Tight exponent for finite field Nikodym

Hung-Hsun Hans Yu

2026-11-04

Table of contents

Some background . . . 1
ChatGPT in the scene . . . 2
The main result . . . 2
The bound . . . 3
The top level . . . 4
The next level . . . 5
Higher dimensions . . . 6
The construction . . . 6
The prototype in $\mathbb{R}^2$ . . . 6
Modification using algebraic integers . . . 7
From weak Nikodym to Nikodym: Cartesian product . . . 9
A two-scaled construction in $\mathbb{Z}^3$ . . . 9
A brief sketch of the construction for $d = 3$ and beyond . . . 10

Some background

A *(weak) Nikodym set* in $\mathbb{R}^2$ is the subset N of $[0,1]^2 \subseteq \mathbb{R}^2$ with the following property: for any $x \in [0,1]^2 \backslash N$, there is a straight line ℓ_x through x so that $(\ell_x \cap [0,1]^2) \backslash \{x\} \subseteq N$. In words, there is a *punctured* line inside N through any point not in N. Intuitively, N would have to be somewhat large. However, Nikodym in 1927 showed that contrary to the intuition, there is in fact a Nikodym set of Lebesgue measure zero.

A related notion is a *Kakeya set*, which is a set containing a unit segment in every direction. Again, contrary to the naïve guess, there is actually a Kakeya set in $\mathbb{R}^2$ (and thus in all higher dimensions) of Lebesgue measure zero by Besicovitch. However, it turns out that any Kakeya set in $\mathbb{R}^2$ has to be 2-dimensional in some rigorous sense, and whether every Kakeya set in $\mathbb{R}^d$ must be d-dimensional is the famous *Kakeya conjecture* that is central to harmonic analysis.

There have been many recent breakthroughs on the Kakeya conjecture, but this is not the focus of this post. I want to shift the focus to the finite field analogue proposed by Wolff, who suggested the problem of determining the minimum size of Kakeya set in $\mathbb{F}_q^d$. Here, a Kakeya set in $\mathbb{F}_q^d$ is a set containing a line in every direction. Wolff noticed that previous lower bounds on the dimensions of Kakeya sets in $\mathbb{R}^d$ worked over $\mathbb{F}_q$ as well. The hope was to understand the finite field Kakeya problem and that some of the proof techniques would transfer to the real setting.

The finite field Kakeya problem turned out to be challenging, and it took nearly a decade until Zeev

Dvir showed a lower bound of $\Omega_d(q^d)$ using polynomial methods. In fact, the early proof uses a connection to Nikodym sets. There are two common definitions for Nikodym sets in the finite field setting: a set $N \subseteq \mathbb{F}_q^d$ is a *Nikodym set* if there is a punctured line $\ell_x \backslash \{x\}$ in N for every $x \in \mathbb{F}_q^d$, and it is a *weak Nikodym set* if the condition only holds for every $x \in \mathbb{F}_q^d \backslash N$. Zeev's earlier proof showed that if $K \subseteq \mathbb{F}_q^d$ is a Kakeya set, then $\mathbb{F}_q K$ is a Nikodym set; separately, it showed that any weak Nikodym set has size $\Omega_d(q^d)$ using polynomial methods.

This work led to the development of polynomial methods in incidence geometry, and we now understand the finite field Kakeya problem fairly well thanks to the work of Boris Bukh and Ting-Wei Chao. In particular, we now know that the minimum size of a Kakeya set in $\mathbb{F}_q^d$ is $(\frac{1}{2^{d-1}} + o(1))q^d$ (for fixed d and growing q). The behavior of the finite field Nikodym problem turned out to be slightly different—all known constructions for weak Nikodym sets are of size $(1 - o(1))q^d$. It was a well-known conjecture in the field that any (weak) Nikodym set in $\mathbb{F}_q^d$ must have size at least $(1 - o(1))q^d$, and this was previously only known when $d = 2$, and when d and the characteristic of $\mathbb{F}_q$ are fixed (by Guo, Kopparty and Sudan).

ChatGPT in the scene

This past summer, when the combinatorics community was reevaluating the shock AI would bring after the disproof of the unit distance conjecture and other results, Ting-Wei was trying to see if ChatGPT was able to solve anything from our very long list of problems. With many prompts, Ting-Wei managed to have ChatGPT 5.6 output a proof that any weak Nikodym set in $\mathbb{F}_q^3$ has size at least $q^3 - O(q^{3-1/5})$. This was already new, and we both jumped straight into reading the proof.

It is common knowledge by now, but at that time I did not know how painful it would be to read through a proof. Thankfully the proof idea was simple enough for me to understand it fairly quickly. I noticed that a step was lossy and tightening it up would give $q^3 - O(q^{3-1/4})$, which was supported by ChatGPT 5.6 after feeding the idea to it. It also generalized the bound to $q^d - O(q^{d-2^{1-d}})$ for general d.

At that time we firmly believed that the answer should be $q^3 - O(q^{3-1/3})$ because of our earlier conjecture, so we kept brainstorming ways to improve it further. Nothing really worked.

Fast forward to the day when Astra came around, Zach Hunter messaged me and told me that Astra one-shot the entire problem, and the true answer was actually (close to) $q^d - O(q^{d-2^{1-d}})$. I was really surprised, and replied that ChatGPT 5.6 actually got a matching lower bound. Zach then told me that he, Cosmin Pohoata and Shengtong Zhang actually got the upper bound with help from ChatGPT 5.6. We decided to join forces and make the result available on Palomar (credit to Shengtong for formalizing and registering it to Palomar). We are also working on a nice writeup of this result, and I will attach that to this post when it is available.

I think it is an interesting case where human inputs were necessary for the earlier versions of ChatGPT but then became unnecessary later…

The main result

As readers have probably noticed, it is more convenient to think about the *complement* of the (weak) Nikodym sets. In fact, when Nikodym first constructed the Nikodym set in $[0,1]^2 \subseteq \mathbb{R}^2$, he actually phrased it in terms of the complement as well.

The main result is that the complement of a (weak) Nikodym set in $\mathbb{F}_d^q$ is of size $O_d(q^{d-2^{1-d}})$, and the exponent is best possible.

Theorem 0.1. *Let $d \geq 2$ be a fixed positive integer. Then any weak Nikodym set $N \subseteq \mathbb{F}_q^d$ satisfies $\left|\mathbb{F}_q^d \backslash N\right| = O_d(q^{d-2^{1-d}})$.*

Moreover, if q is prime, then there exists a Nikodym set $N \subseteq \mathbb{F}_q^d$ with $\left|\mathbb{F}_q^d \backslash N\right| = q^{d-2^{1-d}-o(1)}$.

The upper bound was previously known only for $d = 2$. The construction extends Cosmin's earlier one, which constructs weak Nikodym sets with large complements in $\mathbb{F}_q^2$ for a positive density of primes q.

Both sides of the main result contain some interesting ideas, and I will talk about them separately using $d = 3$ as a demonstration. I personally think it is interesting to think, while going through the proofs, about why we overlooked those proof strategies before.

The bound

Before we begin, let's first reformulate the definition of weak Nikodym sets. Let N be a weak Nikodym set in $\mathbb{F}_q^d$, and let $\mathcal{P} = \mathbb{F}_q^d \backslash N$. The weak Nikodym property is the same as saying that there is a line ℓ_p through each $p \in \mathcal{P}$ that does not pass through any other points in $\mathcal{P}$. If we let $\mathcal{L} = (\ell_p)_{p\in\mathcal{P}}$, this is the same as saying that the incidence graph between $\mathcal{P}$ and $\mathcal{L}$ is exactly a matching! Therefore the maximum complement of a weak Nikodym set is exactly the maximum induced matching of the point-line incidence graph in $\mathbb{F}_q^d$.

This perspective was implicit in many previous arguments, and it was made explicit by Zach, Cosmin, Shengtong and Jacques Verstraete. This change of perspectives turns out to be important later. For now, we say $(\mathcal{P}, \mathcal{L})$ is an *induced matching* if $\mathcal{P} \subseteq \mathbb{F}_q^d$, $\mathcal{L}$ consists of lines defined over $\mathbb{F}_q$, and the lines can be labeled as $\mathcal{L} = (\ell_p)_{p\in\mathcal{P}}$ so that $p' \in \ell_p$ if and only if $p = p'$.

For our convenience, we will work over $\bar{\mathbb{F}}$, the algebraic closure of $\mathbb{F}_q$. This is to distinguish the line ℓ from the set of $\mathbb{F}_q$-points on ℓ.

As we mentioned earlier, we will focus on the case $d = 3$. The basic strategy is simple: given an induced matching $(\mathcal{P}, \mathcal{L})$, we will find a low-degree nonzero polynomial f that vanishes on all lines in $\mathcal{L}$. Passing to the surface $V = Z(f)$ defined by f, we will show that we can find another low-degree nonzero polynomial g vanishing on all lines in $\mathcal{L}$ while not identically on V. Now all lines are in $Z(f, g)$, we can simply apply Bézout's theorem to upper bound the size of $\mathcal{L}$ and thus the size of the induced matching.

There are thus two main steps in $d = 3$, as listed below.

Lemma 0.1 (Top level)**.** *Let $(\mathcal{P}, \mathcal{L})$ be an induced matching in $\mathbb{F}_q^3$ of size at least Cq^2 for some big constant $C > 0$. Then there is a nonzero polynomial $f \in \bar{\mathbb{F}}[x_1, x_2, x_3]$ with*

$$\deg f \lesssim \frac{q^4}{|\mathcal{P}|}$$

and f vanishes identically on all lines in $\mathcal{L}$.

Lemma 0.2 (Next level)**.** *Let V be an irreducible surface in $\bar{\mathbb{F}}^3$. Let $(\mathcal{P}, \mathcal{L})$ be an induced matching contained in V of size at least $Cq \deg V$ for some big constant $C > 0$. Then there is a polynomial*

$g \in \bar{\mathbb{F}}[x_1, x_2, x_3]$ not vanishing identically on V with

$$\deg g \lesssim \frac{q^3 \deg V}{|\mathcal{P}|}$$

and g vanishes identically on all lines in $\mathcal{L}$.

Note that Bézout's theorem and Lemma 0.2 together immediately gives

$$|\mathcal{P}|^2 = |\mathcal{P}|\,|\mathcal{L}| \le |\mathcal{P}| \deg g \deg V \lesssim q^3 (\deg V)^2$$

for any induced matching $(\mathcal{P}, \mathcal{L})$ contained in V. Therefore any induced matching contained in V has size at most $q^{3/2} \deg V$. Plugging this into Lemma 0.1 gives

$$|\mathcal{P}|^2 \lesssim |\mathcal{P}| \cdot q^{3/2} \deg f = q^{11/2}$$

and so $|\mathcal{P}| \lesssim q^{11/4}$, as desired.

I would like to remark that this strategy is very natural in hindsight. In fact Ting-Wei and I have considered similar strategies several times. The difficulty for us is to foresee how the numerology would play out. As one would see later, the leading terms are cancelling everywhere, and without enough faith it is hard to carry out the computation until the very end. Well, that's not an issue for ChatGPT :)

The top level

Lemma 0.1 follows from a fairly standard parameter counting argument. However, even for readers familiar with parameter counting, the bound $q^4/|\mathcal{P}|$ probably looks bizarre. This is related to the fact that leading terms are going to cancel out later.

Let n be a positive integer to be determined. To find f, we will impose the following conditions on f:

1. $\deg f \le (q-1)n - 1$;
2. the multiplicity of f at each $p \in \mathbb{F}_q^3 \backslash \mathcal{P}$ is at least n.

Here, the multiplicity of f at a point p is the degree of the lowest-degree term of f when expanded around p. If f satisfies both conditions, then Bézout's theorem forces f to vanish on each line in $\mathcal{L}$, as each line in $\mathcal{L}$ passes through $q-1$ points in $\mathbb{F}_q^3 \backslash \mathcal{P}$.

Now we carry out the usual parameter counting argument. The space of polynomials of degree at most $(q-1)n-1$ has dimension $\binom{(q-1)n-1+3}{3} = \binom{(q-1)n+2}{3}$. For each $p \in \mathbb{F}_q^3 \backslash \mathcal{P}$, the condition that f vanishes at p to multiplicity at least n imposes $\binom{n-1+3}{3} = \binom{n+2}{3}$ linear conditions. Therefore the space of f satisfying the two conditions has dimension at least

$$\binom{(q-1)n+2}{3} - (q^3 - |\mathcal{P}|)\binom{n+2}{3},$$

and we just need to choose n so that this is positive. A simple computation shows that $\binom{(q-1)n+2}{3} / \binom{n+2}{3} \ge (q-1)^3 - O(\frac{q^3}{n})$. We thus just need to make sure

$$\left((q-1)^3 - O\left(\frac{q^3}{n}\right)\right) - (q^3 - |\mathcal{P}|) > 0.$$

As mentioned earlier, the leading terms q^3 cancel out. As long as $|\mathcal{P}| \ge 4q^2$, it suffices to take n so that $|\mathcal{P}| \ge C' q^3/n$ for some large $C' > 0$. Taking $n = \Theta(q^3/|\mathcal{P}|)$ suffices, producing f with degree $O(q^4/|\mathcal{P}|)$.

The next level

Lemma 0.2 follows essentially the same proof. However, there are two issues that we need to deal with now. First is to enforce that g is not identically zero on V. This is usually done with considering the *Hilbert function* of V.

Definition 0.1. Let $V \subseteq \bar{\mathbb{F}}^3$ be any surface. Let $I_{\leq D}(V)$ be the vector space of polynomials with degree at most D vanishing identically on V. The *Hilbert function*, denoted by h_V, is a function $\mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ sending D to the dimension of the space $P(D)/I_{\leq D}(V)$ where $P(D)$ is the vector space of polynomials with degree at most D.

In other words, we will mod out by polynomials vanishing identically on V. This helps us to ensure that g is not identically zero on V: we just need to make sure that g is a nonzero element in $P(D)/I_{\leq D}(V)$.

However, this gives rise to another issue. Previously, we asked f to vanish to multiplicity n at many points p. This is now ambiguous for g, as shifting by $I_{\leq D}(V)$ might change the Taylor expansion of g with respect to p. A solution to this is to now define the multiplicity of g at p along V to be the maximum multiplicity of $g + h$ at p for any $h \in I(V)$.

Example 0.1. Let's work in $\mathbb{R}^3$ for this example. The polynomial z at $(0, 0, 0)$ has multiplicity 1. If we instead compute this multiplicity along the degree-2 surface $V : xy - z = 0$, then the polynomial z has multiplicity 2 at $(0, 0, 0)$ along V. This is because $h(x, y, z) = xy - z \in I(V)$ and $z + (xy - z) = xy$ has multiplicity 2 at $(0, 0, 0)$.

Geometrically, this is saying that the plane $z = 0$ is tangent to the surface V at $(0, 0, 0)$.

Let $(\mathcal{P}, \mathcal{L})$ be any induced matching in V. Let n be a positive integer to be determined. To find g, we will now impose the following conditions on g:

1. $\deg g \leq (q-1)n - 1$;
2. the multiplicity of g at each $p \in (V \cap \mathbb{F}_q^3) \backslash \mathcal{P}$ along V is at least n.

As before, such g would have to vanish identically on each line in $\mathcal{L}$. Our goal now is to carry out the parameter counting argument within $P(D)/I_{\leq D}(V)$ with $D = (q-1)n - 1$.

To begin the parameter counting argument, we first need to understand the dimension of $P(D)/I_{\leq D}(V)$, which is simply $h_V(D)$. Determining the Hilbert function exactly is generally difficult, but for hypersurfaces this is easy. To see this, suppose that V is defined by a polynomial f, then $\deg f = \deg V$. Moreover, $I_{\leq D}(V)$ is simply the multiples of f that have degree at most D. Therefore the dimension of $I_{\leq D}(V)$ is simply $\binom{D-\deg f+3}{3}$ if $D \geq \deg f$, and 0 if $D < \deg f$. Hence

$$h_V(D) = \binom{D+3}{3} - \binom{\max\{D - \deg f + 3, 0\}}{3}.$$

We also need to understand how many conditions are needed at any $p \in V$ to ensure that g vanishes at p along V with multiplicity at least n. Again, this is usually difficult to determine exactly, but it becomes easier when V is a hypersurface. In our setting, the relevant quantity is the *Hilbert–Samuel multiplicity* $\mu(V, p)$ of V at p. This is the same as the multiplicity of f at p in this case.

Now we want to see how many linear conditions it requires to force g to vanish at p along V with multiplicity at least n. This is the same as requiring the existence of $h \in \bar{\mathbb{F}}[x_1, x_2, x_3]$ so that $g + fh$ vanishes at p with multiplicity at least n. We can take the Taylor expansions of $g + fh$

at p and discard any terms of degrees at least n, and we would like it to be zero. Note that the possible truncated Taylor expansions of fh forms a vector space W of dimension $\binom{n-\mu(V,p)+2}{3}$ if $\mu(V,p) \le n-1$, and 0 otherwise. Therefore we just need to impose

$$\binom{n+2}{3} - \binom{\max\{n-\mu(V,p)+2,0\}}{3}$$

linear conditions to make sure that the truncated Taylor expansion of g lies in W. In some sense, this is the same computation but done locally at p.

For the parameter counting argument to work, we now just need to choose n so that

$$\binom{(q-1)n+2}{3} - \binom{\max\{(q-1)n-\deg f+2,0\}}{3}$$

$$> \sum_{p\in(V\cap\mathbb{F}_q^3)\setminus\mathcal{P}} \left(\binom{n+2}{3} - \binom{\max\{n-\mu(V,p)+2,0\}}{3}\right).$$

Each individual summand on the right hand side is roughly $\mu(V,p)\frac{n^2}{2}$, and so we need to upper bound the sum of the Hilbert–Samuel multiplicities. This is pretty easy to do using the Schwartz–Zippel lemma, which shows that $\sum_{p\in(V\cap\mathbb{F}_q^3)} \mu(V,p) \le q^2 \deg f$. We also have $\mu(V,p) \ge 1$ for every $p \in \mathcal{P} \subseteq V$. Note that if we only look at leading terms, then the left hand side is roughly $q^2 \deg f \frac{n^2}{2}$, and the right hand side can be bounded by roughly $q^2 \deg f \frac{n^2}{2}$. Therefore the leading terms cancel out once again, and we need to be more careful. With some applications of Jensen's inequality and some tedious computation, one can show that $n = O(q^2 \deg f/|\mathcal{P}|) = O(q^2 \deg V/|\mathcal{P}|)$ is enough. This produces g with degree $O(q^3 \deg V/|\mathcal{P}|)$, as desired.

Higher dimensions

The argument above does not immediately generalize to higher dimensions. The difficulty is that once we arrive at some V that is not a hypersurface, we lose control over the Hilbert functions and also the number of conditions needed at each point. To do this, we will need to merge the Schwartz–Zippel lemma and the local counting argument into one step, bounding the sum of the number of linear conditions needed by the Hilbert function of V. This will actually simplify the computation, but it requires many more properties of Hilbert functions (and the *Hilbert–Samuel functions*).

The construction

Now let's turn to the construction. We will assume that q is a prime from now on. Before considering $d = 3$, let's first do the construction for $d = 2$. We will also first construct a weak Nikodym set, and later we will make it a Nikodym set.

The prototype in $\mathbb{R}^2$

Consider a parabola $y = x^2 + c$ in $\mathbb{R}^2$. For each point $(a, a^2 + c)$ in the parabola, its tangent line $y = 2a(x-a) + a^2 + c = 2ax - a^2 + c$ only passes through $(a, a^2 + c)$ and no other points on the parabola. Therefore if we set

$$\mathcal{P} = \{(x,y) \in \mathbb{R}^2 : y = x^2 + c\}$$

and

$$\mathcal{L} = \{\{(x, y) \in \mathbb{R}^2 : y = 2ax - a^2 + c\} : a \in \mathbb{R}\},$$

then $\mathcal{P}$ and $\mathcal{L}$ form an induced matching. As mentioned earlier, this means that the complement of $\mathcal{P}$ forms a weak Nikodym set.

In light of this, for any $p = (x, y) \in \mathbb{R}^2$, let

$$\ell_p = \{(x', y') \in \mathbb{R}^2 : y' = 2x(x - x') + y\}.$$

The previous construction says that if p, p' lie on the same parabola $y = x^2 + c$, then $p' \in \ell_p$ if and only if $p = p'$.

To get a construction in $\mathbb{F}_q^2$, we would like to take a subset of the induced matching and then "modulo q". The most natural attempt is to take the integral points and simply modulo q. Let $\varphi : \mathbb{Z} \to \mathbb{F}_q$ simply be the map modding out by q. For any $p = (x, y) \in \mathbb{F}_q^2$, we also define

$$\ell_p = \{(x', y') \in \mathbb{F}_q^2 : y' = 2x(x - x') + y\}.$$

We would like to take a subset $\mathcal{P} \subseteq \mathbb{Z}^2$ on which φ is injective so that the following two conditions hold.

1. If $p, p' \in \mathcal{P}$ are such that $\varphi(p') \in \ell_{\varphi(p)}$, then $p' \in \ell_p$.
2. If $p, p' \in \mathcal{P}$ are such that $p' \in \ell_p$, then $p = p'$.

Basically the first condition says that there are no wrap-arounds, and the second condition says that we have a weak Nikodym set in $\mathbb{Z}^2$. To avoid wrap-arounds, we simply need to make sure $|x| < cq^{1/2}$ and $|y| < cq$ for some sufficiently small $c > 0$. This gives us $\Theta(q^{3/2})$ choices to work with. Unfortunately, the parabola idea is now too costly: there are $\Theta(q)$ options for $y - x^2$, so even if we pick the most popular parabola, we are left with just $\Theta(q^{1/2})$ choices.

To get around it, we would like essentially the same construction but with the possibilities of $y - x^2$ much more constrained. As in several recent breakthroughs (including disproofs of the unit distance conjecture and the sum-product conjecture), we will replace integers with algebraic integers of higher degrees. In the context of weak Nikodym set, this idea first appeared in Cosmin's earlier work.

Modification using algebraic integers

For demonstration, consider the algebraic integers

$$\mathcal{O} = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} : a, b, c, d \in \mathbb{Z}\}.$$

This is a subring of the ring of algebraic integers in the number field $\mathbb{Q}[\sqrt{2}, \sqrt{3}]$, which is a degree-4 field extension over $\mathbb{Q}$ with four conjugates σ corresponding to $\sigma(\sqrt{2}, \sqrt{3}) = (\pm\sqrt{2}, \pm\sqrt{3})$.

Now suppose that q is a prime for which 2 and 3 are quadratic residues (we will show how to deal with all primes later). Let $s_2, s_3 \in \mathbb{F}_q$ be such that $s_2^2 = 2$ and $s_3^2 = 3$. We will consider the reduction map

$$\varphi : \mathcal{O} \to \mathbb{F}_q \text{ sending } a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} \text{ to } a + bs_2 + cs_3 + ds_2s_3.$$

This map is in fact a ring homomorphism.

Now we would like to find a subset of $\mathcal{O}$ so that within the subset, $\varphi(p') \in \ell_{\varphi(p)}$ implies $p' \in \ell_p$. Suppose $p = (x, y)$ and $p' = (x', y')$. Then we would like

$$\varphi\left((y' - y) - 2x(x' - x)\right) = 0 \Rightarrow (y' - y) - 2x(x' - x) = 0. \tag{1}$$

Note that if $z \in \mathcal{O}$ is nonzero but $\varphi(z) = 0$, we can consider $N(z) = \prod_\sigma \sigma(z)$. This is fixed by all Galois conjugations and thus is an integer. We also know that $\varphi(N(z)) = \prod_\sigma \varphi(\sigma(z)) = 0$. Therefore $|N(z)| \geq q$, and so there is some conjugation σ with $|\sigma(z)| \geq q^{1/4}$. Writing $z = a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}$, this is the same as $|a| + |b|\sqrt{2} + |c|\sqrt{3} + |d|\sqrt{6} \geq q^{1/4}$. Therefore to ensure that Equation 1 holds, we just need to make sure that the $(y' - y) - 2x(x' - x)$ has small coefficients when written as a linear combination of $1, \sqrt{2}, \sqrt{3}, \sqrt{6}$.

In light of this, for any positive real r, let $B_r = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} \in \mathcal{O} : |a|, |b|, |c|, |d| < r\}$, and let $M = cq^{1/4}$ for some sufficiently small c. Then it is clear that Equation 1 holds for any $(x, y), (x', y') \in B_{M^{1/2}} \times B_M$. Note that clearly $|B_r| \gtrsim r^4$. Therefore $|B_{M^{1/2}} \times B_M| = \Theta(q^{3/2})$, similar to the earlier attempt.

The difference now is that instead of taking points (x, y) with the same $y - x^2$, we can do slightly better by considering the *trace map*. The trace map is defined by $\mathrm{tr}(x) = \sum_\sigma \sigma(x)$, which is the same as $4a$ if $x = a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}$. It is a linear map from $\mathcal{O}$ to $\mathbb{Z}$. Moreover, it satisfies that $\mathrm{tr}(x^2) \geq 0$ for any $x \in \mathcal{O}$, and equality holds if and only if $x = 0$. The key claim is the following.

Lemma 0.3. *If $(x, y), (x', y') \in \mathcal{O}^2$ are such that $\mathrm{tr}(y - x^2) = \mathrm{tr}(y' - x'^2)$, then $(x', y') \in \ell_{(x,y)}$ implies $(x, y) = (x', y')$.*

Proof (click to expand)

By the linearity of the trace map, we see that

$$0 = \mathrm{tr}((y' - y) - (x'^2 - x^2)) = \mathrm{tr}\left(2x(x' - x) - (x' + x)(x' - x)\right) = -\mathrm{tr}((x' - x)^2),$$

and so $x' = x$. This forces $y = y'$ as $(x', y') \in \ell_{(x,y)}$.

Therefore we just need to find the most popular $c = \mathrm{tr}(y - x^2)$ among $(x, y) \in B_{M^{1/2}} \times B_M$. Note that for any $(x, y) \in B_{M^{1/2}} \times B_M$, we have $y - x^2 \in B_{CM}$ for some constant C. Thus there are only $O(M) = O(q^{1/4})$ options for $\mathrm{tr}(y - x^2)$. Let c be the most popular trace. Then $\mathcal{P} = \{(x, y) \in B_{M^{1/2}} \times B_M : \mathrm{tr}(y - x^2)\}$ has size $\Omega(q^{3/2}/M) = \Omega(q^{3/2-1/4})$: much better than before! Moreover, $\varphi(\mathcal{P})$ and $\mathcal{L} = \{\ell_{\varphi(p)} : p \in \mathcal{P}\}$ now form an induced matching in $\mathbb{F}_q^2$.

The construction works for primes q for which $2, 3$ are quadratic residues. To extend this to other sufficiently large primes q, we do the following simple trick: note that at least one of $2, 5, 10$ is a quadratic residue modulo q, and at least one of $3, 7, 21$ is a quadratic residue over q. Therefore we will simply consider the same constructions in $\mathbb{Q}[\sqrt{r_1}, \sqrt{r_2}]$ for $r_1 \in \{2, 5, 10\}$ and $r_2 \in \{3, 7, 21\}$. This covers all $q > 7$.

It is not hard to see that if instead of square-roots of two positive integers, we work inside the field $\mathbb{Q}[\sqrt{r_1}, \ldots, \sqrt{r_n}]$ where $r_1, \ldots, r_n \in \mathbb{N}$ are square-free and coprime, then we can take $M = cq^{1/2^n}$ instead, giving $|\mathcal{P}| = \Omega_n(q^{3/2}/M) = \Omega_n(q^{3/2-2^{-n}})$ instead. Taking $n \to \infty$ slowly, this completes the construction of a weak Nikodym set whose complement is of size $q^{3/2-o(1)}$.

From weak Nikodym to Nikodym: Cartesian product

To upgrade the weak Nikodym set to a Nikodym set, we just need a small tweak. We begin with the following observation.

Lemma 0.4. *Let $X, Y \subseteq \mathbb{F}_q$ be such that $\mathbb{F}_q^2 \backslash (X \times Y)$ is a weak Nikodym set. Then $\mathbb{F}_q^2 \backslash (X \times Y)$ is also a Nikodym set.*

> Proof (click to expand)
>
> To show that it is a Nikodym set, it suffices to show that for every $p = (x, y) \notin X \times Y$, there is a line through p that does not intersect $X \times Y$. This is easy: if $x \notin X$, take the vertical line through p; otherwise, we must have $y \notin Y$ and so we can take the horizontal line through p.

Therefore instead of only tracking $\mathrm{tr}(y - x^2)$, we can simply take c_1 to be the most popular $\mathrm{tr}(x^2)$ among $x \in B_{M^{1/2}}$ and c_2 to be the most popular $\mathrm{tr}(y)$ among $y \in B_M$. Since $\mathrm{tr}(x^2)$ and $\mathrm{tr}(y)$ both have $O(M)$ options, we see that $X = \{x \in B_{M^{1/2}} : \mathrm{tr}(x^2) = c_1\}$ has size $\Omega(B_{M^{1/2}}/M)$ and $Y = \{y \in B_M : \mathrm{tr}(y) = c_2\}$ has size $\Omega(B_M/M)$. The complement of $(\varphi(X) \times \varphi(Y))$ is a weak Nikodym set and thus a Nikodym set. Moreover, $|\varphi(X) \times \varphi(Y)| = \Omega_n(q^{3/2}/M^2) = \Omega_n(q^{3/2-2^{-n+1}})$. Therefore we still have a construction whose complement has size $q^{3/2-o(1)}$.

A two-scaled construction in $\mathbb{Z}^3$

Now let's move to $d = 3$ and first think about constructions in $\mathbb{Z}^3$. Previously, our options for x are more limited as we would like to avoid wrap-arounds when we compute $(y' - y) - 2x(x' - x)$. Now, we have an extra coordinate z to work with. To allow more options for x, we will consider the same parabola construction on the xz-plane but at a much coarser scale. The intuition is that even if there are wrap-arounds on the xy-plane, we can already derive the desired (weak) Nikodym property from the xz-plane if x, x' are far away. For x, x' close by, there are no wrap-arounds on the xy-plane anymore, and so we can argue on the xy-plane.

The two-scaled construction looks like the following. Let D be some positive integer to be determined, and let $X = \{s_z D + s_y : s_z \in [L_z], s_y \in [L_y]\}$. To make sure that the coarse scale and the fine scale are separated, assume that $L_y = cD$ for some small $c > 0$. Then for any $x \in X$, we can uniquely decode $s_z(x)$ and $s_y(x)$ from x. Now for any $(x, y, z) \in X \times \mathbb{Z} \times \mathbb{Z}$, let

$$\ell_{(x,y,z)} = \left\{(x', y', z') : (y' - y) = 2s_y(x)(x' - x), (z' - z) = 2s_z(x)(x' - x)\right\}.$$

Suppose that $\mathcal{P} \subseteq X \times \mathbb{Z} \times \mathbb{Z}$ satisfies $y - x^2 = c_y$ and $Dz - x^2 = c_z$. Also suppose that $(x, y, z), (x', y', z')$ are two points in $\mathcal{P}$. I claim that $(x', y', z') \in \ell_{(x,y,z)}$ implies $(x', y', z') = (x, y, z)$. To see this, if $s_z(x) \neq s_z(x')$, we will reach the contradiction by the coarse parabola. Indeed,

$$\begin{aligned} 0 =& (Dz' - x'^2) - (Dz - x^2) = 2s_z(x)D(x' - x) - (x' + x)(x' - x) \\ =& (2(x - s_y(x)) - (x' + x))(x' - x) = -(x' - x + 2s_y(x))(x' - x). \end{aligned}$$

Therefore $x' - x + 2s_y(x) = 0$. However, $|x' - x| \geq D - L_y$ and $\left|s_y(x)\right| \leq L_y$, which is a contradiction when $D > 3L_y$.

If $s_z(x) = s_z(x')$, then we can use the fine parabola $y - x^2 = c_y$:

$$0 = (y' - x'^2) - (y - x^2) = 2s_y(x)(x' - x) - (x' + x)(x' - x) = -(x' - x)(x' - x + 2Ds_z(x)),$$

showing that $x' - x = 0, -2Ds_z(x)$. Since $|x' - x| < L_y < 2D$, this forces $x = x'$, and so $(x, y, z) = (x', y', z')$.

To make sure that this argument is still sound after modding out q, we need to make sure there are no wrap-arounds for $z' - z = 2s_z(x)(x' - x)$. We also need to make sure that when $s_z(x) = s_z(x')$, there are no wrap-arounds for $y' - y = 2s_y(x)(x' - x)$. The latter requirement is the same as $CL_y^2 < q$, and so we will take $L_y, D = \Theta(q^{1/2})$. To make sure that there are no wrap-arounds for $z' - z = 2s_z(x)(x' - x)$, we now need $CDL_z^2 < q$, leading to the choice $L_z = \Theta(q^{1/4})$.

Therefore there are $\Theta(q^{3/4})$ options for x, and $\Theta(q)$ options for both y and z. This gives $\Theta(q^{11/4})$ options for (x, y, z) before fixing c_y and c_z. As in the case of $d = 2$, restricting to $y - x^2 = c_y$ and $Dz - x^2 = c_z$ is still too costly. But the cost can be reduced once again by replacing $\mathbb{Z}$ with algebraic integers of higher degrees.

A brief sketch of the construction for $d = 3$ and beyond

To get a construction for $d = 3$ from the construction above, we still replace $\mathbb{Z}$ with the algebraic integers in $\mathbb{Q}[\sqrt{r_1}, \ldots, \sqrt{r_n}]$. As before, we will select $y, z \in B_M$ where $M = \Theta(q^{1/2^n})$. Now x is chosen from $DB_{M^{1/4}} + B_{M^{1/2}}$ where $D \in \mathbb{N}$ is such that $D/M^{1/2}$ is a sufficiently large constant. One can show that x can still be decoded uniquely as $x = Ds_z(x) + s_y(x)$ with $s_z(x) \in B_{M^{1/4}}$ and $s_y(x) \in B_{M^{1/2}}$. To get an induced matching where the set of points is a Cartesian product, we consider the most popular $(c_x, c_y, c_z) = (\mathrm{tr}(x^2), \mathrm{tr}(y), \mathrm{tr}(Dz))$. There are $O(M^{3/2} \cdot M \cdot M) = O(M^{7/2})$ possibilities for (c_x, c_y, c_z), showing that the number of points we retain is $\Omega(|B_{M^{1/4}}|\,|B_{M^{1/2}}|\,|B_M|^2 / M^{7/2}) = \Omega(q^{11/4 - \frac{7}{2} \cdot 2^{-n}})$. Moreover, those points form an induced matching with the lines $\ell_{\varphi(x,y,z)}$, which can be proven using essentially the same argument discussed above.

For larger d, instead of a two-scaled construction, we now need $(d-1)$ scales. With a correct choice of parameter, one can show a construction of $q^{d - 2^{1-d} - o(1)}$ for all primes q.