Contrastive Identification and Generation in the Limit

Xiaoyu Li
University of New South Wales
xiaoyu.li2@unsw.edu.au

Andi Han
University of Sydney
andi.han@sydney.edu.au

Jiaojiao Jiang
University of New South Wales
jiaojiao.jiang@unsw.edu.au

Junbin Gao
University of Sydney
junbin.gao@sydney.edu.au

Abstract

In the classical *identification in the limit* model of Gold [Inf. Control 1967], a stream of positive examples is presented round by round, and the learner must eventually recover the target hypothesis. Recently, Kleinberg and Mullainathan [NeurIPS 2024] introduced *generation in the limit*, where the learner instead must eventually output novel elements of the target's support. Both lines of work focus on positive-only or fully labeled data. Yet many natural supervision signals are relational rather than singleton, and some reveal only that two examples *differ* without indicating which is positive: cannot-link constraints in constrained clustering, "not the same class" annotations, and dissimilarity feedback all produce such direction-hidden pairs. This motivates us to initiate the learning-theoretic study of *contrastive identification and generation in the limit*, where the learner observes a *contrastive presentation* of data: a stream of unordered pairs $\{x, y\}$ satisfying $h(x) \neq h(y)$ for an unknown target binary hypothesis h, but which element is positive is hidden from the learner. We first present three results in the noiseless setting: an exact characterization of contrastive identifiable classes (a one-line geometric refinement of Angluin's tell-tale condition [Angluin, Inf. Control 1980]), a combinatorial dimension called the *contrastive closure dimension* (a contrastive analogue of the closure dimension in Li et al. [COLT 2025]) that exactly characterizes uniform contrastive generation with tight sample complexity, and a strict hierarchy in which contrastive generation and text identification are mutually incomparable. We then prove a sharp *reversal* under finite adversarial corruption: there exist classes identifiable from contrastive pairs under any finite corruption budget by a single budget-independent algorithm, yet not identifiable from positive examples under even one corrupted observation. More generally, we characterize contrastive identification exactly at every known corruption budget and at unknown finite budgets, with the clean characterization as the budget-zero case. The unifying technical object is the *common crossing graph*, which encodes pairwise ambiguity, family-level generation obstructions, and corruption defects in a single coverage-and-incidence language.

1 Introduction

Identification in the limit, the foundational model introduced by Gold [1967], asks how a learner can recover an unknown target hypothesis h drawn from a known class $\mathcal{H}$ by observing an infinite stream of examples and stabilizing its guesses on the truth. With a fully labeled stream (an *informant*), every countable class is identifiable; with only positive examples (a *text*), even simple

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

classes become unlearnable. Angluin [1980]'s celebrated tell-tale theorem characterized exactly which classes are identifiable from positive data: each hypothesis must be distinguished from its proper sub-hypotheses by a finite "tell-tale" subset of positives, a structural condition that has anchored inductive inference for four decades [Lange et al., 2008].

Recently, Kleinberg and Mullainathan [2024] initiated the parallel study of *generation* in the limit: instead of naming the target, the learner must eventually output novel positives of h not yet seen, and on countable classes with infinite supports they showed this is always possible from a text, in stark contrast to identification. Li et al. [2025] reformulated the model in learning-theoretic notation and sharpened the picture by introducing a combinatorial *closure dimension* that exactly characterizes *uniform* generation, where the learner must succeed within a bounded number of rounds.

A common thread runs through this body of work: the learner observes a stream of *positive examples* of the unknown target. But many natural data sources are inherently *relational* rather than singleton, and a basic subclass reveals only that two examples *differ*, with no indication of which is positive: cannot-link constraints in constrained clustering [Wagstaff et al., 2001], "different class" crowd annotations, and dissimilarity feedback [Dan et al., 2021]. (Directed comparisons, such as A/B outcomes, reveal the orientation and fall outside this model.) We study the limit-learning abstraction of this signal, an adversarial stream of such pairs that eventually covers every positive; it is an idealization in the tradition of Gold, not a model of a deployed generation pipeline. We keep the goals of Gold [1967] and Kleinberg and Mullainathan [2024], including the output of novel *positive* points, and hide only which endpoint of each observed pair is positive, which isolates the value of the missing orientation bit. This raises a basic question: what can a learner accomplish in the limit when its only information is that two examples disagree under the target, with no indication of which is positive?

Contrastive identification and generation in the limit. We introduce *contrastive identification and generation in the limit*[1], where the unknown target $h : \mathcal{X} \to \{0, 1\}$ is a binary hypothesis with positive set (or *support*) $\mathrm{supp}(h) := \{x \in \mathcal{X} : h(x) = 1\}$, and the learner observes a *contrastive presentation* of data: at each round, an unordered pair $\{x, y\}$ of examples that disagree under h, i.e., $h(x) \neq h(y)$, but with no information about which endpoint is positive. This setting is different from Gold's text and informant models since each pair carries an XOR constraint between its endpoints, yet hides the labels themselves. The orientation cannot always be recovered: if h and its complement $\mathbf{1} - h$ both lie in the class, neither task is possible. But the rest of the class may rule one orientation out, and a generator may certify a new positive before orienting any pair: if every hypothesis has a single negative point, then after the pair $\{s, z\}$ every other point is positive under both remaining targets.

It is natural to read a contrastive presentation *geometrically*: take the example space as the vertex set of a graph and each observed pair as an edge, so that the disagreement condition forces every edge to cross the unknown cut $(\mathrm{supp}(h), \mathcal{X} \setminus \mathrm{supp}(h))$ separating positives from negatives. The central technical object that emerges from this lens is the *common crossing graph* of two hypotheses: the pairs that cross both hypotheses' cuts simultaneously, hence look like valid observations under either as the target. This graph captures pairwise ambiguity, family-level generation obstructions, and corruption defects in a single coverage-and-incidence language.

1.1 Our main results

Throughout the paper we work with countable classes of proper nontrivial hypotheses over a countably infinite example space and write CtrId, CtrGen for the families admitting an identifier or a generator in the limit from contrastive presentations, and TxtId, TxtGen for the corresponding text-stream families. Formal definitions are deferred to Section 3; informally, an identifier must eventually output h forever, and a generator must eventually output only unseen positives of h, not pairs.

(1) Exact characterization of contrastive identification. We give an exact combinatorial condition for when a hypothesis class admits an identifier in the limit from contrastive presentations: it must be text-identifiable in the sense of Angluin's tell-tale theorem, and additionally every two

[1] We use "contrastive" in the pair-level data sense of learning from similar and dissimilar pairs [Bao et al., 2018, Dan et al., 2021], *not* in the self-supervised representation-learning sense (SimCLR, InfoNCE).

incomparable hypotheses must form an *overlapping cover*, meaning their supports intersect and together cover the entire example space (Theorem 4.7). The overlapping-cover requirement is the only obstruction that contrastive data introduce beyond positive-only text data.

(2) A combinatorial dimension for contrastive generation. We introduce a closure-style combinatorial dimension that exactly characterizes which hypothesis classes admit a uniform contrastive generator: the edge-set analogue of the closure dimension of Li et al. [2025] for positive-data generation, measuring how long pair constraints can keep an adversary from forcing a novel target element. Finiteness of this dimension is both necessary and sufficient, with tight distinct-edge sample complexity equal to the dimension plus one when the dimension is positive (Theorem 5.4). A non-uniform variant follows by chain decomposition (Theorem 5.5). For ordinary generation on finite classes we prove that generation is possible if and only if every family that admits a shared presentation has infinite common support (Theorem 5.15); on countable classes this condition is necessary but not sufficient, and a characterization remains open.

(3) The clean diamond hierarchy. On countable classes with infinite supports, the four families $\mathsf{CtrId}, \mathsf{TxtId}, \mathsf{CtrGen}, \mathsf{TxtGen}$ form a strict *diamond* rather than a chain (Figure 1): CtrId is strictly smaller than both CtrGen and TxtId, both of which are strictly smaller than TxtGen, but CtrGen and TxtId are mutually incomparable (Theorem 5.13, Theorem 5.14).

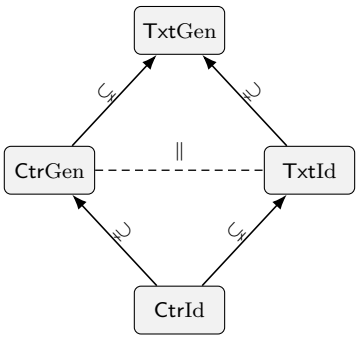


Figure 1: Diamond hierarchy on countable classes with infinite supports. Arrows denote strict containment ($\subsetneq$), oriented from smaller to larger; the dashed line denotes incomparability ($\|$).

(4) Robustness at every corruption budget. The inclusion $\mathsf{CtrId} \subsetneq \mathsf{TxtId}$ fails under finite adversarial corruption, where an adversary plants a bounded number of arbitrary observations into the stream. We isolate the *defect number*, an invariant counting the minimum forced wrong-cut violations in any clean contrastive presentation (Proposition 6.3). An infinite defect number $\kappa(h \to g)$ forces unboundedly many violations of g, which no finite budget can mask; in fact h and g become confusable at budget k exactly when $\max\{\kappa(h \to g), \kappa(g \to h)\} \leq 2k$ (Theorem 6.9). For whole classes we characterize identification at every known budget and at unknown finite budgets (Theorem 6.10); $k = 0$ recovers (1), and for unknown budgets the condition is that the *holes* $\mathcal{X} \setminus \mathrm{supp}(h)$ are pairwise disjoint and any two leave infinitely many common positives (Corollary 6.11). The *co-singleton class* (each hypothesis labels every example positive except a single "hole") is the sharpest instance: it is identifiable under any finite corruption budget by a single budget-independent *absence-count* algorithm, yet fails text identification at one corruption (Theorems 6.5 and 6.6). The construction generalizes to mutual incomparability of k-corrupted contrastive and text identification for every $k \geq 1$ (Theorem 6.8); the budgets form a strict hierarchy (Corollary 6.12), and a budget that grows without bound, however slowly, makes any two hypotheses confusable (Proposition C.6).

1.2 Technical overview

Our results are driven by a single piece of geometry, the *common crossing graph* $\Gamma(h, g) := \Delta(h) \cap \Delta(g)$, where $\Delta(h) := \{\{x, y\} \in [\mathcal{X}]^2 : h(x) \neq h(y)\}$ is the set of pairs crossing h's cut and $\mathrm{V}(E)$ is the vertex set of an edge set E.

Identification: coverage on Γ. Pairwise eliminability has a one-line graph-theoretic translation: g is *not* eliminable from h iff $\mathrm{supp}(h) \subseteq \mathrm{V}(\Gamma(h, g))$ (Proposition 4.2). Case analysis on the four-region partition of $\mathcal{X}$ by membership in $\mathrm{supp}(h), \mathrm{supp}(g)$ identifies three non-eliminable regimes (superset, disjoint, non-covering). The last two are exactly the *new* obstructions contrastive data introduce beyond text, driving the exact characterization (Theorem 4.7).

Generation: edge-induced closure replaces positive-data closure. Contrastive data confirm only the XOR condition $h(x) \neq h(y)$ on each observed pair, so we replace the positive version space by the edge-induced version space $\mathcal{H}_\Delta(E) := \{g \in \mathcal{H} : E \subseteq \Delta(g)\}$, the hypotheses under which every observed pair crosses the cut, and the closure by $\langle E \rangle^{\Delta}_{\mathcal{H}}$ (Definition 5.1).

Robustness: defects as forced cut violations. Corrupted contrastive data are k-close to a clean crossing-edge stream, whose pairs all cross one cut, while a clean text has no such internal structure. The positive-side defect set $D^+_{h\to g} := \mathrm{supp}(h) \setminus \mathrm{V}(\Gamma(h,g))$ lower-bounds the number of pairs violating g's cut in any clean valid presentation of h (Proposition 6.3). Charging each defect to the first pair that covers it shows that k corrupted pairs cancel at most k of these violations, so $\kappa(h \to g) := |D^+_{h\to g}| \geq 2k+1$ rules g out at budget k (Theorems 6.9 and 6.10), and $\kappa(h \to g) = \infty$ rules g out at every known finite budget. Full proofs, discussion, and additional related work are deferred to the appendix.

2 Related Work

Identification in the limit. The paradigm originates with Gold [1967], with Angluin [1980]'s tell-tale theorem giving the canonical positive-data characterization; see Lange et al. [2008] for indexed-family results. Robust and noise-tolerant variants go back to the approximate identification of Wharton [1974] (Section F.1).

Generation in the limit. Kleinberg and Mullainathan [2024] introduced generation in the limit and proved its universality on countable UUS classes. Li et al. [2025] reformulated the model in learning-theoretic notation and introduced the closure dimension that exactly characterizes uniform generation. Subsequent work has examined breadth, density, noise trade-offs, hallucination detection, generation in continuous spaces, agnostic generation, safe generation, differentially private generation, and union closure properties [Kalavasis et al., 2025, Raman and Raman, 2025, Bai et al., 2026, Kleinberg and Wei, 2025, Charikar and Pabbaraju, 2025, Karbasi et al., 2025, Li et al., 2026a, Høgsgaard and Pabbaraju, 2026, Anastasopoulos et al., 2026, Li et al., 2026b, Mehrotra et al., 2026b, Hanneke et al., 2025].

Learning from pair signals. The data model studied here, pairs known to disagree but with the labels stripped, is a special case of pair-level supervision studied in a long line of weakly supervised learning, including learning from similar and dissimilar pairs [Bao et al., 2018, Dan et al., 2021], constrained clustering with cannot-link constraints [Wagstaff et al., 2001], learning from positive and unlabeled data [Elkan and Noto, 2008, Bekker and Davis, 2020], complementary-label learning [Ishida et al., 2017], learning from label proportions [Quadrianto et al., 2009], and multiple-instance learning with the exactly-one-positive constraint [Carbonneau et al., 2018]. A/B outcomes, contrastive examples [Mansouri et al., 2025, Razavi Ravari et al., 2026], and comparison queries [Kane et al., 2017] also use two-point signals, but reveal the orientation, and the last two study query complexity; here it is hidden. Those literatures study finite-sample PAC and risk minimization with stochastic noise; we study the asymptotic identification/generation-in-the-limit regime, in which structural combinatorics (rather than concentration of measure) governs feasibility, and adversarial corruption rather than i.i.d. label noise drives the robustness analysis.

3 Preliminaries

Setting. We adopt the learning-theoretic notations from Li et al. [2025]. Throughout, $\mathcal{X}$ is a countably infinite example space, $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ is a countable class, and $\mathbb{N} = \{0,1,2,\ldots\}$. The support of $h \in \mathcal{H}$ is $\mathrm{supp}(h) := \{x \in \mathcal{X} : h(x) = 1\}$. A hypothesis is *proper nontrivial* if $\varnothing \subsetneq \mathrm{supp}(h) \subsetneq \mathcal{X}$. We work extensionally; $\mathcal{X} = \{u_0, u_1, \ldots\}$ is fixed once for "least element" constructions. We restrict to proper nontrivial targets for contrastive presentation (otherwise no XOR pair exists), and to the standard infiniteness condition for generation:

Definition 3.1 (Uniformly unbounded support [Li et al., 2025])**.** $\mathcal{H}$ satisfies the *uniformly unbounded support* (UUS) property if $|\mathrm{supp}(h)| = \infty$ for every $h \in \mathcal{H}$.

Definition 3.2 (Positive-data closure and version spaces [Li et al., 2025])**.** For $x_{1:n} = (x_1, \ldots, x_n)$, we define the version space $\mathcal{H}(x_{1:n}) := \{g \in \mathcal{H} : \{x_1, \ldots, x_n\} \subseteq \mathrm{supp}(g)\}$ and the positive-data closure $\langle x_{1:n}\rangle_{\mathcal{H}} := \bigcap_{g \in \mathcal{H}(x_{1:n})} \mathrm{supp}(g)$ when nonempty (else $\perp$).

Definition 3.3 (Presentation modes)**.** Let $[\mathcal{X}]^2$ be the set of two-element subsets of $\mathcal{X}$, and let $h \in \mathcal{H}$. A *text presentation* $T = (x_t)_{t\geq 1}$ for h has $x_t \in \mathrm{supp}(h)$ and covers $\mathrm{supp}(h)$; write $\mathrm{Seen}_n(T) := \{x_t\}_{t\leq n}$. An *informant presentation* $J = ((x_t, h(x_t)))_{t\geq 1}$ for h covers all of $\mathcal{X}$;

write $\mathrm{Seen}_n(J) := \{x_t\}_{t\le n}$. A *contrastive presentation* $P = (p_t)_{t\ge 1} \subseteq [\mathcal{X}]^2$ for proper nontrivial h satisfies (i) $\sum_{x\in p_t} h(x) = 1$ (XOR condition), and (ii) $\mathrm{supp}(h) \subseteq \bigcup_t p_t$ (positive-side coverage); pairs are observed as unordered sets and may repeat, and we write $\mathrm{Seen}_n(P) := \bigcup_{t\le n} p_t$ and $\mathrm{Seen}(P) := \bigcup_{t\ge 1} p_t$.

For $\mathsf{X} \in \{\mathsf{Txt}, \mathsf{Inf}, \mathsf{Ctr}\}$ and any presentation Q, let $Q_{\le n}$ denote the length-n prefix.

Definition 3.4 (Identification and generation in the limit [Gold, 1967, Kleinberg and Mullainathan, 2024])**.** An X-*identifier* I is a map from valid X-prefixes to $\mathcal{H}$; it *identifies* h if for every valid presentation Q of h there exists N with $I(Q_{\le n}) = h$ for all $n \ge N$. An X-*generator* G is a map from valid X-prefixes to $\mathcal{X}$; it *generates* h if for every valid presentation Q of h there exists N such that its outputs $\hat{x}_n = G(Q_{\le n})$ satisfy $\hat{x}_n \notin \mathrm{Seen}_n(Q)$ and $\hat{x}_n \in \mathrm{supp}(h)$ for all $n \ge N$. We write $\mathcal{H} \in \mathsf{X}\mathrm{Id}$ (resp. $\mathsf{X}\mathrm{Gen}$) when some identifier (resp. generator) succeeds on every $h \in \mathcal{H}$.

Contrastive generation is not Kleinberg and Mullainathan [2024]'s generation over the lifted sets $\Delta(h) \subseteq [\mathcal{X}]^2$: there the output is a new pair, which names no positive point, and the input must enumerate all of $\Delta(h)$, whereas ours need only cover $\mathrm{supp}(h)$. For disjoint infinite supports A, B the lifted class $\{\Delta(h_A), \Delta(h_B)\}$ is text-generatable (Theorem 3.6), yet $\{h_A, h_B\} \notin \mathsf{Ctr}\mathrm{Gen}$ (Proposition 5.12).

Theorem 3.5 (Angluin's tell-tale theorem [Angluin, 1980])**.** *$\mathcal{H} \in \mathsf{Txt}\mathrm{Id}$ iff for every $g \in \mathcal{H}$ there is a finite "tell-tale" set $T_g \subseteq \mathrm{supp}(g)$ such that no $f \in \mathcal{H}$ with $\mathrm{supp}(f) \subsetneq \mathrm{supp}(g)$ contains T_g.*

Theorem 3.6 (Universality of text generation [Kleinberg and Mullainathan, 2024])**.** *Every countable $\mathcal{H} \subseteq \{0,1\}^{\mathcal{X}}$ satisfying UUS lies in $\mathsf{Txt}\mathrm{Gen}$.*

4 Identification: Eliminability via Common Crossings

This section proves the exact characterization of $\mathsf{Ctr}\mathrm{Id}$ advertised in Section 1.1.

Notation. For $h \in \mathcal{H}$, the *crossing-edge set* is $\Delta(h) := \{\{x, y\} \in [\mathcal{X}]^2 : h(x) \ne h(y)\}$; a contrastive pair for h is precisely an edge in $\Delta(h)$. For two hypotheses h, g, the *common crossing graph* is $\Gamma(h, g) := \Delta(h) \cap \Delta(g)$, viewed as an undirected graph on $\mathcal{X}$ (Figure 2). For an edge set E, $\mathrm{V}(E)$ denotes the vertices incident to some edge of E.

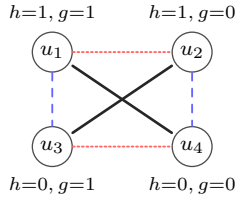


Figure 2: Common crossing graph on a four-vertex example with $\mathrm{supp}(h) = \{u_1, u_2\}$, $\mathrm{supp}(g) = \{u_1, u_3\}$. *Dashed blue*: $\Delta(h)$ only; *dotted red*: $\Delta(g)$ only; *solid dark*: $\Gamma(h, g)$.

4.1 Pairwise eliminability as common-crossing coverage

Definition 4.1 (Pairwise eliminability)**.** For distinct proper nontrivial h, g, we say g is *eliminable from* h if every valid contrastive presentation for h contains a pair outside $\Delta(g)$; equivalently, g is *not* eliminable from h iff there is a valid presentation for h with all pairs in $\Gamma(h, g)$.

Proposition 4.2 (Common-crossing coverage)**.** *For distinct proper nontrivial h, g, g is not eliminable from h iff $\mathrm{supp}(h) \subseteq \mathrm{V}(\Gamma(h, g))$.*

Proof. ($\Rightarrow$) Positive-side coverage forces every $x \in \mathrm{supp}(h)$ to appear in a common-crossing pair, so $\mathrm{supp}(h) \subseteq \mathrm{V}(\Gamma(h, g))$. ($\Leftarrow$) For each $x \in \mathrm{supp}(h)$, pick a partner y_x with $\{x, y_x\} \in \Gamma(h, g)$; enumerate $\mathrm{supp}(h)$ (or list-and-repeat if finite) and emit the corresponding pairs to obtain a valid presentation for h with all pairs in $\Gamma(h, g)$. □

Theorem 4.3 (Eliminability geometry)**.** *For distinct proper nontrivial h, g, partition $\mathcal{X}$ by membership: $A = \mathrm{supp}(h) \cap \mathrm{supp}(g)$, $B = \mathrm{supp}(h) \setminus \mathrm{supp}(g)$, $C = \mathrm{supp}(g) \setminus \mathrm{supp}(h)$, $D = \mathcal{X} \setminus (\mathrm{supp}(h) \cup \mathrm{supp}(g))$. Then*

$$g \text{ is not eliminable from } h \iff (A \ne \varnothing \Rightarrow D \ne \varnothing) \wedge (B \ne \varnothing \Rightarrow C \ne \varnothing).$$

Equivalently, g is not eliminable from h in exactly three regimes: (N1) $\mathrm{supp}(h) \subsetneq \mathrm{supp}(g)$ *(*superset*);* (N2) *incomparable and disjoint (*disjoint*);* (N3) *incomparable, intersecting, and non-covering (*non-covering*).*

Proof sketch. For $x \in A$, a pair $\{x, y\}$ lies in $\Delta(h)$ iff $y \notin \mathrm{supp}(h)$, and additionally in $\Delta(g)$ iff $y \notin \mathrm{supp}(g)$; hence A-vertices are coverable iff $D \neq \varnothing$. Symmetrically for B. Full proofs of this result and the next lemma are in Appendix A. □

Lemma 4.4 (Pairwise shared presentation)**.** *$h, g \in \mathcal{H}$ admit a common contrastive presentation valid for both targets iff* $\mathrm{supp}(h) \cup \mathrm{supp}(g) \subseteq \mathrm{V}(\Gamma(h, g))$. *Mutual non-eliminability implies this confusability.*

4.2 Exact characterization of CtrId

Definition 4.5 (Overlapping cover)**.** h, g with incomparable supports form an *overlapping cover* if $\mathrm{supp}(h) \cap \mathrm{supp}(g) \neq \varnothing$ and $\mathrm{supp}(h) \cup \mathrm{supp}(g) = \mathcal{X}$.

Lemma 4.6 ($\mathsf{CtrId} \subseteq \mathsf{TxtId}$)**.** *On classes of proper nontrivial hypotheses,* $\mathsf{CtrId} \subseteq \mathsf{TxtId}$*, with strict inclusion even on UUS classes.*

Proof sketch. A text identifier simulates a contrastive identifier I by feeding it the synthetic prefix $(\{x_t, z_n\})_{t \leq n}$, where z_n is the least unseen example. For target h, z_n eventually stabilizes at $z^* = \min(\mathcal{X} \setminus \mathrm{supp}(h))$; from that stage onward the synthetic prefix is the prefix of a single fixed valid contrastive presentation $(\{x_t, z^*\})_{t \geq 1}$ for h, on which I converges. The full proof is in Appendix A. □

Theorem 4.7 (Exact characterization of CtrId)**.** *For a countable class $\mathcal{H}$ of proper nontrivial hypotheses, the following are equivalent:* (i) $\mathcal{H} \in \mathsf{CtrId}$*;* (ii) $\mathcal{H} \in \mathsf{TxtId}$ *and the contrastive non-eliminability relation is contained in the positive-data superset relation;* (iii) $\mathcal{H} \in \mathsf{TxtId}$ *and every incomparable pair in $\mathcal{H}$ is an overlapping cover.*

Proof sketch. (ii) $\iff$ (iii) is immediate from Theorem 4.3: among contrastive non-eliminability relations, those not in the superset relation are exactly the disjoint and non-covering incomparable barriers, which are excluded precisely by the overlapping cover condition. (i)⇒(ii) uses Lemma 4.4. (iii)⇒(i): by Theorem 3.5, build an enumerator that outputs the least *eligible* hypothesis (consistent and T_{h_i}-seen). The crucial case $\mathrm{supp}(h) \subsetneq \mathrm{supp}(h_j)$ uses the XOR pair structure: any $t \in T_{h_j} \setminus \mathrm{supp}(h)$ appears in a pair whose other endpoint must lie in $\mathrm{supp}(h) \subsetneq \mathrm{supp}(h_j)$, so both endpoints lie in $\mathrm{supp}(h_j)$, contradicting consistency. Full proof in Appendix A. □

Remark 4.8. The (iii)⇒(i) construction uses the tell-tale family $\{T_g\}_{g \in \mathcal{H}}$ from Theorem 3.5, so the result is information-theoretic. Theorem 6.10 extends it to every corruption budget.

5 Generation: Closure, Cores, and Confusability

Generation asks for eventually-correct novel outputs rather than recovery, and it admits a two-layer theory: a uniform layer governed exactly by a closure dimension, and an ordinary layer with sufficient conditions from safe/eventual cores and obstructions from a confusability complex. We treat the uniform layer first because of its exact characterization, then return to the ordinary layer and the diamond hierarchy. Omitted proofs are in Appendix B.

5.1 Edge-induced closure and the contrastive closure dimension

The proof of Theorem 3.6 relies on confirmed positives. Contrastive data confirm only the XOR condition $h(x) \neq h(y)$ on each observed pair, so we replace the version space $\mathcal{H}(x_{1:n})$ by an edge-induced version space and develop a closure operator over edge sets.

Definition 5.1 (Edge-induced closure)**.** For finite $E \subseteq [\mathcal{X}]^2$, let $\mathcal{H}_\Delta(E) := \{g \in \mathcal{H} : E \subseteq \Delta(g)\}$. The *contrastive closure* is $\langle E \rangle_{\mathcal{H}}^{\Delta} := \bigcap_{g \in \mathcal{H}_\Delta(E)} \mathrm{supp}(g)$ when $\mathcal{H}_\Delta(E) \neq \varnothing$, and $\perp$ otherwise. For a contrastive prefix $P_{\leq n}$, $E_n(P) := \{p_t : t \leq n\}$ is the set of distinct observed pairs and $\mathrm{Safe}_n(P) := \langle E_n(P) \rangle_{\mathcal{H}}^{\Delta}$ when nonempty.

For prefixes $P_{\leq n}$ valid for h, we have $h \in \mathcal{H}_\Delta(E_n(P))$, so $\mathrm{Safe}_n(P) \subseteq \mathrm{supp}(h)$; a point in $\mathrm{Safe}_n(P) \setminus \mathrm{Seen}_n(P)$ is therefore a *certified novel positive*.

Definition 5.2 (Uniform/non-uniform contrastive generation)**.** A generator G is a *uniform contrastive generator with distinct-edge sample complexity* d if for every $h \in \mathcal{H}$ and crossing-edge stream $P \subseteq \Delta(h)$, every prefix length n with $|E_n(P)| \geq d$ yields $G(P_{\leq n}) \in \mathrm{supp}(h) \setminus \mathrm{Seen}_n(P)$. $\mathcal{H}$ is *non-uniformly* contrastively generatable if for one fixed G, some $d_h < \infty$ suffices for each h.

Definition 5.3 (Hollow edge set; contrastive closure dimension)**.** Finite $E \subseteq [\mathcal{X}]^2$ is *contrastively hollow* for $\mathcal{H}$ if $\mathcal{H}_\Delta(E) \neq \varnothing$ and $\langle E \rangle_{\mathcal{H}}^{\Delta} \setminus \mathrm{V}(E) = \varnothing$. The *contrastive closure dimension* is $\mathrm{C}_\Delta(\mathcal{H}) := \sup\{|E| : E \text{ finite contrastively hollow}\} \in \mathbb{N} \cup \{\infty\}$ (empty sup is 0).

A hollow edge set is a finite contrastive prefix after which every currently forced positive already lies in $\mathrm{V}(E)$.

Theorem 5.4 (Uniform contrastive generation)**.** *$\mathcal{H}$ is uniformly contrastively generatable iff $\mathrm{C}_\Delta(\mathcal{H}) < \infty$. Quantitatively, if $1 \leq \mathrm{C}_\Delta(\mathcal{H}) = d < \infty$, then distinct-edge sample complexity $d+1$ is both necessary and sufficient; if $\mathrm{C}_\Delta(\mathcal{H}) = 0$, one distinct edge suffices.*

Proof sketch. Sufficiency: when $|E_n(P)| > d$, $E_n(P)$ is not hollow, so the closure has a point outside $\mathrm{V}(E_n(P)) = \mathrm{Seen}_n(P)$; output the least such, which lies in $\mathrm{supp}(h)$ since $h \in \mathcal{H}_\Delta(E_n(P))$. *Necessity:* presenting the edges of a nonempty hollow E, the generator either violates novelty or outputs a point misclassified by some $h \in \mathcal{H}_\Delta(E)$, and the prefix extends to a valid presentation of h; so every hollow set is smaller than the sample complexity. The full proof is in Appendix B. □

The same exhaustion principle yields the non-uniform variant, where the number of distinct pairs needed may depend on the target.

Theorem 5.5 (Non-uniform contrastive generation)**.** *$\mathcal{H}$ is non-uniformly contrastively generatable iff there is a nondecreasing chain $\mathcal{H}_1 \subseteq \mathcal{H}_2 \subseteq \cdots$ with $\mathcal{H} = \bigcup_m \mathcal{H}_m$ and $\mathrm{C}_\Delta(\mathcal{H}_m) < \infty$ for all m.*

Remark 5.6. The construction underlying Theorem 5.5 is information-theoretic: the generator takes the chain $(\mathcal{H}_m)$ and the per-level dimensions $\{\mathrm{C}_\Delta(\mathcal{H}_m)\}_m$ as inputs.

The dimension does not capture ordinary contrastive generation: classes whose generation relies on *eventual cores* rather than uniform safe sets can have infinite C_Δ, as the next example shows.

Example 5.7 (Punctured-support class)**.** With $A = \{a_m\}_{m \geq 1} \subsetneq \mathcal{X}$ infinite and $b \in \mathcal{X} \setminus A$, define $\mathrm{supp}(h_\infty) = A$ and $\mathrm{supp}(h_m) = A \setminus \{a_m\}$ for $m \geq 1$. The eventual core $(a_m)_{m \geq 1}$ (see Proposition 5.11) gives $\mathcal{H} \in \mathsf{CtrGen}$, yet $E_n = \{\{a_i, b\} : 1 \leq i \leq n\}$ is hollow with $|E_n| = n$, so $\mathrm{C}_\Delta(\mathcal{H}) = \infty$.

5.2 Cores and confusability

Ordinary contrastive generation requires only individual convergence and admits two natural sufficient conditions: an infinite *safe core* (the contrastive closure eventually becomes infinite) and a fixed *eventual core* (a single sequence whose tail eventually enters every target's support). The negative side is governed by *confusability*: families of hypotheses whose supports admit a shared contrastive presentation.

Proposition 5.8 (Safe-core sufficiency)**.** *Suppose that for every $h \in \mathcal{H}$ and every contrastive presentation P valid for h, the safe set $\mathrm{Safe}_n(P)$ is infinite for some n (hence for all larger n, since $\mathrm{Safe}_n(P)$ grows with n). Then $\mathcal{H} \in \mathsf{CtrGen}$.*

Example 5.9 (Augmented-support class)**.** With $A = \{a_m\}_{m \geq 1} \subsetneq \mathcal{X}$ infinite and $\mathcal{X} \setminus A = \{b_m\}_{m \geq 1}$ infinite (without repetition), define $\mathrm{supp}(h_\infty) = A$ and $\mathrm{supp}(h_m) = A \cup \{b_m\}$ for $m \geq 1$. The safe core A certifies $\mathcal{H} \in \mathsf{CtrGen}$ via Proposition 5.8, but the non-covering barrier between h_m and $h_{m'}$ for any $m \neq m'$ (incomparable supports intersecting in A, yet together missing every $b_{m''}$ with $m'' \notin \{m, m'\}$) blocks CtrId; the pairs (h_∞, h_m) are comparable, the harmless superset regime (N1).

Definition 5.10 (Eventual core)**.** An injective sequence $(r_m)_{m \geq 1}$ in $\mathcal{X}$ is an *eventual core* for $\mathcal{H}$ if $\{m : r_m \notin \mathrm{supp}(h)\}$ is finite for every $h \in \mathcal{H}$.

Proposition 5.11 (Eventual-core sufficiency)**.** *A countable class of infinite proper nontrivial hypotheses with an eventual core lies in CtrGen.*

For the obstruction, given a finite $\mathcal{F} \subseteq \mathcal{H}$ let $\Gamma(\mathcal{F}) := \bigcap_{h\in\mathcal{F}} \Delta(h)$ be the family common crossing graph. A *shared contrastive presentation* for $\mathcal{F}$ is a sequence valid for every $h \in \mathcal{F}$; equivalently, $\bigcup_{h\in\mathcal{F}} \mathrm{supp}(h) \subseteq \mathrm{V}(\Gamma(\mathcal{F}))$ (see Proposition B.1 in Appendix B). Writing $\mathcal{F} \Subset \mathcal{H}$ for "$\mathcal{F}$ is a finite subset of $\mathcal{H}$", the *confusability complex* is $\mathfrak{C}(\mathcal{H}) := \{\mathcal{F} \Subset \mathcal{H} : \mathcal{F} \neq \varnothing$ and admits a shared contrastive presentation$\}$, an abstract simplicial complex (downward closed under nonempty subsets); we adopt the convention that the empty face is excluded.

Proposition 5.12 (Finite-family obstruction)**.** *If $\mathcal{F} \in \mathfrak{C}(\mathcal{H})$ with $|\mathcal{F}| \geq 2$ satisfies $|\bigcap_{h\in\mathcal{F}} \mathrm{supp}(h)| < \infty$, then $\mathcal{H} \notin$ CtrGen. (Under UUS the case $|\mathcal{F}| = 1$ is vacuous: a single $h \in \mathcal{H}$ has $\bigcap_{h\in\mathcal{F}} \mathrm{supp}(h) = \mathrm{supp}(h)$, which is infinite.)*

Proof sketch. A deterministic generator on the shared presentation must serve every $h \in \mathcal{F}$ simultaneously, eventually outputting from the finite intersection. But the shared presentation covers each $\mathrm{supp}(h)$, so the (finite) intersection eventually appears in $\mathrm{Seen}_n(P)$, forcing a novelty/precision contradiction. The full proof is in Appendix B. □

Pairwise analysis is incomplete: a family of three hypotheses can lie in $\mathfrak{C}(\mathcal{H})$ with finite triple intersection while every pairwise intersection is infinite (Appendix D).

5.3 The clean hierarchy

Theorem 5.13 (Hierarchy chain)**.** *On countable UUS classes of proper nontrivial hypotheses,* CtrId $\subsetneq$ CtrGen $\subsetneq$ TxtGen *and* CtrId $\subsetneq$ TxtId $\subsetneq$ TxtGen.

Theorem 5.14 (Hierarchy incomparability)**.** CtrGen $\not\subseteq$ TxtId *and* TxtId $\not\subseteq$ CtrGen.

Proof sketch of Theorems 5.13 and 5.14. Three witness classes do all the work. Disjoint-support $\{h_A, h_B\}$ is in TxtId but not in CtrGen (pairwise finite-intersection obstruction). The punctured class $\{h_\infty\} \cup \{h_m : \mathrm{supp}(h_m) = A \setminus \{a_m\}\}$ of Example 5.7 is in CtrGen (eventual core (a_m)) but not in TxtId (no finite tell-tale for h_∞). The augmented-support class of Example 5.9 is in CtrGen (safe core A) but not in CtrId (non-covering barrier). The remaining inclusions follow from Lemma 4.6 and Theorem 3.6 and from generating new points of an identified support. The full proof is in Appendix B. □

On finite classes the family obstruction of Proposition 5.12 is exact; for countable classes, a static characterization of ordinary contrastive generation remains open.

Theorem 5.15 (Ordinary generation on finite classes)**.** *A finite UUS class $\mathcal{H}$ of proper nontrivial hypotheses lies in* CtrGen *if and only if $|\bigcap_{h\in\mathcal{F}} \mathrm{supp}(h)| = \infty$ for every $\mathcal{F} \in \mathfrak{C}(\mathcal{H})$.*

Proof sketch. Necessity is Proposition 5.12. For sufficiency, the consistent hypotheses K eventually stop changing, and the target then lies in the confusable family $\{g \in K : \mathrm{supp}(g) \subseteq \mathrm{V}(\Gamma(K))\}$, whose common support the generator follows (Appendix B). □

On countable classes the condition of Theorem 5.15 stays necessary but is no longer sufficient: Theorem B.2 in Section B.1 gives a countable class all of whose finite subclasses lie in CtrGen while the class itself does not.

6 Robustness under Adversarial Corruption

The clean hierarchy puts contrastive identification strictly below text identification. Adversarial corruption changes the comparison: a corrupted text false positive is indistinguishable from a real one, while a corrupted contrastive pair is structurally a non-edge of $\Delta(h)$, which can break a structure that all honest pairs share. We make this asymmetry rigorous through a defect number that counts forced wrong-cut violations, and then characterize contrastive identification at every corruption budget. Omitted proofs are in Appendix C.

6.1 Corrupted presentations and defect numbers

Definition 6.1 (Corrupted presentations)**.** For $k \geq 0$, a *k-corrupted text* for h has at most k terms outside $\mathrm{supp}(h)$ and covers $\mathrm{supp}(h)$. A *k-corrupted contrastive presentation* for h has $p_t \notin \Delta(h)$ for at most k indices t (repetitions counted), with $\mathrm{supp}(h) \subseteq \mathrm{Seen}(P)$. Write k-TxtId, k-CtrId for the corresponding identification notions when k is known, and Fin-CtrId when a single identifier succeeds for every finite contrastive corruption budget. Corruption affects only the XOR condition; positive-side coverage is preserved.

Definition 6.2 (Defect number)**.** For distinct proper nontrivial h, g, the *positive-side defect set* is $D^+_{h\to g} := \mathrm{supp}(h) \setminus \mathrm{V}(\Gamma(h, g))$, and the *defect number* is $\kappa(h \to g) := |D^+_{h\to g}| \in \mathbb{N} \cup \{\infty\}$.

Proposition 6.3 (Defect number = forced wrong-cut violations)**.** *For any clean valid contrastive presentation P of h, let* $\mathrm{viol}_g(P) := |\{t : p_t \notin \Delta(g)\}|$. *Then* $\inf_{P \text{ valid for } h} \mathrm{viol}_g(P) = \kappa(h \to g)$. *In particular, g is not eliminable from h if and only if $\kappa(h \to g) = 0$.*

Proposition 6.3 is the core mechanism behind the corruption-side reversal: for distinct h, g, each positive of h not covered by any pair valid for both forces a g-violation in any clean valid presentation of h. In the cells A, B, C, D of Theorem 4.3 the defect set has a closed form (Lemma C.1): $D^+_{h\to g} = A$ if h and g cover $\mathcal{X}$, $D^+_{h\to g} = B$ if $\mathrm{supp}(g) \subsetneq \mathrm{supp}(h)$, and $D^+_{h\to g} = \varnothing$ otherwise (proofs in Appendix C).

6.2 The co-singleton reversal

We instantiate infinite defects, in both directions, on the simplest class where they occur: in the *co-singleton class*, each hypothesis labels every example positive except a single "hole", and the common-crossing graph between any two distinct targets is a single edge.

Definition 6.4 (Co-singleton class)**.** Let h_s be the hypothesis with $\mathrm{supp}(h_s) := \mathcal{X} \setminus \{s\}$. We define co-singleton class $\mathcal{H}_{\mathrm{co}} := \{h_s : s \in \mathcal{X}\}$.

Theorem 6.5 (Text fragility)**.** $\mathcal{H}_{\mathrm{co}} \notin 1$-TxtId.

Proof. The enumeration of $\mathcal{X}$ is a one-corrupted text for every h_s (the false positive is s). No identifier can distinguish targets on identical input. □

Algorithm 1: Absence-count for the co-singleton class $\mathcal{H}_{\mathrm{co}}$. On a prefix $P_{\leq n}$, for each $x \in \mathrm{Seen}_n(P)$, compute the *absence count* $a_n(x) := |\{i \leq n : x \notin p_i\}|$; let $s^* \in \mathrm{Seen}_n(P)$ minimize $a_n(x)$, breaking ties by the fixed enumeration of $\mathcal{X}$; return h_{s^*}.

Theorem 6.6 (Contrastive recovery)**.** $\mathcal{H}_{\mathrm{co}} \in$ Fin-CtrId, *hence* $\mathcal{H}_{\mathrm{co}} \in k$-CtrId *for every* $k \geq 0$.

Proof sketch. Algorithm 1 does not depend on the corruption budget. For target h_s: every honest pair has the form $\{s, z\}$, so $a_n(s) \leq k_0$ where k_0 is the (unknown) corruption count. For $t \neq s$, positive-side coverage forces infinitely many honest pairs $\{s, u\}$ with $u \neq t$, each omitting t, so $a_n(t) \to \infty$. Eventually s is the strict minimum. The full proof is in Appendix C. □

Example 6.7 (Trace on $\mathcal{H}_{\mathrm{co}}$ with $k = 1$)**.** Let $\mathcal{X} = \mathbb{N}$, target h_3, and budget $k = 1$. A possible 1-corrupted prefix is $P_{\leq 6} = \big(\{3, 0\}, \{3, 1\}, \underline{\{0, 4\}}, \{3, 2\}, \{3, 4\}, \{3, 5\}\big)$, with the underlined pair corrupted. The absence counts after $n = 6$ are $a_6(0) = 4$, $a_6(1) = 5$, $a_6(2) = 5$, $a_6(3) = 1$, $a_6(4) = 4$, $a_6(5) = 5$, identifying $s = 3$ as the absence-minimizer. As $n \to \infty$ along any 1-corrupted presentation of h_3 extending this prefix, $a_n(3) \leq 1$ stays bounded while each $a_n(t)$ for $t \neq 3$ diverges.

Theorem 6.8 (Corrupted incomparability)**.** *For every $k \geq 1$, k-CtrId and k-TxtId are incomparable.*

The proof, via the co-singleton class and blocks of size $k + 1$, is in Appendix C.

6.3 Two hypotheses: the exact threshold

Theorem 6.9 (Two hypotheses)**.** *Let $h \neq g$ be proper nontrivial, and let $\bar{\kappa}(h, g) := \max\{\kappa(h \to g), \kappa(g \to h)\}$.*

(a) *For $k \in \mathbb{N}$ the following are equivalent:* (i) $\{h, g\} \in k$-CtrId*;* (ii) *no sequence is a k-corrupted contrastive presentation of both h and g;* (iii) $\bar{\kappa}(h, g) \geq 2k + 1$*. Hence the least k for which h and g admit a common k-corrupted contrastive presentation is $\rho(h, g) = \lceil \bar{\kappa}(h, g)/2 \rceil$, and no finite k suffices when $\bar{\kappa}(h, g) = \infty$.*

(b) $\{h, g\} \in$ Fin-CtrId *if and only if $\kappa(h \to g) = \kappa(g \to h) = \infty$, that is, if and only if h and g cover $\mathcal{X}$ and share infinitely many points.*

If $\bar{\kappa} \leq 2k$, the adversary charges half of the joint defect, which lies in one cell, to each target; this is the source of the factor $\frac{1}{2}$. If $\kappa(h \to g) \geq 2k + 1$, outputting h once more than k pairs violate g succeeds, since k corrupted pairs cancel at most k forced violations (Lemma C.2; proofs in Section C.2). Knowing the budget matters already for two hypotheses: if $\mathrm{supp}(h)$ is the set of even numbers and $\mathrm{supp}(g) = \mathbb{N} \setminus \{1\}$, then $\bar{\kappa}(h, g) = \infty$, so $\{h, g\} \in k$-CtrId for every k, yet $\{h, g\} \notin$ Fin-CtrId: the honest pairs $\{c, 1\}$ with c odd, which g needs, look to h exactly like corrupted pairs.

6.4 Identification at every budget

For whole classes, the tell-tale sets must also survive the loss of $2k$ points (Example C.5). Let $\theta_k := 2k + 1$ for $k \in \mathbb{N}$ and $\theta_{\mathrm{Fin}} := \infty$, where "at least ∞ points" means infinitely many.

Theorem 6.10 (Identification at every budget)**.** *Let $\mathcal{H}$ be a countable class of proper nontrivial hypotheses and $k \in \mathbb{N} \cup \{\mathrm{Fin}\}$. Then $\mathcal{H} \in k$-CtrId if and only if* (I) *any two members of $\mathcal{H}$ with incomparable supports cover $\mathcal{X}$ and share at least θ_k points, and* (II) *every $g \in \mathcal{H}$ has a finite set $T_g \subseteq \mathrm{supp}(g)$ such that every $f \in \mathcal{H}$ with $\mathrm{supp}(f) \subsetneq \mathrm{supp}(g)$ misses at least θ_k points of T_g.*

At $k = 0$, (I) is the overlapping-cover condition and (II) is Angluin's condition, so Theorem 6.10 contains the equivalence (i) $\iff$ (iii) of Theorem 4.7. For finite k the learner outputs the least-index hypothesis whose T_g has appeared and which has at most k violations; for $k = \mathrm{Fin}$, the least-index hypothesis with the fewest violations (Section C.3).

Corollary 6.11 (Unknown budgets)**.** *Let $\mathcal{H}$ be a countable class of proper nontrivial hypotheses with at least two members. Then $\mathcal{H} \in$ Fin-CtrId if and only if the holes $\mathcal{X} \setminus \mathrm{supp}(h)$, $h \in \mathcal{H}$, are pairwise disjoint and the union of any two holes misses infinitely many points. If, moreover, two members have holes with at most r points, then $\mathcal{H} \notin r$-TxtId.*

Singleton holes give the co-singleton reversal, on which the least-violation rule is Algorithm 1 ($\mathrm{viol}_{h_x}(P_{\leq n}) = a_n(x)$), and holes of size r give classes in Fin-CtrId $\setminus\, r$-TxtId.

Corollary 6.12 (Budget hierarchy)**.** CtrId $=$ 0-CtrId $\supsetneq$ 1-CtrId $\supsetneq$ 2-CtrId $\supsetneq \cdots \supsetneq \bigcap_{k \in \mathbb{N}} k$-CtrId $\supsetneq$ Fin-CtrId.

Growing corruption. Any unbounded budget, however slowly it grows, gives every two hypotheses a common presentation, so no class with two members is identifiable under sublinear or constant-rate corruption; on the bounded-overlap classes of Section C.4, the least-violation rule succeeds whenever honest coverage outgrows the corruption (Propositions C.6 and C.7).

7 Conclusion

We studied *contrastive identification and generation in the limit*. The common crossing graph Γ unifies pairwise ambiguity, family-level generation obstructions, and corruption defects in a single coverage-and-incidence language; contrastive data are weaker than text in clean settings, but under finite corruption this inclusion fails, because every honest pair carries a negative of the target.

Several natural extensions remain open (Appendix E): *(i)* a *robust closure dimension* for corrupted contrastive generation; *(ii)* phase transitions for statistical contrastive presentations; *(iii)* effective versions of our identifiers; and *(iv)* ordinary contrastive generation on countable classes, where the condition of Theorem 5.15 is necessary but not sufficient (Section B.1).

Acknowledgments and Disclosure of Funding

We thank the anonymous reviewers and the area chair for their helpful comments. This work was supported by the Australian Research Council (ARC) under the DECRA Fellowship DE240101049. The authors declare no competing interests.

References

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

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

Dana Angluin. Inductive inference of formal languages from positive data. *Information and Control*, 45(2):117–135, 1980.

Marcelo Arenas, Pablo Barceló, Luis Cofré, and Alexander Kozachinskiy. Language generation in the limit: Complexity barriers and implications for learning. In *Proceedings of the 43rd International Conference on Machine Learning (ICML)*, volume 306 of *Proceedings of Machine Learning Research*, pages 3423–3433. PMLR, 2026.

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)*. SIAM, 2026.

Han Bao, Gang Niu, and Masashi Sugiyama. Classification from pairwise similarity and unlabeled data. In *Proceedings of the 35th International Conference on Machine Learning (ICML)*, volume 80 of *Proceedings of Machine Learning Research*, pages 452–461. PMLR, 2018.

Jessa Bekker and Jesse Davis. Learning from positive and unlabeled data: a survey. *Machine Learning*, 109:719–760, 2020.

Marc-André Carbonneau, Veronika Cheplygina, Eric Granger, and Ghyslain Gagnon. Multiple instance learning: a survey of problem characteristics and applications. *Pattern Recognition*, 77: 329–353, 2018.

Moses Charikar and Chirag Pabbaraju. Exploring facets of language generation in the limit. In *Proceedings of the Thirty-Eighth Conference on Learning Theory (COLT)*, volume 291 of *Proceedings of Machine Learning Research*, pages 854–887. PMLR, 2025.

Moses Charikar and Chirag Pabbaraju. Pareto-optimal non-uniform language generation. In *Proceedings of the International Conference on Algorithmic Learning Theory (ALT)*, 2026.

Moses Charikar, Chirag Pabbaraju, and Ambuj Tewari. A characterization of list language identification in the limit. In *Proceedings of the Thirty-Ninth Conference on Learning Theory (COLT)*, volume 336 of *Proceedings of Machine Learning Research*, pages 1022–1053. PMLR, 2026.

Soham Dan, Han Bao, and Masashi Sugiyama. Learning from noisy similar and dissimilar data. In *Machine Learning and Knowledge Discovery in Databases (ECML PKDD)*, pages 233–249. Springer, 2021. doi: 10.1007/978-3-030-86520-7_15.

Charles Elkan and Keith Noto. Learning classifiers from only positive and unlabeled data. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2008.

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

Steve Hanneke, Amin Karbasi, Anay Mehrotra, and Grigoris Velegkas. On union-closedness of language generation. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2025.

Mikael Møller Høgsgaard and Chirag Pabbaraju. Agnostic language identification and generation. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2026. To appear; arXiv:2601.23258.

Takashi Ishida, Gang Niu, Weihua Hu, and Masashi Sugiyama. Learning from complementary labels. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.

Alkis Kalavasis, Anay Mehrotra, and Grigoris Velegkas. On the limits of language generation: Trade-offs between hallucination and mode collapse. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing (STOC)*, pages 1732–1743. ACM, 2025.

Alkis Kalavasis, Anay Mehrotra, and Grigoris Velegkas. On characterizations for language generation: Interplay of hallucinations, breadth, and stability. In *Proceedings of the International Conference on Algorithmic Learning Theory (ALT)*, 2026.

Daniel M. Kane, Shachar Lovett, Shay Moran, and Jiapeng Zhang. Active classification with comparison queries. In *58th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 355–366, 2017.

Amin Karbasi, Omar Montasser, John Sous, and Grigoris Velegkas. (Im)possibility of automated hallucination detection in large language models. In *Conference on Language Modeling (COLM)*, 2025.

Jon Kleinberg and Sendhil Mullainathan. Language generation in the limit. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 37, pages 66058–66079, 2024.

Jon Kleinberg and Fan Wei. Density measures for language generation. In *Proceedings of the 66th IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, 2025.

Jon Kleinberg and Fan Wei. Language generation and identification from partial enumeration: Tight density bounds and topological characterizations. In *Proceedings of the 58th Annual ACM Symposium on Theory of Computing (STOC)*. ACM, 2026a.

Jon Kleinberg and Fan Wei. Validity, sparse holes, and breadth in language generation: Banach density, topology, and geometry. *arXiv preprint arXiv:2604.02385*, 2026b.

Steffen Lange, Thomas Zeugmann, and Sandra Zilles. Learning indexed families of recursive languages from positive data: A survey. *Theoretical Computer Science*, 397(1-3):194–232, 2008.

Aaron Li and Ian Zhang. Characterizing the effect of noise in language generation in the limit. In *Proceedings of the 43rd International Conference on Machine Learning (ICML)*, volume 306 of *Proceedings of Machine Learning Research*, pages 72277–72289. PMLR, 2026.

Jiaxun Li, Vinod Raman, and Ambuj Tewari. Generation through the lens of learning theory. In *Proceedings of the Thirty-Eighth Conference on Learning Theory (COLT)*, pages 4740–4776. PMLR, 2025.

Jiaxun Li, Vinod Raman, and Ambuj Tewari. On generation in metric spaces. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2026a. To appear; arXiv:2602.07710.

Xiaoyu Li, Andi Han, Jiaojiao Jiang, and Junbin Gao. On the price of privacy for language identification and generation. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2026b. To appear; arXiv:2604.07238.

Farnam Mansouri, Hans U. Simon, Adish Singla, Yuxin Chen, and Sandra Zilles. Formal models of active learning from contrastive examples. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 38, 2025. doi: 10.52202/085713-3282.

Anay Mehrotra, Grigoris Velegkas, Xifan Yu, and Felix Zhou. Language generation with infinite contamination. In *Proceedings of the Thirty-Ninth Conference on Learning Theory (COLT)*, volume 336 of *Proceedings of Machine Learning Research*, pages 5053–5112. PMLR, 2026a.

Anay Mehrotra, Grigoris Velegkas, Xifan Yu, and Felix Zhou. Differentially private language generation and identification in the limit (extended abstract). In *Proceedings of the Thirty-Ninth Conference on Learning Theory (COLT)*, volume 336 of *Proceedings of Machine Learning Research*, pages 5113–5114. PMLR, 2026b. Full version: arXiv:2604.08504.

Hristo Papazov and Nicolas Flammarion. Learning algorithms in the limit. In *The Thirty Eighth Annual Conference on Learning Theory (COLT)*, pages 4486–4510. PMLR, 2025.

Charlotte Peale, Vinod Raman, and Omer Reingold. Representative language generation. In *Proceedings of the 42nd International Conference on Machine Learning (ICML)*. PMLR, 2025.

Binghui Peng, Amin Saberi, and Grigoris Velegkas. Language identification in the limit with computational trace. In *The Fourteenth International Conference on Learning Representations (ICLR)*, 2026.

Novi Quadrianto, Alex J. Smola, Tibério S. Caetano, and Quoc V. Le. Estimating labels from label proportions. *Journal of Machine Learning Research (JMLR)*, 10:2349–2374, 2009.

Giorgio Racca, Michal Valko, and Amartya Sanyal. Language generation with replay: A learning-theoretic view of model collapse. In *Proceedings of the 43rd International Conference on Machine Learning (ICML)*. PMLR, 2026.

Ananth Raman and Vinod Raman. Generation from noisy examples. In *Proceedings of the 42nd International Conference on Machine Learning (ICML)*, volume 267. PMLR, 2025.

Aryan Alavi Razavi Ravari, Farnam Mansouri, Yuxin Chen, Valentio Iverson, Adish Singla, and Sandra Zilles. Learning half-spaces from perturbed contrastive examples. *arXiv preprint arXiv:2602.02080*, 2026.

Kiri Wagstaff, Claire Cardie, Seth Rogers, and Stefan Schrödl. Constrained k-means clustering with background knowledge. In *Proceedings of the 18th International Conference on Machine Learning (ICML)*, pages 577–584. Morgan Kaufmann, 2001.

R. M. Wharton. Approximate language identification. *Information and Control*, 26(3):236–255, 1974.

Appendix

A Omitted Proofs from Section 4

Theorem 4.3 (Eliminability geometry)**.** *For distinct proper nontrivial h, g, partition $\mathcal{X}$ by membership: $A = \mathrm{supp}(h) \cap \mathrm{supp}(g)$, $B = \mathrm{supp}(h) \setminus \mathrm{supp}(g)$, $C = \mathrm{supp}(g) \setminus \mathrm{supp}(h)$, $D = \mathcal{X} \setminus (\mathrm{supp}(h) \cup \mathrm{supp}(g))$. Then*

$$g \text{ is not eliminable from } h \iff (A \neq \varnothing \Rightarrow D \neq \varnothing) \wedge (B \neq \varnothing \Rightarrow C \neq \varnothing).$$

Equivalently, g is not eliminable from h in exactly three regimes: (N1) $\mathrm{supp}(h) \subsetneq \mathrm{supp}(g)$ *(*superset*);* (N2) *incomparable and disjoint (*disjoint*);* (N3) *incomparable, intersecting, and non-covering (*non-covering*).*

Proof. We use Proposition 4.2. For $x \in A$, a pair $\{x, y\}$ lies in $\Delta(h)$ iff $y \notin \mathrm{supp}(h)$, i.e. $y \in C \cup D$; to additionally lie in $\Delta(g)$ (so the pair is in $\Gamma(h, g)$), since $x \in \mathrm{supp}(g)$ we need $y \notin \mathrm{supp}(g)$, i.e. $y \in D$. Hence A-vertices are coverable iff $D \neq \varnothing$. For $x \in B$, a similar analysis shows the partner must lie in C, so B-vertices are coverable iff $C \neq \varnothing$. This proves the equivalence.

The named regimes are by case analysis on the support relation. If $\mathrm{supp}(h) \subsetneq \mathrm{supp}(g)$: $B = \varnothing$ and the only constraint is $A \neq \varnothing \Rightarrow D \neq \varnothing$, which holds since $\mathrm{supp}(g) \neq \mathcal{X}$ so $D \neq \varnothing$; thus (N1) is non-eliminable. If $\mathrm{supp}(g) \subsetneq \mathrm{supp}(h)$: $C = \varnothing$ but $B \neq \varnothing$, so the second implication fails and g is eliminable. If h, g are incomparable, $B, C \neq \varnothing$, so the second implication is automatic and only $A \neq \varnothing \Rightarrow D \neq \varnothing$ matters: it fails iff $A \neq \varnothing$ and $D = \varnothing$, i.e. iff supports cover $\mathcal{X}$ and intersect (so the pair forms an overlapping cover; g is eliminable). Otherwise non-eliminable, giving (N2) and (N3). □

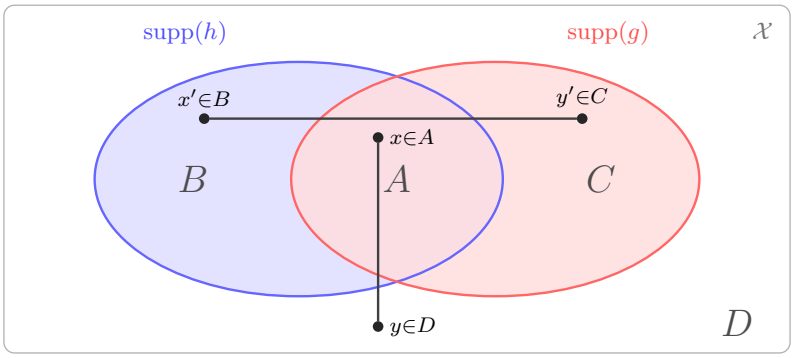


Figure 3: Four-region partition of $\mathcal{X}$ by membership in $\mathrm{supp}(h)$ and $\mathrm{supp}(g)$: A in both, B in $\mathrm{supp}(h)$ only, C in $\mathrm{supp}(g)$ only, D in neither. An A-vertex needs a partner in D to lie in $\Gamma(h, g)$; a B-vertex needs a partner in C (Theorem 4.3).

Lemma 4.4 (Pairwise shared presentation)**.** *$h, g \in \mathcal{H}$ admit a common contrastive presentation valid for both targets iff $\mathrm{supp}(h) \cup \mathrm{supp}(g) \subseteq \mathrm{V}(\Gamma(h, g))$. Mutual non-eliminability implies this confusability.*

Proof. A shared presentation can use only edges in $\Gamma(h, g)$ and must cover both positive supports, giving necessity. Conversely, choose an incident edge in $\Gamma(h, g)$ for each element of $\mathrm{supp}(h) \cup \mathrm{supp}(g)$; enumerate this set if infinite, list-and-repeat if finite, and emit the corresponding chosen edges. Every emitted pair is valid for both hypotheses and both positive sides are covered. The final statement follows because a deterministic identifier produces the same hypothesis on the resulting presentation regardless of which target generated it. □

Lemma 4.6 ($\mathsf{CtrId} \subseteq \mathsf{TxtId}$)**.** *On classes of proper nontrivial hypotheses, $\mathsf{CtrId} \subseteq \mathsf{TxtId}$, with strict inclusion even on UUS classes.*

Proof. Inclusion. Let $\mathcal{H} \in \mathsf{CtrId}$ with contrastive identifier I (extended arbitrarily to all finite sequences in $[\mathcal{X}]^2$). Define a text identifier as follows: on prefix $T_{\le n} = (x_1, \ldots, x_n)$, let z_n be the least element of $\mathcal{X} \setminus \mathrm{Seen}_n(T)$ and feed I the synthetic prefix $(\{x_t, z_n\})_{t \le n}$.

Fix target h and let $z^* := \min(\mathcal{X} \setminus \mathrm{supp}(h))$. The text covers $\mathrm{supp}(h)$, so every example before z^* eventually appears in $\mathrm{Seen}_n(T)$ while z^* does not; therefore $z_n = z^*$ for all sufficiently large n. Beyond that stage, the synthetic prefix is exactly the length-n prefix of the *single* fixed contrastive presentation $P^* := (\{x_t, z^*\})_{t \ge 1}$, which is valid for h (each pair lies in $\Delta(h)$ and positive-side coverage holds). Since I identifies h on P^*, the simulated text identifier converges.

Strictness. Pick disjoint infinite $A = \{a_n\}, B = \{b_n\}$ in $\mathcal{X}$, and let h_A, h_B have supports A, B. Then $\{h_A, h_B\}$ is text-identified from the first positive example, but $(\{a_n, b_n\})_{n \ge 1}$ is a valid contrastive presentation for both h_A and h_B, so no contrastive identifier can distinguish them. □

Theorem 4.7 (Exact characterization of CtrId)**.** *For a countable class $\mathcal{H}$ of proper nontrivial hypotheses, the following are equivalent:* (i) $\mathcal{H} \in \mathsf{CtrId}$*;* (ii) $\mathcal{H} \in \mathsf{TxtId}$ *and the contrastive non-eliminability relation is contained in the positive-data superset relation;* (iii) $\mathcal{H} \in \mathsf{TxtId}$ *and every incomparable pair in $\mathcal{H}$ is an overlapping cover.*

Proof. Write $h \to_\Delta g$ for "g is not eliminable from h" and $h \to_+ g$ for "$\mathrm{supp}(h) \subsetneq \mathrm{supp}(g)$".

(ii) $\iff$ (iii). By Theorem 4.3, the contrastive non-eliminability relations not already in $\to_+$ are precisely the disjoint and non-covering incomparable barriers (N2, N3). These are excluded iff every incomparable pair is an overlapping cover.

(i)⇒(ii). $\mathcal{H} \in \mathsf{TxtId}$ follows from Lemma 4.6. Suppose $h \to_\Delta g$ but $h \not\to_+ g$. Since Theorem 4.3 rules out the case $\mathrm{supp}(g) \subsetneq \mathrm{supp}(h)$, the pair must be incomparable, satisfying (N2) or (N3). Both are symmetric, so $g \to_\Delta h$ as well, and Lemma 4.4 produces a single presentation valid for both targets, contradicting $\mathcal{H} \in \mathsf{CtrId}$.

(iii)⇒(i). By Theorem 3.5, for each $g \in \mathcal{H}$ there is a finite $T_g \subseteq \mathrm{supp}(g)$ such that no proper sub-support inside $\mathcal{H}$ contains T_g. Fix an enumeration $\mathcal{H} = \{h_0, h_1, \ldots\}$. Call h_i *eligible at time* n if $T_{h_i} \subseteq \mathrm{Seen}_n(P)$ and every observed pair lies in $\Delta(h_i)$. The identifier outputs the eligible hypothesis of smallest index (default arbitrary if none).

Let $h = h_i$ be the target. Since T_h is finite and $\mathrm{supp}(h)$ is covered, h is eventually eligible. We show every h_j with $j < i$ is eventually ineligible.

Case 1: $\mathrm{supp}(h_j) \subsetneq \mathrm{supp}(h)$. h_j is eliminable from h by the superset direction (already from positive data), so eventually inconsistent.

Case 2: $\mathrm{supp}(h), \mathrm{supp}(h_j)$ *are incomparable.* Condition (iii) makes them an overlapping cover, so by Theorem 4.3 h_j is eliminable from h, eventually inconsistent.

Case 3: $\mathrm{supp}(h) \subsetneq \mathrm{supp}(h_j)$. We show h_j is never eligible. Suppose for contradiction h_j is eligible at some time n, so $T_{h_j} \subseteq \mathrm{Seen}_n(P)$ and every observed pair lies in $\Delta(h_j)$. The tell-tale property of T_{h_j} rules out $T_{h_j} \subseteq \mathrm{supp}(h)$ (since $\mathrm{supp}(h)$ is a proper sub-support of $\mathrm{supp}(h_j)$). Take $t \in T_{h_j} \setminus \mathrm{supp}(h)$: by eligibility $t \in \mathrm{Seen}_n(P)$, so t appeared in some observed pair $\{t, y\}$, which lies in $\Delta(h)$ by validity. Since $t \notin \mathrm{supp}(h)$, the pair has $y \in \mathrm{supp}(h) \subsetneq \mathrm{supp}(h_j)$ and $t \in T_{h_j} \subseteq \mathrm{supp}(h_j)$, so both endpoints lie in $\mathrm{supp}(h_j)$, contradicting $\{t, y\} \in \Delta(h_j)$.

Thus every earlier h_j is eventually ineligible. Only finitely many indices precede i, so the identifier converges to h. □

B Omitted Proofs from Section 5

Theorem 5.4 (Uniform contrastive generation)**.** *$\mathcal{H}$ is uniformly contrastively generatable iff $\mathrm{C}_\Delta(\mathcal{H}) < \infty$. Quantitatively, if $1 \le \mathrm{C}_\Delta(\mathcal{H}) = d < \infty$, then distinct-edge sample complexity $d + 1$ is both necessary and sufficient; if $\mathrm{C}_\Delta(\mathcal{H}) = 0$, one distinct edge suffices.*

Proof. Sufficiency. Suppose $\mathrm{C}_\Delta(\mathcal{H}) = d < \infty$. On prefix $P_{\le n}$, let $E := E_n(P)$. If $|E| \le d$ or $\mathcal{H}_\Delta(E) = \varnothing$, output anything; otherwise output the least element of $\langle E \rangle_{\mathcal{H}}^{\Delta} \setminus \mathrm{V}(E)$. For target h and any crossing-edge stream P, $h \in \mathcal{H}_\Delta(E_n(P))$, so the version space is nonempty; when

$|E_n(P)| > d$, $E_n(P)$ is not hollow, so the closure has a point outside $\mathrm{V}(E_n(P)) = \mathrm{Seen}_n(P)$, and the chosen point lies in $\mathrm{supp}(h)$ since $h \in \mathcal{H}_\Delta(E_n(P))$.

Necessity. Let G be a uniform contrastive generator with distinct-edge sample complexity d', and let E be a nonempty hollow edge set; suppose for contradiction that $|E| \geq d'$. Present the edges of E in any order as the first $|E|$ pairs of a presentation; then $|E_{|E|}| = |E| \geq d'$, so G is required to output a novel positive at step $|E|$. If G outputs $x \in \mathrm{V}(E)$, it violates novelty; otherwise hollowness gives $x \notin \langle E \rangle^{\Delta}_{\mathcal{H}}$, so there is $h \in \mathcal{H}_\Delta(E)$ with $x \notin \mathrm{supp}(h)$. Since h is proper nontrivial, the prefix extends to a valid contrastive presentation for h by covering remaining positives via any h-crossing edges, and on this extension G errs at step $|E|$, a contradiction. Hence every nonempty hollow edge set has fewer than d' edges. So $\mathrm{C}_\Delta(\mathcal{H}) < \infty$, and if $\mathrm{C}_\Delta(\mathcal{H}) = d \geq 1$ then $d' \geq d + 1$, so sample complexity $d + 1$ is necessary. □

Theorem 5.5 (Non-uniform contrastive generation)**.** *$\mathcal{H}$ is non-uniformly contrastively generatable iff there is a nondecreasing chain $\mathcal{H}_1 \subseteq \mathcal{H}_2 \subseteq \cdots$ with $\mathcal{H} = \bigcup_m \mathcal{H}_m$ and $\mathrm{C}_\Delta(\mathcal{H}_m) < \infty$ for all m.*

Proof. Necessity. Let G be a non-uniform contrastive generator and define $\mathcal{H}_m$ as the set of $h \in \mathcal{H}$ for which G is correct after m distinct edges, on every crossing-edge stream for h. Then $\mathcal{H}_1 \subseteq \mathcal{H}_2 \subseteq \cdots$ and $\bigcup_m \mathcal{H}_m = \mathcal{H}$. If $\mathcal{H}_m$ had a hollow edge set E of size $\geq m$, the necessity argument of Theorem 5.4 applied inside $\mathcal{H}_m$ would produce a target in $\mathcal{H}_m$ on which G errs after m distinct edges. Hence $\mathrm{C}_\Delta(\mathcal{H}_m) < m$.

Sufficiency. Set $d_m := m + \mathrm{C}_\Delta(\mathcal{H}_m) + 1$; then $d_m \to \infty$ and $|E| \geq d_m \Rightarrow |E| > \mathrm{C}_\Delta(\mathcal{H}_m)$. On prefix $P_{\leq n}$ with $E := E_n(P)$, choose the largest m with $d_m \leq |E|$ (default arbitrarily otherwise). If $\mathcal{H}_\Delta(E) \cap \mathcal{H}_m = \varnothing$, output arbitrarily; else output the least element of $\langle E \rangle^{\Delta}_{\mathcal{H}_m} \setminus \mathrm{V}(E)$, which is nonempty because $|E| > \mathrm{C}_\Delta(\mathcal{H}_m)$. For target $h \in \mathcal{H}$, pick m_0 with $h \in \mathcal{H}_{m_0}$; once $|E_n(P)| \geq d_{m_0}$, the chosen m is at least m_0 and $h \in \mathcal{H}_m$, so the output lies in $\mathrm{supp}(h) \setminus \mathrm{Seen}_n(P)$. □

Proposition 5.8 (Safe-core sufficiency)**.** *Suppose that for every $h \in \mathcal{H}$ and every contrastive presentation P valid for h, the safe set $\mathrm{Safe}_n(P)$ is infinite for some n (hence for all larger n, since $\mathrm{Safe}_n(P)$ grows with n). Then $\mathcal{H} \in \mathsf{CtrGen}$.*

Proof. At step n, output the least element of $\mathrm{Safe}_n(P) \setminus (\mathrm{Seen}_n(P) \cup \{\hat{x}_1, \ldots, \hat{x}_{n-1}\})$ if this set is nonempty, and an arbitrary point otherwise. Since P is valid for the target h, every observed pair lies in $\Delta(h)$, so $h \in \mathcal{H}_\Delta(E_n(P))$ and $\mathrm{Safe}_n(P) \subseteq \mathrm{supp}(h)$; every output of the first kind is therefore a positive of h. As n grows, $E_n(P)$ grows and $\mathcal{H}_\Delta(E_n(P))$ shrinks, so $\mathrm{Safe}_n(P)$ grows; by hypothesis it is infinite for all large n, and from then on the excluded set is finite and the output is an unseen positive of h. □

Proposition 5.11 (Eventual-core sufficiency)**.** *A countable class of infinite proper nontrivial hypotheses with an eventual core lies in* CtrGen.

Proof. At step n, output r_m for the least $m \geq n$ with $r_m \notin \mathrm{Seen}_n(P) \cup \{\hat{x}_1, \ldots, \hat{x}_{n-1}\}$. The tail $\{r_m : m \geq n\}$ is infinite (the sequence is injective) while the excluded set is finite, so such m exists, satisfying novelty. For target h, the eventual-core hypothesis states that $\{m : r_m \notin \mathrm{supp}(h)\}$ is finite, so for all sufficiently large m the term r_m lies in $\mathrm{supp}(h)$. Since the chosen index satisfies $m \geq n$ and $n \to \infty$, the output is eventually in $\mathrm{supp}(h)$, giving precision. □

Proposition B.1 (Family shared-presentation criterion)**.** *A finite family $\mathcal{F}$ of proper nontrivial hypotheses admits a shared contrastive presentation iff $\bigcup_{h \in \mathcal{F}} \mathrm{supp}(h) \subseteq \mathrm{V}(\Gamma(\mathcal{F}))$.*

Proof. A shared presentation uses only edges in $\Gamma(\mathcal{F})$ and covers each $\mathrm{supp}(h)$, giving necessity. Sufficiency: pick an incident edge in $\Gamma(\mathcal{F})$ for each $x \in \bigcup_h \mathrm{supp}(h)$, enumerate (or list-and-repeat) and emit the corresponding pairs. □

Proposition 5.12 (Finite-family obstruction)**.** *If $\mathcal{F} \in \mathfrak{C}(\mathcal{H})$ with $|\mathcal{F}| \geq 2$ satisfies $|\bigcap_{h \in \mathcal{F}} \mathrm{supp}(h)| < \infty$, then $\mathcal{H} \notin \mathsf{CtrGen}$. (Under UUS the case $|\mathcal{F}| = 1$ is vacuous: a single $h \in \mathcal{H}$ has $\bigcap_{h \in \mathcal{F}} \mathrm{supp}(h) = \mathrm{supp}(h)$, which is infinite.)*

Proof. Proposition B.1 produces a shared presentation P for $\mathcal{F}$. On P, a deterministic generator outputs the same sequence $(\hat{x}_n)$ regardless of which $h \in \mathcal{F}$ is the target. If it succeeds for every $h \in \mathcal{F}$, then beyond the maximum individual convergence time all outputs lie in $\bigcap_{h\in\mathcal{F}} \mathrm{supp}(h)$. This intersection is finite, and P, being valid for each h, covers $\mathrm{supp}(h) \supseteq \bigcap_{h'\in\mathcal{F}} \mathrm{supp}(h')$, so the (finite) intersection eventually appears in $\mathrm{Seen}_n(P)$. Beyond that stage novelty forbids outputting from the intersection, while eventual precision for every target requires it: contradiction. □

Theorem 5.13 (Hierarchy chain)**.** *On countable UUS classes of proper nontrivial hypotheses,* CtrId $\subsetneq$ CtrGen $\subsetneq$ TxtGen *and* CtrId $\subsetneq$ TxtId $\subsetneq$ TxtGen.

Proof. CtrId $\subseteq$ CtrGen. Given a contrastive identifier I, define $G(P_{\leq n})$ as the least element of $\mathrm{supp}(I(P_{\leq n})) \setminus (\mathrm{Seen}_n(P) \cup \{\hat{x}_1, \ldots, \hat{x}_{n-1}\})$. Once I converges to the (infinite-support) target h, every output lies in $\mathrm{supp}(h)$ and is novel.

Strictness of CtrId $\subsetneq$ CtrGen. Pick disjoint infinite $A, B = \{b_m\}_{m\geq 1} \subseteq \mathcal{X}$ with $\mathcal{X} \setminus (A \cup B) \neq \varnothing$ (possible since $\mathcal{X}$ is countably infinite). Define $\mathrm{supp}(h_\infty) := A$ and $\mathrm{supp}(h_m) := A \cup \{b_m\}$ for $m \geq 1$. For any target and any valid presentation P with edge set $E_n(P)$, the version space $\mathcal{H}_\Delta(E_n(P))$ contains the target itself, and every hypothesis in $\mathcal{H}$ has support $\supseteq A$; hence $\mathrm{Safe}_n(P) = \bigcap_{g\in\mathcal{H}_\Delta(E_n(P))} \mathrm{supp}(g) \supseteq A$ is infinite, and Proposition 5.8 gives $\mathcal{H} \in$ CtrGen. The class is not in CtrId: for $m \neq r$, $\mathrm{supp}(h_m)$ and $\mathrm{supp}(h_r)$ are incomparable, intersect in A, and miss $\mathcal{X} \setminus (A \cup B)$, failing Theorem 4.7.

CtrGen $\subseteq$ TxtGen *and strictness.* Inclusion follows from Theorem 3.6. For strictness, take disjoint infinite $A = \{a_n\}, B = \{b_n\}$ and consider $\{h_A, h_B\}$ with supports A, B. It is in TxtGen by Theorem 3.6, but $(\{a_n, b_n\})_{n\geq 1}$ is valid for both targets while $A \cap B = \varnothing$, so Proposition 5.12 applies.

CtrId $\subseteq$ TxtId *and strictness* are Lemma 4.6.

TxtId $\subseteq$ TxtGen *and strictness* follow from Theorem 3.6 and the punctured-support example below. □

Theorem 5.14 (Hierarchy incomparability)**.** CtrGen $\not\subseteq$ TxtId *and* TxtId $\not\subseteq$ CtrGen.

Proof. TxtId $\not\subseteq$ CtrGen. The class $\{h_A, h_B\}$ above is in TxtId (identified from the first observed positive) but not in CtrGen.

CtrGen $\not\subseteq$ TxtId. Take infinite $A = \{a_m\} \subsetneq \mathcal{X}$ and define $\mathrm{supp}(h_\infty) = A$, $\mathrm{supp}(h_m) = A \setminus \{a_m\}$. Then (a_m) is an eventual core, so $\mathcal{H} \in$ CtrGen by Proposition 5.11. For any finite $T \subseteq A$, pick m with $a_m \notin T$; then $T \subseteq A \setminus \{a_m\} \subsetneq A = \mathrm{supp}(h_\infty)$, so h_∞ has no finite text tell-tale and $\mathcal{H} \notin$ TxtId. □

Proof of Theorem 5.15. Necessity. For $\mathcal{F} \in \mathfrak{C}(\mathcal{H})$ with $|\mathcal{F}| \geq 2$ this is Proposition 5.12; for $|\mathcal{F}| = 1$ it is UUS.

Sufficiency. For a prefix $P_{\leq n}$ let $K_n := \{f \in \mathcal{H} : E_n(P) \subseteq \Delta(f)\}$ and $D(K_n) := \{g \in K_n : \mathrm{supp}(g) \subseteq \mathrm{V}(\Gamma(K_n))\}$. If $D(K_n) \neq \varnothing$, then $\Gamma(K_n) \subseteq \Gamma(D(K_n))$ because $D(K_n) \subseteq K_n$, so $\bigcup_{g\in D(K_n)} \mathrm{supp}(g) \subseteq \mathrm{V}(\Gamma(D(K_n)))$ and $D(K_n) \in \mathfrak{C}(\mathcal{H})$ by Proposition B.1; by assumption its common support is infinite. The generator outputs the least point of $\bigcap_{g\in D(K_n)} \mathrm{supp}(g) \setminus \mathrm{Seen}_n(P)$, and an arbitrary point if $D(K_n) = \varnothing$.

Fix a target h and a valid presentation P of h. The sets K_n decrease and $\mathcal{H}$ is finite, so $K_n = K$ for all $n \geq n_0$, where K is the set of f with every pair of P in $\Delta(f)$; it contains h. Every pair of P lies in $\Gamma(K)$, and P covers $\mathrm{supp}(h)$, so $\mathrm{supp}(h) \subseteq \mathrm{V}(\Gamma(K))$ and $h \in D(K)$. For $n \geq n_0$ the output is therefore an unseen point of $\bigcap_{g\in D(K)} \mathrm{supp}(g) \subseteq \mathrm{supp}(h)$. □

B.1 Countable classes

The proof of Theorem 5.15 uses finiteness once: the set K_n of consistent hypotheses stops changing. On countable classes the condition of Theorem 5.15 stays necessary (Proposition 5.12), but no repair of this step is possible, because the condition is no longer sufficient.

Theorem B.2 (Finite subclasses do not decide generation)**.** *There is a countable UUS class $\mathcal{H}$ of proper nontrivial hypotheses such that $|\bigcap_{h\in\mathcal{F}} \mathrm{supp}(h)| = \infty$ for every $\mathcal{F} \in \mathfrak{C}(\mathcal{H})$, so that every finite subclass of $\mathcal{H}$ lies in* CtrGen *by Theorem 5.15, but $\mathcal{H} \notin$* CtrGen.

Proof. Split $\mathcal{X}$ into an infinite set $C = \{c_1, c_2, \dots\}$, four further infinite sets A_1, A_2, B_1, B_2, and a nonempty set N, and let

$$\mathrm{supp}(a) = C \cup A_1 \cup A_2, \qquad \mathrm{supp}(b) = C \cup B_1 \cup B_2, \qquad \mathrm{supp}(s_m) = \{c_1, \dots, c_m\} \cup A_1 \cup B_1$$

for $m \geq 1$, and $\mathcal{H} := \{a, b, s_1, s_2, \dots\}$. All supports are infinite and miss N.

We first show that every $\mathcal{F} \in \mathfrak{C}(\mathcal{H})$ has infinite common support. If $a \notin \mathcal{F}$, every member of $\mathcal{F}$ contains B_1; if $b \notin \mathcal{F}$, every member contains A_1; and if $\mathcal{F} = \{a, b\}$, the common support is C. In the remaining case $\mathcal{F} = \{a, b\} \cup S$ for a nonempty finite set S of hypotheses s_m, and $\mathcal{F} \notin \mathfrak{C}(\mathcal{H})$. Indeed, take c_k with k larger than every index in S. Its pattern is 1 on a and b and 0 on S. A point with the complementary pattern would lie in every member of S but outside $\mathrm{supp}(a) \cup \mathrm{supp}(b)$, and there is no such point, because every $\mathrm{supp}(s_m)$ lies inside $\mathrm{supp}(a) \cup \mathrm{supp}(b)$. So (4) fails (Proposition D.1).

Now let G be a contrastive generator for $\mathcal{H}$; we derive a contradiction. Fix $z \in N$, and call a pair *admissible* if it has one of the forms

$$\{c_k, z\}, \qquad \{\alpha, \beta'\} \ (\alpha \in A_1,\ \beta' \in B_2), \qquad \{\beta, \alpha'\} \ (\beta \in B_1,\ \alpha' \in A_2).$$

Reading off the supports, every admissible pair crosses the cuts of a and b. The last two forms also cross the cut of every s_m, and $\{c_k, z\}$ crosses the cut of s_m exactly when $k \leq m$. Apply Lemma C.4 to the set $\mathcal{S}$ of all finite sequences of admissible pairs, with $Y := \mathrm{supp}(a) \cup \mathrm{supp}(b)$ and with property E: "G outputs a point of $A_1 \cup B_1$". Let $\sigma \in \mathcal{S}$, and choose $m \geq 1$ with $k \leq m$ for every pair $\{c_k, z\}$ in σ. Let τ_σ be σ followed by admissible pairs that cover $\{c_1, \dots, c_m\} \cup A_1 \cup B_1$, using the form $\{c_k, z\}$ only for $k \leq m$. Every prefix of τ_σ lies in $\mathcal{S}$, and every point of Y lies in an admissible pair, so condition (ii) holds. For (i), note that τ_σ is a clean presentation of s_m. So G eventually outputs only unseen points of $\mathrm{supp}(s_m)$, and once $c_1, \dots, c_m$ have appeared these lie in $A_1 \cup B_1$.

Starting from the empty sequence, the lemma gives a sequence P of admissible pairs that covers $\mathrm{supp}(a) \cup \mathrm{supp}(b)$. It is a clean presentation of both a and b, and G outputs a point of $A_1 \cup B_1$ on infinitely many of its prefixes. Then G outputs infinitely often a point of A_1, which misses $\mathrm{supp}(b)$, or infinitely often a point of B_1, which misses $\mathrm{supp}(a)$. In either case G fails on P. □

For comparison, every countable UUS class is generatable from text (Theorem 3.6); contrastive generation can fail already for a countable class all of whose finite subclasses succeed.

C Omitted Proofs from Section 6

Proposition 6.3 (Defect number = forced wrong-cut violations)**.** *For any clean valid contrastive presentation P of h, let $\mathrm{viol}_g(P) := |\{t : p_t \notin \Delta(g)\}|$. Then $\inf_{P \text{ valid for } h} \mathrm{viol}_g(P) = \kappa(h \to g)$. In particular, g is not eliminable from h if and only if $\kappa(h \to g) = 0$.*

Proof. Lower bound. Fix any valid presentation P of h. For each $x \in D^+_{h\to g}$, positive-side coverage forces x to appear in at least one pair of P; let $t(x)$ be the index of its first appearance. Since $p_{t(x)} \in \Delta(h)$ and x has no incident edge in $\Gamma(h, g)$, the pair $p_{t(x)}$ lies outside $\Delta(g)$, i.e., it is a g-violation. The map $x \mapsto t(x)$ is injective: each h-valid pair has exactly one endpoint in $\mathrm{supp}(h)$ (the other being a non-positive of h), so a single pair can be the first-covering pair of at most one defect. Therefore $\mathrm{viol}_g(P) \geq |\{t(x) : x \in D^+_{h\to g}\}| = |D^+_{h\to g}|$.

Upper bound. If $D^+_{h\to g}$ is infinite, enumerate $\mathrm{supp}(h)$ and pair every non-defect with a common-crossing partner and every defect with any h-negative partner; the result is a valid presentation with infinitely many violations, matching $\kappa(h \to g) = \infty$.

If $D^+_{h\to g}$ is finite, cover each defect once with an h-negative partner (contributing exactly $|D^+_{h\to g}|$ violations) and each non-defect via a chosen common-crossing partner (no violations). For infinite

$\mathrm{supp}(h)$ the enumeration provides infinitely many pairs; for finite $\mathrm{supp}(h)$ we may, after these covering pairs, repeat any non-defect common-crossing pair forever, provided one exists.

Claim. If both h and g are proper nontrivial, there exists $x \in \mathrm{supp}(h) \setminus D^+_{h\to g}$, i.e. a non-defect positive.

Proof of claim. Suppose otherwise: $\mathrm{supp}(h) \subseteq D^+_{h\to g}$, so no vertex of $\mathrm{supp}(h) = A \cup B$ is incident to a common-crossing edge. An A-vertex has a common-crossing partner iff $D \neq \varnothing$; a B-vertex iff $C \neq \varnothing$. Hence $A = \varnothing$ or $D = \varnothing$, and $B = \varnothing$ or $C = \varnothing$. Each of the four resulting sub-cases violates a standing assumption:

- $A = \varnothing$ and $B = \varnothing$: then $\mathrm{supp}(h) = A \cup B = \varnothing$, contradicting h proper nontrivial.
- $A = \varnothing$ and $C = \varnothing$: then $\mathrm{supp}(g) = A \cup C = \varnothing$, contradicting g proper nontrivial.
- $B = \varnothing$ and $D = \varnothing$: then $\mathrm{supp}(h) \subseteq \mathrm{supp}(g)$ and $\mathrm{supp}(h) \cup \mathrm{supp}(g) = \mathcal{X}$, so $\mathrm{supp}(g) = \mathcal{X}$, contradicting g proper nontrivial.
- $C = \varnothing$ and $D = \varnothing$: symmetrically $\mathrm{supp}(g) \subseteq \mathrm{supp}(h)$ and $\mathrm{supp}(h) \cup \mathrm{supp}(g) = \mathcal{X}$, so $\mathrm{supp}(h) = \mathcal{X}$, contradicting h proper nontrivial.

The claim supplies a violation-free pair to repeat, completing the upper bound. □

Theorem 6.6 (Contrastive recovery)**.** $\mathcal{H}_{\mathrm{co}} \in \text{Fin-CtrId}$*, hence* $\mathcal{H}_{\mathrm{co}} \in k\text{-CtrId}$ *for every* $k \geq 0$.

Proof. We show that Algorithm 1 is independent of the corruption budget.

Fix target h_s and let k_0 be the actual (finite, unknown) corruption count. Every honest pair has the form $\{s, z\}$ with $z \neq s$, so s is incident to every honest pair; hence s is absent from at most k_0 pairs, giving $a_n(s) \leq k_0$ for all n.

For $t \neq s$: every honest pair has the form $\{s, u\}$ with $u \neq s$, so the *only* honest pair-set containing t is $\{s, t\}$ (it may appear in the stream any number of times, but every other honest pair omits t). We claim that infinitely many distinct $u \in \mathcal{X} \setminus \{s, t\}$ each appear in at least one honest pair $\{s, u\}$ in the stream. Indeed, positive-side coverage forces $\mathcal{X} \setminus \{s\} \subseteq \mathrm{Seen}(P)$, so each $u \in \mathcal{X} \setminus \{s, t\}$ appears in at least one pair. At most k_0 pairs are corrupted, so at most $2k_0$ distinct vertices appear *only* in corrupted pairs; since $|\mathcal{X} \setminus \{s, t\}| = \infty$, all but finitely many such u appear in some honest pair, which is necessarily $\{s, u\}$ and omits t. Each such honest pair occurrence increments $a_n(t)$, so $a_n(t) \to \infty$.

Therefore only examples appearing in the first $k_0 + 1$ pairs ever have absence count at most k_0; only finitely many false candidates can compete with s, each with divergent absence count. From some stage on, s is the unique minimizer, and the identifier outputs h_s forever. □

Theorem 6.8 (Corrupted incomparability)**.** *For every* $k \geq 1$, $k\text{-CtrId}$ *and* $k\text{-TxtId}$ *are incomparable.*

Proof. $k\text{-CtrId} \not\subseteq k\text{-TxtId}$*:* the co-singleton class is in $\text{Fin-CtrId} \subseteq k\text{-CtrId}$ (Theorem 6.6). For the text side, $k\text{-TxtId} \subseteq 1\text{-TxtId}$ (more corruption can only hurt the learner), so co-singleton $\notin 1\text{-TxtId}$ (Theorem 6.5) implies co-singleton $\notin k\text{-TxtId}$.

$k\text{-TxtId} \not\subseteq k\text{-CtrId}$*:* fix k and pick disjoint infinite $A, B \subseteq \mathcal{X}$ with $\mathcal{X} \setminus (A \cup B) \neq \varnothing$ (possible since $\mathcal{X}$ is countably infinite); partition B into pairwise disjoint blocks $B_0, B_1, \ldots$ with $|B_i| = k + 1$. Set $\mathrm{supp}(h_i) := A \cup B_i$ and $\mathcal{H}_k := \{h_i : i \geq 0\}$.

$\mathcal{H}_k \in k\text{-TxtId}$*:* in any k-corrupted text for h_i, the entire B_i eventually appears; for $j \neq i$, B_j is disjoint from $\mathrm{supp}(h_i)$, so every observed element of B_j is a false positive. Since $|B_j| = k + 1 > k$, no false block is fully observable. The identifier waits until some block has been entirely seen, then outputs the corresponding h_i.

$\mathcal{H}_k \notin k\text{-CtrId}$*:* the class fails already in the clean case. For $i \neq j$, $\mathrm{supp}(h_i)$ and $\mathrm{supp}(h_j)$ are incomparable, intersect in A, and miss $\mathcal{X} \setminus (A \cup B_i \cup B_j) \neq \varnothing$, giving the non-covering barrier (N3) of Theorem 4.3; Lemma 4.4 produces a clean shared presentation, which is in particular a 0-corrupted (hence k-corrupted) shared presentation. □

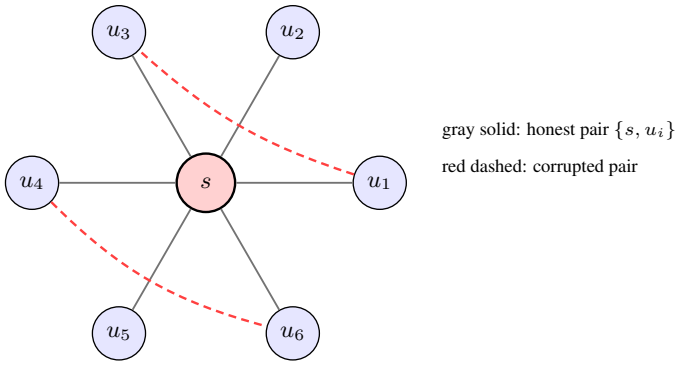


Figure 4: The honest contrastive presentation for target h_s in the co-singleton class $\mathcal{H}_{\text{co}}$. Since s is the unique negative element, every honest pair has the form $\{s, u\}$, forming a star centered at s. An adversary may plant at most $k_0 < \infty$ corrupted pairs not incident to s (red dashed). The absence count $a_n(s) \le k_0$ stays bounded, while for any $t \neq s$ infinitely many honest pairs $\{s, u\}$ omit t, so $a_n(t) \to \infty$. The absence-count algorithm outputs the co-singleton centered at the absence-minimizer.

C.1 Cells, charging, and a diagonal construction

Throughout this section $h \neq g$ are proper nontrivial, with the cells $A = \operatorname{supp}(h) \cap \operatorname{supp}(g)$, $B = \operatorname{supp}(h) \setminus \operatorname{supp}(g)$, $C = \operatorname{supp}(g) \setminus \operatorname{supp}(h)$ and $D = \mathcal{X} \setminus (\operatorname{supp}(h) \cup \operatorname{supp}(g))$ of Theorem 4.3. A pair is *honest* for h if it lies in $\Delta(h)$, and for a finite or infinite sequence $Q = (q_t)$ of pairs we write $\operatorname{viol}_g(Q) := |\{t : q_t \notin \Delta(g)\}|$. We say that h and g *cover* $\mathcal{X}$ if $\operatorname{supp}(h) \cup \operatorname{supp}(g) = \mathcal{X}$, and we call $\mathcal{X} \setminus \operatorname{supp}(h)$ the *hole* of h.

Lemma C.1 (Cells of two hypotheses)**.** (a) *A pair lies in* $\Gamma(h, g)$ *if and only if it joins* A *to* D *or* B *to* C*, and* $\Gamma(h, g) \neq \varnothing$.

(b) *At most one of* B, C, D *is empty. Moreover,* $D = \varnothing$ *iff* h *and* g *cover* $\mathcal{X}$*;* $C = \varnothing$ *iff* $\operatorname{supp}(g) \subsetneq \operatorname{supp}(h)$*; and* $B = \varnothing$ *iff* $\operatorname{supp}(h) \subsetneq \operatorname{supp}(g)$.

(c) $D^+_{h\to g} = A$ *if* $D = \varnothing$*,* $D^+_{h\to g} = B$ *if* $C = \varnothing$*, and* $D^+_{h\to g} = \varnothing$ *otherwise. Symmetrically,* $D^+_{g\to h} = A$ *if* $D = \varnothing$*,* $D^+_{g\to h} = C$ *if* $B = \varnothing$*, and* $D^+_{g\to h} = \varnothing$ *otherwise.*

Consequently the joint defect $\operatorname{Def}(h, g) := D^+_{h\to g} \cup D^+_{g\to h} = (\operatorname{supp}(h) \cup \operatorname{supp}(g)) \setminus \operatorname{V}(\Gamma(h, g))$ *is empty or equal to one of the cells* A, B, C*, and* $|\operatorname{Def}(h, g)| = \bar{\kappa}(h, g)$.

Proof. (a) A pair $\{x, y\}$ crosses both cuts iff $h(y) = 1 - h(x)$ and $g(y) = 1 - g(x)$, that is, iff the membership patterns of x and y are complementary: 11 and 00 (cells A and D), or 10 and 01 (cells B and C). Nonemptiness is the claim in the proof of Proposition 6.3. (b) If $B = C = \varnothing$, then $\operatorname{supp}(h) = \operatorname{supp}(g)$ and $h = g$. If $B = D = \varnothing$, then $\operatorname{supp}(h) \subseteq \operatorname{supp}(g)$ and $\operatorname{supp}(h) \cup \operatorname{supp}(g) = \mathcal{X}$, so $\operatorname{supp}(g) = \mathcal{X}$; symmetrically $C = D = \varnothing$ gives $\operatorname{supp}(h) = \mathcal{X}$. The equivalences restate the definitions of the cells, and the inclusions are strict because $h \neq g$. (c) By (a), a point of A lies in $\operatorname{V}(\Gamma(h, g))$ iff $D \neq \varnothing$, a point of B iff $C \neq \varnothing$, and a point of C iff $B \neq \varnothing$. So $D^+_{h\to g} = [D = \varnothing]A \cup [C = \varnothing]B$ and $D^+_{g\to h} = [D = \varnothing]A \cup [B = \varnothing]C$, and by (b) at most one of the brackets in play is nonzero. Hence either both defect sets equal A, or at most one of them is nonempty; in every case their union is one cell or empty and has $\max\{\kappa(h \to g), \kappa(g \to h)\}$ points. □

So for two distinct proper nontrivial hypotheses, $\bar{\kappa}(h, g) = |\operatorname{supp}(h) \cap \operatorname{supp}(g)|$ if they cover $\mathcal{X}$, $\bar{\kappa}(h, g) = |\operatorname{supp}(h) \triangle \operatorname{supp}(g)|$ if their supports are nested, and $\bar{\kappa}(h, g) = 0$ otherwise.

Lemma C.2 (Charging)**.** *Let* $k \in \mathbb{N}$ *and let* P *be a* k*-corrupted contrastive presentation for* h*. If* $W \subseteq D^+_{h\to g}$ *is finite and* $W \subseteq \operatorname{Seen}_n(P)$*, then* $\operatorname{viol}_g(P_{\le n}) \ge |W| - k$*. Consequently, if* $\kappa(h \to g) \ge 2k + 1$*, then* $\operatorname{viol}_g(P_{\le n}) \ge k + 1$ *for all large* n*; and if* $\kappa(h \to g) = \infty$ *and* P *is* k_0*-corrupted for some finite* k_0*, then* $\operatorname{viol}_g(P_{\le n}) \to \infty$.

Proof. For $x \in W$ let $p_{t(x)}$ be the first pair of P that contains x; then $t(x) \le n$. If $p_{t(x)}$ is honest for h, it lies in $\Delta(h)$ but not in $\Gamma(h,g)$, because x has no common crossing, so it lies outside $\Delta(g)$. An honest pair contains exactly one positive of h, so different points of W have different honest first pairs. A corrupted pair is the first pair of at most two points of W, and if it is the first pair of two points, both lie in one cell by Lemma C.1(c); then g gives them the same label and the pair lies outside $\Delta(g)$. Let y_1 and y_2 count the corrupted pairs in $P_{\le n}$ that are the first pair of exactly one and of exactly two points of W. Then $y_1 + y_2 \le k$, and

$$\mathrm{viol}_g(P_{\le n}) \;\ge\; \big(|W| - y_1 - 2y_2\big) + y_2 \;\ge\; |W| - k.$$

For the consequences, note that P covers $\mathrm{supp}(h) \supseteq W$, so every finite $W \subseteq D^+_{h\to g}$ lies in $\mathrm{Seen}_n(P)$ for all large n, and that $\mathrm{viol}_g(P_{\le n})$ does not decrease in n. □

Lemma C.3 (Charging with disjoint holes)**.** *Suppose the holes of h and g are disjoint, let $p_1, \dots, p_n$ be pairs, and let S_n be the set of positives of h that lie in some $p_t \in \Delta(h)$ with $t \le n$. Then $\mathrm{viol}_g(p_1, \dots, p_n) \ge |S_n \cap \mathrm{supp}(g)|$.*

Proof. A pair in $\Delta(h)$ has the form $\{p, q\}$ with $p \in \mathrm{supp}(h)$ and q in the hole of h. The holes are disjoint, so $q \in \mathrm{supp}(g)$, and the pair lies outside $\Delta(g)$ exactly when $p \in \mathrm{supp}(g)$. Distinct points of $S_n \cap \mathrm{supp}(g)$ come from distinct such pairs. □

Three impossibility results, Theorem 6.9(b) and the necessity of (II) in Theorem 6.10 below, and Theorem B.2 above, use one adversary. It builds a presentation in rounds: in each round it first follows a stream on which the learner is forced into a wrong answer, and then it adds a little progress towards the presentation it finally wants. We isolate this construction once. An enumeration of a finite nonempty set repeats its points, so that every enumeration is an infinite sequence.

Lemma C.4 (Diagonal construction)**.** *Let $\mathcal{S}$ be a set of finite sequences of pairs, let E be a property of finite sequences of pairs, and let $Y \subseteq \mathcal{X}$ be nonempty. Assume that every $\sigma \in \mathcal{S}$ has an infinite extension τ_σ such that*

(i) *every sufficiently long prefix of τ_σ has property E, and*

(ii) *for every prefix ρ of τ_σ that extends σ and every $y \in Y$, there is a finite sequence β of pairs such that $\rho\beta \in \mathcal{S}$ and some pair of β contains y.*

Then every $\sigma_0 \in \mathcal{S}$ extends to an infinite sequence P of pairs such that every finite prefix of P is a prefix of a member of $\mathcal{S}$, every point of Y lies in a pair of P, and infinitely many prefixes of P have property E.

Proof. Enumerate Y as $y_1, y_2, \dots$. Given $\sigma_i \in \mathcal{S}$, use (i) to choose a prefix ρ_i of τ_{σ_i} that is longer than σ_i and has property E, and use (ii) to choose β_i such that $\sigma_{i+1} := \rho_i\beta_i$ lies in $\mathcal{S}$ and some pair of β_i contains y_{i+1}. The sequences $\sigma_0 \sqsubset \rho_0 \sqsubseteq \sigma_1 \sqsubset \rho_1 \sqsubseteq \cdots$ strictly increase, so they have an infinite limit P. Every finite prefix of P is a prefix of some $\sigma_i \in \mathcal{S}$, every y_{i+1} lies in a pair of β_i, and every ρ_i is a prefix of P with property E. □

In each application, $\mathcal{S}$ consists of the streams that the adversary can still extend to a presentation of the hypothesis it finally presents, τ_σ is a presentation of a rival, and E says that the learner's answer is wrong for the hypothesis finally presented.

C.2 Two hypotheses

Theorem 6.9 (Two hypotheses)**.** *Let $h \ne g$ be proper nontrivial, and let $\bar\kappa(h,g) := \max\{\kappa(h \to g), \kappa(g \to h)\}$.*

(a) *For $k \in \mathbb{N}$ the following are equivalent:* (i) $\{h,g\} \in k\text{-}\mathsf{CtrId}$*;* (ii) *no sequence is a k-corrupted contrastive presentation of both h and g;* (iii) $\bar\kappa(h,g) \ge 2k+1$*. Hence the least k for which h and g admit a common k-corrupted contrastive presentation is $\rho(h,g) = \lceil \bar\kappa(h,g)/2 \rceil$, and no finite k suffices when $\bar\kappa(h,g) = \infty$.*

(b) $\{h,g\} \in \mathrm{Fin}\text{-}\mathsf{CtrId}$ *if and only if $\kappa(h \to g) = \kappa(g \to h) = \infty$, that is, if and only if h and g cover $\mathcal{X}$ and share infinitely many points.*

Proof. (a), (i) implies (ii). If one sequence is a k-corrupted presentation of both h and g, an identifier produces one sequence of outputs on it, and these outputs cannot converge both to h and to g.

(a), (ii) implies (iii). Assume $\bar{\kappa}(h, g) \leq 2k$; we build a common k-corrupted presentation. Let $J := \mathrm{Def}(h, g)$. By Lemma C.1, J is empty, or $J = A$ and $D = \varnothing$, or $J = B$ and $C = \varnothing$, or $J = C$ and $B = \varnothing$. Each $x \in J$ lies in a pair e_x^h that h rejects and g accepts, and in a pair e_x^g that g rejects and h accepts: one may take

- $e_x^h = \{x, b\}$ with $b \in B$ and $e_x^g = \{x, c\}$ with $c \in C$ if $J = A$;
- $e_x^h = \{x, a\}$ with $a \in A$ and $e_x^g = \{x, d\}$ with $d \in D$ if $J = B$;
- $e_x^h = \{x, d\}$ with $d \in D$ and $e_x^g = \{x, a\}$ with $a \in A$ if $J = C$.

The cells used are nonempty: by Lemma C.1(b) at most one of B, C, D is empty, and $A = \mathrm{supp}(g)$ when $C = \varnothing$ and $A = \mathrm{supp}(h)$ when $B = \varnothing$. Split J into disjoint sets J_h, J_g with at most k points each. Let P list the pairs e_x^h for $x \in J_h$, then the pairs e_x^g for $x \in J_g$, and then, along an enumeration of the remaining points of $\mathrm{supp}(h) \cup \mathrm{supp}(g)$, one pair of $\Gamma(h, g)$ through each of them; these points lie in $\mathrm{V}(\Gamma(h, g))$ by definition of J, and they form a nonempty set because J is a single cell. Then P covers both supports, h rejects only the at most k pairs e_x^h, and g rejects only the at most k pairs e_x^g.

(a), (iii) implies (i). By Lemma C.1, $\bar{\kappa}(h, g) = \max\{\kappa(h \to g), \kappa(g \to h)\}$; by symmetry let $\kappa(h \to g) \geq 2k + 1$. The identifier outputs h as soon as $\mathrm{viol}_g(P_{\leq n}) \geq k + 1$, and g before that. On a k-corrupted presentation of g the count never exceeds k, so the output is always g. On a k-corrupted presentation of h, Lemma C.2 gives $\mathrm{viol}_g(P_{\leq n}) \geq k + 1$ for all large n, so the output is eventually always h. By the equivalence of (ii) and (iii), the least budget with a common presentation is the least k with $\bar{\kappa}(h, g) \leq 2k$, namely $\lceil \bar{\kappa}(h, g)/2 \rceil$, and there is none if $\bar{\kappa}(h, g) = \infty$.

(b). By Lemma C.1(c), both defect sets are nonempty only when h and g cover $\mathcal{X}$, and then both equal $A = \mathrm{supp}(h) \cap \mathrm{supp}(g)$; this gives the second form of the statement. If both defects are infinite, the identifier that outputs h when $\mathrm{viol}_h(P_{\leq n}) \leq \mathrm{viol}_g(P_{\leq n})$, and g otherwise, succeeds without knowing the budget: on a k_0-corrupted presentation of h the count viol_h stays at most k_0 while $\mathrm{viol}_g \to \infty$ by Lemma C.2, and the case of target g is symmetric. Conversely, let one identifier I succeed on $\{h, g\}$ for every finite budget. Part (a) at every k gives $\bar{\kappa}(h, g) = \infty$. If the supports are not nested, Lemma C.1(c) then shows that h and g cover $\mathcal{X}$ and that both defects are infinite. It remains to exclude nested supports. By symmetry suppose $\mathrm{supp}(h) \subsetneq \mathrm{supp}(g)$, so that $C = \mathrm{supp}(g) \setminus \mathrm{supp}(h)$, and fix $d \in D$, which exists because g is proper. A pair $\{a, d\}$ with $a \in \mathrm{supp}(h)$ crosses both cuts, and a pair $\{c, d\}$ with $c \in C$ crosses the cut of g but not that of h. Apply Lemma C.4 to the set $\mathcal{S}$ of all finite sequences of such pairs, with $Y := \mathrm{supp}(g)$ and with property E: "I outputs h". For $\sigma \in \mathcal{S}$, let τ_σ be σ followed by the pairs $\{a, d\}$ along an enumeration of $\mathrm{supp}(h)$. It is a presentation of h whose only corrupted pairs are the finitely many pairs $\{c, d\}$ of σ, so I outputs h on all its long prefixes and (i) holds. Every prefix of τ_σ consists of such pairs, and so does its extension by one pair $\{y, d\}$ with $y \in \mathrm{supp}(g)$, so (ii) holds. Starting from the empty sequence, the lemma gives a sequence of such pairs that covers $\mathrm{supp}(g)$: a clean presentation of g on which I outputs h infinitely often, a contradiction. □

C.3 Identification at every budget

Theorem 6.10 (Identification at every budget)**.** *Let $\mathcal{H}$ be a countable class of proper nontrivial hypotheses and $k \in \mathbb{N} \cup \{\mathrm{Fin}\}$. Then $\mathcal{H} \in k$-CtrId if and only if* (I) *any two members of $\mathcal{H}$ with incomparable supports cover $\mathcal{X}$ and share at least θ_k points, and* (II) *every $g \in \mathcal{H}$ has a finite set $T_g \subseteq \mathrm{supp}(g)$ such that every $f \in \mathcal{H}$ with $\mathrm{supp}(f) \subsetneq \mathrm{supp}(g)$ misses at least θ_k points of T_g.*

Proof. We prove the necessity of (I), the necessity of (II), and sufficiency for finite k and for $k = \mathrm{Fin}$, in this order. We say that f lies *below* g if $\mathrm{supp}(f) \subsetneq \mathrm{supp}(g)$. A learner for a class also succeeds on every subclass at the same budget.

Necessity of (I). Let $h, g \in \mathcal{H}$ have incomparable supports, so that $B, C \neq \varnothing$. By Lemma C.1(c), $\bar{\kappa}(h, g) > 0$ only if h and g cover $\mathcal{X}$, and then $\bar{\kappa}(h, g) = |\mathrm{supp}(h) \cap \mathrm{supp}(g)|$. So Theorem 6.9 applied to $\{h, g\}$ yields (I): at least $2k + 1$ common points for finite k, and infinitely many for $k = \mathrm{Fin}$.

Necessity of (II) for $k = \mathrm{Fin}$. By Theorem 6.9(b), any two members cover $\mathcal{X}$. Nested supports never cover $\mathcal{X}$, because the larger support would be $\mathcal{X}$. So no member lies below another, and (II) holds with $T_g = \varnothing$.

Necessity of (II) for finite k. Let I identify $\mathcal{H}$ at budget k, and suppose that some $g \in \mathcal{H}$ has no set T_g as in (II). Then for every finite $T \subseteq \mathrm{supp}(g)$ some member below g misses at most $2k$ points of T. We first make this failure uniform: there are a set $L \subseteq \mathrm{supp}(g)$ with at most $2k$ points, finite sets $T_1 \subseteq T_2 \subseteq \cdots$ with union $\mathrm{supp}(g)$ and $L \subseteq T_1$, and members $f_1, f_2, \ldots$ below g with

$$T_r \setminus \mathrm{supp}(f_r) = L \qquad \text{for every } r. \tag{1}$$

To find them, let $U_1 \subseteq U_2 \subseteq \cdots$ be finite sets with union $\mathrm{supp}(g)$, and choose for each s a member f'_s below g for which $\beta_s := U_s \setminus \mathrm{supp}(f'_s)$ has at most $2k$ points. For fixed j, the sets $\beta_s \cap U_j$ with $s \geq j$ are subsets of U_j with at most $2k$ points, so they take finitely many values. Repeated use of the pigeonhole principle gives infinite sets $S_1 \supseteq S_2 \supseteq \cdots$ and sets L_j with $\beta_s \cap U_j = L_j$ for all $s \in S_j$ with $s \geq j$. Then $L_i = L_j \cap U_i$ for $i \leq j$, so $L := \bigcup_j L_j$ has at most $2k$ points and $L \cap U_j = L_j$. Choose j_0 with $L \subseteq U_{j_0}$ and, for each r, an index $s_r \in S_{j_0+r}$ with $s_r \geq j_0 + r$. The sets $T_r := U_{j_0+r}$ and the members $f_r := f'_{s_r}$ satisfy (1), because $T_r \subseteq U_{s_r}$ gives $T_r \setminus \mathrm{supp}(f_r) = \beta_{s_r} \cap T_r = L_{j_0+r} = L$.

Now fix $d \notin \mathrm{supp}(g)$, which exists because g is proper, and a point $p_0 \in \mathrm{supp}(g) \setminus L$, which exists because this set contains the nonempty set $\mathrm{supp}(f_1)$. Dropping finitely many indices r, we may assume $p_0 \in T_1$, so that $p_0 \in \mathrm{supp}(f_r)$ for all r by (1). Split L into sets L', L'' with at most k points each, and let σ_0 list the pairs $\{x, d\}$ for $x \in L'$ and $\{x, p_0\}$ for $x \in L''$. By (1), every $x \in L$ has $g(x) = 1$ and $f_r(x) = 0$; moreover $g(d) = f_r(d) = 0$ and $g(p_0) = f_r(p_0) = 1$. So g rejects only the pairs $\{x, p_0\}$ of σ_0, and each f_r rejects only the pairs $\{x, d\}$; in both cases at most k pairs.

Let $\mathcal{S}$ be the set of sequences that consist of σ_0 followed by finitely many pairs $\{q, d\}$ with $q \in \mathrm{supp}(g) \setminus L$. Apply Lemma C.4 to $\mathcal{S}$, with $Y := \mathrm{supp}(g) \setminus L$ and with property E: "I outputs a hypothesis other than g". For $\sigma \in \mathcal{S}$, choose r so large that T_r contains every point q of a pair $\{q, d\}$ that σ adds to σ_0; these points then lie in $\mathrm{supp}(f_r)$ by (1). Let τ_σ be σ followed by the pairs $\{a, d\}$ along an enumeration of $\mathrm{supp}(f_r)$. Its only pairs outside $\Delta(f_r)$ lie in σ_0, and it covers $\mathrm{supp}(f_r)$, so it is a k-corrupted presentation of $f_r \neq g$, and (i) holds. Every prefix of τ_σ that extends σ lies in $\mathcal{S}$, because $\mathrm{supp}(f_r) \subseteq \mathrm{supp}(g) \setminus L$, and appending the pair $\{y, d\}$ gives (ii). The lemma yields a sequence whose pairs outside $\Delta(g)$ all lie in σ_0 and which covers L and $\mathrm{supp}(g) \setminus L$. It is a k-corrupted presentation of g on which I outputs another hypothesis infinitely often, a contradiction.

Sufficiency for finite k. Enumerate $\mathcal{H} = \{h_0, h_1, \ldots\}$, and call h_i *eligible* at time n if $T_{h_i} \subseteq \mathrm{Seen}_n(P)$ and $\mathrm{viol}_{h_i}(P_{\leq n}) \leq k$. The identifier outputs the eligible hypothesis of least index, and h_0 if none is eligible. Fix a target $h = h_i$ and a k-corrupted presentation P of h. The target is eligible from some time on: its violation count never exceeds k, and the finite set $T_h \subseteq \mathrm{supp}(h)$ eventually appears. It therefore suffices to show that each of the finitely many h_j with $j < i$ is eventually never eligible. There are two cases.

If $\mathrm{supp}(h) \not\subseteq \mathrm{supp}(h_j)$, then $\kappa(h \to h_j) \geq 2k + 1$. Indeed, if h_j lies below h, then $D^+_{h \to h_j} = \mathrm{supp}(h) \setminus \mathrm{supp}(h_j)$ by Lemma C.1(c), and this set contains at least $2k + 1$ points of T_h by (II). Otherwise the supports are incomparable, and by (I) and Lemma C.1(c) the defect is $\mathrm{supp}(h) \cap \mathrm{supp}(h_j)$, with at least $2k + 1$ points. Now Lemma C.2 gives $\mathrm{viol}_{h_j}(P_{\leq n}) \geq k + 1$ for all large n, and h_j is not eligible from then on.

If $\mathrm{supp}(h) \subsetneq \mathrm{supp}(h_j)$, then h_j is never eligible. Let $C_j := \mathrm{supp}(h_j) \setminus \mathrm{supp}(h)$. By (II) for h_j and $f = h$, at least $2k + 1$ points of T_{h_j} lie in C_j, so at a time at which h_j is eligible the stream has shown at least $2k + 1$ points of C_j. But an honest pair contains exactly one negative of h, so at most one point of C_j; if it contains one, its other endpoint is a positive of h, hence of h_j, and h_j rejects the pair. A corrupted pair contains at most two points of C_j, and if it contains two, h_j rejects it too. Let a be the number of honest pairs in $P_{\leq n}$ that meet C_j, and let b_1 and b_2 be the numbers of corrupted pairs with exactly one and with two points in C_j. The number of points of C_j in $\mathrm{Seen}_n(P)$ is then at most

$$a + b_1 + 2b_2 = (a + b_2) + (b_1 + b_2) \;\leq\; \mathrm{viol}_{h_j}(P_{\leq n}) + k, \tag{2}$$

which is at most $2k$ at any time at which h_j is eligible, a contradiction.

Sufficiency for $k = \mathrm{Fin}$. Now the threshold in (II) is infinite, and a finite set T_g cannot miss infinitely many points, so (II) says that no member lies below another: the supports are pairwise incomparable.

By (I), any two members then cover $\mathcal{X}$ and share infinitely many points; in particular the holes are pairwise disjoint. The identifier outputs the least-index hypothesis with the fewest violations, which exists because violation counts are natural numbers. Fix a target h, a budget k_0, and a k_0-corrupted presentation P of h, and let S_n be the set of positives of h that lie in an honest pair among $p_1, \dots, p_n$. By Lemma C.3,

$$\mathrm{viol}_g(P_{\le n}) \;\ge\; |S_n \cap \mathrm{supp}(g)| \qquad \text{for every } g \neq h. \tag{3}$$

Every positive of h lies in a pair of P, and a positive that lies in no honest pair lies in one of the at most k_0 corrupted pairs, so there are at most $2k_0$ such positives. Thus the sets S_n exhaust $\mathrm{supp}(h)$ up to at most $2k_0$ points, and $|S_n \cap \mathrm{supp}(g)| \to \infty$ for every $g \neq h$, because $\mathrm{supp}(h) \cap \mathrm{supp}(g)$ is infinite. In particular $|S_n| \to \infty$. Call a rival g *lagging* at time n if $|S_n \cap \mathrm{supp}(g)| \le k_0$. A lagging rival has at least $|S_n| - k_0$ points of S_n in its hole, and the holes are disjoint, so two lagging rivals would give $|S_n| \ge 2(|S_n| - k_0)$. Thus at most one rival lags once $|S_n| \ge 2k_0 + 1$. A rival that stops lagging never lags again, because S_n only grows; so from some time on at most one fixed rival can lag, and eventually it stops lagging as well. From then on every rival has more than $k_0 \ge \mathrm{viol}_h(P_{\le n})$ violations by (3), and the identifier outputs h. It never uses k_0. □

Condition (II) cannot be dropped, even when every pair passes the threshold of Theorem 6.9.

Example C.5 (Pairwise thresholds do not suffice)**.** Let $Y \subseteq \mathcal{X}$ be infinite with $\mathcal{X} \setminus Y$ infinite, fix two points $\ell_1, \ell_2 \in Y$, list $Y \setminus \{\ell_1, \ell_2\} = \{y_1, y_2, \dots\}$, and let $\mathrm{supp}(g) = Y$ and $\mathrm{supp}(f_j) = \{y_1, \dots, y_{3j}\}$ for $j \ge 1$. All supports are nested and any two differ in at least 3 points, so every pair has $\bar{\kappa} \ge 3$ and is identifiable at budget 1 by Theorem 6.9. The class lies in CtrId, with tell-tale $T_g = \{\ell_1\}$ and $T_{f_j} = \mathrm{supp}(f_j)$. But it is not in 1-CtrId: every finite $T \subseteq Y$ lies, up to the two points ℓ_1, ℓ_2, inside some $\mathrm{supp}(f_j)$, so (II) fails for g. Concretely, the pairs $\{\ell_1, d\}$ and $\{\ell_2, y_1\}$ with $d \notin Y$ cost one violation for g and one for every f_j, and after them the adversary can follow a presentation of g that looks like a presentation of f_j for larger and larger j.

Corollary 6.11 (Unknown budgets)**.** *Let $\mathcal{H}$ be a countable class of proper nontrivial hypotheses with at least two members. Then $\mathcal{H} \in \mathrm{Fin}$-$\mathsf{CtrId}$ if and only if the holes $\mathcal{X} \setminus \mathrm{supp}(h)$, $h \in \mathcal{H}$, are pairwise disjoint and the union of any two holes misses infinitely many points. If, moreover, two members have holes with at most r points, then $\mathcal{H} \notin r$-TxtId.*

Proof. Two members cover $\mathcal{X}$ exactly when their holes are disjoint, and they share infinitely many points exactly when the union of their holes misses infinitely many points. Two nonempty disjoint holes are incomparable, so no support then lies below another. Hence the stated conditions are exactly conditions (I) and (II) of Theorem 6.10 for $k = \mathrm{Fin}$. For the last claim, the enumeration of $\mathcal{X}$ is an r-corrupted text for both members with holes of at most r points, since it lists each point of a hole once; a learner sees the same input under both targets. □

Corollary 6.12 (Budget hierarchy)**.** $\mathsf{CtrId} = 0$-$\mathsf{CtrId} \supsetneq 1$-$\mathsf{CtrId} \supsetneq 2$-$\mathsf{CtrId} \supsetneq \cdots \supsetneq \bigcap_{k \in \mathbb{N}} k$-$\mathsf{CtrId} \supsetneq \mathrm{Fin}$-$\mathsf{CtrId}$.

Proof. A k-corrupted presentation is $(k+1)$-corrupted, so a learner for budget $k+1$ also works for budget k, and a learner for unknown finite budgets works for every k; this gives the inclusions, and 0-$\mathsf{CtrId} = \mathsf{CtrId}$ by definition. For k-$\mathsf{CtrId} \neq (k+1)$-$\mathsf{CtrId}$, take h, g that cover $\mathcal{X}$ and share exactly $2k+1$ points (split $\mathcal{X}$ into a set of $2k+1$ points and two infinite sets). Then $\bar{\kappa}(h, g) = 2k + 1$ by Lemma C.1, so $\{h, g\}$ lies in k-CtrId but not in $(k+1)$-CtrId by Theorem 6.9(a). The last inclusion is strict by the example after Theorem 6.9: there $\mathrm{supp}(h) \subsetneq \mathrm{supp}(g)$ with an infinite difference, so $\bar{\kappa}(h, g) = \infty$ but $\kappa(h \to g) = 0$. □

C.4 Growing corruption

We next let the number of corrupted pairs grow. For a stream P of pairs write $k_n^h(P) := \mathrm{viol}_h(P_{\le n}) = |\{t \le n : p_t \notin \Delta(h)\}|$. For a nondecreasing $\varphi : \mathbb{N} \to \mathbb{N}$, call P a *φ-corrupted presentation* of h if its honest pairs cover $\mathrm{supp}(h)$ and $k_n^h(P) \le \varphi(n)$ for every n, and a *sublinearly corrupted presentation* of h if its honest pairs cover $\mathrm{supp}(h)$ and $k_n^h(P) = o(n)$. A bounded $\varphi \equiv k$ gives k-corrupted presentations whose honest pairs already cover $\mathrm{supp}(h)$, a smaller set than in Definition 6.1.

Proposition C.6 (No unbounded budget)**.** *Let φ be nondecreasing and unbounded. Then any two distinct proper nontrivial hypotheses h, g have a common φ-corrupted presentation. Hence a class of proper nontrivial hypotheses can be identified from all φ-corrupted presentations of its members if and only if it has at most one member. The same holds for sublinearly corrupted presentations, and for presentations with $k_n^h(P) \le \varepsilon n$ for all n, for any fixed $\varepsilon > 0$.*

Proof. Fix $e \in \Gamma(h, g)$ (Lemma C.1(a)) and clean presentations (a_i) of h and (b_i) of g. Choose positions $t_1 < t_2 < \cdots$ with $\varphi(t_j) \ge j$, which exist because φ is unbounded. Put $a_1, b_1, a_2, b_2, \ldots$ at the positions $t_1, t_2, t_3, \ldots$, and put e at every other position. A pair at a position other than the t_j is e, which is honest for both hypotheses. If $t_j \le n$, then $j \le \varphi(t_j) \le \varphi(n)$, so at most $\varphi(n)$ of the first n pairs lie at positions t_j, and only these can be corrupted. The honest pairs a_i cover $\mathrm{supp}(h)$, and the honest pairs b_i cover $\mathrm{supp}(g)$. So the stream is a φ-corrupted presentation of both, and no identifier separates h from g; the converse is trivial. For the last claim take any unbounded φ with $\varphi(n) \le \varepsilon n$ for all n (for instance $\varphi(n) = \lfloor \varepsilon n \rfloor$) and with $\varphi(n) = o(n)$ for the sublinear statement (for instance $\varphi(n) = \lfloor \log_2(n+1) \rfloor$, which is at most εn for all large n): every φ-corrupted presentation is then of the required kind. □

Elapsed time is therefore not a useful progress measure when the adversary may repeat uninformative pairs. Measuring corruption against honest coverage instead gives a positive result. Let $c_h := \sup_{g \in \mathcal{H} \setminus \{h\}} |\mathrm{supp}(h) \cap \mathrm{V}(\Gamma(h, g))|$, and call $\mathcal{H}$ *bounded-overlap* if $c_h < \infty$ for every $h \in \mathcal{H}$; the co-singleton class has $c_h = 1$.

Proposition C.7 (Least-violation identification)**.** *Let $\mathcal{H}$ be a countable bounded-overlap class with a fixed enumeration, and let the identifier output the least-index minimizer of $\mathrm{viol}_g(n) := |\{t \le n : p_t \notin \Delta(g)\}|$ over $g \in \mathcal{H}$. Let P be any stream of pairs, $h \in \mathcal{H}$, let m_n be the number of distinct h-positives in honest pairs among $p_1, \ldots, p_n$, and let $k_n := k_n^h(P)$. If $m_n - k_n \to \infty$, the identifier converges to h on P. In particular, every countable bounded-overlap class of infinite proper nontrivial hypotheses lies in* Fin-CtrId.

Proof. The target has $\mathrm{viol}_h(n) = k_n$. Fix $g \ne h$. Each h-positive x outside $\mathrm{V}(\Gamma(h, g))$ that appears in an honest pair up to time n has a first such pair, which lies in $\Delta(h) \setminus \Gamma(h, g)$ and so violates g; distinct such x give distinct pairs, because an honest pair has exactly one h-positive endpoint. Hence $\mathrm{viol}_g(n) \ge m_n - c_h$ for every $g \ne h$, and once $m_n - k_n > c_h$ the target is the unique minimizer. For a k_0-corrupted presentation, $k_n \le k_0$, and since at most $2k_0$ points appear only in corrupted pairs, $m_n \to \infty$; the rule does not use k_0. □

For infinite supports, $\mathrm{supp}(h) \cap \mathrm{V}(\Gamma(h, g))$ is the hole of g when h and g cover $\mathcal{X}$, and it is infinite otherwise (Lemma C.1(c)). So a bounded-overlap class with at least two members, all of infinite support, has pairwise disjoint holes of uniformly bounded size, and its membership in Fin-CtrId is also a case of Corollary 6.11.

D Membership Patterns and Higher-Order Shared Presentations

The shared-presentation criterion can be restated as a finite combinatorial condition on membership patterns, which is convenient for higher-order obstructions.

Proposition D.1 (Membership-pattern criterion)**.** *For $\mathcal{F} = \{h_1, \ldots, h_r\}$ of proper nontrivial hypotheses and $\alpha \in \{0,1\}^r$, define $R_\alpha := \{x \in \mathcal{X} : (h_i(x))_i = \alpha\}$. Then $\mathcal{F}$ admits a shared contrastive presentation iff*

$$R_\alpha \ne \varnothing \text{ and } \alpha \ne \mathbf{0} \quad \Longrightarrow \quad R_{\mathbf{1}-\alpha} \ne \varnothing. \tag{4}$$

Proof. *Necessity.* For $x \in R_\alpha$ with $\alpha \ne \mathbf{0}$, x belongs to some target support; positive-side coverage forces x to appear in a pair $\{x, y\}$ lying in $\Delta(h_i)$ for every i, so the pattern of y must equal $\mathbf{1} - \alpha$.

Sufficiency. For each $x \in \bigcup_i \mathrm{supp}(h_i)$, pick a witness y_x in the complementary cell (existing by hypothesis). Enumerate (or list-and-repeat) the union and emit the chosen pairs; each pair has patterns summing to $\mathbf{1}$, so it lies in $\Delta(h_i)$ for every i. □

Example D.2 (A higher-order obstruction)**.** Partition $\mathcal{X}$ into six infinite cells with patterns $100, 010, 001$, $110, 101, 011$, with $R_{000} = R_{111} = \varnothing$. Each pairwise intersection $\mathrm{supp}(h_i) \cap \mathrm{supp}(h_j)$ is infinite, but $\mathrm{supp}(h_1) \cap \mathrm{supp}(h_2) \cap \mathrm{supp}(h_3) = R_{111} = \varnothing$. The realized nonzero patterns occur in complementary pairs ($100 \leftrightarrow 011$, $010 \leftrightarrow 101$, $001 \leftrightarrow 110$), so (4) holds and $\{h_1, h_2, h_3\}$ admits a shared contrastive presentation. Proposition 5.12 then rules out contrastive generation. Pairwise analysis is therefore insufficient for CtrGen.

E Discussion and Extensions

This section outlines several natural extensions of the contrastive learning framework that lie beyond the present paper's scope. We give formal definitions where appropriate, articulate the structural difficulties that prevent direct transfer of our techniques, and pose open questions to seed future work.

E.1 Corrupted contrastive generation

The robustness analysis of Section 6 concerns identification. The parallel question for generation is the following.

Definition E.1 (k-corrupted contrastive generator)**.** For $k \geq 0$, a generator G is a *k-corrupted contrastive generator* for $\mathcal{H}$ if for every $h \in \mathcal{H}$ and every k-corrupted contrastive presentation P for h (Definition 6.1), there exists N such that $G(P_{\leq n}) \in \mathrm{supp}(h) \setminus \mathrm{Seen}_n(P)$ for all $n \geq N$. Write $\mathcal{H} \in k\text{-}\mathsf{CtrGen}$ when such a generator exists, and $\mathrm{Fin}\text{-}\mathsf{CtrGen}$ when a single G succeeds for every finite corruption budget.

Whether the closure-dimensional characterization (Theorem 5.4) extends to the corrupted regime is open. The basic difficulty is structural: closure-based generation rests on the edge-induced version space $\mathcal{H}_\Delta(E)$, and a single corrupted pair $p^* \notin \Delta(h)$ can eject the true target h from this version space. The remaining hypotheses in $\mathcal{H}_\Delta(E_n(P))$ need not have supports contained in $\mathrm{supp}(h)$, so the closure $\langle E_n(P)\rangle^{\Delta}_{\mathcal{H}}$ can leak outside $\mathrm{supp}(h)$, and the closure-based generator can output a non-positive of h. By contrast, the identification reversal of Section 6 exploits an incidence invariant (the defect number) whose redundancy is preserved under finitely many inserted edges; the closure operator has no such redundancy.

Remark E.2 (Identify-then-generate)**.** The co-singleton class lies in $\mathrm{Fin}\text{-}\mathsf{CtrGen}$: Algorithm 1 identifies the unique negative s^* in finite time, after which any unseen $x \neq s^*$ is a novel positive of the target. This "identify-then-generate" template lifts $\mathrm{Fin}\text{-}\mathsf{CtrId}$ to $\mathrm{Fin}\text{-}\mathsf{CtrGen}$ on classes with infinite supports: output the least point of the identified support that has been neither seen nor output. By Corollary 6.11, every class with at least two members whose holes are pairwise disjoint and pairwise leave infinitely many common positives therefore lies in $\mathrm{Fin}\text{-}\mathsf{CtrGen}$. The converse direction (whether classes in $\mathsf{CtrGen} \setminus \mathsf{CtrId}$ can achieve $\mathrm{Fin}\text{-}\mathsf{CtrGen}$) may require new tools, since it cannot route through identification.

A natural target is a *robust closure dimension* $\mathrm{C}^{(k)}_\Delta(\mathcal{H})$ measuring how far the closure can be pushed outside the target's support by k adversarially inserted pairs; a quantitative theorem in the spirit of Theorem 5.4 would then express $k\text{-}\mathsf{CtrGen}$ in terms of $\mathrm{C}^{(k)}_\Delta$.

E.2 Statistical contrastive presentations

The presentations of Sections 3 and 6 are adversarial: a contrastive presentation is any sequence satisfying XOR and positive-side coverage, and corruption is adversarial. A natural statistical relaxation samples pairs i.i.d. from a distribution over $[\mathcal{X}]^2$ supported on $\Delta(h)$.

Definition E.3 (μ-random contrastive presentation)**.** Fix a target h and a probability distribution μ on $[\mathcal{X}]^2$ with $\mathrm{supp}(\mu) \subseteq \Delta(h)$. A *$\mu$-random contrastive presentation* is a sequence $(p_t)_{t\geq 1} \overset{\text{i.i.d.}}{\sim} \mu$. Positive-side coverage holds almost surely iff every $x \in \mathrm{supp}(h)$ lies in some pair in $\mathrm{supp}(\mu)$.

In this regime, the observed edge set $E_n(P)$ becomes a random subgraph of $\Delta(h)$. For natural pair distributions (e.g., uniform over a finite edge set, or a product distribution on the unknown bipartition), the induced random graph is a bipartite Erdős–Rényi-type model, conditioned on staying

inside the cut $(\mathrm{supp}(h), \mathcal{X} \setminus \mathrm{supp}(h))$. This connects CtrId and CtrGen to community-detection problems on stochastic block models: recovering the cut from random crossing edges is structurally analogous to recovering the planted bipartition. We expect phase-transition phenomena: a critical edge density below which contrastive identification is statistically impossible, and above which spectral or message-passing algorithms succeed. The development of such a statistical theory is conceptually orthogonal to the asymptotic limit-learning paradigm of the present paper but sits naturally within its geometric framework.

E.3 Effective procedures

The contrastive identifier of Theorem 4.7 and the non-uniform generator of Theorem 5.5 are information-theoretic: they consume the Angluin tell-tale family $\{T_g\}_{g\in\mathcal{H}}$ and the per-level dimensions $\{\mathrm{C}_\Delta(\mathcal{H}_m)\}_m$ as oracle inputs. Whether these constructions can be made effective depends on the available oracles and the computability of the input class.

Oracle hierarchy. Three natural oracles on a hypothesis class are:

- a *consistency* oracle returning whether $E \subseteq \Delta(g)$ for given finite $E \subseteq [\mathcal{X}]^2$ and $g \in \mathcal{H}$;
- a *closure-membership* oracle returning whether $x \in \langle E\rangle^{\Delta}_{\mathcal{H}}$ for given $x \in \mathcal{X}$ and finite E;
- an *ERM* oracle returning some $g \in \mathcal{H}_\Delta(E)$ if such g exists, else $\perp$.

The closure-based generator of Theorem 5.4 requires both the ERM oracle (to detect $\mathcal{H}_\Delta(E) \neq \varnothing$) and the closure-membership oracle (to enumerate $\langle E\rangle^{\Delta}_{\mathcal{H}} \setminus \mathrm{V}(E)$). The eligibility-based identifier of Theorem 4.7 additionally requires the tell-tale family $\{T_g\}$, which Theorem 3.5 asserts to exist but does not construct; so do the robust tell-tale sets of Theorem 6.10 for finite k.

By contrast, our absence-count algorithm is fully constructive: given the contrastive prefix as a finite list of pairs, the absence count $a_n(x)$ is a primitive computable function of the input, and the minimization is over the finite set $\mathrm{Seen}_n(P)$. No oracle on $\mathcal{H}$ is needed, unlike for the general CtrId identifier. The least-violation identifier of Theorem 6.10 for unknown budgets needs only the violation counts, but it minimizes over the whole class. Its proof shows that from some time on every rival has more violations than the target; hence the truncated rule that, at time n, outputs the least-index hypothesis with the fewest violations among the first n hypotheses also converges, and for an indexed family with decidable membership it is computable.

A natural target for future work is to identify combinatorial conditions on $\mathcal{H}$ under which the eligibility-based identifier becomes effective from a closure-membership or ERM oracle alone, without requiring the tell-tale family as a separate input.

F Additional Related Work

F.1 Identification in the limit

The classical paradigm originates with Gold [1967], with Angluin [1980]'s tell-tale theorem giving the canonical positive-data characterization. Earlier work of Angluin [1979] introduces the framework of pattern languages, a concrete subclass that admits positive-data identification despite the negative results of Gold [1967] on broader classes; this work foreshadows much of the structural analysis underlying tell-tale conditions. Wharton [1974] considers an *approximate* variant of identification in which the learner is permitted small deviations from the target language, an early precursor to noise- and corruption-tolerant limit learning. The detailed survey is Lange et al. [2008].

A line of recent work studies relaxed criteria for identification. Charikar et al. [2026] introduce a list-identification model in which the learner is allowed to output a small list of candidate languages, succeeding if the true target appears on the list, and they fully characterize list-identifiability. Papazov and Flammarion [2025] characterize limit-learnability of recursive functions when the learner observes evaluations on every domain point. Peng et al. [2026] augment Gold's model with computational traces of the accepting machine and obtain identifiability across the Chomsky hierarchy with varying corruption tolerance, providing a complementary mechanism for circumventing Gold's negative results.

F.2 Generation in the limit

Kleinberg and Mullainathan [2024] introduced generation in the limit, proving its universality on countable UUS classes; Li et al. [2025] reformulated the model in learning-theoretic notation and introduced the closure dimension that exactly characterizes uniform generation. The closure dimension of Li et al. [2025] is the direct positive-data ancestor of our contrastive closure dimension; we recover the same formalism with finite positive samples replaced by finite sets of pair constraints.

A recent thread examines refinements of the generation criterion. Kalavasis et al. [2026, 2025] characterize generation under various breadth constraints and study trade-offs between hallucination and mode collapse. Peale et al. [2025] introduce representative generation, requiring the generator to cover meaningful sub-collections of the target rather than producing arbitrary novel positives. Charikar and Pabbaraju [2025, 2026] explore facets of language generation in the limit and Pareto-optimal trade-offs in non-uniform generation. Kleinberg and Wei [2026a, 2025] introduce density measures and partial-enumeration variants, providing fine-grained analyses across the space of possible enumeration orderings. Arenas et al. [2026] establish complexity barriers separating different generation modes and draw implications for learning. Kleinberg and Wei [2026b] study Banach density, which measures the breadth of language generation in the limit when strings live in a d-dimensional embedding.

A second thread targets noise tolerance. Raman and Raman [2025] analyze generation from noisy examples, Bai et al. [2026] study generation under noise, loss, and feedback, and Mehrotra et al. [2026a] push noise tolerance to infinite contamination budgets. Li and Zhang [2026] propose quantitative measures of noise in language generation, complementing the qualitative noise-tolerance results. Racca et al. [2026] study generation in a replay-based model that captures forms of model collapse.

A third thread addresses semantic and structural extensions. Hanneke et al. [2025] analyze union closure properties of generation, showing that countable closure can fail. Karbasi et al. [2025] study the (im)possibility of automated hallucination detection. Li et al. [2026a] extend generation to metric spaces, and Høgsgaard and Pabbaraju [2026] introduce agnostic notions of identification and generation. Anastasopoulos et al. [2026] formalize a setting of safe language generation in the limit. None of these works addresses the undirected-pair signal structure we study, but each contributes orthogonally to the broader generation landscape.

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 contributions advertised in the abstract and Section 1.1 (exact characterization of CtrId; closure-dimensional characterization of uniform CtrGen and the finite-class characterization of ordinary CtrGen; strict diamond hierarchy; robustness reversal and exact characterization of identification at every corruption budget) are stated as theorems in Sections 4 to 6 and proved in full in Sections A to C. The introduction states that the adversarial covering stream is a limit-learning abstraction and that ordinary generation on countable classes remains open.

   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 introduction states that the model is a limit-learning abstraction of direction-hidden pair signals, with an adversarial stream that covers every positive, not a model of a deployed pipeline. Section 5 and Section B.1 state that ordinary contrastive generation is characterized only for finite classes and that the natural condition is not sufficient for countable classes. Section 6 and Section C.4 show that corruption budgets growing with elapsed time make any two hypotheses confusable, while budgets measured against honest coverage still allow identification on bounded-overlap classes. Section 7 and Section E list the open problems, including the information-theoretic (non-effective) nature of the identifiers that use tell-tale sets.

   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 formal results state their hypotheses explicitly. Full proofs of every theorem, proposition, lemma, and corollary appear in the main text or in Sections A to D, and each section of the main text points to the appendix that contains its omitted proofs.

   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 presents purely theoretical results in the limit-learning paradigm and contains no 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 is purely theoretical. It has no data or 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: No experiments. The paper is purely theoretical.

   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: No experiments. The paper is purely theoretical.

   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: No experiments. The paper is purely theoretical.

   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 work is foundational theory in inductive inference and learning theory; it conforms with the NeurIPS Code of Ethics, raises no concerns regarding human subjects, fairness, or harmful use, and uses no data or models.

   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: [N/A]

Justification: The paper presents foundational results on the information-theoretic feasibility of identification and generation in the limit from contrastive presentations. It is several layers of abstraction removed from any specific deployed system and has no direct positive or negative societal impact.

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 any data, models, or code.

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 uses no existing data, code, or models.

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 introduces no new code, data, or model assets.

    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 involves no crowdsourcing and no 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 is purely theoretical and does not involve human subjects.

    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: [Yes]

    Justification: An LLM-based assistant (Claude, Anthropic) was used to draft and check parts of the mathematics: the extension of the identification characterization to every corruption budget (Theorems 6.9 and 6.10, Corollaries 6.11 and 6.12, and Lemmas C.1 to C.4), the countable-class example of Theorem B.2, the unbounded-budget result Proposition C.6, and the write-up of proofs of results that the authors stated during the review discussion (Theorem 5.15 and Section C.4). It was also used for editing. All results and proofs were checked by the authors, supported by independent machine checks (brute-force verification on small ground sets and independent reconstructions of the proofs).

    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.