NEURAL NETWORK LEARNING AND QUANTUM GRAVITY

STEFANO LANZA UNIVERSITY OF HAMBURG

BASED ON:
ARXIV:2403.03245,
ARXIV:2311.03437

WHY WE NEED MACHINE LEARNING IN QUANTUM GRAVITY

The number of consistent string theory vacua ranges from $\sim 10^{500}$ to $\sim 10^{272,000}$.

$\Rightarrow$ The string landscape is **too vast** for a systematic exploration!

With Machine Learning we can study **big data sets** of consistent theories and extract relevant information

This could be beneficial for the **Swampland Program**:

- Test existent conjectures with **supervised techniques**
- Uncover new patterns with **unsupervised techniques**

WHAT CAN WE 'LEARN' IN QUANTUM GRAVITY?

But what can we '**learn**' in

Quantum Gravity effective field theories?

WHAT CAN WE 'LEARN' IN QUANTUM GRAVITY?

But what can we '**learn**' in

Quantum Gravity effective field theories?

We address this question with **three key tools**:

- A mathematical definition of Learning
- The shattering dimensions
- o-minimal structures, defining Quantum Gravity interactions

WHAT CAN WE 'LEARN' IN QUANTUM GRAVITY?

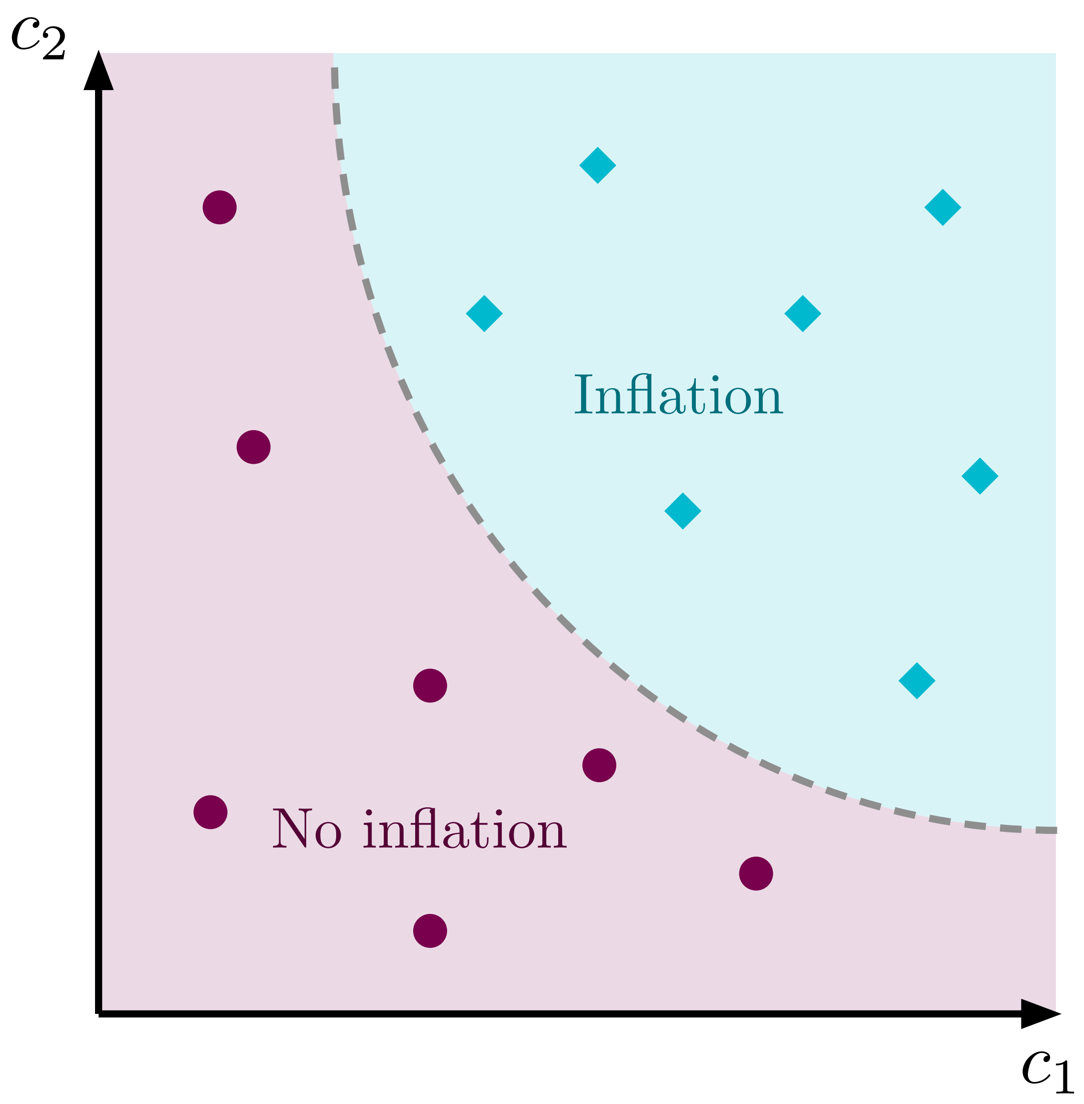
But what can we '**learn**' in

Quantum Gravity effective field theories?

Here, we will focus only on **binary supervised problems**, namely problems that can be answered with a 'yes' or a 'no'.

Example:

Consider an effective theory with a scalar potential $V(\varphi, c)$. Does it accommodate slow-roll inflation for some parameters c^a?



1. A DEFINITION FROM STATISTICAL LEARNING THEORY

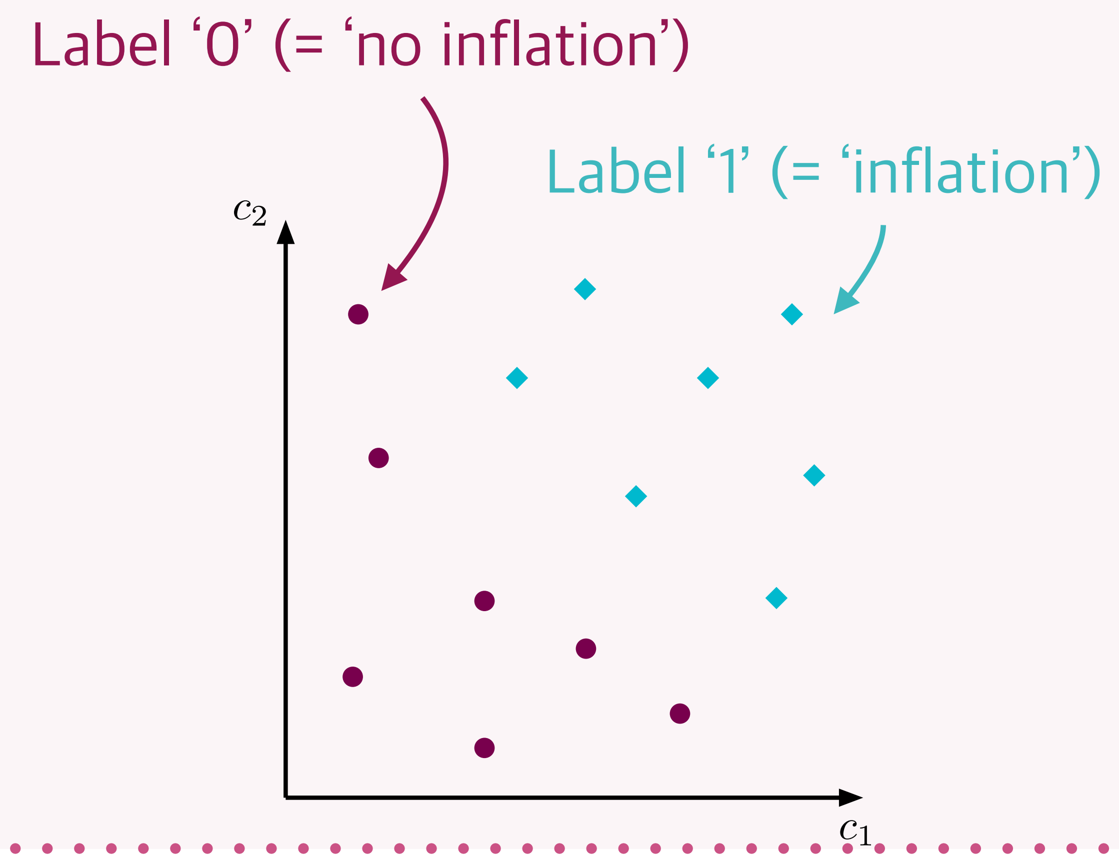
1. A DEFINITION FROM STATISTICAL LEARNING THEORY

- The Data space is $Z = X \times Y$ with

  $X =$ input (coordinate) space $\subset \mathbb{R}^n$, $Y =$ output (label) space $= \{0,1\}$

- The Data set fed to the algorithm is a discrete subset of the Data space:

  Data set $= \{(x_i, y_i)\} \subset X \times Y \subset \mathbb{R}^n \times \{0,1\}$, $\quad i = 1,\ldots,N_{\text{data}}$



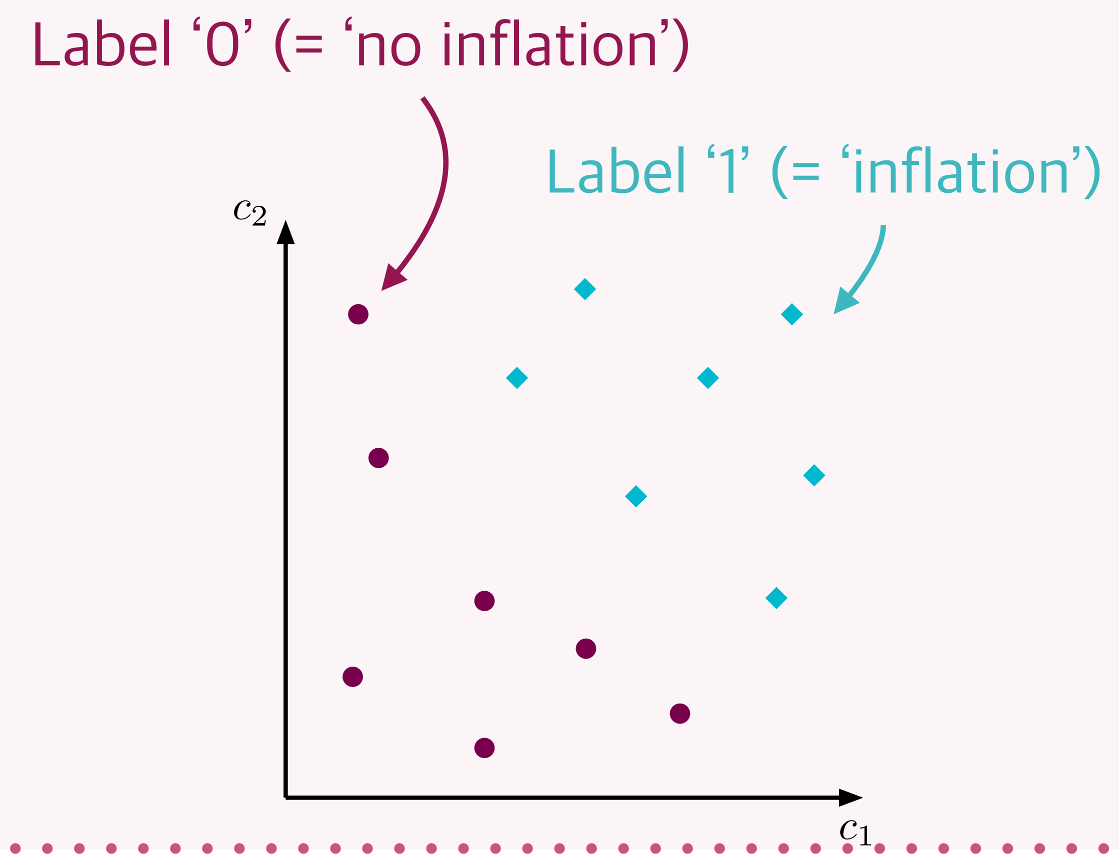
1. A DEFINITION FROM STATISTICAL LEARNING THEORY

➤ The Data space is $Z = X \times Y$ with

$X =$ input (coordinate) space $\subset \mathbb{R}^n$, $Y =$ output (label) space $= \{0,1\}$

➤ The Data set fed to the algorithm is a discrete subset of the Data space:

Data set $= \{(x_i, y_i)\} \subset X \times Y \subset \mathbb{R}^n \times \{0,1\}$, $\quad i = 1,\ldots,N_{\text{data}}$



➤ Introduce a set of functions $\mathscr{F}$ among which we search for a function

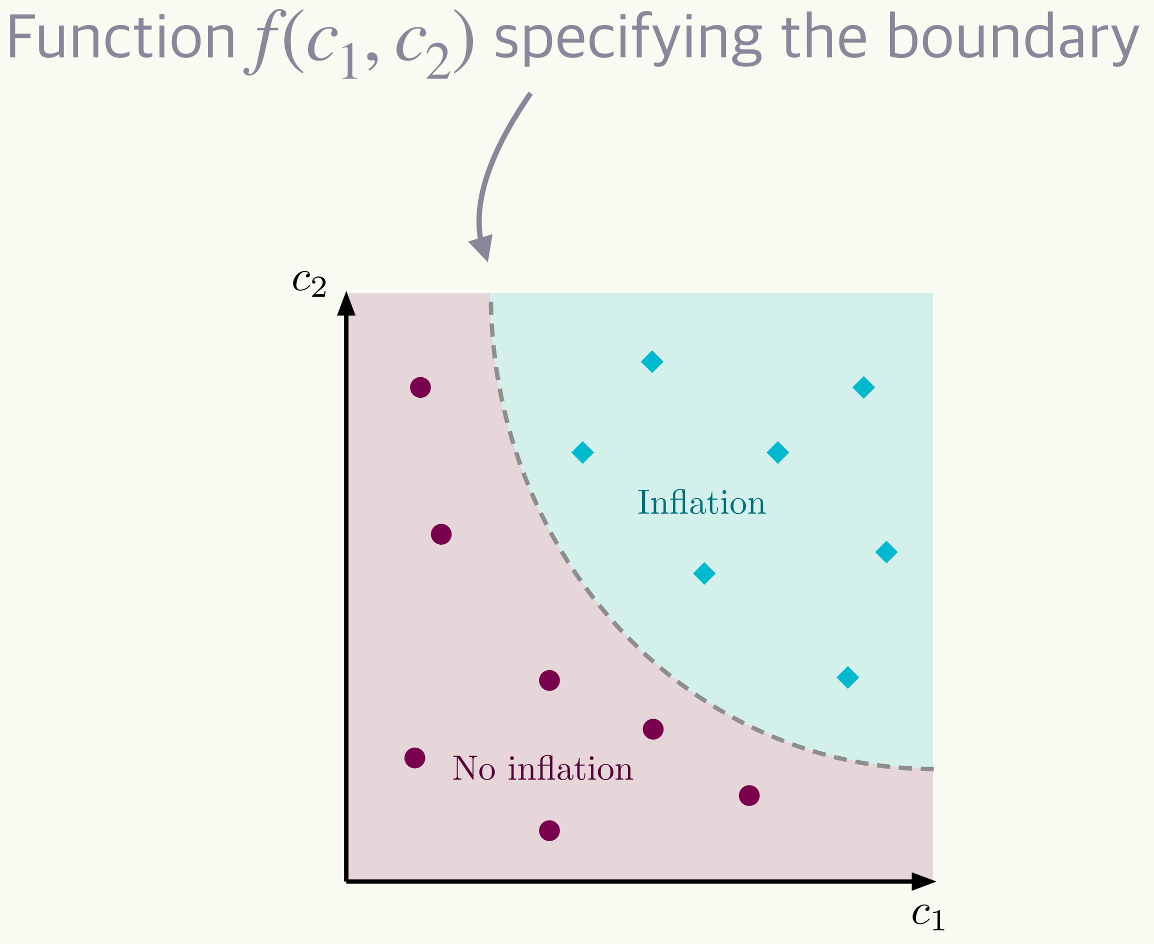
$$f(x) : X \to Y$$

that best models the data, with the smallest possible error.

➤ A **learning algorithm** ℓ is a map

$$\ell \; : \text{Data set}_m \to \mathscr{F}$$

that selects a function f, within $\mathscr{F}$, with the property that, if m is sufficiently large, then the error of $f(x)$ is small enough, $\text{error}_P(\ell) < \text{opt}_P(\mathscr{F}) + \epsilon$, with high enough probability.



2. THE VAPNIK-CHERVONENKIS DIMENSION

Consider a family of sets $\mathscr{C}$ and a set S. We say that the set S is '**shattered**' by the sets of the family $\mathscr{C}$ if all the subsets of S are contained in $\mathscr{C}$, namely:

$$\mathscr{C} \cap S := \{C \cap S \mid C \in \mathscr{C}\} \text{ contains all the subsets of } S$$

The **Vapnik-Chervonenkis dimension** of $\mathscr{C}$ is the cardinality of the largest set that $\mathscr{C}$ can shatter.

If any set can be shattered by the family of sets $\mathscr{C}$, we say that the Vapnik-Chervonenkis dimension of $\mathscr{C}$ is infinite.

2. THE VAPNIK-CHERVONENKIS DIMENSION

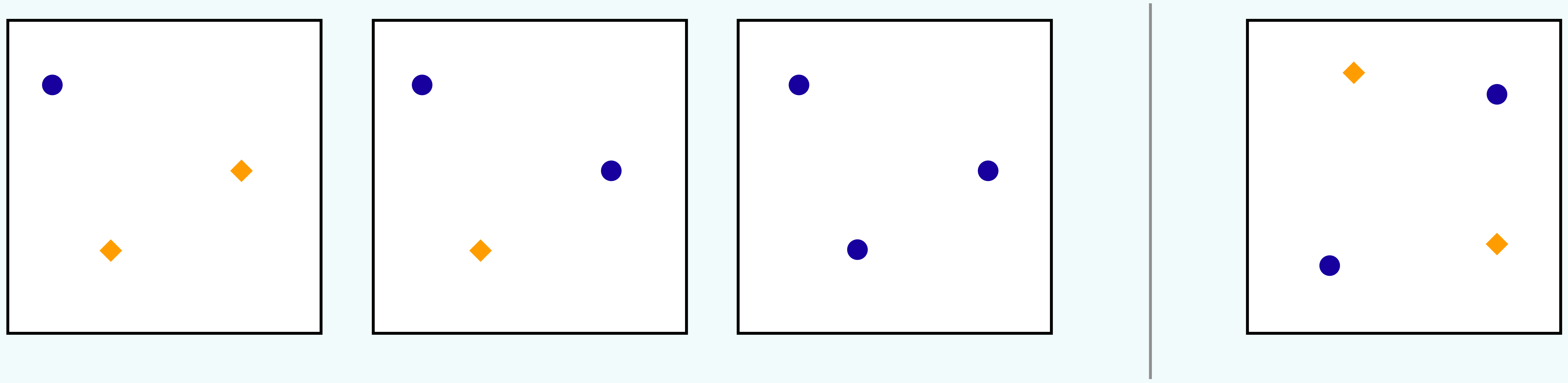
Consider a family of sets $\mathscr{C}$ and a set S. We say that the set S is '**shattered**' by the sets of the family $\mathscr{C}$ if all the subsets of S are contained in $\mathscr{C}$, namely:

$$\mathscr{C} \cap S := \{C \cap S \mid C \in \mathscr{C}\} \text{ contains all the subsets of } S$$

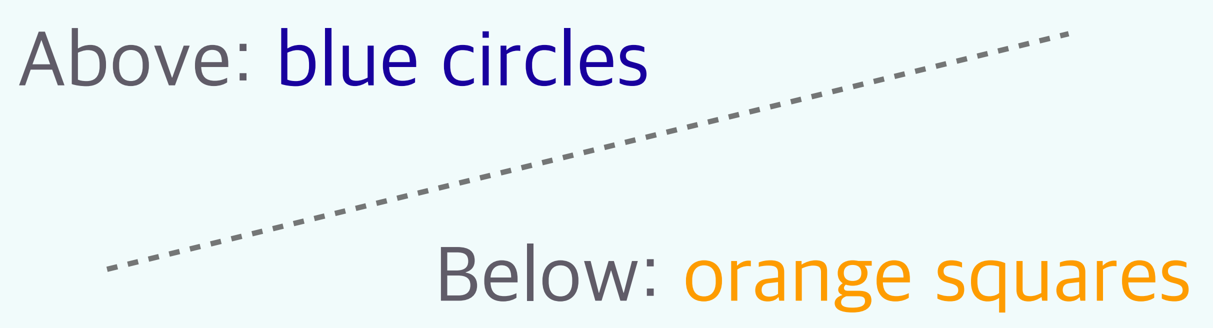
The **Vapnik-Chervonenkis dimension** of $\mathscr{C}$ is the cardinality of the largest set that $\mathscr{C}$ can shatter.

If any set can be shattered by the family of sets $\mathscr{C}$, we say that the Vapnik-Chervonenkis dimension of $\mathscr{C}$ is infinite.

Example: $S =$ set of points, with subsets represented as blue circles and orange squares; we want to 'shatter' its subsets via $\mathscr{C} =$ sets separated by a single line



Our models are lines, such that they 'shatter' the points as



2. THE VAPNIK-CHERVONENKIS DIMENSION

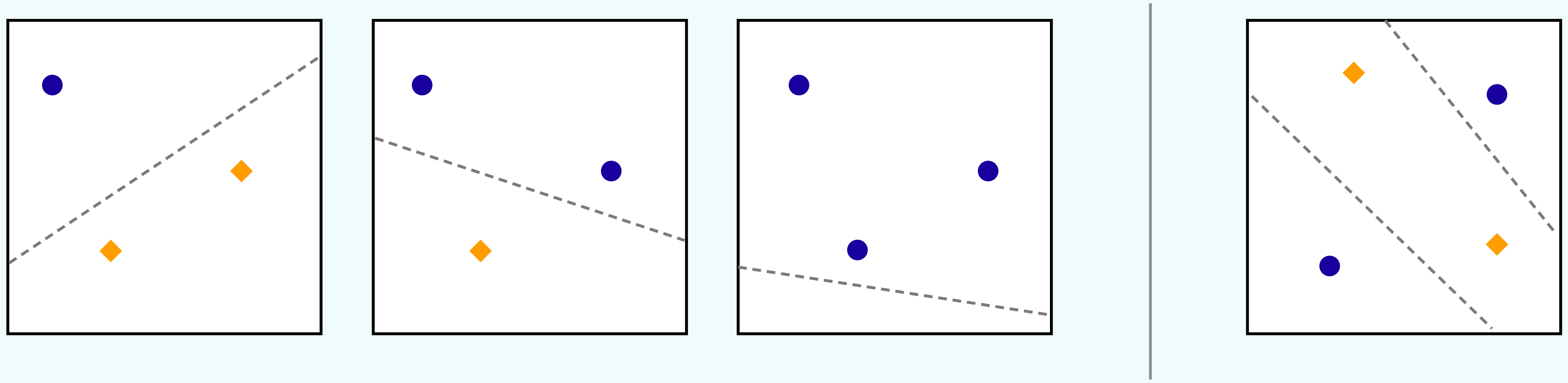
Consider a family of sets $\mathscr{C}$ and a set S. We say that the set S is '**shattered**' by the sets of the family $\mathscr{C}$ if all the subsets of S are contained in $\mathscr{C}$, namely:

$$\mathscr{C} \cap S := \{C \cap S \mid C \in \mathscr{C}\} \text{ contains all the subsets of } S$$

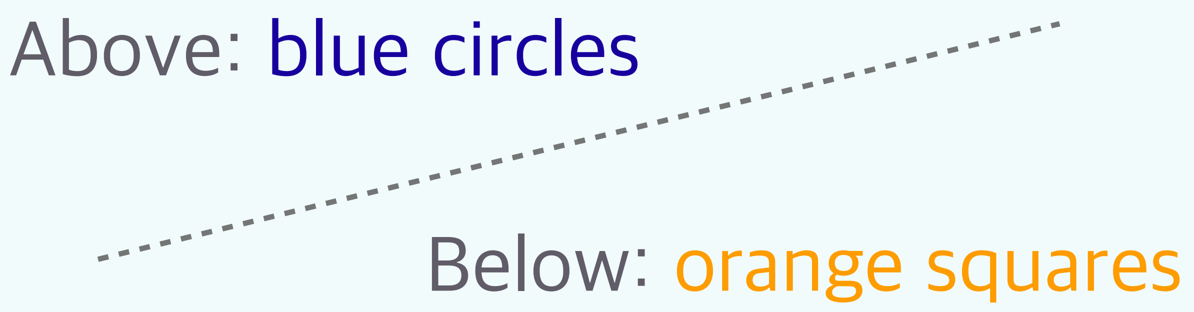
The **Vapnik-Chervonenkis dimension** of $\mathscr{C}$ is the cardinality of the largest set that $\mathscr{C}$ can shatter.

If any set can be shattered by the family of sets $\mathscr{C}$, we say that the Vapnik-Chervonenkis dimension of $\mathscr{C}$ is infinite.

Example: the set of lines can 'shatter' (separate) set of three (non-collinear) points with any label assignment, but not set of four points $\Rightarrow$ The Vapnik–Chervonenkis dimension of the set of lines is three.



Our models are lines, such that they 'shatter' the points as



3. O-MINIMALITY AND QUANTUM GRAVITY

Consider a generic effective field theory of string theory:

$$S^{(D)} = \int \left(\frac{1}{2} R \star 1 - \frac{1}{2} G_{ab}(\varphi, \lambda) \mathrm{d}\varphi^a \wedge \star \mathrm{d}\varphi^b - \frac{1}{2} f_{IJ}(\varphi, \lambda) F^I_{p_I+1} \wedge \star F^J_{p_J+1} - V(\varphi, \lambda) \star 1 + \ldots \right)$$

The **Tameness Conjecture** [Grimm 2021] asserts that all the couplings and the interactions appearing in the effective theory need to be ‘**tame**’, namely definable in a given o-minimal structure.

3. O-MINIMALITY AND QUANTUM GRAVITY

Consider a generic effective field theory of string theory:

$$S^{(D)} = \int \left(\frac{1}{2} R \star 1 - \frac{1}{2} G_{ab}(\varphi, \lambda) \mathrm{d}\varphi^a \wedge \star \mathrm{d}\varphi^b - \frac{1}{2} f_{IJ}(\varphi, \lambda) F^I_{p_I+1} \wedge \star F^J_{p_J+1} - V(\varphi, \lambda) \star 1 + \ldots \right)$$

The **Tameness Conjecture** [Grimm 2021] asserts that all the couplings and the interactions appearing in the effective theory need to be '**tame**', namely definable in a given o-minimal structure.

Namely, any coupling $g(\varphi, \lambda)$ can be written as unions, intersections or complements of a finite number of loci:

$$\exists\, x_1, \ldots, x_l : \qquad P_i(\varphi, \lambda, x, g, f_1, \ldots, f_m) = 0\,, \qquad Q_j(\varphi, \lambda, x, g, f_1, \ldots, f_m) > 0\,,$$

with x_k auxiliary variables, and f_a restricted analytic functions.

3. O-MINIMALITY AND QUANTUM GRAVITY

Consider a generic effective field theory of string theory:

$$S^{(D)} = \int \left(\frac{1}{2} R \star 1 - \frac{1}{2} G_{ab}(\varphi, \lambda) \mathrm{d}\varphi^a \wedge \star \mathrm{d}\varphi^b - \frac{1}{2} f_{IJ}(\varphi, \lambda) F^I_{p_I+1} \wedge \star F^J_{p_J+1} - V(\varphi, \lambda) \star 1 + \dots \right)$$

The **Tameness Conjecture** [Grimm 2021] asserts that all the couplings and the interactions appearing in the effective theory need to be '**tame**', namely definable in a given o-minimal structure.

$\Rightarrow$ Couplings and interactions have regular, monotonic tails and finite critical loci:

COMBINING ALL THE INGREDIENTS

COMBINING ALL THE INGREDIENTS

The tameness of Quantum Gravity effective field theories implies that **the set of functions $\mathscr{F}$** among which the 'true' model resides have to be **definable in a given o-minimal structure**:

$$\Rightarrow f : X \to Y$$ has to have 'regular' tails, and be non-periodic

COMBINING ALL THE INGREDIENTS

The tameness of Quantum Gravity effective field theories implies that **the set of functions $\mathscr{F}$** among which the ‘true’ model resides have to be **definable in a given o-minimal structure**:

$$\Rightarrow f : X \to Y$$ has to have ‘regular’ tails, and be non-periodic

Every o-minimal structure is characterized by **finite Vapnik-Chervonenkis dimension** [Laskowski 1990].

Notice that an infinite shattering dimension would imply an infinite amount of information to learn the function!

THE LEARNABILITY OF QUANTUM GRAVITY

Fundamental Theorem of Statistical Learning

Consider a set of functions $\mathscr{F}$ with binary output, then:

- $\mathscr{F}$ is **learnable** if and only if $\mathscr{F}$ has **finite Vapnik-Chervonenkis dimension**;
- there exist two constants c_1, c_2 such that the **complexity** of the data set is **bounded** as

$$\frac{c_1}{\epsilon^2} \log \frac{1}{\delta} < \min_{\ell} \, m < \frac{c_2}{\epsilon^2} \log \frac{1}{\delta}$$

THE LEARNABILITY OF QUANTUM GRAVITY

Fundamental Theorem of Statistical Learning

Consider a set of functions $\mathscr{F}$ with binary output, then:

- $\mathscr{F}$ is **learnable** if and only if $\mathscr{F}$ has **finite Vapnik-Chervonenkis dimension**;
- there exist two constants c_1, c_2 such that the **complexity** of the data set is **bounded** as

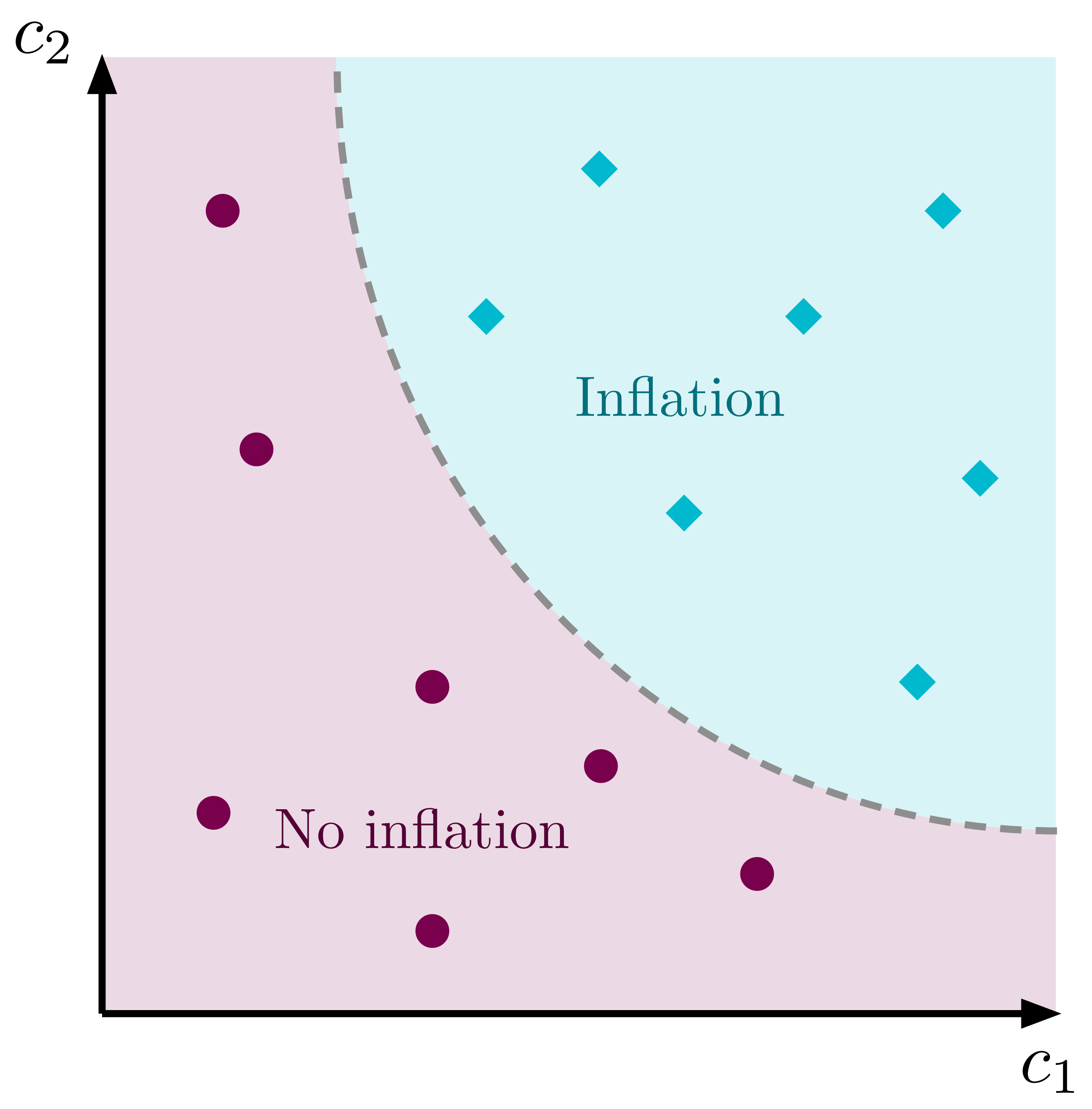
$$\frac{c_1}{\epsilon^2} \log \frac{1}{\delta} < \min_{\ell} m < \frac{c_2}{\epsilon^2} \log \frac{1}{\delta}$$

⇒ In every low-energy effective theory of Quantum Gravity, **any binary classification problem** (involving the tame interactions of the theory) **is learnable**.

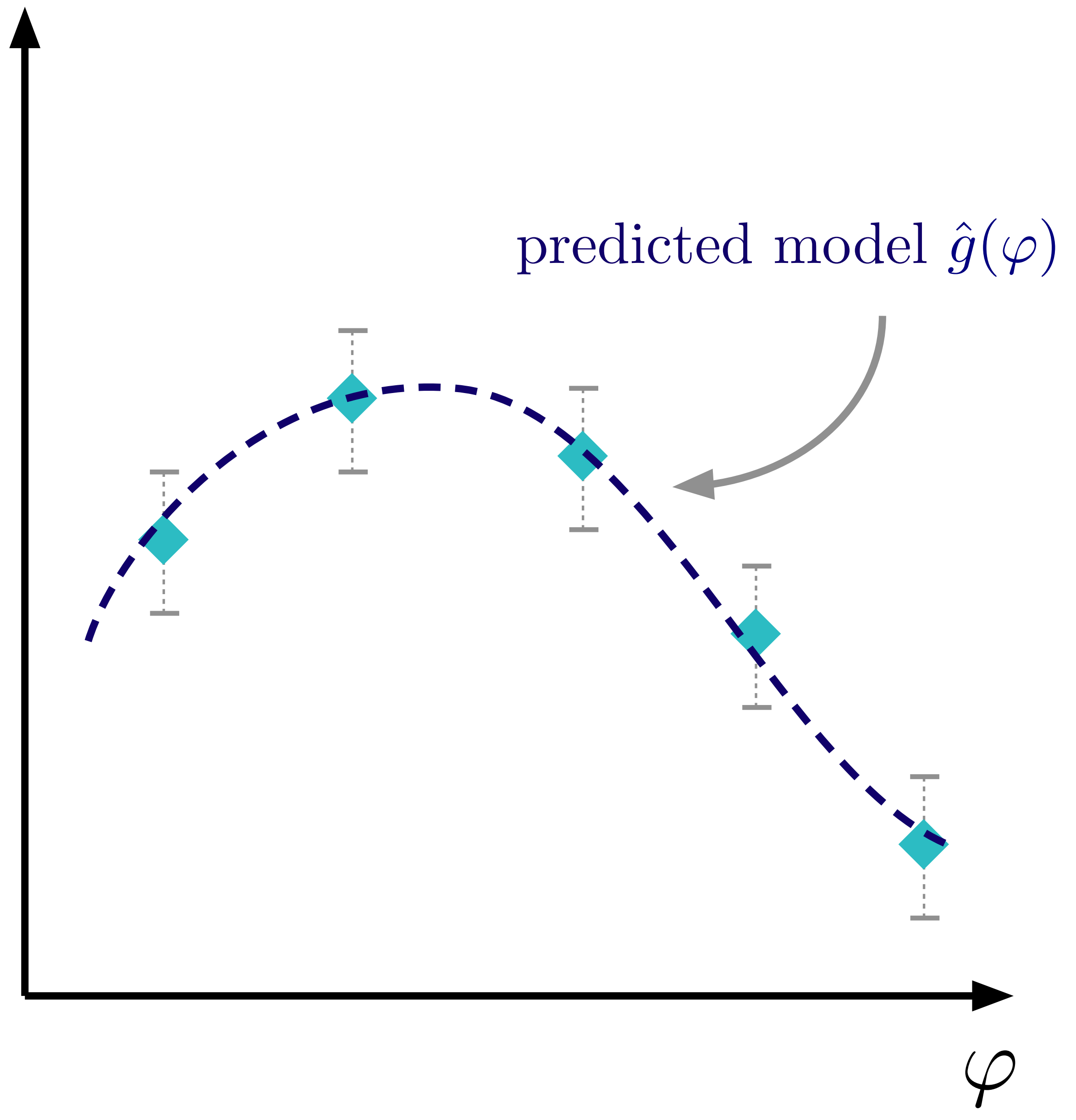
Similarly, also more general Quantum Gravity **regression problems** can be shown to be learnable.

EXAMPLE OF LEARNING PROBLEMS

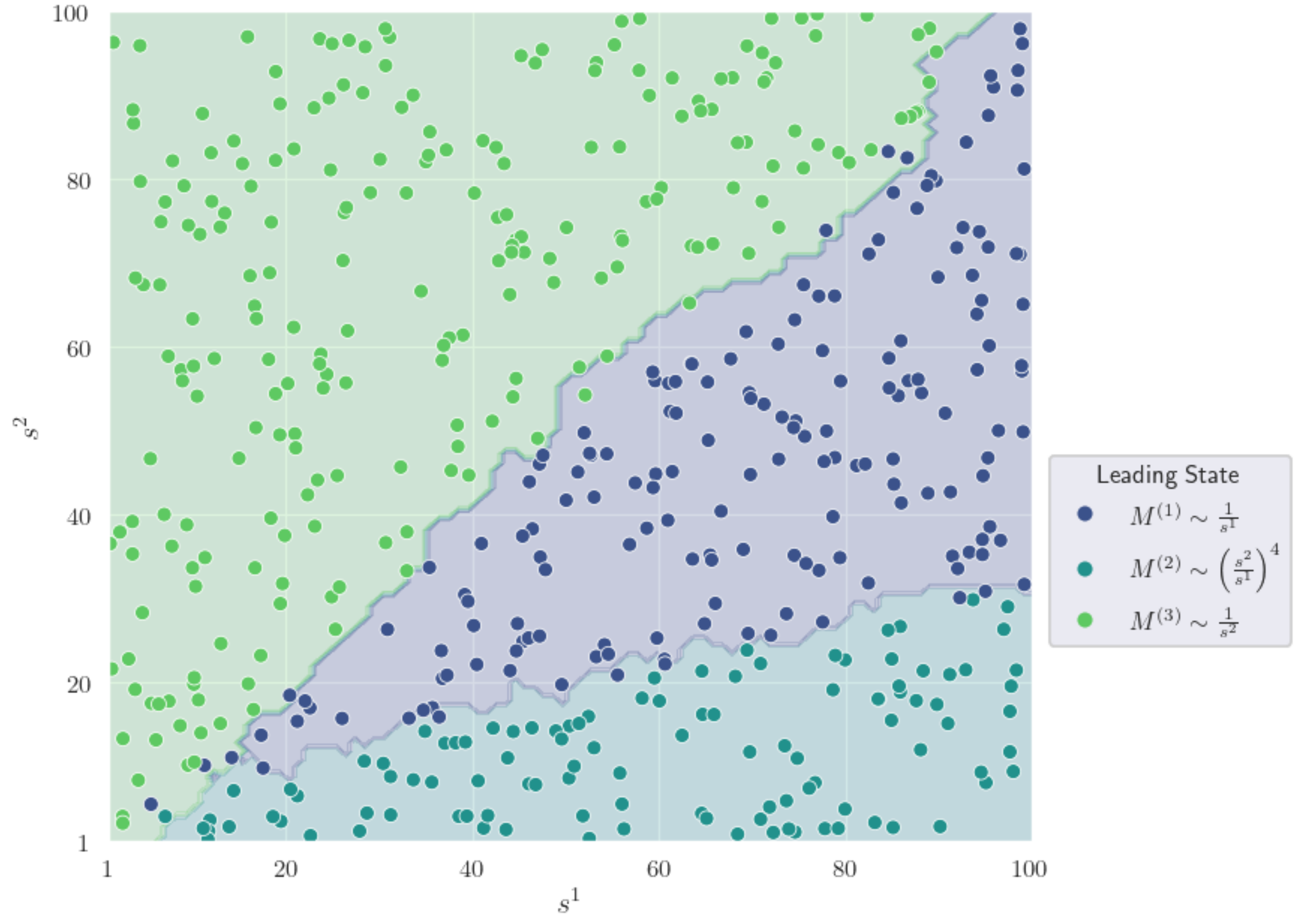
Can an EFT support inflation for some choice of parameters?



What is the predicted value of a gauge coupling?



Which kinds of states first break down an EFT?



ARXIV:2311.03437

CONCLUSIONS AND OUTLOOK

- We have seen that **binary classification problems** formulated within Quantum Gravity effective field theories are **learnable**;
- Similarly, more **general regression** or **interpolation problems** can be shown to be **learnable.**

⇒ The usage of neural network to machine-learn critical properties of Quantum Gravity is justified by the geometrical structures of Quantum Gravity.

It would be interesting to further investigate:

- the learnability via **unsupervised techniques** and the **Swampland Conjectures;**
- relations with **decidability** and the **halting problem**.

Thank you!