New Optima for the Deletion Shadow

Benedict Randall Shaw

Abstract

For a family $\mathcal{F}$ of words of length n drawn from an alphabet $A = [r] = \{1, \dots, r\}$, Danh and Daykin defined the deletion shadow $\Delta\mathcal{F}$ as the family containing all words that can be made by deleting one letter of a word of $\mathcal{F}$. They asked, given the size of such a family, how small its deletion shadow can be, and answered this with a Kruskal-Katona type result when the alphabet has size 2. However, Leck showed that no ordering can give such a result for larger alphabets. The minimal shadow has been known for families of size s^n, where the optimal family has form $[s]^n$. We give the minimal shadow for many intermediate sizes between these levels, showing that families of the form 'all words in $[s]^n$ in which the symbol s appears at most k times' are optimal. This proves a conjecture of Bollobás and Leader. Our proof uses some fractional techniques that may be of independent interest.

Mathematics Subject Classifications: 05D05

1 Introduction

We consider words composed of exactly n symbols from a finite alphabet A, which we will typically take to be $[r] = \{1, \dots, r\}$. We say a family $\mathcal{F} \subset A^n$ has *deletion shadow in direction i*

$$\delta_i\mathcal{F} = \{(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) : (x_1, \dots, x_n) \in \mathcal{F}\},$$

and *deletion shadow*

$$\Delta\mathcal{F} = \bigcup_{i=1}^{n} \delta_i\mathcal{F}.$$

For example, the family $\mathcal{F} = \{112, 123\}$ has deletion shadow $\Delta\mathcal{F} = \{11, 12, 13, 23\}$.

This notion was introduced by Danh and Daykin [4]. Motivated by the Kruskal-Katona theorem, they asked the following question:

Question 1. Given $|\mathcal{F}|$, how small can $|\Delta\mathcal{F}|$ be?

Department of Pure Mathematics and Mathematical Statistics (DPMMS), University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, United Kingdom (`bwr26@cam.ac.uk`).

It is folklore that families of the form $[s]^n$ have minimal shadow given their size (a proof of which we outline later). However, this only gives the minimal deletion shadow for families of size $1^n, 2^n, \ldots, r^n$, which is a very sparse set of values for r fixed and n large.

We define the *simplicial order* for $\{0,1\}^n$ by taking $x < y$ if x contains the symbol 1 fewer times than y, or if x and y each contain the symbol 1 the same number of times and at the first i such that $x_i \neq y_i$, we have $x_i = 1$ and $y_i = 0$. Regarding words in $\{0,1\}^n$ as indicator functions on $[n]$, this definition agrees with the notion of simplicial order on subsets of $[n]$. For an alphabet of two symbols, Danh and Daykin [4, 5] proved an equivalent of the Kruskal-Katona theorem for the deletion shadow.

Theorem 2 ([4])**.** *For $A = \{0,1\}$, across all families $\mathcal{F} \subset A^n$ of a given size, the initial segment of the simplicial order minimises the deletion shadow.*

Danh and Daykin [3] defined a more complex order for general r, the V-order, which they studied hoping a similar result would hold—that every initial segment of the V-order would have minimal deletion shadow for its size. However, Leck [6] proved that for $r \geqslant 3$, such a result is false not only for the V-order, but in fact for any order. This forbids an exact ordering result of this form—in general, no order can have all initial segments minimise their deletion shadow given their size.

However, it leaves open the possibility of an ordering *many* of whose initial segments have minimal shadow given their size. Our main result will be the following, proving a conjecture of Bollobás and Leader (see [8]), which gives the optimal family for a much denser set of values than previously known:

Theorem 3. *For any $s \leqslant r$, $k \leqslant n$, the family $\mathcal{F}$ comprising all words of $[s]^n$ that contain the symbol s at most k times has minimal deletion shadow of all families in $[r]^n$ of the same size.*

We will show this by considering a version of the problem in the setting of fractional families, a more general notion in which each word has a weight between 0 and 1, where weights 0 and 1 correspond respectively to non-membership and membership in discrete families. In this relaxed setting, we will be able to find the family with minimal deletion shadow of each weight. This guarantees that our argument does not come up against Leck's result, in that while we will be able to answer the question exactly in our generalisation, these fractional families will only correspond to discrete families at sizes matching the families in Theorem 3.

We will have to be careful in choosing how to generalise the deletion shadow to fractional families in a way that interacts nicely with the minima we prove. Indeed, there are several natural definitions of a deletion shadow on fractional families. Perhaps counter-intuitively, we choose a definition which in general is minimised not by the families we care about, but rather by the *uniform fractional families*, whose indicator functions are constant.

However, it is known that in the discrete setting, we may assume that our families are down-compressed. We therefore restrict to fractional families obeying a certain weaker

compression condition. Again there are various ways to apply such a condition, and so the choice of this condition is key to the proof. A relatively standard argument then determines the minimal shadow of a fractional family of a given weight, subject to our restrictions. As an immediate consequence, we recover Theorem 3 in the discrete setting.

For other results on the deletion shadow, see Danh and Daykin [3, 4, 5]. Other notions of the deletion shadow have also been studied. For example, Räty [8] proved an exact ordering result for the related shadow obtained by deleting only the character 1.

We finish this section by outlining the folklore proof that $[s]^n$ has minimal shadow given its size. We can view $\delta_i \mathcal{F}$ as a projection of $\mathcal{F}$ onto the hyperplane defined by $x_i = 0$. Then the discrete Loomis-Whitney inequality [7], applied to $\mathcal{F}$, states that

$$|\mathcal{F}| \geqslant \prod_{i=1}^{n} |\delta_i \mathcal{F}|^{\frac{1}{n-1}}.$$

As $\delta_i \mathcal{F} \subset \Delta \mathcal{F}$, this gives the bound $|\Delta \mathcal{F}| \geqslant |\mathcal{F}|^{\frac{n-1}{n}}$. This is attained by $[s]^n$. In particular, this implies that for $s \leqslant r$, the family $\mathcal{F} = [s]^n$ has minimal deletion shadow of all families of that size.

In Section 2, we give a more detailed overview of the proof. In Section 3, we present a version of Danh and Daykin's argument that we may assume discrete families with minimal shadow are down-compressed. In Section 4, we set up the the problem in the context of fractional families, and prove a more general result there which implies Theorem 3. Finally, in Section 5, we conjecture some other discrete families of minimal deletion shadow given their size.

2 Overview

We first note that by an argument of Danh and Daykin [4], we may assume that the family $\mathcal{F}$ is down-compressed—that is to say, if $x_1 \ldots x_n$ is a word in $\mathcal{F}$, then any word $x_1' \ldots x_n'$ with $x_i' \leqslant x_i$ for each i is also in $\mathcal{F}$. We present a version of their argument for completeness, although it will not be necessary to follow the rest of the proof.

Leck [6] proved that there is no exact ordering of words whose initial segments each have minimal deletion shadow for their size. Thus any proof that would give us a result of this form must fail. We therefore work instead in fractional families, introduced by Bollobás and Leader in [1, 2]. In this setting, words cannot just be members and non-members of the family, but may be members of the family to any extent in $[0, 1]$. These allow us to move uniformly between the discrete families that we show to be optimal in Theorem 3 without saying anything at all about the optimal discrete families between those sizes.

We extend the notion of a deletion shadow accordingly. There are several possible ways to define the deletion shadow on fractional families, and we must choose a definition that interacts well with the problem. In fact, the version we use is minimised by the uniform fractional family which contains every set in A^n to the same extent. We would

like to show that certain discrete families are optimal even in the wider setting of fractional families. Unfortunately, the uniform fractional family is not a discrete family.

However, we know we can assume an optimal discrete family is down-compressed, and the uniform fractional family is not down-compressed. Thus we instead minimise the deletion shadow across fractional families satisfying a certain weaker notion of down-compression, which we call *s-compression*. Subject to this condition, we will be able to compute the exact minimal deletion shadow, which will imply Theorem 3.

The proof is an induction on n, essentially by a codimension-1 compression in dimension n. We divide the fractional family f into its layers

$$f_y(x_1, \ldots, x_{n-1}) = f(x_1, \ldots, x_{n-1}, y).$$

We first treat the case where f is zero outside $[s]^n$, where s is minimal such that f has weight at most s^n. We define a notion of a fractional Hamming ball, and consider a fractional family g with the same weight as f so that each layer g_x is a fractional Hamming ball, g_s has the same weight as f_s, and $g_1, \ldots, g_{s-1}$ each have the same weight. Our compression condition will tell us that g_s has weight at most that of g_1. We will then show that the deletion shadow of f has weight at least as large as that of the fractional Hamming ball h with the same weight as f.

We then distinguish three regimes depending on the weight of g_s. If it has weight at most as large as that of h_s, then we find that Δg has the same weight as g's heaviest layer. We observe that by deletion in the last coordinate, f has shadow of weight at least as large as f's largest layer, and thus has shadow of weight at least as large as that of g. We then note that g's shadow has weight at least as large as that of h.

If g_s is heavier than h_s, but g still meets our compression condition, we notice that the shadow of g is exactly the same as that obtained from deletion in non-final coordinates. We note that f's shadow is at least as large as that obtained from f by deletion in non-final coordinates. Our induction on n show that this bound is minimised by g. Since g attains this bound, it is optimal given the weight of g_s. We then calculate that g's shadow has weight at least as large as that of h.

Finally, we consider the case where g_s is heavy enough that g no longer meets our compression condition. Here the bound of the previous case still applies to this one, and moreover its value does not depend on the exact distribution of weight between layers. This allows us to move our weight between until we are in the previous case, where the compression condition is satisfied, and simply apply the result for that case directly.

We finish by checking the case where f is not a subfamily of $[s]^n$. Here we write down an expression for the weight of Δf in terms of the weights of f_x with $x > s$, and find it is a simple optimisation problem.

3 Down-compression

We say $\mathcal{F} \subset A^n$ is *down-compressed* if for any $i \in [n]$, $(x_1, \ldots, x_n) \in \mathcal{F}$ and $y \in A$ such that $y < x_i$, we have $(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) \in \mathcal{F}$. We will need the following lemma, due to Danh and Daykin [4]:

Lemma 4. *For any $\mathcal{F} \subset A^n$, there exists a down-compressed $\mathcal{F}'$ with $|\mathcal{F}'| = |\mathcal{F}|$ such that $|\Delta\mathcal{F}'| \leqslant |\Delta\mathcal{F}|$.*

For completeness, we present a version of their proof for the interested reader. The details will not be needed for the sections following. We make some rather heavy syntactic conventions, for which we apologise—these are not chosen to confuse the reader, but to allow the proof to be written down in a precise way.

Proof of Lemma 4. For $a, b \in A$ with $a < b$, $i \in [n]$, we define the compression $C_i^{a,b}\mathcal{F}$ from $\mathcal{F}$ by replacing any $(x_1, \ldots, x_{i-1}, b, x_{i+1}, \ldots, x_n)$ with $(x_1, \ldots, x_{i-1}, a, x_{i+1}, \ldots, x_n)$ unless it is already in the family $\mathcal{F}$. Then clearly $\left|C_i^{a,b}\mathcal{F}\right| = |\mathcal{F}|$.

Write $C^{a,b}\mathcal{F} = C_1^{a,b} \cdots C_n^{a,b}\mathcal{F}$. Then it will suffice to show that $\left|\Delta C^{a,b}\mathcal{F}\right| \leqslant |\Delta\mathcal{F}|$, as repeatedly applying different compressions of the form $C^{a,b}$ until none of them change $\mathcal{F}$ further eventually gives a down-compressed family. (Note that repeatedly applying these compressions must eventually result in a family that is fixed by all of them, as any such compression that changes $\mathcal{F}$ must decrease the natural number $\sum_{i \in [n], w \in \mathcal{F}} w_i$. Such a family is down-compressed, although it may depend on the order in which compressions are chosen.)

We now define, for each $x \in A$, a quantifier Q_x which will allow us to write a concise condition for a given word to be a member of $C^{a,b}\mathcal{F}$. For a sentence $\phi = \phi(x_1, \ldots, x_n, y_1, \ldots, y_n)$, we will write $Q_a y.\phi(y, x_1, \ldots)$ if for some choice of $y \in \{a, b\}$ the sentence $\phi(y, x_1, \ldots)$ holds—that is to say, $Q_a y$ corresponds to the quantifier $\exists y \in \{a, b\}$. We will write $Q_b y.\phi(y, x_1, \ldots)$ if $\phi(y, x_1, \ldots)$ holds for both choices of $y \in \{a, b\}$—that is to say, $Q_b y$ corresponds to the quantifier $\forall y \in \{a, b\}$. Finally, when $x \in A \setminus \{a, b\}$, we write $Q_x y.\phi(y, x_1, \ldots)$ if $\phi(y, x_1, \ldots)$ holds when $y = x$. Note that, all of these quantifiers imply $\exists y.\phi(y, x_1, \ldots)$.

For example, suppose that $a = 3$ and $b = 4$. Then the statement

$$Q_1 w.Q_3 x.Q_2 y.Q_4 z.\phi(w, x, y, z)$$

is equivalent to the statement

$$\exists x \in \{3, 4\}.\forall z \in \{3, 4\}.\phi(1, x, 2, z).$$

Now notice that $(x_1, \ldots, x_n) \in C^{a,b}\mathcal{F}$ if and only if

$$Q_{x_1} y_1.(y_1, x_2, \ldots, x_n) \in C_2^{a,b} \cdots C_n^{a,b}\mathcal{F}.$$

Inductively, we see that $(x_1, \ldots, x_n) \in C^{a,b}\mathcal{F}$ whenever $Q_{x_1} y_1 \ldots Q_{x_n} y_n.(y_1, \ldots, y_n) \in \mathcal{F}$. However, for any formula ϕ, whenever $Q_x y.\phi(y, x_1, \ldots)$ we also have $\exists y.\phi(y, x_1, \ldots)$.

Now any word in $\Delta C^{a,b}\mathcal{F}$ has the form $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ for some $(x_1, \ldots, x_n) \in C^{a,b}\mathcal{F}$. Then $Q_{x_1} y_1 \ldots Q_{x_n} y_n.(y_1, \ldots, y_n) \in \mathcal{F}$. Hence

$$Q_{x_1} y_1 \ldots Q_{x_{i-1}} y_{i-1}.\exists y_i.Q_{x_{i+1}} y_{i+1} \ldots Q_{x_n} y_n.(y_1, \ldots, y_n) \in \mathcal{F}.$$

And so

$$Q_{x_1}y_1 \ldots Q_{x_{i-1}}y_{i-1}.Q_{x_{i+1}}y_{i+1} \ldots Q_{x_n}y_n.(y_1,\ldots,y_{i-1},y_{i+1},\ldots,y_n) \in \Delta\mathcal{F}.$$

But now we have $(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n) \in C^{a,b}\Delta\mathcal{F}$, as desired.

Thus $\Delta C^{a,b}\mathcal{F} \subset C^{a,b}\Delta\mathcal{F}$, so we must have $\left|\Delta C^{a,b}\mathcal{F}\right| \leqslant |\Delta\mathcal{F}|$, giving the desired result. $\qquad\square$

4 The fractional setting

We can identify a family $\mathcal{F} \subset A^n$ with its *indicator function*, the function from A^n to the set $\{0,1\}$ defined by $f(x) = 1$ for each $x \in \mathcal{F}$, $f(x) = 0$ otherwise. We will work in fractional families, in which an element may be 'partially present'. We define a *fractional family* as any function from A^n to the unit interval $[0,1]$. This will play the role that indicator functions played in discrete (i.e. non-fractional) families. We say a family f has *weight*

$$|f| = \sum_{w \in A^n} f(w).$$

Note that when f is the indicator function of a discrete family $\mathcal{F}$, we have $|f| = |\mathcal{F}|$. We will say a family f is a *subfamily* of g if for every $w \in A^n$, we have $f(w) \leqslant g(w)$. Given a fractional family f, we define the family Δf by

$$(\Delta f)(x_1,\ldots,x_{n-1}) = \max_{y \in A, i \in [n]} f(x_1,\ldots,x_{i-1},y,x_i,\ldots,x_{n-1}).$$

Again, if f is the indicator function of a discrete $\mathcal{F} \subset A^n$, then Δf is the indicator function of $\Delta\mathcal{F}$. For example, where we took a maximum, we could instead have taken a sum capped at 1, obtaining a related shadow $\Delta_\Sigma f$. Another possible definition would be

$$(\Delta_{\Sigma\max}f)(x_1,\ldots,x_{n-1}) = \min\left\{\max_{i \in [n]} \sum_{y \in A} f(x_1,\ldots,x_{i-1},y,x_i,\ldots,x_{n-1}), 1\right\}.$$

Unfortunately, our chosen definition of the deletion shadow is minimised, for a given weight of f, by the *uniform fractional family* which has $f(w)$ constant, which gives no information about discrete families. This is not true of Δ_Σ and $\Delta_{\Sigma\max}$, making them attractive candidates to work with. However, we will see that Δ interacts better with the families we show are optimal.

In discrete families, the *Hamming ball* $\mathcal{B}^{(n,s)}(k)$ is the family of words in $[s]^n$ containing the symbol s at most k times. We extend this notion to *fractional balls*, whose indicator function is defined to be 1 on words containing the symbol s at most k times, some fixed $\alpha \in [0,1]$ on words containing the symbol s at most $k+1$ times, and 0 otherwise. We will call this $b_{k,\alpha}^{(n,s)}$, or sometimes just $b_{k,\alpha}$ where the meaning is clear. Note that $b_{k-1,1}$ and $b_{k,0}$ are the same family, which correspond to the discrete family $\mathcal{B}^{(n,s)}(k)$.

These balls interact extremely well with the definition of the deletion shadow we have chosen, in that $\Delta b_{k,\alpha}^{(n,s)} = b_{k,\alpha}^{(n-1,s)}$. This is not true of our other possible deletion shadows, where we find $\Delta_\Sigma b_{k,\alpha} = b_{k,\min\{n(s-1)\alpha,1\}}$, and $\Delta_{\Sigma\max} b_{k,\alpha} = b_{k,\min\{(s-1)\alpha,1\}}$. (Note that these shadows also decrease the word length n by 1, even though we have not marked this explicitly.) In particular, when $\alpha = \frac{1}{s-1}$, any family f which has $b_{k,\alpha}(w) \leqslant f(w) \leqslant b_{k,1}(w)$ for each word w would have the same shadow $b_{k,1}$ in either Δ_Σ or $\Delta_{\Sigma\max}$, so if the fractional ball is optimal in these shadows, it is far from unique.

In order to work in Δ, we will need to restrict the families we optimise over to avoid the uniform fractional family. We know from Lemma 4 that in discrete families, it suffices to consider down-compressed families. A natural extension of down-compression to fractional families is *solid down-compression*: we say a family f on A^n is *solidly down-compressed* if for any $x_1, \ldots, x_n, y \in A$ and $i \in [n]$ such that $y < x_i$, we have

$$f(x_1, \ldots, x_n) > 0 \quad \implies \quad f(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) = 1.$$

Indeed, restricted to solidly down-compressed families, Δ and $\Delta_{\Sigma\max}$ agree.

However, the fractional ball is not generally down-compressed. Since we aim to show the fractional ball is optimal, we will need to work in a restricted version of compression which the fractional ball satisfies. For $\ell < r$, we say a family f on A^n is *ℓ-compressed* if, for any $x_1, \ldots, x_n, y \in A$ and with $i \in [n]$ such that $y < x_i$ and $\ell \leqslant x_i$, we have

$$f(x_1, \ldots, x_n) > 0 \quad \implies \quad f(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) = 1.$$

We prove the following:

Theorem 5. *Let f be an s-compressed fractional family on A^n with $|f| \leqslant s^n$, and let h be the fractional Hamming ball on A^n with the same weight as f. Then $|\Delta f| \geqslant |\Delta h|$.*

Notice that any down-compressed discrete family is ℓ-compressed for every ℓ. Thus by Lemma 4, this result implies Theorem 3. The compression condition is essential in this theorem, as otherwise the minimal shadow is obtained by a totally uniform fractional family, i.e. one where f is constant.

Proof of Theorem 3 from Theorem 5. Suppose that the discrete family $\mathcal{F}$ has minimal deletion shadow across all families with the same size as $\mathcal{B}^{(n,s)}(k)$. By Lemma 4, we may assume that $\mathcal{F}$ is down-compressed. Consider the indicator function f of $\mathcal{F}$, and notice that as $\mathcal{F}$ is down-compressed, certainly f is s-compressed, as down-compression is a stronger condition. We also have $|f| \leqslant s^n$.

But now Theorem 5 tells us that $|\Delta f|$ is at least as large as $|\Delta h|$, where h is the fractional Hamming ball on A^n with the same weight as f. But then h is the indicator function of $\mathcal{B}^{(n,s)}(k)$. Hence

$$|\Delta\mathcal{F}| = |\Delta f| \geqslant |\Delta h| = \left|\Delta\mathcal{B}^{(n,s)}(k)\right|.$$

Hence $\mathcal{B}^{(n,s)}(k)$ has minimal deletion shadow of all discrete families in $[r]^n$ of the same size, as desired. $\qquad\square$

Proof of Theorem 5. For $x \in A$, define the family $f_x : A^{n-1} \to [0,1]$ by

$$f_x(x_1, \ldots, x_{n-1}) = f(x_1, \ldots, x_{n-1}, x).$$

Note that by considering deletion in the last coordinate, for any x we find that f_x is a subfamily of Δf. In particular, if some $|f_x| > s^{n-1}$, then

$$|\Delta f| \geqslant |f_x| > s^{n-1} \geqslant |\Delta h|,$$

so we may assume that $|f_x| \leqslant s^{n-1}$ for each x.

We first treat the case where $f_x = 0$ for all $x > s$. We will then be able to use the result in this case to solve the remaining case as a simple optimisation problem.

We proceed by induction on n. Notice that for $n = 1$, we have $|\Delta f| = \max\{f(1) \ldots, f(r)\}$, and the condition of s-compression says that if some $x \geqslant s$ has $f(x) > 0$, then $f(y) = 1$ for all $y < x$. In particular, if $|f| > s$, we must have some such $f(x) > 0$, and so $f(0) = 1$, giving $|\Delta f| = 1$, matching that of h. Otherwise, we can only have $|\Delta f| < 1$ if $f(x) = 0$ for each $x \geqslant s$, i.e. $f(0) + \cdots + f(s-1) = |f|$. But here $|\Delta f|$ is minimised by setting $f(0), \ldots, f(s-1)$ equal, which is exactly the ball h. Thus h is optimal, as desired.

We now define $g : A^n \to [0,1]$ by defining each layer

$$g_x(x_1, \ldots, x_{n-1}) = g(x_1, \ldots, x_{n-1}, x).$$

Set $g_1 = \cdots = g_{s-1}$ to be the fractional Hamming ball $b_{k,\alpha} = b_{k,\alpha}^{(n-1,s)}$ on A^{n-1} such that $(s-1)|g_1| = |f_1| + \cdots + |f_{s-1}|$, and g_s to be the fractional Hamming ball $b_{k',\alpha'} = b_{k',\alpha'}^{(n-1,s)}$ on A^{n-1} with the same weight as f_s. We take $g_x = 0$ for every $x > s$. Note that as f is s-compressed, we necessarily have $k + \alpha \leqslant k' + \alpha'$.

If $k' < k$, then certainly g is s-compressed. Note that if g was exactly a ball on A^n, we would have $k' = k - 1, \alpha' = \alpha$. We divide into three cases based on whether g_s has larger weight than $b_{k-1,\alpha}$, and whether g is s-compressed.

Case 1: $k' + \alpha' \leqslant k - 1 + \alpha$ Note that by considering deletion in the last coordinate,

$$|\Delta f| \geqslant \max_{x \in [s-1]} |f_x| \geqslant |b_{k,\alpha}|.$$

Now the family g is a subfamily of $b_{k,\alpha}$, so Δg is a subfamily of $b_{k,\alpha}$. Note that the two uses of $b_{k,\alpha}$ refer to different families according to the word lengths of $g, \Delta g$ respectively. However, $g_1 = b_{k,\alpha}$ is a subfamily of Δg, so indeed Δg is exactly $b_{k,\alpha}$. Hence we have $|\Delta f| \geqslant |\Delta g|$. But now decreasing $k + \alpha$ decreases $|b_{k,\alpha}|$, so the smallest deletion shadow possible in this case is obtained when $k' + \alpha' = k - 1 + \alpha$, in which case $g = h$, as desired.

Case 2: $k - 1 + \alpha \leqslant k' + \alpha' \leqslant k$ Now we have $\Delta g \neq b_{k,\alpha}$, so we show g is optimal with a different lower bound on $|\Delta f|$. By considering deletion in coordinates other than the last one, we note that

$$(\Delta f_x)(w) \leqslant (\Delta f)_x(w)$$

for each $x \in A$ and $w \in A^{n-2}$. Hence

$$|\Delta f| \geqslant \sum_{x=1}^{s} |\Delta f_x|.$$

But now define the family f_{avg} on A^{n-1} by $(s-1)f_{\text{avg}}(w) = f_1(w) + \cdots + f_{s-1}(w)$. Then

$$(\Delta f_{\text{avg}})(x_1, \ldots, x_{n-1}) = \max_{x' \in A, i \in [n]} f_{\text{avg}}(x_1, \ldots, x_{i-1}, x', x_i, \ldots, x_{n-1})$$

$$\leqslant \frac{1}{s-1} \sum_{r=1}^{s-1} \max_{x' \in A, i \in [n]} f_r(x_1, \ldots, x_{i-1}, x', x_i, \ldots, x_{n-1})$$

$$= \frac{1}{s-1} \sum_{r=1}^{s-1} (\Delta f_r)(x_1, \ldots, x_{n-1}).$$

In particular, we find that

$$|\Delta f| \geqslant (s-1)|\Delta f_{\text{avg}}| + |\Delta f_s|.$$

But by induction, these terms are respectively minimised, given the weights of f_{avg} and f_s, when these are both balls, in which case $f_{\text{avg}} = g_1$ and $f_s = g_s$. Now as $k-1+\alpha \leqslant k'+\alpha' \leqslant k$, certainly $b_{k,\alpha}$ is a subfamily of g. The only further contribution to the deletion shadow given by g_s is obtained by deleting non-final coordinates. Hence the deletion shadow of g is given by

$$(\Delta g)(x_1, \ldots, x_{n-1}) = \begin{cases} b_{k,\alpha}(x_1, \ldots, x_{n-2}) & \text{if } x_{n-2} < s, \\ b_{k',\alpha'}(x_1, \ldots, x_{n-2}) & \text{if } x_{n-2} = s. \end{cases}$$

In particular, $|\Delta g|$ attains the bound $(s-1)|\Delta g_1| + |\Delta g_s|$, so is optimal given the weights of g_1 and g_s. We now ask how best to choose these weights to minimise the weight of Δg. Now we have

$$|g| = (s-1)\left(\left|\mathcal{B}^{(n-1,s)}(k)\right| + \alpha \binom{n-1}{k+1}(s-1)^{n-k-2}\right)$$
$$+ \left(\left|\mathcal{B}^{(n-1,s)}(k-1)\right| + \alpha' \binom{n-1}{k}(s-1)^{n-k-1}\right),$$

but

$$|\Delta g| = (s-1)\left(\left|\mathcal{B}^{(n-2,s)}(k)\right| + \alpha \binom{n-2}{k+1}(s-1)^{n-k-3}\right)$$
$$+ \left(\left|\mathcal{B}^{(n-2,s)}(k-1)\right| + \alpha' \binom{n-2}{k}(s-1)^{n-k-2}\right).$$

Then there are positive constants c_1, c_2, c_3, c_4 such that

$$|g| = c_1 + c_2\left(\alpha \frac{n-k-1}{k+1} + \alpha'\right),$$

$$|\Delta g| = c_3 + c_4 \left(\alpha \frac{n-k-2}{k+1} + \alpha' \right).$$

In particular, given the weight of g, the weight of g's shadow is minimised when α is maximised. Within this case, this is attained when $k' + \alpha' = k - 1 + \alpha$, in which case g is exactly h, the Hamming ball on A^n.

Case 3: $k = k'$ As above, we have

$$|\Delta f| \geqslant (s-1) |\Delta f_{\mathrm{avg}}| + |\Delta f_s| \geqslant (s-1) |\Delta g_1| + |\Delta g_s|.$$

However, as above there are also positive constants c_1', c_2', c_3', c_4' such that

$$|f| = c_1' + c_2' \left(\alpha(s-1) + \alpha' \right),$$

$$(s-1) |\Delta g_1| + |\Delta g_s| = c_3' + c_4' \left(\alpha(s-1) + \alpha' \right).$$

But now this bound depends only on the weight of f, and is constant across all of Case 3. It will therefore be sufficient to show that this exceeds the weight of h for a single value of α. Given the weight of f, we choose α, α' such that $\alpha' = 0$ or $\alpha = 1$. Then by the analysis of Case 2, g attains the desired bound. However, we have already shown that this has shadow has weight at least as large as that given by h, as desired.

We now consider the case where we may have $f_x \neq 0$ for $x > s$. Let f' be defined by $f_x' = f_x$ for $x \leqslant s$, and $f_x' = 0$ for any $x > s$. Notice that as f is s-compressed, f_x is a subfamily of f_1 for any $x \geqslant s$, since whenever $f_x(x_1, \ldots, x_{n-1}) > 0$ we in fact have $f_1(x_1, \ldots, x_{n-1}) = 1$. Thus the contribution of $f_{s+1}, f_{s+2}, \ldots$ to the shadow of f is only from deletion of non-final coordinates. Now we have

$$\Delta f(x_1, \ldots, x_n) = \begin{cases} (\Delta f')(x_1, \ldots, x_n) & \text{if } x_n < m, \\ f_{x_n}(x_1, \ldots, x_{n-1}) & \text{if } x_n \geqslant m. \end{cases}$$

In particular,

$$|\Delta f| = |\Delta f'| + \sum_{x=s+1}^{r} |f_x|.$$

But these terms are respectively minimised by balls. Define g by letting g' be the fractional Hamming ball $b_{k,\alpha}^{(n,s)}$ on A^n with the same weight as f', and for each $x > s$, letting g_s be the ball $b_{k_x,\alpha_x}^{(n-1,s)}$ on A^{n-1} with the same weight as f_x. Without loss of generality, we take $\alpha < 1$ whenever $|f'| < s^n$, and $\alpha_x > 0$ whenever $|f_x| \neq 0$. Note that certainly $|\Delta f| \geqslant |\Delta g|$.

Now as f is s-compressed, we have $k_x < k$ for each $x > s$. But now we have

$$|g| = \left| \mathcal{B}^{(n,s)}(k) \right| + \alpha \binom{n}{k+1} (s-1)^{n-k-1}$$

$$+ \sum_{x=s+1}^{r} \left(\left| \mathcal{B}^{(n-1,s)}(k_x) \right| + \alpha_x \binom{n-1}{k_x+1} (s-1)^{n-k_x-2} \right),$$

$$|\Delta g| = \left|\mathcal{B}^{(n-1,s)}(k)\right| + \alpha \binom{n-1}{k+1}(s-1)^{n-k-2}$$

$$+ \sum_{x=s+1}^{r}\left(\left|\mathcal{B}^{(n-2,s)}(k_x)\right| + \alpha_x \binom{n-2}{k_x+1}(s-1)^{n-k_x-3}\right).$$

We now treat this as an optimisation problem, given the weight of g. Write

$$\beta = \alpha\binom{n}{k+1}(s-1)^{n-k-1}, \ \beta_x = \alpha_x\binom{n-2}{k_x+1}(s-1)^{n-k_x-3},$$

so that the constraint on the weight of g becomes a constraint on $\beta + \sum_{x=s+1}^{r}\beta_x$. There is now some constant C such that

$$|\Delta g| = C + \frac{1}{s-1}\left(\frac{n-k-1}{n}\beta + \sum_{x=s+1}^{r}\frac{n-k_x-2}{n-1}\beta_x\right).$$

But $k_x < k$ implies that $\frac{n-k-1}{n} < \frac{n-k_x-2}{n-1}$. So decreasing some β_x to increase β decreases $|\Delta g|$. Equivalently, decreasing some α_x to increase α decreases $|\Delta g|$. Thus the optimum must have either $\alpha = 1$ or all $\alpha_x = 0$. But by our choice of α, this implies that $f_x = 0$ for all $x > s$. $\square$

5 Further questions

Previously, the minimal deletion shadow was only known for families of a given size of the form $[s]^n$, for some $s \leqslant r$. For each $s < r$, Theorem 3 gives $n-1$ new sizes between s^n and $(s+1)^n$ for which the minimal deletion shadow is known. A natural further question would be whether we can construct families of minimal deletion shadow for an even denser set of sizes.

Question 6. Do families of the form

$\big\{w \in [s]^n :$ w contains the symbol s at most k times, or

$\qquad\qquad w$ contains the symbol s exactly $k+1$ times and the symbol $(s-1)$ at most

$\qquad\qquad \ell$ times$\big\}$

have minimal deletion shadow given their size?

Note that this does not conflict with the result of Leck. His proof rests on a counterexample at a single value, and so is very specific in what it forbids. In particular, it does not preclude dense sets of sizes at which initial segments of some ordering give minimal deletion shadow.

More generally, let $c_x(w)$ denote the number of occurrences of the character x in the word w, and let the *signature* of a word be $\sigma(w) = (c_r(w), \ldots, c_1(w))$. Let $<_L$ denote the lexicographic order, where for any two signatures p, q, we say $p <_L q$ if $p_i < q_i$ at the first index i where $p_i \neq q_i$. We pose the following conjecture:

Conjecture 7. For any signature s, the family $\{w \in A^n : \sigma(w) \leqslant_L s\}$ has minimal deletion shadow given its size.

The statement that families of the form $[s]^n$ have minimal deletion shadow corresponds to the special case of this conjecture where s has form $(0, \ldots, 0, n, 0, \ldots, 0)$. Our Theorem 3 is then the special case where s has form $(0, \ldots, 0, k, n-k, 0, \ldots, 0)$.

Acknowledgements

References

[1] B. Bollobás and I. Leader. Isoperimetric inequalities and fractional set systems. *Journal of Combinatorial Theory, Series A*, 56(1):63–74, 1991.

[2] B. Bollobás and I. Leader. Maximal sets of given diameter in the grid and the torus. *Discrete Mathematics*, 122(1):15–35, 1993.

[3] T Danh and D. Daykin. The structure of V-order for integer vectors. *Congressus Numerantium*, 113, 1996.

[4] T. Danh and D. Daykin. Ordering Integer Vectors for Coordinate Deletions. *Journal of the London Mathematical Society*, 55(3):417–426, 1997.

[5] T. Danh and D. Daykin. Sets of 0, 1 vectors with minimal sets of subvectors. *Rostocker Mathematisches Kolloquium*, 50, 1997.

[6] U. Leck. Nonexistence of a Kruskal-Katona Type Theorem for Subword Orders. *Combinatorica*, 24(2):305–312, 2004.

[7] L. H. Loomis and H. Whitney. An inequality related to the isoperimetric inequality. *Bulletin of the American Mathematical Society*, 55(10):961–962, 1949.

[8] E. Räty. Coordinate Deletion of Zeroes. *The Electronic Journal of Combinatorics*, #P3.50, 2019.