Concentration on the Boolean hypercube via pathwise stochastic analysis

Ronen Eldan¹ · Renan Gross¹

Abstract We develop a new technique for proving concentration inequalities which relate the variance and influences of Boolean functions. Using this technique, we

1. Settle a conjecture of Talagrand (Combinatorica 17(2):275–285, 1997), proving that

\[
\int_{\{-1,1\}^n} \sqrt{h_f(x)} d\mu(x) \geq C \cdot \text{Var}(f) \cdot \left( \log \left( \frac{1}{\sum \text{Inf}_i^2(f)} \right) \right)^{1/2},
\]

where \( h_f(x) \) is the number of edges at \( x \) along which \( f \) changes its value, \( \mu(x) \) is the uniform measure on \( \{-1,1\}^n \), and \( \text{Inf}_i(f) \) is the influence of the \( i \)-th coordinate.

2. Strengthen several classical inequalities concerning the influences of a Boolean function, showing that near-maximizers must have large vertex boundaries. An inequality due to Talagrand states that for a Boolean function \( f \), \( \text{Var}(f) \leq C \sum_{i=1}^n \frac{\text{Inf}_i(f)}{1+\log(1/\text{Inf}_i(f))} \). We give a lower bound for the size of the vertex boundary of functions saturating this inequality. As a corollary, we show that for sets that satisfy the edge-isoperimetric

---

¹ Weizmann Institute of Science, Rehovot, Israel

Renan Gross
renan.gross@weizmann.ac.il

Ronen Eldan
ronen.eldan@weizmann.ac.il
inequality or the Kahn–Kalai–Linial inequality up to a constant, a constant proportion of the mass is in the inner vertex boundary.

3. Improve a quantitative relation between influences and noise stability given by Keller and Kindler.

Our proofs rely on techniques based on stochastic calculus, and bypass the use of hypercontractivity common to previous proofs.

Keywords Boolean analysis · Concentration · Isoperimetric inequality · Pathwise analysis

Mathematics Subject Classification 68R05 · 60G44 · 60G55

Contents

1 Introduction ....................................... 937
  1.1 Background ..................................... 937
  1.2 Our results ...................................... 939
    1.2.1 Talagrand’s conjecture ......................... 939
    Direction 1: Theorem 1.4 ⇒ KKL .................... 940
    Direction 2: Theorem 1.4 ⇒ Stability of the Isoperimetric inequality 941
  1.2.2 Talagrand’s influence theorem and its stability 942
    The isoperimetric inequality and vertex boundary ............................................. 942
    The KKL inequality and vertex boundary ......................................................... 943
    An improved quantitative relation between noise sensitivity and influences .......... 944
  1.3 Proof outline ..................................... 945

2 Preliminaries ....................................... 947
  2.1 Boolean functions .................................. 948
  2.2 Stochastic processes and quadratic variation .................................................. 950

3 The main tool: a jump process .............................. 952
  3.1 Construction and basic properties .................. 952
  3.2 The influence process ................................ 955

4 Proof of the strengthening of Talagrand’s conjecture .................. 957
  Case 1: small jump ..................................... 960
  Case 2: large jump ..................................... 962
  4.1 Postponed proofs .................................. 964

5 Talagrand’s influence inequality and its stability ..................... 968
  5.1 Proof of Theorem 1.2 ................................ 971
  5.2 Proof of Theorem 1.5 ................................ 971
  5.3 Postponed proofs .................................. 975

6 Proof of Theorem 1.9 .................................. 979
Appendix A: p-Biased analysis .................................. 985
Appendix B: Postponed proofs .................................. 986

References .......................................... 992
1 Introduction

1.1 Background

The influence of a Boolean function \( f : \{-1, 1\}^n \rightarrow \{-1, 1\} \) in direction \( i = 1, \ldots, n \) is defined as

\[
\Inf_i (f) = \mu \left( \left\{ y \in \{-1, 1\}^n \mid f(y) \neq f(y^{\oplus i}) \right\} \right),
\]

where \( y^{\oplus i} \) is the same as \( y \) but with the \( i \)-th bit flipped, and \( \mu \) is the uniform measure on the discrete hypercube \( \{-1, 1\}^n \). The expectation and variance of a function are given by

\[
\mathbb{E} f = \int_{\{-1,1\}^n} f \, d\mu \quad \text{and} \quad \text{Var}(f) = \mathbb{E}(f - \mathbb{E}f)^2.
\]

The Poincaré inequality gives an immediate relation between the aforementioned quantities, namely,

\[
\text{Var}(f) \leq \sum_{i=1}^n \Inf_i(f).
\]

The total influence \( \sum_i \Inf_i(f) \) on the right hand side is proportional to the number of edges of the hypercube which separate \( f(x) = 1 \) and \( f(x) = -1 \). It can therefore be seen as a type of surface-area of \( f \).

The inequality (1) in fact holds for functions on any domain (for a suitable definition of influence), and it is natural to ask whether it can be improved when Boolean functions are considered. A corollary of the breakthrough paper by Kahn, Kalai and Linial (KKL) [23] shows that this inequality can be strengthened logarithmically: there exists a universal constant \( C > 0 \) such that

\[
\text{Var}(f) \leq C \frac{\sum_i \Inf_i(f)}{\log(1/\max_i(\Inf_i(f)))}.
\]

(The above formulation does not appear explicitly in [23], but follows easily from their methods).

The KKL inequality is tight for the Tribes function (see [28, Section 4.2] for a definition), but is off by a factor of \( \sqrt{n}/\log n \) for the Majority function, whose influences are all of order \( 1/\sqrt{n} \), suggesting that the total influence \( \sum \Inf_i(f) \) may not be the right notion of surface-area for all Boolean functions. In [31,
Theorem 1.1], Talagrand showed that

$$\var(f) \leq \frac{1}{\sqrt{2}} \mathbb{E} \sqrt{h_f},$$

(3)

where

$$h_f(y) := \# \left\{ i \in [n] \mid f(y) \neq f(y \oplus i) \right\}$$

(4)

is the sensitivity of $f$ at point $y$. The value $\mathbb{E} \sqrt{h_f}$ can be seen as another type of surface-area of the function $f$. A sharp tightening of this inequality was given by Bobkov [5]; his inequality gives an elementary proof of the isoperimetric inequality on Gaussian space. Inequality (3) is tight for linear-threshold functions (functions of the form $x \mapsto 1 \sum \alpha_i x_i > \beta (x)$) such as Majority, but not for Tribes. Thus, neither inequality implies the other. This raises the following question:

**Question 1.1** What is the right notion of boundary for Boolean functions? Is there an inequality from which both (2) and (3) can be derived?

As a step in this direction, Talagrand conjectured in [34] that (3) can be strengthened, and that there exists a constant $\beta > 0$ such that

$$\mathbb{E} \sqrt{h_f} \geq \beta \cdot \var(f) \cdot \left( \log \left( 2 + \frac{e}{\sum \inf_i (f)} \right) \right)^{1/2},$$

(5)

Talagrand showed that there exists an $\alpha \leq 1/2$ and a constant $\beta > 0$ such that

$$\int_{\{-1,1\}^n} \sqrt{h_f(x)} d\mu \geq \beta \cdot \var(f) \left( \log \frac{e}{\var(f)} \right)^{1/2-\alpha} \cdot \left( \log \left( 2 + \frac{e}{\sum \inf_i (f)} \right) \right)^{\alpha},$$

but his proof did not yield the conjectured $\alpha = 1/2$.

Another notion of surface-area is the vertex boundary $\partial f$ of $f$, defined as

$$\partial f := \left\{ y \in \{-1,1\}^n \mid \exists i \text{ s.t } f(y) \neq f(y \oplus i) \right\}.$$  

It is the disjoint union of the inner vertex boundary,

$$\partial^+ f := \left\{ y \in \{-1,1\}^n \mid \exists i \text{ s.t } f(y) = 1, f(y \oplus i) = -1 \right\},$$

(6)

and the outer vertex boundary,

$$\partial^- f := \left\{ y \in \{-1,1\}^n \mid \exists i \text{ s.t } f(y) = -1, f(y \oplus i) = 1 \right\}.$$  

(7)
The Cauchy–Schwartz inequality implies that $\mathbb{E} \sqrt{h_f} \leq \sqrt{\mathbb{E}[h_f] \mu(\partial f)} = \sqrt{\sum_i \text{Inf}_i(f) \mu(\partial f)}$, so the above conjecture strengthens the KKL result in the regime $\text{Var}(f) = \Omega(1)$ (see below Theorem 1.4 for a calculation).

The inequality (2) was further generalized in another direction by Talagrand [32], who proved the following:

**Theorem 1.2** There exists an absolute constant $C_T > 0$ such that for every $f : \{-1, 1\}^n \to \{-1, 1\}$,

$$\text{Var}(f) \leq C_T \sum_{i=1}^n \frac{\text{Inf}_i(f)}{1 + \log(1/\text{Inf}_i(f))}. \tag{8}$$

It is known that this inequality is sharp in the sense that for any sequence of influences, there exist examples which saturate it [24].

Inequalities such as (2), (5) and (8), in conjunction with concentration of influence [15] and sharp threshold properties [16], have been widely utilized across many subfields of mathematics and computer science, including learning theory [29], metric embeddings [25], first passage percolation [3], classical and quantum communication complexity [17, 30], and hardness of approximation [9]; and also in social network dynamics [27] and statistical physics [1]. For a general survey, see [26].

Talagrand’s original proof of Theorem 1.2, as well as later proofs (see e.g. [7]), all rely on the hypercontractive principle (see e.g. [28, Chapters 9,10] for more about this fundamental technique).

### 1.2 Our results

In this paper, we develop a new approach in order to prove the aforementioned inequalities. Our proofs are based on pathwise analysis, which bypasses the use of hypercontractivity, and in fact uses classical Boolean Fourier-analysis only sparingly.

#### 1.2.1 Talagrand’s conjecture

Our first result is that Talagrand’s conjecture holds true:

**Theorem 1.3** There exists an absolute constant $C > 0$ such that for all $f : \{-1, 1\}^n \to \{-1, 1\}$,

$$\mathbb{E} \sqrt{h_f} \geq C \cdot \text{Var}(f) \cdot \sqrt{\log \left(2 + \frac{e}{\sum_i \text{Inf}_i(f)^2}\right)}.$$
In fact, we prove a stronger theorem, of which Theorem 1.3 is an immediate corollary:

**Theorem 1.4** There exists an absolute constant $C > 0$ such that the following holds. For all $f : \{-1, 1\}^n \to \{-1, 1\}$, there exists a function $g_f : \{-1, 1\}^n \to [0, 1]$ with $\mathbb{E}g_f^2 \leq 2\text{Var}(f)$ so that for all $1/2 \leq p < 1$,

$$
\mathbb{E}\left[ h_f^pg_f \right] \geq C\text{Var}(f) \cdot \left( \log \left( 2 + \frac{e}{\sum_i \text{Inf}_i(f)^2} \right) \right)^p .
$$

(9)

The above inequality with $p = 1/2$ implies both the KKL inequality in full generality and a new lower bound on total influences in the spirit of the isoperimetric inequality. By the Cauchy–Schwartz inequality,

$$
\mathbb{E}\left[ \sqrt{h_f g_f} \right] \leq \sqrt{\mathbb{E}h_f} \sqrt{\mathbb{E}g_f^2} \leq \sqrt{\text{Inf}(f)} \sqrt{2\text{Var}(f)},
$$

and plugging this into (9), we get

$$
\text{Var}(f) \leq C \cdot \frac{\sum_i \text{Inf}_i(f)}{\log \left( 2 + \frac{e}{\sum_i \text{Inf}_i(f)^2} \right)}
$$

(10)

From this equation we can proceed in two directions.

**Direction 1: Theorem 1.4 $\implies$ KKL**

Denoting $\delta = \max_i \text{Inf}_i(f)$, the above display yields

$$
\text{Var}(f) \leq C \cdot \frac{\sum_i \text{Inf}_i(f)}{\log \left( 2 + \frac{1}{\delta \sum_i \text{Inf}_i(f)} \right)} .
$$

Consider now two cases: if $\delta \sum_i \text{Inf}_i(f) \leq \delta^{1/2}$, we immediately get from the above display that $\text{Var}(f) \leq C \frac{\sum_i \text{Inf}_i(f)}{\log(2 + \frac{1}{\delta^{1/2}})} \leq 2C \frac{\sum_i \text{Inf}_i(f)}{\log(1/\delta)}$. And if $\delta \sum_i \text{Inf}_i(f) \geq \delta^{1/2}$, we have

$$
\sum_i \text{Inf}_i(f) \geq \frac{1}{\delta^{1/2}} \geq \text{Var}(f) \frac{1}{\delta^{1/2}} \geq \frac{1}{2} \text{Var}(f) \log \left( \frac{1}{\delta} \right),
$$

(for $\delta < 1$, otherwise there is nothing to prove), again yielding $\text{Var}(f) \leq 2 \frac{\sum_i \text{Inf}_i(f)}{\log(1/\delta)}$. Thus, Theorem 1.4 implies the KKL inequality.
Hierarchically, the relation between the Poincaré inequality, KKL, Talagrand’s theorem and Theorem 1.4 may be summarized as in Fig. 1.

**Direction 2: Theorem 1.4 \implies Stability of the Isoperimetric inequality**

Assume that $\mathbb{E} f \leq 0$ and let $A = \{x \in \{-1, 1\}^n \mid f(x) = 1\}$ be the support of $f$, so that $\mu(A) \leq 1/2$. The edge-isoperimetric inequality [19, Section 3] states that

$$
\sum_{i=1}^{n} \text{Inf}_i(f) \geq 2\mu(A) \log_2 \frac{1}{\mu(A)},
$$

with equality if and only if $A$ is a subcube. Suppose that $f$ saturates the isoperimetric inequality up to a constant, i.e.

$$
\sum_{i=1}^{n} \text{Inf}_i(f) \leq C\mu(A) \log_2 \frac{1}{\mu(A)}
$$

for some constant $C$. Since $\text{Var}(f) = 4\mu(A)(1 - \mu(A)) \geq 2\mu(A)$, this gives

$$
\sum_{i=1}^{n} \text{Inf}_i(f) \leq C\text{Var}(f) \log_2 \frac{2}{\text{Var}(f)}.
$$

Suppose also that $f$ is monotone; then by standard Fourier analysis, the first level Fourier coefficients are equal to the influences and we have $\text{Var}(f) \geq$
\[ \sum_i \text{Inf}_i(f)^2 \] (see Sect. 2.1). From Eqs. (10) and (12) we get that

\[ \text{Var}(f) \leq C \frac{\sum_i \text{Inf}_i(f)}{\log \left( \frac{2}{\sum_i \text{Inf}_i(f)^2} \right)} \leq C \frac{\sum_i \text{Inf}_i(f)}{\log \left( \frac{2}{\text{Var}(f)} \right)} \leq C' \text{Var}(f). \]

In particular, the two denominators are within a constant factor of each other:

\[ \log \left( \frac{2}{\sum_i \text{Inf}_i(f)^2} \right) = \Theta \left( \log \left( \frac{2}{\text{Var}(f)} \right) \right), \]

implying that the sum of Fourier weights on the first level is proportional to a power of the variance.

1.2.2 Talagrand’s influence theorem and its stability

Next, we reprove Theorem 1.2 using stochastic techniques, and provide a strengthening which can be thought of as a stability version of this bound in terms of the vertex boundary of \( f \): if near-equality is attained in Eq. (8), then both the inner and outer vertex boundaries of \( f \) are large. The theorem reads,

**Theorem 1.5** Let

\[ T(f) := \sum_{i=1}^{n} \frac{\text{Inf}_i(f)}{1 + \log \left( \frac{1}{\text{Inf}_i(f)} \right)} \quad (13) \]

and denote \( r_{\text{Tal}} = \frac{\text{Var}(f)}{T(f)} \). There exists an absolute constant \( C_B > 0 \) such that

\[ \mu(\vartheta^\pm f) \geq \frac{r_{\text{Tal}}}{C_B \log \frac{C_B}{r_{\text{Tal}}}} \text{Var}(f). \]

**Remark 1.6** The proof of Theorem 1.5 shows that \( C_B > C_T \geq r_{\text{Tal}} \), so the right-hand side of the inequality is always positive.

As we now show, Theorem 1.5 can be readily applied to two related functional inequalities—the isoperimetric inequality and the KKL inequality—showing that when either of the inequalities are tight up to a constant, the function must have a large vertex boundary.

**The isoperimetric inequality and vertex boundary**

It is natural to ask about the robustness of the isoperimetric inequality: is it true that if near-equality is attained in (11), then \( A \) is close to a subcube in
some sense? This question was answered in [14] for sets $A$ which are $(1 + \varepsilon)$-close to satisfying the inequality. Conjectures concerning sets for which the inequality is tight only up to a constant multiplicative factor can be found in [21]. We make a step in this direction by giving the first bound which is meaningful when the function is $O(1)$-close to satisfying the inequality (11), showing that in that case, a constant proportion of the set $A$ is in its inner vertex boundary (whereas for the extremizers, the vertex boundary is the entire set $A$).

**Corollary 1.7** Let $f$ have support $A$, with $\mu(A) \leq 1/2$. Let $r_{Iso} = \frac{2\mu(A) \log_2 \frac{1}{\mu(A)}}{\sum_{i=1}^{n} \inf_i(f)}$. Then there exists a constant $c_{Iso} \geq \frac{r_{Iso}}{2C_B \log(\frac{2C_B}{r_{Iso}})}$ depending only on $r_{Iso}$ such that

$$\mu(\partial A) \geq c_{Iso} \mu(A).$$

**Proof** As in Theorem 1.5, denote $r_{Tal} = \frac{\text{Var}(f)}{T(f)}$. Observe that for every index $i$, $\inf_i(f) \leq 2\mu(A)$. Since $\mu(A) \leq 1/2$, we have

$$\text{Var}(f) = 4\mu(A)(1 - \mu(A)) \geq 2\mu(A).$$

This gives a bound on $r_{Tal}$:

$$r_{Tal} = \frac{\text{Var}(f)}{\sum_{i=1}^{n} \frac{\inf_i(f)}{1 + \log(1/\inf_i(f))}} \geq \frac{\text{Var}(f)}{\sum_{i=1}^{n} \frac{\inf_i(f)}{1 + \log(1/2\mu(A))}} = \frac{r_{Iso} \text{Var}(f) \left(1 + \log\left(\frac{1}{2\mu(A)}\right)\right)}{2\log_2 2\mu(A) \log_2 \frac{1}{\mu(A)}} \geq \frac{r_{Iso} \log 2 \cdot \text{Var}(f)}{2\mu(A)} \geq \frac{r_{Iso}}{2}.$$

Thus, by Theorem 1.5, there exists a constant $c_{Iso} \geq \frac{r_{Iso}}{2C_B \log(\frac{2C_B}{r_{Iso}})}$ such that

$$\mu(\partial^\pm f) \geq \frac{c_{Iso}}{2} \text{Var}(f) \geq c_{Iso} \mu(A).$$

\[\square\]

**The KKL inequality and vertex boundary**

In its original formulation, the KKL theorem [23, Theorem 3.1], which follows immediately from (2), states that a Boolean function must have a variable with a relatively large influence: there exists an absolute constant $C > 0$ such that
for every $f : \{-1, 1\}^n \rightarrow \{-1, 1\}$, there exists an index $i \in [n]$ with

$$\text{Inf}_i (f) \geq C \cdot \text{Var} (f) \frac{\log n}{n}.$$ 

Our second corollary states that if all influences are of the order $\text{Var} (f) \frac{\log n}{n}$, then the function must have a large (inner and outer) vertex boundary.

**Corollary 1.8** For every constant $C > 0$, there exists a constant $c_{\text{KKL}} > 0$ and integer $N > 0$ so that the following holds for all $n > N$. If $\text{Inf}_i (f) \leq C \cdot \text{Var} (f) \frac{\log n}{n}$ for all $i$, then

$$\mu (\partial f) \geq c_{\text{KKL}} \text{Var} (f).$$

**Proof** In this case, we have

$$r_{\text{Tal}} = \frac{\text{Var} (f)}{\sum_i \text{Inf}_i (f) \frac{\log (1/\text{Inf}_i (f))}{1+\log (1/\text{Inf}_i (f))}} \geq \frac{\text{Var} (f)}{\sum_i \text{Inf}_i (f) \frac{\log n}{1-\log (C \cdot \text{Var} (f) \frac{\log n}{n})}} \geq \frac{\text{Var} (f) \left( 1 + \log \left( \frac{n}{C \cdot \text{Var} (f) \log n} \right) \right)}{C \cdot \text{Var} (f) \log n} \geq \frac{\log n - \log (C \log n)}{C \log n} \geq \frac{1}{2C}.$$ 

Thus, by Theorem 1.5, there exists a constant $c_{\text{KKL}}$ which depends only on $C$ such that

$$\mu (\partial f) \geq c_{\text{KKL}} \text{Var} (f).$$

An improved quantitative relation between noise sensitivity and influences

The proof of Theorem 1.4 uses an improved version of a bound due to Keller and Kindler [22], which may be of independent interest. Let $S_\varepsilon (f)$ be the noise stability of $f$, i.e.

$$S_\varepsilon (f) := \text{Cov}_{x \sim \mu, y \sim N_\varepsilon (x)} [f (x), f (y)],$$

where $N_\varepsilon (x)$ is a random vector whose $i$-th coordinate is equal to $x_i$ with probability $1 - \varepsilon$ and to a uniformly random bit with probability $\varepsilon$. 

---

The text is from a scientific publication, focusing on theorems and proofs in the field of theoretical computer science, specifically concerning the relationship between noise sensitivity and influences in Boolean functions.
Theorem 1.9 There exists universal constants $C, c > 0$ such that

$$S_\varepsilon (f) \leq C \cdot \operatorname{Var}(f) \left( \sum_{i=1}^{n} \inf_i (f)^2 \right)^{c \varepsilon}.$$  \hspace{1cm} (15)

The bound proved in [22] is the same but with the term $\operatorname{Var}(f)$ is replaced by a constant; thus our result becomes stronger when $\operatorname{Var}(f) = o(1)$. The relation between influences and noise sensitivity was first established in [2], where a qualitative bound of the same nature is proven.

1.3 Proof outline

The core of our proofs is the study of a continuous-time martingale $(B_t)_{t \in [0,1]} = (B_t^{(1)}, \ldots, B_t^{(n)}) \in \mathbb{R}^n$ which satisfies $\|B_t^{(i)}\| = t$ and $B_1 \sim \text{Unif}\{(-1,1)^n\}$ (Proposition 3.1).

Since $B_1$ is uniform on the hypercube, the expected value and variance of $f$ can be obtained by $\mathbb{E} f = \mathbb{E} f (B_1)$ and $\operatorname{Var} f = \operatorname{Var} f (B_1)$, where the expectations in the right hand sides are over the randomness of the process $B_t$. Similarly, the influence of the $i$-th bit is given by $\inf_i (f) = \mathbb{E} \left[ (\partial_i f (B_1))^2 \right]$, where $\partial_i f$ is the discrete derivative of $f$ in direction $i$, and $\mathbb{E} h_f^p$ is given by $\mathbb{E} \|\nabla f (B_1)\|_2^2$.

The strength of the stochastic process approach stems from the fact that the behavior of $f (B_1)$ can be understood by extending the domain of $f$ and $\nabla f$ to the continuous hypercube $[-1,1]^n$ and analysing the processes $B_t$, $f (B_t)$ and $\nabla f (B_t)$ for times smaller than 1. Indeed, using its Fourier decomposition, there is a natural way to extend the domain of a Boolean function $f$ to $[-1,1]^n$ so that the processes $f (B_t)$ and $\partial_i f (B_t)$ become martingales. The variance of $f$ can then be expressed as

$$\operatorname{Var}(f) = 2 \mathbb{E} \sum_{i=1}^{n} \int_0^1 t (\partial_i f (B_t))^2 \, dt$$

(Lemma 3.3 and Corollary 3.4). Bounding the variance is then a matter of bounding this integral, and for this we can utilize tools from real analysis and stochastic processes. Specifically, two well-known inequalities—called the Level-1 and Level-2 inequalities—give us bounds on the speed with which both the individual processes $\partial_i f (B_t)^2$ and their collective sum $\sum_i (\partial_i f (B_t))^2$ are moving in terms of their current value. In the Gaussian setting, somewhat similar ideas of using level inequalities appear in [11].
This points to a significant conceptual difference between existing techniques that use the hypercontractivity of the heat operator and our technique: whereas the former proofs start from a definite value of the function $f$ and analyze the way that it changes by applying the heat semigroup, which corresponds to averaging out the values of $f$, our analysis goes in the other direction. That is, we may think of the process $B_t$ as a way to sample from $\{-1, 1\}^n$ via a continuous filtration, where we add “infinitesimal bits of randomness” as time progresses. The analysis starts from $f(B_0) = \mathbb{E} f$ and considers the way that the martingales evolve as we refine our filtration, or in other words, add more randomness. On a first glance this difference may seem to be only pedagogical, but the strength of our approach is that the pathwise analysis equips us with new tools, such as using stopping times and the optional stopping theorem, and conditioning on the past.

Towards proving Talagrand’s conjecture, we bound $\sum_i (\partial_i f(B_t))^2$ by a time-dependent power of the sum of squares of influences $\sum_i \text{Inf}_i(f)^2$. We do this using the Level-2 inequality (Lemma 2.2), which amounts to the fact that if the first derivative of a process with respect to time is small, then so is its second derivative with respect to time, and thus the first derivative must remain small for a long time. Ideologically, when the sum $\sum_i \text{Inf}_i(f)^2$ is small, this roughly implies that the process $\|\nabla f(B_t)\|_2^2$ makes most of its movement very close to time $t = 1$; this is in fact the essence of Theorem 1.9. This can then be used to show that most of the quadratic variation of $f(B_t)$ comes from paths in which there is a time $t$ such that $\|\nabla f(B_t)\|_2^2$ is larger than $\alpha \left( \log \left( \frac{1}{\sum \text{Inf}_i^2(f)} \right) \right)^{1/2}$ (Lemma 4.2). However, the quadratic variation is itself large with a probability that is a function of the variance (Lemma 2.5). Since $\|\nabla f(B_t)\|_2^{2p}$ is a submartingale, if there was ever a time when $\|\nabla f(B_t)\|_2^{2p}$ is large, then in expectation it continues to be large. Thus $\mathbb{E} \|\nabla f(B_1)\|_2^{2p}$ is larger than $\alpha \text{Var}(f) \left( \log \left( \frac{1}{\sum \text{Inf}_i^2(f)} \right) \right)^{p}$, giving the original Talagrand’s conjecture (Theorem 1.3) when $p = 1/2$. With additional care, it can be shown that $f(B_t)$ itself is large at some time before the gradient was large, which gives the strengthened result (Theorem 1.4).

This is a good place to point out an analogy between our technique and the one demonstrated by Barthe and Maurey [4], who give a stochastic proof of Bobkov’s extension of inequality (3). They use a stochastic argument in order to derive a one-dimensional inequality which implies Bobkov’s inequality via tensorization, and one of the central components in their proof is to establish that a certain process which is analogous to $\|\nabla f(B_t)\|_2$ is a submartingale (this is based on ideas introduced in a paper by Capitaine et al. [8]).
Similarly, for proving Theorem 1.2, we use the Level-1 inequality (Lemma 2.1) to bound each individual \((\partial_i f(B_t))^2\) by a time-dependent power of the influence \(\text{Inf}_i(f)\) (Lemma 5.1). When the influences are small, this roughly implies that the martingale \(f(B_t)\) makes most of its movement very close to time \(t = 1\). Theorem 1.2 then follows by plugging this bound into the integral (Proposition 5.2).

The proof of Theorem 1.5 is more involved, and utilizes the fact that \(f(B_t)\) is both a jump process and a martingale: for such processes, the variance of \(f(B_1)\) is then given by the sum of squares of jumps of \(f(B_t)\) up to time 1:

\[
\text{Var}(f) = \mathbb{E} \sum_{s \in \text{Jump}(B_t)} (\Delta f(B_s))^2 = 2\mathbb{E} \sum_{i=1}^{n} \int_{0}^{1} t (\partial_i f(B_t))^2 dt.
\]

The technical core of the proof (Proposition 5.3 and Lemma 5.4) shows that if \(T(f)\) and \(\text{Var}(f)\) differ only by a multiplicative constant, then with non-negligible probability the process \(f(B_t)\) must make a relatively large jump somewhere along the way. Roughly speaking, this is because if the process \(B_t^{(i)}\) jumps at time \(t\), then the function \(f(B_t)\) also jumps, changing by a value of \(2t\partial_i f(B_t)\). If all the jumps are small, then the expression \(\sum_{s \in \text{Jump}(B_t)} (\Delta f(B_s))^2\) in the left hand side of the above display (which cares only about jumps) must be substantially smaller than the integral in the right hand side (which cares only about the size of the derivatives).

Now, when the process \(f(B_t)\) makes a large jump, it necessarily means that the magnitude of one of the partial derivatives \(\partial_i f\) is large. Since the process \(\partial_i f(B_t)\) is also a martingale, if it is large at some point in time, then it continues to be large with relatively high probability. But at time \(t = 1\), since \(B_1\) takes values in \((-1, 1)^n\), the only possibilities for the values of \(\partial_i f(B_1)\) are \(-1, 0\) and 1. Thus, it is likely that \(|\partial_i f(B_1)| = 1\). This exactly corresponds to the point \(B_1\) being in the vertex boundary, showing that the vertex boundary is large. The distinction between the inner and outer vertex follows by similar arguments, using a symmetrization of \(B_t\) (Proposition 5.5).

In a following work by the first author [12], some of the ideas used in this paper were used to obtain quantitative bounds of Harris’ inequality.

2 Preliminaries

Throughout the text, the letter \(C\) stands for a positive universal constant, whose value may change from line to line.
2.1 Boolean functions

For a general introduction to Boolean functions, see [28]; in what follows, we provide a brief overview of the required background and notation.

Every function \( f : \{-1, 1\}^n \to \mathbb{R} \) may be uniquely written as a sum of monomials:

\[
  f(y) = \sum_{S \subseteq [n]} \hat{f}(S) \prod_{i \in S} y_i,
\]

(16)

where \([n] = \{1, \ldots, n\}\), and the harmonic coefficients (also known as Fourier coefficients) \( \hat{f}(S) \) are given by

\[
  \hat{f}(S) = \mathbb{E}_{y \sim \{-1, 1\}^n} \left[ f(y) \prod_{i \in S} y_i \right].
\]

(17)

The squares of the coefficients, \( \hat{f}(S)^2 \), are known as Fourier weights. For a given \( k \), the set \( \{ \hat{f}(S) \mid |S| = k \} \) is called the set of level-\( k \) coefficients. The variance of a function is given by the sum of its Fourier weights of level greater than 0:

\[
  \text{Var}(f) = \sum_{S \neq \emptyset} \hat{f}(S)^2.
\]

Equation (16) may be used to extend a function’s domain from the discrete hypercube \( \{-1, 1\}^n \) to real space \( \mathbb{R}^n \). We call this the harmonic extension, and denote it also by \( f \). Under this notation, \( f(0) = \mathbb{E} f \). In general, for \( x \in [-1, 1]^n \), the harmonic extension \( f(x) \) is a convex combination of \( f \)'s values on all the points \( y \in \{-1, 1\}^n \):

\[
  f(x) = \sum_{y \in \{-1, 1\}^n} w_x(y) f(y),
\]

(18)

where \( w_x(y) = \prod_{i=1}^n (1 + x_i y_i) / 2 \).

The derivative of a function \( f \) in direction \( i \) is defined as

\[
  \partial_i f(y) = \frac{f(y^{i \to 1}) - f(y^{i \to -1})}{2},
\]

where \( y^{i \to a} \) has \( a \) at coordinate \( i \), and is identical to \( y \) at all other coordinates. The gradient is then defined as \( \nabla f = (\partial_1 f, \ldots, \partial_n f) \). A function is called monotone if \( f(x) \leq f(y) \) whenever \( x_i \leq y_i \) for all \( i \in [n] \). Similar to the
function \( f \), by abuse of notation \( \partial_i f \) will denote the harmonic extension of \( \partial_i f \), and we will treat it as a function on \([-1, 1]^n\).

A short calculation reveals the following properties of the derivative:

1. The harmonic extension of the derivative \( \partial_i f \) is equal to the real-differentiable partial derivative \( \frac{\partial}{\partial x_i} \) of the harmonic extension of \( f \).
2. For functions whose range is \([-1, 1]\), the derivative \( \partial_i f \) takes values in \([-1, 0, 1]\), and the influence of the \( i \)-th coordinate of \( f \) is given by
   \[
   \text{Inf}_i (f) = E (\partial_i f)^2 = E |\partial_i f|.
   \] (19)
3. For monotone functions, the derivative \( \partial_i f \) only takes values in \( \{0, 1\} \), and the influence of the \( i \)-th coordinate is then given by
   \[
   \text{Inf}_i (f) = E \partial_i f = \hat{f} (\{i\}).
   \] (20)

In the definition of the Fourier coefficient in (17), the expectation is over the uniform measure \( \mu (y) = \frac{1}{2^n} \). It is also possible to decompose a function into Fourier coefficients over a biased measure. This type of analysis will be used only in the proof of Theorem 1.9. A brief overview can be found in the appendix.

Finally, we will require two lemmas which effectively relate the weights of the Fourier coefficients at higher levels with those of lower ones; this translates to inequalities between the harmonic extension of a function and its derivatives. See the Remark 2.3 after the lemmas’ statement for a brief explanation.

The first lemma essentially bounds the Fourier weights in the first level by a function of the weights at level zero:

**Lemma 2.1** (Level-1 inequality) There exists a continuous function \( C : [0, 1) \to [0, \infty) \) so that the following holds. Let \( g : [-1, 1]^n \to [0, 1] \) be the harmonic extension of a function \( \tilde{g} : \{-1, 1\}^n \to \{0, 1\} \) on the hypercube, let \( t \in [0, 1) \), and let \( x \in (-1, 1)^n \) be such that \( |x_i| = t \) for all \( i \). Then
   \[
   \| \nabla g (x) \|^2_2 \leq C (t) g (x)^2 \log \frac{e}{g (x)}.
   \] (21)

The second lemma, whose original, uniform case (i.e. \( x = 0 \)) is due to Talagrand [33], essentially bounds the Fourier weights in the second level by those of the first. It is similar to [22, Lemma 6], but for real-valued functions.

**Lemma 2.2** (Level-2 inequality) There exists a continuous function \( C : [0, 1) \to [0, \infty) \) so that the following holds. Let \( g : [-1, 1]^n \to [-1, 1] \) be the harmonic extension of a monotone function, and let \( x \in (-1, 1)^n \) be
such that $|x_i| = t$ for all $i$. Then

$$
\| \nabla^2 g (x) \|_{HS}^2 \leq C (t) \| \nabla g (x) \|_2^2 \cdot \log \left( \frac{C (t)}{\| \nabla g (x) \|_2^2} \right),
$$

(22)

where $\nabla^2 g$ is the Hessian $(\partial_i \partial_j g)_{i,j=1}^n$ of $g$, and $\| X \|_{HS}$ is the Hilbert–Schmidt norm of a matrix.

Remark 2.3 The Level-1 and Level-2 inequalities are usually stated in terms of the level-0, level-1 and level-2 Fourier weights, and so the above formulation is non-standard. However, the $k$-th derivatives of a function are related to its $k$-level Fourier weights; see Proposition 2.3 in the appendix. Similarly, the value $g (x)$ itself corresponds to $\hat{g} (\emptyset)$ under a $\frac{1+x}{2}$-biased Fourier analysis.

Remark 2.4 In both lemmas, the requirement that $|x_i| = t$ for all $i$ is not crucial, and can be replaced by $|x_i| \leq t$ for all $i$.

The proofs of both lemmas are found in the appendix.

2.2 Stochastic processes and quadratic variation

For a general introduction to stochastic processes and Poisson processes, see [10, 20].

A Poisson point process $N_t$ with rate $\lambda (t)$ is an integer-valued process such that $N_0 = 0$, and for every $0 \leq a < b$, the difference $N_b - N_a$ distributes as a Poisson random variable with rate $\int_a^b \lambda (t) \, dt$. If $\int_a^b \lambda (t) \, dt < \infty$ for all $0 \leq a < b$, then there is a version of the Poisson point process for which the sample-paths are right-continuous almost surely. The (random) set of times at which the sample-path is discontinuous is denoted by $\text{Jump} (N_t)$.

Let $\lambda (t)$ be such that $\int_a^b \lambda (t) \, dt < \infty$ for all $0 \leq a < b$ and let $N_t$ be a right-continuous Poisson point process with rate $\lambda (t)$. The set $\text{Jump} (N_t) = \{ t_1, t_2, \ldots \}$ is then almost surely discrete. A process $X_t$ is said to be a piecewise-smooth jump process with rate $\lambda (t)$ if $X_t$ is right-continuous and is smooth in the interval $[t_i, t_{i+1})$ for every $i = 1, 2, \ldots$. This definition can be extended to the case where $\int_0^b \lambda (t) \, dt = \infty$ but $\int_a^b \lambda (t) \, dt < \infty$ for all $0 < a < b$ (this happens, for example, when $\lambda = 1/t$): in this case $\text{Jump} (N_t)$ has only a single accumulation point at 0, and intervals between successive jump times are still well defined.

An important notion in the analysis of stochastic processes is quadratic variation. Intuitively, the quadratic variation of a process $X_t$, denoted $[X]_t$,
describes how wildly the process $X_t$ fluctuates; formally, it is defined as

$$\lbrack X \rbrack_t = \lim_{\|P\| \to 0} \sum_{k=1}^{n} (X_{t_k} - X_{t_{k-1}})^2,$$

if the limit exists; here $P$ is an $n$-part partition of $[0, t]$, and the notation $\lim_{\|P\| \to 0}$ indicates that the size of the largest part goes to 0. Not all processes have a (finite) quadratic variation, but piecewise-smooth jump processes do; in fact, it can be seen from definition that if $X_t$ is a piecewise-smooth jump process then

$$\lbrack X \rbrack_t = \sum_{s \in \text{Jump}(X_t) \cap [0, t]} (\Delta X_s)^2,$$

(23)

where $\Delta X_s = \lim_{\varepsilon \to 0^+} (X_{s+\varepsilon} - X_{s-\varepsilon})$ is the size of the jump at time $s$.

The quadratic variation is especially useful for martingales due to its relation with the variance: if $X_t$ is a martingale, then

$$\text{Var} (X_t) = \mathbb{E} (\lbrack X \rbrack_t).$$

(24)

In general, the quadratic variation may have a heavy tail. However, for bounded martingales with large variance, there is a non-negligible probability that the quadratic variation itself is large (and not just its expectation), as is shown by the following lemma. This lemma reflects one of the crucial ways in which we exploit the “pathwise” approach; namely, we analyze the behavior of the quadratic variation process beyond its expectation.

**Lemma 2.5** Let $a \in [0, 1)$ and let $c_1 \in [0, 1/3)$. Then there exists a constant $c_2$, depending only on $a$ and $c_1$, so that the following holds. Let $(X_t)_{t=0}^{1}$ be a right-continuous martingale such that:

- $X_0 \in [-a, a]$.
- $X_1 \in \{-1, 1\}$.
- $X_t \in [-1, 1]$ \(\forall t \in [0, 1]\).

Then

$$\mathbb{P} \left[ \lbrack X \rbrack_1 \geq c_1 (1 - a)^2 \right] \geq c_2 (1 - a)^2.$$

The constant $c_2$ can be taken to be larger than $c_1 \frac{c_1}{2(2 - \log(1-a)^2 - \log c_1)}$.

The proof, which relies on the fact that for such martingales the quadratic variation behaves like a sub-exponential random variable, is postponed to the appendix.
3 The main tool: a jump process

3.1 Construction and basic properties

The proof of Theorems 1.2 and 1.5 relies on the properties of a piecewise-smooth jump process martingale $B_t$, described below. One of its key properties is that it will allow us to express quantities such as the variance of $f$ in terms of derivatives of its harmonic extension, e.g.:

$$\text{Var}(f) = 2\mathbb{E} \sum_{i=1}^{n} \int_0^1 t (\partial_i f (B_t))^2 dt.$$  

The process $(B_t)_{t \geq 0}$ is characterized by the following properties:

1. $B_t \in \mathbb{R}^n$, with $B_t^{(i)}$ independent and identically distributed for all $i \in [n]$.
2. $B_t^{(i)}$ is a martingale for all $i$.
3. $|B_t^{(i)}| = t$ almost surely for all $i \in [n]$ and $t \geq 0$.

**Proposition 3.1** There exists a right continuous martingale with the above properties. This process satisfies, for all $t, h > 0$,

$$\mathbb{P}\left[\text{sign} B_{t+h}^{(i)} \neq \text{sign} B_t^{(i)} \mid B_t\right] = \frac{h}{2(t+h)}. \tag{25}$$

**Proof** Let $W_s$ be a standard Brownian motion. Consider the family of stopping times

$$\tau (t) = \inf \left\{ s > 0 \left| |W_s| > t \right. \right\}$$

and define $X_t = W_{\tau(t)}$. Then by definition, $|X_t| = t$, and $X_t$ is a martingale due to the optional stopping theorem. Observe that by definition $X_t$, the paths of $X_t$ are always right-continuous. The process $B_t$ is defined as $B_t = \left( X_t^{(1)}, \ldots, X_t^{(n)} \right)$, where $X_t^{(i)}$ are independent copies of $X_t$.

To prove Eq. (25), set $p = \mathbb{P}(\text{sign} X_{t+h} \neq \text{sign} X_t \mid X_t)$ and use the martingale property:

$$t \text{sign} X_t = X_t$$

$$= \mathbb{E} \left[ X_{t+h} \mid X_t \right]$$

$$= (-t-h) \text{sign} X_t \cdot p + (1-p) (t+h) \text{sign} X_t.$$  

Rearranging gives $p = \frac{h}{2(t+h)}$ as needed. \(\square\)
It can be readily seen that $B_t^{(i)}$ is a piecewise-smooth jump process with rate $\lambda(t) = 1/(2t)$. Denote its set of discontinuities by $J_i = \text{Jump}(B_t^{(i)})$.

As described in (16), the harmonic extension of a function $f: \{-1, 1\}^n \rightarrow \mathbb{R}$ is a multilinear polynomial. Since the product of two independent martingales is also a martingale with respect to its natural filtration, by independence of the coordinates of $B_t$, we conclude that

**Fact 3.2** *For a function $f: \{-1, 1\}^n \rightarrow \mathbb{R}$, the process $f(B_t)$ is a martingale.*

We denote this process by $f_t = f(B_t)$, and by slight abuse of notation, write $\partial_i f_t = \partial_i f(B_t)$ and $\nabla f_t = \nabla f(B_t)$. Since $B_t$ is right-continuous, these processes are right-continuous also; when referring to the left limit at jump discontinuities, we write $f_t^-, \partial_i f_t^-$ and $\nabla f_t^-$, with $f_t^- = \lim_{\varepsilon \downarrow 0} f_{t-\varepsilon}$. Some example sample paths of $f_t$ for the 15-bit Majority function are given in Fig. 2.

Since $f_t$ is a piecewise-smooth jump process, by (23) its quadratic variation is equal to the sum of squares of its jumps. Now, almost surely, $B_t$ can make a jump only in one coordinate at a time, and when the $i$-th coordinate jumps, the value of $f_t$ changes by $2t \partial_i f_t$, since $f$ is multi-linear. The quadratic variation of $f_t$ is therefore

$$[f]_t = \sum_{i=1}^n \sum_{s \in J_i \cap [0,t]} (2s \cdot \partial_i f_s)^2. \quad (26)$$

A crucial property of $B_t$ is that the expected value of these jumps behaves smoothly, as the next lemma shows:
Lemma 3.3 Let $0 \leq t_1 < t_2 \leq 1$, and let $g_t$ be a bounded process is left-continuous and measurable with respect to the filtration generated by $\{B_s\}_{0 \leq s < t}$. Then for every $i \in [n]$,

$$\mathbb{E} \sum_{t \in J_i \cap [t_1, t_2]} 4t^2 g_t = 2\mathbb{E} \int_{t_1}^{t_2} t \cdot g_t \, dt.$$  \hfill (27)

The proof is essentially a change in the order of summation, and involves going over all points in $[t_1, t_2]$ and calculating the jump rate at each point. It is postponed to the appendix.

Corollary 3.4 Let $f : \{-1, 1\}^n \to \mathbb{R}$. Then for all $t_0 > 0$,

$$\text{Var}(f_{t_0}) = 2\mathbb{E} \sum_{i=1}^{n} \int_{0}^{t_0} t (\partial_i f_t)^2 \, dt.$$  \hfill (28)

Proof Since $f_{t_0}$ is a martingale, by Eq. (24) its variance is the expected value of the quadratic variation, and hence by Eq. (26) we have:

$$\text{Var}(f_{t_0}) = \mathbb{E} \sum_{i=1}^{n} \sum_{t \in J_i \cap [0, t_0]} (2t \partial_i f_t)^2.$$  

Setting $g_t = (\partial_i f_t)^2$ for every $i \in [n]$ in (27) completes the proof. \hfill \square

Corollary 3.5 Let $f : \{-1, 1\}^n \to \mathbb{R}$. Then

$$\frac{d}{dt} \mathbb{E} f_t^2 = 2t \mathbb{E} \sum_{i=1}^{n} (\partial_i f_t)^2 = 2t \mathbb{E} \| \nabla f_t \|^2.$$  

Proof By the martingale property of $f_t$,

$$\frac{d}{dt} \mathbb{E} f_t^2 = \frac{d}{dt} (\mathbb{E} f_t^2 - \mathbb{E} f_0^2) = \frac{d}{dt} (\mathbb{E} (f_t - f_0)^2) = \frac{d}{dt} \text{Var}(f_t).$$

Taking the derivative of Eq. (28) and using the fundamental theorem of calculus on the right hand side gives the desired result. \hfill \square
3.2 The influence process

It is a basic fact (see e.g. [18, section 4.3]) that if $f$ has Fourier expansion $f(x) = \sum_S \hat{f}(S) \chi_S(x)$, its noise stability (defined in (14)) is given by

$$S_\varepsilon(f) = \sum_{S \neq \emptyset} \hat{f}(S)^2 (1 - \varepsilon)^{|S|}.$$  

On the other hand, recalling $f_t = f(B_t)$, a short calculation reveals that

$$\text{Var}(f_t) = \sum_{S \neq \emptyset} \hat{f}(S)^2 t^{2|S|}.$$  

Thus $S_\varepsilon(f) = \text{Var}(f_{\sqrt{1-\varepsilon}})$, and the inequality (15) in the statement of Theorem 1.9 becomes

$$\text{Var}(f_{\sqrt{1-\varepsilon}}) \leq C\text{Var}(f) \left( \sum_{i=1}^n \text{Inf}_i(f)^2 \right)^{c_\varepsilon}.$$  

Together with Eq. (28), this turns into

$$\mathbb{E} \sum_{i=1}^n \int_0^{\sqrt{1-\varepsilon}} t (\partial_i f_t)^2 dt \leq C\text{Var}(f) \left( \sum_{i=1}^n \text{Inf}_i(f)^2 \right)^{c_\varepsilon}. \quad (29)$$  

We will use this formulation rather than the original statement of (15).

For every index $i$, let $f^{(i)}$ be the harmonic extension of $|\partial_i f|$, and let

$$f^{(i)}_t := f^{(i)}(B_t). \quad (30)$$

If $f$ is monotone then $f^{(i)} = \partial_i f$, since the derivatives are positive, but in general,

$$f^{(i)}(x) \geq |\partial_i f(x)| \quad \forall x \in [-1, 1]^n \quad (31)$$

by convexity. In particular, plugging (31) into Corollary 3.4, we have

$$\text{Var}(f_0) \leq 2 \sum_{i=1}^n \int_0^t \mathbb{E} \left( f^{(i)}_t \right)^2 dt. \quad (32)$$

We call the process $f^{(i)}_t$ the “influence process”, because of how the expectation of its square relates to the influence of $f$: observe that, by (19),

$$f^{(i)}_0 = \mathbb{E} f^{(i)} = \mathbb{E} |\partial_i f| = \text{Inf}_i(f). \quad (33)$$
Thus, at time 0, we have $E \left( f_0^{(i)} \right)^2 = \left( f_0^{(i)} \right)^2 = \text{Inf}_i (f)^2$, while at time 1, since $f^{(i)} (y)^2 = f^{(i)} (y)$ for $y \in \{-1, 1\}^n$, we have $E \left( f_1^{(i)} \right)^2 = \mathbb{E} f_1^{(i)} = \mathbb{E} f^{(i)} = \text{Inf}_i (f)$. The expected value $E \left( f_1^{(i)} \right)^2$ increases from $\text{Inf}_i (f)^2$ to $\text{Inf}_i (f)$ as $t$ goes from 0 to 1. We denote this expected value by $\psi_i (t) := E \left( f_t^{(i)} \right)^2$. Equation (32) then becomes

$$\text{Var} \left( f \left( B_{t_0} \right) \right) \leq 2 \sum_{i=1}^{n} \int_{0}^{t_0} t \psi_i (t) \, dt.$$ 

The integral $\mathbb{E} \int_{0}^{1} t \psi_i (t) \, dt$ may be more easily handled using a time-change which makes $\psi_i (t)$ log-convex; we can then bound it by a power of the influence. For this purpose, for $s \in (0, \infty)$, denote

$$\varphi_i (s) := \psi_i \left( e^{-s} \right) = E \left( f_{e^{-s}}^{(i)} \right)^2. \tag{34}$$

**Lemma 3.6** Let $g : [-1, 1]^n \to \mathbb{R}$ be the harmonic extension of a function on the hypercube, and let $h (s) = \mathbb{E} g \left( B_{e^{-s}} \right)^2$. Then $h (s)$ is a log-convex function of $s$.

**Proof** Expanding $g$ as a Fourier polynomial, we have

$$h (s) = \mathbb{E} g \left( B_{e^{-s}} \right)^2 = \mathbb{E} \left[ \left( \sum_{S \subseteq [n]} \hat{g} (S) \prod_{i \in S} \left( B_{e^{-s}}^{(i)} \right) \right)^2 \right]$$

$$= \sum_{S \subseteq [n]} \hat{g} (S)^2 \prod_{i \in S} \mathbb{E} \left( B_{e^{-s}}^{(i)} \right)^2$$

$$= \sum_{S \subseteq [n]} \hat{g} (S)^2 e^{-2s|S|}. \tag{35}$$

This is a positive linear combination of log convex-functions $e^{-2s|S|}$, and is therefore also log-convex [6, section 3.5.2].

Finally, we will need the following easy technical lemma, whose short proof is postponed to the appendix.
Lemma 3.7 Let \( g : [0, \infty) \rightarrow [0, \infty) \) be a differentiable function satisfying

\[
g'(t) \leq C \cdot g(t) \log \frac{K}{g(t)}, \tag{36}
\]

where \( C, K \) are some positive constants. Suppose that \( g(0) \leq K/2 \). Then there exists a time \( t_0 \) which depends only on \( C \) and \( K \), such that for all \( t \in [0, t_0] \),

\[
g(t) \leq K^{Ct} g(0)^{1-Ct}.
\]

4 Proof of the strengthening of Talagrand’s conjecture

We prove Theorem 1.4 assuming that inequality (29) holds; the proof of (29) is found in Sect. 6. Without loss of generality, we assume that \( \mathbb{E} f \leq 0 \).

The function \( g_f : \{-1, 1\}^n \rightarrow [0, 1] \) used in Theorem 1.4 will be defined by

\[
g_f(y) := \mathbb{E} \left[ \sup_{0 \leq s \leq 1} \frac{1 + f_s}{2} \mid B_1 = y \right].
\]

It follows from Doob’s martingale inequality (see e.g. [10, Theorem 4.4.4]) that for a martingale \( (X_t)_{t=0}^1 \),

\[
\mathbb{E} \left[ \left( \sup_{0 \leq s \leq 1} X_s \right)^2 \right] \leq 4 \mathbb{E} \left[ X_1^2 \right].
\]

Using this inequality for the martingale \( X_t = \frac{1 + f_t}{2} \), we thus have

\[
\mathbb{E} g_f^2 = \mathbb{E} \left[ \mathbb{E} \left[ \sup_{0 \leq s \leq 1} \frac{1 + f_s}{2} \mid B_1 \right]^2 \right] \leq \mathbb{E} \left[ \mathbb{E} \left[ \sup_{0 \leq s \leq 1} \left( \frac{1 + f_s}{2} \right)^2 \mid B_1 \right] \right] = \mathbb{E} \left[ \sup_{0 \leq s \leq 1} \left( \frac{1 + f_s}{2} \right)^2 \right] \leq 2 \left( 1 + \mathbb{E} f \right) = 2 \frac{1 - (\mathbb{E} f)^2}{1 - \mathbb{E} f} \leq 2 \text{Var}(f).
\]

as required by the theorem.

We now turn to prove the inequality (9) using \( g_f \). To this end, we can relate the product \( h_f^p(y) g_f(y) \) to the stochastic constructions in the previous section. By definition of the discrete derivative, for any \( y \in \{-1, 1\}^n \), \( \partial_i f(y) = 0 \)
if \( f(y) = f(y^\oplus i) \), and \( \partial_i f(y) = \pm 1 \) if \( f(y) \neq f(y^\oplus i) \), so

\[
h_f(y) = \sum_{i=1}^{n} \partial_i f(y)^2 = \|\nabla f(y)\|_2^2.
\]

We thus have \( h_f^p(y) = \|\nabla f(y)\|_2^{2p} \), and using this relation we define the stochastic process

\[
\Psi_t := \|\nabla f_t\|_2^{2p} \sup_{0 \leq s \leq t} \frac{1 + f_s}{2}, \tag{37}
\]

noting that

\[
\mathbb{E}\Psi_1 = \mathbb{E}\left[h_f^p g_f\right].
\]

Our goal is therefore to bound \( \mathbb{E}\Psi_1 \) from below.

Since the function \( \|\cdot\|_2^{2p} : \mathbb{R}^n \rightarrow \mathbb{R} \) is convex, the process \( \|\nabla f_t\|_2^{2p} \) is a submartingale by Jensen’s inequality. A simple calculation then shows that \( \mathbb{E}\Psi_t \) is an increasing function of \( t \). For the rest of this section, we therefore assume that

\[
\mathbb{E}\Psi_t \leq \text{Var}(f) \cdot \left( \log \left( 2 + \frac{e}{\sum_i \text{Inf}_i(f)^2} \right) \right)^p; \quad 0 \leq t \leq 1, \tag{38}
\]

otherwise there is nothing to prove.

Our proof relies on the existence of a stopping time \( \tau_\alpha \) such that with high probability \( \Psi_{\tau_\alpha} \) is large, as is shown by the following proposition. Fix \( \alpha > 0 \) whose value is to be chosen later, and define

\[
\tau_\alpha := \inf \left\{ 0 \leq t \leq 1 \mid \Psi_t > \frac{1}{8} \alpha^{2p} \left( \log \left( 2 + \frac{e}{\sum_i \text{Inf}_i(f)^2} \right) \right)^p \right\} \land 1. \tag{39}
\]

**Proposition 4.1** Assume that (38) holds. Then there exist universal constants \( C > 0 \), and \( \alpha > 0 \) such that

\[
\mathbb{P}[\tau_\alpha < 1] \geq C \text{Var}(f).
\]

Assuming the above proposition holds, we can prove our main result:

**Proof of Theorem 1.4** By conditioning on the event \( \tau_\alpha < 1 \), we have
\[ E \Psi_1 \geq E \Psi_{\tau_\alpha} \geq E \left[ \Psi_{\tau_\alpha} \bigg| \tau_\alpha < 1 \right] \mathbb{P} [ \tau_\alpha < 1 ] \geq \frac{1}{8} \alpha^2 p \left( \log \left( 2 + \frac{e}{\sum_i \inf_i (f)^2} \right) \right)^p \cdot C \text{Var} (f). \]

The rest of this section is devoted to proving Proposition 4.1. The main idea is to see how different sample paths contribute to the quadratic variation of \( f_t \). The lion’s share of the quadratic variation is gained from paths where the gradient’s norm \( \| \nabla f_t \|_2 \) is large. Since the quadratic variation has a relatively high probability to be large, \( \| \nabla f_t \|_2 \) must be large with relatively high probability as well. This argument takes care of the gradient’s contribution to \( \Psi_t \); to deal with the sup \( \sup_{s \leq t} \frac{1 + f_s}{2} \) term, we show that with high enough probability, either \( f_t \) makes a large jump (which causes both sup \( f_t \) and the gradient’s norm to be large at the same time) or there is a time where \( f_t \)’s position is bounded away from the endpoints \( \{-1, 1\} \), allowing its gradient to be large later on.

**Proof of Proposition 4.1** Let

\[ \theta := \inf \{ t \geq 0 \mid f_t > 0 \} \land 1. \tag{40} \]

Since \( f_t \) is a martingale,

\[ f_0 = \mathbb{E}[f_1] = 2 \mathbb{P} [ f_1 = 1 ] - 1, \]

and since \( \{ f_1 = 1 \} \subseteq \{ \theta < 1 \} \),

\[ \mathbb{P} [ \theta < 1 ] \geq \mathbb{P} [ f_1 = 1 ] = \frac{1 + f_0}{2} = \frac{1 - f_0^2}{2 (1 - f_0)} = \frac{\text{Var}(f)}{2 (1 - f_0)} \geq \frac{1}{4} \text{Var}(f). \tag{41} \]

By conditioning on \( \theta < 1 \) we have

\[ \mathbb{P} [ \tau_\alpha < 1 ] \geq \mathbb{P} \left[ \tau_\alpha < 1 \bigg| \theta < 1 \right] \mathbb{P} [ \theta < 1 ] \geq \frac{1}{4} \text{Var}(f) \cdot \mathbb{P} \left[ \tau_\alpha < 1 \bigg| \theta < 1 \right]. \]

It remains only to show that for small enough (but fixed) \( \alpha \), \( \mathbb{P} [ \tau_\alpha < 1 \mid \theta < 1 ] \) is larger than some constant.

Since we assume \( \mathbb{E}[f] \leq 0 \), the process \( f_t \) starts at \( f_0 \leq 0 \). There are two different ways for \( f_t \) to cross above the value 0: it could either move across it continuously, or it could jump from some value smaller than 0 to some value larger than 0. We now divide the analysis into two cases, depending on the probability that \( f_t \) makes a very large jump over 0 at time \( \theta \).
Case 1: small jump

Suppose that $\mathbb{P} \left[ f_\theta \in [0, 1/2] \mid \theta < 1 \right] \geq 1/2$, i.e. with constant probability, no large jump was made at time $\theta$. The next two propositions show that with high probability the quadratic variation gained from time $\theta$ onward must be large, and that most of the quadratic variation is gained at times when, just before the process jumps, the gradient $\nabla f_i$ is large (recall that $\nabla f_i = \lim_{\varepsilon \to 0} \nabla f_{i-\varepsilon}$). To make these notions precise, we will need the following definitions. Let

$$F_{\alpha,t} := \left\{ \| \nabla f_i \|_2 > \alpha \sqrt{\log \left( 2 + \frac{e}{\sum_i \text{Inf}_i (f)^2} \right)} \right\}$$  \hspace{1cm} (42)

be the event that the norm of the gradient is large just before time $t$; let

$$E_t := \left\{ \sup_{0 \leq s < t} f_s \geq 0 \right\}$$  \hspace{1cm} (43)

be the event that sup $f_s$ is large strictly before time $t$; let

$$V_\alpha := \sum_{i=1}^{n} \sum_{t \in J_i \cap [0,1]} (2t \partial_i f_t)^2 1_{F_{\alpha,t}} 1_{E_t}$$  \hspace{1cm} (44)

be the quadratic variation accumulated at times when the supremum is large but the gradient is small; and let

$$\mathcal{V}^{t_1 \to t_2} := \sum_{i=1}^{n} \sum_{t \in J_i \cap [t_1, t_2]} (2t \partial_i f_t)^2$$  \hspace{1cm} (45)

be the gain in quadratic variation from time $t_1$ up to and including time $t_2$. We will need the following lemma, whose proof is found in the next subsection.

**Lemma 4.2** Let $0 < \alpha < 1/e$, and assume that (38) holds. There exists a function $\rho : [0, 1] \to \mathbb{R}$ with $\lim_{s \to 0} \rho (s) = 0$ such that

$$\mathbb{E} [V_\alpha] \leq \rho (\alpha) \text{Var} (f).$$  \hspace{1cm} (46)

Using this lemma, we show that $\mathbb{P} \left[ \tau_\alpha < 1 \mid \theta < 1 \right] \geq C$. On one hand, by Lemma 4.2, the expected increase in the quadratic variation from time $\theta$ to time 1 of paths with small gradient must be small: denoting
Concentration on the Boolean hypercube

\[ A_{\alpha,\theta} := \bigcap_{t \in (\theta,1]} F_{\alpha,t}^\theta \]  

(47)

to be the event that the gradient was small at all times from \( \theta \) to 1, we have

\[
\mathbb{E}\left[ V^{\theta \to 1} 1_{A_{\alpha,\theta}} \mid \theta < 1 \right] \leq \mathbb{E}\left[ \sum_i \sum_{t \in J_i \cap (\theta,1]} (2\partial_i f_t)^2 1_{F_{\alpha,t}^\theta} \mid \theta < 1 \right]
\]

\[
= \mathbb{E}\left[ \sum_i \sum_{t \in J_i \cap (\theta,1]} (2\partial_i f_t)^2 1_{F_{\alpha,t}^\theta} 1_{E_t} \mid \theta < 1 \right]
\]

\[
= \frac{\mathbb{E}\left[ \sum_i \sum_{t \in J_i \cap (\theta,1]} (2\partial_i f_t)^2 1_{F_{\alpha,t}^\theta} 1_{E_t} \right]}{\mathbb{P}[\theta < 1]} \leq \mathbb{E}\left[ V_{\alpha} \right] \overset{(46),(41)}{\leq} 4\rho(\alpha).
\]

(48)

On the other hand, consider the martingale \( X_t = f_{\theta+t,1-\theta} \). Conditioned on the event \( \{ f_\theta \in [0,1/2] \} \), \( X_t \) is still a martingale, and in this case it satisfies \( X_0 \in [-1/2,1/2] \), \( X_1 \in \{-1,1\} \), and \( X_t \in [-1,1] \) for all \( t \in [0,1] \). We can therefore invoke Lemma 2.5 with \( a = 1/2 \) and \( c_1 = 1/4 \), yielding

\[
\mathbb{P}\left[ [X]_1 \geq \frac{1}{16} \mid f_\theta \in [0,1/2] \right] \geq \frac{1}{640}.
\]

Since we are dealing with the case of small jumps, i.e. \( \mathbb{P}[f_\theta \in [0,1/2] \mid \theta < 1] \geq 1/2 \), this gives a bound on the gain in the quadratic variation from time \( \theta \):

\[
\mathbb{P}\left[ V^{\theta \to 1} \geq \frac{1}{16} \mid \theta < 1 \right] = \mathbb{P}\left[ [X]_1 \geq \frac{1}{16} \mid \theta < 1 \right]
\]

\[
\geq \mathbb{P}\left[ [X]_1 \geq \frac{1}{16} \mid f_\theta \in [0,1/2] \right] \mathbb{P}[f_\theta \in [0,1/2]]
\]

\[
\geq \frac{1}{1280}.
\]

We then have

\[
\mathbb{E}\left[ V^{\theta \to 1} 1_{A_{\alpha,\theta}} \mid \theta < 1 \right] \geq \frac{1}{16} \mathbb{P}\left[ \left\{ V^{\theta \to 1} > \frac{1}{16} \right\} \cap A_{\alpha,\theta} \mid \theta < 1 \right]
\]

\[
\geq \frac{1}{16} \left( \mathbb{P}\left[ V^{\theta \to 1} > \frac{1}{16} \mid \theta < 1 \right] - \mathbb{P}\left[ A_{\alpha,\theta}^C \mid \theta < 1 \right] \right).
\]
Solving for $\mathbb{P}\left[A_{\alpha,\theta}^C \mid \theta < 1\right]$ gives

$$\mathbb{P}\left[A_{\alpha,\theta}^C \mid \theta < 1\right] \geq \mathbb{P}\left[V^{\theta \rightarrow 1} > \frac{1}{16} \mid \theta < 1\right] - 16\mathbb{E}\left[V^{\theta \rightarrow 1}1_{A_{\alpha,\theta}} \mid \theta < 1\right] \geq \frac{1}{1280} - C\rho(\alpha).$$

Since $\lim_{\alpha \to 0} \rho(\alpha) = 0$, this last expression is larger than some constant $c$ for small enough $\alpha$. But the event $A_{\alpha,\theta}^C$ means that at some time $t^* \geq \theta$, the gradient $\|\nabla f_t\|_2$ was larger than $\alpha \sqrt{\log\left(2 + \frac{e}{\sum_i \text{Inf}_i(f)^2}\right)}$, while the event $\theta < 1$ implies that $\sup_{0\leq s \leq \theta} \frac{1 + f_s}{2} \geq 1/2$, yielding $\Psi_{t^*} \geq \frac{1}{2}\alpha^2 p \left(\log\left(2 + \frac{e}{\sum_i \text{Inf}_i(f)^2}\right)\right)^p$. Thus

$$\mathbb{P}\left[\tau_{\alpha} < 1 \mid \theta < 1\right] \geq \mathbb{P}\left[A_{\alpha,\theta}^C \mid \theta < 1\right] > c,$$

as needed.

**Case 2: large jump**

Suppose now that $\mathbb{P}\left[f_\theta \in [0, \frac{1}{2}] \mid \theta < 1\right] < 1/2$, meaning that with large probability $f_i$ makes a large jump at time $\theta$ (defined in (40)) to some value greater than $1/2$.

$$\mathbb{P}\left[\Delta f_{\theta} \geq \frac{1}{2} \mid \theta < 1\right] \geq \frac{1}{2}. \quad (49)$$

The next lemma, parallel to Lemma 4.2, shows that most of the quadratic variation is gained at times when, just after the process jumped, the gradient was large. For a constant $\alpha > 0$ whose value is to be chosen later, let

$$H_{\alpha, t} := \left\{\|\nabla f_t\|_2 > \alpha \sqrt{\log\left(2 + \frac{e}{\sum_i \text{Inf}_i(f)^2}\right)}\right\} \quad (50)$$

1 We know of no specific Boolean function which satisfies this condition. In fact, we believe that the “large jump” case essentially never happens, i.e. that there exist global constants $c_1$ and $c_2$ such that $\mathbb{P}\left[f_\theta \in [0, c_1] \mid \theta < 1\right] > c_2$. However, we currently have no proof for such a statement.
be the event that the norm of the gradient is large exactly at time \( t \), and let

\[
U_\alpha := \sum_{i=1}^{n} \sum_{t \in I_i \cap [0,1]} (2t \partial_t f_i)^2 1_{H_{\alpha,t}^c, t \{ f_i \geq 0 \}}
\]  

(51)

be the quadratic variation accumulated at times when \( f \)'s value is large but the gradient is small.

**Lemma 4.3** Let \( 0 < \alpha < 1/e \), and assume that (38) holds. There exists a function \( \rho : [0, 1] \rightarrow \mathbb{R} \) with \( \lim_{x \to 0} \rho (x) = 0 \) such that

\[
\mathbb{E} [U_\alpha] \leq \rho (\alpha) \text{Var} (f) .
\]  

(52)

The proof of this lemma is found in the next subsection. Assuming it holds, we now show that \( \mathbb{P} [\tau_\alpha < 1] \geq C \text{Var} (f) \). Consider the event that at time \( \theta \), both \( \| \nabla f_\theta \|_2 < \alpha \sqrt{\log \left( 2 + \frac{e}{\sum_i \text{Inf}_i (f)^2} \right)} \) and \( \Delta f_\theta \geq 1/2 \). Since \( \theta \) is the first time that \( f_i \geq 0 \), and since with probability 1 there is no jump at time 1, this event contributes at least \( 1/4 \) to \( U_\alpha \); thus

\[
\mathbb{P} \left[ U_\alpha \geq \frac{1}{4} \right] 
\geq \mathbb{P} \left[ \| \nabla f_\theta \|_2 \leq \alpha \sqrt{\log \left( 2 + \frac{e}{\sum_i \text{Inf}_i (f)^2} \right)} \cap \Delta f_\theta \geq \frac{1}{2} \right]
\geq \mathbb{P} \left[ \| \nabla f_\theta \|_2 \leq \alpha \sqrt{\log \left( 2 + \frac{e}{\sum_i \text{Inf}_i (f)^2} \right)} \big| \Delta f_\theta \geq \frac{1}{2} \right] \mathbb{P} \left[ \Delta f_\theta \geq \frac{1}{2} \right]
\geq \mathbb{P} \left[ \| \nabla f_\theta \|_2 \leq \alpha \sqrt{\log \left( 2 + \frac{e}{\sum_i \text{Inf}_i (f)^2} \right)} \big| \Delta f_\theta \geq \frac{1}{2} \right] \frac{1}{8} \text{Var} (f) .
\]  

On the other hand, by Markov’s inequality and Lemma 4.3, this probability is upper-bounded by

\[
\mathbb{P} \left[ U_\alpha \geq \frac{1}{4} \right] \leq \frac{\mathbb{E} [U_\alpha]}{1/4} \leq 4 \rho (\alpha) \text{Var} (f) .
\]

Combining the two displays, we get
Taking $\alpha$ small enough so that the right hand side is smaller than $1/2$, we have

$$P \left[ H_{\alpha, \theta} \big| \Delta f_{\theta} \geq \frac{1}{2} \right] > \frac{1}{2}.$$ 

For this $\alpha$, under the event $H_{\alpha, \theta} \cap \{ \theta < 1 \}$, $\Psi_{\theta}$ (defined in (37)) is large:

$$\Psi_{\theta} = \| \nabla f_{\theta} \|_2^{2p} \sup_{0 \leq s \leq \theta} \frac{1 + f_s}{2} \geq \| \nabla f_{\theta} \|_2^{2p} \frac{1 + f_{\theta}}{2} \geq \frac{1}{2} \alpha^{2p} \left( \log \left( 2 + \frac{e}{\sum_i \inf_i (f)^2} \right) \right)^p,$$

and so

$$P \left[ \tau_{\alpha} < 1 \big| \theta < 1 \right] \geq P \left[ \tau_{\alpha} < 1 \big| \theta < 1 \text{ and } \Delta f_{\theta} \geq \frac{1}{2} \right] P \left[ \Delta f_{\theta} \geq \frac{1}{2} \big| \theta < 1 \right]$$

$$= P \left[ \tau_{\alpha} < 1 \big| \Delta f_{\theta} \geq \frac{1}{2} \right] P \left[ \Delta f_{\theta} \geq \frac{1}{2} \big| \theta < 1 \right]$$

$$\geq P \left[ H_{\alpha, \theta} \big| \Delta f_{\theta} \geq \frac{1}{2} \right] \cdot \frac{1}{2} \geq \frac{1}{4},$$

as needed, where the middle inequality uses the fact that the event $\{ \Delta f_{\theta} \geq \frac{1}{2} \}$ contains $\{ \theta < 1 \}$ up to a set of measure 0, since with probability 1 there is no jump at time 1. \hfill \Box

4.1 Postponed proofs

**Proof of Lemma 4.2** We first express $E \left[ V_\alpha \right] = \sum_{i=1}^n \sum_{t \in J_i \cap [0,1]} (2t \partial_i f_t)^2 \mathbf{1}_{F_a} \mathbf{1}_{E_i}$ as an integral, rather than a sum over jumps. Since $\partial_i f_t$ is independent of coordinate $i$, we have that for $t \in J_i$, $\partial_i f_t = \partial_i f_t$. Thus

$$E \left[ V_\alpha \right] = E \left[ \sum_{i=1}^n \sum_{t \in J_i \cap [0,1]} (2t \partial_i f_t)^2 \mathbf{1}_{F_a} \mathbf{1}_{E_i} \right].$$
The process \( g_t = (\partial_i f_t^-)^2 1_{F_{\alpha}^c} 1_{E_t} \) is measurable with respect to the filtration generated by \( \{B_s\}_{0 \leq s < t} \) and is left-continuous. Invoking Lemma 3.3, we have

\[
\mathbb{E}[V_\alpha] = \mathbb{E} \sum_{i=1}^{n} \sum_{t \in J_i \cap [0,1]} 4t^2 g_t = 2 \int_0^1 t \mathbb{E}[\|\nabla f_t^-\|_2^2 1_{F_{\alpha}^c} 1_{E_t}] dt
\]

(Lemma 3.3)

\[
= 2 \int_0^1 t \mathbb{E}[\|\nabla f_t\|_2^2 1_{F_{\alpha}^c} 1_{E_t}] dt,
\]

where the last equality is because \( \|\nabla f_t^-\|_2 1_{F_{\alpha}^c} 1_{E_t} \) can differ from \( \|\nabla f_t\|_2 1_{F_{\alpha}^c} 1_{E_t} \) only at finitely many discontinuities.

Let \( \delta' > 0 \) be defined as

\[
\delta' := \begin{cases} 
\frac{c^{-1} \log(1/\alpha)}{\log(2 + \frac{e^e}{\sum_i \text{Inf}_i(f)^2})} & \text{if } \sum_i \text{Inf}_i(f)^2 \leq \frac{1}{2} \\
1 & \text{otherwise},
\end{cases}
\]

where \( c \) is the universal constant from Theorem 1.9, and set

\[
\delta := \min \{\delta', 1\}.
\]

Consider the integral

\[
\int_0^1 t \mathbb{E}[\|\nabla f_t\|_2^2 1_{F_{\alpha,t}^c} 1_{E_t}] dt = \int_0^{1-\delta} t \mathbb{E}[\|\nabla f_t\|_2^2 1_{F_{\alpha,t}^c} 1_{E_t}] dt + \int_{1-\delta}^1 t \mathbb{E}[\|\nabla f_t\|_2^2 1_{F_{\alpha,t}^c} 1_{E_t}] dt.
\]

The first integral on the right hand side is equal to 0 if \( \delta = 1 \). Otherwise, we necessarily have that \( \sum_i \text{Inf}_i(f)^2 \leq 1/2 \), in which case this integral can be bounded using Eq. (29): since \( 1 - \delta \leq \sqrt{1 - \delta} \) for all \( \delta \in [0, 1] \), we have

\[
\int_0^{1-\delta} t \mathbb{E}[\|\nabla f_t\|_2^2 1_{F_{\alpha,t}^c} 1_{E_t}] dt \leq \mathbb{E} \int_0^{1-\delta} t \|\nabla f_t\|_2^2 dt 
\]

\[
\leq \mathbb{E} \int_0^{\sqrt{1-\delta}} t \|\nabla f_t\|_2^2 dt
\]

\( \text{Springer} \)
\[
\begin{align*}
\leq C_1 \text{Var} \left( f \right) \left( \sum_i \inf_i \left( f \right)^2 \right)^{c_0} \\
\leq C_1 \text{Var} \left( f \right) \alpha \log \left( \frac{\sum_i \inf_i \left( f \right)^2}{\log 2 \sum_i \inf_i \left( f \right)^2 + e} \right)
\end{align*}
\]

for some constant $C_1 > 0$. Since $\sum_i \inf_i \left( f \right)^2 \leq 1/2$, the exponent $\log \left( \sum_i \inf_i \left( f \right)^2 \right) / \log \left( \frac{2 \sum_i \inf_i \left( f \right)^2 + e}{\log 2 \sum_i \inf_i \left( f \right)^2 + e} \right)$ is bounded below by $1/3$. Since $\alpha < 1$, we thus have that regardless of the value of $\sum_i \inf_i \left( f \right)^2$,

\[
\int_{1-\delta}^1 t \mathbb{E} \left[ \| \nabla f_t \|^2_2 \mathbf{1}_{F_{\alpha,t}^C} \mathbf{1}_{E_t} \right] dt \leq C_1 \alpha^{1/3} \text{Var} \left( f \right).
\] \text{(56)}

The second integral on the right hand side can be bounded using the definitions of $F_{\alpha,t}$ and $E_t$: by definition of $F_{\alpha,t}$ we have

\[
\mathbb{E} \left[ \| \nabla f_t \|^2_2 \| \nabla f_t \|^2_2 \mathbf{1}_{F_{\alpha,t}^C} \mathbf{1}_{E_t} \right] = \mathbb{E} \left[ \left( \| \nabla f_t \|^2_2 \mathbf{1}_{F_{\alpha,t}^C} \right) \| \nabla f_t \|^2_2 \mathbf{1}_{E_t} \right] \\
\leq \alpha^{2-2p} \left( \log \left( 2 + \frac{e}{\sum_i \inf_i \left( f \right)^2} \right) \right)^{1-p} \times \mathbb{E} \left[ \| \nabla f_t \|^2_2 \mathbf{1}_{E_t} \right],
\] \text{(57)}

whereas by the definition of $E_t$,

\[
\mathbb{E} \left[ \| \nabla f_t \|^2_2 \mathbf{1}_{E_t} \right] \leq 2 \mathbb{E} \left[ \| \nabla f_t \|^2_2 \sup_{s < t} \frac{1 + f_s}{2} \mathbf{1}_{E_t} \right] \\
\leq 2 \mathbb{E} \left[ \| \nabla f_t \|^2_2 \sup_{s < t} \frac{1 + f_s}{2} \right] \\
\leq 2 \mathbb{E} \Psi_t \leq 2 \text{Var} \left( f \right) \left( \log \left( 2 + \frac{e}{\sum_i \inf_i \left( f \right)^2} \right) \right)^{p}.
\]

Plugging the above display into (57), the second integral in (55) can be bounded by

\[
\int_{1-\delta}^1 t \mathbb{E} \left[ \| \nabla f_t \|^2_2 \mathbf{1}_{F_{\alpha,t}^C} \mathbf{1}_{E_t} \right] dt = \int_{1-\delta}^1 t \mathbb{E} \left[ \| \nabla f_t \|^2_2 \| \nabla f_t \|^2_2 \mathbf{1}_{F_{\alpha,t}^C} \mathbf{1}_{E_t} \right] dt \\
\leq \delta \alpha^{2-2p} \text{Var} \left( f \right) \log \left( 2 + \frac{e}{\sum_i \inf_i \left( f \right)^2} \right).
\]
Concentration on the Boolean hypercube 967

Since \( \alpha < 1/e \), there exists a constant \( C_2 > 0 \) so that any case, \( \delta \leq C_2 \cdot \frac{c^{-1} \log(1/\alpha)}{\log\left(2 + \frac{e}{\sum_{i} \inf(f_i)^2}\right)} \). We thus have

\[
\int_{1-\delta}^{1} t \mathbb{E} \left[ \| \nabla f_t \|_2^2 1_{F_{\alpha,t} \cap E_i} \right] dt \leq C_2 c^{-1} \alpha^{2-2p} \log \left(1/\alpha\right) \text{Var}(f) .
\] (58)

Combining (56) and (58), there exists an absolute constant \( C := C_1 + C_2 c^{-1} > 0 \) such that

\[
\int_{0}^{1} t \mathbb{E} \left( \| \nabla f_t \|_2^2 1_{F_{\alpha,t} \cap E_i} \right) dt \leq C \left( \alpha^{1/3} + \alpha^{2-2p} \right) \log \left(1/\alpha\right) \text{Var}(f) .
\] (59)

Plugging this into (53) finishes the proof, with \( \rho(x) = C \left( x^{1/3} + x^{2-2p} \right) \log \left(1/x\right) \).

\( \square \)

Proof of Lemma 4.3 For \( i \in [n] \), let \( s_i : \mathbb{R}^n \to \mathbb{R}^n \) be the function which “flips” the \( i \)-th coordinate. i.e.

\[
(s_i(x))^{(j)} = \begin{cases} -x & j = i, \\ x & j \neq i. \end{cases}
\]

Observe that if \( B_t^{(i)} \) makes a jump at time \( t \), then \( B_t^{-} = s_i(B_t) \). Thus, for such times, we have \( f(B_t^{-}) = f(s_i(B_t)) \) and \( \| \nabla f(B_t^{-}) \|_2 = \| \nabla f(s_i(B_t)) \|_2 \).

Denote \( \beta = \alpha \sqrt{\log \left(2 + \frac{e}{\sum_{i} \inf(f_i)^2}\right)} \). Since \( B_t \) and \( s_i(B_t) \) have the same distribution for every \( i \in [n] \), and since \( \partial_i f(x) \) does not depend on the \( i \)-th coordinate of \( x \), we have that

\[
\mathbb{E} [U_\alpha] \overset{\text{def. (51)}}{=} \mathbb{E} \left[ \sum_{i=1}^{n} \sum_{t \in J_i \cap [0,1]} (2t \partial_i f_t)^2 1_{\| \nabla f(B_t) \|_2 \leq \beta} 1_{\{f(B_t) \geq 0\}} \right] \\
= \mathbb{E} \left[ \sum_{i=1}^{n} \sum_{t \in J_i \cap [0,1]} (2t \partial_i f_t)^2 1_{\| \nabla f(s_i(B_t)) \|_2 \leq \beta} 1_{\{f(s_i(B_t)) \geq 0\}} \right] \\
= \mathbb{E} \left[ \sum_{i=1}^{n} \sum_{t \in J_i \cap [0,1]} (2t \partial_i f_t)^2 1_{\| \nabla f_t^- \|_2 \leq \beta} 1_{\{f_t^- \geq 0\}} \right]
\]
\[
\leq \mathbb{E} \left[ \sum_{i=1}^{n} \sum_{t \in J_i \cap [0,1]} (2t \partial_t f_t)^2 1_{\|\nabla f_t - \|_2 \leq \beta} 1_{\{\sup_{0 \leq s < t} f_s \geq 0\}} \right]
\]
def. (44) = \mathbb{E} [V_\alpha].

The lemma then follows from Lemma 4.2.

\[\square\]

5 Talagrand’s influence inequality and its stability

Our proofs of Theorems 1.2 and 1.5 are similar in spirit to that of Theorem 1.3, and again require bounding the gain in quadratic variation. However, extra care is needed to bound the size of the individual influence processes \(f_t^{(i)}\).

We first define several quantities which will be central to our proofs. For a fixed \(0 < \alpha \leq 1\) whose value will be chosen later, let

\[
F_\alpha := \left\{ \exists t \in [0, 1], \exists i \in [n] \mid t \in J_i \text{ and } f_t^{(i)} \geq \alpha \right\}
\]

be the event that a coordinate had large derivative at the time it jumped; let

\[
Q_\alpha^{(i)} := 2 \int_0^1 t \left( f_t^{(i)} \right)^2 1_{f_t^{(i)} < \alpha} dt,
\]

and

\[
Q_\alpha := \sum_{i=1}^{n} Q_\alpha^{(i)};
\]

and let

\[
V_\alpha^{(i)} := \sum_{t \in J_i \cap [0,1]} (2t \partial_t f_t)^2 1_{f_t^{(i)} < \alpha}
\]

and

\[
V_\alpha := \sum_{i=1}^{n} V_\alpha^{(i)}.
\]

Note that for \(\alpha = 1\), \(Q_1 = V_1\) are just the quadratic variation of \(f_t\). For \(\alpha < 1\), \(V_\alpha\) can be thought of as the quadratic variation of the process \(f_t\), but where big jumps (i.e. those larger than \(t\alpha\)) are excluded. \(Q_\alpha\) is the integral analogue.

Finally, define

\[
\rho (x) := x \left( \log \frac{1}{x} + 2 \right).
\]

Instead of using Theorem 1.9 to bound influences, we use the following lemma.
Lemma 5.1 There exists a universal constant $\gamma > 1$ so that for all $0 \leq s \leq \gamma$,

$$\varphi_i (s) \leq \gamma \inf_i (f)^{1+s/(2\gamma)},$$

(64)

where $\varphi_i (s) = \mathbb{E} \left( f^{(i)}_{e-s} \right)^2$ is as defined in (34).

This lemma can be derived from the hypercontractivity principle (see e.g. [7] and [28, Cor. 9.25]). However, we give a different proof based on the analysis of the stochastic process $f_t$; this analysis can be pushed further to obtain the stability results. On an intuitive level and in light of Eq. (28) the lemma shows that the essential contribution of the quadratic variation to the variance of the function happens very close to time 1.

Proof of Lemma 5.1 Let $\gamma > 1$, to be chosen later. We start by showing that there exists a constant $c'_\gamma > 0$ such that

$$\varphi_i (\gamma) \leq \gamma \varphi_i (0)^{1+c'_\gamma}.$$  

(65)

Recall that by definition, $\psi_i (t) = \varphi_i (\log 1/t)$; by applying Corollary 3.5 to the function $f_t^{(i)}$ (defined in (30)), we see that $\psi_i$ satisfies

$$\frac{d\psi_i}{dt} = 2t \mathbb{E} \left| \nabla f_t^{(i)} \right|^2.$$  

(66)

The right hand side of Eq. (66) can be bounded using Lemma 2.1: taking $g = f^{(i)}$ and $x = B_t$ in Eq. (21) and substituting this in Eq. (66), we have

$$\frac{d\psi_i}{dt} \leq 2t C (t) \mathbb{E} \left[ \left( f_t^{(i)} \right)^2 \log \frac{e}{\left( f_t^{(i)} \right)^2} \right].$$

Since $C (t)$ is continuous, there is an $L > 0$ such that $C (t) \leq L$ for all $t \in [0, 1/2]$, yielding

$$\frac{d\psi_i}{dt} \leq L \mathbb{E} \left[ \left( f_t^{(i)} \right)^2 \log \frac{e}{\left( f_t^{(i)} \right)^2} \right].$$

(Jensen’s inequality)

$$\leq L \mathbb{E} \left[ \left( f_t^{(i)} \right)^2 \right] \log \frac{e}{\mathbb{E} \left[ \left( f_t^{(i)} \right)^2 \right]}$$

$$= L \psi_i (t) \log \frac{e}{\psi_i (t)}.$$  

(67)
By invoking Lemma 3.7 with $C = L$ and $K = e$, there exists a time $t_0 \leq 1/2$ and a constant $\tilde{K}$ such that for all $t \in [0, t_0]$, 

$$\psi_i (t) \leq \tilde{K} \psi_i (0)^{1-Lt} = \tilde{K} \inf_i (f)^{2-2Lt},$$

where in the last equality we used Eq. (33) and the fact that $\psi_i (0) = (f_0^{(i)})^2$. Since $\psi_i (s) = \psi_i (e^{-s})$, if $\gamma \geq \log (1/t_0)$ then

$$\varphi_i (\gamma) \leq \tilde{K} \inf_i (f)^{2-2Le^{-\gamma}}$$

$$= \tilde{K} \inf_i (f)^{1+(1-2Le^{-\gamma})}$$

Set $c'_\gamma = 1 - 2Le^{-\gamma}$ and take $\gamma > \log 2L$ to ensure that $c'_\gamma > 0$. Taking $\gamma$ to also be larger than $\tilde{K}$ then gives the desired result: Eq. (65) follows because $\varphi_i (0) = E (f_1^{(i)})^2 = E (f^{(i)})^2 = \inf_i (f)$ by Eq. (19).

Using Eq. (65) together with the log-convexity from Lemma 3.6, for all $0 \leq s \leq \gamma$ we can bound $\varphi_i (s)$ by

$$\varphi_i (s) = \varphi_i \left( \left( 1 - \frac{s}{\gamma} \right) \cdot 0 + \frac{s}{\gamma} \cdot \gamma \right)$$

$$\leq \varphi_i (0) \left( 1 - s/\gamma \right) \varphi_i (\gamma)^{s/\gamma}$$

$$\leq \varphi_i (0) \left( 1 - s/\gamma \right) \left( \gamma \varphi_i (0) \left( 1+c'_\gamma \right)^{s/\gamma} \right)$$

$$\leq \gamma \inf_i (f)^{1+c'_\gamma s}$$

as needed, with $c_\gamma = c'_\gamma / \gamma = (1 - 2Le^{-\gamma}) / \gamma$. The lemma then follows by taking $\gamma$ large enough so that $\gamma \geq \log 2L$ and $c'_\gamma \geq 1/2$. $\square$

The following two lemmas are somewhat analogous to Lemmas 4.2 and 2.5.

**Lemma 5.2** Let $0 < \alpha \leq 1$. Then

$$\mathbb{E} [Q_\alpha] \leq 4\gamma^2 \rho (\alpha) T (f),$$

where $\gamma$ is the universal constant from Lemma 5.1 and $T (f)$ is as defined in (13).

**Lemma 5.3** If $0 < \alpha < 1/8$ and $\mathbb{P} [F_\alpha] < 2^{-14} \text{Var} (f)$, then

$$\mathbb{P} \left[ V_\alpha \geq \frac{1}{16} \right] > 2^{-14} \text{Var} (f).$$

(68)
The proof of Lemma 5.2 is very similar to that of Lemma 4.2, while the proof of Lemma 5.3 uses Lemma 2.5. Both proofs are found in Sect. 5.3.

5.1 Proof of Theorem 1.2

Proof of Theorem 1.2 Let \( \gamma \) be the constant from the statement of Lemma 5.1. By Lemma 5.2, for every \( 0 < \alpha \leq 1 \), we have
\[
\mathbb{E}[Q_\alpha] \leq 4\gamma^2 \rho(\alpha) T(f).
\]
Choosing \( \alpha = 1 \) just gives \( Q_1 = 2 \sum_{i=1}^{n} \int_{0}^{1} t (f_i^{(i)})^2 dt \), since the derivatives are bounded by 1; the expectation of this expression, as seen in (32), is larger than \( \text{Var}(f) \). We thus have
\[
\text{Var}(f) \leq \mathbb{E}[Q_1] \leq 4\gamma^2 \rho(1) T(f) = 8\gamma^2 \cdot T(f). \quad \square
\]

5.2 Proof of Theorem 1.5

Using Lemmas 5.2 and 5.3, we can obtain the following lemma.

Lemma 5.4 Let \( \gamma \) be the constant from Lemma 5.1, let \( F_\alpha \) be the event defined in Eq. (60), and assume that \( 0 < \alpha < 1/16 \) is small enough so that \( 4\gamma^2 \rho(\alpha) T(f) \leq 2^{-18} \text{Var}(f) \), where \( T \) is defined in (13). Then
\[
\mathbb{P}[F_\alpha] \geq \frac{1}{2^{14}} \text{Var}(f).
\]

Proof Since neither \( \partial_i f(x) \) nor \( f_i^{(i)}(x) \) depends on the \( i \)-th coordinate of \( x \), we may write
\[
\mathbb{E}\left[V^{(i)}_\alpha\right] = \mathbb{E}\sum_{t \in J_i \cap [0,1]} (2t \partial_i f_t)^2 1_{f_t^{(i)} < \alpha} = \mathbb{E}\sum_{t \in J_i \cap [0,1]} (2t \partial_i f_{t^-})^2 1_{f_{t^-}^{(i)} < \alpha}.
\]
Invoking Lemma 3.3 with \( g_t = (\partial_i f_{t^-})^2 1_{f_{t^-}^{(i)} < \alpha} \), we get
\[
\mathbb{E}\left[V^{(i)}_\alpha\right] = \mathbb{E}\sum_{t \in J_i \cap [0,1]} (2t \partial_i f_{t^-})^2 1_{f_{t^-}^{(i)} < \alpha}
\leq 2\mathbb{E}\int_{0}^{1} t (\partial_i f_{t^-})^2 1_{f_{t^-}^{(i)} < \alpha} dt
= 2\mathbb{E}\int_{0}^{1} t (\partial_i f_{t^-})^2 1_{f_{t^-}^{(i)} < \alpha} dt = \mathbb{E}\left[Q^{(i)}_\alpha\right].
\]
Using Lemma 5.2, this means that

\[ \mathbb{E} [V_\alpha] \leq 4\gamma^2 \rho (\alpha) T (f) . \]

Suppose by contradiction that \( \mathbb{P} [F_\alpha] < 2^{-14} \text{Var} (f) \). Then we can invoke Lemma 5.3, and together with Markov’s inequality this gives

\[ 4\gamma^2 \rho (\alpha) T (f) \geq \mathbb{E} [V_\alpha] \geq \mathbb{P} \left[ V_\alpha \geq \frac{1}{16} \right] > \frac{1}{218} \text{Var} (f) , \]

contradicting the assumption that \( 4\gamma^2 \rho (\alpha) T (f) \leq 2^{-18} \text{Var} (f) \).

The main assertion involved in proving Theorem 1.5 makes a connection between the vertex boundary and the probability that the function makes a large jump.

**Proposition 5.5** For \( 0 < \alpha \leq 1 \), let \( F_\alpha \) be the event defined in Eq. (60). Then

\[ \mu (\partial^{\pm} f) \geq \frac{1}{2} \alpha \mathbb{P} [F_\alpha] . \quad (69) \]

To prove this proposition, we will construct a modification \( \tilde{B}_t \) of \( B_t \), which can be thought of as a “hesitant” version of \( B_t \). For each coordinate \( i \), let \( \tilde{J}_i \) be the jump set of a Poisson point process on \((0, 1]\) with intensity \( 1/(2t) \), independent from \( B_t \) (and in particular, independent from the jump process \( J_i = \text{Jump} \left( B_t^{(i)} \right) \)). Define \( \tilde{B}_t = \left( \tilde{B}_t^{(1)}, \ldots, \tilde{B}_t^{(n)} \right) \) to be the process such that for every \( i \),

\[ \tilde{B}_t^{(i)} = \begin{cases} 0 & \text{if } t \in J_i \cup \tilde{J}_i \\ B_t^{(i)} & \text{otherwise.} \end{cases} \]

Loosely speaking, there are several ways of thinking about \( \tilde{B}_t^{(i)} \):

1. The process \( \tilde{B}_t^{(i)} \) can be seen as a “hesitant” variation of \( B_t^{(i)} \): it jumps with twice the rate (since its set of discontinuities is the union of two Poisson processes with rate \( 1/(2t) \)), but half of those times, it returns to the original sign rather than inverting it. We refer to this view as the “standard coupling” of \( \tilde{B}_t^{(i)} \) with \( B_t^{(i)} \): the process \( \tilde{B}_t^{(i)} \) is a copy of \( B_t^{(i)} \), but with additional independent hesitant jumps.

2. The process \( \tilde{B}_t^{(i)} \) is equal to 0 at a discrete set of times which follows the law of a Poisson point process with intensity \( 1/t \) (this is the union \( J_i \cup \tilde{J}_i \)); between two successive zeros it chooses randomly to be either \( t \) or \(-t\), each with probability \( 1/2 \).
Similarly to the notation using $B_t$, we write $\tilde{f}_t = f\left(\tilde{B}_t\right)$, and analogously $\partial_if, \nabla \tilde{f}_t$ and $f_t^{(i)}$.

**Lemma 5.6** The process $\tilde{B}_t$ is a martingale.

**Proof** Let $0 \leq s < t \leq 1$. For $s = 0$, since $\tilde{B}_0^{(i)} = 0$ always, we trivially have $\mathbb{E}\left[\tilde{B}_0^{(i)}\right] = 0$, so assume $s > 0$.

If $\tilde{B}_s^{(i)} = 0$, then $s \in J_i \cup \tilde{J}_i$. Being independent Poisson point processes, almost surely we have $J_i \cap \tilde{J}_i = \emptyset$, and $\mathbb{P}\left[s \in J_i \mid \tilde{B}_s^{(i)}\right] = \mathbb{P}\left[s \in \tilde{J}_i \mid \tilde{B}_s^{(i)}\right] = 1/2$. Since $\tilde{B}_t^{(i)} = B_t^{(i)}$ almost surely, we thus have that under the event $\tilde{B}_s^{(i)} = 0$,

$$\mathbb{E}\left[\tilde{B}_t^{(i)} \mid \tilde{B}_s^{(i)}\right] = \frac{1}{2} \mathbb{E}\left[B_t^{(i)} \mid B_s^{(i)}, s \in J_i\right] + \frac{1}{2} \mathbb{E}\left[B_t^{(i)} \mid B_s^{(i)}, s \in \tilde{J}_i\right].$$

It is evident by the definition of the process $B_t$ that when $\tilde{B}_s^{(i)} = 0$,

$$\mathbb{E}\left[B_t^{(i)} \mid B_s^{(i)}, s \in J_i\right] = -\mathbb{E}\left[B_t^{(i)} \mid B_s^{(i)}, s \in \tilde{J}_i\right],$$

so that $\mathbb{E}\left[\tilde{B}_t^{(i)} \mid \tilde{B}_s^{(i)}\right] = 0 = \tilde{B}_s^{(i)}$.

Finally, under the event $\tilde{B}_s^{(i)} \neq 0$, since $\tilde{B}_t^{(i)} \neq 0$ almost surely, we have by (25) that

$$\mathbb{P}\left[\text{sign} \tilde{B}_t^{(i)} \neq \text{sign} \tilde{B}_s^{(i)} \mid \tilde{B}_s^{(i)}\right] = \frac{t - s}{2t}.$$ 

Thus

$$\mathbb{E}\left[\tilde{B}_t^{(i)} \mid \tilde{B}_s^{(i)}\right] = \text{sign} \tilde{B}_s^{(i)} \cdot t \cdot \frac{t - s}{2t} + \text{sign} \tilde{B}_s^{(i)} \cdot (-t) \cdot \frac{t - s}{2t}$$

$$= \text{sign} \tilde{B}_s^{(i)} \cdot s$$

$$= \tilde{B}_s^{(i)}.$$

$\square$

**Proof of Proposition 5.5** In order to distinguish between the vertex boundaries, we will use the hesitant jump process $\tilde{f}_t$ defined above. We prove (69) for the inner vertex boundary $\partial^+$; the proof for $\partial^-$ follows by considering $-f$. Let $\tau = \inf \left\{ t > 0 \mid \exists i \in [n] \text{ s.t. } \tilde{B}_t^{(i)} = 0 \right\} \wedge 1$. Note that for any $t_0 > 0$, we almost surely have that $\tilde{B}_t^{(i)} = 0$ only finitely many times.
for $t \in [t_0, 1]$. Thus, if $0 < \tau < 1$, then the infimum in the definition of $\tau$ is attained as a minimum, and there exists an $i_0$ such that $\tilde{B}^{(i_0)}_{\tau} = 0$ and $f^{(i_0)}_{\tau} \geq \alpha$. In fact, this holds true if $\tau = 0$ as well: in this case, there is a sequence of times $t_k \to 0$ and indices $i_k$ such that $f^{(i_k)}_{t_k} \geq \alpha$ and $\tilde{B}^{(i_k)}_{t_k} = 0$. Since there are only finitely many different indices $i_k$, there is a subsequence $k_{\ell}$ so that $i_{k_{\ell}}$ are all the same index $i_0$, and the claim follows by continuity of $f^{(i_0)}$ and the fact that $\tilde{B}_0 = 0$.

When $F_\alpha$ occurs, we necessarily have $\tau < 1$, since $\tilde{B}^{(i)}_t = 0$ whenever $B^{(i)}_t$ is discontinuous. Since $\tilde{B}_1$ is uniform on the hypercube,

$$\mu(\partial^+ f) = \mathbb{P}[\tilde{B}_1 \in \partial^+ f] \geq \mathbb{P}[\tilde{B}_1 \in \partial^+ f \mid \tau < 1] \mathbb{P}[\tau < 1] \geq \mathbb{P}[\tilde{B}_1 \in \partial^+ f \mid \tau < 1] \mathbb{P}[F_\alpha],$$

and so it suffices to show that

$$\mathbb{P}[\tilde{B}_1 \in \partial^+ f \mid \tau < 1] \geq \frac{1}{2} \alpha. \quad (70)$$

Supposing that $\tau < 1$, denote by $i_0$ a coordinate for which $\tilde{B}^{(i_0)}_{\tau} = 0$ and $f^{(i_0)}_{\tau} \geq \alpha$. If $\tau = 0$ then $\tilde{B}_\tau = B_\tau$; otherwise, almost surely $i_0$ is the only coordinate of $\tilde{B}_\tau$ which is 0, and so $\tilde{B}^{(j)}_{\tau} = B^{(j)}_{\tau}$ for all $j \neq i_0$ almost surely. Since the function $f^{(i)}(x)$ does not depend on the $i$-th coordinate of $x$, we deduce that $\tilde{f}^{(i_0)}_{\tau} = f^{(i_0)}_{\tau} \geq \alpha$ almost surely. Thus, under $\tau < 1$, by the martingale property of $f^{(i_0)}_{\tau}$, we have

$$\mathbb{P}[\tilde{f}^{(i_0)}_{1} = 1 \mid \tilde{B}_\tau] \geq \alpha. \quad (71)$$

Similarly, using the martingale property of $\tilde{B}^{(i_0)}_{\tau}$, we have $\mathbb{E}[\tilde{B}^{(i_0)}_{\tau} \mid \tilde{B}_\tau] = \tilde{B}^{(i_0)}_{\tau} = 0$, and so

$$\mathbb{P}[\tilde{B}^{(i_0)}_{1} = 1 \mid \tilde{B}_\tau] = \mathbb{P}[\tilde{B}^{(i_0)}_{1} = -1 \mid \tilde{B}_\tau] = \frac{1}{2}.$$

Since $\partial i_0 \tilde{f}_t$ is independent of $\tilde{B}^{(i_0)}_{\tau}$, we finally obtain

$$\mathbb{P}[\tilde{B}_1 \in \partial^+ f \mid \tilde{B}_\tau] = \mathbb{P}[\tilde{B}^{(i_0)}_{1} = 1 \wedge \partial i_0 \tilde{f}_1 = 1 \mid \tilde{B}_\tau] + \mathbb{P}[\tilde{B}^{(i_0)}_{1} = -1 \wedge \partial i_0 \tilde{f}_1 = -1 \mid \tilde{B}_\tau]$$

$$= \mathbb{P}[\tilde{B}^{(i_0)}_{1} = 1 \mid \tilde{B}_\tau] = \frac{1}{2} \mathbb{P}[\partial i_0 \tilde{f}_1 = 1 \mid \tilde{B}_\tau] + \frac{1}{2} \mathbb{P}[\partial i_0 \tilde{f}_1 = -1 \mid \tilde{B}_\tau].$$
\[
\mathbb{P}\left[ f_1^{(i_0)} = 1 \right] = \frac{1}{2}\mathbb{P}\left[ \tilde{B}_\tau \right] \geq \frac{1}{2}\alpha.
\]

(71)

\[\mu\left( \delta^\pm f \right) \geq \frac{1}{2}\alpha\mathbb{P}\left[ F_\alpha \right] \geq \frac{1}{2^{14}}\alpha\text{Var}(f).\] (72)

**Proof of Theorem 1.5** Let $\gamma$ be the constant from the statement of Lemma 5.1. Let $\alpha$ be such that $\alpha \log \frac{1}{\alpha} = \frac{1}{2^{30}\gamma^2}r_{Tal}$. Then the condition $4\gamma^2 \rho(\alpha) T(f) \leq 2^{-18}\text{Var}(f)$ is satisfied in Lemma 5.4, implying that $\mathbb{P}[F_\alpha] \geq 2^{-14}\text{Var}(f)$. Together with Proposition 5.5, we have

All that remains is to obtain a lower bound on $\alpha$. To this end, observe that $\alpha \log \frac{1}{\alpha} \leq \sqrt{\alpha}$ for all $\alpha \in [0, 1]$, and so $\alpha \geq \frac{1}{2^{30}\gamma^2}r_{Tal}$. Thus $\log \frac{1}{\alpha} \leq \log \left( \frac{2^{30}\gamma^2}{r_{Tal}^2} \right) \leq \log \left( \frac{2^{30}\gamma^4(1+C_T)^2}{r_{Tal}^2} \right)$, where $C_T$ is the constant in Theorem 1.2. Thus there exists a constant $C'_B$, which can be taken to be larger than $C_T$, such that

\[
\alpha = \frac{r_{Tal}}{2^{30}\gamma^2 \log \frac{1}{\alpha}} \geq \frac{r_{Tal}}{C'_B \log \frac{C'_B}{r_{Tal}}}.
\]

Plugging this into (72) gives the desired result. \[\square\]

### 5.3 Postponed proofs

**Proof of Lemma 5.2** As defined in (61), since $Q_\alpha = \sum_{i=1}^{n} Q_\alpha^{(i)}$, it is enough to show that

\[
\mathbb{E}\left[ Q_\alpha^{(i)} \right] \leq 4\gamma^2 \rho(\alpha) \frac{\text{Inf}_i(f)}{1 + \log(1/\text{Inf}_i(f))}.
\]

Denoting $\tilde{\varphi}_i(s) = \mathbb{E}\left[ f_i^{(i)} e^{-s} \mathbb{1}_{f_i^{(i)} < \alpha} \right]$, by a change of variables we get

\[
\mathbb{E}\left[ Q_\alpha^{(i)} \right] \leq 2 \int_0^\infty e^{-2s} \tilde{\varphi}_i(s)\, ds.
\] (73)
Let $\tau = \frac{\log(1/\alpha)}{2 + \frac{1}{2\gamma} \log(1/\text{Inf}_i(f))}$. Assume first that $\tau \leq \frac{1}{2} \gamma$. The integral in Eq. (73) then splits up into three parts:

$$
\mathbb{E} \left[ Q^{(i)}_\alpha \right] \leq 2 \int_0^\tau e^{-2s} \tilde{\varphi}_i(s) \, ds + 2 \int_\tau^\gamma e^{-2s} \tilde{\varphi}_i(s) \, ds + 2 \int_\gamma^\infty e^{-2s} \tilde{\varphi}_i(s) \, ds.
$$

For the first integral on the right hand side, we write

$$
\tilde{\varphi}_i(s) \leq \alpha \mathbb{E} \left| f^{(i)}_{e^{-s}} \right| = \alpha \mathbb{E} f^{(i)}_{e^{-s}} = \alpha f^{(i)}_0 = \alpha \text{Inf}_i(f).
$$

Thus

$$
\int_0^\tau e^{-2s} \tilde{\varphi}_i(s) \, ds \leq \alpha \tau \text{Inf}_i(f) \leq 2\gamma \frac{\text{Inf}_i(f)}{4\gamma + \log(1/\text{Inf}_i(f))} - \alpha \log(1/\alpha) \quad \text{by choice of } \tau \quad (\gamma > 1)
$$

$$
\leq 2\gamma \frac{\text{Inf}_i(f)}{1 + \log(1/\text{Inf}_i(f))} - \alpha \log(1/\alpha). \quad (75)
$$

For the second and third integrals, we use the fact that trivially, $\varphi_i(s) \leq \text{Inf}_i(f)$ for all $s$. By Lemma 5.1, for $s \in [\tau, \gamma]$ we then have $\tilde{\varphi}_i(s) \leq \gamma \text{Inf}_i(f)^{1+s/2\gamma}$. The second integral is therefore bounded by

$$
\int_\tau^\gamma e^{-2s} \tilde{\varphi}_i(s) \, ds \leq \gamma \int_\tau^\gamma e^{-2s} \text{Inf}_i(f)^{1+s/(2\gamma)} \, ds
$$

$$
\leq \gamma \int_\tau^\infty e^{-2s} \text{Inf}_i(f)^{1+s/(2\gamma)} \, ds
$$

$$
= \gamma \text{Inf}_i(f) \int_\tau^\infty e^{s\left(\frac{1}{2\gamma} \log\text{Inf}_i(f) - 2\right)} \, ds
$$

$$
\leq \gamma \frac{\text{Inf}_i(f)}{2 + \frac{1}{2\gamma} \log(1/\text{Inf}_i(f))} e^{\tau\left(\frac{1}{2\gamma} \log\text{Inf}_i(f) - 2\right)}
$$

$$
\leq 2\gamma^2 \frac{\text{Inf}_i(f)}{1 + \log(1/\text{Inf}_i(f))} \alpha. \quad (76)
$$

For the third integral, since $\gamma > 1$ and $\tau \leq \frac{1}{2} \gamma$, we have $\int_\gamma^\infty e^{-2s} \, ds \leq \int_{\frac{1}{2} \gamma}^\gamma e^{-2s} \, ds$. Using the fact that $\varphi_i(s)$ is a decreasing function in $s$ (as can be
Concentration on the Boolean hypercube

seen from Eq. (35)), we immediately get

$$\int_{\gamma}^{\infty} e^{-2s \tilde{\varphi}_i(s)} ds \leq \int_{\tau}^{\gamma} e^{-2s \tilde{\varphi}_i(s)} ds. \tag{77}$$

Putting the bounds (75), (76) and (77) together, when $\tau < \frac{1}{2} \gamma$ we get that

$$\mathbb{E} \left[ Q^{(i)}_{\alpha} \right] \leq 2 \left( 2\gamma \frac{\text{Inf}_i(f)}{1 + \log(1/\text{Inf}_i(f))} \alpha \log(1/\alpha) 
+ (2 + 2) \gamma^2 \frac{\text{Inf}_i(f)}{1 + \log(1/\text{Inf}_i(f))} \frac{\text{Inf}_i(f)}{\text{Inf}_i(f)} \right)
\leq 4\gamma^2 \rho(\alpha) \frac{\text{Inf}_i(f)}{1 + \log(1/\text{Inf}_i(f))}.$$

Now assume that $\tau \geq \frac{1}{2} \gamma$. The integral in Eq. (73) then splits up into two parts:

$$\mathbb{E} \left[ Q^{(i)}_{\alpha} \right] \leq 2 \int_{0}^{\tau} e^{-2s \tilde{\varphi}_i(s)} ds + 2 \int_{\tau}^{\infty} e^{-2s \tilde{\varphi}_i(s)} ds.$$

Again, since $\varphi_i(s)$ is decreasing as a function of $s$ and since $\tau \geq \frac{1}{2} \gamma > \frac{1}{2}$, the second integral is smaller than the first, and so by (75),

$$\mathbb{E} \left[ Q^{(i)}_{\alpha} \right] \leq 2 \cdot 2\gamma^2 \frac{\text{Inf}_i(f)}{1 + \log(1/\text{Inf}_i(f))} \alpha \log(1/\alpha)
\leq 4\gamma^2 \rho(\alpha) \frac{\text{Inf}_i(f)}{1 + \log(1/\text{Inf}_i(f))}$$

in this case as well. \hfill \Box

**Proof of Lemma 5.3** Assume without loss of generality that $f_0 = \mathbb{E}_f f \leq 0$ (if not, use $-f$ instead of $f$; the variances and the probability $\mathbb{P}[V_{\alpha} \geq x]$ are the same for both functions). Let $\tau = \inf \{0 \leq t \leq 1 \mid f_t \in (0, 2\alpha)\} \land 1$. By conditioning on the event $\{\tau < 1\}$, we have

$$\mathbb{P}[\lfloor f \rfloor_1 \geq 1/16] \geq \mathbb{P}[\lfloor f \rfloor_1 \geq 1/16 \mid \tau < 1] \mathbb{P}[\tau < 1]. \tag{78}$$

We start by bounding the probability $\mathbb{P}[\tau < 1]$. Let $A = \{\exists t \in [0, 1] \text{ s.t. } f_t > 0\}$, and observe that $\{\tau < 1\} \subseteq A$. Under the event $A \setminus \{\tau < 1\}$, the process $f_t$ never visited the interval $(0, 2\alpha)$ and yet at some point reached a value larger than 0, and so necessarily had a jump discontinuity of size at least $2\alpha$. But a jump occurring at time $t$ due to a discontinuity in $B_t^{(i)}$ is of
size $2t \mid \partial_i f_t \mid$, and so $2t \mid \partial_i f_t \mid \geq 2\alpha$, implying that $f_t^{(i)} \geq \mid \partial_i f_t \mid \geq \alpha$. Thus, $A \cap \{ \tau = 1 \} \subseteq A \cap F_{\alpha}$, and so $A \cap F_{\alpha}^C \subseteq A \cap \{ \tau < 1 \} = \{ \tau < 1 \}$. Hence

$$P[\tau < 1] \geq P[A \setminus F_{\alpha}] \geq P[A] - P[F_{\alpha}].$$

To bound $P[A]$, note that $\{ f_1 = 1 \} \subseteq A$. By the martingale property of $f_t$,

$$f_0 = \mathbb{E} f_1 = 2P[f_1 = 1] - 1,$$

and so

$$P[A] \geq P[f_1 = 1] = \frac{1 + f_0}{2} = \frac{1 - f_0^2}{2(1 - f_0)} = \frac{\text{Var}(f)}{2(1 - f_0)} \geq \frac{1}{4} \text{Var}(f).$$

Putting this together with the assumption that $P[F_{\alpha}] < 2^{-14} \text{Var}(f)$ gives

$$P[\tau < 1] \geq \frac{1}{8} \text{Var}(f). \quad (79)$$

Next we bound the probability $P\left[ [f]_1 \geq 1/16 \mid \tau < 1 \right]$, by relating the quadratic variation to the variance of $f_t$.

Denote $X_t = f_{t+\tau(1-t)}$. Conditioned on the event $\tau < 1$, $X_t$ is still a martingale, with $X_0 \in (0, 2\alpha) \subseteq [-1/2, 1/2]$, $X_1 \in (-1, 1)$, and $X_t \in [-1, 1]$ for all $t \in [0, 1]$. We can therefore invoke Lemma 2.5 with $a = 1/2$ and $c_1 = 1/4$, yielding

$$P\left[ [f]_1 \geq 1/16 \mid \tau < 1 \right] \geq \frac{1}{2^{10}}.$$

Together with (78) and (79), this gives

$$P\left[ [f]_1 \geq x \right] \geq \frac{1}{2^{13}} \text{Var}(f).$$

Now, by the definitions of $F_{\alpha}$ (60) and $V_{\alpha}$ (62), under the event $F_{\alpha}^C$, we have $V_{\alpha} = [f]_1$, i.e. all of the quadratic variation comes from small jumps. Thus

$$P\left[ V_{\alpha} \geq \frac{1}{16} \right] \geq P\left[ \left\{ [f]_1 \geq \frac{1}{16} \right\} \cap F_{\alpha}^C \right] \geq P\left[ [f]_1 \geq \frac{1}{16} \right] - P[F_{\alpha}]$$
Concentration on the Boolean hypercube

\[ \geq \frac{1}{214} \text{Var} (f). \]

6 Proof of Theorem 1.9

As explained above in Eq. (29), our goal is to show that

\[ S_{\varepsilon} (f) = \mathbb{E} \sum_{i=1}^{n} \int_{0}^{\sqrt{1-\varepsilon}} t (\partial_{i} f)^{2} \, dt \leq C \text{Var} (f) \left( \sum_{i=1}^{n} \text{Inf}_{i} (f)^{2} \right)^{CE}. \] (80)

We first show that we may assume that \( f \) is monotone. For an index \( i = 1, \ldots, n \), define an operator \( \kappa_{i} \) by

\[(\kappa_{i} f) (y) := \begin{cases} \max \{ f (y), f (y^{\oplus i}) \} & y_{i} = 1, \\ \min \{ f (y), f (y^{\oplus i}) \} & y_{i} = 0. \end{cases}\]

The following lemma relates between the influences and sensitivities of \( \kappa_{i} f \) and \( f \):

**Lemma 6.1** [2, Lemma 2.7] \( \kappa_{1} \kappa_{2} \ldots \kappa_{n} f \) is monotone, and for every pair of indices \( i, j \), \( \text{Inf}_{i} (\kappa_{j} f) \leq \text{Inf}_{i} (f) \) and \( S_{\varepsilon} (\kappa_{i} f) \geq S_{\varepsilon} (f) \).

Thus, if Eq. (80) holds for \( \tilde{f} = \kappa_{1} \ldots \kappa_{n} f \), then it holds for \( f \) as well, since \( S_{\varepsilon} (f) \leq S_{\varepsilon} (\tilde{f}) \) and \( \sum_{i=1}^{n} \text{Inf}_{i} (f)^{2} \geq \sum_{i=1}^{n} \text{Inf}_{i} (\tilde{f})^{2} \). So it is enough to verify (80) for monotone functions.

In order to prove (80), we may also assume that for any fixed universal constant \( K < 1 \),

\[ \sum_{i=1}^{n} \text{Inf}_{i} (f)^{2} \leq K. \] (81)

For if \( \sum_{i=1}^{n} \text{Inf}_{i} (f)^{2} \geq K \) for some \( K \), then since \( f \) is monotone,

\[ \text{Var} (f) = \sum_{S \subseteq [n], S \neq \emptyset} \hat{f} (S)^{2} \geq \sum_{i=1}^{n} \hat{f} (\{i\})^{2} \overset{20}{=} \sum_{i=1}^{n} \text{Inf}_{i} (f)^{2} \geq K, \]

and so \( \text{Var} \cdot (\sum_{i} \text{Inf}_{i} (f)^{2}) \geq K^{2} \). Equation (80) then holds trivially with \( C = 1/K^{2} \) and \( c = 1 \), since \( S_{\varepsilon} \leq 1 \) for all \( \varepsilon \).

Similarly, we may assume that for any fixed, universal constant \( K \),

\[ \text{Var} (f) \leq K; \] (82)

\[ \square \] Springer
otherwise Theorem 1.9 would be equivalent (up to constants) to the original theorem proved in [22].

**Remark 6.2** Our proof actually recovers the original theorem proved in [22], but we make the assumption \( \text{Var}(f) \leq K \) since it simplifies some bounds.

Define

\[
R(t) := \mathbb{E} \sum_i (\partial_i f_i)^2 = \mathbb{E} \| \nabla f_t \|_2^2.
\]

At time 0, we have

\[
R(0) = \sum_{i=1}^n (\partial_i f_0)^2 = \sum_{i=1}^n (\hat{\partial}_i f(\emptyset))^2 = \sum_{i=1}^n \hat{f}(\{x_i\})^2 = \sum_{i=1}^n \text{Inf}_i(f)^2.
\]

The function \( R(t) \) is monotone in \( t \): since \( \partial_i f_t \) is a martingale, \( (\partial_i f_t)^2 \) is a submartingale and so \( \mathbb{E} (\partial_i f_t)^2 \) is increasing.

By invoking Corollary 3.5 on \( \partial_i f \), for every index \( i \) we have

\[
\frac{d}{dt} \mathbb{E} (\partial_i f_t)^2 = 2t \mathbb{E} \sum_{j=1}^n (\partial_j \partial_i f_t)^2.
\]

Thus

\[
\frac{d}{dt} R(t) = 2t \mathbb{E} \sum_{i=1}^n \sum_{j=1}^n (\partial_i \partial_j f_t)^2 (t \leq 1) \leq 2 \mathbb{E} \| \nabla^2 f_t \|_{HS}^2,
\]

where \( \| X \|_{HS} \) is the Hilbert–Schmidt norm of a matrix. By Lemma 2.2, there exists a continuous positive function \( C(t) \) such that

\[
\frac{d}{dt} R(t) \leq 2C(t) \mathbb{E} \left[ \| \nabla f_t \|_2^2 \log \frac{C(t)}{\| \nabla f_t \|_2^2} \right] \leq 2C(t) \mathbb{E} \left[ \| \nabla f_t \|_2^2 \right] \log \frac{C(t)}{\mathbb{E} \left[ \| \nabla f_t \|_2^2 \right]} = 2C(t) R(t) \log \left( \frac{C(t)}{R(t)} \right).
\]

Since \( C(t) \) is continuous, it is bounded in \([0, 1/2]\), so there exists a constant \( c > 0 \), independent of \( R \), such that for all \( t \in [0, 1/2] \),

\[
\frac{d}{dt} R(t) \leq c \cdot R(t) \log \left( \frac{c}{R(t)} \right).
\]
By (81), we can assume that \( R (0) \leq c/2 \). Using (85) together with Lemma 3.7, there exist constants \( C, L > 0 \) and a time \( t_0 \), all of which depend only on \( c \) (and so are independent of \( R \)), such that for all \( t \in [0, t_0] \),

\[
R (t) \leq L \cdot R (0) e^{-Ct}.
\]

In particular, there exists a constant \( K > 0 \) such that

\[
R (e^{-K}) \leq L \cdot R (0)^{5/6},
\]

and since \( R \) is increasing, we can always assume that \( K > 1 \). Denote

\[
G (s) := R (e^{-s}).
\]

By Lemma 3.6, every summand of the form \( \mathbb{E} \left( \partial_i f_{-s} \right)^2 \) is log-convex, and so \( G (s) \) itself, being a positive linear combination of log-convex functions, is also log-convex in \( s \) [6, Section 3.5.2].

**Lemma 6.3** Let \( K \geq 1 \) and let \( G (s) \) be a log-convex decreasing function. Denote \( v = \int_0^K e^{-2s} G (s) \, ds \) and assume that \( v > G (K) \). Then for all \( r < K \),

\[
\int_0^r e^{-2s} G (s) \, ds \geq v \cdot \left( 1 - \left( \frac{G (K)}{v} \right)^{r/K} \right).
\]

**Proof** Consider the function

\[
h_{\ell} (s) = \frac{v \ell}{1 - e^{-\ell K} e^{-(\ell-2)s}},
\]

where \( \ell \) is the largest solution to the equation \( h_{\ell} (K) = G (K) \). By choice of \( h_{\ell} \), we have

\[
\int_0^K e^{-2s} h_{\ell} (s) \, ds = v = \int_0^K e^{-2s} G (s) \, ds.
\]

Since \( e^{-2s} h_{\ell} (s) \) is log-linear on \([0, K]\), \( e^{-2s} G (s) \) is log-convex on \([0, K]\), they have the same integral on \([0, K]\) and \( G (K) = h_{\ell} (K) \), we must have one of two cases:

1. \( h_{\ell} (s) = G (s) \)
2. The functions intersect at most once in the interval \([0, K]\) at some point \( s_0 \) such that \( G (s) \geq h_{\ell} (s) \) for all \( s < s_0 \).
In either case, for all \( r \in [0, K] \), we have
\[
\int_0^r e^{-2s} G(s) \, ds \geq \int_0^r e^{-2s} h_{\ell}(s) \, ds = \frac{v}{1 - e^{-K\ell}} \left( 1 - e^{-r\ell} \right) \geq v \left( 1 - e^{-r\ell} \right). \tag{88}
\]

On the other hand, we chose \( \ell \) to be such that \( h_{\ell}(K) = G(K) \), and so
\[
\frac{\ell}{1 - e^{-\ell K}} e^{-(\ell-2)K} = \frac{h_{\ell}(K)}{v} = \frac{G(K)}{v} < 1,
\]
where the last inequality is by assumption on \( v \). The function \( \frac{x}{1 - e^{-xK}} e^{-(x-2)K} \) is decreasing as a function of \( x \) in the interval \([2, \infty)\), but is greater than 1 at \( x = 2 \); hence, since \( \ell \) is the largest number for which \( h_{\ell}(K) = G(K) \), we must have \( \ell > 2 \). We then have
\[
\ell e^{-(\ell-2)K} = \frac{G(K)}{v} \left( 1 - e^{-\ell K} \right) \leq \frac{G(K)}{v},
\]
and after rearranging, since \( \ell > 2 \),
\[
e^{-\ell K} \leq \frac{G(K) e^{-2K}}{v} \leq \frac{G(K)}{\ell v}.
\]
Thus
\[
\ell \geq \frac{1}{K} \log \frac{v}{G(K)}.
\]
Putting this into the right hand side of (88) gives
\[
\int_0^r e^{-2s} G(s) \, ds \geq v \left( 1 - e^{-r\ell} \right) \geq v \left( 1 - e^{-K \log \frac{v}{G(K)}} \right) = v \left( 1 - \left( \frac{G(K)}{v} \right)^{\frac{v}{G(K)}} \right).
\]

Proof of Theorem 1.9 By Corollary 3.4,
\[
\text{Var}(f) = \text{Var}(f_1) = 2\mathbb{E} \sum_{i=1}^n \int_0^1 t (\partial_i f_i)^2 \, dt \overset{(83)}{=} 2 \int_0^1 t \cdot R(t) \, dt,
\]
\[\text{Springer}\]
and by a change of variables this becomes

\[
\text{Var} (f) \stackrel{(87)}{=} 2 \int_0^\infty e^{-2s} G (s) \, ds.
\]

Define \( v = \int_0^K e^{-2s} G (s) \, ds \), where \( K \) is the constant from Eq. (86). Note that since \( K \geq 1 \) and \( G \) is decreasing,

\[
\text{Var} (f) - v = \int_K^\infty e^{-2s} G (s) \, ds \leq G (K) \stackrel{(86)}{\leq} L \cdot R (0)^{5/6}. \tag{89}
\]

Rearranging, this gives

\[
\frac{v}{G (K)} \geq \frac{\text{Var} (f) - LR (0)^{5/6}}{LR (0)^{5/6}}. \tag{90}
\]

Set \( g(x) = \frac{1 + f(x)}{2} \), and assume without loss of generality that \( \mathbb{E} f = f(0) \leq 0 \), so that \( \mathbb{E} g = g(0) \leq \frac{1}{2} \) (if \( f(0) > 0 \), we can take \( g(x) = (1 - f(x)) / 2 \)). This implies that

\[
\text{Var} (g) = g(0) (1 - g(0)) \geq \frac{1}{2} g(0).
\]

Invoking Lemma 2.1 with \( g \) and \( t = 0 \), there exists a constant \( C \) such that

\[
R (0) = \mathbb{E} \sum_{i=1}^n (\partial_i f (0))^2 = \| \nabla f (0) \|_2^2 = 4 \| \nabla g (0) \|_2^2 \leq C g(0)^2 \log \frac{e}{g(0)}
\]

\[
\leq C' \left( \text{Var} (g) \right)^2 \log \frac{C'}{\text{Var} (g)} \leq C'' \text{Var} (f)^2 \log \frac{4C''}{\text{Var} (f)}. \tag{91}
\]

By (82), we can assume that \( \text{Var} (f) \) is small enough so (91) implies

\[
R (0)^{2/3} \leq \text{Var} (f). \tag{92}
\]

Plugging this into (90), we get

\[
\frac{v}{G (K)} \geq \frac{R (0)^{2/3} - LR (0)^{5/6}}{LR (0)^{5/6}}.
\]
By (81), we can assume that \( R(0) \) is smaller than any fixed constant. Since \( L \) and \( K \) do not depend on \( R \), we can choose small enough \( R(0) \) so that \( LR(0)^{5/6} \leq \frac{1}{2} R(0)^{2/3} \), and so

\[
\frac{v}{G(K)} \geq \frac{1}{2L} R(0)^{-1/6}.
\] (93)

For small enough \( R(0) \), the above display implies that \( v > G(K) \), and so we can use Lemma 6.3. Together with Eqs. (89) and (90), we have

\[
\int_0^r e^{-2s} G(s) ds \geq v \left( 1 - \frac{G(K)}{v} \right)^{r/K} \geq \left( \text{Var}(f) - LR(0)^{5/6} \right) \left( 1 - (2L \cdot R(0)^{1/6})^{r/K} \right).
\] (94)

This allows us to prove (80):

\[
S_\varepsilon(f) = \mathbb{E} \sum_{i=1}^n \int_0^{\sqrt{1-\varepsilon}} t (\partial_i f_i)^2 dt = \int_0^{\sqrt{1-\varepsilon}} t R(t) dt \leq \int_0^{e^{-\varepsilon/2}} t R(t) dt = \text{Var}(f) - \int_0^{-\varepsilon/2} e^{-2s} G(s) ds \leq LR(0)^{5/6} + \text{Var}(f) \left( 2L \cdot R(0)^{1/6} \right)^{\varepsilon/(2K)} \leq L \text{Var}(f) R(0)^{1/6} + \text{Var}(f) \left( 2L \cdot R(0)^{1/6} \right)^{\varepsilon/(2K)} \leq C \cdot \text{Var}(f) R(0)^{\varepsilon/(12K)}
\]

for some universal constant \( C \).

\[\square\]

Acknowledgements The first author would like to thank Noam Lifshitz for useful discussions and in particular for pointing out the possible application to stability of the isoperimetric inequality. We are also thankful to Ramon Van Handel, Itai Benjamini and Gil Kalai for an enlightening discussion, and to Gregory Rosenthal for his comments. We are grateful to the anonymous referees, whose insightful comments have greatly improved the paper. We thank Mark Sellke for a simplification of the proofs of Lemmas 2.5 and 5.3. R.E. is an incumbent of the Elaine Blond Career Development Chair, and is supported by a European Research Council.
Starting Grant (ERC StG) and by an Israel Science Foundation Grant No. 718/19. R.G. is supported by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities, the European Research Council and by the Israeli Science Foundation.

Appendix A: \( p \)-Biased analysis

For \( p = (p_1, \ldots, p_n) \in [0, 1]^n \), let \( \mu_p \) be the measure

\[
\mu_p (y) = \prod_{i=1}^{n} \frac{1 + y_i (2 p_i - 1)}{2} =: w_{(2p-1)} (y) ,
\]

which sets the \( i \)-th bit to 1 with probability \( p_i \). Let

\[
\omega_i (y) = \frac{1}{2} \left( \frac{1 - 2p_i}{\sqrt{p_i (1 - p_i)}} + y_i \frac{1}{\sqrt{p_i (1 - p_i)}} \right),
\]

and for a set \( S \subseteq [n] \), define \( \omega_S (y) = \prod_{i \in S} \omega_i (y) \). Then every function \( f \) can be written as

\[
f (y) = \sum_{S \subseteq [n]} \hat{f}_p (S) \omega_S (y)
\]

\[
= \sum_{S \subseteq [n]} \left( \mathbb{E}_{\mu_p} [f \cdot \omega_S] \right) \omega_S (y) \tag{96}
\]

\[
= \sum_{S \subseteq [n]} \left( \sum_{y \in \{-1,1\}^n} f (y) \omega_S (y) w_{2p-1} (y) \right) \omega_S (y) \tag{97}
\]

The coefficients \( \hat{f}_p (S) := \mathbb{E}_{\mu_p} [f \cdot \omega_S] \) are called the “\( p \)-biased” Fourier coefficients of \( f \).

The \( p \)-biased influence of the \( i \)-th bit is

\[
\text{Inf}_i^p (f) = 4 p_i (1 - p_i) \mathbb{P}_{y \sim \mu_p} \left[ f (x) \neq f (x^{\oplus i}) \right].
\]

If \( f \) is monotone, then

\[
\text{Inf}_i^p (f) = 2 \sqrt{p_i \sqrt{1 - p_i}} \hat{f}_p ([i]) .
\]

The \( p \)-biased Fourier coefficients are related to the derivatives of \( f \) by the following proposition, whose proof (using slightly different notation) can be found in [28, Section 8].
Proposition A.1 Let $S = \{i_1, \ldots, i_k\} \subseteq [n]$ be a set of indices, $x \in (-1, 1)^n$, and $p = \frac{1+x}{2}$. Then
\[
\partial_{i_1} \ldots \partial_{i_k} f (x) = \left( \prod_{i \in S} \frac{4}{\sqrt{1 - x_i^2}} \right) \hat{f}_p (S).
\]

Appendix B: Postponed proofs

Proof of Lemma 2.1 We prove this lemma by using the analogous Level-1 inequality for Gaussian sets, which states that among all sets with a fixed Gaussian measure, the center of mass of the Gaussian half-plane is the furthest from the origin. More formally, let $\gamma_n$ denote the $n$-dimensional Gaussian measure on $\mathbb{R}^n$, and let $A \subseteq \mathbb{R}^n$. Then
\[
\left\| \int_A z \gamma_n (dz) \right\|_2 \leq \int_{\Phi^{-1}_n(A)} t \gamma_1 (dt),
\]
where $\Phi (s) = (2\pi)^{-1/2} \int_{-\infty}^{s} e^{-z^2/2} dz$ is the one-dimensional Gaussian CDF. For a proof, see [13, Fact 10]. Through a technical calculation, it can be shown [11, Lemma 21] that there exists a constant $C > 0$ such that for all $s \in (0, 1)$,
\[
\int_{\Phi^{-1}_n(s)} t \gamma_1 (dt) \leq C s \sqrt{\log \frac{e}{s}}.
\]
We now relate the derivative $\nabla g (x)$ to integrals of the form $\int_A z \gamma_n (dz)$. Let $p = \frac{1+x}{2}$, let $\mu_p$ be the $p$-biased measure on $\{-1, 1\}^n$, and let $\alpha \in \mathbb{R}^n$ be defined by $\alpha_i = \Phi^{-1} (1 - p_i)$. This means that $1 - p_i = \Phi (\alpha_i) = (2\pi)^{-1/2} \int_{-\infty}^{\alpha_i} e^{-x^2/2} dx$, and for every function $f : \{-1, 1\}^n \to \mathbb{R}$,
\[
\int_{\mathbb{R}^n} f (\text{sign}_n (z + \alpha)) d \gamma_n (z) = \mathbb{E}_{\mu_p} [f] = f (x).
\]
Since $g$ takes the value $\{0, 1\}$ on $\{-1, 1\}^n$, plugging in $f = g$ gives $\gamma_n (A) = g (x)$, where $A = \{z \in \mathbb{R}^n \mid g (\text{sign}_n (z + \alpha)) = 1\}$. For this choice of $A$, a calculation shows that for all $i \in [n],$
\[
\int_A z_i \gamma_n (dz) = \frac{1}{\sqrt{2\pi}} e^{-\alpha_i^2/2} \int_{\mathbb{R}^n} \partial_i g (\text{sign}_n (z + \alpha)) d \gamma_n (z)
\]
\[
\overset{(101)}{=} \frac{1}{\sqrt{2\pi}} e^{-\alpha_i^2/2} \partial_i g (x)
\]
\[ \geq \frac{1}{\sqrt{2\pi}} e^{-\Phi^{-1}\left(\frac{1+t}{2}\right)^2/2} \partial_i g (x). \]  
(102)

We then have

\[ \| \nabla g (x) \|_2^2 \leq 2 \pi e^{-\Phi^{-1}\left(\frac{1+t}{2}\right)^2} \left( \int_A z \gamma_n (z) \right) \]
(102)

\[ \leq 2 \pi e^{-\Phi^{-1}\left(\frac{1+t}{2}\right)^2} \int_{-\Phi^{-1}(g(x))}^{\infty} t \gamma_1 (dt) \]
(99)

\[ \leq Ce^{-\Phi^{-1}\left(\frac{1+t}{2}\right)^2} g (x)^2 \log \left( \frac{e}{g (x)} \right) \]
(100)

\[ \leq \sum_{S \subseteq \{1, \ldots, n\}, |S| = 2} \hat{g}_p (S)^2. \]
(103)

The following lemma, which bounds the sum of squares of \(p\)-biased Fourier coefficients, is immediately obtained from the work of Keller and Kindler [22, Lemma 6]. While that work uses hypercontractivity, the lemma can also be proved without hypercontractivity, by following the proof of the non-biased case by Talagrand [33, Theorem 2.4], or Section 3 by Eldan [12].

Lemma B.1 Let \(0 \leq t < 1\), let \(p \in (0, 1)^n\) be such that \(p_i \in \{1+t, 1-t\}\) for all \(i\), and let \(c_p = p_i (1-p_i) = (1-t^2)/4\). For a function \(g : \{-1, 1\}^n \rightarrow \{-1, 1\}\), let

\[ W (f) = c_p \sum_{i=1}^n \inf_{p_i}^p (g)^2. \]

There exists a function \(C (t)\) such that for every \(g\),

\[ \sum_{S \subseteq \{1, \ldots, n\}, |S| = 2} \hat{g}_p (S)^2 \leq C (t) W (g) \cdot \log \left( \frac{2}{W (g)} \right). \]
(104)
Combining (103) and (104), we get
\[
\| \nabla^2 g(x) \|^2_{HS} \leq C(t) \mathcal{W}(g) \cdot \log \left( \frac{2}{\mathcal{W}(g)} \right) .
\] (105)

As stated in Eq. (98), for monotone functions the influence of the \( i \)-th bit is given by
\[
\text{Inf}_i^p(g) = 2\sqrt{p_i} \sqrt{1 - p_i} \hat{g}_p \{ i \} = 2\sqrt{c_p} \hat{g}_p \{ i \},
\]
and so,
\[
\mathcal{W}(f) = c_p \sum_{i=1}^n \text{Inf}_i^p(g)^2 = 4c_p^2 \sum_{i=1}^n \hat{g}_p \{ i \}^2.
\]

On the other hand, using Proposition A.1 with \( S = \{ i \} \),
\[
\| \nabla g(x) \|^2_2 = \sum_{i=1}^n (\partial_i g(x))^2
\]
\[
= \frac{16}{1 - t^2} \sum_{i=1}^n \hat{g}_p \{ i \}^2
\]
\[
= \frac{4}{(1 - t^2) c_p^2} \mathcal{W}(g) := C'(t) \mathcal{W}(g).
\]

Plugging this into (105), we see that for some \( C(t) \) we have
\[
\| \nabla^2 g(x) \|^2_{HS} \leq C(t) \| \nabla g(x) \|^2_2 \log \left( \frac{C(t)}{\| \nabla g(x) \|^2_2} \right).
\]

\(\Box\)

**Proof of Lemma 2.5** Let \( s \) be an integer to be chosen later. Let \( c_2 = \frac{c_1}{2(2 - \log(1-a)^2 - \log c_1)} \), and suppose for the sake of contradiction that
\[
P \left[ |X_1| \geq c_1 (1 - a)^2 \right] < c_2 (1 - a)^2.
\]
Then,
\[
\mathbb{E} [ |X_1| ] \leq \mathbb{E} [ |X_1| \mathbf{1}_{|X_1| \leq c_1 (1 - a)^2} ] + \mathbb{E} [ |X_1| \mathbf{1}_{|X_1| \in [c_1 (1-a)^2, s]} ] + \mathbb{E} [ |X_1| \mathbf{1}_{|X_1| \geq s} ]
\]
\[ \leq c_1 (1 - a)^2 + c_2 (1 - a)^2 s + \int_s^\infty P[X_t \geq t] \, dt. \]

To bound the third term, we use the fact that the quadratic variation is sub-exponential, as follows. Denote \( \tau_0 = 0 \), and for every \( k \in \mathbb{N} \), let \( \tau_k := \inf \{ t > 0 \mid [X]_t \geq 2k \} \). Recall that the infimum of an empty set is \( \infty \). Since the total variation is finite almost surely, the random variable \( k^* = \sup \{ k \in \mathbb{N} \cup \{0\} \mid \tau_k < \infty \} \) is finite almost surely. For every measurable stopping time \( \tau \in [0, 1] \), we have \( E ([X]_1 - [X]_\tau) \leq 1 \). Thus

\[ P [k^* \geq k + 1 \mid k^* \geq k] = P [X_1 - [X]_{\tau_k} \geq 2] \leq \frac{\mathbb{E} [X_1 - [X]_{\tau_k}]}{2} \leq \frac{1}{2}. \]

By induction we get that

\[ P [X_1 \geq 2k] = \Pr [k^* \geq k] \leq 2^{-k}. \]

Thus, if \( s \geq 2 \) is even, we have

\[ \int_s^\infty P[X_1 \geq t] \, dt \leq 2 \cdot \sum_{k=s/2}^\infty 2^{-k} = 4 \cdot 2^{-s/2}. \]

The expected value of the quadratic variation can then be bounded by

\[ \mathbb{E} [X_1] \leq c_1 (1 - a)^2 + c_2 (1 - a)^2 s + 4 \cdot 2^{-s/2}. \]

Choosing \( s \) to be the smallest even integer larger than \( 2 \left( 2 - \log_2 \left( (1 - a)^2 \right) \right) - \log_2 c_1 \), we get

\[ \mathbb{E} [X_1] \leq c_1 (1 - a)^2 + 4c_2 (1 - a)^2 \left( 2 - \log_2 \left( (1 - a)^2 \right) - \log_2 c_1 \right) + (1 - a)^2 c_1 \]

\[ = (1 - a)^2 \left( 2c_1 + 4c_2 \left( 2 - \log (1 - a)^2 - \log c_1 \right) \right). \]

By choice of \( c_2 = \frac{c_1}{2(2 - \log(1 - a)^2 - \log c_1)} \), we get \( \mathbb{E} [X_1] \leq 3c_1 (1 - a)^2 \), which, since \( c_1 < 1/3 \), contradicts the fact that \( \mathbb{E} [X_1] = \text{Var}(X_1) \geq (1 - a)^2 \).

**Proof of Lemma 3.3** To prove (27), assume first that \( t_1 > 0 \), so that the number of jumps that \( B_t \) makes in the time interval \([t_1, t_2]\) is almost surely finite. For any integer \( N > 0 \), partition the interval \([t_1, t_2]\) into \( N \) equal parts, setting

\[ \square \]
\( t_k^N = t_1 + \frac{k}{N} (t_2 - t_1) \) for \( k = 0, \ldots, N \). Almost surely, none of the jumps of \( B_t \) occur at any \( t_k^N \). Since there are only finitely many jumps, there exists an almost surely finite \( N_0 \) so that for all \( N > N_0 \), every sub-interval in the partition contains at most one jump point. Since \( g_t \) is left-continuous, we therefore almost surely have

\[
\sum_{t \in J_i \cap [t_1, t_2]} 4t^2 g_t = \lim_{N \to \infty} \sum_{k=0}^{N-1} 4 \left( t_k^N \right)^2 g_{t_k^N} 1_{J_i \cap [t_k^N, t_{k+1}^N] \neq \emptyset}.
\]

Since \( g_{t_k^N} \) is bounded, the expression

\[
\sum_{k=0}^{N-1} 4 \left( t_k^N \right)^2 g_{t_k^N} 1_{J_i \cap [t_k^N, t_{k+1}^N] \neq \emptyset}
\]

is bounded in absolute value by a constant times the number of jumps of \( B_t \) in the interval \([t_1, t_2]\), which is integrable. By the dominated convergence theorem, we then have

\[
\mathbb{E} \sum_{t \in J_i \cap [t_1, t_2]} 4t^2 g_t = \lim_{N \to \infty} \mathbb{E} \sum_{k=0}^{N-1} 4 \left( t_k^N \right)^2 g_{t_k^N} 1_{J_i \cap [t_k^N, t_{k+1}^N] \neq \emptyset}.
\]

Since \( g_{t_k^N} \) is measurable with respect to \( \{B_s\}_{0 \leq s < t_k^N} \), it is independent of whether or not a jump occurred in the interval \([t_k^N, t_{k+1}^N]\), and the expectation breaks up into

\[
\mathbb{E} \sum_{t \in J_i \cap [t_1, t_2]} 4t^2 g_t = \lim_{N \to \infty} \sum_{k=0}^{N-1} \mathbb{E} \left[ 4 \left( t_k^N \right)^2 g_{t_k^N} \right] \mathbb{E} \left[ 1_{J_i \cap [t_k^N, t_{k+1}^N] \neq \emptyset} \right].
\]

The set \( J_i = \text{Jump}(B_t^{(i)}) \) is a Poisson process with rate \( 1/2t \), and so the number of jumps in the interval \([t_k^N, t_{k+1}^N]\) distributes as \( \text{Pois} (\lambda) \), where

\[
\lambda = \int_{t^N_k}^{t^N_{k+1}} \frac{1}{2t} dt = \frac{1}{2} \log \frac{t^N_{k+1}}{t^N_k}.
\]
The probability of having at least one jump is then equal to
\[
P\left[ J_i \cap [t_k^N, t_{k+1}^N] \neq \emptyset \right] = 1 - e^{-\lambda} = 1 - \sqrt{\frac{t_k^N}{t_{k+1}^N}} = 1 - \sqrt{1 - \frac{(t_2 - t_1) / N}{t_{k+1}^N}}.
\]

Plugging this into display (106), we get
\[
E \sum_{t \in J_i \cap [t_1, t_2]} 4t^2 g_t = \lim_{N \to \infty} E \sum_{k=0}^{N-1} 4 \left( t_k^N \right)^2 g_{t_k^N} \left( \frac{(t_2 - t_1) / N}{2t_k^N} + O \left( \frac{1}{N^2} \right) \right).
\]

The factor \( O \left( \frac{1}{N^2} \right) \) is negligible in the limit \( N \to \infty \), since the sum contains only \( N \) bounded terms. We are left with
\[
\lim_{N \to \infty} E \sum_{k=0}^{N-1} 4 \left( t_k^N \right)^2 g_{t_k^N} \frac{(t_2 - t_1) / N}{2t_k^N} = \lim_{N \to \infty} E \sum_{k=0}^{N-1} \left[ 2t_k^N g_{t_k^N} \right] \frac{t_2 - t_1}{N}.
\]

Since \( g_t \) is continuous almost everywhere, by the definition of the Riemann integral, the limit is equal to \( 2E \int_{t_1}^{t_2} t \cdot g_t \, dt \), and we get
\[
E \sum_{t \in J_i \cap [t_1, t_2]} 4t^2 g_t = 2E \int_{t_1}^{t_2} t \cdot g_t \, dt
\]
for all \( t_1 > 0 \). Taking the limit \( t_1 \to 0 \) gives the desired result for \( t_1 = 0 \) by continuity of the right hand side in \( t_1 \).

**Proof of Lemma 3.7** Let \( t_1 = \inf \{ t \mid g(t) \geq K \} \), and denote \( L = \max \{ x \log \frac{K}{x} \mid x \in [0, K] \} \). Note that \( L \) depends only on \( K \). Then, for all \( t \leq t_1 \), we have
\[
g'(t) \leq C \cdot L.
\]

Integrating, this means that for all \( t \leq t_1 \)
\[
g(t) \leq g(0) + tC L \leq \frac{K}{2} + tC L.
\]
In particular, \( t_1 \geq \frac{K}{2CL} \), otherwise we would have \( g(t_1) < K \), contradicting the definition of \( t_1 \) and continuity of \( g \). Denoting \( t_0 = \frac{K}{4CL} \), we must have \( g(t) < K \) for all \( t \in [0, t_0] \). This ensures that \( \log \frac{K}{g(t)} \) is positive in this interval, which means we can rearrange the differential inequality (36) to give

\[
\frac{g'(t)}{g(t) \log \frac{K}{g(t)}} \leq C
\]

for all \( t \in [0, t_0] \). A short calculation reveals that the left hand side is the derivative of \(- \log \log (K/g)\). Integrating from 0 to \( t \), we get

\[
\log \log \frac{K}{g(0)} - \log \log \frac{K}{g(t)} \leq Ct.
\]

Rearranging gives

\[
\log \log \frac{K}{g(t)} \geq \log \log \frac{K}{g(0)} - Ct,
\]

and exponentiating twice gives the inequality

\[
g(t) \leq \left( \frac{1}{K} \right)^{e^{-Ct} - 1} g(0)^{e^{-Ct}}.
\]

By the Taylor expansion of \( e^{-Ct} \), we can write \( e^{-Ct} := 1 - Ct + R(Ct) \) with \( R(x) = O(x^2) \) and \( R(x) > 0 \) for \( x > 0 \). This gives

\[
g(t) \leq g(0)^{1-Ct+R(Ct)} \left( \frac{1}{K} \right)^{-Ct+R(Ct)}
\]

\[
= g(0)^{1-Ct} K^{Ct} \left( \frac{g(0)}{K} \right)^{R(Ct)}
\]

\[
\leq g(0)^{1-Ct} K^{Ct} \left( \frac{1}{2} \right)^{R(Ct)}
\]

\[
\leq g(0)^{1-Ct} K^{Ct}
\]

as needed. \( \square \)

References

1. Beffara, V., Duminil-Copin, H.: The self-dual point of the two-dimensional random-cluster model is critical for \( q \geq 1 \). Probab. Theory Relat. Fields 153(3–4), 511–542 (2012)
2. Benjamini, I., Kalai, G., Schramm, O.: Noise sensitivity of Boolean functions and applications to percolation. Inst. Hautes Études Sci. Publ. Math. 90, 5–43 (1999)
3. Benjamini, I., Kalai, G., Schramm, O.: First passage percolation has sublinear distance variance. Ann. Probab. 31(4), 1970–1978 (2003)
4. Barthe, F., Maurey, B.: Some remarks on isoperimetry of Gaussian type. Ann. Inst. H. Poincaré Probab. Statist. 36(4), 419–434 (2000)
5. Bobkov, S.G.: An isoperimetric inequality on the discrete cube, and an elementary proof of the isoperimetric inequality in Gaussian space. Ann. Probab. 25(1), 206–214 (1997)
6. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
7. Cordero-Erausquin, D., Ledoux, M.: Hypercontractive Measures, Talagrand’s Inequality, and Influences, pp. 169–189. Springer, Berlin (2012)
8. Capitaine, M., Hsu, E.P., Ledoux, M.: Martingale representation and a simple proof of logarithmic Sobolev inequalities on path spaces. Electron. Commun. Probab. 2, 71–81 (1997)
9. Dinur, I., Safra, S.: On the hardness of approximating minimum vertex cover. Ann. Math. (2) 162(1), 439–485 (2005)
10. Durrett, R.: Probability—theory and examples. In: Cambridge Series in Statistical and Probabilistic Mathematics, vol. 49. Cambridge University Press, Cambridge (2019)
11. Eldan, R.: A two-sided estimate for the Gaussian noise stability deficit. Invent. Math. 201(2), 561–624 (2015)
12. Eldan, R.: Second-order bounds on correlations between increasing families
13. Eldan, R.: Analysis of high-dimensional distributions using pathwise methods. In: 2022 ICM Proceedings (2021)
14. Ellis, D.: Almost isoperimetric subsets of the discrete cube. Combin. Probab. Comput. 20(3), 363–380 (2011)
15. Friedgut, E.: Boolean functions with low average sensitivity depend on few coordinates. Combinatorica 18(1), 27–35 (1998)
16. Friedgut, E.: Sharp thresholds of graph properties, and the k-sat problem. J. Am. Math. Soc. 12(4), 1017–1054 (1999). (With an appendix by Jean Bourgain)
17. Gavinsky, D., Kempe, J., Kerenidis, I., Raz, R., de Wolf, R.: Exponential separation for one-way quantum communication complexity, with applications to cryptography. SIAM J. Comput. 38(5), 1695–1708 (2008/2009)
18. Garban, C., Steif, J.E.: Noise sensitivity of Boolean functions and percolation. In: Institute of Mathematical Statistics Textbooks, vol. 5. Cambridge University Press, New York (2015)
19. Hart, S.: A note on the edges of the n-cube. Discrete Math. 14(2), 157–163 (1976)
20. Kingman, J.F.C.: Poisson processes. In: Oxford Studies in Probability, vol. 3. The Clarendon Press, Oxford University Press, New York, Oxford Science Publications (1993)
21. Kahn, J., Kalai, G.: Thresholds and expectation thresholds. Combin. Probab. Comput. 16(3), 495–502 (2007)
22. Keller, N., Kindler, G.: Quantitative relation between noise sensitivity and influences. Combinatorica 33(1), 45–71 (2013)
23. Kahn, J., Kalai, G., Linial, N.: The influence of variables on Boolean functions (extended abstract). pp. 68–80 (1988)
24. Klein, S., Levi, A., Safra, M., Shikhelman, C., Spinka, Y.: On the converse of Talagrand’s influence inequality. arXiv:1506.06325 (2015)
25. Krauthgamer, R., Rabani, Y.: Improved lower bounds for embeddings into L1. SIAM J. Comput. 38(6), 2487–2498 (2009)
26. Kalai, G., Safra, S.: Threshold phenomena and influence: perspectives from mathematics, computer science, and economics. In: Computational Complexity and Statistical Physics, St. Fe Inst. Stud. Sci. Complex., Oxford University Press, New York, pp. 25–60 (2006)
27. Mossel, E., Neeman, J., Tamuz, O.: Majority dynamics and aggregation of information in social networks. Auton. Agent. Multi-Agent Syst. 28(3), 408–429 (2014)
28. O’Donnell, R.: Analysis of Boolean Functions. Cambridge University Press, New York (2014)
29. O’Donnell, R., Servedio, R.A.: Learning monotone decision trees in polynomial time. SIAM J. Comput. 37(3), 827–844 (2007)
30. Raz, R.: Fourier analysis for probabilistic communication complexity. Comput. Complex. 5(3–4), 205–221 (1995)
31. Talagrand, M.: Isoperimetry, logarithmic Sobolev inequalities on the discrete cube, and Margulis’ graph connectivity theorem. Geom. Funct. Anal. 3(3), 295–314 (1993)
32. Talagrand, M.: On Russo’s approximate zero-one law. Ann. Probab. 22(3), 1576–1587 (1994)
33. Talagrand, M.: How much are increasing sets positively correlated? Combinatorica 16(2), 243–258 (1996)
34. Talagrand, M.: On boundaries and influences. Combinatorica 17(2), 275–285 (1997)

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