The Journal of Geometric Analysis (2025) 35:256
https://doi.org/10.1007/s12220-025-01965-x

The One-Dimensional Centred Maximal Function Diminishes the Variation of Indicator Functions

Constantin Bilz[1] · **Julian Weigt**[2]

Received: 6 October 2021 / Accepted: 5 February 2024 / Published online: 7 July 2025

Abstract
We prove sharp local and global variation bounds for the centred Hardy–Littlewood maximal functions of indicator functions in one dimension. We characterise maximisers, treat both the continuous and discrete settings and extend our results to a larger class of functions.

Keywords Maximal function · Bounded variation · Indicator function · Sharp estimate

Mathematics Subject Classification Primary 42B25; Secondary 26A45

1 Introduction

We are concerned with sharp variation bounds for the centred Hardy-Littlewood maximal function Mf on the real line $\mathbb{R}$ defined by

$$Mf(x) = \sup_{r>0} \fint_{x-r}^{x+r} |f(y)|\,\mathrm{d}y = \sup_{r>0} \frac{1}{2r} \int_{x-r}^{x+r} |f(y)|\,\mathrm{d}y.$$

The variation of a function $f\colon \mathbb{R} \to \mathbb{R}$ on an interval $I \subseteq \mathbb{R}$ is

$$\mathrm{var}_I(f) = \sup_{\phi\colon \mathbb{Z}\to I \text{ monotone}} \sum_{i\in\mathbb{Z}} |f(\phi(i)) - f(\phi(i+1))|.$$

We write $\mathrm{var}(f) = \mathrm{var}_{\mathbb{R}}(f)$ and say that f is of *bounded variation* if $\mathrm{var}(f) < \infty$.

✉ Julian Weigt
julian.weigt@warwick.ac.uk

Constantin Bilz
c.bilz@pgr.bham.ac.uk

[1] School of Mathematics, University of Birmingham, Edgbaston, Birmingham B15 2TT, England

[2] Department of Mathematics and Systems Analysis, Aalto University, P.O. Box 11100, FI-00076 Aalto, Finland

Kurka [8] proved that for any such function it holds that

$$\operatorname{var}(Mf) \leq C \operatorname{var}(f) \tag{1.1}$$

for some large constant C independent of f. It is an open conjecture that the optimal constant in this inequality is $C = 1$, see e.g. [4, 8]. The following main result proves this in the case of indicator functions.

Theorem 1.1 *Let $f : \mathbb{R} \to \{0, 1\}$ be a function of bounded variation. Then (1.1) holds with $C = 1$. Equality is attained if and only if f is constant or the set $\{x \in \mathbb{R} | f(x) = 1\}$ is a bounded interval of positive length.*

Note that an indicator function is of bounded variation precisely if it has at most finitely many jumps. This directly implies that $f(x) = 0$ or $f(x) = Mf(x)$ for Lebesgue/almost every $x \in \mathbb{R}$. Our methods only require this weaker assumption, allowing us to prove the following more general result for nonnegative functions.

Theorem 1.2 *Let $f : \mathbb{R} \to [0, \infty)$ be a function of bounded variation such that for almost every $x \in \mathbb{R}$ we have that $f(x) = 0$ or $f(x) = Mf(x)$. Then (1.1) holds with $C = 1$. Equality is attained if and only if f is constant or the set $\{x \in \mathbb{R} | f(x) > 0\}$ is a bounded interval of positive length and for any $x \in \mathbb{R}$,*

$$\liminf_{y \to x} f(y) \leq f(x) \leq \limsup_{y \to x} f(y).$$

The regularity of maximal functions was first studied by Kinnunen [6] who proved that the d-dimensional centred Hardy-Littlewood maximal operator is bounded on the Sobolev space $W^{1,p}(\mathbb{R}^d)$ when $1 < p \leq \infty$ and $d \geq 1$. Hajłasz and Onninen [5] later asked whether the endpoint inequality

$$\|\nabla Mf\|_{L^1(\mathbb{R}^d)} \leq C \|\nabla f\|_{L^1(\mathbb{R}^d)} \tag{1.2}$$

also holds and Kurka's inequality (1.1) provides a positive answer to this question in the one/dimensional case. The higher/dimensional case remains completely open.

In comparison to the one/dimensional *centred* Hardy–Littlewood maximal function, its *uncentred* counterpart

$$\widetilde{M} f(x) = \sup_{x_0 < x < x_1} \frac{1}{x_1 - x_0} \int_{x_0}^{x_1} |f(y)| \, \mathrm{d}y$$

allows averages over a larger class of intervals and hence may be expected to be smoother. Indeed, Tanaka [13] gave a short proof of the uncentred version of (1.1) with $C = 2$ and later Aldaz and Pérez Lázaro [2] showed that the optimal constant is $C = 1$. Ramos [12] studied the sharp version of (1.1) for a family of *nontangential* maximal functions interpolating between the centred and uncentred Hardy–Littlewood maximal functions.

Similarly, higher/dimensional partial results are available for the uncentred maximal function where the corresponding results are not known in the centred case. The first

such result is due to Aldaz and Pérez Lázaro [1] who proved the uncentred version of (1.2) for so/called block decreasing functions. Later, Luiro [10] proved the same for radial functions and the second author [15] proved the corresponding inequality for indicator functions.

1.1 Discrete Setting

Our methods also imply discrete analogues of Theorems 1.1 and 1.2. The discrete centred Hardy-Littlewood maximal function $Mf : \mathbb{Z} \to \mathbb{R}$ of a bounded function $f : \mathbb{Z} \to \mathbb{R}$ is defined by

$$Mf(n) = \sup_{r \in \mathbb{Z}_{\geq 0}} \sum_{m=n-r}^{n+r} |f(m)| = \sup_{r \in \mathbb{Z}_{\geq 0}} \frac{1}{2r+1} \sum_{m=n-r}^{n+r} |f(m)|.$$

For a *discrete interval* $I \subseteq \mathbb{Z}$, i.e. the intersection of $\mathbb{Z}$ and a real interval, the variation of f on I is

$$\mathrm{var}_I(f) = \sum_{n, n+1 \in I} |f(n) - f(n+1)|.$$

We say that f is of *bounded variation* if $\mathrm{var}_{\mathbb{Z}}(f) < \infty$.

Bober, Carneiro, Hughes and Pierce [4] proved that

$$\mathrm{var}_{\mathbb{Z}}(Mf) \leq C \sum_{n \in \mathbb{Z}} |f(n)|$$

for $C = 2 + \frac{146}{315}$. They asked whether the optimal constant in this inequality is $C = 2$ and whether the stronger inequality

$$\mathrm{var}_{\mathbb{Z}}(Mf) \leq C \, \mathrm{var}_{\mathbb{Z}}(f) \tag{1.3}$$

analogous to (1.1) holds. Madrid [11] affirmatively answered the first question and Temur [14] adapted Kurka's method to prove (1.3) with a non/optimal constant. We improve a special case of Temur's result by establishing the optimal constant $C = 1$ in the case of indicator functions.

Theorem 1.3 *Let $f : \mathbb{Z} \to \{0, 1\}$ be a function of bounded variation. Then (1.3) holds with $C = 1$. Equality is attained if and only if f is constant or the set $\{n \in \mathbb{Z} | f(n) = 1\}$ is a bounded nonempty discrete interval.*

In fact this result quickly follows from the continuous Theorem 1.1 and an embedding argument. In the same way, we also establish the following relationship between the optimal constants in the continuous and discrete variation bounds for general functions of bounded variation.

Proposition 1.4 *If (1.1) holds for all functions of bounded variation, then the same is true for (1.3) with the same constant.*

However, we do not know whether a similar embedding argument can be used to prove the following discrete analogue of the stronger Theorem 1.2. This is mainly because of the additional assumptions in these theorems. We circumvent this issue by adapting the proof of Theorem 1.2 to the discrete setting.

Theorem 1.5 *Let $f : \mathbb{Z} \to [0, \infty)$ be a function of bounded variation such that for any $n \in \mathbb{Z}$ we have $f(n) = 0$ or $f(n) = Mf(n)$. Then (1.3) holds with $C = 1$. Equality is attained if and only if f is constant or the set $\{n \in \mathbb{Z} | f(n) > 0\}$ is a bounded nonempty discrete interval.*

Although the proofs of Theorems 1.2 and 1.5 are quite similar, different technical difficulties arise in each case. In the continuous setting, we have to deal with compactness issues and exceptional sets of measure zero. In the discrete setting, one inconvenience is that not every integer interval has an integer midpoint.

1.2 Proof Strategy

Let us explain some ideas of the proofs using the example of the continuous setting. Our main observation is that for a function $f : \mathbb{R} \to [0, \infty)$ satisfying the assumptions of Theorem 1.2, the local variation bound

$$\operatorname{var}_{[a,b]}(Mf) \le \operatorname{var}_{[a,b]}(f) \tag{1.4}$$

holds for any real numbers $a < b$ such that $f(a) = Mf(a)$ and $f(b) = Mf(b)$, i.e. such that Mf is *attached* to f at a and b. Our proof of Theorem 1.2 heavily relies on this property. The following example shows a typical situation. Denote

$$\chi_{[a,b]}(x) = \begin{cases} 1 & \text{if } a < x < b, \\ 1/2 & \text{if } x = a \text{ or } x = b, \\ 0 & \text{otherwise.} \end{cases}$$

Example 1.6 Let $c \in (1, 3)$ and $f = \chi_{[-c,-1]} + \chi_{[1,c]}$. Then Mf is attached to f at any point x with $1 \le |x| \le c$ and

$$\operatorname{var}_{[-1,1]}(Mf) = c^{-1} < 1 = \operatorname{var}_{[-1,1]}(f).$$

The maximal function Mf has a strict local maximum of value $(c-1)/c$ at 0 and two strict local minima of value $(3c-3)/(4c)$ at $\pm c/3$, see Fig. 1.

The calculations leading to Example 1.6 and Fig. 1, as well as to Example 1.8 Figs. 2 and 3 below, are straightforward because for step functions it holds that

$$Mf(x) = \sup_{y \neq x \text{is a jump of} f} \fint_{x-|x-y|}^{x+|x-y|} |f(z)| \, \mathrm{d}z.$$

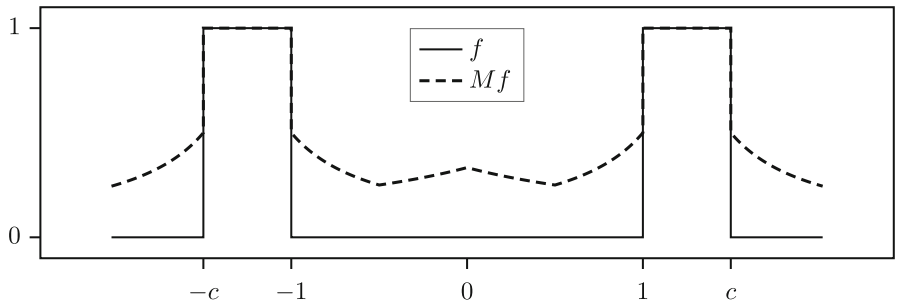

Fig. 1 The functions f and Mf in Example 1.6 with $c = 3/2$

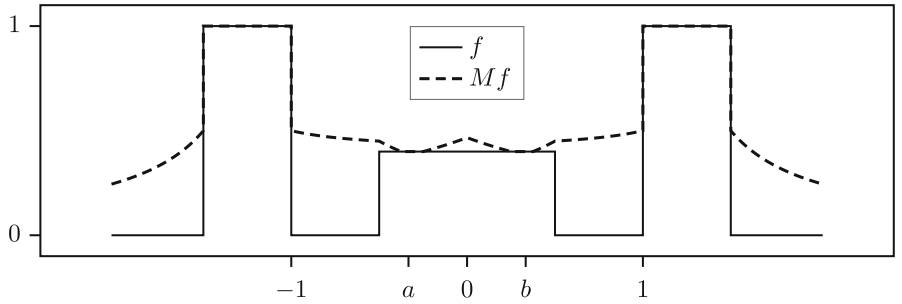

Fig. 2 The functions f and Mf in Example 1.8

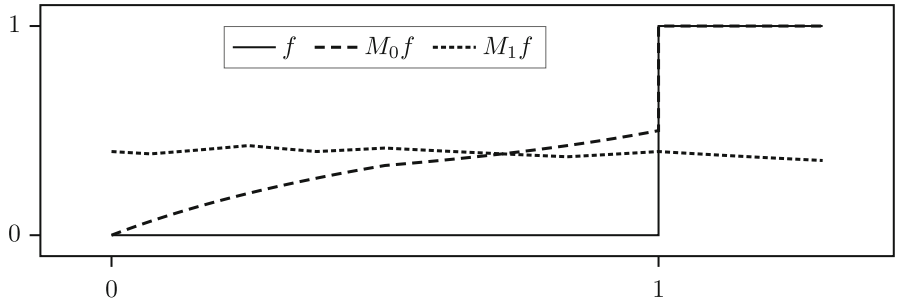

Fig. 3 The auxiliary maximal functions $M_0 f$ and $M_1 f$ on $[0, 1]$ when $f = \chi_{[-5/2,-2]} + \chi_{[-3/2,-1]} + \chi_{[1,2]} + \chi_{[3,7/2]}$

The local variation bound (1.4) will follow from part ((1)) of the following result. An analogue for unbounded intervals is contained in part ((2)).

Proposition 1.7 *Let $f : \mathbb{R} \to [0, \infty)$ be a bounded Lebesgue/measurable function and let $I \subseteq \mathbb{R}$ be an interval such that $f(x) = 0$ for almost every $x \in I$. Then the following holds:*

(1) If $I = [a, b]$ for some real numbers $a < b$, then $\mathrm{var}_{[a,(a+b)/2]}(Mf) \leq Mf(a)$ and $\mathrm{var}_{[(a+b)/2,b]}(Mf) \leq Mf(b)$. Both of these inequalities are strict unless f vanishes almost everywhere on $\mathbb{R}$.

(2) *If* $I = (-\infty, a]$ *or* $I = [a, \infty)$ *for some real* a*, then* Mf *is monotone on* I *and* $\operatorname{var}_I(Mf) = Mf(a) - \inf_{x\in I} Mf(x)$.

We prove this local variation bound in Example 2 and we apply it in Example 3 to show Theorem 1.2 and hence Theorem 1.1. In Sect. 4.2 we prove an analogous discrete local variation bound which we then apply in Sects. 4.3 and 4.4 to show the discrete Theorem 1.5. These proofs can be read mostly independently from Sects. 2 and 3. Section 4.1 contains the embedding argument leading to Proposition 1.4 and the derivation of Theorem 1.3 from Theorem 1.1.

Our approach can be compared to the strategy in [2] for the uncentred Hardy--Littlewood maximal function $\widetilde{M}f$. They show that if $f\colon \mathbb{R} \to \mathbb{R}$ is of bounded variation and satisfies $f(x) = \limsup_{y\to x} f(y)$ for any $x \in \mathbb{R}$, then $\widetilde{M}f \geq f$ and $\widetilde{M}f$ is attached to f at any strict local maximum point of $\widetilde{M}f$. This can be used to show (1.4) when Mf is replaced by $\widetilde{M}f$ and a and b are neighbouring strict local maximum points of $\widetilde{M}f$.

However, in the centred case, Mf is not necessarily attached to f at strict local maxima of Mf, see Example 1.6 above. We overcome this obstruction by making use of a gradient bound for Mf in the proof of Proposition 1.7. On the other hand, this bound becomes less useful for our purposes if a function fails to satisfy the assumptions of Theorem 1.2. In fact, for general functions of bounded variation, the local variation bound (1.4) can fail between some points of attachment. This prevents us from generalising our results to a substantially larger class of functions than in Theorem 1.2.

Example 1.8 Let $h = 2/5$ and $f = \chi_{[-3/2,-1]} + h \cdot \chi_{[-1/2,1/2]} + \chi_{[1,3/2]}$. Then f is constant in $(-1/2, 1/2)$ and Mf is attached to f at any point x with $2 \leq 8|x| \leq 3$, but Mf has a strict local maximum of value $7/15 > h$ at 0. In particular, (1.4) fails between the points of attachment $a = -1/3$ and $b = 1/3$, see Fig. 2.

1.3 Further Remarks

Maximisers and Maximising Sequences

By Theorem 1.1, all indicator functions f satisfy the inequality $\operatorname{var}(Mf) \leq \operatorname{var}(f)$ and equality is attained for example for $f = \chi_{(0,1)}$. This means that $\chi_{(0,1)}$ is a maximiser of (1.1) in the class of indicator functions and thus also in the larger class from Theorem 1.2. However, not every sequence of indicator functions f_n with $\operatorname{var}(Mf_n)/\operatorname{var}(f_n) \to 1$ converges pointwise modulo symmetries to a nonzero maximiser, e.g. take $c \to 1$ in Example 1.6.

Sobolev Variation

Another common notion of variation is given by the total variation $|\mathrm{D}f|(\mathbb{R}^d)$ of the distributional derivative $\mathrm{D}f$, i.e. the measure satisfying the integration by parts rule

$$\int_{\mathbb{R}^d} f\varphi' \,\mathrm{d}x = -\int_{\mathbb{R}^d} \varphi \,\mathrm{d}(\mathrm{D}f)$$

for all functions $\varphi \in C_c^\infty(\mathbb{R}^d)$. The variation of a function on $\mathbb{R}^d$ with $d > 1$ is usually defined in this way. For any function $f\colon \mathbb{R} \to \mathbb{R}$ of bounded variation it holds that $|\mathrm{D}f|(\mathbb{R}) \le \mathrm{var}(f)$. Conversely, if $|\mathrm{D}f|(\mathbb{R}) < \infty$, then there exists a function $\bar{f}$ equal to f almost everywhere such that $\mathrm{var}(\bar{f}) = |\mathrm{D}f|(\mathbb{R})$, see e.g. [9, Theorem 7.2]. If f satisfies the hypotheses of Theorem 1.2, then it follows that

$$|\mathrm{D}Mf|(\mathbb{R}) \le \mathrm{var}(Mf) = \mathrm{var}(M\bar{f}) \le \mathrm{var}(\bar{f}) = |\mathrm{D}f|(\mathbb{R}).$$

Hence Theorem 1.2 remains true for this definition of the variation.

2 Proof of Proposition 1.7

Throughout this section, let $f\colon \mathbb{R} \to [0, \infty)$ be a bounded measurable function. The following result proves the unbounded case in Proposition 1.7((2)). By symmetry, it suffices to take $I = [a, \infty)$.

Lemma 2.1 *Let $a \in \mathbb{R}$ be such that $f(x) = 0$ for almost every $x \ge a$. Then Mf is nonincreasing on $[a, \infty)$ and hence*

$$\mathrm{var}_{[a,\infty)}(Mf) = Mf(a) - \inf_{x\in[a,\infty)} Mf(x).$$

Proof Let $a \le x \le y$. By the definition of Mf and the assumptions on f,

$$Mf(x) = \sup_{r>x-a} \fint_{x-r}^{x+r} f(z)\,\mathrm{d}z \ge \sup_{r>x-a} \fint_{x-r}^{x+r+2(y-x)} f(z)\,\mathrm{d}z = Mf(y).$$

This completes the proof. □

The rest of this section is devoted to the proof of Proposition 1.7((1)), i.e. the case that $I = [a, b]$ for some real numbers $a < b$. It suffices to consider the special case that $a = -1$ and $b = 1$ and to prove the strict inequality

$$\mathrm{var}_{[0,1]}(Mf) < Mf(1) \tag{2.1}$$

under the assumption that $f(x) = 0$ for almost every $x \in [-1, 1]$ and that f does not vanish almost everywhere on $\mathbb{R}$. The general case follows from this because for any nonconstant affine map $\phi\colon \mathbb{R} \to \mathbb{R}$ we have that $M(f \circ \phi)(1) = Mf(\phi(1))$ and

$$\mathrm{var}_{\phi([0,1])}(Mf) = \mathrm{var}_{[0,1]}((Mf) \circ \phi) = \mathrm{var}_{[0,1]}(M(f \circ \phi)).$$

For the proof of (2.1) we first note that Mf restricted to $[0, \infty)$ is the pointwise maximum of the auxiliary maximal functions $M_0 f, M_1 f\colon [0, \infty) \to [0, \infty)$ defined by

$$M_0 f(x) = \sup_{r\le 1+x} \fint_{x-r}^{x+r} f(y)\,\mathrm{d}y, \qquad M_1 f(x) = \sup_{r\ge 1+x} \fint_{x-r}^{x+r} f(y)\,\mathrm{d}y,$$

see Fig. 3 for an example.

Of these, $M_1 f$ only permits averages over large radii. Based on this, our first lemma bounds the difference quotients of $M_1 f$.

Lemma 2.2 *Let $x, y \geq 0$ be distinct and let $r \geq 1 + x$ be such that*

$$M_1 f(x) = \sup_{s \geq r} \fint_{x-s}^{x+s} f(z)\, dz.$$

Then,

$$\frac{M_1 f(x) - M_1 f(y)}{|x - y|} \leq \frac{M_1 f(x)}{r + |x - y|} \leq \frac{M_1 f(y)}{r}.$$

Note that by the definition of $M_1 f$, we can always take r to be at least $1 + x$. The lemma also holds for Mf instead of $M_1 f$, but then we are not guaranteed a good lower bound on r.

Proof We have that $M_1 f(x) < \infty$ since f is bounded. Hence, for any $\epsilon > 0$ there exists an $s \geq r$ such that $(1 - \epsilon) M_1 f(x) \leq \fint_{x-s}^{x+s} f(z)\, \mathrm{d}z$ and therefore,

$$\begin{aligned}
(1 - \epsilon) M_1 f(x) - M_1 f(y) &\leq \fint_{x-s}^{x+s} f(z)\, \mathrm{d}z - \fint_{y-s-|x-y|}^{y+s+|x-y|} f(z)\, \mathrm{d}z \\
&\leq \left(\frac{1}{2s} - \frac{1}{2s + 2|x - y|} \right) \int_{x-s}^{x+s} f(z)\, \mathrm{d}z \\
&= \frac{|x - y|}{s + |x - y|} \fint_{x-s}^{x+s} f(z)\, \mathrm{d}z \\
&\leq \frac{|x - y|}{r + |x - y|} M_1 f(x).
\end{aligned}$$

The first inequality uses the definition of $M_1 f(y)$ together with the fact that $s + |x - y| \geq 1 + y$. In the second inequality, we use the nonnegativity of f to reduce the domain of integration of the second integral. The last two relations follow from definitions. Now the first inequality in the lemma follows by letting $\epsilon \to 0$. The second inequality follows after rearranging terms. □

Bounds similar to Lemma 2.2 have frequently appeared in the literature, including in higher dimensions. The related inequality $|\nabla M_\alpha f(x)| \leq C M_{\alpha-1} f(x)$ for the fractional maximal function $M_\alpha f$ with $1 \leq \alpha \leq d$ was proved in [7]. A generalisation to the range $0 \leq \alpha \leq d$ can be found in [3, Section 2.5].

We now employ the previous result to prove a local variation bound for $M_1 f$. The strictness of this inequality will be crucial to the characterisation of maximisers in our results.

Lemma 2.3 *It holds that $\mathrm{var}_{[0,1]}(M_1 f) \leq M_1 f(1)$ and this inequality is strict if $f(x) = 0$ for almost every $x \in [-1, 1]$ and $f(x) > 0$ for any x in some set of positive measure.*

Proof First assume that $f(x) = 0$ for almost every $x \in [-1, 1]$ and $f(x) > 0$ for any x in some set of positive measure. Then $M_1 f(x) > 0$ for any $x \geq 0$. By Lemma 2.2, $M_1 f$ is continuous. Since the map $(x, s) \mapsto \fint_{x-s}^{x+s} f(y)\,\mathrm{d}y$ is continuous at $(x, s) = (0, 1)$ and $f(y) = 0$ for almost every $y \in [-1, 1]$, this implies the existence of a $\delta \in (0, 1)$ such that for any $x \in [0, \delta)$,

$$M_1 f(x) = \sup_{s \geq 1+x+\delta} \fint_{x-s}^{x+s} f(y)\,\mathrm{d}y$$

and hence f and x satisfy the hypotheses of Lemma 2.2 with $r = 1+x+\delta$. Without the additional assumptions that $f(x) = 0$ for almost every $x \in [-1, 1]$ and that $f(x) > 0$ for any x in some set of positive measure, this remains true for $\delta = 0$.

In order to estimate the variation of $M_1 f$ on $[0, 1]$, we let $k \geq 1$ and

$$0 = x_0 < x_1 < \ldots < x_k = 1$$

We write $\delta_i = \delta$ if $x_i < \delta$ and $\delta_i = 0$ otherwise. By the two inequalities in Lemma 2.2 and the monotonicity of the sequences x_i and δ_i,

$$\begin{aligned}
&\sum_{i=0}^{k-1} |M_1 f(x_i) - M_1 f(x_{i+1})| \\
&\quad \leq \sum_{i=0}^{k-1} \max\Big(\frac{M_1 f(x_i)}{1 + x_{i+1} + \delta_i}, \frac{M_1 f(x_{i+1})}{1 + 2x_{i+1} - x_i + \delta_{i+1}}\Big)(x_{i+1} - x_i) \\
&\quad \leq \sum_{i=0}^{k-1} \max(M_1 f(x_i), M_1 f(x_{i+1})) \frac{x_{i+1} - x_i}{1 + x_i + \delta_{i+1}} \\
&\quad \leq \sum_{i=0}^{k-1} \max\Big(\frac{2 + \delta_i}{1 + x_i + \delta_i}, \frac{2 + \delta_{i+1}}{1 + x_{i+1} + \delta_{i+1}}\Big) \frac{x_{i+1} - x_i}{1 + x_i + \delta_{i+1}} M_1 f(1) \\
&\quad \leq \sum_{i=0}^{k-1} \frac{(2 + \delta_i)(x_{i+1} - x_i)}{(1 + x_i + \delta_{i+1})^2} M_1 f(1).
\end{aligned}$$

Except for the summand corresponding to the case $x_i < \delta \leq x_{i+1}$, this is a Riemann sum and therefore

$$\begin{aligned}
\mathrm{var}_{[0,1]}(M_1 f) &\leq \Big(\int_0^\delta \frac{2 + \delta}{(1 + x + \delta)^2}\,\mathrm{d}x + \int_\delta^1 \frac{2}{(1 + x)^2}\,\mathrm{d}x\Big) M_1 f(1) \\
&\leq \int_0^1 \frac{2}{(1 + x)^2}\,\mathrm{d}x \cdot M_1 f(1) \\
&= M_1 f(1)
\end{aligned}$$

and the second inequality is strict if $\delta > 0$. This completes the proof. □

Remark 2.4 Let us sketch a shorter but less elementary version of the second part of the above proof. By Lemma 2.2, the auxiliary maximal function $M_1 f$ is Lipschitz continuous. Hence it is differentiable almost everywhere and

$$\operatorname{var}_{[0,1]}(M_1 f) = \int_0^1 |(M_1 f)'(x)| \, \mathrm{d}x.$$

At any point of differentiability $x \in (0, 1)$ we have by Lemma 2.2 that

$$|(M_1 f)'(x)| \leq \frac{M_1 f(x)}{1+x} \leq \frac{2M_1 f(1)}{(1+x)^2}$$

and the first inequality is strict in some neighbourhood of 0. Plugging this into the above formula for $\operatorname{var}_{[0,1]}(M_1 f)$ yields Lemma 2.3.

The next lemma concerns the other auxiliary maximal function $M_0 f$.

Lemma 2.5 *Let $f(x) = 0$ for almost every $x \in [-1, 1]$. Then $M_0 f$ is nondecreasing on $[0, 1]$.*

Proof This is similar to the proof of Lemma 2.1. Let $0 < x \leq y \leq 1$. Then,

$$M_0 f(x) = \sup_{1-x<r\leq 1+x} \fint_{x-r}^{x+r} f(z) \, \mathrm{d}z \leq \sup_{1-x<r\leq 1+x} \fint_{x-r+2(y-x)}^{x+r} f(z) \, \mathrm{d}z \leq M_0 f(y).$$

Since $M_0 f(0) = 0$ and $M_0 f$ is nonnegative, this completes the proof. □

We have established the monotonicity of $M_0 f$ in Lemma 2.5 and a variation bound for $M_1 f$ in Lemma 2.3. The next result will allow us to deduce a variation bound for the pointwise maximum $Mf = \max(M_0 f, M_1 f)$.

Lemma 2.6 *Let $g, h \colon [0, 1] \to \mathbb{R}$ be functions such that $g(1) \leq h(1)$ and let g be nondecreasing. Then $\operatorname{var}_{[0,1]}(\max(g, h)) \leq \operatorname{var}_{[0,1]}(h)$.*

Proof Write $u = \max(g, h)$. We need to show that for any $k \geq 1$ and any

$$0 = x_0 < x_1 < \ldots < x_k = 1$$

it holds that

$$\sum_{i=0}^{k-1} |u(x_i) - u(x_{i+1})| \leq \operatorname{var}_{[0,1]}(h).$$

Write $x_{k+1} = 1$ and let $p(-1) < p(0) < \ldots < p(\ell)$ be the elements of the set

$$P = \{-1\} \cup \{i \in \{0, \ldots, k\} | h(x_i) \geq u(x_{i+1})\}.$$

Clearly, $p(-1) = -1$. Since $x_k = x_{k+1} = 1$ and by assumption, $h(x_k) = u(x_{k+1})$ and hence $p(\ell) = k$. If $i \in \{0, \dots, k\} \setminus P$, then $h(x_i) < u(x_{i+1})$. Since g is nondecreasing, we also have that $g(x_i) \le g(x_{i+1})$ and hence $u(x_i) \le u(x_{i+1})$. On the other hand, if $i \in P \setminus \{-1\}$, then

$$h(x_i) \ge u(x_{i+1}) \ge g(x_{i+1}) \ge g(x_i).$$

Hence, $h(x_i) = u(x_i)$ and $u(x_i) \ge u(x_{i+1})$. This shows that for any $0 \le j \le \ell$,

$$u(x_{p(j-1)+1}) \le u(x_{p(j-1)+2}) \le \dots \le u(x_{p(j)}) = h(x_{p(j)}) \ge u(x_{p(j)+1}).$$

We conclude that

$$\begin{aligned}
\sum_{i=0}^{k-1} |u(x_i) - u(x_{i+1})| &= \sum_{j=0}^{\ell} \sum_{i=p(j-1)+1}^{p(j)} |u(x_i) - u(x_{i+1})| \\
&= \sum_{j=0}^{\ell} 2h(x_{p(j)}) - u(x_{p(j-1)+1}) - u(x_{p(j)+1}) \\
&\le \sum_{j=0}^{\ell} 2h(x_{p(j)}) - h(x_{p(j-1)+1}) - h(x_{p(j)+1}) \\
&\le \operatorname{var}_{[0,1]}(h).
\end{aligned}$$

This completes the proof. □

We are now ready to prove (2.1). Let $f(x) = 0$ for almost every $x \in [-1, 1]$ and let $h\colon [0, 1] \to [0, \infty)$ be the function defined by $h(x) = M_1 f(x)$ for $0 \le x < 1$ and $h(1) = Mf(1)$. Then $M_0 f(1) \le h(1)$ and Mf restricted to $[0, 1]$ is the pointwise maximum of $M_0 f$ and h. Hence by an application of Lemmas 2.5 and 2.6 and then Lemma 2.3,

$$\operatorname{var}_{[0,1]}(Mf) \le \operatorname{var}_{[0,1]}(h) \le \operatorname{var}_{[0,1]}(M_1 f) + Mf(1) - M_1 f(1) \le Mf(1).$$

The last inequality is strict if f does not vanish almost everywhere on $\mathbb{R}$. This shows (2.1) and hence completes the proof of Proposition 1.7.

Remark 2.7 One can show that $M_0 f$ is also Lipschitz continuous on $[0, 1]$ and that $M_0 f$ and $M_1 f$ do not coincide at more than one point in $[0, 1]$ if f does not vanish almost everywhere on $\mathbb{R}$. Let us only sketch a proof of the fact that if $y \in [0, 1]$ is such that $M_0 f(y) \ge M_1 f(y)$, then $M_0 f(x) > M_1 f(x)$ for any $x \in (y, 1]$. By Lemma 2.2,

$$\frac{M_1 f(x) - M_1 f(y)}{x - y} \le \frac{M_1 f(y)}{1 + x}.$$

Similarly as in the proofs of Lemmas 2.2 and 2.5, one can show that

$$\frac{M_0 f(x) - M_0 f(y)}{x - y} \geq \frac{M_0 f(y)}{1 - x + 2y}.$$

Since $M_0 f(y) \geq M_1 f(y) > 0$ and $y < x$ it follows that $M_0 f(x) > M_1 f(x)$.

3 Proof of Theorem 1.2

Throughout this section, let $f : \mathbb{R} \to [0, \infty)$ be a function of bounded variation such that for almost every $x \in \mathbb{R}$ we have that $f(x) = 0$ or $f(x) = Mf(x)$. In order to prove Theorem 1.2, we need to show the inequality

$$\operatorname{var}(Mf) \leq \operatorname{var}(f) \tag{3.1}$$

and determine its cases of equality. We will accomplish this by using a certain canonical representative $\bar{f}$ whose properties facilitate the application of Proposition 1.7. In Sect. 3.1, we define $\bar{f}$, show that f and $\bar{f}$ agree almost everywhere and that

$$\operatorname{var}(\bar{f}) \leq \operatorname{var}(f). \tag{3.2}$$

There, we also establish some further properties of $\bar{f}$. In Sect. 3.2, we apply Proposition 1.7 to show (3.1) for $\bar{f}$, i.e. we show that

$$\operatorname{var}(Mf) \leq \operatorname{var}(\bar{f}). \tag{3.3}$$

Inequalities (3.2) and (3.3) together imply (3.1). In Sect. 3.3, we characterise the cases of equality in (3.1) by characterising the cases of equality in (3.3) and then characterising the cases of equality in (3.2) under the assumption that equality holds in (3.3).

3.1 Canonical Representative

Let us define a function $\bar{f} : \mathbb{R} \to [0, \infty)$ as follows. If $x \in \mathbb{R}$ is such that

$$\limsup_{r \searrow 0} \fint_{x-r}^{x+r} f(y)\,\mathrm{d}y = 0, \tag{3.4}$$

then we let $\bar{f}(x) = 0$ and otherwise we let $\bar{f}(x) = Mf(x)$. This *canonical representative* is related to but distinct from the homonymous object in [2]. By the Lebesgue differentiation theorem and the assumption on f, we have that $f(x) = \bar{f}(x)$ for almost every $x \in \mathbb{R}$ and hence $Mf(x) = M\bar{f}(x)$ for any $x \in \mathbb{R}$. Since f is of bounded variation, its one/sided limits exist at any point. It follows that (3.4) can be rewritten without the use of an integral, but we will not need this.

The following lemma will be used multiple times throughout this section.

Lemma 3.1 *The maximal function Mf is lower semi/continuous, i.e. for any $x \in \mathbb{R}$ it holds that* $\liminf_{y\to x} Mf(y) \geq Mf(x)$.

Proof By definition, Mf is the pointwise supremum of the continuous functions

$$\mathbb{R} \ni x \mapsto \fint_{x-r}^{x+r} f(y)\,\mathrm{d}y, \quad r > 0.$$

The lemma follows from this. □

We now show that the canonical representative does not increase the variation.

Lemma 3.2 *Inequality (3.2) holds.*

Proof We first claim that it suffices to show that for any $x \in \mathbb{R}$ and $\epsilon > 0$ there exist $y_1, y_2 \in (x-\epsilon, x+\epsilon)$ such that $f(y_1) - \epsilon \leq \bar{f}(x) \leq f(y_2) + \epsilon$. Let $k \geq 1$ and let

$$-\infty < x_0 < x_1 < \ldots < x_k < \infty.$$

By iteratively removing any points x_i with $1 \leq i \leq k-1$ for which $\bar{f}(x_i)$ lies in the convex hull of $\{\bar{f}(x_{i-1}), \bar{f}(x_{i+1})\}$, we obtain a subsequence $x_0' < \ldots < x_\ell'$ such that

$$\sum_{i=0}^{k-1} |\bar{f}(x_i) - \bar{f}(x_{i+1})| = \sigma \sum_{i=0}^{\ell-1} (-1)^i \bar{f}(x_i') + (-1)^{i+1} \bar{f}(x_{i+1}')$$

for some $\sigma \in \{-1, 1\}$. Let $\epsilon > 0$. By the claim made at the beginning of the proof, there exist points $y_i \in (x_i' - \epsilon, x_i' + \epsilon)$ such that

$$\sigma(-1)^i \bar{f}(x_i') \leq \sigma(-1)^i f(y_i) + \epsilon$$

for any $0 \leq i \leq \ell$. If ϵ is small enough, then y_i is increasing in i and hence

$$\sum_{i=0}^{k-1} |\bar{f}(x_i) - \bar{f}(x_{i+1})| - 2\ell\epsilon \leq \sigma \sum_{i=0}^{\ell-1} (-1)^i f(y_i) + (-1)^{i+1} f(y_{i+1}) \leq \mathrm{var}(f).$$

Let $\epsilon \to 0$ and then take the supremum over all k and x_i as above to show (3.2).

It remains to show that for any $x \in \mathbb{R}$ and $\epsilon > 0$ there exist points y_1 and y_2 as above. Let $r \in (0, \epsilon)$. We start with the existence of y_1. By the definitions of $\bar{f}(x)$ and $Mf(x)$,

$$\bar{f}(x) \geq \limsup_{r \searrow 0} \fint_{x-r}^{x+r} f(y)\,\mathrm{d}y.$$

Hence if r is sufficiently small, then the integral on the right/hand side is at most $\bar{f}(x) + \epsilon$ and so there exists a $y_1 \in (x-r, x+r)$ with $f(y_1) - \epsilon \leq \bar{f}(x)$, as required.

We complete the proof by showing the existence of y_2. If $\bar{f}(x) = 0$, then we may simply choose $y_2 = x$ because f is nonnegative. So we assume that $\bar{f}(x) = Mf(x) > 0$. Since (3.4) fails, $f(y) > 0$ for any y in some subset of positive measure of $(x - r, x + r)$. As f and $\bar{f}$ are equal almost everywhere, it follows that $f(y_2) = \bar{f}(y_2) = Mf(y_2)$ for some $y_2 \in (x - r, x + r)$. Hence if r is small enough, then Lemma 3.1 implies that $f(y_2) + \epsilon \geq \bar{f}(x)$, as required. □

In particular, (3.2) shows that $\bar{f}$ is of bounded variation. Together with the definition of $\bar{f}$, this implies some topological properties of the vanishing set

$$V = \{x \in \mathbb{R} | \bar{f}(x) = 0\}.$$

Lemma 3.3 *The set V is open and its boundary has no limit points in $\mathbb{R}$.*

This can be stated equivalently as follows: There exists a finite or countably infinite nondecreasing sequence of points $a_i \in \mathbb{R} \cup \{\pm\infty\}$ without accumulation points in $\mathbb{R}$ such that $V = \bigcup_i (a_{2i}, a_{2i+1})$.

Proof If f vanishes almost everywhere, then $V = \mathbb{R}$ and the lemma follows. Since f is nonnegative, we may therefore assume that f is positive in a set of positive measure. Let $x \in \mathbb{R}$. Then $Mf(x) > 0$ and by Lemma 3.1 there exists an $\epsilon > 0$ such that $Mf(y) > \epsilon$ for any $y \in (x - \epsilon, x + \epsilon)$.

We first show that x is not a limit point of the boundary of V. Let $k \geq 0$ and let

$$x - \epsilon < x_0 < x_1 < \ldots < x_{2k+1} < x + \epsilon$$

be a sequence of points with $\bar{f}(x_{2i}) = 0$ and $\bar{f}(x_{2i+1}) = M\bar{f}(x_{2i+1})$ for any $0 \leq i \leq k$. It suffices to show that k is bounded by a constant that only depends on $\bar{f}$ and ϵ. Such a bound holds because

$$(2k + 1)\epsilon < \sum_{i=0}^{2k} |\bar{f}(x_{i+1}) - \bar{f}(x_i)| \leq \mathrm{var}(\bar{f}) < \infty.$$

The first inequality above holds by the properties of ϵ and x_k. The second inequality holds by definition. Hence x is not a limit point of the boundary of V.

It remains to show that V is open. To this end, let x be a boundary point of V. We need to show that $\bar{f}(x) > 0$. By the first part of the proof, $f(y) > \epsilon$ for any y in some one/sided neighbourhood of x, i.e. for any y in $(x - r, x)$ or $(x, x + r)$ for some $r > 0$. Since f is nonnegative, we see that (3.4) fails and hence $\bar{f}(x) = Mf(x) > 0$. This completes the proof. □

3.2 Global Variation Bound

In Sect. 3.1, we proved (3.2). Together with the following result, this implies (3.1), proving the first part of Theorem 1.2.

Proposition 3.4 *Inequality (3.3) holds.*

Proof By Lemma 3.3 and a subdivision of $\mathbb{R}$ we see that (3.3) holds if

$$\mathrm{var}_I(Mf) \leq \mathrm{var}_I(\bar{f}) \tag{3.5}$$

whenever I is a connected component of $\mathbb{R}\setminus V$ or the closure of a connected component of V. If I is a connected component of $\mathbb{R}\setminus V$, then $\bar{f}$ and Mf agree on I, so that (3.5) holds with equality. Now let I be the closure of a connected component of V. If $I=\mathbb{R}$, then both sides of (3.5) are zero. On the other hand, if $I\neq\mathbb{R}$, then by Lemma 3.3, $\bar{f}$ and Mf agree on the boundary of I and therefore (3.5) follows from either ((1)) or ((2)) in Proposition 1.7. This completes the proof. □

3.3 Cases of Equality

It remains to characterise the cases of equality in (3.1). We first establish certain regularity properties of $\bar{f}$.

Lemma 3.5 *Any connected component of V or $\mathbb{R}\setminus V$ has positive length.*

Proof Let $x\in\mathbb{R}$. If $\bar{f}(y)>0$ for any $y\neq x$ in some compact neighbourhood of x, then by Lemma 3.1 there exists an $\epsilon>0$ such that $\bar{f}(y)=Mf(y)>\epsilon$ for any such y. Hence (3.4) fails and $\bar{f}(x)=Mf(x)>\epsilon$. This shows that $\{x\}$ is not a connected component of V.

On the other hand, if $\bar{f}(y)=0$ for any $y\neq x$ in some neighbourhood of x, then (3.4) holds and hence $\bar{f}(x)=0$. This shows that $\{x\}$ is not a connected component of $\mathbb{R}\setminus V$. Since $x\in\mathbb{R}$ was arbitrary, it follows that any connected component of V or $\mathbb{R}\setminus V$ has positive length. □

Now we investigate the behaviour of the canonical representative $\bar{f}$ on connected components of its support

$$\mathbb{R}\setminus V=\{x\in\mathbb{R}|\,\bar{f}(x)>0\}.$$

This set is closed by Lemma 3.3. Our next result will only be applied in the case of an unbounded connected component, but its proof is identical in the bounded and unbounded cases.

Lemma 3.6 *The function $\bar{f}$ is concave on any connected component of $\mathbb{R}\setminus V$.*

Proof Suppose for a contradiction that $\bar{f}$ is not concave on some connected component I of $\mathbb{R}\setminus V$. Then there exist points $x_0<x_1<x_2$ in I such that $\bar{f}(x_1)<L(x_1)$ where $L:\mathbb{R}\to\mathbb{R}$ is the affine linear function defined by $L(x_0)=\bar{f}(x_0)$ and $L(x_2)=\bar{f}(x_2)$. Hence for $g=\bar{f}-L$ we have $g(x_1)<0$ and $g(x_0)=g(x_2)=0$.

Since $\bar{f}$ and Mf are equal in I, Lemma 3.1 and the continuity of L imply that there exists a smallest $x_1'\in[x_0,x_2]$ such that

$$g(x_1')=\inf_{x_0\leq y\leq x_2} g(y)<0.$$

Since $g(x_0) = 0$, there exists an $r > 0$ such that $[x_1' - r, x_1' + r] \subseteq [x_0, x_2]$. We have that $g(y) \geq g(x_1')$ for any $y \in [x_0, x_2]$ and the inequality is strict if $y < x_1'$. Hence by the mean value property for L,

$$Mf(x_1') \geq \fint_{x_1'-r}^{x_1'+r} \bar{f}(y)\,\mathrm{d}y = \fint_{x_1'-r}^{x_1'+r} g(y)\,\mathrm{d}y + L(x_1') > g(x_1') + L(x_1') = \bar{f}(x_1').$$

This is a contradiction to the fact that $x_1' \in \mathbb{R} \setminus V$. Therefore $\bar{f}$ is concave on I. □

The following result is a consequence of Lemma 3.6.

Lemma 3.7 *Let I be an unbounded connected component of $\mathbb{R} \setminus V$. Then,*

$$\lim_{|x|\to\infty;\, x\in I} \bar{f}(x) > 0.$$

Furthermore, if $I = \mathbb{R}$, then $\bar{f}$ is constant.

Proof Suppose for a contradiction that one of the conclusions of the lemma is false. Let $x_0 \in I$, meaning that $\bar{f}(x_0) > 0$. Then by symmetry, we may assume that $[x_0, \infty) \subseteq I$ and that there exists a $x_1 > x_0$ such that $\bar{f}(x_1) < \bar{f}(x_0)$. By Lemma 3.6, it follows that $\bar{f}(x_2) \leq L(x_2)$ for any $x_2 \geq x_1$ where $L\colon \mathbb{R} \to \mathbb{R}$ is the affine linear function defined by $L(x_0) = \bar{f}(x_0)$ and $L(x_1) = \bar{f}(x_1)$. Notice that L is strictly decreasing and hence $\bar{f}(x_2) < 0$ if x_2 is large enough. This is a contradiction to the nonnegativity of $\bar{f}$. □

We can now characterise the cases of equality in the intermediate inequality (3.3).

Proposition 3.8 *Equality holds in (3.3) if and only if $\bar{f}$ is constant or $\mathbb{R} \setminus V$ is a compact interval of positive length.*

Proof It suffices to consider the case that $\bar{f}$ is not constant since otherwise both sides of (3.3) are zero. Then $\mathbb{R} \setminus V$ is nonempty. By the second part of Lemma 3.7, we also have that V is nonempty.

By the proof of Lemma 3.4, equality holds in (3.3) if and only if (3.5) holds with equality whenever I is the closure of some connected component of V. Any such I has positive length by Lemma 3.5. By the strictness in Proposition 1.7((1)), this means that (3.3) can only hold with equality if all connected components of V are unbounded, i.e. if $\mathbb{R} \setminus V$ is a nonempty interval. This interval is closed by Lemma 3.3 and has positive length by Lemma 3.5.

Now let $I \neq \mathbb{R}$ be an unbounded connected component of V. Since the function $\bar{f}$ is of bounded variation, its limits at $\pm\infty$ exist and for any $x \in \mathbb{R}$,

$$Mf(x) \geq \lim_{r\to\infty} \fint_{x-r}^{x+r} \bar{f}(y)\,\mathrm{d}y = \lim_{y\to\infty} \frac{\bar{f}(y) + \bar{f}(-y)}{2}.$$

Furthermore, if the right/hand side is zero, then $\lim_{|x|\to\infty} Mf(x) = 0$. By Proposition 1.7((2)), it follows that (3.5) holds with equality if and only if $\lim_{|x|\to\infty} \bar{f}(x) = 0$. By

Lemma 3.7, this is the case precisely when $\mathbb{R} \setminus V$ has no unbounded components. We conclude that if $\bar{f}$ is not constant, then (3.3) holds with equality if and only if $\mathbb{R} \setminus V$ is a compact interval of positive length. This completes the proof. □

We can now characterise the cases of equality in (3.1). We first assume that equality holds and show that f is constant or the set $\{x \in \mathbb{R} | f(x) > 0\}$ is a bounded interval of positive length and for any $x \in \mathbb{R}$,

$$\liminf_{y \to x} f(y) \leq f(x) \leq \limsup_{y \to x} f(y). \tag{3.6}$$

If $\bar{f}$ is constant, then Mf is constant and equality in (3.1) implies that f is constant. Now consider the case that $\bar{f}$ is not constant. By Lemma 3.2 and 3.4, equality in (3.1) implies equality in (3.3) and (3.2). Hence by Proposition 3.8, it follows that $\mathbb{R} \setminus V = [a, b]$ for some real numbers $a < b$. By Lemma 3.6, the canonical representative $\bar{f}$ is concave on $[a, b]$ and hence continuous on (a, b) with

$$0 = \lim_{y \nearrow a} \bar{f}(y) \leq \bar{f}(a) \leq \lim_{y \searrow a} \bar{f}(y) \quad \text{and} \quad 0 = \lim_{y \searrow b} \bar{f}(y) \leq \bar{f}(b) \leq \lim_{y \nearrow b} \bar{f}(y) \tag{3.7}$$

Since f and $\bar{f}$ are equal almost everywhere and $\bar{f}$ is continuous in $\mathbb{R} \setminus \{a, b\}$, equality in (3.2) now implies that $f(x) = \bar{f}(x)$ for any $x \in \mathbb{R} \setminus \{a, b\}$. It follows that $\{x \in \mathbb{R} | f(x) > 0\}$ contains (a, b) and is contained in $[a, b]$, verifying that this set is a bounded interval of positive length. Furthermore, (3.6) holds if $x \notin \{a, b\}$. At $x = a$, using (3.7) together with the equality in (3.2) we obtain that

$$0 = \lim_{y \nearrow a} \bar{f}(y) \leq f(a) \leq \lim_{y \searrow a} \bar{f}(y)$$

and a similar statement holds at $x = b$. Since we already established that $f(y) = \bar{f}(y)$ for $y \in \mathbb{R} \setminus \{a, b\}$, we can replace $\bar{f}(y)$ by $f(y)$ in the above limits. Hence, (3.6) is true also for $x \in \{a, b\}$, and we conclude that f has the properties stated in the previous paragraph.

Conversely, assume that f is constant or the set $\{x \in \mathbb{R} | f(x) > 0\}$ is a bounded interval of positive length and (3.6) holds for any $x \in \mathbb{R}$. We need to show equality in (3.1). This is immediate if f is constant, so it remains to consider the case that there exist real numbers $a < b$ such that $\{x \in \mathbb{R} | f(x) > 0\}$ contains (a, b) and is contained in $[a, b]$. Using the Lebesgue differentiation theorem and Lemma 3.3 it follows that $\mathbb{R} \setminus V = [a, b]$. Hence equality holds in (3.3) by Proposition 3.8 and $\bar{f}$ is continuous in (a, b) by Lemma 3.6. It remains to show equality in (3.2). Since f and $\bar{f}$ agree almost everywhere and are of bounded variation, we have that

$$\liminf_{y \to x} f(y) = \liminf_{y \to x} \bar{f}(y) = \bar{f}(x)$$

for every $x \in (a, b)$ and the same holds with lim inf replaced by lim sup. Now $f(x) = \bar{f}(x)$ holds for any $x \in (a, b)$ by (3.6) and for any $x \in \mathbb{R} \setminus [a, b]$ by assumption on f, i.e. $f(x)$ and $\bar{f}(x)$ may only disagree if $x \in \{a, b\}$. However, Lemma 3.6 implies (3.7)

as before and (3.7) continues to hold if $\bar{f}$ is replaced by f, which is a consequence of (3.6). We can conclude equality in (3.2), which implies equality in (3.1) because we already showed equality in (3.3). This completes the proof of Theorem 1.2.

4 Discrete Setting

In this section, we first use an embedding argument to prove the conditional result Proposition 1.4 and to derive the discrete Theorem 1.3 from the continuous Theorem 1.1. Afterwards, we adapt the arguments in Sects. 2 and 3 to show the more general discrete Theorem 1.5.

4.1 Embedding

Let $f\colon \mathbb{Z} \to \mathbb{R}$ be a function of bounded variation and let $Mf\colon \mathbb{Z} \to \mathbb{R}$ be the discrete maximal function as defined in Sect. 1.1. We define an associated step function $f_c\colon \mathbb{R} \to \mathbb{R}$ by setting $f_c(x) = f(n)$ for any integer n and any $x \in [n-1/2, n+1/2)$. Let $Mf_c\colon \mathbb{R} \to \mathbb{R}$ be the continuous maximal function as defined in Sect. 1.

For any monotone map $\phi\colon \mathbb{Z} \to \mathbb{Z}$ there exists a monotone map $\psi\colon \mathbb{Z} \to \mathbb{R}$ such that $f \circ \phi = f_c \circ \psi$ and vice versa. Hence $\operatorname{var}_{\mathbb{Z}}(f) = \operatorname{var}(f_c)$. Our next claim is that $\operatorname{var}_{\mathbb{Z}}(Mf) \le \operatorname{var}(Mf_c)$. This is an immediate consequence of the following result.

Lemma 4.1 *We have that $Mf(n) = Mf_c(n)$ for any integer n.*

Proof For any nonnegative integer m, the step function f_c is constant on the intervals $[n-m-1/2, n-m+1/2)$ and $[n+m-1/2, n+m+1/2)$. Thus for any positive radius r with $|r-m| \le 1/2$ we have that

$$\frac{1}{2r}\int_{n-r}^{n+r} f_c(y)\,\mathrm{d}y = \frac{1}{2r}\int_{n-m}^{n+m} f_c(y)\,\mathrm{d}y + \frac{r-m}{2r}(f_c(n-m) + f_c(n+m)).$$

The right/hand side is of the form $A + B/r$ for some constants A and B independent of r, where $B = 0$ if $m = 0$. It follows that the map $r \mapsto \fint_{n-r}^{n+r} f_c(y)\,\mathrm{d}y$ is constant on $(0, 1/2]$ and monotone on $[m-1/2, m+1/2]$ for any positive integer m. Hence,

$$Mf_c(n) = \sup_{r\in\mathbb{Z}_{\ge 0}} \fint_{n-r-1/2}^{n+r+1/2} f_c(y)\,\mathrm{d}y = \sup_{r\in\mathbb{Z}_{\ge 0}} \sum_{m=n-r}^{n+r} f(m) = Mf(n).$$

This completes the proof. □

If f_c satisfies (1.1) for some constant C, then it follows from the above that

$$\operatorname{var}_{\mathbb{Z}}(Mf) \le \operatorname{var}(Mf_c) \le C \operatorname{var}(f_c) = C \operatorname{var}_{\mathbb{Z}}(f) \tag{4.1}$$

and hence f satisfies (1.3) with the same constant. This proves Proposition 1.4 and enables us to derive Theorem 1.3 from Theorem 1.1.

Proof of Theorem 1.3 By assumption, f is $\{0, 1\}$-valued and of bounded variation and so the same is true for f_c. Hence by (4.1) and Theorem 1.1, we see that f satisfies (1.3) with $C = 1$. Equality can only hold if equality holds in Theorem 1.1. For a nonconstant f, this implies that the set $\{x \in \mathbb{R} | f_c(x) = 1\}$ is a bounded interval of positive length and hence the set $\{n \in \mathbb{Z} | f(n) = 1\}$ is a bounded nonempty discrete interval. On the other hand, if f is of this form, then equality is attained because for any integer n with $f(n) = 1$ we have that

$$\mathrm{var}_{\mathbb{Z}}(Mf) \geq 2Mf(n) - \lim_{m\to\infty} Mf(m) + Mf(-m) = 2 - 0 = \mathrm{var}_{\mathbb{Z}}(f).$$

This completes the proof. □

4.2 Discrete Local Variation Bound

The following result is the discrete analogue of Proposition 1.7. We will use it to derive Theorem 1.5 similarly as Theorem 1.2 in the continuous setting, but without any of the technical difficulties related to compactness issues or exceptional sets of measure zero.

Proposition 4.2 *Let $f : \mathbb{Z} \to [0, \infty)$ be a bounded function and let $I \subseteq \mathbb{R}$ be an interval such that $f(n) = 0$ for any integer n in the interior of I. Then the following holds:*

(1) If $I = [a, b]$ for some integers $a < b$, then $\mathrm{var}_{I\cap\mathbb{Z}}(Mf) \leq Mf(a) + Mf(b)$. The inequality is strict unless f vanishes everywhere on $\mathbb{Z}$.
(2) If $I = (-\infty, a]$ or $I = [a, \infty)$ for some integer a, then Mf is monotone on $I \cap \mathbb{Z}$ and $\mathrm{var}_{I\cap\mathbb{Z}}(Mf) = Mf(a) - \inf_{n\in I\cap\mathbb{Z}} Mf(n)$.

The proof of this result goes along similar lines of the proof of Proposition 1.7, although some details differ. In particular we have to work around the fact that not all integer intervals have integer midpoints.

We first prove the unbounded case in Proposition 4.2((2)). By symmetry, it suffices to take $I = [a, \infty)$.

Lemma 4.3 *Let $f : \mathbb{Z} \to [0, \infty)$ be a bounded function and let $a \in \mathbb{Z}$ be such that $f(n) = 0$ for every integer $n > a$. Then Mf is nonincreasing on $[a, \infty) \cap \mathbb{Z}$ and hence*

$$\mathrm{var}_{[a,\infty)\cap\mathbb{Z}}(Mf) = Mf(a) - \inf_{n\in[a,\infty)\cap\mathbb{Z}} Mf(n).$$

Proof This is similar to the proof of Lemma 2.1. Let $n, m \in \mathbb{Z}$ be such that $a \leq n \leq m$. Then,

$$Mf(n) = \sup_{r\geq n-a} \sum_{k=n-r}^{n+r} f(k) \geq \sup_{r\geq n-a} \sum_{k=n-r}^{n+r+2(m-n)} f(k) = Mf(m).$$

This completes the proof. □

The rest of this section is devoted to the proof of Proposition 4.2((1)), i.e. the case that $I = [a, b]$ for some integers $a < b$. We start with a reduction using translation invariance. We also insert a midpoint in the case that $a + b$ is odd. For this, let f be as in Proposition 4.2. Set

$$S = \begin{cases} \mathbb{Z} & \text{if } a+b \text{ is even,} \\ \mathbb{Z} + \frac{1}{2} = \{\ldots, -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}, \ldots\} & \text{if } a+b \text{ is odd} \end{cases}$$

and write $S_0 = S \cup \{0\}$. We define a translated function $\tilde{f}\colon S \to [0, \infty)$ by

$$\tilde{f}(n) = f\Big(n + \frac{a+b}{2}\Big)$$

and we define its centred maximal function $M\tilde{f}\colon S_0 \to [0, \infty)$ by

$$M\tilde{f}(n) = \sup_{v \in S;\, v \le n} \sum_{m=v}^{2n-v} \tilde{f}(m).$$

Given a domain $T \in \{S, S_0\}$, a function $g : T \to [0, \infty)$ and an interval $I \subseteq \mathbb{R}$ we define the variation of g on the *discrete interval* $I \cap T$ by

$$\mathrm{var}_{I\cap T}(g) = \sup_{\phi\colon \mathbb{Z} \to I\cap T \text{ monotone}} \sum_{i\in\mathbb{Z}} |g(\phi(i)) - g(\phi(i+1))|.$$

If $S = \mathbb{Z}$, then these definitions agree with those in Sect. 1.1. Note that

$$\mathrm{var}_{[a,b]\cap\mathbb{Z}}(Mf) = \mathrm{var}_{[-(b-a)/2,(b-a)/2]\cap S}(M\tilde{f}) \le \mathrm{var}_{[-(b-a)/2,(b-a)/2]\cap S_0}(M\tilde{f})$$

and

$$M\tilde{f}\Big(-\frac{b-a}{2}\Big) = Mf(a), \qquad M\tilde{f}\Big(\frac{b-a}{2}\Big) = Mf(b).$$

From now on and for the rest of the proof of Proposition 4.2((1)), let $f\colon S \to [0, \infty)$ be a bounded nonzero function. By the above relations and by symmetry, it is enough to show the strict inequality

$$\mathrm{var}_{[0,a]\cap S_0}(Mf) < Mf(a) \tag{4.2}$$

for any positive $a \in S$ such that $f(n) = 0$ for all $n \in S$ with $-a < n < a$. This is analogous to (2.1).

Similarly as in the continuous setting, Mf restricted to $[0, a] \cap S_0$ is the pointwise maximum of the auxiliary maximal functions $M_0 f, M_1 f\colon [0, a] \cap S_0 \to [0, \infty)$

defined by

$$M_0 f(n) = \max_{v \in S;\ -a < v \le n} \sum_{m=v}^{2n-v} f(m), \qquad M_1 f(n) = \sup_{v \in S;\ v \le -a} \sum_{m=v}^{2n-v} f(m).$$

The following gradient bound for $M_1 f$ is analogous to the continuous Lemma 2.2. Since admissible radii in the above discrete setting are separated by a distance of 1, an additional term $1/2$ appears in this bound. Because of this, we also dispense with the additional lower bound on the radii in Lemma 2.2. Except for these differences, the proof is similar to the continuous case.

Lemma 4.4 *Let $n, m \in [0, \infty) \cap S_0$ be distinct. Then,*

$$\frac{M_1 f(n) - M_1 f(m)}{|n-m|} \le \frac{M_1 f(n)}{n + a + 1/2 + |n-m|} \le \frac{M_1 f(m)}{n + a + 1/2}.$$

Proof We have $M_1 f(n) < \infty$ since f is bounded. Hence for any $\epsilon > 0$ there exists a $v \in S$ with $v \le -a$ such that $(1-\epsilon) M_1 f(n) \le \sum_{k=v}^{2n-v} f(k)$. Let w be such that

$$m - w = n - v + |n-m|.$$

Then $w \in S$ because $v - w$ is an integer. Since $w \le v < 2n - v \le 2m - w$,

$$\begin{aligned}
(1-\epsilon) M_1 f(n) - M_1 f(m) &\le \sum_{k=v}^{2n-v} f(k) - \sum_{k=w}^{2m-w} f(k) \\
&\le \Big(\frac{1}{2(n-v)+1} - \frac{1}{2(m-w)+1} \Big) \sum_{k=v}^{2n-v} f(k) \\
&= \frac{2|n-m|}{2(m-w)+1} \sum_{k=v}^{2n-v} f(k) \\
&\le \frac{|n-m|}{n + a + 1/2 + |n-m|} M_1 f(n).
\end{aligned}$$

The first, third and fourth relations follow from definitions and the fact that $w \le -a$. In the second line, we use that f is nonnegative to reduce the range of summation of the second sum. Now the first inequality in the lemma follows by letting $\epsilon \to 0$. The second inequality follows after rearranging terms. □

Our next result is a local variation bound for $M_1 f$ analogous to the continuous Lemma 2.3. Here the proof is somewhat simplified due to a telescoping argument. Furthermore, due to the additional term $1/2$ in Lemma 4.4 above, we are able to show a slightly stronger inequality than in the continuous setting. This artefact already allows us to obtain a strict inequality, whereas in the continuous setting we have to work a little harder to get the strict inequality in Lemma 2.3.

Lemma 4.5 *Let $a \in S$ be nonnegative. Then,*

$$\operatorname{var}_{[0,a]\cap S_0}(M_1 f) \le \frac{2a}{2a+1} M_1 f(a).$$

Proof Let $n < m$ be elements of $[0, a] \cap S_0$. By the two inequalities in Lemma 4.4,

$$\begin{aligned} |M_1 f(n) - M_1 f(m)| &\le \frac{m-n}{m+a+1/2} \max(M_1 f(n), M_1 f(m)) \\ &\le \frac{(m-n)(2a+1/2)}{(n+a+1/2)(m+a+1/2)} M_1 f(a) \\ &= \Big(\frac{2a+1/2}{n+a+1/2} - \frac{2a+1/2}{m+a+1/2}\Big) M_1 f(a). \end{aligned}$$

Now let $0 = n_0 < n_1 < \ldots < n_k = a$ be an enumeration of $[0, a] \cap S_0$. We use the above estimate and evaluate the resulting telescoping sum to obtain that

$$\operatorname{var}_{[0,a]\cap S_0}(M_1 f) = \sum_{i=0}^{k-1} |M_1 f(n_i) - M_1 f(n_{i+1})| \le \Big(\frac{2a+1/2}{a+1/2} - 1\Big) M_1 f(a).$$

This completes the proof. □

Regarding the other auxiliary maximal function $M_0 f$, the following result similar to Lemmas 2.5 and 4.3 holds.

Lemma 4.6 *Let $a \in S$ be nonnegative and let $f(n) = 0$ for any $n \in S$ with $-a < n < a$. Then $M_0 f$ is nondecreasing on $[0, a] \cap S_0$.*

Proof Let $n, m \in S$ be such that $0 < n \le m \le a$. Then,

$$M_0 f(n) = \max_{v\in S;\ -a<v\le 2n-a} \sum_{k=v}^{2n-v} f(k) \le \max_{v\in S;\ -a<v\le 2n-a} \sum_{k=v+2(m-n)}^{2n-v} f(k) \le M_0 f(m).$$

Since $M_0 f(0) = 0$ and $M_0 f$ is nonnegative, this completes the proof. □

Having established the monotonicity of $M_0 f$ and a variation bound for $M_1 f$ similarly as in the continuous setting, the next step is to combine these results using the following analogue of Lemma 2.6. We omit the proof because it is the same.

Lemma 4.7 *Let $a \in S$ be nonnegative. Let $g, h\colon [0, a] \cap S_0 \to \mathbb{R}$ be functions such that $g(a) \le h(a)$ and let g be nondecreasing. Then,*

$$\operatorname{var}_{[0,a]\cap S_0}(\max(g, h)) \le \operatorname{var}_{[0,a]\cap S_0}(h).$$

We are now ready to prove (4.2). Let $a \in S$ be positive such that $f(n) = 0$ for any $n \in S$ with $-a < n < a$ and let $h\colon [0, a] \cap S_0 \to [0, \infty)$ be the function defined

by $h(n) = M_1 f(n)$ for $n < a$ and $h(a) = Mf(a)$. Then $M_0 f(a) \le h(a)$ and Mf restricted to $[0, a] \cap S_0$ is the pointwise maximum of $M_0 f$ and h. Hence we can apply Lemmas 4.6 and 4.7 and then Lemma 4.5 to obtain that

$$\mathrm{var}_{[0,a]\cap S_0}(Mf) \le \mathrm{var}_{[0,a]\cap S_0}(h) \le \mathrm{var}_{[0,a]\cap S_0}(M_1 f) + Mf(a) - M_1 f(a) < Mf(a).$$

This proves (4.2) and thus completes the proof of Proposition 4.2.

4.3 Discrete Global Variation Bound

We now prove the inequality in Theorem 1.5. Throughout this section and the next section, let $f : \mathbb{Z} \to [0, \infty)$ be a function of bounded variation such that for any $n \in \mathbb{Z}$ we have $f(n) = 0$ or $f(n) = Mf(n)$. For possibly infinite endpoints $a \le b$ we write

$$[a, b] \cap \mathbb{Z} = \{n \in \mathbb{Z} | a \le n \le b\}.$$

There exists a possibly unbounded discrete interval $\mathcal{I} \subseteq \mathbb{Z}$ with at least two elements and a nondecreasing sequence $(a_i)_{i\in\mathcal{I}}$ of points in $\mathbb{Z} \cup \{\pm\infty\}$ such that

$$\{n \in \mathbb{Z} | f(n) > 0\} = \bigcup_{i, i+1\in\mathcal{I};\, i \text{ odd}} [a_i, a_{i+1}] \cap \mathbb{Z} = \mathbb{Z} \setminus \bigcup_{i, i+1\in\mathcal{I};\, i \text{ even}} (a_i, a_{i+1})$$

and $a_i + 2 \le a_{i+1}$ for any even $i \in \mathcal{I}$ such that $i + 1 \in \mathcal{I}$. We may further assume that the points $\pm\infty$ each occur at most once in the sequence $(a_i)_{i\in\mathcal{I}}$.

Let $i \in \mathcal{I}$ be such that $i + 1 \in \mathcal{I}$. If i is even, then by Proposition 4.2,

$$\mathrm{var}_{[a_i, a_{i+1}]\cap\mathbb{Z}}(Mf) \le \mathrm{var}_{[a_i, a_{i+1}]\cap\mathbb{Z}}(f). \tag{4.3}$$

On the other hand, if i is odd, then by assumption it holds that $f(n) = Mf(n)$ for all $n \in [a_i, a_{i+1}] \cap \mathbb{Z}$ and thus (4.3) holds with equality. We can conclude that

$$\mathrm{var}_{\mathbb{Z}}(Mf) = \sum_{i, i+1\in\mathcal{I}} \mathrm{var}_{[a_i, a_{i+1}]\cap\mathbb{Z}}(Mf) \le \sum_{i, i+1\in\mathcal{I}} \mathrm{var}_{[a_i, a_{i+1}]\cap\mathbb{Z}}(f) = \mathrm{var}_{\mathbb{Z}}(f) \tag{4.4}$$

This proves the inequality in Theorem 1.5.

4.4 Cases of Equality

For the characterisation of the cases of equality in Theorem 1.5, we may assume that f is not constant since otherwise both sides of (4.4) are zero. By the last subsection, equality holds in (4.4) if and only if for every even $i \in \mathcal{I}$ with $i + 1 \in \mathcal{I}$ we have equality in (4.3). We need the following concavity result whose proof we omit because it is similar to the proof of Lemma 3.7. The conclusion of this result slightly differs from Lemma 3.7 because here we already assume f to be nonconstant.

Lemma 4.8 *Let $i \in \mathcal{I}$ be odd and such that $i+1 \in \mathcal{I}$ and $a_{i+1} = \infty$. Then $\lim_{n\to\infty} f(n) > 0$ and $a_i > -\infty$.*

Since f is not constant, it is not the zero function. Hence $\mathcal{I}$ is not of the form $\{i, i+1\}$ for any even i. By Lemma 4.8, it is also not of this form for any odd i. Hence $\mathcal{I}$ has at least three elements. If there exists an even $i \in \mathcal{I}$ with $i+1 \in \mathcal{I}$ and $a_i, a_{i+1} \in \mathbb{Z}$, then (4.3) is a strict inequality by Proposition 4.2 and hence (4.4) is strict. It remains to consider the case that no such i exists. After re/indexing and up to symmetry, this means that $\mathcal{I}$ is either $\{0, 1, 2, 3\}$ or $\{0, 1, 2\}$.

In the first case, f is finitely supported and hence, by Proposition 4.2((2)), equality holds in (4.3) for the even indices $i = 0$ and $i = 2$. Thus (4.4) holds with equality. In the second case, by Lemma 4.8,

$$Mf(n) \geq \lim_{r\to\infty} \sum_{m=n-r}^{n+r} f(m) = \lim_{m\to\infty} \frac{f(m)}{2} > 0.$$

for any integer n. By Proposition 4.2((2)), this means that (4.3) is strict for $i = 0$ and hence (4.4) is strict. We conclude that equality holds in (4.4) if and only if f is constant or $\{n \in \mathbb{Z} | f(n) > 0\} = [a, b] \cap \mathbb{Z}$ for some integers $a \leq b$. This completes the proof of Theorem 1.5.

Acknowledgements The first author was partially supported by the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy, Project 390685813 – EXC-2047/1, during a stay at the Hausdorff Research Institute for Mathematics. The second author has been partially supported by the Vilho, Yrjö and Kalle Väisälä Foundation of the Finnish Academy of Science and Letters, and the Magnus Ehrnrooth foundation. We would like to thank the supervisors of the first author, Diogo Oliveira e Silva and Jonathan Bennett, and the supervisor of the second author, Juha Kinnunen, for all their support.

Funding Open Access funding provided by Aalto University.

References

1. Aldaz, J.M., Pérez Lázaro, F.J.: Regularity of the Hardy-Littlewood maximal operator on block decreasing functions. Studia Math. **194**(3), 253–277 (2009)
2. Aldaz, J.M., Pérez Lázaro, J.: Functions of bounded variation, the derivative of the one dimensional maximal function, and applications to inequalities. Trans. Amer. Math. Soc. **359**(5), 2443–2461 (2007)
3. Beltran, D., González-Riquelme, C., Madrid, J., Weigt, J.: Continuity of the gradient of the fractional maximal operator on $W^{1,1}(\mathbb{R}^d)$. (2021). arXiv:2102.10206 [math.CA]
4. Bober, J., Carneiro, E., Hughes, K., Pierce, L.B.: On a discrete version of Tanaka's theorem for maximal functions. Proc. Amer. Math. Soc. **140**(5), 1669–1680 (2012)
5. Hajłasz, P., Onninen, J.: On boundedness of maximal functions in Sobolev spaces. Ann. Acad. Sci. Fenn. Math. **29**, 167–176 (2004)

6. Kinnunen, J.: The Hardy-Littlewood maximal function of a Sobolev function. Israel J. Math. **100**, 117–124 (1997)
7. Kinnunen, J., Saksman, E.: Regularity of the fractional maximal function. Bull. London Math. Soc. **35**(4), 529–535 (2003)
8. Kurka, O.: On the variation of the Hardy-Littlewood maximal function. Ann. Acad. Sci. Fenn. Math. **40**, 109–133 (2015)
9. Leoni, G.: A first course in Sobolev spaces. Vol. 105. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, (2009), pp. xvi+607
10. Luiro, H.: The variation of the maximal function of a radial function. Ark. Mat. **56**(1), 147–161 (2018)
11. Madrid, J.: Sharp inequalities for the variation of the discrete maximal function. Bull. Aust. Math. Soc. **95**(1), 94–107 (2017)
12. Ramos, J.P.G.: Sharp total variation results for maximal functions. Ann. Acad. Sci. Fenn. Math. **44**(1), 41–64 (2019)
13. Tanaka, H.: A remark on the derivative of the one-dimensional Hardy-Littlewood maximal function. Bull. Austral. Math. Soc. **65**(2), 253–258 (2002)
14. Temur, F.: On regularity of the discrete Hardy–Littlewood maximal function. (2013). arXiv:1303.3993 [math.CA]
15. Weigt, J.: Variation of the uncentered maximal characteristic function. (2020). arXiv:2004.10485