

Sketching Approximability of (Weak) Monarchy Predicates

Chi-Ning Chou

School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA

Alexander Golovnev

Department of Computer Science, Georgetown University, Washington, D.C., USA

Amirbehshad Shahrabi

Microsoft, Redmond, WA, USA

Madhu Sudan

School of Engineering and Applied Sciences, Harvard University, Cambridge, MA USA

Santhoshini Velusamy

School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA

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.

2012 ACM Subject Classification Theory of computation \rightarrow Sketching and sampling; Theory of computation \rightarrow Approximation algorithms analysis

Keywords and phrases sketching algorithms, approximability, linear threshold functions

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2022.35

Category APPROX

Related Version *Full Version:* <https://arxiv.org/abs/2205.02345>

Funding *Chi-Ning Chou:* Partially supported by NSF grants DMS-2134157 and CCF-1565264, DOE grant DE-SC0022199, and the Simons foundation.

Amirbehshad Shahrabi: The author was with Harvard University and supported by CRA-CCC Computing Innovations Fellowship (CIFellowship 2020) during this work.

Madhu Sudan: Supported in part by a Simons Investigator Award and NSF Awards CCF 1715187 and CCF 2152413.

Santhoshini Velusamy: Supported in part by a Google Ph.D. Fellowship, a Simons Investigator Award to Madhu Sudan, and NSF Awards CCF 1715187 and CCF 2152413.

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



© Chi-Ning Chou, Alexander Golovnev, Amirbehshad Shahrabi, Madhu Sudan, and Santhoshini Velusamy;

licensed under Creative Commons License CC-BY 4.0

Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2022).

Editors: Amit Chakrabarti and Chaitanya Swamy; Article No. 35; pp. 35:1–35:17



Leibniz International Proceedings in Informatics

LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

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 $f : \{-1, 1\}^k \rightarrow \{0, 1\}$, the Boolean CSP associated with f , denoted $\text{Max-CSP}(f)$ is the following optimization problem. Given m constraints C_1, \dots, C_m on n Boolean variables X_1, \dots, X_n , where each constraint applies f to a sequence of k distinct literals from the set $\{X_1, \dots, X_n, -X_1, \dots, -X_n\}$, find the maximum fraction of constraints that can be satisfied by an assignment to the n variables. For an instance Ψ of $\text{Max-CSP}(f)$ 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 $\text{Max-CSP}(f)$. For $0 \leq \beta < \gamma \leq 1$ we define the (γ, β) -gapped version of $\text{Max-CSP}(f)$, abbreviated to (γ, β) - $\text{Max-CSP}(f)$, to be the following promise decision problem: Given an instance Ψ satisfying $\text{val}_\Psi \geq \gamma$ or $\text{val}_\Psi < \beta$ 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 its 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 $\text{Max-CSP}(f)$ if the following holds: on every input instance Ψ , the algorithm outputs v such that $\alpha \cdot \text{val}_\Psi \leq v \leq \text{val}_\Psi$ 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) - $\text{Max-CSP}(f)$ for every $\gamma, \beta \in [0, 1]$ with $\beta \leq \alpha \cdot \gamma$.

For a function $f : \{-1, 1\}^k \rightarrow \{0, 1\}$, define $\rho(f) = 2^{-k} \cdot |\{x \in \{-1, 1\}^k \mid f(x) = 1\}|$. For every f and every instance Ψ of $\text{Max-CSP}(f)$, a random assignment satisfies $\rho(f)$ fraction of the constraints in expectation and so every Ψ satisfies $\text{val}_\Psi \geq \rho(f)$. Thus the $(1, \rho(f))$ - $\text{Max-CSP}(f)$ problem is trivially solvable by the algorithm that always outputs $\text{val}_\Psi \geq 1$ (since the set $\{\Psi \mid \text{val}_\Psi < \rho(f)\}$ is empty). We say $\text{Max-CSP}(f)$ 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 $(1 - \varepsilon, \rho(f) + \varepsilon)$ - $\text{Max-CSP}(f)$. We say that $\text{Max-CSP}(f)$ is approximation resistant to space $s(n)$ if for every $\varepsilon > 0$, every sketching algorithm for $(1, \rho(f) + \varepsilon)$ - $\text{Max-CSP}(f)$ 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 [16, 13, 14, 8, 7, 15, 6, 4, 5, 19, 2, 3]. In particular, recently Chou, Golovnev, Sudan, and Velusamy [4, 5] gave a dichotomy result for sketching approximability of all finite CSPs. Specifically, they proved the following theorem.

► **Theorem 1** ([5]). *For every k , every predicate $f : \{-1, 1\}^k \rightarrow \{0, 1\}$ and every $0 \leq \beta < \gamma \leq 1$ one of the following holds: (1) (γ, β) -Max-CSP(f) is solvable by an $O(\log(n))$ -space sketching algorithm, or (2) for every $\varepsilon > 0$, $(\gamma - \varepsilon, \beta + \varepsilon)$ -Max-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 [4] 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 [5] 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 [5] 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 [2] 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 2** (Linear threshold function). *A linear threshold function, or LTF, is a Boolean function $f : \{-1, 1\}^k \rightarrow \{0, 1\}$ of the form*

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

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

► **Definition 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 [18] 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 [5, 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 $\text{Max-CSP}(f)$ 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 4.** *For every balanced LTF f on $k \leq 4$ variables, $\text{Max-CSP}(f)$ 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 $k \in \mathbb{N}$, $\text{MON}_k : \{-1, 1\}^k \rightarrow \{0, 1\}$ is given by $\text{MON}_k(x_1, \dots, x_k) = \text{sign}((k-2)x_1 + x_2 + \dots + x_k)$. It may be easily verified that MON_k is a balanced LTF. We have the following theorem.

► **Theorem 5.** *For every $k \geq 5$, $\text{Max-CSP}(\text{MON}_k)$ 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 [18] 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 $\text{WMON}_{k,j} : \{-1, 1\}^k \rightarrow \{0, 1\}$ be the function given by $\text{WMON}_{k,j}(x_1, \dots, x_k) = \text{sign}(j \cdot x_1 + x_2 + \dots + x_k)$. It may be easily verified that when $j+k$ is even, then $\text{WMON}_{k,j}$ is a balanced LTF. We have

► **Theorem 6.** *For all integers $j \geq 2$ and $k \geq 7j^3$ such that $k+j$ is even, $\text{Max-CSP}(\text{WMON}_{k,j})$ is sketching approximable in $O(\log(n))$ space. In particular, for every j , there exist infinitely many k such that $\text{Max-CSP}(\text{WMON}_{k,j})$ is sketching approximable.*

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

² Note that Max-DICUT (shown to be sketching approximable in [6, 4]) is not considered a Boolean CSP in [5] since the Max-DICUT constraints are applied on variables and not on literals.

³ Such functions are also sometimes called presidential type predicates [10].

Comparison to the polynomial time regime

Hast [9] proves that (a generalization of) Theorem 23 holds in the polynomial time regime (thus, implying an analogue of Theorem 6 in the polynomial time regime). Austrin, Benabbas, and Magen [1] prove that MON_k is approximable in polynomial time, which is in sharp contrast to the result of Theorem 5 in the sketching setting. Huang and Potechin [10] show that almost all WMON predicates are approximable in polynomial time. Finally, Potechin [18] 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 $\{1, \dots, n\}$. We write vector variables in boldface, e.g., \mathbf{x} , and we use x_i to denote their i th 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 $p \in [0, 1]$, $\text{Bern}(p)$ 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 $f: \{-1, 1\}^k \rightarrow \{0, 1\}$, 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 7** (Sketching approximation resistance). *For a function $f: \{-1, 1\}^k \rightarrow \{0, 1\}$, 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 8** (Sketching approximability). *For a function $f: \{-1, 1\}^k \rightarrow \{0, 1\}$, 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, [5] proved that every f is either approximable or approximation resistant.⁴

2.2 Characterization of approximability from [5]

In this work, we focus on CSPs that use a single function f applied to literals. Thus, we will use the machinery from [5] instead of the more general (and more notationally-heavy) version in [4]. For a distribution $\mathcal{D} \in \Delta(\{-1, 1\}^k)$, by $\boldsymbol{\mu}(\mathcal{D})$ we denote its marginals, i.e., $\boldsymbol{\mu}(\mathcal{D}) = (\mu_1, \dots, \mu_k)$ 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 10), 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.

35:6 Sketching Approximability of (Weak) Monarchy Predicates

► **Definition 9** ([5, 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

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

and

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

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 10** ([5, 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 [5].⁵

► **Theorem 11** ([5, Corollary 1.2]). For every $k \in \mathbb{N}$ and every function $f : \{-1, 1\}^k \rightarrow \{0, 1\}$, 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 12.** A monarchy predicate on $k \geq 2$ variables $MON_k : \{-1, 1\}^k \rightarrow \{0, 1\}$ is defined as

$$MON_k(x_1, \dots, x_k) = \text{sign} \left((k-2)x_1 + \sum_{i=2}^k x_i \right).$$

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

► **Definition 13** (Weak monarchy functions). A weak monarchy predicate of order j on $k \geq 2$ variables $WMON_{k,j} : \{-1, 1\}^k \rightarrow \{0, 1\}$ is defined as

$$WMON_{k,j}(x_1, \dots, x_k) = \text{sign} \left(j \cdot x_1 + \sum_{i=2}^k x_i \right).$$

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

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

⁵ Strictly speaking the statement in Corollary 1.2 in [5] 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, [17]).

► **Definition 14** (Characteristic functions). *For every $S \subseteq [k]$ such that $|S| \geq 1$, the characteristic function $\chi_S : \{-1, 1\}^k \rightarrow \{-1, 1\}$ 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 15** (Fourier expansions). *The Fourier expansion of a Boolean function $f : \{-1, 1\}^k \rightarrow \{0, 1\}$ is given by*

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

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

► **Definition 16** (Chow parameters). *The Chow parameters of a Boolean function $f : \{-1, 1\}^k \rightarrow \{0, 1\}$ are the degree-0 Fourier coefficient and the k degree-1 Fourier coefficients of f , i.e., $\widehat{f}(\emptyset), \widehat{f}(\{1\}), \dots, \widehat{f}(\{k\})$.*

► **Proposition 17.** *For every Boolean function $f : \{-1, 1\}^k \rightarrow \{0, 1\}$,*

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 $\rho(f)$: $\rho(f) = \mathbb{E}_{x \sim \text{Unif}\{-1, 1\}^k} [f(x)] = \widehat{f}(\emptyset)$. For the second statement, observe that for all $S \subseteq [k]$,

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

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

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

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 $\text{Max-CSP}(\text{MON}_4)$ is approximable.

3.1 Approximability of MON_4

Recall that by Theorem 11, it suffices to show that $K_1^Y(\text{MON}_4) \cap K_{1/2}^N(\text{MON}_4) = \emptyset$. 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

35:8 Sketching Approximability of (Weak) Monarchy Predicates

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 $\mathbf{u} = (u_0, \dots, u_{k-1}), \mathbf{v} = (v_0, \dots, v_{k-1}) \in \mathbb{R}_{\geq 0}^k$ with $\sum_i u_i + v_i = 1$, where for $0 \leq i \leq k-1$,

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

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

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

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

► **Definition 18.** For $\gamma, \beta \in \mathbb{R}$ and $k \geq 2$,

$$\begin{aligned} \tilde{K}_\gamma^Y(\text{MON}_k) &= \{(\mu_1, \mu') \mid (\mu_1, \mu', \dots, \mu') \in K_\gamma^Y(\text{MON}_k)\} \\ \text{and } \tilde{K}_\beta^N(\text{MON}_k) &= \{(\mu_1, \mu') \mid (\mu_1, \mu', \dots, \mu') \in K_\beta^N(\text{MON}_k)\}. \end{aligned}$$

► **Lemma 19.** For $\gamma, \beta \in \mathbb{R}$ and $k \geq 2$,

$$K_\gamma^Y(\text{MON}_k) \cap K_\beta^N(\text{MON}_k) = \emptyset \text{ if and only if } \tilde{K}_\gamma^Y(\text{MON}_k) \cap \tilde{K}_\beta^N(\text{MON}_k) = \emptyset.$$

Proof. First, if $(\mu_1, \mu', \dots, \mu') \in \tilde{K}_\gamma^Y(\text{MON}_k) \cap \tilde{K}_\beta^N(\text{MON}_k)$, then by Definition 18, $(\mu_1, \mu', \dots, \mu') \in K_\gamma^Y(\text{MON}_k) \cap K_\beta^N(\text{MON}_k)$.

For the other direction. Assume that there is a vector $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_k) \in K_\gamma^Y(\text{MON}_k) \cap K_\beta^N(\text{MON}_k)$. Consider two distributions $\mathcal{D}_Y \in S_\gamma^Y$ and $\mathcal{D}_N \in S_\beta^N$ yielding the vector $\boldsymbol{\mu} = \boldsymbol{\mu}(\mathcal{D}_Y) = \boldsymbol{\mu}(\mathcal{D}_N)$. Given that the variables x_2, \dots, x_k are symmetric, any distribution that is yielded by permuting x_2, \dots, x_k in \mathcal{D}_Y (or \mathcal{D}_N) is also in S_γ^Y (or S_β^N). Note that the marginals of these distributions are also permutations of $\boldsymbol{\mu}$. By Lemma 10, K_γ^Y and K_β^N are convex, so they also contain the averages of these vectors: $(\mu_1, \mu', \dots, \mu') \in K_\gamma^Y(\text{MON}_k) \cap K_\beta^N(\text{MON}_k)$ for $\mu' = (\mu_2 + \dots + \mu_k)/(k-1)$. Finally, by Definition 18, $(\mu_1, \mu') \in \tilde{K}_\gamma^Y(\text{MON}_k) \cap \tilde{K}_\beta^N(\text{MON}_k)$. ◀

Next, we characterize the set $\tilde{K}_1^Y(\text{MON}_k)$.

► **Lemma 20.** For every $k \geq 2$, $\tilde{K}_1^Y(\text{MON}_k) = \{(\mu_1, \mu') \in [-1, 1]^2 : \mu_1(k-2) + \mu'(k-1) \geq 1\}$.

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 $(\mu_1, \mu') \in \tilde{K}_1^Y(\text{MON}_k)$.

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

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

► **Lemma 21.** *Max-CSP(MON_4) is sketching approximable in $O(\log(n))$ space.*

Proof. Note that Lemma 20 gives that $\tilde{K}_1^Y(\text{MON}_4) = \{(\mu_1, \mu') \in [-1, 1]^2 : 2\mu_1 + 3\mu' \geq 1\}$. We show that \tilde{K}_1^Y and $\tilde{K}_{1/2}^N$ are disjoint, and then Lemma 19 and Theorem 11 imply that $\text{Max-CSP}(\text{MON}_4)$ 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 \tilde{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:

$$\mathbb{E}_{\mathbf{b} \sim \mathcal{D}_N} \mathbb{E}_{\mathbf{a} \sim \text{Bern}(p)^k} [f(\mathbf{b} \odot \mathbf{a})] \leq \beta, \forall p. \quad (2)$$

Let the function $h_{\mathcal{D}}(p)$ denote the probability of an assignment from \mathcal{D} that has undergone bit flips with respect to $\text{Bern}(p)^k$ 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

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

35:10 Sketching Approximability of (Weak) Monarchy Predicates

This gives that

$$\begin{aligned} h_{\mathcal{D}}(p) &= \sum_{i=0}^{k-1} u_i [p(1 - (1-p)^i p^{k-1-i}) + (1-p)^{k-i} p^i] \\ &\quad + \sum_{i=0}^{k-1} v_i [(1-p)(1 - (1-p)^i p^{k-1-i}) + (1-p)^{k-1-i} p^{i+1}] \end{aligned} \quad (3)$$

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

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

Every distribution (whose marginals are) in $\tilde{K}_{1/2}^N(\text{MON}_4)$ must satisfy the following system of equations and inequalities, where (4)–(7) are equivalent to $h_{\mathcal{D}}(p) = \frac{1}{2}$, and (8)–(10) guarantee that u_i s and v_i s describe a distribution.

$$u_0 + v_0 + v_1 + v_2 = \frac{1}{2} \quad (4)$$

$$-3u_0 + 2u_1 + u_2 - v_1 - 2v_2 + 3v_3 = 0 \quad (5)$$

$$6u_0 - 3u_1 + 3u_3 - 3v_0 + 3v_2 - 6v_3 = 0 \quad (6)$$

$$-4u_0 + 2u_1 - 2u_3 + 2v_0 - 2v_2 + 4v_3 = 0 \quad (7)$$

$$\sum_{i=0}^3 (u_i + v_i) = 1 \quad (8)$$

$$u_i \geq 0, \quad \forall 0 \leq i \leq 3 \quad (9)$$

$$v_i \geq 0, \quad \forall 0 \leq i \leq 3 \quad (10)$$

Summing up (5) multiplied by 3, (7) multiplied by $-13/6$, and (8) 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 (1). By Lemma 20, $\tilde{K}_1^Y(\text{MON}_4) = \{(\mu_1, \mu') \in [-1, 1]^2 : 2\mu_1 + 3\mu' \geq 1\}$, and from the above inequality every vector $(\mu_1, \mu') \in \tilde{K}_{1/2}^N(\text{MON}_4)$ satisfies $2\mu_1 + 3\mu' \leq 2/3$. This implies that $\tilde{K}_1^Y(\text{MON}_4) \cap \tilde{K}_{1/2}^N(\text{MON}_4) = \emptyset$, and finishes the proof. \blacktriangleleft

3.2 Balanced LTFs on 4 variables

In this section, we prove Theorem 4.

► **Theorem 4.** *For every balanced LTF f on $k \leq 4$ variables, $\text{Max-CSP}(f)$ 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 $f(x_1, x_2) = x_1 \text{ OR } x_2$, then $\text{Max-CSP}(f)$ is approximation resistant to space $o(n)$ even in the larger class of streaming algorithms (see, e.g., Corollary 4.2 in [6]).

Proof of Theorem 4. After relabeling and negating some of the variables of f , we can assume that $f(x_1, x_2, x_3, x_4) = \text{sign}(w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4)$, where $w_1 \geq w_2 \geq w_3 \geq w_4 \geq 0$ (if f depends on $i < 4$ variables, then we set $w_{i+1} = \dots = w_4 = 0$). 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 $f = \text{sign}(x_1)$ is a dictator function, so $\text{Max-CSP}(f)$ can be trivially $(1 - \varepsilon)$ -approximated in $O(\log(n)/\varepsilon^2)$ space by an ℓ_1 -sketch algorithm [11, 12].
- If $w_2 + w_3 - w_4 < w_1 < w_2 + w_3 + w_4$, then $f = \text{MON}_4$ 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 21, $\text{Max-CSP}(f)$ is sketching approximable in $O(\log(n))$ space.
- If $w_1 < w_2 + w_3 - w_4$, then $f = \text{MAJ}(x_1, x_2, x_3)$ 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, $\text{Max-CSP}(f)$ is known to be sketching approximable in space $O(\log(n))$ (this follows from the characterization of sketching approximable symmetric functions in [5, 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 25. 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., [17, 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 25, $\text{Max-CSP}(f)$ is sketching approximable in space $O(\log(n))$. ◀

4 Approximation resistance of Monarchy Functions

In this section, we prove Theorem 5: we show that for $k \geq 5$, the MON_k function is approximation resistant. Recall that by Lemma 19 it suffices to show that $\tilde{K}_1^Y(\text{MON}_k) \cap \tilde{K}_{1/2}^N(\text{MON}_k) \neq \emptyset$ 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 $\tilde{K}_1^Y(\text{MON}_k) \cap \tilde{K}_{1/2}^N(\text{MON}_k) \neq \emptyset$.

We defer the proof of Lemma 22 to the full version of the paper.

► **Lemma 22.** *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 $(\mu_1, \mu', \dots, \mu')$ where $\mu_1 = \sum_i (u_i - v_i)$, and $\mu' = \sum_i (\frac{2i}{k-1} - 1)(u_i + v_i)$.

35:12 Sketching Approximability of (Weak) Monarchy Predicates

(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 $p' = \Pr_{\mathbf{x} \sim \mathcal{D}}[x_2 = 1] = \frac{1}{k-1} (\sum_i i u_i + \sum_i i v_i)$ 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 $\boldsymbol{\mu}(\mathcal{D}_Y) = (\mu_1, \mu', \dots, \mu')$.

Now, we are ready to prove Theorem 5 using Lemma 22 and Theorem 11.

► **Theorem 5.** For every $k \geq 5$, $\text{Max-CSP}(\text{MON}_k)$ 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 22. Note that condition (i) guarantees that \mathbf{u}, \mathbf{v} define a distribution \mathcal{D} with marginal $(\mu_1, \mu', \dots, \mu')$.

First, we show that condition (ii) is a sufficient condition for $(\mu_1, \mu') \in \tilde{K}_{1/2}^N$. Recall that $\mathcal{D}_N \in S_{1/2}^N(\text{MON}_k)$ if for every $\delta \in [-1/2, 1/2]$, $\mathbb{E}_{\mathbf{b} \in \mathcal{D}_N} \mathbb{E}_{\mathbf{a} \sim \text{Bern}(1/2+\delta)}[\text{MON}_k(\mathbf{b} \odot \mathbf{a})] = 1/2$. Since $\Pr_{\mathbf{x}}[\text{MON}_k(\mathbf{x}) = 1] = \Pr_{\mathbf{x}}[x_1 = 1] - \Pr_{\mathbf{x}}[\mathbf{x} = 10^{k-1}] + \Pr_{\mathbf{x}}[\mathbf{x} = 01^{k-1}]$, we have that $\mathbb{E}_{\mathbf{b} \in \mathcal{D}_N} \mathbb{E}_{\mathbf{a} \sim \text{Bern}(1/2+\delta)}[\text{MON}_k(\mathbf{b} \odot \mathbf{a})] = \Pr_{\mathbf{b}, \mathbf{a}}[\mathbf{b}_1 \odot \mathbf{a}_1 = 1] - \Pr_{\mathbf{b}, \mathbf{a}}[\mathbf{b} \odot \mathbf{a} = 1(-1)^{k-1}] + \Pr_{\mathbf{b}, \mathbf{a}}[\mathbf{b} \odot \mathbf{a} = (-1)1^{k-1}]$.

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

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

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 $\mathcal{D} \in S_{1/2}^N(\text{MON}_k)$. As condition (i) gives $\boldsymbol{\mu}(\mathcal{D}_N) = (\mu_1, \mu', \dots, \mu')$, we have $(\mu_1, \mu') \in \tilde{K}_{1/2}^N$ 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 20, this implies that $(\mu_1, \mu') \in \tilde{K}_1^Y(\text{MON}_k)$ as desired.

To sum up, Lemma 22 gives us $(\mu_1, \mu') \in \tilde{K}_1^Y \cap \tilde{K}_{1/2}^N$ for every $k \geq 5$ and Lemma 19 implies $(\mu_1, \mu', \dots, \mu') \in K_1^Y \cap K_{1/2}^N$. By Theorem 11, we conclude that MON_k is sketching approximation resistant to space $o(\sqrt{n})$ and, hence, complete the proof of Theorem 5. ◀

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 Section 5.1 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 5.2, 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 23.** *For every Boolean function $f : \{-1, 1\}^k \rightarrow \{0, 1\}$ of the form*

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

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

► **Definition 24.** *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 23.

► **Theorem 25.** *For every Boolean function $f : \{-1, 1\}^k \rightarrow \{0, 1\}$ and every $\varepsilon > 0$, there exists an $O(\log(n))$ space $(\rho(f) + \varepsilon^*(f) - \varepsilon)$ -approximation algorithm for Max-CSP(f).*

First we show how to prove Theorem 23 using Theorem 25.

Proof of Theorem 23. If $f(x)$ is the constant zero function, then it's trivially approximable in $O(\log(n))$ space. Otherwise, when $f(x) = \text{sign}(\sum_{i=1}^k \widehat{f}(\{i\}) \cdot x_i)$, 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 25 implies that there is a $(\rho(f) + \varepsilon^*(f)/2)$ -approximation algorithm for Max-CSP(f), and finishes the proof. ◀

Before we prove Theorem 25, we will describe some useful definitions and lemmas from [5].

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

► **Definition 26 (Bias (vector)).** *For $\lambda = (\lambda_1, \dots, \lambda_k) \in \mathbb{R}^k$, and instance $\Psi = (C_1, \dots, C_m; w_1, \dots, w_m)$ of Max-CSP(f) where $C_i = (\mathbf{j}(i), \mathbf{b}(i))$ and $w_i \geq 0$, we let the λ -bias vector of Ψ , denoted $\text{bias}_\lambda(\Psi)$, be the vector in \mathbb{R}^n given by*

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

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

35:14 Sketching Approximability of (Weak) Monarchy Predicates

► **Lemma 27** ([5, Lemma 4.7]). *For every $\lambda \in \mathbb{R}^k$, we have $B_\lambda(\Psi) = \max_{a \in \{-1,1\}^n} \langle a, \text{bias}_\lambda(\Psi) \rangle$.*

► **Lemma 28** ([5, Lemma 4.4]). *For every vector $\lambda \in \mathbb{R}^k$ and $\varepsilon > 0$, there exists a $O(\log(n))$ space sketching algorithm \mathcal{A} that on input a stream $\sigma_1, \dots, \sigma_\ell$, representing an instance $\Psi = (C_1, \dots, C_m; w_1, \dots, w_m)$, outputs a $(1 \pm \varepsilon)$ -approximation to $B_\lambda(\Psi)$, i.e., for every Ψ , $(1 - \varepsilon)B_\lambda(\Psi) \leq \mathcal{A}(\Psi) \leq (1 + \varepsilon)B_\lambda(\Psi)$, 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 $\text{Max-CSP}(f)$.

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

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

- 1: Let $\lambda = (\widehat{f}(\{1\}), \dots, \widehat{f}(\{k\})) \in \mathbb{R}^k$ and $\varepsilon' = \varepsilon/8$.
- 2: Use the algorithm \mathcal{A} from Lemma 28 to compute \tilde{B} to be a $(1 \pm \varepsilon')$ approximation to $B_\lambda(\Psi)$, i.e., $(1 - \varepsilon')B_\lambda(\Psi) \leq \tilde{B} \leq (1 + \varepsilon')B_\lambda(\Psi)$ 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 28 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 val_Ψ in terms of $B_\lambda(\Psi)$ where $\lambda = (\widehat{f}(\{1\}), \dots, \widehat{f}(\{k\}))$.

► **Lemma 29** (Lower bound on val_Ψ). *Let $f : \{-1, 1\}^k \rightarrow \{0, 1\}$ be a Boolean function, and Ψ be an instance of $\text{Max-CSP}(f)$. Then*

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

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

► **Lemma 30** (Upper bound on val_Ψ). *Let $f : \{-1, 1\}^k \rightarrow \{0, 1\}$ be a Boolean function, $\varepsilon_0(f)$ be as defined in Definition 24, and Ψ be an instance of $\text{Max-CSP}(f)$. Then*

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

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

We defer the proofs of Lemma 29 and Lemma 30 to the full version of the paper. We now show the correctness of Algorithm 1 using these lemmas.

5.1.2 Proof of Theorem 25

Proof of Theorem 25. First, by Lemma 28, with probability at least $2/3$, \tilde{B} is a $(1 \pm \varepsilon')$ approximation to $B_\lambda(\Psi)$, i.e., $(1 - \varepsilon')B_\lambda(\Psi) \leq \tilde{B} \leq (1 + \varepsilon')B_\lambda(\Psi)$. Next, we show that with probability at least $2/3$, (i) $v \leq \text{val}_\Psi$ and (ii) $v \geq (\rho(f) + \varepsilon^*(f) - \varepsilon) \cdot \text{val}_\Psi$.

(i) $v \leq \text{val}_\Psi$

We have

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

where the last inequality follows from Lemma 29.

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

We have

$$v = \rho(f) + \frac{\tilde{B}\tilde{\delta}}{(1 + \varepsilon')^2} \geq \rho(f) + B_\lambda(\Psi)\delta(\Psi) \left(\frac{1 - \varepsilon'}{1 + \varepsilon'} \right)^2 \geq \rho(f) + B_\lambda(\Psi)\delta(\Psi)(1 - \varepsilon), \quad (11)$$

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

$$B_\lambda(\Psi)\delta(\Psi) \geq \varepsilon_0(f) \cdot \min \left\{ \frac{1}{3k}, \frac{2\varepsilon_0(f)}{9\rho(f)k^2} \right\} \geq \varepsilon^*, \quad (12)$$

where the last equality follows from the definition of $\varepsilon^*(f)$ in Definition 24.

Combining Equation (11) and Equation (12), we get

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

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

Now, let us consider the case when $B_\lambda(\Psi) < \varepsilon_0(f)$. It follows from Proposition 17 that $\varepsilon_0(f) \leq \rho(f)k$. Therefore,

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

and so $\delta(\Psi) = \frac{2B_\lambda(\Psi)}{9\rho(f)k^2}$. Combining Equation (11) and Lemma 30, we have

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

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

$$\frac{\rho(f) + \frac{2B_\lambda(\Psi)^2}{9\rho(f)k^2}}{\rho(f) + \frac{B_\lambda(\Psi)}{k}} \geq \frac{\rho(f) + \frac{2\varepsilon_0(f)^2}{9\rho(f)k^2}}{\rho(f) + \frac{\varepsilon_0(f)}{k}}. \quad (13)$$

This immediately implies that

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

Consider the function $g(p) = \frac{\rho(f) + \frac{2p^2}{9\rho(f)}}{\rho(f) + p}$. In order to show Equation (13), 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 25. \blacktriangleleft

5.2 Approximability of weak monarchy functions

In this section, we analyze the streaming approximability of $\text{Max-CSP}(f)$ where f is a weak monarchy function. Note that in order for $\text{WMON}_{k,j}$ 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. We defer the proof of Lemma 31 to the full version of the paper.

► **Lemma 31.** *For all integers $j \geq 2$ and $k \geq 7j^3$ such that $k + j$ is even,*

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

Note that Lemma 31 along with Theorem 23 directly conclude Theorem 6 restated below.

► **Theorem 6.** *For all integers $j \geq 2$ and $k \geq 7j^3$ such that $k + j$ is even, $\text{Max-CSP}(\text{WMON}_{k,j})$ is sketching approximable in $O(\log(n))$ space. In particular, for every j , there exist infinitely many k such that $\text{Max-CSP}(\text{WMON}_{k,j})$ is sketching approximable.*

References

- 1 Per Austrin, Siavosh Benabbas, and Avner Magen. On quadratic threshold csps. In *LATIN 2010*, pages 332–343. Springer, 2010.
- 2 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.
- 3 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.
- 4 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. doi:10.1109/FOCS52979.2021.00117.
- 5 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.
- 6 Chi-Ning Chou, Alexander Golovnev, and Santhoshini Velusamy. Optimal Streaming Approximations for all Boolean Max-2CSPs and Max- k SAT. In *FOCS 2020*, pages 330–341. IEEE, 2020. doi:10.1109/FOCS46700.2020.00039.
- 7 Venkatesan Guruswami and Runzhou Tao. Streaming Hardness of Unique Games. In *APPROX 2019*, pages 5:1–5:12. Schloss Dagstuhl, 2019.
- 8 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.
- 9 Gustav Hast. *Beating a random assignment: Approximating constraint satisfaction problems*. PhD thesis, KTH, 2005.
- 10 Neng Huang and Aaron Potechin. On the approximability of presidential type predicates. In *APPROX 2020*, pages 58:1–58:20. Schloss Dagstuhl, 2020.
- 11 Piotr Indyk. Stable distributions, pseudorandom generators, embeddings and data stream computation. In *FOCS 2000*, pages 189–197. IEEE, 2000.
- 12 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.
- 13 Michael Kapralov, Sanjeev Khanna, and Madhu Sudan. Streaming lower bounds for approximating MAX-CUT. In *SODA 2015*, pages 1263–1282. SIAM, 2015.

- 14 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.
- 15 Michael Kapralov and Dmitry Krachun. An optimal space lower bound for approximating MAX-CUT. In *STOC 2019*, pages 277–288. ACM, 2019. doi:10.1145/3313276.3316364.
- 16 Dmitry Kogan and Robert Krauthgamer. Sketching cuts in graphs and hypergraphs. In *ITCS 2015*, pages 367–376. ACM, 2015.
- 17 Ryan O’Donnell. *Analysis of Boolean functions*. Cambridge University Press, 2014.
- 18 Aaron Potechin. On the approximation resistance of balanced linear threshold functions. In *STOC 2019*, pages 430–441. ACM, 2019.
- 19 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.