

Noise sensitivity of Boolean functions and percolation, by Christophe Garban and Jeffrey E. Steif, Institute of Mathematical Statistics Textbooks, Vol. 5, Cambridge University Press, New York, 2015, xvii+203 pp., ISBN 978-1-107-07643-3

The notion of *noise sensitivity* was introduced by Benjamini, Kalai, and Schramm in 1998 [3], in the context of percolation theory. Since then noise sensitivity has found applications in many fields, including concentration of measure, social choice theory, and theoretical computer science.

Say we have n people voting for one of two candidates, \mathcal{D} or \mathcal{R} . One can write their choices as $\omega = (\omega(1), \dots, \omega(n)) \in \{\mathcal{D}, \mathcal{R}\}^n$, and the election is decided by some voting rule $f(\omega) \in \{\mathcal{D}, \mathcal{R}\}$. Errors occur during the ballot counting process, and the election is actually decided according to slightly perturbed data ω_ϵ . The subject of noise sensitivity studies how susceptible the voting rule is to the noise as a function of the number of voters n and the level of noise ϵ . Intuitively, $f(\cdot)$ is noise sensitive if even for a small level of noise, given enough voters, $f(\omega_\epsilon)$ gives us very little information on the true outcome of the election $f(\omega)$. On the other hand, $f(\cdot)$ is noise stable if for a small enough level of noise the election is decided by the unperturbed data ω , regardless of the number of voters.

A *Boolean function* is a function from the hypercube $\Omega_n = \{-1, 1\}^n$ into $\{-1, 1\}$ (or $\{0, 1\}$). We call the elements of Ω_n *configurations*. Consider the hypercube Ω_n endowed with the uniform measure $\mathbf{P}^n = (\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)^{\otimes n}$. Let $\omega \in \Omega_n$ be a configuration sampled according to \mathbf{P}^n , and let ω' be an independent configuration. For every $\epsilon > 0$, denote by ω_ϵ the “noisy” configuration obtained from ω by re-sampling each bit independently with probability ϵ , i.e., independently for every $x \in [n] := \{1, 2, \dots, n\}$,

$$\omega_\epsilon(x) = \begin{cases} \omega(x) & \text{with probability } 1 - \epsilon, \\ \omega'(x) & \text{with probability } \epsilon. \end{cases}$$

We denote by \mathbf{P} the joint distribution of ω and ω_ϵ , and by \mathbf{E} the expectation with respect to \mathbf{P} .

Next we introduce two main concepts regarding sensitivity to noise.

Definition 1. Let m_n be an increasing sequence of integers, and let $\{f_n\}$ be a sequence of functions $f_n : \Omega_{m_n} \rightarrow \{-1, 1\}$. The sequence $\{f_n\}$ is called *noise sensitive* if for every $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} (\mathbf{E}[f_n(\omega)f_n(\omega_\epsilon)] - \mathbf{E}[f_n(\omega)]^2) = 0.$$

Note that since f_n are Boolean, zero covariance implies independence. Next we present a trivial but important example of a noise sensitive sequence of functions.

Example 2. Define the parity function, $\mathbf{PAR}_n(\omega(1), \dots, \omega(n)) = \prod_{i=1}^n \omega(i)$. Note that for every i , with probability $1 - \epsilon$, $\omega(i) = \omega_\epsilon(i)$, and thus their product is 1. With probability ϵ , $\omega(i)$ and $\omega_\epsilon(i)$ are independent and thus their product has zero

2010 *Mathematics Subject Classification*. Primary 60-02, 60K35; Secondary 60D05.

Define the sequence of events $A_n = \{\text{there is a left right crossing}\}$, and the sequence of Boolean functions $f_n : \{0, 1\}^{\mathcal{E}(B_n)} \rightarrow \{0, 1\}$,

$$f_n(\omega) = \mathbb{1}_{A_n}(\omega) = \begin{cases} 1 & \text{if } \omega \in A_n, \\ 0 & \text{else.} \end{cases}$$

A fundamental result of Benjamini, Kalai, and Schramm [3, Theorem 1.2] states that $\{f_n\}$ is noise sensitive.

In the heart of the theory of noise sensitivity lies the notion of *influence*. This concept first arose in political science to measure the power of individual voters in a voting scheme. For a configuration $\omega \in \Omega_n$, denote by ω^k the configuration obtained from ω by flipping the k th coordinate,

$$\omega^k(x) = \begin{cases} \omega(x) & \text{if } x \neq k, \\ -\omega(x) & \text{if } x = k. \end{cases}$$

Abbreviate the discrete partial derivatives $\forall k \in [n], \nabla_k f(\omega) = f(\omega) - f(\omega^k)$.

Definition 5. The *influence* of the k th coordinate is defined as $\mathbf{I}_k(f) := \|\nabla_k f\|_1 = \mathbf{E}[\|\nabla_k f\|]$.

For Boolean functions (to $\{-1, 1\}$), $\nabla_k f \in \{-2, 0, 2\}$, and thus

$$\mathbf{I}_k(f) = 2\mathbf{P}[f(\omega) \neq f(\omega^k)].$$

This relates to the classical concept of pivotal edges from percolation theory. Say that $k \in [n]$ is *pivotal* for f in the configuration ω if $f(\omega) \neq f(\omega^k)$. Note that the event $\{f(\omega) \neq f(\omega^k)\}$ is independent of $\omega(k)$. If we denote by

$$\mathcal{P} = \mathcal{P}_f(\omega) := \{i \in [n] : i \text{ is pivotal for } f \text{ in } \omega\},$$

we can write $\mathbf{I}_k(f) = 2\mathbf{P}[k \in \mathcal{P}]$.

Definition 6. The *total influence*, $\mathbf{I}(f)$, is defined by

$$\mathbf{I}(f) := \sum_{k=1}^n \mathbf{I}_k(f) = 2\mathbf{E}[\|\mathcal{P}\|].$$

We are now in a position to state one of the main theorems in the theory of noise sensitivity:

Theorem 7 ([3, Theorem 1.3]). *Let $\{f_n\}$ be a sequence of Boolean functions. If*

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n \mathbf{I}_k(f_n)^2 = 0,$$

then $\{f_n\}$ is noise sensitive.

The converse is not true in general, as can be seen by Example 2. However, it is true for a sequence of monotonic functions.

Definition 8. A function $f : \Omega_n \rightarrow \mathbb{R}$ is *monotone* if $f(x) \leq f(y)$ whenever $\forall j \in [n], x(j) \leq y(j)$.

Theorem 9 ([3, Theorem 1.4]). *Let $\{f_n\}$ be a sequence of monotone Boolean functions. If*

$$\inf_n \sum_{k=1}^n \mathbf{I}_k(f_n)^2 > 0,$$

then $\{f_n\}$ is not noise sensitive (but not necessarily noise stable).

In recent years, Theorem 7 and other results such as the KKL theorem (Theorem 11 below), were generalized by Keller, Mossel, and Sen [13, 14] to continuous distributions such as the Gaussian measure.

TOOLS FROM DISCRETE HARMONIC ANALYSIS

The main tool in the theory we are discussing is the *Fourier–Walsh* expansion. It is useful to define the expansion for non-Boolean functions, thus we consider the space $L^2(\Omega_n)$ of real valued functions. The inner product is defined relative to the uniform measure \mathbf{P}^n , $\langle f, g \rangle := \sum_{\omega} 2^{-n} f(\omega)g(\omega) = \mathbf{E}[fg]$. As the complete orthonormal Fourier base we take the Walsh functions: for any subset $S \subset [n]$, define

$$\chi_S(\omega) := \prod_{i \in S} \omega(i),$$

with the convention that $\chi_{\emptyset} \equiv 1$. Thus any function can be decomposed according to the Fourier–Walsh series

$$f = \sum_{S \subset [n]} \hat{f}(S)\chi_S := \sum_{S \subset [n]} \langle f, \chi_S \rangle \chi_S.$$

This allows us to study concentration and noise sensitivity of functions by their Fourier–Walsh spectrum. The catch phrase one can learn from Benjamini, Kalai, and Schramm [3] is that functions of high frequency are noise sensitive and functions of low frequencies are noise stable.

In the case of a monotonic function $f : \Omega_n \rightarrow \{-1, 1\}$, we get

$$\begin{aligned} \mathbf{I}_k(f) &= \mathbf{E}[|\nabla_k f|] = \mathbf{E}[|\nabla_k f| \mathbb{1}_{k \in \mathcal{D}}] \\ &= \mathbf{E}[(f(\omega) - f(\omega^k))\omega(k) \mathbb{1}_{k \in \mathcal{D}}] = 2\mathbf{E}[f\chi_{\{k\}} \mathbb{1}_{k \in \mathcal{D}}], \end{aligned}$$

where the third equality is due to the monotonicity of f . For every function

$$\mathbf{E}[f \cdot \omega(k) \mathbb{1}_{k \notin \mathcal{D}}] = 0,$$

we get for monotonic Boolean functions $\hat{f}(\{k\}) = \mathbf{I}_k/2$. The Cauchy–Schwarz inequality yields for monotonic functions that $\mathbf{I}(f) \leq \sqrt{n}$. A calculation shows the majority function has total influence equal to \sqrt{n} , hinting that the majority function is maximal among all monotonic functions.

INFLUENCE, CONCENTRATION OF MEASURE, AND HYPERCONTRACTIVITY

The study of influences is very important in many scientific disciplines, e.g., learning, information theory, and social choice theory.

For a Boolean function (to $\{-1, 1\}$), a calculation shows that

$$\widehat{\nabla_k f}(S) = \begin{cases} 2\hat{f}(S) & \text{if } k \in S, \\ 0 & \text{otherwise.} \end{cases}$$

Since $\nabla_k f \in \{-2, 0, 2\}$, $\|\nabla_k f\|_2^2 = \mathbf{E}[|\nabla_k f|^2] = 4\mathbf{E}[|\nabla_k f|] = 4\|\nabla_k f\|_1$, and together with Parseval’s formula we obtain that

$$(1) \quad \mathbf{I}_k(f) = \sum_{S \subset [n]: k \in S} \hat{f}(S)^2 \quad \text{and} \quad \mathbf{I}(f) = \sum_{S \subset [n]} |S|\hat{f}(S)^2.$$

The variance of a function f can be easily represented with Parseval’s formula:

$$(2) \quad \mathbf{Var}[f] = \mathbf{E}[f^2] - \mathbf{E}[f]^2 = \langle f, f \rangle - \langle f, \chi_{\emptyset} \rangle^2 = \sum_{S \subset [n]} \hat{f}(S)^2 - \hat{f}(\emptyset)^2.$$

This together with (1) immediately gives us the *discrete Poincaré inequality* for Boolean functions.

Theorem 10. *For any Boolean function $f : \Omega_n \rightarrow \{-1, 1\}$,*

$$\mathbf{Var}[f] \leq \sum_k \mathbf{I}_k(f).$$

This means that

$$\max_i \mathbf{I}_i(f) \geq \frac{\mathbf{Var}[f]}{n}.$$

The celebrated KKL theorem of Kahn, Kalai, and Linial [11] gives a logarithmic correction to the above inequality. Though it might seem that a *log* correction is not significant, the fact that it diverges was enough for Kahn, Kalai, and Linial to show that in an idealized voting system with two candidates it is enough to bribe $o(1)$ fraction of the voters in order to control the elections with high probability.

Theorem 11 ([11]). *There is a universal constant $c > 0$ such that for any Boolean function*

$$\max_i \mathbf{I}_i(f) \geq c \mathbf{Var}[f] \frac{\log n}{n}.$$

Remarkably, the logarithmic correction is sharp, as shown by the tribes example of Ben-Or and Linial [2]: Partition $[n]$ into blocks of length $\log_2(n) - \log_2(\log_2(n))$, and connect the leftover arbitrarily. Define f_n to be 1 if there is a block of all 1's and 0 otherwise.

An inequality of Talagrand [20], which improves upon some results of Kahn, Kalai, and Linial [11], proved to be instrumental in the study of fluctuations in *first passage percolation*. For $p \in [0, 1]$, consider the measure $\mathbf{P}_p^n = ((1 - p)\delta_{-1} + p\delta_1)^{\otimes n}$ on Ω_n .

Theorem 12 ([20, Theorem 1.5]). *Let $f : \Omega_n \rightarrow \mathbb{R}$, then for any $p \in [0, 1]$ there is a $K = K(p) \in (0, \infty)$ such that*

$$\mathbf{Var}_p(f) \leq K \cdot \sum_{k=1}^n \frac{\|\nabla_k f\|_2^2}{1 + \log(\|\nabla_k f\|_2 / \|\nabla_k f\|_1)}.$$

Talagrand's inequality also proved useful in the study of percolation isoperimetry [5, 9, 19], showing concentration of the Cheeger constant and proving a weak version of a conjecture by Benjamini.

Talagrand's approach for concentration estimates, such as Theorem 12, is based on *hypercontractivity*. In [7], Sourav Chatterjee presents a modern general approach to hypercontractivity. A semigroup P_t is called hypercontractive if for any $p > 1$ and $t > 0$ there exists a $q(t, p) > p$ such that

$$\|P_t f\|_{L^q} \leq \|f\|_{L^p}.$$

This approach began with Nelson's [18] proof of the hypercontractivity of the Ornstein–Uhlenbeck semigroup and Gross's proof [10] that a semigroup is hypercontractive if the associated Dirichlet form satisfies a logarithmic Sobolev inequality.

First passage percolation (FPP) is a model for a random metric on a graph; see [1] for a thorough review of FPP. Consider the edge set \mathcal{E} of \mathbb{Z}^d . For every edge $e \in \mathcal{E}$ we associate a weight $\omega(e)$ distributed according to some measure μ , such that $\{\omega(e)\}_e$ are i.i.d. For any two vertices $x, y \in \mathbb{Z}^d$ and any path connecting

them $\gamma : x \mapsto y$, $\gamma = (e_1, \dots, e_l)$, we associate a length $\lambda(\gamma) = \sum_{i=1}^l \omega(e_i)$. This defines in general a semimetric $d_\omega(x, y) = \inf_{\gamma: x \mapsto y} \lambda(\gamma)$. By Kingman's subadditive ergodic theorem [16, 17], under weak assumptions on μ , one obtains that balls in the metric converge in the Hausdorff sense to a ball \mathcal{D} , of a deterministic norm on \mathbb{R}^d ,

$$\lim_{n \rightarrow \infty} D_n := \lim_{n \rightarrow \infty} \frac{1}{n} \{x \in \mathbb{Z}^d : d_\omega(0, x) < n\} = \mathcal{D}.$$

This can be thought of as a geometric version of the law of large numbers. Many open questions remain about the asymptotic shape \mathcal{D} ; e.g., if μ is a continuous distribution, is \mathcal{D} strictly convex? What are the conditions for μ under which corners appear in \mathcal{D} ? The question relevant to our discussion is that of the fluctuations around the limit shape.

Benjamini, Kalai, and Schramm [4] studied the simple case in which $\mu = p\delta_a + (1 - p)\delta_b$, where $0 < a < b$. This allows us to study the metric d_ω in the framework of Boolean functions. They prove that the variance of the FPP metric is sublinear and thus is also little o of the expectation.

Theorem 13 ([4, Theorem 1]). *There is a constant $C = C(d, a, b)$ such that for any $v \in \mathbb{Z}^d$, $\|v\|_1 > 2$,*

$$\text{Var}_p[d_\omega(0, v)] \leq C \frac{\|v\|_1}{\log \|v\|_1}.$$

This theorem was generalized by Damron, Hanson, and Sosoe [8] under weak assumptions on μ . This is a good point to demonstrate a simple but powerful idea developed by Benjamini, Kalai, and Schramm. Since we took as our weight distribution a measure supported on two points $\{a, b\}$, any path γ connecting 0 and v for which $\lambda(\gamma) = d_\omega(0, v)$ must satisfy

$$a|\gamma| \leq d_\omega(0, v) \leq b\|v\|_1.$$

Thus we get $|\gamma| \leq \frac{b}{a}\|v\|_1$. This means that we need to only look at a finite box around the origin in order to find $d_\omega(0, v)$. Moreover, only edges contained in all minimizing paths (not necessarily unique) can have nonzero influence. Changing the weight of a path from a to b or from b to a can change d_ω by no more than $|b - a|$. This leads to

$$\sum_e \|\nabla_e d_\omega(0, v)\|_2^2 \leq |\mathcal{P}|(b - a)^2 \leq \frac{b}{a}(b - a)^2 \|v\|_1,$$

providing justification to the nominator in Theorem 13. This simple idea of controlling the sum of influences by some geometric a priori knowledge has proved useful in many applications, e.g., concentration of maximal cardinality matchings for general graphs [6].

Note that if we choose μ to be the exponential distribution, \mathcal{D} is the limiting shape of the *Eden model*, which is the aggregation process one gets by adding edges according to the uniform distribution on the boundary of the aggregate. Kardar, Parisi, and Zhang [12] explain with their KPZ equation that D_n should have $n^{1/3}$ fluctuations around the limiting shape \mathcal{D} , which is very far from the known bound given by Benjamini, Kalai, and Schramm.

ABOUT THE BOOK

The book under review is a refined version of lecture notes that have been circulating for a few years. Considerable effort was made to make the book as thorough and concise as possible but still readable and friendly. There are many interesting and important subjects covered in the book that we did not discuss in this review. We believe this book is worthy for any departmental or scholarly library. It is clear that it will turn out to be the “go to” source for studying the subject of noise sensitivity of Boolean functions.

REFERENCES

- [1] Antonio Auffinger, Jack Hanson, and Michael Damron, *50 years of first passage percolation*, 2015. [arXiv:1511.03262](#), 50.
- [2] Michael Ben-Or and Nathan Linial, *Collective coin flipping*. Leibniz Center for Research in Computer Science, Department of Computer Science, Hebrew University of Jerusalem, 1987.
- [3] Itai Benjamini, Gil Kalai, and Oded Schramm, *Noise sensitivity of Boolean functions and applications to percolation*, Inst. Hautes Études Sci. Publ. Math. **90** (1999), 5–43 (2001). MR1813223
- [4] Itai Benjamini, Gil Kalai, and Oded Schramm, *First passage percolation has sublinear distance variance* [MR2016607], Selected Works of Oded Schramm. Volume 1, 2, Sel. Works Probab. Stat., Springer, New York, 2011, pp. 779–787, DOI 10.1007/978-1-4419-9675-6_26. MR2883392
- [5] Marek Biskup, Oren Luidor, Eviatar B. Procaccia, and Ron Rosenthal, *Isoperimetry in two-dimensional percolation*, Comm. Pure Appl. Math. **68** (2015), no. 9, 1483–1531, DOI 10.1002/cpa.21558. MR3378192
- [6] Avrim Blum, Ioannis Caragiannis, Nika Haghtalab, Ariel D Procaccia, Eviatar B Procaccia, and Rohit Vaish, *Opting into optimal matchings*, in Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 2351–2363, SIAM, Philadelphia, PA, 2017.
- [7] Sourav Chatterjee, *Superconcentration and related topics*, Springer Monographs in Mathematics, Springer, Cham, 2014. MR3157205
- [8] Michael Damron, Jack Hanson, and Philippe Sosoe, *Sublinear variance in first-passage percolation for general distributions*, Probab. Theory Related Fields **163** (2015), no. 1-2, 223–258, DOI 10.1007/s00440-014-0591-7. MR3405617
- [9] Julian Gold, *Isoperimetry in supercritical bond percolation in dimensions three and higher*. [arXiv:1602.05598](#), 2016.
- [10] Leonard Gross, *Logarithmic Sobolev inequalities*, Amer. J. Math. **97** (1975), no. 4, 1061–1083, DOI 10.2307/2373688. MR0420249
- [11] Jeff Kahn, Gil Kalai, and Nathan Linial, *The influence of variables on boolean functions*, in 29th Annual Symposium on Foundations of Computer Science, 1988, pp. 68–80, IEEE, 1988.
- [12] Mehran Kardar, Giorgio Parisi, and Yi-Cheng Zhang, *Dynamic scaling of growing interfaces*, Phys. Rev. Lett. **56** (1986), no. 9, 889.
- [13] Nathan Keller, Elchanan Mossel, and Arnab Sen, *Geometric influences*, Ann. Probab. **40** (2012), no. 3, 1135–1166, DOI 10.1214/11-AOP643. MR2962089
- [14] Nathan Keller, Elchanan Mossel, and Arnab Sen, *Geometric influences II: Correlation inequalities and noise sensitivity* (English, with English and French summaries), Ann. Inst. Henri Poincaré Probab. Stat. **50** (2014), no. 4, 1121–1139, DOI 10.1214/13-AIHP557. MR3269987
- [15] Harry Kesten, *The critical probability of bond percolation on the square lattice equals $\frac{1}{2}$* , Comm. Math. Phys. **74** (1980), no. 1, 41–59. MR575895
- [16] J. F. C. Kingman, *Subadditive ergodic theory*, Ann. Probab. **1** (1973), 883–909. With discussion by D. L. Burkholder, Daryl Daley, H. Kesten, P. Ney, Frank Spitzer and J. M. Hammersley, and a reply by the author. MR0356192
- [17] Thomas M. Liggett, *An improved subadditive ergodic theorem*, Ann. Probab. **13** (1985), no. 4, 1279–1285. MR806224
- [18] Edward Nelson, *The free Markoff field*, J. Funct. Anal. **12** (1973), 211–227. MR0343816

- [19] Eviatar B. Procaccia and Ron Rosenthal, *Concentration estimates for the isoperimetric constant of the supercritical percolation cluster*, *Electron. Commun. Probab.* **17** (2012), no. 30, 11, DOI 10.1214/ECP.v17-2185. MR2955495
- [20] Michel Talagrand, *On Russo's approximate zero-one law*, *Ann. Probab.* **22** (1994), no. 3, 1576–1587. MR1303654

EVIATAR B. PROCACCIA

DEPARTMENT OF MATHEMATICS, TEXAS A&M UNIVERSITY

E-mail address: eviatarp@gmail.com

www.math.tamu.edu/~procaccia