Electronic Colloquium on Computational Complexity, Revision 2 of Report No. 62 (2026)

Toward a Characterization of Simulation Between Arithmetic Theories

Hunter Monroe

Abstract

We study when a sound arithmetic theory $\mathcal{S}\supseteq S_2^1$ with polynomial-time decidable axioms efficiently proves the bounded consistency statements $Con_{\mathcal{S}+\phi}(n)$ for a true sentence ϕ. Our main structural result characterizes the nonexistence of a length-optimal propositional proof system: this is equivalent to every such $\mathcal{S}$ failing to simulate $S_2^1+\mathsf{Exp}+\phi_{BB}(k)$ for all sufficiently large k, where $\phi_{BB}(k)$ asserts the exact value of the k-state Busy Beaver function and Exp asserts total exponentiation. We obtain this characterization by combining the Krajíček–Pudlák correspondence with our transfer theorem: for each fixed $\mathcal{S}$, the existence of a true extension not simulated by $\mathcal{S}$ is equivalent to this eventual Busy Beaver non-simulation property. A finite version transfers individual lower bounds to Busy Beaver instances with explicit parameter changes and proof-length bounds. For finitely axiomatized sequential $\mathcal{S}$, we also prove that $S_2^1\vdash Con_{\mathcal{S}}\rightarrow Con_{\mathcal{S}+\phi}$ yields polynomial-size $\mathcal{S}$-proofs of $Con_{\mathcal{S}+\phi}(n)$. We refute our earlier conjecture that $EA\vdash Con_{\mathcal{S}}\rightarrow Con_{\mathcal{S}+\phi}$ characterizes simulation.

In this finitely axiomatized sequential setting, a new Kolmogorov Blindness conjecture predicts hardness of random tautologies: true logarithmic-deficiency randomness axioms $x\in R^{\log}$ yield bounded-consistency statements requiring exponentially long proofs, at a rate depending on the fixed proof encoding, beyond an onset common to all strings of each sufficiently large length. It further predicts this hardness after adjoining a true randomness axiom $y\in R^{\log}$ with $|x|=|y|$, provided x remains random conditional on y. For fixed Π_1 axiom families with computable proof-length bounds and a common computable onset, we prove a quantitative equivalence between dense hardness with polynomially small exceptional fractions and hardness above a corresponding Kolmogorov-complexity threshold.

1 Introduction

"...the essence of the big open problems in complexity theory could be logical, rather than combinatorial" — Pudlák [22, p. 581]

ISSN 1433-8092

Let $\mathcal{S} \supseteq S_2^1$ be a sound arithmetic theory with polynomial-time decidable axioms, and let ϕ be a true sentence. We study when $\mathcal{S}$ efficiently proves the bounded consistency statements $Con_{\mathcal{S}+\phi}(n)$, where $Con_{\mathcal{S}+\phi}(n)$ asserts that $\mathcal{S}+\phi$ has no proof of 0=1 within n bits under a fixed proof encoding. The question of how long these consistency proofs must be goes back to Gödel, as reported by Kreisel (1967, p. 241); see Pudlák [21, p. 628].

The finite-consistency comparison asks whether $\mathcal{S}$ has proofs of $Con_{\mathcal{S}+\phi}(n)$ of size polynomial in n (Krajíček and Pudlák [18]); we abbreviate this by saying that $\mathcal{S}$ simulates $\mathcal{S}+\phi$. The literature provides sufficient conditions for $\mathcal{S}$ to simulate $\mathcal{S}+\phi$: Hrubeš's self-reference construction and Jeřábek's interpretation argument give true independent extensions with short bounded-consistency proofs under their respective hypotheses (Pudlák [23, Theorem 3.4 and Lemma 3.6]; see Theorems 4.1 and 3.1).

On the other hand, if no length-optimal propositional proof system exists, then for each $\mathcal{S}$ as above there is a true sentence ϕ such that $\mathcal{S} \nvdash^{n^c} Con_{\mathcal{S}+\phi}(n)$ for every constant $c>0$; see Krajíček and Pudlák [18, Theorem 2.1] and Corollary 3.13 below. Thus the problem is to identify a criterion for ϕ that governs simulation and non-simulation. We formulate this as follows.

> **Characterization Problem.** For a sound theory $\mathcal{S} \supseteq S_2^1$ with polynomial-time decidable axioms and a true sentence ϕ, determine when $\mathcal{S}$ has polynomial-size proofs of $Con_{\mathcal{S}+\phi}(n)$ and when no polynomial bounds their lengths.

The paper's main structural result is an exact Busy Beaver characterization of the nonexistence of a length-optimal propositional proof system (Corollary 3.13): such a system does not exist if and only if every sound $\mathcal{S} \supseteq S_2^1$ with polynomial-time decidable axioms fails to simulate $S_2^1+\mathsf{Exp}+\phi_{BB}(k)$ for all sufficiently large k. This combines the Krajíček–Pudlák correspondence [18, Theorem 2.1] with our transfer theorem. Under this equivalent no-optimality hypothesis, the propositional encodings of these Busy Beaver consistency statements give a canonical collection of tautology families: every Cook–Reckhow proof system has no polynomial bound on proof lengths as n varies in each family with sufficiently large k, with the threshold depending on the proof system. Classification of individual true extensions remains open.

On the positive side, we give a sufficient condition for simulation requiring only relative consistency provable in S_2^1. If S_2^1 proves $Con_{\mathcal{S}} \to Con_{\mathcal{S}+\phi}$, Parikh's theorem (Buss [7]) gives a polynomial bounded-consistency transfer;

combined with short self-consistency proofs, this yields simulation (Theorem 3.4). Finitely axiomatized sequential theories have the required short self-consistency proofs (Pudlák [20]). This includes the verified interpretation mechanism of Jeřábek, as presented by Pudlák [23, Lemma 3.6]. On the hardness side, Pudlák [20, Problem 1, p. 193] conjectures that simulation fails for $\phi{=}Con_{\mathcal{S}}$.

The Busy Beaver transfer identifies exact Busy Beaver values as canonical witnesses to non-simulation. For each sound $\mathcal{S}{\supseteq}S^1_2$ with polynomial-time decidable axioms, a hard true extension exists if and only if $\mathcal{S}$ fails to simulate $S^1_2{+}\mathsf{Exp}{+}\phi_{BB}(k)$ for all sufficiently large k (Theorem 3.10). Here $\phi_{BB}(k)$ specifies the maximum halting time of a halting k-state Turing machine (Radó [25]). Exponentiation is used inside the witness theory; $\mathcal{S}$ need only extend S^1_2. A finite version transfers an individual lower bound to a Busy Beaver instance, and conversely to a suitable true extension, with explicit changes of parameter and proof length (Theorem 3.11). For any class of such theories each having a hard true extension, a nonconstructive diagonal choice also gives one Busy Beaver consistency family hard throughout that class (Corollary 3.14); no hypothesis is imposed on theories outside the class. Alternative canonical witnesses over PA are exact prefixes of Chaitin's halting probability and complete finite randomness lists (Proposition 3.15).

We propose Kolmogorov randomness as an obstruction to simulation: our Kolmogorov Blindness conjecture predicts exponential lower bounds for bounded-consistency proofs associated with sufficiently long true randomness axioms (Conjecture 5.1). It further asserts that the bounds hold after adjoining another true randomness axiom of the same string length, provided the target string remains random conditional on the added string. For sound finitely axiomatized sequential $\mathcal{S}{\supseteq}S^1_2$, it implies Strong KH, requiring non-simulation for every sufficiently long true randomness axiom, and Density KH, requiring such true hard axioms for a fraction tending to one of strings (Theorem 5.3). Its universal form also implies Conditional Factual No Access (Proposition 5.6).

The density contribution makes the connection between randomness and dense hardness precise: for fixed Π_1 axiom families with computable proof-length bounds and a computable common onset, we prove a quantitative equivalence between dense hardness and hardness above a corresponding randomness threshold (Section 6). One direction is counting: the logarithmic-deficiency randomness set $R^{\log}$ defined in Section 5 has density tending to one, so hardness for every sufficiently long member gives dense hardness.

Conversely, if the fraction of exceptions to truth or any prescribed bound is $O(m^{-a})$, every sufficiently long string with $K_U(x) \geq m - a \log_2 m + c$, for a suitable constant c, gives a true instance satisfying every bound after the common onset. Exceptions can be effectively enumerated under the stated hypotheses; their sparsity remains a premise. The pairwise version uses joint complexity and requires both bounds to hold when the other axiom is adjoined. Density is measured among indices, with the same strings satisfying the whole prescribed tail.

Together, these results constrain how fine-grained and how coarse-grained the ordering of theories by simulation can be. Verified relative consistency prevents a separation by simulation, while any failure to simulate a true extension must also appear among canonical Busy Beaver theories. The density equivalence further constrains which random indices can be exceptions to prescribed hardness bounds.

The first version of this paper proposed that, for sound finitely axiomatized sequential $\mathcal{S} \supseteq S_2^1$, simulation of a true extension $\mathcal{S}+\phi$ is equivalent to $EA \vdash Con_{\mathcal{S}} \to Con_{\mathcal{S}+\phi}$. Section 4 refutes its necessary direction, Higher Relative Consistency (HRC), using Hrubeš's self-reference construction. Replacing EA by S_2^1 does not repair the characterization.

A standard folklore argument shows that a consistent $\mathcal{T} \supseteq EA$ with polynomial-time decidable axioms cannot prove the uniform assertion of its own non-simulation of a fixed extension: that assertion implies $Con_{\mathcal{T}}$. Section 7 gives the precise statement and discusses why randomness relative to a stronger theory suggests an obstacle to proving hardness, without establishing such unprovability. For the corresponding propositional argument, see Pudlák [24, Proposition 4.5 and Corollary 4.7].

Section 2 provides preliminaries. Section 3 proves the relative-consistency and Busy Beaver results; Section 4 refutes HRC. Sections 5 and 6 develop Kolmogorov Blindness and the effective density equivalence, respectively. Section 7 discusses provability, and Section 8 concludes.

2 Preliminaries

We work with sound arithmetical theories $\mathcal{S} \supseteq S_2^1$ with polynomial-time decidable axioms, and true sentences ϕ, unless stated otherwise. Fix a standard arithmetization in the language of bounded arithmetic, extended by explicit definitions. Use ordinary Hilbert proofs with reusable earlier lines, equivalently

a directed acyclic graph calculus with cut, and binary numerals (Pudlák [23, Sections 2.4–2.5]). Length is the bit length of an ordinary local binary serialization: formulas are written explicitly, and numeral digits, variable identifiers, line references and headers are counted. The encoding adds no padding depending on the total proof length. Polynomial simulation is preserved by verified polynomial proof translations; a numerical exponential rate requires this particular encoding to remain fixed.

We fix a finite axiomatization of S^1_2 (Krajíček [17, Section 2]). A theory extending S^1_2 either contains these axioms literally or comes with a polynomial-time proof embedding verified in S^1_2. Fixed proofs of the finitely many base axioms supply such an embedding. Each theory with polynomial-time decidable axioms has a *polynomial-time ordinary proof predicate* $Prf_{\mathcal{T}}$: its fixed clocked verifier is S^1_2-provably equivalent to this predicate, from which $Con_{\mathcal{T}}$ and $Con_{\mathcal{T}}(n)$ are formed.

The sentence $Con_{\mathcal{S}}(n)$ asserts that no $\mathcal{S}$-proof of 0=1 has length at most n (Pudlák [20]).[1] We write $\mathcal{S}\vdash^{n^c}\phi(n)$ when, for every sufficiently large n, there is an $\mathcal{S}$-proof of $\phi(n)$ of length at most n^c, and $\mathcal{S}\vdash^{n^{O(1)}}\phi(n)$ when a polynomial bounds these lengths. For a single sentence ψ, $\mathcal{S}\vdash^{s}\psi$ means that a proof of length at most s exists.

We abbreviate *finite-consistency simulation* by *simulation*: $\mathcal{S}$ simulates $\mathcal{T}$ when $\mathcal{S}\vdash^{n^{O(1)}}Con_{\mathcal{T}}(n)$ (Krajíček and Pudlák [18]). The target $\mathcal{T}$ need not be a single-sentence extension of $\mathcal{S}$. For propositional systems, P simulates Q in the length sense if every tautology with a Q-proof of length s has a P-proof of length at most $p(s)$ for one polynomial p. Cook and Reckhow [11, Definition 1.5] additionally require a polynomial-time translation for p-simulation. An *optimal*, or *length-optimal*, system simulates every Cook–Reckhow system in the length sense (Pudlák [23, Section 5]). Write $\mathrm{RFN}(P)$ for the arithmetic sentence asserting that every formula with a P-proof is a tautology.

Let Exp assert total exponentiation: $\forall x\,\exists y\,E_2(x,y)$, where the bounded formula $E_2(x,y)$ expresses $y{=}2^x$ and S^1_2 verifies its elementary defining properties, including $|y|{=}x{+}1$. We treat $S^1_2{+}\mathsf{Exp}{+}\psi$ as extension by the single sentence $\mathsf{Exp}{\wedge}\psi$. Elementary Arithmetic EA uses the common language fixed in this section, with the usual explicit definitions of $|x|$, $\lfloor x/2\rfloor$ and $x\#y$. Under this translation EA proves the axioms of S^1_2.

[1] Friedman's 1979 work and Pudlák's subsequent results [23, Introduction] developed the quantitative theory.

A *uniform feasible relative-consistency certificate* is an S^1_2-proof of $\forall n\,(Con_{\mathcal{S}}(p(n)) \to Con_{\mathcal{S}+\phi}(n))$ for a fixed polynomial p. In contrast, *instance-wise feasible relative consistency* means the external family $\mathcal{S} \vdash^{n^{O(1)}} Con_{\mathcal{S}}(p(n)) \to Con_{\mathcal{S}+\phi}(n)$. With short self-consistency proofs the latter is equivalent to simulation (Theorem 3.2); it supplies no uniform certificate. Neither notion means $EA{+}Con_{\mathcal{S}} \vdash \phi$.

For Jeřábek's argument, a polynomial-time interpretation includes a polynomial proof translation whose correctness is verified in S^1_2, including uniform translated axiom proofs for infinite presentations. The standard formalized Σ^b_1-completeness bound supplies short proofs of accepted computations (Khaniki [14, Theorem 2.2]). We additionally use its syntactic computation-writing construction: for each fixed clocked verifier V, a polynomial-time function takes an input and an accepting computation tableau to an S^1_2-proof that V accepts that input, of size polynomial in the input and tableau lengths. Its correctness is verified in S^1_2 by checking each computation step. The same convention applies to rejecting computations. This explicit construction, rather than an inference from mere existence of short proofs, is used in the proof transformations below. Replacing an axiom by a fixed derivation is another syntactic transformation verified in S^1_2.

3 Relative Consistency and Busy Beaver Transfer

This section presents the paper's first two unconditional results. The first shows in certain settings that relative consistency implies feasible relative consistency and hence simulation. The second shows that any hard true ϕ can, in a precise sense, be replaced by one of Busy Beaver form. Together, these results isolate both a positive mechanism underlying simulation and a canonical source of hardness, constraining any possible characterization.

The literature shows that simulation is possible in special cases. In particular, when $\mathcal{S}$ has polynomial-size proofs of its own bounded consistency, polynomial-time interpretability implies efficient provability of $Con_{\mathcal{S}+\phi}(n)$; see Theorem 3.1.[2] By contrast, Pudlák [20, Problem 1, p. 193] conjec-

[2]Freund and Pakhomov [12, Theorem 3.6] prove that PA has polynomial-size proofs of the bounded consistency of PA+Con^*(PA), where Con^*(PA) is their slow-consistency sentence. This example uses full PA, rather than a finitely axiomatized base.

tures that simulation fails already for $\phi{=}Con_{\mathcal{S}}$ in the finitely axiomatized setting (Conjecture 3.5).[3] The assertion that every sound theory $\mathcal{S}{\supseteq}S^1_2$ with polynomial-time decidable axioms has a true extension $\mathcal{S}{+}\phi$ such that $\mathcal{S} \not\vdash^{n^{O(1)}} Con_{\mathcal{S}+\phi}(n)$ remains open. This universal assertion is equivalent to the nonexistence of a length-optimal propositional proof system, and hence would imply **NP**$\neq$**coNP** (Corollary 3.13).

3.1 A Known Sufficient Condition for Simulation

Jeřábek's proof-translation argument, presented by Pudlák [23, Lemma 3.6] for finitely axiomatized sequential theories, gives the following sufficient condition. Our explicit feasible-translation convention also permits infinite presentations when the required translation is supplied.

Theorem 3.1. *(Proof-translation argument of Jeřábek, via Pudlák [23, Lemma 3.6]) Suppose $\mathcal{S}{\supseteq}S^1_2$ and $\mathcal{S}\vdash^{n^{O(1)}} Con_{\mathcal{S}}(n)$. If $\mathcal{S}$ polynomial-time interprets $\mathcal{S}{+}\phi$, then $\mathcal{S}\vdash^{n^{O(1)}} Con_{\mathcal{S}+\phi}(n)$.*

Proof. The interpretation translates an $(\mathcal{S}{+}\phi)$-refutation into an $\mathcal{S}$-proof of $(0{=}1)^i$ with polynomial growth. Translating a fixed proof of $0{\neq}1$ supplies an $\mathcal{S}$-proof of $\neg(0{=}1)^i$, hence an $\mathcal{S}$-refutation with polynomial growth. By our interpretation convention, S^1_2 verifies this construction, so for a polynomial p it proves $\forall n\,(Con_{\mathcal{S}}(p(n)){\rightarrow}Con_{\mathcal{S}+\phi}(n))$. Instantiating this fixed proof and combining it with the assumed short proofs of $Con_{\mathcal{S}}(p(n))$ gives the result. □

The uniform certificate here is stronger than the existence of a short-proof family; Section 4 shows that the latter need not supply the former.

3.2 Stronger Results on Simulation

The interpretation argument uses a uniform bounded-consistency implication. Its instance-wise analogue has a simple characterization.

[3] A computable jump operator assigns to each propositional proof system P a proof system that P cannot simulate. Khaniki [15, Theorem 5.11] proves that the existence of such an operator is equivalent to a weakened form of Pudlák's conjecture, requiring the consistency-extension lower bound only for consistent finitely axiomatized theories proving a fixed sound finitely axiomatized extension of S^1_2. The existence of such operators remains open.

Theorem 3.2. *Let $\mathcal{S}\supseteq S_2^1$ be sound with polynomial-time decidable axioms and $\mathcal{S}\vdash^{n^{O(1)}} Con_{\mathcal{S}}(n)$. For any fixed polynomial p taking natural-number values, $\mathcal{S}$ simulates $\mathcal{S}+\phi$ if and only if $\mathcal{S}\vdash^{n^{O(1)}} Con_{\mathcal{S}}(p(n))\to Con_{\mathcal{S}+\phi}(n)$.*

Proof. Short proofs of the implication combine with short proofs of $Con_{\mathcal{S}}(p(n))$ to give simulation. Conversely, a proof of $Con_{\mathcal{S}+\phi}(n)$ gives the implication by propositional logic with polynomial overhead. Finitely many small values of $p(n)$ can be covered by fixed proofs of their true finite-consistency statements. □

The corresponding equivalence for non-simulation follows immediately.

Corollary 3.3. *Let $\mathcal{S}\supseteq S_2^1$ be sound with polynomial-time decidable axioms, and suppose $\mathcal{S}\vdash^{n^{O(1)}} Con_{\mathcal{S}}(n)$. For any sentence ϕ and any fixed polynomial p taking natural-number values, $\mathcal{S}$ fails to simulate $\mathcal{S}+\phi$ if and only if $\mathcal{S}\not\vdash^{n^{O(1)}} Con_{\mathcal{S}}(p(n))\to Con_{\mathcal{S}+\phi}(n)$.*

Proof. Negate both sides of the equivalence in Theorem 3.2. □

The proof of Theorem 3.1 uses only a relative-consistency implication that S_2^1 proves uniformly in n. The next theorem shows that such an implication follows whenever S_2^1 proves the unbounded implication $Con_{\mathcal{S}}\to Con_{\mathcal{S}+\phi}$; no interpretation is needed. Note that the hypothesis $\mathcal{S}\vdash^{n^{O(1)}} Con_{\mathcal{S}}(n)$ holds whenever $\mathcal{S}$ is finitely axiomatized and sequential (Pudlák [20, Theorem 5.4 and p. 189]).

Theorem 3.4. *Let $\mathcal{S}\supseteq S_2^1$ have a polynomial-time ordinary proof predicate, and suppose $\mathcal{S}\vdash^{n^{O(1)}} Con_{\mathcal{S}}(n)$. Let ϕ be a sentence. If $S_2^1\vdash Con_{\mathcal{S}}\to Con_{\mathcal{S}+\phi}$, then there is a polynomial q with $q(n)\geq n$ such that $S_2^1\vdash \forall n\,\big(Con_{\mathcal{S}}(q(n))\to Con_{\mathcal{S}+\phi}(n)\big)$. Its instances have polynomial-size $\mathcal{S}$-proofs, and $\mathcal{S}$ simulates $\mathcal{S}+\phi$.*

Proof. Write $\bot$ for $\ulcorner 0{=}1\urcorner$. By the conventions of Section 2, $Prf_{\mathcal{S}}$ and $Prf_{\mathcal{S}+\phi}$ are computed by clocked polynomial-time verifiers whose computations are S_2^1-provably equivalent to them, so each may be taken to be a bounded formula. Classical logic turns the hypothesis into $S_2^1\vdash \forall\pi\,\exists\rho\,\big(\neg Prf_{\mathcal{S}+\phi}(\pi,\bot)\vee Prf_{\mathcal{S}}(\rho,\bot)\big)$: if π is not an $(\mathcal{S}+\phi)$-refutation, any ρ will do, and otherwise $\neg Con_{\mathcal{S}+\phi}$ holds, so the hypothesis supplies an $\mathcal{S}$-refutation ρ. The matrix is a bounded formula, so Parikh's theorem, in the form stated by Buss [7,

Section 4.7, Theorem 11, p. 83], gives a term $t(\pi)$ of bounded arithmetic such that S_2^1 proves the same sentence with the quantifier restricted to $\rho{\leq}t(\pi)$.

The value of $t(\pi)$ need not be polynomial in π, but its binary length is polynomial in $|\pi|$. By induction on the structure of t, using $|x{+}y|{\leq}\max(|x|,|y|){+}1$, $|x{\cdot}y|{\leq}|x|{+}|y|$, $|x\#y|{=}|x||y|{+}1$ and the corresponding bounds for the remaining function symbols, there is a polynomial q with nonnegative integer coefficients, which we may take with $q(n){\geq}n$, such that $S_2^1{\vdash}\forall\pi\,\forall n\,\forall\rho\,\big(|\pi|{\leq}n\,\wedge\,\rho{\leq}t(\pi)\,\rightarrow\,|\rho|{\leq}q(n)\big)$. Reasoning in S_2^1, let π be an $(\mathcal{S}{+}\phi)$-refutation with $|\pi|{\leq}n$. The bounded witness $\rho{\leq}t(\pi)$ then satisfies $Prf_{\mathcal{S}}(\rho,\bot)$ and $|\rho|{\leq}q(n)$, so $\neg Con_{\mathcal{S}}(q(n))$. Hence $S_2^1{\vdash}\forall n\,\big(Con_{\mathcal{S}}(q(n)){\rightarrow}Con_{\mathcal{S}+\phi}(n)\big)$.

For the external consequence, fix this S_2^1-proof once and for all. Since $\mathcal{S}$ extends S_2^1, it yields an $\mathcal{S}$-proof of the same sentence. For a standard n, instantiate the universal quantifier at the binary numeral $\bar{n}$ and replace the term $q(\bar{n})$ by the numeral $\overline{q(n)}$, using an S_2^1-proof of $q(\bar{n}){=}\overline{q(n)}$ of size polynomial in n. This gives $\mathcal{S}$-proofs of $Con_{\mathcal{S}}(q(n)){\rightarrow}Con_{\mathcal{S}+\phi}(n)$ of size polynomial in n. By hypothesis there are constants c and N such that $Con_{\mathcal{S}}(m)$ has an $\mathcal{S}$-proof of size at most m^c for every $m{\geq}N$. For $n{\geq}N$ we have $q(n){\geq}N$, and modus ponens with that proof at $m{=}q(n)$ gives an $\mathcal{S}$-proof of $Con_{\mathcal{S}+\phi}(n)$ of size polynomial in n. □

Whether $EA{\vdash}Con_{\mathcal{S}}{\rightarrow}Con_{\mathcal{S}+\phi}$ suffices for simulation remains open here. Parikh's term-bounding argument does not apply to EA, which proves total exponentiation. The Friedman characterization (Visser [27, Theorem 4.3]) uses consistency restricted by formula complexity, not the ordinary consistency statements here, so it supplies no replacement argument. Theorem 3.4 subsumes the verified interpretation mechanism, whose bounded implication yields the unbounded S_2^1 implication.

By contrast, Pudlák [20, Problem 1, p. 193] conjectures that even the benchmark case $\phi{=}Con_{\mathcal{S}}$ already yields non-simulation; see also Pudlák [23, Section 3.1, p. 12]. We use the following fixed-theory formulation:

Conjecture 3.5. *(Pudlák's Conjecture for $\mathcal{S}$) Let $\mathcal{S}{\supseteq}S_2^1$ be sound with polynomial-time decidable axioms and its standard proof predicate. For every constant c, $\mathcal{S}\not\vdash^{n^c} Con_{\mathcal{S}+Con_{\mathcal{S}}}(n)$.*

Under Pudlák's Conjecture for $\mathcal{S}$, any true extension that raises absolute consistency (proves $Con_{\mathcal{S}}$) is hard:

Theorem 3.6. *Assume Pudlák's Conjecture for $\mathcal{S}$, and suppose $\mathcal{S}+\phi \vdash Con_{\mathcal{S}}$. Then for every constant c, $\mathcal{S} \not\vdash^{n^c} Con_{\mathcal{S}+\phi}(n)$.*

Proof. Replacing each use of $Con_{\mathcal{S}}$ by its fixed $(\mathcal{S}+\phi)$-proof gives an S^1_2-verified polynomial refutation translation. Hence, for a polynomial p, $\mathcal{S}$ has polynomial-size proofs of $Con_{\mathcal{S}+\phi}(p(n)) \to Con_{\mathcal{S}+Con_{\mathcal{S}}}(n)$. Simulation of $\mathcal{S}+\phi$ would therefore give simulation of $\mathcal{S}+Con_{\mathcal{S}}$, contrary to the conjecture. □

In particular, the conjecture gives hardness for strictly stronger true extensions as well.

Corollary 3.7. *Assume Pudlák's Conjecture for a sound theory $\mathcal{S} \supseteq S^1_2$ with polynomial-time decidable axioms. Let θ be true but unprovable in $\mathcal{S}+Con_{\mathcal{S}}$, and put $\phi := Con_{\mathcal{S}} \wedge \theta$. Then $\mathcal{S}+\phi$ is strictly stronger than $\mathcal{S}+Con_{\mathcal{S}}$, and $\mathcal{S}$ does not simulate $\mathcal{S}+\phi$.*

Proof. Since $\mathcal{S}+\phi \vdash Con_{\mathcal{S}}$, non-simulation follows from Theorem 3.6. The implication $\phi \to Con_{\mathcal{S}}$ is provable in $\mathcal{S}$, whereas $Con_{\mathcal{S}} \to \phi$ is not: such a proof would make $\mathcal{S}+Con_{\mathcal{S}}$ prove θ, contrary to the hypothesis. □

Corollary 3.7 shows upward closure of the conjectural hardness obstruction; it does not establish a converse characterization.

3.3 Non-Simulation Implies Non-Simulation for Busy Beavers

A sound theory $\mathcal{S} \supseteq S^1_2$ with polynomial-time decidable axioms for which, for every true sentence ϕ, there exists a polynomial bound on $\mathcal{S}$-proofs of $Con_{\mathcal{S}+\phi}(n)$ would simulate every true extension, even though $\mathcal{S}$ does not prove all true sentences. We show that any failure of this universal simulation is already witnessed by a canonical family of true sentences, namely exact Busy Beaver value statements adjoined to $S^1_2+\mathsf{Exp}$ (Theorem 3.10).

For each k, let $t_k = BB(k)$ be the true k-state Busy Beaver value, and let $\phi_{BB}(k)$ denote the true sentence asserting the exact value $BB(k)=t_k$. Here $BB(k)$ is the maximum halting time of any halting k-state machine on blank input, for a fixed deterministic machine convention with padding by unreachable states. The sentence $\phi_{BB}(k)$ implies that every halting k-state machine halts within t_k steps; we use this bounded-halting consequence in the proof of Lemma 3.9.

We begin by proving that, for any fixed consistent theory with a polynomial-time ordinary proof predicate, sufficiently large Busy Beaver axioms imply its consistency over the base S_2^1+Exp; the contradiction-search idea is taken from the proof of Aaronson's Proposition 3 [1]. The relevant threshold is the number of states of an explicitly compiled contradiction-search machine for the theory. Once this is in place, any hard true extension of a sound theory $\mathcal{S}\supseteq S_2^1$ with polynomial-time decidable axioms yields eventual hardness throughout the Busy Beaver family over S_2^1+Exp.

Definition 3.8. (**Contradiction Search Threshold**) Let $\mathcal{T}$ have a polynomial-time ordinary proof predicate, and let $E_{\mathcal{T}}$ be the fixed Turing machine that, on blank input, searches for a $\mathcal{T}$-proof of contradiction. Define $k_{\mathcal{T}}^{\mathrm{cs}}$ to be the number of states of $E_{\mathcal{T}}$. For each $k\geq k_{\mathcal{T}}^{\mathrm{cs}}$, let $E_{\mathcal{T},k}$ be the k-state machine obtained by adding unreachable states.

Lemma 3.9. *Let $\mathcal{T}$ be a consistent theory with a polynomial-time ordinary proof predicate. Then for every standard $k\geq k_{\mathcal{T}}^{\mathrm{cs}}$, one has S_2^1+Exp+$\phi_{BB}(k)\vdash Con_{\mathcal{T}}$.*

Proof. Implement $E_{\mathcal{T}}$ by enumerating candidate $\mathcal{T}$-proofs in length-lexicographic order, running the clocked verifier for $Prf_{\mathcal{T}}(\cdot,\ulcorner 0{=}1\urcorner)$ on each candidate, and halting exactly when it accepts.

Use a canonical binary tableau encoding: a history of ℓ steps contains exactly $\ell{+}1$ configurations, recording the state, head position and the tape segment $[-\ell,\ell]$, with self-delimiting headers and no additional padding. Write $\ell(h)$ for the step count and $\mathrm{HaltTrace}_M(h)$ for the polynomial-time check that this is a complete halting history of M on blank input. Its bounded-arithmetic representation includes the format check. Put $\mathrm{Halt}_M \equiv \exists h\,\mathrm{HaltTrace}_M(h)$ and $\mathrm{HaltBy}_M(t) \equiv \exists h\,(\mathrm{HaltTrace}_M(h)\wedge \ell(h){\leq}t)$. The format gives a fixed polynomial q_M such that S_2^1 proves $\mathrm{HaltTrace}_M(h)\wedge \ell(h){\leq}t \;\longrightarrow\; |h|{\leq}q_M(t)$.

For each fixed standard t, choose the standard integer $H_{M,t}{=}2^{q_M(t)+1}$. Then S_2^1 proves the fixed-parameter equivalence $\mathrm{HaltBy}_M(\bar{t}) \;\leftrightarrow\; \exists h{<}\overline{H_{M,t}}\,(\mathrm{HaltTrace}_M(h)\wedge \ell(h){\leq}\bar{t})$. Here $\overline{H_{M,t}}$ is a fixed numeral; no bounded-arithmetic term for $2^{q_M(t)+1}$ with variable t, and no uniformly bounded formula for $\mathrm{HaltBy}_M(t)$, is asserted.

To formalize $\phi_{BB}(k)$, let $\mathrm{HaltTrace}(M,h)$ be the corresponding tableau formula uniform in the machine code, so that $\mathrm{HaltTrace}_M(h)$ abbreviates $\mathrm{HaltTrace}(\ulcorner M\urcorner,h)$. Writing Mach_k for the set of codes of k-state machines, $\phi_{BB}(k)$ is the conjunction of $\forall M{\in}\mathrm{Mach}_k\,\forall h\,(\mathrm{HaltTrace}(M,h){\rightarrow}\ell(h){\leq}t_k)$ and

$\exists M{\in}\mathrm{Mach}_k\,\exists h\,(\mathrm{HaltTrace}(M,h){\wedge}\ell(h){=}t_k)$. The first clause bounds every halting computation; the second asserts attainment of the bound. The numeral t_k is fixed before the proof-length parameter n, and no uniformly feasible computation of t_k from k is asserted.

First, $S^1_2{+}\mathsf{Exp}$ proves that every $\mathcal{T}$-refutation yields a halting history of $E_\mathcal{T}$. Let q be a polynomial bounding the running time of the clocked verifier. If p is a $\mathcal{T}$-refutation, then $E_\mathcal{T}$ examines fewer than $2^{|p|+1}$ candidates before it reaches p and spends at most $q(|p|)$ steps on each, up to a fixed coding factor, so it halts within $2^{|p|+1}q(|p|)$ steps up to that factor; the complete history up to halting has size bounded by a fixed polynomial in this quantity. Choose a fixed polynomial r such that $2^{r(|p|)}$ exceeds this size bound. There is a fixed polynomial-time function $F_\mathcal{T}$ which, given p and a padding string z, writes out the history of $E_\mathcal{T}$ up to halting, and S^1_2 proves

$$Prf_\mathcal{T}(p,\ulcorner 0{=}1\urcorner)\wedge E_2(r(|p|),a)\wedge E_2(a,z)\;\rightarrow\;\mathrm{HaltTrace}_{E_\mathcal{T}}\big(F_\mathcal{T}(p,z)\big).$$

Indeed, $E_2(a,z)$ gives $|z|{=}a{+}1$. Let $F_\mathcal{T}(p,z)$ run $E_\mathcal{T}$ on blank input for at most $|z|$ steps, recording the successive configurations, and output a fixed dummy string if $E_\mathcal{T}$ has not halted by then; its running time and output length are polynomial in the length of its padded input. S^1_2 proves by Σ^b_1 length induction (Buss [7, Chapter 2]) on the step counter $i{\le}|z|$ that the first i recorded configurations form a valid computation of $E_\mathcal{T}$, and by the same induction on the index j of the candidate under examination, which stays below $2^{|p|+1}{\le}a{<}|z|$, that the search has either halted at an earlier candidate or reaches p within the step bound above. Since $Prf_\mathcal{T}(p,\ulcorner 0{=}1\urcorner)$ holds and the clocked verifier is S^1_2-provably equivalent to $Prf_\mathcal{T}$, the verifier accepts at p at the latest. Hence the clock does not expire, and the output is a complete halting history. Already in S^1_2 the value $a{=}2^{r(|p|)}$ exists: its binary length is polynomial in $|p|$, and it is obtained by a polynomial-time function, with a term bound supplied by smash. The axiom Exp is used to obtain $z{=}2^a$, whose length is exponential in $|p|$. Hence $S^1_2{+}\mathsf{Exp}\vdash\forall p\,\big(Prf_\mathcal{T}(p,\ulcorner 0{=}1\urcorner)\rightarrow\mathrm{Halt}_{E_\mathcal{T}}\big)$.

The induction invariant asserts either that a valid halting prefix has already been produced, or that the search prefix is valid, with its current candidate and verifier configuration, and every completed verifier run has rejected. Prefixes and verifier tableaux have size polynomial in $|p|{+}|z|$ and therefore have bounded-arithmetic term bounds supplied by smash. Their polynomial-time format and transition checks have Σ^b_1 representations, so the invariant may be taken in that class. All loop indices are bounded by the

length of the padding parameter z; the argument uses length induction with that parameter, rather than unrestricted induction through an exponential numerical bound.

Now fix a standard $k \geq k^{\mathrm{cs}}_{\mathcal{T}}$ and let $U := S^1_2 + \mathsf{Exp} + \phi_{BB}(k)$. Three further facts hold:

$$S^1_2 \vdash \mathrm{Halt}_{E_{\mathcal{T}}} \to \mathrm{Halt}_{E_{\mathcal{T},k}},$$
$$S^1_2 + \phi_{BB}(k) \vdash \mathrm{Halt}_{E_{\mathcal{T},k}} \to \mathrm{HaltBy}_{E_{\mathcal{T},k}}(t_k),$$
$$S^1_2 \vdash \neg\mathrm{HaltBy}_{E_{\mathcal{T},k}}(t_k).$$

The first follows because padding adds only unreachable states, and S^1_2 verifies the corresponding map of computation histories. The second follows by instantiating the upper-bound clause of $\phi_{BB}(k)$ at $E_{\mathcal{T},k}$: S^1_2 proves the true bounded fact $\ulcorner E_{\mathcal{T},k} \urcorner \in \mathrm{Mach}_k$. For the third, consistency of $\mathcal{T}$ implies that $E_{\mathcal{T},k}$ does not halt within its first t_k steps; the fixed-parameter equivalence established above makes $\neg\mathrm{HaltBy}_{E_{\mathcal{T},k}}(\overline{t_k})$ equivalent in S^1_2 to a true closed bounded sentence. Exhaustive verification of that bounded sentence gives an S^1_2-proof, possibly of enormous length. This uses the fixed standard parameters k, t_k, not a claim that every true Π_1 sentence is provable. Reasoning in U, a $\mathcal{T}$-refutation would yield a halting history of $E_{\mathcal{T}}$, hence of $E_{\mathcal{T},k}$, hence a halting history with at most t_k steps, contradicting the third fact. Therefore $U \vdash Con_{\mathcal{T}}$. □

Choice of base. We use $S^1_2 + \mathsf{Exp}$ as the base of the Busy Beaver witness theories; the simulating theory need only extend S^1_2. Whether Lemma 3.9 holds over S^1_2 alone remains open. The base EA also suffices under the usual definitional translation.

If a consistent $\mathcal{T} \supseteq S^1_2$ with a polynomial-time ordinary proof predicate also proves Exp, the second incompleteness theorem therefore makes $\phi_{BB}(k)$ unprovable in $\mathcal{T}$ for every $k \geq k^{\mathrm{cs}}_{\mathcal{T}}$.

With this result, we can show that the existence of some hard true extension $\mathcal{S} + \phi$ is equivalent to eventual hardness throughout the Busy Beaver family over $S^1_2 + \mathsf{Exp}$:

Theorem 3.10. *Let $\mathcal{S} \supseteq S^1_2$ be a sound theory with polynomial-time decidable axioms. The following are equivalent:*

1. *There exists a true sentence ϕ such that for every constant c, $\mathcal{S} \not\vdash^{n^c} Con_{\mathcal{S}+\phi}(n)$.*

2. For all sufficiently large k, $\mathcal{S}$ does not simulate $S_2^1{+}\mathsf{Exp}{+}\phi_{BB}(k)$.

Proof. (1)$\to$(2) Put $\mathcal{T}{:=}\mathcal{S}{+}\phi$. It is consistent and has a polynomial-time ordinary proof predicate. For each fixed $k{\geq}k^{\mathrm{cs}}_{\mathcal{T}}$, Lemma 3.9 gives a fixed proof π_k of $Con_{\mathcal{T}}$ in $U_k{:=}S_2^1{+}\mathsf{Exp}{+}\phi_{BB}(k)$.

Given a $\mathcal{T}$-refutation ρ, the verifier-to-proof construction of Section 2 supplies a U_k-proof of $Prf_{\mathcal{T}}(\bar{\rho},\ulcorner 0{=}1\urcorner)$ of size polynomial in $|\rho|$. Appending π_k yields a U_k-refutation. The construction is verified in S_2^1, so for a polynomial $p_k(n){\geq}n$, $\mathcal{S}$ has polynomial-size proofs of $Con_{U_k}(p_k(n)){\to}Con_{\mathcal{T}}(n)$. Simulation of U_k would then imply simulation of $\mathcal{T}$, contradicting (1). Exponentiation enters through the fixed proof π_k; it is not required in $\mathcal{S}$.

(2)$\to$(1) For any k supplied by (2), take the true sentence $\phi{:=}\mathsf{Exp}{\wedge}\phi_{BB}(k)$. The specified embedding of S_2^1 into $\mathcal{S}$ extends to an S_2^1-verified polynomial embedding of U_k-proofs into $(\mathcal{S}{+}\phi)$-proofs. Thus for a polynomial $p(n){\geq}n$, $\mathcal{S}$ has polynomial-size proofs of $Con_{\mathcal{S}+\phi}(p(n)){\to}Con_{U_k}(n)$. Simulation of $\mathcal{S}{+}\phi$ would contradict (2). □

In Theorem 3.10, hardness holds for every sufficiently large k and, for each such k and each exponent, infinitely many consistency parameters n.

We next ask whether the asymptotic Busy Beaver transfer also holds when hardness is assumed at just one finite parameter. The answer is yes: for a fixed proving theory, an exponential lower bound at a single sufficiently large N transfers to a Busy Beaver consistency statement at a polynomially larger parameter, with a lower bound still exponential in the original N.

Theorem 3.11 (Finite Busy Beaver transfer)**.** *Let $\mathcal{S}{\supseteq}S_2^1$ be sound with a polynomial-time ordinary proof predicate, let ϕ be true, and put $\mathcal{T}{:=}\mathcal{S}{+}\phi$. Write $\ell_{\mathcal{S}}(\sigma)$ for the length of a shortest $\mathcal{S}$-proof of σ, with value ∞ if no proof exists. Fix $k{\geq}k^{\mathrm{cs}}_{\mathcal{T}}$ and put $U_k{:=}S_2^1{+}\mathsf{Exp}{+}\phi_{BB}(k)$. There is an integer polynomial $p_k(N){\geq}N$ such that, for every $a{>}0$, there are $b{>}0$ and N_0 for which every $N{\geq}N_0$ satisfies*

$$\ell_{\mathcal{S}}\big(Con_{\mathcal{T}}(N)\big){>}2^{aN} \quad\Longrightarrow\quad \ell_{\mathcal{S}}\big(Con_{U_k}(p_k(N))\big){>}2^{bN}.$$

If $\mathcal{S}{\vdash}\mathsf{Exp}$, the same assertion holds with U_k replaced by $\mathcal{S}{+}\phi_{BB}(k)$, with possibly different parameters.

Proof. Fix the U_k-proof π_k of $Con_{\mathcal{T}}$ supplied by Lemma 3.9. As in the proof of Theorem 3.10, the verifier-to-proof transformation, with π_k appended, gives an S_2^1-verified polynomial refutation translation from $\mathcal{T}$ to U_k. Hence

for a polynomial p_k there are $\mathcal{S}$-proofs of $Con_{U_k}(p_k(N))\to Con_{\mathcal{T}}(N)$ of size polynomial in N. Enlarge the bound to cover every standard N, including the finitely many small values. Composing such a proof with a proof of its antecedent of length at most L has length bounded by an integer polynomial $r_k(N, L)$, chosen nondecreasing in its nonnegative arguments. Contraposition gives the general finite bound: for every $N, L\in\mathbb{N}$, if $\ell_{\mathcal{S}}(Con_{\mathcal{T}}(N))>r_k(N, L)$, then $\ell_{\mathcal{S}}(Con_{U_k}(p_k(N)))>L$. For an append-only proof encoding, concatenating the implication proof and one application of modus ponens gives $r_k(N, L)=L+q_k(N)$ for a polynomial q_k; otherwise the polynomial bound suffices.

Choose constants $C, d\geq 1$ such that $r_k(N, L)\leq C(N+L+1)^d$. For any $a>0$, choose $0<b<a/d$. Then $r_k(N, \lfloor 2^{bN}\rfloor)\leq 2^{aN}$ for all sufficiently large N. Apply the finite bound with $L=\lfloor 2^{bN}\rfloor$; since proof lengths are integers, the conclusion is the asserted exponential lower bound.

For the same-base assertion, substitute fixed $\mathcal{S}$-proofs of the base axioms, including Exp, in π_k. This gives an $(\mathcal{S}+\phi_{BB}(k))$-proof of $Con_{\mathcal{T}}$, to which the same argument applies. Without the hypothesis $\mathcal{S}\vdash\mathsf{Exp}$, use $\mathcal{S}+(\mathsf{Exp}\wedge\phi_{BB}(k))$ instead. □

The theorem applies to a single $N\geq N_0$ at which the stated lower bound holds; it assumes no lower bounds at other parameters. Its conclusion is exponential in the original N. If p_k has degree D, the guarantee in the new parameter $M=p_k(N)$ is in general only $2^{\Omega(M^{1/D})}$. Thus polynomial reindexing need not preserve a fully exponential bound in M.

Conversely, fix any standard k and put $\phi_k:=\mathsf{Exp}\wedge\phi_{BB}(k)$. The feasible embedding of U_k into $\mathcal{S}+\phi_k$ gives nondecreasing integer polynomials $p'_k(N)\geq N$ and $r'_k(N, L)$ such that $\ell_{\mathcal{S}}(Con_{\mathcal{S}+\phi_k}(p'_k(N)))\leq L$ implies $\ell_{\mathcal{S}}(Con_{U_k}(N))\leq r'_k(N, L)$: translate refutations, instantiate the resulting bounded-consistency implication, and compose proofs. Contraposition transfers a finite Busy Beaver lower bound to the true extension $\mathcal{S}+\phi_k$, which need not be the original $\mathcal{S}+\phi$.

The forward constants and N_0 depend on the fixed k and its consistency proof π_k. Thus the transferred bound depends on the chosen Busy Beaver instance, and a single finite lower bound gives no asymptotic hardness conclusion.

Under Pudlák's Conjecture, one can give a stronger statement indicating just how large k must be:

Theorem 3.12. *Assume Pudlák's Conjecture for $\mathcal{S}$, and suppose that $\mathcal{S}\vdash\mathsf{Exp}$. Then for every $k\geq k^{\mathrm{cs}}_{\mathcal{S}}$, for every constant c, $\mathcal{S}\not\vdash^{n^c} Con_{\mathcal{S}+\phi_{BB}(k)}(n)$.*

Proof. Fix $k\geq k^{\mathrm{cs}}_{\mathcal{S}}$. Lemma 3.9 gives $S^1_2+\mathsf{Exp}+\phi_{BB}(k)\vdash Con_{\mathcal{S}}$. Since $\mathcal{S}$ extends S^1_2 and proves Exp, fixed proof substitution gives $\mathcal{S}+\phi_{BB}(k)\vdash Con_{\mathcal{S}}$. Theorem 3.6, using the assumed instance of Pudlák's conjecture for $\mathcal{S}$, gives the result. □

Without assuming $\mathcal{S}\vdash\mathsf{Exp}$, Pudlák's conjecture still gives non-simulation of $S^1_2+\mathsf{Exp}+\phi_{BB}(k)$ for every $k\geq k^{\mathrm{cs}}_{\mathcal{S}+Con_{\mathcal{S}}}$, by applying Theorem 3.10 to $\phi{=}Con_{\mathcal{S}}$. These are sufficient thresholds; the results do not classify the smaller Busy Beaver instances.

Combining Busy Beaver transfer with the Krajíček–Pudlák correspondence yields the paper's main structural characterization of the nonexistence of a length-optimal propositional proof system.

Corollary 3.13. *The following are equivalent:*

1. *No optimal propositional proof system exists.*
2. *For every sound theory $\mathcal{S}\supseteq S^1_2$ with polynomial-time decidable axioms, there is a true sentence ϕ such that for every constant c, $\mathcal{S}\not\vdash^{n^c} Con_{\mathcal{S}+\phi}(n)$.*
3. *For every sound theory $\mathcal{S}\supseteq S^1_2$ with polynomial-time decidable axioms, there is $K_{\mathcal{S}}$ such that for every $k\geq K_{\mathcal{S}}$ and every constant c, $\mathcal{S}\not\vdash^{n^c} Con_{S^1_2+\mathsf{Exp}+\phi_{BB}(k)}(n)$.*

The equivalence is unchanged if the theories $\mathcal{S}$ quantified in (2) and (3) are restricted to sound finitely axiomatized sequential extensions of S^1_2.

Proof. Conditions (2) and (3) are equivalent by Theorem 3.10. For (1)→(2), the Krajíček–Pudlák correspondence [18, Theorem 2.1], in the formulation of Khaniki [14, Theorem 2.3], gives a consistent $\mathcal{T}\supseteq S^1_2$ with polynomial-time decidable axioms whose bounded consistency has no polynomial-size $\mathcal{S}$-proofs. Take $\phi{:=}Con_{\mathcal{T}}$. Verifying a $\mathcal{T}$-refutation and applying this axiom gives an $(\mathcal{S}+\phi)$-refutation with polynomial overhead, so simulation of $\mathcal{S}+\phi$ would imply simulation of $\mathcal{T}$.

Conversely, an optimal proof system P gives the sound theory $\mathcal{S}_P{:=}S^1_2+\mathrm{RFN}(P)$, which simulates every consistent theory in the class

by the classical reflection construction (Krajíček and Pudlák [18, construction (a2) and proof of Theorem 2.1]). Specifically, for each fixed clocked verifier of a target theory $\mathcal{T}$, its Cook tableau encodings of $Con_{\mathcal{T}}(n)$ have polynomial size. The encoding construction gives polynomial-size S^1_2-proofs that their truth under every assignment implies $Con_{\mathcal{T}}(n)$. Optimality gives polynomial-size P-proofs of these tautologies when $\mathcal{T}$ is consistent. Verifying each such proof and using RFN(P) gives the required $\mathcal{S}_P$-proofs. In particular $\mathcal{S}_P$ simulates all its true single-sentence extensions, contradicting (2). Since $\mathcal{S}_P$ is finitely axiomatized and sequential, the same argument proves the assertion for the restricted class. □

The propositional consequence applies to every Cook–Reckhow proof system P, under the equivalent hypotheses of Corollary 3.13. Let $\tau_{k,n}$ be the standard Cook encoding of $Con_{U_k}(n)$, where $U_k{:=}S^1_2{+}\mathsf{Exp}{+}\phi_{BB}(k)$. For each fixed k this is a family of tautologies of size polynomial in n. Take the sound finite theory $\mathcal{S}_P{:=}S^1_2{+}\mathrm{RFN}(P)$. A P-proof of $\tau_{k,n}$ gives an $\mathcal{S}_P$-proof of $Con_{U_k}(n)$ with polynomial overhead, by verifying the proof and applying RFN(P). Condition (3) therefore implies that, for every sufficiently large k, the family $(\tau_{k,n})_n$ has no polynomial bound on its P-proof lengths. The collection uses the true exact Busy Beaver values; neither these values nor the threshold in k are supplied effectively. This consequence uses the global no-optimality hypothesis, whereas Theorem 3.11 transfers individual bounded-consistency lower bounds in a fixed arithmetic theory.

A diagonal argument using the Busy Beaver transfer theorem yields a single Busy Beaver family hard for every theory in a specified class:

Corollary 3.14. *Let $\mathfrak{C}$ be any nonempty class of sound theories $\mathcal{S}{\supseteq}S^1_2$ with polynomial-time decidable axioms, each having a true extension not simulated by $\mathcal{S}$. Then there is a nonconstructively chosen function $k{:}\mathbb{N}{\to}\mathbb{N}$ with unbounded range such that, for every $\mathcal{S}{\in}\mathfrak{C}$ and every constant $c{>}0$, $\mathcal{S} \nvdash^{n^c} Con_{U_{k(n)}}(n)$, where $U_j{:=}S^1_2{+}\mathsf{Exp}{+}\phi_{BB}(j)$. The hypothesis concerns only the theories in $\mathfrak{C}$; taking all eligible theories gives the universal hypothesis of Corollary 3.13.*

Proof. List all pairs $(\mathcal{S}_e, d_e)$ with $\mathcal{S}_e{\in}\mathfrak{C}$ and positive integer d_e, repeating each pair infinitely often. The list need not be effective. By Theorem 3.10, choose strictly increasing k_e such that the fixed family $(Con_{U_{k_e}}(n))_n$ has no polynomial bound on its $\mathcal{S}_e$-proof lengths. Then choose strictly increasing n_e, with n_e at least the written length of $\phi_{BB}(k_e)$, such that $\ell_{\mathcal{S}_e}(Con_{U_{k_e}}(n_e)){>}en_e^{d_e}$.

Put $k(n_e){:=}k_e$ and $k(n){:=}1$ elsewhere. An eventual proof bound An^c for a theory $\mathcal{S}{\in}\mathfrak{C}$ is contradicted at a repeated pair $(\mathcal{S}, d)$ with integer $d{\geq}c$, stage $e{>}A$, and n_e beyond the proposed onset. The strictly increasing k_e give unbounded range. □

This conclusion applies to classes of theories with hard true extensions even when other theories are universal simulators. It does not identify these classes with all non-optimal Cook–Reckhow proof systems: adjoining RFN(P) supplies a theory that proves P sound, but does not in general preserve P's proof lengths in the reverse direction (Pudlák [24, Section 4.1]).

3.4 Alternative Canonical Witnesses

Theorem 3.10 can also be stated with exact Busy Beaver values replaced by exact prefixes of Chaitin's halting probability or complete finite lists of Kolmogorov-random strings, using PA as the witness base. Fix optimal plain and prefix-free machines U_0, V with compiler simulations verified in PA. Let $\chi_\Omega(r)$ assert the first r bits of $\Omega_V{:=}\sum_{V(p)\downarrow} 2^{-|p|}$; they determine halting for programs of length at most r (Calude, Dinneen and Shu [8, Theorem 4]). For a fixed positive rational d, put $g_d(m){:=}\max\{0, \lceil m{-}d\log_2 m\rceil\}$. Let $\chi_R(r)$ conjoin the true statements $K_{U_0}(x){\geq}g_d(|x|)$ or their negations for every x with $2{\leq}|x|{\leq}r$, specifying a complete correct list. For related uses of finite complexity information, see Bienvenu et al. [6, Section 3.1 of the preprint].

Proposition 3.15 (Alternative canonical witnesses)**.** *Let $\mathcal{S}{\supseteq}S^1_2$ be sound with polynomial-time decidable axioms. Some true extension of $\mathcal{S}$ is not simulated if and only if $\mathcal{S}$ fails to simulate* PA$+\chi_\Omega(r)$ *for every sufficiently large r, and equivalently if and only if it fails to simulate* PA$+\chi_R(r)$ *for every sufficiently large r.*

Proof. It suffices to show that, for each consistent theory $\mathcal{T}$ with a polynomial-time ordinary proof predicate, either witness theory $\mathcal{V}_r$ proves $Con_{\mathcal{T}}$ for all sufficiently large r. The proof transformation in Theorem 3.10 then gives the forward implications. Conversely, if $\mathcal{V}_r$ is not simulated, take the true sentence $\phi{:=}Con_{\mathcal{V}_r}$. Verifying a $\mathcal{V}_r$-refutation and applying this axiom gives an $(\mathcal{S}{+}\phi)$-refutation of polynomially larger size, so simulation of $\mathcal{S}{+}\phi$ would imply simulation of $\mathcal{V}_r$.

For the Ω family, put $q_r{:=}2^{-r}\lfloor 2^r\Omega_V\rfloor$ and let ω_s be the increasing rational approximation obtained by enumerating halting V-programs. Using coded

rational comparisons, formalize $\chi_\Omega(r)$ as $(\exists s\ \omega_s{>}q_r){\wedge}(\exists j\ \forall s\ \omega_s{\leq}q_r{+}2^{-r}{-}2^{-j})$, with q_r fixed data. Compile $E_\mathcal{T}$ into a V-program $p_\mathcal{T}$ with the simulation verified in PA. For $r{\geq}|p_\mathcal{T}|$, choose a standard stage s with $\omega_s{>}q_r$. Consistency of $\mathcal{T}$ ensures that $p_\mathcal{T}$ is absent at that stage; both finite facts are verifiable in PA. Its later halting would add at least 2^{-r} to the sum, contradicting $\chi_\Omega(r)$. Thus $\mathsf{PA}{+}\chi_\Omega(r){\vdash}Con_\mathcal{T}$.

For the list family, let $D_t(m)$ contain the length-m outputs produced within t steps by U_0-programs shorter than $g_d(m)$. Fix a standard stage s_m by which all genuinely non-random strings of length m have appeared. The axiom $\chi_R(m)$ and these finite computations prove that, for $t{\geq}s_m$, $D_t(m)$ is exactly the set of non-random strings of length m. Given m, run $E_\mathcal{T}$; if it halts after t steps, output the first length-m string outside $D_t(m)$. Such a string exists because fewer than $2^{g_d(m)}{\leq}2^m$ programs are considered. This algorithm has a compiled description of length at most $2\lceil\log_2 m\rceil{+}c_\mathcal{T}{<}g_d(m)$ for all sufficiently large m. Halting at $t{\geq}s_m$ would therefore produce a short description of a string asserted random by $\chi_R(m)$, a contradiction. Consistency also gives a PA-proof that $E_\mathcal{T}$ does not halt before the fixed stage s_m. Formalizing these arguments in PA gives $\mathsf{PA}{+}\chi_R(m){\vdash}Con_\mathcal{T}$. □

Either family can therefore replace Busy Beaver values in Corollary 3.13, with witness base PA. The list argument also works for fixed deficiency. Finite transfers have bounds depending on the fixed witness and its consistency proof; no feasible bound uniform in r or reduction of the witness base is established. Hardness of a complete list does not imply hardness of its individual axioms.

4 Refutation of the Higher Relative Consistency Conjecture

In the first version of this paper, we conjectured under the name Higher Relative Consistency (HRC) that, for sound finitely axiomatized sequential $\mathcal{S}{\supseteq}S^1_2$, simulation of a true extension $\mathcal{S}{+}\phi$ requires $EA{\vdash}Con_\mathcal{S}{\rightarrow}Con_{\mathcal{S}+\phi}$. Hrubeš's self-reference construction refutes this necessary condition. We use his result as presented in Pudlák's survey [23, Theorem 3.4, p. 13].

Theorem 4.1 (Hrubeš)**.** *Let $\mathcal{S}$ be a consistent, finitely axiomatized sequential extension of S_2^1, and let $\mathcal{T} \supseteq \mathcal{S}$ be consistent with polynomial-time decidable axioms and standard proof predicates. There is a true sentence ρ, equivalent over EA to a Π_1 sentence, such that $\mathcal{T} \nvdash \rho$ while $\mathcal{S}$ has polynomial-size proofs of $Con_{\mathcal{S}+\rho}(n)$.*

To refute HRC, choose a sound finitely axiomatized sequential $\mathcal{S}$ containing EA and S_2^1 in the common definitional language of Section 2. For example, use a finite axiomatization of $I\Sigma_1$ with the chosen finite S_2^1 axioms adjoined (Hájek and Pudlák [13, Chapter I, Theorem 2.52]). Put $\mathcal{T} := \mathcal{S} + Con_{\mathcal{S}}$, which is sound and has polynomial-time decidable axioms. Theorem 4.1 gives a true sentence ρ such that $\mathcal{S}$ simulates $\mathcal{S}+\rho$ but $\mathcal{T} \nvdash \rho$. Since $\mathcal{S}$ contains EA, we may replace ρ by its Π_1 form, preserving these properties by fixed proof substitution.

Suppose $EA \vdash Con_{\mathcal{S}} \to Con_{\mathcal{S}+\rho}$. Then $\mathcal{T} \vdash Con_{\mathcal{S}+\rho}$. Formalized Σ_1-completeness gives $\mathcal{S} \vdash \neg\rho \to Prov_{\mathcal{S}}(\ulcorner \neg\rho \urcorner)$, hence $\mathcal{S} \vdash Con_{\mathcal{S}+\rho} \to \rho$ (Artemov and Beklemishev [5, Proposition 16]). Thus $\mathcal{T} \vdash \rho$, a contradiction. We have a simulated true extension without the EA relative-consistency certificate required by HRC.

The same example has no such certificate in S_2^1, since every S_2^1-proof yields an EA-proof under our language conventions. The same contradiction follows if the implication is provable in $\mathcal{S}$ itself, so Hrubeš's construction also rules out the characterization that $\mathcal{S}$ simulates $\mathcal{S}+\phi$ if and only if $\mathcal{S} \vdash Con_{\mathcal{S}} \to Con_{\mathcal{S}+\phi}$. The sufficient condition in Theorem 3.4 is unaffected: S_2^1-provable relative consistency yields simulation when $\mathcal{S}$ has short self-consistency proofs.

5 Kolmogorov Blindness

A natural refinement of the Busy Beaver phenomenon is obtained by considering Kolmogorov-random strings. Instead of an axiom specifying an exact Busy Beaver value, we consider axioms asserting that particular strings have no substantially shorter program descriptions. The set of exact Busy Beaver pairs $\{(k, BB(k)) : k \geq 1\}$ and the randomness set $R^{\log}$ defined below are both *immune*: each is infinite but contains no infinite computably enumerable subset.[4] Consequently, a fixed sound effectively axiomatized arithmetic theory

[4] An infinite enumeration of exact Busy Beaver pairs would give a computable upper bound on $BB(n)$ by finding a pair with first coordinate at least n, contradicting Aaronson [1, Proposition 1]. An infinite enumeration of $R^{\log}$ would allow a program given L to output a

can prove membership for only finitely many elements of either set. This supports the intuition that a theory has limited information about $R^{\log}$ and motivates a hardness conjecture. Immunity alone gives no proof-length lower bound.

5.1 Randomness and the Blindness Conjecture

Kolmogorov Blindness predicts hardness at individual finite scales, paralleling finite Busy Beaver transfer: exponential bounds beyond an onset common to true randomness axioms of each sufficiently large string length, and hence dense hardness because $R^{\log}$ has density tending to one. The bounds are also conjectured to hold after adjoining a true randomness axiom of the same length when the target remains random conditional on the added string. This is stronger than dense hardness obtained by padding $\phi_{BB}(k)$ with $|x|>0$: all those padded axioms are equivalent, so adjoining one gives short bounded-consistency proofs for the others over a finitely axiomatized sequential base.

The motivation comes from random DNF hardness candidates (Khaniki [15, Introduction]) and proof-complexity generators (Alekhnovich et al. [2]; Krajíček [16]). Here incompressibility specifies the indices and the bounds concern finite consistency parameters. The conditional-randomness requirement is a further conjectural strengthening proposed here.

Fix an optimal machine U and a positive rational constant d. For each string $x\in\{0,1\}^*$, let $K_U(x)$ denote the plain Kolmogorov complexity of x, that is, the length of the shortest program p such that $U(p)=x$. Logs are to base two. Put $g(m):=\max\{0,\lceil m-d\log_2 m\rceil\}$ for $m\geq 2$ and $R^{\log}:=\{x:|x|\geq 2,\ K_U(x)\geq g(|x|)\}$. For rational d, the integer threshold g is computable by integer comparisons; no real-number parameter enters the arithmetic sentence. The rounding affects only constant factors in counting estimates. Fewer than $2^{g(m)}$ programs have length below $g(m)$, so the fraction outside $R^{\log}$ at length m is $O(m^{-d})$. A fixed-deficiency randomness predicate need not have density tending to one, which is why the density property below uses logarithmic deficiency.

For true $x\in R^{\log}$, let $H_{\mathcal{S}}(x)$ mean that no polynomial bounds the $\mathcal{S}$-proof lengths of $Con_{\mathcal{S}+(x\in R^{\log})}(n)$. This fixes x before varying n. Truth is essential:

member of length at least L, with description length $O(\log L)$, contradicting its logarithmic-deficiency randomness for large L; this is the compression argument underlying Chaitin's incompleteness theorem (see Shen [26, Theorem 4]).

a false randomness axiom makes its extension inconsistent, so a sound $\mathcal{S}$ eventually has no proofs of its finite consistency, giving vacuous hardness.

Fix a compatible universal conditional machine. For $|x|=|y|=m$, say that x remains random conditional on y when $K_U(x|y)\geq g(m)$. Write $x\perp_{R^{\log}} y$ when $x, y\in R^{\log}$ and each remains random conditional on the other. For the empty string ε, set $K_U(x|\varepsilon):=K_U(x)$ and let $\Delta(\varepsilon)$ add no axiom; for a nonempty string y, let $\Delta(y)$ be $y\in R^{\log}$.

Conjecture 5.1. *(**Kolmogorov Blindness**) For every sound finitely axiomatized sequential $\mathcal{S}\supseteq S^1_2$ with its fixed proof presentation, there exist a positive rational $\eta_{\mathcal{S}}$ and a cutoff $k^{\mathcal{S}}_R$ such that for every $m>k^{\mathcal{S}}_R$ there exists $N^{\mathcal{S}}_{R,m}$ such that, for every true $x\in R^{\log}\cap\{0,1\}^m$, every y that is either ε or a string of length m with $y\in R^{\log}$ true, and every $n>N^{\mathcal{S}}_{R,m}$,*

$$\mathcal{S}+\Delta(y)\Big|\!\!\overline{}^{2^{\eta_{\mathcal{S}} n}} Con_{\mathcal{S}+(x\in R^{\log})}(n) \quad\Longrightarrow\quad K_U(x|y)<g(m).$$

The rate $\eta_{\mathcal{S}}$ may depend on the fixed proof encoding, the machine U and the deficiency coefficient d. The onset $N^{\mathcal{S}}_{R,m}$ is common to all strings of length m, but no small or computable bound for it is asserted. Encoding dependence is essential: repeating every proof bit r times reduces the unpadded consistency parameter to approximately n/r, while exhaustive finite verification (Pudlák [24, Section 4.3]) gives proofs of size $2^{Cn/r}$ up to polynomial factors, for a fixed C. Thus a rate arbitrarily close to one cannot hold for all such encodings. General polynomial recodings can also change an exponential scale to a stretched exponential one. We use the ordinary local encoding of Section 2; no implication from SETH for satisfiability is asserted.

5.2 Hardness and Density Consequences

Fix a sound finitely axiomatized sequential $\mathcal{S}\supseteq S^1_2$, with the machines and proof presentation above.

Definition 5.2 (Density KH and Strong KH)**. Density KH** holds if $\lim_{m\to\infty} 2^{-m}|\{x\in\{0,1\}^m : x\in R^{\log} \text{ and } H_{\mathcal{S}}(x)\}|=1$. **Strong KH** holds if there is $k^{\mathcal{S},U,d}_{\mathrm{KH}}$ such that every true $x\in R^{\log}$ with $|x|>k^{\mathcal{S},U,d}_{\mathrm{KH}}$ satisfies $H_{\mathcal{S}}(x)$.

Density KH permits a vanishing fraction of easy true axioms and specifies no convergence rate or computable onset. The next theorem derives both properties and the single-axiom and pairwise finite exponential bounds from

Blindness. The pairwise target is the consistency of $\mathcal{S}+(x\in R^{\log})$, not of the joint extension; the assertion allows proof shortening above the prescribed lower bound.

Theorem 5.3 (Hardness consequences of Blindness)**.** *Assume Kolmogorov Blindness at $\mathcal{S}$. Let $m>k_R^{\mathcal{S}}$ and $n>N_{R,m}^{\mathcal{S}}$.*

1. *For every true $x\in R^{\log}\cap\{0,1\}^m$, $\mathcal{S}\not\vdash^{2^{\eta_{\mathcal{S}}n}} Con_{\mathcal{S}+(x\in R^{\log})}(n)$. At least $(1-O(m^{-d}))2^m$ strings of length m satisfy membership and this bound.*
2. *For $x,y\in\{0,1\}^m$ with $x\perp_{R^{\log}}y$, $\mathcal{S}+(y\in R^{\log})\not\vdash^{2^{\eta_{\mathcal{S}}n}} Con_{\mathcal{S}+(x\in R^{\log})}(n)$. The same bound holds with x,y interchanged. At least $(1-O(m^{-d}))2^{2m}$ ordered pairs satisfy the mutual randomness conditions and both bounds for all sufficiently large such m.*
3. *Strong KH and Density KH hold. One may take $k_{\mathrm{KH}}^{\mathcal{S},U,d}:=k_R^{\mathcal{S}}$.*

Proof. For (1), take $y=\varepsilon$ in Blindness. A proof within the stated bound would give $K_U(x)<g(m)$, contradicting $x\in R^{\log}$. Counting programs shorter than $g(m)$ bounds the exceptional fraction by $O(m^{-d})$.

For (2), apply Blindness to (x,y) and (y,x), using the two conditional incompressibility bounds. For each fixed y, fewer than $2^{g(m)}$ strings x violate $K_U(x|y)\geq g(m)$. The same estimate holds in the opposite direction. The two unconditional membership failures also affect only an $O(m^{-d})$ fraction of ordered pairs. A union bound proves the density estimate. Symmetry supplies both hardness bounds with the same onset.

For (3), fix a true $x\in R^{\log}$ above the threshold. Its eventual lower bound $2^{\eta_{\mathcal{S}}n}$ exceeds every polynomial because $\eta_{\mathcal{S}}>0$, so Strong KH holds. Counting the true members of $R^{\log}$ then gives Density KH. □

These density conclusions use the same strings or pairs for every n beyond the common onset. The encoding of a randomness assertion uses the fixed-template observation of Calude and Jürgensen [9, proof of Theorem 5.2]: the assertion adds only constant overhead to the written representation of x. Here $x\in R^{\log}$ uses a fixed predicate that computes its threshold from $|x|$, so no separate threshold numeral is needed. By *random tautologies* we mean the standard propositional consistency encodings indexed by $x\in R^{\log}$ (Krajíček and Pudlák [18, Section 3]). Density is among these indices. The Cook encodings have size polynomial in $m+n$, and S_2^1 has proofs of polynomial size that their truth under every assignment implies the corresponding consistency

statements. For a fixed Cook–Reckhow proof system P whose soundness $\mathcal{S}$ proves, verifying a P-proof and applying RFN(P) therefore gives an $\mathcal{S}$-proof of the consistency statement with polynomial overhead. Blindness at such an $\mathcal{S}$ therefore yields hardness in P, with the lower bounds adjusted for this overhead. Universal Blindness allows $\mathcal{S}{:=}S^1_2{+}\mathrm{RFN}(P)$. The onsets may be large and need not be computable in m.

Universal Blindness therefore implies the universal Density KH property and, by Corollary 3.13, the nonexistence of a length-optimal propositional proof system and **NP≠coNP**. This conclusion uses all eligible finite bases; a property at one fixed theory is not enough. No converse from Strong KH, Density KH or the single-axiom exponential property to Blindness is established. Applying a single-axiom assertion at $\mathcal{S}{+}(y{\in}R^{\log})$ changes the target to the joint extension and does not give the pairwise conclusion above.

5.3 Factual Nonaccess and Certificate Obstructions

Conditional incompressibility says that the bits of y do not describe x briefly; it does not by itself control the information in $Con_{\mathcal{S}+(y\in R^{\log})}$. Universal Blindness supplies a nonaccess consequence through an auxiliary finite theory, even when the original $\mathcal{S}$ has infinitely many polynomial-time decidable axioms.

Definition 5.4 (Conditional Factual No Access property)**.** Fix a sound theory $\mathcal{S}{\supseteq}S^1_2$ with polynomial-time decidable axioms and the proof presentation of Section 2. We say that **Conditional Factual No Access** holds for $\mathcal{S}, R^{\log}$ if there exists $k^{\mathcal{S}}_{R^{\log},\mathrm{access}}$ such that, for every $m{>}k^{\mathcal{S}}_{R^{\log},\mathrm{access}}$ and every $x, y{\in}\{0,1\}^m$ satisfying $x{\perp_{R^{\log}}}y$, one has $EA{+}Con_{\mathcal{S}+(y\in R^{\log})}{\nvdash}x{\in}R^{\log}$ and $EA{+}Con_{\mathcal{S}+(x\in R^{\log})}{\nvdash}y{\in}R^{\log}$.

For the next lemma, let R denote either $R^{\log}$ or fixed-deficiency randomness $K_U(x){\geq}|x|{-}c_U$, with its fixed presentation.

Lemma 5.5. *For each fixed x, $EA{\vdash}Con_{\mathcal{S}+(x\in R)}{\to}x{\in}R$. Consequently, for any sentence η, if $EA{\vdash}\eta{\to}Con_{\mathcal{S}+(x\in R)}$, then $EA{+}\eta{\vdash}x{\in}R$.*

Proof. If $R{=}R^{\log}$ and $|x|{<}2$, its definition gives a fixed S^1_2-proof of nonmembership; combining it with the added axiom gives a refutation. Otherwise let $h_R(x)$ be the integer threshold in the randomness assertion. Falsity is witnessed by a program p of length below $h_R(x)$ and a complete halting

computation h of $U(p)$ with output x. Checking the format and transitions of h, the output, and the length of p is a polynomial-time computation. The computation-writing construction of Section 2 gives an S^1_2-proof of the verified facts and hence, by existential introduction, of $\neg(x{\in}R)$. Its correctness is formalizable in EA, which can represent the computation and the resulting proof. Embed this proof into $\mathcal{S}{+}(x{\in}R)$ by the fixed proof embedding of Section 2, and use the added axiom to obtain a refutation. Thus EA proves $\neg(x{\in}R){\to}\neg Con_{\mathcal{S}+(x\in R)}$, without assuming that $\mathcal{S}$ contains EA. Contraposition proves the first assertion; composition proves the second. □

Taking $\eta{=}Con_{\mathcal{S}}$, or $\eta{=}Con_{\mathcal{S}+(y\in R)}$, gives the single-axiom and pairwise certificate obstructions by contraposition. These implications alone give no proof-length lower bounds. Chaitin's theorem gives a finite cutoff for randomness facts provable in the sound theory $EA{+}Con_{\mathcal{S}}$, but no comparison with a hardness threshold is established.

Proposition 5.6 (Universal Blindness implies factual nonaccess)**.** *Suppose* ***Kolmogorov Blindness*** *holds for every sound finitely axiomatized sequential theory extending S^1_2. Then every sound $\mathcal{S}{\supseteq}S^1_2$ with polynomial-time decidable axioms and the fixed proof presentation satisfies* ***Conditional Factual No Access*** *for $R^{\log}$. Hence, for all sufficiently long mutually random pairs, neither $EA{\vdash}Con_{\mathcal{S}+(y\in R^{\log})}{\to}Con_{\mathcal{S}+(x\in R^{\log})}$ nor the implication with x, y interchanged holds.*

Proof. Fix $\mathcal{S}$. Choose a fixed sound finitely axiomatized sequential base A containing EA and S^1_2 in the common definitional language, as in Section 4. Using the uniform arithmetization of the axiom family, put $\chi_{\mathcal{S}}{:=}\forall z\,(z{\in}R^{\log}{\to}Con_{\mathcal{S}+(z\in R^{\log})})$ and $\mathcal{T}{:=}A{+}\chi_{\mathcal{S}}$. The sentence $\chi_{\mathcal{S}}$ is true because $\mathcal{S}$ is sound. Thus $\mathcal{T}$ is sound, finitely axiomatized and sequential, and satisfies Blindness by hypothesis.

Set $k^{\mathcal{S}}_{R^{\log},\mathrm{access}}{:=}k^{\mathcal{T}}_{R}$. Suppose $x{\perp_{R^{\log}}}y$ has length above this threshold and $EA{+}Con_{\mathcal{S}+(y\in R^{\log})}{\vdash}(x{\in}R^{\log})$. The finite theory $Q{:=}\mathcal{T}{+}(y{\in}R^{\log})$ proves $Con_{\mathcal{S}+(y\in R^{\log})}$ using $\chi_{\mathcal{S}}$, and therefore proves $(x{\in}R^{\log})$. Replacing the added axiom $(x{\in}R^{\log})$ by this fixed Q-proof translates $(\mathcal{T}{+}(x{\in}R^{\log}))$-refutations into Q-refutations with polynomial growth, verified in S^1_2. Since Q is finitely axiomatized and sequential, its polynomial-size self-consistency proofs (Pudlák [23, Theorem 3.1(2)]) then give polynomial-size Q-proofs of $Con_{\mathcal{T}+(x\in R^{\log})}(n)$. This contradicts Theorem 5.3(2) for the ambient theory $\mathcal{T}$, since its positive exponential rate exceeds every polynomial for sufficiently large n. Interchanging

x, y proves the other nonaccess assertion. The relative-consistency conclusions follow by contraposition of Lemma 5.5. □

This implication uses Blindness for the auxiliary theory, not just at the original $\mathcal{S}$. Conversely, factual nonaccess alone gives certificate obstructions, with no established implication to exponential hardness. At a fixed $\mathcal{S}$, Blindness also excludes S_2^1-proofs of $Con_{\mathcal{S}+(y\in R^{\log})} \to Con_{\mathcal{S}+(y\in R^{\log})+(x\in R^{\log})}$ for mutually random x, y above its threshold: Theorem 3.4 would give short consistency proofs for the joint extension, hence for $\mathcal{S}+(x\in R^{\log})$, contradicting the pairwise bound.

Changes of randomness convention preserve KH under suitable provable translations. Fix a sound finitely axiomatized sequential $\mathcal{S}\supseteq S_2^1$ with its standard proof presentation. If $\mathcal{S}\vdash\beta_x\to\alpha_x$ for each sufficiently long x, non-simulation of $\mathcal{S}+\alpha_x$ implies non-simulation of $\mathcal{S}+\beta_x$: replacing uses of α_x by its fixed derivation from β_x gives a polynomial proof translation. Thus Strong KH transfers to stronger true axioms, and Density KH transfers when the stronger axioms are true for a fraction of strings tending to one; provably equivalent axiom families have equivalent KH properties. Standard comparisons between optimal plain machines and between plain and prefix-free complexity require constant or logarithmic threshold adjustments (Shen [26, Sections 3–4 and 35]). When $\mathcal{S}$ verifies the corresponding program translations, the schema asserting either KH property for every positive rational logarithmic-deficiency coefficient is therefore invariant under these changes. This does not establish equivalence at an identical fixed threshold or extend automatically to conditional, time-bounded or oracle-relative randomness; finite-scale bounds additionally require proof translations uniform in x.

6 Effective Density and Randomness

We prove a quantitative equivalence between dense hardness and hardness above a corresponding randomness threshold, for fixed Π_1 axiom families with computable proof-length bounds and a computable common onset. Counting gives one direction; in the other, an exception's position in an effective enumeration gives a short description. The coding argument is the standard incompressibility method (Shen [26, Theorems 2(d) and 9]); its application here covers whole hard tails and pairs whose bounds hold after adjoining the other axiom. All complexities below are plain complexities for the fixed optimal machines.

Theorem 6.1 (Effective density and incompressibility)**.** *Let $L, f{:}\mathbb{N}\to\mathbb{N}$ be total computable functions with $0\le f(m)\le L(m)$, and let $G_m\subseteq\{0,1\}^{L(m)}$. Suppose the complements $E_m{:=}\{0,1\}^{L(m)}\setminus G_m$ are uniformly computably enumerable from m. The following are equivalent:*

1. *There is a constant C such that $|E_m|\le C2^{L(m)-f(m)}$ for all sufficiently large m.*

2. *There is an integer constant c such that, for all sufficiently large m and every $z\in\{0,1\}^{L(m)}$, $K_U(z|m)\ge L(m)-f(m)+c \implies z\in G_m$.*

Proof. For (1)⇒(2) choose an integer $b\ge 0$ with $C\le 2^b$. Given m, enumerate E_m without repetition. An element of E_m is described by its zero-based position in this enumeration, using exactly $L(m)-f(m)+b$ bits. The decoder computes this length from m and outputs the element at that position when it appears. It need not detect when the enumeration is complete. Optimality of the fixed conditional machine gives $z\in E_m \implies K_U(z|m)\le L(m)-f(m)+b+a$ for a fixed decoder constant a. Take $c{=}b{+}a{+}1$. Conversely, (2) puts every exception below complexity $L(m)-f(m)+c$. Fewer than $2^{L(m)-f(m)+c}$ strings have such descriptions, proving (1). This reverse direction does not require enumerability. □

The loss from conditioning on m can be avoided when the description length determines m up to bounded ambiguity.

Proposition 6.2 (Unconditional form)**.** *Let L, f, G_m and E_m be as in Theorem 6.1, and suppose that there are integers r and m' such that at most r integers $m\ge m'$ share each value of $L(m)-f(m)$. Then condition (1) of Theorem 6.1 is equivalent to:*

(3) There is an integer constant c such that, for all sufficiently large m and every $z\in\{0,1\}^{L(m)}$, $K_U(z)\ge L(m)-f(m)+c \implies z\in G_m$.

Proof. Assume (1) for $m\ge m_0$ with constant C, where we may take $m_0\ge m'$, and choose an integer $b\ge 0$ with $rC\le 2^b$. For each j let V_j be the union of the sets E_m with $m\ge m_0$ and $L(m)-f(m)+b{=}j$. The relation $z\in V_j$ is enumerable, because L and f are computable and the E_m are uniformly computably enumerable, and $|V_j|\le rC2^{j-b}\le 2^j$. By Shen [26, Theorem 2(d)] there is a constant a with $K_U(z)\le j+a$ for all $z\in V_j$. Hence every $z\in E_m$ with $m\ge m_0$ satisfies $K_U(z)\le L(m)-f(m)+b+a$; take $c{=}b{+}a{+}1$. Conversely, (3) puts every exception below unconditional complexity $L(m)-f(m)+c$, and fewer than $2^{L(m)-f(m)+c}$ strings have such descriptions. □

For the hardness application, fix a sound, finitely axiomatized sequential $\mathcal{S}\supseteq S_2^1$ with its ordinary proof predicate, a fixed Π_1 formula $\psi(x)$, a total computable bound $b{:}\mathbb{N}\to\mathbb{N}$, and a total computable onset $H(m)$. Counterexamples to the truth of ψ can be enumerated; this applies to the randomness predicates above. For example, $b(n){:=}\lfloor 2^{\eta n}\rfloor$ for a fixed positive rational η prescribes an exponential scale. The result also applies to polynomial or other computable bounds; it does not prove that the chosen bounds are hard. Let A_m consist of the strings $x{\in}\{0,1\}^m$ such that $\psi(x)$ is true and $\forall n{>}H(m)\quad \ell_{\mathcal{S}}\big(Con_{\mathcal{S}+\psi(x)}(n)\big){>}b(n)$. Let B_m consist of concatenations xy with $x,y{\in}A_m$ for which, for every $n{>}H(m)$, both

$$\ell_{\mathcal{S}+\psi(y)}\big(Con_{\mathcal{S}+\psi(x)}(n)\big){>}b(n),$$
$$\ell_{\mathcal{S}+\psi(x)}\big(Con_{\mathcal{S}+\psi(y)}(n)\big){>}b(n).$$

Thus A_m requires truth and every bound after the common onset; B_m additionally requires both bounds when the other axiom is adjoined. Concatenations have equal-length halves, and these sets are fixed before n varies.

Corollary 6.3 (Dense hard tails and random instances)**.** *For the fixed data just specified and any total computable f with $0{\le}f(m){\le}m$, the following are equivalent:*

1. *The exceptional fractions for both A_m and B_m are $O(2^{-f(m)})$.*
2. *There are constants c_1,c_2 such that, for all sufficiently large m and all $x,y{\in}\{0,1\}^m$,*

$$\begin{aligned} K_U(x|m){\ge}m{-}f(m){+}c_1 \quad &\Longrightarrow \quad x{\in}A_m,\\ K_U(xy|m){\ge}2m{-}f(m){+}c_2 \quad &\Longrightarrow \quad xy{\in}B_m. \end{aligned}$$

In particular the first implication gives the truth of $\psi(x)$, and the second gives the truth of both $\psi(x)$ and $\psi(y)$ together with both no-help bounds.

Proof. Both complements are uniformly computably enumerable from m. Write $\psi(x)$ in the form $\forall z\, R(x,z)$ with decidable R. For each string of length m, search for a counterexample to $R(x,z)$ and, in parallel, search over integers $n{>}H(m)$ and candidate $\mathcal{S}$-proofs of $Con_{\mathcal{S}+\psi(x)}(n)$ of length at most $b(n)$. Output the string when either search finds a witness. The functions H and b are computable, and the ordinary proof predicate is decidable uniformly

in x, so every exception is eventually listed and no other string is listed. For pairs, run the corresponding searches for both truth assertions, both base-theory bounds and both bounds after adjoining the other axiom. These searches need not be efficient or recognize when an enumeration is complete. Apply Theorem 6.1 with $L(m)=m$ and $L(m)=2m$, respectively. □

The bound for A_m in (1) follows from the bound for B_m with the same f: each $x\notin A_m$ gives 2^m concatenations $xy\notin B_m$, so $|\{0,1\}^m\setminus A_m|\leq 2^{-m}|\{0,1\}^{2m}\setminus B_m|$. Condition (1) may therefore be stated for B_m alone.

Enumerability follows from the hypotheses, but sparsity remains a premise. For randomness axioms, counting bounds the false instances by $O(m^{-d})$; the fraction of true instances violating a hardness bound is unknown. For example, $\psi(x)\equiv 0=0$ and $b(n)=2^{\lfloor n/2\rfloor}$ make every string exceptional, since short self-consistency proofs give polynomial-size proofs for this trivial extension.

Polynomially small exceptions. An exceptional fraction $O(m^{-a})$, for positive rational a, is equivalent to coverage of every sufficiently long x with $K_U(x)\geq m-a\log_2 m+c$, for a suitable constant c. For pairs the threshold is $K_U(xy)\geq 2m-a\log_2 m+c'$. Each equivalence applies separately to A_m or B_m.

To obtain unconditional complexity, apply Proposition 6.2 with $f(m)=\min\{m,\lfloor a\log_2 m\rfloor\}$. For sufficiently large m, $f(m+j)-f(m)\leq 1$ for $j\in\{1,2\}$, so $m-f(m)$ is nondecreasing with fibers of size at most two, while $2m-f(m)$ is strictly increasing. Both meet the multiplicity hypothesis. In particular, for $\delta<a$, all sufficiently long strings with $K_U(x)\geq m-\delta\log_2 m$ lie in A_m, and the analogous joint threshold gives membership in B_m. Joint incompressibility is distinct from the separate conditions $x\perp_{R^{\log}} y$. For randomness axioms the attainable rate has the following limitation.

Proposition 6.4 (Limitation of the density rate)**.** *If $\{x\in\{0,1\}^m : x\notin R^{\log}\}$ has density $O(m^{-a})$ for a real $a>0$, then $a\leq d$. Consequently the preceding density argument, with $\psi(x):=x\in R^{\log}$, guarantees coverage by this rate comparison for $\delta<a\leq d$; the endpoint $\delta=d$ and Strong KH are not supplied by the rate comparison alone.*

Proof. Suppose $a>d$ and choose a rational b with $d<b<a$. For $j\geq 2$, a fixed machine maps each j-bit string y to $x=y0^{\lceil b\log_2 j\rceil}$, of length $m_j:=j+\lceil b\log_2 j\rceil$. For a constant c_b, plain complexity satisfies $K_U(x)\leq j+c_b$: the description is y, whose length tells the decoder how much padding to add (Shen [26, Theorem 2(c)]). But $g(m_j)-j=(b-d)\log_2 j+O(1)\longrightarrow\infty$. Thus, for all sufficiently

large j, all 2^j distinct outputs lie outside $R^{\log}$. Their fraction at length m_j is $2^{-\lceil b \log_2 j \rceil} = \Omega(m_j^{-b})$, contradicting an $O(m_j^{-a})$ bound because $a > b$. For the consequence, every x outside $R^{\log}$ also lies outside A_m. □

The equivalence supplies neither the density estimate nor a computable onset. Blindness supplies common onsets but does not assert they are computably bounded; density separately at each n need not give a dense set satisfying all $n > H(m)$. Density KH also specifies no polynomial rate, so the criterion does not turn it into Strong KH. Effectiveness is essential: without it, even density $1-2^{-m}$ need not cover all incompressible strings, as shown by omitting one string of complexity at least m at each length.

7 Remarks on Provability

Corollary 3.13 characterizes the nonexistence of a length-optimal proof system by universal Busy Beaver non-simulation. What can theories prove about these assertions and Kolmogorov hardness? Non-simulation requires lower-bound instances beyond every onset for every polynomial exponent; Blindness asserts the stronger eventual exponential bounds.

The familiar self-unprovability argument concerns uniform assertions. It cannot prohibit individual lower-bound proofs: Hrubeš's lemma gives polynomial-size proofs, for each standard n, that a chosen non-theorem has no proof shorter than $\log_2 n$, including when the proving theory is the theory whose proofs are bounded. The following proposition makes the uniform restriction precise.

Proposition 7.1. *Let $\mathcal{T} \supseteq EA$ be consistent, with polynomial-time decidable axioms.*

1. *Let* $\mathrm{Blind}(\mathcal{T})$ *be the sentence* $\exists b \forall x \neg Prov_{\mathcal{T}}(\ulcorner K_U(\bar{x}) > \bar{b} \urcorner)$. *Then* $EA \vdash \mathrm{Blind}(\mathcal{T}) \to Con_{\mathcal{T}}$, *and therefore* $\mathcal{T} \nvdash \mathrm{Blind}(\mathcal{T})$. *If* $\mathcal{T}$ *is sound,* $\mathrm{Blind}(\mathcal{T})$ *is true.*

2. *Let f be a clocked polynomial-time function and r a polynomial such that EA proves that, for every n, $f(n)$ is the code of a sentence of size at most $r(n)$. Let* $\mathrm{NonSim}(\mathcal{T}, f)$ *be the arithmetized statement that for every c and every N there is $n \geq N$ such that the sentence coded by $f(n)$ has no $\mathcal{T}$-proof of size at most n^c. Then* $EA \vdash \mathrm{NonSim}(\mathcal{T}, f) \to Con_{\mathcal{T}}$, *and therefore* $\mathcal{T} \nvdash \mathrm{NonSim}(\mathcal{T}, f)$.

Proof. Both parts are explosion arguments formalized in EA. For part (1), a $\mathcal{T}$-proof of 0=1 extends to a $\mathcal{T}$-proof of $K_U(\bar{0})>\bar{b}$ for every b, so $\neg Con_{\mathcal{T}}$ refutes Blind($\mathcal{T}$). Truth for sound $\mathcal{T}$ is Chaitin's theorem [10]; see also Shen [26, Theorem 4]. For part (2), a $\mathcal{T}$-proof w of 0=1 extends, for every n, to a $\mathcal{T}$-proof of the sentence coded by $f(n)$ of size at most $q(|w|+n)$ for a fixed polynomial q bounding proof composition and the explosion step. The theory EA verifies this construction because it proves that $f(n)$ codes a sentence of size at most $r(n)$. For fixed w, choose c and N such that $q(|w|+n)\leq n^c$ for every $n\geq N$; then every such n has a proof of size at most n^c, so $\neg Con_{\mathcal{T}}$ refutes NonSim($\mathcal{T}, f$). In both parts the second incompleteness theorem for $\mathcal{T}$ gives the unprovability. □

For $\mathcal{T}=\mathcal{S}\supseteq EA$, the function $f(n)=\ulcorner Con_{\mathcal{S}+\phi}(\bar{n})\urcorner$ satisfies part (2)'s internal well-formedness hypothesis by fixed-formula numeral substitution. Thus $\mathcal{S}$ cannot prove its own uniform non-simulation assertion for a fixed extension, including any fixed randomness axiom. The same holds for $f(n)=\ulcorner Con_{S^1_2+\mathsf{Exp}+\phi_{BB(k)}}(\bar{n})\urcorner$ at each fixed k. This leaves open provability in a stronger theory. The classical, non-quantitative sentence Blind($\mathcal{T}$) in part (1) is distinct from Conjecture 5.1.

A stronger ambient theory nevertheless has its own randomness limitations. For a sound effectively axiomatized $\mathcal{S}$, let A_i be the set of true Π_i sentences and $\mathcal{T}_i:=\mathcal{S}+A_i$, with $A_0:=\emptyset$. Its theorems are enumerable relative to A_i. Chaitin's search argument therefore shows that $\mathcal{T}_i$ proves only finitely many assertions $K^{A_i}(x)\geq g(|x|)$: otherwise an oracle program given L could find such an x of length above L with description length $O(\log L)$, contradicting soundness. Counting still gives density one for these oracle-random strings. This suggests an obstacle to proving hardness for a weaker theory, but proves no such unprovability: inability to establish individual randomness facts does not exclude a uniform theorem about the weaker theory's proof lengths.

Self-unprovability does not establish that the universal Busy Beaver property, or equivalently the nonexistence of a length-optimal proof system, is unprovable in a specified ambient theory such as PA. Nor is inability to prove unprovability a distinctive feature of these hardness assertions. For any $\mathcal{T}$ as in Proposition 7.1 and any fixed sentence θ, formalized explosion gives $EA\vdash\neg Prov_{\mathcal{T}}(\ulcorner\theta\urcorner)\rightarrow Con_{\mathcal{T}}$. Hence $\mathcal{T}\nvdash\neg Prov_{\mathcal{T}}(\ulcorner\theta\urcorner)$ by the second incompleteness theorem (Artemov and Beklemishev [5, Theorem 19]), regardless of the subject matter of θ.

Unifying hardness criteria. These characterizations parallel connections between circuit size and KT complexity (Allender et al. [3, Theorem 11], with the authors' clarification [4]) and between one-way functions and the average-case hardness of time-bounded Kolmogorov complexity K^t (Liu and Pass [19, Theorem 1.1]). It would be desirable to formulate a single criterion in terms of information constraints whose instances imply proof-length lower bounds, circuit lower bounds, and average-case hardness of inverting candidate one-way functions. Such a criterion would need precise resource bounds and, for average-case claims, specified distributions and success probabilities.

A criterion based on the absence of an efficiency certificate encounters a proof-theoretic obstacle. Fix a certificate theory B satisfying the hypotheses of Proposition 7.1, and let E_τ and H_τ be arithmetical assertions of efficiency and hardness for a precisely specified task τ. Consider a proposed rule $\neg Prov_B(\ulcorner E_\tau \urcorner) \to H_\tau$. In any ambient theory proving this rule, establishing its premise already suffices to prove the desired hardness. If that ambient theory is B, the preceding argument prevents it from proving the premise; a stronger theory may do so. Mechanically formulating such a schema remains possible, but its usefulness depends on how its instances are justified and their premises established. The refutation of HRC in Section 4 gives a concrete caution: failure of a specified relative-consistency certificate need not imply non-simulation. A proposed unification must therefore justify both the passage from a missing certificate to hardness and a way to establish that the certificate is missing.

8 Conclusion

The paper's main structural result characterizes the nonexistence of a length-optimal propositional proof system by universal Busy Beaver non-simulation (Corollary 3.13). It combines Busy Beaver transfer with the Krajíček–Pudlák correspondence. The transfer equates, at each fixed theory, the existence of a hard true extension with eventual Busy Beaver non-simulation; finite lower bounds transfer with explicit changes of scale. Relative consistency provable in S^1_2 yields simulation when the proving theory has short self-consistency proofs. The effective density theorem characterizes coverage of incompressible indices by specified true hard tails and pairs, under computable bounds and a common computable onset.

Kolmogorov Blindness proposes exponential hardness for sufficiently long randomness axioms, including after adjoining another true random axiom when conditional randomness is preserved. It implies Strong KH, Density KH and their finite-scale forms; universally, it also implies factual nonaccess. The density theorem does not prove the sparsity premise or the computable onset needed to derive these hardness assertions. Pudlák's consistency-strength conjecture remains separate.

Hrubeš's construction refutes the earlier characterization by weak-base relative consistency. Whether ordinary EA relative consistency suffices for simulation remains open here, as does a characterization of individual true extensions.

References

[1] Scott Aaronson, *The busy beaver frontier*, SIGACT News **51** (2020), no. 3, 32–54.

[2] Michael Alekhnovich, Eli Ben-Sasson, Alexander A. Razborov, and Avi Wigderson, *Pseudorandom generators in propositional proof complexity*, SIAM Journal on Computing **34** (2004), no. 1, 67–88.

[3] Eric Allender, Harry Buhrman, Michal Koucký, Dieter van Melkebeek, and Detlef Ronneburger, *Power from random strings*, SIAM Journal on Computing **35** (2006), no. 6, 1467–1493.

[4] Eric Allender and Dieter van Melkebeek, *A comment regarding KT and circuit size*, Author's clarification.

[5] Sergei N. Artemov and Lev D. Beklemishev, *Provability logic*, Handbook of Philosophical Logic (Dov M. Gabbay and Franz Guenthner, eds.), vol. 13, Springer, 2 ed., 2005, pp. 189–360.

[6] Laurent Bienvenu, Andrei Romashchenko, Alexander Shen, Antoine Taveneaux, and Stijn Vermeeren, *The axiomatic power of Kolmogorov complexity*, Annals of Pure and Applied Logic **165** (2014), no. 9, 1380–1402.

[7] Samuel R. Buss, *Bounded arithmetic*, Lecture notes, Bibliopolis, 1986.

[8] Cristian S. Calude, Michael J. Dinneen, and Chi-Kou Shu, *Computing a glimpse of randomness*, CDMTCS Research Report 167, University of Auckland, 2002, November 2001, revised January 2002. `https://www.cs.auckland.ac.nz/~mjd/misc/167omega.pdf`.

[9] Cristian S. Calude and Helmut Jürgensen, *Is complexity a source of incompleteness?*, Advances in Applied Mathematics **35** (2005), no. 1, 1–15.

[10] Gregory J. Chaitin, *Information-theoretic limitations of formal systems*, JACM **21** (1974), no. 3, 403–424.

[11] Stephen Cook and Robert Reckhow, *The relative efficiency of propositional proof systems*, J. Symb. Log. **44** (1979), 36–50.

[12] Anton Freund and Fedor Pakhomov, *Short proofs for slow consistency*, Notre Dame Journal of Formal Logic **61** (2020), no. 1, 31–49.

[13] Petr Hájek and Pavel Pudlák, *Metamathematics of first-order arithmetic*, Perspectives in Mathematical Logic, Springer, Berlin, 1993, Second printing 1998.

[14] Erfan Khaniki, *New relations and separations of conjectures about incompleteness in the finite domain*, The Journal of Symbolic Logic **87** (2022), no. 3, 912–937.

[15] ———, *Jump operators, interactive proofs and proof complexity generators*, 2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS), IEEE, 2024, pp. 573–593.

[16] Jan Krajíček, *On the existence of strong proof complexity generators*, Bulletin of Symbolic Logic **30** (2024), no. 1, 20–40.

[17] ———, *A proof complexity conjecture and the incompleteness theorem*, Journal of Symbolic Logic **90** (2025), no. 3, 1206–1210.

[18] Jan Krajíček and Pavel Pudlák, *Propositional proof systems, the consistency of first order theories and the complexity of computations*, The Journal of Symbolic Logic **54** (1989), no. 3, 1063–1079.

[19] Yanyi Liu and Rafael Pass, *On one-way functions and kolmogorov complexity*, 61st Annual IEEE Symposium on Foundations of Computer Science (FOCS 2020), 2020, arXiv:2009.11514.

[20] Pavel Pudlák, *On the length of proofs of finitistic consistency statements in first order theories*, Studies in Logic and the Foundations of Mathematics, vol. 120, Elsevier, 1986, pp. 165–196.

[21] ———, *The lengths of proofs*, Handbook of Proof Theory (Samuel R. Buss, ed.), Elsevier, 1998.

[22] ———, *Logical foundations of mathematics and computational complexity: A gentle introduction*, Springer, 2013.

[23] ———, *Incompleteness in the finite domain*, Author's manuscript, 2017, Version of November 21, 2017. https://users.math.cas.cz/~pudlak/inco.pdf.

[24] ———, *Reflection principles, propositional proof systems, and theories*, arXiv:2007.14835, 2020, https://arxiv.org/abs/2007.14835.

[25] Tibor Rado, *On non-computable functions*, The Bell System Technical Journal **41** (1962), no. 3, 877–884.

[26] Alexander Shen, *Around Kolmogorov complexity: Basic notions and results*, 2015, https://arxiv.org/abs/1504.04955.

[27] Albert Visser, *The interpretation existence lemma*, Feferman on Foundations: Logic, Mathematics, Philosophy (Gerhard Jäger and Wilfried Sieg, eds.), Springer International Publishing, Cham, 2017, pp. 101–144.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il