$.

The solutions of this system of constraints are another "mathematical landscape" which can be classified and understood. Physicists then ask questions such as

- Can we realize the Standard Model gauge group and matter, *i.e.* $N_1 = 1$, $N_2 = 2$, $N_3 = 3$, $15N_{gen} + 4$ specific matter fields, limits on exotic matter charged in any $(N_{1,2,3}, \bar{N}_{i>3})$.
- What fraction of configurations realize it? What is the distribution of the number of generations?
- What is the distribution of extra matter? Is the extra matter consistent with experimental constraints? If so, can we propose a new experiment to detect it?

And it can be generalized, for example T^6 can be replaced by a Calabi-Yau threefold.

Following up this idea, and again to make a long story short, one obtains the $T^6/\mathbb{Z}_2 \times \mathbb{Z}_2$ compactifications of type IIa superstring theory. This was worked out in a series of papers starting in the 90's and by now there is a list of data and constraints which uniquely specify such a compactification, and ways to compute the spectrum and couplings of the resulting four-dimensional theory.

Without going into all the details, here is the flavor of the problem. The bulk of the data is a list of vectors of six integers with multiplicities N_a,

$$\oplus_a N_a(n_a^1, m_a^1, n_a^2, m_a^2, n_a^3, m_a^3).$$

Each vector specifies a 3-dimensional hyperplane in T^6, into which we embed a D6-brane. These vectors must satisfy a fixed set of nonlinear constraints, some with undetermined parameters. They are linear in terms of cubic monomials $X^0 \sim n^1 n^2 n^3$, $X^1 \sim n^1 m^2 m^3$, *etc.*, so

$$T^I = \sum_a N_a X_a^I \,\forall I; \qquad \exists U^I \; s.t. \sum_I X_a^I U^I = 0 \,\forall a; \quad \cdots .$$

The 4D Yang-Mills gauge group and matter is then (up to some $\mathbb{Z}_2$'s)

$$\otimes_a U(N_a); \qquad \oplus_{a,b} v_a \times v_b(N_a, \bar{N}_b)$$

where the cross product is $v_a \times v_b = \prod_{i=1,2,3}(n_a^i m_b^i - m_a^i n_b^i)$.

The solutions of this system of constraints are another "mathematical landscape" which can be classified and understood. Physicists then ask questions such as

- Can we realize the Standard Model gauge group and matter, *i.e.* $N_1 = 1$, $N_2 = 2$, $N_3 = 3$, $15N_{gen} + 4$ specific matter fields, limits on exotic matter charged in any $(N_{1,2,3}, \bar{N}_{i>3})$.
- What fraction of configurations realize it? What is the distribution of the number of generations?
- What is the distribution of extra matter? Is the extra matter consistent with experimental constraints? If so, can we propose a new experiment to detect it?

And it can be generalized, for example T^6 can be replaced by a Calabi-Yau threefold.

The 4D Yang-Mills gauge group and matter is then (up to some $\mathbb{Z}_2$'s)

$$\otimes_a U(N_a); \qquad \oplus_{a,b} v_a \times v_b (N_a, \bar{N}_b)$$

where the cross product is $v_a \times v_b = \prod_{i=1,2,3}(n_a^i m_b^i - m_a^i n_b^i)$.

The solutions of this system of constraints are another "mathematical landscape" which can be classified and understood. Physicists then ask questions such as

- Can we realize the Standard Model gauge group and matter, *i.e.* $N_1 = 1$, $N_2 = 2$, $N_3 = 3$, $15N_{gen} + 4$ specific matter fields, limits on exotic matter charged in any $(N_{1,2,3}, \bar{N}_{i>3})$.
- What fraction of configurations realize it? What is the distribution of the number of generations?
- What is the distribution of extra matter? Is the extra matter consistent with experimental constraints? If so, can we propose a new experiment to detect it?

And it can be generalized, for example T^6 can be replaced by a Calabi-Yau threefold.

Thus we want to sample solutions of a complicated combinatorial optimization problem. A simple approach would be to postulate transitions between configurations, say which vary the (n, m) data, and do a random walk. But it turns out that reinforcement learning, with a reward function encoding the constraints and the Standard Model target, is a much more efficient technique for doing this (Halverson, Nelson and Ruehle, `hep-th/1903.11616`).

We leave the details for other talks, and just raise two questions:

- Is the sampling unbiased? If it is biased, could this be meaningful, *i.e.* somehow suggesting preferred configurations?
- In quantum cosmology there is a "correct" transition network. We don't know it, but it only visits configurations which satisfy the tadpole constraints (but are not necessarily supersymmetric). It might also go through additional states. Do these details affect the dynamics?

Thus we want to sample solutions of a complicated combinatorial optimization problem. A simple approach would be to postulate transitions between configurations, say which vary the (n, m) data, and do a random walk. But it turns out that reinforcement learning, with a reward function encoding the constraints and the Standard Model target, is a much more efficient technique for doing this (Halverson, Nelson and Ruehle, `hep-th/1903.11616`).

We leave the details for other talks, and just raise two questions:

- Is the sampling unbiased? If it is biased, could this be meaningful, *i.e.* somehow suggesting preferred configurations?
- In quantum cosmology there is a "correct" transition network. We don't know it, but it only visits configurations which satisfy the tadpole constraints (but are not necessarily supersymmetric). It might also go through additional states. Do these details affect the dynamics?

The 4D Yang-Mills gauge group and matter is then (up to some $\mathbb{Z}_2$'s)

$$\otimes_a U(N_a); \qquad \oplus_{a,b} v_a \times v_b (N_a, \bar{N}_b)$$

where the cross product is $v_a \times v_b = \prod_{i=1,2,3}(n^i_a m^i_b - m^i_a n^i_b)$.

The solutions of this system of constraints are another "mathematical landscape" which can be classified and understood. Physicists then ask questions such as

- Can we realize the Standard Model gauge group and matter, *i.e.* $N_1 = 1$, $N_2 = 2$, $N_3 = 3$, $15N_{gen} + 4$ specific matter fields, limits on exotic matter charged in any $(N_{1,2,3}, \bar{N}_{i>3})$.
- What fraction of configurations realize it? What is the distribution of the number of generations?
- What is the distribution of extra matter? Is the extra matter consistent with experimental constraints? If so, can we propose a new experiment to detect it?

And it can be generalized, for example T^6 can be replaced by a Calabi-Yau threefold.

Thus we want to sample solutions of a complicated combinatorial optimization problem. A simple approach would be to postulate transitions between configurations, say which vary the (n, m) data, and do a random walk. But it turns out that reinforcement learning, with a reward function encoding the constraints and the Standard Model target, is a much more efficient technique for doing this (Halverson, Nelson and Ruehle, `hep-th/1903.11616`).

We leave the details for other talks, and just raise two questions:

- Is the sampling unbiased? If it is biased, could this be meaningful, *i.e.* somehow suggesting preferred configurations?
- In quantum cosmology there is a “correct” transition network. We don’t know it, but it only visits configurations which satisfy the tadpole constraints (but are not necessarily supersymmetric). It might also go through additional states. Do these details affect the dynamics?

Special holonomy manifolds

While there are M which are flat (tori) or almost flat (orbifolds), these might only be a tiny subset of the possibilities, not preferred *a priori*. But, the Einstein equations are very complicated nonlinear equations, and no other closed form solutions on compact manifolds are known.

Happily, there is another class of Ricci-flat manifolds, the special holonomy manifolds. The holonomy group of a manifold is the group of rotations generated by parallel transport of a vector, and "special" means that it is a proper subgroup of the full group $SO(D)$. Although not obvious, this condition generally determines the metric in terms of a single function, so the Einstein equations reduce to a single non-linear PDE, which one can prove has solutions.

One can also argue that low energy supersymmetry requires M to be special holonomy. In six dimensions these are the Calabi-Yau "threefolds," while in seven they are G_2 manifolds. Elliptically fibered Calabi-Yau "fourfolds" are also relevant for F theory.

Special holonomy manifolds

While there are M which are flat (tori) or almost flat (orbifolds), these might only be a tiny subset of the possibilities, not preferred *a priori*. But, the Einstein equations are very complicated nonlinear equations, and no other closed form solutions on compact manifolds are known.

Happily, there is another class of Ricci-flat manifolds, the special holonomy manifolds. The holonomy group of a manifold is the group of rotations generated by parallel transport of a vector, and "special" means that it is a proper subgroup of the full group $SO(D)$. Although not obvious, this condition generally determines the metric in terms of a single function, so the Einstein equations reduce to a single non-linear PDE, which one can prove has solutions.

One can also argue that low energy supersymmetry requires M to be special holonomy. In six dimensions these are the Calabi-Yau "threefolds," while in seven they are G_2 manifolds. Elliptically fibered Calabi-Yau "fourfolds" are also relevant for F theory.

The basic example is the quintic Calabi-Yau threefold, defined as the solutions to the equation

$$0 = \sum_{i=1}^{5} Z_i^5$$

in complex projective space $\mathbb{CP}^4$. $\mathbb{CP}^4$ is defined by identifying points $\vec{Z} \sim \lambda\vec{Z}, \forall\lambda$ (in words, we ignore the overall scale of $\vec{Z}$).
This is the "Fermat quintic," a particularly simple and symmetric case.
In fact one can use any polynomial $P(\vec{Z})$ of degree 5 such that $\nabla P \neq 0$ where $P = 0$. Such polynomials have 126 independent coefficients, while $\mathbb{CP}^4$ has a 25-dimensional symmetry group, leaving 101 parameters for the general quintic. Finally there is one overall size parameter (Kähler modulus).

By combining sophisticated techniques in algebraic geometry with some numerical PDE, one can do amazing computations, such as the metric on moduli space, the flux superpotential, *etc.*

The basic example is the quintic Calabi-Yau threefold, defined as the solutions to the equation

$$0 = \sum_{i=1}^{5} Z_i^5$$

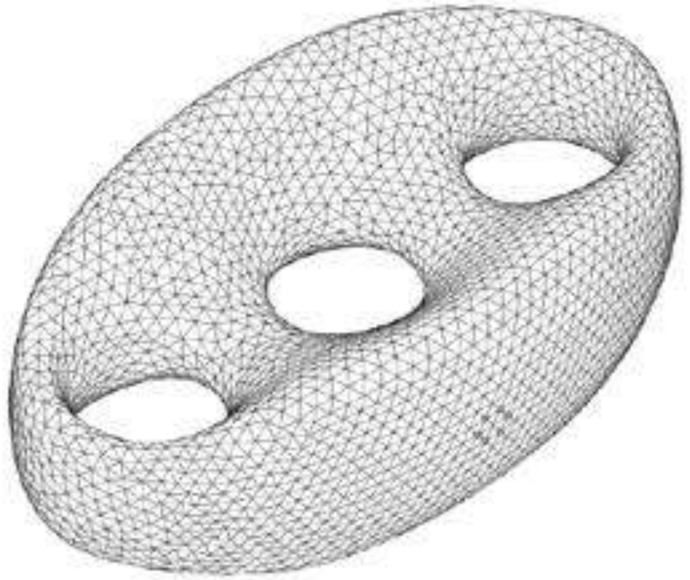
in complex projective space $\mathbb{CP}^4$. $\mathbb{CP}^4$ is defined by identifying points $\vec{Z} \sim \lambda\vec{Z}, \forall\lambda$ (in words, we ignore the overall scale of $\vec{Z}$).
This is the "Fermat quintic," a particularly simple and symmetric case.
In fact one can use any polynomial $P(\vec{Z})$ of degree 5 such that $\nabla P \neq 0$ where $P = 0$. Such polynomials have 126 independent coefficients, while $\mathbb{CP}^4$ has a 25-dimensional symmetry group, leaving 101 parameters for the general quintic. Finally there is one overall size parameter (Kähler modulus).

By combining sophisticated techniques in algebraic geometry with some numerical PDE, one can do amazing computations, such as the metric on moduli space, the flux superpotential, *etc.*

Geometry of space

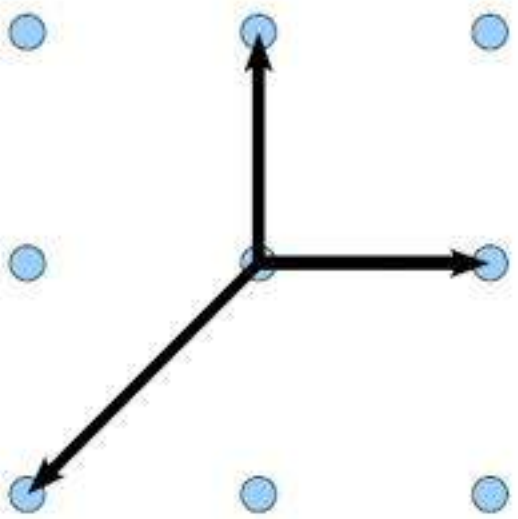
There are many mathematical formalizations of the concept of "space."

- algebraic – as solutions of equations, *e.g.* $\sum_i x_i^2 = 1$.
- by patching together overlapping neighborhoods – manifolds
- by gluing together simple spaces – triangulations, complexes
- in terms of functions on the space – sheaves, toric geometry
- in terms of algebras associated to the space, *etc.*.

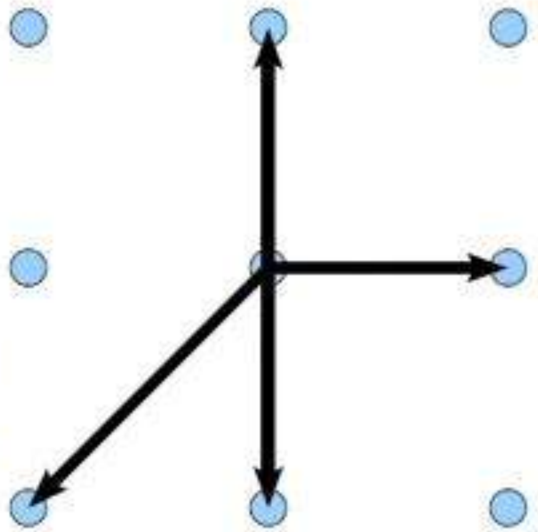


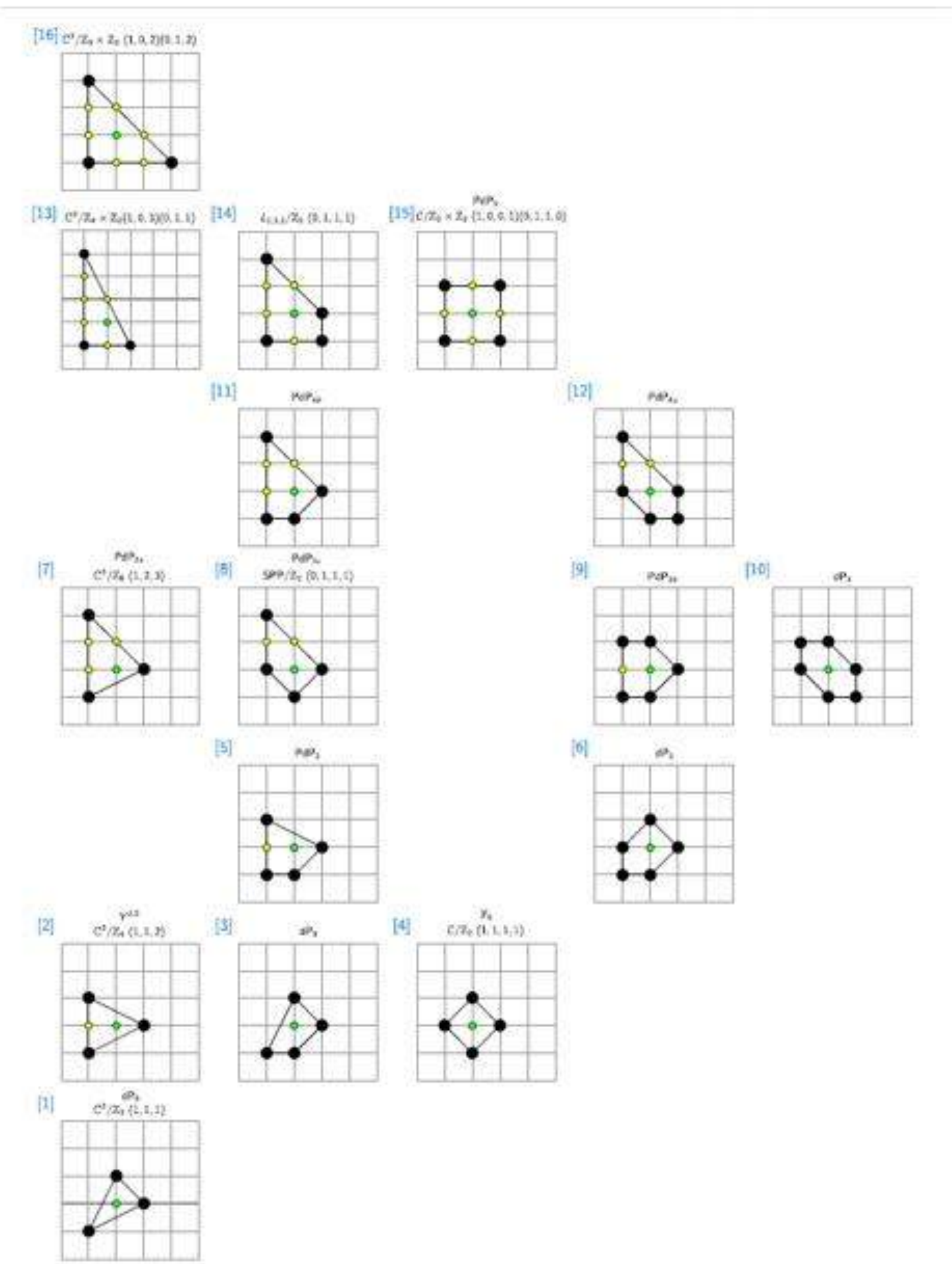
Toric manifolds

A good way to get more CY3's is to keep the single defining equation (a hypersurface), but generalize the ambient space $\mathbb{CP}^4$. A rather general construction of analogous spaces is provided by toric geometry. This combines many of the ideas on the previous slide.



In concrete terms, a toric space is specified by a fan. Each ray of the fan corresponds to a coordinate Z^i, while linear relations between rays correspond to symmetries. Thus, the top figure represents $\mathbb{CP}^2$, and the relation $v_0 + v_1 + v_2 = 0$ corresponds to the action $Z^i \to \lambda Z^i$. The bottom figure has an extra ray and an extra relation. It represents the del Pezzo surface dP_1.





These are the $d = 2$ reflexive polytopes, from Y.-H. He, R. K. Seong and S.-T. Yau, `arXiv:1704.03462`. The symmetry under reflection across the x-axis relates pairs of dual polytopes.

Kreuzer-Skarke database

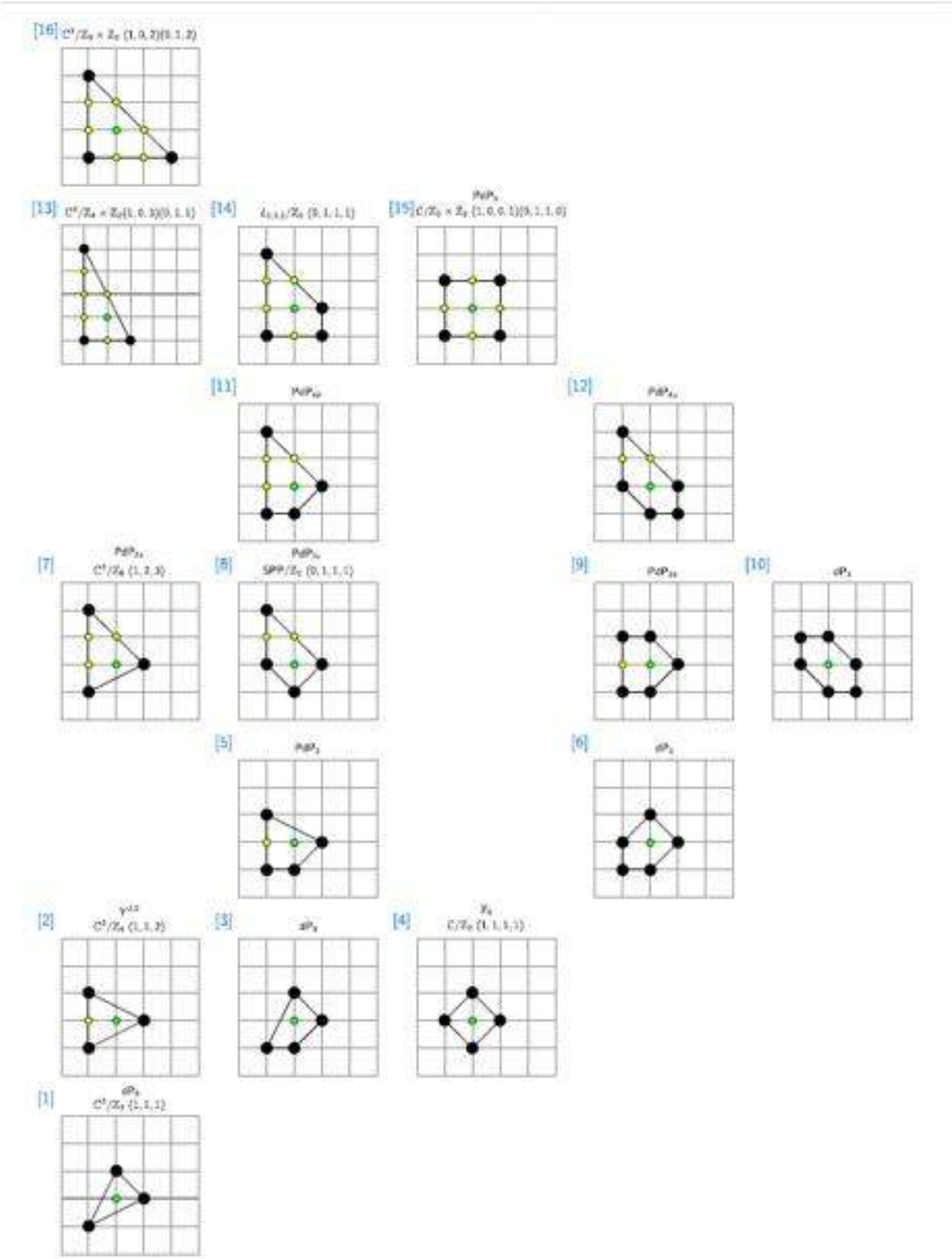
An equivalent presentation is to give the convex hull of an integer vertex along each ray. This is a lattice polytope Δ. Its **dual polytope** is the set

$$\Delta^* \equiv \{y \in \mathbb{R}^{d+1} : x \cdot y \geq -1 \; \forall x \in \Delta\}. \tag{1}$$

A lattice polytope is **reflexive** if its dual is also a lattice polytope.

It turns out (as first shown by Batyrev) that reflexivity is precisely the condition for there to be a Calabi-Yau hypersurface in Δ. Clearly there will be one in Δ^* as well and thus these CY manifolds come in pairs (unless $\Delta \cong \Delta^*$). This is mirror symmetry.

For fixed d, the set of reflexive lattice polytopes is finite. Starting in the 90's, Kreuzer and Skarke made a database of the 473,800,776 instances for $d = 3$. (See `http://hep.itp.tuwien.ac.at/~kreuzer/CY/`)



These are the $d = 2$ reflexive polytopes, from Y.-H. He, R. K. Seong and S.-T. Yau, arXiv:1704.03462. The symmetry under reflection across the x-axis relates pairs of dual polytopes.

Kreuzer-Skarke database

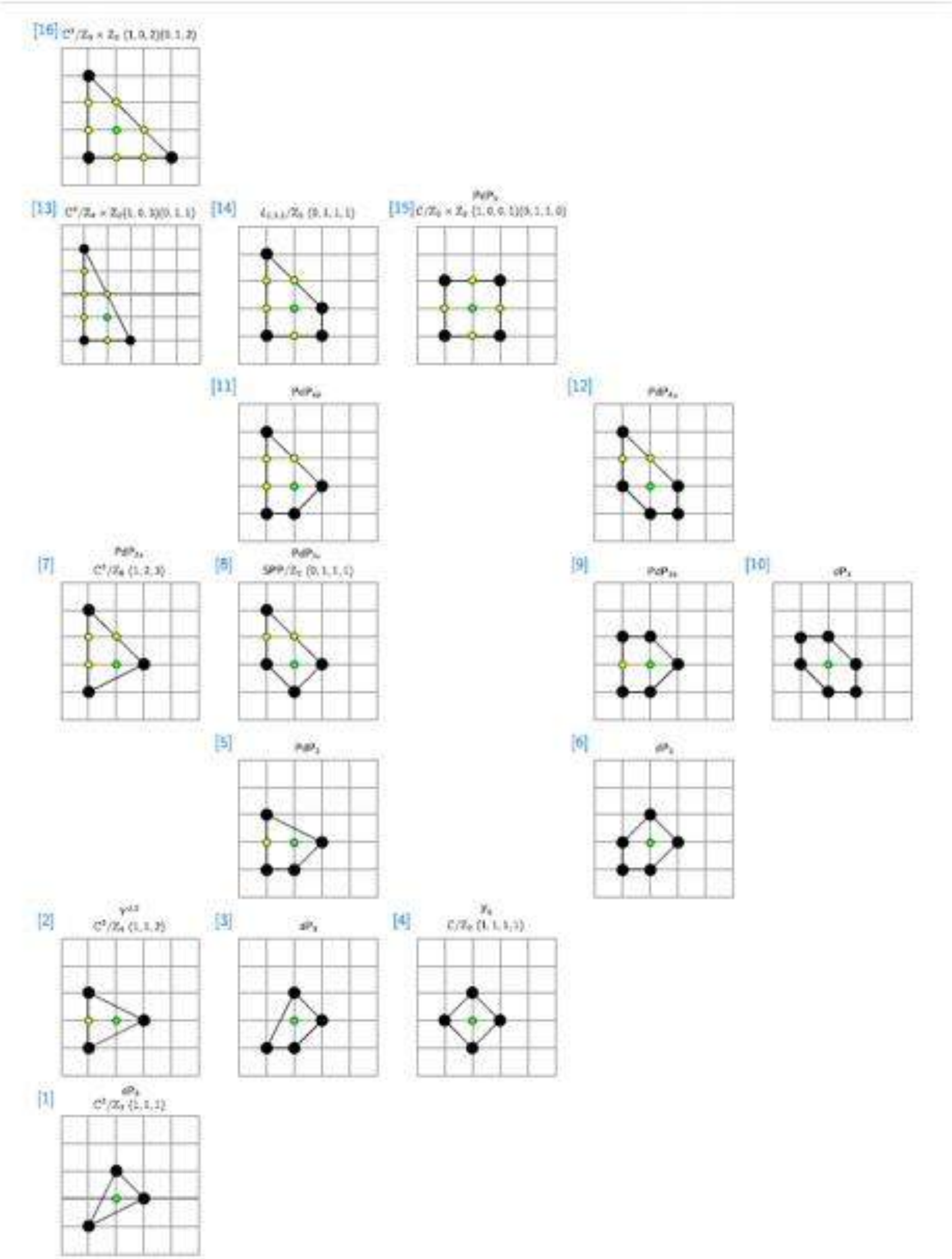
An equivalent presentation is to give the convex hull of an integer vertex along each ray. This is a lattice polytope Δ. Its **dual polytope** is the set

$$\Delta^* \equiv \{y \in \mathbb{R}^{d+1} : x \cdot y \geq -1 \ \forall x \in \Delta\}. \quad (1)$$

A lattice polytope is **reflexive** if its dual is also a lattice polytope.

It turns out (as first shown by Batyrev) that reflexivity is precisely the condition for there to be a Calabi-Yau hypersurface in Δ. Clearly there will be one in Δ^* as well and thus these CY manifolds come in pairs (unless $\Delta \cong \Delta^*$). This is mirror symmetry.

For fixed d, the set of reflexive lattice polytopes is finite. Starting in the 90's, Kreuzer and Skarke made a database of the 473,800,776 instances for $d = 3$. (See `http://hep.itp.tuwien.ac.at/~kreuzer/CY/`)



These are the $d = 2$ reflexive polytopes, from Y.-H. He, R. K. Seong and S.-T. Yau, `arXiv:1704.03462`. The symmetry under reflection across the x-axis relates pairs of dual polytopes.

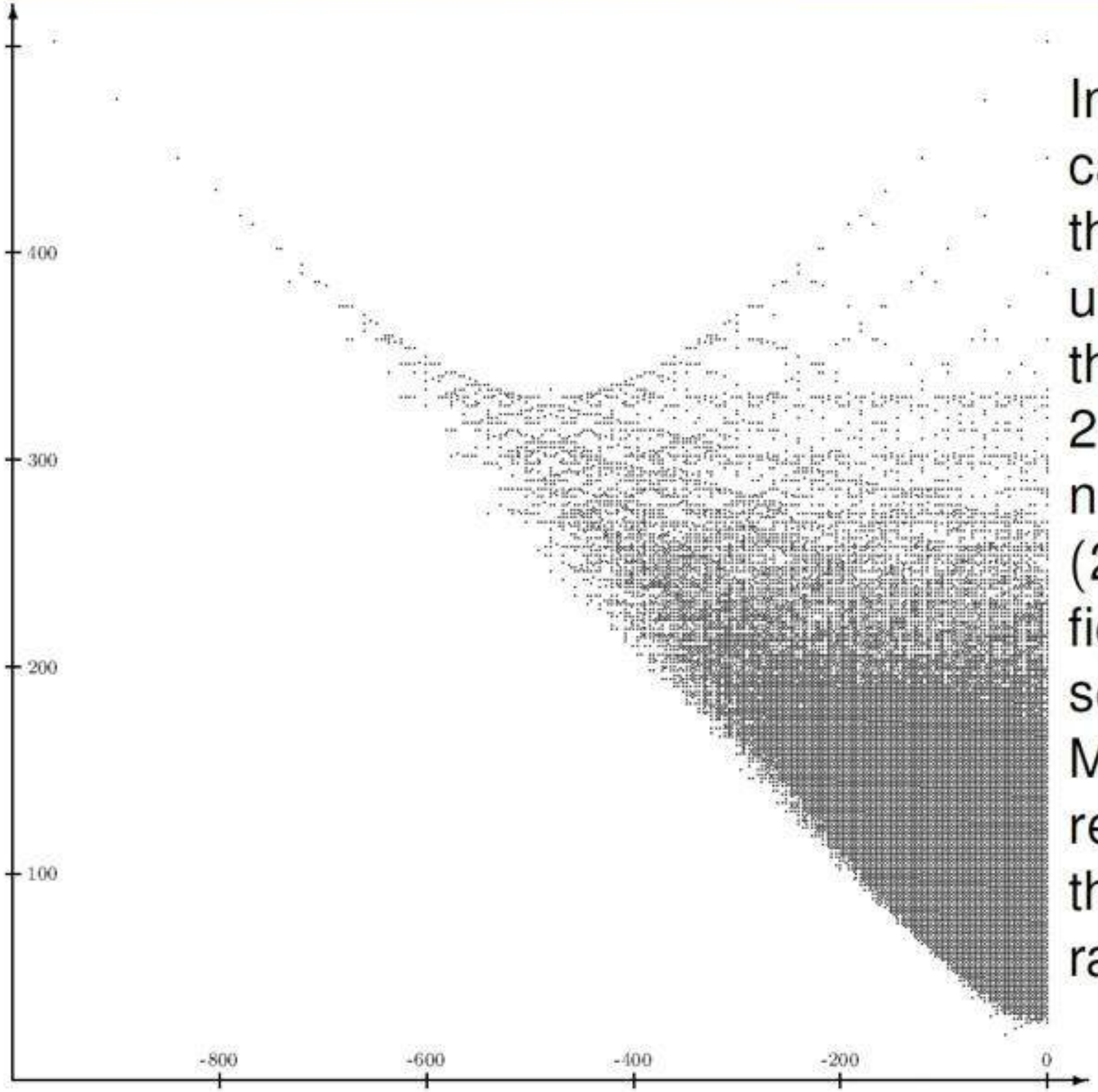
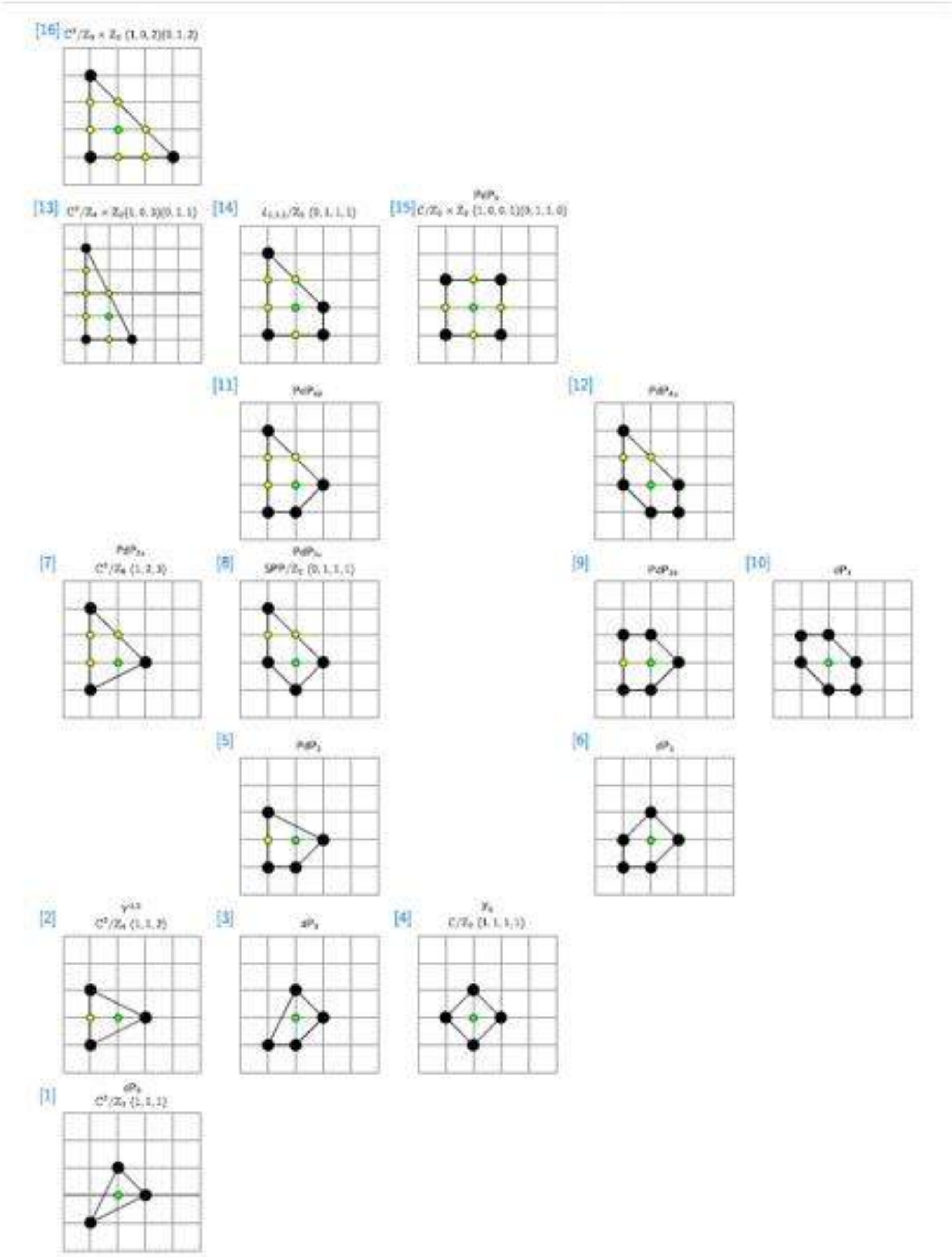


Fig. 1: $h_{11} + h_{12}$ vs. Euler number $\chi = 2(h_{11} - h_{12})$ for all pairs (h_{11}, h_{12}) with $h_{11} \leq h_{12}$.

In this figure, the vertical axis is $b^{1,1} + b^{2,1}$, the number of CY moduli. The horizontal axis is the Euler character $\chi = 2b^{1,1} - 2b^{2,1}$, twice the number of generations in $(2,2)$ heterotic compactification. Each dot represents one or more CY_3's. Mirror symmetry is the reflection symmetry with the (omitted) $\chi > 0$ quadrant.



These are the $d = 2$ reflexive polytopes, from Y.-H. He, R. K. Seong and S.-T. Yau, arXiv:1704.03462. The symmetry under reflection across the x-axis relates pairs of dual polytopes.

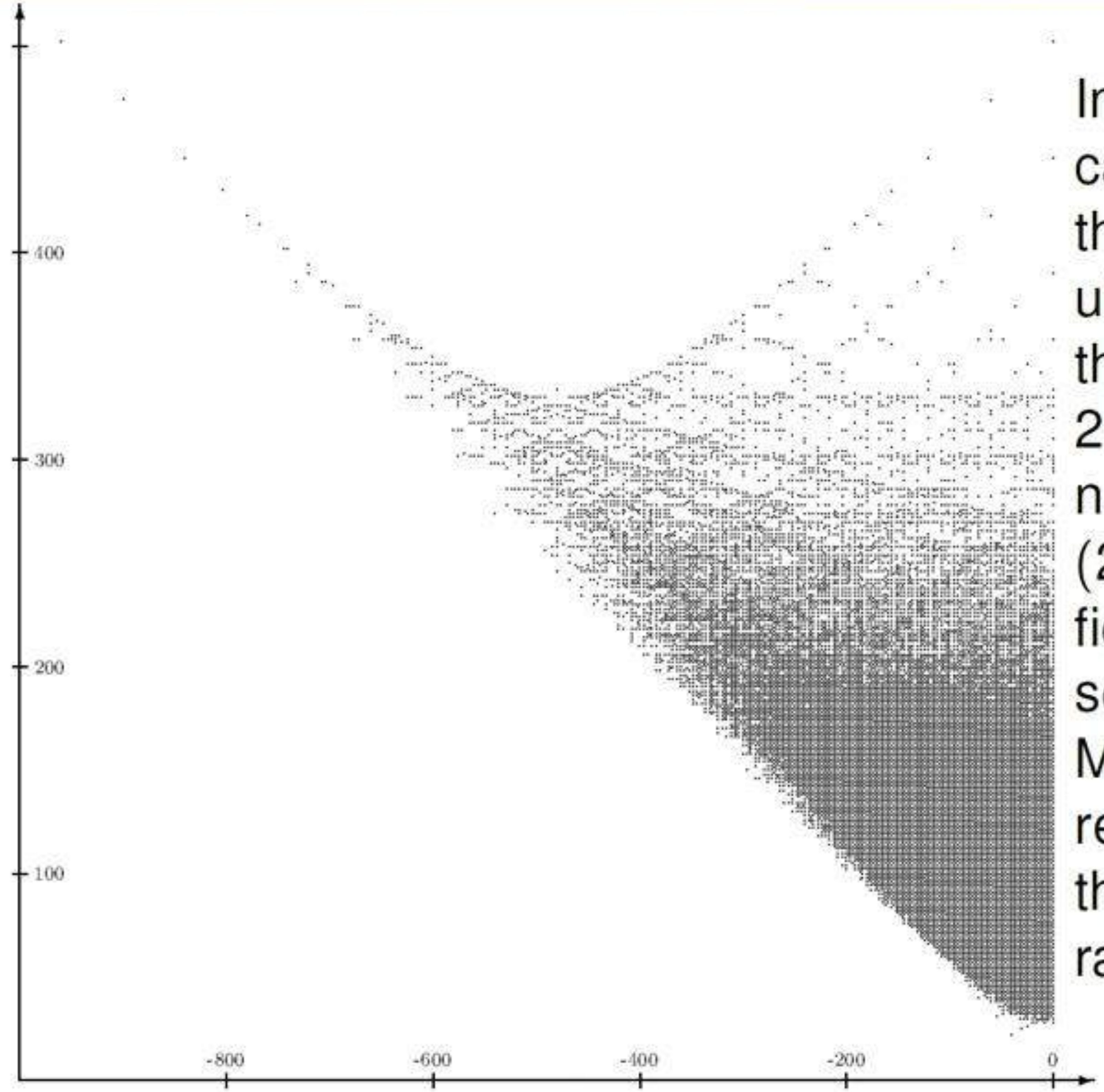


Fig. 1: $h_{11} + h_{12}$ vs. Euler number $\chi = 2(h_{11} - h_{12})$ for all pairs (h_{11}, h_{12}) with $h_{11} \le h_{12}$.

In this figure, the vertical axis is $b^{1,1} + b^{2,1}$, the number of CY moduli. The horizontal axis is the Euler character $\chi = 2b^{1,1} - 2b^{2,1}$, twice the number of generations in $(2,2)$ heterotic compactification. Each dot represents one or more CY_3's. Mirror symmetry is the reflection symmetry with the (omitted) $\chi > 0$ quadrant.

While Kreuzer and Skarke were inspired by mirror symmetry, their dataset displayed a more subtle pattern, the simple shape of its boundary. This is usually called the "shield" after the shape of the full boundary with the mirror symmetry.

While the finiteness of the set was known, the shield was unexpected and deserves to be called a computer-aided discovery. We might ask:

> *What is the shape of the boundary for higher-dimensional reflexive polytopes?*

Since their number grows superexponentially with d, this question will require sampling and probably ML to study. I offer it as a challenge to the audience.

To what extent can we understand the shield? The upper boundary is (largely?) reproduced by a related but different construction of CY3's, as elliptic fibrations. Does this help?

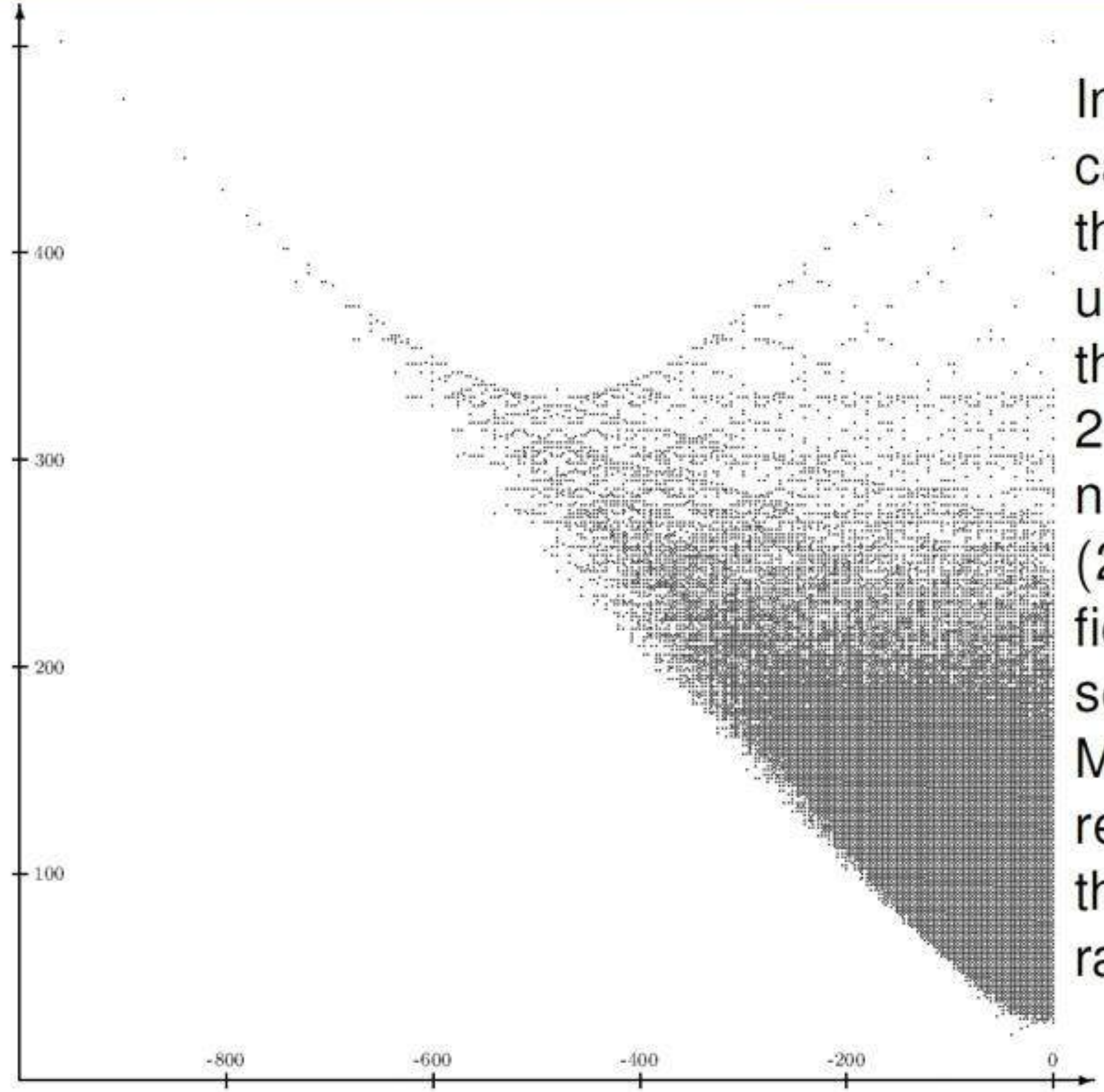


Fig. 1: $h_{11} + h_{12}$ vs. Euler number $\chi = 2(h_{11} - h_{12})$ for all pairs (h_{11}, h_{12}) with $h_{11} \leq h_{12}$.

In this figure, the vertical axis is $b^{1,1} + b^{2,1}$, the number of CY moduli. The horizontal axis is the Euler character $\chi = 2b^{1,1} - 2b^{2,1}$, twice the number of generations in $(2,2)$ heterotic compactification. Each dot represents one or more CY_3's. Mirror symmetry is the reflection symmetry with the (omitted) $\chi > 0$ quadrant.

While Kreuzer and Skarke were inspired by mirror symmetry, their dataset displayed a more subtle pattern, the simple shape of its boundary. This is usually called the “shield” after the shape of the full boundary with the mirror symmetry.

While the finiteness of the set was known, the shield was unexpected and deserves to be called a computer-aided discovery. We might ask:

> *What is the shape of the boundary for higher-dimensional reflexive polytopes?*

Since their number grows superexponentially with d, this question will require sampling and probably ML to study. I offer it as a challenge to the audience.

To what extent can we understand the shield? The upper boundary is (largely?) reproduced by a related but different construction of CY3's, as elliptic fibrations. Does this help?

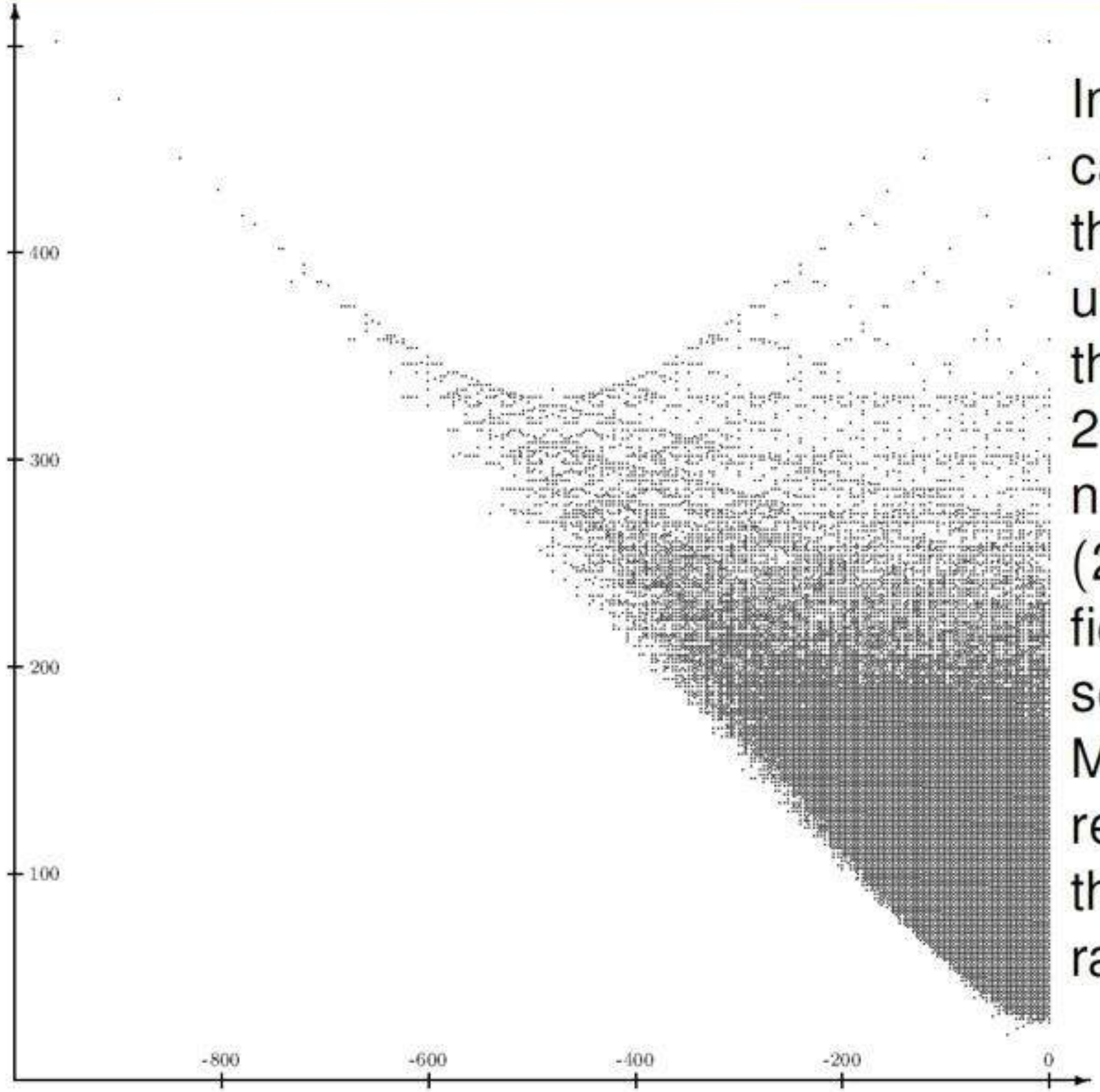


Fig. 1: $h_{11} + h_{12}$ vs. Euler number $\chi = 2(h_{11} - h_{12})$ for all pairs (h_{11}, h_{12}) with $h_{11} \leq h_{12}$.

In this figure, the vertical axis is $b^{1,1} + b^{2,1}$, the number of CY moduli. The horizontal axis is the Euler character $\chi = 2b^{1,1} - 2b^{2,1}$, twice the number of generations in $(2,2)$ heterotic compactification. Each dot represents one or more CY_3's. Mirror symmetry is the reflection symmetry with the (omitted) $\chi > 0$ quadrant.

While Kreuzer and Skarke were inspired by mirror symmetry, their dataset displayed a more subtle pattern, the simple shape of its boundary. This is usually called the "shield" after the shape of the full boundary with the mirror symmetry.

While the finiteness of the set was known, the shield was unexpected and deserves to be called a computer-aided discovery. We might ask:

> *What is the shape of the boundary for higher-dimensional reflexive polytopes?*

Since their number grows superexponentially with d, this question will require sampling and probably ML to study. I offer it as a challenge to the audience.

To what extent can we understand the shield? The upper boundary is (largely?) reproduced by a related but different construction of CY3's, as elliptic fibrations. Does this help?

While Kreuzer and Skarke were inspired by mirror symmetry, their dataset displayed a more subtle pattern, the simple shape of its boundary. This is usually called the “shield” after the shape of the full boundary with the mirror symmetry.

While the finiteness of the set was known, the shield was unexpected and deserves to be called a computer-aided discovery. We might ask:

> *What is the shape of the boundary for higher-dimensional reflexive polytopes?*

Since their number grows superexponentially with d, this question will require sampling and probably ML to study. I offer it as a challenge to the audience.

To what extent can we understand the shield? The upper boundary is (largely?) reproduced by a related but different construction of CY3’s, as elliptic fibrations. Does this help?

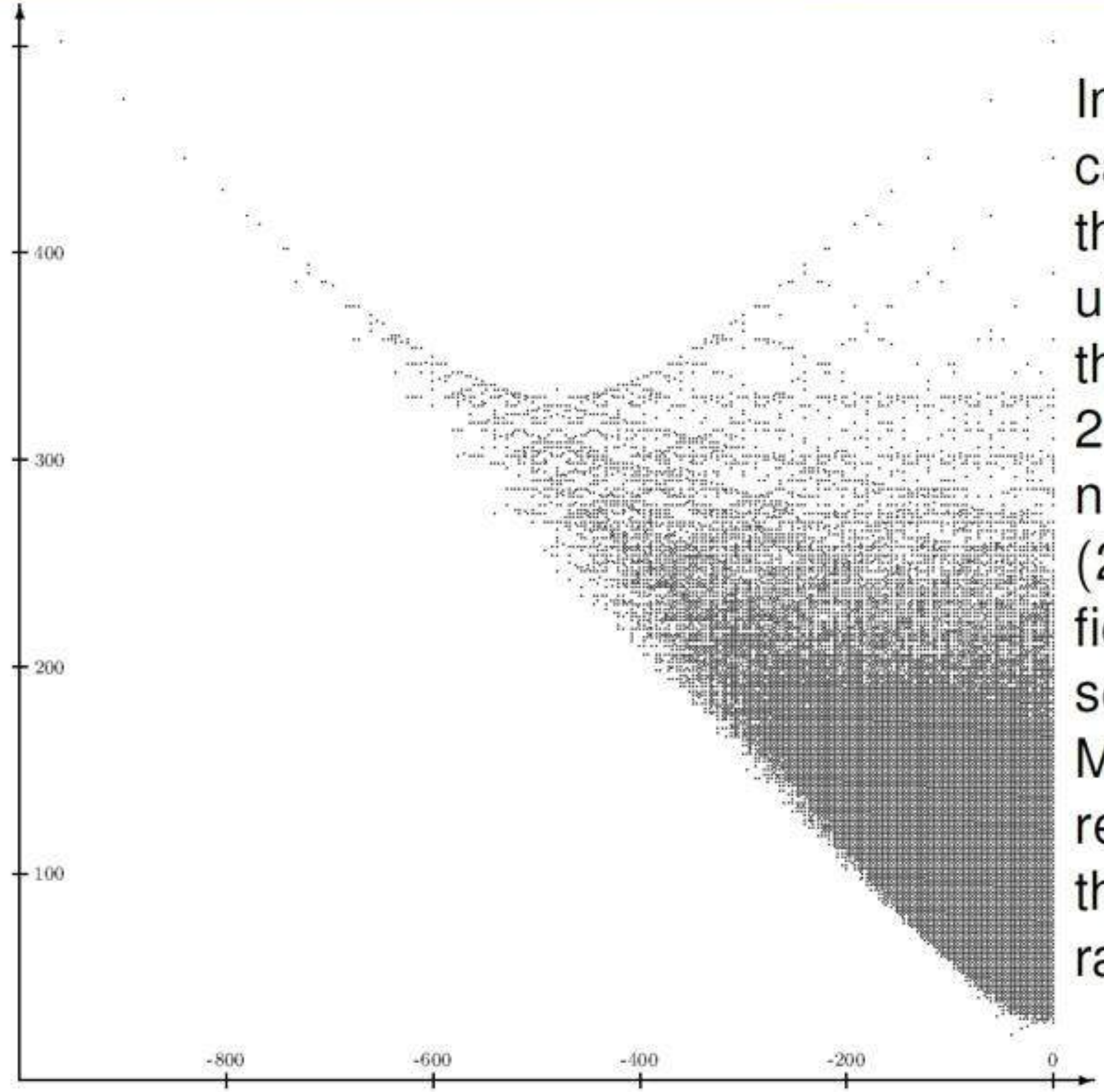


Fig. 1: $h_{11} + h_{12}$ vs. Euler number $\chi = 2(h_{11} - h_{12})$ for all pairs (h_{11}, h_{12}) with $h_{11} \leq h_{12}$.

In this figure, the vertical axis is $b^{1,1} + b^{2,1}$, the number of CY moduli. The horizontal axis is the Euler character $\chi = 2b^{1,1} - 2b^{2,1}$, twice the number of generations in $(2,2)$ heterotic compactification. Each dot represents one or more CY_3's. Mirror symmetry is the reflection symmetry with the (omitted) $\chi > 0$ quadrant.

While Kreuzer and Skarke were inspired by mirror symmetry, their dataset displayed a more subtle pattern, the simple shape of its boundary. This is usually called the "shield" after the shape of the full boundary with the mirror symmetry.

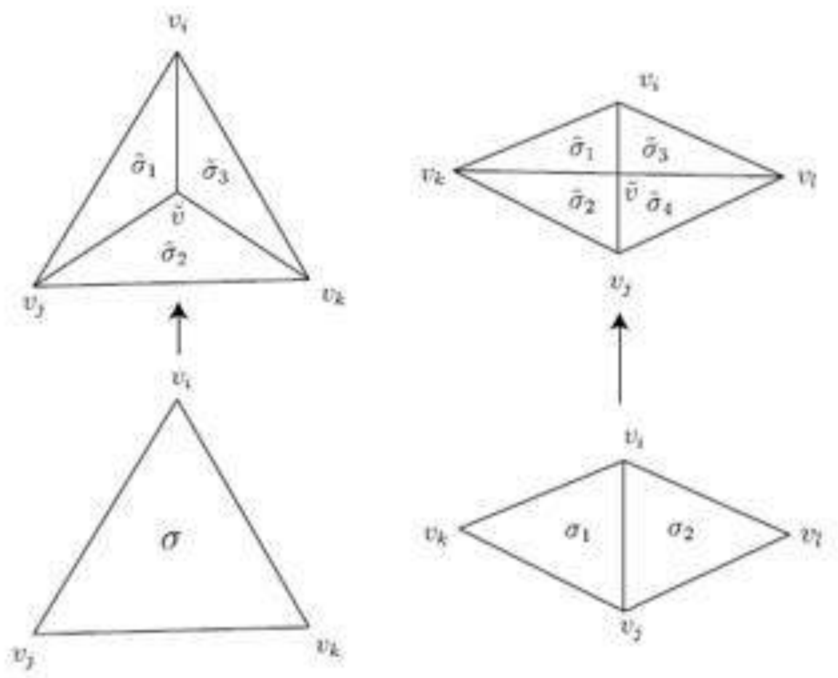
While the finiteness of the set was known, the shield was unexpected and deserves to be called a computer-aided discovery. We might ask:

> *What is the shape of the boundary for higher-dimensional reflexive polytopes?*

Since their number grows superexponentially with d, this question will require sampling and probably ML to study. I offer it as a challenge to the audience.

To what extent can we understand the shield? The upper boundary is (largely?) reproduced by a related but different construction of CY3's, as elliptic fibrations. Does this help?

The set of CY's also comes with natural transitions, blowups and extremal transitions. These are also simple in the toric language – they correspond to subdivisions of the fan as in the figures, or rearrangements which are compositions of these. This network is connected, and the dynamics on it has been studied, both to sample the CYs and to give toy models of quantum cosmology in string theory.



Landscape of finite groups

Let us come back to the set of finite groups. Could we work with it in the same manner? Many points to settle and differences to address:

- We would like a state-action network with transitions between groups. Restricting to finite simple groups may be too sparse for this. All finite groups are connected by group extensions/quotients, but does this lead to a well connected network?
- The set is infinite and we need a measure of complexity, to keep the search from running away.
- Need to represent the group data in a way that fits with above choices.
- Need simple features of group as input to ML to guide search.

Here is a challenge for AI/ML: can a computer, by exploring the set of finite groups,

- Generate a (short) list including all of the finite simple groups ?
- Produce evidence that there are no more finite simple groups ?
- Give us an "understandable picture" of the set of finite groups ?

We are not asking for it to do this starting only from the axioms (yet). The idea is to program some sort of search algorithm, which can take advantage of known facts in group theory, such that running a program of length $\ll$ 5000 pages generates this information.

Ultimately, by combining the facts generated by such an automated search, with automated theorem verification technology (more later), one might come up with a proof of the classification which while longer in absolute terms, could start from far less human input. And to the extent that the computer can help us navigate and work with this information, the longer proof might be more understandable and useful.

Here is a challenge for AI/ML: can a computer, by exploring the set of finite groups,

- Generate a (short) list including all of the finite simple groups ?
- Produce evidence that there are no more finite simple groups ?
- Give us an “understandable picture” of the set of finite groups ?

We are not asking for it to do this starting only from the axioms (yet). The idea is to program some sort of search algorithm, which can take advantage of known facts in group theory, such that running a program of length $\ll$ 5000 pages generates this information.

Ultimately, by combining the facts generated by such an automated search, with automated theorem verification technology (more later), one might come up with a proof of the classification which while longer in absolute terms, could start from far less human input. And to the extent that the computer can help us navigate and work with this information, the longer proof might be more understandable and useful.

Conclusions

We discussed various examples of mathematical landscapes: of lattice polytopes, of toroidal compactifications, and of finite groups. Even if they can be understood without the computer, they are intricate enough that its help is very welcome. And a long but conceptually simple description of a set – perhaps discovered using ML – might sometimes be more helpful than a short but conceptually more complex description.

Indeed, some landscapes remain too intricate for humans alone. Examples are the sets of unavoidable/reducible graphs which arise in the Appel-Haken and Robertson *et al* proofs of the Four Color Theorem. These are verified by computer using "discharge rules" which were postulated by the authors. But could a computer come up with these rules? The Robertson *et al* proof was formalized and machine verified by Gonthier *et al*; could the same be done for the computer-generated rules ? If so, is this a "simple" proof ?

Conclusions

We discussed various examples of mathematical landscapes: of lattice polytopes, of toroidal compactifications, and of finite groups. Even if they can be understood without the computer, they are intricate enough that its help is very welcome. And a long but conceptually simple description of a set – perhaps discovered using ML – might sometimes be more helpful than a short but conceptually more complex description.

Indeed, some landscapes remain too intricate for humans alone. Examples are the sets of unavoidable/reducible graphs which arise in the Appel-Haken and Robertson *et al* proofs of the Four Color Theorem. These are verified by computer using “discharge rules” which were postulated by the authors. But could a computer come up with these rules? The Robertson *et al* proof was formalized and machine verified by Gonthier *et al*; could the same be done for the computer-generated rules ? If so, is this a “simple” proof ?

Even if ML can find patterns in mathematical data, if the patterns are complicated, we will want more computer help to work with them and prove theorems about them. Thus, to realize the potential of this direction, we need a synthesis between regarding mathematics as data and regarding mathematics as a deductive system. Automated theorem verification and proof does the latter and we can hope that it will eventually share in the recent advances of AI.

Conclusions

We discussed various examples of mathematical landscapes: of lattice polytopes, of toroidal compactifications, and of finite groups. Even if they can be understood without the computer, they are intricate enough that its help is very welcome. And a long but conceptually simple description of a set – perhaps discovered using ML – might sometimes be more helpful than a short but conceptually more complex description.

Indeed, some landscapes remain too intricate for humans alone. Examples are the sets of unavoidable/reducible graphs which arise in the Appel-Haken and Robertson *et al* proofs of the Four Color Theorem. These are verified by computer using "discharge rules" which were postulated by the authors. But could a computer come up with these rules? The Robertson *et al* proof was formalized and machine verified by Gonthier *et al*; could the same be done for the computer-generated rules ? If so, is this a "simple" proof ?

Physics and Machine Learning, Microsoft Research, Seattle, 25 April 2019
AdS/CFT workshop at OIST, Okinawa, 3 April 2019
"Machine learning landscape" workshop, ICTP Trieste, 10 Dec 2018
East Asian string workshop, KIAS, Seoul, 7 Nov 2018
KMI colloquium, Nagoya U, 25 Oct 2018
"QG meets lattice QCD workshop", ECT*, Trento, 3 Sep 2018
APCTP focus program, Hanyang U, Seoul, 15 Aug 2018
QFT workshop, YITP, Kyoto, 31 July 2018
Machine learning workshop, TSiMF, China, 15 June, 2018
Paris QCD workshop, 11 June, 2018
DLAP2018 workshop, Osaka, 1 June, 2018
MPI, AEI, 13 Apr, 2018
MIT, CTP, 4 Apr, 2018
CQuest, Seoul, 29 March, 2018
KIAS, Seoul, 26 March, 2018

Holography, Matter and Deep Learning

Koji Hashimoto (Osaka Univ.)

ArXiv:1903.04951
ArXiv:1802.08313, 1809.10536
w/ S. Sugishita (Kentucky), A. Tanaka (RIKEN), A. Tomiya (RIKEN)

8 pages

0. Networks in matter/gravity

5 pages

1. Holography as a deep learning

3 pages

2. Application: Quark matter

0-1 Networks in matter and gravity

Quantum Matter
Wave function

- Matrix product states
- Tensor network states
  - MERA
- Neural network states
  - Boltzmann machine
  - Feedforward

Quantum Gravity

- String theory
- Holographic principle
  - Quantum codes
- Regge calculus
  - Causal dynamical triangulation

0-1

Networks in matter and gravity

Quantum Matter
Wave function

- Matrix product states

- Tensor network states

  MERA

- Neural network states

  Boltzmann machine
  Feedforward

0-2 Quantum matter : ground state?

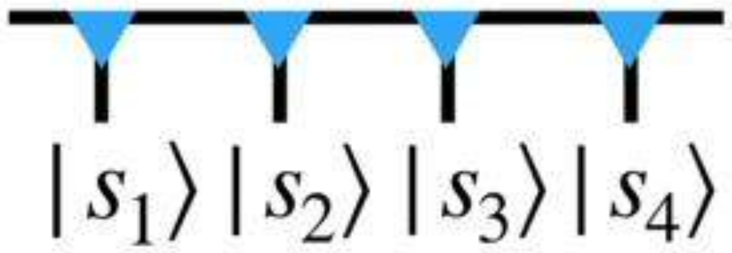
Ground state wave function for *N* qubits $\psi(s_1, s_2, \cdots, s_N)$

Question: minimize its energy *E* for a given Hamiltonian *H*,

$$E \equiv \sum_{s_1, s_2, \cdots} H(s_1, s_2, \cdots, s_N) \, |\psi(s_1, s_2, \cdots, s_N)|^2$$

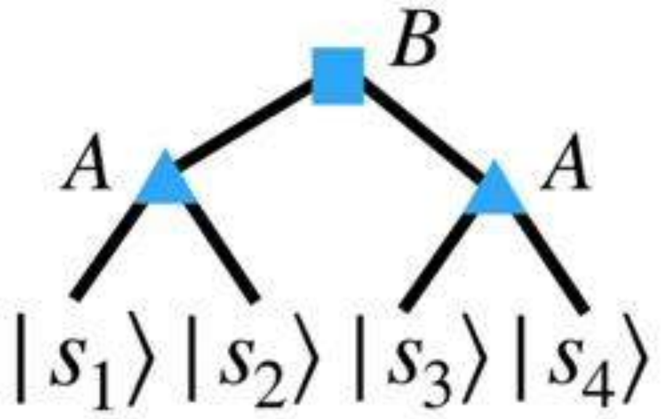
- Matrix product states

$$\psi(s_1, s_2, \cdots) = \mathrm{tr}[A^{(s_1)} A^{(s_2)} \cdots]$$



- Tensor network states

$$\psi(s_1, s_2, \cdots) = \sum_{m,n} B_{mn} A_{m s_1 s_2} A_{n s_3 s_4}$$



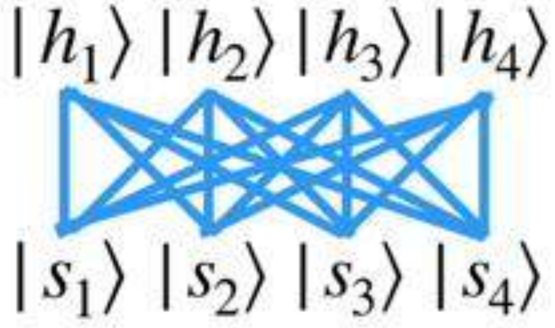
0-3 Neural network states

- Boltzmann machine states

  [Carleo Troyer `17],

  [Nomura, Darmawan, Yamaji, Imada `17], ..

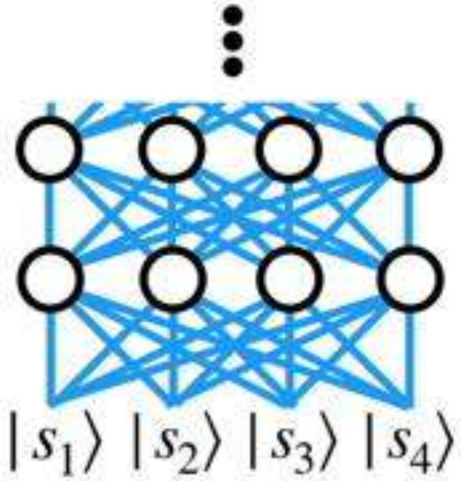
$$\psi(s_1, \cdots, s_N) = \sum_{h_A} \exp\left[\sum_a a_a s_a + \sum_A b_A h_A + \sum_{a,A} J_{aA} s_a h_A\right]$$



- Deep Boltzmann machine states

  [Carleo, Nomura, Imada `18], ..

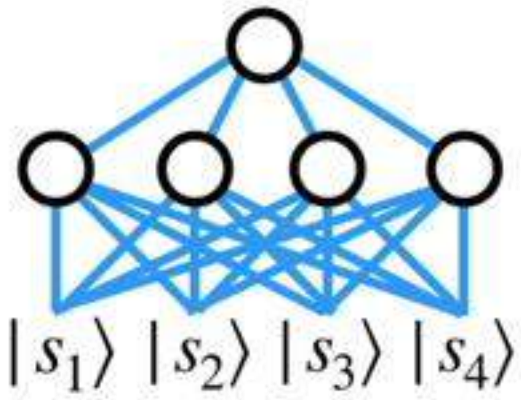
$$|\psi\rangle = \lim_{\tau\to\infty} e^{-\tau H}|\text{any}\rangle = e^{-\Delta\tau H}e^{-\Delta\tau H}\cdots|\text{any}\rangle$$



- Feedfoward network states

  [Saito `18], ..

$$\psi(s_1, \cdots, s_N) = \sum_i f_i\, \sigma\left(\sum_j W_{ij} s_j + b_i\right)$$



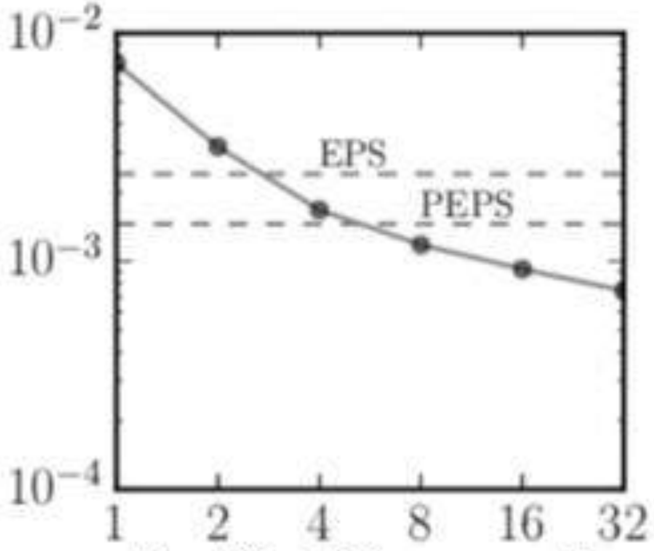
0-3 Better ansatz?

Neural states may beat conventional ones.

Ex) 2-dimensional antiferromagnetic Heisenberg model

[Carleo Troyer `17]

Energy with RBM states



# of hidden units

Relations among them:

1) Boltzmann machine states are tensor network states

[Chen, Cheng, Xie, Wang, Xiang `18]

2) Tensor states are deep Boltzmann

[Gao, Duan `17] [Huang, Moore `17]

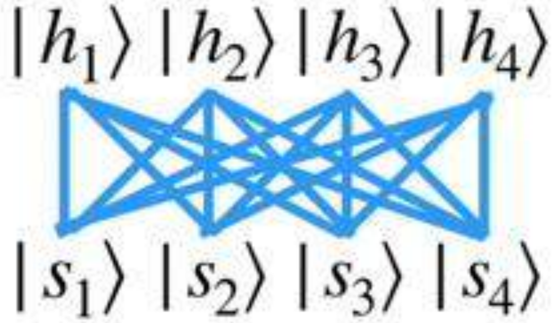
3) Tensor states are feedforward with "product pooling"

[Cohen, Shashua `18]

0-3 Neural network states

- Boltzmann machine states

  [Carleo Troyer `17],
  [Nomura, Darmawan, Yamaji, Imada `17], ..

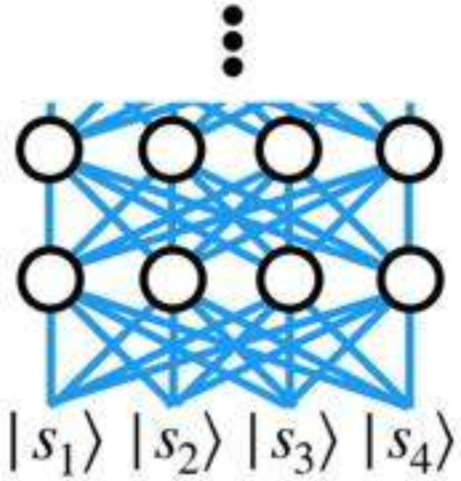


$$\psi(s_1,\cdots,s_N)=\sum_{h_A}\exp\left[\sum_a a_a s_a+\sum_A b_A h_A+\sum_{a,A} J_{aA} s_a h_A\right]$$

- Deep Boltzmann machine states

  [Carleo, Nomura, Imada `18], ..

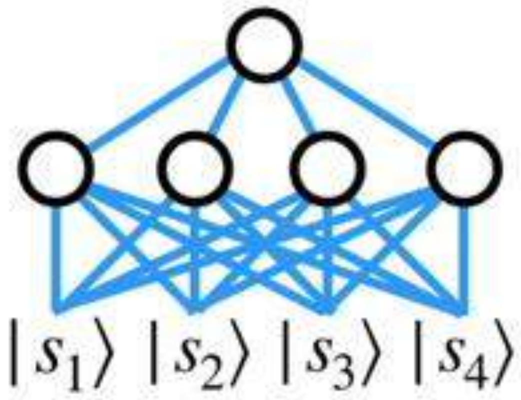
$$|\psi\rangle=\lim_{\tau\to\infty}e^{-\tau H}|\text{any}\rangle=e^{-\Delta\tau H}e^{-\Delta\tau H}\cdots|\text{any}\rangle$$



- Feedfoward network states

  [Saito `18], ..

$$\psi(s_1,\cdots,s_N)=\sum_i f_i\,\sigma\left(\sum_j W_{ij}s_j+b_i\right)$$



(0-2) Quantum matter : ground state?

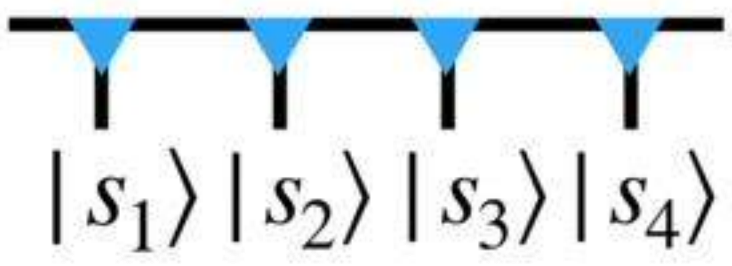
Ground state wave function for N qubits $\quad \psi(s_1, s_2, \cdots, s_N)$

Question: minimize its energy E for a given Hamiltonian H,

$$E \equiv \sum_{s_1, s_2, \cdots} H(s_1, s_2, \cdots, s_N) \, |\psi(s_1, s_2, \cdots, s_N)|^2$$

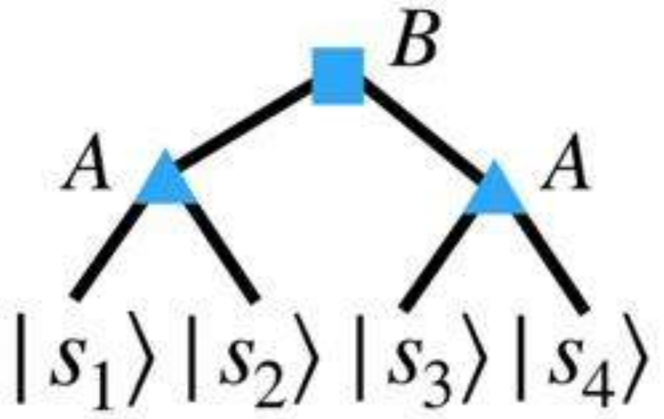
- Matrix product states

$$\psi(s_1, s_2, \cdots) = \mathrm{tr}[A^{(s_1)} A^{(s_2)} \cdots]$$



- Tensor network states

$$\psi(s_1, s_2, \cdots) = \sum_{m,n} B_{mn} A_{m s_1 s_2} A_{n s_3 s_4}$$



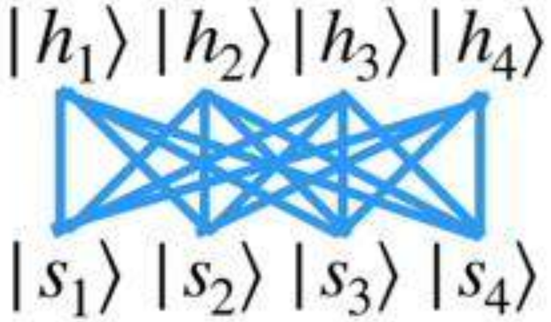
0-3 Neural network states

- Boltzmann machine states

  [Carleo Troyer `17],

  [Nomura, Darmawan, Yamaji, Imada `17], ..

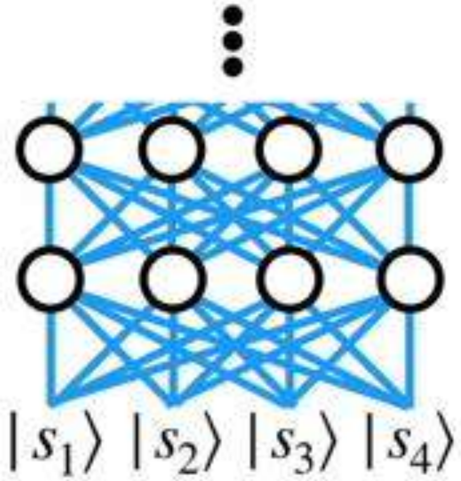
$$\psi(s_1, \cdots, s_N) = \sum_{h_A} \exp\left[\sum_a a_a s_a + \sum_A b_A h_A + \sum_{a,A} J_{aA} s_a h_A\right]$$



- Deep Boltzmann machine states

  [Carleo, Nomura, Imada `18], ..

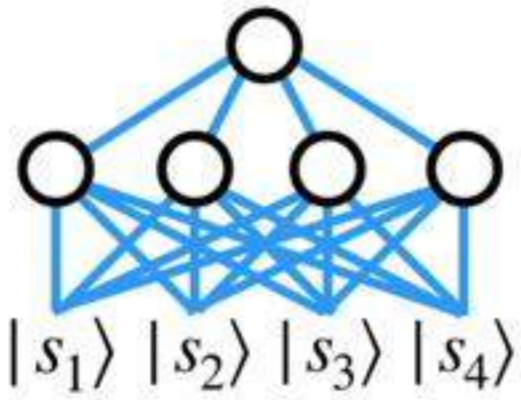
$$|\psi\rangle = \lim_{\tau\to\infty} e^{-\tau H}|\text{any}\rangle = e^{-\Delta\tau H}e^{-\Delta\tau H}\cdots|\text{any}\rangle$$



- Feedfoward network states

  [Saito `18], ..

$$\psi(s_1, \cdots, s_N) = \sum_i f_i\, \sigma\left(\sum_j W_{ij} s_j + b_i\right)$$

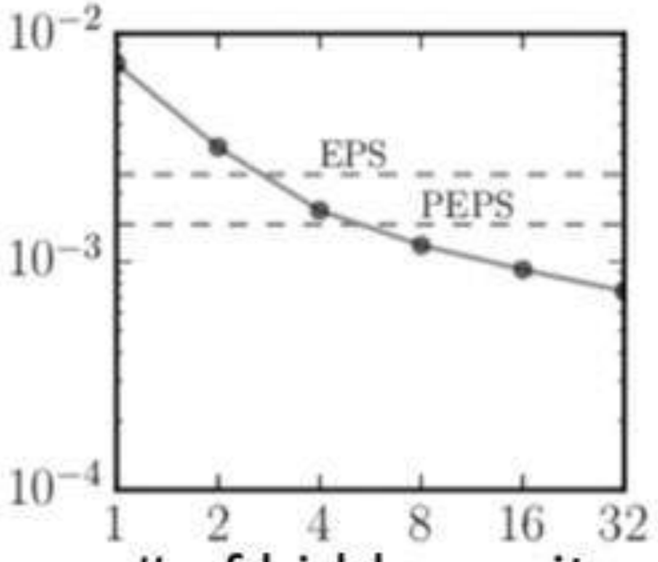


0-3 Better ansatz?

Neural states may beat conventional ones.

Ex) 2-dimensional antiferromagnetic Heisenberg model [Carleo Troyer `17]

Energy with RBM states



# of hidden units

Relations among them:

1) Boltzmann machine states are tensor network states

[Chen, Cheng, Xie, Wang, Xiang `18]

2) Tensor states are deep Boltzmann

[Gao, Duan `17] [Huang, Moore `17]

3) Tensor states are feedforward with "product pooling"

[Cohen, Shashua `18]

0-1 Networks in matter and gravity

Quantum Matter Wave function

- Matrix product states
- Tensor network states
  - MERA
- Neural network states
  - Boltzmann machine
  - Feedforward

Quantum Gravity

- String theory
- Holographic principle
  - Quantum codes
- Regge calculus
  - Causal dynamical triangulation

0-1 Networks in matter and gravity

Quantum Gravity

- String theory

- Holographic principle

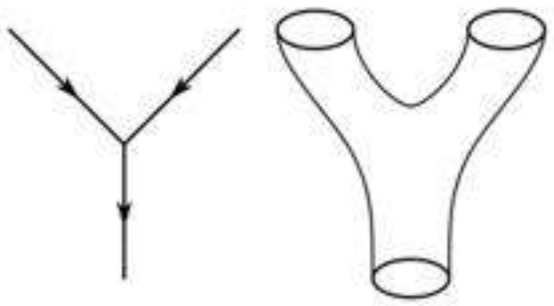
Quantum codes

- Regge calculus

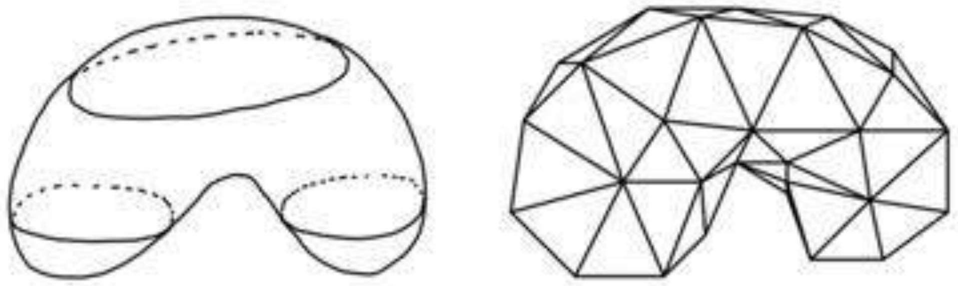
Causal dynamical triangulation

0-5 Quantum gravity : the final frontier

String motion derives Einstein eq. around flat geometry
[Scherk,Schwarz 1974] [Yoneya 1974]



Spacetime is discretized to collected simplices
[Regge 1961] [Ambjorn Loll 1998]



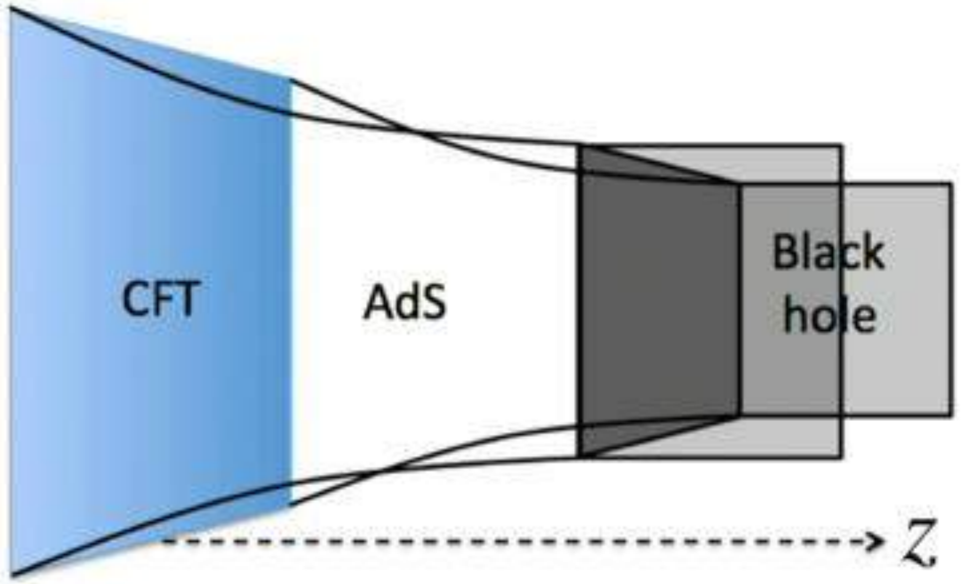
Quantum Gravity

- String theory
- Holographic principle
  Quantum codes
- Regge calculus
  Causal dynamical triangulation

0-6 Quantum gravity : the final frontier

Holographic principle
("AdS/CFT correspondence")
defines quantum gravity by
quantum matter

[Maldacena 1997]



$$Z_{\mathrm{QFT}}[J] = \int_{\phi(z=0)=J} \mathcal{D}\phi \, \exp\left(-S_{\mathrm{gravity}}[\phi]\right)$$

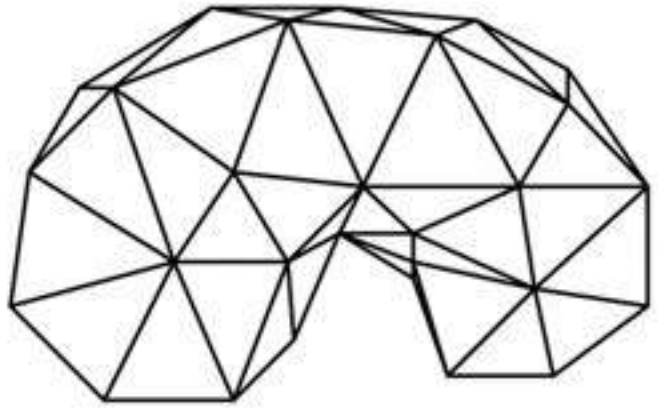
Quantum Gravity

- String theory
- Holographic principle
  Quantum codes
- Regge calculus
  Causal dynamical triangulation

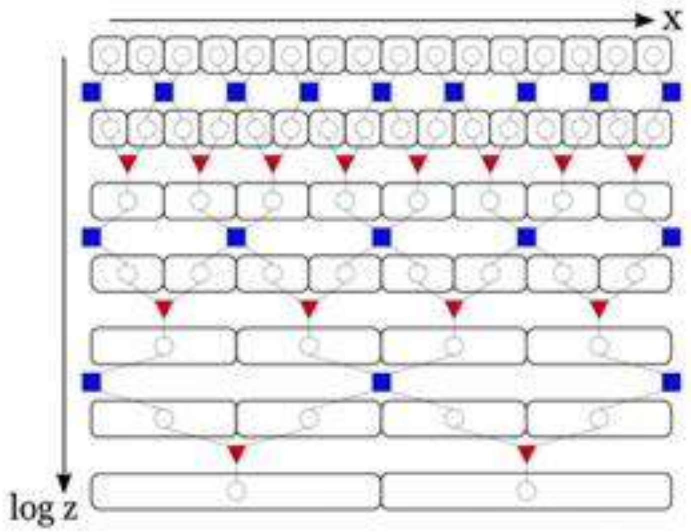
0-7 Quantum gravity is network

Popular examples of networking quantum gravity

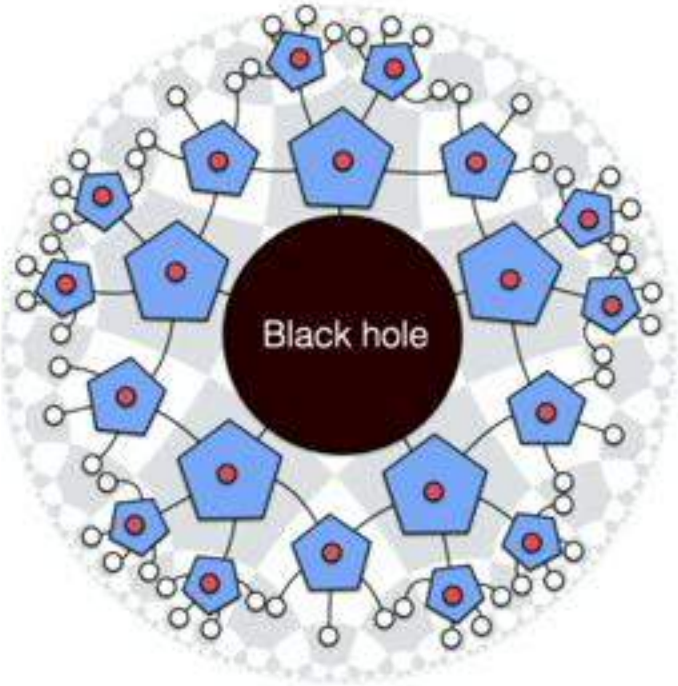
Regge calculus
[Regge 1961]
Causal dynamical triangulation
[Ambjorn, Loll 1998]



AdS/MERA
[Swingle `09]



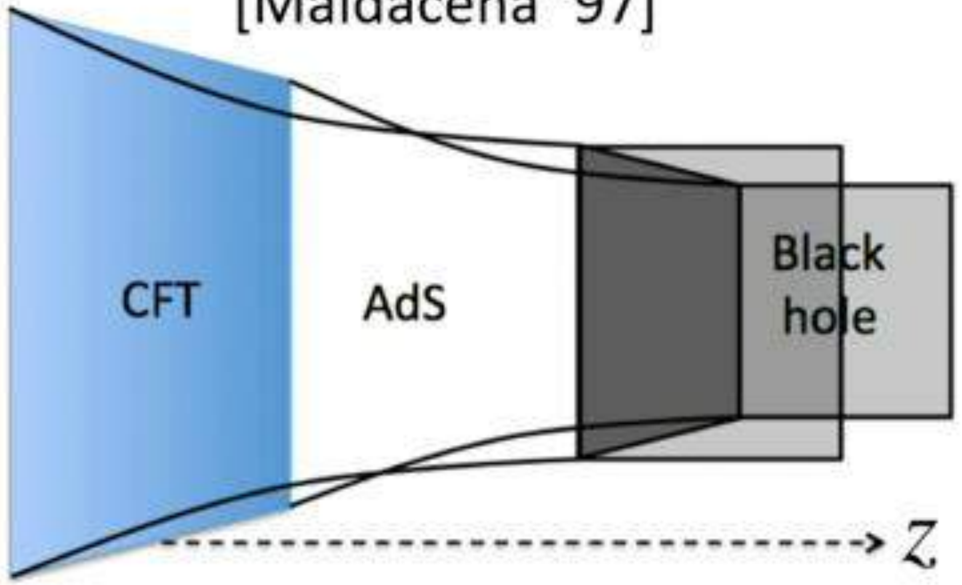
Quantum codes for holography
[Pastawski, Yoshida, Harlow, Preskill `15]



0-8 Neural network as a spacetime?

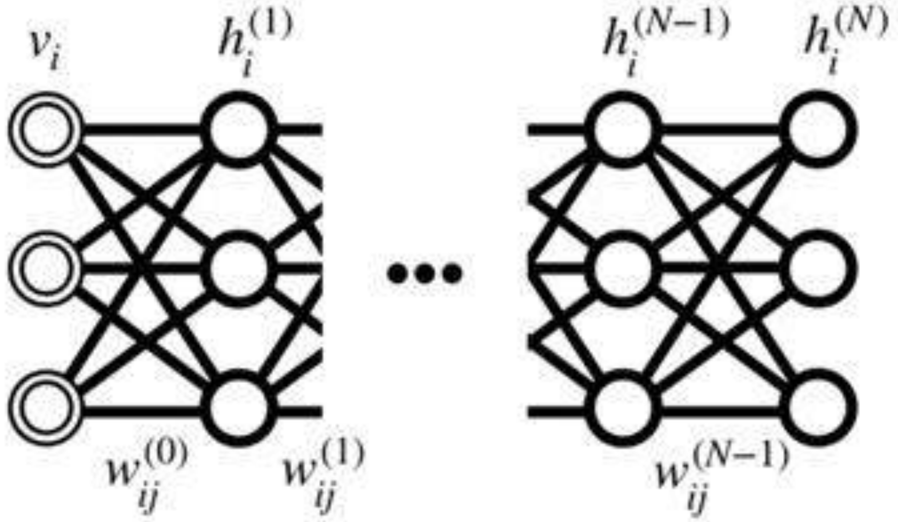
Holographic principle

[Maldacena `97]



Deep Boltzmann machine

[Salakhutdinov, Hinton `09]



$$Z_{\mathrm{QFT}}[J] = \int_{\phi(z=0)=J} \mathcal{D}\phi \, \exp\left(-S_{\mathrm{gravity}}[\phi]\right)$$

$$P(v_i) = \sum_{h_i \in \{0,1\}} \exp\left[-\mathcal{E}(v_i, h_i)\right]$$

AdS/CFT	Deep Boltzmann machine
Bulk coordinate z	Hidden layer label k
QFT source $J(x)$	Input value v_i
Bulk field $\phi(x, z)$	Hidden variables $h_i^{(k)}$
QFT generating function $Z[J]$	Provability distribution $P(v_i)$
Bulk action $S[\phi]$	Energy function $\mathcal{E}(v_i, h_i^{(k)})$

[KH `19]
[Hu,Li,Wang, You `19]

Day 2 | Friday, April 26 1:30 PM–3:00 PM **Breakout group**

Breakout Group

Machine Learning Holography

Koji Hashimoto, Yi-Zhuang You

8 pages

0. Networks in matter/gravity

5 pages

1. Holography as a deep learning

3 pages

2. Application: Quark matter

1-1 Wishes : Using NN for matter/gravity?

Machine learning may be good because …

- Many many network ansatz, proposed.
- Optimization of network parameters.
  (Efficient computational methods, invented)
- Solving inverse problems.

$$A = f(B)$$

Normal problems:
System "f" given, calculate A for given B.
Inverse problems:
For given many data (A,B), find "f".

Emergent spacetime for given matter?

1-3 Holography: matter response from gravity

[Klebanov, Witten 1998]

Classical scalar field probing 5-dim. curved spacetime

$$S = \int d\eta dx^4 \sqrt{\det g} \left[(\partial_\eta \phi)^2 - V(\phi) \right]$$

$$ds^2 = -f(\eta)dt^2 + d\eta^2 + g(\eta)(dx_1^2 + \cdots + dx_{d-1}^2)$$

- AdS boundary ($\eta \sim \infty$) : $f \sim g \sim \exp[2\eta/L]$
- Black hole horizon ($\eta \sim 0$) : $f \sim \eta^2,\ g \sim \text{const.}$

Solve eq. of motion to get response $\langle \bar{\psi}\psi \rangle_{m_q}$

- AdS boundary ($\eta \sim \infty$) : $\phi = m_q e^{-\eta} + \langle \bar{\psi}\psi \rangle e^{-3\eta}$
- Black hole horizon ($\eta \sim 0$) : $\partial_\eta \phi \big|_{\eta=0} = 0$

1-4 Neural network for the spacetime

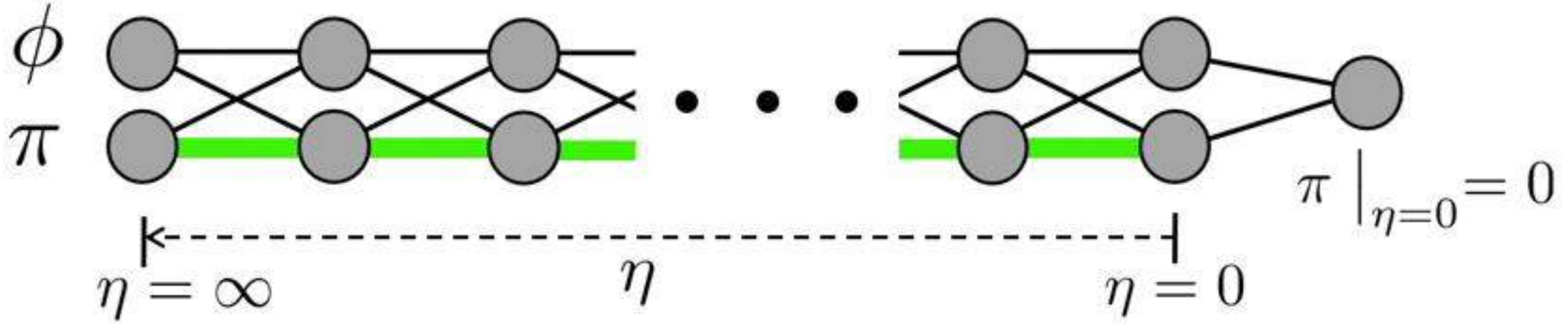
Eq. of motion $\partial_\eta^2 \phi + \underline{\underline{h(\eta)}} \partial_\eta \phi - \frac{\delta V[\phi]}{\delta \phi} = 0$

metric $h(\eta) \equiv \partial_\eta \left[\log \sqrt{f(\eta) g(\eta)^{d-1}} \right]$

Discretization, Hamilton form

$$\begin{cases} \phi(\eta + \Delta\eta) = \phi(\eta) + \Delta\eta \, \pi(\eta) \\ \pi(\eta + \Delta\eta) = \pi(\eta) + \Delta\eta \left(h(\eta)\pi(\eta) - \frac{\delta V(\phi(\eta))}{\delta \phi(\eta)} \right) \end{cases}$$

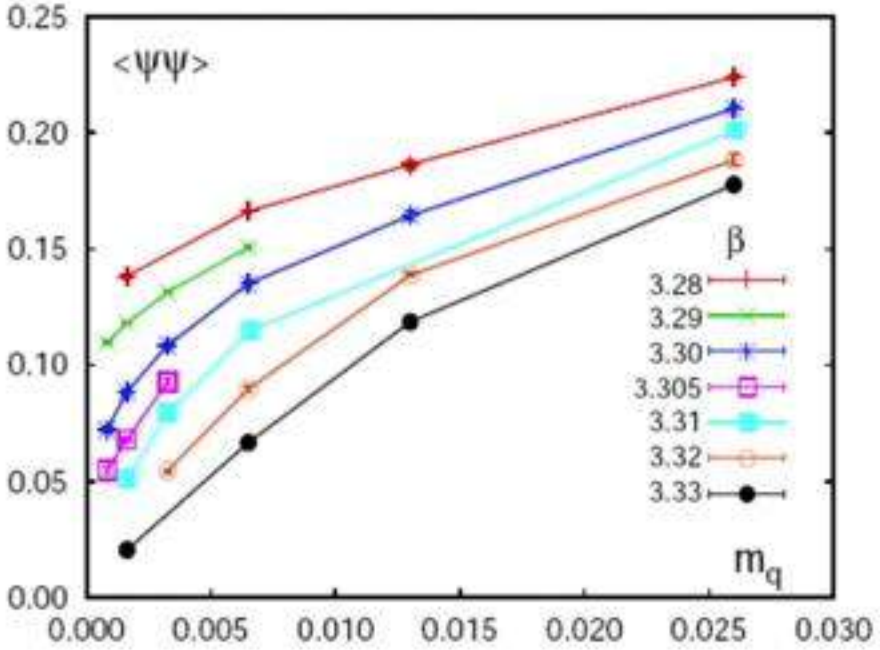
Feedfoward neural network representation



2-1 Application : Quark matter

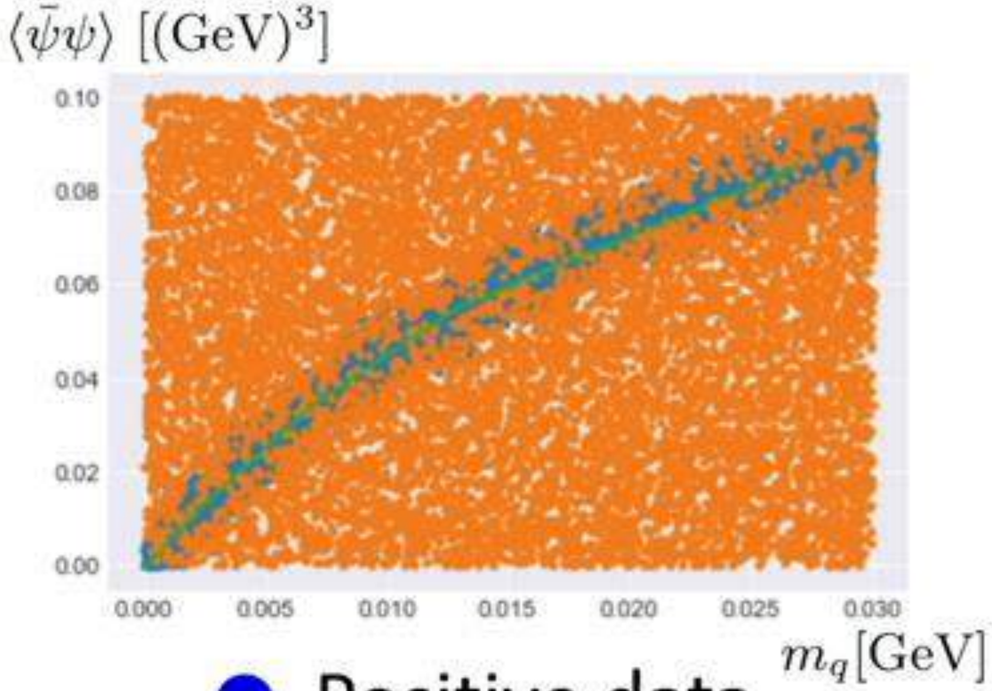
1) Use a QCD data.
2) Let the network learn the metric.
3) Calculate other physical quantities.

Chiral condensate VS quark mass.



β=3.30 ⇔ T=196[MeV]
[RBC-Bielefeld collaboration, 2008]
(Courtesy of W.Unger)

Pick up
β=3.33
data

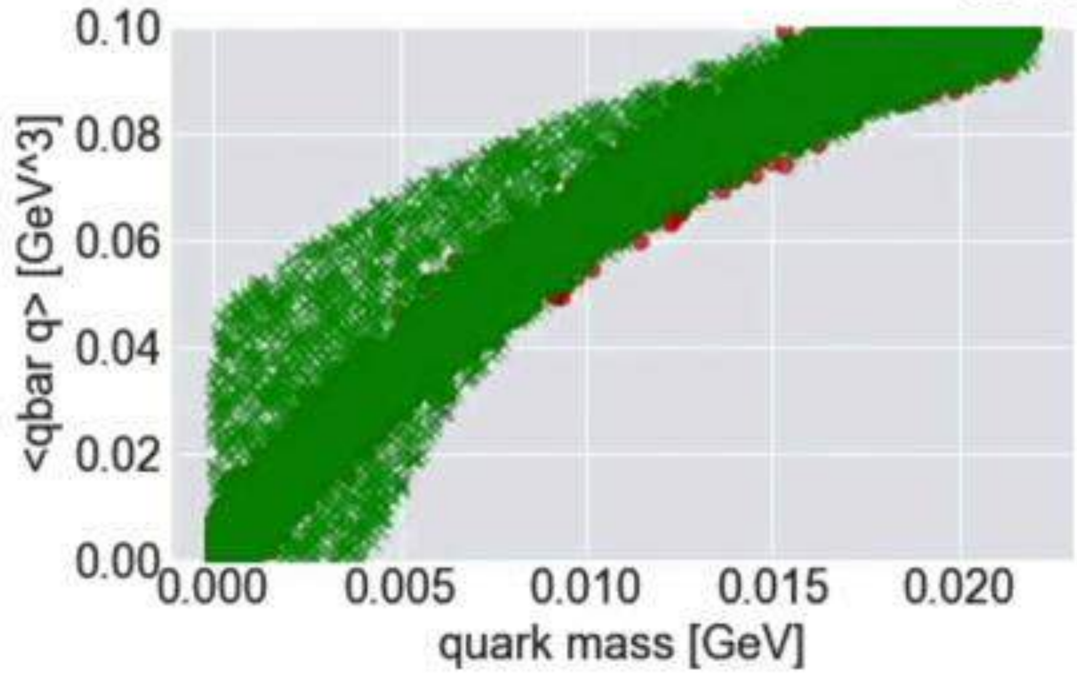


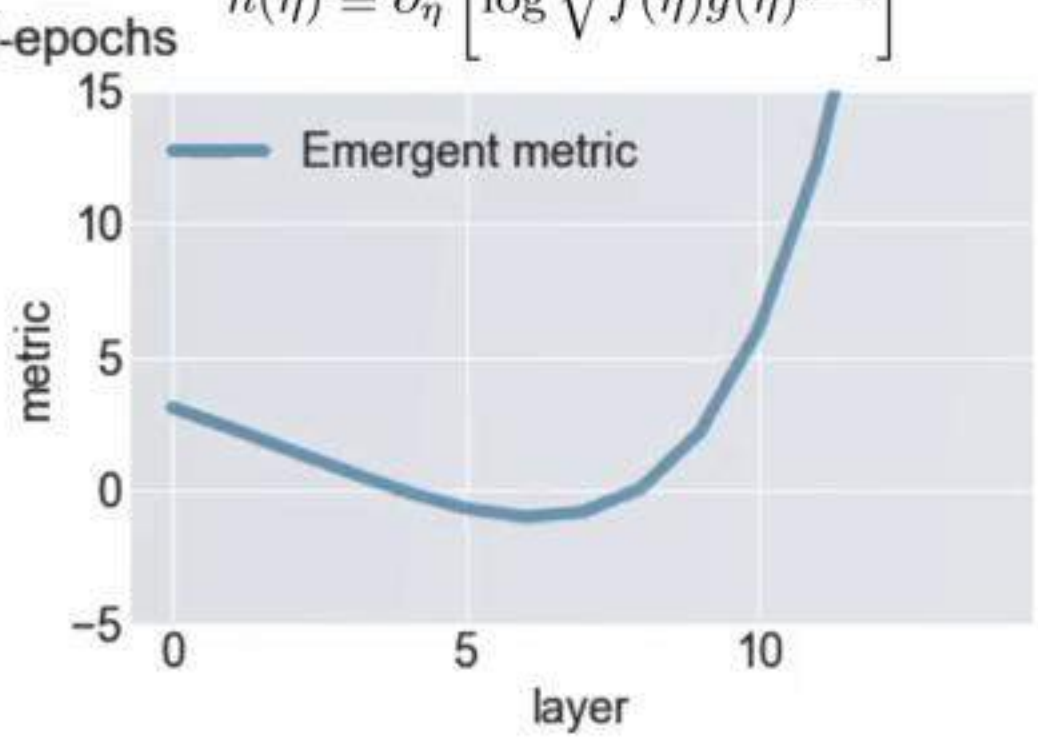
Positive data
Negative data

2-2 Application : Quark matter

1) Use a QCD data.
2) Let the network learn the metric.
3) Calculate other physical quantities.

$$h(\eta) \equiv \partial_\eta \left[\log \sqrt{f(\eta) g(\eta)^{d-1}} \right]$$

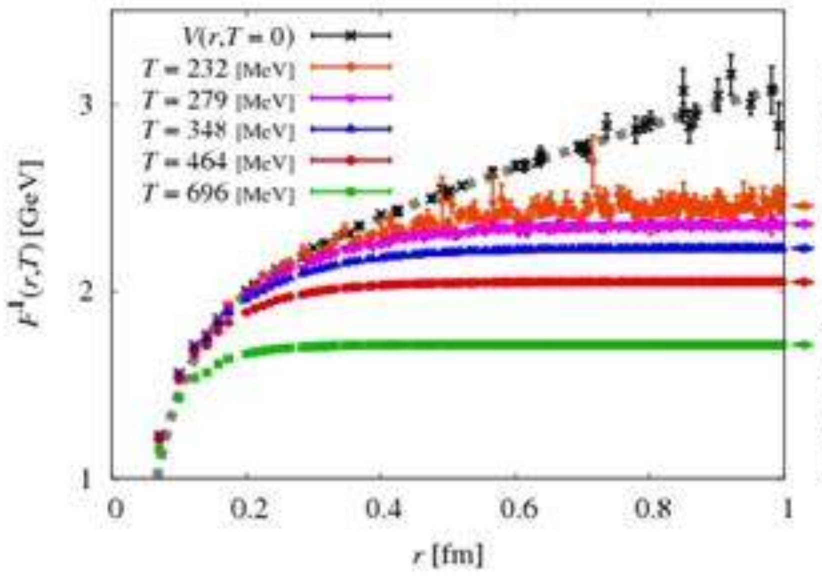




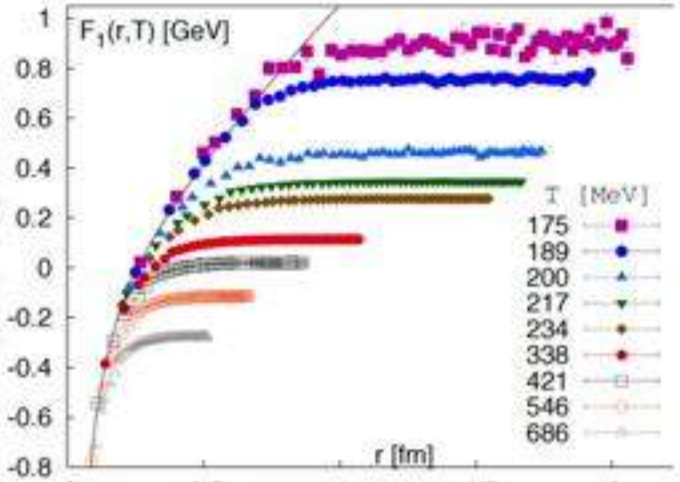
2-3 Application : Quark matter

1) Use a QCD data.
2) Let the network learn the metric.
3) Calculate other physical quantities.

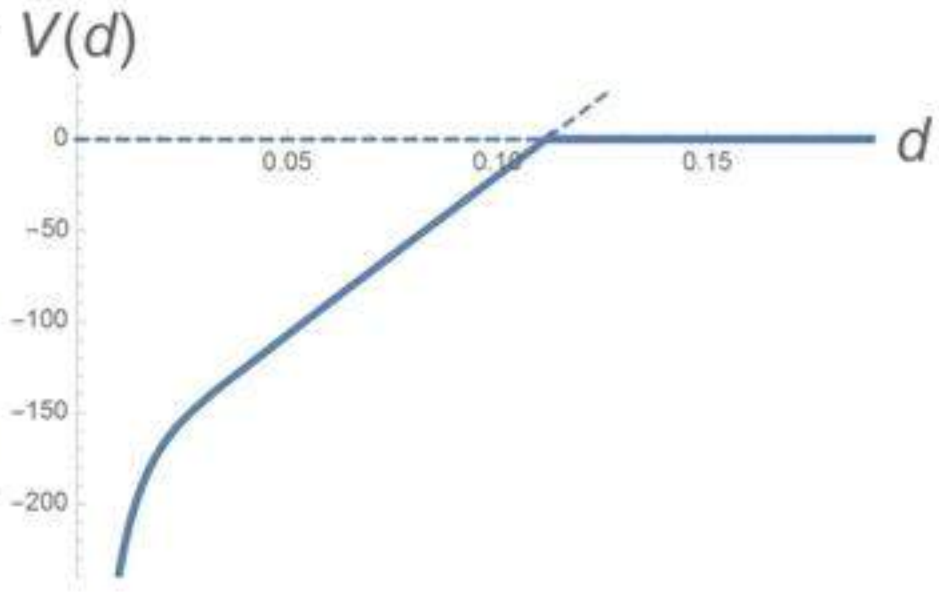
Quark potential



[T.Ishikawa et al., `08,
CPPACS + JLQCD collaboration]



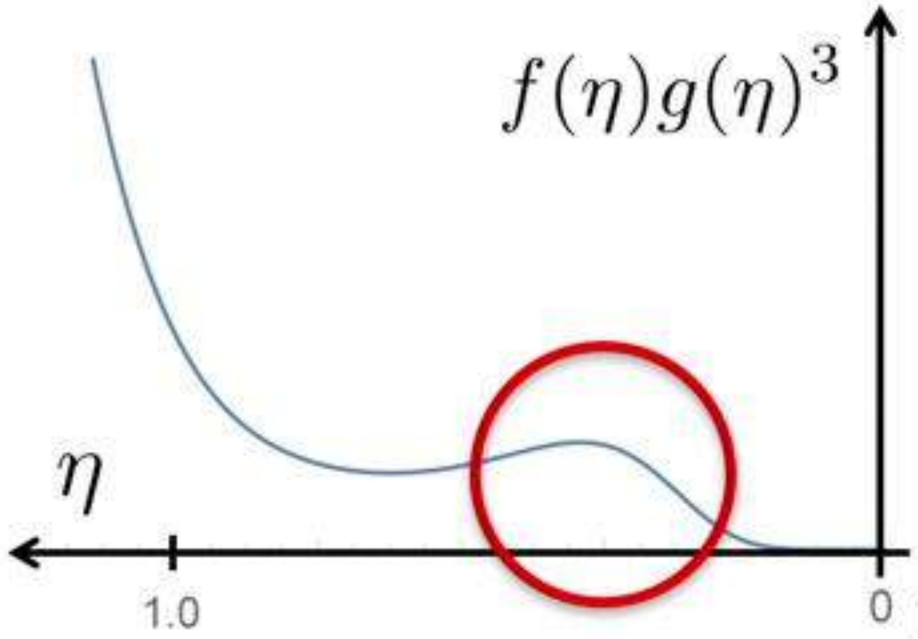
[Petreczky, `10]



2-3 Application : Quark matter

1) Use a QCD data.
2) Let the network learn the metric.
3) Calculate other physical quantities.

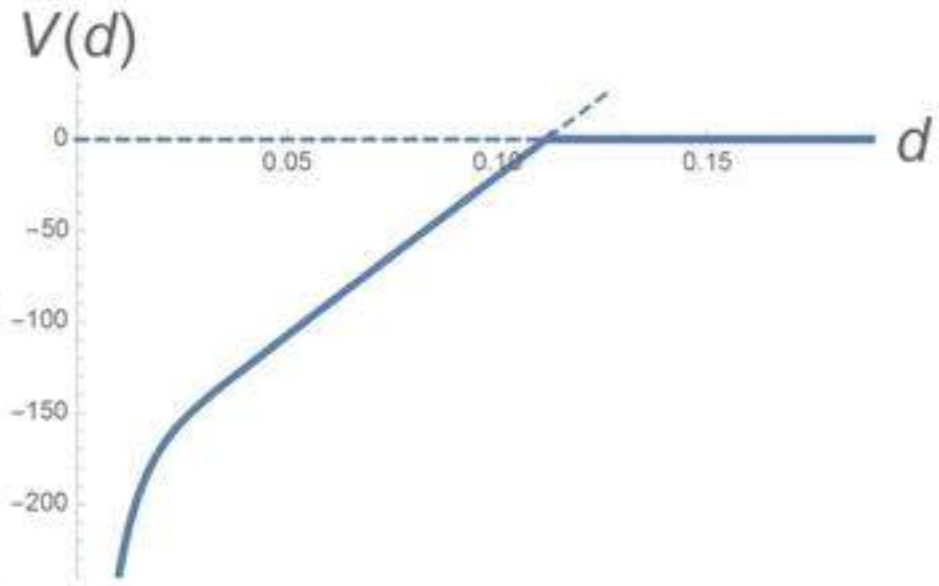
Learned metric



Bump

Procedures based on [Maldacena '98] [Rey,Theisen,Yee '98]

Quark potential



Quantum gravity effect? Cf [Hyakutake 2014]