APPROXIMATION VIA CORRELATION DECAY WHEN STRONG SPATIAL MIXING FAILS*

IVONA BEZÁKOVÁ†, ANDREAS GALANIS‡, LESLIE ANN GOLDBERG‡, HENG GUO§, AND DANIEL ŠTEFANKOVIČ¶

Abstract. Approximate counting via correlation decay is the core algorithmic technique used in the sharp delineation of the computational phase transition that arises in the approximation of the partition function of antiferromagnetic 2-spin models. Previous analyses of correlation-decay algorithms implicitly depended on the occurrence of *strong spatial mixing*. This, roughly, means that one uses worst-case analysis of the recursive procedure that creates the subinstances. In this paper, we develop a new analysis method that is more refined than the worst-case analysis. We take the shape of instances in the computation tree into consideration and we amortize against certain "bad" instances that are created as the recursion proceeds. This enables us to show correlation decay and to obtain a fully polynomial-time approximation scheme (FPTAS) even when strong spatial mixing fails. We apply our technique to the problem of approximately counting independent sets in hypergraphs with degree upper bound Δ and with a lower bound k on the arity of hyperedges. Liu and Lin gave an FPTAS for $k \geq 2$ and $\Delta \leq 5$ (lack of strong spatial mixing was the obstacle preventing this algorithm from being generalized to $\Delta = 6$). Our technique gives a tight result for $\Delta = 6$, showing that there is an FPTAS for $k \geq 3$ and $\Delta \leq 6$. The best previously known approximation scheme for $\Delta = 6$ is the Markov-chain simulation based fully polynomial-time randomized approximation scheme (FPRAS) of Bordewich, Dyer, and Karpinski, which only works for $k \geq 8$. Our technique also applies for larger values of k, giving an FPTAS for $k \geq \Delta$. This bound is not substantially stronger than existing randomized results in the literature. Nevertheless, it gives the first deterministic approximation scheme in this regime. Moreover, unlike existing results, it leads to an FPTAS for counting dominating sets in regular graphs with sufficiently large degree. We further demonstrate that in the hypergraph independent set model, approximating the partition function is NP-hard even within the uniqueness regime. Also, approximately counting dominating sets of bounded-degree graphs (without the regularity restriction) is NP-hard.

Key words. approximate counting, hypergraph independent set, correlation decay

AMS subject classifications. 68W20, 68W40, 68Q87

DOI. 10.1137/16M1083906

1. Introduction.

We develop a new method for analyzing correlation decays in spin systems. In particular, we take the shape of instances in the computation tree into consideration and we amortize against certain "bad" instances that are created as the recursion proceeds. This enables us to show correlation decay and to obtain

*Received by the editors July 8, 2016; accepted for publication (in revised form) January 30, 2019; published electronically March 28, 2019.

http://www.siam.org/journals/sicomp/48-2/M108390.html

Funding: The first author was supported by NSF grant CCF-1319987. The second and third authors have received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) ERC grant agreement 334828. The paper reflects only the authors' views and not the views of the ERC or the European Commission. The European Union is not liable for any use that may be made of the information contained therein. The fifth author was supported by NSF grant CCF-1563757.

†Department of Computer Science, Rochester Institute of Technology, Rochester, NY 14623 (ib@cs.rit.edu).

‡Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford, OX1 3QD, UK (andreas.galanis@cs.ox.ac.uk, leslie.goldberg@cs.ox.ac.uk).

§School of Informatics, University of Edinburgh, Informatics Forum, Edinburgh, EH8 9AB, UK (hguo@inf.ed.ac.uk).

¶Department of Computer Science, University of Rochester, Rochester, NY 14627 (stefanko@cs.rochester.edu).

a fully polynomial-time approximation scheme (FPTAS) even when strong spatial mixing fails. To the best of our knowledge, strong spatial mixing is a requirement for all previous correlation-decay based algorithms. To illustrate our technique, we focus on the computational complexity of approximately counting independent sets in *hypergraphs*, or equivalently on counting the satisfying assignments of monotone conjunctive normal form (CNF) formulas.

The problem of counting independent sets in *graphs* (denoted #IS) is extensively studied. A beautiful connection has been established, showing that approximately counting independent sets in graphs of maximum degree Δ undergoes a computational transition which coincides with the uniqueness phase transition from statistical physics on the infinite Δ-regular tree. The computational transition can be described as follows. Weitz [22] designed an FPTAS for counting independent sets on graphs with maximum degree at most $\Delta = 5$. On the other hand, Sly [19] proved that there is no fully polynomial-time randomized approximation scheme (FPRAS) for approximately counting independent sets on graphs with maximum degree at most $\Delta = 6$ (unless NP = RP). The same connection has been established in the more general context of approximating the partition function of the hard-core model [22, 16, 19, 4, 6, 20] and in the even broader context of approximating the partition functions of generic antiferromagnetic 2-spin models [18, 6, 20, 11] (which includes, for example, the antiferromagnetic Ising model). As a consequence, the boundary for the existence of efficient approximation algorithms for these models has been mapped out.[1]

Approximate counting via correlation decay is the core technique in the algorithmic developments which enabled the sharp delineation of the computational phase transition. Another standard approach for approximate counting, namely Markov chain Monte Carlo (MCMC) simulation, is also conjectured to work up to the uniqueness threshold, but the current analysis tools that we have do not seem to be powerful enough to show that. For example, sampling independent sets via MCMC simulation is known to have fast mixing only for graphs with degree at most 4 [14, 3], rather than obtaining the true threshold of 5.

In this work, we consider counting independent sets in hypergraphs with upper-bounded vertex degree and lower-bounded hyperedge size. A hypergraph $H = (V, \mathcal{F})$ consists of a vertex set V and a set $\mathcal{F}$ of hyperedges, each of which is a subset of V. A hypergraph is said to be *k-uniform* if every hyperedge contains exactly k vertices. Thus, a 2-uniform hypergraph is the same as a graph. We will consider the more general case where each hyperedge has arity at least k, rather than exactly k.

An independent set in a hypergraph H is a subset of vertices that does not contain a hyperedge as a subset. We will be interested in computing Z_H, which is the total number of independent sets in H (also referred to as the partition function of H). Formally, the problem of counting independent sets has two parameters—a degree upper bound Δ and a lower bound k on the arity of hyperedges. The problem is defined as follows.[2]

Name #HyperIndSet(k, Δ).

[1] Note that there are nonmonotonic examples of antiferromagnetic 2-spin systems where the boundary is more complicated because the uniqueness threshold fails to be monotonic in Δ [10]. However, this can be cleared up by stating the uniqueness condition as uniqueness for all $d \leq \Delta$. See [10, 11] for details.

[2] Equivalently, one may think of this problem as the problem of counting satisfying assignments of a monotone CNF formula, where vertices are variables and hyperedges are clauses. Being out of the independent set (as a vertex) corresponds to being true (as a variable).

Instance A hypergraph H with maximum degree at most Δ where each hyperedge has cardinality (arity) at least k.
Output The number Z_H of independent sets in H.

Previously, #HyperIndSet(k, Δ) has been studied using the MCMC technique by Bordewich, Dyer, and Karpinski [1, 2] (see also [3]). They give an FPRAS for all $k \geq \Delta + 2 \geq 5$ and for $k \geq 2$ and $\Delta = 3$. Despite equipping path coupling with optimized metrics obtained using linear programming, these bounds are not tight for small k. Liu and Lu [12] showed that there exists an FPTAS for all $k \geq 2$ and $\Delta \leq 5$ using the correlation decay technique.

Thus, the situation seems to be similar to the graph case—given the analysis tools that we have, correlation-decay brings us closer to the truth than the best-tuned analysis of MCMC simulation algorithms. On the other hand, the technique of Liu and Lu [12] does not extend beyond $\Delta = 5$. To explain the reason why it does not, we need to briefly describe the correlation-decay based algorithm framework introduced by Weitz [22]. The main idea is to build a recursive procedure for computing the marginal probability that any given vertex is in the independent set. The recursion works by examining subinstances with "boundary conditions" which require certain vertices to be in, or out, of the independent set. The recursion structure is called a "computation tree." Nodes of the tree correspond to intermediate instances, and boundary conditions are different in different branches. The computation tree allows one to compute the marginal probability exactly, but the time needed to do so may be exponentially large since, in general, the tree is exponentially large. Typically, an approximate marginal probability is obtained by truncating the computation tree to logarithmic depth so that the (approximation) algorithm runs in polynomial time. If the correlation between boundary conditions at the leaves of the (truncated) computation tree and the marginal probability at the root decays exponentially with respect to the depth, then the error incurred from the truncation is small and the algorithm succeeds in obtaining a close approximation.

All previous instantiations under this framework require a property called *strong spatial mixing* (SSM),[3] which roughly states that, conditioned on *any* boundary condition on intermediate nodes, the correlation decays. In other words, SSM guards against the worst-case boundary conditions that might be created by the recursive procedure.

Let the $(\Delta - 1)$-ary k-uniform hypertree $\mathbb{T}_{k,\Delta}$ be the recursively defined hypergraph in which each vertex has $\Delta - 1$ "descending" hyperedges, each containing $k - 1$ new vertices.

Observation 1.1. Let $k \geq 2$. For $\Delta \geq 6$, strong spatial mixing *does not hold* on $\mathbb{T}_{k,\Delta}$.

Observation 1.1 follows from the fact that the infinite $(\Delta-1)$-ary tree $\mathbb{T}_{2,\Delta}$ can be embedded in the hypertree $\mathbb{T}_{k,\Delta}$, and from well-known facts about the phase transition on $\mathbb{T}_{2,\Delta}$.

Observation 1.1 prevents the generalization of Liu and Lu's algorithm [12] so that it applies for $\Delta \geq 6$, even with an edge-size lower bound k. The problem is that the construction of the computation tree involves constructing intermediate instances in which the arity of a hyperedge can be as small as 2. So, even if we start with a k-uniform hypergraph, the computation tree will contain instances with small hyperedges. Without SSM, these small hyperedges cause problems in the analysis.

[3]See section 2.1 for a definition.

Lu, Yang, and Zhang [13] discuss this problem and say "How to avoid this effect is a major open question whose solution may have applications in many other problems." This question motivates our work.

To overcome this difficulty, we introduce a new amortization technique in the analysis. Since lack of correlation decay is caused primarily by the presence of small-arity hyperedges within the intermediate instances, we keep track of such hyperedges. Thus, we track not only the correlation but also combinatorial properties of the intermediate instances in the computation tree. Using this idea, we obtain the following result.

THEOREM 1.2. *There is an FPTAS for* $\#\mathsf{HyperIndSet}(3, 6)$.

Note that $\#\mathsf{HyperIndSet}(2, 6)$ is NP-hard to approximate due to [19], so our result is tight for $\Delta = 6$. This also shows that $\Delta = 6$ is the first case where the complexity of approximately counting independent sets differs on hypergraphs and graphs, as for $\Delta \leq 5$ both admit an FPTAS [12]. Moreover, Theorem 1.2 is stronger than the best MCMC algorithm [2] when $\Delta = 6$ as [2] only works for $k \geq 8$.

We also apply our technique to large k.

THEOREM 1.3. *Let* k *and* Δ *be two integers such that* $k \geq \Delta$ *and* $\Delta \geq 200$. *Then there is an FPTAS for the problem* $\#\mathsf{HyperIndSet}(k, \Delta)$.

In the large k case, our result is not substantially stronger than that obtained by analysis of the MCMC algorithm [2] ($k \geq \Delta$ rather than $k \geq \Delta + 2$), but it is incomparable since our algorithm is deterministic rather than randomized. Perhaps more importantly, the bound $k \geq \Delta$ allows us to connect the problem of counting independent sets in hypergraphs with the problem of counting dominating sets in Δ-regular graphs and show that the latter admits an FPTAS when Δ is sufficiently large. Recall that a dominating set in a graph G is a subset S of the vertices such that every vertex not in S is adjacent to at least one vertex in S. We then consider the following problem.

Name $\#\mathsf{RegDomSet}(\Delta)$.
Instance A Δ-regular graph G.
Output The number of dominating sets in G.

Our theorems have the following corollary.

COROLLARY 1.4. *For all positive integers* Δ *satisfying either* $\Delta \leq 5$ *or* $\Delta \geq 199$, *there is an FPTAS for the problem* $\#\mathsf{RegDomSet}(\Delta)$.

We remark that Corollary 1.4 cannot be obtained using the result of [2]; i.e., the seemingly small difference between $k \geq \Delta$ and $k \geq \Delta + 2$ does matter in deriving Corollary 1.4 from Theorems 1.2 and 1.3. We should also emphasize that it is necessary to consider Δ-regular graphs as inputs to the dominating set problem, since otherwise for graphs of maximum degree Δ (not necessarily regular), we show that the problem is NP-hard to approximate for $\Delta \geq 18$ (Theorem 6.5). It is relevant to remark here that we believe that Corollary 1.4 should hold for all Δ; to do this, it would be sufficient to remove the restriction $\Delta \geq 200$ from Theorem 1.3. Note that, while we do not know how to remove the restriction, it would at least be possible to improve "200" to some smaller number. However, we have chosen to stick with "200" in order to keep the proof accessible. Next we explain the difficulties in obtaining Theorems 1.2 and 1.3.

The main technical difficulty in correlation-decay analysis is bounding a function that we call the "decay rate." This boils down to solving an optimization problem

with $(k-1)(\Delta-1)$ variables. In previous work (e.g., [17]), this optimization has been solved using a so-called *symmetrization* argument, which reduces the problem to a univariate optimization via convexity. However, the many variables represent different branches in the computation tree. Since our analysis takes the shape of intermediate instances in the tree into consideration, the symmetrization argument does not work for us, and different branches take different values at the maximum. This problem is compounded by the fact that the shape of the subtree consisting of "bad" intermediate instances is heavily lopsided, and the assignment of variables achieving the maximum is far from uniform. Given these problems, there does not seem to be a clean solution to the optimization in our analysis. Instead of optimizing, we give an upper bound on the maximum decay rate. In Theorem 1.2, as k and Δ are small, the number of variables is manageable, and our bounds are much sharper than those in Theorem 1.3. On the other hand, because of this, the proof of Theorem 1.3 is much more accessible, and we will use Theorem 1.3 as a running example to demonstrate our technique.

We also provide some insight on the hardness side. Recall that for graphs it is NP-hard to approximate #IS beyond the uniqueness threshold ($\Delta = 6$) [19].[4] We prove that it is NP-hard to approximate #HyperIndSet$(6, 22)$ (Corollary 6.3). In contrast, we show that uniqueness holds on the 6-uniform Δ-regular hypertree iff $\Delta \leq 28$ (Corollary 7.6). Thus, efficient approximation schemes cease to exist well below the uniqueness threshold on the hypertree. In fact, we show that this discrepancy grows exponentially in k: for large k, it is NP-hard to approximate #HyperIndSet(k, Δ) when $\Delta \geq 5 \cdot 2^{k/2}$ (Theorem 6.2 and Corollary 6.4), despite the fact that uniqueness holds on the hypertree for all $\Delta \leq 2^k/(2k)$ (Lemma 7.7). Theorem 6.2 follows from a rather standard reduction to the hard-core model on graphs. Nevertheless, it demonstrates that the computational-threshold phenomena in the hypergraph case ($k > 2$) are substantially different from those in the graph case ($k = 2$).

As mentioned earlier, there are models where efficient (randomized) approximation schemes exist (based on MCMC simulation) even though SSM does not hold. In fact, this can happen even when uniqueness does not hold. A striking example is the ferromagnetic Ising model (with external field). As [18] shows, there are parameter regimes where uniqueness holds but SSM fails. It is easy to modify the parameters so that even uniqueness fails. Nevertheless, Jerrum and Sinclair [9] gave an MCMC-based FPRAS that applies for all parameters and for general graphs (with no degree bounds). It is still an open question to give a correlation-decay based FPTAS for the ferromagnetic Ising model.

1.1. Note added in revision—recent work. After our work, Hermon, Sly, and Zhang [8] have shown that the Glauber dynamics for sampling independent sets on n-vertex k-uniform hypergraphs has mixing time $O(n \log n)$ if $\Delta \leq c2^{k/2}$ for some constant c. Combined with our hardness result, Corollary 6.4, this establishes a sharp computational complexity phase transition, up to constants. Moreover, Moitra [15] has given a deterministic algorithm for approximately counting and sampling the satisfying assignments of k-CNF formulas, provided that the variable degree $\Delta \lesssim 2^{k/60}$. Moitra's work relies on connections to the Lovász local lemma. His work improves our $\Delta \leq k$ bound in Theorem 1.3 and removes the monotonicity assumption. Despite this recent progress, Theorem 1.2 remains the best algorithmic result for $k = 3$.

[4] For graphs the uniqueness and SSM thresholds coincide, but for hypergraphs they differ.

1.2. Outline of paper. In section 2, we first give some preliminaries. We give a formal definition of SSM (section 2.1) and a reformulation of $\#\mathsf{HyperIndSet}(k,\Delta)$ as the problem of counting satisfying assignments in monotone CNF formulas (section 2.2). This will allow us to use the computation tree used by Liu and Lu [12]. A formal description of the computation tree of [12] is given in section 2.3.

In section 3, we give an overview of our proof approach, i.e., the main idea behind our new amortization technique. Section 3.2 concludes the proof of Theorems 1.2 and 1.3, using two (not yet proved) technical lemmas (Lemma 3.9 and 3.10) which solve a complicated multivariate optimization problem and represent the bulk of the technical work of the paper. Section 3.3 gives the proof of Corollary 1.4.

In section 4, we give the proof of Lemma 3.9, which applies to the large-Δ setting of Theorem 1.3 and is by far the technically simpler of the two lemmas. Section 5 contains the proof of the technically more challenging Lemma 3.10 which applies to the $k=3,\Delta=6$ setting of Theorem 1.2.

Section 6 gives the formal statements and proofs of the hardness results stated in the introduction; section 6.1 has the hardness results for independent sets in hypergraphs, and section 6.2 has the hardness results for dominating sets in graphs. Also, section 7 studies the uniqueness threshold on the k-uniform Δ-regular hypertree (and gives the proofs of the uniqueness statements made in the introduction). Finally, section 8 gives the proof for several technical inequalities used in section 5.

Throughout the paper we use computer algebra to prove multivariate polynomial inequalities over the field of real numbers (the coefficients of the polynomials are rational). More specifically, we use the Resolve command in Mathematica. The underlying quantifier elimination algorithm (described in [21]) provides a rigorous decision procedure that determines feasibility of a collection of polynomial inequalities.

2. Preliminaries.

2.1. Strong spatial mixing. For the purposes of this section, it will be convenient to view the independent set model as a 2-spin model. Namely, if $H=(V,\mathcal{F})$ is a hypergraph, each independent set I can be viewed as a $\{0,1\}$-assignment σ to the vertices in V, where a vertex v is assigned the spin 1 under σ if $v\in I$ and 0 otherwise.

We denote by Ω_H the set of all independent sets in H. The Gibbs distribution $\mu_H(\cdot)$ is the uniform distribution over Ω_H. The Gibbs distribution of H can clearly be viewed as the uniform distribution over those assignments $\sigma: V\to\{0,1\}$ which encode a valid independent set of H. For an assignment $\sigma: V\to\{0,1\}$ and a subset $\Lambda\subset V$, we denote by σ_Λ the restriction of σ to the subset Λ.

For a hypergraph $H=(V,\mathcal{F})$ and a subset $\Lambda\subset V$, we denote by H_Λ the subgraph of H induced by Λ, i.e., $H_\Lambda:=(\Lambda,\bigcup_{e\in\mathcal{F}}(e\cap\Lambda))$. Also, for a vertex $v\in V$ and $\Lambda\subset V$, we denote by $\mathrm{dist}(v,\Lambda)$ the length of the shortest path[5] between v and a vertex of Λ.

Definition 2.1. *Let $\delta:\mathbb{Z}_+\to[0,1]$. The independent set model exhibits* strong spatial mixing *on a hypergraph $H=(V,\mathcal{F})$ with decay rate $\delta(\cdot)$ iff for every $v\in V$, for every $\Lambda\subset V$, and for any two configurations $\eta,\eta':\Lambda\to\{0,1\}$ encoding independent sets of H_Λ, it holds that*

$$\Big|\mu_H(\sigma(v)=1\mid\sigma_\Lambda=\eta)-\mu_H(\sigma(v)=1\mid\sigma_\Lambda=\eta')\Big|\le\delta\big(\mathrm{dist}(v,\Lambda')\big),$$

where Λ' denotes the set of vertices in Λ such that η and η' differ.

[5] A path in a hypergraph with hyperedge set $\mathcal{F}$ is a sequence of edges $e_0,\dots,e_\ell\in\mathcal{F}$ such that $e_i\cap e_{i+1}\neq\emptyset$ for all $i=0,\dots,\ell-1$.

2.2. Reformulation in terms of monotone CNF formulas. The problem of counting the independent sets of a hypergraph has an equivalent formulation in terms of monotone CNF formulas. In order to describe the equivalent formulation, we first describe the problem of counting satisfying assignments of monotone CNF formulas.

A monotone CNF formula C consists of a set of variables V and a set of clauses $\{c_1, c_2, \ldots\}$. Each clause c_i is associated with some subset S_i of V and is the disjunction of all variables in S_i. The *arity* of a clause c_i, denoted $|c_i|$, is defined to be $|S_i|$. For a variable $x \in V$, its *degree* $d_x(C)$ is the number of clauses where x appears. The *maximum degree* of C is given by $\max_{x \in V} d_x(C)$.

DEFINITION 2.2. *Let $\mathcal{C}_{k,\Delta}$ be the set of all monotone CNF formulas which have maximum degree at most Δ and whose clauses have arity at least k.*

Note that a formula in $\mathcal{C}_{k,\Delta}$ may have some clauses with arbitrarily large arities. A *satisfying assignment* of the formula is an assignment of truth values to the variables that makes the formula evaluate to "true."

Suppose that $H = (V, \mathcal{F})$ is a hypergraph with maximum degree at most Δ where each hyperedge has arity at least k. Let C be the corresponding formula in $\mathcal{C}_{k,\Delta}$ with variable set V. The correspondence is that each hyperedge S_i of H is associated with exactly one clause c_i of C. Independent sets of H are in one-to-one correspondence with satisfying assignments of C—a variable is assigned value "true" in an assignment iff it is out of the corresponding independent set.

Going the other direction, any monotone CNF formula can be viewed as a hypergraph. In the technical sections of this paper, we use the monotone CNF formulation.

In this article, when we consider a monotone CNF formula C we will typically use n to denote $|V|$. Variables in V will be denoted by $x_1, x_2, \ldots$. When x and C are clear from context, we will sometimes use d to denote $d_x(C)$. When C is clear from context, we will sometimes use Δ to denote $\max_{x \in V} d_x(C)$.

2.3. The computation tree. In this section, we set up relevant notation and give an exposition of the computation tree of Liu and Lu [12] which will also be used in our proof (though our analysis will be different). The computation tree of Liu and Lu is given in terms of the monotone CNF version of the problem. Below we give the relevant definitions and notation; our notation aligns as much as possible with that of [12].

Our goal is to approximately count the number of satisfying assignments of a formula $C \in \mathcal{C}_{k,\Delta}$, which we denote by $Z(C)$. Since C is monotone, an assignment $\sigma : V \to \{0,1\}$ is satisfying if, for every clause in C, there is at least one variable $x \in c$ with $\sigma(x) = 1$. Note that $Z(C) > 0$ since the all-1 assignment satisfies every monotone CNF formula. For convenience, we will use the simplified notation "$x = 1$" to denote (the set of) satisfying assignments of C in which x is set to 1, and we similarly use "$x = 0$." We associate the formula C with a probability distribution in which each satisfying assignment has probability mass $1/Z(C)$. We will denote probabilities with respect to this distribution by $\Pr_C(\cdot)$.

Let x be a variable in V. Define $R(C, x) := \frac{\Pr_C(x=0)}{\Pr_C(x=1)}$; this is well defined since $\Pr_C(x = 1) > 0$ by the monotonicity of C. In fact, the monotonicity of C also implies that $0 \le R(C, x) \le 1$, where the upper bound follows from the fact that, for every satisfying assignment with $x = 0$, flipping the assignment of x to 1 does not affect satisfiability. Our interest in the quantity $R(C, x)$ stems from the following simple lemma from [12].

LEMMA 2.3 (see [12]). *Let k and Δ be positive integers. Suppose that there is a polynomial-time algorithm (in n and $1/\varepsilon$) that takes an n-variable formula $C \in \mathcal{C}_{k,\Delta}$, a variable x of C, and an $\varepsilon > 0$ and computes a quantity $\widehat{R}(C,x)$ satisfying $|\widehat{R}(C,x) - R(C,x)| \leq \varepsilon$. Then, there is an FPTAS which approximates $Z(C)$ for every $C \in \mathcal{C}_{k,\Delta}$.*

Proof. The proof is actually identical to the argument in [12, Appendix A]. We include the proof for completeness and also because an examination of the proof is necessary to check that the FPTAS for approximating $Z(C)$ invokes the algorithm that computes $\widehat{R}(C,x)$ only on formulas C whose clauses have arity at least k (and maximum degree Δ).

Let $\varepsilon > 0$ and C be a monotone CNF formula C with maximum degree Δ whose clauses have arity at least k. Let $x_1, \ldots, x_n$ be the variables in C. Let C_i be the formula obtained from C by setting $x_1 = \cdots = x_i = 1$ and removing all the clauses that are satisfied (i.e., all clauses that contain a variable from $x_1, \ldots, x_i$). We have

$$\begin{aligned} \frac{1}{Z(C)} &= \Pr_C(x_1 = \cdots = x_n = 1) = \prod_{i=0}^{n-1} \Pr_C(x_{i+1} = 1 \mid x_1 = \cdots = x_i = 1) \\ &= \prod_{i=0}^{n-1} \Pr_{C_i}(x_{i+1} = 1) = \prod_{i=0}^{n-1} \frac{1}{1 + R(C_i, x_{i+1})}. \end{aligned} \tag{2.1}$$

Note that every C_i is a monotone CNF formula with maximum degree Δ whose clauses have arity at least k. By the assumption in the lemma, we can compute (in $\mathrm{poly}(n, 1/\varepsilon)$ time) quantities $\widehat{R}(C_i, x_{i+1})$ such that

$$\left|\widehat{R}(C_i, x_{i+1}) - R(C_i, x_{i+1})\right| \leq \varepsilon/(100n) \text{ for all } i = 0, \ldots, n-1. \tag{2.2}$$

Let

$$\widehat{Z}(C) = \prod_{i=0}^{n-1} \left(1 + \widehat{R}(C_i, x_{i+1})\right). \tag{2.3}$$

It is not hard to conclude from (2.1), (2.2), and (2.3) that $(1-\varepsilon)Z(C) \leq \widehat{Z}(C) \leq (1+\varepsilon)Z(C)$. This completes the proof. ∎

Liu and Lu [12] established that a computation tree approach gives a recursive procedure for *exactly* calculating $R(C,x)$ for any monotone CNF formula C and any variable $x \in C$. We next give the details of this recursive procedure (see [12, Lemma 5]). First, we introduce the following definitions.

DEFINITION 2.4. *Let C be a monotone CNF formula and let x be a variable in C. We call the variable x* forced *(in C) if x appears in a clause of arity 1 in C (note that in every satisfying assignment of C it must be the case that $x = 1$ and hence $R(C,x) = 0$). We call the variable x* free *if x does not appear in any clause of C (note that $R(C,x) = 1$ in this case).*

DEFINITION 2.5. *Let C be a monotone CNF formula and let c be a clause in C. We call the clause c* redundant *(in C) if there is a clause c' in C such that c is a (strict) superset of c' (note that removing c from C does not affect the set of satisfying assignments of C).*

We next give the details of the computation tree. The nodes in the computation tree will be pairs (C, x) such that

(2.4) C is a monotone CNF formula and x is a variable which is not forced in C.

Let (C, x) satisfy (2.4). We first perform a preprocessing step on C which involves (i) initially removing all of the redundant clauses, and (ii) then removing all clauses of arity 1. Note that part (ii) of the preprocessing step removes all forced variables that were present in C; at the time of the removal, forced variables appear only in clauses of arity 1 since part (i) of the preprocessing step has already removed all redundant clauses in C (and hence all clauses of arity greater than 1 that contain forced variables). Denote the formula after the completion of the preprocessing step by $\widetilde{C}$. Note that every clause in $\widetilde{C}$ is also a clause in the initial formula C. It follows that x is not forced in $\widetilde{C}$. Further, since removing redundant clauses does not change the set of satisfying assignments of C and x is not forced in C, we have that $R(\widetilde{C}, x) = R(C, x)$.

If x is free in $\widetilde{C}$ (the formula after the preprocessing step), then the start node (C, x) is (declared) a leaf of the computation tree (note that in this case $R(C, x) = 1$). In what follows, we assume that x is not free in $\widetilde{C}$. Denote by $\{c_i\}_{i\in[d]}$ the clauses where x occurs in $\widetilde{C}$ and let $w_i = |c_i| - 1$ (note that $d \geq 1$). We will use $\mathbf{w}$ to denote the vector $(w_1, \ldots, w_d)$. The variables in clause c_i other than x will be denoted by $x_{i,1}, \ldots, x_{i,w_i}$. For the pair (C, x), we next construct pairs $(C_{i,j}, x_{i,j})$ for $i \in [d]$ and $j \in [w_i]$, where $C_{i,j}$ is an appropriate subformula obtained from $\widetilde{C}$, roughly, by hard-coding (some of) the occurrences of the variables in $\widetilde{C}$ to either 1 or 0 (this will be explained below and will henceforth be referred to as pinning).[6]

Precisely, for $i \in [d]$, let C_i be the formula obtained from $\widetilde{C}$ by removing clauses $c_1, \ldots, c_{i-1}$ (note that this has the same effect as pinning the occurrences of x in these clauses to 1) and pinning the occurrences of x in $c_{i+1}, \ldots, c_d$ to 0 (this corresponds to removing x from these clauses, and thus reducing their arities). For $j \in [w_i]$, the formula $C_{i,j}$ is obtained from C_i by further removing clause c_i and pinning all the occurrences of $x_{i,1}, \ldots, x_{i,j-1}$ to 0.

Before proceeding, let us argue that the pairs $(C_{i,j}, x_{i,j})$ satisfy (2.4) for all $i \in [d]$ and $j \in [w_i]$. For such i, j, we first prove that $C_{i,j}$ is a (satisfiable) monotone CNF formula. That is, we prove that the various pinnings in the construction of $C_{i,j}$ from $\widetilde{C}$ do not pin all variables of some clause of $\widetilde{C}$ to 0. For the sake of contradiction, assume otherwise. Observe that $C_{i,j}$ is obtained from $\widetilde{C}$ by either removing some clauses or pinning some occurrences of the variables to 0. Clearly, removal of clauses does not affect satisfiability, so we may focus on the effect of pinning. For $i \in [d]$ and $j \in [w_i]$, the only variables whose (some of the) occurrences in $\widetilde{C}$ get pinned to 0 are $x, x_{i,1}, \ldots, x_{i,j-1}$. Since we assumed (for contradiction) that $C_{i,j}$ is unsatisfiable, it must be the case that there exists a clause c' in $\widetilde{C}$ all of whose variables are (a subset of) $x, x_{i,1}, \ldots, x_{i,j-1}$. It follows that c_i is redundant in $\widetilde{C}$ since it is a strict superset of clause c'. This gives a contradiction since the preprocessing operation ensures that $\widetilde{C}$ has no redundant clauses. Thus, $C_{i,j}$ is satisfiable, as wanted. Next, we show that $x_{i,j}$ is not forced in $C_{i,j}$. First, observe that $x_{i,j}$ is not forced in $\widetilde{C}$ since the second part of the preprocessing step ensures that $\widetilde{C}$ does not contain forced variables. Thus,

[6]Note that our notation for i, j is different from that in [12]; there, the roles of i, j are interchanged.

the only way that $x_{i,j}$ can be forced in $C_{i,j}$ is if there existed a clause c' in $\widetilde{C}$ whose variables were $x_{i,j}$ together with a subset of $x, x_{i,1}, \ldots, x_{i,j-1}$. Since $\widetilde{C}$ includes c_i and $\widetilde{C}$ does not have redundant clauses, it must be the case that $c' = c_i$. It remains to observe that $C_{i,j}$ does not include (any subclause of) c_i, from which it follows that $x_{i,j}$ is not forced in $C_{i,j}$.

We are now ready to state the relation between $R(C, x)$ and the quantities $R(C_{i,j}, x_{i,j})$ with $i \in [d]$ and $j \in [w_i]$.

LEMMA 2.6 (see [12, Lemma 5]). *It holds that*

$$R(C,x) = \prod_{i=1}^{d} \left(1 - \prod_{j=1}^{w_i} \frac{R(C_{i,j}, x_{i,j})}{1 + R(C_{i,j}, x_{i,j})}\right). \tag{2.5}$$

Proof. The proof is identical to the proof of [12, Lemma 5] (which in turn builds on the technique of [22]), we give the proof for completeness.

Recall that $\widetilde{C}$ is the formula after the preprocessing step and that $R(\widetilde{C}, x) = R(C, x)$. We may assume that x is not free in $\widetilde{C}$ (otherwise, it holds that $R(\widetilde{C}, x) = 1$, which coincides with the evaluation of the right-hand side of (2.5) under the standard convention that the empty product evaluates to 1).

Equation (2.5) follows immediately from the following two equalities:

$$R(\widetilde{C}, x) = \prod_{i \in [d]} R(C_i, x), \quad R(C_i, x) = 1 - \prod_{j=1}^{w_i} \frac{R(C_{i,j}, x_{i,j})}{1 + R(C_{i,j}, x_{i,j})}. \tag{2.6}$$

The first equality in (2.6) is a consequence of a telescoping expansion of $\frac{\Pr_{\widetilde{C}}(x=0)}{\Pr_{\widetilde{C}}(x=1)}$. To see this, let $\widetilde{C}'$ be the formula obtained from $\widetilde{C}$ by replacing, for all $i \in [d]$, the occurrence of the variable x in clause c_i by a new variable x'_i. We have that

$$\begin{aligned} R(\widetilde{C}, x) &= \frac{\Pr_{\widetilde{C}}(x=0)}{\Pr_{\widetilde{C}}(x=1)} = \frac{\Pr_{\widetilde{C}'}(x'_1 = 0, \ldots, x'_d = 0)}{\Pr_{\widetilde{C}'}(x'_1 = 1, \ldots, x'_d = 1)} \\ &= \prod_{i \in [d]} \frac{\Pr_{\widetilde{C}'}(x'_1 = 1, \ldots, x'_{i-1} = 1, x'_i = 0, \ldots x'_d = 0)}{\Pr_{\widetilde{C}'}(x'_1 = 1, \ldots, x'_i = 1, x'_{i+1} = 0, \ldots x'_d = 0)} = \prod_{i \in [d]} \frac{\Pr_{C_i}(x=0)}{\Pr_{C_i}(x=1)}, \end{aligned}$$

which yields the first equality in (2.6) after substituting $R(C_i, x) = \frac{\Pr_{C_i}(x=0)}{\Pr_{C_i}(x=1)}$.

For the second equality in (2.6), observe that x appears only in clause c_i of the formula C_i, and thus (denoting by $C_i \backslash c_i$ the formula which is obtained from C_i by deleting clause c_i)

$$R(C_i, x) = \frac{\Pr_{C_i}(x=0)}{\Pr_{C_i}(x=1)} = 1 - \Pr_{C_i \backslash c_i}(x_{i,1} = 0, \ldots, x_{i,w_i} = 0) = 1 - \prod_{j=1}^{w_i} \Pr_{C_{i,j}}(x_{i,j} = 0),$$

which proves the desired equality after substituting $\Pr_{C_{i,j}}(x_{i,j} = 0) = \frac{R(C_{i,j}, x_{i,j})}{1 + R(C_{i,j}, x_{i,j})}$. ∎

By applying (2.5) recursively, it is not hard to see that one can compute the quantity $R(C, x)$ *exactly*. Of course, exact computation using this scheme will typically require exponential time, so as in [12] we will stop the recursion at some (small) depth L to keep the computations feasible within polynomial time. This will yield a quantity $R(C, x, L)$, and the hope is that, by choosing L appropriately, the error $|R(C, x, L) - R(C, x)|$ will be sufficiently small.

In light of (2.5), a natural way to define $R(C,x,L)$ for integer $L \geq 0$ is as follows.[7]

$$R(C,x,L) = \begin{cases} 1 & \text{if } x \text{ is free in } \widetilde{C} \text{ or } L = 0, \\ \prod_{i=1}^{d} \left(1 - \prod_{j=1}^{w_i} \frac{R(C_{i,j},x_{i,j},L-1)}{1+R(C_{i,j},x_{i,j},L-1)}\right) & \text{otherwise.} \end{cases}$$

It is immediate that when the formula C has maximum degree bounded by a constant and, further, every clause has arity also bounded above by a constant, one can compute $R(C,x,L)$ in time polynomial in n whenever $L = O(\log n)$. To account for formulas where the arities of the clauses can be arbitrarily large (but still where the degrees of variables are bounded by a constant), one needs to be more careful with clauses of large arity (i.e., when their arity as a function of n is $\omega(1)$, say $\log n$). As in [12], we will account for this more general setting by pruning the recursion earlier whenever we encounter clauses with large arity.

DEFINITION 2.7. *For any integer w, let $l_w := \lceil \log_6(w+1) \rceil$.*

Note that $l_1 = \cdots = l_5 = 1$. For integer L, we set

(2.7)
$$R(C,x,L) = \begin{cases} 1 & \text{if } x \text{ is free in } \widetilde{C} \text{ or } L \leq 0, \\ \prod_{i=1}^{d} \left(1 - \prod_{j=1}^{w_i} \frac{R(C_{i,j},x_{i,j},L-l_{w_i})}{1+R(C_{i,j},x_{i,j},L-l_{w_i})}\right) & \text{otherwise.} \end{cases}$$

The particular choice of the logarithm base in Definition 2.7 is not very important as long as it is a big enough constant. The quantity $R(C,x,L)$ is typically called a "message" (because it gets passed up the computation tree).

Remark 2.8. For formulas C with a variable x which is not forced in C, we have the lower bound $R(C,x) \geq (1/2)^{d_x(C)}$. The bound is simple to see using (2.5) and the fact that $R(C_{i,j},x_{i,j}) \leq 1$ for all $i \in [d_x(C)]$ and $j \in [w_i]$. Similarly, for all integers L and all nodes (C,x) in the computation tree we have the bound $R(C,x,L) \geq (1/2)^{d_x(C)}$.

3. Proof outline. We want to guarantee that the error $|R(C,x,L) - R(C,x)|$ is exponentially small in L. Notice that if we run the recursion long enough, it computes the true value; namely, $R(C,x,\infty) = R(C,x)$. More precisely, we will prove the following two lemmas, which correspond to the settings of Theorem 1.2 and Theorem 1.3, respectively. Recall that $\mathcal{C}_{k,\Delta}$ is the set of all monotone CNF formulas which have maximum degree at most Δ and whose clauses have arity at least k. Our proof will use the following constant.

DEFINITION 3.1. *Let $\alpha = 1 - 10^{-4}$.*

LEMMA 3.2. *There exists a constant $\tau > 0$ such that for every $C \in \mathcal{C}_{3,6}$, every variable $x \in C$, and every integer L,*

$$|R(C,x,L) - R(C,x,\infty)| \leq \tau\alpha^L.$$

LEMMA 3.3. *Let k and Δ be two integers such that $k \geq \Delta$ and $\Delta \geq 200$. There exists a constant $\tau > 0$ such that for every $C \in \mathcal{C}_{k,\Delta}$, every variable x in C, and every integer L,*

$$|R(C,x,L) - R(C,x,\infty)| \leq \tau\alpha^L.$$

[7]Note that the value 1 of $R(C,x,L)$ when $L \leq 0$ is somewhat arbitrary since $L \leq 0$ corresponds to stopping the recursion. Our choice of the value 1 will be convenient for technical reasons that will become apparent in the proof of the upcoming Lemma 3.8.

Our proof uses correlation-decay techniques together with a new method which accounts for the shape of the computation tree. Lemma 3.3 is technically simpler and is proved in section 4 to better illustrate the idea. Lemma 3.2 is proved in section 5. In the rest of this section, we give an overview of our overall proof strategy.

To analyze the error of the recursion, the standard approach so far in the literature has been to show that, for a node (C, x) in the computation tree, the quantity $|R(C, x, L) - R(C, x, \infty)|$ is bounded by $\alpha \max_{i,j} |R(C_{i,j}, x_{i,j}, L-1) - R(C_{i,j}, x_{i,j}, \infty)|$ for some constant $0 < \alpha < 1$. This allows one to inductively deduce that $|R(C, x, L) - R(C, x, \infty)|$ decays exponentially in L. This approach has been extremely successful when SSM holds [18, 11, 12, 17, 23, 13].

In fact, this stepwise decay seldom holds if we track $R(C, x, L)$ directly. Instead, the analysis is usually done by tracking $\Phi(R(C, x, L))$ for an appropriate potential function Φ. In particular, let $\Phi : (0, 1] \to \mathbb{R}$ be a potential function that satisfies

(3.1)
Φ is continuously differentiable on $(0, 1]$, and $\varphi := \Phi'$ satisfies $\varphi(z) > 0$ for $z \in (0, 1]$.

The usual approach is to show that $|\Phi(R(C, x, L)) - \Phi(R(C, x, \infty))|$ decays exponentially in L, which is sufficient to imply lemmas like Lemmas 3.2 and 3.3.

In our setting, this inductive approach is problematic since, inside the computation tree, we are faced with the possibility that the formula at the root of a subtree could have many arity-2 clauses. For $\Delta \geq 6$, these subtrees prohibit the application of the above proof scheme since they are in nonuniqueness and hence the desired step-by-step decay is no longer present, regardless of the choice of the potential function.

While arity-2 clauses are problematic, clauses with larger arities do at least lead to good decay of correlation in a single step. In general, as the arity gets larger, the decay gets better. Thus, our approach is to do an amortized analysis. In a single step, we track both the one-step decay of correlation and the number of variables in the current formula that are pinned to 0. These 0-pinnings will decrease the effective arity of clauses and will later lead to worse decay.

More formally, we will track a specific quantity $m(C, x, L)$ which is assigned to each node in the computation tree. Let C^* be the original monotone CNF formula and let (C, x) be a node in the computation tree. As explained in section 2.3, each clause c' of C is obtained from a clause c of C^* by pinning a certain number of variables to 0 (possibly none), which effectively is the same as removing those variables from the clause. We call these 0-pinnings *deficits* and let $\max\{0, k - |c'|\}$ be the number of deficits of c'. Note that a clause of arity larger than k is considered to have no deficit, although some variables of it may have been pinned to 0.

DEFINITION 3.4. *Let* $D(C) = \sum_{c' \in C} \max\{0, k - |c'|\}$ *denote the sum of the deficits of the clauses in* C.

Observe that if a clause c of C^* does not show up in C, it does not contribute any deficits. For any node (C, x) in the computation tree, let

$$m(C, x, L) := \delta^{D(C)} \Phi(R(C, x, L)),$$

where $\delta \in (0, 1)$ is a constant that we will choose later, and the potential function Φ will be specified shortly in Definition 3.5. Crucially, the root formula C^* satisfies

$$|m(C^*, x, L) - m(C^*, x, \infty)| = |\Phi(R(C^*, x, L)) - \Phi(R(C^*, x, \infty))|,$$

since at the root no variable is pinned yet and $D(C^*) = 0$. Thus, the key step is to show that the quantity $|m(C, x, L) - m(C, x, \infty)|$ decays exponentially with L; we will

show that, for an arbitrary node (C, x) in the computation tree, it holds that

$$|m(C,x,L) - m(C,x,\infty)| \le \alpha \max_{i,j} |m(C_{i,j}, x_{i,j}, L-1) - m(C_{i,j}, x_{i,j}, \infty)|, \tag{3.2}$$

where $\alpha = 1 - 10^{-4}$ is from Definition 3.1.

In previous applications of the correlation decay technique, the ordering of the children of each node (C, x) is usually arbitrary. Since we want to take the shape of the computation tree into consideration, this ordering becomes important to us. We will order clauses in order of increasing size, except that we leave arity-2 clauses to the end.

Unfortunately, the quantity $m(C, x, L)$ is more complicated than the plain message $R(C, x, L)$ and it is even more complicated than $\Phi(R(C, x, L))$, since it incorporates combinatorial information about the formula C and thus it does not satisfy a simple recursion (unlike $R(C, x, L)$). Nevertheless, we are able to define a multivariable quantity κ_* (see (3.5)), and we will show (see Lemma 3.8) that when $\kappa_* \le 1$, inequality (3.2) holds.

We will use the following potential function, which satisfies (3.1) as required. In general, the choice of an appropriate potential function is guided by an "educated guess."

DEFINITION 3.5. *Let* $\chi = 1/2$ *and* $\psi = 13/10$. *Define*

$$\Phi(z) := \frac{1}{\chi\psi} \log\left(\frac{z^\chi}{\psi - z^\chi}\right).$$

Let $\varphi(z)$ *denote* $\Phi'(z)$ *so that*

$$\varphi(z) := \Phi'(z) = \frac{1}{z(\psi - z^\chi)}.$$

Remark 3.6. The exact values of χ and ψ do not matter at this stage, but it is important that $0 < \chi \le 1$ and $\psi > 1$. For such values of χ and ψ, Φ^{-1} exists and is uniquely defined over the range of $\Phi(z)$ for $z \in (0, 1]$.

3.1. A general framework to bound the error. First let us calculate how the number of deficits changes in one step of the recursion. To avoid trivialities, we assume $k \ge 2$. Let (C, x) be a node in the computation tree. As in section 2.3, we first perform a preprocessing step on C which removes redundant clauses and removes clauses of arity 1, producing a new formula $\widetilde{C}$. Every clause in $\widetilde{C}$ is a clause of C, so $D(\widetilde{C}) \le D(C)$. Also, x is neither forced nor free in $\widetilde{C}$ (otherwise, (C, x) is a leaf in the computation tree). Let $d = d_x(\widetilde{C})$ and let $c_1, \ldots, c_d$ be the clauses where x occurs in $\widetilde{C}$. Recall that $w_i = |c_i| - 1$. As we mentioned in section 3, the order of $c_1, \ldots, c_d$ is important. We will order clauses in order of increasing size, except that we leave arity-2 clauses to the end. Here is some notation to describe the ordering. Let b_ℓ denote the number of clauses amongst $c_1, \ldots, c_d$ with arity ℓ. We will use the variables b'_ℓ to denote cumulative sums for $\ell > 2$, so $b'_2 = 0$ and, for $\ell \ge 3$, $b'_\ell = b'_{\ell-1} + b_\ell$. We order the clauses so that, for $\ell \ge 3$, clauses $c_{b'_{\ell-1}+1}, \ldots, c_{b'_\ell}$ have arity ℓ. Finally, clauses $c_{d-b_2+1}, \ldots, c_d$ have arity 2. Let s_i be the sum of the deficits of clauses $c_1, \ldots, c_i$. Thus,

$$s_i = \sum_{t=1}^{i} \max(0, k - w_t - 1).$$

We will now consider how the deficits change when we construct the node $(C_{i,j}, x_{i,j})$ from $(\widetilde{C}, x)$ according to the method described in section 2.3, where $j \in [w_i]$.

- The arity-2 clauses in $\widetilde{C}$ are always removed in the construction of $C_{i,j}$, resulting in a loss of deficit of $b_2(k-2)$.
- The clauses $c_1, \ldots, c_i$ are also removed in the construction of $C_{i,j}$ resulting in an additional loss of deficit of $s_{\min(i,d-b_2)}$. (The minimum is to avoid double-counting if $i > d-b_2$ since in that case some of these clauses have arity 2 and have already been counted.)
- The occurrences of x are pinned to 0 in clauses $c_{i+1}, \ldots, c_d$. Consider some $t \in \{i+1, \ldots, d\}$. If the arity of c_t is greater than k, then this pinning does not cause any increase in deficit. Also, if the arity of c_t is 2, then the clause will be removed, so there is no increase in deficit. Thus, the increase in deficit from these pinnings is at most $\max(0, b'_k - i)$.
- If $i \le d - b_2$, then $w_i > 1$ and all occurrences of $x_{i,1}, \ldots, x_{i,j-1}$ are pinned to 0, resulting in an increase in deficit of at most $(j-1)(\Delta-1)$.

Let $\mathbf{1}_{i \le d-b_2}$ be zero-one indicator variable for the event that i is at most $d-b_2$. Then, putting these observations together, we conclude that

$$D(C_{i,j}) \le D(\widetilde{C}) - b_2(k-2) - s_{\min(i,d-b_2)} + \max(0, b'_k - i) + (j-1)(\Delta-1)\mathbf{1}_{i \le d-b_2}.$$

Since $D(\widetilde{C}) \le D(C)$, we have

$$D(C_{i,j}) \le D(C) - b_2(k-2) - s_{\min(i,d-b_2)} + \max(0, b'_k - i) + (j-1)(\Delta-1)\mathbf{1}_{i \le d-b_2}. \tag{3.3}$$

Recall that the recursion for $R(C, x, L)$ depends on the function $F^{d,\mathbf{w}}(\mathbf{r})$ implicitly defined by (2.7), i.e.,

$$F^{d,\mathbf{w}}(\mathbf{r}) := \prod_{i=1}^{d} \left(1 - \prod_{j=1}^{w_i} \frac{r_{i,j}}{1 + r_{i,j}}\right), \tag{3.4}$$

where $r_{i,j} \in [0,1]$ for all $i \in [d]$ and $j \in [w_i]$. In particular, unless x is free in $\widetilde{C}$ or $L \le 0$, we have

$$R(C, x, L) = F^{d,\mathbf{w}}(\{R(C_{i,j}, x_{i,j}, L - l_{w_i})\}).$$

The $m(C, x, L)$ variables also satisfy a recursion which could be made explicit by mapping them back to the $R(C, x, L)$ variables, though we will not directly analyze this recursion. Instead, we will define a quantity $\kappa_*^{d,\mathbf{w}}(\mathbf{r})$ which tracks the rate at which $|m(C, x, L) - m(C, x, \infty)|$ decays in the recursion. Specifically, define $\kappa_*^{d,\mathbf{w}}(\mathbf{r})$ as follows:

$$\kappa_*^{d,\mathbf{w}}(\mathbf{r}) := \sum_{i=1}^{d} \sum_{j=1}^{w_i} \alpha^{-l_{w_i}} \delta^{(b_2(k-2) + s_{\min(i,d-b_2)} - \max(0, b'_k - i) - (j-1)(\Delta-1)\mathbf{1}_{i \le d-b_2})} \frac{\varphi(F^{d,\mathbf{w}}(\mathbf{r}))}{\varphi(r_{i,j})} \left| \frac{\partial F^{d,\mathbf{w}}(\mathbf{r})}{\partial r_{i,j}} \right|. \tag{3.5}$$

The main step in the proofs of Lemmas 3.2 and 3.3 will be to bound $\kappa_*^{d,\mathbf{w}}(\mathbf{r})$. By construction, the elements in $\mathbf{w}$ are in increasing order, apart from the 1's at the end, and the bound on $\kappa_*^{d,\mathbf{w}}(\mathbf{r})$ will use this fact, so we give the following definition.

DEFINITION 3.7. *Let $w_0 = 2$. A vector $\mathbf{w} = w_1, \ldots, w_d$ is* suitable *if its entries are positive integers and there is a $t \in \{0, \ldots, d\}$ such for all j in $\{1, \ldots, t\}$, $w_j \ge$*

w_{j-1}, and for all j in $\{t+1,\dots,d\}$, $w_j = 1$. Given a suitable vector $\mathbf{w}$ we use the following global notation (which depends implicitly on $\mathbf{w}$): b_ℓ is the number of entries of $\mathbf{w}$ which are equal to $\ell-1$. $b'_k = b_3+\cdots+b_k$. Finally, $s_i = \sum_{t=1}^{i} \max(0, k-w_t-1)$.

LEMMA 3.8. *Suppose that Δ and k are integers with $\Delta \geq 2$ and $k \geq 3$. Suppose that there are constants $0 < \delta < 1$ and $U > 0$ such that, for all $1 \leq d \leq \Delta$, all suitable $\mathbf{w} = w_1,\dots,w_d$, and all $\mathbf{r}$ satisfying $(1/2)^{\Delta-1}\mathbf{1} \leq \mathbf{r} \leq \mathbf{1}$, it holds that*

$$\kappa_*^{d,\mathbf{w}}(\mathbf{r}) \leq \begin{cases} 1 & \text{when } d \leq \Delta - 1, \\ U & \text{when } d = \Delta. \end{cases} \tag{3.6}$$

Then there exists a constant $\tau > 0$ such that, for every $C \in \mathcal{C}_{k,\Delta}$, every variable $x \in C$, and every integer L, it holds that

$$|R(C,x,L) - R(C,x,\infty)| \leq \tau\alpha^L. \tag{3.7}$$

Proof. We will show that there is a constant $\hat{\tau} > 0$ such that for all such C, x, and L, it holds that

$$|m(C,x,L) - m(C,x,\infty)| \leq \hat{\tau}\alpha^L. \tag{3.8}$$

Assuming (3.8) for the moment, let us conclude (3.7). Consider $C \in \mathcal{C}_{k,\Delta}$. We may assume that $L > 0$ since for $L = 0$ the inequality (3.7) holds for all sufficiently large τ (any $\tau \geq 1$ works). Consider any $x \in C$ and consider the computation tree rooted at (C,x). By the definition of $m(\cdot,\cdot,\cdot)$ and since by assumption $D(C) = 0$, we have that

$$|m(C,x,L) - m(C,x,\infty)| = \big|\Phi(R(C,x,L)) - \Phi(R(C,x,\infty))\big|. \tag{3.9}$$

Let $\eta = (1/2)^{\Delta-1}$ and let

$$K_\Phi^{\min} := \min_{x\in[\eta,1]} \Phi'(x) = \min_{x\in[\eta,1]} \varphi(x) \qquad \text{and} \qquad K_\Phi^{\max} := \max_{x\in[\eta,1]} \Phi'(x) = \max_{x\in[\eta,1]} \varphi(x).$$

Since φ is continuous and $\varphi(x) > 0$ for all $x \in [\eta,1]$, we have that $K_\Phi^{\min}$ and $K_\Phi^{\max}$ are positive. We have that

$$|R(C,x,L) - R(C,x,\infty)| \leq \frac{1}{K_\Phi^{\min}} |\Phi(R(C,x,L)) - \Phi(R(C,x,\infty))|. \tag{3.10}$$

To see (3.10), we may assume that $R(C,x,L) \neq R(C,x,\infty)$ (otherwise the inequality holds at equality), in which case the inequality follows by an immediate application of the Mean Value Theorem to the function Φ. Combining (3.9), (3.10) with (3.8) yields (3.7) with $\tau = \hat{\tau}/K_\Phi^{\min}$, as desired.

To prove (3.8), we will first show a slightly weaker claim. Namely, for all nodes (C,x) in the computation tree where $C \in \mathcal{C}_{k,\Delta}$ and $x \in C$ is a variable with degree $\leq \Delta - 1$, for all integers L, it holds that

$$|m(C,x,L) - m(C,x,\infty)| \leq K_\Phi^{\max}\alpha^L. \tag{3.11}$$

For $L \leq 0$, we have that

$$\begin{aligned} |m(C,x,L) - m(C,x,\infty)| &= \delta^{D(C)} |\Phi(R(C,x,L)) - \Phi(R(C,x,\infty))| \\ &\leq K_\Phi^{\max} |R(C,x,L) - R(C,x,\infty)| \leq K_\Phi^{\max}, \end{aligned}$$

where in the first inequality we used that $\delta \in (0,1]$, $D(C) \geq 0$ and an application of the Mean Value Theorem analogous to that used in (3.10), while in the second inequality we used that for $L \leq 0$ it holds that $R(C,x,L) = 1$ (by definition) and $0 \leq R(C,x,\infty) \leq 1$. Since $0 < \alpha < 1$, this proves (3.11) for $L \leq 0$.

To prove (3.11) for integer $L > 0$, we proceed by induction on L. Namely, we assume that $L > 0$ and that (3.11) holds for all values smaller than L (the base cases $L \leq 0$ have already been shown).

Recall from (2.4) that x is not forced in C and that $\widetilde{C}$ is the formula obtained by removing the redundant clauses in C. We may assume that x is not free in $\widetilde{C}$—otherwise, observe that $R(C,x,L) = R(C,x,\infty) = 1$ and thus $m(C,x,L) = m(C,x,\infty)$ so that (3.11) holds. We will thus focus on x's which appear only in (a nonzero number of) clauses in $\widetilde{C}$ of arity ≥ 2. Let $d = d_x(\widetilde{C})$. For $i \in [d]$ and $j \in [w_i]$, it holds that

$$\begin{aligned} m(C_{i,j}, x_{i,j}, L - l_{w_i}) &= \delta^{D(C_{i,j})}\, \Phi(R(C_{i,j}, x_{i,j}, L - l_{w_i})), \\ m(C_{i,j}, x_{i,j}, \infty) &= \delta^{D(C_{i,j})}\, \Phi(R(C_{i,j}, x_{i,j}, \infty)). \end{aligned} \tag{3.12}$$

Denote by $\mathbf{r}^{(1)}$ the vector whose coordinates are given by $R(C_{i,j}, x_{i,j}, L - l_{w_i})$ for $i \in [d]$ and $j \in [w_i]$. Denote also by $\mathbf{r}^{(2)}$ the vector whose coordinates are given by $R(C_{i,j}, x_{i,j}, \infty)$ for $i \in [d]$ and $j \in [w_i]$. Observe that

$$R(C,x,L) = R(\widetilde{C},x,L) = F^{d,\mathbf{w}}\big(\mathbf{r}^{(1)}\big) \text{ and } R(C,x,\infty) = R(\widetilde{C},x,\infty) = F^{d,\mathbf{w}}\big(\mathbf{r}^{(2)}\big).$$

Note also that $\eta\mathbf{1} \leq \mathbf{r}^{(1)}, \mathbf{r}^{(2)} \leq \mathbf{1}$ (cf. property (2.4) for the nodes of the computation tree, (2.7), footnote 7, and Remark 2.8).

We next bound $|\Phi(F^{d,\mathbf{w}}(\mathbf{r}^{(1)})) - \Phi(F^{d,\mathbf{w}}(\mathbf{r}^{(2)}))|$ in terms of $\max_{i,j} |\Phi(r^{(1)}_{i,j}) - \Phi(r^{(2)}_{i,j})|$. For convenience, denote $F := F^{d,\mathbf{w}}$, and for $i \in [d]$ and $j \in [w_i]$, let

$$z^{(1)}_{i,j} := \Phi\big(r^{(1)}_{i,j}\big), \quad z^{(2)}_{i,j} := \Phi\big(r^{(2)}_{i,j}\big).$$

For $\theta \in [0,1]$, let $z_{i,j}(\theta) := \theta\, z^{(1)}_{i,j} + (1-\theta)\, z^{(2)}_{i,j}$, and let $r_{i,j}(\theta) := \Phi^{-1}(z_{i,j}(\theta))$ (note that the inverse Φ^{-1} exists and is uniquely defined in the interval $\Phi([\eta,1])$; cf. Remark 3.6 and (3.1)). Denote by $\mathbf{r}(\theta)$ the vector whose coordinates are $r_{i,j}(\theta)$ and note that $\eta\mathbf{1} \leq \mathbf{r}(\theta) \leq \mathbf{1}$. Finally, let $h(\theta) := \Phi(F(\mathbf{r}(\theta)))$. Observe that h is differentiable for all values of $\theta \in [0,1]$. By applying the Mean Value Theorem to the function $h(\theta)$, we obtain that there exists $\theta_0 \in (0,1)$ such that

$$\Phi\big(F\big(\mathbf{r}^{(1)}\big)\big) - \Phi\big(F\big(\mathbf{r}^{(2)}\big)\big) = \left.\frac{\partial h}{\partial \theta}\right|_{\theta=\theta_0}.$$

We have that

$$\begin{aligned} \frac{\partial h}{\partial \theta} &= \frac{\partial \Phi(F(\mathbf{r}(\theta)))}{\partial \theta} = \Phi'(F(\mathbf{r}(\theta)))\frac{\partial F(\mathbf{r}(\theta))}{\partial \theta} = \Phi'(F(\mathbf{r}(\theta)))\left(\sum_{i=1}^{d}\sum_{j=1}^{w_i} \frac{\partial F(\mathbf{r}(\theta))}{\partial r_{i,j}}\frac{\partial r_{i,j}(\theta)}{\partial \theta}\right) \\ &= \Phi'(F(\mathbf{r}(\theta)))\left(\sum_{i=1}^{d}\sum_{j=1}^{w_i} \frac{1}{\Phi'(r_{i,j}(\theta))}\frac{\partial F(\mathbf{r}(\theta))}{\partial r_{i,j}}\frac{\partial z_{i,j}(\theta)}{\partial \theta}\right) \\ &= \Phi'(F(\mathbf{r}(\theta)))\left(\sum_{i=1}^{d}\sum_{j=1}^{w_i} \frac{1}{\Phi'(r_{i,j}(\theta))}\frac{\partial F(\mathbf{r}(\theta))}{\partial r_{i,j}}\big(z^{(1)}_{i,j} - z^{(2)}_{i,j}\big)\right). \end{aligned}$$

It follows that

$$
\begin{aligned}
&|m(C,x,L) - m(C,x,\infty)| \\
&= \delta^{D(C)} |\Phi(R(C,x,L)) - \Phi(R(C,x,\infty))| \\
&\le \max_{\mathbf{r}} \left\{ \delta^{D(C)} \left(\sum_{i=1}^{d} \sum_{j=1}^{w_i} \frac{\varphi(F(\mathbf{r}))}{\varphi(r_{i,j})} \Big| \frac{\partial F(\mathbf{r})}{\partial r_{i,j}} \Big| \left| z_{i,j}^{(1)} - z_{i,j}^{(2)} \right| \right) \right\} \\
&= \max_{\mathbf{r}} \left(\sum_{i=1}^{d} \sum_{j=1}^{w_i} \delta^{D(C)-D(C_{i,j})} \frac{\varphi(F(\mathbf{r}))}{\varphi(r_{i,j})} \Big| \frac{\partial F(\mathbf{r})}{\partial r_{i,j}} \Big| \left| m(C_{i,j}, x_{i,j}, L - l_{w_i}) - m(C_{i,j}, x_{i,j}, \infty) \right| \right),
\end{aligned} \tag{3.13}
$$

where the inequality follows by the triangle inequality and the fact (see Definition 3.5) that $\varphi = \Phi'$ satisfies (3.1). Now note that, for all $i \in [d]$ and $j \in [w_i]$, we have by induction that

$$
\left| m(C_{i,j}, x_{i,j}, L - l_{w_i}) - m(C_{i,j}, x_{i,j}, \infty) \right| \le K_{\Phi}^{\max} \, \alpha^{L - l_{w_i}}. \tag{3.14}
$$

From (3.13) and the bounds (3.14), we obtain

$$
|m(C,x,L) - m(C,x,\infty)| \le K_{\Phi}^{\max} \max_{\mathbf{r}} \left(\sum_{i=1}^{d} \sum_{j=1}^{w_i} \delta^{D(C)-D(C_{i,j})} \frac{\varphi(F(\mathbf{r}))}{\varphi(r_{i,j})} \Big| \frac{\partial F(\mathbf{r})}{\partial r_{i,j}} \Big| \, \alpha^{-l_{w_i}} \right) \alpha^{L}. \tag{3.15}
$$

The lower bounds on $D(C) - D(C_{i,j})$, (3.3), imply that

$$
\sum_{i=1}^{d} \sum_{j=1}^{w_i} \delta^{D(C)-D(C_{i,j})} \frac{\varphi(F(\mathbf{r}))}{\varphi(r_{i,j})} \Big| \frac{\partial F(\mathbf{r})}{\partial r_{i,j}} \Big| \, \alpha^{-l_{w_i}} \le \kappa_*^{d,\mathbf{w}}(\mathbf{r}), \tag{3.16}
$$

where in the inequality we used that $\delta \in (0,1]$.

Combining (3.15), (3.16), and the assumption (3.6) that $\kappa_*^{d,\mathbf{w}}(\mathbf{r}) \le 1$ for $d \le \Delta - 1$ yields (3.11), as wanted.

Finally, we prove (3.8) with $\hat{\tau} := K_{\Phi}^{\max} \cdot \max\{U, 1\}$, where U is the constant in assumption (3.6). For x of degree $\le \Delta - 1$, we get (3.8) immediately from (3.11). Now suppose that (C,x) are such that x has degree Δ in C. Note that, for a node (C',x') in the computation tree, x' may have degree $d = \Delta$ in C' only if (C',x') is the root of the tree. It follows that the children of the node (C,x), say $(C_{i,j}, x_{i,j})$ with $i \in [d]$ and $j \in [w_i]$, are such that $x_{i,j}$ has degree at most $\Delta - 1$ in $C_{i,j}$. Hence, by applying (3.11), we obtain that (3.14) holds for all i,j, and hence (as before) we deduce that (3.15), (3.16) hold as well. Inequality (3.8) now follows since by assumption (3.6) we have that $\kappa_*^{d,\mathbf{w}}(\mathbf{r}) \le \max\{U,1\}$ for $d \le \Delta$.

This concludes the proof of Lemma 3.8. ∎

We now state two technical lemmas which will be proved later in the paper. These lemmas verify the premise of Lemma 3.8 in the settings of Theorems 1.2 and 1.3. Note from (3.5) that $\kappa_*^{d,\mathbf{w}}(\mathbf{r})$ depends on the global quantity k and on various quantities (depending on $\mathbf{w}$) which are defined in Definition 3.7.

LEMMA 3.9. *Let k and Δ be two integers such that $k \ge \Delta$ and $\Delta \ge 200$. There are constants $0 < \delta < 1$ and $U > 0$ such that, for all $1 \le d \le \Delta$, all suitable $\mathbf{w} = w_1, \dots, w_d$, and all $\mathbf{r}$ satisfying $\mathbf{0} < \mathbf{r} \le \mathbf{1}$, it holds that*

$$
\kappa_*^{d,\mathbf{w}}(\mathbf{r}) \le \begin{cases} 1 & \text{when } d \le \Delta - 1, \\ U & \text{when } d = \Delta. \end{cases}
$$

LEMMA 3.10. *Let $\Delta = 6$ and $k = 3$. There are constants $0 < \delta < 1$ and $U > 0$ such that, for all $1 \le d \le \Delta$, all suitable $\mathbf{w} = w_1, \ldots, w_d$, and all $\mathbf{r}$ satisfying $(1/2)^{\Delta-1}\mathbf{1} \le \mathbf{r} \le \mathbf{1}$, it holds that*

$$\kappa_*^{d,\mathbf{w}}(\mathbf{r}) \le \begin{cases} 1 & \text{when } d \le \Delta - 1, \\ U & \text{when } d = \Delta. \end{cases}$$

Lemma 3.9 will be proved in section 4, and Lemma 3.10 will be proved in section 5. Using these lemmas and Lemma 3.8, we can prove Lemmas 3.2 and 3.3.

LEMMA 3.2. *There exists a constant $\tau > 0$ such that for every $C \in \mathcal{C}_{3,6}$, every variable $x \in C$, and every integer L,*

$$|R(C, x, L) - R(C, x, \infty)| \le \tau\alpha^L.$$

Proof. The lemma follows immediately from Lemmas 3.10 and 3.8. ∎

LEMMA 3.3. *Let k and Δ be two integers such that $k \ge \Delta$ and $\Delta \ge 200$. There exists a constant $\tau > 0$ such that for every $C \in \mathcal{C}_{k,\Delta}$, every variable x in C, and every integer L,*

$$|R(C, x, L) - R(C, x, \infty)| \le \tau\alpha^L.$$

Proof. The lemma follows immediately from Lemmas 3.9 and 3.8. ∎

3.2. Proof of the main theorems. In this section, we give the proofs of Theorems 1.2 and 1.3, which we restate here for convenience.

THEOREM 1.2. *There is an FPTAS for* #HyperIndSet$(3, 6)$.

Proof. First, reformulate #HyperIndSet$(3, 6)$ as the monotone CNF problem with instances in $\mathcal{C}_{3,6}$, following the reformulation in section 2.2. Then, just invoke Lemmas 2.3 and 3.2. ∎

THEOREM 1.3. *Let k and Δ be two integers such that $k \ge \Delta$ and $\Delta \ge 200$. Then there is an FPTAS for the problem* #HyperIndSet(k, Δ).

Proof. Once again, reformulate #HyperIndSet(k, Δ) as the monotone CNF problem with instances in $\mathcal{C}_{k,\Delta}$. Then invoke Lemmas 2.3 and 3.3. ∎

We have now finished the proofs of Theorems 1.2 and 1.3, except that we have not yet proved Lemmas 3.9 and 3.10, which we used to bound $\kappa_*^{d,\mathbf{w}}(\mathbf{r})$ in the proofs of Lemmas 3.2 and 3.3. Lemma 3.9 will be proved in section 4, and Lemma 3.10 will be proved in section 5. The proofs bound the multivariate decay rate function $\kappa_*^{d,\mathbf{w}}(\mathbf{r})$. This is an optimization problem that is quite complicated to solve. Moreover, we need to solve it for all possible suitable vectors $\mathbf{w}$. The analysis of bounding $\kappa_*^{d,\mathbf{w}}(\mathbf{r})$ is the technical core of our proof.

3.3. Application to counting dominating sets. In this section, we use Theorems 1.2 and 1.3 to obtain Corollary 1.4, which we restate here for convenience.

COROLLARY 1.4. *For all positive integers Δ satisfying either $\Delta \le 5$ or $\Delta \ge 199$, there is an FPTAS for the problem* #RegDomSet(Δ).

The corollary follows from the observation that a dominating set in a Δ-regular graph is defined by a collection of constraints of arity $\Delta+1$ with each variable occurring in $\Delta + 1$ constraints. The details are spelled out in the following proof.

Proof of Corollary 1.4. Let Δ be an integer satisfying either $2 \le \Delta \le 5$ or $\Delta \ge 199$. Let G be a Δ-regular graph, and denote by $\#\mathrm{DomSets}(G)$ the number of dominating sets in G.

Let $k' := \Delta + 1, \Delta' := \Delta + 1$. We will construct a k'-uniform hypergraph H where every vertex of H has degree Δ' such that the number Z_H of independent sets in H satisfies $Z_H = \#\mathrm{DomSets}(G)$. To conclude the corollary we then only have to check that, for the relevant range of Δ, we can invoke one of the FPTASs of Theorems 1.2 and 1.3 to approximate Z_H. Indeed, when $\Delta \ge 199$, we have $\Delta' \ge 200$ and $k' \ge \Delta'$. Thus, by Theorem 1.3, $\#\mathsf{HyperIndSet}(k', \Delta')$ admits an FPTAS and thus so does $\#\mathsf{RegDomSet}(\Delta)$. Similarly, when $2 \le \Delta \le 5$, we have $\Delta' \le 6$ and $k' \ge 3$. By Theorem 1.2, $\#\mathsf{HyperIndSet}(3, 6)$ admits an FPTAS and thus so does $\#\mathsf{RegDomSet}(\Delta)$.

We conclude the proof by showing the construction of H. For a vertex $v \in V$ of G, denote by $v_1, \dots, v_\Delta$ its neighbors in G (note that there are exactly Δ of those since G is Δ-regular), and let $e_v = \{v, v_1, \dots, v_\Delta\}$. Then H is the hypergraph with vertex set V and hyperedge set $\mathcal{F} = \cup_{v \in V} e_v$. It is clear from the construction that H is k'-uniform with $k' = \Delta + 1$. Further, every vertex v of H has degree $\Delta' = \Delta + 1$ since it appears in the hyperedges $e_v, e_{v_1}, \dots, e_{v_\Delta}$. This completes the construction of H.

It remains to show that the number of dominating sets in the graph G is equal to the number of independent sets in the hypergraph H. It suffices to show that $S \subseteq V$ is an independent set of H iff $V \backslash S$ is a dominating set of G. Indeed, if S is an independent set of H, then for every vertex $v \in S$, at least one of $v_1, \dots, v_\Delta$ is not in S (since e_v is a hyperedge of H), and hence $V \backslash S$ is a dominating set of G. Similarly, if $V \backslash S$ is a dominating set of G, then for every vertex $v \in S$, at least one of $v_1, \dots, v_\Delta$ is in $V \backslash S$, and hence S is an independent set of H (since each hyperedge e_v of H contains at least one vertex which does not belong to S).

This completes the proof. ∎

4. Bounding the decay rate for large Δ. This section is devoted to proving Lemma 3.9. Let k and Δ be two integers such that $k \ge \Delta$ and $\Delta \ge 200$. We start by setting up some upper bounds on the function $\kappa_*^{d,\mathbf{w}}(\mathbf{r})$, which is defined in (3.5) using notation from Definition 3.7. Consider the following definitions, which apply to suitable vectors $\mathbf{w}$:

(4.1)
$$\rho^{\mathbf{w},i}(\mathbf{r}) := \begin{cases} \alpha^{-l_{w_i}} \delta^{s_i + i - d + b_2} \sum_{j=1}^{w_i} \delta^{-(j-1)(\Delta-1)} \frac{1}{\varphi(r_{i,j})} \left| \frac{\partial F^{d,\mathbf{w}}}{\partial r_{i,j}} \right| & \text{if } 1 \le i \le d - b_2, \\ \alpha^{-l_{w_i}} \delta^{s_{d-b_2}} \frac{1}{\varphi(r_{i,1})} \left| \frac{\partial F^{d,\mathbf{w}}}{\partial r_{i,1}} \right| & \text{if } d - b_2 + 1 \le i \le d, \end{cases}$$

(4.2)
$$\kappa^{d,\mathbf{w}}(\mathbf{r}) := \varphi(F^{d,\mathbf{w}}(\mathbf{r})) \delta^{b_2(k-2)} \sum_{i=1}^{d} \rho^{\mathbf{w},i}(\mathbf{r}).$$

We first argue that $\kappa_*^{d,\mathbf{w}}(\mathbf{r}) \le \kappa^{d,\mathbf{w}}(\mathbf{r})$. To see this, note that $s_{\min(i,d-b_2)}$ is s_i for $1 \le i \le d - b_2$ and is s_{d-b_2} for $d - b_2 + 1 \le i \le d$. Also, $b'_k \le d - b_2$, so $\max(0, b'_k - i) \le d - b_2 - i$, and if $i \ge d - b_2 + 1$, then $\max(0, b'_k - i) = 0$. Finally, $(j-1)(\Delta-1)\mathbf{1}_{i \le d-b_2}$ is equal to $(j-1)(\Delta-1)$ when $1 \le i \le d - b_2$ and to 0 when $d - b_2 + 1 \le i$.

Recall from Definition 3.1 that $\alpha = 1 - 10^{-4}$. For brevity, we will denote $F^{d,\mathbf{w}}(\mathbf{r})$ by $F(\mathbf{r})$. Finally, we define the constant δ (which depends on Δ).

DEFINITION 4.1. *Let $c = 0.7$. Given Δ, define δ by $\delta^\Delta = c$.*

In order to bound $\kappa^{d,\mathbf{w}}(\mathbf{r})$, an important special case is when $b_2 = 0$. Indeed, as we will see soon, handling this special case implies an upper bound of $\kappa^{d,\mathbf{w}}(\mathbf{r})$ for the general case. Define the following quantity for suitable vectors $\mathbf{w}$:

(4.3)
$$\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) := \frac{1}{\psi - F(\mathbf{r})^\chi} \sum_{i=1}^{d} \delta^{s_i+i-d-(w_i-1)(\Delta-1)} \alpha^{-l_{w_i}} \frac{\prod_{j=1}^{w_i} \frac{r_{i,j}}{1+r_{i,j}}}{1 - \prod_{j=1}^{w_i} \frac{r_{i,j}}{1+r_{i,j}}} \sum_{j=1}^{w_i} \frac{\psi - r_{i,j}^\chi}{1+r_{i,j}}.$$

The next lemma gives an upper bound on the quantity $\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r})$.

LEMMA 4.2. *Let k and Δ be two integers such that $k \geq \Delta$ and $\Delta \geq 200$. Let d be a positive integer such that $d \leq \Delta - 1$. Let $\mathbf{w} = w_1, \ldots, w_d$ be a suitable vector with $b_2 = 0$. Then, for all $\mathbf{r}$ satisfying $\mathbf{0} < \mathbf{r} \leq \mathbf{1}$, $\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) \leq 1$.*

In the case $d = \Delta$ for $\mathbf{w} = w_1, \ldots, w_d$, a suitable vector with $b_2 = 0$, and all $\mathbf{r}$ satisfying $\mathbf{0} < \mathbf{r} \leq \mathbf{1}$, $\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) \leq 1/\delta$.

Note that in the statement of Lemma 4.2, the w_i's are positive integers in non-decreasing order, all of which are at least 2. Lemma 4.2 will be proved in the remainder of section 4. First we use it to prove Lemma 3.9, which we restate for convenience.

LEMMA 3.9. *Let k and Δ be two integers such that $k \geq \Delta$ and $\Delta \geq 200$. There are constants $0 < \delta < 1$ and $U > 0$ such that, for all $1 \leq d \leq \Delta$, all suitable $\mathbf{w} = w_1, \ldots, w_d$, and all $\mathbf{r}$ satisfying $\mathbf{0} < \mathbf{r} \leq \mathbf{1}$, it holds that*

$$\kappa_*^{d,\mathbf{w}}(\mathbf{r}) \leq \begin{cases} 1 & \text{when } d \leq \Delta - 1, \\ U & \text{when } d = \Delta. \end{cases}$$

Proof. First note that

$$\left| \frac{\partial F}{\partial r_{i,j}} \right| = \frac{F(\mathbf{r})}{r_{i,j}(1+r_{i,j})} \cdot \frac{\prod_{t=1}^{w_i} \frac{r_{i,t}}{1+r_{i,t}}}{1 - \prod_{t=1}^{w_i} \frac{r_{i,t}}{1+r_{i,t}}}.$$

Hence, for $1 \leq i \leq d - b_2$,

$$\begin{aligned} \rho^{\mathbf{w},i}(\mathbf{r}) &= F(\mathbf{r}) \delta^{s_i+i-d+b_2} \alpha^{-l_{w_i}} \frac{\prod_{j=1}^{w_i} \frac{r_{i,j}}{1+r_{i,j}}}{1 - \prod_{j=1}^{w_i} \frac{r_{i,j}}{1+r_{i,j}}} \sum_{j=1}^{w_i} \delta^{-(j-1)(\Delta-1)} \frac{1}{\varphi(r_{i,j})} \frac{1}{r_{i,j}(1+r_{i,j})} \\ &\leq F(\mathbf{r}) \delta^{s_i+i-d+b_2-(w_i-1)(\Delta-1)} \alpha^{-l_{w_i}} \frac{\prod_{j=1}^{w_i} \frac{r_{i,j}}{1+r_{i,j}}}{1 - \prod_{j=1}^{w_i} \frac{r_{i,j}}{1+r_{i,j}}} \sum_{j=1}^{w_i} \frac{\psi - r_{i,j}^\chi}{1+r_{i,j}}; \end{aligned}$$

and for $d - b_2 < i \leq d$,

$$\begin{aligned} \rho^{\mathbf{w},i}(\mathbf{r}) &= \frac{\alpha^{-l_1} F(\mathbf{r})}{\varphi(r_{i,1})} \delta^{s_{d-b_2}} \frac{r_{i,1}}{1+r_{i,1}} \\ &\leq F(\mathbf{r}) \delta^{s_{d-b_2}} \alpha^{-l_1} \frac{r_{i,1}(\psi - r_{i,1}^\chi)}{1+r_{i,1}}. \end{aligned}$$

Recall from Definition 2.7 that $l_{w_i} = \lceil \log_6(w_i + 1) \rceil$ and in particular that $l_1 = \cdots =$

$l_5 = 1$. This implies that

$$\kappa^{d,\mathbf{w}}(\mathbf{r}) \leq \frac{\delta^{b_2(k-2)}}{\psi - F(\mathbf{r})^\chi} \left(\sum_{i=1}^{d-b_2} \delta^{s_i+i-d+b_2-(w_i-1)(\Delta-1)} \alpha^{-l_{w_i}} \frac{\prod_{j=1}^{w_i} \frac{r_{i,j}}{1+r_{i,j}}}{1-\prod_{j=1}^{w_i} \frac{r_{i,j}}{1+r_{i,j}}} \sum_{j=1}^{w_i} \frac{\psi - r_{i,j}^\chi}{1+r_{i,j}} \right.$$

$$\left. + \sum_{i=d-b_2+1}^{d} \delta^{s_{d-b_2}} \frac{r_{i,1}(\psi - r_{i,1}^\chi)}{\alpha(1+r_{i,1})} \right). \tag{4.4}$$

We want to move the term in front of the parentheses in (4.4) inside. Let $\mathbf{w}' = \{w_1, \ldots, w_{d-b2}\}$ be the prefix of $\mathbf{w}$. Then

$$F^{d-b_2,\mathbf{w}'}(\mathbf{r}) = \prod_{i=1}^{d-b_2} \left(1 - \prod_{j=1}^{w_i} \frac{r_{i,j}}{1+r_{i,j}} \right).$$

First, suppose that $d \leq \Delta - 1$. By (4.4),

$$\begin{aligned}
&\kappa^{d,\mathbf{w}}(\mathbf{r}) \\
&\leq \delta^{b_2(k-2)} \left(\frac{1}{\psi - F^{d-b_2,\mathbf{w}'}(\mathbf{r})^\chi} \sum_{i=1}^{d-b_2} \delta^{s_i+i-d+b_2-(w_i-1)(\Delta-1)} \alpha^{-l_{w_i}} \frac{\prod_{j=1}^{w_i} \frac{r_{i,j}}{1+r_{i,j}}}{1-\prod_{j=1}^{w_i} \frac{r_{i,j}}{1+r_{i,j}}} \sum_{j=1}^{w_i} \frac{\psi - r_{i,j}^\chi}{1+r_{i,j}} \right. \\
&\qquad \left. + \frac{1}{\psi - \prod_{i=d-b_2+1}^{d} (1+r_{i,1})^{-\chi}} \sum_{i=d-b_2+1}^{d} \delta^{s_{d-b_2}} \frac{r_{i,1}(\psi - r_{i,1}^\chi)}{\alpha(1+r_{i,1})} \right) \\
&\leq \delta^{b_2(k-2)} \left(\widehat{\kappa}^{d-b_2,\mathbf{w}'}(\mathbf{r}) + \sum_{i=d-b_2+1}^{d} \frac{1}{\psi - (1+r_{i,1})^{-\chi}} \cdot \frac{r_{i,1}(\psi - r_{i,1}^\chi)}{\alpha(1+r_{i,1})} \right) \\
&\leq \delta^{b_2(k-2)} \left(1 + \sum_{i=d-b_2+1}^{d} \frac{1}{\psi - (1+r_{i,1})^{-\chi}} \cdot \frac{r_{i,1}(\psi - r_{i,1}^\chi)}{\alpha(1+r_{i,1})} \right),
\end{aligned} \tag{4.5}$$

where we use the definition of $\widehat{\kappa}^{d,\mathbf{w}}$ from (4.3) and Lemma 4.2 (using the fact that $d \leq \Delta - 1$) in the last step. In section 8.1.1, we verify using Mathematica's RESOLVE function that, for any $0 \leq r \leq 1$, it holds that

$$\frac{1}{\psi - (1+r)^{-\chi}} \cdot \frac{r(\psi - r^\chi)}{\alpha(1+r)} \leq 0.42. \tag{4.6}$$

Therefore, (4.5) simplifies to

$$\kappa^{d,\mathbf{w}}(\mathbf{r}) \leq \delta^{b_2(k-2)} \left(1 + 0.42 b_2\right).$$

Since $\delta^{k-2} \leq \delta^{\Delta-2} \leq c^{1-2/200}$ and $c = 0.7$, we have that

$$\kappa_*^{d,\mathbf{w}}(\mathbf{r}) \leq \kappa^{d,\mathbf{w}}(\mathbf{r}) \leq 1$$

for any integer $b_2 \geq 0$. This finishes the proof.

The bound in the case $d = \Delta$ follows the same argument; the only difference is that in (4.5) we use the $d = \Delta$ case of Lemma 4.2 and hence obtain a weaker (constant) bound on $\kappa^{d,\mathbf{w}}(\mathbf{r})$. ∎

TABLE 1
A lookup table for various constants.

Name	Value	Definition
α	$1-10^{-4}$	Definition 3.1
ψ	13/10	Definition 3.5
χ	1/2	Definition 3.5
c	0.7	Definition 4.1
δ	$\delta = c^{1/\Delta}$	Definition 4.1
c_5	$c^{1-6/200}$	Definition 4.9
K_2	1.11614	Lemma 4.5
K_6	1	Lemma 4.5
τ_2	4.5932	Lemma 4.7
τ_6	2.7805	Lemma 4.7
Y_0	≈ 0.933133	(4.31)
Y_1	≈ 0.988369	(4.31)

TABLE 2
A lookup table for various functions.

Name	Defining equation
$\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r})$	(4.3)
$\xi(w,r)$	(4.12)
$\zeta(\mathbf{w},d)$	(4.18)
$h_1(y)$	(4.30)
$h(y)$	(4.32)
$\sigma_{t,6}(\mathbf{y})$	(4.34)
$\sigma_{8,2}(\mathbf{y})$	(4.36)

4.1. Useful lemmas for the proof of Lemma 4.2. We have now finished the proof of Theorem 1.3 apart from the proof of Lemma 4.2, and the remainder of section 4 is devoted to the proof of Lemma 4.2. First, in this section, we prove some useful lemmas. To make them easier to read, we list some useful constants and functions in Tables 1 and 2.

The first two lemmas are merely technical.

LEMMA 4.3. *Let* $f(r) := \frac{\psi - r^\chi}{1+r}$ *for* $r \in (0,1]$, *and parameterize* t *in terms of* r *via* $e^t = \frac{r}{1+r}$. *Then,* f, *viewed as a function of* t, *is concave.*

Proof. Note that since r ranges from 0 to 1, we have that t ranges from $-\infty$ to $\ln(1/2)$. Further, we have that $r = \frac{e^t}{1-e^t}$, so we want to show that the function

$$\hat{f}(t) := f\Big(\frac{e^t}{1-e^t}\Big)$$

is concave for all $t \le \ln(1/2)$. We do this by verifying that $\hat{f}''(t) < 0$ for $t \le \ln(1/2)$ using Mathematica's RESOLVE function; see section 8.1.2 for details. ∎

LEMMA 4.4. *Let* x, y *be such that* $0 < y < x < 1$. *For integers* $d, w \ge 1$, *consider the functions* $f_1(d) := \frac{1-x^d}{1-y^d}$ *and* $f_2(w) := \frac{1-x^w}{1-x^{w+1}} / \frac{1-y^w}{1-y^{w+1}}$. *Then, for all integers* $d, w \ge 1$, *it holds that*

$$f_1(d) \le f_1(d+1), \quad f_2(w) \ge f_2(w+1).$$

Proof. To show that $f_1(d) \le f_1(d+1)$ for integer $d \ge 1$, we only need to show that $\bar{f}_1(y) \le \bar{f}_1(x)$ where $\bar{f}_1(t) := \frac{1-t^{d+1}}{1-t^d}$, i.e., that $\bar{f}_1(t)$ is increasing for $t \in (0,1)$.

We calculate

$$\bar{f}_1{}'(t) = \frac{t^{d-1}\left(t^{d+1} - (d+1)t + d\right)}{(1-t^d)^2},$$

which is nonnegative for all $t \in (0,1)$ since the quantity $t^{d+1}-(d+1)t+d$ is nonnegative for $t \in (0,1)$ (it vanishes at $t=1$ and is a decreasing function in t for $t \in (0,1]$ since the first derivative equals $(d+1)(t^d-1) \le 0$).

To show that $f_2(w) \ge f_2(w+1)$ for integer $w \ge 1$, we only need to show that $\bar{f}_2(x) \le \bar{f}_2(y)$ where $\bar{f}_2(t) := \frac{(1-t^w)(1-t^{w+2})}{(1-t^{w+1})^2}$, i.e., that $\bar{f}_2(t)$ is decreasing for $t \in (0,1)$. We calculate

$$\bar{f}_2{}'(t) = \frac{(1-t)t^{w-1}\left(wt^{w+2} - (w+2)t^{w+1} + t(w+2) - w\right)}{(1-t^{w+1})^3},$$

which is at most 0 for all $t \in (0,1)$ since the quantity $wt^{w+2}-(w+2)t^{w+1}+t(w+2)-w$ is at most 0 for $t \in [0,1]$ (it vanishes at $t=1$, is negative at $t=0$, and is a concave function in t for $t \in [0,1]$ since the second derivative equals $w(w+2)(w+1)(t^w - t^{w-1}) \le 0$). ∎

We now turn to the problem of upper bounding $\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r})$, which is what we need to do to prove Lemma 4.2. Recall from Definition 3.7 that $s_i = \sum_{t=1}^{i} \max(0, k - w_t - 1)$. Thus, for $i \ge 1$ and $d \le \Delta - 1$, we have

$$\begin{aligned} s_i + i - d - (w_i - 1)(\Delta - 1) &\ge s_i + i - w_i(\Delta - 1) \\ &\ge s_{i-1} + k - w_i - 1 + i - w_i(\Delta - 1) \\ &\ge s_{i-1} + k - w_i\Delta. \end{aligned}$$

Hence, from the definition of $\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r})$, we have

$$\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) \le \frac{\delta^k}{\psi - F(\mathbf{r})^\chi}\left(\sum_{i=1}^{d} \delta^{s_{i-1}-w_i\Delta}\alpha^{-l_{w_i}} \frac{\prod_{j=1}^{w_i}\frac{r_{i,j}}{1+r_{i,j}}}{1-\prod_{j=1}^{w_i}\frac{r_{i,j}}{1+r_{i,j}}} \sum_{j=1}^{w_i} \frac{\psi - r_{i,j}^\chi}{1+r_{i,j}}\right). \tag{4.7}$$

Recall from (3.4) that $F(\mathbf{r}) = \prod_{i=1}^{d}(1 - \prod_{j=1}^{w_i}\frac{r_{i,j}}{1+r_{i,j}})$. Define $t_{i,j}$ by $e^{t_{i,j}} = r_{i,j}/(1+r_{i,j})$, and let $u_i = \sum_{j=1}^{w_i} t_{i,j}$. Then we can express the right-hand side of (4.7) as a function of the $t_{i,j}$'s as follows:

$$\frac{\delta^k}{\psi - \left(\prod_{i=1}^{d}(1-e^{u_i})\right)^\chi}\left(\sum_{i=1}^{d} \delta^{s_{i-1}-w_i\Delta}\alpha^{-l_{w_i}} \frac{e^{u_i}}{1-e^{u_i}} \sum_{j=1}^{w_i} \hat{f}(t_{i,j})\right),$$

where $\hat{f}$ is as defined in Lemma 4.3. By Lemma 4.3, $\hat{f}$ is concave, so Jensen's inequality applies, showing that the quantity is at most

$$\frac{\delta^k}{\psi - \left(\prod_{i=1}^{d}(1-e^{u_i})\right)^\chi}\left(\sum_{i=1}^{d} \delta^{s_{i-1}-w_i\Delta}\alpha^{-l_{w_i}} \frac{e^{u_i}}{1-e^{u_i}} w_i \hat{f}\left(\frac{u_i}{w_i}\right)\right).$$

So we can replace each $t_{i,j}$ with u_i/w_i without decreasing the right-hand side. Equivalently, we can replace each $r_{i,j}$ with a quantity r_i so that $\prod_{j=1}^{w_i}\frac{r_{i,j}}{1+r_{i,j}} = (\frac{r_i}{1+r_i})^{w_i}$.

Define y_i such that $y_i = (1 - (\frac{r_i}{1+r_i})^{w_i})^{1/2}$. Also notice that $\delta^k \leq \delta^\Delta = c = 0.7$. From (4.7), we thus obtain that

$$\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) \leq \frac{\delta^k}{\psi - \prod_{i=1}^d y_i} \sum_{i=1}^d \delta^{s_{i-1}} g(y_i, w_i), \tag{4.8}$$

where

$$g(y,w) := c^{-w}\alpha^{-l_w} w \frac{1-y^2}{y^2}\left(1-(1-y^2)^{1/w}\right)\left(\psi - \left(\frac{(1-y^2)^{1/w}}{1-(1-y^2)^{1/w}}\right)^\chi\right), \tag{4.9}$$

and the range of y_i is $y_i \in [(1-2^{-w_i})^{1/2}, 1)$. The next lemma gives an upper bound on $g(y,w)$.

LEMMA 4.5. *For all $w \geq 2$, for all $y \in [(1-2^{-w})^{1/2}, 1)$, it holds that*

$$g(y,w) \leq K_w\, g((1-2^{-w})^{1/2}, w) = 0.15 K_w \alpha^{-l_w} w (2c)^{-w}\left(1-2^{-w}\right)^{-1}, \tag{4.10}$$

where $K_2 = 1.11614$, $K_3 = 1.03$, $K_4 = 1.01$, and $K_w = 1$ for all $w \geq 5$. In particular, $K_{w+1} \leq K_w$ for any $w \geq 2$.

Proof. The equality in (4.10) is immediate by just substituting $y = (1-2^{-w})^{1/2}$ into $g(y,w)$, so we only need to argue for the inequality in (4.10).

We make the change of variable $t = (1-y^2)^{1/w}$ so that the range of t is $(0, 1/2]$. The inequality can then be written as

$$\frac{\hat{g}(t,w)}{\hat{g}(1/2,w)} \leq K_w \text{ where } \hat{g}(t,w) := \frac{t^w}{1-t^w}(1-t)\left(\psi - \left(\frac{t}{1-t}\right)^\chi\right). \tag{4.11}$$

We verify (4.11) for $w = 2,3,4,5$ and all $t \in (0,1/2]$ using Mathematica's RESOLVE function; see section 8.1.3 for details. To obtain the lemma for $w \geq 6$, it then suffices to show that

$$\frac{\hat{g}(t,w+1)}{\hat{g}(1/2,w+1)} \leq \frac{\hat{g}(t,w)}{\hat{g}(1/2,w)} \text{ for all } t \in (0,1/2].$$

This can be massaged into

$$\overline{g}(t) \leq \overline{g}(1/2), \text{ where } \overline{g}(t) := \frac{\hat{g}(t,w+1)}{\hat{g}(t,w)} = \frac{t(1-t^w)}{1-t^{w+1}},$$

so it suffices to show that $\overline{g}$ is increasing for $t \in (0,1/2]$. We calculate

$$\overline{g}'(t) = \frac{wt^{w+1} - (w+1)t^w + 1}{(1-t^{w+1})^2},$$

which is positive for all $t \in (0,1)$ since the function $f(t) := wt^{w+1} - (w+1)t^w + 1$ satisfies $f(1) = 0$ and $f'(t) = w(w+1)t^w(1-\frac{1}{t}) < 0$ for all $t \in (0,1)$. ∎

For every positive integer w, define $\xi(w,0) = 0$. For every positive integer d, define

$$\xi(w,d) := \alpha^{-l_w}(2c)^{-w} w \left(1-2^{-w}\right)^{-1} \sum_{i=0}^{d-1} \delta^{\max\{0,k-w-1\}i}. \tag{4.12}$$

In order to upper bound $\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r})$, it is going to be useful to know when $\xi(w,d)$ is decreasing in w. This is captured by the following lemma.

LEMMA 4.6. *Let k, Δ be integers satisfying $k \geq \Delta$, $\Delta \geq 200$. Then for all integers d such that $1 \leq d \leq \Delta$, for all integers w satisfying either $2 \leq w \leq k-3$ or $w \geq k-1$, it holds that $\xi(w+1,d) \leq \xi(w,d)$.*

Proof. Recall from Definition 2.7 that $l_w = \lceil \log_6(w+1) \rceil$. Thus, $l_w - l_{w+1} \in \{0, -1\}$. Using this fact and the fact that $0 < \alpha < 1$, we obtain that

$$\frac{\xi(w+1,d)}{\xi(w,d)} = \alpha^{l_w - l_{w+1}} (2c)^{-1} (1 + w^{-1}) \frac{1 - 2^{-w}}{1 - 2^{-w-1}} \frac{\sum_{i=0}^{d-1} \delta^{\max\{0,k-w-2\}i}}{\sum_{i=0}^{d-1} \delta^{\max\{0,k-w-1\}i}}$$

$$\leq (2\alpha c)^{-1} (1 + w^{-1}) \frac{1 - 2^{-w}}{1 - 2^{-w-1}} D_{w,d}, \tag{4.13}$$

where

$$D_{w,d} := \frac{\sum_{i=0}^{d-1} \delta^{\max\{0,k-w-2\}i}}{\sum_{i=0}^{d-1} \delta^{\max\{0,k-w-1\}i}}.$$

Note that when $w \geq k-1$, $D_{w,d}$ is trivially equal to 1, so for this range of w the right-hand side of (4.13) is upper bounded by 1 (since $w \geq k-1 \geq 199$ and, from $\alpha = 1 - 10^{-4}$ and $c = 0.7$, we have $(2\alpha c)^{-1} < 3/4$).

Henceforth, we will thus focus on the range $2 \leq w \leq k-3$. To bound the right-hand side of (4.13), we will use the following bound on $D_{w,d}$:

$$D_{w,d} \leq M_w, \text{ where } M_w := \begin{cases} 2/(1+c) & \text{if } k-3 \geq w \geq 6, \\ 1.05 & \text{if } 2 \leq w \leq 5. \end{cases} \tag{4.14}$$

We will justify the bound in (4.14) shortly; for now note that the desired inequality $\xi(w+1,d)/\xi(w,d) \leq 1$ will follow from

$$M_w (2ac)^{-1} (1 + w^{-1}) \frac{1 - 2^{-w}}{1 - 2^{-w-1}} \leq 1. \tag{4.15}$$

We verify (4.15) for $w = 2, 3, 4, 5$ and also show that for $w = 6$ we have $M_w(2ac)^{-1}(1 + w^{-1}) < 1$ (for $w \geq 7$ the result follows by monotonicity); see section 8.1.4 for the code.

We finish the proof by justifying the bound in (4.14). Note that for all $2 \leq w \leq k-3$, we have that

$$D_{w,d} = \frac{1 - \delta^{(k-w-2)d}}{1 - \delta^{(k-w-1)d}} \cdot \frac{1 - \delta^{k-w-1}}{1 - \delta^{k-w-2}}.$$

To prove (4.14) for $2 \leq w \leq 5$, we use the fact that $k \geq \Delta \geq 200$, the definition $\delta^\Delta = c$, and the fact that the function $x/(1-x)$ is increasing for $x \in (0,1)$ to obtain

$$D_{w,d} \leq \frac{1 - \delta^{k-w-1}}{1 - \delta^{k-w-2}} = 1 + \frac{\delta^{k-w-2}}{1 - \delta^{k-w-2}} (1 - \delta) \leq 1 + \frac{c^{1-7/\Delta}}{1 - c^{1-7/\Delta}} (1 - c^{1/\Delta})$$

$$\leq 1 + \frac{c^{1-7/200}}{1 - c^{1-7/200}} (1 - c^{1/200}) \leq 1.05; \tag{4.16}$$

see section 8.1.4 for the calculation in the last inequality. To prove (4.14) for $6 \leq w \leq k-3$, we will show that

$$D_{w,d} \leq D_{w,\Delta} \leq D_{k-3,\Delta} = \frac{1+\delta}{1+c} < 2/(1+c). \tag{4.17}$$

To justify (4.17), note that the equality is immediate using that $\delta^\Delta = c$. Also, the very last (strict) inequality in (4.17) is immediate using $0 < \delta < 1$. In the following,

we may thus focus on proving the first two inequalities in (4.17). To justify the first inequality in (4.17), i.e., $D_{w,d} \leq D_{w,\Delta}$ for $1 \leq d \leq \Delta$, just use Lemma 4.4 for the function f_1 with $x = \delta^{k-w-2}$ and $y = \delta^{k-w-1}$. For the second inequality in (4.17), i.e., $D_{w,\Delta} \leq D_{k-3,\Delta}$ for $6 \leq w \leq k-3$, note that, by $\delta^{\Delta} = c$,

$$D_{w,\Delta} = \frac{1-c^{(k-w-2)}}{1-c^{(k-w-1)}} \cdot \frac{1-\delta^{k-w-1}}{1-\delta^{k-w-2}} = \frac{1-c^{w'}}{1-c^{w'+1}} / \frac{1-\delta^{w'}}{1-\delta^{w'+1}}, \text{ where } w' := k-w-2.$$

Then $D_{w,\Delta} \leq D_{k-3,\Delta}$ when $w \leq k-3$ (i.e., $w' \geq 1$) follows from Lemma 4.4 for the function f_2 with $x = \delta$ and $y = c$.

This concludes the proof of the bound in (4.14), which completes the proof of Lemma 4.6. ∎

For $\mathbf{w} = \{w_1, \ldots, w_d\}$ define the quantity

$$\zeta(\mathbf{w}, d) := \sum_{i=1}^{d} \delta^{s_{i-1}} \alpha^{-l_{w_i}} (2c)^{-w_i} w_i \left(1 - 2^{-w_i}\right)^{-1}. \tag{4.18}$$

In order to upper bound $\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r})$, it is going to be useful to have an upper bound on $\zeta(\mathbf{w}, d)$, and this is done in the following lemma.

LEMMA 4.7. *Let k, Δ be integers satisfying $k \geq \Delta$, $\Delta \geq 200$. Let $\tau_2 := 4.5932$ and $\tau_6 := 2.7805$. Consider $d \leq \Delta$ and let $\mathbf{w} = w_1, \ldots, w_d$ be a vector of integers satisfying $w_1 \leq \cdots \leq w_d$.*

1. *If $w_1 \geq 2$, then $\zeta(\mathbf{w}, d) \leq \tau_2$.*
2. *If $w_1 \geq 6$, then $\zeta(\mathbf{w}, d) \leq \tau_6$.*

Proof. We wish to find an upper bound for $\zeta(\mathbf{w}, d)$. The reason that the task is difficult is that the vector $\mathbf{w}$ may have up to d different entries. We say that a vector $\mathbf{w}' = \{w'_1, \ldots, w'_d\}$ of integers *dominates* $\mathbf{w}$ if the following are true.

- $w'_1 = w_1$,
- $w'_1 \leq \cdots \leq w'_d$,
- for all i, $w'_i \leq w_i$, and
- $\zeta(\mathbf{w}, d) \leq \zeta(\mathbf{w}', d)$.

So a good way to find an upper bound for $\zeta(\mathbf{w}, d)$ is to find a "simple" vector $\mathbf{w}'$ which dominates $\mathbf{w}$ and then find an upper bound for $\zeta(\mathbf{w}', d)$. To do this, we define several classes of vectors $\mathbf{w}$, depending on how "desirable" they are for proving upper bounds.

- A vector $\mathbf{w}$ of integers is "partly good" if $2 \leq w_1 \leq \cdots \leq w_d$.
- A vector $\mathbf{w}$ is "fairly good" if it is partly good and $w_1 = \cdots = w_d$ or $w_d \leq k-1$.
- A vector $\mathbf{w}$ is "very good" if it is fairly good and every w_i is in $\{w_1, k-1\}$.

The vector $\mathbf{w}$ that we start with (in the statement of the lemma) is partly good, but it will be easiest to prove upper bounds on $\zeta(\mathbf{w}, d)$ for very good vectors $\mathbf{w}$. Thus, we will define two transformations.

1. The first transformation starts with a partly good vector $\mathbf{w}$. If $\mathbf{w}$ is fairly good, then the transformation does nothing. Otherwise, it produces a partly good vector $\mathbf{w}' \neq \mathbf{w}$ which dominates $\mathbf{w}$.
2. The second transformation starts with a fairly good vector $\mathbf{w}$. If $\mathbf{w}$ is very good, then the transformation does nothing. Otherwise, it produces a fairly good vector $\mathbf{w}' \neq \mathbf{w}$ which dominates $\mathbf{w}$.

Both transformations make progress in the sense that there exists an i such that $w'_i < w_i$.

If we start with any partly good vector $\mathbf{w}$ and repeatedly apply the first transformation, then, after a finite number of transformations, we must obtain a fairly good vector $\mathbf{w}'$ which dominates $\mathbf{w}$. (The reason that a finite number of transformations suffices is that each individual transformation makes progress, but the entries stay sorted, and the first coordinate never changes.) Next, we apply the second transformation repeatedly, starting from $\mathbf{w}'$. Again, after a finite number of transformations, we end up with a very good vector $\mathbf{w}''$ which dominates $\mathbf{w}'$ and therefore dominates $\mathbf{w}$. An upper bound on $\zeta(\mathbf{w}'', d)$ gives an upper bound on $\zeta(\mathbf{w}, d)$. So to finish the proof, we must show that the two transformations are possible. Then we must show that for every very good vector $\mathbf{w}$ with $w_1 \geq 2$, $\zeta(\mathbf{w}, d) \leq \tau_2$, and we must show that for every very good vector $\mathbf{w}$ with $w_1 \geq 6$, $\zeta(\mathbf{w}, d) \leq \tau_6$.

Transformation 1. Start with a partly good vector $\mathbf{w}$. If $\mathbf{w}$ is fairly good, do nothing. Otherwise, $w_d > \max(w_1, k-1)$. Choose the integer t to be as small as possible, subject to the constraint that, for all i in $t+1, \ldots, d$, we have $w_i = w_d$. Note that $1 \leq t < d$ and $w_t < w_d$. Recall from Definition 3.7 that $s_i = \sum_{j=1}^{i} \max(0, k - w_j - 1)$. Since $w_{t+1} = \cdots = w_d$, we have for any $j \in \{0, \ldots, d-t\}$ that $s_{t+j} = s_t + \max\{0, k - w_d - 1\}j$. This means that $\sum_{i=t+1}^{d} \delta^{s_{i-1}} = \delta^{s_t} \sum_{i=0}^{d-t-1} \delta^{\max\{0,k-w_d-1\}i}$. So,

$$\zeta(\mathbf{w}, d) = \sum_{i=1}^{t} \delta^{s_{i-1}} \alpha^{-l_{w_i}} (2c)^{-w_i} w_i \left(1 - 2^{-w_i}\right)^{-1} + \delta^{s_t} \xi(w_d, d-t). \tag{4.19}$$

By Lemma 4.6, $\xi(w_d, d-t) \leq \xi(\max(w_t, k-1), d-t)$, so decreasing $w_{t+1} = \cdots = w_d$ from w_d to $\max(w_t, k-1)$ does not decrease $\zeta(\mathbf{w}, d)$. Thus, the transformation sets $w'_{t+1} = \cdots = w'_d = \max(w_t, k-1)$ and $w'_1, \ldots, w'_t = w_1, \ldots, w_t$.

Transformation 2. Start with a fairly good vector $\mathbf{w}$. If $\mathbf{w}$ is very good, do nothing. Otherwise, $w_1 < w_d \leq k-1$. Choose the integer t to be as small as possible, subject to the constraint that, for all i in $t+1, \ldots, d$, we have $w_i = k-1$. Clearly, $t \leq d$. We defined $\xi(w, 0)$ to be 0, so from (4.19) we have

$$\zeta(\mathbf{w}, d) = \sum_{i=1}^{t} \delta^{s_{i-1}} \alpha^{-l_{w_i}} (2c)^{-w_i} w_i \left(1 - 2^{-w_i}\right)^{-1} + \delta^{s_t} \xi(k-1, d-t). \tag{4.20}$$

Now, choose the integer t' to be as small as possible, subject to the constraint that, for all $i \in \{t'+1, \ldots, t\}$, $w_i = w_t$. Since $\mathbf{w}$ is not very good, $1 \leq t' < t$ and $w_{t'} < w_t < k-1$. Since $w_{t'+1} = \cdots = w_t$, we can decompose the right-hand side of (4.20) as

$$\zeta(\mathbf{w}, d) = \sum_{i=1}^{t'} \delta^{s_{i-1}} \alpha^{-l_{w_i}} (2c)^{-w_i} w_i \left(1 - 2^{-w_i}\right)^{-1} + \delta^{s_{t'}} \xi(w_t, t-t') + \delta^{s_t} \xi(k-1, d-t).$$

Now $w_{t'} < w_t < k-1$ implies $w_{t'} \leq k-3$ and $w_t \leq k-2$, so by Lemma 4.6, $\xi(w_t, d-t) \leq \xi(w_{t'}, d-t)$, so decreasing $w_{t'+1} = \cdots = w_t$ from w_t to $w_{t'}$ does not decrease $\zeta(\mathbf{w}, d)$. Thus, the transformation sets $w'_{t'+1} = \cdots = w'_t = w_{t'}$ and $w'_1, \ldots, w'_{t'} = w_1, \ldots, w_{t'}$ and $w'_{t+1} = \cdots = w'_d = w_{t+1} = \cdots = w_d = k-1$.

To finish the proof, we must show that for every very good vector $\mathbf{w}$ with $w_1 \geq 2$, $\zeta(\mathbf{w}, d) \leq \tau_2$, and we must show that for every very good vector $\mathbf{w}$ with $w_1 \geq 6$, $\zeta(\mathbf{w}, d) \leq \tau_6$. Now if $\mathbf{w}$ is a very good vector, then choose t as small as possible so that $w_{t+1} = \cdots = w_d = k-1$. Note that $0 \leq t \leq d$ and $w_1 = \cdots = w_t$ and

$w_{t+1} = \cdots = w_d = k-1$. Thus,

$$\zeta(\mathbf{w}, d) \le \xi(w_1, t) + \xi(k-1, d-t). \tag{4.21}$$

To bound the terms in (4.21), note that $\xi(k-1,0) = 0$ and for any $1 \le d-t \le \Delta \le k$ it holds that

(4.22)
$$\xi(k-1,d-t) = \alpha^{-l_{k-1}}(2c)^{-(k-1)}(k-1)(1-2^{-(k-1)})^{-1}(d-t) \le 2k^2(2\alpha c)^{-(k-1)} < 10^{-10},$$

where the last inequality follows from $(2\alpha c)^{-1} < 3/4$ and $k \ge 200$.

For $t = 0$, we have $\xi(w_1, t) = 0$ and $\xi(k-1, d-t) < 10^{-10}$ by (4.22). Thus assume $t > 0$ so that $w_1 < k-1$. Then, by Lemma 4.6, for $k-1 > w_1 \ge 2$, we have $\xi(w_1, t) \le \xi(2, t)$ and

$$\begin{aligned}
\xi(2,t) &= \alpha^{-1}(2c)^{-2}2\left(1-2^{-2}\right)^{-1}\sum_{i=0}^{t-1}\delta^{(k-3)i} \\
&\le \alpha^{-1}(2c)^{-2}2\left(1-2^{-2}\right)^{-1}\frac{1}{1-\delta^{k-3}} \\
&= \alpha^{-1}(2c)^{-2}2\left(1-2^{-2}\right)^{-1}\frac{1}{1-c^{(k-3)/\Delta}} \\
&\le \alpha^{-1}(2c)^{-2}2\left(1-2^{-2}\right)^{-1}\frac{1}{1-c^{1-3/200}}. \qquad (4.23)
\end{aligned}$$

The numerical calculation in section 8.1.5 shows that this is at most 4.5931. Since this plus 10^{-10} is less than τ_2, we obtain the first part of the lemma. Similarly,

$$\xi(6,t) \le \alpha^{-2}(2c)^{-6}6\left(1-2^{-6}\right)^{-1}\frac{1}{1-c^{1-7/200}} \le 2.78045; \tag{4.24}$$

see section 8.1.5 for the calculation in the last inequality. Since $2.78045 + 10^{-10} < \tau_6$, we obtain the second part of the lemma, and we have finished the proof of the lemma. □

4.2. A quick proof of a weaker theorem. Our goal is to prove Lemma 4.2, but the proof, which will be given in section 4.3, is a little bit technical. In order to give the intuition, without getting into technical details, we first state and prove a weaker version of the lemma. Lemma 4.8 below is identical to Lemma 4.2, except that the condition $k \ge \Delta$ has been strengthened to $k \ge 2.64\Delta$. Using Lemma 4.8 in place of Lemma 4.2 strengthens the condition to $k \ge 2.64\Delta$ in Lemmas 3.9 and 3.3. Thus, Lemma 4.8 immediately gives a weaker version of Theorem 1.3 where the condition $k \ge \Delta$ is replaced with $k \ge 2.64\Delta$.

LEMMA 4.8. *Let k and Δ be two integers such that $k \ge 2.64\Delta$ and $\Delta \ge 200$. Let d be a positive integer such that $d \le \Delta - 1$. Let $\mathbf{w} = w_1, \ldots, w_d$ be a suitable vector with $b_2 = 0$. Then, for all $\mathbf{r}$ satisfying $\mathbf{0} < \mathbf{r} \le \mathbf{1}$, $\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) \le 1$.*

In the case $d = \Delta$ for $\mathbf{w} = w_1, \ldots, w_d$, a suitable vector with $b_2 = 0$, and all $\mathbf{r}$ satisfying $\mathbf{0} < \mathbf{r} \le \mathbf{1}$, $\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) \le 1/\delta$.

Proof. Recall that for $d \le \Delta - 1$,

$$\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) \le \frac{\delta^k}{\psi - \prod_{i=1}^d y_i}\sum_{i=1}^d \delta^{s_{i-1}} g(y_i, w_i), \tag{4.8}$$

where $y_i = (1 - (\frac{r_i}{1+r_i})^{w_i})^{1/2}$ and $y_i \in [(1-2^{-w_i})^{1/2}, 1)$. From Lemma 4.5 we have $g(y_i, w_i) \le 0.15 K_2 \alpha^{-l_{w_i}} w_i (2c)^{-w_i} (1-2^{-w_i})^{-1}$, so, from the definition of $\zeta(\mathbf{w}, d)$ (see (4.18)),

$$\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) \le \frac{\delta^k}{\psi - 1} 0.15 K_2 \zeta(\mathbf{w}, d) \le \frac{c^{2.64}}{\psi - 1} 0.15 K_2 \tau_2, \tag{4.25}$$

where in the last inequality we used the fact that $\zeta(\mathbf{w}, d) \le \tau_2$ from the first part of Lemma 4.7 and the fact that $\delta^k \le \delta^{2.64\Delta} = c^{2.64}$. It is a matter of a simple numerical calculation to check that the right-hand side of (4.25) is less than 1; see section 8.1.6 for details. Thus, we have shown that $\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) < 1$.

The case $d = \Delta$ has the same proof; the only difference is that in (4.7) and (4.8) we replace δ^k by δ^{k-1}, losing a factor of $1/\delta$ (the upper bound is valid using the same argument as before, now using inequality $s_i + i - d - (w_i - 1)(\Delta - 1) \ge s_{i-1} + (k - 1) - w_i \Delta$). ∎

4.3. The proof of Lemma 4.2. Recall that our actual goal is to prove Lemma 4.2, which is stronger than Lemma 4.8 because it only assumes $k \ge \Delta$, not $k \ge 2.64\Delta$. Recall that

$$\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) \le \frac{\delta^k}{\psi - \prod_{i=1}^d y_i} \sum_{i=1}^d \delta^{s_{i-1}} g(y_i, w_i), \tag{4.8}$$

where $y_i = (1 - (\frac{r_i}{1+r_i})^{w_i})^{1/2}$ so that $y_i \in [(1-2^{-w_i})^{1/2}, 1)$ and the function $g(y, w)$ is given by (4.9). Using the fact that $k \ge \Delta$ and $\delta^\Delta = c$, we obtain the inequality

$$\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) \le \frac{c}{\psi - \prod_{i=1}^d y_i} \sum_{i=1}^d \delta^{s_{i-1}} g(y_i, w_i). \tag{4.26}$$

Lemma 4.5 gives an upper bound on $g(y, w)$ in terms of the constants K_w. Since $K_w = 1$ for all $w \ge 5$, we want to split the sum in (4.26) for $w \le 5$ and $w \ge 6$. More generally, we split the summation in the bound (4.26) at an index $t \le d$ using Lemma 4.5 as follows:

$$\begin{aligned}
&\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) \\
&\le \frac{c}{\psi - \prod_{i=1}^d y_i} \left(\sum_{i=1}^t \delta^{s_{i-1}} g(y_i, w_i) + \sum_{i=t+1}^d \delta^{s_{i-1}} g(y_i, w_i) \right) \\
&\le \frac{c}{\psi - \prod_{i=1}^t y_i} \left(\sum_{i=1}^t \delta^{s_{i-1}} g(y_i, w_i) + 0.15 \delta^{s_t} K_{w_{t+1}} \sum_{i=t+1}^d \delta^{s_{i-1} - s_t} \alpha^{-l_{w_i}} (2c)^{-w_i} w_i \left(1 - 2^{-w_i}\right)^{-1} \right) \\
&= \frac{c}{\psi - \prod_{i=1}^t y_i} \left(\sum_{i=1}^t \delta^{s_{i-1}} g(y_i, w_i) + 0.15 \delta^{s_t} K_{w_{t+1}} \sum_{i=t+1}^d \delta^{s'_{i-1}} \alpha^{-l_{w_i}} (2c)^{-w_i} w_i \left(1 - 2^{-w_i}\right)^{-1} \right), \\
&= \frac{c}{\psi - \prod_{i=1}^t y_i} \left(\sum_{i=1}^t \delta^{s_{i-1}} g(y_i, w_i) + 0.15 \delta^{s_t} K_{w_{t+1}} \zeta(\mathbf{w}', d-t) \right),
\end{aligned} \tag{4.27}$$

where $s'_{i-1} = s_{i-1} - s_t = \sum_{j=t+1}^{i-1} \max(0, k - w_j - 1)$, $\mathbf{w}' = \{w_{t+1}, \ldots, w_d\}$, and the function $\zeta(\mathbf{w}, d)$ is as defined in (4.18).

Intuitively, the term $\zeta(\mathbf{w}', d-t)$ bounds a tail sum coming from the last $d - t$ clauses corresponding to the variables $w_{t+1}, \ldots, w_d$. Recall from the statement of Lemma 4.2 and the sentences following its statement that the w_j's are in increasing

order. Preferably, we want to choose the index t to split the sum in (4.27) so that $w_t \le 5$ and $w_{t+1} \ge 6$. However, we also do not want too many terms to be in the first sum (since each of these will cause us work), so we insist that $t \le 8$. When $t = 8$, we will use the bound $w_{t+1} \ge 2$. If $t < 8$, we will be able to use the stronger bound $w_{t+1} \ge 6$.

The first step in the proof of Lemma 4.2 is to find a good way to control $g(y, w)$ for $2 \le w \le 5$. To this end, we use a piecewise linear function to upper bound $g(y, w)$. In particular, in section 8.1.7, we verify using Mathematica's RESOLVE function that, for $w = 2, 3, 4, 5$ and $y \in [(1 - 2^{-w})^{1/2}, 1)$, it holds that

$$g(y, w) \le h_1(y), \text{ where } h_1(y) := \min\{0.2279, -1.5y + 1.6276, -8y + 8.052\}. \tag{4.28}$$

Moreover, let $t_6 + 1$ be the first index where $w_{t_6+1} \ge 6$. Therefore, $w_i \le 5$ for all $2 \le i \le t_6$. Then applying (4.27) with $t = t_6$ and Lemma 4.7,

$$\begin{aligned} \widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) &\le \frac{c}{\psi - \prod_{i=1}^{t_6} y_i} \left(\sum_{i=1}^{t_6} \delta^{s_{i-1}} g(y_i, w_i) + 0.15\delta^{s_{t_6}} \tau_6 \right) \\ &\le \frac{c}{\psi - \prod_{i=1}^{t_6} y_i} \left(\sum_{i=1}^{t_6} \delta^{(i-1)(k-6)} h_1(y_i) + 0.15\delta^{t_6(k-6)} \tau_6 \right), \end{aligned} \tag{4.29}$$

where, for $i \in \{1, \dots, t_6\}$, y_i is in the range $[(1 - 2^{-w_i})^{1/2}, 1)$. Since each w_i is at least 2, we have $(1 - 2^{-w_i})^{1/2} \ge \sqrt{3}/2$, so each $y_i \in [\sqrt{3}/2, 1]$.

The function $h_1(y)$ is nonincreasing, and in fact

$$h_1(y) = \begin{cases} 0.2279 & \text{if } \sqrt{3}/2 \le y \le Y_0, \\ -1.5y + 1.6276 & \text{if } Y_0 \le y \le Y_1, \\ -8y + 8.052 & \text{if } Y_1 \le y \le 1, \end{cases} \tag{4.30}$$

where $Y_0 \approx 0.933133$ and $Y_1 \approx 0.988369$ are two constants such that

$$\begin{aligned} -1.5Y_0 + 1.6276 &= 0.2279, \\ -1.5Y_1 + 1.6276 &= -8Y_1 + 8.052. \end{aligned} \tag{4.31}$$

Notice that $h_1(y_i)$ is a constant for $y_i \in [\sqrt{3}/2, Y_0]$. Moreover, $\frac{c}{\psi - \prod_{i=1}^{t} y_i}$ is an increasing function of y_i for each i. Hence, to upper bound the right-hand side of (4.29), we may restrict the y_i's to be in the interval $[Y_0, 1]$. More precisely, let

$$h(y) := \begin{cases} -1.5y + 1.6276 & \text{if } Y_0 \le y \le Y_1, \\ -8y + 8.052 & \text{if } Y_1 \le y \le 1. \end{cases} \tag{4.32}$$

DEFINITION 4.9. *Let* $c_5 := c^{1-6/200}$.

Observe that $c_5 = \delta^{\Delta(1-6/200)} \ge \delta^{\Delta-6} \ge \delta^{k-6}$ as $k \ge \Delta \ge 200$ and $\delta < 1$. By (4.29) we have that

$$\begin{aligned} \widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) &\le \frac{c}{\psi - \prod_{i=1}^{t_6} y_i} \left(\sum_{i=1}^{t_6} \delta^{(i-1)(k-6)} h(y_i) + 0.15 c_5^{t_6} \tau_6 \right) \\ &\le \frac{c}{\psi - \prod_{i=1}^{t_6} y_i} \left(\sum_{i=1}^{t_6} c_5^{i-1} h(y_i) + 0.15 c_5^{t_6} \tau_6 \right), \end{aligned} \tag{4.33}$$

where each $y_i \in [Y_0, 1]$.

What is left is to bound the right-hand side of (4.33). For $y_i \in [Y_0, 1]$, define

$$\sigma_{t,6}(\mathbf{y}) := \frac{c}{\psi - \prod_{i=1}^{t} y_i} \left(\sum_{i=1}^{t} c_5^{i-1} h(y_i) + 0.15 c_5^t \tau_6 \right), \tag{4.34}$$

where τ_6 can be found in Table 1. The following lemma, which is proved in section 4.4, bounds $\sigma_{t,6}(\mathbf{y})$ when $0 \le t \le 7$.

LEMMA 4.10. *Let t be an integer such that $0 \le t \le 7$. For any $(y_1, \dots, y_t)$ where each y_i is in the range $[Y_0, 1]$, we have $\sigma_{t,6}(\mathbf{y}) \le 1$.*

Combining (4.33), (4.34), and Lemma 4.10, we get an upper bound on $\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r})$ provided $\mathbf{w}$ is such that $t_6 \le 7$. If $t_6 \ge 8$, we will set $t = 8$ in (4.27). Similarly to deriving (4.33), we get that

$$\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) \le \frac{c}{\psi - \prod_{i=1}^{8} y_i} \left(\sum_{i=1}^{8} c_5^{i-1} h(y_i) + 0.15 c_5^8 K_2 \tau_2 \right), \tag{4.35}$$

where each $y_i \in [Y_0, 1]$ and K_2 and τ_2 can be found in Table 1. Similarly to (4.34), define $\sigma_{8,2}(\mathbf{y})$ to be the right-hand side of (4.35). Namely,

$$\sigma_{8,2}(\mathbf{y}) := \frac{c}{\psi - \prod_{i=1}^{8} y_i} \left(\sum_{i=1}^{8} c_5^{i-1} h(y_i) + 0.15 c_5^8 K_2 \tau_2 \right). \tag{4.36}$$

The next lemma bounds $\sigma_{8,2}(\mathbf{y})$ and is proved in section 4.4.

LEMMA 4.11. *For any $(y_1, \dots, y_8)$ where each y_i is in the range $[Y_0, 1]$, we have $\sigma_{8,2}(\mathbf{y}) \le 1$.*

We can now prove Lemma 4.2, which we restate here for convenience.

LEMMA 4.2. *Let k and Δ be two integers such that $k \ge \Delta$ and $\Delta \ge 200$. Let d be a positive integer such that $d \le \Delta - 1$. Let $\mathbf{w} = w_1, \dots, w_d$ be a suitable vector with $b_2 = 0$. Then, for all $\mathbf{r}$ satisfying $\mathbf{0} < \mathbf{r} \le \mathbf{1}$, $\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) \le 1$.*

In the case $d = \Delta$ for $\mathbf{w} = w_1, \dots, w_d$, a suitable vector with $b_2 = 0$, and all $\mathbf{r}$ satisfying $\mathbf{0} < \mathbf{r} \le \mathbf{1}$, $\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) \le 1/\delta$.

Proof. Consider t_6.

1. If $t_6 \le 7$, then by (4.33), (4.34), and Lemma 4.10, $\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) \le \sigma_{t_6,6}(\mathbf{y}) \le 1$.
2. Otherwise $t_6 \ge 8$. By (4.35), (4.36), and Lemma 4.11, $\widehat{\kappa}^{d,\mathbf{w}}(\mathbf{r}) \le \sigma_{8,2}(\mathbf{y}) \le 1$.

The case $d = \Delta$ has the same proof. Like the proof of Lemma 4.8, the only difference in the $d = \Delta$ case is that in (4.7) and (4.8) we replace δ^k by δ^{k-1}, losing a factor of $1/\delta$ (the upper bound is valid using the same argument as before, now using inequality $s_i + i - d - (w_i - 1)(\Delta - 1) \ge s_{i-1} + (k-1) - w_i \Delta$). ∎

4.4. Remaining proofs. In this section we provide technical details of Lemmas 4.10 and 4.11, which we restate for convenience.

LEMMA 4.10. *Let t be an integer such that $0 \le t \le 7$. For any $(y_1, \dots, y_t)$ where each y_i is in the range $[Y_0, 1]$, we have $\sigma_{t,6}(\mathbf{y}) \le 1$.*

Proof. Recall that

$$\sigma_{t,6}(\mathbf{y}) := \frac{c}{\psi - \prod_{i=1}^{t} y_i} \left(\sum_{i=1}^{t} c_5^{i-1} h(y_i) + 0.15 c_5^t \tau_6 \right), \tag{4.34}$$

where $c_5 = c^{1-6/200}$, $c = 0.7$, and $\tau_6 = 2.7805$ are as in Table 1 and the function $h(y)$ is as given by (4.32). For $t = 0$, we have

$$\sigma_{0,6}(\mathbf{y}) = \frac{0.15c\tau_6}{\psi - 1} \approx 0.973044 < 1; \tag{4.37}$$

see section 8.1.8 for the calculation. For $t = 1$, we have

$$\sigma_{1,6}(\mathbf{y}) = \frac{c}{\psi - y_1}\big(h(y_1) + 0.15c_5\tau_6\big) \le 1; \tag{4.38}$$

see section 8.1.8 for the verification using Mathematica's RESOLVE function. Thus, we may assume that $t \ge 2$ henceforth.

For the sake of contradiction, suppose that there exists $\mathbf{y}'$ such that $\sigma_{t,6}(\mathbf{y}') > 1$ for some $2 \le t \le 7$. We will gradually adjust the variables y_i' without decreasing $\sigma_{t,6}(\mathbf{y}')$ until there is only one variable left, in which case we will be able to exclude the possibility that $\sigma_{t,6}(\mathbf{y}') > 1$.

We first observe that the partial derivative of $\sigma_{t,6}(\mathbf{y})$ with respect to y_i is

$$\frac{\partial\sigma_{t,6}(\mathbf{y})}{\partial y_i} = \frac{1}{\psi - \prod_{j=1}^t y_j}\left(\frac{\sigma_{t,6}(\mathbf{y})\prod_{j=1}^t y_j}{y_i} - 1.5cc_5^{i-1}\right) \tag{4.39}$$

if $Y_0 \le y_i \le Y_1$ and

$$\frac{\partial\sigma_{t,6}(\mathbf{y})}{\partial y_i} = \frac{1}{\psi - \prod_{j=1}^t y_j}\left(\frac{\sigma_{t,6}(\mathbf{y})\prod_{j=1}^t y_j}{y_i} - 8cc_5^{i-1}\right) \tag{4.40}$$

if $Y_1 \le y_i \le 1$.

Suppose that there exists an index $3 \le i \le t$ such that $y_i' \le Y_1$. Using our initial assumption that $\sigma_{t,6}(\mathbf{y}') > 1$, we then have (from (4.39) and $y_j' \ge Y_0$) that

$$\left.\frac{\partial\sigma_{t,6}(\mathbf{y})}{\partial y_i}\right|_{\mathbf{y}=\mathbf{y}'} > \frac{1}{\psi - \prod_{j=1}^t y_j'}\left(Y_0^{t-1} - 1.5cc_5^{i-1}\right) > 0 \tag{4.41}$$

for any $2 \le t \le 7$ and $3 \le i \le t$; see section 8.1.8 for the verification of the last inequality. Hence $\sigma_{t,6}(\mathbf{y})$ is increasing in this y_i', and we may thus assume $y_i' \ge Y_1$ for all $3 \le i \le t$.

Suppose that $y_2' \le Y_1$. Then, using that $y_i' \ge Y_1$ for all $3 \le i \le t$ and $y_1' \ge Y_0$, together with our assumption that $\sigma_{t,6}(\mathbf{y}') > 1$, we have (from (4.39)) that

$$\left.\frac{\partial\sigma_{t,6}(\mathbf{y})}{\partial y_2}\right|_{\mathbf{y}=\mathbf{y}'} > \frac{1}{\psi - \prod_{j=1}^t y_j'}\left(Y_0Y_1^{t-2} - 1.5cc_5\right) > 0 \tag{4.42}$$

for any $2 \le t \le 7$; see section 8.1.8 for the verification of the last inequality. Arguing as before, we may therefore assume $y_i' \ge Y_1$ for all $2 \le i \le t$.

Suppose that there exists an index $2 \le i \le \min\{5, t\}$ such that $y_i' > Y_1$. Since $y_i' \ge Y_1$ for all $2 \le i \le t$ and $y_1' \ge Y_0$, we obtain (from (4.34)) the following upper bound on $\sigma_{t,6}(\mathbf{y}')$ (using also the fact that h is decreasing):

$$\sigma_{t,6}(\mathbf{y}') \le \frac{c}{\psi - 1}\left(h(Y_0) + \sum_{i=2}^t c_5^{i-1}h(Y_1) + 0.15c_5^t\tau_6\right). \tag{4.43}$$

Plugging (4.43) into (4.40) and using the fact that the y_j's are at most 1, we obtain that

(4.44)
$$\left.\frac{\partial\sigma_{t,6}(\mathbf{y})}{\partial y_i}\right|_{\mathbf{y}=\mathbf{y}'} \le \frac{1}{\psi-\prod_{j=1}^{t} y_j'}\left(\frac{c}{\psi-1}\left(h(Y_0)+c_5h(Y_1)\frac{1-c_5^{t-1}}{1-c_5}+0.15c_5^t\tau_6\right)-8cc_5^{i-1}\right)<0;$$

see section 8.1.8 for the verification of the last inequality. Therefore, $\sigma_{t,6}(\mathbf{y})$ is decreasing in y_i' for any $2 \le i \le \min\{5,t\}$, and we may therefore assume that $y_i' = Y_1$ for any $2 \le i \le \min\{5,t\}$.

For $2 \le t \le 5$, using the fact that $y_2' = \cdots = y_t' = Y_1$, we thus have that

$$\sigma_{t,6}(\mathbf{y}') = \frac{c}{\psi - y_1' Y_1^{t-1}}\left(h(y_1')+c_5h(Y_1)\frac{1-c_5^{t-1}}{1-c_5}+0.15c_5^t\tau_6\right) > 1, \tag{4.45}$$

which is false for all $y_1' \in [Y_0,1]$; see section 8.1.8 for the proof using Mathematica's RESOLVE function. Similarly, for $t=6$, using that $y_2' = \cdots = y_5' = Y_1$, $Y_1 \le y_6' \le 1$ and that h is decreasing, we have that

$$1 < \sigma_{6,6}(\mathbf{y}') \le \frac{c}{\psi - y_1' Y_1^4}\left(h(y_1')+c_5h(Y_1)\frac{1-c_5^{5}}{1-c_5}+0.15c_5^6\tau_6\right), \tag{4.46}$$

which is false for all $y_1' \in [Y_0,1]$; see section 8.1.8 for the proof using Mathematica's RESOLVE function. Finally, for $t=7$, we have that $y_2' = y_3' = y_4' = y_5' = Y_1$. Using this and $Y_1 \le y_6', y_7' \le 1$, we obtain that

$$1 < \sigma_{7,6}(\mathbf{y}') \le \frac{c}{\psi - y_1' Y_1^4}\left(h(y_1')+c_5h(Y_1)\frac{1-c_5^{6}}{1-c_5}+0.15c_5^7\tau_6\right), \tag{4.47}$$

which is false for all $y_1' \in [Y_0,1]$; see section 8.1.8 for the proof using Mathematica's RESOLVE function.

Thus, for all $2 \le t \le 7$, the assumption that there exists $\mathbf{y}'$ such that $\sigma_{t,6}(\mathbf{y}') > 1$ has lead to a contradiction, thus completing the proof of Lemma 4.10 for all $0 \le t \le 7$. □

The proof of Lemma 4.11 is very similar.

LEMMA 4.11. *For any $(y_1,\dots,y_8)$ where each y_i is in the range $[Y_0,1]$, we have $\sigma_{8,2}(\mathbf{y}) \le 1$.*

Proof. Recall that

$$\sigma_{8,2}(\mathbf{y}) := \frac{c}{\psi - \prod_{i=1}^{8} y_i}\left(\sum_{i=1}^{8} c_5^{i-1}h(y_i) + 0.15c_5^t K_2\tau_2\right), \tag{4.36}$$

where $c_5 = c^{1-6/200}$, $c = 0.7$, $K_2 = 1.11614$, and $\tau_2 = 2.7805$ are as in Table 1 and the function $h(y)$ is given by (4.32). For the sake of contradiction, suppose that there exists $\mathbf{y}'$ such that $\sigma_{8,2}(\mathbf{y}') > 1$. We will gradually adjust the variables y_i' without decreasing $\sigma_{8,2}(\mathbf{y}')$ until there is only one variable left, in which case we can directly verify that $\sigma_{8,2}(\mathbf{y}') > 1$ is impossible.

Identically to Lemma 4.10, the partial derivative of $\sigma_{8,2}(\mathbf{y})$ with respect to y_i is

$$\frac{\partial\sigma_{8,2}(\mathbf{y})}{\partial y_i} = \frac{1}{\psi - \prod_{j=1}^{8} y_j}\left(\frac{\sigma_{8,2}(\mathbf{y})\prod_{j=1}^{8} y_j}{y_i} - 1.5cc_5^{i-1}\right) \tag{4.39}$$

if $Y_0 \le y_i \le Y_1$ and

$$\frac{\partial \sigma_{8,2}(\mathbf{y})}{\partial y_i} = \frac{1}{\psi - \prod_{j=1}^{8} y_j}\left(\frac{\sigma_{8,2}(\mathbf{y})\prod_{j=1}^{8} y_j}{y_i} - 8cc_5^{i-1}\right) \tag{4.40}$$

if $Y_1 \le y_i \le 1$. We may thus use the same line of arguments as in Lemma 4.10 to conclude that we may assume that $y_i' \ge Y_1$ for $3 \le i \le 8$ (by verifying (4.41) for $t = 8$ and $3 \le i \le 8$) and then bootstrap that to $y_i' \ge Y_i$ for $2 \le i \le 8$ (by verifying (4.42) for $t = 8$ and $i = 2$); see section 8.1.9 for the details of the verification.

We thus obtain the following upper bound for $\sigma_{8,2}(\mathbf{y}')$ (this is an analogue of (4.43) which is obtained using the fact that h is decreasing):

$$\sigma_{8,2}(\mathbf{y}') \le \frac{c}{\psi - 1}\left(h(Y_0) + \sum_{i=2}^{8} c_5^{i-1} h(Y_1) + 0.15c_5^8 K_2\tau_2\right). \tag{4.48}$$

Now suppose that there exists an index $2 \le i \le 5$ such that $y_i' > Y_1$. We plug (4.48) into (4.40) and obtain the following analogue of (4.44):

$$\left.\frac{\partial \sigma_{8,2}(\mathbf{y})}{\partial y_i}\right|_{\mathbf{y}=\mathbf{y}'} \le \frac{1}{\psi - \prod_{j=1}^{8} y_j'}\left(\frac{c}{\psi - 1}\left(h(Y_0) + c_5 h(Y_1)\frac{1 - c_5^7}{1 - c_5} + 0.15c_5^8 K_2\tau_2\right) - 8cc_5^{i-1}\right) < 0; \tag{4.49}$$

see section 8.1.9 for the details of the verification of the last inequality. Thus, we may assume that $y_i' = Y_1$ for $2 \le i \le 5$.

Then we can bootstrap our bound on $\sigma_{8,2}(\mathbf{y}')$ in (4.48) to

$$\sigma_{8,2}(\mathbf{y}') \le \frac{c}{\psi - y_1' Y_1^4}\left(h(y_1') + \sum_{i=2}^{8} c_5^{i-1} h(Y_1) + 0.15c_5^8 K_2\tau_2\right), \tag{4.50}$$

which gives that

$$\left.\frac{\partial \sigma_{8,2}(\mathbf{y})}{\partial y_6}\right|_{\mathbf{y}=\mathbf{y}'} \le \frac{1}{\psi - \prod_{j=1}^{t} y_j'}\left(\frac{cY_1^4}{\psi - y_1' Y_1^4}\left(h(y_1') + c_5 h(Y_1)\frac{1 - c_5^7}{1 - c_5} + 0.15c_5^8 K_2\tau_2\right) - 8cc_5^5\right) < 0, \tag{4.51}$$

where the last inequality holds for all $y_1' \in [Y_0, 1]$; see section 8.1.9 for the verification using Mathematica's RESOLVE function. This implies that we can set $y_6' = Y_1$ as well.

Using $y_2' = \cdots = y_5' = y_6' = Y_1$, $Y_1 \le y_7', y_8' \le 1$, and the fact that h is decreasing, we obtain that

$$1 < \sigma_{8,2}(\mathbf{y}') \le \frac{c}{\psi - y_1' Y_1^5}\left(h(y_1') + c_5 h(Y_1)\frac{1 - c_5^7}{1 - c_5} + 0.15c_5^8 K_2\tau_2\right), \tag{4.52}$$

which is false for all $y_1' \in [Y_0, 1]$; see section 8.1.9 for the proof using Mathematica's RESOLVE function.

Thus, the assumption that there exists $\mathbf{y}'$ such that $\sigma_{8,2}(\mathbf{y}') > 1$ has led to a contradiction. This completes the proof of the lemma. ∎

5. Bounding the decay rate for $k = 3$, $\Delta = 6$. In this section, we fix $k = 3$ and $\Delta = 6$. This section is devoted to proving Lemma 3.10, which we restate here for convenience.

LEMMA 3.10. *Let $\Delta = 6$ and $k = 3$. There are constants $0 < \delta < 1$ and $U > 0$ such that, for all $1 \le d \le \Delta$, all suitable $\mathbf{w} = w_1, \dots, w_d$, and all $\mathbf{r}$ satisfying $(1/2)^{\Delta-1}\mathbf{1} \le \mathbf{r} \le \mathbf{1}$, it holds that*

$$\kappa_*^{d,\mathbf{w}}(\mathbf{r}) \le \begin{cases} 1 & \text{when } d \le \Delta - 1, \\ U & \text{when } d = \Delta. \end{cases}$$

In the statement of Lemma 3.10, d is an integer between 1 and 6. The vector $\mathbf{w}$ is a suitable vector, which is defined in Definition 3.7. This means that the entries $w_1, \dots, w_d$ are in nondecreasing order, except that any "1" entries are left to the end. The definition of "suitable" also includes some global notation which depends implicitly on $\mathbf{w}$: For all positive integers ℓ, b_ℓ is the number of entries amongst $w_1, \dots, w_d$ which are equal to $\ell - 1$. Hence, $\sum_{\ell=2}^{\infty} b_\ell = d$. Also, $w_1, \dots, w_{b_3}$ are all equal to 2, whereas for $i > b_3$, w_i is either 1 or at least 3. We have $k = 3$, so $b'_3 = b_3$. Finally, $s_i = \sum_{t=1}^{i} \max(0, 2 - w_t)$. Recall the definition of $\kappa_*^{d,\mathbf{w}}(\mathbf{r})$ from (3.5), which we have specialized here to $k = 3$:

$$\kappa_*^{d,\mathbf{w}}(\mathbf{r}) := \sum_{i=1}^{d} \sum_{j=1}^{w_i} \alpha^{-l_{w_i}} \delta^{(b_2 + s_{\min(i,d-b_2)} - \max(0, b_3 - i) - (j-1)(\Delta-1)\mathbf{1}_{i \le d-b_2})} \frac{\varphi(F^{d,\mathbf{w}}(\mathbf{r}))}{\varphi(r_{i,j})} \left| \frac{\partial F^{d,\mathbf{w}}}{\partial r_{i,j}} \right|.$$

Consider the following definitions, which apply to all suitable $\mathbf{w}$.

$$\rho^{\mathbf{w},i}(\mathbf{r}) := \alpha^{-l_{w_i}} \sum_{j=1}^{w_i} \left(\frac{1}{\delta}\right)^{(j-1)(\Delta-1)} \frac{1}{\varphi(r_{i,j})} \left| \frac{\partial F^{d,\mathbf{w}}}{\partial r_{i,j}} \right|, \tag{5.1}$$

$$\kappa^{d,\mathbf{w}}(\mathbf{r}) := \varphi(F(\mathbf{r}))\, \delta^{b_2} \left(\sum_{i=1}^{b_3} \left(\frac{1}{\delta}\right)^{b_3 - i} \rho^{\mathbf{w},i}(\mathbf{r}) + \sum_{i=b_3+1}^{d} \rho^{\mathbf{w},i}(\mathbf{r}) \right). \tag{5.2}$$

We first argue that $\kappa_*^{d,\mathbf{w}}(\mathbf{r}) \le \kappa^{d,\mathbf{w}}(\mathbf{r})$. To see this, note that $s_{\min(i,d-b_2)} \ge 0$. Also, the term $\max(0, b_3 - i)$ is $b_3 - i$ if $1 \le i \le b_3$ and is 0 if $i > b_3$. Finally, $\mathbf{1}_{i \le d-b_2} \le 1$.

DEFINITION 5.1. *Let $\delta = 9789/10000$.*

We can now state a lemma which immediately implies Lemma 3.10 since $\kappa_*^{d,\mathbf{w}}(\mathbf{r}) \le \kappa^{d,\mathbf{w}}(\mathbf{r})$.

LEMMA 5.2. *Let $\Delta = 6$ and $k = 3$. There is constant $U > 0$ such that, for all $1 \le d \le \Delta$, all suitable $\mathbf{w} = w_1, \dots, w_d$, and all $\mathbf{r}$ satisfying $(1/2)^{\Delta-1}\mathbf{1} \le \mathbf{r} \le \mathbf{1}$, it holds that*

$$\kappa^{d,\mathbf{w}}(\mathbf{r}) \le \begin{cases} 1 & \text{when } d \le \Delta - 1, \\ U & \text{when } d = \Delta. \end{cases}$$

The rest of the section contains the proof of Lemma 5.2. This is a more involved optimization problem than the one that arose in the proof of Lemma 3.9. Before delving into the details, we set up some convenient notation and then give a roadmap of the argument.

5.1. Outline of the proof. For convenience, let $F := F^{d,\mathbf{w}}$, $\kappa(\mathbf{r}) := \kappa^{d,\mathbf{w}}(\mathbf{r})$, which is defined in (5.2), and for $i \in [d]$, let $\rho_i(\mathbf{r}) := \rho^{\mathbf{w},i}(\mathbf{r})$, which is defined in (5.1). Our goal is to show that, for all $\mathbf{r}$ such that $\frac{1}{2^{\Delta-1}}\mathbf{1} \le \mathbf{r} \le \mathbf{1}$, it holds that $\kappa(\mathbf{r}) \le 1$ when $d \le \Delta - 1$ and that $\kappa(\mathbf{r})$ is bounded by a constant when $d = \Delta$.

Here is a rough outline of our analysis.

1. The first part of the proof will be to bound the quantities $\rho_i(\mathbf{r})$ appropriately for each $i \in [d]$. Namely, the main goal here will be to replace the $\{r_{i,j}\}_{j\in[w_i]}$ by a suitable quantity $\hat{r}_i$. In fact, for this part of the proof, rather than working with the $r_{i,j}$'s, it will be easier to work with $t_{i,j} = \frac{r_{i,j}}{1+r_{i,j}}$. (See Lemma 5.3.)
2. After the first part, we will have significantly reduced the dimensionality of the optimization problem: from the initial number of $\sum_{i\in[d]} w_i$ variables $\{r_{i,j}\}_{i\in[d],j\in[w_i]}$, we will be left with just d "representative" variables (one for each i).
3. Despite having reduced the number of variables quite a lot, the w_i's so far can be arbitrarily large integers. It will be convenient for us to restrict the range of the w_i's. Using a rather crude argument, we will be able to restrict our attention to i's such that $1 \le w_i \le 5$. Intuitively, the reason is that large w_i's make κ smaller. We quantify this effect in an appropriate way for our analysis. (See Lemma 5.4.)
4. The next step is a further reduction of the number of variables. In particular, recall from item 2 that we have reduced to the case where the number of variables is d (one for each i) and, from item 3, for each $i \in [d]$ it holds that $1 \le w_i \le 5$. We will further group these variables together according to their values. That is, for integer $1 \le w \le 5$ we will be able to use a single variable (indexed by w) to capture the aggregate contribution of the variables w_i with $w_i = w$. (See (5.23) and the inequality in Lemma 5.7 and the relevant Lemmas 5.6 and 5.7.)
5. At this point, we are able to do the final steps of the optimization. The most important case turns out to be when all of the w_i's are either 1 or 2 (Lemma 5.8). In this case, we obtain quite sharp bounds for κ for each value of d (i.e., $d = 1, 2, \ldots, \Delta - 1 = 5$). This facilitates the application of cruder arguments to handle the cases where there exist $i \in [d]$ with $w_i \neq 1, 2$ (Lemmas 5.9, 5.10, and 5.11).

5.2. The details of the argument. In this section, we expand in detail the outline of section 5.1 and give the necessary technical ingredients needed to complete the proof of Lemma 5.2. Later subsections contain the leftover technical proofs which would significantly interrupt the flow.

The first part of the proof will be to bound the quantities $\rho_i(\mathbf{r})$ appropriately. For $i \in [d]$ and $j \in [w_i]$, we have (by differentiating $\ln F(\mathbf{r})$ as in the proof of Lemma 3.9)

$$\frac{\partial F}{\partial r_{i,j}} = -F(\mathbf{r}) \frac{\prod_{j'=1}^{w_i} \frac{r_{i,j'}}{1+r_{i,j'}}}{1 - \prod_{j'=1}^{w_i} \frac{r_{i,j'}}{1+r_{i,j'}}} \frac{1}{r_{i,j}(1+r_{i,j})}.$$

Let $g(z) := \frac{1}{\varphi(z)} \frac{1}{z(1+z)} = \frac{\psi - z^\chi}{1+z}$. For $i \in [d]$, the quantity $\rho_i(\mathbf{r})$ can then be written as

$$\rho_i(\mathbf{r}) := F(\mathbf{r}) \frac{\prod_{j=1}^{w_i} \frac{r_{i,j}}{1+r_{i,j}}}{1 - \prod_{j=1}^{w_i} \frac{r_{i,j}}{1+r_{i,j}}} \alpha^{-l_{w_i}} \sum_{j=1}^{w_i} \left(\frac{1}{\delta}\right)^{(j-1)(\Delta-1)} g(r_{i,j}). \tag{5.3}$$

Let $t_{i,j} := \frac{r_{i,j}}{1+r_{i,j}}$, and let $\hat{t}_i$ be the geometric mean of the $t_{i,j}$'s, i.e., $(\hat{t}_i)^{w_i} := \prod_{j=1}^{w_i} t_{i,j} = \prod_{j=1}^{w_i} \frac{r_{i,j}}{1+r_{i,j}}$. As we shall see soon, $\hat{t}_i$ will be used to capture the "aggregate" effect of w_i. Let $\mathbf{t}$ be the vector whose entries are given by $t_{i,j}$ with $i \in [d]$ and $j \in [w_i]$. Note that $\frac{1}{2^{\Delta-1}+1}\mathbf{1} \le \mathbf{t} \le (1/2)\mathbf{1}$.

We will view the quantities $\kappa(\mathbf{r})$ and $\rho_i(\mathbf{r})$ as a function of $\mathbf{t}$. For that, it will be convenient to consider the function

$$h(t) := g\left(\frac{t}{1-t}\right) = (1-t)\left[\psi - \left(\frac{t}{1-t}\right)^{\chi}\right] \tag{5.4}$$

for $t \in [1/(2^{\Delta-1}+1), 1/2]$. With this preprocessing, for $i \in [d]$, the quantities $\rho_i(\mathbf{r})$, $\kappa(\mathbf{r})$ as a function of $\mathbf{t}$ become

$$\begin{aligned}
\rho_i(\mathbf{t}) &= \overline{F}(\mathbf{t})\frac{(\hat{t}_i)^{w_i}}{1-(\hat{t}_i)^{w_i}}\alpha^{-l_{w_i}}\sum_{j=1}^{w_i}\left(\frac{1}{\delta}\right)^{(j-1)(\Delta-1)}h(t_{i,j}),\\
\bar{\kappa}(\mathbf{t}) &= \varphi\big(\overline{F}(\mathbf{t})\big)\,\delta^{b_2}\left(\sum_{i=1}^{b_3}\left(\frac{1}{\delta}\right)^{b_3-i}\rho_i(\mathbf{t}) + \sum_{i=b_3+1}^{d}\rho_i(\mathbf{t})\right),
\end{aligned} \tag{5.5}$$

where

$$\overline{F}(\mathbf{t}) := \prod_{i=1}^{d}\Big(1-\prod_{j=1}^{w_i}t_{i,j}\Big).$$

After this preliminary step, for $i \in [d]$, we will now pursue the task of substituting the variables $t_{i,j}$ with $j \in [w_i]$ with a single variable $\hat{t}_i$. Let $\hat{\mathbf{t}} = \{\hat{t}_i\}_{i=1,\dots,d}$, and note that $\frac{1}{2^{\Delta-1}+1}\mathbf{1} \le \hat{\mathbf{t}} \le (1/2)\mathbf{1}$. As a starting point, observe that

$$\overline{F}(\mathbf{t}) = \widehat{F}(\hat{\mathbf{t}}), \text{ where } \widehat{F}(\hat{\mathbf{t}}) = \prod_{i=1}^{d}\big(1-(\hat{t}_i)^{w_i}\big). \tag{5.6}$$

Recall that $\Delta = 6$, $\delta = 9789/10000$, $\chi = 1/2$, and $\psi = 13/10$. The following technical lemma, proved in section 5.3, will be crucial in reducing the number of the variables $t_{i,j}$.

LEMMA 5.3. *Define the following constants.*

$$K_\delta^{(1)} = 1, \quad K_\delta^{(2)} = 1069/1000, \quad K_\delta^{(3)} = 1160/1000, \quad K_\delta^{(4)} = 1225/1000, \text{ and}$$
$$K_\delta^{(w)} = \left(\frac{1}{\delta}\right)^{(w-1)(\Delta-1)} \text{ for } w \ge 5.$$

Then for all positive integers w, the following inequality holds for all $t_1, \dots, t_w \in [0, 1/2]$:

$$\sum_{j=1}^{w}\frac{1}{\delta^{(j-1)(\Delta-1)}}h(t_j) \le w\,K_\delta^{(w)}\,h(t),$$

where t is the geometric mean of the t_i's, i.e., $t = (t_1\cdots t_w)^{1/w}$.

Applying Lemma 5.3 to the quantity $\rho_i(\mathbf{t})$ in (5.5) yields that for all $i \in [d]$ it holds that

$$\rho_i(\mathbf{t}) \le \rho_i^{(1)}\big(\hat{\mathbf{t}}\big), \text{ where } \rho_i^{(1)}\big(\hat{\mathbf{t}}\big) = \widehat{F}\big(\hat{\mathbf{t}}\big)w_i\,K_\delta^{(w_i)}\alpha^{-l_{w_i}}\frac{(\hat{t}_i)^{w_i}}{1-(\hat{t}_i)^{w_i}}h(\hat{t}_i), \tag{5.7}$$

where $\widehat{F}(\hat{\mathbf{t}})$ is as given by (5.6).

The next part of the proof will be to bound the contribution of w_i's with $w_i \ge 6$ by small quantities so that we will eliminate those i with $w_i \ge 6$ (and hence the respective variables $\hat{t}_i$) from consideration. This will be accomplished by the following lemma (proved in section 5.4).

LEMMA 5.4. *Let $M = 25/1000$. Recall that $\widehat{F}(\hat{\mathbf{t}})$ is given by (5.6). Let i be such that $w_i \geq 6$. Then, for all $\hat{\mathbf{t}}$ such that $\mathbf{0} \leq \hat{\mathbf{t}} \leq (1/2)\mathbf{1}$, it holds that*

$$\rho_i^{(1)}(\hat{\mathbf{t}}) \leq \frac{1}{\alpha}\widehat{F}(\hat{\mathbf{t}})M.$$

Recall from the beginning of section 5 (based on Definition 3.7) that b_ℓ is the number of entries amongst $w_1, \ldots, w_d$ which are equal to $\ell - 1$. Using Lemma 5.4, we will now be able to eliminate those i such that $w_i \geq 6$. In order to do that easily, we will first reorder the entries in $\mathbf{w}$. Let $B = b_2 + b_3 + b_4 + b_5 + b_6$ and note that $0 \leq B \leq d$. Note that, in the context of (5.5), the ordering of the w_i's with $w_i \neq 2$ does not matter as long as we maintain the invariant that their index i satisfies $i \geq b_3 + 1$. Thus, from now on, without loss of generality, we will assume that $w_i \geq 6$ implies that $i \geq B + 1$. That is, w_i's with $w_i \geq 6$ have indices larger than w_i's with $w_i \leq 5$.

Using (5.5), (5.7), Lemma 5.4, and $l_1 = \cdots = l_5 = 1$, we thus obtain

$$\bar{\kappa}(\mathbf{t}) \leq \frac{1}{\alpha}\kappa^{(1)}(\hat{\mathbf{t}}), \tag{5.8}$$

where

(5.9)
$$\kappa^{(1)}(\hat{\mathbf{t}}) := \varphi\big(\widehat{F}(\hat{\mathbf{t}})\big)\widehat{F}(\hat{\mathbf{t}})\delta^{b_2} \left(2K_\delta^{(2)}\sum_{i=1}^{b_3}\left(\frac{1}{\delta}\right)^{b_3-i}\frac{(\hat{t}_i)^2}{1-(\hat{t}_i)^2}h(\hat{t}_i) + \sum_{i=b_3+1}^{B} w_i\, K_\delta^{(w_i)}\frac{(\hat{t}_i)^{w_i}}{1-(\hat{t}_i)^{w_i}}h(\hat{t}_i) + (d-B)M\right),$$

where $\widehat{F}(\hat{\mathbf{t}})$ is given by (5.6) and $M = 25/1000$ (cf. Lemma 5.4).

To complete the program of eliminating those variables $\hat{t}_i$ where i is such that $w_i \geq 6$, observe that $z\varphi(z) = 1/(\psi - z^\chi)$, so

(5.10)
$$\varphi\big(\widehat{F}(\hat{\mathbf{t}})\big)\widehat{F}(\hat{\mathbf{t}}) = \frac{1}{\psi - \big(\widehat{F}(\hat{\mathbf{t}})\big)^\chi} = \frac{1}{\psi - \prod_{i=1}^{d}\big(1-(\hat{t}_i)^{w_i}\big)^\chi} \leq \frac{1}{\psi - \prod_{i=1}^{B}\big(1-(\hat{t}_i)^{w_i}\big)^\chi}.$$

Using (5.10), we thus obtain that

$$\kappa^{(1)}(\hat{\mathbf{t}}) \leq \kappa^{(2)}(\hat{\mathbf{t}}), \tag{5.11}$$

where

(5.12)
$$\kappa^{(2)}(\hat{\mathbf{t}}) = \delta^{b_2}\frac{2K_\delta^{(2)}\sum_{i=1}^{b_3}\left(\frac{1}{\delta}\right)^{b_3-i}\frac{(\hat{t}_i)^2}{1-(\hat{t}_i)^2}h(\hat{t}_i) + \sum_{i=b_3+1}^{B} w_i\, K_\delta^{(w_i)}\frac{(\hat{t}_i)^{w_i}}{1-(\hat{t}_i)^{w_i}}h(\hat{t}_i) + (d-B)M}{\psi - \prod_{i=1}^{B}\big(1-(\hat{t}_i)^{w_i}\big)^\chi}.$$

The following quantity $\kappa^{(3)}(\hat{\mathbf{t}})$ is similar to $\kappa^{(2)}(\hat{\mathbf{t}})$:

(5.13)
$$\kappa^{(3)}(\hat{\mathbf{t}}) = \delta^{b_2}\frac{2K_\delta^{(2)}\sum_{i=1}^{b_3}\left(\frac{1}{\delta}\right)^{b_3-i}\frac{(\hat{t}_i)^2}{1-(\hat{t}_i)^2}h(\hat{t}_i) + \sum_{i=b_3+1}^{B} w_i\, K_\delta^{(w_i)}\frac{(\hat{t}_i)^{w_i}}{1-(\hat{t}_i)^{w_i}}h(\hat{t}_i)}{\psi - \prod_{i=1}^{B}\big(1-(\hat{t}_i)^{w_i}\big)^\chi}.$$

The only difference between $\kappa^{(2)}(\hat{\mathbf{t}})$ and $\kappa^{(3)}(\hat{\mathbf{t}})$ is that the term $(d-B)M$ is not present in the numerator of the latter. We therefore have

$$\kappa^{(2)}(\hat{\mathbf{t}}) = \kappa^{(3)}(\hat{\mathbf{t}}) + \frac{\delta^{b_2}(d-B)M}{\psi - \prod_{i=1}^{B}\big(1-(\hat{t}_i)^{w_i}\big)^\chi} \leq \kappa^{(3)}(\hat{\mathbf{t}}) + \frac{(d-B)M}{\psi - 1}, \tag{5.14}$$

where the last inequality follows from $b_2 \geq 0$, $\delta \in (0,1]$, and the fact that the $\hat{t}_i$'s are positive.

The following lemma, proved later in this section, will allow us to conclude Lemma 5.2.

LEMMA 5.5. *Let $\Delta = 6$ and B be a nonnegative integer less than or equal to $\Delta - 1 = 5$. Recall that $\alpha = 1 - 10^{-4}$. There exists a constant $\varepsilon_B \leq \alpha$ such that for all nonnegative integers b_2, b_3, b_4, b_5, b_6 with $b_2 + b_3 + b_4 + b_5 + b_6 = B$, it holds that $\kappa^{(3)}(\hat{\mathbf{t}}) \leq \varepsilon_B$.*

In particular, we will show that

(5.15)
$$\varepsilon_0 = 0, \quad \varepsilon_1 = 6/10, \quad \varepsilon_2 = 7/10, \quad \varepsilon_3 = 83/100, \quad \varepsilon_4 = 91/100, \quad \varepsilon_5 = \alpha = 1 - 10^{-4}.$$

Assuming Lemma 5.5 for the moment, we give the proof of Lemma 5.2, which we restate here for convenience.

LEMMA 5.2. *Let $\Delta = 6$ and $k = 3$. There is constant $U > 0$ such that, for all $1 \leq d \leq \Delta$, all suitable $\mathbf{w} = w_1, \ldots, w_d$, and all $\mathbf{r}$ satisfying $(1/2)^{\Delta-1}\mathbf{1} \leq \mathbf{r} \leq \mathbf{1}$, it holds that*

$$\kappa^{d,\mathbf{w}}(\mathbf{r}) \leq \begin{cases} 1 & \text{when } d \leq \Delta - 1, \\ U & \text{when } d = \Delta. \end{cases}$$

Proof of Lemma 5.2. We first derive the bound $\kappa(\mathbf{r}) \leq 1$ when $d \leq \Delta - 1 = 5$. Recall that the quantity $\kappa(\mathbf{r})$ (as given in (5.2)) is equal to the quantity $\bar{\kappa}(\mathbf{t})$ (as given in (5.5)). Also, we have shown that

$$\bar{\kappa}(\mathbf{t}) \leq \frac{1}{\alpha}\kappa^{(1)}\big(\hat{\mathbf{t}}\big), \tag{5.8}$$

where $\kappa^{(1)}\big(\hat{\mathbf{t}}\big)$ is as in (5.9). We have also shown that

$$\kappa^{(1)}\big(\hat{\mathbf{t}}\big) \leq \kappa^{(2)}\big(\hat{\mathbf{t}}\big), \tag{5.11}$$

where $\kappa^{(2)}\big(\hat{\mathbf{t}}\big)$ is as in (5.12). Moreover, we showed that

$$\kappa^{(2)}(\hat{\mathbf{t}}) \leq \kappa^{(3)}(\hat{\mathbf{t}}) + \frac{(d-B)M}{\psi - 1}, \tag{5.14}$$

where $\kappa^{(3)}(\hat{\mathbf{t}})$ is as in (5.13), B is a nonnegative integer less than or equal to $\Delta - 1 = 5$ satisfying $B = b_2 + b_3 + b_4 + b_5 + b_6$, and $M = 25/1000$ is as in Lemma 5.4. Finally, by Lemma 5.5, we have

$$\kappa^{(3)}(\hat{\mathbf{t}}) \leq \varepsilon_B,$$

where the constants ε_B are as in (5.15). Combining all of the above, we obtain that

$$\kappa(\mathbf{r}) \leq \frac{1}{\alpha}\Big(\varepsilon_B + \frac{(d-B)M}{\psi - 1}\Big). \tag{5.16}$$

It is a matter of numerical calculations to check that

$$\frac{1}{\alpha}\Big(\varepsilon_B + \frac{(d-B)M}{\psi - 1}\Big) \leq 1 \tag{5.17}$$

for all $B = 0, 1, \ldots, 5$ and $d = \Delta - 1 = 5$; see section 8.2.1 for the explicit calculations. This completes the proof of the lemma for $d \leq \Delta - 1 = 5$.

We next consider the case $d = \Delta$; i.e., we show that there exists a constant $U > 0$ such that $\kappa(\mathbf{r}) \le U$. This will follow by continuity arguments. More precisely, first note that inequalities (5.8), (5.11), and (5.14) still hold in the case where $d = \Delta$, with the minor modification that in (5.14), the integer B (which, recall, is equal to $b_2 + b_3 + b_4 + b_5 + b_6$) can be as large as (but not bigger than) Δ. Let us fix B to be a nonnegative integer which is at most $\Delta = 6$. Observe that there are finitely many possibilities for the nonnegative integers b_2, b_3, b_4, b_5, b_6 such that $b_2+b_3+b_4+b_5+b_6 = B$. For each such choice, the quantity $\kappa^{(3)}(\hat{\mathbf{t}})$ is a continuous function of the (finitely many) variables $\hat{t}_1, \dots, \hat{t}_B$ and hence is bounded above by an absolute constant when $\mathbf{0} \le \hat{\mathbf{t}} \le (1/2)\mathbf{1}$. It follows that for every nonnegative integer $B \le \Delta = 6$, there exists an absolute constant $U_B > 0$ such that $\kappa^{(3)}(\hat{\mathbf{t}}) \le U_B$. Thus, analogously to (5.16), we obtain the bound

$$\kappa(\mathbf{r}) \le \max_{B=0,\dots,6} \left\{ \frac{1}{\alpha} \Big(U_B + \frac{(\Delta - B)M}{\psi - 1} \Big) \right\} =: U. \tag{5.18}$$

Note that U, as defined in (5.18), is a constant. The desired bound on $\kappa(\mathbf{r})$ when $d = \Delta$ follows.

This concludes the proof of Lemma 5.2. ∎

The remainder of this section will focus on the proof of Lemma 5.5. We begin our considerations by reducing the number of variables. We first need the following transformation. Namely, for $w = 1, 2, \dots$ and all $i \in [B]$ such that $w_i = w$, we set $y_i := 1 - t_i^w$ (note that $y_i \in [1-(1/2)^w, 1-1/(2^{\Delta-1}+1)^w]$). For $w = 1, 2, \dots$, consider the functions $g_w(y)$ defined for $y \in [1 - (1/2)^w, 1 - 1/(2^{\Delta-1} + 1)^w]$.

$$g_w(y) := \frac{(1-y)}{y} h\big((1-y)^{1/w}\big). \tag{5.19}$$

The quantity $\kappa^{(3)}(\hat{\mathbf{t}})$ as a function of $\mathbf{y} = \{y_i\}_{i=1}^B$ and the functions g_w then become

$$\kappa^{(4)}(\mathbf{y}) = \delta^{b_2} \frac{2K_\delta^{(2)} \sum_{i=1}^{b_3} \big(\frac{1}{\delta}\big)^{b_3 - i} g_2(y_i) + \sum_{i=b_3+1}^{B} w_i \, K_\delta^{(w_i)} \, g_{w_i}(y_i)}{\psi - \big(\prod_{i=1}^B y_i\big)^\chi}. \tag{5.20}$$

Let $\hat{y}_w$ be the geometric mean of those y_i's with $w_i = w$ (note that the number of such i's is equal to b_{w+1}). More precisely, for $b_{w+1} > 0$, let

$$\hat{y}_w := \Big(\prod_{i; w_i = w} y_i \Big)^{1/b_{w+1}},$$

and when $b_{w+1} = 0$, let $\hat{y}_w = 1$. Note that

$$\prod_{i=1}^{B} y_i = \prod_{w=1}^{5} (\hat{y}_w)^{b_{w+1}}. \tag{5.21}$$

Let $\widehat{\mathbf{y}} = \{\hat{y}_i\}_{i=1,\dots,5}$. Our goal will be to bound $\kappa^{(4)}(\mathbf{y})$ by a function of $\widehat{\mathbf{y}}$.

LEMMA 5.6. *For $w = 1, 2, \dots, 5$, the function $g_w(e^z)$ is a concave function of z in the interval $\big[\ln\big(1 - (1/2)^w\big), 0\big]$.*

Proof. For $z \in \big[\ln\big(1 - (1/2)^w\big), 0\big]$, let $f(z) = g_w(e^z)$. Our goal is to show that for $w = 1, \dots, 5$, it holds that

$$f''(z) \le 0 \text{ for all } z \in \big[\ln\big(1 - (1/2)^w\big), 0\big]. \tag{5.22}$$

For convenience, we use Mathematica's RESOLVE function; see section 8.2.3 for details. ∎

Lemma 5.6 and Jensen's inequality yield that

(5.23)
$$\sum_{i;w_i=w} g_w(y_i) = \sum_{i;w_i=w} g_w(e^{\ln y_i}) = b_{w+1}\, g_w\big(e^{\frac{1}{b_{w+1}}\sum_{i;w_i=w}\ln y_i}\big) \le b_{w+1}\, g_w\big(\hat{y}_w\big).$$

To bound $\sum_{i=1}^{b_3}\big(\frac{1}{\delta}\big)^{b_3-i} g_2(y_i)$ by a function of $\hat{y}_2$, we will use the following lemma (proved in section 5.5).

LEMMA 5.7. *Let $\Delta = 6$ and b_3 be a nonnegative integer less than or equal to $\Delta - 1 = 5$. There exists a constant $C_\delta^{(b_3)} \ge 0$ so that for $y_1, \ldots, y_{b_3} \in [3/4, 1 - 1/(2^d+1)^2]$ it holds that*

$$\sum_{i=1}^{b_3}\left(\frac{1}{\delta}\right)^{b_3-i} g_2(y_i) \le b_3\, C_\delta^{(b_3)}\, g_2\big(\sqrt[b_3]{y_1 \cdots y_{b_3}}\big).$$

In particular, we will show that the inequality holds with $C_\delta^{(b_3)} = \frac{1}{\delta^{b_3-1}}\big(\frac{1-\delta^{b_3 p}}{b_3(1-\delta^p)}\big)^{1/p}$, where $p = 27/2$. For $\delta = 9789/10000$, we have the following upper bounds on the values of $C_\delta^{(b_3)}$:

$$C_\delta^{(0)} = 0,\ \ C_\delta^{(1)} = 1,\ \ C_\delta^{(2)} = 102/100,\ \ C_\delta^{(3)} = 103/100,\ \ C_\delta^{(4)} = 104/100,\ \ C_\delta^{(5)} = 105/100.$$

Using (5.21), (5.23), and the inequality in Lemma 5.7, we obtain that

$$\kappa^{(4)}(\mathbf{y}) \le \kappa^{(5)}(\widehat{\mathbf{y}}),$$

where

$$\kappa^{(5)}(\widehat{\mathbf{y}}) := \delta^{b_2} \cdot \frac{2b_3 K_\delta^{(2)}\, C_\delta^{(b_3)}\, g_2(\hat{y}_2) + \sum_{w;w\in[5],w\ne 2} w\, b_{w+1} K_\delta^{(w)} g_w(\hat{y}_w)}{\psi - \prod_{w=1}^{5}(\hat{y}_w)^{b_{w+1}\chi}},$$

with the values of $K_\delta^{(w)}$ as in Lemma 5.3 and the values of $C_\delta^{(b_3)}$ as in Lemma 5.7.

We next define the following constants τ_{b_2,b_3} for nonnegative integers b_2, b_3 satisfying $b_2 + b_3 \le \Delta - 1 = 5$ (they are all at most α and we will show that they bound $\kappa^{(5)}(\widehat{\mathbf{y}})$):

(5.24)
$$\tau_{0,0} = 0, \quad \tau_{0,1} = \tau_{1,0} = 42/100,$$
$$\tau_{0,2} = 54/100, \quad \tau_{1,1} = 59/100, \quad \tau_{2,0} = 63/100,$$
$$\tau_{0,3} = 72/100,\ \ \tau_{1,2} = 74/100,\ \ \tau_{2,1} = 76/100,\ \ \tau_{3,0} = 79/100,$$
$$\tau_{0,4} = 864/1000,\ \ \tau_{1,3} = 868/1000,\ \ \tau_{2,2} = 876/1000,\ \ \tau_{3,1} = 886/1000,\ \ \tau_{4,0} = 901/1000,$$
$$\tau_{b_2,b_3} = \alpha \text{ when } b_2 + b_3 = 5.$$

LEMMA 5.8. *Let $b_4 = b_5 = b_6 = 0$. For all nonnegative integers b_2, b_3 such that $b_2 + b_3 \le \Delta - 1 = 5$, it holds that $\kappa^{(5)}(\widehat{\mathbf{y}}) \le \tau_{b_2,b_3}$.*

Proof of Lemma 5.8. For $b_4 = b_5 = b_6 = 0$, the quantity $\kappa^{(5)}(\widehat{\mathbf{y}})$ simplifies to

$$\kappa^{(5)}(\widehat{\mathbf{y}}) := \delta^{b_2} \cdot \frac{2b_3 K_\delta^{(2)}\, C_\delta^{(b_3)}\, g_2(\hat{y}_2) + b_2\, g_1(\hat{y}_1)}{\psi - (\hat{y}_1)^{b_2\chi}(\hat{y}_2)^{b_3\chi}}, \tag{5.25}$$

where we used that $K_\delta^{(1)} = 1$ (note that the values of the variables $\hat{y}_3, \hat{y}_4, \hat{y}_5$ do not affect the value of $\kappa^{(5)}$ when $b_4 = b_5 = b_6 = 0$).

To bound $\kappa^{(5)}(\widehat{\mathbf{y}})$, we need a couple of transformations. The first one is simple: set $v_1 := 1 - \hat{y}_1$ and $v_2 := (1 - \hat{y}_2)^{1/2}$ and note that $v_1, v_2 \in [1/(2^{\Delta-1}+1), 1/2]$. From the definition of the functions g_1, g_2 (cf. (5.19)), we have that

$$\begin{aligned} g_1(\hat{y}_1) &= \frac{1-\hat{y}_1}{\hat{y}_1} h(1-\hat{y}_1) = \frac{v_1}{1-v_1} h(v_1) = v_1\Big(\psi - \Big(\frac{v_1}{1-v_1}\Big)^\chi\Big), \\ g_2(\hat{y}_2) &= \frac{1-\hat{y}_2}{\hat{y}_2} h\big((1-\hat{y}_2)^{1/2}\big) = \frac{v_2^2}{1-v_2^2} h(v_2) = \frac{v_2^2}{1+v_2}\Big(\psi - \Big(\frac{v_2}{1-v_2}\Big)^\chi\Big). \end{aligned} \tag{5.26}$$

It follows that the quantity $\kappa^{(5)}(\widehat{\mathbf{y}})$ as a function of v_1, v_2 becomes

$$\kappa^{(6)}(v_1, v_2) := \delta^{b_2} \frac{2b_3 K_\delta^{(2)} C_\delta^{(b_3)} \frac{v_2^2}{1+v_2}\Big(\psi - \big(\frac{v_2}{1-v_2}\big)^\chi\Big) + b_2 v_1\Big(\psi - \big(\frac{v_1}{1-v_1}\big)^\chi\Big)}{\psi - (1-v_1)^{b_2\chi}(1-v_2^2)^{b_3\chi}}.$$

We will show that for all $v_1, v_2 \in [0, 1/2]$, it holds that $\kappa^{(6)}(v_1, v_2) \le \tau_{b_2,b_3}$. To do this, we will use Mathematica's RESOLVE function. We first need to rationalize the expressions to keep the computations feasible (this can be achieved for $\chi = 1/2$). This brings us to the second (and final) transformation. In particular, set

$$v_1 = \frac{4z_1^2}{(1+z_1^2)^2}, \quad v_2 = \frac{4z_2^2}{1+4z_2^4}$$

for $0 \le z_1 \le \sqrt{2}-1$ and $0 \le z_2 \le \frac{1}{2}(\sqrt{3}-1)$. Under these transformations, for $\chi = 1/2$, we obtain that

$$\Big(\frac{v_1}{1-v_1}\Big)^{1/2} = \frac{2z_1}{1-z_1^2}, \quad (1-v_1)^{1/2} = \frac{1-z_1^2}{1+z_1^2}, \quad \Big(\frac{v_2}{1-v_2}\Big)^{1/2} = \frac{2z_2}{1-2z_2^2}, \quad (1-v_2^2)^{1/2} = \frac{1-4z_2^4}{1+4z_2^4}.$$

The quantity $\kappa^{(6)}(v_1, v_2)$ in terms of z_1, z_2 thus becomes

$$\kappa^{(7)}(z_1, z_2) := \delta^{b_2} \frac{b_2 \frac{4z_1^2}{(1+z_1^2)^2}\big(\psi - \frac{2z_1}{1-z_1^2}\big) + 2b_3 K_\delta^{(2)} C_\delta^{(b_3)} \frac{16z_2^4}{(1+2z_2^2)^2(1+4z_2^4)}\big(\psi - \frac{2z_2}{1-2z_2^2}\big)}{\psi - \Big(\frac{1-z_1^2}{1+z_1^2}\Big)^{b_2}\Big(\frac{1-4z_2^4}{1+4z_2^4}\Big)^{b_3}}.$$

Our goal is to show that, for $b_2, b_3 \ge 0$ satisfying $b_2 + b_3 \le \Delta - 1 = 5$, there do not exist z_1, z_2 in the range $0 \le z_1 \le \sqrt{2}-1$ and $0 \le z_2 \le \frac{1}{2}(\sqrt{3}-1)$ such that $\kappa^{(7)}(z_1, z_2) > \tau_{b_2,b_3}$, where the constants τ are as in (5.24). This can be done symbolically using Mathematica. We give the code in section 8.2.5. ∎

We use Lemma 5.8 to show the following.

LEMMA 5.9. *Let $b_5 = b_6 = 0$ and B be a nonnegative integer less than or equal to $\Delta - 1 = 5$. For all nonnegative integers b_2, b_3, b_4 such that $b_2 + b_3 + b_4 = B$, it holds that $\kappa^{(5)}(\widehat{\mathbf{y}}) \le \tau_{B,0}$, where the constants $\tau_{B,0}$ are as given by* (5.24).

Proof of Lemma 5.9. We may assume that $b_4 \ge 1$ (when $b_4 = 0$, the bounds in the lemma follow immediately from Lemma 5.8). For $b_5 = b_6 = 0$, the quantity $\kappa^{(5)}(\widehat{\mathbf{y}})$ simplifies to

$$\kappa^{(5)}(\widehat{\mathbf{y}}) := \delta^{b_2} \cdot \frac{b_2\, g_1(\hat{y}_1) + 2b_3 K_\delta^{(2)} C_\delta^{(b_3)}\, g_2(\hat{y}_2) + 3b_4 K_\delta^{(3)}\, g_3(\hat{y}_3)}{\psi - (\hat{y}_1)^{b_2\chi}(\hat{y}_2)^{b_3\chi}(\hat{y}_3)^{b_4\chi}},$$

where we used that $K_\delta^{(1)} = 1$ (note that the values of the variables $\hat{y}_4, \hat{y}_5$ do not affect the value of $\kappa^{(4)}$ when $b_5 = b_6 = 0$). The proof splits into two cases depending on whether b_3 is zero.

Case I. $b_3 \geq 1$. Let $A := (\hat{y}_1)^{b_2\chi}(\hat{y}_2)^{b_3\chi}$. Since $\hat{y}_1, \hat{y}_2 \in [0,1]$, we have the crude bound $0 \leq A \leq 1$. By Lemma 5.8 (see also (5.25)), we have that

$$\delta^{b_2} \cdot \frac{b_2\, g_1(\hat{y}_1) + 2b_3 K_\delta^{(2)}\, C_\delta^{(b_3)}\, g_2(\hat{y}_2)}{\psi - A} \leq \tau_{b_2,b_3},$$

where the values of the constants τ_{b_2,b_3} are as in Lemma 5.8 (cf. (5.24)). It follows that

$$\kappa^{(5)}(\widehat{\mathbf{y}}) \leq \kappa^{(6)}(A, \hat{y}_3), \text{ where } \kappa^{(6)}(A, \hat{y}_3) := \frac{\tau_{b_2,b_3}(\psi - A) + 3\delta^{b_2}\, b_4\, K_\delta^{(3)}\, g_3(\hat{y}_3)}{\psi - A(\hat{y}_3)^{b_4\chi}}.$$

We next perform a transformation for the variable $\hat{y}_3$ (similar to the one used in the proof of Lemma 5.8); namely, we set $v_3 := (1 - \hat{y}_3)^{1/3}$ so that $v_3 \in [0, 1/2]$. From the definition of the function g_3 (cf. (5.19)), we have that

$$g_3(\hat{y}_3) = \frac{1 - \hat{y}_3}{\hat{y}_3} h\big((1 - \hat{y}_3)^{1/3}\big) = \frac{v_3^3}{1 - v_3^3} h(v_3) = \frac{v_3^3}{1 + v_3 + v_3^2}\Big(\psi - \big(\frac{v_3}{1 - v_3}\big)^\chi\Big). \tag{5.27}$$

It follows that the quantity $\kappa^{(6)}(A, \hat{y}_3)$ as a function of A, v_3 can now be written as

$$\kappa^{(7)}(A, v_3) := \frac{\tau_{b_2,b_3}(\psi - A) + 3\delta^{b_2}\, b_4\, K_\delta^{(3)}\, \frac{v_3^3}{1+v_3+v_3^2}\Big(\psi - \big(\frac{v_3}{1-v_3}\big)^\chi\Big)}{\psi - A(1 - v_3^3)^{b_4\chi}}.$$

We will show that

$$\kappa^{(7)}(A, v_3) \leq \tau_{B,0} \text{ for all } 0 \leq A \leq 1,\ 0 \leq v_3 \leq 1/2 \text{ (recall } B = b_2 + b_3 + b_4). \tag{5.28}$$

We use Mathematica's RESOLVE function; see section 8.2.6 for details.

Case II. $b_3 = 0$. For $b_3 = b_5 = b_6 = 0$, the quantity $\kappa^{(4)}(\widehat{\mathbf{y}})$ simplifies to

$$\kappa^{(5)}(\widehat{\mathbf{y}}) := \delta^{b_2} \cdot \frac{b_2\, g_1(\hat{y}_1) + 3b_4 K_\delta^{(3)}\, g_3(\hat{y}_3)}{\psi - (\hat{y}_1)^{b_2\chi}(\hat{y}_3)^{b_4\chi}}.$$

We next perform a transformation on the variables $\hat{y}_1, \hat{y}_3$ (similar to the one used in the proof of Lemma 5.8); namely, we set $v_1 = 1 - \hat{y}_1$ and $v_3 := (1 - \hat{y}_3)^{1/3}$ so that $v_1, v_3 \in [0, 1/2]$. Using (5.26) and (5.27), we obtain the following expression for $\kappa^{(5)}$ in terms of v_1, v_3:

$$\kappa^{(8)}(v_1, v_3) := \delta^{b_2} \frac{b_2 v_1\Big(\psi - \big(\frac{v_1}{1-v_1}\big)^\chi\Big) + 3b_4 K_\delta^{(3)}\, \frac{v_3^3}{1+v_3+v_3^2}\Big(\psi - \big(\frac{v_3}{1-v_3}\big)^\chi\Big)}{\psi - (1 - v_1)^{b_2\chi}(1 - v_3^3)^{b_4\chi}}. \tag{5.29}$$

This quantity is still too complicated for Mathematica to resolve efficiently, so we will need one more transformation. In particular, let u_1, u_3 be positive reals defined by

$$v_1 = \frac{u_1^2}{1 + u_1^2}, \qquad v_3 = \frac{u_3^2}{1 + u_3^2},$$

and note that $0 \le u_1, u_3 \le 1$. The quantity $\kappa^{(8)}$ in terms of u_1, u_3 becomes

$$\kappa^{(9)}(u_1,u_3) := \delta^{b_2}\frac{b_2\frac{u_1^2}{1+u_1^2}(\psi-u_1)+3b_4K_\delta^{(3)}\,\frac{u_3^6}{3u_3^6+6u_3^4+4u_3^2+1}(\psi-u_3)}{\psi-\left(\frac{1}{1+u_1^2}\right)^{b_2\chi}\left(1-\left(\frac{u_3^2}{1+u_3^2}\right)^3\right)^{b_4\chi}}.$$

We will show that

$$\kappa^{(9)}(u_1,u_3) \le \tau_{B,0} \text{ for all } 0 \le u_1,u_3 \le 1 \text{ (note that } B = b_2+b_4). \tag{5.30}$$

We use Mathematica's RESOLVE function; see section 8.2.6 for details.

This completes the case analysis and therefore the proof of Lemma 5.9. ∎

LEMMA 5.10. *Let $b_6 = 0$ and B be a nonnegative integer less than or equal to $\Delta - 1 = 5$. For all nonnegative integers b_2, b_3, b_4, b_5 such that $b_2+b_3+b_4+b_5 = B$, it holds that $\kappa^{(5)}(\widehat{\mathbf{y}}) \le \tau_{B,0}$, where the constants $\tau_{B,0}$ are as given by* (5.24).

Proof of Lemma 5.10. We may assume that $b_5 \ge 1$ (when $b_5 = 0$, the bounds in the lemma follow immediately from Lemma 5.9). For $b_6 = 0$ (note that the value of the variable $\hat{y}_5$ does not affect the value of $\kappa^{(5)}$), the quantity $\kappa^{(5)}(\widehat{\mathbf{y}})$ becomes

$$\kappa^{(5)}(\widehat{\mathbf{y}}) := \delta^{b_2}\cdot\frac{b_2\,g_1(\hat{y}_1)+2b_3K_\delta^{(2)}\,C_\delta^{(b_3)}\,g_2(\hat{y}_2)+3b_4K_\delta^{(3)}\,g_3(\hat{y}_3)+4b_5K_\delta^{(4)}\,g_4(\hat{y}_4)}{\psi-(\hat{y}_1)^{b_2\chi}(\hat{y}_2)^{b_3\chi}(\hat{y}_3)^{b_4\chi}(\hat{y}_4)^{b_5\chi}}.$$

Let $A := (\hat{y}_1)^{b_2\chi}(\hat{y}_2)^{b_3\chi}(\hat{y}_3)^{b_4\chi}$. Since $\hat{y}_1, \hat{y}_2, \hat{y}_3 \in [0,1]$, we have the crude bound $0 \le A \le 1$. By Lemma 5.9, we have that

$$\delta^{b_2}\cdot\frac{b_2\,g_1(\hat{y}_1)+2b_3K_\delta^{(2)}\,C_\delta^{(b_3)}\,g_2(\hat{y}_2)+3b_4K_\delta^{(3)}\,g_3(\hat{y}_3)}{\psi-A} \le \tau_{B',0}, \text{ where } B' = b_2+b_3+b_4,$$

where the values of the constants $\tau_{B',0}$ are given by (5.24). Using that $\delta^{b_2} \le 1$, it follows that

$$\kappa^{(5)}(\widehat{\mathbf{y}}) \le \kappa^{(6)}(A,\hat{y}_4), \text{ where } \kappa^{(6)}(A,\hat{y}_4) := \frac{\tau_{B',0}(\psi-A)+4\,b_5\,K_\delta^{(4)}\,g_4(\hat{y}_4)}{\psi-A(\hat{y}_4)^{b_5\chi}}.$$

We next perform a transformation on the variable $\hat{y}_4$ (similar to the one used in the proof of Lemma 5.8); namely, we set $v_4 := (1-\hat{y}_4)^{1/4}$ so that $v_4 \in [0,1/2]$. From the definition of the function g_4 (cf. (5.19)), we have that

$$g_4(\hat{y}_4) = \frac{1-\hat{y}_4}{\hat{y}_4}h\big((1-\hat{y}_4)^{1/4}\big) = \frac{v_4^4}{1-v_4^4}h(v_4) = \frac{v_4^4}{1+v_4+v_4^2+v_4^3}\Big(\psi-\Big(\frac{v_4}{1-v_4}\Big)^\chi\Big).$$

It follows that the quantity $\kappa^{(6)}(A,\hat{y}_4)$ as a function of A, v_4 can now be written as

$$\kappa^{(7)}(A,v_4) := \frac{\tau_{B',0}(\psi-A)+4b_5\,K_\delta^{(4)}\,\frac{v_4^4}{1+v_4+v_4^2+v_4^3}\Big(\psi-\big(\frac{v_4}{1-v_4}\big)^\chi\Big)}{\psi-A(1-v_4^4)^{b_5\chi}}.$$

We will show that

$$\kappa^{(7)}(A,v_4) \le \tau_{B,0} \text{ for all } 0 \le A \le 1,\ 0 \le v_4 \le 1/2 \text{ (recall } B = b_2+b_3+b_4+b_5 = B'+b_5). \tag{5.31}$$

We use Mathematica's RESOLVE function; see section 8.2.7 for details. ∎

LEMMA 5.11. *Let B be a nonnegative integer less than or equal to $\Delta-1=5$. For all nonnegative integers b_2,b_3,b_4,b_5,b_6 such that $b_2+b_3+b_4+b_5+b_6=B$, it holds that $\kappa^{(5)}(\widehat{\mathbf{y}})\le\tau_{B,0}$, where the constants $\tau_{B,0}$ are given by* (5.24).

Proof of Lemma 5.11. We may assume that $b_6\ge 1$ (when $b_6=0$, the bounds in the lemma follow immediately from Lemma 5.10). Recall that the quantity $\kappa^{(5)}(\widehat{\mathbf{y}})$ is given by

$$\kappa^{(5)}(\widehat{\mathbf{y}})=\delta^{b_2}\cdot\frac{b_2\,g_1(\hat y_1)+2b_3K_\delta^{(2)}\,C_\delta^{(b_3)}\,g_2(\hat y_2)+3b_4K_\delta^{(3)}\,g_3(\hat y_3)+4b_5K_\delta^{(4)}\,g_4(\hat y_4)+5b_6K_\delta^{(5)}\,g_5(\hat y_5)}{\psi-(\hat y_1)^{b_2\chi}(\hat y_2)^{b_3\chi}(\hat y_3)^{b_4\chi}(\hat y_4)^{b_5\chi}(\hat y_5)^{b_6\chi}}.$$

Let $A:=(\hat y_1)^{b_2\chi}(\hat y_2)^{b_3\chi}(\hat y_3)^{b_4\chi}(\hat y_4)^{b_5\chi}$. Since $\hat y_1,\hat y_2,\hat y_3,\hat y_4\in[0,1]$, we have the crude bound $0\le A\le 1$. By Lemma 5.10, we have that

$$\delta^{b_2}\cdot\frac{b_2\,g_1(\hat y_1)+2b_3K_\delta^{(2)}\,C_\delta^{(b_3)}\,g_2(\hat y_2)+3b_4K_\delta^{(3)}\,g_3(\hat y_3)+4b_5K_\delta^{(4)}\,g_4(\hat y_4)}{\psi-A}\le\tau_{B',0},$$

where $B'=b_2+b_3+b_4+b_5$ and the values of the constants $\tau_{B',0}$ are given by (5.24). Using that $\delta^{b_2}\le 1$, it follows that

$$\kappa^{(5)}(\widehat{\mathbf{y}})\le\kappa^{(6)}(A,\hat y_5),\ \text{where}\ \kappa^{(6)}(A,\hat y_5):=\frac{\tau_{B',0}(\psi-A)+5\,b_6\,K_\delta^{(5)}\,g_5(\hat y_5)}{\psi-A(\hat y_5)^{b_6\chi}}.$$

We next perform a transformation on the variable $\hat y_5$ (similar to the one used in the proof of Lemma 5.8); namely, we set $v_5:=(1-\hat y_5)^{1/5}$ so that $v_5\in[0,1/2]$. From the definition of the function g_5 (cf. (5.19)), we have that

$$g_5(\hat y_5)=\frac{1-\hat y_5}{\hat y_5}h\big((1-\hat y_5)^{1/5}\big)=\frac{v_5^5}{1-v_5^5}h(v_5)=\frac{v_5^5}{1+v_5+v_5^2+v_5^3+v_5^4}\Big(\psi-\Big(\frac{v_5}{1-v_5}\Big)^\chi\Big).$$

It follows that the quantity $\kappa^{(6)}(A,\hat y_5)$ as a function of A,v_5 can now be written as

$$\kappa^{(7)}(A,v_5):=\frac{\tau_{B',0}(\psi-A)+5b_6\,K_\delta^{(5)}\,\frac{v_5^5}{1+v_5+v_5^2+v_5^3+v_5^4}\Big(\psi-\big(\frac{v_5}{1-v_5}\big)^\chi\Big)}{\psi-A(1-v_5^5)^{b_6\chi}}.$$

We will show that

$$\kappa^{(7)}(A,v_5)\le\tau_{B,0}\ \text{for all}\ 0\le A\le 1,\ 0\le v_5\le 1/2\ (\text{recall } B=B'+b_6).\tag{5.32}$$

We use Mathematica's RESOLVE function; see section 8.2.8 for details. ∎

The proof of Lemma 5.5, which was important in proving Lemma 5.2, is now immediate.

LEMMA 5.5. *Let $\Delta=6$ and B be a nonnegative integer less than or equal to $\Delta-1=5$. Recall that $\alpha=1-10^{-4}$. There exists a constant $\varepsilon_B\le\alpha$ such that for all nonnegative integers b_2,b_3,b_4,b_5,b_6 with $b_2+b_3+b_4+b_5+b_6=B$, it holds that $\kappa^{(3)}(\hat{\mathbf{t}})\le\varepsilon_B$.*

In particular, we will show that

$$\varepsilon_0=0,\quad \varepsilon_1=6/10,\quad \varepsilon_2=7/10,\quad \varepsilon_3=83/100,\quad \varepsilon_4=91/100,\quad \varepsilon_5=\alpha=1-10^{-4}.\tag{5.15}$$

Proof of Lemma 5.5. By the definition of $\widehat{\mathbf{y}}$, we have that $\kappa^{(3)}(\hat{\mathbf{t}})=\kappa^{(5)}(\widehat{\mathbf{y}})$. Now, just use Lemma 5.11 and observe that the ε_B in (5.15) and the constants $\tau_{B,0}$ in (5.24) satisfy $\tau_{B,0}\le\varepsilon_B$. ∎

5.3. Simplifying the optimization using geometric means. In this section, we prove Lemma 5.3, which we restate here for convenience. Roughly, the lemma bounds the contribution of a w_i with $w_i = w$ to κ. The main accomplishment here is the significant reduction of the number of variables; initially the contribution is a function of w variables $t_1, \ldots, t_w$. The lemma shows that we can reduce the number of variables to 1 by considering the geometric mean of the t_j's. The challenge here is to deal with the asymmetry caused by the δ terms without introducing too much slack into the argument, especially for small values of w (say, $w \le 4$).

LEMMA 5.3. *Define the following constants:*

$$K_\delta^{(1)} = 1, \quad K_\delta^{(2)} = 1069/1000, \quad K_\delta^{(3)} = 1160/1000, \quad K_\delta^{(4)} = 1225/1000, \text{ and}$$
$$K_\delta^{(w)} = \left(\frac{1}{\delta}\right)^{(w-1)(\Delta-1)} \text{ for } w \ge 5.$$

Then for all positive integers w, the following inequality holds for all $t_1, \ldots, t_w \in [0, 1/2]$:

$$\sum_{j=1}^{w} \frac{1}{\delta^{(j-1)(\Delta-1)}} h(t_j) \le w\, K_\delta^{(w)}\, h(t),$$

where t is the geometric mean of the t_i's, i.e., $t = (t_1 \cdots t_w)^{1/w}$.

Recall that the function $h(t)$ is given by

$$h(t) = (1-t)\left[\psi - \left(\frac{t}{1-t}\right)^{\chi}\right] \text{ for } t \in [0, 1/2], \tag{5.4}$$

where $\chi = 1/2$, $\psi = 13/10$. We begin with the following lemma.

LEMMA 5.12. *The function $h(e^y)$ is a concave function of y in the interval $(-\infty, \ln(1/2)]$.*

Proof of Lemma 5.12. Let $f(y) := h(e^y)$ for $y \in (-\infty, \ln(1/2)]$. We will show that

$$f''(y) \le 0 \text{ for all } y \in (-\infty, \ln(1/2)]. \tag{5.33}$$

For convenience, we use Mathematica's RESOLVE function; see section 8.2.9 for the code. ∎

As an immediate consequence of Lemma 5.12 and Jensen's inequality, we obtain the following inequality for $w = 2, 3, \ldots$, for all $t_1, \ldots, t_w \in [0, 1/2]$:

$$\sum_{j=1}^{w} h(t_j) = \sum_{j=1}^{w} h(e^{\ln t_j}) \le w\, h\big(e^{\frac{1}{w}\sum_{j=1}^{w} \ln t_j}\big) = w\, h(t), \tag{5.34}$$

where t is the geometric mean of the t_i's, i.e., $t = (t_1 \cdots t_w)^{1/w}$. Using that $\delta \in (0, 1]$ and inequality (5.34), it follows that

$$\sum_{j=1}^{w} \frac{1}{\delta^{(j-1)(\Delta-1)}} h(t_j) \le \frac{1}{\delta^{(w-1)(\Delta-1)}} \sum_{j=1}^{w} h(t_j) \le w\, \frac{1}{\delta^{(w-1)(\Delta-1)}}\, h\big(t\big).$$

This proves the bounds on $K_\delta^{(w)}$ stated in the lemma for all integers $w \ge 5$.

For the bounds on $K_\delta^{(w)}$ stated in the lemma for $w = 2, 3, 4$ we will have to work harder. Our goal is to prove the following inequalities for $t_i \in [0, 1/2]$ $(i = 1, 2, \ldots)$:

$$h(t_1) + \frac{1}{\delta^5}\, h(t_2) \le 2K_\delta^{(2)} h(\sqrt{t_1 t_2}), \tag{5.35}$$

$$h(t_1) + \frac{1}{\delta^5}\, h(t_2) + \frac{1}{\delta^{10}}\, h(t_3) \le 3K_\delta^{(3)} h(\sqrt[3]{t_1 t_2 t_3}), \tag{5.36}$$

$$h(t_1) + \frac{1}{\delta^5}\, h(t_2) + \frac{1}{\delta^{10}}\, h(t_3) + \frac{1}{\delta^{15}}\, h(t_4) \le 4K_\delta^{(4)} h(\sqrt[4]{t_1 t_2 t_3 t_4}). \tag{5.37}$$

To prove these, we will need the following inequalities.

LEMMA 5.13. *Let $A_1, A_2 > 0$ be real numbers. There exists $A > 0$ such that for $t_1, t_2 \in [0, 1/2]$, it holds that*

$$A_1 h(t_1) + A_2 h(t_2) \le A\, h(\sqrt{t_1 t_2}). \tag{5.38}$$

In particular, inequality (5.38) *holds for the following values of A_1, A_2, A:*

$$A_1 = 1, \qquad A_2 = \frac{1}{\delta^5}, \qquad A = 2K_\delta^{(2)}, \tag{5.39}$$

$$A_1 = 2, \qquad A_2 = \frac{2}{\delta^5} K_\delta^{(2)}, \qquad A = 4 \cdot \frac{1120}{1000}, \tag{5.40}$$

$$A_1 = 1, \qquad A_2 = \frac{1}{\delta^{15}}, \qquad A = \frac{5}{2}, \tag{5.41}$$

$$A_1 = \frac{2}{\delta^5} K_\delta^{(2)}, \qquad A_2 = \frac{5}{2}, \qquad A = 4 \cdot K_\delta^{(4)}, \tag{5.42}$$

where $K_\delta^{(2)} = 1069/1000$ and $K_\delta^{(4)} = 1225/1000$ are as in Lemma 5.3.

Proof. The existence of such an A follows by standard continuity and compactness arguments. The positivity of A is also easy to prove. We thus focus on the more intricate task of verifying (5.38) for the values of A_1, A_2, A given in the statement of the lemma.

We will use Mathematica's RESOLVE function. To do this, we first need to rationalize the expressions which can be achieved for $\chi = 1/2$. In particular, we will use the transformations

$$t_1 = \frac{4x_1^2}{(1 + x_1^2)^2}, \quad t_2 = \frac{4x_2^2}{(1 + x_2^2)^2} \tag{5.43}$$

for $0 \le x_1, x_2 \le \sqrt{2} - 1$. Under these transformations, for $\chi = 1/2$, we obtain that

$$\Big(\frac{t_1}{1 - t_1}\Big)^\chi = \frac{2x_1}{1 - x_1^2}, \quad \Big(\frac{t_2}{1 - t_2}\Big)^\chi = \frac{2x_2}{1 - x_2^2}, \quad \sqrt{t_1 t_2} = \frac{4x_1 x_2}{(1 + x_1^2)(1 + x_2^2)}. \tag{5.44}$$

We are quite close to rationalizing the desired inequality; we only have to address the rationalization of $\big(\frac{\sqrt{t_1 t_2}}{1 - \sqrt{t_1 t_2}}\big)^\chi$. Unfortunately, we will have to explicitly eradicate the radical for this expression.

In particular, inequality (5.38) is equivalent to

$$\frac{A_1(1 - t_1)\Big[\psi - \Big(\frac{t_1}{1 - t_1}\Big)^\chi\Big] + A_2(1 - t_2)\Big[\psi - \Big(\frac{t_2}{1 - t_2}\Big)^\chi\Big]}{A(1 - \sqrt{t_1 t_2})} \le \psi - \Big(\frac{\sqrt{t_1 t_2}}{1 - \sqrt{t_1 t_2}}\Big)^\chi. \tag{5.45}$$

Inequality (5.45) will follow from the following inequalities:

$$\frac{A_1(1-t_1)\left[\psi-\left(\frac{t_1}{1-t_1}\right)^{\chi}\right]+A_2(1-t_2)\left[\psi-\left(\frac{t_2}{1-t_2}\right)^{\chi}\right]}{A(1-\sqrt{t_1t_2})}\le\psi \tag{5.46}$$

and

$$\frac{\sqrt{t_1t_2}}{1-\sqrt{t_1t_2}}\le\left(\psi-\frac{A_1(1-t_1)\left[\psi-\left(\frac{t_1}{1-t_1}\right)^{\chi}\right]+A_2(1-t_2)\left[\psi-\left(\frac{t_2}{1-t_2}\right)^{\chi}\right]}{A(1-\sqrt{t_1t_2})}\right)^2. \tag{5.47}$$

Note that (5.46) allows us to take square roots in (5.47), and thus (5.45) follows.

It remains to prove (5.46) and (5.47). Using the substitutions (5.43) and (5.44), inequalities (5.46) and (5.47) are equivalent to

$$0\le\psi-\frac{A_1\frac{(1-x_1^2)^2}{(1+x_1^2)^2}\left(\psi-\frac{2x_1}{1-x_1^2}\right)+A_2\frac{(1-x_2^2)^2}{(1+x_2^2)^2}\left(\psi-\frac{2x_2}{1-x_2^2}\right)}{A\left(1-\frac{4x_1x_2}{(1+x_1^2)(1+x_2^2)}\right)} \tag{5.48}$$

and

$$\frac{4x_1x_2}{(1+x_1^2)(1+x_2^2)-4x_1x_2}\le\left(\psi-\frac{A_1\frac{(1-x_1^2)^2}{(1+x_1^2)^2}\left(\psi-\frac{2x_1}{1-x_1^2}\right)+A_2\frac{(1-x_2^2)^2}{(1+x_2^2)^2}\left(\psi-\frac{2x_2}{1-x_2^2}\right)}{A\left(1-\frac{4x_1x_2}{(1+x_1^2)(1+x_2^2)}\right)}\right)^2, \tag{5.49}$$

respectively. The last inequalities involve rational expressions and can be resolved using Mathematica for the values of A_1, A_2, A given in the statement of the lemma; the code can be found in section 8.2.10. ∎

We now return to the task of proving the inequalities (5.35), (5.36), and (5.37). Inequality (5.35) is an immediate consequence of Lemma 5.13 (cf. the values in (5.39)).

To prove (5.36), we will use the following inequality:

$$h(t_1)+\frac{1}{\delta^5}h(t_2)+\frac{1}{\delta^{10}}h(t_3)+h(t_4)\le 4\cdot\frac{1120}{1000}h(\sqrt[4]{t_1t_2t_3t_4}). \tag{5.50}$$

Applying (5.50) with $t_4=\sqrt[3]{t_1t_2t_3}$ (note that with this value of t_4 it holds that $\sqrt[4]{t_1t_2t_3t_4}=\sqrt[3]{t_1t_2t_3}$) yields

$$h(t_1)+\frac{1}{\delta^5}h(t_2)+\frac{1}{\delta^{10}}h(t_3)\le 3\left(\frac{4\cdot\frac{1120}{1000}-1}{3}\right)h(\sqrt[3]{t_1t_2t_3})=3K_\delta^{(3)}h(\sqrt[3]{t_1t_2t_3}),$$

which proves (5.36). It remains to prove (5.50), which follows by adding the following inequalities:

$$h(t_1)+h(t_4)\le 2h(\sqrt{t_1t_4}), \tag{5.51}$$

$$\frac{1}{\delta^5}h(t_2)+\frac{1}{\delta^{10}}h(t_3)\le\frac{2}{\delta^5}K_\delta^{(2)}h(\sqrt{t_2t_3}), \tag{5.52}$$

$$2h(\sqrt{t_1t_4})+\frac{2}{\delta^5}K_\delta^{(2)}h(\sqrt{t_2t_3})\le 4\cdot\frac{1120}{1000}h(\sqrt[4]{t_1t_2t_3t_4}). \tag{5.53}$$

Inequality (5.51) is an immediate consequence of Lemma 5.12. Inequality (5.52) is an immediate consequence of inequality (5.35) (multiplied by $1/\delta^5$). For inequality

(5.53), we use the transformations $u_1 = \sqrt{t_1 t_4}$ and $u_2 = \sqrt{t_2 t_3}$, so that we need to show

$$2h(u_1) + \frac{2}{\delta^5} K_\delta^{(2)} h(u_2) \le 4 \cdot \frac{1120}{1000} h(\sqrt{u_1 u_2}) \tag{5.54}$$

for $u_1, u_2 \in [0, 1/2]$, which follows from Lemma 5.13 (cf. the values (5.40)).

Finally, we conclude with the proof of inequality (5.37). This is obtained by adding the following three inequalities:

$$\frac{1}{\delta^5}\, h(t_2) + \frac{1}{\delta^{10}}\, h(t_3) \le \frac{2}{\delta^5} K_\delta^{(2)} h(\sqrt{t_2 t_3}), \tag{5.55}$$

$$h(t_1) + \frac{1}{\delta^{15}} h(t_4) \le \frac{5}{2} h(\sqrt{t_1 t_4}), \tag{5.56}$$

$$\frac{2}{\delta^5}\, K_\delta^{(2)} h(\sqrt{t_2 t_3}) + \frac{5}{2} h(\sqrt{t_1 t_4}) \le 4 K_\delta^{(4)} h(\sqrt[4]{t_1 t_2 t_3 t_4}). \tag{5.57}$$

Inequality (5.55) is an immediate consequence of (5.35) (again, multiplied by $1/\delta^5$). Inequality (5.56) follows from Lemma 5.13 (cf. the values (5.41)). Finally, inequality (5.57) can be proved using a transformation analogous to the one used to prove (5.53); the required analogue of inequality (5.54) has been proved in Lemma 5.13 (cf. the values (5.42)).

This concludes the proof of Lemma 5.3.

5.4. Eliminating large arity clauses from consideration. In this section, we prove Lemma 5.4, which we restate here for convenience. Recall that in the construction of the optimization problem from the original correlation-decay argument, w_i is the arity of the ith clause containing x minus one. Intuitively, clauses with large arity should not significantly affect the correlation decay. The following lemma captures this in a quantitative way which is sufficient for our needs (for clauses with $w_i \ge 6$).

LEMMA 5.4. *Let $M = 25/1000$. Recall that $\widehat{F}(\hat{\mathbf{t}})$ is given by* (5.6). *Let i be such that $w_i \ge 6$. Then, for all $\hat{\mathbf{t}}$ such that $\mathbf{0} \le \hat{\mathbf{t}} \le (1/2)\mathbf{1}$, it holds that*

$$\rho_i^{(1)}(\hat{\mathbf{t}}) \le \frac{1}{\alpha} \widehat{F}(\hat{\mathbf{t}}) M.$$

Proof. We will show that for all integers $w \ge 6$ and all $t \in [0, 1/2]$, it holds that

$$\frac{t^w}{1 - t^w} \le \frac{63}{2^w - 1} \frac{t^6}{1 - t^6}. \tag{5.58}$$

We will also show that for $\chi = 1/2, \psi = 13/10$, it holds that

$$\max_{t \in [0,1/2]} \frac{t^6}{1 - t^6} h(t) \le M_1, \text{ where } M_1 = \frac{1}{410}. \tag{5.59}$$

Finally, we will show that for integer $w \ge 6$, $l_w = \lceil \log_6(w+1) \rceil$, $\delta = 9789/10000$, it holds that

$$\frac{63 M_1}{2^w - 1} w\, K_\delta^{(w)}\, \alpha^{-l_w} \le \frac{M}{\alpha}, \tag{5.60}$$

where $K_\delta^{(w)}$ for $w \ge 6$ is as given by Lemma 5.3 and $M = 25/1000$ is as in the statement of the lemma. The lemma then follows by multiplying (5.58), (5.59), and (5.60).

We start with the verification of (5.58). Note that (5.58) holds at equality for $t = 1/2$, so it suffices to show that for all integers $w \geq 6$, the function

$$f(t) := \frac{t^w(1-t^6)}{t^6(1-t^w)} = \frac{t^{w-6}(1-t^6)}{1-t^w}$$

is increasing for $t \in [0, 1/2]$. For $w = 6$, there is nothing to show, so we may assume that $w \geq 7$. We then calculate (see section 8.2.2 for the calculation)

$$f'(t) = \frac{t^{w-7}\, p(t)}{(t^w-1)^2}, \text{ where } p(t) := 6t^w - t^6 w + w - 6, \tag{5.61}$$

so we only need to show that $p(t) \geq 0$ for $t \in [0, 1/2]$. Note that $p'(t) = 6w(t^{w-1} - t^5) \leq 0$ for all $t \in [0, 1/2]$ since $w \geq 7$. It follows that

$$p(t) \geq p(1/2) = 6(1/2)^w + (63/64)w - 6 \geq (63/64)w - 6 \geq 1/2,$$

where in the last inequality we again used that $w \geq 7$. This completes the verification of (5.58).

We next verify (5.59). For convenience, we use Mathematica's RESOLVE function; see section 8.2.2.

Finally, we verify (5.60). Since $l_w \leq \log_6(w+1) + 1$ and $\alpha = 1 - 10^{-4} < 1$, it suffices to show that for all $w \geq 6$ it holds that

$$\frac{63M_1}{2^w - 1} w \left(\frac{1}{\delta}\right)^{(w-1)(\Delta-1)} \alpha^{-\log_6(w+1)} \leq M. \tag{5.62}$$

It is a matter of numerical calculations to show that $\alpha^{-1} \leq \exp(11 \cdot 10^{-5})$. Thus, to show (5.62), it suffices to show that

$$\frac{63M_1}{2^w - 1} w \left(\frac{1}{\delta}\right)^{(w-1)(\Delta-1)} (w+1)^{11\cdot 10^{-5}/\ln 6} \leq M. \tag{5.63}$$

We view the left-hand side in (5.63) as a function of w, say $f(w)$. We will prove that

$$f(6) \leq M \text{ and } f(w+1)/f(w) \leq 1 \text{ for all } w \geq 6, \tag{5.64}$$

from which inequality (5.63) follows. The first inequality in (5.64) follows by a numerical calculation; see section 8.2.2 for details. For the second inequality in (5.64), we have

$$\frac{f(w+1)}{f(w)} = \left(\frac{1}{\delta}\right)^{\Delta-1} \frac{2^w - 1}{2^{w+1} - 1} (1 + w^{-1})\left(1 + (w+1)^{-1}\right)^{11\cdot 10^{-5}/\ln 6}.$$

Note that $\frac{2^w-1}{2^{w+1}-1} \leq 1/2$, and hence, for $w \geq 6$, we have the bound

$$\frac{f(w+1)}{f(w)} \leq \frac{1}{2}\left(\frac{1}{\delta}\right)^{\Delta-1} (1 + 6^{-1})(1 + 7^{-1})^{11\cdot 10^{-5}/\ln 6} < 1, \tag{5.65}$$

where the last inequality follows by a numerical calculation; see section 8.2.2 for details. This concludes the proof of (5.64) and thus the proof of Lemma 5.4. ∎

5.5. The contribution of arity-3 clauses. In this section, we give the proof of Lemma 5.7. Roughly, the lemma bounds the aggregate contribution of arity-3 clauses along with the effect of the creation of arity-2 clauses (due to the pinnings when processing arity-3 clauses). This was used to further reduce the number of variables.

To prove Lemma 5.7, we will use Hölder's inequality. Let $p > 1$. For $q = p/(p-1)$ and positive real numbers α_i, β_i for $i \in [b_3]$, it holds that

$$\sum_{i=1}^{b_3} \alpha_i \beta_i \le \Big(\sum_{i=1}^{b_3} \alpha_i^p \Big)^{1/p} \Big(\sum_{i=1}^{b_3} \beta_i^q \Big)^{1/q}.$$

This yields

$$\sum_{i=1}^{b_3} \Big(\frac{1}{\delta}\Big)^{b_3-i} g_2(y_i) \le \Big(\sum_{i=1}^{b_3} \Big(\frac{1}{\delta}\Big)^{(b_3-i)p} \Big)^{1/p} \Big(\sum_{i=1}^{b_3} (g_2(y_i))^q \Big)^{1/q}. \tag{5.66}$$

For $p = 27/2$, we have $q = 27/25$. Recall from (5.19) that the function $g_2(y)$ is given by

$$g_2(y) := \frac{(1-y)}{y} h\big((1-y)^{1/2}\big) = \frac{(1-y)\big(1-(1-y)^{1/2}\big)}{y} \Big(\psi - \Big(\frac{(1-y)+(1-y)^{1/2}}{y} \Big)^{\chi} \Big). \tag{5.19}$$

LEMMA 5.14. *For $q = 27/25$ and $\Delta = 6$, the function $(g_2(e^t)))^q$ is a concave function of t in the interval $[\ln(3/4), \ln(1 - \frac{1}{(2^{\Delta-1}+1)^2})]$.*

Proof of Lemma 5.14. Let $\bar{g}_2(t) := (g_2(e^t))^q$. Our goal is to show that $\bar{g}_2''(t) \le 0$ for all $t \in [\ln(3/4), \ln(1 - \frac{1}{(2^{\Delta-1}+1)^2})]$. We have

$$\bar{g}_2'(t) = q \left(g_2(e^t)\right)^{q-1} g_2'(e^t) e^t,$$
$$\bar{g}_2''(t) = q(q-1) \left(g_2(e^t)\right)^{q-2} \big(g_2'(e^t)\big)^2 e^{2t} + q\big(g_2(e^t)\big)^{q-1} g_2''(e^t) e^{2t} + q\big(g_2(e^t)\big)^{q-1} g_2'(e^t) e^t.$$

Observe that $g_2(e^t)e^t > 0$ for all $t \in [\ln(3/4), \ln(1 - \frac{1}{(2^{\Delta-1}+1)^2})]$. Thus, using the transformation $y = e^t$, it suffices to show that

$$(q-1) y \left(g_2'(y)\right)^2 + g_2(y)\big(g_2'(y) + y\, g_2''(y)\big) \le 0 \tag{5.67}$$

for all $y \in [3/4, 1 - \frac{1}{(2^{\Delta-1}+1)^2}]$. For convenience, we verify (5.67) using Mathematica's RESOLVE function; see section 8.2.11 for the code. ∎

We are now ready to prove Lemma 5.7, which we restate here for convenience.

LEMMA 5.7. *Let $\Delta = 6$ and b_3 be a nonnegative integer less than or equal to $\Delta - 1 = 5$. There exists a constant $C_\delta^{(b_3)} \ge 0$ so that for $y_1, \ldots, y_{b_3} \in [3/4, 1 - 1/(2^d+1)^2]$ it holds that*

$$\sum_{i=1}^{b_3} \Big(\frac{1}{\delta}\Big)^{b_3-i} g_2(y_i) \le b_3\, C_\delta^{(b_3)}\, g_2\big(\sqrt[b_3]{y_1 \cdots y_{b_3}}\big).$$

In particular, we will show that the inequality holds with $C_\delta^{(b_3)} = \frac{1}{\delta^{b_3-1}} \big(\frac{1-\delta^{b_3 p}}{b_3(1-\delta^p)}\big)^{1/p}$, where $p = 27/2$. For $\delta = 9789/10000$, we have the following upper bounds on the values of $C_\delta^{(b_3)}$:

$$C_\delta^{(0)} = 0,\ C_\delta^{(1)} = 1,\ C_\delta^{(2)} = 102/100,\ C_\delta^{(3)} = 103/100,\ C_\delta^{(4)} = 104/100,\ C_\delta^{(5)} = 105/100.$$

Proof. For $p = 27/2$ and $q = 27/25$, inequality (5.66) and Lemma 5.14 yield that

$$\sum_{i=1}^{b_3} \left(\frac{1}{\delta}\right)^{b_3-i} g_2(y_i) \le b_3^{1/q} \left(\sum_{i=1}^{b_3} \left(\frac{1}{\delta}\right)^{(b_3-i)p} \right)^{1/p} g_2(\sqrt[b_3]{y_1 \cdots y_{b_3}}).$$

Note that for $b_3 = 1, \ldots, 5$ it holds that

$$\sum_{i=1}^{b_3} \left(\frac{1}{\delta}\right)^{(b_3-i)p} = \frac{\frac{1}{\delta^{b_3 p}} - 1}{\frac{1}{\delta^p} - 1} = \frac{1}{\delta^{(b_3-1)p}} \cdot \frac{1-\delta^{b_3 p}}{1-\delta^p}.$$

Using this and $q = p/(p-1)$, we obtain that

$$\sum_{i=1}^{b_3} \left(\frac{1}{\delta}\right)^{b_3-i} g_2(y_i) \le b_3 \, C_\delta^{(b_3)} g_2(\sqrt[b_3]{y_1 \cdots y_{b_3}}).$$

where $C_\delta^{(b_3)} = \frac{1}{\delta^{b_3-1}} (\frac{1-\delta^{b_3 p}}{b_3(1-\delta^p)})^{1/p}$, as desired. The numerical bounds on the values of $C_\delta^{(b_3)}$ given in the statement of Lemma 5.7 for $b_3 = 1, \ldots, 5$ can be verified by a direct calculation using Mathematica; see section 8.2.4. This concludes the proof of Lemma 5.7. ∎

6. Hardness for approximate counting. In this section, we prove the hardness results stated in the introduction for the problem of counting independent sets in hypergraphs and the problem of counting dominating sets in graphs.

6.1. Counting independent sets in hypergraphs. In this section, we prove inapproximability results for the #HyperIndSet(k, Δ) problem. For this section, it will be convenient to return to the original hypergraph independent set formulation of the problem (instead of the monotone CNF formulation). The proof is via a reduction to the independent set model on graphs which was used by Bordewich, Dyer, and Karpinski [2]. The precise inapproximability results for the hard-core model had not yet been proved at the time [2] was written, so we carry out the details explicitly to obtain the bound that their reduction gives.

Namely, we will use the inapproximability result of Sly and Sun [20] for the hard-core model. We first remind the reader of the relevant definitions. Let $\lambda > 0$. For a graph $G = (V, E)$, the hard-core model with parameter λ is a probability distribution over the set of independent sets of G; each independent set I of G has weight proportional to $\lambda^{|I|}$. The normalizing factor of this distribution is the partition function $Z_G(\lambda)$, formally defined as $Z_G(\lambda) := \sum_I \lambda^{|I|}$, where the sum ranges over all independent sets I of G.

THEOREM 6.1 (see [20]). *For $\Delta \ge 3$, let $\lambda_c(\Delta) := (\Delta-1)^{\Delta-1}/(\Delta-2)^\Delta$. For all $\lambda > \lambda_c(\Delta)$, it is* NP*-hard to approximate $Z_G(\lambda)$ on Δ-regular graphs G, even within an exponential factor.*

THEOREM 6.2. *Let $k \ge 2,\ \Delta \ge 3$ be integers. Suppose that $2^{\lceil k/2 \rceil} - 1 < \frac{(\Delta-2)^\Delta}{(\Delta-1)^{\Delta-1}}$. Then, it is* NP*-hard to approximate* #HyperIndSet(k, Δ)*, even within an exponential factor.*

Proof. Let $k \ge 2,\ \Delta \ge 3$ be integers satisfying $2^{\lceil k/2 \rceil} - 1 < \frac{(\Delta-2)^\Delta}{(\Delta-1)^{\Delta-1}}$, and let

$$\lambda := 1/(2^{\lceil k/2 \rceil} - 1).$$

Note that $\lambda > \lambda_c(\Delta)$ where $\lambda_c(\Delta)$ is as in Theorem 6.1. For convenience, let $k' := \lceil k/2 \rceil$ in what follows.

Let $G = (V, E)$ be a Δ-regular graph, and set $n := |V|$. We will construct a $(2k')$-uniform hypergraph $H = (U, \mathcal{F})$ with maximum degree Δ such that $|U| = k'|V|$, $|\mathcal{F}| = |E|$, and

$$Z_H = (2^{k'} - 1)^n \, Z_G(\lambda). \tag{6.1}$$

Note that the size of H is larger than the size of G only by a constant factor. It thus follows that if we could approximate $\#\mathsf{HyperIndSet}(k, \Delta)$ within an arbitrarily small exponential factor, we could also approximate $Z_G(\lambda)$ within an (arbitrarily small) exponential factor for all Δ-regular graphs G, contradicting Theorem 6.1.

It remains to construct the hypergraph $H = (U, \mathcal{F})$. Let

$$U = \bigcup_{v \in V} \{u_{v,1}, \ldots, u_{v,k'}\}, \quad \mathcal{F} = \bigcup_{(v,w) \in E} \{\{u_{v,1}, \ldots, u_{v,k'}, u_{w,1}, \ldots, u_{w,k'}\}\}.$$

In words, every vertex v of G maps to a (distinct) set of k' vertices in H, the set $\{u_{v,1}, \ldots, u_{v,k'}\}$, which we will henceforth denote as S_v. Further, each edge (v, w) in G maps to a hyperedge in H which is given by $S_v \cup S_w$. It is clear from the construction that every vertex of H has degree Δ (since G is a Δ-regular graph) and, further, that every hyperedge of H has arity $2k' \geq k$. Also, note that $|U| = k'|V|$ and $|\mathcal{F}| = |E|$.

We complete the proof by showing (6.1). To do this, we will map independent sets of the hypergraph H to independent sets of the graph G as follows. Let I_H be an independent set of H. Define I_G to be the subset of vertices of G such that $v \in I_G$ iff $S_v \cap I_H = S_v$. It is immediate that I_G is an independent set of G. In fact, it is not hard to see that for every independent set I_G of G there are exactly $(2^{k'} - 1)^{n - |I_G|}$ independent sets of H that map to I_G. From this, (6.1) follows, thus completing the proof. ∎

Corollary 6.3. *Let $k = 6$ and $\Delta = 22$. It is* NP*-hard to approximate $\#\mathsf{HyperIndSet}(k, \Delta)$, even within an exponential factor.*

Proof. Just plug in the values of k, Δ to check that the inequality $2^{\lceil k/2 \rceil} - 1 < \frac{(\Delta-2)^\Delta}{(\Delta-1)^{\Delta-1}}$ holds. Then, apply Theorem 6.2. ∎

The following corollary is a crude estimate of the range of Δ in which $\#\mathsf{HyperIndSet}(k, \Delta)$ is hard to approximate (by applying Theorem 6.2).

Corollary 6.4. *Let $k \geq 2$. For all integers $\Delta \geq 5 \cdot 2^{k/2}$, it is* NP*-hard to approximate $\#\mathsf{HyperIndSet}(k, \Delta)$, even within an exponential factor.*

Proof. For $\Delta \geq 5 \cdot 2^{k/2}$, we have that

$$\frac{(\Delta-2)^\Delta}{(\Delta-1)^{\Delta-1}} = (\Delta-2)\Big(1 - \frac{1}{\Delta-1}\Big)^{\Delta-1} \geq \left(\frac{4}{5}\right)^5 (\Delta-2) > \sqrt{2} \cdot 2^{k/2} - 1 \geq 2^{\lceil k/2 \rceil} - 1.$$

The first inequality follows from the fact that $(1 - \frac{1}{\Delta-1})^{\Delta-1}$ is an increasing function of Δ and the (trivial) absolute bound $\Delta \geq 6$ (since $k \geq 2$). The second inequality follows from the fact that $\Delta \geq 5 \cdot 2^{k/2}$ and $k \geq 2$. Finally, the last inequality is trivial.

Now apply Theorem 6.2. ∎

6.2. Counting dominating sets in graphs. In this section, we prove inapproximability results for the problem of counting dominating sets in graphs of maximum degree Δ. In contrast to Corollary 1.4, where we showed algorithmic results for Δ-regular graphs, here we consider graphs which are not necessarily regular but only have bounded maximum degree. Formally, we are interested in the following problem:

Name $\#\mathsf{DomSet}(\Delta)$.
Instance A graph G with maximum degree at most Δ.
Output The number of dominating sets in G.

For unbounded degree graphs, it was shown by Goldberg, Gysel, and Lapinskas [7, Theorem 4] that it is #SAT-hard to approximate the number of dominating sets. We refine this result in a bounded degree setting. More precisely, we show the following.

THEOREM 6.5. *For all integers* $\Delta \geq 18$*, it is* NP*-hard to approximate* $\#\mathsf{DomSet}(\Delta)$*, even within an exponential factor.*

To prove Theorem 6.5, we will utilize inapproximability results for the partition function of the antiferromagnetic 2-spin system on graphs. We give a quick overview of the relevant definitions and results, following [5, 11]. A 2-spin system on a graph is specified by three parameters $\beta, \gamma \geq 0$ and $\lambda > 0$. For a graph $G = (V, E)$, configurations of the system are all possible assignments $\sigma : V \to \{0, 1\}$, and the partition function is given by

$$Z_G(\beta, \gamma, \lambda) = \sum_{\sigma: V \to \{0,1\}} \lambda^{|\sigma^{-1}(0)|} \prod_{(u,v) \in E} \beta^{\mathbf{1}\{\sigma(u)=\sigma(v)=0\}} \gamma^{\mathbf{1}\{\sigma(u)=\sigma(v)=1\}},$$

with the convention that $0^0 \equiv 1$ when one of the parameters β, γ is equal to zero. The case $\beta = \gamma$ corresponds to the Ising model, while the case $\beta = 0$ and $\gamma = 1$ corresponds to the hard-core model (which we already encountered in section 6.1).

The 2-spin system with parameters β, γ, λ is called *antiferromagnetic* if $\beta\gamma < 1$. In [20], the following inapproximability result was shown for antiferromagnetic 2-spin systems on Δ-regular graphs.[8]

THEOREM 6.6 (see [20, Theorems 2 and 3]). *Let* $\beta, \gamma \geq 0$ *with* $\beta\gamma < 1$*,* $\gamma > 0$*,* $\lambda > 0$*, and* $\Delta \geq 3$*. If the 2-spin system specified by the parameters* β, γ, λ *is in the nonuniqueness regime of the infinite* Δ*-regular tree, then there is a* $c > 1$ *such that it is* NP*-hard to approximate* $Z_{\beta,\gamma,\lambda;G}$ *within a factor of* c^n *on the class of* Δ*-regular graphs* G*.*

To apply Theorem 6.6, we will need the following characterization of the uniqueness regime on the infinite Δ-regular tree (see, e.g., [11, Lemma 21] or [5, section 3] for more details). For a 2-spin system with parameters β, γ, λ, nonuniqueness on the infinite Δ-regular tree holds iff the system of equations

$$x = \lambda\Big(\frac{\beta y + 1}{y + \gamma}\Big)^{\Delta - 1}, \quad y = \lambda\Big(\frac{\beta x + 1}{x + \gamma}\Big)^{\Delta - 1} \tag{6.2}$$

has multiple (i.e., more than one) positive solutions (x, y).

We are now ready to give the proof of Theorem 6.5.

[8] Theorems 2 and 3 in [20] are about the hard-core and the antiferromagnetic Ising model on Δ-regular graphs. It is standard to derive from those Theorem 6.6 (which applies to general antiferromagnetic 2-spin systems), since it is well known (see, e.g., [20]) that antiferromagnetic 2-spin systems on Δ-regular graphs can be expressed in terms of either the Ising model or the hard-core model. The detailed derivation can be found in [5, Corollary 21].

Proof of Theorem 6.5. Let $\Delta := 17$, $\Delta' := \Delta + 1 = 18$. Consider the 2-spin system with

$$\lambda = 1/2,\ \beta = 1/2,\ \gamma = 1.$$

In section 8.3.1, we use Mathematica to find that (6.2) has multiple positive solutions (x, y). Thus, we have that the 2-spin system is in the nonuniqueness regime of the infinite Δ-regular tree, and hence Theorem 6.6 applies.

Let $G = (V, E)$ be a Δ-regular graph for which we want to compute $Z_G(\beta, \gamma, \lambda)$. Set $n := |V|$ and $m := |E|$. We will construct a graph $G' = (V', E')$ with maximum degree $\Delta' = 18$ such that $|V'| = 2n + m$, $|E'| = |2m + n|$ and

$$\#\mathrm{DomSets}(G') = 2^n 2^m\, Z_G(\beta, \gamma, \lambda), \tag{6.3}$$

where $\#\mathrm{DomSets}(G')$ denotes the number of dominating sets of G'. Note that the size of G' is larger than the size of G only by a constant factor. It thus follows that if we could approximate $\#\mathsf{DomSet}(\Delta')$ within an arbitrarily small exponential factor, we could also approximate $Z_G(\beta, \gamma, \lambda)$ within an (arbitrarily small) exponential factor for all Δ-regular graphs G, contradicting Theorem 6.6.

It remains to construct the graph $G' = (V', E')$. The graph G' is obtained as follows from G. Denote $V = \{v_1, \ldots, v_n\}$ and $E = \{e_1, \ldots, e_m\}$. For each vertex $v_i \in V$, add a new vertex u_i and connect it to v_i. Further, for each edge $e_t = (v_i, v_j)$ add a new vertex w_t, connect it to both v_i and v_j, and delete the edge e_t. In particular, we have that

$$V' = \{v_1, \ldots, v_n\} \bigcup \{u_1, \ldots, u_n\} \bigcup \{w_1, \ldots, w_m\},$$
$$E' = \{(v_1, u_1), \ldots, (v_n, u_n)\} \bigcup \bigcup_{e_t = (v_i, v_j) \in E} \{(v_i, w_t), (w_t, v_j)\}.$$

It is clear from the construction that every vertex of H has maximum degree at most $\Delta' = \Delta + 1$ (since G is a Δ-regular graph, in G' each of $v_1, \ldots, v_n$ has degree Δ', each of $u_1, \ldots, u_n$ has degree 1, and each of $w_1, \ldots, w_m$ has degree 2). Further, it is clear that $|V'| = 2n + m$ and $|E'| = 2m + n$.

We complete the proof by showing (6.3). To do this, we will map dominating sets of G' to configurations $\sigma : V \to \{0, 1\}$. In particular, let S be a dominating set of G'. For a vertex $v \in V$, we set $\sigma(v) = 1$ iff $v \in S$. Then, it remains to observe that for every $\sigma : V \to \{0, 1\}$, there are exactly

$$2^{|\sigma^{-1}(1)|} \prod_{e=(u,v)\in E} 2^{\mathbf{1}\{\sigma(u)=1 \vee \sigma(v)=1\}} = 2^n 2^m\, (1/2)^{|\sigma^{-1}(0)|} \prod_{e=(u,v)\in E} (1/2)^{\mathbf{1}\{\sigma(u)=\sigma(v)=0\}} \tag{6.4}$$

dominating sets of the graph G' that map to σ. To see this, fix $\sigma : V \to \{0, 1\}$. We will consider the possibilities for a dominating set S of G' that maps to σ. For every vertex $v_i \in V$, if $\sigma(v_i) = 1$, we have that $v_i \in S$ and hence the vertex u_i can either belong to S or not belong to S (two choices). In contrast, if $\sigma(v_i) = 0$, we have that $v_i \notin S$, and hence the vertex u_i must belong to S in order to be dominated since its only neighbor is vertex v_i (one choice). Similarly, for every $e_t = (v_i, v_j) \in E$, if $\sigma(v_i) = \sigma(v_j) = 0$, then $v_i, v_j \notin S$, and hence the vertex w_t must belong to S in order to be dominated since its only neighbors are the vertices v_i, v_j (one choice). In all other cases, w_t can either belong to S or not belong to S (two choices). This justifies (6.4), thus completing the justification of (6.3).

Note that the purpose of the "bristle" vertices $u_1, \ldots, u_n$ is to make the interactions of the edges of G' independent of each other. If G has edges $e_t = (v_i, v_j)$ and $e_{t'} = (v_i, v_{j'})$ and amongst $\{v_i, w_t, v_j\}$, only $v_j \in S$, then u_i has to be in S, so $w_{t'}$ can either be in S or not, independently of w_t and v_j.

This concludes the proof. ∎

7. The uniqueness threshold on the infinite hypertree. We denote by $\mathbb{T}_{k,\Delta}$ the infinite $(\Delta-1)$-ary k-uniform hypertree with root vertex ρ. Also, for $n = 0, 1, 2, \ldots$, denote by $\mathbb{T}_{k,\Delta}(n)$ the subtree of $\mathbb{T}_{k,\Delta}$ obtained by the first n levels; i.e., $\mathbb{T}_{k,\Delta}(n)$ is the tree induced by the set of vertices at distance $\leq n$ from ρ in $\mathbb{T}_{k,\Delta}$. We denote by V_n the vertex set of $\mathbb{T}_{k,\Delta}(n)$ and by L_n the leaves of the tree, i.e., vertices with degree 1 in $\mathbb{T}_{k,\Delta}(n)$.

Denote by μ_n the Gibbs distribution of the independent set model on $\mathbb{T}_{k,\Delta}(n)$ (see section 2.1). For a configuration $\sigma : V_n \to \{0,1\}$, we denote by σ_{L_n} the restriction of σ to the set L_n and by σ_ρ the spin of the root ρ.

DEFINITION 7.1. *Let $k \geq 2, \Delta \geq 2$ be integers. The independent set model has uniqueness on $\mathbb{T}_{k,\Delta}$ iff*

$$\limsup_{n\to\infty} \max_{\eta,\eta' : L_n \to \{0,1\}} \left|\mu_n(\sigma_\rho = 1 \mid \sigma_{L_n} = \eta) - \mu_n(\sigma_\rho = 1 \mid \sigma_{L_n} = \eta')\right| = 0. \tag{7.1}$$

We will use $\sigma_{L_n} = 1$ to denote that, in the configuration σ, all vertices in L_n are assigned the spin 1. For $n = 0, 1, 2, \ldots$, define

$$p_n = \mu_n(\sigma_\rho = 1 \mid \sigma_{L_n} = 1). \tag{7.2}$$

When $\sigma_\rho = 1$, in each of the $\Delta - 1$ hyperedges that include ρ, at least one of the $k-1$ vertices (other than ρ) must have spin 0. When $\sigma_\rho = 0$, any configuration on the neighbors of ρ is allowed. By considering the (normalized) weight of such configurations on $\mathbb{T}_{k,\Delta}(n+1)\backslash\rho$, it is not hard to see that the sequence p_n satisfies the following recursion for every integer $n \geq 0$:

$$p_{n+1} = f(p_n), \text{ where } f(z) := \frac{(1-z^{k-1})^{\Delta-1}}{1+(1-z^{k-1})^{\Delta-1}}. \tag{7.3}$$

For any configuration $\eta : L_n \to \{0,1\}$, we will see that $\mu_n(\sigma_\rho = 1 \mid \sigma_{L_n} = \eta)$ is sandwiched between p_n and p_{n+1}. This yields the following.

LEMMA 7.2. *Let $k \geq 2, \Delta \geq 2$ be integers. The independent set model has uniqueness on $\mathbb{T}_{k,\Delta}$ iff*

$$\limsup_{n\to\infty} |p_{n+1} - p_n| = 0. \tag{7.4}$$

Proof. Let n be a nonnegative integer, and let η be an arbitrary configuration on L_n, i.e., $\eta : L_n \to \{0,1\}$. We will show that

$$\begin{aligned} &p_{n+1} \leq \mu_n(\sigma_\rho = 1 \mid \sigma_{L_n} = \eta) \leq p_n \text{ for even integers } n, \\ &p_n \leq \mu_n(\sigma_\rho = 1 \mid \sigma_{L_n} = \eta) \leq p_{n+1} \text{ for odd integers } n. \end{aligned} \tag{7.5}$$

Let us first conclude the lemma assuming (7.5). From (7.5), we obtain that for all n, it holds that

$$\max_{\eta,\eta' : L_n \to \{0,1\}} \left|\mu_n(\sigma_\rho = 1 \mid \sigma_{L_n} = \eta) - \mu_n(\sigma_\rho = 1 \mid \sigma_{L_n} = \eta')\right| = |p_{n+1} - p_n|.$$

It follows that the conditions in (7.1) and (7.4) are equivalent, which yields the statement in the lemma.

We next show (7.5). The proof is by induction on n. The claim is trivial for $n = 0$ since $p_0 = 1$ and $p_1 = 0$. So assuming that the claim holds for all nonnegative integers less than n, we will show it for n.

Set $d := \Delta - 1$. Let $e_1, \dots, e_d$ be the d hyperedges containing ρ, and for $i \in [d]$ denote by $v_{i,1}, \dots, v_{i,k-1}$ the vertices in e_i other than ρ, i.e.,

$$e_1 = \{\rho, v_{1,1}, \dots, v_{1,k-1}\}, \dots, e_d = \{\rho, v_{d,1}, \dots, v_{d,k}\}.$$

For $i \in [d]$ and $j \in [k-1]$, let $T_{i,j}$ be the subtree of $\mathbb{T}_{k,\Delta}(n)$ rooted at $v_{i,j}$. Denote by $S_{i,j}$ the leaves of $T_{i,j}$ and by $\eta_{i,j}$ the restriction of η on $S_{i,j}$. Let $\mu_{T_{i,j}}$ be the Gibbs distribution of $T_{i,j}$ in the independent set model. Note that $\cup_{i\in[d],j\in[k-1]} S_{i,j} = L_n$. Finally, let

$$q_{i,j} := \mu_{T_{i,j}}(\sigma_{v_{i,j}} = 1 \mid \sigma_{S_{i,j}} = \eta_{i,j}),$$
$$q := \mu_n(\sigma_\rho = 1 \mid \sigma_{L_n} = \eta).$$

It is simple to see that

$$q = \frac{\prod_{i=1}^{d}(1 - \prod_{j=1}^{k-1} q_{i,j})}{1 + \prod_{i=1}^{d}(1 - \prod_{j=1}^{k-1} q_{i,j})}, \text{ or equivalently that } \frac{q}{1-q} = \prod_{i=1}^{d}\left(1 - \prod_{j=1}^{k-1} q_{i,j}\right). \tag{7.6}$$

For $i \in [d]$ and $j \in [k-1]$, note that $T_{i,j}$ is isomorphic to $\mathbb{T}_{k,\Delta}(n-1)$, and hence we can use the induction hypothesis to bound $q_{i,j}$. Let us consider first the case where n is odd. Then, we have that

$$p_n \le q_{i,j} \le p_{n-1}. \tag{7.7}$$

It follows from (7.6) and (7.7) that

$$\left(1 - (p_{n-1})^{k-1}\right)^d \le \frac{q}{1-q} \le \left(1 - (p_n)^{k-1}\right)^d, \text{ so that } p_n \le q \le p_{n+1}, \tag{7.8}$$

where to derive the last inequality we used that the sequence p_n satisfies the recursion in (7.3) and that the function $\frac{x}{1-x}$ is increasing in x. The proof for odd n is completely analogous, modulo that the inequalities in (7.7) and (7.8) hold in the opposite direction.

This concludes the induction step and hence the proof of (7.5). The proof of the lemma is thus complete. ∎

LEMMA 7.3. *Let $k \ge 2$, $\Delta \ge 2$ be integers. Let $f(z) = \frac{(1-z^{k-1})^{\Delta-1}}{1+(1-z^{k-1})^{\Delta-1}}$ be as in* (7.3). *The function f is strictly decreasing in the interval $[0,1]$. Also, there is unique $x \in [0,1]$ such that*

$$f(x) = x, \tag{7.9}$$

which further satisfies $|f'(x)| = \frac{(\Delta-1)(k-1)x^{k-1}(1-x)}{1-x^{k-1}}$.

Finally, if $|f'(x)| < 1$, the equation

$$f(f(z)) = z \text{ for } z \in [0,1] \tag{7.10}$$

is uniquely *satisfied by $z = x$.*

Proof. To see that the function f is decreasing, we calculate

$$f'(z) = -\frac{(\Delta-1)(k-1)\big(1-z^{k-1}\big)^{\Delta-2}z^{k-2}}{\big(1+(1-z^{k-1})^{\Delta-1}\big)^2}, \tag{7.11}$$

which clearly shows that f is decreasing for $z \in [0,1]$. Note that $f'(z) = 0$ iff $z = 0$ or $z = 1$, so in fact f is strictly decreasing over the interval [0,1].

We next show the second part of the lemma. We can rewrite $z = f(z)$ as

$$g(z) = 0, \text{ where } g(z) := z - (1-z)(1-z^{k-1})^{\Delta-1}.$$

For the function g, we have that $g(0) = -1$, $g(1) = 1$, and g is continuous on $[0,1]$. It thus follows that there exists x such that $g(x) = 0$, which implies that $f(x) = x$. We next prove that x is unique; i.e., for all $z \neq x$ it holds that $g(z) \neq 0$. For this, it suffices to show that g is increasing on $[0,1]$ or that $g'(z) > 0$ for all $z \in [0,1]$. We calculate

$$g'(z) = 1 + (1-z^{k-1})^{\Delta-1} + (k-1)(\Delta-1)(1-z)(1-z^{k-1})^{\Delta-2}z^{k-2},$$

which clearly shows that $g'(z) \geq 1$ for all $z \in [0,1]$. Finally, to see the expression for $|f'(x)|$, just use (7.11) and use that

$$x = \frac{(1-x^{k-1})^{\Delta-1}}{1+(1-x^{k-1})^{\Delta-1}}, \quad 1-x = \frac{1}{1+(1-x^{k-1})^{\Delta-1}}$$

to simplify. This proves the second assertion in the lemma.

Finally, we show the last part of the lemma. Namely, suppose that $|f'(x)| < 1$. Consider the function

$$h(z) := f(f(z)) - z \text{ for } z \in [0,1].$$

Clearly, we have that $h(x) = 0$. By the assumption $|f'(x)| < 1$, we have that $h'(x) < 0$, and hence for small $\varepsilon > 0$ it holds that $h(x-\varepsilon) > 0$ and $h(x+\varepsilon) < 0$. Note also that $h(0) > 0$ and $h(1) < 0$.

For the sake of contradiction, assume that h has a zero other than x, say at $z = \rho_1$, where $\rho_1 \neq x$. Let $\rho_2 := f(\rho_1)$, and note that $f(\rho_2) = \rho_1$. Observe also that h has another zero at $z = \rho_2$. Also, using that f is decreasing and $f(x) = x$, we have $\rho_2 \neq \rho_1, x$. In fact, we have that x is between ρ_1 and ρ_2. Without loss of generality we may thus assume that $\rho_1 > x > \rho_2$ (otherwise, we may swap ρ_1 and ρ_2).

We may also assume that $h'(\rho_1) \geq 0$. Otherwise, we claim that there exists $\rho \in (x, \rho_1)$ such that $h(\rho) = 0$ and $h'(\rho) \geq 0$, so that if $h'(\rho_1) < 0$, we could switch our focus from ρ_1 to ρ (the role of $\rho_2 = f(\rho_1)$ would be played by $f(\rho)$). To see the claim, assume that $h'(\rho_1) < 0$, and let $\varepsilon > 0$ be sufficiently small. Then, it holds that $h(\rho_1 - \varepsilon) > 0$. Recall also that $h(x+\varepsilon) < 0$. It follows that there must exist a *crossing* of h in the interval (x, ρ_1), i.e., a real number $\rho \in (x, \rho_1)$ such that $h(\rho) = 0$ and, for all sufficiently small ε', it holds that $h(\rho+\varepsilon') < 0$ and $h(\rho-\varepsilon') > 0$. It must thus be the case that $h'(\rho) \geq 0$, as claimed.

Thus, if $|f'(x)| < 1$, we have concluded the existence of ρ_1, ρ_2 such that $1 > \rho_1 > \rho_2 > 0$, $\rho_1 = f(\rho_2)$, $\rho_2 = f(\rho_1)$, and $h'(\rho_1) \geq 0$. We obtain a contradiction by showing that $h'(\rho_1) < 0$. We have

$$h'(\rho_1) = f'(\rho_1)f'(\rho_2) - 1 = \frac{(\Delta-1)^2(k-1)^2\rho_1^{k-2}\rho_2^{k-2}\big(1-\rho_1^{k-1}\big)^{\Delta-2}\big(1-\rho_2^{k-1}\big)^{\Delta-2}}{\big(1+(1-\rho_1^{k-1})^{\Delta-1}\big)^2\big(1+(1-\rho_2^{k-1})^{\Delta-1}\big)^2} - 1. \tag{7.12}$$

For $k = 2, \Delta = 2$, (7.12) is trivial, so henceforth we may assume that at least one of k, Δ is greater than or equal to 3.

We can rewrite $\rho_1 = f(\rho_2)$ and $\rho_2 = f(\rho_1)$ as

$$1 - \rho_1 = \frac{1}{1 + \left(1 - \rho_2^{k-1}\right)^{\Delta-1}}, \quad 1 - \rho_2 = \frac{1}{1 + \left(1 - \rho_1^{k-1}\right)^{\Delta-1}}. \tag{7.13}$$

Subtracting the equations in (7.13) gives

$$\rho_1 - \rho_2 = \frac{\left(1 - \rho_2^{k-1}\right)^{\Delta-1} - \left(1 - \rho_1^{k-1}\right)^{\Delta-1}}{\left(1 + (1 - \rho_1^{k-1})^{\Delta-1}\right)\left(1 + (1 - \rho_2^{k-1})^{\Delta-1}\right)}. \tag{7.14}$$

Let $E := \left(1-\rho_2^{k-1}\right)^{\Delta-1} - \left(1-\rho_1^{k-1}\right)^{\Delta-1}$, and assume for now that $k \geq 3$, $\Delta \geq 3$. We use the inequality $x^d - y^d > d(x-y)(xy)^{(d-1)/2}$, which holds for all $d \geq 2$ and real numbers $x > y > 0$ (see, for example, [6, Claim 35]), to obtain the following bound for E:

$$\begin{aligned} E &> (\Delta-1)(\rho_1^{k-1} - \rho_2^{k-1})(1-\rho_1^{k-1})^{(\Delta-2)/2}(1-\rho_2^{k-1})^{(\Delta-2)/2} \\ &> (\Delta-1)(k-1)(\rho_1 - \rho_2)\rho_1^{(k-2)/2}\rho_2^{(k-2)/2}(1-\rho_1^{k-1})^{(\Delta-2)/2}(1-\rho_2^{k-1})^{(\Delta-2)/2}. \end{aligned} \tag{7.15}$$

Note that the strict inequality in (7.15) holds even in the cases where $k = 2$, $\Delta \geq 3$ or $\Delta = 2, k \geq 3$.

Combining (7.14) and (7.15), we obtain that

$$1 > \frac{(\Delta-1)(k-1)\rho_1^{(k-2)/2}\rho_2^{(k-2)/2}(1-\rho_1^{k-1})^{(\Delta-2)/2}(1-\rho_2^{k-1})^{(\Delta-2)/2}}{\left(1 + (1 - \rho_1^{k-1})^{\Delta-1}\right)\left(1 + (1 - \rho_2^{k-1})^{\Delta-1}\right)}.$$

Squaring the last inequality and using (7.12), we obtain $h'(\rho_1) < 0$, as desired.

This concludes the proof of Lemma 7.3. ∎

LEMMA 7.4. *Let $k \geq 2, \Delta \geq 2$ be integers. Let $f(z) = \frac{(1-z^{k-1})^{\Delta-1}}{1+(1-z^{k-1})^{\Delta-1}}$ be as in* (7.3) *and x be as in Lemma* 7.3.

If $|f'(x)| < 1$, the independent set model has uniqueness on $\mathbb{T}_{k,\Delta}$.

If $|f'(x)| > 1$, the independent set model has nonuniqueness on $\mathbb{T}_{k,\Delta}$.

Proof. Recall the sequence p_n defined in (7.3). Let $p_n^+ = p_{2n}$ and $p_n^- = p_{2n+1}$. As a consequence of the fact that f is decreasing (cf. Lemma 7.3) and $p_0^+ = 1$, $p_0^- = 0$, we have that

$$p_n^+ \downarrow p^+, \quad p_n^- \uparrow p^-, \tag{7.16}$$

where p^+, p^- are real numbers in $[0, 1]$. To see the existence of these limits, note that $p_0^+ = 1 \geq p_1^+$ and $p_0^- = 0 \leq p_1^-$. Since $p_{n+1}^{\pm} = f(f(p_n^{\pm}))$, a simple induction shows that p_n^+ is a decreasing sequence and p_n^- is an increasing sequence. Since both sequences are bounded, we obtain the existence of the limits in (7.16). For later use, we remark here that the continuity of f and the recursions $p_{n+1}^{\pm} = f(p_n^{\mp})$ and $p_{n+1}^{\pm} = f(f(p_n^{\pm}))$ imply that p^+, p^- satisfy the equalities

$$p^+ = f(p^-) \text{ and } p^- = f(p^+), \tag{7.17}$$

$$p^+ = f(f(p^+)) \text{ and } p^- = f(f(p^-)). \tag{7.18}$$

As a consequence of the existence of the limits in (7.16), we can conclude that the condition $\limsup_{n\to\infty} |p_{n+1} - p_n| = 0$ is equivalent to

$$p^+ = p^-. \tag{7.19}$$

We are now ready to show the equivalence in the lemma. For the first part, assume that $|f'(x)| < 1$. To show that uniqueness holds on $\mathbb{T}_{k,\Delta}$, it suffices to show that (7.19) holds, i.e., $p^+ = p^-$. We have that p^+, p^- satisfy (7.18). From the second part of Lemma 7.3 and the assumption $|f'(x)| < 1$, we thus obtain that $p^+ = x = p^-$, as wanted.

For the second part, it suffices to show the contrapositive. So, assuming that uniqueness holds on $\mathbb{T}_{k,\Delta}$, we will show that $|f'(x)| \leq 1$, where we recall that x is specified by the relation $x = f(x)$ (cf. the second part of Lemma 7.3). From Lemma 7.2 we have that $\limsup_{n\to\infty} |p_{n+1} - p_n| = 0$, and hence by our previous arguments, we obtain that (7.19) also holds. From (7.17) and using the uniqueness of x (cf. Lemma 7.3), we obtain that the common value of p^+ and p^- is x and thus $p_n \to x$.

By the Mean Value Theorem we have that there exists ξ_n between p_{n+1} and p_n such that $|f'(\xi_n)| = |p_{n+1} - p_n|/|p_n - p_{n-1}|$. From $p_n \to x$, we also have that $\xi_n \to x$. Since $p_n \to x$, we have that for infinitely many n it holds that $|p_{n+1} - p_n| \leq |p_n - p_{n-1}|$. For all such n, it holds that $|f'(\xi_n)| \leq 1$. Using that f' is continuous and $\xi_n \to x$, we obtain that $|f'(x)| \leq 1$, as desired.

This concludes the proof of Lemma 7.4. ∎

The following lemma establishes the intuitive fact that, for the independent set model, uniqueness on $\mathbb{T}_{k,\Delta}$ is a monotone property with respect to Δ.

LEMMA 7.5. *Let $k \geq 2$ be an integer. There exists $\Delta_c(k) \geq 3$ such that the following holds for all integer $\Delta \geq 2$. The independent set model has uniqueness on $\mathbb{T}_{k,\Delta}$ whenever $\Delta < \Delta_c(k)$ and nonuniqueness whenever $\Delta > \Delta_c(k)$.*

Proof. Fix an integer $k \geq 2$. For integer $\Delta \geq 2$, parameterize the function $f(z)$ in (7.3) by Δ; i.e., set

$$f_\Delta(z) = \frac{(1 - z^{k-1})^{\Delta-1}}{1 + (1 - z^{k-1})^{\Delta-1}}.$$

Let x_Δ be the unique solution of $z = f_\Delta(z)$ (cf. the second part of Lemma 7.3). Recall also from Lemma 7.3 that

$$|f'_\Delta(x_\Delta)| = (\Delta - 1)(k - 1)h(x_\Delta), \text{ where } h(z) := \frac{z^{k-1}(1 - z)}{1 - z^{k-1}}. \tag{7.20}$$

We will show that $|f'_\Delta(x_\Delta)|$ is a (strictly) increasing function of Δ, i.e., $|f'_{\Delta+1}(x_{\Delta+1})| > |f'_\Delta(x_\Delta)|$. This follows immediately by multiplying the following inequalities:

$$\Delta(x_{\Delta+1})^{k-1} > (\Delta - 1)(x_\Delta)^{k-1}, \tag{7.21}$$

$$\frac{(x_\Delta)^{k-1} - 1}{x_\Delta - 1} \geq \frac{(x_{\Delta+1})^{k-1} - 1}{x_{\Delta+1} - 1}. \tag{7.22}$$

Thus, to show the desired monotonicity, we only need to argue for the validity of (7.21) and (7.22). We will need the following simple fact:

$$\text{for all } z \in (0, 1) \text{ and all } \Delta \geq 2, \text{ it holds that } f_{\Delta+1}(z) < f_\Delta(z). \tag{7.23}$$

To see (7.23), fix $z \in (0,1)$, and let $\Delta \geq 2$. Then we have

$$\frac{f_{\Delta+1}(z)}{1-f_{\Delta+1}(z)} = (1-z^{k-1})^{\Delta} < (1-z^{k-1})^{\Delta-1} = \frac{f_{\Delta}(z)}{1-f_{\Delta}(z)},$$

and thus (7.23) follows.

We next proceed to show (7.21) and (7.22). The crucial step will be to show that x_Δ is a decreasing function of Δ, i.e., $x_{\Delta+1} < x_\Delta$. Suppose for the sake of contradiction that $x_\Delta \leq x_{\Delta+1}$ for some Δ. Using (7.23) for $z = x_\Delta$ and the fact that $f_{\Delta+1}(z)$ is decreasing (by Lemma 7.3), we have

$$x_\Delta = f_\Delta(x_\Delta) > f_{\Delta+1}(x_\Delta) \geq f_{\Delta+1}(x_{\Delta+1}) = x_{\Delta+1},$$

a contradiction. It thus follows that $x_\Delta > x_{\Delta+1}$. From this, it is simple to conclude (7.22):

$$\frac{(x_\Delta)^{k-1}-1}{x_\Delta - 1} = \sum_{j=0}^{k-2}(x_\Delta)^j \geq \sum_{j=0}^{k-2}(x_{\Delta+1})^j = \frac{(x_{\Delta+1})^{k-1}-1}{x_{\Delta+1}-1}.$$

We next show the slightly harder (7.21). From $x_\Delta > x_{\Delta+1}$ and the facts $x_\Delta = f_\Delta(x_\Delta)$ and $x_{\Delta+1} = f_{\Delta+1}(x_{\Delta+1})$, it follows that $f_\Delta(x_\Delta) > f_{\Delta+1}(x_{\Delta+1})$. This in turn yields

$$(1-x_\Delta^{k-1})^{\Delta-1} = \frac{f_\Delta(x_\Delta)}{1-f_\Delta(x_\Delta)} > \frac{f_{\Delta+1}(x_{\Delta+1})}{1-f_{\Delta+1}(x_{\Delta+1})} = (1-x_{\Delta+1}^{k-1})^{\Delta}.$$

Using Bernoulli's inequality, we thus obtain

$$1-x_\Delta^{k-1} > (1-x_{\Delta+1}^{k-1})^{\frac{\Delta}{\Delta-1}} \geq 1 - \frac{\Delta}{\Delta-1}x_{\Delta+1}^{k-1},$$

and, by rearranging, we obtain (7.21).

Since $|f'_\Delta(x_\Delta)|$ is increasing for $\Delta \geq 2$, to complete the proof of the lemma, it suffices to show that for $\Delta = 2$ it holds that $|f'_\Delta(x_\Delta)| < 1$ and that for some Δ it holds that $|f'_\Delta(x_\Delta)| > 1$.

Note that for all $z \in (0,1)$, we have $(k-1)z^k + 1 > kz^{k-1}$ (the function $(k-1)z^k + 1 - kz^{k-1}$ is decreasing for $z \in [0,1]$ and it has a zero at $z = 1$). Rearranging, we obtain that $h(z) < 1/(k-1)$, so that for $\Delta = 2$ we have from (7.20) that $|f'_\Delta(x_\Delta)| < 1$, as needed.

For large Δ, since x_Δ is decreasing and bounded, we have that x_Δ converges. From $x_\Delta = f_\Delta(x_\Delta)$, we thus obtain that $x_\Delta \downarrow 0$ as $\Delta \to \infty$. We claim also that $\Delta(x_\Delta)^{k-1} \to \infty$ as $\Delta \to \infty$, from which it clearly follows that $|f'_\Delta(x_\Delta)| \to \infty$ (cf. (7.20)) and hence $|f'_\Delta(x_\Delta)| > 1$ for large Δ. To see the claimed limit, assume for the sake of contradiction that there was $M > 0$ such that $\Delta(x_\Delta)^{k-1} < M$ for infinitely many Δ. For all such Δ, it holds that

$$x_\Delta = f_\Delta(x_\Delta) = \frac{(1-x_\Delta^{k-1})^{\Delta-1}}{1+(1-x_\Delta^{k-1})^{\Delta-1}} \geq \frac{1}{2}(1-x_\Delta^{k-1})^{\Delta-1} \geq \frac{1}{2}\Big(1-\frac{M}{\Delta}\Big)^{\Delta-1}.$$

It follows that $\limsup_{\Delta\to\infty} x_\Delta \geq \frac{1}{2}\exp(-M) > 0$, contradicting that $x_\Delta \downarrow 0$ as $\Delta \to \infty$.

This completes the proof of Lemma 7.5. ∎

COROLLARY 7.6. *Let $k = 6$. The independent set model has uniqueness on $\mathbb{T}_{k,\Delta}$ iff $\Delta \leq 28$.*

Proof. We will show that for $\Delta = 28$, the independent set model has uniqueness on $\mathbb{T}_{k,\Delta}$ and nonuniqueness for $\Delta = 29$. The result then follows from Lemma 7.5.

For $\Delta = 28$, we have that $|f'(x)| < 0.996$, where $x, f(x)$ are as in Lemma 7.3. For $\Delta = 29$, we have that $|f'(x)| > 1.01$. Lemma 7.4 thus shows that the independent set model has uniqueness for $\Delta = 28$ and does not have uniqueness for $\Delta = 29$, as needed. ∎

The following lemma asserts that, asymptotically in k, the critical value $\Delta_c(k)$ in Lemma 7.5 satisfies $\Delta_c(k) = (1 + o_k(1))2^k/k$. The precise statement is as follows.

LEMMA 7.7. *Let $\Delta_c(k)$ be as in Lemma 7.5. For all $\varepsilon > 0$, there exists an integer $k_0(\varepsilon) > 0$ such that the following holds. For all integers $k \geq k_0(\varepsilon)$,*

$$(1 - \varepsilon)2^k/k \leq \Delta_c(k) \leq (1 + \varepsilon)2^k/k.$$

Proof. Let $\varepsilon > 0$, and let $k_0(\varepsilon) \geq 2$ be a large constant (depending only on ε). We will henceforth assume that k is an integer satisfying $k \geq k_0(\varepsilon)$.

Let $\Delta_k^+ := \lfloor (1+\varepsilon)2^k/k \rfloor$ and $\Delta_k^- := \lceil (1-\varepsilon)2^k/k \rceil$. Note that for all sufficiently large k it holds that $\Delta_k^\pm \geq 3$. Let $f_{k,\pm}(z)$ be the function $f(z)$ in (7.3) when $\Delta = \Delta_k^\pm$, i.e.,

$$f_{k,+}(z) = \frac{(1 - z^{k-1})^{\Delta_k^+ - 1}}{1 + (1 - z^{k-1})^{\Delta_k^+ - 1}} \text{ and } f_{k,-}(z) = \frac{(1 - z^{k-1})^{\Delta_k^- - 1}}{1 + (1 - z^{k-1})^{\Delta_k^- - 1}}.$$

Further, let $x_k^\pm$ be the unique solution of $x = f_{k,\pm}(x)$. We will show that for all sufficiently large k, it holds that

$$|f'_{k,-}(x_k^-)| < 1 \text{ and } |f'_{k,+}(x_k^+)| > 1. \tag{7.24}$$

Together with Lemmas 7.4 and 7.5, (7.24) yields that $\Delta_k^- \leq \Delta_c(k) \leq \Delta_k^+$, as desired.

The key to showing (7.24) is that $x_k^\pm \to 1/2$ as $k \to \infty$. Assuming this for the moment, let us conclude (7.24). Recall from Lemma 7.3 that

$$|f'_{k,-}(x_k^-)| = \frac{(\Delta_k^- - 1)(k-1)(x_k^-)^{k-1}(1 - x_k^-)}{1 - (x_k^-)^{k-1}},\ |f'_{k,+}(x_k^+)| = \frac{(\Delta_k^+ - 1)(k-1)(x_k^+)^{k-1}(1 - x_k^+)}{1 - (x_k^+)^{k-1}}.$$

Now just using that $x_k^\pm \to 1/2$, we obtain that for $k \to \infty$ it holds that

$$|f'_{k,-}(x_k^-)| \to 1 - \varepsilon, \qquad |f'_{k,+}(x_k^+)| \to 1 + \varepsilon.$$

This shows that (7.24) holds for all sufficiently large k.

We next show that $x_k^\pm \to 1/2$ as $k \to \infty$. From $x_k^\pm = f_{k,\pm}(x_k^\pm)$, we obtain that

$$\frac{x_k^\pm}{1 - x_k^\pm} = \left(1 - (x_k^\pm)^{k-1}\right)^{\Delta - 1}, \tag{7.25}$$

from which it clearly follows that $x_k^\pm \leq 1/2$. Thus, it suffices to show that for every $\varepsilon' > 0$, for all sufficiently large k, it holds that $x_k^\pm > (1/2) - \varepsilon'$.

From (7.25), by taking logarithms and then the kth root, we obtain

$$(\Delta_k^\pm - 1)^{1/k} = g(x_k^\pm)^{1/k}, \text{ with } g(z) := \frac{\ln\left(\frac{z}{1-z}\right)}{\ln\left(1 - z^{k-1}\right)}. \tag{7.26}$$

We will use that the function $g(z)$ is decreasing for $z \in (0, 1/2]$, since

$$g'(z) = \frac{(k-1)z^{k-2}\ln\left(\frac{z}{1-z}\right)}{(1-z^{k-1})\ln^2(1-z^{k-1})} + \frac{\left(\frac{z}{(1-z)^2}+\frac{1}{1-z}\right)(1-z)}{z\ln(1-z^{k-1})} \le 0.$$

We will also use that for any $z \in (0, 1/2)$ it holds that

$$\lim_{k\to\infty}(g(z))^{1/k} = 1/z, \text{ since } \lim_{k\to\infty}\Big(-\ln\Big(\frac{z}{1-z}\Big)\Big)^{1/k} \to 1 \text{ and } \lim_{k\to\infty}\big(-\ln(1-z^{k-1})\big)^{1/k} \to z.$$

Now, for the sake of contradiction, assume that there was $\varepsilon' > 0$ such that $x_k \le (1/2) - \varepsilon'$ for infinitely many k. For all such k, we would have that $g(x_k^{\pm}) \ge g((1/2)-\varepsilon')$, and thus, by taking the lim sup in (7.26), we obtain

$$2 = \limsup_{k\to\infty}(\Delta_k^{\pm}-1)^{1/k} = \limsup_{k\to\infty}\big(g(x_k^{\pm})\big)^{1/k} \ge \limsup_{k\to\infty}\big(g\big((1/2)-\varepsilon'\big)\big)^{1/k} = \frac{1}{\frac{1}{2}-\varepsilon'} > 2,$$

a contradiction. This proves that $x_k^{\pm} \to 1/2$ as $k \to \infty$, thus concluding the proof of Lemma 7.7. ∎

8. Computer assisted proofs. The code in each subsection below can be executed by copying it into a Mathematica cell. It is safest to quit the local kernel before executing each subsection to avoid interference between them.

8.1. Analysis for $k \ge \Delta \ge 200$.

8.1.1. Mathematica code for proof of (4.6).

```
psi = 13/10; chi = 1/2; alpha = 1 - 10^(-4);
f = 1/(psi - (1 + r)^-chi)*(r (psi - r^chi))/(alpha (1 + r));
Resolve[Exists[r, f > 42/100 && 0 <= r <= 1]]
```

8.1.2. Mathematica code for proof of Lemma 4.3.

```
psi = 13/10; chi = 1/2;
f = (psi - r^chi)/(1 + r);
hatf = f /. {r -> Exp[t]/(1 - Exp[t])};
SECDER = (FullSimplify[D[hatf, {t, 2}]]) /. {Exp[t]->T,Exp[-t] -> 1/T};
Resolve[Exists[T, 0 <= T <= 1/2 && SECDER > 0]]
```

8.1.3. Mathematica code for proof of Lemma 4.5. (** Verification of (4.11) for w=2,3,4,5 **)

```
psi = 13/10; chi = 1/2;
K2 = 111614/10^5; K3 = 103/100; K4 = 101/100; K5 = 1;
g[w_, t_] := t^w/(1 - t^w) (1 - t) (psi - (t/(1 - t))^chi)
Resolve[Exists[t, g[2, t] >  K2 * g[2, 1/2] && 0 < t <= 1/2]]
Resolve[Exists[t, g[3, t] >  K3 * g[3, 1/2] && 0 < t <= 1/2]]
Resolve[Exists[t, g[4, t] >  K4 * g[4, 1/2] && 0 < t <= 1/2]]
Resolve[Exists[t, g[5, t] >  K5 * g[5, 1/2] && 0 < t <= 1/2]]
```

8.1.4. Mathematica code for proof of Lemma 4.6.

```
alpha=1-10^(-4); c=7/10;
(** Verification of (4.15) **)
For[w = 2, w <= 5, w++,
  Print[105/
```

```
      100 (2*alpha*c)^(-1) (1 +
       w^(-1))*(1 - 2^(-w))/(1 - 2^(-w - 1)) < 1]];

(2/(1 + c) (2 alpha*c)^(-1) (1 + w^(-1)) /. w -> 6) < 1
(** Verification of (4.16) **)
1+(c^(1-7/200)/(1-c^(1-7/200)))*(1-c^(1/200))<105/100
```

8.1.5. Mathematica code for proof of Lemma 4.7.

```
alpha=1-10^(-4); c=7/10;
(** Verification of (4.23) **)
alpha^(-1)*(2c)^(-2)*2*(1-2^(-2))^(-1)/(1-c^(1-3/200))<45931/10000
(** Verification of (4.24) **)
alpha^(-2)*(2c)^(-6)*6*(1-2^(-6))^(-1)/(1-c^(1-7/200))<278045/100000
```

8.1.6. Mathematica code for proof of Lemma 4.8.

```
(** Verification that
the r.h.s. in (4.25) is less than 1 **)
psi = 13/10; K2 = 111614/100000; c=7/10; tau2= 45932/10000;

c^(264/100)/(psi-1)*(15/100)*K2*tau2<1
```

8.1.7. Mathematica code for proof of (4.28).

```
alpha = 1 - 10^(-4); c = 7/10;
psi = 13/10; chi = 1/2;
l[2] = 1; l[3] = 1; l[4] = 1; l[5] = 1;

g[y_, w_] :=  c^(-w) alpha^(-l[w])*w*(1 - y^2)/y^2*(1 - (1 - y^2)^(1/w))*
               (psi - ((1 - y^2)^(1/w)/(1 - (1 - y^2)^(1/w)))^chi);
ylow[w_] := (1 - 2^(-w))^(1/2); yup[w_] := 1;

For[ww = 2, ww <= 5, ww++,
  Print[Resolve[
     Exists[y, g[y, ww] > 2279/10^4 && ylow[ww] <= y < yup[ww]]]];
  ];

For[ww = 2, ww <= 5, ww++,
  Print[Resolve[
     Exists[y,
      g[y, ww] > -15/10 y + 16276/10^4 && ylow[ww] <= y < yup[ww]]]];
  ];

For[ww = 2, ww <= 5, ww++,
  Print[Resolve[
     Exists[y,
      g[y, ww] > -8 y + 8052/10^3 && ylow[ww] <= y < yup[ww]]]];
  ];
```

8.1.8. Mathematica code for proof of Lemma 4.10.

```
(** The exact values
of Y0 and Y1 in (4.31) **)
SOL=Flatten[Solve[{-15/10 YY0 + 16276/10^4 == 2279/10^4,
          -15/10 YY1 + 16276/10^4 == -8 YY1 + 8052/10^3}, {YY0, YY1}]];
Y0=YY0/.SOL; Y1=YY1/.SOL;
```

```
c=7/10; c5=c^(1-6/200); tau6=27805/10^4; psi=13/10;
h[y_] := Min[-15/10 y + 16276/10^4, -8 y + 8052/10^3]
(** Verification of (4.37) **)
15/100*c*tau6/(psi-1)<1
(** Verification of (4.38) **)
Resolve[Exists[y1,
  Y0 <= y1 <= 1 && c/(psi - y1)*(h[y1] + 15/100*c5*tau6) > 1 ] ]
(** Verification of (4.41) for 2<=t<=7, 3<=i<=7 **)
Print[ Table[ Y0^(t-1)-(15/10)*c*c5^(i-1)>0, {t,2,7}, {i,3,t} ] ]
(** Verification of (4.42) for 2<=t<=7 **)
Print[ Table[ Y0*Y1^(t-2)-(15/10)*c*c5>0, {t,2,7} ] ]
(** Verification of (4.44) for t=2,..,7 and 2<=i<=Min[5,t] **)
Print[ Table[
c/(psi-1)*( h[Y0]+ c5*h[Y1]*(1-c5^(t-1))/(1-c5)+15/100*c5^t*tau6 )
      -8c*c5^(i-1)>=0, {t,2,7}, {i,2,Min[5,t]}
] ]
(** Proof that (4.45) is false for t=2,3,4,5 **)
For[tt = 2, tt <= 5, tt++,
Print[Resolve[Exists[y1,
  Y0 <= y1 <= 1 &&
  c/(psi - y1*Y1^(tt-1) ) *
   (h[y1] + c5*h[Y1]*(1-c5^(tt-1))/(1-c5)  +15/100*c5^tt*tau6) > 1
   ] ] ];
]
(** Proof that (4.46) is false **)
Resolve[Exists[y1,
  Y0 <= y1 <= 1 &&
  c/(psi - y1*Y1^4 ) *
   (h[y1] + c5*h[Y1]*(1-c5^5)/(1-c5)  +15/100*c5^6*tau6) > 1
   ] ]
(** Proof that (4.47) is false **)
Resolve[Exists[y1,
  Y0 <= y1 <= 1 &&
  c/(psi - y1*Y1^4 ) *
   (h[y1] + c5*h[Y1]*(1-c5^6)/(1-c5)  +15/100*c5^7*tau6) > 1
   ] ]
```

8.1.9. Mathematica code for proof of Lemma 4.11.

```
(** The exact values of Y0 and Y1 in (4.31) **)
SOL=Flatten[Solve[{-15/10 YY0 + 16276/10^4 == 2279/10^4,
          -15/10 YY1 + 16276/10^4 == -8 YY1 + 8052/10^3}, {YY0, YY1}]];
Y0=YY0/.SOL; Y1=YY1/.SOL;

c=7/10; c5=c^(1-6/200); tau2=45932/10^4; K2=111614/10^5; psi=13/10;
h[y_] := Min[-15/10 y + 16276/10^4, -8 y + 8052/10^3]
(** Verification of (4.41) for t=8, 3<=i<=8 **)
Print[ Table[ Y0^(t-1)-(15/10)*c*c5^(i-1)>0, {t,8,8}, {i,3,t} ] ]
(** Verification of (4.42) for t=8, i=2 **)
Print[ Table[ Y0*Y1^(t-2)-(15/10)*c*c5>0, {t,8,8} ] ]
(** Verification of (4.49) for 2<=i<=5 **)
```

```
Print[ Table[
c/(psi-1)*( h[Y0]+ c5*h[Y1]*(1-c5^(7))/(1-c5)+15/100*c5^8*K2*tau2 )
-8c*c5^(i-1)<0, {i,2,5}
] ]
(** Verification of (4.51) **)
Resolve[Exists[y1,
  Y0 <= y1 <= 1 &&
   c Y1^4/(psi - y1*Y1^4)*
          (h[y1] + c5*h[Y1]*(1 - c5^7)/(1 - c5) + 15/100*c5^8*K2*tau2)
       - 8 c*c5^5 >= 0]]
(** Proof that (4.52) is false **)
Resolve[Exists[y1,
Y0 <= y1 <= 1 &&
c/(psi - y1*Y1^5 ) *
(h[y1] + c5*h[Y1]*(1-c5^7)/(1-c5) +15/100*c5^8*K2*tau2) > 1
] ]
```

8.2. Analysis for $k \geq 3$, $\Delta = 6$.

8.2.1. Mathematica code for proof of Lemma 5.2. (** Verification of (5.17) **)

```
alpha = 1 - 10^(-4); psi = 13/10; Delta = 6; d0 = Delta - 1;
eps[0]=0; eps[1]=6/10; eps[2]=7/10; eps[3]=83/100; eps[4]=91/100;  eps[5]=alpha;
M = 25/1000;

For[BB = 0, BB<=d0, BB++,
  Print["Running for BB=", BB];
  Print[Resolve[1/alpha (eps[BB] + (d0 - BB) M/(psi-1)) > 1]];
  ];
```

8.2.2. Mathematica code for Lemma 5.4. (** Verification of (5.59) **)

```
M1 = 1/410;
h= (1-t) (psi- (t/(1-t))^chi)/.{chi -> 1/2, psi -> 13/10};
f = t^6/(1 - t^6) h;
Resolve[Exists[t, f > M1 && 0 <= t <= 1/2]]
```

(** Verification of (5.61) **)

```
f = (t^(w - 6) (1 - t^6))/(1 - t^w);
p = 6 t^w - t^6 w + w - 6;
Simplify[D[f, t] - (t^(w - 7) p)/(1 - t^w)^2]
```

(** Verification of the first inequality in (5.64) **)

```
M = 25/1000; M1 = 1/410; Delta = 6; delta = 9789/10000;
Kw = (1/delta)^((w - 1) (Delta - 1));
f = (63 M1/(2^w - 1))* Kw * w * (w + 1)^(11*10^(-5)/Log[6]);
(f /. {w -> 6}) <= M
```

(** Verification of (5.65) **)

```
Delta = 6; delta = 9789/10000;
(1/2)*(1/delta)^(Delta-1)(1+6^(-1))(1+7^(-1))^(11*10^(-5)/Log[6])<1
```

8.2.3. Mathematica code for proof of Lemma 5.6. (** Verification of (5.22) **)

```
h= (1-t) (psi- (t/(1-t))^chi)/.{chi -> 1/2, psi -> 13/10};
```

```
For[w = 1, w <= 5, w++,
  Print["Running for w=", w];
  g[w] = (1 - y)/y * (h /. {t -> (1 - y)^(1/w)});
  f[w] = g[w] /. {y -> Exp[z]};
  f2[w] = (D[f[w], {z, 2}])/. {Exp[z] -> T, Exp[-z] -> 1/T, Exp[2*z] -> T^2};
  Print[Resolve[Exists[T, f2[w] > 0 && 1 - (1/2)^w <= T <= 1]]];
];
```

8.2.4. Mathematica code for proof of Lemma 5.7.

```
(** Verification of values C_δ^(b3)  **)
CC[1]=1; CC[2]=102/100; CC[3]=103/100; CC[4]=104/100; CC[5]=105/100;
Delta=6; delta=9789/10000; p=27/2;
Cproof=(1/delta)^(b3-1) * ( (1-delta^(b3*p))/(b3 (1-delta^p)) )^(1/p);
For[bb3=1, bb3<=Delta-1, bb3++,
   Print["Checking for b3=",bb3];
   Print[Resolve[ (Cproof/.{b3->bb3}) > CC[bb3] ]];
]
```

8.2.5. Mathematica code for proof of Lemma 5.8.

```
SUBS1 = {(v1/(1 - v1))^chi -> (2 z1)/(1 - z1^2),
         (1 - v1)^(b2 chi) -> ((1 - z1^2)/(1 + z1^2))^b2,
         v1 -> (4 z1^2)/(1 + z1^2)^2
        };
SUBS2 = {(v2/(1 - v2))^chi -> (2z2)/(1 - 2z2^2),
         (1-v2^2)^(b3 chi) -> ((1 - 4z2^4)/(1+4z2^4))^b3,
         v2 -> (4z2^2)/(1+4z2^4)
        };
SUBS = Join[SUBS1, SUBS2];
kappa = delta^b2 *
        (b2 v1(psi-(v1/(1-v1))^chi)
           +2 b3 K2 CC[b3]((v2^2)/(1+v2))(psi-(v2/(1-v2))^chi))/
        (psi - (1-v1)^(b2 chi)(1-v2^2)^(b3 chi))/. SUBS

delta0=9789/10000; psi0=13/10; KK2=1069/1000; d0=5; alpha=1-10^(-4);
CC[0]=0; CC[1]=1; CC[2]=102/100; CC[3]=103/100; CC[4]=104/100; CC[5]=105/100;
tau[0, 0] = 0; tau[0, 1] = 42/100; tau[1, 0] = 42/100;
tau[0, 2] = 54/100; tau[1, 1] = 59/100; tau[2, 0] = 63/100;
tau[0, 3] = 72/100; tau[1, 2] = 74/100; tau[2, 1] = 76/100; tau[3, 0] = 79/100;
tau[0, 4] = 864/1000; tau[1, 3] = 868/1000; tau[2, 2] = 876/1000;
tau[3, 1] = 886/1000; tau[4, 0] = 901/1000;
For[bb2 = 0, bb2 <= d0, bb2++,
 tau[bb2, d0 - bb2] = alpha;
 ]

kappa = kappa/.{d->d0, delta->delta0, psi->psi0, K2->KK2, M->M0};
z1up = 2^(1/2) - 1; z2up = 1/2(3^(1/2)-1);
For[bb2 = 0, bb2 <= d0, bb2++,
  For[bb3 = 0, bb3 <= d0 - bb2, bb3++,
    Print["b2=", bb2, ", b3=", bb3];
    EXP = kappa /. {b2 -> bb2, b3 -> bb3};
    Print[
     Resolve[Exists[{z1, z2},
       EXP > tau[bb2,bb3] && 0 <= z1 <= z1up && 0 <= z2 <= z2up]]];
    ];
```

```
  ];
```

8.2.6. Mathematica code for proof of Lemma 5.9. (** Verification of (5.28) **)

```
kappa = (tau[b2, b3] (psi-A) +
       3* delta^b2 * K3 * b4 * (v3^3)/(1+v3+v3^2)*(psi-(v3/(1-v3))^chi))/
        (psi - A(1-v3^3)^(b4*chi))

Delta=6; dd0=Delta-1; delta0=9789/10000; psi0=13/10; chi0=1/2; alpha=1-10^(-4);
tau[0, 0]=0; tau[0, 1]=42/100; tau[1, 0]=42/100;
tau[0, 2]=54/100; tau[1, 1]=59/100; tau[2, 0]=63/100;
tau[0, 3]=72/100; tau[1, 2]=74/100; tau[2, 1]=76/100; tau[3, 0]=79/100;
tau[0, 4]=864/1000; tau[1, 3]=868/1000; tau[2, 2]=876/1000;
tau[3, 1]=886/1000; tau[4, 0]=901/1000;
tau[5, 0]=alpha;

kappa = kappa/.{delta->delta0, psi->psi0, chi->chi0, K3->1160/1000}
For[bb2 = 0, bb2 <= dd0, bb2++,
  For[bb3 = 1, bb3 <= dd0 - bb2, bb3++,
    For[bb4 = 1, bb4 <= dd0 - bb2 - bb3, bb4++,
      Print["Running for b2=", bb2, " b3=", bb3, " and b4=", bb4];
      EXP = kappa /. {b2 -> bb2, b3 -> bb3, b4 -> bb4};
      B = bb2 + bb3 + bb4;
      Print[
       Resolve[Exists[{A, v3},
         EXP > tau[B, 0] && 0 <= v3 <= 1/2  && 0 <= A <= 1]]
        ];
      ];
    ];
  ];
```

(** Verification of (5.30) **)

```
SUBS = {(v1/(1 - v1))^chi->u1, v1->(u1^2)/(1+u1^2), (v3/(1-v3))^chi->u3,
   v3->(u3^2)/(1+u3^2)};
kappa = delta^b2 *(  b2 v1 (psi - (v1/(1 - v1))^chi)
       +3*K3*b4*(v3^3)/(1+v3+v3^2)*(psi - (v3/(1 - v3))^chi))/
       (psi - (1-v1)^(b2*chi) (1-v3^3)^(b4*chi)) /. SUBS

Delta=6; dd0=Delta-1; delta0=9789/10000; psi0=13/10; chi0=1/2; alpha= 1-10^(-4);
tau[1,0]=42/100; tau[2,0]=63/100; tau[3,0]=79/100; tau[4,0]=901/1000; tau[5,0]=alpha;

kappa = kappa /.{delta->delta0, psi->psi0, chi->chi0, K3->1160/1000}
For[bb2 = 0, bb2 <= dd0, bb2++,
  For[bb4 = 1, bb4 <= dd0 - bb2, bb4++,
    Print["Running for b2=", bb2, " and b4=", bb4];
    EXP = kappa/.{b2 -> bb2, b3 -> bb3, b4 -> bb4};
    B = bb2 + bb4;
    Print[
     Resolve[Exists[{u1, u3},
       EXP > tau[B, 0] && 0 <= u1 <= 1 && 0 <= u3 <= 1]]];
    ];
  ];
```

8.2.7. Mathematica code for proof of Lemma 5.10. The following code can be executed by copying it in a Mathematica cell.

(** Verification of (5.31) **)

```
kappa = (tau[Bp, 0] (psi-A) +
       4 * K4 * b5 * (v4^4)/(1+v4+v4^2+v4^3)*(psi-(v4/(1-v4))^chi))/
```

```
       (psi - A(1-v4^4)^(b5*chi))

Delta=6; dd0=Delta-1; delta0=9789/10000; psi0=13/10; chi0=1/2; alpha=1-10^(-4);
tau[0,0]=0; tau[1,0]=42/100; tau[2,0]=63/100; tau[3,0]=79/100; tau[4,0]=901/1000;
tau[5,0]=alpha;

kappa = kappa/.{delta->delta0, psi->psi0, chi->chi0, K4->1225/1000}
For[BBp = 0, BBp <= dd0, BBp++,
  For[bb5 = 1, bb5 <= dd0 - BBp, bb5++,
      Print["Running for Bp=", BBp, " and b5=", bb5];
      EXP = kappa /. {Bp -> BBp, b5 -> bb5};
      B = BBp + bb5;
      Print[
       Resolve[Exists[{A, v4},
         EXP > tau[B, 0] && 0 <= v4 <= 1/2  && 0 <= A <= 1]]
       ];
      ];
    ];
```

8.2.8. Mathematica code for proof of Lemma 5.11. (** Verification of (5.32) **)

```
kappa = (tau[Bp, 0] (psi-A) +
       5 * K5 * b6 * (v5^5)/(1+v5+v5^2+v5^3+v5^4)*(psi-(v5/(1-v5))^chi))/
        (psi - A(1-v5^5)^(b6*chi))

Delta=6; dd0=Delta-1; delta0=9789/10000; psi0=13/10; chi0=1/2; alpha=1-10^(-4);
tau[0,0]=0; tau[1,0]=42/100; tau[2,0]=63/100; tau[3,0]=79/100; tau[4,0]=901/1000;
tau[5,0]=alpha;
Resolve[(1/delta0)^(4*dd0)>1532/1000]

kappa = kappa/.{delta->delta0, psi->psi0, chi->chi0, K5->1532/1000}
For[BBp = 0, BBp <= dd0, BBp++,
  For[bb6 = 1, bb6 <= dd0 - BBp, bb6++,
      Print["Running for Bp=", BBp, " and b6=", bb6];
      EXP = kappa /. {Bp -> BBp, b6 -> bb6};
      B = BBp + bb6;
      Print[
       Resolve[Exists[{A, v5},
         EXP > tau[B, 0] && 0 <= v5 <= 1/2  && 0 <= A <= 1]]
       ];
      ];
    ];
```

8.2.9. Mathematica code for proof of Lemma 5.12. (** Verification of (5.33) **)

```
f = (1 - t) (psi - (t/(1 - t))^chi) /. {psi->13/10, chi->1/2, t -> Exp[y]}
f2 = FullSimplify[D[f, {y, 2}]]
Resolve[Exists[y, f2 > 0 && y <= Log[1/2]]]
```

8.2.10. Mathematica code for proof of Lemma 5.13. (** Verification of (5.48) and (5.49) **)

```
psi=13/10; delta=9789/10000; K2=1069/1000; K4=1225/1000;
AA[1]={A1->1, A2->1/delta^5, A->2K2};
AA[2]={A1->2, A2->2K2/delta^5, A->4*1120/1000};
AA[3]={A1->1, A2->1/delta^15, A->5/2};
AA[4]={A1->2K2/delta^5, A2->5/2, A->4K4};

LHS = (4 x1 x2)/((1 + x1^2) (1 + x2^2) - 4 x1 x2);
RHS = psi - (A1*(1 - x1^2)^2/(1 + x1^2)^2 (psi - (2 x1)/(1 - x1^2)) +
      A2*(1 - x2^2)^2/(1 + x2^2)^2 (psi - (2 x2)/(1 - x2^2)))/
      (A (1 - (4 x1 x2)/((1 + x1^2) (1 + x2^2))))
```

```
EXPR = RHS^2 - LHS

xup = Sqrt[2] - 1;
For[i=1, i<= 4, i++,
    Print["Running for A1=", A1/.AA[i], " A2=", A2/.AA[i], " A=", A/.AA[i]];
    RHS0=RHS/.AA[i]; EXPR0 = EXPR /. AA[i];
    Print[Resolve[Exists[{x1, x2},
            RHS0 < 0 && 0 <= x1 <= xup && 0 <= x2 <= xup]]
    ];
    Print[Resolve[Exists[{x1, x2},
            EXPR0 < 0 && 0 <= x1 <= xup && 0 <= x2 <= xup]]
    ];
];
```

8.2.11. Mathematica code for proof of Lemma 5.14.

```
(** Verification of (5.67) **)
q = 27/25; Delta = 6;
h = (1-t) (psi - (t/(1 - t))^chi) /. {psi->13/10, chi->1/2, t->(1-y)^(1/2)};
g2 = ((1-y)/y) * h;
FUN = Simplify[(q - 1) y*(D[g2, y])^2 + g2 * ((D[g2, y]) + y*(D[g2, {y, 2}]))];
Resolve[Exists[y, FUN > 0 && 3/4 <= y <= 1 - 1/(2^(Delta-1) + 1)^2]]
```

8.3. Remaining proofs.

8.3.1. Proof of Theorem 6.5.

```
(** Verification that (6.2) has multiple solutions **)
Delta = 17;
lambda = 1/2; beta = 1/2; gamma = 1;
Resolve[Exists[{x, y},
  x == lambda ((beta y + 1)/(y + gamma))^(Delta - 1) &&
   y == lambda ((beta x + 1)/(x + gamma))^(Delta - 1) && x > 0 &&
   y > 0 && x > y]]
```

Acknowledgments. This work was done in part while the authors were visiting the Simons Institute for the Theory of Computing. We thank the anonymous referees for their valuable comments.

REFERENCES

[1] M. Bordewich, M. E. Dyer, and M. Karpinski, *Path coupling using stopping times and counting independent sets and colorings in hypergraphs*, Random Struct. Algorithms, 32 (2008), pp. 375–399.

[2] M. Bordewich, M. E. Dyer, and M. Karpinski, *Stopping times, metrics and approximate counting*, in Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP), Lecture Notes in Comput. Sci. 4051, Springer, Berlin, 2006, pp. 108–119.

[3] M. Dyer and C. Greenhill, *On Markov chains for independent sets*, J. Algorithms, 35 (2000), pp. 17–49.

[4] A. Galanis, Q. Ge, D. Štefankovič, E. Vigoda, and L. Yang, *Improved inapproximability results for counting independent sets in the hard-core model*, Random Struct. Algorithms, 45 (2014), pp. 78–110.

[5] A. Galanis and L. A. Goldberg, *The complexity of approximately counting in 2-spin systems on k-uniform bounded-degree hypergraphs*, Inform. Comput., 251 (2016), pp. 36–66; full version available at https://arxiv.org/abs/1505.06146.

[6] A. Galanis, D. Štefankovič, and E. Vigoda, *Inapproximability of the partition function for the antiferromagnetic Ising and hard-core models*, Combin. Probab. Comput., 25 (2016), pp. 500–559.

[7] L. A. Goldberg, R. Gysel, and J. Lapinskas, *Approximately counting locally-optimal structures*, J. Comput. System Sci., 82 (2016), pp. 1144–1160.
[8] J. Hermon, A. Sly, and Y. Zhang, *Rapid mixing of hypergraph independent set*, Random Struct. Algorithms, to appear; full version available at https://arxiv.org/abs/1610.07999.
[9] M. Jerrum and A. Sinclair, *Polynomial-time approximation algorithms for the Ising model*, SIAM J. Comput., 22 (1993), pp. 1087–1116, https://doi.org/10.1137/0222066.
[10] L. Li, P. Lu, and Y. Yin, *Approximate counting via correlation decay in spin systems*, in Proceedings of the 23rd Annual ACM–SIAM Symposium on Discrete Algorithms (SODA), ACM, New York, SIAM, Philadelphia, 2012, pp. 922–940, https://doi.org/10.1137/1.9781611973099.74.
[11] L. Li, P. Lu, and Y. Yin, *Correlation decay up to uniqueness in spin systems*, in Proceedings of the 24th Annual ACM–SIAM Symposium on Discrete Algorithms (SODA), ACM, New York, SIAM, Philadelphia, 2013, pp. 67–84; full version available at https://arxiv.org/abs/1111.7064.
[12] J. Liu and P. Lu, *FPTAS for counting monotone CNF*, in Proceedings of the 26th Annual ACM–SIAM Symposium on Discrete Algorithms (SODA), ACM, New York, SIAM, Philadelphia, 2015, pp. 1531–1548; full version available at https://arxiv.org/abs/1311.3728.
[13] P. Lu, K. Yang, and C. Zhang, *FPTAS for hardcore and Ising models on hypergraphs*, in Proceedings of the 33rd Symposium on Theoretical Aspects of Computer Science (STACS), 2016, 51; full version available at https://arxiv.org/abs/1509.05494.
[14] M. Luby and E. Vigoda, *Fast convergence of the Glauber dynamics for sampling independent sets*, Random Struct. Algorithms, 15 (1999), pp. 229–241.
[15] A. Moitra, *Approximate counting, the Lovász local lemma, and inference in graphical models*, J. ACM, to appear; full version available at https://arxiv.org/abs/1610.04317.
[16] E. Mossel, D. Weitz, and N. Wormald, *On the hardness of sampling independent sets beyond the tree threshold*, Probab. Theory Related Fields, 143 (2009), pp. 401–439.
[17] A. Sinclair, P. Srivastava, D. Štefankovič, and Y. Yin, *Spatial mixing and the connective constant: Optimal bounds*, in Proceedings of the 26th Annual ACM–SIAM Symposium on Discrete Algorithms (SODA), ACM, New York, SIAM, Philadelphia, 2015, pp. 1549–1563, https://doi.org/10.1137/1.9781611973730.102.
[18] A. Sinclair, P. Srivastava, and M. Thurley, *Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs*, in Proceedings of the 23rd Annual ACM–SIAM Symposium on Discrete Algorithms (SODA), ACM, New York, SIAM, Philadelphia, 2012, pp. 941–953, https://doi.org/10.1137/1.9781611973099.75.
[19] A. Sly, *Computational transition at the uniqueness threshold*, in Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS), IEEE, Washington, DC, 2010, pp. 287–296.
[20] A. Sly and N. Sun, *Counting in two-spin models on d-regular graphs*, Ann. Probab., 42 (2014), pp. 2383–2416.
[21] A. W. Strzebonski, *Cylindrical algebraic decomposition using validated numerics*, J. Symbolic Comput., 41 (2006), pp. 1021–1038.
[22] D. Weitz, *Counting independent sets up to the tree threshold*, in Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC), ACM, New York, 2006, pp. 140–149.
[23] Y. Yin and C. Zhang, *Sampling in Potts model on sparse random graphs*, in APPROX-RANDOM, Paris, France, 2016, 47.