Noise Sensitivity of Boolean Functions and Applications to Percolation: Report

Manuel Paez
map2332@columbia.edu

April 28, 2025

1 Introduction

The paper [BKS99] showcases a theoretical analysis of noise-sensitivity and noise-stability for Boolean functions. In particular, a large class of events in a product probability space exhibit noise sensitivity: if a configuration is slightly perturbed by noise, the perturbed configuration offers little to no information about whether the original event occurred. Conversely, the paper shows that the weighted majority functions are noise-stable, which is a class of events that are robust to small amounts of noise. This paper presents several theorems and conditions necessary and sufficient for determining noise sensitivity and stability. They illustrate the examples of bond percolation and weighted majority to showcase noise-sensitivity and noise-stability by proving the major result in statistical physics: the crossing event in 2D bond percolation is asymptotically noise sensitive, revealing interesting insights into the critical phenomena. Essentially, this paper bridges ideas from probability theory (noise processes), analysis, combinatorics, statistical physics (percolation), and more to showcase a fundamental way to classify functions based on how robust or fragile they are to random noise in their inputs.

1.1 Noise Sensitivity

Let us introduce the general concepts and definitions. Let $\Omega_n = \{0,1\}^n$ be the Hamming cube (Boolean cube) with the uniform probability measure P. Define an event $A \subset \Omega_n$. Given a random $x = (x_1, ..., x_n) \in \Omega_n$, consider $N_\epsilon(x)$ to be the random perturbation of x. This means the following: for every $j \in \{1, ..., n\}$, $N_\epsilon(x)_j = x_j$ with probability $1-\epsilon$ and $N_\epsilon(x)_j \neq x_j$ with probability ϵ for $\epsilon \in (0,1)$. In informal terms, $N_\epsilon(x)$ is created by flipping each bit x_j independently with probability ϵ. Given $N_\epsilon(x)$, we would like to predict the event $x \in A$. Let us define the Noise Sensitivity of the event A as follows:

Definition 1.1 (Noise Sensitivity)**.** An event A is noise sensitive if for some small $\delta > 0$,

$$\gamma(A, \epsilon, \delta) := P(x : |P(N_\epsilon(x) \in A|x) - P(A)| > \delta) < \delta$$

Essentially, an event A is noise sensitive if the knowledge of x gives almost no information (predictive power) about whether $N_\epsilon(x) \in A$. For even more informal terms, if flipping is small, random fraction ϵ of these inputs makes it almost impossible to predict the original outcome based on the noisy input, the event is noise sensitive. Let us now define the sensitivity gauge of A, which measures the noise sensitivity.

Definition 1.2 (Sensitivity gauge of A)**.** Given the event A, the sensitivity gauge of A is the infimum of all $\delta > 0$ such that the Noise Sensitivity holds:

$$\phi(A, \epsilon) = \inf\{\delta > 0 : \gamma(A, \epsilon, \delta) < \delta\}$$

Essentially, the sensitivity gauge of A measures how unpredictable A becomes under noise. The concept of Noise-Sensitivity becomes particularly interesting when considering sequences of events A_m defined on spaces of increasing dimension n_m:

Definition 1.3 (Asymptotically noise sensitive)**.** A sequence of events $A_m \subset \Omega_{n_m}$ is asymptotically noise sensitive.

$$\lim_{m\to\infty} \phi(A_m, \epsilon) = 0 \quad \forall \epsilon \in (0, 1/2)$$

This is equivalent to the variance of the conditional probability $P(N_\epsilon(x) \in A_M|x)$ vanishing to zero:

$$\lim_{m\to\infty} \mathrm{VAR}[P(N_\epsilon(x) \in A_m|x)] = 0$$

Think of ϵ as any fixed noise level between 0 and 1/2.

The paper considers examples in which a sequence of events is not asymptotically noise sensitive and examples in which the sequence of events is asymptotically noise sensitive. In particular, they show that the Weighted Majority, i.e. event which depends on the sum of the bits, is not asymptotically noise sensitive but has stability, and that the Bond percolation crossing event C_m on a $(m + 1) \times m$ grid is asymptotically noise sensitive.

Before these examples are considered, the paper establishes key conditions equivalent or related to asymptotic noise sensitivity.

1.2 Conditions for Asymptotic Noise Sensitivity

Influences of variables The concept of "influence" quantifies how much a single input variable affects the output of a Boolean function or the occurrence of an event. Let us now define the influence of variables and show how it relates to asymptotic noise-sensitivity.

Given $x \in \Omega = \{0, 1\}^n$, consider $\sigma_j x = (x'_1, ..., x'_n)$ to be the configuration x with the k-th bit flipped, where $x'_k = x_k$ when $k \neq j$ and $x'_k = 1 - x_j$ when $k = j$. For a function $f : \Omega \to \mathbb{R}$ (often the indicator function χ_A of an event A), the influence of the k-th variable is

$$I_k = \|f(\sigma_k x) - f(x)\|_1$$

Essentially, $I_k(f)$ is the expected absolute value of the change in f when the k-th bit x_k is flipped. When f is the indicator function of an event A, $I_k(A) = I_k(\chi_A)$ is precisely the probability that flipped the k-th bit changes the outcome, i.e exactly one of the two elements $x, \sigma_k x$ is in A. A variable with high influence is critical for determining the outcome for a significant fraction of inputs. Let us now define the sum of influences (total influence) $I(f)$:

$$I(f) := \sum_k I_k(f)$$

This represents the total sensitivity to single-bit flips. In percolation, this corresponds to the expected number of pivotal edges. Let us also define the sum of squared influences:

$$II(f) := \sum_k I_k(f)^2$$

From this, the paper proves the following theorems, which establish conditions for asymptotic noise-sensitivity.

Theorem 1.4. *Let $A_m \subset \Omega_{n,m}$ be a sequence of events. If $\lim_{m\to\infty} II(A_m) = 0$, then $\{A_m\}$ is asymptotically noise sensitive.*

This theorem shows that small total squared influence implies noise-sensitivity. Specifically, it connects the collective "diffuseness" of influence to the event's fragility under noise. If the influence is spread out thinly across many variables such that the sum of squares is small, then perturbing a small fraction ϵ of variables is unlikely to provide much information about the original outcome.

Let us now shift to monotone functions. Let Ω_n have the following lattice order: $(x_1, ..., x_n) \leq (y_1, ..., y_n)$ if and only if $x_j \leq y_j$ for all $j \in [n]$. Consider a monotone function $f : \Omega_n \to \mathbb{R}$, i.e. $f(x) \leq f(y)$ whenever $x \leq y$ and a monotone event $A \subset \Omega_n$, i.e. its characteristic function χ_A is monotone. For monotone events, we have the converse result of 1.4.

Theorem 1.5. *Let $A_m \subset \Omega_{n_m}$ be a sequence of monotone events with $\inf_m II(A_m) > 0$ (bounded away from 0), then $\{A_m\}$ is not asymptotically noise sensitive.*

Essentially, for a sequence of monotone events, if the influence squared does not vanish, the event retains some predictability even under noise. The paper explicitly states that monotonicity is necessary here, giving parity as a counter-example. This theorem links noise insensitivity (or stability) for monotone events to having a non-vanishing sum of squared influences, which we will discuss in depth later. Afterwards, the paper shows a quantitative bound for 1.4 given the assumption that $II(A_m)$ goes to zero fast enough:

Theorem 1.6. *Let $A \subset \Omega_n$, and suppose that $II(A) \leq n^{-a}$, where $a \in (0, 1/2]$. Then there exist $c_1, c_2 > 0$, depending only on a so that*

$$\phi(A, \epsilon) \leq c_1 n^{-c_2 \epsilon} \quad \forall \epsilon \in (0, 1/4)$$

Essentially, if the sum of squared influences $II(A)$ decays polynomially, i.e. $II(A_m) \leq (n_m)^{-a}$, the noise sensitivity is strong, meaning the vent is sensitive to noise level ϵ_m as long as $\epsilon_m \log n_m \to \infty$, which provides a rate at which the noise level can decrease while still destroying predictability.

Weighted Majority Noise sensitivity can be connected to the correlation of a function f with majority functions, both weighted and unweighted. In particular, for monotone events, noise insensitivity is closely related to having a significant correlation with majority functions. This suggests that functions that behave somewhat like a majority vote are more likely to be robust against noise. For $K \subset [n]$, define the majority function on K to be $M_K(x) = \text{sign}(\sum_{j\in K}(2x_j - 1))$:

$$M_K(x) = \begin{cases} -1, & \text{for } \sum_{j\in K} x_j < |K|/2 \\ 0, & \text{for } \sum_{j\in K} x_j = |K|/2 \\ 1, & \text{for } \sum_{j\in K} x_j > |K|/2 \end{cases}$$

and for $f : \Omega_n \to \mathbb{R}$, define $\Lambda(f)$ to be the correlation function of f as

$$\Lambda(f) = \max\{|\mathbb{E}(fM_K))| : K \subset [n]\}$$

which is the maximum absolute correlation between the function f, which represents the event A, and any such unweighted majority function M_K. With this, we have the following theorem:

Theorem 1.7. *For monotone $f : \Omega_n \to [0, 1]$, then*

$$II(f) \leq C\Lambda(f)^2\Big(1 - \log \Lambda(f)\Big) \log n$$

where C is some universal constant. For a sequence of monotone events $A_m \subset \Omega_{n_m}$ with

$$\lim_{m\to\infty} \Lambda(A_m)^2\Big(1 - \log \Lambda(A_m)\Big) \log n_m$$

then the sequence $\{A_m\}$ is asymptotically noise sensitive.

Combined with Theorem 1.4 (small II implies sensitivity), Theorem 1.7 implies that if a sequence of monotone events $\{A_m\}$ has a correlation $\Lambda(A_m)$ that vanishes sufficiently faster, then the sequence must be asymptotically noise sensitive. For a sharper characterization, let us consider the connection between noise-sensitivity and weighted majority. For positive weights $w = (w_1, ..., w_n)$, consider the weighted majority function $M_w(x) = \text{sign}(\sum_{j=1}^n (2x_j - 1)w_j)$ and correlation function

$$\widetilde{\Lambda}(A) = \max_{w\in[0,1]^n}\{|\mathbb{E}(\chi_A M_w)|\}$$

The authors show the following theorem:

Theorem 1.8. *For a sequence of monotone events $A_m \subset \Omega_{n_m}$, $\{A_m\}$ is asymptotically noise sensitive if and only if $\lim_{m\to\infty} \widetilde{\Lambda}(A_m) = 0$.*

Theorem 1.8 establishes weighted majority functions as the benchmark for noise stability in the realm of monotone functions. Specifically, if a monotone event is not noise sensitive, it must have a non-vanishing correlation with some weighted majority function. Conversely, if a monotone event is uncorrelated with all weighted majorities in the limit, it must be noise sensitive.

1.3 Noise Stability

Noise stability is defined to be the opposite concept of noise sensitivity. While noise sensitivity describes events whose outcomes become unpredictable under small random perturbations, noise stability describes events that are robust to such noise. Essentially, if an event A is noise stable, applying a small amount of noise ϵ (flipping each bit independently with probability ϵ) is unlikely to change whether the event occurs or not. For a more formal definition, for a sequence of events $A_i \subset \Omega_{n_i}$ and $x \in \Omega_n$ randomly and uniformly chosen, we can define the uniform stability as follows:

Definition 1.9 (Uniform Stable)**.** A sequence of events $\{A_i\}$ is uniformly stable if $\lim_{\epsilon\to 0} P[x \in A_i \Delta N_\epsilon A_i] = 0$ is uniform in i.

Essentially, the probability of the noise changing the outcome $P[x \in A_i \Delta N_\epsilon A_i]$ goes to zero as ϵ goes to zero. The faster $P[x \in A_i \Delta N_\epsilon A_i]$ goes to zero, the more noise-stable A is.

The primary examples of noise-stable functions discussed in the paper are the (generalized) weighted majority functions. These are events of the form $\mathcal{M}_{w,s} = \{x \in \{0,1\}^n : \sum_{j=1}^n (2x_j - 1)w_j > s\}$ for some weights $w \in \mathbb{R}^n$ and threshold $s \in \mathbb{R}$. Let $\mathcal{M}$ be the collection of all such (generalized) weighted majority events. We will now state the theorem that $\mathcal{M}$ is stable.

Theorem 1.10. *$\mathcal{M}$ is uniformly stable. Specifically, for every $M \in \mathcal{M}$, $P[M - N_\epsilon M] \leq C\epsilon^{1/4}$ for C being a universal constant independent of M.*

It is important to note that asymptotically noise-sensitive and uniformly stable sequences are asymptotically uncorrelated. This will be shown in the later results.

Fourier-Walsh Expansion For a boolean function f on the Hamming cube $\Omega_n = \{0,1\}^n$, it can be uniquely expressed as a linear combination of Walsh functions u_s :

$$f = \sum_{s \subset [n]} \widehat{f}(S) u_s$$

where $u_S(T) = (-1)^{|S \cap T|}$ for an input $T \subset [n]$ which represents the configuration of x where $x_j = 1$ iff $j \in T$, and $\widehat{f}(S)$ are the Fourier-Walsh coefficients, which measure the correlation of the function f with the corresponding Walsh function u_s. For this, we identify any vector $x \in \Omega_n$ with the subset $\{j \in [n] : x_j = 1\}$ of $[n] = \{1, .., n\}$. It is important to know that the Walsh functions form an orthonormal basis with respect to the uniform measure on Ω_n. Note that $|x| = \|x\|_1$ is the cardinality of the set. We have the following theorem:

Theorem 1.11. *Let $A_m \subset \Omega_{n_m}$ be a sequence of events and set $g_m = \chi_{A_m}$.*

- *$\{A_m\}$ is asymptotically noise sensitive if and only if for every finite k*

$$\lim_m \sum \{\widehat{g}_m(S)^2 : S \subset [n], 1 \leq |S| \leq k\} = 0 \qquad \textit{(low-order coefficients vanish)}$$

- *$\{A_m\}$ is uniformly stable if and only if*

$$\lim_{k \to \infty} \sup_m \sum \{\widehat{g}_m(S)^2 : S \subset [n], |S| \geq k\} = 0 \qquad \textit{(high-order coefficients vanish)}$$

The connection with total influence follows:

$$I(f) = 4 \sum_{S \subset [n]} \widehat{f}(S)^2 |S|$$

Let us define $J(f) = \sum_{S \neq \emptyset} \widehat{f}(S)^2 / |S|$. Let us define the parameter $\alpha(A) = \log I(A) / \log(n)$ to relate to the expected number of pivotal elements and $\beta(A) = -\log J(A) / \log n$ to relate to the absence of low-frequency coefficients. Both of these parameters are based on the distribution of the Fourier mass, and are given conjectures for their values in percolation.

1.4 Past Work and Literature

This paper is the first paper to connect asymptotic noise sensitivity and uniform stability with Fourier analysis. However, previous and future works have established connections with noise and other tools. [Tal94, Tal95] are isoperimetric inequalities that are helpful for the noise analysis provided in the paper. Some of the isoperimetric inequalities are further developed in first passage percolation problems in this paper [BKS11]. The notion of Noise has been studied in [Tsi99, WS99]. In [Tsi99], Tsirelson showcases that a two-dimensional array of independent random signs produces coalescing random walks starting at the origin, with its position after n steps being a nonlinear, noise-sensitive function of the signs. "Noise", in Tsirelson's view, is a σ-field filtration: a way to mathematically model the information available through a random process as time unfolds.

2 Sensitivity to Noise and Influences of Variables

In this section, we formulate the analytic framework for noise sensitivity and the tools used to analyze noise sensitivity. In particular, we connect noise sensitivity to the Fourier-Walsh expansion of Boolean functions defined on the Hamming cube $\Omega_n = \{0,1\}^n$. We want to determine if knowing x helps predict whether $N_\epsilon(x)$ satisfies a certain property (event A).

Generalized Noise Framework For the challenge of integrating different types of random error and noise into this analysis, we introduce a generalized noise model $x = N(\omega, q)$ that goes beyond the standard ϵ-noise bit-flipping $N_\epsilon(x)$. This serves two purposes:

1. It provides a single framework that encompasses various noise types, i.e. standard N_ϵ noise (flipping bits independently) or asymmetric noise.

2. It allows results to be expressed in terms of fundamental properties of the noise parameter, such as their variances.

We start off by generalizing the noise operator N_ϵ by first selecting a random $x \in \Omega_n$ in two stages that still follows the uniform distribution overall:

- Consider a random vector $q = (q_1, ..., q_n)$ with each coordinate q_j is an independent random variable in $[0,1]$ drawn independently according to some measure ν_j. It is required that $\mathbb{E}[q_j] = \frac{1}{2}$ for $j = [n]$

- Given q, generate x from the product measure P_q where $P_q\{x : x_j = 1\} = q_j$ and $P_q\{x : x_j = 0\} = 1 - q_j$ independently for each j. This final configuration of x is determined by comparing q_j to a uniform random $\omega_j \in [0,1]$: $x_j = 1$ if $1 - \omega_j < q_j$ and $x_j = 0$ otherwise, ensuring that x is uniformly distributed on Ω_n. Thus, by the construction of $N(\omega, q)$ covering the initial $N_\epsilon(x)$ model, $N_\epsilon(x)$ corresponds to a specific choice of q based on z and ϵ.

From this, the overall distribution of x (averaging over q) is the uniform measure on Ω_n because $\mathbb{E}[P_q\{x_j = 1\}] = E[q_j] = 1/2$.

Predictability The core idea of this chapter is to predict $f(x)$ from q: this means we want to see how much information knowing q (first-stage) gives about the outcome $f(x)$ for a function $f : \Omega_n \to \mathbb{R}$. With this, let us define the conditional expectation

$$G(f, z) = \mathbb{E}[f(x)|q = z]$$

to be the best prediction of $f(x)$ given the specific noise parameter $q = z$. Noise sensitivity means that for most z, $G(f, z)$ is close to the overall average $E(f)$. We also define the second moment

$$Z(f, \nu) = \mathbb{E}_q[G(f, q)^2] = \int G(f, z)^2 d\nu(z)$$

which measures the average squared prediction. From this, the variance of this prediction

$$VAR_q(G(f, q)) = Z(f, v) - (\mathbb{E}[f])^2 = \mathbb{E}_q[G(f, q)^2] - (\mathbb{E}[f])^2$$

measures how much, on average, knowing the noise parameter q helps in predicting $f(x)$. Low variance implies that knowing q does not significantly improve the prediction beyond the average $\mathbb{E}[f]$. From this, we showcase a lemma that shows $Z(f, \nu)$ depends only on f's Fourier coefficients $\widehat{f}(S)$ and the variance $\zeta_j = \mathrm{VAR}(q_j)$.

2.1 Fourier Analysis and Noise Sensitivity

Lemma 2.1. *The second moment $Z(f, \nu)$ depends only on the Fourier-Walsh coefficients $\widehat{f}(S)$ of the function f and the variance $\zeta_j = VAR(q_j)$ of the variable q_j:*

$$Z(f, \nu) = \sum_{S \subset [n]} \widehat{f}(S)^2 \prod_{j \in S} 4\zeta_j$$

Sketch of Proof. Firstly, express $G(f, z) = \mathbb{E}[f(x)|q] = \sum_x f(x)P_q(x)$ using the Fourier expansion $f = \sum_{S\subset[n]} \widehat{f}(S)u_S$, which implies to:

$$G(f, z) = \sum_{S \subset [n]} \widehat{f}(S)^2 \prod_{j \in S} (1 - 2z_j)$$

From this, we calculate $Z(f, \nu) = \mathbb{E}_q[G(f, q)^2]$. We obtain

$$Z(f, \nu) = \mathbb{E}_q[(\sum_S \widehat{f}(S) \prod_{j \in S} (1 - 2q_j))(\sum_{S'} \widehat{f}(S') \prod_{k \in S'} (1 - 2q_k)$$

Afterwards, expand the square $G(f, q)^2$. We know that q_j are independent and $\mathbb{E}[1 - 2q_j] = 0$ as $\mathbb{E}[q_j] = 1/2$, the expectation cross-terms for $\prod_{j\in S}(1 - 2q_j) \prod_{j\in S'}(1 - 2q_j)$ vanishes unless $S = S'$. Thus, we have

$$Z(f, \nu) = \sum_{S \subset \Omega_n} \widehat{f}(S)^2 \mathbb{E}[\prod_{j \in S} (1 - 2q_j)^2]$$

From a few calculations that shows $\mathbb{E}[(1 - 2q_j)^2] = 4\zeta_j$. Substitute this into the result. □

Our focus now shifts back to the standard noise model N_ϵ with the variance denoted as $\mathrm{VAR}(f, \epsilon)$. Given the noise operator Q_ϵ, let us $Q_\epsilon f(x) = \mathbb{E}[f(N_\epsilon(x))]$ where the expectation of

f is over the noise applied to x. A crucial calculation shows that Q_ϵ acts diagonally on the Fourier basis:

$$Q_\epsilon u_S = (1-2\epsilon)^{|S|} u_S$$

This means that Q_ϵ acts as a multiplier on each Fourier coefficient, with the multiplier decaying exponentially with the size of the set S. By linearity, $Q_\epsilon f = \sum_{S\subset[n]} \widehat{f}(S)(1-2\epsilon)^{|S|}u_S$. The variance relevant for noise sensitivity (as defined in Chapter 1) is

$$\mathrm{VAR}(f,\epsilon) = \mathrm{VAR}(Q_\epsilon f) = \mathbb{E}[(Q_\epsilon f)^2] - (\mathbb{E}f)^2$$

With Parseval's identity, i.e. $\mathbb{E}[g^2] = \sum \widehat{g}(S)^2$ and the formula for $Q_\epsilon f$ and noting that $\mathbb{E}[f] = \widehat{f}(\emptyset)$, the variance becomes

$$\mathrm{VAR}(f,\epsilon) = \sum_{\emptyset\neq S\subset[n]} \widehat{f}(S)^2(1-2\epsilon)^{2|S|}$$

This formula is vital: it explicitly shows how the variance under ϵ-noise depends on the (squared) Fourier coefficients, weighted by a factor that decays with $|S|$.

From this, we now formally establish a quantitative linked between variance $\mathrm{VAR}(A,\epsilon)$ and the sensitivity gauge $\phi(A,\epsilon)$ by showing that the analytic quantity $\mathrm{VAR}(A,\epsilon)\to 0$ is equivalent to the definition of asymptotic noise sensitivity, i.e. $\phi(A_m,\epsilon)\to 0$.

Proposition 2.2. *Providing bounds related $VAR(A,\epsilon)$ to the sensitivity gauge $\phi(A,\epsilon)$: $\frac{1}{2}VAR(A,\epsilon) \leq \phi(A,\epsilon) \leq VAR(A,\epsilon)^{1/3}$.*

Proof Sketch. For the lower bound, let $\delta = \phi(A,\epsilon)$ and $Y = \{y : |Q_\epsilon\chi_A(y) - P(A)| > \delta\}$. By definition of ϕ, $P[Y]\geq\delta$. Then from applying Chebyshev inequality, $\mathrm{VAR}(A,\epsilon) \geq \delta^2 P[Y] \geq \delta^3$ follows. For the upper bound, let $Y' = \{y : |Q_\epsilon\chi_A(y) - P(A)| \geq \delta$ so that $P[Y']\leq\delta$. From this, we have

$$\mathrm{VAR} = \mathbb{E}[(Q_\epsilon\chi_A - P[A])^2] \leq 1^2\cdot P[Y'] + \delta^2\cdot P[\Omega_n\setminus Y'] \leq \delta+\delta^2$$

For $\delta\leq 1$, then this bounds VAR: $\mathrm{VAR}\leq 2\delta = 2\phi(A,\epsilon)$, and the conclusion follows. ☐

Proof sketch of 1.11. This proof will formally establish the necessary and sufficient conditions for sensitivity and stability using the Fourier techniques developed.

1. **Noise Sensitivity:** From Proposition 2.2, A_m is asymptotically noise sensitive $\Leftrightarrow \phi(A_m,\epsilon)\to 0 \Leftrightarrow \mathrm{VAR}(A_m,\epsilon)\to 0$. Using $\mathrm{VAR}(f,\epsilon) = \sum_{S\neq\emptyset}\widehat{f}(S)^2(1-2\epsilon)^{2|S|}$, for fixed $\epsilon\in(0,1/2)$, the factor $(1-2\epsilon)^{2|S|}$ is bounded away from 0 for small $|S|$ and decays for large $|S|$. The sum goes to 0 if and only if the contribution from any finite number of low-order terms vanishes, i.e., $\sum_{1\leq|S|\leq k}\widehat{f}(S)^2\to 0$ for all fixed k.

2. **Stability:** We know that uniform stability $\Leftrightarrow \|\chi_{A_m} - Q_\epsilon\chi_{A_m}\|_1 \to 0$ uniformly as $\epsilon\to 0$. For bounded functions, L_1 convergence is equivalent to L_2 convergence. Calculate:

$$\|\chi_{A_m} - Q_\epsilon\chi_{A_m}\|_2^2 = \sum_{S\neq\emptyset}(\widehat{f}(S) - \widehat{Q_\epsilon f}(S))^2 = \sum_{S\neq\emptyset}\widehat{f}(S)^2(1-(1-2\epsilon)^{|S|})^2$$

This sum goes to 0 uniformly as $\epsilon\to 0$ if and only if the tail sum $\sum_{|S|\geq k}\widehat{f}(S)^2$ can be made arbitrarily small for large enough k, independent of m, which is precisely the condition stated in the theorem.

☐

2.2 Sensitivity Theorems

Now, we switch to proving the main theorems for sensitivity (Theorems 1.4 and 1.6). In particular, the main goal is to show that if the sum of squared influences $II(f) = \sum_k I_k(f)^2$ goes to zero, then the event must be noise-sensitive. The strategy is to show that small $II(f)$ implies that the low-order Fourier coefficients (small $|S|$) must be small. We turn to several tools and techniques for proving this:

Operator T_η: Defined as $T_\eta f = \sum \widehat{f}(S)\eta^{|S|}u_S$ for $\eta \in [0,1]$, it is the heat operator on the hypercube. It relates to Q_ϵ via $T_{1-2\epsilon} = Q_\epsilon$. Note that $T_1 f = f$ and $T_0 f = \mathbb{E}[f]$. The variance relates as

$$\|T_{1-2\epsilon}f\|_2^2 = \mathrm{VAR}(f,\epsilon) + (\mathbb{E}[f])^2$$

Lemma 2.3. *[Bonami, Beckner] Hypercontractivity: $\|T_\eta f\|_2 \leq \|f\|_{1+\eta^2}$ this relates the L_2 norm of the smoothed function $T_\eta f$ to a lower L_p norm where $p = 1+\eta^2 \leq 2$ of the original function f.*

Now, we begin with the main proofs of the noise-sensitivity theorem.

Theorem 2.4 (Weaker Sensitivity Condition of 1.4, 1.6)**.** *If $\lim_{m\to\infty} \frac{\log II(A_m)}{\log\log n_m} = -\infty$, then $\{A_m\}$ is asymptotically noise sensitive.*

This theorem shows that sensitivity of $II(A_m)$ decays faster than any power of $1/\log\log n_m$.

Proof Sketch. Use Bonami-Beckner to bound $\mathrm{VAR}(A,\epsilon) = F_A(\eta)$ in terms of $II(A)$ and n. This is done by defining the difference function $f_j(x) = f(x) - f(\sigma_j x)$. Using $\|f\|_p = I_j(f)^{1/p}$, apply the Bonami-Beckner inequality to f_j: $\|T_\eta f_j\|_2 \leq \|f_j\|_{1+\eta^2}$, and summing over j, we obtain

$$F_A(\eta) \leq n^{\eta^2/(1+\eta^2)} II(A)^{1/(1+\eta^2)}$$

If $II(A)$is small, then the variance must also be small. Thus, $F_A(\eta) \to 0$ for fixed $\eta < 1$. ☐

Proof Sketch of Theorem 1.6. Using Theorem 2.4 and Proposition 2.2, if we assume that $II(A) \leq n^{-a}, a \in (0, 1/2]$, we have

$$\phi(A,\epsilon) \leq \mathrm{VAR}(A,\epsilon)^{1/3} = F_A(1-2\epsilon)^{1/3} \leq 2^{1/3}\Big(1-2\epsilon\Big)^{\frac{a\log n}{9\log(1/a)}}$$

for $\epsilon \in (0, 1/4)$. The theorem follows. ☐

While Bonami-Beckner provides a link between II and VAR, proving Theorem 1.4 (sensitivity when $II(A) \to 0$) requires a more direct connection showing that small II forces the low-order Fourier spectrum to vanish. We introduce Talagrand's Inequality [Tal94]:

Theorem 2.5 (Talagrand's Inequality)**.** *For each $k = 1, 2, ...$ there is a constant $C_k < \infty$ with the following property. For monotone event $A \subset \Omega_n$ and function $f = \chi_A$, the sum of squares of Fourier coefficients for a fixed size k is bound: $\sum_{|S|=k} \widehat{f}(S)^2 \leq C_K II(A)(-\log II(A))^{k-1}$.*

We can now obtain a proof of Theorem 1.4 for monotone functions.

Proof of 1.4 for monotone events. By combining Theorem 2.5 ($II \to 0 \implies$ low-order coefficients vanish) and 1.11 (low-order coefficients vanish $\implies$ sensitivity), one obtains the result. ☐

However, we want to consider any general function $f : \Omega_n \to \mathbb{R}$. Since there is no monotonicity condition for f, we introduce the shifting operator to allow us to transform f into a monotone function.

Definition 2.6 (Shifting Operator)**.** Let $j \in [n]$ and let $f : \Omega_n \to \mathbb{R}$. For $x \in \Omega_n$, the operator κ_j is such that

$$\kappa_j f(x) = \left\{ \begin{array}{ll} \max\{f(x), f(\sigma_j x)\}, & \text{if } x_j = 1, \\ \min\{f(x), f(\sigma_j x)\}, & \text{if } x_j = 0. \end{array} \right\}$$

From this, we introduce several shifting properties for the operator κ_j:

Lemma 2.7 (Shifting properties)**.** *Let $f : \Omega_n \to \mathbb{R}$ and let $j, i \in [n]$. Then*

1. *Apply all shifts $\kappa_1 ... \kappa_n$ sequentially transformers any function f into a monotone function g.*
2. *$I_i(\kappa_j f) \leq I_i(f)$, i.e. shifting does not increase influences*
3. *$VAR(k_j f, \epsilon) \geq VAR(f, \epsilon)$, i.e. shifting does increase variance.*

Using these tools, we have the proof of 1.4 and for 1.5.

Proof sketch for Theorem 1.4. Let $f = \chi_A$. Let $g = \kappa_1 ... \kappa_n f$. By lemma 2.7, g is monotone with $II(g) \leq II(f)$, and $\mathrm{VAR}(g, \epsilon) \geq \mathrm{VAR}(f, \epsilon)$. Then, one applies the monotone case result to g, which allows one to conclude its low-order Fourier coefficients vanish: since $II(A) \to 0 \implies II(g) \to 0$, g is sensitive, meaning $\mathrm{VAR}(g, \epsilon) \to 0$. Therefore, $\mathrm{VAR}(A, \epsilon) \to 0$ and thus A is sensitive. □

Proof sketch for Theorem 1.5. For monotone functions f, it has influence $I_j(f) = 2|\widehat{f}(\{j\})|$, so $II(f) = 4 \sum_j |\widehat{f}(\{j\})|^2$. If $\inf II(F) > 0$, then the sum of squares of the first-order coefficients $\sum_{|S|=1} \widehat{f}(S)^2$ is bounded away from 0 as there is at least one $\widehat{f}(\{j\})^2$ must be significant. Thus, by theorem 1.11, this implies that the sequence is not asymptotically noise sensitive. □

3 Majority function

This section establishes the connection between noise (in)sensitivity and stability of monotone functions and their correlation with majority functions (both weighted and unweighted). It essentially argues that functions robust against noise (noise-stable) behave like weighed majorities, while functions that are highly sensitive to noise might be those that differ significantly from any majority-like computation. This provides interesting connections between majority-like structures in functions and noise (in)sensitivity.

3.1 Correlation with Unweighted majority

The initial focus is on the standard majority function (unweighted) $M = M_n = M_{[n]} = \mathrm{sign}(\sum_{j=1}^{n}(2x_j - 1))$ on all n variables and its variants $M_K(x)$ restricted to subset $K \subset [n]$. Particularly, the central aim is to establish a quantitative connection between how much influence the variables have on a monotone function (measured by $I(f)$ and $II(f)$ and how well this function f correlates with the outcome of a simple majority vote among its inputs.

We start by expressing both the total influence $I(f)$ and the correlation $\mathbb{E}[fM]$ in terms of $\overline{f}(k)$, the average value of f over inputs x with exactly k ones:

$$\overline{f}(k) = \binom{n}{k}^{-1} \sum_{|x|=k} f(x)$$

Think of $\overline{f}(k)$ as the average value of f on different levels of the Hamming cube. For monotone f, let $I(f)$ be

$$I(f) = 2^{-n} \sum_{k} \binom{n}{k} \overline{f}(k)(2k-n)$$

a sum weighted by $(2k-n)$. This relates total influence to a weighted version of the asymmetry, where levels further from the middle contribute more. Let $\mathbb{E}[fM]$ be the sum of differences $(\widehat{f}(k) - \widehat{f}(n-k) > n/2$:

$$\mathbb{E}[fM] = 2^{-n} \sum_{k>n/2} \binom{n}{k} (\overline{f}(k) - \overline{f}(n-k))$$

This shows the correlation depends on the asymmetry of the function's average value above and below the middle layer ($k = n/2$) of the Boolean cube.

Theorem 3.1. *Let $f : \Omega_n \to [0,1]$ be monotone. Then*

$$I(f) \leq \sqrt{n}\mathbb{E}[fM]\Big(1 + \sqrt{-\log(\mathbb{E}[fM])}\Big)$$

for some universal constant C.

This is the main theoretical result of this subsection, which proves a bound on the total influence of the correlation with the majority function. Essentially, if a monotone function has a very low correlation with the majority, its total influence must also be relatively small, and conversely, a high total influence $I(f)$ requires a substantial correlation $E[fM]$ with the majority function.

Proof. (Proof Sketch): Given $I(f)$ and $\mathbb{E}[fM]$, we introduce a cut-off level $k(\lambda) = (n + \lambda\sqrt{n})/2$ to split the sum for $I(f)$. Bound the sum $k \leq k(\lambda)$ using $E(fM)$ and the weight $(2k(\lambda) - n) = \lambda\sqrt{n}$: $\lambda\sqrt{n}\mathbb{E}(fM)$. Bound the tail sum $k > k(\lambda)$ using the fact that $0 \leq \overline{f}(k) - \overline{f}(n-k) \leq 1$ and a standard Gaussian approximation for the tail of the binomial distribution weighted by $(2k-n)$. They obtain this:

$$2^{-n}\binom{n}{k}(2k-n) \leq C_1 \exp(-\frac{(2k-n)^2}{C_2 n})$$

for constants $C_1, C_2 > 0$. After this, choose $\lambda \propto \sqrt{-\log \mathbb{E}(fM)}$ to make the tail bound comparable, which allows one to obtain $I(f) \leq C\sqrt{n}E(fM)(1 + \sqrt{-\log \mathbb{E}(fM)}$. □

The corollary 3.2 extends this result to arbitrary subsets $K \subset [n]$ and M_K.

Corollary 3.2. *(Generalized 3.1) Let $K \subset [n]$ and suppose that $f : \Omega_n \to [0,1]$ is monotone. Then*

$$I_k(f) \leq C\sqrt{|K|}\mathbb{E}(fM_K)(1 + \sqrt{-\log(\mathbb{E}[fM_K])})$$

where C is some universal constant.

The proof of 3.2 is done by defining a new function f_K on the sub-cube Ω_K by averaging f over the variables not in K. Afterwards, you apply theorem 3.1 to f, letting $I(f_K)$ correspond to $I_K(f)$. We can now establish a proof for Theorem 1.8:

Proof Sketch for 1.8. Sort the influences $I_j(f)$ in decreasing order, then apply 3.2 to the set $K = \{1, ..., k\}$ of variables with the k largest influences to bound the partial sum

$$S_k = \sum_{j=1}^{k} I_j(f) \leq C\sqrt{k}\Lambda(f)\Big(1 + \sqrt{-\log \Lambda(f)}\Big)$$

Then $I_k = S_k - S_{k-1}$. One can bound $II(f) = \sum_k I_k(f)^2$ by summing the squares based on the bounds derived for the partial sums, i.e., $\approx \sum(S_k - S_{k-1})^2$. Since $S_k \approx \sqrt{k}$, the difference I_k behaves roughly like $1/\sqrt{k}$, which means to $\sum I_k(f)^2 \approx \log n$. From this, one obtains the $\Lambda(f)^2(1 - \log \Lambda(f)) \log n$ bound. The noise sensitivity goes to 0 because the RHS goes to 0. □

The consequences of Theorem 1.7 link low correlation with all unweighted majority ($\Lambda(f)$ small) to a small sum of squared influences $II(f)$, which via Theorem 1.4 provides a sufficient condition for noise sensitivity in monotone functions. This section concludes with the following theorem that the standard majority function M_n actually maximizes the total influence $I(f)$ among all monotone events $A \subset \Omega_n$.

Theorem 3.3. *Majority function maximizes the total influence $I(f)$ among all monotone events.*

3.2 Correlation with Weighted Majority

Now, we will bring back the concept of uniform stability, which is introduced as the opposite of sensitivity: for an event $A \subset \Omega_n$ and configuration $x \in \Omega_n$, $P[A\Delta N_\epsilon A] \to 0$ uniformly as $\epsilon \to 0$. As shown in Theorem 1.11, this corresponds to the Fourier mass being concentrated on high-degree coefficients. We will now show the connection between uniform stability and weighted majority in Theorem 1.10 and their correlation with monotone events in Theorem 1.8. Consider the weighted majority functions $M_w(x) = \text{sign}(\sum_{j=1}^{n} w_j(2x_j - 1))$ and their generalized forms $M_{w,s}$ (comparing their weighted sum to a threshold s). We will now prove Theorem 1.10:

Proof sketch of Theorem 1.10. For $f(x) = \sum_{j=1}^{n} w_j(2x_j - 1)$, we first analyze the noise effect. Applying ϵ-noise N_ϵ to x is equivalent to flipping the sign $(2x_j - 1)$ for a random set $J \subset \{1, ..., n\}$ where each j is included in J independently with probability ϵ. The difference in the weighted sum $f(x) - f(N_\epsilon(x))$ has roughly the same distribution as $2Y(J)$ for a random variable $Y(J) = w_j(2x_j - 1)$. The event $M\Delta N_\epsilon M$ occurs if $f(x)$ is close to the threshold s_0 and $|f(x) - f(N_\epsilon(x))| \approx |2Y(J)|$ is large enough to cross the threshold. From this, the core idea is to show that $f(x)$ is unlikely to be very close to the threshold s_0. We can bound the probability by choosing the right ϵ: for any $a > 0$

$$P[M\Delta N_\epsilon M] \leq P[|f(x) - s_0| \leq 2a] + P[|2Y(j)| \geq 2a]$$

The remaining parts of the proof use concentration and anti-concentration bounds to partition the probability space and exploit independence and negative correlation properties. From this, they relate $P[|f(x) - s_0| \leq 2a]$ to the probability $P[|Y(j)| \geq a] \leq \delta$ using the variance $W(J) = \sum_{j \in J} w_J^2$

and decomposing the noise process, and end up showing that $P[|f(x) - s_0| \leq 2a]$ is bounded by a function of ϵ and δ. Letting $\delta = \epsilon^{1/4}$ leads us to bound

$$P[|f(x) - s_0| \leq 2a] + P[|2Y(j)| \geq 2a] = O(\epsilon^{1/4})$$

and thus the final result for any weighted majority M. Essentially, it shows that the accumulated noise is unlikely to bridge the gap between the original function value $f(x)$ and the threshold s_0. □

Orthogonality of Sensitive and Stable Functions

Lemma 3.4. *If $\{A_m\}$ is asymptotically noise-sensitive and $\{B_m\}$ is uniformly noise-stable, they are asymptotically uncorrelated:* $\lim_{m\to\infty} Cov(\chi_{A_m}, \chi_{B_m}) = \lim_{m\to\infty} P[A_m \cap B_m] - P[A_m]P[B_m] = 0$.

Proof Sketch. Using Theorem 1.11, let $f_m = \chi_{A_m}$ and $g_m = \chi_{B_m}$. Noise sensitivity means $\sum_{1\leq|S|\leq k} \widehat{f_m}(S)^2 \to 0$ for fixed k. Uniform stability means $\sum_{S\neq 0} \widehat{g}_m(S)^2 \to 0$ uniformly as $k \to \infty$. The covariance is $Cov(f_m g_m) = \sum_{S\neq\emptyset} \widehat{f}_m(S), \widehat{g}_m(S)$. Split the sum at k. From this, the sum for $|S| \leq k$ is small because $\widehat{f}_m(S)$ terms are small and the sum of $|S| > k$ can be made small (by choosing large k) because $\widehat{g}_m(S)$ terms are small. Thus, the sum goes to zero. □

Knowing that weighted majorities are stable by Theorem 1.10, Lemma 3.4 shows that any noise-sensitive sequence must be asymptotically uncorrelated with all weighted majority functions. This leads us to the proof of Theorem 1.8.

Proof sketch of 1.8. For the $\Longleftarrow$ direction, if $\{A_m\}$ is sensitivity, since weighted majorities M_w are stable from 1.10 (stability of weight majorities), then 3.4 (uncorrelation of sensitive and stable events) implies that their correlation goes to zero. For the $\Longrightarrow$ direction, assume $\lim_{m\to\infty} \overline{\Lambda}(A_m) = 0$. Suppose for the sake of contradiction that $\{A_m\}$ is non-sensitive. Since it's monotone, Theorem 1.5 implies that $\inf_m II(A_M) > 0$. This means that the influence vector $I^{A_m} = (I_1(A_m), ..., I_N(A_m))$ has norm $\|I^{A_m}\|_2$ bounded below by some $\delta > 0$. From this, we use a theorem from [Tal94] which states that for monotone events, low correlation implies low inner product of influence vectors. The following proposition allows us to combine our previous results:

Proposition 3.5. *There is an absolute constant $c > 0$ such that the influence vector of the corresponding majority M_w has a non-negligible projection onto w: $\langle w, I^w\rangle \geq c$. This holds for every $n = 1, 2, ...$ and every $w \in \mathbb{R}^n$ with non-negative coordinates and $\|w\|_2 = 1$*

Proof Sketch. This connects the inner product to the expectation of the absolute value of the linear function $f(x) = \sum_{j=1}^n w_j(2x_j - 1)$ via Fourier analysis: $\langle w, I^w\rangle = \langle \widehat{f}, \widehat{M_w}\rangle = \mathbb{E}[fM_w] = \mathbb{E}[|f(x)|]$. Then, one can find a lower bound for $\mathbb{E}[|f(x)|]$. □

From here, we combine: if A_m is not sensitive, let $w_m = I^{A_m}/\|I^{A_m}\|_2$. By proposition 3.5, we have $\langle w_m, I^{w_m}\rangle \geq c$. Then

$$\langle I^{A_m}, I^{w_m}\rangle = \|I^{A_m}\|\langle w_m, I^{w_m}\rangle \geq \delta c > 0$$

By Talagrand's theorem, this implies the correlation $\mathbb{E}[\chi A_m M_{w_m}]$ is bounded away form zero, contradicting the assumptions hat $\widetilde{\Lambda}(A_m) = \max_w |E(\chi_{A_m} M_w)| \to 0$. Thus, the initial supposition (not sensitive) must be false. □

4 Bond Percolation Application

This section applies the concepts of noise sensitivity developed earlier to the specific example of bond percolation on a rectangular grid. Consider the rectangle R to be $(m+1) \times m$ in the grid in $\mathbb{Z}^2$. Let Ω be the set of all functions from the set of edges of R (bonds) to $\{0,1\}$, i.e., $\{0,1\}^E$. Let $x \in \Omega$ be a random configuration selected based on the uniform measure, which involves assigning each edge e in the grid as either 'open' with $x(e) = 1$ or 'closed' with $x(e) = 0$ with probability 1/2. The event of interest is $C = C_m \subset \Omega$, which is the event that there is a left-right crossing of R, i.e., the event that there exists a path of open edges connecting the left and right boundaries of R. From duality, $P(C_m) = 1/2$. The main goal of this section is to prove the following theorem:

Theorem 4.1. *The crossing events C_m are asymptotically noise sensitive:* $\lim_{m\to\infty} \phi(C_m, \epsilon) = 0$

This means that knowing the state (open/closed) of most edges but with a small fraction subject to random noise (flipping state with small probability ϵ) gives almost no information about whether a crossing exists or not for large m.

Proof Sketch of Theorem 4.1. The proof strategy relies on showing that the sum of squared influences $II(C_m) = \sum_{e\in E} I_e(C_m)^2$ goes to zero as the grid size $m \to \infty$, which by Theorem 2.4 or Theorem 1.4 implies asymptotic noise sensitivity. This shows that bounding $II(C_m)$ can be achieved by bounding the maximum correlation $\Lambda(C_m) = \max_K E[\chi_C M_K]$ for an arbitrary subset of edges K between the crossing event C_m and unweighted majority functions M_K over subset of edges K. Showing $\Lambda(C_m)$ vanishes sufficiently fast will imply $II(C_m) \to 0$.

Algorithmic Sampling The main technical part of the proof is to estimate correlation $E[\chi_C M_K]$ for K being subsets of edges in the right half of the rectangle E_r. Firstly, to tackle the challenge of analyzing a crossing event based on random open edges in the rectangle, we devise an algorithm to generate a random configuration x. It explores open clusters starting from the left boundary V_1. When an edge e is encountered, its state $y(e)$ is determined using a random bit from pre-defined sequences ω^K if $e \in K$ and $\widehat{w}^K$ otherwise. Edges never encountered ($E - VISITED$) are assigned states randomly and independently. This algorithm produces a random configuration x that is uniformly distributed, which validates the method. Crucially, the crossing event $x \in C$ depends only on the explored part y (reveal information). The formal algorithm is as follows. Note that an edge $e \in K$ only enters the explored set $VISITED$ if it is connected back to the left boundary via an open path within the configuration x.

Percolation Theory and Correlation The proof makes essential use of a result from Russo-Seymour-Walsh (RSW) Theorem, which provides bounds on crossing probabilities in rectangles:

Theorem 4.2 (Russo-Seymour-Welsh Theorem (Informal))**.** *The probability that a vertex in R is connected to another vertex at Euclidean distance r within the configuration is bounded by $Cr^{1/\rho}$ for some constants $C, \rho > 0$.*

Using the RSW result, for each edge $e \in K$ (right half E_r) has a small probability of being visited: $\leq Cm^{-1/\rho}$. Thus, the expected number of explored edges within K is also small:

$$\mathbb{E}[|K \cap \text{VISITED}|] \leq C|K|m^{-1/\rho}$$

Algorithm 1 Algorithmic method of randomized configuration

Require: w^K, $\widehat{w}^k$:= be two independent elements of $\Omega_{|K|}$ and $\Omega_{n-|K|}$ respectively. V_1 := set of vertices on the left boundary of R, VISITED = $\emptyset$
Ensure:
  while edge $e = (v, u) \notin$ VISITED exists for $v \in V_1$, $u \notin V_1$ **do**
    Choose edge $e = (v, u)$
    VISITED $\leftarrow$ VISITED $\cup \{e\}$
    if $e \in K$ **then**
      $y(e)$:= first bit in the sequence w^K that has not been previously used by the algorithm
    else
      $y(e)$:= first bit in the sequence $\widehat{w}^K$ that has not been previously used by the algorithm
    end if
    if $y(e) = 1$ **then**
      $V_1 \leftarrow V_1 \cup \{u\}$
    end if
  end while

The next step of the proof is to show that it is unlikely that many edges in K are explored or that the partial sums determining M_K deviate significantly from their mean during exploration. Let A_1 be the event that

$$A_1 := \{x \in \Omega : |K \cap \text{VISITED}| \geq C|K|m^{-2/(3\rho)}\}$$

By Markov's Inequality, $P[A_1] \leq C|K|m^{-1/\rho}$ with high probability, meaning that the actual number of visited edges is small. Our next challenge is to connect small $|K \cap VISITED|$ to small correlation. We know that the majority function $M_K(x)$ depends on the values $x(e)$ for all $e \in K$, however, the algorithm only reveals $y(e) = x(e)$ for $e \in K \cap VISITED$. To rectify this, let A_2 be the event that the sum $\sum_{e \in K \cap VISITED} y(e)$ deviates significantly from its mean $|K \cap VISITED|/2$. By standard concentration bounds, $P[A_2]$ decays super-polynomial with m and is thus small. Finally, if neither A_1 nor A_2 occurs, then the number of explore edges in K is small and the sum $\sum_{e \in VISITED \cap K} y(e)$ is close to its mean $|VISITED \cap K|/2$. Thus, the overall majority function $M_K(x)$ depends on both the revealed $y(e)$ for $e \in K \cap VISITED$ and the unrevealed $z(e)$ for $e \in K \setminus VISITED$. Since $|K \cap VISITED|$ is small, the revealed part gives little information about the final outcome of M_K. This is formalized by bounding the condition expectation given the explored edges y: $\mathbb{E}[M_K(x)|y] \leq O(1)m^{-1/(3\rho)} \log m$ when A_1, A_2 do not occur. Since χ_C depends only on y, combining these bounds obtains:

$$\mathbb{E}[\chi_C M_K] = \mathbb{E}[\chi_C \mathbb{E}[M_K(x)|y]] \leq O(1)m^{-1/(3\rho)} \log m$$

Since C_m is monotone, we can apply Corollary 3.2 to yield:

$$I_k(C) \leq O(1)\sqrt{|K|}m^{-1/(3\rho)}(\log m)^{3/2}$$

Following the derivation in the proof of Theorem 1.7's proof:

$$II(C_m) \leq O(1)m^{-2/(3\rho)}(\log m)^4$$

Since $\rho > 0$, $II(C_m) \to 0$ as $m \to \infty$. Thus, by Theorem 2.4 or 1.4, the crossing event C_m is asymptotically noise sensitive. □

References

[BKS99] Itai Benjamini, Gil Kalai, and Oded Schramm. Noise sensitivity of boolean functions and applications to percolation. *Publications Mathématiques de l'Institut des Hautes Études Scientifiques*, 90:5–43, 1999.

[BKS11] Itai Benjamini, Gil Kalai, and Oded Schramm. First passage percolation has sublinear distance variance. *Selected Works of Oded Schramm*, pages 779–787, 2011.

[Tal94] Michel Talagrand. On russo's approximate zero-one law. *The Annals of Probability*, pages 1576–1587, 1994.

[Tal95] Michel Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Publications Mathématiques de l'Institut des Hautes Etudes Scientifiques*, 81:73–205, 1995.

[Tsi99] Boris Tsirelson. Fourier-walsh coefficients for a coalescing flow (discrete time). *arXiv preprint math/9903068*, 1999.

[WS99] John W Wenzel and Mark E Siddall. Noise. *Cladistics*, 15(1):51–64, 1999.