A variational restriction theorem

Kova, Vjekoslav; Oliveira E Silva, Diogo

[Link to publication on Research at Birmingham portal](#)

A variational restriction theorem

Vjekoslav Kovač and Diogo Oliveira e Silva

Abstract. We establish variational estimates related to the problem of restricting the Fourier transform of a three-dimensional function to the two-dimensional Euclidean sphere. At the same time, we give a short survey of the recent field of maximal Fourier restriction theory.

Mathematics Subject Classification. 42B10.

Keywords. Fourier restriction, Variational estimates, Gaussian domination.

1. Introduction. This short note serves primarily as a commentary on the very recent topic of pointwise estimates for the operator that restricts the Fourier transform to a hypersurface. We will concentrate exclusively on the case of the two-dimensional unit sphere $\mathbb{S}^2$ in three-dimensional Euclidean space $\mathbb{R}^3$. This both simplifies the exposition and enables the formulation of more general results.

Classical Fourier restriction theory seeks for *a priori* L^p-estimates for the operator $f \mapsto \widehat{f}|_S$, where S is a hypersurface in the Euclidean space. In the case of $\mathbb{S}^2 \subset \mathbb{R}^3$, the endpoint Tomas–Stein inequality [17,19] reads as follows:

$$\big\|\widehat{f}\big|_{\mathbb{S}^2}\big\|_{\mathrm{L}^2(\mathbb{S}^2,\sigma)} \lesssim \|f\|_{\mathrm{L}^{4/3}(\mathbb{R}^3)}. \tag{1.1}$$

Here σ denotes the standard surface measure on $\mathbb{S}^2$. It is well-known that $4/3$ is the largest exponent that can appear on the right-hand side of (1.1) provided that we stick to the L^2-norm on the left-hand side. Estimate (1.1) enables the *Fourier restriction operator*

$$\mathcal{R} \colon \mathrm{L}^{4/3}(\mathbb{R}^3) \to \mathrm{L}^2(\mathbb{S}^2,\sigma) \tag{1.2}$$

to be defined via an approximation of identity argument as follows. Fix a complex-valued Schwartz function χ such that $\int_{\mathbb{R}^3} \chi(x)\,\mathrm{d}x = 1$, and write χ_ε for the L^1-normalized dilate of χ, defined as

$$\chi_\varepsilon(x) := \varepsilon^{-3}\chi(\varepsilon^{-1}x)$$

for $x \in \mathbb{R}^3$ and $\varepsilon > 0$. Given $f \in L^{4/3}(\mathbb{R}^3)$, then, thanks to (1.1), $\mathcal{R}f$ can be defined as the limit

$$\lim_{\varepsilon \to 0^+} (\widehat{f} * \chi_\varepsilon)\big|_{\mathbb{S}^2}$$

in the norm of the space $L^2(\mathbb{S}^2, \sigma)$.

Maximal restriction theorems were recently inaugurated by Müller, Ricci, and Wright [13]. In that work, the authors considered general C^2 planar curves with nonnegative signed curvature equipped with the affine arclength measure, and established a maximal restriction theorem in the full range of exponents where the usual restriction estimate is known to hold. Shortly thereafter, Vitturi [20] provided an elementary argument which leads to a partial generalization to higher-dimensional spheres. In $\mathbb{R}^3$, Vitturi's result covers the full Tomas–Stein range, whose endpoint estimate amounts to

$$\left\| \sup_{\varepsilon > 0} \left| (\widehat{f} * \chi_\varepsilon)(\omega) \right| \right\|_{L^2_\omega(\mathbb{S}^2, \sigma)} \lesssim_\chi \|f\|_{L^{4/3}(\mathbb{R}^3)}. \tag{1.3}$$

An easy consequence of (1.3) and of obvious convergence properties in the dense class of Schwartz functions is the fact that the limit

$$\lim_{\varepsilon \to 0^+} (\widehat{f} * \chi_\varepsilon)(\omega) \tag{1.4}$$

exists for each $f \in L^{4/3}(\mathbb{R}^3)$ and for σ-almost every $\omega \in \mathbb{S}^2$. This enables us to recover the operator (1.2) also in the pointwise sense, and not only in the L^2-norm, which was the main motivation behind the paper [13].

For the elegant proof of (1.3), Vitturi [20] used the following equivalent, non-oscillatory reformulation of the ordinary restriction estimate (1.1):

$$\left| \int_{(\mathbb{S}^2)^2} g(\omega)\overline{g(\omega')}h(\omega - \omega') \, d\sigma(\omega) \, d\sigma(\omega') \right| \lesssim \|g\|_{L^2(\mathbb{S}^2, \sigma)}^2 \|h\|_{L^2(\mathbb{R}^3)}. \tag{1.5}$$

The proof of the equivalence between (1.1) and (1.5) amounts to passing to the adjoint operator (i.e., to a Fourier extension estimate) and expanding out the L^4-norm using Plancherel's identity. These steps make the choice of exponents $4/3$ and 2 into the most convenient one. The advantage of the expanded adjoint formulation (1.5) is that one can easily insert the iterated maximal function operator in h, and simply invoke its boundedness on $L^2(\mathbb{R}^3)$. We refer the reader to [20] for details. In some sense, we will be following a similar step below.

In this paper, we quantify the pointwise convergence result (1.4) in terms of the so-called variational norms. These were introduced by Lépingle [11] in the context of probability theory. Various modifications were then defined and used by Bourgain for numerous problems in harmonic analysis and ergodic theory; see for instance [5]. Recall that, given a function $a \colon (0, \infty) \to \mathbb{C}$ and an exponent $\varrho \in [1, \infty)$, the ϱ-variation seminorm of a is defined as

$$\|a\|_{\widetilde{V}^\varrho} := \sup_{\substack{m\in\mathbb{N}\\ 0<\varepsilon_0<\varepsilon_1<\cdots<\varepsilon_m}} \Big(\sum_{j=1}^m |a(\varepsilon_j) - a(\varepsilon_{j-1})|^\varrho \Big)^{1/\varrho}.$$

In order to turn it into a *ϱ-variation norm*, we simply add the term $|a(\varepsilon_0)|^\varrho$ as follows:

$$\|a\|_{V^\varrho} := \sup_{\substack{m\in\mathbb{N}\\ 0<\varepsilon_0<\varepsilon_1<\cdots<\varepsilon_m}} \Big(|a(\varepsilon_0)|^\varrho + \sum_{j=1}^m |a(\varepsilon_j) - a(\varepsilon_{j-1})|^\varrho \Big)^{1/\varrho}.$$

Since this is not the standard convention in the literature, we make a distinction as we are going to use both quantities. Clearly, $\|a\|_{V^\varrho}$ controls both $\sup_{\varepsilon>0}|a(\varepsilon)|$ and the number of "jumps" of $a(\varepsilon)$ as $\varepsilon \to 0^+$ (and even as $\varepsilon \to \infty$, which is not the interesting case here). A particular instance of the main result of this paper (which is Theorem 1 below) is the following variational generalization of estimate (1.3):

$$\Big\| \big\| (\widehat{f} * \chi_\varepsilon)(\omega) \big\|_{V_\varepsilon^\varrho} \Big\|_{L^2_\omega(\mathbb{S}^2, \sigma)} \lesssim_{\chi, \varrho} \|f\|_{L^{4/3}(\mathbb{R}^3)} \tag{1.6}$$

when $\varrho \in (2, \infty)$ and $f \in L^{4/3}(\mathbb{R}^3)$. The reader can still consider χ to be a fixed Schwartz function, but we are just about to discuss more general possible choices. Variational estimates like (1.6) for various averages and truncations of integral operators have been extensively studied; the papers [2, 6, 7, 9, 12] are just a sample from the available literature. In addition to quantifying the mere convergence, such estimates establish convergence in the whole L^p-space in an explicit and quantitative manner, without the need for pre-existing convergence results on a dense subspace. Later in the text, we will also use the *biparameter ϱ-variation seminorm*, defined for a function of two variables $b \colon (0,\infty) \times (0,\infty) \to \mathbb{C}$ as

$$\|b\|_{\widetilde{W}^\varrho} := \sup_{\substack{m,n\in\mathbb{N}\\ 0<\varepsilon_0<\varepsilon_1<\cdots<\varepsilon_m\\ 0<\eta_0<\eta_1<\cdots<\eta_n}} \Big(\sum_{\substack{1\leq j\leq m\\ 1\leq k\leq n}} \big| b(\varepsilon_j, \eta_k) - b(\varepsilon_j, \eta_{k-1})$$
$$- b(\varepsilon_{j-1}, \eta_k) + b(\varepsilon_{j-1}, \eta_{k-1}) \big|^\varrho \Big)^{1/\varrho}.$$

It is natural to consider more general averaging functions χ; this has already been suggested (albeit somewhat implicitly) in the papers [13, 20]. It is clear from the proof in [20] that the function χ does not need to be smooth. One can, for instance, take χ to be the L^1-normalized indicator function of the unit ball in $\mathbb{R}^3$, in which case the $(\widehat{f} * \chi_\varepsilon)(\omega)$ become the usual Hardy–Littlewood averages of $\widehat{f}$ over Euclidean balls $B(\omega, \varepsilon)$,

$$\frac{1}{|B(\omega, \varepsilon)|} \int_{B(\omega, \varepsilon)} \widehat{f}(y) \mathrm{d}y.$$

Moreover, Ramos [15] concluded that, for each $f \in L^p(\mathbb{R}^3)$ and $1 \leq p \leq 4/3$, almost every point on the sphere is a Lebesgue point of $\widehat{f}$, i.e., for σ-a.e. $\omega \in \mathbb{S}^2$,

we have that

$$\lim_{\varepsilon \to 0+} \frac{1}{|B(\omega, \varepsilon)|} \int_{B(\omega,\varepsilon)} |\widehat{f}(y) - (\mathcal{R}f)(y)| \, dy = 0.$$

Prior to [15], this had been confirmed by Vitturi [20] for functions $f \in L^p(\mathbb{R}^3)$, $1 \le p \le 8/7$, who repeated the two-dimensional argument of Müller, Ricci, and Wright [13].

Subsequent papers [10] and [14], which appeared after the first version of the present paper, generalize the averaging procedure even further, by convolving $\widehat{f}$ with certain averaging measures μ. In light of this more recent research, we now take the opportunity to both generalize (1.6) and complete our short survey of maximal and variational Fourier restriction theories with papers that appeared in the meantime. In what follows, μ will be a finite complex measure defined on the Borel sets in $\mathbb{R}^3$; its dilates μ_ε are now defined as

$$\mu_\varepsilon(E) := \mu(\varepsilon^{-1}E)$$

for every Borel set $E \subseteq \mathbb{R}^3$ and $\varepsilon > 0$. For reasons of elegance, one can additionally assume that μ is *normalized* by $\mu(\mathbb{R}^3) = 1$ and that it is *even*, i.e., centrally symmetric with respect to the origin, which means that

$$\mu(-E) = \mu(E)$$

for each Borel set $E \subseteq \mathbb{R}^3$.

In [10], one of the present authors showed analogues of (1.3) and (1.6) when χ is replaced by a measure μ whose Fourier transform $\widehat{\mu}$ is C^∞ and satisfies the decay condition

$$|\nabla\widehat{\mu}(x)| \lesssim (1 + |x|)^{-1-\delta} \tag{1.7}$$

for some $\delta > 0$. It is interesting to make the following observation, which applies to the cases of two or three dimensions only as things improve in higher dimensions. If one takes μ to be the normalized spherical measure, i.e., $\mu = \sigma/\sigma(\mathbb{S}^2)$, then the decay of $|\nabla\widehat{\mu}(x)|$ as $|x| \to \infty$ is only $O(|x|^{-1})$; see [1]. Consequently, the results from [10] do not apply. This was one of the sources of motivation for Ramos [14], who reused Vitturi's argument from [20] to conclude that the maximal estimate

$$\left\| \sup_{\varepsilon>0} \left| (\widehat{f} * \mu_\varepsilon)(\omega) \right| \right\|_{L^2_\omega(\mathbb{S}^2, \sigma)} \lesssim_\mu \|f\|_{L^{4/3}(\mathbb{R}^3)}$$

holds as soon as the maximal operator $h \mapsto \sup_{\varepsilon>0}(h * \mu_\varepsilon)$ is bounded on $L^2(\mathbb{R}^3)$. Relating this to the work of Rubio de Francia [16], he further deduced the following sufficient condition in terms of the decays of $\widehat{\mu}$ and $\nabla\widehat{\mu}$:

$$|\widehat{\mu}(x)| \lesssim (1 + |x|)^{-\alpha} \text{ and } |\nabla\widehat{\mu}(x)| \lesssim (1 + |x|)^{-\beta} \text{ with } \alpha + \beta > 1.$$

This condition includes the spherical measure as it satisfies $|\widehat{\mu}(x)| = O(|x|^{-1})$ and $|\nabla\widehat{\mu}(x)| = O(|x|^{-1})$ as $|x| \to \infty$.

The main result of this note is a variational estimate which generalizes (1.6) slightly beyond the previously covered cases of the averaging functions χ or measures μ.

Theorem 1. *Suppose that $\varrho \in (1, \infty)$, and that μ is a normalized, even complex measure, defined on the Borel subsets of $\mathbb{R}^3$, satisfying any of the following three conditions:*

(a) $-x \cdot \nabla \widehat{\mu}(x) \geq 0$ *for each* $x \in \mathbb{R}^3$, *and*

$$\left\| \left\| (h * \mu_\varepsilon)(x) \right\|_{\widetilde{V}_\varepsilon^\varrho} \right\|_{L_x^2(\mathbb{R}^3)} \lesssim_{\mu, \varrho} \|h\|_{L^2(\mathbb{R}^3)} \tag{1.8}$$

holds for each Schwartz function h;

(b) $\varrho \in (2, \infty)$, *while* $\widehat{\mu}$ *is* C^2 *and satisfies the decay condition* (1.7);

(c) the inequality

$$\left\| \left\| (h * \mu_\varepsilon * \overline{\mu}_\eta)(x) \right\|_{\widetilde{W}_{\varepsilon, \eta}^\varrho} \right\|_{L_x^2(\mathbb{R}^3)} \lesssim_{\mu, \varrho} \|h\|_{L^2(\mathbb{R}^3)} \tag{1.9}$$

holds for each Schwartz function h.

Then, for each Schwartz function f, the following estimate holds:

$$\left\| \left\| (\widehat{f} * \mu_\varepsilon)(\omega) \right\|_{V_\varepsilon^\varrho} \right\|_{L_\omega^2(\mathbb{S}^2, \sigma)} \lesssim_{\mu, \varrho} \|f\|_{L^{4/3}(\mathbb{R}^3)}. \tag{1.10}$$

Let us immediately clarify one minor technical issue. Theorem 1 claims estimate (1.10) for Schwartz functions f only, but it immediately extends to all $f \in L^{4/3}(\mathbb{R}^3)$ whenever μ is absolutely continuous with respect to the Lebesgue measure. Otherwise, we could run into measurability issues on the left-hand side of (1.10) for singular measures μ.

Condition (a) in Theorem 1 is quite restrictive, but it is satisfied at least when $\varrho > 2$ and the Radon–Nikodym density of μ is a radial Gaussian function. Indeed, if $d\mu(x) = \alpha^3 e^{-\pi\alpha^2|x|^2} dx$ for some $\alpha \in (0, \infty)$, then

$$- x \cdot \nabla \widehat{\mu}(x) = 2\pi\alpha^{-2}|x|^2 e^{-\pi\alpha^{-2}|x|^2} \tag{1.11}$$

is nonnegative, and (1.8) is a standard estimate by Bourgain [5, Lemma 3.28]. In fact, Bourgain formulated (1.8) for one-dimensional Schwartz averaging functions in [5], but the proof carries over to higher dimensions. Alternatively, one can invoke more general results from the subsequent literature, such as the work of Jones, Seeger, and Wright [9], which covered higher-dimensional convolutions, more general dilation structures, and both strong and weak-type variational estimates in a range of L^p-spaces.

Theorem 1 under condition (b) was covered by the paper [10], up to minor technicalities, such as the fact that here we do not need $\widehat{\mu}$ to be smoother than C^2. However, [10] was concerned with more general surfaces and more general measures σ on them, while here we are able to give a more direct proof that is specific to the sphere and to the stated choice of the Lebesgue space exponents. In fact, as we have already noted, the present proof predates [10].

Condition (c) above is somewhat artificial and difficult to verify, but we include it since the proof that it implies (1.10) will be the most straightforward.

Maximal restriction estimates have found a nice application in the very recent work of Bilz [4], who used them to show that there exists a compact set $E \subset \mathbb{R}^3$ of full Hausdorff dimension that does not allow any nontrivial a priori Fourier restriction estimates for any nontrivial Borel measure on E. We do not discuss the details here, but rather refer an interested reader to [4].

1.1. Notation. If $A, B \colon X \to [0, \infty)$ are two functions (or functional expressions) such that, for each $x \in X$, $A(x) \leq CB(x)$ for some unimportant constant $0 \leq C < \infty$, then we write $A(x) \lesssim B(x)$ or $A(x) = O(B(x))$. If the constant C depends on a set of parameters P, we emphasize it notationally by writing $A(x) \lesssim_P B(x)$ or $A(x) = O_P(B(x))$.

We write a variable in the subscript of the letter denoting a function space whenever we need to emphasize with respect to which variable the corresponding (semi)norm is taken. For instance, $\|g(\omega)\|_{L^p_\omega}$ can be written in place of $\|g\|_{L^p}$, while $\|g(\varepsilon)\|_{V^\varrho_\varepsilon}$ can be written in place of $\|g\|_{V^\varrho}$, whenever the variables ω, ε need to be written explicitly.

The *Fourier transform* of a function $f \in L^1(\mathbb{R}^3)$ is normalized as

$$\widehat{f}(y) := \int_{\mathbb{R}^3} f(x) e^{-2\pi i x \cdot y} \mathrm{d}x$$

for each $y \in \mathbb{R}^3$. Here $x \cdot y$ denotes the standard scalar product of vectors $x, y \in \mathbb{R}^3$, and integration is performed with respect to the Lebesgue measure. The map $f \mapsto \widehat{f}$ is then extended, as usual, by continuity to bounded linear operators $L^p(\mathbb{R}^3) \to L^{p'}(\mathbb{R}^3)$ for each $p \in (1, 2]$ and $p' = p/(p-1)$.

More generally, the Fourier transform of a complex measure μ is the function $\widehat{\mu}$ defined as

$$\widehat{\mu}(y) := \int_{\mathbb{R}^3} e^{-2\pi i x \cdot y} \mathrm{d}\mu(x)$$

for each $y \in \mathbb{R}^3$.

The set of complex-valued Schwartz functions on $\mathbb{R}^3$ will be denoted by $\mathcal{S}(\mathbb{R}^3)$.

The remainder of this paper is devoted to the proof of Theorem 1.

2. Proof of Theorem 1. We need to establish (1.10) assuming any one of the three conditions from the statement of Theorem 1. Let us start with condition (a).

Proof of Theorem 1 assuming condition (a). Start by observing that

$$\sup_{\varepsilon_0 > 0} \left|\widehat{f} * \mu_{\varepsilon_0}\right| \leq \left|\widehat{f} * \mu\right| + \sup_{\varepsilon_0 > 0} \left|\widehat{f} * (\mu_{\varepsilon_0} - \mu_1)\right|,$$

that $\widehat{f} * \mu = (f\widehat{\mu})\check{}$, and that the ordinary restriction estimate (1.1) applies to $f\widehat{\mu}$ and yields

$$\left\|\widehat{f} * \mu\right\|_{L^2(\mathbb{S}^2, \sigma)} \lesssim \|f\widehat{\mu}\|_{L^{4/3}(\mathbb{R}^3)} \lesssim_\mu \|f\|_{L^{4/3}(\mathbb{R}^3)}.$$

Thus, inequality (1.10) reduces to two applications of

$$\left\|\left\|(\widehat{f} * \mu_\varepsilon)(\omega)\right\|_{\widetilde{V}^\varrho_\varepsilon}\right\|_{L^2_\omega(\mathbb{S}^2, \sigma)} \lesssim_{\mu, \varrho} \|f\|_{L^{4/3}(\mathbb{R}^3)}, \tag{2.1}$$

which we proceed to establish. The desired estimate (2.1) unfolds as

$$\left\|\sup_{\substack{m\in\mathbb{N}\\0<\varepsilon_0<\varepsilon_1<\cdots<\varepsilon_m}}\left(\sum_{j=1}^{m}\left|(\widehat{f}*(\mu_{\varepsilon_{j-1}}-\mu_{\varepsilon_j}))(\omega)\right|^{\varrho}\right)^{1/\varrho}\right\|_{\mathrm{L}^2_\omega(\mathbb{S}^2,\sigma)}\lesssim_{\mu,\varrho}\|f\|_{\mathrm{L}^{4/3}(\mathbb{R}^3)}.$$

(2.2)

The numbers ε_j in the above supremum can be restricted to a fixed interval $[\varepsilon_{\min},\varepsilon_{\max}]$ with $0<\varepsilon_{\min}<\varepsilon_{\max}$, but the estimate needs to be established with a constant independent of $\varepsilon_{\min}$ and $\varepsilon_{\max}$. Afterwards one simply applies the monotone convergence theorem letting $\varepsilon_{\min}\to0^+$ and $\varepsilon_{\max}\to\infty$. Moreover, by only increasing the left-hand side of (2.2), we can also achieve $\varepsilon_0=\varepsilon_{\min}$ and $\varepsilon_m=\varepsilon_{\max}$.

Next, by continuity, one may further restrict the attention to rational numbers in the interval $[\varepsilon_{\min},\varepsilon_{\max}]$, and, by yet another application of the monotone convergence theorem, one may consider only finitely many values in that interval. In this way, no generality is lost in assuming that the supremum in (2.2) is achieved for some $m\in\mathbb{N}$ and for some measurable functions $\varepsilon_k\colon\mathbb{S}^2\to[\varepsilon_{\min},\varepsilon_{\max}]$, $k\in\{0,1,\ldots,m\}$, such that $\varepsilon_0(\omega)\equiv\varepsilon_{\min}$, $\varepsilon_m(\omega)\equiv\varepsilon_{\max}$. Estimate (2.2) then becomes

$$\left\|\left(\sum_{j=1}^{m}\left|(\widehat{f}*(\mu_{\varepsilon_{j-1}(\omega)}-\mu_{\varepsilon_j(\omega)}))(\omega)\right|^{\varrho}\right)^{1/\varrho}\right\|_{\mathrm{L}^2_\omega(\mathbb{S}^2,\sigma)}\lesssim_{\mu,\varrho}\|f\|_{\mathrm{L}^{4/3}(\mathbb{R}^3)}.$$

Once again, the implicit constant needs to be independent of m and the functions ε_k. The reduction we just performed is an instance of the Kolmogorov–Seliverstov–Plessner linearization method used in [13].

Dualizing the mixed $\mathrm{L}^2_\omega(\ell^\varrho_j)$-norm, see [3], we turn the latter estimate into

$$\left|\Lambda(f,\mathbf{g})\right|\lesssim_{\mu,\varrho}\|f\|_{\mathrm{L}^{4/3}(\mathbb{R}^3)}\left\|\left(\sum_{j=1}^{m}|g_j|^{\varrho'}\right)^{1/\varrho'}\right\|_{\mathrm{L}^2(\mathbb{S}^2,\sigma)},$$

(2.3)

where the bilinear form Λ is defined via

$$\Lambda(f,\mathbf{g}):=\int_{\mathbb{S}^2}\sum_{j=1}^{m}\left(\widehat{f}*(\mu_{\varepsilon_{j-1}(\omega)}-\mu_{\varepsilon_j(\omega)})\right)(\omega)\overline{g_j(\omega)}\,\mathrm{d}\sigma(\omega).$$

Here, $\varrho'=\varrho/(\varrho-1)$ denotes the exponent conjugate to ϱ as usual, and $g_j\colon\mathbb{S}^2\to\mathbb{C}$ are arbitrary measurable functions, $j\in\{1,2,\ldots,m\}$, gathered in a single vector-valued function $\mathbf{g}=(g_j)_{j=1}^{m}$. By elementary properties of the Fourier transform, Λ can be rewritten as

$$\Lambda(f,\mathbf{g})=\int_{\mathbb{R}^3}f(x)\overline{\mathcal{E}(\mathbf{g})(x)}\mathrm{d}x,$$

where $\mathcal{E}$ is a certain extension-type operator given by

$$\mathcal{E}(\mathbf{g})(x):=\int_{\mathbb{S}^2}\sum_{j=1}^{m}\left(\widehat{\mu}(\varepsilon_{j-1}(\omega)x)-\widehat{\mu}(\varepsilon_j(\omega)x)\right)g_j(\omega)e^{2\pi ix\cdot\omega}\,\mathrm{d}\sigma(\omega).$$

(2.4)

By Hölder's inequality, (2.3) is in turn equivalent to

$$\|\mathcal{E}(\mathbf{g})\|_{L^4(\mathbb{R}^3)} \lesssim_{\mu,\varrho} \left\|\left(\sum_{j=1}^m |g_j|^{\varrho'}\right)^{1/\varrho'}\right\|_{L^2(\mathbb{S}^2,\sigma)}. \tag{2.5}$$

If we denote

$$\vartheta(x) := -x \cdot (\nabla\widehat{\mu})(x), \tag{2.6}$$

then we also have

$$-t\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mu}(tx) = -(tx)\cdot(\nabla\widehat{\mu})(tx) = \vartheta(tx)$$

for any $x \in \mathbb{R}^3$, which in turn implies

$$\widehat{\mu}(ax) - \widehat{\mu}(bx) = \int_a^b \vartheta(tx)\frac{\mathrm{d}t}{t} \tag{2.7}$$

for any $0 < a < b$. Substituting this into the definition of $\mathcal{E}$ yields

$$\mathcal{E}(\mathbf{g})(x) = \int_{\varepsilon_{\min}}^{\varepsilon_{\max}} \vartheta(tx)\int_{\mathbb{S}^2} g_{j(t,\omega)}(\omega)e^{2\pi ix\cdot\omega}\,\mathrm{d}\sigma(\omega)\frac{\mathrm{d}t}{t},$$

where, for each $t \in [\varepsilon_{\min}, \varepsilon_{\max})$ and each $\omega \in \mathbb{S}^2$, we denote by $j(t,\omega)$ the unique index $j \in \{1, 2, \ldots, m\}$ such that $t \in [\varepsilon_{j-1}(\omega), \varepsilon_j(\omega))$. Since $\vartheta(x) \geq 0$, by the standing assumption in (a), we can apply the Cauchy–Schwarz inequality in the variable t to estimate

$$|\mathcal{E}(\mathbf{g})(x)|^2 \lesssim_\mu \mathcal{A}(x)\mathcal{B}(\mathbf{g})(x), \tag{2.8}$$

where

$$\mathcal{A}(x) := \int_{\varepsilon_{\min}}^{\varepsilon_{\max}} \vartheta(tx)\frac{\mathrm{d}t}{t}$$

and

$$\mathcal{B}(\mathbf{g})(x) := \int_{\varepsilon_{\min}}^{\varepsilon_{\max}} \vartheta(tx)\left|\int_{\mathbb{S}^2} g_{j(t,\omega)}(\omega)e^{2\pi ix\cdot\omega}\,\mathrm{d}\sigma(\omega)\right|^2\frac{\mathrm{d}t}{t}. \tag{2.9}$$

By (2.7),

$$\mathcal{A}(x) = \widehat{\mu}\big(\varepsilon_{\min}x\big) - \widehat{\mu}\big(\varepsilon_{\max}x\big) \lesssim 1. \tag{2.10}$$

From (2.8) and (2.10), we see that (2.5) will be established once we prove

$$\|\mathcal{B}(\mathbf{g})\|_{L^2(\mathbb{R}^3)} \lesssim_{\mu,\varrho} \left\|\left(\sum_{j=1}^m |g_j|^{\varrho'}\right)^{1/\varrho'}\right\|_{L^2(\mathbb{S}^2,\sigma)}^2. \tag{2.11}$$

Expanding out the square in the definition of $\mathcal{B}(\mathbf{g})(x)$ yields

$$\mathcal{B}(\mathbf{g})(x) = \int_{(\mathbb{S}^2)^2}\int_{\varepsilon_{\min}}^{\varepsilon_{\max}} g_{j(t,\omega)}(\omega)\overline{g_{j(t,\omega')}(\omega')}\,e^{2\pi ix\cdot(\omega-\omega')}\,\vartheta(tx)\frac{\mathrm{d}t}{t}\,\mathrm{d}\sigma(\omega)\,\mathrm{d}\sigma(\omega').$$

$$\tag{2.12}$$

For fixed $\omega, \omega' \in \mathbb{S}^2$, consider

$$
J(\omega, \omega')
$$
$$
:= \{(j, j') \in \{1, 2, \ldots, m\}^2 \; : \; [\varepsilon_{j-1}(\omega), \varepsilon_j(\omega)) \cap [\varepsilon_{j'-1}(\omega'), \varepsilon_{j'}(\omega')) \neq \emptyset\}.
$$

The intersection of two half-open intervals is either the empty set or again a half-open interval. For each pair $(j, j') \in J(\omega, \omega')$, it follows that there exist unique real numbers $a(j, j', \omega, \omega')$ and $b(j, j', \omega, \omega')$, such that

$$
[\varepsilon_{j-1}(\omega), \varepsilon_j(\omega)) \cap [\varepsilon_{j'-1}(\omega'), \varepsilon_{j'}(\omega')) = [a(j, j', \omega, \omega'), b(j, j', \omega, \omega')). \quad (2.13)
$$

Clearly the intervals (2.13) constitute a finite partition of $[\varepsilon_{\min}, \varepsilon_{\max})$. Using (2.7), we can rewrite (2.12) as

$$
\mathcal{B}(\mathbf{g})(x) = \int\limits_{(\mathbb{S}^2)^2} \sum_{(j,j') \in J(\omega, \omega')} g_j(\omega) \overline{g_{j'}(\omega')} \, e^{2\pi i x \cdot (\omega - \omega')}
$$
$$
\times \left(\int\limits_{a(j,j',\omega,\omega')}^{b(j,j',\omega,\omega')} \vartheta(tx) \frac{\mathrm{d}t}{t} \right) \mathrm{d}\sigma(\omega) \, \mathrm{d}\sigma(\omega')
$$
$$
= \int\limits_{(\mathbb{S}^2)^2} \sum_{(j,j') \in J(\omega, \omega')} g_j(\omega) \overline{g_{j'}(\omega')} \, e^{2\pi i x \cdot (\omega - \omega')}
$$
$$
\times \left(\widehat{\mu}\big(a(j, j', \omega, \omega')x\big) - \widehat{\mu}\big(b(j, j', \omega, \omega')x\big) \right) \mathrm{d}\sigma(\omega) \, \mathrm{d}\sigma(\omega').
$$

Taking $h \in \mathcal{S}(\mathbb{R}^3)$ and dualizing with $\widehat{h}$ leads to the form

$$
\Theta(\mathbf{g}, h) := \int\limits_{\mathbb{R}^3} \mathcal{B}(\mathbf{g})(x) \widehat{h}(x) \, \mathrm{d}x.
$$

By Plancherel's identity, we have

$$
\|\mathcal{B}(\mathbf{g})\|_{\mathrm{L}^2(\mathbb{R}^3)} = \sup \left\{ |\Theta(\mathbf{g}, h)| \; : \; h \in \mathcal{S}(\mathbb{R}^3), \; \|h\|_{\mathrm{L}^2(\mathbb{R}^3)} = 1 \right\},
$$

so (2.11) will follow from

$$
|\Theta(\mathbf{g}, h)| \lesssim_{\mu, \varrho} \left\| \left(\sum_{j=1}^m |g_j|^{\varrho'} \right)^{1/\varrho'} \right\|_{\mathrm{L}^2(\mathbb{S}^2, \sigma)}^2 \|h\|_{\mathrm{L}^2(\mathbb{R}^3)}, \quad (2.14)
$$

which we proceed to establish.

Using basic properties of the Fourier transform, the form Θ can be rewritten as

$$
\Theta(\mathbf{g}, h) = \int\limits_{(\mathbb{S}^2)^2} \sum_{(j,j') \in J(\omega, \omega')} g_j(\omega) \overline{g_{j'}(\omega')}
$$
$$
\times \big(h * \mu_{a(j,j',\omega,\omega')} - h * \mu_{b(j,j',\omega,\omega')} \big)(\omega - \omega') \, \mathrm{d}\sigma(\omega) \, \mathrm{d}\sigma(\omega')
$$

for any Schwartz function h. Applying Hölder's inequality to the sum in (j, j'), and recalling the definition of the ϱ-variation yields

$$|\Theta(\mathbf{g}, h)| \leq \int\limits_{(\mathbb{S}^2)^2} \Big(\sum_{j=1}^m |g_j(\omega)|^{\varrho'} \Big)^{1/\varrho'} \Big(\sum_{j=1}^m |g_j(\omega')|^{\varrho'} \Big)^{1/\varrho'}$$
$$\times \|(h * \mu_\varepsilon)(\omega - \omega')\|_{\widetilde{V}_\varepsilon^\varrho} \, d\sigma(\omega) \, d\sigma(\omega').$$

By the usual Tomas–Stein restriction theorem in the formulation (1.5), applied with g replaced by $\big(\sum_{j=1}^m |g_j|^{\varrho'}\big)^{1/\varrho'}$ and with h replaced by $\|h * \mu_\varepsilon\|_{\widetilde{V}_\varepsilon^\varrho}$, we obtain

$$|\Theta(\mathbf{g}, h)| \lesssim \Big\| \Big(\sum_{j=1}^m |g_j|^{\varrho'} \Big)^{1/\varrho'} \Big\|_{L^2(\mathbb{S}^2, \sigma)}^2 \big\| \|h * \mu_\varepsilon\|_{\widetilde{V}_\varepsilon^\varrho} \big\|_{L^2(\mathbb{R}^3)}.$$

Invoking assumption (1.8) completes the proof of estimate (2.14), and therefore also that of (1.10). $\qquad\square$

Next, we will impose condition (b) and reduce the proof to the previous one by replacing μ with a superposition of "nicer" measures.

Proof of Theorem 1 assuming condition (b). We can repeat the same steps as before, reducing (1.10) to (2.11), where $\mathcal{B}$ is as in (2.9) and ϑ is defined by (2.6).

We have already noted that (1.8) is satisfied for measures μ with Gaussian densities, i.e., when $d\mu(x) = \alpha^3 e^{-\pi \alpha^2 |x|^2} \, dx$ for some $\alpha \in (0, \infty)$, and that in this case (1.11) equals $\psi(x/\alpha)$, where

$$\psi(x) := 2\pi |x|^2 e^{-\pi |x|^2}.$$

Therefore, the previous proof of (2.11) specialized to this measure yields

$$\Big\| \int\limits_{\varepsilon_{\min}}^{\varepsilon_{\max}} \psi\Big(\frac{tx}{\alpha}\Big) \Big| \int\limits_{\mathbb{S}^2} g_{j(t, \omega)}(\omega) e^{2\pi i x \cdot \omega} \, d\sigma(\omega) \Big|^2 \frac{dt}{t} \Big\|_{L_x^2(\mathbb{R}^3)} \lesssim_\varrho \Big\| \Big(\sum_{j=1}^m |g_j|^{\varrho'} \Big)^{1/\varrho'} \Big\|_{L^2(\mathbb{S}^2, \sigma)}^2. \tag{2.15}$$

Also note that the L^1-normalization of the above Gaussians guarantees that the left-hand side of (1.8) does not depend on the parameter α. Consequently, the previous proof makes the constant in the bound (2.15) independent of α as well. In this way, estimate (2.11) for a general measure μ satisfying condition (b) will be a consequence of (2.15) and Minkowski's inequality for integrals if we can only dominate $\vartheta(x)$ pointwise as follows:

$$|\vartheta(x)| \lesssim_{\mu, \delta} \int\limits_1^\infty \psi\Big(\frac{x}{\alpha}\Big) \frac{d\alpha}{\alpha^{1+\delta}} \tag{2.16}$$

for each $x \in \mathbb{R}^3$.

Denote by $\Psi(x)$ the right-hand side of (2.16), and observe that $\Psi(0) = 0$ and $\Psi(x) > 0$ for each $x \neq 0$. By continuity and compactness, it suffices to show that the ratio $|\vartheta(x)|/\Psi(x)$ remains bounded as $|x| \to \infty$ or $|x| \to 0^+$.

Substituting $r = \pi\alpha^{-2}|x|^2$, we may rewrite Ψ as

$$\Psi(x) = 2\pi|x|^2 \int\limits_1^\infty e^{-\pi\alpha^{-2}|x|^2} \frac{\mathrm{d}\alpha}{\alpha^{3+\delta}} \tag{2.17}$$

$$= \int\limits_0^{\pi|x|^2} e^{-r} \left(\frac{\sqrt{r}}{\sqrt{\pi}|x|}\right)^\delta \mathrm{d}r. \tag{2.18}$$

From (2.18), we see that $\lim_{|x|\to\infty} \Psi(x)/|x|^{-\delta} \in (0, \infty)$, while decay condition (1.7) gives $|\vartheta(x)| = O(|x|^{-\delta})$ as $|x| \to \infty$.

On the other hand, using Taylor's formula for the function $x \mapsto e^{-\pi\alpha^{-2}|x|^2}$ and substituting into (2.17), we easily obtain

$$\Psi(x) = 2\pi|x|^2 \left(\frac{1}{2+\delta} + O_\delta(|x|^2)\right)$$

on a neighborhood of the origin. Moreover, $\widehat{\mu}$ is C^2 and even since μ is even, and so we have that $(\nabla\widehat{\mu})(0) = 0$. Taylor's formula then yields

$$\vartheta(x) = O_\mu(|x|^2)$$

on a neighborhood of the origin. It follows that $|\vartheta(x)|/\Psi(x) = O_{\mu,\delta}(1)$ for sufficiently small nonzero $|x|$, and this completes the proof of (2.16). □

The Gaussian domination trick which we have just used can be attributed to Stein, see [18, Chapter V, §3.1]. It was generalized and used in a slightly different context by Durcik [8].

Proof of Theorem 1 assuming condition (c). We can repeat the same steps as before that reduce (1.10) to (2.5). This time we define the form Θ differently via

$$\Theta(\mathbf{g}, h) := \int\limits_{\mathbb{R}^3} |\mathcal{E}(\mathbf{g})(x)|^2 \, \widehat{h}(x) \, \mathrm{d}x,$$

where $\mathbf{g}$ is as before and $h \in \mathcal{S}(\mathbb{R}^3)$. Again, by duality, we only need to establish (2.14).

Squaring out (2.4) and inserting that into the above definition of Θ, we obtain

$$\Theta(\mathbf{g}, h) = \int\limits_{\mathbb{R}^3} \int\limits_{(\mathbb{S}^2)^2} \sum_{\substack{1 \le j \le m \\ 1 \le k \le m}} g_j(\omega)\overline{g_k(\omega')} \left(\widehat{\mu}\big(\varepsilon_{j-1}(\omega)x\big) - \widehat{\mu}\big(\varepsilon_j(\omega)x\big)\right)$$

$$\times \left(\overline{\widehat{\mu}\big(\varepsilon_{k-1}(\omega')x\big)} - \overline{\widehat{\mu}\big(\varepsilon_k(\omega')x\big)}\right) \widehat{h}(x) \, e^{2\pi i x \cdot (\omega - \omega')} \, \mathrm{d}\sigma(\omega) \, \mathrm{d}\sigma(\omega') \, \mathrm{d}x,$$

i.e.,

$$
\Theta(\mathbf{g}, h) = \int\limits_{(\mathbb{S}^2)^2} \sum_{\substack{1 \leq j \leq m \\ 1 \leq k \leq m}} g_j(\omega)\overline{g_k(\omega')}
$$
$$
\times \big(h * (\mu_{\varepsilon_{j-1}(\omega)} - \mu_{\varepsilon_j(\omega)}) * \overline{(\mu_{\varepsilon_{k-1}(\omega)} - \mu_{\varepsilon_k(\omega)})}\big)(\omega - \omega')\, \mathrm{d}\sigma(\omega)\, \mathrm{d}\sigma(\omega').
$$

Applying Hölder's inequality to the sum in (j, k), and recalling the definition of the biparameter ϱ-variation seminorm yields

$$
|\Theta(\mathbf{g}, h)| \leq \int\limits_{(\mathbb{S}^2)^2} \Big(\sum_{j=1}^{m} |g_j(\omega)|^{\varrho'}\Big)^{1/\varrho'} \Big(\sum_{k=1}^{m} |g_k(\omega')|^{\varrho'}\Big)^{1/\varrho'}
$$
$$
\times \|(h * \mu_\varepsilon * \overline{\mu}_\eta)(\omega - \omega')\|_{\widetilde{\mathrm{W}}^{\varrho}_{\varepsilon,\eta}}\, \mathrm{d}\sigma(\omega)\, \mathrm{d}\sigma(\omega').
$$

Using (1.5) we obtain

$$
|\Theta(\mathbf{g}, h)| \lesssim \Big\|\Big(\sum_{j=1}^{m} |g_j|^{\varrho'}\Big)^{1/\varrho'}\Big\|^2_{\mathrm{L}^2(\mathbb{S}^2, \sigma)} \big\|\|h * \mu_\varepsilon * \overline{\mu}_\eta\|_{\widetilde{\mathrm{W}}^{\varrho}_{\varepsilon,\eta}}\big\|_{\mathrm{L}^2(\mathbb{R}^3)},
$$

and it remains to invoke the assumption (1.9). This proves (2.14) and, thus, also completes the proof of the theorem assuming condition (c). □

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

[1] Abramowitz, M., Stegun, I.A.: Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover Publications Inc, New York (1992)

[2] Akcoglu, M.A., Jones, R.L., Schwartz, P.O.: Variation in probability, ergodic theory and analysis. Illinois J. Math. **42**(1), 154–177 (1998)

[3] Benedek, A., Panzone, R.: The space L^p, with mixed norm. Duke Math. J. **28**, 301–324 (1961)

[4] Bilz, C.: Large sets without Fourier restriction theorems. arXiv:2001.10016 (2020)

[5] Bourgain, J.: Pointwise ergodic theorems for arithmetic sets. With an appendix by the author, H. Furstenberg, Y. Katznelson, and D. S. Ornstein. Inst. Hautes Études Sci. Publ. Math. **69**, 5–45 (1989)

[6] Bourgain, J., Mirek, M., Stein, E.M., Wróbel, B.: On dimension-free variational inequalities for averaging operators in $\mathbb{R}^d$. Geom. Funct. Anal. **28**(1), 58–99 (2018)

[7] Campbell, J.T., Jones, R.L., Reinhold, K., Wierdl, M.: Oscillation and variation for the Hilbert transform. Duke Math. J. **105**(1), 59–83 (2000)

[8] Durcik, P.: An L^4 estimate for a singular entangled quadrilinear form. Math. Res. Lett. **22**(5), 1317–1332 (2015)

[9] Jones, R.L., Seeger, A., Wright, J.: Strong variational and jump inequalities in harmonic analysis. Trans. Amer. Math. Soc. **360**(12), 6711–6742 (2008)

[10] Kovač, V.: Fourier restriction implies maximal and variational Fourier restriction. J. Funct. Anal. **277**(10), 3355–3372 (2019)

[11] Lépingle, D.: La variation d'ordre p des semi-martingales. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **36**(4), 295–316 (1976)

[12] Mirek, M., Stein, E.M., Zorin-Kranich, P.: A bootstrapping approach to jump inequalities and their applications. Anal. PDE **13**(2), 527–558 (2020)

[13] Müller, D., Ricci, F., Wright, J.: A maximal restriction theorem and Lebesgue points of functions in $\mathcal{F}(L^p)$. Rev. Mat. Iberoam. **35**(3), 693–702 (2019)

[14] Ramos, J. P. G.: Low-dimensional maximal restriction principles for the Fourier transform. arXiv:1904.10858 (2019)

[15] Ramos, J.P.G.: Maximal restriction estimates and the maximal function of the Fourier transform. Proc. Amer. Math. Soc. **148**(3), 1131–1138 (2020)

[16] de Francia, J.L.R.: Maximal functions and Fourier transforms. Duke Math. J. **53**(2), 395–404 (1986)

[17] Stein, E.M.: Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals. Princeton Math. Ser., vol. 43. Princeton Univ. Press, Princeton (1993)

[18] Stein, E.M.: Singular Integrals and Differentiability Properties of Functions. Princeton Math. Ser., vol. 30. Princeton Univ. Press, Princeton (1970)

[19] Tomas, P.: A restriction theorem for the Fourier transform. Bull. Amer. Math. Soc. **81**(2), 477–478 (1975)

[20] Vitturi, M.: A note on maximal Fourier restriction for spheres in all dimensions. arXiv:1703.09495 (2017)

Vjekoslav Kovač
Department of Mathematics, Faculty of Science
University of Zagreb
Bijenička cesta 30
10000 Zagreb
Croatia
e-mail: vjekovac@math.hr

Diogo Oliveira e Silva
School of Mathematics
University of Birmingham
Edgbaston
Birmingham B15 2TT
England
e-mail: d.oliveiraesilva@bham.ac.uk