Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 68 (2022)

Sketching Approximability of (Weak) Monarchy Predicates

Chi-Ning Chou[*] Alexander Golovnev[†] Amirbehshad Shahrasbi[‡] Madhu Sudan[§]
Santhoshini Velusamy[¶]

Abstract

We analyze the sketching approximability of constraint satisfaction problems on Boolean domains, where the constraints are balanced linear threshold functions applied to literals. In particular, we explore the approximability of monarchy-like functions where the value of the function is determined by a weighted combination of the vote of the first variable (the president) and the sum of the votes of all remaining variables. The pure version of this function is when the president can only be overruled by when all remaining variables agree. For every $k \geq 5$, we show that CSPs where the underlying predicate is a pure monarchy function on k variables have no non-trivial sketching approximation algorithm in $o(\sqrt{n})$ space. We also show infinitely many weaker monarchy functions for which CSPs using such constraints are non-trivially approximable by $O(\log(n))$ space sketching algorithms. Moreover, we give the first example of sketching approximable asymmetric Boolean CSPs. Our results work within the framework of Chou, Golovnev, Sudan, and Velusamy (FOCS 2021) that characterizes the sketching approximability of all CSPs. Their framework can be applied naturally to get a computer-aided analysis of the approximability of any specific constraint satisfaction problem. The novelty of our work is in using their work to get an analysis that applies to *infinitely* many problems simultaneously.

[*]School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, USA. Partially supported by NSF grants DMS-2134157 and CCF-1565264, DOE grant DE-SC0022199, and the Simons foundation. Email: `chiningchou@g.harvard.edu`.

[†]Department of Computer Science, Georgetown University. Email: `alexgolovnev@gmail.com`.

[‡]Microsoft, USA. Email: `ashahrasbi@microsoft.com`. The author was with Harvard University and supported by CRA-CCC Computing Innovations Fellowship (CIFellowship 2020) during this work.

[§]School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, USA. Supported in part by a Simons Investigator Award and NSF Awards CCF 1715187 and CCF 2152413. Email: `madhu@cs.harvard.edu`.

[¶]School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, USA. Supported in part by a Google Ph.D. Fellowship, a Simons Investigator Award to Madhu Sudan, and NSF Awards CCF 1715187 and CCF 2152413. Email: `svelusamy@g.harvard.edu`.

ISSN 1433-8092

1 Introduction

In this paper we consider the sketching complexity of solving constraint satisfaction problems (CSPs) approximately where the constraints are given by linear threshold functions over a collection of Boolean literals. We introduce these terms below.

CSPs: Given a Boolean function $f : \{-1,1\}^k \to \{0,1\}$, the Boolean CSP associated with f, denoted $\mathsf{Max\text{-}CSP}(f)$ is the following optimization problem. Given m constraints $C_1, \ldots, C_m$ on n Boolean variables $X_1, \ldots, X_n$, where each constraint applies f to a sequence of k distinct literals from the set $\{X_1, \ldots, X_n, -X_1, \ldots, -X_n\}$, find the maximum fraction of constraints that can be satisfied by an assignment to the n variables. For an instance Ψ of $\mathsf{Max\text{-}CSP}(f)$ we use val_Ψ to denote this maximum value. We are interested in approximating val_Ψ and this task is known to be equivalent to solving a gapped decision version of $\mathsf{Max\text{-}CSP}(f)$. For $0 \leq \beta < \gamma \leq 1$ we define the (γ, β)-gapped version of $\mathsf{Max\text{-}CSP}(f)$, abbreviated to (γ, β)-$\mathsf{Max\text{-}CSP}(f)$, to be the following promise decision problem: Given an instance Ψ satisfying $\mathsf{val}_\Psi \geq \gamma$ or $\mathsf{val}_\Psi < \beta$ decide which one of the two conditions holds.

Sketching algorithms: The class of algorithms we consider (and rule out) are randomized sketching algorithms. Inputs to these algorithms arrive as a stream of elements, in our case a stream of constraints. We consider algorithms that use some bounded amount of space, denoted $s(n)$, to process the stream and maintain a sketch of their output. When the stream ends the algorithm outputs it verdict based on the current sketch. A key restriction of a sketching algorithm is that its sketch should satisfy the following composability property. Given two streams σ and τ and a fixing of the randomness, the sketch of their concatenation $S(\sigma \circ \tau)$ should be determined by their sketches $S(\sigma)$ and $S(\tau)$ alone.[1] Most existing algorithms for streaming CSPs are sketching algorithms. We say a sketching algorithm solves a (gapped) decision problem if on every input its answer is correct with probability at least 2/3.

Approximability and approximation resistance: For $\alpha \in [0,1]$, we say an algorithm is an α-approximation algorithm for $\mathsf{Max\text{-}CSP}(f)$ if the following holds: on every input instance Ψ, the algorithm outputs v such that $\alpha \cdot \mathsf{val}_\Psi \leq v \leq \mathsf{val}_\Psi$ with probability at least 2/3. Note that the existence of an α-approximation algorithm is equivalent to the existence of an algorithm for solving (γ, β)-$\mathsf{Max\text{-}CSP}(f)$ for every $\gamma, \beta \in [0,1]$ with $\beta \leq \alpha \cdot \gamma$.

For a function $f : \{-1,1\}^k \to \{0,1\}$, define $\rho(f) = 2^{-k} \cdot |\{x \in \{-1,1\}^k | f(x) = 1\}|$. For every f and every instance Ψ of $\mathsf{Max\text{-}CSP}(f)$, a random assignment satisfies $\rho(f)$ fraction of the constraints in expectation and so every Ψ satisfies $\mathsf{val}_\Psi \geq \rho(f)$. Thus the $(1, \rho(f))$-$\mathsf{Max\text{-}CSP}(f)$ problem is trivially solvable by the algorithm that always outputs $\mathsf{val}_\Psi \geq 1$ (since the set $\{\Psi | \mathsf{val}_\Psi < \rho(f)\}$ is empty). We say $\mathsf{Max\text{-}CSP}(f)$ is sketching approximable within space $s(n)$ if there is an $\varepsilon > 0$ and a sketching algorithm using at most $s(n)$ space that solves $(1-\varepsilon, \rho(f)+\varepsilon)$-$\mathsf{Max\text{-}CSP}(f)$. We say that $\mathsf{Max\text{-}CSP}(f)$ is approximation resistant to space $s(n)$ if for every $\varepsilon > 0$, every sketching algorithm for $(1, \rho(f) + \varepsilon)$-$\mathsf{Max\text{-}CSP}(f)$ requires $\Omega(s(n))$ space.

[1] In contrast, a general streaming algorithm maintains a state $S(\sigma \circ \tau)$ that may depend on $S(\sigma)$ and all of τ.

1.1 Motivation and related work

There has been an increasing interest in studying the approximability of CSPs in the streaming setting [KK15, KKS15, KKSV17, GVV17, GT19, KK19, CGV20, CGSV21, CGSV22, SSV21, BHP+22, CGS+22]. In particular, recently Chou, Golovnev, Sudan, and Velusamy [CGSV21, CGSV22] gave a dichotomy result for sketching approximability of all finite CSPs. Specifically, they proved the following theorem.

Theorem 1.1 ([CGSV22]). *For every k, every predicate $f : \{-1, 1\}^k \to \{0, 1\}$ and every $0 \leq \beta < \gamma \leq 1$ one of the following holds: (1) (γ, β)-$\mathsf{Max\text{-}CSP}(f)$ is solvable by an $O(\log(n))$-space sketching algorithm, or (2) for every $\varepsilon > 0$, $(\gamma - \varepsilon, \beta + \varepsilon)$-$\mathsf{Max\text{-}CSP}(f)$ is not solvable by any $o(\sqrt{n})$-space sketching algorithm. Furthermore there is a decidable procedure that determines, given $\mathcal{F}$, γ and β, which of the two conditions hold.*

We note that a followup paper by the same authors [CGSV21] extends the result to a more general setting: Specifically they allow non-Boolean variables, allow a set of predicates rather than a single function; and allow the predicates to be applied to variables rather than literals. While their result is more general all results in this paper work in the more restricted setting of [CGSV22] and so we will describe our results in their language (which can be somewhat simpler for problems that are expressible in their setting).

While the results of [CGSV22] imply a dichotomy, to explicitly get the optimal sketching approximation ratio for a given predicate f, they need to solve an optimization problem which in general needs computer-aided analysis. In order to get more explicit results one needs to restrict the families of functions considered, and even then it is unclear if there can be a closed-form expression. In the only example we are aware of, Boyland, Hwang, Prasad, Singer, and Velusamy [BHP+22] gave closed-form expressions for the optimal sketching approximation ratio of some *symmetric* Boolean CSPs. This still leaves the question of exploring the sketching approximability of other subfamilies of CSPs and extracting some qualitative results yielding necessary or sufficient conditions for non-trivial approximability.

1.2 Main results

In this paper we study sketching approximability of CSPs on linear threshold functions. Below we define the classes of linear threshold functions and balanced linear threshold functions.

Definition 1.2 (Linear threshold function). *A linear threshold function, or LTF, is a Boolean function $f : \{-1, 1\}^k \to \{0, 1\}$ of the form*

$$f(x) = \mathsf{sign}\left(\sum_{i=1}^{k} w_i x_i + \theta\right),$$

where $w_1, \ldots, w_k, \theta \in \mathbb{R}$. The function $\mathsf{sign}(z)$ has value 1 if $z > 0$ and 0 if $z \leq 0$; $w_1, \ldots, w_k$ are called the weights of f and θ is the threshold.

Definition 1.3 (Balanced linear threshold function). *A balanced linear threshold function, or balanced LTF, is an LTF with threshold 0 and the additional restriction that for every $x \in \{-1, 1\}^k$, we have $\sum_{i=1}^{k} w_i x_i \neq 0$. Specifically, a balanced LTF f satisfies $f(-x) = 1 - f(x)$ for every x.*

Note that for a balanced LTF f, $\rho(f) = 1/2$, and the goal of approximability is to beat this factor. Balanced LTFs form a technically important class of functions to study visavis CSP approximability. For instance Potechin [Pot19] studies them in the polynomial time regime giving a (somewhat complex) approximation-resistant function in this class. In the sketching setting, interest in this class of functions comes from [CGSV22, Theorem 1.3] which shows that if a function f supports one-wise independence (i.e., f^{-1} supports a distribution on $\{-1,1\}^k$ that is uniform on each of the k marginals) then $\textsf{Max-CSP}(f)$ is approximation resistant to $o(\sqrt{n})$ space streaming algorithms. Balanced LTFs are the most basic class of functions that *do not* support one-wise independence and hence are not covered by this theorem. Studying this class thus offers the possibility of finding new classes of CSPs that are approximation resistant to $o(\sqrt{n})$-space streaming algorithms.

Our first result shows that every balanced LTF on up to 4 variables is sketching approximable. (So to search for new approximation resistant functions we need to look at functions on more variables!) We note that there are only finitely many such LTFs, but already this theorem gives the first example of an asymmetric Boolean CSP which is approximable by sketching algorithms.[2]

Theorem 1.4. *For every balanced LTF f on $k \leq 4$ variables, $\textsf{Max-CSP}(f)$ is sketching approximable in $O(\log(n))$ space.*

Our next result shows that there do exist balanced LTFs functions on 5 or more variables that are sketching approximation resistant. The specific family of functions we show this for are the "Monarchy" functions. For $k \in \mathbb{N}$, $\textsf{MON}_k : \{-1,1\}^k \to \{0,1\}$ is given by $\textsf{MON}_k(x_1, \ldots, x_k) = \textsf{sign}\left((k-2)x_1 + x_2 + \cdots + x_k\right)$. It may be easily verified that $\textsf{MON}_k$ is a balanced LTF. We have the following theorem.

Theorem 1.5. *For every $k \geq 5$, $\textsf{Max-CSP}(\textsf{MON}_k)$ is sketching approximation resistant to space $o(\sqrt{n})$.*

Thus we get the first examples of functions that do not support one-wise independence that are approximation resistant to space $o(\sqrt{n})$ sketching algorithms. In fact, the theorem gives infinitely many such examples. We suspect that the Balanced LTF constructed in [Pot19] should also be approximation-resistant but so far we don't have a proof. The monarchy functions, by virtue of the simplicity allow a simpler analytic proof, though admittedly even in this case we do not have great intuition for the proof and do not know how to extend it to other classes of functions.

Finally we also give an infinite subclass of balanced LTFs that are approximable using $O(\log(n))$ space. The functions we consider here are what we call "weak monarchy" functions.[3] For $j \leq k \in \mathbb{N}$, let $\textsf{WMON}_{k,j} : \{-1,1\}^k \to \{0,1\}$ be the function given by $\textsf{WMON}_{k,j}(x_1, \ldots, x_k) = \textsf{sign}\left(j \cdot x_1 + x_2 + \cdots + x_k\right)$. It may be easily verified that when $j + k$ is even, then $\textsf{WMON}_{k,j}$ is a balanced LTF. We have

Theorem 1.6. *For all integers $j \geq 2$ and $k \geq 7j^3$ such that $k + j$ is even, $\textsf{Max-CSP}(\textsf{WMON}_{k,j})$ is sketching approximable in $O(\log(n))$ space. In particular, for every j, there exist infinitely many k such that $\textsf{Max-CSP}(\textsf{WMON}_{k,j})$ is sketching approximable.*

The results above give the first examples of asymmetric Boolean CSPs for which $\textsf{Max-CSP}(f)$ is sketching approximable. Again we get an infinite family of such functions.

[2] Note that Max-DICUT (shown to be sketching approximable in [CGV20, CGSV21]) is not considered a Boolean CSP in [CGSV22] since the Max-DICUT constraints are applied on variables and not on literals.

[3] Such functions are also sometimes called presidential type predicates [HP20].

Comparison to the polynomial time regime. Hast [Has05] proves that (a generalization of) Theorem 5.1 holds in the polynomial time regime (thus, implying an analogue of Theorem 1.6 in the polynomial time regime). Austrin, Benabbas, and Magen [ABM10] prove that MON_k is approximable in polynomial time, which is in sharp contrast to the result of Theorem 1.5 in the sketching setting. Huang and Potechin [HP20] show that almost all WMON predicates are approximable in polynomial time. Finally, Potechin [Pot19] gives a balanced LTF which is (conditionally) approximation resistant in the polynomial time regime.

Organization of the paper. We start with giving formal definitions and stating relevant previous results in Section 2. The three main theorems are proved in Section 3, Section 4, and Section 5, respectively.

2 Preliminaries

We use $\mathbb{N}, \mathbb{R}$, and $\mathbb{R}_{\geq 0}$ to denote the sets of all natural, real, and non-negative real numbers, respectively. We use $[n]$ to denote the set $\{1, \ldots, n\}$. We write vector variables in boldface, e.g., $\mathbf{x}$, and we use x_i to denote their ith entry. For two vectors of the same length $\mathbf{x}, \mathbf{y} \in \mathbb{R}^k$, $\mathbf{x} \odot \mathbf{y} \in \mathbb{R}^k$ denotes the entry-wise product of $\mathbf{x}$ and $\mathbf{y}$. For $p \in [0, 1]$, $\mathsf{Bern}(p)$ denotes the Bernoulli distribution taking value 1 with probability p, and value -1 with probability $1 - p$. We adopt the convention that $\binom{n}{k} = 0$ for $k < 0$ or $k > n$. By $\binom{n}{\leq k}$ we denote the sum $\sum_{i=0}^{k} \binom{n}{i}$.

2.1 Sketching approximability and approximation resistance

For a function $f \colon \{-1, 1\}^k \to \{0, 1\}$, let $\rho(f) = 2^{-k} \cdot |\{\mathbf{a} \in \{-1, 1\}^k \mid f(\mathbf{a}) = 1\}|$ denote the probability that a uniformly random assignment of the variables satisfies f.

Definition 2.1 (Sketching approximation resistance)**.** *For a function $f \colon \{-1, 1\}^k \to \{0, 1\}$, we say that f is* sketching approximation resistant *to space $s(n)$ if for every $\varepsilon > 0$, every sketching algorithm for $(1, \rho(f) + \varepsilon)$-Max-CSP$(f)$ requires $\Omega(\sqrt{n})$ space.*

Definition 2.2 (Sketching approximability)**.** *For a function $f \colon \{-1, 1\}^k \to \{0, 1\}$, we say that f is* sketching approximable *in space $s(n)$ if there exist $\varepsilon > 0$ and a sketching algorithm that solves $(1 - \varepsilon, \rho(f) + \varepsilon)$-Max-CSP$(f)$ using space $s(n)$.*

At first glance, it seems that if f is not sketching approximation resistant then it's not necessarily sketching approximable. Nonetheless, [CGSV22] proved that every f is either approximable or approximation resistant.[4]

2.2 Characterization of approximability from [CGSV22]

In this work, we focus on CSPs that use a single function f applied to literals. Thus, we will use the machinery from [CGSV22] instead of the more general (and more notationally-heavy) version in [CGSV21]. For a distribution $\mathcal{D} \in \Delta(\{-1, 1\}^k)$, by $\boldsymbol{\mu}(\mathcal{D})$ we denote its marginals, i.e., $\boldsymbol{\mu}(\mathcal{D}) = (\mu_1, \ldots, \mu_k)$ where $\mu_i = \mathbb{E}_{\mathbf{b} \sim \mathcal{D}}[b_i]$ for all $i \in [k]$.

[4] Concretely, as the sets K^Y, K^N are closed (see Lemma 2.4), an algorithm for $(1, \rho(f) + \varepsilon)$-Max-CSP$(f)$ implies an algorithm for $(1 - \varepsilon', \rho(f) + \varepsilon)$-Max-CSP$(f)$ for some $\varepsilon' > 0$, which in turn implies that Max-CSP(f) is approximable.

Definition 2.3 ([CGSV22, Definitions 2.1 and 2.2])**.** *For $\gamma, \beta \in \mathbb{R}$, we define the sets of distributions S_γ^Y and S_β^N as*

$$S_\gamma^Y = S_\gamma^Y(f) = \{\mathcal{D}_Y \in \Delta(\{-1,1\}^k) \mid \mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}_Y}[f(\mathbf{b})] \geq \gamma\}$$

and

$$S_\beta^N = S_\beta^N(f) = \{\mathcal{D}_N \in \Delta(\{-1,1\}^k) \mid \mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}_N} \mathop{\mathbb{E}}_{\mathbf{a}\sim\mathsf{Bern}(p)^k}[f(\mathbf{b}\odot\mathbf{a})] \leq \beta, \forall p \in [0,1]\},$$

and the sets of marginals of these distributions

$$K_\gamma^Y = K_\gamma^Y(f) = \{\ \boldsymbol{\mu}(\mathcal{D}_Y) \mid \mathcal{D}_Y \in S_\gamma^Y\}$$

and

$$K_\beta^N = K_\beta^N(f) = \{\ \boldsymbol{\mu}(\mathcal{D}_N) \mid \mathcal{D}_N \in S_\beta^N\}.$$

We will use the following properties of the sets K_γ^Y and K_β^N.

Lemma 2.4 ([CGSV22, Lemma 2.4])**.** *For every $\gamma, \beta \in [0,1]$ the sets K_γ^N and K_β^Y are bounded, closed and convex.*

With these definitions, we are ready to present the approximability criteria from [CGSV22].[5]

Theorem 2.5 ([CGSV22, Corollary 1.2])**.** *For every $k \in \mathbb{N}$ and every function $f : \{-1,1\}^k \to \{0,1\}$, if $K_1^Y(f) \cap K_{\rho(f)}^N(f) = \emptyset$, then f is sketching approximable within space $O(\log(n))$, if $K_1^Y(f) \cap K_{\rho(f)}^N(f) \neq \emptyset$, then f is sketching approximation resistant to space $o(\sqrt{n})$.*

2.3 (Weak) Monarchy functions

Definition 2.6. *A monarchy predicate on $k \geq 2$ variables $\mathsf{MON}_k \colon \{-1,1\}^k \to \{0,1\}$ is defined as*

$$\mathsf{MON}_k(x_1, \ldots, x_k) = \mathsf{sign}\left((k-2)x_1 + \sum_{i=2}^k x_i\right).$$

Here x_1 is commonly referred to as the president and the rest of x_is are called citizens.

Definition 2.7 (Weak monarchy functions)**.** *A weak monarchy predicate of order j on $k \geq 2$ variables $\mathsf{WMON}_{k,j} \colon \{-1,1\}^k \to \{0,1\}$ is defined as*

$$\mathsf{WMON}_{k,j}(x_1, \ldots, x_k) = \mathsf{sign}\left(j \cdot x_1 + \sum_{i=2}^k x_i\right).$$

Similar to ordinary monarchy functions, x_1 is commonly referred to as the president and the rest of x_is are called citizens.

It is straightforward to see that MON_k is a balanced LTF for every $k \geq 2$ and $\mathsf{WMON}_{k,j}$ is a balanced LTF whenever $k + j$ is even.

[5] Strictly speaking the statement in Corollary 1.2 in [CGSV22] is somewhat different, but their proof of Corollary 1.2 asserts this explicitly.

2.4 Fourier analysis of Boolean functions

We will need the following basic notions from Fourier analysis over the Boolean hypercube (see, for instance, [O'D14]).

Definition 2.8 (Characteristic functions)**.** *For every $S \subseteq [k]$ such that $|S| \geq 1$, the characteristic function $\chi_S : \{-1,1\}^k \to \{-1,1\}$ is defined as $\chi_S(x) = \prod_{i \in S} x_i$. The characteristic function corresponding to the empty set is defined as the constant function $\chi_\emptyset(x) = 1$ for all $x \in \{-1,1\}^k$.*

Definition 2.9 (Fourier expansions)**.** *The Fourier expansion of a Boolean function $f : \{-1,1\}^k \to \{0,1\}$ is given by*

$$f = \sum_{S \subseteq [k]} \widehat{f}(S) \cdot \chi_S \,,$$

where $\widehat{f}(S) = \mathbb{E}_{x \sim \mathsf{Unif}\{-1,1\}^k}[f(x) \cdot \chi_S(x)]$ and $\mathsf{Unif}(\{-1,1\}^k)$ denotes the uniform distribution on $\{-1,1\}^k$.

Definition 2.10 (Chow parameters)**.** *The Chow parameters of a Boolean function $f : \{-1,1\}^k \to \{0,1\}$ are the degree-0 Fourier coefficient and the k degree-1 Fourier coefficients of f, i.e., $\widehat{f}(\emptyset), \widehat{f}(\{1\}), \ldots, \widehat{f}(\{k\})$.*

Proposition 2.11. *For every Boolean function $f : \{-1,1\}^k \to \{0,1\}$,*

1. *$\rho(f) = \widehat{f}(\emptyset)$,*
2. *for every $S \subseteq [k]$, $|\widehat{f}(S)| \leq \widehat{f}(\emptyset)$, and*
3. *for every $x \in \{-1,1\}^k$, $-\widehat{f}(\emptyset) \cdot k \leq \sum_{i=1}^k \widehat{f}(\{i\}) \cdot x_i \leq \widehat{f}(\emptyset) \cdot k$.*

Proof. The first statement of the proposition follows directly from the definition of $\rho(f)$: $\rho(f) = \mathbb{E}_{x \sim \mathsf{Unif}(\{-1,1\}^k)}[f(x)] = \widehat{f}(\emptyset)$. For the second statement, observe that for all $S \subseteq [k]$,

$$\begin{aligned} |\widehat{f}(S)| &= |\mathbb{E}_{x \sim \mathsf{Unif}(\{-1,1\}^k)}[f(x) \cdot \chi_S(x)]| \\ &\leq \mathbb{E}_{x \sim \mathsf{Unif}(\{-1,1\}^k)}[|f(x) \cdot \chi_S(x)|] \\ &= \mathbb{E}_{x \sim \mathsf{Unif}(\{-1,1\}^k)}[f(x)] \\ &= \widehat{f}(\emptyset) \,. \end{aligned}$$

It immediately follows that for all $x \in \{-1,1\}^k$,

$$\left| \sum_{i=1}^k \widehat{f}(\{i\}) \cdot x_i \right| \leq \sum_{i=1}^k |\widehat{f}(\{i\}) \cdot x_i| \leq \widehat{f}(\emptyset) \cdot k \,.$$

□

3 Approximability of Balanced LTFs on 4 variables

In this section, we show that all balanced LTFs on at most 4 variables are sketching approximable in $O(\log(n))$ space. We start by proving that $\mathsf{Max\text{-}CSP}(\mathsf{MON}_4)$ is approximable.

3.1 Approximability of MON_4

Recall that by Theorem 2.5, it suffices to show that $K^Y_1(\mathsf{MON}_4) \cap K^N_{1/2}(\mathsf{MON}_4) = \emptyset$. For $k \geq 2$, the inputs $x_2, \dots, x_k$ are symmetric, and we will only consider distributions $\mathcal{D} \in \Delta(\{-1,1\}^k)$ where all vectors having the same sum of coordinates and the same value in the first coordinate have the same probability masses. Concretely, for $\mathbf{x}, \mathbf{y} \in \{-1,1\}^k$, if $x_1 = y_1$ and $\sum_i x_i = \sum_i y_i$, then $\mathcal{D}(x) = \mathcal{D}(y)$. Such a distribution $\mathcal{D}$ is uniquely specified by a pair of vectors $\mathbf{u} = (u_0, \dots, u_{k-1}), \mathbf{v} = (v_0, \dots, v_{k-1}) \in \mathbb{R}^k_{\geq 0}$ with $\sum_i u_i + v_i = 1$, where for $0 \leq i \leq k-1$,

$$
\begin{aligned}
u_i &= \Pr\{x_1 = 1 \quad \text{and exactly } i \text{ of the rest of } x_i\text{s are } 1\} \,, \\
v_i &= \Pr\{x_1 = -1 \text{ and exactly } i \text{ of the rest of } x_i\text{s are } 1\} \,.
\end{aligned}
$$

Note that when $\sum_i u_i + v_i = 1$, $\mathbf{u}, \mathbf{v}$ define a distribution $\mathcal{D}$ with marginals $\boldsymbol{\mu}(\mathcal{D}) = (\mu_1, \mu', \dots, \mu')$ where

$$
\mu_1 = \sum_{i=0}^{k-1} (u_i - v_i) \text{ and } \mu' = \sum_{i=0}^{k-1} (\frac{2i}{k-1} - 1)(u_i + v_i) \,. \tag{3.1}
$$

Next we show that for MON_k functions, restricting our attention to this class of distributions is without loss of generality.

Definition 3.2. *For $\gamma, \beta \in \mathbb{R}$ and $k \geq 2$,*

$$
\begin{aligned}
\widetilde{K}^Y_\gamma(\mathit{MON}_k) &= \{\ (\mu_1, \mu') \mid (\mu_1, \mu', \dots, \mu') \in K^Y_\gamma(\mathit{MON}_k)\} \\
\text{and } \widetilde{K}^N_\beta(\mathit{MON}_k) &= \{\ (\mu_1, \mu') \mid (\mu_1, \mu', \dots, \mu') \in K^N_\beta(\mathit{MON}_k)\} \,.
\end{aligned}
$$

Lemma 3.3. *For $\gamma, \beta \in \mathbb{R}$ and $k \geq 2$,*

$$
K^Y_\gamma(\mathit{MON}_k) \cap K^N_\beta(\mathit{MON}_k) = \emptyset \text{ if and only if } \widetilde{K}^Y_\gamma(\mathit{MON}_k) \cap \widetilde{K}^N_\beta(\mathit{MON}_k) = \emptyset \,.
$$

Proof. First, if $(\mu_1, \mu', \dots, \mu') \in \widetilde{K}^Y_\gamma(\mathsf{MON}_k) \cap \widetilde{K}^N_\beta(\mathsf{MON}_k)$, then by Definition 3.2, $(\mu_1, \mu', \dots, \mu') \in K^Y_\gamma(\mathsf{MON}_k) \cap K^N_\beta(\mathsf{MON}_k)$.

For the other direction. Assume that there is a vector $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_k) \in K^Y_\gamma(\mathsf{MON}_k) \cap K^N_\beta(\mathsf{MON}_k)$. Consider two distribution $\mathcal{D}_Y \in S^Y_\gamma$ and $\mathcal{D}_N \in S^N_\beta$ yielding the vector $\boldsymbol{\mu} = \boldsymbol{\mu}(\mathcal{D}_Y) = \boldsymbol{\mu}(\mathcal{D}_N)$. Given that the variables $x_2, \cdots, x_k$ are symmetric, any distribution that is yielded by permuting $x_2, \cdots, x_k$ in $\mathcal{D}_Y$ (or $\mathcal{D}_N$) is also in S^Y_γ (or S^N_β). Note that the marginals of these distributions are also permutations of $\boldsymbol{\mu}$. By Lemma 2.4, K^Y_γ and K^N_β are convex, so they also contain the averages of these vectors: $(\mu_1, \mu', \dots, \mu') \in K^Y_\gamma(\mathsf{MON}_k) \cap K^N_\beta(\mathsf{MON}_k)$ for $\mu' = (\mu_2 + \dots + \mu_k)/(k-1)$. Finally, by Definition 3.2, $(\mu_1, \mu') \in \widetilde{K}^Y_\gamma(\mathsf{MON}_k) \cap \widetilde{K}^N_\beta(\mathsf{MON}_k)$. □

Next, we characterize the set $\widetilde{K}^Y_1(\mathsf{MON}_k)$.

Lemma 3.4. *For every $k \geq 2$, $\widetilde{K}^Y_1(\mathit{MON}_k) = \{(\mu_1, \mu') \in [-1,1]^2 \colon \mu_1(k-2) + \mu'(k-1) \geq 1\}$.*

Proof. For $\mu_1, \mu' \in [-1,1]$ satisfying $\mu_1(k-2) + \mu'(k-1) \geq 1$, consider the distribution $\mathcal{D}_Y$ given by $u_1 = \frac{(k-1)(1-\mu')}{2(k-2)}, u_{k-1} = \frac{(k-1)\mu' + (k-2)\mu_1 - 1}{2(k-2)}, v_{k-1} = (1-\mu_1)/2$, and $u_i = 0$ for $i \notin \{1, k-1\}$

and $v_j = 0$ for $j \neq k-1$. Note that $u_1, v_{k-1} \geq 0$ from $\mu_1, \mu' \in [-1,1]$, and $u_{k-1} \geq 0$ from $\mu_1(k-2) + \mu'(k-1) \geq 1$. It is also easy to check that $u_1 + u_{k-1} + v_{k-1} = 1$ which implies that $\mathcal{D}_Y$ is a distribution, and that it is supported on the preimages of 1 under MON_k. Therefore $(\mu_1, \mu') \in \widetilde{K}_1^Y(\mathsf{MON}_k)$.

For the other direction, a distribution $\mathcal{D}_Y$ supported on the preimages of 1 under MON_k satisfies $u_1 + \ldots + u_{k-1} + v_{k-1} = 1$. Then, from (3.1),

$$\begin{aligned}
\mu_1(k-2) + \mu'(k-1) &= (k-2)\sum_{i=0}^{k-1}(u_i - v_i) + \sum_{i=0}^{k-1}(2i-k+1)(u_i+v_i) \\
&= \sum_{i=1}^{k-1}(2i-1)u_i + v_{k-1} \\
&\geq \sum_{i=1}^{k-1} u_i + v_{k-1} = 1 \,,
\end{aligned}$$

where the second equality uses that $u_0 = 0$ and $v_j = 0$ for $j < k-1$. This concludes the proof of the lemma. □

Now we show that for the MON_4 function, $\widetilde{K}_1^Y$ and $\widetilde{K}_{1/2}^N$ are disjoint, and, thus, MON_4 is approximable in $O(\log(n))$ space.

Lemma 3.5. *Max-CSP(MON_4) is sketching approximable in $O(\log(n))$ space.*

Proof. Note that Lemma 3.4 gives that $\widetilde{K}_1^Y(\mathsf{MON}_4) = \{(\mu_1, \mu') \in [-1,1]^2 \colon 2\mu_1 + 3\mu' \geq 1\}$. We show that $\widetilde{K}_1^Y$ and $\widetilde{K}_{1/2}^N$ are disjoint, and then Lemma 3.3 and Theorem 2.5 imply that Max-CSP(MON_4) is sketching approximable in space $O(\log(n))$. Next, we prove that no distribution $\mathcal{D} \in S_{1/2}^N$ has marginals that lie in $\widetilde{K}_1^Y$.

We start by characterizing $K_{1/2}^N$ (for general MON_k). Take a distribution $\mathcal{D} \in \Delta(\{-1,1\}^k)$. In order for $\mathcal{D}$ to lie within $S_{1/2}^N$, the following needs to be satisfied:

$$\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}_N} \mathop{\mathbb{E}}_{\mathbf{a}\sim\mathsf{Bern}(p)^k}[f(\mathbf{b}\odot\mathbf{a})] \leq \beta, \forall p \,. \tag{3.6}$$

Let the function $h_{\mathcal{D}}(p)$ denote the probability of an assignment from $\mathcal{D}$ that has undergone bit flips with respect to $\mathsf{Bern}(p)^k$ to satisfy the monarchy predicate with the probability of $\beta = 1/2$ or less. With this definition, $\mathcal{D} \in S_{1/2}^N$ if and only if $h_{\mathcal{D}}(p) \leq \frac{1}{2}$ for all $0 \leq p \leq 1$. Note that negating all variables x_i flips the output of the monarchy predicate. Therefore, the negation of a "true" assignment is "false" and vice versa. This gives that $h_{\mathcal{D}}(p) = 1 - h_{\mathcal{D}}(1-p)$ for all $0 \leq p \leq 1$ which implies that $\mathcal{D} \in S_{1/2}^N$ if and only if for all $0 \leq p \leq 1$

$$h_{\mathcal{D}}(p) = \frac{1}{2} \,.$$

We now write down the coefficients of the polynomial $h_{\mathcal{D}}(p)$ in terms of u_i and v_i describing the distribution (as used earlier in this section).

If one draws an assignment from $\mathcal{D}$ where $x_1 = 1$ and exactly i of the rest of the variables are 1, the probability of the resulting assignment satisfying the monarchy predicate after the Bernoulli flipping is

$$p(1-(1-p)^i p^{k-1-i}) + (1-p)^{k-i} p^i \ .$$

Similarly, if $x_1 = -1$ and exactly i of the rest of the variables are 1, the probability of the resulting assignment satisfying the monarchy predicate after the Bernoulli flipping is

$$(1-p)(1-(1-p)^i p^{k-1-i}) + (1-p)^{k-1-i} p^{i+1} \ .$$

This gives that

$$\begin{aligned} h_{\mathcal{D}}(p) &= \sum_{i=0}^{k-1} u_i \left[p(1-(1-p)^i p^{k-1-i}) + (1-p)^{k-i} p^i \right] \\ &\quad + \sum_{i=0}^{k-1} v_i \left[(1-p)(1-(1-p)^i p^{k-1-i}) + (1-p)^{k-1-i} p^{i+1} \right] \end{aligned} \tag{3.7}$$

To prove this lemma, we form the polynomial $h_{\mathcal{D}}(p)$ for $k = 4$ and show that no set of u_is and v_is satisfy both $h_{\mathcal{D}}(p) = \frac{1}{2}$ and $2\mu_1 + 3\mu' \geq 1$ (where, by (3.1), $\mu_1 = \sum_{i=0}^{3}(u_i - v_i)$ and $\mu' = \sum_{i=0}^{3}(\frac{2i}{3} - 1)(u_i + v_i)$.)

$$\begin{aligned} h_{\mathcal{D}}(p) &= u_0 \left[p(1-p^3) + (1-p)^4 \right] \\ &\quad + u_1 \left[p(1-(1-p)p^2) + (1-p)^3 p \right] \\ &\quad + u_2 \left[p(1-(1-p)^2 p) + (1-p)^2 p^2 \right] \\ &\quad + u_3 \left[p(1-(1-p)^3) + (1-p)p^3 \right] \\ &\quad + v_0 \left[(1-p)(1-p^3) + (1-p)^3 p \right] \\ &\quad + v_1 \left[(1-p)(1-(1-p)p^2) + (1-p)^2 p^2 \right] \\ &\quad + v_2 \left[(1-p)(1-(1-p)^2 p) + (1-p)p^3 \right] \\ &\quad + v_3 \left[(1-p)(1-(1-p)^3) + p^4 \right] \\ &= u_0 + v_0 + v_1 + v_2 \\ &\quad + p \cdot (-3u_0 + 2u_1 + u_2 - v_1 - 2v_2 + 3v_3) \\ &\quad + p^2 \cdot (6u_0 - 3u_1 + 3u_3 - 3v_0 + 3v_2 - 6v_3) \\ &\quad + p^3 \cdot (-4u_0 + 2u_1 - 2u_3 + 2v_0 - 2v_2 + 4v_3) \end{aligned}$$

Every distribution (whose marginals are) in $\widetilde{K}^N_{1/2}(\mathsf{MON}_4)$ must satisfy the following system of equations and inequalities, where (3.8)–(3.11) are equivalent to $h_{\mathcal{D}}(p) = \frac{1}{2}$, and (3.12)–(3.14) guarantee that u_is and v_is describe a distribution.

$$u_0 + v_0 + v_1 + v_2 = \frac{1}{2} \tag{3.8}$$

$$-3u_0 + 2u_1 + u_2 - v_1 - 2v_2 + 3v_3 = 0 \tag{3.9}$$

$$6u_0 - 3u_1 + 3u_3 - 3v_0 + 3v_2 - 6v_3 = 0 \tag{3.10}$$

$$-4u_0 + 2u_1 - 2u_3 + 2v_0 - 2v_2 + 4v_3 = 0 \tag{3.11}$$

$$\sum_{i=0}^{3}(u_i + v_i) = 1 \tag{3.12}$$

$$u_i \geq 0, \quad \forall 0 \leq i \leq 3 \tag{3.13}$$

$$v_i \geq 0, \quad \forall 0 \leq i \leq 3 \tag{3.14}$$

Summing up (3.9) multiplied by 3, (3.11) multiplied by $-13/6$, and (3.12) multiplied by $2/3$, we have that

$$\begin{aligned} 2/3 &= u_0/3 + 7u_1/3 + 11u_2/3 + 5u_3 - 11v_0/3 - 7v_1/3 - v_2 + v_3 \\ &\geq -u_0 + u_1 + 3u_2 + 5u_3 - 5v_0 - 3v_1 - v_2 + v_3 \\ &= 2\mu_1 + 3\mu' , \end{aligned}$$

where the last equality uses (3.1). By Lemma 3.4, $\widetilde{K}_1^Y(\mathsf{MON}_4) = \{(\mu_1, \mu') \in [-1,1]^2 : 2\mu_1 + 3\mu' \geq 1\}$, and from the above inequality every vector $(\mu_1, \mu') \in \widetilde{K}_{1/2}^N(\mathsf{MON}_4)$ satisfies $2\mu_1 + 3\mu' \leq 2/3$. This implies that $\widetilde{K}_1^Y(\mathsf{MON}_4) \cap \widetilde{K}_{1/2}^N(\mathsf{MON}_4) = \emptyset$, and finishes the proof. □

3.2 Balanced LTFs on 4 variables

In this section, we prove Theorem 1.4.

Theorem 1.4. *For every balanced LTF f on $k \leq 4$ variables, Max-CSP(f) is sketching approximable in $O(\log(n))$ space.*

We remark that there are non-balanced LTFs on fewer than four variables that are approximation resistant. For example, if $f(x_1, x_2) = x_1$ OR x_2, then Max-CSP(f) is approximation resistant to space $o(n)$ even in the larger class of streaming algorithms (see, e.g., Corollary 4.2 in [CGV20]).

Proof of Theorem 1.4. After relabeling and negating some of the variables of f, we can assume that $f(x_1, x_2, x_3, x_4) = \mathsf{sign}(w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4)$, where $w_1 \geq w_2 \geq w_3 \geq w_4 \geq 0$ (if f depends on $i < 4$ variables, then we set $w_{i+1} = \ldots = w_4 = 0$). Since f is balanced, $\xi_1w_1 + \xi_2w_2 + \xi_3w_3 + \xi_4w_4 \neq 0$ for all $\xi_i \in \{-1, 1\}$. Now consider the following three cases.

- If $w_1 > w_2 + w_3 + w_4$, then $f = \mathsf{sign}(x_1)$ is a dictator function, so Max-CSP(f) can be trivially $(1-\varepsilon)$-approximated in $O(\log(n)/\varepsilon^2)$ space by an ℓ_1-sketch algorithm [Ind00, KNW10].

- If $w_2 + w_3 - w_4 < w_1 < w_2 + w_3 + w_4$, then $f = \mathsf{MON}_4$ is a monarchy function on $k = 4$ variables. Indeed, in this case only the sum of the votes of the three last variables overrules the vote of the first variable. By Lemma 3.5, Max-CSP(f) is sketching approximable in $O(\log(n))$ space.

- If $w_1 < w_2 + w_3 - w_4$, then $f = \mathsf{MAJ}(x_1, x_2, x_3)$ is the majority function on 3 variables. Indeed, the sum of any two weights of the first three variables outweighs the sum of the remaining weights. In this case, Max-CSP(f) is known to be sketching approximable in space $O(\log(n))$ (this follows from the characterization of sketching approximable symmetric functions in [CGSV22, Lemma 2.14] and the fact that a balanced LTF doesn't support one-wise independent distributions).

Another way to see that the majority function is sketching approximable is via Theorem 5.3. Indeed, since majority is a symmetric function, the (non-empty) Chow parameters of the majority function are all equal and non-zero (see, e.g., [O'D14, Theorem 5.19] for the exact values of the Fourier coefficients of the majority function). Then the Chow parameters define the majority function itself, and, by Theorem 5.3, Max-CSP(f) is sketching approximable in space $O(\log(n))$. □

4 Approximation resistance of Monarchy Functions

In this section, we prove Theorem 1.5: we show that for $k \geq 5$, the MON_k function is approximation resistant. Recall that by Lemma 3.3 it suffices to show that $\widetilde{K}_1^Y(\mathsf{MON}_k) \cap \widetilde{K}_{1/2}^N(\mathsf{MON}_k) \neq \emptyset$ for $k \geq 5$.

In the following we show that for $k \geq 5$, there exist vectors $(\mathbf{u}, \mathbf{v})$ with certain properties that will be useful in showing that $\widetilde{K}_1^Y(\mathsf{MON}_k) \cap \widetilde{K}_{1/2}^N(\mathsf{MON}_k) \neq \emptyset$.

Lemma 4.1. *For every $k \geq 5$, there exists $\mathbf{u}, \mathbf{v} \in \mathbb{R}_{\geq 0}^k$ satisfying the following conditions.*

(i) $\sum_i (u_i + v_i) = 1$*, i.e.,* $\mathbf{u}, \mathbf{v}$ *define a distribution* $\mathcal{D}$*. In particular, the marginals of* $\mathcal{D}$ *is* $(\mu_1, \mu', \ldots, \mu')$ *where* $\mu_1 = \sum_i (u_i - v_i)$*, and* $\mu' = \sum_i (\frac{2i}{k-1} - 1)(u_i + v_i)$*.*

(ii) $\mathbf{u}$ *and* $\mathbf{v}$ *satisfy*

$$\begin{aligned}
&(1/2-\delta)\sum_{i=0}^{k-1} u_i + (1/2+\delta)\sum_{i=0}^{k-1} v_i \\
&+\sum_{i=0}^{k-1} u_i \left(-(1/2+\delta)^i(1/2-\delta)^{k-i} + (1/2-\delta)^i(1/2+\delta)^{k-i}\right) \\
&+\sum_{i=0}^{k-1} v_i \left(-(1/2+\delta)^{i+1}(1/2-\delta)^{k-1-i} + (1/2-\delta)^{i+1}(1/2+\delta)^{k-1-i}\right) \\
&=1/2
\end{aligned}$$

for every $\delta \in [-1/2, 1/2]$*. In particular, this implies that* $\mathcal{D} \in S_{1/2}^N$*.*

(iii) $p' \geq 1 - \frac{k-2}{k-1} p_1$ *where* $p' = \Pr_{\mathbf{x}\sim\mathcal{D}}[x_2 = 1] = \frac{1}{k-1}\left(\sum_i i u_i + \sum_i i v_i\right)$ *and* $p_1 = \Pr_{\mathbf{x}\sim\mathcal{D}}[x_1 = 1] = \sum_i u_i$*. In particular, this implies the existence of* $\mathcal{D}_Y \in S_1^Y$ *and* $\boldsymbol{\mu}(\mathcal{D}_Y) = (\mu_1, \mu', \ldots, \mu')$*.*

Now, we are ready to prove Theorem 1.5 using Lemma 4.1 and Theorem 2.5.

Theorem 1.5. *For every $k \geq 5$, Max-CSP(MON_k) is sketching approximation resistant to space $o(\sqrt{n})$.*

Proof. For every $k \geq 5$, let $\mathbf{u}, \mathbf{v} \in \mathbb{R}_{\geq 0}^k$, and $\mu_1, \mu' \in [-1, 1]$ be the vectors given by Lemma 4.1. Note that condition (i) guarantees that $\mathbf{u}, \mathbf{v}$ define a distribution $\mathcal{D}$ with marginal $(\mu_1, \mu', \ldots, \mu')$.

First, we show that condition (ii) is a sufficient condition for $(\mu_1, \mu') \in \widetilde{K}_{1/2}^N$. Recall that $\mathcal{D}_N \in S_{1/2}^N(\mathsf{MON}_k)$ if for every $\delta \in [-1/2, 1/2]$, $\mathsf{E}_{\mathbf{b}\in\mathcal{D}_N}\mathsf{E}_{\mathbf{a}\sim\mathsf{Bern}(1/2+\delta)}[\mathsf{MON}_k(\mathbf{b}\odot\mathbf{a})] = 1/2$. Since $\Pr_{\mathbf{x}}[\mathsf{MON}_k(\mathbf{x}) = 1] = \Pr_{\mathbf{x}}[x_1 = 1] - \Pr_{\mathbf{x}}[\mathbf{x} = 10^{k-1}] + \Pr_{\mathbf{x}}[\mathbf{x} = 01^{k-1}]$, we have that

$$\mathsf{E}_{\mathbf{b}\in\mathcal{D}_N}\mathsf{E}_{\mathbf{a}\sim\mathsf{Bern}(1/2+\delta)}[\mathsf{MON}_k(\mathbf{b}\odot\mathbf{a})] = \Pr_{\mathbf{b},\mathbf{a}}[\mathbf{b}_1\odot\mathbf{a}_1 = 1] - \Pr_{\mathbf{b},\mathbf{a}}[\mathbf{b}\odot\mathbf{a} = 1(-1)^{k-1}] + \Pr_{\mathbf{b},\mathbf{a}}[\mathbf{b}\odot\mathbf{a} = (-1)1^{k-1}].$$

We compute these three probabilities in terms of $\mathbf{u}, \mathbf{v}, \delta$.

$$\Pr_{\mathbf{b},\mathbf{a}}[\mathbf{b}_1 \odot \mathbf{a}_1 = 1] = (1/2-\delta)\sum_{i=0}^{k-1} u_i + (1/2+\delta)\sum_{i=0}^{k-1} v_i \,,$$

$$\Pr_{\mathbf{b},\mathbf{a}}[\mathbf{b} \odot \mathbf{a} = 1(-1)^{k-1}] = \sum_{i=0}^{k-1} u_i(1/2+\delta)^i(1/2-\delta)^{k-i} + \sum_{i=0}^{k-1} v_i(1/2+\delta)^{i+1}(1/2-\delta)^{k-1-i} \,,$$

$$\Pr_{\mathbf{b},\mathbf{a}}[\mathbf{b} \odot \mathbf{a} = (-1)1^{k-1}] = \sum_{i=0}^{k-1} u_i(1/2-\delta)^i(1/2+\delta)^{k-i} + \sum_{i=0}^{k-1} v_i(1/2-\delta)^{i+1}(1/2+\delta)^{k-1-i} \,.$$

Note that condition (ii) implies that

$$\Pr_{\mathbf{b},\mathbf{a}}[\mathbf{b}_1 \odot \mathbf{a}_1 = 1] + \Pr_{\mathbf{b},\mathbf{a}}[\mathbf{b} \odot \mathbf{a} = 1(-1)^{k-1}] + \Pr_{\mathbf{b},\mathbf{a}}[\mathbf{b} \odot \mathbf{a} = (-1)1^{k-1}] = \frac{1}{2}$$

for every $\delta \in [-1/2, 1/2]$ as desired. This implies that $\mathcal{D} \in S^N_{1/2}(\mathsf{MON}_k)$. As condition (i) gives $\boldsymbol{\mu}(\mathcal{D}_N) = (\mu_1, \mu', \ldots, \mu')$, we have $(\mu_1, \mu') \in \widetilde{K}^N_{1/2}$ as desired.

Next, as $p' = \frac{\mu'+1}{2}$ and $1 - \frac{k-2}{k-1}p_1 = 1 - \frac{(k-2)(\mu_1+1)}{2(k-1)}$, condition (iii) implies $\mu_1(k-2) + \mu'(k-1) \geq 1$. By Lemma 3.4, this implies that $(\mu_1, \mu') \in \widetilde{K}^Y_1(\mathsf{MON}_k)$ as desired.

To sum up, Lemma 4.1 gives us $(\mu_1, \mu') \in \widetilde{K}^Y_1 \cap \widetilde{K}^N_{1/2}$ for every $k \geq 5$ and Lemma 3.3 implies $(\mu_1, \mu', \ldots, \mu') \in K^Y_1 \cap K^N_{1/2}$. By Theorem 2.5, we conclude that MON_k is sketching approximation resistant to space $o(\sqrt{n})$ and, hence, complete the proof of Theorem 1.5. □

4.1 Proof of Lemma 4.1

In the proof of Lemma 4.1 we will use the following combinatorial identity.

Lemma 4.2. *For every $\delta \in [-1/2, 1/2]$ and $m \in \mathbb{N}$,*

$$\begin{aligned}
&\sum_{i=\lceil m/2 \rceil}^{m} (1/2+\delta)^{i+1}(1/2-\delta)^{m-i}\left(\binom{m}{i} - \binom{m}{i+1}\right) \\
&- \sum_{i=\lceil m/2 \rceil}^{m} (1/2-\delta)^{i+1}(1/2+\delta)^{m-i}\left(\binom{m}{i} - \binom{m}{i+1}\right) \\
&= 2\delta \,.
\end{aligned}$$

Proof. Let $X_1, \ldots, X_{m+1}$ be independent identically distributed random variables, each having the distribution $\mathsf{Bern}(1/2+\delta)$. For $j \in \{0, \ldots, m+1\}$, let $\mathbb{1}_j$ be the indicator of the event that exactly j variables from $X_1, \ldots, X_{m+1}$ are ones. First observe that for $i \in \{0, \ldots, m+1\}$,

$$\mathsf{E}[x_1 \cdot \mathbb{1}_i] = (1/2+\delta)^i(1/2-\delta)^{m+1-i}\left(\binom{m}{i-1} - \binom{m}{i}\right) \,.$$

Using the above, we are going to show that the left hand side of the equation in Lemma 4.1 equals to $\sum_{i=0}^{m+1} \mathsf{E}[x_1 \cdot \mathbb{1}_i] = \mathsf{E}[x_1] = 2\delta$. By changing summations' limits and updating the binomial coefficients accordingly, we have

$$
\begin{aligned}
&\sum_{i=\lceil m/2 \rceil}^{m} (1/2+\delta)^{i+1}(1/2-\delta)^{m-i} \left(\binom{m}{i} - \binom{m}{i+1}\right) \\
&- \sum_{i=\lceil m/2 \rceil}^{m} (1/2-\delta)^{i+1}(1/2+\delta)^{m-i} \left(\binom{m}{i} - \binom{m}{i+1}\right) . \\
&= \sum_{i=0}^{\lfloor m/2 \rfloor} (1/2+\delta)^{m-i+1}(1/2-\delta)^{i} \left(\binom{m}{i} - \binom{m}{i-1}\right) \\
&- \sum_{i=\lceil m/2 \rceil+1}^{m+1} (1/2-\delta)^{i}(1/2+\delta)^{m-i+1} \left(\binom{m}{i-1} - \binom{m}{i}\right) .
\end{aligned}
$$

Using $\binom{m}{\lfloor m/2 \rfloor} = \binom{m}{\lceil m/2 \rceil}$, we update the first summation's limits:

$$
\begin{aligned}
&= \sum_{i=0}^{\lceil m/2 \rceil} (1/2+\delta)^{m-i+1}(1/2-\delta)^{i} \left(\binom{m}{i} - \binom{m}{i-1}\right) \\
&- \sum_{i=\lceil m/2 \rceil+1}^{m+1} (1/2-\delta)^{i}(1/2+\delta)^{m-i+1} \left(\binom{m}{i-1} - \binom{m}{i}\right) \\
&= \sum_{i=0}^{m} (1/2-\delta)^{i}(1/2+\delta)^{m-i+1} \left(\binom{m}{i} - \binom{m}{i-1}\right) \\
&= \sum_{i=0}^{m+1} \mathsf{E}[x_1 \cdot \mathbb{1}_i] = \mathsf{E}[x_1] = 2\delta \,,
\end{aligned}
$$

which concludes the proof. □

We are now ready to prove Lemma 4.1.

Proof of Lemma 4.1. We prove this lemma by considering three cases: $k = 5$, $k > 5$ is even, and $k > 5$ is odd.

Case I: $k = 5$. In this case, we consider the following pair of vectors

$$
\begin{aligned}
\mathbf{u} = (u_0, u_1, u_2, u_3, u_4) &= \left(0, 0, 0, 0, \frac{1}{3}\right) , \\
\mathbf{v} = (v_0, v_1, v_2, v_3, v_4) &= \left(0, 0, \frac{1}{3}, \frac{1}{6}, \frac{1}{6}\right) .
\end{aligned}
$$

(i) It's straightforward to verify that $\sum_i (u_i + v_i) = 1$.

(ii) For $\delta \in [-1/2, 1/2)$, using the substitution $y = (1/2+\delta)/(1/2-\delta)$, we have

$$
\begin{aligned}
&(1/2-\delta)\sum_{i=0}^{k-1} u_i + (1/2+\delta)\sum_{i=0}^{k-1} v_i \\
&+\sum_{i=0}^{k-1} u_i \left(-(1/2+\delta)^i(1/2-\delta)^{k-i} + (1/2-\delta)^i(1/2+\delta)^{k-i}\right) \\
&+\sum_{i=0}^{k-1} v_i \left(-(1/2+\delta)^{i+1}(1/2-\delta)^{k-1-i} + (1/2-\delta)^{i+1}(1/2+\delta)^{k-1-i}\right) \\
=&1/2+\delta/3 \\
&+1/3(-(1/2+\delta)^4(1/2-\delta)+(1/2-\delta)^4(1/2+\delta)) \\
&+1/3(-(1/2+\delta)^3(1/2-\delta)^2+(1/2-\delta)^3(1/2+\delta)^2) \\
&+1/6(-(1/2+\delta)^4(1/2-\delta)+(1/2-\delta)^4(1/2+\delta)) \\
&+1/6(-(1/2+\delta)^5+(1/2-\delta)^5) \\
=&1/2+\delta/3+(1/2-\delta)^5(-y^4/3+y/3-y^3/3+y^2/3-y^4/6+y/6-y^5/6+1/6) \\
=&1/2+\delta/3-(1/2-\delta)^5(y-1)(y+1)^4/6 \\
=&1/2+\delta/3-(1/2-\delta)^5\left(\frac{2\delta}{1/2-\delta}\right)\left(\frac{1}{1/2-\delta}\right)^4/6 \\
=&1/2+\delta/3-2\delta/6 = 1/2\,.
\end{aligned}
$$

For $\delta = 1/2$, it's easy to see that the sum above equals $1/2$, too.

(iii) Since $p_1 = \sum_i u_i = 1/3$ and $p' = \frac{1}{k-1}\left(\sum_i iu_i + \sum_i iv_i\right) = 19/24$, the inequality $p' \geq 1 - \frac{k-2}{k-1}p_1$ holds.

Case II: $k > 5$ is even. Let $T = \binom{k}{k/2} - 2$. Consider the vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^k_{\geq 0}$ as follows.

$$
u_i = \begin{cases} \frac{T-2}{2T} & \text{, if } i = k/2 \\ 0 & \text{, otherwise.} \end{cases} \quad \text{and} \quad v_i = \begin{cases} \frac{\binom{k-1}{i} - \binom{k-1}{i+1}}{T} & \text{, if } i \geq k/2 \\ 0 & \text{, otherwise.} \end{cases}
$$

(i) Note that

$$
\sum_{i=0}^{k-1} v_i = \frac{1}{T}\sum_{i=k/2}^{k}\left(\binom{k-1}{i} - \binom{k-1}{i+1}\right) = \frac{1}{T}\binom{k-1}{k/2} = \frac{1}{2T}\binom{k}{k/2} = \frac{T+2}{2T}\,.
$$

Thus,

$$
\sum_{i=0}^{k-1} u_i + \sum_{i=0}^{k-1} v_i = \frac{T-2}{2T} + \frac{T+2}{2T} = 1\,.
$$

(ii) From the definition of $\mathbf{u}$ and $\mathbf{v}$, using $\sum_{i=0}^{k-1} v_i = \frac{T+2}{2T}$ and applying Lemma 4.2 with $m = k-1$, we have that for every $\delta \in [-1/2, 1/2]$,

$$
(1/2-\delta)\sum_{i=0}^{k-1} u_i + (1/2+\delta)\sum_{i=0}^{k-1} v_i
$$

$$
\begin{aligned}
&+\sum_{i=0}^{k-1} u_i \left(-(1/2+\delta)^i(1/2-\delta)^{k-i} + (1/2-\delta)^i(1/2+\delta)^{k-i}\right) \\
&+\sum_{i=0}^{k-1} v_i \left(-(1/2+\delta)^{i+1}(1/2-\delta)^{k-1-i} + (1/2-\delta)^{i+1}(1/2+\delta)^{k-1-i}\right) \\
&= \left(\frac{T-2}{2T}(1/2-\delta) + \frac{T+2}{2T}(1/2+\delta)\right) + 0 \\
&+\frac{1}{T}\sum_{i=k/2}^{k-1} \left(\binom{k-1}{i} - \binom{k-1}{i+1}\right)\left(-(1/2+\delta)^{i+1}(1/2-\delta)^{k-1-i} + (1/2-\delta)^{i+1}(1/2+\delta)^{k-1-i}\right) \\
&= (1/2 + 2\delta/T) - 2\delta/T = 1/2\ .
\end{aligned}
$$

(iii) From the definition of $\mathbf{u}$ and $\mathbf{v}$, we have that $p_1 = \sum_0^{k-1} u_i = \frac{T-2}{2T}$.

$$
\begin{aligned}
p' &= \frac{1}{k-1}\left(\sum_0^{k-1} iu_i + \sum_0^{k-1} iv_i\right) \\
&= \frac{1}{T(k-1)}\left((T-2)k/4 + \sum_{k/2}^{k-1} i\left(\binom{k-1}{i} - \binom{k-1}{i+1}\right)\right) \\
&= \frac{1}{T(k-1)}\left((T-2)k/4 + (k/2-1)\binom{k-1}{k/2} + \sum_{i=k/2}^{k-1} i\binom{k-1}{i} - \sum_{i=k/2-1}^{k-1} i\binom{k-1}{i+1}\right) \\
&= \frac{1}{T(k-1)}\left((T-2)k/4 + (k/2-1)\binom{k-1}{k/2} + \sum_{i=k/2}^{k-1} i\binom{k-1}{i} - \sum_{i=k/2}^{k} (i-1)\binom{k-1}{i}\right) \\
&= \frac{1}{T(k-1)}\left((T-2)k/4 + (k/2-1)\binom{k-1}{k/2} + \sum_{i=k/2}^{k-1} \binom{k-1}{i}\right) \\
&= \frac{1}{T(k-1)}\left((T-2)k/4 + (k/2-1)\binom{k-1}{k/2} + 2^{k-2}\right)\ .
\end{aligned}
$$

Using $\binom{k-1}{k/2} = \frac{1}{2}\binom{k}{k/2} = (T+2)/2$

$$
= \frac{1}{T(k-1)}\left((T-2)k/4 + (k/2-1)(T+2)/2 + 2^{k-2}\right)\ .
$$

Using $2^{k-2} \geq k + (\binom{k}{k/2} - 2)/2 = k + T/2$ for $k \geq 6$

$$
\begin{aligned}
&\geq \frac{Tk/2 + k - 1}{T(k-1)} \\
&= 1 - \frac{Tk - 2T - 2k + 2}{2T(k-1)} \\
&> 1 - \frac{k-2}{k-1}\cdot\frac{T-2}{2T}
\end{aligned}
$$

$$= 1 - \frac{k-2}{k-1} p_1 \,.$$

Case III: $k > 5$ is odd. Let $T = 2\binom{k-1}{\frac{k-1}{2}} - 2$. Consider the vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^k_{\geq 0}$ as follows.

$$u_i = \begin{cases} \frac{T-2}{4T} & \text{, if } i = \frac{k-1}{2} \text{ or } i = \frac{k+1}{2} \\ 0 & \text{, otherwise.} \end{cases} \quad \text{and} \quad v_i = \begin{cases} \frac{\binom{k-1}{i} - \binom{k-1}{i+1}}{T} & \text{, if } i \geq \frac{k-1}{2} \\ 0 & \text{, otherwise.} \end{cases}$$

(i) Similarly to Case II, $\sum_{i=0}^{k-1} v_i = \frac{T+2}{2T}$ and $\sum_{i=0}^{k-1} u_i + \sum_{i=0}^{k-1} v_i = 1$.

(ii) Using $\sum_{i=0}^{k-1} v_i = \frac{T+2}{2T}$ and Lemma 4.2 with $m = k-1$, we conclude that for every $\delta \in [-1/2, 1/2]$,

$$\begin{aligned}
&(1/2 - \delta) \sum_{i=0}^{k-1} u_i + (1/2 + \delta) \sum_{i=0}^{k-1} v_i \\
&+ \sum_{i=0}^{k-1} u_i \left(-(1/2+\delta)^i (1/2-\delta)^{k-i} + (1/2-\delta)^i (1/2+\delta)^{k-i} \right) \\
&+ \sum_{i=0}^{k-1} v_i \left(-(1/2+\delta)^{i+1} (1/2-\delta)^{k-1-i} + (1/2-\delta)^{i+1} (1/2+\delta)^{k-1-i} \right) \\
&= \left(\frac{2(T-2)}{4T} (1/2 - \delta) + \frac{T+2}{2T} (1/2 + \delta) \right) \\
&+ \frac{T-2}{4T} \left(-(1/2+\delta)^{\frac{k-1}{2}} (1/2-\delta)^{\frac{k+1}{2}} + (1/2-\delta)^{\frac{k-1}{2}} (1/2+\delta)^{\frac{k+1}{2}} \right) \\
&+ \frac{T-2}{4T} \left(-(1/2+\delta)^{\frac{k+1}{2}} (1/2-\delta)^{\frac{k-1}{2}} + (1/2-\delta)^{\frac{k+1}{2}} (1/2+\delta)^{\frac{k-1}{2}} \right) \\
&+ \frac{1}{T} \sum_{i=(k-1)/2}^{k-1} \left(\binom{k-1}{i} - \binom{k-1}{i+1} \right) \left(-(1/2+\delta)^{i+1} (1/2-\delta)^{k-1-i} + (1/2-\delta)^{i+1} (1/2+\delta)^{k-1-i} \right) \\
&= (1/2 + 2\delta/T) - 2\delta/T = 1/2 \,.
\end{aligned}$$

(iii) Similarly to the previous case, $p_1 = \sum_0^{k-1} u_i = 2\frac{T-2}{4T} = \frac{T-2}{2T}$, and

$$\begin{aligned}
p' &= \frac{1}{k-1} \left(\sum_{i=0}^{k-1} i u_i + \sum_{i=0}^{k-1} i v_i \right) \\
&= \frac{1}{T(k-1)} \left(\frac{(T-2)}{4} \times \left(\frac{k-1}{2} + \frac{k+1}{2} \right) + \sum_{i=(k-1)/2}^{k-1} i \left(\binom{k-1}{i} - \binom{k-1}{i+1} \right) \right) \\
&= \frac{1}{T(k-1)} \left((T-2)k/4 + \left(\frac{k-1}{2} - 1 \right) \binom{k-1}{(k-1)/2} \right. \\
&\qquad \left. + \sum_{i=(k-1)/2}^{k-1} i \binom{k-1}{i} - \sum_{i=(k-1)/2-1}^{k-1} i \binom{k-1}{i+1} \right)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{T(k-1)} \left((T-2)k/4 + \left(\frac{k-1}{2} - 1\right)\binom{k-1}{(k-1)/2} \right. \\
&\qquad \left. + \sum_{i=(k-1)/2}^{k-1} i\binom{k-1}{i} - \sum_{i=(k-1)/2}^{k} (i-1)\binom{k-1}{i} \right) \\
&= \frac{1}{T(k-1)} \left((T-2)k/4 + \left(\frac{k-1}{2} - 1\right)\binom{k-1}{(k-1)/2} + \sum_{i=(k-1)/2}^{k-1} \binom{k-1}{i} \right) \\
&= \frac{1}{T(k-1)} \left((T-2)k/4 + \left(\frac{k-1}{2} - 1\right)\binom{k-1}{(k-1)/2} + \binom{k-1}{\leq (k-1)/2} \right) .
\end{aligned}
$$

Using $\binom{k-1}{(k-1)/2} = (T+2)/2$

$$
\begin{aligned}
&= \frac{1}{T(k-1)} \left((T-2)k/4 + \left(\frac{k-1}{2} - 1\right)(T+2)/2 + \binom{k-1}{\leq (k-1)/2} \right) \\
&= \frac{1}{T(k-1)} \left(Tk/2 - 3T/4 - 3/2 + \binom{k-1}{\leq (k-1)/2} \right)
\end{aligned}
$$

Using $\binom{k-1}{\leq (k-1)/2} \geq \frac{3}{2}\binom{k-1}{\frac{k-1}{2}} + k - 2 = 3T/4 + k - 1/2$ which holds for every $k \geq 7$

$$
\begin{aligned}
&\geq \frac{Tk/2 + k - 2}{T(k-1)} \\
&= 1 - \frac{Tk - 2T - 2k + 4}{2T(k-1)} \\
&= 1 - \frac{k-2}{k-1} \cdot \frac{T-2}{2T} \\
&= 1 - \frac{k-2}{k-1} p_1 .
\end{aligned}
$$

This concludes the proof of Lemma 4.1. □

5 Chow parameters and the approximability of weak monarchies

In this section, we prove that infinitely many weak monarchy functions are sketching approximable within $O(\log(n))$ space. We first prove in Sections 5.1 and 5.2 that every LTF defined by its Chow parameters (i.e., degree-1 Fourier coefficients as weights and threshold 0) is sketching approximable within $O(\log(n))$ space. And later in Section 5.3, we prove that infinitely many weak monarchy functions are balanced LTFs defined by their Chow parameters.

5.1 Approximability of LTFs defined by their Chow parameters

Theorem 5.1. *For every Boolean function $f : \{-1,1\}^k \to \{0,1\}$ of the form*

$$
f(x) = \mathsf{sign}\left(\sum_{i=1}^{k} \widehat{f}(\{i\}) x_i \right) ,
$$

$\mathsf{Max\text{-}CSP}(f)$ is sketching approximable in $O(\log(n))$ space.

Definition 5.2. *Define* $\varepsilon_0(f) = \min\{\sum_{i=1}^k \widehat{f}(\{i\}) \cdot x_i : f(x) = 1\}$*. Define* $\varepsilon^*(f) = \min\{\frac{\varepsilon_0(f)}{3k}, \frac{2\varepsilon_0(f)^2}{9\rho(f)k^2}\}$*.*

We will use the following theorem to prove Theorem 5.1.

Theorem 5.3. *For every Boolean function* $f : \{-1,1\}^k \to \{0,1\}$ *and every* $\varepsilon > 0$*, there exists an* $O(\log(n))$ *space* $(\rho(f) + \varepsilon^*(f) - \varepsilon)$*-approximation algorithm for* $\mathsf{Max\text{-}CSP}(f)$*.*

First we show how to prove Theorem 5.1 using Theorem 5.3.

Proof of Theorem 5.1. If $f(x)$ is the constant zero function, then it's trivially approximable in $O(\log(n))$ space. Otherwise, when $f(x) = \mathsf{sign}\left(\sum_{i=1}^k \widehat{f}(\{i\}) \cdot x_i\right)$, we have $\varepsilon_0(f) = \min\{\sum_{i=1}^k \widehat{f}(\{i\}) \cdot x_i : f(x) = 1\} > 0$ and hence $\varepsilon^*(f) > 0$ by their definitions. Now for $\varepsilon = \varepsilon^*(f)/2$, Theorem 5.3 implies that there is a $(\rho(f) + \varepsilon^*(f)/2)$-approximation algorithm for $\mathsf{Max\text{-}CSP}(f)$, and finishes the proof. □

Before we prove Theorem 5.3, we will describe some useful definitions and lemmas from [CGSV22].

Let $f : \{-1,1\}^k \to \{0,1\}$ be a Boolean constraint function of arity k and $X_1, \ldots, X_n$ be variables. A constraint C consists of $\mathbf{j} = (j_1, \ldots, j_k) \in [n]^k$ and $\mathbf{b} = (b_1, \ldots, b_k) \in \{-1,1\}^k$ where the j_i's are distinct. The constraint C reads as requiring $f(\mathbf{b} \odot \mathbf{X}|_{\mathbf{j}}) = f(b_1 X_{j_1}, \ldots, b_k X_{j_k}) = 1$. A $\mathsf{Max\text{-}CSP}(f)$ instance Ψ contains m constraints $C_1, \ldots, C_m$ with non-negative weights $w_1, \ldots, w_m$ where $C_i = (\mathbf{j}(i), \mathbf{b}(i))$ and $w_i \in \mathbb{R}$ for each $i \in [m]$. For an assignment $\boldsymbol{\sigma} \in \{-1,1\}^n$, the value $\mathsf{val}_\Psi(\boldsymbol{\sigma})$ of $\boldsymbol{\sigma}$ on Ψ is the fraction of weight of constraints satisfied by $\boldsymbol{\sigma}$, i.e., $\mathsf{val}_\Psi(\boldsymbol{\sigma}) = \frac{1}{W}\sum_{i\in[m]} w_i \cdot f(\mathbf{b}(i) \odot \boldsymbol{\sigma}|_{\mathbf{j}(i)})$, where $W = \sum_{i=1}^m w_i$. The optimal value of Ψ is defined as $\mathsf{val}_\Psi = \max_{\boldsymbol{\sigma}\in\{-1,1\}^n} \mathsf{val}_\Psi(\boldsymbol{\sigma})$.

Definition 5.4 (Bias (vector))**.** *For* $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_k) \in \mathbb{R}^k$*, and instance* $\Psi = (C_1, \ldots, C_m; w_1, \ldots, w_m)$ *of* $\mathsf{Max\text{-}CSP}(f)$ *where* $C_i = (\mathbf{j}(i), \mathbf{b}(i))$ *and* $w_i \geq 0$*, we let the* $\boldsymbol{\lambda}$-bias vector *of* Ψ*, denoted* $\mathsf{bias}_{\boldsymbol{\lambda}}(\Psi)$*, be the vector in* $\mathbb{R}^n$ *given by*

$$\mathsf{bias}_{\boldsymbol{\lambda}}(\Psi)_\ell = \frac{1}{W} \cdot \sum_{i\in[m], t\in[k]: j(i)_t = \ell} \lambda_t w_i \cdot b(i)_t\,,$$

for $\ell \in [n]$*, where* $W = \sum_{i\in[m]} w_i$*. The* $\boldsymbol{\lambda}$*-bias of* Ψ*, denoted* $B_{\boldsymbol{\lambda}}(\Psi)$*, is the* ℓ_1 *norm of* $\mathsf{bias}_{\boldsymbol{\lambda}}(\Psi)$*, i.e.,* $B_{\boldsymbol{\lambda}}(\Psi) = \sum_{\ell=1}^n |\mathsf{bias}_{\boldsymbol{\lambda}}(\Psi)_\ell|$*.*

Lemma 5.5 ([CGSV22, Lemma 4.7])**.** *For every* $\boldsymbol{\lambda} \in \mathbb{R}^k$*, we have* $B_{\boldsymbol{\lambda}}(\Psi) = \max_{a\in\{-1,1\}^n} \langle a, \mathsf{bias}_{\boldsymbol{\lambda}}(\Psi)\rangle$*.*

Lemma 5.6 ([CGSV22, Lemma 4.4])**.** *For every vector* $\boldsymbol{\lambda} \in \mathbb{R}^k$ *and* $\varepsilon > 0$*, there exists a* $O(\log(n))$ *space sketching algorithm* $\mathcal{A}$ *that on input a stream* $\sigma_1, \ldots, \sigma_\ell$*, representing an instance* $\Psi = (C_1, \ldots, C_m; w_1, \ldots, w_m)$*, outputs a* $(1 \pm \varepsilon)$*-approximation to* $B_{\boldsymbol{\lambda}}(\Psi)$*, i.e., for every* Ψ*,* $(1-\varepsilon)B_{\boldsymbol{\lambda}}(\Psi) \leq \mathcal{A}(\Psi) \leq (1+\varepsilon)B_{\boldsymbol{\lambda}}(\Psi)$*, with probability at least* $2/3$*.*

Below, we describe Algorithm 1 and show that it is an $O(\log(n))$ space $(\rho(f) + \varepsilon^*(f) - \varepsilon)$-approximation algorithm for $\mathsf{Max\text{-}CSP}(f)$.

Algorithm 1 A sketching $(\rho(f) + \varepsilon^*(f) - \varepsilon)$-approximation algorithm for $\textsf{Max-CSP}(f)$

Input: a stream $\sigma_1, \ldots, \sigma_\ell$ representing an instance Ψ of $\textsf{Max-CSP}(f)$ where $\sigma_i = ((\mathbf{j}(i), \mathbf{b}(i)), w_i)$.

1: Let $\boldsymbol{\lambda} = (\widehat{f}(\{1\}), \ldots, \widehat{f}(\{k\})) \in \mathbb{R}^k$ and $\varepsilon' = \varepsilon/8$.

2: Use the algorithm $\mathcal{A}$ from Lemma 5.6 to compute $\tilde{B}$ to be a $(1 \pm \varepsilon')$ approximation to $B_{\boldsymbol{\lambda}}(\Psi)$, i.e., $(1 - \varepsilon')B_{\boldsymbol{\lambda}}(\Psi) \leq \tilde{B} \leq (1 + \varepsilon')B_{\boldsymbol{\lambda}}(\Psi)$ with probability at least 2/3.

3: Let $\tilde{\delta} = \min\{\frac{1}{3k}, \frac{2\tilde{B}}{9\rho(f)k^2}\}$.

4: **Output:** $v = \rho(f) + \frac{\tilde{B}\tilde{\delta}}{(1+\varepsilon')^2}$.

It is clear that the algorithm above runs in $O(\log(n))$ space (in particular by Lemma 5.6 for Step 2). We now turn to analyzing the correctness of the algorithm.

5.1.1 Analysis of the correctness of Algorithm 1

Before we analyse Algorithm 1, we establish some upper and lower bounds on $\textsf{val}_\Psi$ in terms of $B_{\boldsymbol{\lambda}}(\Psi)$ where $\boldsymbol{\lambda} = (\widehat{f}(\{1\}), \ldots, \widehat{f}(\{k\}))$.

Lemma 5.7 (Lower bound on $\textsf{val}_\Psi$)**.** *Let $f : \{-1, 1\}^k \to \{0, 1\}$ be a Boolean function, and Ψ be an instance of Max-CSP(f). Then*

$$\textsf{val}_\Psi \geq \rho(f) + B_{\boldsymbol{\lambda}}(\Psi)\delta(\Psi)\,,$$

where $\boldsymbol{\lambda} = (\widehat{f}(\{1\}), \ldots, \widehat{f}(\{k\}))$ and $\delta(\Psi) = \min\{\frac{1}{3k}, \frac{2B_{\boldsymbol{\lambda}}(\Psi)}{9\rho(f)k^2}\}$.

Lemma 5.8 (Upper bound on $\textsf{val}_\Psi$)**.** *Let $f : \{-1, 1\}^k \to \{0, 1\}$ be a Boolean function, $\varepsilon_0(f)$ be as defined in Definition 5.2, and Ψ be an instance of Max-CSP(f). Then*

$$\textsf{val}_\Psi \leq \frac{B_{\boldsymbol{\lambda}}(\Psi) + \rho(f) \cdot k}{\varepsilon_0(f) + \rho(f) \cdot k}\,,$$

where $\boldsymbol{\lambda} = (\widehat{f}(\{1\}), \ldots, \widehat{f}(\{k\}))$.

We defer the proofs of Lemma 5.7 and Lemma 5.8 to Section 5.2. We now show the correctness of Algorithm 1 using these lemmas.

5.1.2 Proof of Theorem 5.3

Proof of Theorem 5.3. First, by Lemma 5.6, with probability at least 2/3, $\tilde{B}$ is a $(1 \pm \varepsilon')$ approximation to $B_{\boldsymbol{\lambda}}(\Psi)$, i.e., $(1 - \varepsilon')B_{\boldsymbol{\lambda}}(\Psi) \leq \tilde{B} \leq (1 + \varepsilon')B_{\boldsymbol{\lambda}}(\Psi)$. Next, we show that with probability at least 2/3, (i) $v \leq \textsf{val}_\Psi$ and (ii) $v \geq (\rho(f) + \varepsilon^*(f) - \varepsilon) \cdot \textsf{val}_\Psi$.

(i) $v \leq \textsf{val}_\Psi$. We have

$$v = \rho(f) + \frac{\tilde{B}\tilde{\delta}}{(1 + \varepsilon')^2} \leq \rho(f) + B_{\boldsymbol{\lambda}}(\Psi)\delta(\Psi) \leq \textsf{val}_\Psi\,,$$

where the last inequality follows from Lemma 5.7.

(ii) $v \geq (\rho(f) + \varepsilon^*(f) - \varepsilon) \cdot \mathsf{val}_\Psi$. We have

$$v = \rho(f) + \frac{\tilde{B}\tilde{\delta}}{(1+\varepsilon')^2} \geq \rho(f) + B_{\boldsymbol{\lambda}}(\Psi)\delta(\Psi)\left(\frac{1-\varepsilon'}{1+\varepsilon'}\right)^2 \geq \rho(f) + B_{\boldsymbol{\lambda}}(\Psi)\delta(\Psi)(1-\varepsilon)\,, \tag{5.9}$$

where the last inequality follows from the choice of ε'. Let us first consider the case when $B_{\boldsymbol{\lambda}}(\Psi) \geq \varepsilon_0(f)$. We have

$$B_{\boldsymbol{\lambda}}(\Psi)\delta(\Psi) \geq \varepsilon_0(f) \cdot \min\left\{\frac{1}{3k}, \frac{2\varepsilon_0(f)}{9\rho(f)k^2}\right\} \geq \varepsilon^*\,, \tag{5.10}$$

where the last equality follows from the definition of $\varepsilon^*(f)$ in Definition 5.2.

Combining Eq. (5.9) and Eq. (5.10), we get

$$v \geq \rho(f) + \varepsilon^*(f)(1-\varepsilon) \geq (\rho(f) + \varepsilon^*(f) - \varepsilon)\mathsf{val}_\Psi\,,$$

where the last inequality follows from $\mathsf{val}_\Psi \leq 1$.

Now, let us consider the case when $B_{\boldsymbol{\lambda}}(\Psi) < \varepsilon_0(f)$. It follows from Proposition 2.11 that $\varepsilon_0(f) \leq \rho(f)k$. Therefore,

$$\frac{2B_{\boldsymbol{\lambda}}(\Psi)}{9\rho(f)k^2} \leq \frac{2\varepsilon_0(f)}{9\rho(f)k^2} \leq \frac{2}{9k} < \frac{1}{3k}\,,$$

and so $\delta(\Psi) = \frac{2B_{\boldsymbol{\lambda}}(\Psi)}{9\rho(f)k^2}$. Combining Eq. (5.9) and Lemma 5.8, we have

$$\frac{v}{\mathsf{val}_\Psi} \geq (1-\varepsilon)\left(\frac{\rho(f) + \frac{2B_{\boldsymbol{\lambda}}(\Psi)^2}{9\rho(f)k^2}}{\rho(f) + \frac{B_{\boldsymbol{\lambda}}(\Psi)}{k}}\right)\left(\rho(f) + \frac{\varepsilon_0(f)}{k}\right).$$

We show that for $0 \leq B_{\boldsymbol{\lambda}}(\Psi) \leq \varepsilon_0(f)$,

$$\frac{\rho(f) + \frac{2B_{\boldsymbol{\lambda}}(\Psi)^2}{9\rho(f)k^2}}{\rho(f) + \frac{B_{\boldsymbol{\lambda}}(\Psi)}{k}} \geq \frac{\rho(f) + \frac{2\varepsilon_0(f)^2}{9\rho(f)k^2}}{\rho(f) + \frac{\varepsilon_0(f)}{k}}\,. \tag{5.11}$$

This immediately implies that

$$\frac{v}{\mathsf{val}_\Psi} \geq (1-\varepsilon)\left(\rho(f) + \frac{2\varepsilon_0(f)^2}{9\rho(f)k^2}\right) \geq (1-\varepsilon)(\rho(f) + \varepsilon^*(f)) > \rho(f) + \varepsilon^*(f) - \varepsilon\,.$$

Consider the function $g(p) = \frac{\rho(f) + \frac{2p^2}{9\rho(f)}}{\rho(f)+p}$. In order to show Eq. (5.11), it suffices to show that in the range $p \in [0, \frac{\varepsilon_0(f)}{k}]$, $g(p)$ attains the minimum value at $p = \frac{\varepsilon_0(f)}{k}$, i.e, $g'(p) < 0$ in this range. We have $g'(p) = \frac{\left(\frac{2(p+\rho(f))^2}{9\rho(f)} - \frac{11\rho(f)}{9}\right)}{(\rho(f)+p)^2}$ and for $p \in [0, \frac{\varepsilon_0(f)}{k}]$, we have

$$\left(\frac{2(p+\rho(f))^2}{9\rho(f)} - \frac{11\rho(f)}{9}\right) \leq \left(\frac{2(\varepsilon_0(f)/k + \rho(f))^2}{9\rho(f)} - \frac{11\rho(f)}{9}\right) \leq \frac{8\rho(f)}{9} - \frac{11\rho(f)}{9} = -\frac{\rho(f)}{3} < 0\,.$$

This completes the proof of Theorem 5.3. □

5.2 Proofs of Lemma 5.7 and Lemma 5.8

In this section, we prove Lemma 5.7 and Lemma 5.8.

Proof of Lemma 5.7. Let $\mathsf{Bern}(p) \in \Delta(\{-1,1\})$ denote the Bernoulli distribution where 1 is sampled with probability p. Given an instance $\Psi = (C_1, \ldots, C_m; w_1, \ldots, w_m)$ of $\mathsf{Max\text{-}CSP}(f)$ where $C_i = (\mathbf{j}(i), \mathbf{b}(i))$ and $w_i \geq 0$, let $\gamma = 3 \cdot \delta(\Psi) = \min\{\frac{1}{k}, \frac{2B_{\boldsymbol{\lambda}}(\psi)}{3\rho(f)k^2}\}$. Let $\boldsymbol{\sigma} = \arg\max_{\mathbf{a}\in\{-1,1\}^n} \langle \mathbf{a}, \mathsf{bias}_{\boldsymbol{\lambda}}(\Psi)\rangle$. It follows from Lemma 5.5 that $B_{\boldsymbol{\lambda}}(\Psi) = \langle \boldsymbol{\sigma}, \mathsf{bias}_{\boldsymbol{\lambda}}(\Psi)\rangle$. In order to prove the lemma, we will show that

$$\mathbb{E}_{\mathbf{a}\sim\left(\mathsf{Bern}(\frac{1+\gamma}{2})\right)^n}[\mathsf{val}_\Psi(\mathbf{a}\odot\boldsymbol{\sigma})] \geq \rho(f) + B_{\boldsymbol{\lambda}}(\Psi)\delta(\Psi)\,.$$

The lemma then directly follows from the fact that $\mathsf{val}_\Psi \geq \mathbb{E}_{\mathbf{a}\sim\left(\mathsf{Bern}(\frac{1+\gamma}{2})\right)^n}[\mathsf{val}_\Psi(\mathbf{a}\odot\boldsymbol{\sigma})]$.

We have

$$\begin{aligned}
\mathbb{E}_{\mathbf{a}\sim\left(\mathsf{Bern}(\frac{1+\gamma}{2})\right)^n}[\mathsf{val}_\Psi(\mathbf{a}\odot\boldsymbol{\sigma})] &= \mathbb{E}_{\mathbf{a}\sim\left(\mathsf{Bern}(\frac{1+\gamma}{2})\right)^n}\left[\frac{1}{W}\sum_{i=1}^m w_i\cdot f(\mathbf{a}|_{\mathbf{j}(i)}\odot\boldsymbol{\sigma}|_{\mathbf{j}(i)}\odot\mathbf{b}(i))\right]\\
&= \mathbb{E}_{\mathbf{a}\sim\left(\mathsf{Bern}(\frac{1+\gamma}{2})\right)^n}\left[\frac{1}{W}\sum_{i=1}^m w_i\cdot \sum_{S\subseteq[k]}\widehat{f}(S)\cdot\chi_S(\mathbf{a}|_{\mathbf{j}(i)}\odot\boldsymbol{\sigma}|_{\mathbf{j}(i)}\odot\mathbf{b}(i))\right]\\
&\qquad\text{(Fourier expansion of } f)\\
&= \frac{1}{W}\sum_{i=1}^m w_i\cdot\sum_{S\subseteq[k]}\widehat{f}(S)\cdot\mathbb{E}_{\mathbf{a}\sim\left(\mathsf{Bern}(\frac{1+\gamma}{2})\right)^n}\left[\chi_S(\mathbf{a}|_{\mathbf{j}(i)}\odot\boldsymbol{\sigma}|_{\mathbf{j}(i)}\odot\mathbf{b}(i))\right]\\
&\qquad\text{(Linearity of expectation)}\\
&= \frac{1}{W}\sum_{i=1}^m w_i\cdot\sum_{S\subseteq[k]}\widehat{f}(S)\cdot\chi_S(\mathbf{b}(i)\odot\boldsymbol{\sigma}|_{\mathbf{j}(i)})\cdot\mathbb{E}_{\mathbf{a}\sim\left(\mathsf{Bern}(\frac{1+\gamma}{2})\right)^n}\left[\chi_S(\mathbf{a}|_{\mathbf{j}(i)})\right]\\
&\qquad\text{(Since } \chi_S(a\odot b) = \chi_S(a)\cdot\chi_S(b))\\
&= \frac{1}{W}\sum_{i=1}^m w_i\cdot\sum_{S\subseteq[k]}\widehat{f}(S)\cdot\chi_S(\mathbf{b}(i)\odot\boldsymbol{\sigma}|_{\mathbf{j}(i)})\cdot\gamma^{|S|}\\
&\qquad\text{(Since } \mathbb{E}_{\mathbf{a}\sim\left(\mathsf{Bern}(\frac{1+\gamma}{2})\right)^n}[a_\ell] = \gamma \text{ for all } \ell\in[n])\\
&= \widehat{f}(\emptyset) + \frac{1}{W}\sum_{i\in[m]} w_i\sum_{t\in[k]}\widehat{f}(\{t\})\cdot b(i)_t\cdot\sigma_{j(i)_t}\cdot\gamma\\
&\quad + \frac{1}{W}\sum_{i=1}^m w_i\cdot\sum_{S\subseteq[k]:|S|\geq 2}\widehat{f}(S)\cdot\chi_S(\mathbf{b}(i)\odot\boldsymbol{\sigma}|_{\mathbf{j}(i)})\cdot\gamma^{|S|}\\
&= \widehat{f}(\emptyset) + \sum_{\ell\in[n]}\left(\frac{1}{W}\sum_{i\in[m],t\in[k]:j(i)_t=\ell}\widehat{f}(\{t\})\cdot w_i\cdot b(i)_t\right)\sigma_\ell\cdot\gamma\\
&\quad + \frac{1}{W}\sum_{i=1}^m w_i\cdot\sum_{S\subseteq[k]:|S|\geq 2}\widehat{f}(S)\cdot\chi_S(\mathbf{b}(i)\odot\boldsymbol{\sigma}|_{\mathbf{j}(i)})\cdot\gamma^{|S|}
\end{aligned}$$

(Rearranging the summations)

$$\geq \widehat{f}(\emptyset) + \gamma\langle \mathsf{bias}_{\boldsymbol{\lambda}}(\Psi), \boldsymbol{\sigma}\rangle - \sum_{S\subseteq[k]:|S|\geq 2} |\widehat{f}(S)| \cdot \gamma^{|S|}$$

(By the definition of $\boldsymbol{\lambda}$, $\mathsf{bias}_{\boldsymbol{\lambda}}(\Psi)$), and $|\chi_S(\cdot)| \leq 1$)

$$\geq \rho(f) + \gamma \cdot B_{\boldsymbol{\lambda}}(\Psi) - \rho(f) \sum_{S\subseteq[k]:|S|\geq 2} \gamma^{|S|}$$

(By the definition of $\boldsymbol{\sigma}$ and Proposition 2.11)

$$= \rho(f) + \gamma \cdot B_{\boldsymbol{\lambda}}(\Psi) - \rho(f) \sum_{r=2}^{k} \binom{k}{r} \cdot \gamma^r .$$

We now prove that $\rho(f)\sum_{r=2}^{k}\binom{k}{r} \cdot \gamma^r \leq \frac{2\gamma}{3} \cdot B_{\boldsymbol{\lambda}}(\Psi)$. Consider the combinatorial identity $\binom{k}{r} = \frac{k\cdot\binom{k-1}{r-1}}{r}$. Since $\gamma \leq \frac{1}{k}$ and $r \geq 2$, we have

$$\binom{k}{r}\gamma^r = \frac{k \cdot \binom{k-1}{r-1}}{r} \cdot \gamma^r \leq \frac{1}{2} \cdot \binom{k-1}{r-1} \cdot \gamma^{r-1} < \frac{1}{2} \cdot \binom{k}{r-1} \cdot \gamma^{r-1} .$$

Hence $\sum_{r=2}^{k}\binom{k}{r} \cdot \gamma^r \leq 2 \cdot \binom{k}{2} \cdot \gamma^2$. Since $\gamma \leq \frac{2B_{\boldsymbol{\lambda}}(\Psi)}{3\rho(f)k^2}$, we have

$$\rho(f) \cdot \sum_{r=2}^{k} \binom{k}{r} \cdot \gamma^r \leq \rho(f) \cdot 2 \cdot \binom{k}{2} \cdot \gamma^2 \leq \frac{2B_{\boldsymbol{\lambda}}(\Psi)\cdot\gamma}{3} .$$

Recall that $\gamma = 3\delta(\Psi)$. Finally, we conclude that

$$\mathsf{val}_\Psi \geq \mathbb{E}_{\mathbf{a}\sim\left(\mathsf{Bern}(\frac{1+\gamma}{2})\right)^n}[\mathsf{val}_\Psi(\mathbf{a}\odot\boldsymbol{\sigma})] \geq \rho(f) + \frac{\gamma}{3} \cdot B_{\boldsymbol{\lambda}}(\Psi) = \rho(f) + B_{\boldsymbol{\lambda}}(\Psi)\delta(\Psi) .$$

□

Proof of Lemma 5.8. Let $\Psi = (C_1, \ldots, C_m; w_1, \ldots, w_m)$ be an instance of $\mathsf{Max\text{-}CSP}(f)$ where $C_i = (\mathbf{j}(i), \mathbf{b}(i))$ and $w_i \geq 0$. Let $\mathbf{a}^* \in \{-1,1\}^n$ denote the assignment that satisfies the maximum weight of constraints in Ψ, i.e., $\mathbf{a}^* = \arg\max_{\mathbf{a}\in\{-1,1\}^n} \mathsf{val}_\Psi(\mathbf{a})$. It follows from Lemma 5.5 that $B_{\boldsymbol{\lambda}}(\Psi) \geq \langle \mathbf{a}^*, \mathsf{bias}_{\boldsymbol{\lambda}}(\Psi)\rangle$. Let S be the set of indices corresponding to constraints of Ψ satisfied by $\mathbf{a}^*$, i.e., $S = \{i \in [m] : f(\mathbf{a}^*|_{\mathbf{j}(i)} \odot \mathbf{b}(i)) = 1\}$. We have

$$\langle \mathbf{a}^*, \mathsf{bias}_{\boldsymbol{\lambda}}(\Psi)\rangle = \sum_{\ell\in[n]} a^*_\ell \cdot \frac{1}{W} \cdot \sum_{i\in[m],t\in[k]:j(i)_t=\ell} \lambda_t w_i b(i)_t$$

$$= \frac{1}{W}\sum_{i\in[m]} w_i \sum_{t\in[k]} \lambda_t \cdot b(i)_t \cdot a^*_{j(i)_t}$$

(Exchanging the summations)

$$= \frac{1}{W}\sum_{i\in[m]} w_i \sum_{t\in[k]} \widehat{f}(\{t\}) \cdot b(i)_t \cdot a^*_{j(i)_t}$$

$$
\begin{aligned}
&\qquad\qquad (\ \boldsymbol{\lambda} = (\widehat{f}(\{1\}), \dots, \widehat{f}(\{k\}))\ ) \\
&= \frac{1}{W}\sum_{i\in S} w_i \sum_{t\in[k]} \widehat{f}(\{t\})\cdot b(i)_t \cdot a^*_{j(i)_t} + \frac{1}{W}\sum_{i\notin S} w_i \sum_{t\in[k]} \widehat{f}(\{t\})\cdot b(i)_t \cdot a^*_{j(i)_t} \\
&\geq \frac{1}{W}\sum_{i\in S} w_i \cdot \varepsilon_0(f) - \frac{1}{W}\sum_{i\notin S} w_i \cdot \rho(f)\cdot k \\
&\qquad\qquad \text{(By the definition of } S \text{ and } \varepsilon_0(f)\text{, and Proposition 2.11)} \\
&= \mathsf{val}_\Psi \cdot \varepsilon_0(f) - (1-\mathsf{val}_\Psi)\rho(f)\cdot k \\
&\qquad\qquad \text{(By the definition of } \mathbf{a}^*)\,.
\end{aligned}
$$

Therefore, we get

$$
B_{\boldsymbol{\lambda}}(\Psi) \geq \mathsf{val}_\Psi \cdot \varepsilon_0(f) - (1-\mathsf{val}_\Psi)\rho(f)\cdot k\,.
$$

Rearranging the terms, we get

$$
\mathsf{val}_\Psi \leq \frac{B_{\boldsymbol{\lambda}}(\Psi) + \rho(f)\cdot k}{\varepsilon_0(f) + \rho(f)\cdot k}\,.
$$

□

5.3 Approximability of weak monarchy functions

In this section, we analyze the streaming approximability of $\mathsf{Max\text{-}CSP}(f)$ where f is a weak monarchy function. Note that in order for $\mathsf{WMON}_{k,j}$ to be a balanced LTF, the total number of votes, i.e., $j+k-1$, needs to be odd. Therefore, we make such assumption throughout the rest of this section.

Lemma 5.12. *For all integers $j \geq 2$ and $k \geq 7j^3$ such that $k+j$ is even,*

$$
\mathsf{WMON}_{k,j}(x) = \mathsf{sign}\left(\sum_{i=1}^{k} \widehat{\mathsf{WMON}}_{k,j}(\{i\})x_i\right).
$$

Note that Lemma 5.12 along with Theorem 5.1 directly conclude Theorem 1.6 restated below.

Theorem 1.6. *For all integers $j \geq 2$ and $k \geq 7j^3$ such that $k+j$ is even, $\mathsf{Max\text{-}CSP}(\mathsf{WMON}_{k,j})$ is sketching approximable in $O(\log(n))$ space. In particular, for every j, there exist infinitely many k such that $\mathsf{Max\text{-}CSP}(\mathsf{WMON}_{k,j})$ is sketching approximable.*

Proof of Lemma 5.12. We start by finding the Chow parameters of $\mathsf{WMON}_{k,j}$. As mentioned earlier, we only consider the case where $k+j$ is even. For the president,

$$
\begin{aligned}
\widehat{\mathsf{WMON}}_{k,j}(\{1\}) \;&=\; \Pr\{x_1 = 1, \mathsf{WMON}_{k,j}(x) = 1\}\times 1 \\
&\quad\; \Pr\{x_1 = -1, \mathsf{WMON}_{k,j}(x) = 1\}\times (-1) \\
&\quad\; \Pr\{x_1 = 1, \mathsf{WMON}_{k,j}(x) = 0\}\times 0 \\
&\quad\; \Pr\{x_1 = -1, \mathsf{WMON}_{k,j}(x) = 0\}\times 0 \\
&=\; \frac{1}{2^k}\left(\binom{k-1}{\geq \frac{k+j}{2} - j} - \binom{k-1}{\geq \frac{k+j}{2}}\right)
\end{aligned}
$$

$$= \frac{1}{2^k}\left(\binom{k-1}{\geq \frac{k-j}{2}} - \binom{k-1}{\geq \frac{k+j}{2}}\right)$$
$$= \frac{1}{2^k}\left(2^{k-1} - 2\binom{k-1}{< \frac{k-j}{2}}\right).$$

For citizen x_i ($i > 1$),

$$\begin{aligned}
\widehat{\mathsf{WMON}}_{k,j}(\{i\}) &= \Pr\{x_1 = 1, x_i = 1, \mathsf{WMON}_{k,j}(x) = 1\} \times 1 \\
&\quad \Pr\{x_1 = 1, x_i = -1, \mathsf{WMON}_{k,j}(x) = 1\} \times (-1) \\
&\quad \Pr\{x_1 = -1, x_i = 1, \mathsf{WMON}_{k,j}(x) = 1\} \times 1 \\
&\quad \Pr\{x_1 = -1, x_i = -1, \mathsf{WMON}_{k,j}(x) = 1\} \times (-1) \\
&\quad \Pr\{x_1 = 1, x_i = 1, \mathsf{WMON}_{k,j}(x) = 0\} \times 0 \\
&\quad \Pr\{x_1 = 1, x_i = -1, \mathsf{WMON}_{k,j}(x) = 0\} \times 0 \\
&\quad \Pr\{x_1 = -1, x_i = 1, \mathsf{WMON}_{k,j}(x) = 0\} \times 0 \\
&\quad \Pr\{x_1 = -1, x_i = -1, \mathsf{WMON}_{k,j}(x) = 0\} \times 0 \\
&= \frac{1}{2^k}\left(\binom{k-2}{\geq \frac{k+j}{2} - j - 1} - \binom{k-2}{\geq \frac{k+j}{2} - j} + \binom{k-2}{\geq \frac{k+j}{2} - 1} - \binom{k-2}{\geq \frac{k+j}{2}}\right) \\
&= \frac{1}{2^k}\left(\binom{k-2}{\frac{k+j}{2} - j - 1} + \binom{k-2}{\frac{k+j}{2} - 1}\right) \\
&= \frac{\binom{k-2}{\frac{k-j}{2}-1}}{2^{k-1}}.
\end{aligned}$$

Note that in order for functions $\mathsf{WMON}_{k,j}(x)$ and $\mathsf{sign}\left(\sum_{i=1}^{k} \widehat{\mathsf{WMON}}_{k,j}(\{i\})x_i\right)$ to be the same, it suffices to have

$$j - 1 < \frac{\widehat{\mathsf{WMON}}_{k,j}(\{1\})}{\widehat{\mathsf{WMON}}_{k,j}(\{i\})} < j + 1.$$ [6]

Thus, in the rest of the proof, we find values for k that guarantee the bounds above. We start with the upper-bound:

$$\frac{\widehat{\mathsf{WMON}}_{k,j}(\{1\})}{\widehat{\mathsf{WMON}}_{k,j}(\{i\})} = \frac{2^{k-2} - \binom{k-1}{< \frac{k-j}{2}}}{\binom{k-2}{\frac{k-j}{2}-1}} \leq \frac{\frac{j}{2} \cdot \binom{k-1}{\lfloor \frac{k-1}{2} \rfloor}}{\binom{k-2}{\frac{k-j}{2}-1}}. \tag{5.13}$$

The last inequality holds as below:

- If $k-1$ is odd: $2^{k-2} = \sum_{i=0}^{\frac{k-2}{2}} \binom{k-1}{i} \Rightarrow 2^{k-2} - \binom{k-1}{< \frac{k-j}{2}} = \sum_{i=\frac{k-j}{2}}^{\frac{k-2}{2}} \binom{k-1}{i} \leq \frac{j}{2} \cdot \binom{k-1}{\frac{k-2}{2}}$
- If $k-1$ is even: $2^{k-2} = \frac{1}{2}\binom{k-1}{\frac{k-1}{2}} + \sum_{i=0}^{\frac{k-3}{2}} \binom{k-1}{i} \Rightarrow 2^{k-2} - \binom{k-1}{< \frac{k-j}{2}} \leq \frac{j}{2} \cdot \binom{k-1}{\frac{k-1}{2}}$

[6]Indeed, letting $w = \frac{\widehat{\mathsf{WMON}}_{k,j}(\{1\})}{\widehat{\mathsf{WMON}}_{k,j}(\{i\})}$, if $j - 1 < w < j + 1$, then we have that $\mathsf{sign}\left(\sum_{i=1}^{k} \widehat{\mathsf{WMON}}_{k,j}(\{i\})x_i\right) = \mathsf{sign}\left(wx_1 + \sum_{i=2}^{k} x_i\right) = \mathsf{sign}\left(jx_1 + \sum_{i=2}^{k} x_i\right) = \mathsf{WMON}_{k,j}(x)$.

Therefore,

$$\begin{aligned}
\frac{\widehat{\mathsf{WMON}}_{k,j}(\{1\})}{\widehat{\mathsf{WMON}}_{k,j}(\{i\})} &\le \frac{j}{2}\cdot\frac{\frac{(k-1)!}{(\lfloor\frac{k-1}{2}\rfloor)!(\lceil\frac{k-1}{2}\rceil)!}}{\frac{(k-2)!}{(\frac{k-j}{2}-1)!(\frac{k+j}{2}-1)!}} = \frac{j}{2}\cdot\frac{k-1}{\lfloor\frac{k-1}{2}\rfloor}\cdot\frac{(\frac{k+j}{2}-1)\cdots(\lceil\frac{k-1}{2}\rceil+1)}{(\lfloor\frac{k-1}{2}-1\rfloor)\cdots(\frac{k-j}{2})} \\
&\le \frac{j}{2}\cdot 2\left(1+\frac{1}{k-2}\right)\cdot\left(\frac{\lceil\frac{k-1}{2}\rceil+1}{\frac{k-j}{2}}\right)^{\lfloor\frac{j-1}{2}\rfloor} \\
&\le j\cdot\left(1+\frac{1}{k-2}\right)\cdot\left(\frac{k+2}{k-j}\right)^{\frac{j-1}{2}} \\
&= j\cdot\left(1+\frac{1}{k-2}\right)\cdot\left(1+\frac{j+2}{k-j}\right)^{\frac{j-1}{2}} \le j\cdot\left(1+\frac{j+2}{k-j}\right)^{j}.
\end{aligned}$$

For any given j, $\left(1+\frac{j+2}{k-j}\right)^j$ tends to 1 as k goes to ∞. Therefore, there exits some K_0 such that for all $k \ge K_0$, $\frac{\widehat{\mathsf{WMON}}_{k,j}(\{1\})}{\widehat{\mathsf{WMON}}_{k,j}(\{i\})} < j+1$. More precisely, we take k to be at least $K_0 = 2j^3+4j^2+j \le 7j^3$. This way,

$$\begin{aligned}
\frac{\widehat{\mathsf{WMON}}_{k,j}(\{1\})}{\widehat{\mathsf{WMON}}_{k,j}(\{i\})} &\le j\cdot\left(1+\frac{1}{2j^2}\right)^j \le j\cdot\left(1+j\cdot\frac{1}{2j^2}+j^2\cdot\frac{1}{(2j^2)^2}+j^3\cdot\frac{1}{(2j^2)^3}+\cdots\right) \\
&= j\cdot\left(1+\frac{1}{2j}+\frac{1}{(2j)^2}+\frac{1}{(2j)^3}+\cdots\right) \\
&< j\cdot\left(1+\frac{1}{j}\right).
\end{aligned}$$

We now proceed to the lower bound.

$$\frac{\widehat{\mathsf{WMON}}_{k,j}(\{1\})}{\widehat{\mathsf{WMON}}_{k,j}(\{i\})} = \frac{2^{k-2}-\binom{k-1}{<\frac{k-j}{2}}}{\binom{k-2}{\frac{k-j}{2}-1}} \ge \frac{\frac{j}{2}\cdot\binom{k-1}{\frac{k-j}{2}}}{\binom{k-2}{\frac{k-j}{2}-1}}. \tag{5.14}$$

Similar to the upper-bound case, the last inequality can be observed as follows:

- If $k-1$ is odd: $2^{k-2} = \sum_{i=0}^{\frac{k-2}{2}}\binom{k-1}{i} \Rightarrow 2^{k-2}-\binom{k-1}{<\frac{k-j}{2}} = \sum_{i=\frac{k-j}{2}}^{\frac{k-2}{2}}\binom{k-1}{i} \ge \frac{j}{2}\cdot\binom{k-1}{\frac{k-j}{2}}$
- If $k-1$ is even: $2^{k-2} = \frac{1}{2}\binom{k-1}{\frac{k-1}{2}} + \sum_{i=0}^{\frac{k-3}{2}}\binom{k-1}{i} \Rightarrow 2^{k-2}-\binom{k-1}{<\frac{k-j}{2}} \ge \frac{j}{2}\cdot\binom{k-1}{\frac{k-j}{2}}$

Therefore,

$$\begin{aligned}
\frac{\widehat{\mathsf{WMON}}_{k,j}(\{1\})}{\widehat{\mathsf{WMON}}_{k,j}(\{i\})} &\ge \frac{j}{2}\cdot\frac{\binom{k-1}{\frac{k-j}{2}}}{\binom{k-2}{\frac{k-j}{2}-1}} = \frac{j}{2}\cdot\frac{\binom{k-2}{\frac{k-j}{2}}+\binom{k-2}{\frac{k-j}{2}-1}}{\binom{k-2}{\frac{k-j}{2}-1}} = \frac{j}{2}\cdot\left(1+\frac{\binom{k-2}{\frac{k-j}{2}}}{\binom{k-2}{\frac{k-j}{2}-1}}\right) \\
&= \frac{j}{2}\cdot\left(1+\frac{\frac{k+j}{2}-1}{\frac{k-j}{2}}\right) = \frac{j}{2}\cdot\left(1+\frac{k+j-2}{k-j}\right)
\end{aligned}$$

$$= \quad j \cdot \left(1 + \frac{j-1}{k-j}\right) .$$

This lower bound is larger than j for every $k > j$. Thus, for every $k \geq 2j^3 + 4j^2 + j$, $j \leq \frac{\widehat{\mathsf{WMON}}_{k,j}(\{1\})}{\widehat{\mathsf{WMON}}_{k,j}(\{i\})} < j+1$ which implies that $\mathsf{WMON}_{k,j}(x) = \mathsf{sign}\left(\sum_{i=1}^{k} \widehat{\mathsf{WMON}}_{k,j}(\{i\}) x_i\right)$, and concludes the proof. □

Acknowledgments

We thank the anonymous reviewers for their helpful and constructive comments.

References

[ABM10] Per Austrin, Siavosh Benabbas, and Avner Magen. On quadratic threshold csps. In *LATIN 2010*, pages 332–343. Springer, 2010.

[BHP+22] Joanna Boyland, Michael Hwang, Tarun Prasad, Noah Singer, and Santhoshini Velusamy. Closed-form expressions for the sketching approximability of (some) symmetric Boolean CSPs. *CoRR*, abs/2112.06319, February 2022.

[CGS+22] Chi-Ning Chou, Alexander Golovnev, Madhu Sudan, Ameya Velingker, and Santhoshini Velusamy. Linear Space Streaming Lower Bounds for Approximating CSPs. In *STOC 2022*, 2022. To appear.

[CGSV21] Chi-Ning Chou, Alexander Golovnev, Madhu Sudan, and Santhoshini Velusamy. Approximability of all finite CSPs with linear sketches. In *FOCS 2021*, pages 1197–1208. IEEE, 2021.

[CGSV22] Chi-Ning Chou, Alexander Golovnev, Madhu Sudan, and Santhoshini Velusamy. Approximability of all Boolean CSPs with linear sketches. *CoRR*, abs/2102.12351v8, 11th February 2022.

[CGV20] Chi-Ning Chou, Alexander Golovnev, and Santhoshini Velusamy. Optimal Streaming Approximations for all Boolean Max-2CSPs and Max-kSAT. In *FOCS 2020*, pages 330–341. IEEE, 2020.

[GT19] Venkatesan Guruswami and Runzhou Tao. Streaming Hardness of Unique Games. In *APPROX 2019*, pages 5:1–5:12. Schloss Dagstuhl, 2019.

[GVV17] Venkatesan Guruswami, Ameya Velingker, and Santhoshini Velusamy. Streaming Complexity of Approximating Max 2CSP and Max Acyclic Subgraph. In *APPROX 2017*, pages 8:1–8:19. Schloss Dagstuhl, 2017.

[Has05] Gustav Hast. *Beating a random assignment: Approximating constraint satisfaction problems.* PhD thesis, KTH, 2005.

[HP20] Neng Huang and Aaron Potechin. On the approximability of presidential type predicates. In *APPROX 2020*, pages 58:1–58:20. Schloss Dagstuhl, 2020.

[Ind00] Piotr Indyk. Stable distributions, pseudorandom generators, embeddings and data stream computation. In *FOCS 2000*, pages 189–197. IEEE, 2000.

[KK15] Dmitry Kogan and Robert Krauthgamer. Sketching cuts in graphs and hypergraphs. In *ITCS 2015*, pages 367–376. ACM, 2015.

[KK19] Michael Kapralov and Dmitry Krachun. An optimal space lower bound for approximating MAX-CUT. In *STOC 2019*, pages 277–288. ACM, 2019.

[KKS15] Michael Kapralov, Sanjeev Khanna, and Madhu Sudan. Streaming lower bounds for approximating MAX-CUT. In *SODA 2015*, pages 1263–1282. SIAM, 2015.

[KKSV17] Michael Kapralov, Sanjeev Khanna, Madhu Sudan, and Ameya Velingker. $(1 + \omega(1))$-approximation to MAX-CUT requires linear space. In *SODA 2017*, pages 1703–1722. SIAM, 2017.

[KNW10] Daniel M. Kane, Jelani Nelson, and David P. Woodruff. On the exact space complexity of sketching and streaming small norms. In *SODA 2010*, pages 1161–1178. SIAM, 2010.

[O'D14] Ryan O'Donnell. *Analysis of Boolean functions*. Cambridge University Press, 2014.

[Pot19] Aaron Potechin. On the approximation resistance of balanced linear threshold functions. In *STOC 2019*, pages 430–441. ACM, 2019.

[SSV21] Noah Singer, Madhu Sudan, and Santhoshini Velusamy. Streaming approximation resistance of every ordering CSP. In *APPROX 2021*, pages 17:1–17:19. Schloss Dagstuhl, 2021.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il