On Generation in Metric Spaces

Jiaxun Li
Department of Statistics
University of Michigan
Ann Arbor, MI 48109
jasonli@umich.edu

Vinod Raman
Department of Statistics
University of Michigan
Ann Arbor, MI 48109
vkraman@umich.edu

Ambuj Tewari
Department of Statistics
University of Michigan
Ann Arbor, MI 48109
tewaria@umich.edu

Abstract

We study generation in separable metric instance spaces. We extend the generation framework from Kleinberg and Mullainathan [13] beyond countable domains by defining novelty through metric separation and allowing asymmetric novelty parameters for the adversary and the generator. We introduce the $(\varepsilon_a, \varepsilon_g)$-closure dimension, a scale-sensitive analogue of closure dimension, which yields characterizations of uniform and non-uniform generatability. We also show that the union of two uniformly generatable classes can fail to be generatable in the limit, even on the real line. Along the way, we identify a sharp geometric contrast. Namely, in doubling spaces, including all finite-dimensional normed spaces, generatability is stable across novelty scales and invariant under equivalent metrics. In general metric spaces, however, generatability can be highly scale-sensitive and metric-dependent; even in the natural infinite-dimensional Hilbert space ℓ^2, all notions of generation may fail abruptly as the novelty parameters vary.

1 Introduction

Generative models are trained on massive amounts of data from which they learn to generate new, previously unseen, well-formed objects such as text, images, code, and molecules. A clean, minimal framework has emerged in recent work [13, 21] that captures the essence of generation. In this model, the learner receives a sequence of objects from an adversary drawn from a set in a known set collection. The learner has to generate novel valid objects from the true set that is initially unknown to the learner. This framework is inspired by the seminal work of Gold [7] on language identification in the limit. In Gold [7]'s model, an adversary enumerates strings from a language in a collection. The learner's task is to identify the true language from only the positive examples shown by the adversary. Kleinberg and Mullainathan [13] changed the identification requirement to generation: the learner has to generate novel, unseen objects from the true set.

This simple change, from identification to generation, completely changes the nature of provable guarantees. On the one hand, since the seminal characterization by Angluin [3] of language collections identifiable in the Gold [7] model, we know that very few natural collections are identifiable in the limit. On the other hand, Kleinberg and Mullainathan [13] showed that *any* countable collection is generatable in the limit. They also connected their theoretical findings to empirical phenomena such as under-generation (mode collapse) and over-generation (hallucination). Their elegant framework and surprising results have led to a flurry of work on extending their basic framework in various interesting directions (see Sec. 1.2).

However, most work in this area assumes a countable example space with a discrete 0-1 metric. This implies that the generated object is considered novel if it is distinct from previously seen objects. This notion is often too weak: in image generation, changing a single pixel of an image will not necessarily produce a genuinely new image, and in text generation, replacing one word in a sentence will not necessarily produce a genuinely new sentence.

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

In this paper we initiate a theoretical study of *generation in metric spaces*. A metric space structure is general enough to encompass not only traditional modalities like text and images but also emerging applications such as SDE solutions over a spatial domain [16] and protein generation in 3D space that respects translational and rotational symmetries [9].

Motivated by practice, we operationalize novelty by requiring that the distance of generated objects to seen objects must lie above a given threshold ε_g. For example, in protein structure generation, a generated structure is often classified as novel only if its maximum structural similarity to all training structures falls below a fixed threshold, corresponding to a distance-based novelty notion that is relative to the training set [17].

Furthermore, empirical studies of generative model generalization have observed that generated samples may lie closer to the training set than independently held-out test samples in a feature space, reflecting a separation between the scale at which training data are revealed and the scale at which novelty is required of generated outputs [10]. Our framework makes this distinction explicit by allowing separate novelty parameters ε_a, ε_g: one for the adversary, and one for the generator.

1.1 Main Contributions

Our main contributions are as follows:

- We extend the language generation model to separable metric instance spaces $\mathcal{X}$, while retaining binary hypothesis classes $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$. We introduce a notion of novelty based on metric separation rather than mere distinctness, and define two independent novelty parameters $\varepsilon_a, \varepsilon_g$, for the adversary and the generator, respectively. This leads to a strictly richer and nontrivial generalization of the countable setting.
- We introduce the $(\varepsilon_a, \varepsilon_g)$-closure dimension, a metric analogue of closure dimension that explicitly depends on scale. Using this notion, we obtain complete characterizations of uniform and non-uniform generatability. We also show that the union of two uniformly generatable classes can fail to be generatable in the limit, even on the real line. Moreover, we generalize the algorithmic framework from the countable case to the metric space setting under appropriate oracle access.
- In doubling spaces, including all finite-dimensional normed spaces, we show that generatability behaves in a well-controlled manner: uniform and non-uniform generatability is invariant to the choice of novelty scales and preserved under equivalent metrics.
- Beyond doubling spaces, we demonstrate sharp, non-intuitive behaviors. Generatability can be highly sensitive to novelty parameters and metric choice. Our constructions show that all notions of generation can fail abruptly as scales change even in the simple infinite-dimensional space ℓ^2, highlighting a qualitative gap between finite and infinite-dimensional spaces.

Extending generation to metric spaces presents several challenges. First, unlike in the countable setting, novelty can no longer be defined by distinctness alone and must reflect geometric separation. This motivates asymmetric novelty parameters for the adversary and the generator. Second, several characterization results in the countable setting rely explicitly on countability, for example through enumeration arguments, and do not extend to general metric spaces. Finally, due to the asymmetry between the adversary's and generator's novelty parameters, constructing hypothesis classes and counterexamples requires careful consideration of their relative scales.

1.2 Related Work

Our work builds on the framework of language generation in the limit, introduced by Kleinberg and Mullainathan [13]. Since then, subsequent work has investigated a range of questions surrounding this language generation model. Here, we discuss the most relevant prior works.

Uniform and Non-Uniform Generation Raman et al. [21] introduced stronger notions of generation: uniform and non-uniform generation, and provided characterization for uniform and non-uniform generation. Charikar and Pabbaraju [5] also, independently studied non-uniform generation and showed that all countable collections can be non-uniformly generated. Charikar and Pabbaraju [6] proposed an almost Pareto-optimal algorithm for non-uniform language generation. Hanneke et al. [8] studied unions of generatable classes and solved two open problems posed by Raman et al. [21].

Extensions of Language Generation Several extensions of language generation have been proposed. Raman and Raman [20], Mehrotra et al. [18] and Bai et al. [4] investigated language generation with contamination, in which the adversary is allowed to omit instances or generate instances outside the language. Peale et al. [19] studied representative generation, in which the generator needs to output a distribution that is representative of the seen data. Anastasopoulos et al. [2] investigated safe language generation, in which there exists a harmful language that the generator needs to avoid.

Other Related Work Several follow-up studies have proposed and studied different definitions of breadth in the language generation model, and showed that requiring the generator to generate diverse instances from the target language makes generation significantly harder [11, 12, 14, 15].

Alves [1] study language identification in metric spaces by revisiting classical identification in the limit and allowing distances between languages to be measured by a metric. In contrast, our work takes a fundamentally different perspective: we study *generation* rather than identification, and impose metric structure on the *instance space* rather than on the space of languages.

2 Preliminaries

2.1 Example Space and Hypothesis Class

Throughout the paper, we let $(\mathcal{X}, \rho)$ be a *separable metric space*, meaning that $\mathcal{X}$ admits a countable *dense* subset with respect to ρ. Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ be a binary hypothesis class. For a hypothesis $h \in \mathcal{H}$, let $\text{supp}(h) := \{x \in \mathcal{X} : h(x) = 1\}$, that is, its collection of positive examples. For any class $\mathcal{H}$ and a finite sequence of examples $x_1, \ldots, x_n$, define the version space as:

$$\mathcal{H}(x_1, \ldots, x_n) := \{h \in \mathcal{H} : \{x_1, \ldots, x_n\} \subseteq \text{supp}(h)\}.$$

For any class $\mathcal{H}$, define its induced closure operator $\langle \cdot \rangle_{\mathcal{H}}$ as:

$$\langle x_1, \ldots, x_n \rangle_{\mathcal{H}} := \begin{cases} \bigcap_{h \in \mathcal{H}(x_{1:n})} \text{supp}(h), & \text{if } |\mathcal{H}(x_{1:n})| \geq 1 \\ \bot, & \text{if } |\mathcal{H}(x_{1:n})| = 0 \end{cases}.$$

Intuitively, $\langle x_1, \ldots, x_n \rangle_{\mathcal{H}}$ captures the set that must belong to the target hypothesis given consistency with the observed examples. We will also need several other definitions about metric spaces.

Definition 2.1 (Closed Ball)**.** *Let $(\mathcal{X}, \rho)$ be a metric space. For a point $x \in \mathcal{X}$ or a set $A \subseteq \mathcal{X}$, we denote by $B(x, \varepsilon)$ or $B(A, \varepsilon)$ the closed ε-balls centered at x or A, respectively, defined as*

$$B(x, \varepsilon) := \{y \in \mathcal{X} : \rho(x, y) \leq \varepsilon\}, \qquad B(A, \varepsilon) := \{y \in \mathcal{X} : \inf_{x \in A} \rho(x, y) \leq \varepsilon\}.$$

Definition 2.2 (Covering Number)**.** *Let $(\mathcal{X}, \rho)$ be a metric space and $\varepsilon > 0$. An ε-covering of a set $A \subseteq \mathcal{X}$ is a finite set $\{\theta_1, \ldots, \theta_M\} \subset \mathcal{X}$ such that $A \subset B(\{\theta_i\}_{i=1}^M, \varepsilon)$. The ε-covering number $N(\varepsilon; A, \rho)$ is defined as*

$$N(\varepsilon; A, \rho) := \min \{M \in \mathbb{N} : \exists\ \varepsilon\text{-covering of } A \text{ with } M \text{ points}\}.$$

If no such finite covering exists, we write $N(\varepsilon; A, \rho) = \infty$.

To prevent the adversary from exhausting all positive examples and thus making the task of generation impossible, we make the following assumption for the hypothesis set. This assumption extends a similar condition used in previous work on generation in the countable example space setting (see Definition A.1).

Definition 2.3 (r-Uniformly Unbounded Support (r-UUS))**.** *Let $(\mathcal{X}, \rho)$ be a metric space and let $r > 0$. A hypothesis class $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ satisfies the* r-Uniformly Unbounded Support (r-UUS) *property if every $h \in \mathcal{H}$ has nonempty support and, for every sequence $(x_i)_{i \geq 1}$ in $\text{supp}(h)$ satisfying*

$$\text{supp}(h) \subseteq B(\{x_i : i \geq 1\}, r),$$

we have

$$\lim_{t \to \infty} N(r; \{x_1, \ldots, x_t\}, \rho) = \infty.$$

The r-UUS property satisfies the following monotonicity result, which we prove in Appendix D.

Lemma 2.1. *Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ be any class satisfying the r-UUS property for an $r > 0$. Then for any positive $\delta < r$, $\mathcal{H}$ satisfies the δ-UUS property.*

2.2 Generatability

Consider the following two-player game for a hypothesis set $\mathcal{H}$. At the start, the adversary picks a hypothesis $h \in \mathcal{H}$ and does not reveal it to the learner. The game then proceeds over rounds $t = 1, 2, \ldots$. In each round $t \in \mathbb{N}$, the adversary reveals a *new* example $x_t \in \text{supp}(h)$. The generator, after observing enough examples of $x_1, \ldots, x_t$, must output some *new* example inside $\text{supp}(h)$. Crucially, the generator *never* observes its loss as it does not know h. The goal of the generator is to eventually perfectly generate new, positive examples $x \in \text{supp}(h)$. To make this formal, we first define a generator as a mapping from a finite sequence of examples to a new example.

Definition 2.4 (Generator)**.** *A generator is a map $\mathcal{G} : \mathcal{X}^\star \to \mathcal{X}$ that takes a finite sequence of examples $x_1, x_2, \ldots$ and outputs a new example x.*

The challenge of extending the framework to metric spaces lies in the definition of *novelty*. In a countable example space, novelty can simply be defined as *distinct instances*. In a metric space, however, distinctness is not enough: one does not expect an instance that is very close to the previous ones to be a *new* instance.

As a result, we define a *new* instance to be ε far away from the previous instances. To make the definition general, the ε is allowed to be different on the adversary and the generator sides. That is, we define two novelty parameters for the adversary and the generator, denoted as ε_a and ε_g, respectively.

Definition 2.5 (Generatability in the Limit)**.** *Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ be any hypothesis class satisfying the r-*UUS *property for some $r > 0$ and let $0 < \varepsilon_a, \varepsilon_g \leq r$. If there exists a generator $\mathcal{G}$ such that for every $h \in \mathcal{H}$, and any sequence $\{x_i\}_{i=1}^\infty \subseteq \text{supp}(h)$ such that $\text{supp}(h) \subseteq B(\{x_i\}_{i=1}^\infty, \varepsilon_a)$ (i.e., the adversary eventually reveals an ε_a-cover of the support), there exists a $t^\star \in \mathbb{N}$ such that $\mathcal{G}(x_{1:s}) \in \text{supp}(h) \setminus B(\{x_i\}_{i=1}^s, \varepsilon_g)$ for all $s \geq t^\star$, then we say $\mathcal{H}$ is* $(\varepsilon_a, \varepsilon_g)$-generatable in the limit*.*

Remark 2.1. *The assumption $\text{supp}(h) \subseteq B(\{x_i\}_{i=1}^\infty, \varepsilon_a)$ is an analogous to the enumeration assumption in the countable example space setting (see Definition A.2). It requires the adversary to eventually show the whole structure of the hypothesis set to the generator.*

Our definition above is intentionally general and does not specify the relative magnitudes of ε_a and ε_g. This is because the relative magnitudes of novelty on the adversary and generator sides are application dependent. For example, consider authors preparing a paper submission to a conference. An experienced author may read several papers that are closely related in topics (corresponding to a small ε_a for adversary) yet generate a paper that is substantially different, for example, by introducing a new technique or applying known methods to a new problem. In this case the generator's ε_g is larger than than the adversary's ε_a. Alternatively, a beginning researcher may read a very diverse set of papers spanning many topics (a large ε_a), yet generate a paper closely aligned with one of them, such as by extending an existing line of work. Here, the generator's ε_g is smaller than the adversary's ε_a. Our definition of generatability in the limit is designed to encompass both behaviors.

Next, we define uniform generation, which requires a common covering-number threshold independent of both the target hypothesis and the positive sequence. The round at which this threshold is reached may depend on the sequence.

Definition 2.6 (Uniform Generatability)**.** *Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ be any hypothesis class satisfying the r-*UUS *property for some $r > 0$ and let $0 < \varepsilon_a, \varepsilon_g \leq r$. If there exists a generator $\mathcal{G}$ and $d^\star \in \mathbb{N}$, such that for every $h \in \mathcal{H}$ and any sequence $\{x_i\}_{i=1}^\infty \subseteq \text{supp}(h)$, if there exists t such that $N(\varepsilon_a; \{x_i\}_{i=1}^t, \rho) \geq d^\star$, then $\mathcal{G}(x_{1:s}) \in \text{supp}(h) \setminus B(\{x_i\}_{i=1}^s, \varepsilon_g)$ for all $s \geq t$, then we say $\mathcal{H}$ is* $(\varepsilon_a, \varepsilon_g)$-uniformly generatable*.*

For non-uniform generation, the covering-number threshold may depend on the target hypothesis, but not on the positive sequence.

Definition 2.7 (Non-Uniform Generatability)**.** *Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ be any hypothesis class satisfying the r-*UUS *property for some $r > 0$ and let $0 < \varepsilon_a, \varepsilon_g \leq r$. If there exists a generator $\mathcal{G}$ such that for every $h \in \mathcal{H}$ there exists a $d_h \in \mathbb{N}$, such that for any sequence $\{x_i\}_{i=1}^\infty \subseteq \text{supp}(h)$, if there exists t such that $N(\varepsilon_a; \{x_i\}_{i=1}^t, \rho) \geq d_h$, then $\mathcal{G}(x_{1:s}) \in \text{supp}(h) \setminus B(\{x_i\}_{i=1}^s, \varepsilon_g)$ for all $s \geq t$, then we say $\mathcal{H}$ is* $(\varepsilon_a, \varepsilon_g)$-non-uniformly generatable*.*

It is clear that our framework strictly generalizes the countable example space setting considered in Raman et al. [21]. This is achieved by equipping a countable example space with the discrete $\{0,1\}$-metric. We include a formal proposition in Appendix D for completeness sake.

3 Characterizations of Generatability

In this section, we characterize hypothesis classes that are uniformly and non-uniformly generatable, and we show that the union of two uniformly generatable classes can fail to be generatable in the limit. We show that, despite the additional metric structure, the resulting characterizations closely parallel those obtained in the countable example space setting (see Appendix A for more details). Consequently, we state the main results in this section and defer the proofs to Appendix B. In Appendix E, we also have a discussion about the computability of generation. Although these characterizations are straightforward, generation in metric spaces exhibits several counterintuitive phenomena that do not arise in the countable example space setting. We explore these phenomena in more detail in Section 4.

We first extend the techniques from Raman et al. [21] to account for metric space structure. As a result, we first define a metric analog of the Closure dimension originally defined for countable example space (see Definition A.5).

Definition 3.1 ($(\varepsilon_a, \varepsilon_g)$-closure dimension)**.** *The* $(\varepsilon_a, \varepsilon_g)$-closure dimension *of* $\mathcal{H}$*, denoted* $\mathrm{C}_{\varepsilon_a,\varepsilon_g}(\mathcal{H})$*, is the largest* $d \in \mathbb{N}$ *such that there exist* $h \in \mathcal{H}$ *and* $\{x_i\}_{i=1}^t \subseteq \mathrm{supp}(h)$ *satisfying*

$$N(\varepsilon_a; \{x_i\}_{i=1}^t, \rho) = d,$$

and a finite set $F \subseteq \langle x_1, \ldots, x_t \rangle_{\mathcal{H}}$ *such that*

$$\langle x_1, \ldots, x_t \rangle_{\mathcal{H}} \subseteq B(F, \varepsilon_g).$$

If this holds for arbitrarily large $d \in \mathbb{N}$*, we set* $\mathrm{C}_{\varepsilon_a,\varepsilon_g}(\mathcal{H}) = \infty$*. If no such sequence exists for any* $d \in \mathbb{N}$*, we set* $\mathrm{C}_{\varepsilon_a,\varepsilon_g}(\mathcal{H}) = 0$*.*

Unlike the Closure dimension, the $(\varepsilon_a, \varepsilon_g)$-Closure dimension is *scale-sensitive*. In Section 4 we will further explore its property for different scales of ε_a and ε_g. Our first theorem uses this dimension to provide a complete characterization of which classes are uniform generatable.

Theorem 3.1 (Characterization of Uniform Generatability)**.** *Let* $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ *be any hypothesis class satisfying the* r-UUS *property for some* $r > 0$ *and let* $0 < \varepsilon_a, \varepsilon_g \leq r$*. Then,* $\mathcal{H}$ *is* $(\varepsilon_a, \varepsilon_g)$*-uniformly generatable if and only if* $\mathrm{C}_{\varepsilon_a,\varepsilon_g}(\mathcal{H}) < \infty$*.*

Similarly to the countable example space setting, for every finite hypothesis class $\mathcal{H}$ satisfying the r-UUS property with $\varepsilon_a \geq \varepsilon_g$, we have $\mathrm{C}_{\varepsilon_a,\varepsilon_g}(\mathcal{H}) < \infty$. In contrast, when $\varepsilon_a < \varepsilon_g$, there exists a finite hypothesis class that is not non-uniformly generatable. Intuitively, if the generator is required to produce outputs that are more novel than those revealed by the adversary, the generation task may become impossible. Further discussion is provided in Appendix B.

Corollary 3.2 (Finite Classes are Uniformly Generatable)**.** *Let* $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ *be any finite hypothesis class satisfying the* r-UUS *property for some* $r > 0$ *and let* $0 < \varepsilon_a, \varepsilon_g \leq r$ *such that* $\varepsilon_g \leq \varepsilon_a$*. Then,* $\mathcal{H}$ *is* $(\varepsilon_a, \varepsilon_g)$*-uniformly generatable.*

We next move to characterize non-uniform generatability. We show that our characterization of uniform generatability leads to a characterization of non-uniform generatability.

Theorem 3.3 (Characterization of Non-uniform Generatability)**.** *Let* $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ *be any hypothesis class with the* r-UUS *property for some* $r > 0$ *and let* $0 < \varepsilon_a, \varepsilon_g \leq r$*. The following statements are equivalent.*

(i) $\mathcal{H}$ *is* $(\varepsilon_a, \varepsilon_g)$*-non-uniformly generatable.*
(ii) There exists a non-decreasing *sequence of classes* $\mathcal{H}_1 \subseteq \mathcal{H}_2 \subseteq \ldots$ *such that* $\mathcal{H} = \bigcup_{i \in \mathbb{N}} \mathcal{H}_i$ *and* $\mathrm{C}_{\varepsilon_a,\varepsilon_g}(\mathcal{H}_n) < \infty$ *for every* $n \in \mathbb{N}$*.*

An immediate consequence of combing Corollary 3.2 and Theorem 3.3 is the following corollary.

Corollary 3.4 (Countable Classes are Non-uniformly Generatable)**.** *Let* $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ *be any countable hypothesis class satisfying the* r-UUS *property for some* $r > 0$ *and let* $0 < \varepsilon_a, \varepsilon_g \leq r$ *such that* $\varepsilon_g \leq \varepsilon_a$*. Then,* $\mathcal{H}$ *is* $(\varepsilon_a, \varepsilon_g)$*-non-uniformly generatable.*

We next show that uniform generatability of the component classes does not ensure generation in the limit for their union. This obstruction occurs even on the real line.

Theorem 3.5 (Failure of Finite-Union Closure)**.** *Let $r > 0$ and $0 < \varepsilon_a, \varepsilon_g \leq r$. On $\mathbb{R}$ with its usual metric, there exist hypothesis classes $\mathcal{H}_1, \mathcal{H}_2 \subseteq \{0,1\}^{\mathbb{R}}$, each satisfying r-UUS and each $(\varepsilon_a, \varepsilon_g)$-uniformly generatable, such that $\mathcal{H}_1 \cup \mathcal{H}_2$ is not $(\varepsilon_a, \varepsilon_g)$-generatable in the limit.*

In contrast, Raman et al. [21] showed that, in the countable example space setting, every finite union of uniformly generatable classes is generatable in the limit. Theorem 3.5 shows that this finite-union property fails in the metric setting, even for two classes on the real line.

4 Properties of Generatability

Beyond the structure of the hypothesis class itself, generatability depends on two additional parameters: the novelty thresholds with respect to the adversary and the generator, denoted by ε_a and ε_g, respectively. In this section, we study how generatability varies as these novelty parameters change. We also investigate generatability under different metrics defined on the same underlying space.

4.1 Generatability is Well Behaved in Doubling Spaces

We first establish a robustness result for generatability in doubling spaces. A doubling space is a metric space in which balls at one scale can be covered by finitely many balls at any smaller scale. In such spaces, we show that, both uniform and non-uniform generatability are invariant across all choices of scale parameters ε_a and ε_g, and are preserved under any change to an equivalent metric. Throughout, we say that two metrics are equivalent if they are bi-Lipschitz equivalent; see Appendix D.3 for details.

Definition 4.1. *A metric space $(\mathcal{X}, \rho)$ is said to be* doubling *if there exists a constant $M > 0$ such that for any $x \in \mathcal{X}$ and any $r > 0$, the ball $B(x, r)$ can be covered by at most M balls of radius $r/2$.*

We first show a stronger result than Lemma 2.1: in doubling spaces, the r-UUS property is invariant to the choice of the scale parameter r. The proof is deferred to Appendix D for completeness.

Theorem 4.1. *Let $(\mathcal{X}, \rho)$ be a doubling metric space. Suppose $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ is a class satisfying r-UUS property for some $r > 0$. Then it also satisfies the r'-UUS property for any $r' > 0$. In this case, we simply say that $\mathcal{H}$ has the UUS property.*

We are now ready to present robustness results for uniform and non-uniform generatability in doubling spaces.

Theorem 4.2. *Let $(\mathcal{X}, \rho)$ be a doubling metric space. Let $\varepsilon_a, \varepsilon_g, \delta_a, \delta_g > 0$ satisfy $0 < \delta_a < \varepsilon_a$ and $0 < \delta_g < \varepsilon_g$. Suppose $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ satisfies the UUS property. Then $\mathcal{H}$ is $(\varepsilon_a, \varepsilon_g)$-uniformly (resp. non-uniformly) generatable if and only if it is (δ_a, δ_g)-uniformly (resp. non-uniformly) generatable.*

Proof. By Theorem 3.1 and 3.3, we only need to show that $\mathrm{C}_{\varepsilon_a, \varepsilon_g}(\mathcal{H}) < \infty$ if and only if $\mathrm{C}_{\delta_a, \delta_g}(\mathcal{H}) < \infty$. Since $\mathcal{X}$ is a doubling space, there exists C such that for any $x \in \mathcal{X}$, $B(x, \varepsilon_a)$ can be covered by the union of at most C balls of radius δ_a and $B(x, \varepsilon_g)$ can be covered by the union of at most C balls of radius δ_g. Then under our assumption, for any subset $A \subseteq \mathcal{X}$,

$$N(\varepsilon_a; A, \rho) = \infty \iff N(\delta_a; A, \rho) = \infty.$$

And if $N(\varepsilon_a; A, \rho) < \infty$, we have

$$N(\varepsilon_a; A, \rho) \leq N(\delta_a; A, \rho) \leq C N(\varepsilon_a; A, \rho).$$

We also need to compare covers with centers in the closure. If a nonempty set K has a finite ε_g-cover with centers in K, use doubling to cover each of these balls by finitely many balls of radius $\delta_g/2$. From each smaller ball meeting K, choose one point of K as a new center. The triangle inequality shows that the chosen points form a finite δ_g-cover of K. Conversely, every δ_g-cover is a ε_g-cover. Thus the same generatable samples have closures admitting the required finite internal covers at the two generator scales. The displayed covering-number inequalities then give

$$\mathrm{C}_{\varepsilon_a, \varepsilon_g}(\mathcal{H}) \leq \mathrm{C}_{\delta_a, \delta_g}(\mathcal{H}) \leq C\, \mathrm{C}_{\varepsilon_a, \varepsilon_g}(\mathcal{H}).$$

This proves the uniform case. Applying the same comparison to every subclass in Theorem 3.3 proves the non-uniform case. □

It is well known that Euclidean space $\mathbb{R}^d$ is a doubling metric space. As a result, we have the following corollary:

Corollary 4.3. *In any finite-dimensional normed vector space $\mathcal{X}$, if $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ satisfies the UUS property, then for any $\varepsilon_a, \varepsilon_g, \delta_a, \delta_g > 0$, $\mathcal{H}$ is $(\varepsilon_a, \varepsilon_g)$-uniformly (resp. non-uniformly) generatable if and only if it is (δ_a, δ_g)-uniformly (resp. non-uniformly) generatable.*

For arbitrary parameter pairs, choose a positive pair strictly smaller in each coordinate than both pairs and apply Theorem 4.2 twice; finite-dimensional normed spaces are doubling. We next study how generatability behaves under changes of the underlying metric. The proof is similar to that of Theorem 4.2 so we defer it to Appendix D.

Theorem 4.4. *Let $(\mathcal{X}, \rho_1)$ be a doubling metric space, and let ρ_2 be a metric on $\mathcal{X}$ bi-Lipschitz equivalent to ρ_1. Suppose $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ is a hypothesis class. Then, it satisfies the UUS property with respect to ρ_1 if and only if it satisfies the UUS property with respect to ρ_2. If these equivalent UUS conditions hold, then $\mathcal{H}$ is uniformly (resp. non-uniformly) generatable with respect to ρ_1 if and only if it is uniformly (resp. non-uniformly) generatable with respect to ρ_2.*

So far, we have established the desired stability properties of uniform and non-uniform generation in doubling metric spaces. What about generation in the limit? Unfortunately, pathological phenomena can occur: generatability can change abruptly when changing the scale parameters ε_a or ε_g. We present such an example in Example 4.5 here and leave the construction and proof to Appendix C.

Example 4.5. *On the real line $\mathbb{R}$, there is a hypothesis class $\mathcal{H}$ that satisfies the UUS property, and is $(\varepsilon, 1)$-generatable in the limit for any $\varepsilon \in (0,1)$, but not $(1,1)$-generatable in the limit.*

These examples illustrate phenomena that are specific to generation in metric spaces and are examined in more detail in the next section.

4.2 Generatability is Scale-Sensitive

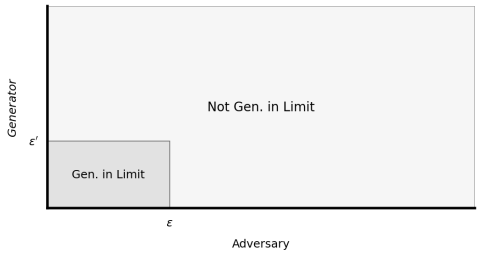
Outside doubling spaces, these robustness properties can fail: generatability may depend sharply on ε_a, ε_g, and the underlying metric. The adversary reveals a sequence of positive examples from a hypothesis h, while the generator must eventually produce points in $\mathrm{supp}(h)$ at distance strictly greater than ε_g from every previously observed point. For uniform and non-uniform generation, the guarantee is triggered when the observed sample's ε_a-covering number reaches a prescribed threshold. Increasing ε_a can only decrease this covering number, making the guarantee easier to satisfy. Decreasing ε_g also makes generation easier, since the generator is required to produce less novel instances.

In this section, we establish monotonicity results with respect to both ε_a and ε_g. For *uniform* and *non-uniform* generation, the results align with this intuition: if a hypothesis class $\mathcal{H}$ is $(\varepsilon_a, \varepsilon_g)$-uniformly (resp. non-uniformly) generatable, then it remains so when ε_a is increased or ε_g is decreased. We use Example 4.9 to illustrate this behavior in the ℓ^2 space , where increasing ε_a or decreasing ε_g makes generation easier. Notably, this example also shows a separation between finite-dimensional normed spaces, and the infinite-dimensional space ℓ^2. Depending on the choice of scale parameters ε_a and ε_g, the same hypothesis class can exhibit all possible regimes of generatability, ranging from not generatable in the limit to uniformly generatable. This phenomenon cannot occur in a doubling space.

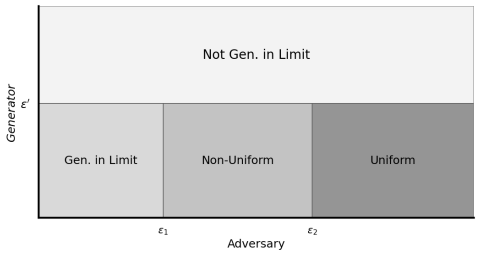
The situation for *generation in the limit* is more subtle, as already illustrated by Example 4.5. When considering generation in the limit, monotonicity holds only when ε_a is decreased. In contrast, increasing ε_a can cause generatability to fail: a hypothesis class may be $(\varepsilon_a, \varepsilon_g)$-generatable in the limit for small ε_a, yet lose this property once ε_a exceeds a threshold. Example 4.8 demonstrates that, in ℓ^2, generatability in the limit can disappear when ε_a and ε_g exceed such thresholds. Moreover, in contrast to Example 4.5 in the doubling-space setting, Example 4.8 is more general: the threshold values of ε_a and ε_g at which generatability fails can be chosen arbitrarily.

Theorem 4.6. *Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ be any hypothesis class satisfying the r-UUS property for some $r > 0$ and let $0 < \varepsilon_a, \varepsilon_g \leq r$. If $\mathcal{H}$ is $(\varepsilon_a, \varepsilon_g)$-generatable in the limit, then for any positive $\delta_a \leq \varepsilon_a$, $\delta_g \leq \varepsilon_g$, $\mathcal{H}$ is (δ_a, δ_g)-generatable in the limit.*

Theorem 4.7. *Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ be any hypothesis class satisfying the r-UUS property for some $r > 0$ and let $0 < \varepsilon_a, \varepsilon_g \leq r$. If $\mathcal{H}$ is $(\varepsilon_a, \varepsilon_g)$-uniformly (or non-uniformly) generatable, then for any $\varepsilon_a \leq \delta_a \leq r$ and $0 < \delta_g \leq \varepsilon_g$, $\mathcal{H}$ is (δ_a, δ_g)-uniformly (or non-uniformly) generatable.*



(a) Illustration of Example 4.8, showing generatability in the limit can disappear when ε_a and ε_g are larger than a threshold.



(b) Illustration of Example 4.9, showing increasing ε_a or decreasing ε_g makes generation easier.

Figure 1: Illustration of two generation regimes in a metric space. The horizontal axis represents the adversary parameter, and the vertical axis represents the generator parameter.

Theorems 4.6 and 4.7 establish monotonicity with respect to both ε_a and ε_g for all three notions of generation, which we prove in Appendix D.

Example 4.8. *In ℓ^2-space, for any $r > 0$ and any $0 < \varepsilon_a, \varepsilon_g \leq r$, there exists a hypothesis class $\mathcal{H} \subseteq \{0,1\}^{\ell^2}$ satisfying the r-*UUS *property such that, for every $0 < \gamma_a, \gamma_g \leq r$, it is (γ_a, γ_g)-generatable in the limit if and only if $\gamma_a < \varepsilon_a$ and $\gamma_g < \varepsilon_g$.*

Figure 1a gives an illustration of this example. We now present the construction and the underlying intuition here, and defer the full proof to Appendix C.

Recall that ℓ^2 is the space that contains all sequences $(x_i)_{i=1}^{\infty}$ such that $\sum_i x_i^2 < \infty$. Let $e_k \in \ell^2$ denote the k-th standard basis vector. Define

$$u_k := 4r \cdot e_k, \qquad a_k := \varepsilon_a \cdot e_k, \qquad g_k := \varepsilon_g \cdot e_k.$$

Let $\mathcal{I}$ be the set of all infinite subsets of $\mathbb{N}_+$, i.e., $\mathcal{I} = \{I | I \in 2^{\mathbb{N}_+}, |I| = \infty\}$. For $I \in \mathcal{I}$, let

$$U_I = \{u_k | k \in I\}, \quad O_I = \{a_{2k} | k \in I\} \cup \{g_{2k^n+1} | k \in I, n \in \mathbb{N}\}, \quad D_I = \{g_{2k+1} | k \in I\}$$

and construct the hypothesis class $\mathcal{H}$ such that

$$\mathcal{H} = \{h_{I_1,I_2,I_3} | I_1, I_2, I_3 \in \mathcal{I}, \quad \operatorname{supp}(h_{I_1,I_2,I_3}) = \{\mathbf{0}\} \cup U_{I_1} \cup O_{I_2} \cup D_{I_3}\}$$

Here, u_k and U_I are constructed to ensure that $\mathcal{H}$ satisfies the r-UUS property. As will be seen in the proof, the points a_k are constructed for the adversary, while the points g_k are constructed for the generator. In fact, the g_k are the only instances with an explicit structure, which enables the generator to repeatedly produce them. The set O_I is the *optimal set*: when $\gamma_g < \varepsilon_g$, revealing $a_{2k} \in O_I$ with $k \geq 2$ identifies an infinite tower $\{g_{2k^n+1} : n \geq 1\}$ from which novel outputs can be chosen. In contrast, D_I is the *disturbing set*: if the adversary continues to reveal $g_k \in D_I$, then the generator fails to generate in the limit.

Intuitively, when $\gamma_a < \varepsilon_a$ and $\gamma_g < \varepsilon_g$, since the adversary must choose a sequence $\{x_i\}_{i=1}^{\infty}$ such that $\operatorname{supp}(h) \subseteq B(\{x_i\}_{i=1}^{\infty}, \gamma_a)$, there must exist some time t at which $x_t = a_{2k}$ for some $k \geq 2$. From that point on, the generator can choose a point g_{2k^n+1} whose coordinate has not appeared in the finite observed sample; this point belongs to the target support and is more than γ_g from every observation. When $\gamma_g \geq \varepsilon_g$, however, as long as the adversary reveals $\mathbf{0}$, the generator cannot generate any g_k, since each g_k lies within the γ_g-ball centered at $\mathbf{0}$. Finally, when $\gamma_a \geq \varepsilon_a$, the adversary may choose to reveal only instances of the form $\mathbf{0}$, g_k, and u_k. The generator's task is to uncover the structure of O_I and generate from it. However, due to the interference caused by D_I, the generator cannot distinguish whether the observed instances originate from O_I or from D_I, rendering generation in the limit impossible.

Example 4.9. *In ℓ^2-space, for any $r > 0$ and any $0 < \varepsilon_{a1} \leq \varepsilon_{a2} \leq r$ and $0 < \varepsilon_g \leq r$, there exists a hypothesis class $\mathcal{H} \subseteq \{0,1\}^{\ell^2}$ satisfying the r-*UUS *property with the following generation behavior. For $\gamma_g \in (0, \varepsilon_g)$,*

$\gamma_a \in [\varepsilon_{a2}, r]$, *then (γ_a, γ_g)-uniformly generatable*,
$\gamma_a \in [\varepsilon_{a1}, \varepsilon_{a2})$, *then (γ_a, γ_g)-non-uniformly but not uniformly generatable*,
$\gamma_a \in (0, \varepsilon_{a1})$, *then (γ_a, γ_g)-generatable in the limit but not non-uniformly generatable*.

Moreover, for $\gamma_g \in [\varepsilon_g, r]$ and any $\gamma_a \in (0, r]$, $\mathcal{H}$ is not (γ_a, γ_g)-generatable in the limit.

Figure 1b illustrates this example. We give the construction and intuition here and prove the claims in Appendix C. Let $K = \{2, 3, \dots\}$. For $k \in K$, define points on four disjoint coordinate sets:

$$u_k := 2r\, e_{4k}, \qquad a_{k,1} := \varepsilon_{a1} e_{4k+1},$$
$$a_{k,2} := \varepsilon_{a2} e_{4k+2}, \qquad g_k := \varepsilon_g e_{4k+3}.$$

Let $\mathcal{I}$ and $\mathcal{I}_{\text{fin}}$ be the infinite and finite subsets of K, respectively. Set

$$O_I = \{u_k : k \in I\} \cup \{g_{k^n} : k \in I,\ n \geq 1\},$$
$$D_{I,J} = \{a_{k,1} : k \in I\} \cup \{a_{k,2} : k \in J\},$$

and define $\mathcal{H} = \{h_{I_1,I_2,J} : I_1, I_2 \in \mathcal{I},\ J \in \mathcal{I}_{\text{fin}}\}$ by

$$\text{supp}(h_{I_1,I_2,J}) = \{0\} \cup O_{I_1} \cup D_{I_2,J}.$$

The u-points ensure r-UUS. Observing either u_k or g_k identifies an infinite tower $\{g_{k^n} : n \geq 1\}$ in the target support. When $\gamma_g < \varepsilon_g$, a tower point on an unobserved coordinate is always novel.

If $\gamma_a \geq \varepsilon_{a2}$, the origin and all disturbing points in $D_{I_2,J}$ have covering number at most 1, so a sample covering number of at least 2 forces a tower observation. If $\varepsilon_{a1} \leq \gamma_a < \varepsilon_{a2}$, these points have covering number at most $1 + |J|$, giving a target-dependent threshold; arbitrarily large finite sets of second-family points rule out a uniform threshold. Below ε_{a1}, arbitrary infinite subsets of first-family points obstruct non-uniform generation. Nevertheless, every valid covering sequence must reveal a u-point, so generation in the limit is possible whenever $\gamma_g < \varepsilon_g$. Finally, when $\gamma_g \geq \varepsilon_g$, observing 0 blocks all g-points; a diagonal choice of the remaining support defeats generation in the limit.

4.3 Generatability is Metric-Dependent

In Section 4.1, we showed that uniform generatability is invariant under changes to equivalent metrics. In contrast, in general metric spaces, generatability can depend on the underlying metric: even equivalent metrics on the same instance space may induce different generatability properties.

We begin by establishing a stability result. Theorem 4.10 shows that if one metric is bounded by another, then both the UUS property and generatability in the limit transfer in a controlled manner, with the relevant parameters scaling linearly with the bounding constant. In this sense, generatability remains stable under sufficiently mild distortions of the metric. We present the proof in Appendix D.

Theorem 4.10. *Let $\mathcal{X}$ be a metric space with metric ρ_1 and ρ_2. Assume there exists $M > 0$ such that for any $x, y \in \mathcal{X}$, $\rho_2(x, y) \leq M\rho_1(x, y)$. Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ be any class satisfying the r-UUS property for an $r > 0$ with respect to ρ_2, and $0 < \varepsilon_a, \varepsilon_g \leq r$. Then $\mathcal{H}$ satisfies the r/M-UUS property with respect to ρ_1. Moreover, if $\mathcal{H}$ is $(\varepsilon_a, \varepsilon_g)$-generatable in the limit with respect to ρ_2, then $\mathcal{H}$ is also $(\varepsilon_a/M, \varepsilon_g/M)$-generatable in the limit with respect to ρ_1.*

Our next example highlights a sharp contrast. Even when two metrics are uniformly equivalent, generatability need not be preserved. In particular, we construct an example in ℓ^2 showing that a hypothesis class can be uniformly generatable under the standard metric, yet fail to be non-uniformly generatable under an equivalent metric. We state the example and leave the full proof in Appendix C.

Example 4.11. *In ℓ^2-space, for any $r > 0$, there exists a hypothesis class $\mathcal{H}$ that satisfies the r-UUS property and a metric ρ' that is equivalent with the ordinary metric ρ such that $\mathcal{H}$ is $(\varepsilon, \varepsilon)$-uniform generatable with respect to ρ for every $\varepsilon \in (0, r)$, but $\mathcal{H}$ is not $(\varepsilon, \varepsilon)$-non-uniformly generatable with respect to ρ' for any $\varepsilon \in (0, r)$.*

5 Discussion

We initiated a theoretical study of generation in metric spaces, extending prior frameworks that were restricted to countable domains. Our results show that introducing geometric structure leads to new scale-dependent and metric-dependent phenomena, while preserving several older characterizations in modified form. In particular, we identified doubling metric spaces as a broad regime in which generatability behaves in a stable and robust manner, and demonstrated that outside this regime, generatability can be highly sensitive to both the novelty parameters and the choice of metric.

Several directions remain open, including studying generation when outputs are considered equivalent under natural symmetries, identifying geometric conditions beyond doubling that ensure stable behavior, and strengthening our understanding of generation under additional structural assumptions.

References

[1] Fernando C. Alves. Language learnability in the limit for general metrics: a gold-angluin result, 2021. URL https://arxiv.org/abs/2103.13166.

[2] Antonios Anastasopoulos, Giuseppe Ateniese, and Evgenios M Kornaropoulos. Safe language generation in the limit. *arXiv preprint arXiv:2601.08648*, 2026.

[3] Dana Angluin. Finding patterns common to a set of strings. In *Proceedings of the eleventh annual ACM Symposium on Theory of Computing*, pages 130–141, 1979.

[4] Yannan Bai, Debmalya Panigrahi, and Ian Zhang. Language generation in the limit: Noise, loss, and feedback. In *Proceedings of the 2026 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 794–816. SIAM, 2026.

[5] Moses Charikar and Chirag Pabbaraju. Exploring facets of language generation in the limit, 2024. URL https://arxiv.org/abs/2411.15364.

[6] Moses Charikar and Chirag Pabbaraju. Pareto-optimal non-uniform language generation. *arXiv preprint arXiv:2510.02795*, 2025.

[7] E Mark Gold. Language identification in the limit. *Information and control*, 10(5):447–474, 1967.

[8] Steve Hanneke, Amin Karbasi, Anay Mehrotra, and Grigoris Velegkas. On union-closedness of language generation. *arXiv preprint arXiv:2506.18642*, 2025.

[9] Emiel Hoogeboom, Víctor Garcia Satorras, Clément Vignac, and Max Welling. Equivariant diffusion for molecule generation in 3D. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvari, Gang Niu, and Sivan Sabato, editors, *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 8867–8887. PMLR, 17–23 Jul 2022. URL https://proceedings.mlr.press/v162/hoogeboom22a.html.

[10] Marco Jiralerspong, Joey Bose, Ian Gemp, Chongli Qin, Yoram Bachrach, and Gauthier Gidel. Feature likelihood score: Evaluating the generalization of generative models using samples. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023. URL https://openreview.net/forum?id=l2VKZkolT7.

[11] Alkis Kalavasis, Anay Mehrotra, and Grigoris Velegkas. Characterizations of language generation with breadth. *arXiv preprint arXiv:2412.18530*, page 3, 2024.

[12] Alkis Kalavasis, Anay Mehrotra, and Grigoris Velegkas. On the limits of language generation: Trade-offs between hallucination and mode collapse. *arXiv preprint arXiv:2411.09642*, 2024.

[13] Jon Kleinberg and Sendhil Mullainathan. Language generation in the limit. *Advances in Neural Information Processing Systems*, 37:66058–66079, 2024.

[14] Jon Kleinberg and Fan Wei. Density measures for language generation. *arXiv preprint arXiv:2504.14370*, 2025.

[15] Jon Kleinberg and Fan Wei. Language generation and identification from partial enumeration: Tight density bounds and topological characterizations. *arXiv preprint arXiv:2511.05295*, 2025.

[16] Nikola Kovachki, Zongyi Li, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: learning maps between function spaces with applications to pdes. *J. Mach. Learn. Res.*, 24(1), January 2023. ISSN 1532-4435.

[17] Yeqing Lin and Mohammed AlQuraishi. Generating novel, designable, and diverse protein structures by equivariantly diffusing oriented residue clouds. In *Proceedings of the 40th International Conference on Machine Learning*, ICML'23. JMLR.org, 2023.

[18] Anay Mehrotra, Grigoris Velegkas, Xifan Yu, and Felix Zhou. Language generation with infinite contamination. *arXiv preprint arXiv:2511.07417*, 2025.

[19] Charlotte Peale, Vinod Raman, and Omer Reingold. Representative language generation. *arXiv preprint arXiv:2505.21819*, 2025.

[20] Ananth Raman and Vinod Raman. Generation from noisy examples. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=ouaiJlE1hJ.

[21] Vinod Raman, Jiaxun Li, and Ambuj Tewari. Generation through the lens of learning theory. In Nika Haghtalab and Ankur Moitra, editors, *Proceedings of Thirty Eighth Conference on Learning Theory*, volume 291 of *Proceedings of Machine Learning Research*, pages 4740–4776. PMLR, 30 Jun–04 Jul 2025. URL https://proceedings.mlr.press/v291/raman25a.html.

A Summary of Results for Countable Example Space Generation

We recall the formal definitions of generation in countable example spaces and the main structural results established in prior work, which motivate and inform our development of generation in metric spaces.

Through this section, $\mathcal{X}$ is a countable example space and $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ is a binary hypothesis class.

A.1 Definitions of Generation

Definition A.1 (Uniformly Unbounded Support (UUS)[21])**.** *A hypothesis class $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ satisfies the* Uniformly Unbounded Support (UUS) property *if $|\operatorname{supp}(h)| = \infty$ for every $h \in \mathcal{H}$.*

Let $\mathcal{X}$ be an countable example space and let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ be a hypothesis class satisfying the UUS property.

Definition A.2 (Generatability in the Limit [21])**.** *The hypothesis class $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ is* generatable in the limit *if there exists a generator $\mathcal{G}$ such that for every $h \in \mathcal{H}$ and every enumeration $\{x_1, x_2, \dots\}$ of $\operatorname{supp}(h)$, there exists $t^\star \in \mathbb{N}$ such that $\mathcal{G}(x_{1:s}) \in \operatorname{supp}(h) \setminus \{x_1, \dots, x_s\}$ for all $s \geq t^\star$.*

Definition A.3 (Uniform Generatability [21])**.** *The hypothesis class $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ is* uniformly generatable *if there exists a generator $\mathcal{G}$ and an integer $d^\star \in \mathbb{N}$ such that for every $h \in \mathcal{H}$ and every sequence $\{x_1, x_2, \dots\} \subseteq \operatorname{supp}(h)$, if there exists $t \in \mathbb{N}$ such that $|\{x_1, \dots, x_t\}| = d^\star$ then $\mathcal{G}(x_{1:s}) \in \operatorname{supp}(h) \setminus \{x_1, \dots, x_s\}$ for all $s \geq t$.*

Definition A.4 (Non-uniform Generatability [21])**.** *The hypothesis class $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ is* non-uniformly generatable *if there exists a generator $\mathcal{G}$ such that for every $h \in \mathcal{H}$, there exists an integer $d_h^\star \in \mathbb{N}$ such that for every sequence $\{x_1, x_2, \dots\} \subseteq \operatorname{supp}(h)$, if there exists $t \in \mathbb{N}$ such that $|\{x_1, \dots, x_t\}| = d_h^\star$, then $\mathcal{G}(x_{1:s}) \in \operatorname{supp}(h) \setminus \{x_1, \dots, x_s\}$ for all $s \geq t$.*

A.2 Closure Dimension and Characterizations

Definition A.5 (Closure Dimension [21])**.** *The* closure dimension *of a hypothesis class $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$, denoted $\mathrm{C}(\mathcal{H})$, is the largest integer $d \in \mathbb{N}$ such that there exist distinct points $x_1, \dots, x_d \in \mathcal{X}$ with $\langle x_1, \dots, x_d \rangle_{\mathcal{H}} \neq \perp$ and $|\langle x_1, \dots, x_d \rangle_{\mathcal{H}}| < \infty$. If such d exists for arbitrarily large values, we write $\mathrm{C}(\mathcal{H}) = \infty$.*

Theorem A.1 (Uniform Generatability Characterization [21])**.** *A hypothesis class $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ is uniformly generatable if and only if $\mathrm{C}(\mathcal{H}) < \infty$.*

Theorem A.2 (Non-uniform Generatability Characterization [21])**.** *A hypothesis class $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ is non-uniformly generatable if and only if there exists a non-decreasing sequence of subclasses $\mathcal{H}_1 \subseteq \mathcal{H}_2 \subseteq \cdots$ such that $\mathcal{H} = \bigcup_{i=1}^{\infty} \mathcal{H}_i$ and $\mathrm{C}(\mathcal{H}_i) < \infty$ for all $i \in \mathbb{N}$.*

Theorem A.3 (Sufficient Condition for Generatability in the Limit [21]). *A hypothesis class $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ is generatable in the limit if there exists a finite sequence of subclasses $\mathcal{H}_1, \ldots, \mathcal{H}_n \subseteq \mathcal{H}$ such that $\mathcal{H} = \bigcup_{i=1}^{n} \mathcal{H}_i$ and $\mathrm{C}(\mathcal{H}_i) < \infty$ for all $i \in [n]$.*

B Proofs in Section 3

To simplify the proof, we give a definition of generatable sequence, which is a sequence that belongs to some hypothesis class h.

Definition B.1 (Generatable Sequence). *Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ be a hypothesis class satisfying the r-*UUS *property for some $r > 0$. A (finite or infinite) sequence $\{x_i\}_{i \in I} \subseteq \mathcal{X}$ is called* generatable *if there exists $h \in \mathcal{H}$ such that*

$$\{x_i\}_{i \in I} \subseteq \mathrm{supp}(h),$$

where $I = \{1, \ldots, t\}$ for some $t \in \mathbb{N}$ or $I = \mathbb{N}$.

B.1 Proofs for Uniform Generation

Lemma B.1 (Necessity in Theorem 3.1). *Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ be any hypothesis class satisfying the r-*UUS *property for some $r > 0$ and and let $0 < \varepsilon_a, \varepsilon_g \leq r$. If $\mathrm{C}_{\varepsilon_a, \varepsilon_g}(\mathcal{H}) = \infty$, then $\mathcal{H}$ is* not *$(\varepsilon_a, \varepsilon_g)$-uniformly generatable.*

Proof. Let $\mathcal{G}$ be any generator and fix any proposed uniform threshold $d \in \mathbb{N}$. Since $\mathrm{C}_{\varepsilon_a, \varepsilon_g}(\mathcal{H}) = \infty$, there exists a generatable sequence $z_1, \ldots, z_t$ such that

$$N(\varepsilon_a; \{z_i\}_{i=1}^{t}, \rho) \geq d$$

and, writing $A = \langle z_1, \ldots, z_t \rangle_{\mathcal{H}}$, there is a finite set $F \subseteq A$ with $A \subseteq B(F, \varepsilon_g)$.

Append the points of F as $z_{t+1}, \ldots, z_p$. Since every point of F belongs to every hypothesis in $\mathcal{H}(z_{1:t})$, we have

$$\mathcal{H}(z_{1:p}) = \mathcal{H}(z_{1:t}).$$

Let $\hat{x}_p = \mathcal{G}(z_{1:p})$. If $\hat{x}_p \in A$, then $\hat{x}_p \in B(F, \varepsilon_g)$, so it is not novel; in this case, choose any $h^\star \in \mathcal{H}(z_{1:t})$. Otherwise, by the definition of the closure, choose $h^\star \in \mathcal{H}(z_{1:t})$ such that $\hat{x}_p \notin \mathrm{supp}(h^\star)$.

In either case, the entire prefix lies in $\mathrm{supp}(h^\star)$ and

$$\mathcal{G}(z_{1:p}) \notin \mathrm{supp}(h^\star) \setminus \bigcup_{i=1}^{p} B(z_i, \varepsilon_g).$$

Extend this prefix by any infinite sequence of positive examples from $h^\star$. The generator fails at round $p \geq t$, after the covering-number threshold d has been reached. Since $\mathcal{G}$ and d were arbitrary, $\mathcal{H}$ is not $(\varepsilon_a, \varepsilon_g)$-uniformly generatable. □

Lemma B.2 (Sufficiency in Theorem 3.1). *Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ be any hypothesis class satisfying the r-*UUS *property for some $r > 0$ and and let $0 < \varepsilon_a, \varepsilon_g \leq r$. When $\mathrm{C}_{\varepsilon_a, \varepsilon_g}(\mathcal{H}) < \infty$, there exists a generator $\mathcal{G}$, such that for every $h \in \mathcal{H}$ and any sequence $(x_i)_{i \in \mathbb{N}}$ with $\{x_1, x_2, \ldots\} \subseteq \mathrm{supp}(h)$, if there exists t such that $N(\varepsilon_a; \{x_i\}_{i=1}^{t}, \rho) \geq \mathrm{C}_{\varepsilon_a, \varepsilon_g}(\mathcal{H}) + 1$, then $\mathcal{G}(x_{1:s}) \in \mathrm{supp}(h) \setminus \bigcup_{i=1}^{s} B(x_i, \varepsilon_g)$ for all $s \geq t$.*

Proof. Suppose $\mathrm{C}_{\varepsilon_a, \varepsilon_g}(\mathcal{H}) = d < \infty$. For every generatable sample $x_1, \ldots, x_s$ satisfying

$$N(\varepsilon_a; \{x_i\}_{i=1}^{s}, \rho) \geq d + 1,$$

its closure $A_s = \langle x_1, \ldots, x_s \rangle_{\mathcal{H}}$ has no finite ε_g-cover with centers in A_s. Since $\{x_i\}_{i=1}^{s} \subseteq A_s$, we obtain

$$A_s \setminus \bigcup_{i=1}^{s} B(x_i, \varepsilon_g) \neq \emptyset;$$

otherwise, the observed points themselves would form such a cover.

Define $\mathcal{G}$ to choose any point in this difference whenever the sample is generatable and its covering number is at least $d + 1$, and to output an arbitrary point of $\mathcal{X}$ otherwise.

Fix a target $h \in \mathcal{H}$ and a sequence of its positive examples. If the covering number is at least $d+1$ at round t, then monotonicity of $N(\varepsilon_a; \cdot, \rho)$ under set inclusion ensures that it remains at least $d+1$ at every round $s \geq t$. Moreover, $A_s \subseteq \operatorname{supp}(h)$. Therefore,

$$\mathcal{G}(x_{1:s}) \in \operatorname{supp}(h) \setminus \bigcup_{i=1}^{s} B(x_i, \varepsilon_g) \qquad \text{for all } s \geq t.$$

This proves the desired uniform guarantee. □

Composing Lemmas B.1 and B.2 gives a characterization of uniform generatability. Next we prove Corollary 3.2.

Proof of Corollary 3.2. We only need to show that $\mathrm{C}_{\varepsilon_a,\varepsilon_g}(\mathcal{H}) < \infty$. To do this, suppose $\mathcal{H} = \{h_1, \ldots, h_s\}$. For any $A \subset [s]$ such that $A \neq \emptyset$, define

$$N_{A,\varepsilon_a} = N(\varepsilon_a; \bigcap_{i \in A} \operatorname{supp}(h_i), \rho), \quad N_{A,\varepsilon_g} = N(\varepsilon_g; \bigcap_{i \in A} \operatorname{supp}(h_i), \rho).$$

Since $\varepsilon_g \leq \varepsilon_a$, we have $N_{A,\varepsilon_a} \leq N_{A,\varepsilon_g}$. Define

$$M := \max\Big(\{0\} \cup \{N_{A,\varepsilon_a} : \emptyset \neq A \subseteq [s],\ N_{A,\varepsilon_g} < \infty\}\Big) < \infty.$$

This maximum is finite because there are only finitely many nonempty subsets of $[s]$. Consider any generatable sample $x_1, \ldots, x_t$ whose closure has a finite ε_g-cover with centers in the closure. Let $A = \{i \in [s] : h_i \in \mathcal{H}(x_1, \ldots, x_t)\}$. Then $A \neq \emptyset$, the closure equals $\bigcap_{i \in A} \operatorname{supp}(h_i)$, and $N_{A,\varepsilon_g} < \infty$. Since every observed point belongs to this intersection,

$$N(\varepsilon_a; \{x_i\}_{i=1}^t, \rho) \leq N_{A,\varepsilon_a} \leq M.$$

Thus $\mathrm{C}_{\varepsilon_a,\varepsilon_g}(\mathcal{H}) \leq M < \infty$, as required. □

We also provide an example of a hypothesis class $\mathcal{H}$ consisting of only two hypotheses that is not non-uniformly generatable in ℓ^2.

Example B.3. *For any $r > 0$ and $0 < \varepsilon_a < \varepsilon_g \leq r$, there exists a two-hypothesis class $\mathcal{H} \subseteq \{0,1\}^{\ell^2}$ satisfying the r-*UUS *property that is not $(\varepsilon_a, \varepsilon_g)$-non-uniformly generatable.*

Proof. Let e_k be the k-th standard basis vector of ℓ^2, and choose α with $\varepsilon_a < \alpha < \varepsilon_g$. Define

$$\begin{aligned} a_k &:= \alpha e_{3k}, & S &:= \{0\} \cup \{a_k : k \geq 1\}, \\ T_1 &:= \{2re_{3k+1} : k \geq 1\}, & T_2 &:= \{2re_{3k+2} : k \geq 1\}. \end{aligned}$$

Let $\mathcal{H} = \{h_1, h_2\}$, where

$$\operatorname{supp}(h_1) = S \cup T_1, \qquad \operatorname{supp}(h_2) = S \cup T_2.$$

The r-UUS property is clear from the definition. For every $t \geq 1$, both hypotheses contain $a_1, \ldots, a_t$, and their supports intersect exactly in S. Hence

$$\langle a_1, \ldots, a_t \rangle_{\mathcal{H}} = S \subseteq B(0, \varepsilon_g), \qquad 0 \in S.$$

This is a finite cover with its center in the closure. But our definition of a_i gives

$$N(\varepsilon_a; \{a_1, \ldots, a_t\}, \rho) \longrightarrow \infty, \qquad t \longrightarrow \infty.$$

These samples witness $\mathrm{C}_{\varepsilon_a,\varepsilon_g}(\mathcal{H}) = \infty$, so Theorem 3.1 rules out uniform generation. If a non-uniform generator existed, the maximum of its two target-dependent thresholds would be a uniform threshold for the same generator, a contradiction. □

B.2 Proofs for Non-Uniform Generation

We now prove Theorem 3.3 across two lemmas, starting with the necessity direction.

Lemma B.4 (Necessity in Theorem 3.3)**.** *Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ be any hypothesis class satisfying the r-*UUS *property for some $r > 0$ and and let $0 < \varepsilon_a, \varepsilon_g \leq r$. If $\mathcal{H}$ is $(\varepsilon_a, \varepsilon_g)$-non-uniformly generatable, then there exists a sequence of* non-decreasing, *$(\varepsilon_a, \varepsilon_g)$-*uniformly generatable *classes $\mathcal{H}_1 \subseteq \mathcal{H}_2 \subseteq \cdots$ such that $\mathcal{H} = \bigcup_{n\in\mathbb{N}} \mathcal{H}_n$.*

Proof. Suppose $\mathcal{H}$ is $(\varepsilon_a, \varepsilon_g)$-non-uniformly generatable and $\mathcal{G}$ is a $(\varepsilon_a, \varepsilon_g)$-non-uniform generator for $\mathcal{H}$. For every $h \in \mathcal{H}$, let $d_h \in \mathbb{N}$ be the smallest natural number such that for any sequence $\{x_i\}_{i=1}^{\infty} \subset \operatorname{supp}(h)$, if there exists t such that $N(\varepsilon_a; \{x_i\}_{i=1}^{t}, \rho) \geq d_h$, then $\mathcal{G}(x_{1:s}) \in \operatorname{supp}(h) \setminus \bigcup_{i=1}^{s} B(x_i, \varepsilon_g)$ for all $s \geq t$. Let $\mathcal{H}_n := \{h \in \mathcal{H} : d_h \leq n\}$ for all $n \in \mathbb{N}$. Then, by the definition of $\mathcal{G}$, we know for every $n \in \mathbb{N}$, $\mathcal{G}$ is a $(\varepsilon_a, \varepsilon_g)$-uniform generator for $\mathcal{H}_n$, and therefore $\mathcal{H}_n$ is $(\varepsilon_a, \varepsilon_g)$-uniformly generatable. The proof is complete after noting that $\mathcal{H}_1 \subseteq \mathcal{H}_2 \subseteq \cdots$ and $\mathcal{H} = \bigcup_{n\in\mathbb{N}} \mathcal{H}_n$. □

Lemma B.5 (Sufficiency in Theorem 3.3)**.** *Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ be any hypothesis class satisfying the r-*UUS *property for some $r > 0$ and and let $0 < \varepsilon_a, \varepsilon_g \leq r$. If there exists a sequence of* non-decreasing, *$(\varepsilon_a, \varepsilon_g)$-*uniformly generatable *classes $\mathcal{H}_1 \subseteq \mathcal{H}_2 \subseteq \ldots$ such that $\mathcal{H} = \bigcup_{n\in\mathbb{N}} \mathcal{H}_n$, then $\mathcal{H}$ is $(\varepsilon_a, \varepsilon_g)$-non-uniformly generatable.*

Proof. Suppose $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ is a class satisfying the r-UUS property for which there exists a sequence of non-decreasing, $(\varepsilon_a, \varepsilon_g)$-uniformly generatable classes $\mathcal{H}_1 \subseteq \mathcal{H}_2 \subseteq \ldots$ with $\mathcal{H} = \bigcup_{n=1}^{\infty} \mathcal{H}_n$. By definition of uniform generatability, for every $n \in \mathbb{N}$, there exists a uniform generator $\mathcal{G}_n$ for $\mathcal{H}_n$. Let $d_{\mathcal{G}_n}$ denote the $(\varepsilon_a, \varepsilon_g)$-uniform generation sample complexity of $\mathcal{G}_n$ with respect to $\mathcal{H}_n$.

Consider the following generator $\mathcal{G}$. On any sample inconsistent with every hypothesis in $\mathcal{H}$, let $\mathcal{G}$ output an arbitrary point of $\mathcal{X}$. Otherwise, fix $t \in \mathbb{N}$ and a sample $x_1, \ldots, x_t$ such that $|\mathcal{H}(x_1, \ldots, x_t)| \geq 1$. Let $d_t := N(\varepsilon_a; \{x_i\}_{i=1}^{t}, \rho)$. $\mathcal{G}$ first computes $n_t = \max\Big(\{n \in [t] : d_{\mathcal{G}_n} \leq d_t\} \cup \{0\}\Big)$. If $n_t = 0$, $\mathcal{G}$ plays any $\hat{x}_t \in \mathcal{X}$. If $n_t > 0$, $\mathcal{G}$ uses $\mathcal{G}_{n_t}$ to generate new instances, which means $\mathcal{G}(x_1, \ldots, x_t) = \mathcal{G}_{n_t}(x_1, \ldots, x_t)$.

We now prove that such a $\mathcal{G}$ is a $(\varepsilon_a, \varepsilon_g)$-non-uniform generator for $\mathcal{H}$. To that end, let $h^\star$ be the hypothesis chosen by the adversary and suppose that $h^\star$ belongs to $\mathcal{H}_{n^\star}$. Let $d^\star = \max\{d_{\mathcal{G}_{n^\star}}, n^\star\}$. We show that for any sequence $\{x_i\}_{i=1}^{\infty} \subseteq \operatorname{supp}(h^\star)$, if there exists $t^\star \in \mathbb{N}$ such that $N(\varepsilon_a; \{x_i\}_{i=1}^{t^\star}, \rho) \geq d^\star$, then $\mathcal{G}(x_{1:s}) \in \operatorname{supp}(h^\star) \setminus \bigcup_{i=1}^{s} B(x_i, \varepsilon_g)$ for all $s \geq t^\star$.

Fix any valid sequence $\{x_i\}_{i=1}^{\infty} \subseteq \operatorname{supp}(h^\star)$, and suppose that $N(\varepsilon_a; \{x_i\}_{i=1}^{t^\star}, \rho) \geq d^\star$ for some $t^\star \in \mathbb{N}$. Fix any $s \geq t^\star$. By definition, $\mathcal{G}$ first computes

$$n_s = \max\Big(\{n \in [s] : d_{\mathcal{G}_n} \leq d_s\} \cup \{0\}\Big).$$

Note that $n_s \geq n^\star$ since $s \geq t^\star \geq d^\star \geq n^\star$ (the covering number of a sample is at most its length) and $d_s \geq d^\star \geq d_{\mathcal{G}_{n^\star}}$ by monotonicity of covering numbers. Thus, $|\mathcal{H}_{n_s}(x_{1:s})| \geq 1$ since $h^\star \in \mathcal{H}_{n^\star}$. Accordingly, by construction of $\mathcal{G}$, it uses $\mathcal{G}_{n_s}$ to generate a new instance. The proof is complete by noting that $h^\star \in \mathcal{H}_{n_s}$ and $d_s \geq d_{\mathcal{G}_{n_s}}$ which guarantees that $\mathcal{G}_{n_s}(x_1, \ldots, x_s) \in \operatorname{supp}(h^\star) \setminus \bigcup_{i=1}^{s} B(x_i, \varepsilon_g)$.

□

We now show that every countable hypothesis class are non-uniformly generatable with $\varepsilon_a \geq \varepsilon_g$.

Proof of Corollary 3.4. Suppose $\mathcal{H}$ is a countable hypothesis class satisfying the r-UUS property. If $\mathcal{H}$ is finite, Corollary 3.2 gives uniform, and hence non-uniform, generatability. Otherwise, $\mathcal{H}$ is countably infinite; fix an enumeration $h_1, h_2, \ldots$ without repetition. Let $\mathcal{H}_n = \{h_1, \ldots, h_n\}$ for all $n \in \mathbb{N}$. Then, $\mathcal{H}_1, \mathcal{H}_2, \ldots$ is a non-decreasing sequence of classes such that $\mathcal{H} = \bigcup_{n\in\mathbb{N}} \mathcal{H}_n$. Moreover, since for every $n \in \mathbb{N}$, we have that $|\mathcal{H}_n| = n < \infty$, Corollary 3.2 gives that $\mathcal{H}_n$ is $(\varepsilon_a, \varepsilon_g)$-uniformly generatable, completing the proof. □

B.3 Proofs for Generation in the Limit

Proof of Theorem 3.5. Fix $r > 0$ and $0 < \varepsilon_a, \varepsilon_g \leq r$. Define

$$A = \{3rn : n \geq 1\}, \qquad B = \{3rn + \varepsilon_a : n \geq 1\},$$

and let $\mathcal{S}$ consist of all sets $S_s = \{s_n : n \geq 1\}$ with

$$s_n \in (3rn, 3rn + \varepsilon_a) \qquad (n \geq 1).$$

Let $\mathcal{H}_1 = \{h_s^A : S_s \in \mathcal{S}\}$ and $\mathcal{H}_2 = \{h_s^B : S_s \in \mathcal{S}\}$, where

$$\operatorname{supp}(h_s^A) = A \cup S_s, \qquad \operatorname{supp}(h_s^B) = B \cup S_s.$$

Clearly both classes and their union satisfies the r-UUS property. To see both classes are uniformly generatable, for $\mathcal{H}_1$, after any nonempty finite sample, consider the generator that outputs the first point of A exceeding the largest observation by more than ε_g. This output belongs to every target in $\mathcal{H}_1$ and is novel. Use B in the same way for $\mathcal{H}_2$. Thus, both classes are uniformly generatable with threshold 1.

Now fix any generator $\mathcal{G}$ for the union. Choose s_n recursively in $(3rn, 3rn + \varepsilon_a)$ so that

$$s_n \notin \{\mathcal{G}(s_1, \ldots, s_t) : 1 \leq t < n\}.$$

Only finitely many points are excluded at each step, so this choice is possible. Set $S = \{s_n : n \geq 1\}$. The sequence $(s_n)_{n \geq 1}$ is positive for both fixed targets h_s^A and h_s^B. Moreover,

$$|s_n - 3rn| < \varepsilon_a, \qquad |s_n - (3rn + \varepsilon_a)| < \varepsilon_a,$$

so it ε_a-covers both supports. In fact, distinct observations are more than $3r - \varepsilon_a \geq 2r$ apart, and hence

$$N(\varepsilon_a; \{s_1, \ldots, s_t\}, |\cdot|) = t.$$

The sets A, B, and S are pairwise disjoint, so the intersection of the two target supports is exactly S. At any round t, the output $\mathcal{G}(s_1, \ldots, s_t)$ cannot equal any future s_n, by the recursive choice. If it belongs to S, it is therefore already observed and not novel. Otherwise it is outside at least one of the two target supports. Thus at every round the generator fails for at least one target. One of these two fixed targets consequently has infinitely many failure rounds along this admissible sequence. Since $\mathcal{G}$ was arbitrary, the union is not $(\varepsilon_a, \varepsilon_g)$-generatable in the limit. □

C Constructing Examples in Section 4

C.1 Useful Lemmas for Constructing Hypothesis Classes

This subsection gives a threshold-transfer lemma for the non-uniform lower bounds, a covering-number estimate for orthogonal points, and a criterion for failure of generation in the limit.

Lemma C.1 (Transfer of non-uniform thresholds)**.** *Let $\mathcal{H}$ satisfy r-*UUS *and let $0 < \varepsilon_a, \varepsilon_g \leq r$. Let $\mathcal{S}$ be a family of nonempty subsets of $\mathcal{X}$, and let $T_1, \ldots, T_q \subseteq \mathcal{X}$ satisfy $\bigcap_{i=1}^q T_i = \emptyset$. Suppose that, for every $S \in \mathcal{S}$ and $i \in [q]$, $\mathcal{H}$ contains a hypothesis $h_{S,i}$ with support $S \cup T_i$. If $\mathcal{G}$ is an $(\varepsilon_a, \varepsilon_g)$-non-uniform generator for $\mathcal{H}$, then for every $S \in \mathcal{S}$ there is a finite threshold d_S such that, for every sequence $(x_j)_{j \geq 1}$ in S, whenever*

$$N(\varepsilon_a; \{x_1, \ldots, x_t\}, \rho) \geq d_S,$$

we have

$$\mathcal{G}(x_{1:s}) \in S \setminus B(\{x_1, \ldots, x_s\}, \varepsilon_g) \qquad \textit{for every } s \geq t.$$

Note that no UUS or coverage assumption is imposed on S. This Lemma will be useful in later constructions of examples

Proof. Let $d_{S,i}$ be a threshold for $\mathcal{G}$ on $h_{S,i}$ and set $d_S = \max_{i \in [q]} d_{S,i}$. A sequence in S is positive for every $h_{S,i}$. Once its covering number reaches d_S, the guarantee for each of these finitely many targets applies at every subsequent round. Intersecting their required output sets gives the conclusion, since

$$\bigcap_{i=1}^q (S \cup T_i) = S \cup \bigcap_{i=1}^q T_i = S.$$

□

We will also use the following elementary estimate. For mutually orthogonal points $v_1, \ldots, v_t$ of norm R in a Hilbert space and $0 < \eta < R$,

$$N(\eta; \{v_1, \ldots, v_t\}, \rho) \geq t \left(1 - \frac{\eta^2}{R^2}\right). \tag{1}$$

Indeed, if a radius-η ball with arbitrary center z contains q of these points, their mean $\bar{v}$ satisfies

$$\eta^2 \geq \frac{1}{q} \sum_{j=1}^{q} \|v_{i_j} - z\|^2 = R^2(1 - 1/q) + \|\bar{v} - z\|^2.$$

Thus each ball contains at most $R^2/(R^2 - \eta^2)$ points, proving the estimate. It applies both to the ordinary and to the weighted Hilbert norms used below.

The next lemma gives a way to defeat generation in the limit. The idea is to make the generator repeatedly commit to a target that is consistent with the observations so far, and then continue the sequence in a way that defeats that commitment. Thinking of the example spaces having infinitely many rows, and each hypothesis contains an entire row and some points from other rows. At stage n, condition 2 provides a target that guide the generator to reveal the entire n-th row. If the generator is tricked to commit to the n-th row, we stop revealing this row and move to the next one. We also reveal one point of the zeroth row at each stage. Condition 1 guarantees that all the observations accumulated this way belong to, and cover, a single final target with the zeroth row. For this final target, condition 3 makes sure that the generator will make a mistake at each stage n, and therefore infinitely many errors eventually.

Lemma C.2. *Let $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ satisfy the r-UUS property for some $r > 0$, and let $0 < \varepsilon_a, \varepsilon_g \leq r$. Suppose there exist $k \in \mathbb{N} \cup \{0\}$, points $y_1, \ldots, y_k \in \mathcal{X}$, a sequence $(x_{0i})_{i \geq 1} \subseteq \mathcal{X}$, and an array $(x_{ij})_{i \geq 1,\, j \geq 1} \subseteq \mathcal{X}$ such that the following hold.*

1. *There exists a set $A_0 \subseteq \mathcal{X}$ such that for every infinite sequence of positive integers $(j_i)_{i \geq 1}$, there exists $h \in \mathcal{H}$ containing all the specified observations:*

$$\{y_\ell : 1 \leq \ell \leq k\} \cup \{x_{0i} : i \geq 1\} \cup \bigcup_{i \geq 1} \{x_{i\ell} : 1 \leq \ell \leq j_i\} \subseteq \mathrm{supp}(h),$$

   and satisfying

$$\mathrm{supp}(h) \subseteq B(\{y_\ell\}_{\ell=1}^{k}, \varepsilon_a) \,\cup\, B(\{x_{0i}\}_{i \geq 1}, \varepsilon_a) \,\cup\, B\Big(\bigcup_{i \geq 1} \{x_{i\ell} : 1 \leq \ell \leq j_i\},\, \varepsilon_a\Big),$$

   and

$$\mathrm{supp}(h) \subseteq B(\{y_\ell\}_{\ell=1}^{k}, \varepsilon_g) \,\cup\, A_0 \,\cup\, B\Big(\bigcup_{i \geq 1} \{x_{i\ell} : 1 \leq \ell \leq j_i\},\, \varepsilon_g\Big).$$

2. *For every $n \geq 1$, there exists a set $A_n \subseteq \mathcal{X}$ such that for every finite sequence of positive integers $(j_i)_{i=0}^{n-1}$, there exists $h \in \mathcal{H}$ containing all the specified observations:*

$$\{y_\ell : 1 \leq \ell \leq k\} \cup \{x_{ni} : i \geq 1\} \cup \bigcup_{i=0}^{n-1} \{x_{i\ell} : 1 \leq \ell \leq j_i\} \subseteq \mathrm{supp}(h),$$

   and satisfying

$$\mathrm{supp}(h) \subseteq B(\{y_\ell\}_{\ell=1}^{k}, \varepsilon_a) \,\cup\, B(\{x_{ni}\}_{i \geq 1}, \varepsilon_a) \,\cup\, B\Big(\bigcup_{i=0}^{n-1} \{x_{i\ell} : 1 \leq \ell \leq j_i\},\, \varepsilon_a\Big),$$

   and

$$\mathrm{supp}(h) \subseteq B(\{y_\ell\}_{\ell=1}^{k}, \varepsilon_g) \,\cup\, A_n \,\cup\, B\Big(\bigcup_{i=0}^{n-1} \{x_{i\ell} : 1 \leq \ell \leq j_i\},\, \varepsilon_g\Big).$$

3. *For every $n \geq 1$,*

$$A_n \,\cap\, \Big(A_0 \,\cup\, B(\{x_{ij}\}_{i \geq n+1,\, j \geq 1}, \varepsilon_g)\Big) \,=\, \emptyset.$$

Then $\mathcal{H}$ is not $(\varepsilon_a, \varepsilon_g)$-generatable in the limit.

Proof. Let $\mathcal{G}$ be a generator that ensures $\mathcal{H}$ is generatable in the limit. We show that there exists a hypothesis $h \in \mathcal{H}$ and a sequence $\{z_i\}_{i=1}^{\infty}$, both depending on $\mathcal{G}$, such that

$$\{z_i : i \geq 1\} \subseteq \mathrm{supp}(h) \subseteq B(\{z_i : i \geq 1\}, \varepsilon_a),$$

which leads to a contradiction.

We construct the sequence $\{z_i\}_{i=1}^{\infty}$ inductively. Intuitively, we want the adversary first reveals the points $y_1, \ldots, y_k$, and then at stage m reveals x_{0m} together with a prefix of the m-th row $\{x_{m1}, x_{m2}, \ldots\}$ until the generator outputs a point in A_m.

At each stage m, condition 2 ensures that the already played prefix, the next point x_{0m}, and the entire m-th row belong to a single witnessing hypothesis h_m. Continuing with that row therefore gives a positive sequence satisfying the required ε_a-coverage condition. The second inclusion in condition 2 places every point of $\mathrm{supp}(h_m) \setminus A_m$ within ε_g of the fixed observations already played before continuing the row. Thus eventual correct generation along this continuation forces an output in A_m after a finite, nonempty row prefix.

Let $n_0 = k$ and $z_1 = y_1, z_2 = y_2, \ldots, z_{n_0} = y_k$. From our second assumption, there exists $h_1 \in \mathcal{H}$ such that

$$\mathrm{supp}(h_1) \subseteq B\big(\{y_i\}_{i=1}^{k}, \varepsilon_a\big) \cup B\big(\{x_{1i}\}_{i=1}^{\infty}, \varepsilon_a\big) \cup B(x_{01}, \varepsilon_a).$$

Since $\mathcal{G}$ is generatable in the limit, there exists $j_1 \in \mathbb{N}$ such that

$$\mathcal{G}(z_1, \ldots, z_{n_0}, x_{01}, x_{11}, \ldots, x_{1j_1}) \in A_1.$$

Let $n_1 = n_0 + 1 + j_1$ and $z_{n_0+1} = x_{01}, z_{n_0+2} = x_{11}, z_{n_0+3} = x_{12}, \ldots, z_{n_1} = x_{1j_1}$, so

$$\mathcal{G}(z_1, \ldots, z_{n_1}) \in A_1.$$

Similarly, there exists $h_2 \in \mathcal{H}$ such that

$$\mathrm{supp}(h_2) \subseteq B\big(\{y_i\}_{i=1}^{k}, \varepsilon_a\big) \cup B\big(\{x_{2i}\}_{i=1}^{\infty}, \varepsilon_a\big) \cup B\Big(\{x_{0i}\}_{i=1}^{2}, \varepsilon_a\Big) \cup B\Big(\{x_{1i}\}_{i=1}^{j_1}, \varepsilon_a\Big).$$

Since $\mathcal{G}$ is generatable in the limit, there exists $j_2 \in \mathbb{N}$ such that

$$\mathcal{G}(z_1, \ldots, z_{n_1}, x_{02}, x_{21}, \ldots, x_{2j_2}) \in A_2.$$

Let $n_2 = n_1 + 1 + j_2$ and $z_{n_1+1} = x_{02}, z_{n_1+2} = x_{21}, z_{n_1+3} = x_{22}, \ldots, z_{n_2} = x_{2j_2}$, so

$$\mathcal{G}(z_1, \ldots, z_{n_2}) \in A_2.$$

Now, for some $m \geq 1$, assume we have defined $h_1, \ldots, h_{m-1}, j_1, \ldots, j_{m-1}$ and $n_1, \ldots, n_{m-1}$. Inductively, there exists $h_m \in \mathcal{H}$ such that

$$\mathrm{supp}(h_m) \subseteq B\big(\{y_i\}_{i=1}^{k}, \varepsilon_a\big) \cup B\big(\{x_{mi}\}_{i=1}^{\infty}, \varepsilon_a\big) \cup B\Big(\{x_{0i}\}_{i=1}^{m}, \varepsilon_a\Big) \cup B\Big(\bigcup_{i=1}^{m-1} \{x_{i\ell}\}_{\ell=1}^{j_i}, \varepsilon_a\Big).$$

Since $\mathcal{G}$ is generatable in the limit, there exists $j_m \in \mathbb{N}$ such that

$$\mathcal{G}(z_1, \ldots, z_{n_{m-1}}, x_{0m}, x_{m1}, \ldots, x_{mj_m}) \in A_m.$$

Let $n_m = n_{m-1} + 1 + j_m$ and $z_{n_{m-1}+1} = x_{0m}, z_{n_{m-1}+2} = x_{m1}, z_{n_{m-1}+3} = x_{m2}, \ldots, z_{n_m} = x_{mj_m}$, so

$$\mathcal{G}(z_1, \ldots, z_{n_m}) \in A_m.$$

Finally, condition 1 gives $h_0 \in \mathcal{H}$ containing the entire constructed sequence and such that

$$\mathrm{supp}(h_0) \subseteq B\big(\{y_i\}_{i=1}^{k}, \varepsilon_a\big) \cup B\big(\{x_{0i}\}_{i=1}^{\infty}, \varepsilon_a\big) \cup B\Big(\bigcup_{i=1}^{\infty} \{x_{i\ell}\}_{\ell=1}^{j_i}, \varepsilon_a\Big),$$

which means

$$\mathrm{supp}(h_0) \subseteq B\big(\{z_i\}_{i=1}^{\infty}, \varepsilon_a\big).$$

However, from our construction and first assumption, for any $m \geq 1$,

$$\mathrm{supp}(h_0) \setminus B(\{z_i\}_{i=1}^{n_m}, \varepsilon_g) \subseteq \Big(A_0 \cup B\Big(\{x_{ij}\}_{i \geq m+1, j \geq 1}, \varepsilon_g\Big)\Big)$$

but from our third assumption,

$$A_m \cap \Big(A_0 \cup B\Big(\{x_{ij}\}_{i \geq m+1, j \geq 1}, \varepsilon_g\Big)\Big) = \emptyset,$$

thus when observing $\{z_i\}_{i=1}^{\infty}$, $\mathcal{G}$ will make mistakes infinite times. Thus $\mathcal{H}$ is not generatable in the limit, which completes the proof. □

C.2 Proof of Example 4.5

Let $\rho : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ be the metric defined by

$$\rho(x, y) = |x - y|, \qquad x, y \in \mathbb{R}.$$

Let P denote the set of all positive odd prime integers. For each $p \in P$, define

$$A_p = \{p^n : n \in \mathbb{N}\} \cup \{p^n - 1 : n \in \mathbb{N}\}.$$

Let $2\mathbb{N} = \{2n : n \in \mathbb{N}\}$ denote the set of positive even integers. Define the hypothesis class

$$\mathcal{H} = \Big\{ h \in \{0,1\}^{\mathbb{R}} \;\Big|\; \operatorname{supp}(h) = A_p \cup B, \text{ for some } p \in P \text{ and } B \subseteq 2\mathbb{N} \Big\}.$$

Every support contains the unbounded set A_p. For any $r > 0$, a positive sequence covering the support at radius r must have unbounded range. This range contains arbitrarily large finite sets with pairwise distances greater than $2r$, so its prefix r-covering numbers tend to infinity. Hence $\mathcal{H}$ satisfies r-UUS for every $r > 0$. We claim that $\mathcal{H}$ is $(\varepsilon, 1)$-generatable in the limit for every $\varepsilon \in (0, 1)$, but is not $(1, 1)$-generatable in the limit.

We first show that it is $(\varepsilon, 1)$-generatable in the limit for any $\varepsilon \in (0, 1)$. For any $h \in \mathcal{H}$, $\{x_i\}_{i=1}^{\infty} \subseteq \operatorname{supp}(h)$, if $\operatorname{supp}(h) \subseteq B(\{x_i\}_{i=1}^{\infty}, \varepsilon)$, then every support point must occur among the observations: all support points and observations are integers, so an omitted support point would have distance at least $1 > \varepsilon$ from the observed set. Thus the sequence enumerates $\operatorname{supp}(h)$. As a result, there exists some $t \in \mathbb{N}$ such that $x_t = p^k$ for some $p \in P, k \in \mathbb{N}$. Define $\mathcal{G}$ such that if no observed point is a power of an odd prime, the $\mathcal{G}(x_{1:s}) = x_1$, otherwise, once a p^k is observed, let $\mathcal{G}(x_{1:s}) \in \{p^n | n \in \mathbb{N}\} \setminus B(\{x_i\}_{i=1}^{s}, 1)$. Such an output always exists because the prime powers are unbounded and finitely many radius-1 balls form a bounded set. This guarantees generation in the limit.

Next, we use Lemma C.2 to show that it is not $(1, 1)$-generatable in the limit.

We show that it is not $(1, 1/2)$ generatable, combine with Theorem 4.6, we then know it is not $(1, 1)$-generatable in the limit. Let $\{p_i\}_{i \geq 0}$ denote the sequence of odd prime numbers listed in increasing order, where p_i is the $(i + 1)$-st odd prime; in particular,

$$p_0 = 3, \quad p_1 = 5, \quad p_2 = 7, \ \ldots$$

Let

$$x_{ij} = p_i^j - 1, \qquad i \geq 0, j \geq 1.$$

Apply Lemma C.2 with no initial y-points ($k = 0$) and any UUS radius greater than 1. We verify its three conditions.

1. For any infinite sequence of positive integers $\{j_i\}_{i=1}^{\infty}$, let $h \in \mathcal{H}$ be such that

$$\operatorname{supp}(h) = A_{p_0} \cup \left(\bigcup_{i=1}^{\infty} \{x_{i\ell}\}_{\ell=1}^{j_i} \right)$$

   This support contains the entire current row and all the indicated row prefixes, as required by the membership condition of the lemma. It also satisfies

$$\operatorname{supp}(h) \subseteq B\big(\{x_{0i}\}_{i=1}^{\infty}, 1\big) \cup B\Big(\bigcup_{i=1}^{\infty} \{x_{i\ell}\}_{\ell=1}^{j_i}, 1\Big).$$

   and

$$\operatorname{supp}(h) \subseteq A_{p_0} \cup B\Big(\bigcup_{i=1}^{\infty} \{x_{i\ell}\}_{\ell=1}^{j_i}, 1/2\Big).$$

2. For any $n > 0$ and any finite sequence of positive integers $\{j_i\}_{i=0}^{n-1}$, let $h \in \mathcal{H}$ be such that

$$\operatorname{supp}(h) = A_{p_n} \cup \left(\bigcup_{i=0}^{n-1} \{x_{i\ell}\}_{\ell=1}^{j_i} \right)$$

This support contains the entire current row and all the indicated row prefixes, as required by the membership condition of the lemma. It also satisfies

$$\mathrm{supp}(h) \subseteq B\big(\{x_{ni}\}_{i=1}^{\infty}, 1\big) \cup B\Big(\bigcup_{i=0}^{n-1} \{x_{i\ell}\}_{\ell=1}^{j_i}, 1\Big).$$

and

$$\mathrm{supp}(h) \subseteq A_{p_n} \cup B\Big(\bigcup_{i=0}^{n-1} \{x_{i\ell}\}_{\ell=1}^{j_i}, 1/2\Big).$$

3. For any $n \in \mathbb{N}$, $A_{p_n} \cap A_{p_0} = \emptyset$. Moreover, every point of A_{p_n} is an integer distinct from every $x_{ij} = p_i^j - 1$ with $i \geq n+1$ and $j \geq 1$: odd powers cannot equal these even integers, and equality between the even points would contradict uniqueness of prime factorization. Thus the distance between any such pair is at least 1, so

$$A_{p_n} \cap B(\{x_{ij}\}_{i \geq n+1, j \geq 1}, 1/2) = \emptyset.$$

Thus all three conditions of Lemma C.2 holds, makes $\mathcal{H}$ not $(1, 1/2)$-generatable in the limit.

C.3 Proof of Example 4.8

Proof. Recall that ℓ^2 is the space that contains all sequences $(x_i)_{i=1}^{\infty}$ such that $\sum_i x_i^2 < \infty$. Let $e_k \in \ell^2$ denote the k-th standard basis vector. Define

$$u_k := 4r \cdot e_k, \qquad a_k := \varepsilon_a \cdot e_k, \qquad g_k := \varepsilon_g \cdot e_k.$$

Let $\mathcal{I}$ be the set of all infinite subsets of $\mathbb{N}$, i.e., $\mathcal{I} = \{I | I \in 2^{\mathbb{N}}, |I| = \infty\}$. For $I \in \mathcal{I}$, let

$$U_I = \{u_k | k \in I\}, \quad O_I = \{a_{2k} | k \in I\} \cup \{g_{2k^n+1} | k \in I, n \in \mathbb{N}\}, \quad D_I = \{g_{2k+1} | k \in I\}$$

and construct the hypothesis class $\mathcal{H}$ such that

$$\mathcal{H} = \{h_{I_1,I_2,I_3} | I_1, I_2, I_3 \in \mathcal{I}, \quad \mathrm{supp}(h_{I_1,I_2,I_3}) = \{\mathbf{0}\} \cup U_{I_1} \cup O_{I_2} \cup D_{I_3}\}$$

Here, u_k and U_I are constructed to ensure that $\mathcal{H}$ satisfies the r-UUS property. As will be seen in the proof, the points a_k are constructed for the adversary, while the points g_k are constructed for the generator. In fact, the g_k are the only instances with an explicit structure, which enables the generator to repeatedly produce them. The set O_I is the *optimal set*: when $\gamma_g < \varepsilon_g$, revealing $a_{2k} \in O_I$ with $k \geq 2$ identifies an infinite tower $\{g_{2k^n+1} : n \geq 1\}$ from which novel outputs can be chosen. In contrast, D_I is the *disturbing set*: if the adversary continues to reveal $g_k \in D_I$, then the generator fails to generate in the limit.

Every u-point in the support has distance at least $3r$ from every other support point: the a- and g-points have norm at most r, whereas distinct u-points are at distance $4\sqrt{2}r$. Thus every positive sequence covering the support at radius r must observe all points of the infinite set U_{I_1}. Since each radius-r ball contains at most one u-point, the prefix r-covering numbers tend to infinity. This verifies r-UUS.

Case 1: $\gamma_a \in (0, \varepsilon_a)$ and $\gamma_g \in (0, \varepsilon_g)$. Fix a target h and a positive sequence with $\mathrm{supp}(h) \subseteq B(\{x_i : i \geq 1\}, \gamma_a)$. Each a_{2k} in the support has distance at least ε_a from every other support point. Indeed, its distance from 0 is ε_a, other a-points are at distance $\sqrt{2}\varepsilon_a$, all g-points have odd coordinates and are at distance $\sqrt{\varepsilon_a^2 + \varepsilon_g^2}$, and a u-point on the same coordinate is at distance $4r - \varepsilon_a \geq \varepsilon_a$ (other u-points are farther away). Therefore every such a_{2k} must itself be observed, even under the infimum-based coverage definition.

Since I_2 is infinite, eventually an observation a_{2K} with $K \geq 2$ occurs. Before the first such observation, let the generator output an arbitrary point. Afterwards, fix that K. Its tower $\{g_{2K^n+1} : n \geq 1\}$ is infinite and lies in the target support. At each round s, choose the least $n \geq 1$ for which the coordinate $2K^n + 1$ has not occurred among the nonzero observations $x_1, \ldots, x_s$, and output g_{2K^n+1}. Such an n exists because the observed sample is finite. The output has distance $\varepsilon_g > \gamma_g$ from 0 and, by orthogonality, distance at least $\varepsilon_g > \gamma_g$ from every other observation. Thus the generator is eventually always correct and novel.

We next prove the two negative regimes, using Lemma C.2 in Case 2 and a direct diagonal argument in Case 3.

Case 2: $\gamma_a \in [\varepsilon_a, r]$ and $\gamma_g \in (0, \varepsilon_g)$. Let $\{p_i\}_{i\geq 1} = \{3, 5, 7, \dots\}$ be the sequence of odd prime numbers. Using Lemma C.2, let $k = 1$, $y_1 = \mathbf{0}$, for any $i \geq 1, j \geq 1$ let

$$x_{ij} = \begin{cases} u_{p_i^{(j+1)/2}} & j \text{ is odd,} \\ g_{2\cdot p_i^{j/2}+1} & j \text{ is even.} \end{cases}$$

Moreover, let

$$x_{0j} = \begin{cases} u_{2^{(j+1)/2}} & j \text{ is odd,} \\ g_{2\cdot 2^{j/2}+1} & j \text{ is even.} \end{cases}$$

We verify that the conditions of Lemma C.2 holds.

1. For any infinite sequence of positive integers $\{j_i\}_{i=1}^{\infty}$, from our construction of x_{ij}, we can find $J_1, J_2 \in 2^{\mathbb{N}}$ such that

$$\bigcup_{i=1}^{\infty} \{x_{i\ell}\}_{\ell=1}^{j_i} = \{u_k | k \in J_1\} \cup \{g_{2k+1} | k \in J_2\}.$$

   Then, if we let $I_2 = \{2^n | n \in \mathbb{N}\}$, $I_1 = I_2 \cup J_1$, $I_3 = I_2 \cup J_2$, since $\gamma_a \geq \varepsilon_a$, $B(\mathbf{0}, \gamma_a)$ can cover all the points of the form a_k. Thus, $h = h_{I_1,I_2,I_3}$ satisfies

$$\operatorname{supp}(h) \subseteq B(\mathbf{0}, \gamma_a) \cup B(\{x_{0i}\}_{i=1}^{\infty}, \gamma_a) \cup B(\bigcup_{i=1}^{\infty} \{x_{i\ell}\}_{\ell=1}^{j_i}, \gamma_a)$$

   Moreover, if

$$A_0 = \bigcup_{k\in\mathbb{N}} \left\{u_{2^k}, a_{2\cdot 2^k}, g_{2\cdot 2^k+1}\right\},$$

   that is, A_0 contains the "protected" part for the $n = 0$ hypothesis (a whole copy of a $u/a/g$-pattern indexed by powers of 2). We then know

$$\operatorname{supp}(h) \subseteq B(\mathbf{0}, \gamma_g) \cup A_0 \cup B(\bigcup_{i=1}^{\infty} \{x_{i\ell}\}_{\ell=1}^{j_i}, \gamma_g).$$

   The support also contains 0, the entire zeroth row, and every selected row prefix. Thus both the membership and covering requirements of the first condition hold.

2. For $n \geq 1$ and any prescribed prefixes of rows $0, \dots, n-1$, write their union as

$$\bigcup_{i=0}^{n-1} \{x_{i\ell} : 1 \leq \ell \leq j_i\} = \{u_k : k \in J_1\} \cup \{g_{2k+1} : k \in J_2\}.$$

   Set $Q_n = \{p_n^k : k \geq 1\}$ and choose

$$I_1 = Q_n \cup J_1, \qquad I_2 = Q_n, \qquad I_3 = Q_n \cup J_2.$$

   All three sets are infinite. The resulting support contains 0, all prescribed prefixes, and the entire n-th row. Since powers of powers of p_n remain powers of p_n, its remaining points are exactly the a-points in

$$A_n = \bigcup_{k\geq 1} \{u_{p_n^k}, a_{2p_n^k}, g_{2p_n^k+1}\}.$$

   These a-points lie in $B(0, \gamma_a)$, so the adversarial covering requirement holds. The whole support is contained in $\{0\} \cup A_n$ together with the prescribed prefixes, which gives the generator-scale covering requirement.

3. Uniqueness of prime factorization gives $A_n \cap A_0 = \emptyset$: within each point type the underlying prime powers differ; a-points have even coordinates whereas g-points have odd coordinates; and u-points have norm $4r$, different from either other type. Consider a point of A_n and a point in a later row $i > n$, which contains only u- and g-points. Distinct u-points are at distance $4\sqrt{2}r$, and distinct g-points are at distance $\sqrt{2}\varepsilon_g$. A u-point and a g-point, including when they share a coordinate, have distance at least $4r - \varepsilon_g \geq 3r$. An a-point

in A_n has an even coordinate, whereas both types in a later row have odd coordinates; its distances to them are $\sqrt{\varepsilon_a^2 + 16r^2}$ and $\sqrt{\varepsilon_a^2 + \varepsilon_g^2}$. All these distances are at least $\varepsilon_g > \gamma_g$. This uniform bound also excludes membership in the infimum-based neighborhood of all future rows. Hence condition 3 holds.

All three conditions of Lemma C.2 hold, so $\mathcal{H}$ is not generatable in the limit when $\gamma_a \geq \varepsilon_a$ and $0 < \gamma_g < \varepsilon_g$.

Case 3: $\gamma_g \in [\varepsilon_g, r]$. Fix any $0 < \gamma_a \leq r$ and any generator $\mathcal{G}$. Let

$$D = \{g_{2k+1} : k \geq 1\}.$$

For every infinite set P of odd primes, the choices

$$I_1 = \{2p : p \in P\}, \qquad I_2 = P, \qquad I_3 = \mathbb{N}_+$$

define a hypothesis $h_P \in \mathcal{H}$ with

$$\operatorname{supp}(h_P) = \{0\} \cup D \cup \{u_{2p}, a_{2p} : p \in P\}.$$

Indeed, all g-towers required by $I_2 = P$ already lie in D.

We choose P and a sequence of positive observations inductively. First reveal 0. Maintain finite sets of selected and forbidden odd primes, initially empty. At stage $m \geq 1$, choose an odd prime p_m in neither set, add it to the selected set, and reveal g_{2m+1}, u_{2p_m}, and a_{2p_m}. Inspect the output of $\mathcal{G}$ at the end of this stage. If it is u_{2q} or a_{2q} for an odd prime q not yet selected, add q to the forbidden set. Only finitely many primes are unavailable at any finite stage, so the construction continues indefinitely.

Let $P = \{p_m : m \geq 1\}$. This set is infinite, and the resulting sequence enumerates exactly $\operatorname{supp}(h_P)$, hence is positive and satisfies the coverage requirement for every $\gamma_a > 0$. At the end of each stage, an output in $\{0\} \cup D$ is not novel because all such points lie in $B(0, \gamma_g)$. An output from a pair belonging to an already selected prime has itself been observed. An output from a pair belonging to an unselected odd prime is excluded from the final support by the forbidden-set rule. Every other output lies outside $\operatorname{supp}(h_P)$. Thus $\mathcal{G}$ fails at the end of every stage. Since $\mathcal{G}$ was arbitrary, $\mathcal{H}$ is not (γ_a, γ_g)-generatable in the limit throughout this regime, including $\gamma_g = \varepsilon_g$. □

C.4 Proof of Example 4.9

Proof. Use the construction in the main text: $K = \{2, 3, \dots\}$,

$$u_k = 2r\, e_{4k}, \qquad a_{k,1} = \varepsilon_{a1} e_{4k+1},$$
$$a_{k,2} = \varepsilon_{a2} e_{4k+2}, \qquad g_k = \varepsilon_g e_{4k+3},$$

and

$$O_I = \{u_k : k \in I\} \cup \{g_{k^n} : k \in I,\ n \geq 1\},$$
$$D_{I,J} = \{a_{k,1} : k \in I\} \cup \{a_{k,2} : k \in J\},$$
$$\operatorname{supp}(h_{I_1,I_2,J}) = \{0\} \cup O_{I_1} \cup D_{I_2,J},$$

where $I_1, I_2 \subseteq K$ are infinite and $J \subseteq K$ is finite. The four coordinate sets are disjoint even when some scale parameters coincide. Every u-point has distance at least $2r$ from every other support point. Hence every positive sequence covering the support at radius r must observe all its infinitely many u-points. Their pairwise distances are $2\sqrt{2}r > 2r$, so each radius-r ball contains at most one of them. The sample's r-covering number is therefore at least the number of observed u-points and tends to infinity. This proves r-UUS.

For all the positive results below, use the following single generator. Until an observation of the form u_k or g_k occurs, output 0. After the first such observation, retain its index $k \geq 2$ and, at each round, output the first point of $\{g_{k^n} : n \geq 1\}$ whose coordinate has not appeared among the nonzero observations. This defines the generator on positive samples; on other inputs it may output arbitrarily. The tower belongs to the target: if g_k was observed, then $k = j^b$ for some $j \in I_1$ and $b \geq 1$, and $k^n = j^{bn}$. The tower is infinite because $k \geq 2$. Its selected point has distance ε_g from 0 and, by orthogonality, at least ε_g from every observation on another coordinate. Thus every output after the first tower observation is correct and novel whenever $\gamma_g < \varepsilon_g$.

Case 1: $\gamma_a \in [\varepsilon_{a2}, r]$ and $\gamma_g \in (0, \varepsilon_g)$. The set $\{0\} \cup D_{I_2,J}$ lies in $B(0, \gamma_a)$. Therefore a positive sample with γ_a-covering number at least 2 must contain a point of O_{I_1}. The generator just defined is consequently uniform with threshold 2.

Case 2: $\gamma_a \in [\varepsilon_{a1}, \varepsilon_{a2})$ and $\gamma_g \in (0, \varepsilon_g)$. The origin and all first-family a-points lie in $B(0, \gamma_a)$, and the $|J|$ second-family points can each be covered by a ball centered at that point. Thus

$$N(\gamma_a; \{0\} \cup D_{I_2,J}, \rho) \le 1 + |J|.$$

A sample covering number of at least $|J| + 2$ forces an observation from O_{I_1}. The same generator is therefore non-uniform.

To rule out uniform generation, let $F_t = \{a_{k,2} : 2 \le k \le t+1\}$. Its closure is exactly $\{0\} \cup F_t$. Indeed, choose $J = \{2, \dots, t+1\}$ to exclude every other second-family point; any first-family point can be excluded by choosing I_2 to omit its index; and the two choices $I_1 = \{2^j : j \ge 1\}$ and $I_1 = \{3^j : j \ge 1\}$ have disjoint O-sets. This finite closure has a finite internal γ_g-cover, whereas (1) gives

$$N(\gamma_a; F_t, \rho) \ge t \left(1 - \frac{\gamma_a^2}{\varepsilon_{a2}^2}\right) \longrightarrow \infty.$$

Hence $\mathrm{C}_{\gamma_a,\gamma_g}(\mathcal{H}) = \infty$, and Theorem 3.1 rules out uniform generation. If $\varepsilon_{a1} = \varepsilon_{a2}$, this case is empty.

Case 3: $\gamma_a \in (0, \varepsilon_{a1})$ and $\gamma_g \in (0, \varepsilon_g)$. First, generation in the limit holds. Every u_k has distance at least $2r$ from every other support point: its coordinate set is disjoint from the other point types, and different u-points are farther apart. Since $\gamma_a \le r < 2r$, a positive sequence covering the support at radius γ_a must observe each u_k in the support. In particular it eventually identifies a tower, after which the generator is always correct and novel.

For the non-uniform lower bound, set

$$S_I = \{0\} \cup \{a_{k,1} : k \in I\}, \qquad T_1 = O_{\{2^j : j \ge 1\}}, \quad T_2 = O_{\{3^j : j \ge 1\}}.$$

Here I ranges over infinite subsets of K. The sets T_1, T_2 are disjoint, and $S_I \cup T_i$ is a support in $\mathcal{H}$ (take $J = \emptyset$). Suppose a non-uniform generator $\mathcal{G}$ existed. Lemma C.1 would give it a finite threshold on every S_I.

We construct a contradiction to that conclusion. Reveal 0, and maintain finite sets of selected and forbidden indices in K, initially empty. At stage t, choose k_t in neither set, select it, and reveal $a_{k_t,1}$. If the resulting output is $a_{q,1}$ for an unselected $q \in K$, forbid q for all subsequent stages. Let $I = \{k_t : t \ge 1\}$. The sequence is positive for S_I. At the end of each stage, outputs equal to 0 or to already selected points are not novel, outputs at unselected first-family indices are excluded from S_I, and every other output is outside S_I. Nevertheless, after t stages its sample covering number is at least

$$t \left(1 - \frac{\gamma_a^2}{\varepsilon_{a1}^2}\right) \longrightarrow \infty$$

by (1). It therefore eventually exceeds the finite threshold supplied by the lemma, a contradiction. This argument does not assume any coverage or generator-scale UUS property for S_I.

Case 4: $\gamma_g \in [\varepsilon_g, r]$. Fix any $0 < \gamma_a \le r$ and any generator $\mathcal{G}$. Restrict to targets with $I_1 = I_2 = P$, $J = \emptyset$, where P is an infinite set of odd primes. Their supports are

$$\{0\} \cup \{u_p, a_{p,1} : p \in P\} \cup \{g_{p^n} : p \in P,\ n \ge 1\}.$$

Reveal 0. At stage t, choose a previously unselected and unforbidden odd prime p_t, and reveal u_{p_t} and $a_{p_t,1}$. Also reveal all not-yet-observed $g_{p_i^n}$ with $1 \le i, n \le t$. At the end of the stage, if the generator outputs u_q or $a_{q,1}$ for an unselected odd prime q, forbid q. Only finitely many primes have been selected or forbidden at any finite stage, so the construction continues.

For $P = \{p_t : t \ge 1\}$, the resulting sequence enumerates exactly the target support and is valid for every $\gamma_a > 0$. At every stage end, all g-points and 0 are non-novel because they lie in $B(0, \gamma_g)$. Points in pairs for selected primes have already been observed, and predictions of pairs for unselected primes are excluded from the final support. All other outputs are outside the support. Thus $\mathcal{G}$ fails infinitely often, proving the claimed impossibility, including the boundary $\gamma_g = \varepsilon_g$. □

C.5 Proof of Example 4.11

Proof. Fix $r > 0$. Retain the weighted Hilbert metric

$$\rho'(x,y) = \left(\sum_{j\geq 1} \delta_j (x_j - y_j)^2\right)^{1/2}, \qquad \delta_j = \begin{cases} 1, & j \text{ odd}, \\ 1/4, & j \text{ even}. \end{cases}$$

Writing ρ for the ordinary ℓ^2 metric, we have $\rho/2 \leq \rho' \leq \rho$, so the metrics are bi-Lipschitz equivalent. For $m, k \geq 1$, define

$$\alpha_m = 2^{3-m} r, \qquad a_{m,k} = \alpha_m e_k.$$

Let $\mathcal{I}$ be the infinite subsets of $\mathbb{N}_+ = \{1, 2, \dots\}$, and set

$$A_{m,I} = \{a_{m,\, 2^n(2k+1)} : k \in I,\ n \geq 0\}.$$

Define $\mathcal{H} = \{h_{\mathbf{I}} : \mathbf{I} = (I_m)_{m\geq 1} \in \mathcal{I}^{\mathbb{N}_+}\}$ by

$$\operatorname{supp}(h_{\mathbf{I}}) = \{0\} \cup \bigcup_{m\geq 1} A_{m,I_m}.$$

Distinct layers consist of points with different ordinary norms, so their point sets are disjoint. Within a layer, different odd indices generate disjoint doubling orbits.

Every support contains infinitely many first-layer odd-coordinate points, each of norm $\alpha_1 = 4r$ under both metrics. Each such point has distance at least $2r$ from every other support point: later-layer points have norm at most $2r$, and distinct first-layer points occupy orthogonal coordinates. Thus a positive sequence covering the support at radius r, under either metric, must observe all these odd-coordinate points. Their pairwise distances are $4\sqrt{2}r > 2r$ under both metrics, so every radius-r ball contains at most one of them. The prefix r-covering numbers therefore tend to infinity. This proves r-UUS under both metrics.

Uniform generation under ρ. Fix $0 < \varepsilon < r$. If a positive sample has ε-covering number at least 2, it contains a point of ordinary norm greater than ε: otherwise the sample would lie in $B_\rho(0, \varepsilon)$. Write such an observation as $a_{m,k}$, where $\alpha_m > \varepsilon$. Its full forward orbit

$$\{a_{m,2^n k} : n \geq 1\}$$

belongs to the target support. Indeed, multiplying its coordinate by another power of 2 preserves the odd part indexing that orbit.

Define the generator to output 0 until the first observation of norm greater than ε. Thereafter retain that observation $a_{m,k}$ and at each round output the first point in its forward orbit whose coordinate is absent from the nonzero observations. The orbit is infinite, so this is always possible. The output has distance $\alpha_m > \varepsilon$ from 0 and distance at least $\alpha_m > \varepsilon$ from every observation on another coordinate. Thus the generator is correct and novel as soon as the sample covering number reaches 2, and at every later round. Define its output arbitrarily on other inputs. This proves $(\varepsilon, \varepsilon)$-uniform generation with threshold 2.

Failure of non-uniform generation under ρ'. Fix any $0 < \varepsilon < r$. Choose the unique $m \geq 3$ such that

$$\alpha_m/2 \leq \varepsilon < \alpha_m.$$

These intervals partition $(0, r)$, including their lower endpoints. In layer m, odd-coordinate points have weighted norm $\alpha_m > \varepsilon$, while all even-coordinate descendants have weighted norm $\alpha_m/2 \leq \varepsilon$.

For $I \in \mathcal{I}$, let $S_I = \{0\} \cup A_{m,I}$. Choose the two disjoint infinite index sets

$$E_1 = \{2j : j \geq 1\}, \qquad E_2 = \{2j - 1 : j \geq 1\},$$

and define completions

$$T_b = \bigcup_{\ell\geq 1,\ \ell\neq m} A_{\ell,E_b}, \qquad b = 1, 2.$$

The sets T_1, T_2 are disjoint: different layers have different norms, and within any layer the disjoint sets of odd parts give disjoint orbits. For every I, both $S_I \cup T_1$ and $S_I \cup T_2$ are supports in $\mathcal{H}$.

Suppose that $\mathcal{G}$ were an $(\varepsilon, \varepsilon)$-non-uniform generator under ρ'. Lemma C.1, applied with this metric, would give a finite threshold for $\mathcal{G}$ on each S_I. We contradict this using only the origin and odd-coordinate observations from layer m.

Reveal 0. Maintain finite sets of selected and forbidden indices in $\mathbb{N}_+$, initially empty. At stage t, select an index k_t in neither set and reveal $a_{m,2k_t+1}$. If the generator outputs an odd-coordinate point $a_{m,2q+1}$ with $q \geq 1$ unselected, forbid q. Let $I = \{k_t : t \geq 1\}$. This defines an infinite set and a sequence positive for S_I. At every stage end, an output at a selected odd-coordinate point has already been observed; an output at an unselected odd-coordinate point is excluded from S_I. Every even-coordinate point in S_I lies in $B_{\rho'}(0, \varepsilon)$ and is therefore non-novel, including when $\alpha_m/2 = \varepsilon$. The origin is also non-novel, and all other outputs are outside S_I.

Finally, the observed odd-coordinate points are mutually orthogonal with weighted norm α_m. By (1),

$$N(\varepsilon; \{0, a_{m,2k_1+1}, \ldots, a_{m,2k_t+1}\}, \rho') \geq t \left(1 - \frac{\varepsilon^2}{\alpha_m^2}\right) \longrightarrow \infty.$$

Thus the sequence eventually crosses the finite threshold supplied by the transfer lemma while the generator continues to fail, a contradiction. Since $\varepsilon \in (0, r)$ was arbitrary, the same class fails non-uniform generation under ρ' at every scale required by the statement. □

D Other Technical Details

D.1 Proofs Related to the UUS Property

This subsection collects short proofs establishing a basic monotonicity result for the UUS property in general metric spaces, as well as a stronger scale-invariance result in doubling spaces. We first show that the UUS property is monotone in the scale parameter.

Proof of Lemma 2.1. Fix $0 < \delta < r$, $h \in \mathcal{H}$, and a sequence $(x_i)_{i \geq 1}$ in $\operatorname{supp}(h)$ such that

$$\operatorname{supp}(h) \subseteq B(\{x_i : i \geq 1\}, \delta).$$

Since $\delta < r$, this sequence also covers the support at radius r. Hence r-UUS and monotonicity of covering numbers give

$$N(\delta; \{x_1, \ldots, x_t\}, \rho) \geq N(r; \{x_1, \ldots, x_t\}, \rho) \longrightarrow \infty.$$

The supports are nonempty by assumption, so $\mathcal{H}$ satisfies δ-UUS. □

The previous lemma shows monotonicity of the UUS property in the scale parameter. Under the additional assumption that $(\mathcal{X}, \rho)$ is doubling, we now show that UUS is in fact scale-invariant.

Proof of Theorem 4.1. Fix $h \in \mathcal{H}$. Its support is nonempty and unbounded. Indeed, if it were bounded, doubling would give a finite radius-r cover of the support. A countable dense sequence in the support would then cover it at radius r while having uniformly bounded prefix r-covering numbers, contradicting r-UUS.

Now fix $r' > 0$ and any positive sequence covering the support at radius r'. Its range must be unbounded, and therefore contains arbitrarily large finite sets with pairwise distances greater than $2r'$. Each such set occurs in a finite prefix, and each radius-r' ball contains at most one of its points. Consequently,

$$N(r'; \{x_1, \ldots, x_t\}, \rho) \longrightarrow \infty.$$

Thus $\mathcal{H}$ satisfies r'-UUS for every $r' > 0$. □

D.2 Recovering the Countable Framework

This subsection shows that our metric-space formulation strictly generalizes the countable example space model studied in [21]. In particular, when the instance space is equipped with the discrete metric at a sufficiently small scale, our notions of novelty, coverage, and generatability reduce exactly to their counterparts in the countable setting. Thus, all generatability notions considered in this paper are consistent with the classical framework.

Proposition D.1. *Let $\mathcal{X}$ be a countable set equipped with the discrete metric*

$$\rho(x,y) = \begin{cases} 1, & x \neq y, \\ 0, & x = y. \end{cases}$$

Let $r < 1$ and $0 < \varepsilon_a, \varepsilon_g \leq r$. Suppose $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ satisfies the r-UUS property. Then $\mathcal{H}$ is $(\varepsilon_a, \varepsilon_g)$-generatable in the limit (respectively, $(\varepsilon_a, \varepsilon_g)$-non-uniformly generatable, $(\varepsilon_a, \varepsilon_g)$-uniformly generatable) if and only if $\mathcal{H}$ is generatable in the limit (respectively, non-uniformly generatable, uniformly generatable) in the sense of [21] (see Appendix A for the definitions).

Proof sketch. Since $0 < r < 1$, every radius-r ball in the discrete metric is a singleton. A positive sequence covers the support at radius r if and only if it enumerates the entire support, and each prefix covering number equals its number of distinct observations. Thus r-UUS is equivalent to every support being infinite.

Likewise, since $0 < \varepsilon_a, \varepsilon_g \leq r < 1$, the sample's ε_a-covering number equals its number of distinct observations, ε_a-coverage requires enumeration of the support, and ε_g-novelty means not having been observed. The three generation notions therefore coincide with their countable counterparts. □

D.3 Definition of Metric Equivalence

We formally define equivalent metric here for completeness sake.

Definition D.1 (Equivalent metrics)**.** *Let ρ_1 and ρ_2 be two metrics on the same space $\mathcal{X}$. We say that ρ_1 and ρ_2 are* equivalent *if there exist constants $0 < c \leq C < \infty$ such that for all $x, y \in \mathcal{X}$,*

$$c\,\rho_1(x,y) \leq \rho_2(x,y) \leq C\,\rho_1(x,y).$$

D.4 Other Proofs in Section 4

This subsection collects proofs in Section 4, showing how generatability properties behave under changes of scale and metric. We first establish that, in doubling metric spaces, generatability is invariant under equivalent metrics via the closure dimension. We then show that, in general metric spaces, generatability notions are monotone with respect to the novelty parameters and are stable under Lipschitz changes of the metric, with appropriate rescaling of parameters.

Proof of Theorem 4.4. In a doubling space, the argument for Theorem 4.1 shows that UUS holds precisely when every support is nonempty and unbounded. Bi-Lipschitz equivalent metrics have the same bounded sets, and both are doubling. Hence UUS is preserved. Assume these UUS conditions hold. Choose any $\varepsilon_a, \varepsilon_g > 0$, denote $\mathrm{C}_{\varepsilon_a,\varepsilon_g,1}(\mathcal{H})$ and $\mathrm{C}_{\varepsilon_a,\varepsilon_g,2}(\mathcal{H})$ as the $(\varepsilon_a, \varepsilon_g)$-closure dimension of $\mathcal{H}$ with respect to ρ_1 and ρ_2. By Theorem 3.1 and 3.3, we only need to show that $\mathrm{C}_{\varepsilon_a,\varepsilon_g,1}(\mathcal{H}) < \infty$ if and only if $\mathrm{C}_{\varepsilon_a,\varepsilon_g,2}(\mathcal{H}) < \infty$. Since $\mathcal{X}$ is a doubling space and ρ_1 is equivalent with ρ_2, for any subset $A \subseteq \mathcal{X}$,

$$N(\varepsilon_a; A, \rho_1) = \infty \iff N(\varepsilon_a; A, \rho_2) = \infty.$$

and there exist constants $C_1, C_2 > 0$, independent of A, such that whenever $N(\varepsilon_a; A, \rho_1) < \infty$,

$$C_1 N(\varepsilon_a; A, \rho_1) \leq N(\varepsilon_a; A, \rho_2) \leq C_2 N(\varepsilon_a; A, \rho_1).$$

For completeness, finite internal covers at the generator scale are also preserved. Choose $L > 0$ such that $\rho_2 \leq L\rho_1$. A finite ρ_1-cover of a nonempty set K at radius ε_g, with centers in K, gives a finite ρ_2-cover at radius $L\varepsilon_g$ with the same centers. By doubling for ρ_2, refine these balls to finitely many balls of radius $\varepsilon_g/2$, and choose a point of K from each smaller ball that meets K. These points form a finite internal ρ_2-cover at radius ε_g. The converse follows by interchanging the metrics. Here doubling holds for both metrics because they are bi-Lipschitz equivalent. Thus the same samples have closures with finite internal covers under the two metrics, and the displayed bounds imply equivalence of finite closure dimension. Theorem 3.1 proves the uniform case, and applying the comparison to every subclass in Theorem 3.3 proves the non-uniform case. □

Proof of Theorem 4.6. Suppose $\mathcal{H}$ is $(\varepsilon_a, \varepsilon_g)$-generatable in the limit, and let $\mathcal{G}$ be an $(\varepsilon_a, \varepsilon_g)$-generator. We show that $\mathcal{G}$ is also a (δ_a, δ_g)-generator, which implies that $\mathcal{H}$ is (δ_a, δ_g)-generatable in the limit.

Let $\{x_i\}_{i=1}^{\infty} \subseteq \mathrm{supp}(h)$ be a sequence such that

$$\mathrm{supp}(h) \subseteq B(\{x_i\}_{i=1}^{\infty}, \delta_a).$$

Since $\delta_a \leq \varepsilon_a$, we also have

$$\mathrm{supp}(h) \subseteq B(\{x_i\}_{i=1}^{\infty}, \varepsilon_a)\,.$$

By the definition of an $(\varepsilon_a, \varepsilon_g)$-generator, there exists s such that for all $t \geq s$,

$$\mathcal{G}(x_{1:t}) \in \mathrm{supp}(h) \setminus B\big(\{x_i\}_{i=1}^{t}, \varepsilon_g\big)\,.$$

Since $\varepsilon_g \geq \delta_g$, it follows that

$$\mathcal{G}(x_{1:t}) \in \mathrm{supp}(h) \setminus B\big(\{x_i\}_{i=1}^{t}, \delta_g\big)\,.$$

This completes the proof. □

Proof of Theorem 4.7. We prove the result for uniform generatability; the non-uniform case follows by an analogous argument.

Suppose $\mathcal{H}$ is $(\varepsilon_a, \varepsilon_g)$-uniformly generatable, and let $\mathcal{G}$ be an $(\varepsilon_a, \varepsilon_g)$-uniform generator. Let $d \in \mathbb{N}$ be the smallest integer such that for any sequence $\{x_i\}_{i=1}^{\infty} \subseteq \mathrm{supp}(h)$, if there exists t with

$$N(\varepsilon_a; \{x_i\}_{i=1}^{t}, \rho) \geq d,$$

then

$$\mathcal{G}(x_{1:s}) \in \mathrm{supp}(h) \setminus B(\{x_i\}_{i=1}^{s}, \varepsilon_g) \quad \text{for all } s \geq t.$$

We show that $\mathcal{G}$ is also a (δ_a, δ_g)-uniform generator.

Fix any sequence $\{x_i\}_{i=1}^{\infty} \subseteq \mathrm{supp}(h)$ and suppose that for some t,

$$N(\delta_a; \{x_i\}_{i=1}^{t}, \rho) \geq d.$$

Since $\delta_a \geq \varepsilon_a$, it follows that

$$N(\varepsilon_a; \{x_i\}_{i=1}^{t}, \rho) \geq d.$$

By the definition of $\mathcal{G}$, for all $s \geq t$ we therefore have

$$\mathcal{G}(x_{1:s}) \in \mathrm{supp}(h) \setminus B(\{x_i\}_{i=1}^{s}, \varepsilon_g)\,.$$

Because $\varepsilon_g \geq \delta_g$, this implies

$$\mathcal{G}(x_{1:s}) \in \mathrm{supp}(h) \setminus B(\{x_i\}_{i=1}^{s}, \delta_g)\,.$$

This completes the proof. □

Proof of Theorem 4.10. We write $B_1(x, \varepsilon)$ and $B_2(x, \varepsilon)$ for the ε-balls centered at x with respect to ρ_1 and ρ_2, respectively. First observe that for any $x \in \mathcal{X}$ and $\gamma > 0$,

$$B_1\Big(x, \frac{\gamma}{M}\Big) \subseteq B_2(x, \gamma).$$

To prove (r/M)-UUS under ρ_1, fix $h \in \mathcal{H}$ and a sequence $(x_i)_{i \geq 1}$ in its support satisfying

$$\mathrm{supp}(h) \subseteq B_1(\{x_i : i \geq 1\}, r/M).$$

Taking infima in $\rho_2 \leq M\rho_1$ gives

$$\mathrm{supp}(h) \subseteq B_2(\{x_i : i \geq 1\}, r).$$

By r-UUS under ρ_2,

$$N(r; \{x_1, \ldots, x_t\}, \rho_2) \longrightarrow \infty.$$

Every ρ_1-cover at radius r/M is a ρ_2-cover at radius r, so

$$N(r/M; \{x_1, \ldots, x_t\}, \rho_1) \geq N(r; \{x_1, \ldots, x_t\}, \rho_2) \longrightarrow \infty.$$

The supports are nonempty by assumption. Thus $\mathcal{H}$ satisfies (r/M)-UUS under ρ_1. Now suppose $\mathcal{G}$ is an $(\varepsilon_a, \varepsilon_g)$-generator for $\mathcal{H}$ with respect to ρ_2. We show that $\mathcal{G}$ is also an $(\varepsilon_a/M, \varepsilon_g/M)$-generator with respect to ρ_1.

Let $\{x_i\}_{i=1}^{\infty} \subseteq \mathrm{supp}(h)$ be a sequence such that

$$\mathrm{supp}(h) \subseteq B_1(\{x_i\}_{i=1}^{\infty}, \varepsilon_a/M)\,.$$

For every nonempty $A \subseteq \mathcal{X}$ and $x \in \mathcal{X}$, taking infima in the metric inequality gives

$$\inf_{a\in A} \rho_2(x,a) \le M \inf_{a\in A} \rho_1(x,a).$$

Thus $B_1(A, \varepsilon_a/M) \subseteq B_2(A, \varepsilon_a)$ under the neighborhood definition used here, so the coverage assumption implies

$$\mathrm{supp}(h) \subseteq B_2(\{x_i\}_{i=1}^{\infty}, \varepsilon_a)\,.$$

Therefore, there exists s such that for all $t \ge s$,

$$\mathcal{G}(x_{1:t}) \in \mathrm{supp}(h) \setminus B_2\big(\{x_i\}_{i=1}^{t}, \varepsilon_g\big)\,.$$

Since

$$B_1(x, \varepsilon_g/M) \subseteq B_2(x, \varepsilon_g),$$

it follows that

$$\mathcal{G}(x_{1:t}) \in \mathrm{supp}(h) \setminus B_1\big(\{x_i\}_{i=1}^{t}, \varepsilon_g/M\big)\,.$$

This completes the proof. □

E Remarks on Computability

There are few remarks on the algorithms on generation in the limit in the countable example space setting [13, 21, 6]. We point out that the algorithm by [13] for generation in the limit which only uses the oracle "is $x \in \mathrm{supp}(h)$" is not directly extendable to the metric space setting. Because it explicitly uses the countable structure of the example space, as well as the countable structure of the hypothesis class. We leave the extension as an open problem. The ERM-based closure-membership test in Appendix I of [21] does extend to the metric setting. Finding a novel closure point requires an additional search mechanism, as discussed below.

For a generatable sample $x_1, \ldots, x_n$, the closure $\langle x_1, \ldots, x_n \rangle_{\mathcal{H}}$ consists of the positive examples common to all hypotheses consistent with the sample. An Empirical Risk Minimization (ERM) oracle tests membership in this closure. Formally, let

$$\mathcal{O}_{\mathrm{ERM}} : \big(2^{\{0,1\}^{\mathcal{X}}} \setminus \{\emptyset\}\big) \times (\mathcal{X} \times \{0,1\})^{\star} \longrightarrow \mathbb{N} \cup \{0\}$$

return $\min_{h\in\mathcal{H}} \sum_{(x,y)\in S} \mathbb{1}\{h(x) \neq y\}$ on a nonempty class $\mathcal{H}$ and a finite labeled sample S. Given a candidate x, query this oracle on the labeled sample $S_x = ((x_1,1), \ldots, (x_n,1), (x,0))$. Since the positive sample is realizable, the answer r_x is at most 1. It equals 0 precisely when some hypothesis contains all observations but excludes x. Hence

$$r_x = \mathbb{1}\{x \in \langle x_1, \ldots, x_n \rangle_{\mathcal{H}}\}.$$

The uniform guarantee in Lemma B.2 can be implemented using one call per round to a max–min oracle. Fix $\varepsilon_g > 0$, a nonempty instance space $\mathcal{X}$, and a default point $x_0 \in \mathcal{X}$. Define

$$\mathcal{O}_{\textit{max-min}} : \big(2^{\{0,1\}^{\mathcal{X}}} \setminus \{\emptyset\}\big) \times \mathcal{X}^{\star} \to \mathcal{X}.$$

Given a nonempty hypothesis class $\mathcal{H}$ and a finite sample $x_1, \ldots, x_t$, the oracle returns a chosen element of

$$\operatorname*{arg\,max}_{x\in\mathcal{X}\setminus B(\{x_i\}_{i=1}^{t}, \varepsilon_g)} \min_{h\in\mathcal{H}} \left(\sum_{i=1}^{t} \mathbb{1}\{h(x_i) \neq 1\} + \mathbb{1}\{h(x) \neq 0\} \right).$$

If the feasible set is empty, the oracle returns x_0. For a nonempty feasible set, a maximizer exists because the objective takes values in the finite set $\{0, \ldots, t+1\}$. On the empty sample, take the novelty exclusion set to be empty. These conventions define the oracle on its full domain.

The inner minimization is the ERM value defined above. Thus, whenever the feasible set is nonempty,

$$\mathcal{O}_{\textit{max-min}}(\mathcal{H}, x_{1:t}) \in \operatorname*{arg\,max}_{x\in\mathcal{X}\setminus B(\{x_i\}_{i=1}^{t}, \varepsilon_g)} \mathcal{O}_{\mathrm{ERM}}\big(\mathcal{H}, \{(x_1,1), \ldots, (x_t,1), (x,0)\}\big).$$

Algorithm 1 Conditional search over a candidate enumeration using ERM

Input: A generatable sample $x_1, \dots, x_t$ and candidate enumeration $(y_i)_{i \geq 1}$

for $i = 1, 2, 3, \dots$ **do**
  if $y_i \notin B(\{x_j\}_{j=1}^t, \varepsilon_g)$ **then**
    if $\mathcal{O}_{\mathrm{ERM}}\big(\mathcal{H}, \{(x_1, 1), \dots, (x_t, 1), (y_i, 0)\}\big) \geq 1$ **then**
      return y_i
    end
  end
end

Returning the oracle's output at each round gives the uniform guarantee. To see this, fix a nonempty class as in Lemma B.2, a generatable sample, and suppose that $N(\varepsilon_a; \{x_i\}_{i=1}^t, \rho) \geq \mathrm{C}_{\varepsilon_a, \varepsilon_g}(\mathcal{H}) + 1$. The closure then has no finite ε_g-cover with centers in the closure. Since every observed point belongs to the closure, it follows that

$$\langle x_1, \dots, x_t \rangle_{\mathcal{H}} \setminus B(\{x_i\}_{i=1}^t, \varepsilon_g) \neq \emptyset,$$

and moreover, for every $x \in \langle x_1, \dots, x_t \rangle_{\mathcal{H}} \setminus B(\{x_i\}_{i=1}^t, \varepsilon_g)$,

$$\mathcal{O}_{\mathrm{ERM}}\big(\mathcal{H}, \{(x_1, 1), \dots, (x_t, 1), (x, 0)\}\big) \geq 1.$$

As a result,

$$\max_{x \in \mathcal{X} \setminus B(\{x_i\}_{i=1}^t, \varepsilon_g)} \mathcal{O}_{\mathrm{ERM}}\big(\mathcal{H}, \{(x_1, 1), \dots, (x_t, 1), (x, 0)\}\big) \geq 1,$$

and thus the oracle $\mathcal{O}_{\textit{max-min}}(\mathcal{H}, x_{1:t})$ returns a point $\hat{x}_t \in \mathcal{X} \setminus B(\{x_i\}_{i=1}^t, \varepsilon_g)$ such that

$$\mathcal{O}_{\mathrm{ERM}}\big(\mathcal{H}, \{(x_1, 1), \dots, (x_t, 1), (\hat{x}_t, 0)\}\big) \geq 1.$$

Such a point $\hat{x}_t$ must therefore belong to $\langle x_1, \dots, x_t \rangle_{\mathcal{H}}$, completing the proof.

The max–min oracle supplies a search operation beyond the ERM membership test; the preceding argument does not implement it using finitely many ERM queries. For an ERM-only search, suppose a candidate enumeration $D = \{y_i : i \geq 1\}$ is supplied, together with exact tests of whether $y_i \notin B(\{x_j\}_{j=1}^t, \varepsilon_g)$. On a generatable sample, the following procedure returns a novel closure point whenever

$$D \cap \big(\langle x_1, \dots, x_t \rangle_{\mathcal{H}} \setminus B(\{x_j\}_{j=1}^t, \varepsilon_g)\big) \neq \emptyset.$$

Under the displayed intersection condition, a successful candidate occurs at a finite index, so the procedure terminates after finitely many tests and ERM calls. Above the covering-number threshold in Lemma B.2, the novel closure is nonempty. A sufficient additional condition for the displayed intersection is that $D \cap C_t$ be dense in $C_t = \langle x_1, \dots, x_t \rangle_{\mathcal{H}}$, since $C_t \setminus B(\{x_j\}_{j=1}^t, \varepsilon_g)$ is relatively open in C_t. Density of D in $\mathcal{X}$ alone is insufficient: a dense subset can miss the entire closure. Neither the required intersection property nor a supplied candidate enumeration with that property follows from separability alone.

This conditional search need not terminate on other samples, including samples below the threshold for which no novel closure point exists. It therefore does not by itself define a generator on every finite input. A complete implementation would also need a terminating fallback rule; for example, a decidable threshold test would allow an arbitrary output below the threshold if the candidate-intersection condition were guaranteed above it. We do not assume these additional capabilities here. In general, a finite-ERM implementation per round remains unresolved by this argument, including the search for a novel closure point above the threshold.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction state the paper's main contributions : extending generation to separable metric spaces, characterizing uniform and non-uniform generatability, identifying stability of generation in doubling spaces and instability results in general metrc spaces.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: The paper discusses limitations and open directions in the Discussion section, including the need to study generation under natural symmetries, geometric conditions beyond doubling, and additional structural assumptions. The results are theoretical and depend on the stated metric-space and hypothesis-class assumptions.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: All main results state their assumptions explicitly, including separability, the r-UUS condition, novelty parameters, doubling assumptions, and metric-equivalence assumptions. Proofs are provided in the appendices, with the main paper giving statements and intuition for the central results.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [N/A]

Justification: The paper is theoretical and does not include empirical experiments.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some

way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [N/A]

   Justification: The paper does not report experimental results requiring code.

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [N/A]

   Justification: The paper does not report experimental results.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [N/A]

   Justification: The paper does not report experimental results.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
   - The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).

- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [N/A]

   Justification: The paper does not include computational experiments.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The research is theoretical and does not involve data collection, human subjects, deployment of models, or other procedures that would raise apparent conflicts with the NeurIPS Code of Ethics.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: The paper is foundational theoretical work on generative models and novelty in metric spaces. It may help clarify when generation is possible or unstable, but it does not introduce a deployable system, dataset, or model; any societal impacts are indirect.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.

- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: The paper does not release models, datasets, or other artifacts with a high risk of misuse.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [N/A]

    Justification: The paper does not use existing datasets, codebases, pretrained models, or other assets. Prior work is cited in the bibliography.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated

licenses for some datasets. Their licensing guide can help determine the license of a dataset.

- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: The paper does not introduce or release new datasets, code, pretrained models, or benchmarks.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing or research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The paper does not involve human subjects, so IRB approval is not applicable.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.

- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: The core methodology and theoretical results do not use LLMs as an important, original, or non-standard component.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.