STRONG NOISE SENSITIVITY AND RANDOM GRAPHS

EYAL LUBETZKY AND JEFFREY E. STEIF

Abstract. The noise sensitivity of a Boolean function describes its likelihood to flip under small perturbations of its input. Introduced in the seminal work of Benjamini, Kalai and Schramm (1999), it was there shown to be governed by the first level of Fourier coefficients in the central case of monotone functions at a constant critical probability p_c.

Here we study noise sensitivity and a natural stronger version of it, addressing the effect of noise given a specific witness in the original input. Our main context is the Erdős-Rényi random graph, where already the property of containing a given graph is sufficiently rich to separate these notions. In particular, our analysis implies (strong) noise sensitivity in settings where the BKS criterion involving the first Fourier level does not apply, e.g., when $p_c \to 0$ polynomially fast in the number of variables.

1. Introduction

The concept of noise sensitivity, introduced by Benjamini, Kalai and Schramm [4], captures the notion that the value of a Boolean function of many i.i.d. variables would change under small perturbations of its input. Roughly put, it corresponds to the case where a small perturbation of the input variables via i.i.d. noise suffices to make the new value of the function asymptotically independent of its original value.

Formally, consider a sequence of functions $f_n : \Omega_n \to \{0,1\}$ paired with a sequence of probabilities p_n, where each domain $\Omega_n = \{0,1\}^{\Lambda_n}$ is a product space of Bernoulli(p_n) variables, and the sets Λ_n are finite and increasing with n. Further assume that the sequence (p_n) is *non-degenerate* in the sense that $\mathbb{P}(f_n = 1)$ is uniformly bounded away from 0 and 1. Given $\omega \in \Omega_n$ and some $\varepsilon \in (0,1)$, let ω^ε denote the result of resampling the Bernoulli(p_n) variable ω_x independently with probability ε for each $x \in \Lambda_n$. The sequence (f_n) is said to be *noise sensitive* (Sens) w.r.t. p_n if for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(f_n(\omega^\varepsilon) = 1 \mid f_n(\omega) = 1\right) - \mathbb{P}\left(f_n = 1\right) = 0\,, \tag{1.1}$$

or equivalently (recall that (f_n) is non-degenerate), $\mathrm{Cov}\left(f_n(\omega), f_n(\omega^\varepsilon)\right) \to 0$. When a function (f_n) is Sens it is natural to further discuss *quantitative noise sensitivity*, i.e., how fast can $\varepsilon \to 0$ with n such that (1.1) still holds.

2010 *Mathematics Subject Classification.* 82C43, 82B43, 60K35.

Key words and phrases. noise sensitivity of Boolean functions, random graphs.

In the setting where $p_n \equiv 1/2$ and the functions f_n are *monotone* w.r.t. the natural partial order on the hypercube Ω_n (as is notably the case for critical 2D percolation), a beautiful argument of [4] gave a criterion for noise sensitivity in terms of the first level of Fourier coefficients of f_n. Namely, (f_n) is noise sensitive if and only if $\lim_{n\to\infty} \sum_{x\in\Lambda_n} \hat{f}_n(x)^2 = 0$, where $\hat{f}_n(x)$ is the Fourier coefficient corresponding to the singleton $\{x\}$, and is also one half the probability that x is *pivotal*, i.e., flipping its value would flip the value of f_n. For more on noise sensitivity in this case, see [8] and the references therein. Unfortunately, this criterion becomes invalid when $p_n \to 0$ (e.g., formal definitions postponed, the indicator of a random graph being triangle-free satisfies the above condition and yet it is *not* noise sensitive; see [4, §6.4]), and determining noise sensitivity without it can prove to be a challenging task already for fairly simple monotone functions enjoying many symmetries.

1.1. Strong noise sensitivity.

Going back to (1.1), this is known (see §2.2) to be equivalent to having the average of $\left|\mathbb{P}\left(f_n(\omega^\varepsilon) = 1 \mid \omega\right) - \mathbb{P}\left(f_n = 1\right)\right|$ over $\{\omega : f_n(\omega) = 1\}$ tend to 0 as $n \to \infty$. That is, if (f_n) is noise sensitive then most inputs $\omega \in \Omega_n$ with $f_n(\omega) = 1$ are such that conditioning on ω will not give any substantial information on the probability that $f_n(\omega^\varepsilon) = 1$. When dealing with monotone functions, however, it is in many cases more natural and useful to condition on a *witness* for $f_n(\omega) = 1$ (for instance, a particular crossing in 2D percolation) instead of the entire configuration ω.

Definition 1.1. A 1-*witness* for a monotone function $f : \{0,1\}^\Lambda \to \{0,1\}$ is a minimal subset $W \subset \Lambda$ such that $\omega_W \equiv 1$ implies $f(\omega) = 1$.

Let $\mathcal{W}_1 = \mathcal{W}_1(f)$ denote the set of 1-witnesses of a monotone Boolean function f, and let $\mathcal{W}_0 = \mathcal{W}_0(f)$ denote its analogously defined 0-witnesses.

Perhaps surprisingly, it can be the case that (f_n) is noise sensitive and yet the probability that $f_n(\omega^\varepsilon) = 1$ substantially increases when we condition on *any particular* 1-witness in ω. This motivates the following definition.

Definition 1.2. A sequence (f_n) of monotone increasing Boolean functions is said to be *1-strongly noise sensitive* ($\textsc{StrSens}_1$) if for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \max_{W\in\mathcal{W}_1} \mathbb{P}(f_n(\omega^\varepsilon) = 1 \mid \omega_W \equiv 1) - \mathbb{P}(f_n = 1) = 0\,. \tag{1.2}$$

The notion of *0-strong noise sensitivity* ($\textsc{StrSens}_0$) is defined analogously. (Note that a sequence of increasing functions (f_n) is $\textsc{StrSens}_0$ if and only if its complement $(\overline{f_n})$ is $\textsc{StrSens}_1$, where $\overline{f_n}(\omega) = \overline{f_n(\overline{\omega})}$ with $\overline{x} = 1 - x$.)

As we will later see (and as suggested by its name), the notion of strong noise sensitivity, which addresses the subtler effect of conditioning on any *particular* witness (cf. (1.1) vs. (1.2)), indeed implies (even when $\varepsilon \to 0$) the standard noise sensitivity but not vice versa.

We now demonstrate this concept through two examples of monotone noise sensitive functions discussed by Benjamini, Kalai and Schramm in [4], both of which trace back to Ben-Or and Linial in the related work [5].

(i) *Tribes*: partition $\Lambda_n = \{x_1, \dots, x_n\}$ into blocks of $\log_2 n - \log_2 \log_2 n$ variables, let $p_n \equiv 1/2$ and set f_n to be 1 if there is an all-1 block.

It is known [4, §6.1] that this function is non-degenerate and Sens. A 1-witness W in ω is a full block, which the noise will destroy with probability approaching 1, and the probability of encountering another in ω^ε should be asymptotically $\mathbb{P}(f_n = 1)$. Indeed, tribes is $\textsc{StrSens}_1$.

(ii) *Recursive 3-Majority*: Index $n = 3^k$ variables by the leaves of a ternary tree, and iteratively set the value of each node to be the majority of its children. Take $p_n \equiv 1/2$ and define f_n to be the value at the root.

Clearly non-degenerate, this function is known [4, §6.2] to be Sens, i.e., $\mathbb{P}\left(f_n(\omega^\varepsilon) = 1 \mid f_n(\omega) = 1\right) \to 1/2$ as $n \to \infty$. A 1-witness W is a set of 2^k leaves (positioned in the obvious way to force the majority). It is then easy to verify that $\mathbb{P}\left(f_n(\omega^\varepsilon) = 1 \mid \omega_W \equiv 1\right) = 1 - \varepsilon/2$, and therefore this function is *not* $\textsc{StrSens}_1$ (nor $\textsc{StrSens}_0$ by symmetry).

It is important to emphasize the potentially different behaviors of 0-witnesses and 1-witnesses w.r.t. strong noise sensitivity, vs. standard noise sensitivity which is closed under taking complements. Indeed, by a general principle, the tribes function, mentioned above as being $\textsc{StrSens}_1$, is *not* $\textsc{StrSens}_0$ (conditioning on a particular 0-witness in ω does affect $f_n(\omega^\varepsilon)$ in the limit).

The above examples all featured $p_n \equiv 1/2$. Indeed, as noted in [4, §6.4],

> "*When p tends to zero with n, new phenomena occur. Consider, for example, random graphs on n vertices with edge probability $p = n^{-a}$...*"

Many key features of the Erdős-Rényi random graph are non-degenerate at such p and yet the BKS criterion for Sens is then no longer applicable.

1.2. **Properties of random graphs.** The Erdős-Rényi random graph, $\mathcal{G}(n,p)$, is a probability distribution over graphs on n labeled vertices, where each undirected edge appears independently with probability $p = p(n)$. A monotone increasing graph property is a collection of graphs closed under isomorphism and the addition of edges, and we will often identify it with its indicator function (a monotone Boolean function on the $\binom{n}{2}$ edge variables).

As a first example, consider $\mathcal{G}(n,p)$ at its famous critical window centered at $p = 1/n$, where the longest cycle is typically of order $n^{1/3}$ (see, e.g., [9]).

Theorem 1.3. *Fix $0 < a < b$ and let f_n be the property that the critical random graph $\mathcal{G}(n, 1/n)$ contains a cycle of length $\ell \in (an^{1/3}, bn^{1/3})$. Then (f_n) is non-degenerate and noise sensitive, and furthermore, it is* $\textsc{StrSens}_1$.

Moreover, the analogue of this conclusion for quantitative noise sensitivity holds if and only if the noise parameter $\varepsilon = \varepsilon(n)$ satisfies $\varepsilon \gg n^{-1/3}$.

Theorem 1.3 in fact holds throughout the critical window $p = \frac{1\pm\xi}{n}$ with $\xi = O(n^{-1/3})$, around which the longest cycle grows from constant to linear (e.g., taking $\xi^3 n \to \infty$ still with $\xi = o(1)$, the maximum length of a cycle is $\Theta_{\mathrm{P}}(1/\xi)$ at $p = \frac{1-\xi}{n}$ and $\Theta_{\mathrm{P}}(\xi^2 n)$ at $p = \frac{1+\xi}{n}$; see [9, Theorems 5.17, 5.18]).

Revisiting the quantitative conclusion of Theorem 1.3 now highlights an interesting phenomenon, where the $\varepsilon \gg n^{-1/3}$ threshold for noise sensitivity coincides with the boundary of the critical window ($p = \frac{1\pm\xi}{n}$ for $\xi \gg n^{-1/3}$). This phenomenon is best explained through the following equivalent process:

- Let ω be a uniform set of $N \sim \mathrm{Bin}(\binom{n}{2}, p)$ edges.
- Obtain $\bar\omega$ by deleting a uniform set of $\mathrm{Bin}(N, \varepsilon(1-p))$ edges from ω.
- Add a uniform set of $\mathrm{Bin}(\binom{n}{2} - N, \varepsilon p)$ edges missing from ω to get ω^ε.

As the edge probability in $\bar\omega$ is $p(1-\varepsilon) + \varepsilon p^2$, on a heuristic level we have:

(a) If $\varepsilon \lesssim n^{-1/3}$ then $\bar\omega$ remains in the critical window, where (f_n) is non-degenerate, so $f_n(\omega), f_n(\bar\omega)$ (thus $f_n(\omega), f_n(\omega^\varepsilon)$) should be correlated.

(b) If $\varepsilon \gg n^{-1/3}$ then $\bar\omega$ is subcritical whence $f_n(\bar\omega)$ is degenerate, effectively decorrelating $f_n(\bar\omega)$ from $f_n(\omega)$ (thus also $f_n(\omega), f_n(\omega^\varepsilon)$) yielding Sens.

Although plausible, it is unclear that in general the degeneracy of $f_n(\bar\omega)$ will indeed result in the decorrelation of $f_n(\omega)$ and $f_n(\omega^\varepsilon)$.

Intuitively, we expect a random graph property to be noise sensitive when it has no bounded-size witnesses (thus none will survive the noise in fact) and distinct witnesses are essentially independent (so surviving fragments of a witness will have negligible impact), as is the case in the theorem above.

However, for various important graph properties the witnesses happen to be highly correlated, foiling this intuition. For instance, containing a Hamilton cycle is non-degenerate at $p \sim \frac{\log n}{n}$ yet the expected number of witnesses becomes exponentially large in n already at $p = O(1/n)$, and similarly for perfect matchings. Nevertheless, both are in fact noise sensitive:

Theorem 1.4. *Let f_n be the property that the minimum degree of $\mathcal{G}(n,p)$ is at least k for some fixed $k \geq 1$, and suppose $p = p(n)$ is such that (f_n) is non-degenerate. Then (f_n) is noise sensitive, and moreover, it is* StrSens$_0$.

As a result, the following properties of $\mathcal{G}(n,p)$ are noise sensitive:

(i) containing a Hamilton cycle,

(ii) containing a perfect matching (in general, an r-factor[1] for r fixed),

(iii) connectivity (in general, k-vertex and k-edge connectivity for k fixed),

(iv) having an isoperimetric constant[2] of at least γ for some fixed $\gamma > 0$.

Furthermore, each of these is quantitatively noise sensitive iff $\varepsilon \gg \frac{1}{\log n}$.

[1] An r-factor of a graph is a spanning r-regular subgraph

[2] The isoperimetric constant of a graph is the minimum of $\frac{e(S,S^c)}{|S|\wedge|S^c|}$ over all subsets S of the vertices, where $e(S,S^c)$ is the number of edges between S and its complement.

It is worthwhile noting that not even the (non-strong) noise sensitivity in Theorems 1.3 or 1.4 can be obtained from the best known generalizations of the BKS criterion for varying p (see [11]), as these all require $1/p = n^{o(1)}$.

We turn our attention to the well-studied family of properties of the form "$\mathcal{G}(n,p)$ contains a copy of a given graph H_n". Obviously, if the size of H_n is uniformly bounded then this property is *not* noise sensitive, since a copy of H_n will survive the noise with positive probability (as noted in [4, §6.4], it is *noise stable*, a notion basically the opposite of being noise sensitive). Note that having the number of edges in H_n grow with n is a necessary but not sufficient condition for noise sensitivity (e.g., take $\log n$ disjoint edges).

The case where H_n is a clique concerns the maximum clique size in $\mathcal{G}(n,p)$. It is well-known (see, e.g., [1]) that at $p = 1/2$ this concentrates on a single point $k_n \sim 2\log_2 n$ for most values of n, while for exceptional values of n it is either k_n or $k_n + 1$ with high probability. In the latter case, one can ask whether the property that k_n is the maximum clique size is noise sensitive. Indeed it is, as implied by the BKS criterion (see §2.5). However, one would expect there to be a direct proof of this fact that does not employ the machinery of Fourier analysis and hyper-contractive estimates.

Here we provide a direct proof of strong noise sensitivity for this property.

Theorem 1.5. *Let f_n be the property that $\mathcal{G}(n,p)$ has a clique of size k_n for $k_n = n^{o(1)}$ such that $k_n \to \infty$ with n, and suppose $p = p(n)$ is such that (f_n) is non-degenerate. Then (f_n) is noise sensitive. Moreover, it is* $\textsc{StrSens}_1$.

Consider the above theorem for $1 \ll k_n \lesssim \log n$. When H_n is a clique of size k_n, containing H_n in $\mathcal{G}(n,p)$ is $\textsc{Sens}$. However, if H_n consists of k_n disjoint edges for the same sequence k_n then the property is noise stable (essentially as a majority function). In light of these two opposite behaviors, one wishes to understand which features of the given graph H_n dictate $\textsc{Sens}$.

While determining noise sensitivity for graphs H_n whose size grows rapidly with n can be delicate, the picture is fairly well-understood when the graph sizes are at most a certain poly-log of n. In that case, it turns out that a single feature of H_n — being *strictly balanced* — governs noise sensitivity. A graph is *balanced* if its average degree is at least that of any of its proper subgraphs, and it is *strictly balanced* if these inequalities are all strict (e.g., a clique is strictly balanced whereas a collection of disjoint edges is balanced).

Theorem 1.6. *Let H_n be a sequence of graphs and let f_n be the property that the random graph $\mathcal{G}(n,p)$ contains a copy of H_n. The following holds:*

1. *If H_n is strictly balanced with $1 \ll \ell_n \le \big(\frac{\log n}{\log\log n}\big)^{1/2}$ edges then (f_n) is noise sensitive, and furthermore, it is* $\textsc{StrSens}_1$.
2. *There exists a sequence of strictly balanced graphs H_n with $\ell_n \asymp \log n$ edges for which (f_n) is not noise sensitive.*

We stress that the assumption that H_n is strictly balanced is necessary in the sense that, without it, one could take H_n to be ℓ_n disjoint copies of any fixed strictly balanced graph (e.g., a clique or a tree) for any $\ell_n \ll \sqrt{n}$, whence containing H_n is not SENS (in fact it is noise stable). However, it is not that having H_n be strictly balanced is a necessary condition for SENS, e.g., we will see that containing a disjoint union of two cliques is STRSENS_1.

The last two theorems will be obtained as a consequence of a general tool (Proposition 4.1) which deduces STRSENS_1 from an appropriate Poisson approximation of the number of copies of H_n in G.

We note that each of the properties shown in Theorems 1.3–1.6 to be STRSENS_1 is *not* STRSENS_0, and the properties that were shown to be STRSENS_0 are *not* STRSENS_1. Indeed, a general principle (Lemma 5.1) will yield that, if we let X_n denote the number of 1-witnesses W for which $\omega_W \equiv 1$, then having $\mathbb{E}[X_n] = O(1)$ precludes STRSENS_0 (and similarly for 0-witnesses). At the same time, there can be monotone Boolean functions that are both STRSENS_0 and STRSENS_1, as we demonstrate in §5.

1.3. **Organization.** The rest of the paper is outlined as follows. In §2 we provide prerequisites on noise sensitivity. Section 3 demonstrates the use of strong noise sensitivity towards establishing noise sensitivity, including the proof of Theorems 1.3 and 1.4. Section 4 looks into the dependencies between witnesses for a sufficient condition for strong noise sensitivity. This condition is then applied in the context of containing a given graph in $\mathcal{G}(n, p)$ and in particular towards the proofs of Theorems 1.5 and 1.6. Finally, §5 compares the 0-strong and 1-strong noise sensitivity of a function, as well as the validity of these properties under varying levels of noise.

2. Preliminaries

This section includes background on noise sensitivity, both for constant p and when the probabilities p are allowed to vary with n (see, e.g., [11] for additional information on this topic). We first set some standard notation.

2.1. **Notation.** Throughout the paper, a sequence of events A_n is said to hold with high probability (w.h.p.) if $\mathbb{P}(A_n) \to 1$ as $n \to \infty$. We use the notation $f = O_{\mathrm{P}}(g)$ to denote that the ratio f/g is bounded in probability, and the analogous $f = \Theta_{\mathrm{P}}(g)$ to denote that $f = O_{\mathrm{P}}(g)$ and $g = O_{\mathrm{P}}(f)$. At times we use $f \ll g$ and $f \lesssim g$ to abbreviate $f = o(g)$ and $f = O(g)$, resp., as well as the converse form of these. We will often omit the subscript n from the probabilities p_n under consideration in this paper (though these will typically tend to 0 as $n \to \infty$) for simplicity.

2.2. **Influences and the pivotal set.** The notion of influence, defined next, is fundamental in the study of noise sensitivity of functions.

Definition 2.1. Given a Boolean function f from $\Omega = \{0,1\}^\Lambda$ into $\{0,1\}$, $p \in (0,1)$ and $i \in \Lambda$, the *influence of* i *with respect to* p is defined to be

$$\mathbf{I}_i(f) = \mathbb{P}(f(\omega) \neq f(\omega^i)) \tag{2.1}$$

where ω^i is ω flipped in the i-th coordinate.

(As usual, the above definition implicitly depends on p through $\mathbb{P}$.) The following theorem of [4] is one of the central results on noise sensitivity.

Theorem 2.2 ([4])**.** *Let $p_n \equiv p$ for some fixed $0 < p < 1$. If*

$$\lim_{n\to\infty} \sum_i \mathbf{I}_i(f_n)^2 = 0 \tag{2.2}$$

for a sequence of Boolean functions (f_n), then (f_n) is Sens*.*

As we will see below, for monotone functions and constant p the converse is also true, while what occurs when $p_n \to 0$ is more subtle.

Consider the random set of *pivotal* variables defined as

$$\mathcal{P}(\omega) := \mathcal{P}_f(\omega) := \{i \in \Lambda : f(\omega) \neq f(\omega^i)\}$$

(Notice $\mathbb{P}(i \in \mathcal{P}) = \mathbf{I}_i$.) The following easy lemma will be used in this paper.

Lemma 2.3. *Every monotone Boolean function f satisfies*

$$\mathbb{E}[|\mathcal{P}| \mid f = 1] = \frac{p}{\mathbb{P}(f=1)} \mathbb{E}|\mathcal{P}|$$

Proof. Note that $\{f(\omega) \neq f(\omega^i)\}$ and $\{\omega_i = 1\}$ are independent and so the left-hand-side of the desired equality is easily seen to be equal to

$$\sum_i \mathbb{P}(f(\omega) \neq f(\omega^i) \mid f = 1) = \sum_i \frac{\mathbb{P}(f(\omega) \neq f(\omega^i),\, \omega_i = 1)}{\mathbb{P}(f=1)} = \frac{p}{\mathbb{P}(f=1)} \mathbb{E}|\mathcal{P}|$$

where the first equality uses monotonicity and the second equality uses the earlier stated independence. ∎

Remark. The above also holds for non-monotone functions when $p = 1/2$.

We now indicate that the equivalence holding for monotone functions and constant p between $\sum_i \mathbf{I}_i(f_n)^2 = o(1)$ and Sens in fact fails for varying p in either direction. Let f_n be the indicator function of a random graph containing a copy of K_4 with $p = n^{-2/3}$. Clearly $\mathbb{E}[|\mathcal{P}| \mid f = 1] \leq 6$ which by Lemma 2.3 implies that $\mathbb{E}|\mathcal{P}| = O(n^{2/3})$. By symmetry, this yields $I_i = O(n^{-4/3})$ for each i, which easily yields (2.2), and yet this sequence is clearly stable. On the other hand, if f_n is the indicator function of a random graph with $p = \frac{\log n}{n}$ having minimal degree 1, then $\{f_n\}$ is Sens (see Theorem 1.4). However, it is easy to verify that $\mathbb{E}[|\mathcal{P}| \mid f = 0] \gtrsim n$, which by Lemma 2.3 yields $\mathbb{E}|\mathcal{P}| \gtrsim n$ and so $\sum_i \mathbf{I}_i(f_n)^2 \gtrsim 1$.

We will see in the next subsection that asking about a possible equivalence of $\sum_i \mathbf{I}_i(f_n)^2 = o(1)$ and Sens is in fact not really the right question: instead one should ask about a possible equivalence of $p \sum_i \mathbf{I}_i(f_n)^2 = o(1)$ and Sens.

2.3. **Fourier analysis.** Fourier analysis is usually a crucial tool in studying noise sensitivity. We give a quick presentation of this. From it, one readily sees some of the basic properties of noise sensitivity.

For a set Λ, $\omega \in \{0,1\}^\Lambda$ and $i \in \Lambda$, we define

$$\chi_i(\omega) = \begin{cases} \sqrt{(1-p)/p} & \text{if } \omega_i = 1\,, \\ -\sqrt{p/(1-p)} & \text{if } \omega_i = 0\,. \end{cases}$$

Furthermore, for $S \subseteq \Lambda$, let $\chi_S(\omega) := \prod_{i \in S} \chi_i(\omega)$. (In particular, $\chi_\emptyset$ is the constant function 1.) The set $\{\chi_S\}_{S \subseteq \Lambda}$ forms an orthonormal basis for the set of functions $f : \{0,1\}^\Lambda \mapsto \mathbb{R}$ when the latter is equipped with the inner product $\langle f, g \rangle := \mathbb{E}[fg]$ (recall there is always an implicit p when we write $\mathbb{P}$ or $\mathbb{E}$). We can therefore expand such functions $f(\omega) = \sum_{S \subseteq \Lambda} \hat{f}(S)\chi_S(\omega)$, where $\hat{f}(S) := \mathbb{E}[f\chi_S]$ is the Fourier-Walsh coefficient of f. Note that $\hat{f}(\emptyset)$ is the average $\mathbb{E}f$ and by Parseval's formula $\mathbb{E}[f^2] = \sum_{S \subseteq \Lambda} \hat{f}(S)^2$. This orthogonal basis turns out to be an extremely useful one for studying noise sensitivity, as the following easily verified formula demonstrates:

$$\mathbb{E}[f(\omega)f(\omega^\varepsilon)] = \sum_S \hat{f}(S)^2(1-\varepsilon)^{|S|}\,. \tag{2.3}$$

This yields

$$\operatorname{Cov}\left(f_n(\omega), f_n(\omega^\varepsilon)\right) = \sum_{S \neq \emptyset} \hat{f}(S)^2(1-\varepsilon)^{|S|}\,.$$

The following theorem now follows immediately; note importantly how it shows that if the appropriate covariance goes to 0 for one value of ε, then it does so for all ε. Note that there is no condition on the sequence (p_n).

Theorem 2.4. *Let (f_n) be a sequence of Boolean functions. Then (f_n) is* Sens *if and only if any one of the following conditions holds:*
(1) For some $0 < \varepsilon < 1$ we have $\lim_{n\to\infty} \sum_{S \neq \emptyset} \hat{f}_n(S)^2(1-\varepsilon)^{|S|} = 0$.
(2) For every $0 < \varepsilon < 1$ we have $\lim_{n\to\infty} \sum_{S \neq \emptyset} \hat{f}_n(S)^2(1-\varepsilon)^{|S|} = 0$.
(3) For every k we have $\lim_{n\to\infty} \sum_{0<|S|<k} \hat{f}_n(S)^2 = 0$.

A very useful mnemonic device is the so-called spectral sample $\mathcal{S} = \mathcal{S}_f$ of a Boolean function f, defined distributionally by

$$\mathbb{P}(\mathcal{S} = S) := \hat{f}(S)^2 \qquad (S \subset \Lambda)\,.$$

The total weight of this distribution is less than 1 (unless $f \equiv 1$). Note that the terms in Items (1) and (3) in Theorem 2.4 respectively become

$$\mathbb{E}\left[(1-\varepsilon)^{|\mathcal{S}_n|}\mathbb{1}_{\{\mathcal{S}\neq\emptyset\}}\right] \text{ and } \mathbb{P}(0<|\mathcal{S}_n|<k)\,.$$

It turns out that SENS is equivalent to another condition — appearing perhaps stronger at first glance — according to which for most ω with $f_n(\omega)=1$, the conditional probability that $f_n(\omega^\varepsilon)=1$ given ω is close to the unconditional probability.

Proposition 2.5. *Let (f_n) be a sequence of Boolean functions. Then (f_n) is* SENS *if and only if any one of the following conditions holds:*
(1) $[\mathbb{P}(f_n(\omega^\varepsilon)=1\mid\omega)-\mathbb{P}(f_n(\omega)=1)]\xrightarrow{\mathrm{P}}0$.
(2) $[\mathbb{P}(f_n(\omega^\varepsilon)=1\mid\omega)-\mathbb{P}(f_n(\omega)=1)]\,\mathbb{1}_{\{f_n(\omega)=1\}}\xrightarrow{\mathrm{P}}0$.

Proof. It is immediate that (1) implies (2). To see that (2) implies SENS as per (1.1), simply write the expression appearing in (1.1) as

$$\sum_{\omega:f_n(\omega)=1}\left[\mathbb{P}\left(f_n(\omega^\varepsilon)=1\mid\omega\right)-\mathbb{P}\left(f_n=1\right)\right]\frac{\mathbb{P}(\omega)}{\mathbb{P}\left(f_n=1\right)}\,.$$

It remains to show that SENS implies (1). It is easy to verify that

$$\mathrm{Var}\left(\mathbb{P}\left(f_n(\omega^\varepsilon)=1\mid\omega\right)\right)=\sum_{S\neq\emptyset}\hat{f}_n(S)^2(1-\varepsilon)^{2|S|}\,.$$

Therefore, by Theorem 2.4, if (f_n) is SENS we can infer that

$$\lim_{n\to\infty}\mathrm{Var}(\mathbb{P}\left(f_n(\omega^\varepsilon)=1\mid\omega\right))=0\,.$$

Since $\mathbb{E}[\mathbb{P}\left(f_n(\omega^\varepsilon)=1\mid\omega\right)]=\mathbb{P}(f(\omega)=1)$, this immediately gives (1). ∎

While Theorem 2.4 is quite easy, Theorem 2.2 is much deeper. It turns out that the converse of Theorem 2.2 with constant p is true for monotone functions as we now explain. First, for a monotone Boolean function f mapping into $\{0,1\}$, one can easily check that

$$\hat{f}(\{i\})=\sqrt{p(1-p)}\mathbf{I}_i(f)\,. \tag{2.4}$$

This formula together with Theorem 2.4 immediately yields the converse of Theorem 2.2 for fixed p. This reinterprets Theorem 2.2 in the monotone case as saying that, for constant p, if the "sum of the squares of the level 1 Fourier coefficients" $\sum_{|S|=1}\hat{f}_n(S)^2$ approaches 0, then the sequence in SENS.

We now consider Theorem 2.2 in the context of varying p, in particular for p tending to 0 with n. As above, for monotone functions, (2.4) and Theorem 2.4 yield the fact that, for arbitrary (p_n), SENS implies

$$\lim_{n\to\infty}p(1-p)\sum_i\mathbf{I}_i(f_n)^2=0\,. \tag{2.5}$$

From this discussion, it follows that the version of Theorem 2.2 that one might hope for, for arbitrary (p_n), is that (2.5) implies SENS; equivalently, for monotone functions, convergence of the level 1 Fourier coefficients implies SENS. Unfortunately, this is not true as we saw in the previous subsection for the event "containing a K_4". Alternatively, if we let $p_n = 1/n$ and consider the indicator function of containing a triangle, then it is easy to see that this sequence is not SENS (and in fact noise stable, see this definition below) although (2.5) is of order $1/n$. The stability of the indicator function f_n for containing a triangle implies that $\lim_{k\to\infty} \sup_n \sum_{|S|\geq k} \hat{f}_n(S)^2 = 0$. In addition, in [7] it is shown that for any $k \not\equiv 0 \pmod 3$ this f_n satisfies

$$\lim_{n\to\infty} \sum_{|S|=k} \hat{f}_n(S)^2 = 0\,,$$

i.e., the Fourier weights are concentrated on levels $0, 3, 6, \ldots$ but stay near 0. (Such a thing cannot occur for monotone functions with constant p.)

We end this subsection by defining the closely related (but opposite) concept to SENS, namely *noise stability*.

Definition 2.6. The sequence of functions $f_n : \{0,1\}^{\Lambda_n} \to \{0,1\}$ is noise stable (STAB) if for any $\delta > 0$ there exists an $\varepsilon > 0$ such that

$$\sup_n \mathbb{P}(f_n(\omega) \neq f_n(\omega^\varepsilon)) \leq \delta\,.$$

If $\varepsilon_n \to 0$ with n, one can talk about STAB with respect to $\{\varepsilon_n\}$ in the obvious way. Note that while STRSENS$_1$ and SENS with respect to a sequence $\{\varepsilon_n\}$ going to 0 is stronger than ordinary STRSENS$_1$ and SENS, STAB with respect to such a sequence is weaker than ordinary STAB.

2.4. **Relation to coarse and sharp thresholds.** It is natural to wonder where the important results in [7] concerning sharp thresholds fall into the context of this paper. In short, they occur in a very different regime. To explain this, consider for the moment $p = 1/2$. There are three common scenarios that can occur (as well as various combinations).

(1) $\mathbb{E}|\mathcal{S}_n| = O(1)$.
(2) $\mathbb{E}|\mathcal{S}_n| \to \infty$ and yet $|\mathcal{S}_n|$ is bounded in probability.
(3) For every fixed k we have $\mathbb{P}(0 < |\mathcal{S}_n| < k) \to 0$, i.e., (f_n) is SENS.

The first scenario occurs for example if f_n only depends on a fixed finite number of variables independent of n. An example where the second scenario occurs is the sequence of majority functions. Similar to (2.5), there is another relationship between influences and the Fourier picture which does not require monotonicity. This states that $\sum_S \hat{f}(S)^2|S| = p(1-p)\sum_i \mathbf{I}_i(f)$, or equivalently,

$$\mathbb{E}|\mathcal{S}| = p(1-p)\mathbb{E}|\mathcal{P}| \tag{2.6}$$

(as was established for $p = 1/2$ in [10]; the case of general p follows similarly).

In [7], results of the form that, if you are in the first scenario, then for graph properties, the function can be well approximated by functions which depend on a fixed number of graphs. Since the context of [7] was $p = o(1)$, in view of (2.6), the assumptions in [7] are of the form $p \sum_i \mathbf{I}_i(f) \le C$.

2.5. **Maximum cliques in random graphs.** As mentioned above, the maximum clique of $\mathcal{G}(n,p)$ for $p = 1/2$ concentrates on 1 point for most values of n, yet for infinitely many values of n it is concentrated on 2 points. It is for the latter values of n that we have a non-degenerate indicator function corresponding to the event that we contain a clique of size about $k_n \sim 2\log_2 n$. We describe here how Theorem 2.2 yields Sens, as was indicated by Jeff Kahn. Consider the expected size of $\mathcal{P}_n$ (the set of pivotal edges). Since $p = 1/2$, Lemma 2.3 gives

$$\mathbb{E}|\mathcal{P}_n| = 2\mathbb{P}(f_n = 1)\mathbb{E}[|\mathcal{P}_n| \mid f_n = 1] \, .$$

Hence, for the non-degenerate n we focus on, $\mathbb{E}|\mathcal{P}_n|$ and $\mathbb{E}[|\mathcal{P}_n| \mid f_n = 1]$ are of the same order. Clearly whenever $f_n = 1$ necessarily $|\mathcal{P}_n| = O(\log^2 n)$ since if there is at least one clique, one can choose such a clique arbitrarily and then observe that any pivotal edge must belong to it. This shows that $\mathbb{E}|\mathcal{P}_n| = O(\log^2 n)$ and hence the influence of each edge is of order at most $(\frac{\log n}{n})^2$. Squaring this and multiplying by the number of edges, one obtains that $\sum_i \mathbf{I}_i(f_n)^2 \lesssim (\log n)^4/n^2$. Since this approaches 0 with n, Theorem 2.2 yields noise sensitivity.

3. From witnesses to noise sensitivity

In this section we relate noise sensitivity to strong noise sensitivity. Via this connection we prove quantitative versions of Theorems 1.3 and 1.4.

3.1. **Strong noise sensitivity.** We begin with a straightforward lemma showing that strong noise sensitivity indeed implies the standard one.

Lemma 3.1. *Let (f_n) be a non-degenerate sequence of monotone Boolean functions. If (f_n) is* StrSens$_1$ *then it is noise sensitive. Furthermore,* StrSens$_1$ *w.r.t. $\varepsilon = \varepsilon(n) \to 0$ implies quantitative* Sens *w.r.t. the same ε.*

Proof. By the definition of noise sensitivity in (1.1), we aim to show that

$$\mathbb{P}\left(f_n(\omega^\varepsilon) = 1 \mid f_n(\omega) = 1\right) - \mathbb{P}(f_n = 1) \to 0$$

as $n \to \infty$, where $\varepsilon = \varepsilon(n)$ is allowed to tend to 0 with n. By the FKG inequality we have $\mathbb{P}\left(f_n(\omega^\varepsilon) = 1 \mid f_n(\omega) = 1\right) \ge \mathbb{P}(f_n = 1)$ and it remains to provide the corresponding upper bound. Let $\mathcal{W}_1 = \{W_1, \ldots, W_{m_n}\}$ be the 1-witnesses for f_n (arbitrarily ordered), and define the variable J to be

$$J = \min\{1 \le j \le m_n \;:\; \omega_{W_j} \equiv 1\}$$

or ∞ in case $f_n(\omega) = 0$. With this notation,

$$\mathbb{P}(f_n(\omega^\varepsilon) = 1 \mid f_n(\omega) = 1) = \sum_{j=1}^{m_n} \mathbb{P}(f_n(\omega^\varepsilon) = 1 \mid J = j)\mathbb{P}(J = j \mid f_n(\omega) = 1) \tag{3.1}$$

and again by FKG we see that

$$\mathbb{P}\left(f_n(\omega^\varepsilon) = 1 \mid J = j\right) \le \mathbb{P}\left(f_n(\omega^\varepsilon) = 1 \mid \omega_{W_j} \equiv 1\right)$$

since we can condition on $\{J = j\}$ by first conditioning on $\{\omega_{W_j} \equiv 1\}$ (obtaining a positively associated measure which enjoys the FKG inequality) and then further conditioning on the decreasing event $\bigcap_{j'<j}\{\omega_{W_{j'}} \not\equiv 1\}$. The latter can only decrease the probability of the increasing event $\{f_n(\omega^\varepsilon) = 1\}$, thus the last display is established, and altogether we obtain that

$$\mathbb{P}\left(f_n(\omega^\varepsilon) = 1 \mid f_n(\omega) = 1\right) \le \max_{W \in \mathcal{W}_1} \mathbb{P}\left(f_n(\omega^\varepsilon) = 1 \mid \omega_W \equiv 1\right). \tag{3.2}$$

Subtracting $\mathbb{P}(f_n = 1)$ and taking $n \to \infty$ now completes the proof by the definition of $\textsc{StrSens}_1$ in (1.2). ■

Remark 3.2. The proof that strong noise sensitivity implies the standard one in fact required a slightly weaker condition than the one stated in (1.2). Instead of having $\max_W[\mathbb{P}\left(f_n(\omega^\varepsilon) = 1 \mid \omega_W \equiv 1\right) - \mathbb{P}(f_n = 1)] \to 0$, we only need an expectation over this quantity w.r.t. a certain distribution over the witnesses (the first W to appear according to some ordering) to vanish.

In particular, Lemma 3.1 remains valid under the analogue of (1.2) for all witnesses W except some subset $\mathcal{W}_1^* \subset \mathcal{W}_1$ with $\mathbb{P}(\cup_{W \in \mathcal{W}_1^*}\{\omega_W \equiv 1\}) \to 0$.

Example (Tribes)**.** Recalling the definition of the tribes function from the introduction, a 1-witness $W \in \mathcal{W}_1$ is a full block. Writing

$$\begin{aligned}\mathbb{P}\left(f_n(\omega^\varepsilon) = 1 \mid \omega_W \equiv 1\right) \le \mathbb{P}\bigg(&\bigcup_{W' \neq W} \{\omega^\varepsilon_{W'} \equiv 1\} \;\bigg|\; \omega_W \equiv 1\bigg) \\ &+ \mathbb{P}\left(\omega^\varepsilon_W \equiv 1 \mid \omega_W \equiv 1\right),\end{aligned}$$

the last term is equal to $(1 - \varepsilon/2)^{|W|} \to 0$ as we have $|W| \sim \log_2 n \to \infty$ with n, while the first term on the right hand side is equal to

$$\mathbb{P}\bigg(\bigcup_{W' \neq W} \{\omega_{W'} \equiv 1\}\bigg) \le \mathbb{P}(f_n = 1)$$

since any two distinct witnesses W, W' are disjoint and thus $\{\omega_W \equiv 1\}$ and $\{\omega_{W'} \equiv 1\}$ are independent. This establishes that

$$\limsup_{n\to\infty} \max_W [\mathbb{P}\left(f_n(\omega^\varepsilon) = 1 \mid \omega_W \equiv 1\right) - \mathbb{P}(f_n = 1)] \le 0\,,$$

and since it is always nonnegative (by a monotonicity argument) we conclude that the tribes function is $\textsc{StrSens}_1$.

Example (Recursive majority)**.** Consider first the canonical 1-witness W for the recursive 3-majority of $n = 3^k$ variables (i.e., W repeatedly reveals the first 2 of the 3 children of a vertex). Recalling that $p = 1/2$, the quantity

$$\zeta_k^\varepsilon = \mathbb{P}\left(f_n(\omega^\varepsilon) = 1 \mid \omega_W \equiv 1\right) \tag{3.3}$$

is easily seen (by the nature of this recursive definition) to satisfy

$$\zeta_k^\varepsilon = (\zeta_{k-1}^\varepsilon)^2 + 2\zeta_{k-1}^\varepsilon(1 - \zeta_{k-1}^\varepsilon)p = \zeta_{k-1}^\varepsilon\,,$$

thus $\zeta_k^\varepsilon = \zeta_0^\varepsilon = 1 - \varepsilon/2$ for any k. In particular, recursive 3-majority is not $\textsc{StrSens}_1$ despite the fact that it is noise sensitive (indeed, it is easy to see that the influence of a variable is 2^{-k} and so the sum of squared influences is $(3/4)^k$ which vanishes as $k \to \infty$, satisfying the BKS criterion for $\textsc{Sens}$).

We emphasize that for this function not only is $\mathbb{P}(f_n(\omega^\varepsilon) = 1 \mid \omega_W \equiv 1)$ bounded away from $\mathbb{P}(f_n = 1) = 1/2$ (enough in itself to preclude $\textsc{StrSens}_1$) but rather it is $1 - \delta(\varepsilon)$ where $\delta(\varepsilon) \to 0$ with ε. This resembles the notion of noise stability (where $\mathbb{P}(f_n(\omega^\varepsilon) = 1 \mid f_n(\omega) = 1)$ approaches 1 as $\varepsilon \to 0$).

Interestingly, further increasing the size of the majority yields an even stronger witness dependency. As before $\mathbb{P}(f_n(\omega^\varepsilon) = 1 \mid \omega_W \equiv 1) \ge 1 - \delta(\varepsilon)$, but instead of $\delta(\varepsilon) = \varepsilon/2$ (the case for 3-majority) we now have $\delta(\varepsilon) = o(1)$.

Claim 3.3. *Let f_n be the recursive 5-majority function on $n = 5^k$ vertices. Then for every $0 < \varepsilon < 1$,*

$$\lim_{n\to\infty} \inf_{W\in\mathcal{W}_1} \mathbb{P}(f_n(\omega^\varepsilon) = 1 \mid \omega_W \equiv 1) = 1\,.$$

Proof. As before, consider the canonical 1-witness W which repeatedly specifies 3 of 5 children of a vertex, and define ζ_k^ε as in (3.3). In this way, conditioned on W the root has 3 children each of which is a Bernoulli(ζ_{k-1}) and 2 other children which are Bernoulli(1/2). It is then easy to check that

$$\zeta_k^\varepsilon = -\frac{1}{2}(\zeta_{k-1}^\varepsilon)^3 + \frac{3}{4}(\zeta_{k-1}^\varepsilon)^2 + \frac{3}{4}\zeta_{k-1}^\varepsilon\,,$$

and as before $\zeta_0^\varepsilon = 1 - \frac{\varepsilon}{2}$. Letting

$$h(x) = -\frac{1}{2}x^3 + \frac{3}{4}x^2 + \frac{3}{4}x \tag{3.4}$$

we thus have $\zeta_k^\varepsilon = h(\zeta_{k-1}^\varepsilon)$, and the proof follows from the easily verifiable facts that h maps $[0, 1]$ to itself with fixed points at $\{0, 1/2, 1\}$, out of which

$1/2$ is a repelling fixed point since $h'(1/2) = 9/8 > 1$. Hence, $\zeta_k^\varepsilon \to 1$ as long as $\zeta_0^\varepsilon > 1/2$, which is indeed the case by the hypothesis $0 < \varepsilon < 1$. ∎

We note in passing that the analogue of Claim 3.3 for noise sensitivity (rather than strong noise sensitivity) is not possible for any non-degenerate sequence (f_n), since $\mathbb{P}\left(f_n(\omega^\varepsilon) = 1 \mid f_n(\omega) = 1\right) \leq 1 - g(\varepsilon)$ for $g(\varepsilon) \gtrsim \varepsilon$.

3.2. **Quantitative noise sensitivity for cycles at criticality.** In this section we prove the following stronger form of Theorem 1.3, offering a more detailed examination of the phase transition for noise sensitivity around the point where the noise parameter ε is of order $n^{-1/3}$.

Theorem 3.4. *Fix $0 < a < b$ and let f_n be the property that $\mathcal{G}(n,p)$ with $p = (1 + O(n^{-1/3}))/n$ contains a cycle of length $\ell \in (an^{1/3}, bn^{1/3})$. Then (f_n) is non-degenerate and according to the noise parameter $\varepsilon(n)$ we have:*

(i) If $\varepsilon \gg n^{-1/3}$ then (f_n) is Sens *and furthermore* StrSens$_1$ *w.r.t. ε.*
(ii) If $\varepsilon \ll n^{-1/3}$ then (f_n) is Stab *w.r.t. ε.*
(iii) If $\varepsilon \asymp n^{-1/3}$ then (f_n) is neither Sens *w.r.t. ε nor* Stab *w.r.t. ε.*

Proof. Let $G \sim \mathcal{G}(n,p)$ and let ω denote its edge configuration (i.e., ω_{uv} is set to 1 if the edge uv is present in G and it is 0 otherwise). Let $\lambda_1, \lambda_2 > 0$ be such that $1 - \lambda_1 n^{-1/3} \leq np \leq 1 + \lambda_2 n^{-1/3}$ for all n and let X_ℓ count the number of cycles of length ℓ in G. Put $\mathcal{I} = (an^{1/3}, bn^{1/3})$ and define

$$X = \sum_{\ell \in \mathcal{I}} X_\ell = \#\{W \in \mathcal{W}_1 : \omega_W \equiv 1\}\,.$$

As the number of potential cycles notwithstanding automorphisms in G (that is, the cardinality of $\mathcal{W}_1$) is $\frac{1}{2}\binom{n}{\ell}(\ell-1)!$ we see that $\mathbb{E}X_\ell \sim (np)^\ell/(2\ell)$ uniformly over $\ell \in \mathcal{I}$ and so

$$(1-o(1))e^{-\lambda_1 b} \leq \frac{\mathbb{E}X}{\frac{1}{2}\log(b/a)} \leq (1+o(1))e^{\lambda_2 b}\,. \tag{3.5}$$

At this point, the FKG inequality immediately implies that

$$\mathbb{P}(X=0) \geq \prod_{\ell \in \mathcal{I}} \left(1 - p^\ell\right)^{\frac{1}{2}\binom{n}{\ell}(\ell-1)!} \geq e^{-(1+o(1))\mathbb{E}X}\,, \tag{3.6}$$

(where the second inequality used the fact that $1 - x = e^{-(1+o(1))x}$ as $x \to 0$) which is bounded away from 0 thanks to (3.5).

Next, we examine $\mathrm{Var}(X)$. For any two cycles $W \neq W'$, let $\kappa(W,W')$ count the number of nontrivial connected components in the intersection of the edges of W and W' (each of which is a simple path), and define

$$\zeta_m := \sum_{\substack{W,W' \in \mathcal{W}_1 \\ \kappa(W,W')=m}} \mathbb{P}\left(\omega_W \equiv 1\,,\, \omega_{W'} \equiv 1\right)$$

for each $m \geq 1$. With this notation,

$$\mathrm{Var}(X) \leq \mathbb{E}X + \sum_{m\geq 1} \zeta_m \,,$$

prompting the task of estimating the ζ_m's. In what follows, let ℓ, ℓ' run over the potential lengths of W, W', resp., while s will run over the total number of edges in the intersection of W and W'. We then have

$$\zeta_m \leq \sum_{\ell\in\mathcal{I}}\sum_{\ell'\in\mathcal{I}}\sum_{m\leq s<\ell} \binom{s}{m-1} (2\ell\ell')^m \frac{n^\ell p^\ell}{2\ell} \frac{n^{\ell'-(s+m)} p^{\ell'-s}}{2\ell'} \,,$$

where the first term accounts for the partitioning of the s total edges into the m intersection paths (with room to spare), the second one accounts for selecting the paths within W (starting point and direction per path) as well as their position within W', and the final two terms correspond to selecting W and W' with this intersection pattern. The fact that $np \leq 1 + \lambda_2 n^{-1/3}$ translates into having $(np)^{\ell+\ell'-s} < C$ for $C = e^{2b\lambda_2}$, thus

$$\zeta_m \leq \frac{C}{2n} \sum_{\ell}\sum_{\ell'}\sum_{s} \frac{(2\ell\ell' s/n)^{m-1}}{(m-1)!} \leq \frac{C}{2}(b-a)^2 b \frac{(2b^3)^{m-1}}{(m-1)!} \,.$$

and

$$\sum_{m\geq 1} \zeta_m \leq \frac{C}{2}(b-a)^2 b e^{2b^3} = O(1) \,.$$

In particular we get that $\mathbb{E}[X^2] = O(1)$.

An immediate consequence of Cauchy-Schwarz is that any non-negative random variable X satisfies $\mathbb{P}(X > 0) \geq (\mathbb{E}X)^2/\mathbb{E}[X^2]$, thus in particular $\mathbb{P}(X > 0)$ is bounded away from 0. Combining this with (3.6), it now follows that (f_n) is non-degenerate.

Remark. Using similar moment analysis, one can infer that the limiting distribution of X is not Poisson; for instance, already ζ_1 is uniformly bounded away from 0 (as it is apparent that $\zeta_1 \geq (\frac{1}{2} - o(1))(b-a)^2 a$ from the argument above) and consequently $\mathrm{Var}(X)$ is bounded away from $\mathbb{E}X$ as $n \to \infty$.

- *Noise sensitivity iff* $\varepsilon \gg n^{-1/3}$:

The strong noise sensitivity of (f_n) when $\varepsilon \gg n^{-1/3}$ will be derived from a calculation akin to the second moment analysis given above, yet this time it will incorporate the noise in the following prominent way. For any $W \in \mathcal{W}_1$ of some length ℓ, define

$$\zeta'_m := \sum_{\substack{W'\in\mathcal{W}_1 \\ \kappa(W,W')=m}} \mathbb{P}\left(\omega^\varepsilon_{W'} \equiv 1 \mid \omega_W \equiv 1\right) \,.$$

By the same line of arguments presented above for ζ_m we have

$$\begin{aligned}\zeta'_m &\leq \sum_{\ell'}\sum_{s}\binom{s}{m-1}(2\ell\ell')^m\frac{n^{\ell'-(s+m)}p^{\ell'-s}}{2\ell'}(1-\varepsilon(1-p))^s\\ &\leq \frac{C\ell}{n}\sum_{\ell'}\sum_{s}\frac{(2\ell\ell's/n)^{m-1}}{(m-1)!}(1-\varepsilon(1-p))^s\,,\end{aligned}$$

again using the fact that $(np)^{\ell'-s} < C$ for $C = e^{\lambda_2 b}$. Thanks to the crucial last term, accounting for the probability of retaining the s edges in the intersection paths, it follows that

$$\begin{aligned}\zeta'_m &\leq \frac{Cb(b-a)}{n^{1/3}}\frac{(2b^3)^{m-1}}{(m-1)!}\sum_{s}(1-\varepsilon(1-p))^s\\ &\leq \frac{Cb(b-a)}{n^{1/3}\varepsilon(1-p)}\frac{(2b^3)^{m-1}}{(m-1)!}\,,\end{aligned}$$

and so

$$\sum_{m\geq 1}\zeta'_m \leq \frac{Cb(b-a)e^{2b^3}}{n^{1/3}\varepsilon(1-p)} = O\left(\frac{1}{\varepsilon n^{1/3}}\right)\,. \tag{3.7}$$

In particular, when $\varepsilon \gg n^{-1/3}$ (Part (i)) we can infer that $\sum_{m\geq 1}\zeta'_m = o(1)$. To deduce that (f_n) is $\textsc{StrSens}_1$ in this case, argue as follows. Fix in what follows some $W \in \mathcal{W}_1$. Partitioning $\mathcal{W}_1 = \{W\}\cup\mathcal{W}'_1\cup\mathcal{W}''_1$ where $\mathcal{W}'_1 := \{W' \neq W : \kappa(W,W') > 0\}$ (and $\mathcal{W}''_1$ contains cycles that are edge-disjoint from W, thus independent) gives

$$\begin{aligned}\mathbb{P}(f_n(\omega^\varepsilon)=1 \mid \omega_W\equiv 1) &\leq \mathbb{P}\Big(\bigcup_{W'\in\mathcal{W}'_1}\{\omega^\varepsilon_{W'}\equiv 1\}\;\Big|\;\omega_W\equiv 1\Big)\\ &\quad+\mathbb{P}\Big(\bigcup_{W''\in\mathcal{W}''_1}\{\omega^\varepsilon_{W''}\equiv 1\}\;\Big|\;\omega_W\equiv 1\Big)+\mathbb{P}\left(\omega^\varepsilon_W\equiv 1\mid\omega_W\equiv 1\right)\,.\end{aligned}$$

By the definition of ζ'_m and Eq. (3.7) in the case of $\varepsilon \gg n^{-1/3}$,

$$\mathbb{P}\Big(\bigcup_{W'\in\mathcal{W}'_1}\{\omega^\varepsilon_{W'}\equiv 1\}\;\Big|\;\omega_W\equiv 1\Big) \leq \sum_{m\geq 1}\zeta'_m = o(1)\,,$$

while clearly

$$\mathbb{P}\Big(\bigcup_{W''\in\mathcal{W}''_1}\{\omega^\varepsilon_{W''}\equiv 1\}\;\Big|\;\omega_W\equiv 1\Big) = \mathbb{P}\Big(\bigcup_{W''\in\mathcal{W}''_1}\{\omega_{W''}\equiv 1\}\Big) \leq \mathbb{P}(f_n=1)$$

and

$$\mathbb{P}\left(\omega^\varepsilon_W\equiv 1\mid\omega_W\equiv 1\right) = (1-\varepsilon(1-p))^\ell \leq e^{-\varepsilon(1-p)an^{1/3}} = o(1)$$

again thanks to the assumption that $\varepsilon \gg n^{-1/3}$. Altogether, this yields

$$\mathbb{P}(f_n(\omega^\varepsilon) = 1 \mid \omega_W \equiv 1) \leq \mathbb{P}(f_n = 1) + o(1)\,,$$

thus establishing that (f_n) is $\textsc{StrSens}_1$ when $\varepsilon \gg n^{-1/3}$.

We will now show that (f_n) is not $\textsc{Sens}$ w.r.t. ε whenever $\varepsilon = O(n^{-1/3})$, to which end we will appeal to the Fourier representation described in §2. The first observation, using Lemma 2.3, is that the set of pivotals $\mathcal{P}_n$ satisfies

$$\mathbb{E}|\mathcal{P}_n| = p^{-1}\mathbb{P}(f_n = 1)\mathbb{E}[|\mathcal{P}_n| \mid f_n = 1] \leq p^{-1} b n^{1/3}\,,$$

where the last inequality relied on the fact that given that there exists some cycle C_ℓ with $\ell \in \mathcal{I}$ in G, every pivotal edge must in particular belong to C_ℓ and so there can be at most $\ell \leq bn^{1/3}$ such edges. By (2.6), the spectral sample $\mathcal{S}_n$ satisfies

$$\mathbb{E}|\mathcal{S}_n| = p(1-p)\mathbb{E}|\mathcal{P}_n| \leq bn^{1/3}\,,$$

which will rule out noise sensitivity for (f_n) w.r.t. ε by a standard argument. As we have established above that (f_n) is non-degenerate, let $\theta < 1$ be some constant such that $\mathbb{P}(f_n = 1) < \theta$ for any sufficiently large n, and set

$$M = 2b/(1-\theta)\,.$$

Since $\mathbb{P}(\mathcal{S}_n = \emptyset) = \mathbb{P}(f_n = 1) < \theta$ while $\mathbb{P}(|\mathcal{S}_n| > Mn^{1/3}) \leq (1-\theta)/2$ by Markov's inequality, we deduce that

$$\mathbb{P}\left(0 < |\mathcal{S}_n| < Mn^{1/3}\right) > 1 - \theta - \frac{1-\theta}{2} = \frac{1-\theta}{2}\,,$$

and in particular this probability is bounded away from 0. Due to the hypothesis $\varepsilon = O(n^{-1/3})$, we further have

$$(1-\varepsilon)^{|\mathcal{S}_n|}\mathbb{1}_{\{0<|\mathcal{S}_n|<Mn^{1/3}\}} \geq e^{-(1-o(1))\varepsilon M n^{1/3}} \geq c$$

for some fixed $c > 0$, and altogether we obtain that

$$\liminf_{n\to\infty} \operatorname{Cov}\left(f_n(\omega), f_n(\omega^\varepsilon)\right) = \liminf_{n\to\infty} \mathbb{E}\left[(1-\varepsilon)^{|\mathcal{S}_n|}\mathbb{1}_{\{\mathcal{S}_n \neq \emptyset\}}\right] > 0\,,$$

i.e., (f_n) is not $\textsc{Sens}$ w.r.t. ε in this regime.

- *Noise stability iff* $\varepsilon = o(n^{-1/3})$:

Let ω be any configuration corresponding to a graph for which $f_n = 1$, where by definition there exists some cycle W of length $\ell \in (an^{1/3}, bn^{1/3})$ such that $\omega_W \equiv 1$. Under the assumption $\varepsilon \ll n^{-1/3}$ we have that $\mathbb{P}(\omega^\varepsilon_W \equiv 1 \mid \omega) \geq 1 - \varepsilon b n^{1/3} = 1 - o(1)$. In other words, for any ω such that $f_n(\omega) = 1$ we have $\mathbb{P}\left(f_n(\omega^\varepsilon) = 1 \mid \omega\right) = 1 - o(1)$, implying that (f_n) is $\textsc{Stab}$ w.r.t. ε.

To see that (f_n) is not $\textsc{Stab}$ w.r.t. ε whenever $\varepsilon \gtrsim n^{-1/3}$, observe first that if W corresponds to a cycle of length $\ell \in \mathcal{I}$ then

$$\mathbb{P}\left(\omega^\varepsilon_W \not\equiv 1 \mid \omega_W \equiv 1\right) = 1 - (1 - \varepsilon(1-p))^\ell \geq c_0$$

for some fixed $c_0 > 0$ which depends on a as well as the implicit constant in the assumption $\varepsilon \gtrsim n^{-1/3}$. At the same time, with the same notation as above,

$$\mathbb{P}\Big(\bigcap_{W''\in\mathcal{W}_1''}\{\omega^\varepsilon_{W''}\not\equiv 1\}\;\Big|\;\omega_W\equiv 1\Big)\geq \mathbb{P}(f_n=0)>c_1$$

for some fixed $c_1 > 0$ thanks to the above established fact that (f_n) is non-degenerate, whereas by FKG

$$\begin{aligned}\mathbb{P}\Big(\bigcap_{W'\in\mathcal{W}_1'}\{\omega^\varepsilon_{W'}\not\equiv 1\}\;\Big|\;\omega_W\equiv 1\Big)&\geq \prod_{W'\in\mathcal{W}_1'}\mathbb{P}\left(\omega^\varepsilon_{W'}\not\equiv 1\;\Big|\;\omega_W\equiv 1\right)\\ &\geq e^{-(1-o(1))\sum_{m\geq 1}\zeta_m'}\geq c_2\end{aligned}$$

for some fixed $c_2 > 0$ which depends on a, b and the constant in the hypothesis $\varepsilon \gtrsim n^{-1/3}$ as specified in (3.7). Combining the last three inequalities, again by virtue of FKG, we deduce that

$$\mathbb{P}\left(f_n(\omega^\varepsilon)=1\mid \omega_W\equiv 1\right)\leq 1-c_0c_1c_2\,,$$

which by Eq. (3.2) implies that $\mathbb{P}\left(f_n(\omega^\varepsilon)=1\mid f_n(\omega)=1\right)$ is bounded away from 1, precluding noise stability.
This completes the proof. ∎

Remark 3.5. One can construct a function which exhibits a phase transition at the critical window of $\mathcal{G}(n,p)$, and yet not only is a noise of $\varepsilon \gg n^{-1/3}$ (effectively moving ω^ε to the subcritical degenerate regime and then back into the critical window) insufficient for decorrelating $f_n(\omega), f_n(\omega^\varepsilon)$, neither does any fixed $\varepsilon > 0$. The following example demonstrates this.

For some constants $0 < a < b$ to be determined below, let f_n the property that the largest component of G, denoted by $\mathcal{C}_1$, either satisfies $|\mathcal{C}_1| > bn^{2/3}$, or alternatively $an^{2/3} < |\mathcal{C}_1| \leq bn^{2/3}$ while G further contains a triangle.

Clearly, $\mathbb{P}(f_n = 1) = o(1)$ when $G \sim \mathcal{G}(n,p)$ for $p = (1-\xi)/n$ with $\xi \gg n^{-1/3}$ as in that case $|\mathcal{C}_1| = o(n^{2/3})$, whereas $\mathbb{P}(f_n = 1) = 1-o(1)$ when $p = (1+\xi)/n$ for the same ξ since $|\mathcal{C}_1|$ then concentrates around $2\xi n \gg n^{2/3}$ (see, e.g., [6, Chapter 6] and [9, Chapter 5]).

At $p = (1 \pm \xi)/n$ for $\xi = O(n^{-1/3})$ the sequence (f_n) is non-degenerate. An immediate way to ensure this would be to select a sufficiently small and b sufficiently large. Indeed, it is well-known that $|\mathcal{C}_1|/n^{2/3}$ converges in probability to a nontrivial distribution with full support on $\mathbb{R}_+$, and in particular for any small $\delta > 0$ we can select a sufficiently small and b sufficiently large so that $\mathbb{P}(a < |\mathcal{C}_1|n^{-2/3} < b) > 1-\delta$. On this event, f_n identifies with the property g_n of containing a triangle, which is known to be noise stable. In particular,

$$\mathbb{P}\left(f_n(\omega^\varepsilon)=f_n(\omega)\right)\geq \mathbb{P}\left(g_n(\omega^\varepsilon)=g_n(\omega)\right)-2\delta\geq 1-\delta'$$

for some $\delta'(\varepsilon, a, b)$ which can be made arbitrarily small for suitable ε, a, b. This precludes the noise sensitivity of f_n for any fixed $\varepsilon > 0$, as claimed.

We note in passing that f_n satisfies $\sum_x \hat{f}_n(x)^2 = O(n^{-2/3}) = o(1)$, i.e., the BKS criterion for Sens is met, and nevertheless (f_n) is not Sens.

3.3. **Quantitative noise sensitivity for minimum degree.** Analogously to the previous section, here we prove a stronger version of Theorem 1.4, which addresses the noise stability vs. sensitivity at the critical noise level.

Theorem 3.6. *Let f_n be the property that the minimum degree of $\mathcal{G}(n,p)$ is at least k for some fixed $k \geq 1$, and suppose $p = p(n)$ is such that (f_n) is non-degenerate. The following holds depending on the noise parameter $\varepsilon(n)$:*

(i) If $\varepsilon \gg \frac{1}{\log n}$ then (f_n) is Sens *and furthermore* StrSens$_0$ *w.r.t. ε.*
(ii) If $\varepsilon \ll \frac{1}{\log n}$ then (f_n) is Stab *w.r.t. ε.*
(iii) If $\varepsilon \asymp \frac{1}{\log n}$ then (f_n) is neither Sens *w.r.t. ε nor* Stab *w.r.t. ε.*

Moreover, the classification into Sens *w.r.t. ε in (i),* Stab *w.r.t. ε in (ii) or neither in (iii) holds for all graph properties listed in Theorem 1.4.*

Proof. Let $G \sim \mathcal{G}(n,p)$ and let ω denote its edge configuration. Fix $k \geq 1$ and let D_n be the graphs (or corresponding configurations ω) with minimum degree at least k, so that $f_n(\omega) = \mathbb{1}_{\{\omega \in D_n\}}$. The assumption that (f_n) is non-degenerate is well-known (see, e.g., [6,9]) to correspond to

$$p = \frac{\log n + (k-1)\log\log n + O(1)}{n}\,. \tag{3.8}$$

Consider first the range $\frac{1}{\log n} \ll \varepsilon < 1$. In this regime, we wish to compare $\mathbb{P}(\omega^\varepsilon \in D_n^c \mid \omega_W \equiv 0)$ to $\mathbb{P}(\omega \in D_n^c)$ for any 0-witness W for D_n. Clearly, such a 0-witness W is precisely a set of $n-k$ edges incident to a vertex. Denoting the vertices by $v_1, v_2, \ldots, v_n$, assume without loss of generality that this W consists of the edges $\{v_1 v_i : i = 2, \ldots, n-k+1\}$. By the symmetry of witnesses, it is enough to show that for each $\varepsilon > 0$,

$$\liminf_{n\to\infty} \mathbb{P}\left(\omega^\varepsilon \in D_n \mid \omega_W \equiv 0\right) - \mathbb{P}\left(\omega \in D_n\right) \geq 0\,. \tag{3.9}$$

Let A_n be the event that the induced subgraph on the vertices $\{v_2, \ldots, v_n\}$ has minimum degree at least k. We claim that

$$\liminf_{n\to\infty} \mathbb{P}(\omega \in A_n) - \mathbb{P}(\omega \in D_n) \geq 0\,. \tag{3.10}$$

(The limit is in fact 0 but that will not be needed.) It suffices to show that

$$\lim_{n\to\infty} \mathbb{P}(\omega \in A_n^c \cap D_n) = 0\,.$$

Any graph in $A_n^c \cap D_n$ has some vertex v_i with $2 \le i \le n$ such that the degree of v_i is precisely k and $v_1 v_i$ is an edge. By a union bound, the probability that ω satisfies the latter is at most

$$(n-1)\binom{n-2}{k-1} p^k (1-p)^{n-1-k} \le (np)^k e^{-p(n-1-k)} \lesssim \frac{\log n}{n} = o(1)\,,$$

having plugged in the expression for p from (3.8). This establishes (3.10).

Next, let B_n be the set of graphs where the degree of v_1 is at least k. We claim that

$$\lim_{n\to\infty} \mathbb{P}\left(\omega^\varepsilon \in B_n \mid \omega_W \equiv 0\right) = 1\,. \tag{3.11}$$

Indeed, if C_n is the set of graphs where v_1 is isolated then $\mathbb{P}(\omega \in \cdot \mid \omega_W \equiv 0)$ stochastically dominates $\mathbb{P}(\omega \in \cdot \mid \omega \in C_n)$ where $\mathbb{P}(\omega \in \cdot \mid A)$ denotes the conditional distribution of ω conditioned on A. Thus, as B_n is increasing, by FKG we have

$$\mathbb{P}\left(\omega^\varepsilon \in B_n \mid \omega_W \equiv 0\right) \ge \mathbb{P}\left(\omega^\varepsilon \in B_n \mid \omega \in C_n\right) = \mathbb{P}\left(\mathrm{Bin}(n-1, \varepsilon p) \ge k\right)\,. \tag{3.12}$$

Since $p \sim \frac{\log n}{n}$ and $\varepsilon \gg \frac{1}{\log n}$, the above binomial variable concentrates on $(n-1)\varepsilon p \gg k$, hence the last expression is $1-o(1)$. This demonstrates (3.11).

To put it all together, observe that

$$\begin{aligned}\mathbb{P}\left(\omega^\varepsilon \in D_n \mid \omega_W \equiv 0\right) &\ge \mathbb{P}\left(\omega^\varepsilon \in A_n \cap B_n \mid \omega_W \equiv 0\right)\\ &= \mathbb{P}\left(\omega^\varepsilon \in A_n \mid \omega_W \equiv 0\right) \mathbb{P}\left(\omega^\varepsilon \in B_n \mid \omega_W \equiv 0\right)\,,\end{aligned}$$

since the events A_n and B_n are (conditionally) independent. Plugging in (3.11) and using the independence of $\{\omega^\varepsilon \in A_n\}$ and $\{\omega_W \equiv 0\}$ we conclude that

$$\mathbb{P}\left(\omega^\varepsilon \in D_n \mid \omega_W \equiv 0\right) \ge \mathbb{P}\left(\omega^\varepsilon \in A_n\right) - o(1)\,,$$

and the required inequality (3.9) now follows from (3.10) and completes the proof of Part (i).

For Part (ii) consider any $\omega \in D_n^c$, whereby the corresponding graph G contains some vertex v_i of degree less than k. Since $\varepsilon = o(1/\log n)$, the probability that the degree of v_i increases due to the noise is at most $(n-1)\varepsilon p = o(1)$, and so $\mathbb{P}\left(\omega^\varepsilon \in D_n^c \mid \omega\right) = 1 - o(1)$. Translating this in terms of f_n, for any ω such that $f_n(\omega) = 0$ we have $\mathbb{P}\left(f_n(\omega^\varepsilon) = 0 \mid \omega\right) = 1 - o(1)$, which establishes noise stability w.r.t. ε.

We next proceed to Part (iii), addressing the critical regime of $\varepsilon \asymp \frac{1}{\log n}$. To show (f_n) is not STAB w.r.t. ε, note first that the binomial variable in the right-hand-side of (3.12) is now approximately Poisson with mean bounded away from 0 and ∞, implying (by the same line of arguments as above) that

$$\mathbb{P}\left(\omega^\varepsilon \in D_n \mid \omega_W \equiv 0\right) \ge \delta \mathbb{P}(\omega \in D_n)$$

for some fixed $\delta > 0$ and all n, or equivalently,

$$\mathbb{P}\left(\omega^{\varepsilon} \in D_n^c \mid \omega_W \equiv 0\right) \leq 1 - \delta\mathbb{P}(\omega \in D_n)\,.$$

Appealing to Eq. (3.2) from the proof of Lemma 3.1, and using the symmetry of 0-witnesses, we now deduce that

$$\mathbb{P}\left(f_n(\omega^{\varepsilon}) = 0 \mid f_n(\omega) = 0\right) \leq 1 - \delta\mathbb{P}(f_n = 1)\,,$$

which precludes noise stability w.r.t. ε as (f_n) is non-degenerate.

To rule out noise sensitivity for $\varepsilon \asymp \frac{1}{\log n}$, as in the proof of Theorem 3.4 we appeal to the Fourier representation of $f_n(\omega^{\varepsilon})$. For any ω such that $f_n(\omega) = 0$, an edge uv can only be pivotal if every $w \neq u, v$ has degree at least k in ω. Moreover, if both u, v have degree $k-1$ in ω then this would be the unique pivotal edge, and otherwise $|\mathcal{P}_n| = n - k$. In particular, using (2.6) and Lemma 2.3, we see that

$$\mathbb{E}|\mathcal{S}_n| = p(1-p)\mathbb{E}|\mathcal{P}_n| = p\mathbb{P}(f_n = 0)\mathbb{E}[|\mathcal{P}_n| \mid f_n = 0] \leq (1 + o(1))\log n\,.$$

As (f_n) is non-degenerate by hypothesis, let $\theta < 1$ be some constant such that $\mathbb{P}(f_n = 1) < \theta$ for large enough n, and set $M = 2/(1-\theta)$. Since the spectral sample $\mathcal{S}_n$ satisfies $\mathbb{P}(\mathcal{S}_n = \emptyset) = \mathbb{P}(f_n = 1)$, Markov's inequality implies that

$$\mathbb{P}\left(0 < |\mathcal{S}_n| < M\log n\right) > 1 - \theta - \frac{1-\theta}{2} - o(1) = \frac{1-\theta}{2} - o(1)\,.$$

Consequently, when $\varepsilon = O(1/\log n)$ there exists some $c > 0$ such that

$$(1-\varepsilon)^{|\mathcal{S}_n|}\mathbb{1}_{\{0 < |\mathcal{S}_n| < M\log n\}} \geq e^{-(1-o(1))\varepsilon M\log n} \geq c > 0\,,$$

and so

$$\liminf_{n\to\infty} \operatorname{Cov}\left(f_n(\omega), f_n(\omega^{\varepsilon})\right) = \liminf_{n\to\infty} \mathbb{E}\left[(1-\varepsilon)^{|\mathcal{S}_n|}\mathbb{1}_{\{\mathcal{S}_n \neq \emptyset\}}\right] > 0\,,$$

i.e., (f_n) is not Sens w.r.t. ε in this regime.

Finally, it remains to extend the classification of either Sens or Stab w.r.t. ε to the graph properties listed in Theorem 1.4. To this end, recall the well-known facts (see [3,6,9]) that each such property (g_n) is asymptotically equal to the property (f_n) of having minimum degree at least k (for an appropriate k), in the sense that $\lim_{n\to\infty} \mathbb{P}(f_n \neq g_n) = 0$. It is elementary that if (f_n) is noise sensitive (noise stable) and (g_n) is asymptotically equal to (f_n) then (g_n) is noise sensitive (noise stable), since

$$|\mathbb{E}[f_n(\omega^{\varepsilon})f_n(\omega)] - \mathbb{E}[g_n(\omega^{\varepsilon})g_n(\omega)]| \leq 2\mathbb{P}(f_n \neq g_n)\,,$$

thus translating the quantitative statements on (f_n) to (g_n), as required. ∎

Remark 3.7. As an alternative way to obtain noise sensitivity for $\mathcal{G}(n,p)$ having minimum degree at least k, one could appeal to [15, Theorem 1.8] and present a randomized algorithm for this event whose probability of querying any given edge tends to 0. This would imply a quantitative noise sensitivity result, albeit weaker than the sharp one obtained above.

4. Noise sensitivity of witness-transitive functions

Let f be a monotone Boolean function on a domain Ω. We say that f is 1-*witness-transitive* if the set of automorphisms of f (the set of permutations π on Ω under which f is invariant, i.e., $f \equiv f \circ \pi$) is such that for any two witnesses $W, W' \in \mathcal{W}_1(f)$ there exists an automorphism of f mapping W to W'. That is to say, any two 1-witnesses for f are equivalent.

For instance, the classical examples for noise sensitive functions which were mentioned in the introduction, tribes and recursive majority, are both 1-witness-transitive, as is the property of containing an unlabeled copy of a certain graph H in a random graph $G \sim \mathcal{G}(n,p)$.

4.1. A Poissonization tool for strong noise sensitivity.

Our goal in this section is the prove a sufficient condition for strong noise sensitivity of 1-witness-transitive functions. This condition will be in the form of a Poisson approximation of the total number of occurring 1-witnesses, as stated next.

Proposition 4.1. *Let (f_n) be a sequence of 1-witness-transitive monotone Boolean functions. Let $W_\star = W_\star(n)$ be a canonical 1-witness for f_n, and suppose that $(1-p_n)|W_\star| \to \infty$ with n. Let $X_n = \sum_{W \in \mathcal{W}_1(f_n)} \mathbb{1}_{\{\omega_W \equiv 1\}}$ count the occurring 1-witnesses, and assume that for some $\lambda \in \mathbb{R}_+$ we have:*

$$\lim_{n\to\infty} \mathbb{E}[X_n] = \lambda \quad \text{and} \quad \lim_{n\to\infty} \mathrm{Var}(X_n) = \lambda\,, \tag{4.1}$$

$$\lim_{n\to\infty} \mathbb{E}\left[X_n \mid \omega_{W_\star} \equiv 0\right] = \lambda\,. \tag{4.2}$$

Then $X_n \xrightarrow{\mathrm{d}} \mathrm{Po}(\lambda)$ as $n \to \infty$ and (f_n) is Sens *and moreover* StrSens$_1$*. Furthermore, quantitative* Sens *(as well as* StrSens$_1$*) holds w.r.t. $\varepsilon(n)$ iff*

$$\varepsilon \gg \left[(1-p_n)|W_\star|\right]^{-1}\,. \tag{4.3}$$

Proof. The fact that the X_n converges in distribution to a Poisson random variable under the given assumptions follows from a standard application of the Chen-Stein method (see, e.g., [2, Theorem 1] and [9, Theorem 6.24]). Indeed, writing $I_W = \mathbb{1}_{\{\omega_W \equiv 1\}}$ for $W \in \mathcal{W}_1$ we see that $\mathbb{P}(I_W) = p^{|W|} = o(1)$ thanks to the assumption $(1-p)|W_\star| \to \infty$. As these indicators are positively related by FKG, we can invoke a simplified form of the Chen-Stein method (see [9, Theorem 6.24]), at which point the assumptions (4.1) imply that

$$\|X_n - \mathrm{Po}(\lambda)\|_{\mathrm{TV}} \le \frac{\mathrm{Var}(X_n)}{\mathbb{E}[X_n]} - 1 + 2 \max_{W\in\mathcal{W}_1} \mathbb{P}(I_W) = o(1)\,.$$

Linking the above to strong noise sensitivity will be achieved by the next key definition, which we phrase for general monotone Boolean functions (not necessarily witness-transitive) as it may be of independent interest. The proof of Proposition 4.1 will be continued after this detour.

Definition 4.2. A sequence (f_n) of monotone increasing Boolean functions is said to be *1-witness-disjoint* if

$$\lim_{n\to\infty} \max_{W\in\mathcal{W}_1} \mathbb{P}\Bigg(\bigcup_{\substack{W'\in\mathcal{W}_1\setminus\{W\} \\ W'\cap W\neq\emptyset}} \{\omega_{W'}\equiv 1\} \;\Bigg|\; \omega_W \equiv 1 \Bigg) = 0\,.$$

Note that the above condition would trivially hold if every pair of distinct 1-witnesses were disjoint (as is the case for instance for the tribes function, where the 1-witnesses are full blocks). In a sense, Definition 4.2 provides an approximation to such a situation, which, as we show next, is powerful enough to imply (quantitative) strong noise sensitivity.

Lemma 4.3. *Let (f_n) be a sequence of monotone Boolean functions that is 1-witness-disjoint. Let $\varepsilon(n)$ be such that $\varepsilon(1-p_n)\ell_n \to \infty$ with n, where ℓ_n is the minimum size of a 1-witness for f_n. Then (f_n) is $\textsc{StrSens}_1$ w.r.t. ε.*

Proof. Thanks to our assumption on ε we have that for any 1-witness W,

$$\mathbb{P}\left(\omega^\varepsilon_W \equiv 1 \mid \omega_W\equiv 1\right) = (1-\varepsilon(1-p))^{|W|} \le e^{-\varepsilon(1-p_n)\ell_n} = o(1)\,,$$

and therefore

$$\begin{aligned}\mathbb{P}(f_n(\omega^\varepsilon)=1 \mid \omega_W\equiv 1) &= \mathbb{P}\Bigg(\bigcup_{W'\in\mathcal{W}_1} \{\omega^\varepsilon_{W'}\equiv 1\} \;\Bigg|\; \omega_W\equiv 1\Bigg) \\ &\le \mathbb{P}\Bigg(\bigcup_{W'\in\mathcal{W}_1\setminus\{W\}} \{\omega^\varepsilon_{W'}\equiv 1\} \;\Bigg|\; \omega_W\equiv 1\Bigg) + o(1)\,.\end{aligned} \tag{4.4}$$

Define the events A_n and B_n by

$$A_n = \bigcup_{\substack{W'\in\mathcal{W}_1 \\ W'\cap W=\emptyset}} \{\omega^\varepsilon_{W'}\equiv 1\} \quad,\quad B_n = \bigcup_{\substack{W'\in\mathcal{W}_1\setminus\{W\} \\ W'\cap W\neq\emptyset}} \{\omega^\varepsilon_{W'}\equiv 1\}\,.$$

Of course, $\mathbb{P}(A_n \mid \omega_W\equiv 1) \le \mathbb{P}(f_n=1)$ as the events A_n and $\{\omega_W\equiv 1\}$ are mutually independent, and together with (4.4) this yields

$$\mathbb{P}(f_n(\omega^\varepsilon)=1 \mid \omega_W\equiv 1) - \mathbb{P}\left(f_n=1\right) \le \mathbb{P}\left(B_n \mid \omega_W\equiv 1\right) + o(1)\,. \tag{4.5}$$

Next, since the distribution of ω^ε conditioned on $\omega_W \equiv 1$ is stochastically dominated by the distribution of ω conditioned on $\omega_W \equiv 1$,

$$\mathbb{P}\left(B_n \mid \omega_W \equiv 1\right) \leq \mathbb{P}\bigg(\bigcup_{\substack{W' \in \mathcal{W}_1 \setminus \{W\} \\ W \cap W' \neq \emptyset}} \{\omega_{W'} \equiv 1\} \;\bigg|\; \omega_W \equiv 1 \bigg) .$$

Now take a supremum over $W \in \mathcal{W}_1$, under which the final expression goes to 0 by Definition 4.2. Combined with (4.5), this concludes the proof. ∎

Returning to the proof of Proposition 4.1, we claim that under the hypotheses $\mathbb{E}X_n \to \lambda$ and $\mathbb{E}[X_n \mid \omega_{W_\star} \equiv 0] \to \lambda$ given there, the extra assumption $\mathrm{Var}(X_n) \to \lambda$ in (4.1) is equivalent to having

$$\lim_{n\to\infty} \sum_{\substack{W \in \mathcal{W}_1 \setminus \{W_\star\} \\ W \cap W_\star \neq \emptyset}} \mathbb{P}\left(\omega_W \equiv 1 \mid \omega_{W_\star} \equiv 1\right) = 0 \,. \tag{4.6}$$

As per Definition (4.2), this would imply (thanks to the witness-transitivity) that (f_n) is 1-witness-disjoint, and in light of Lemma 4.3 we will thereafter arrive at strong noise sensitivity w.r.t. ε assuming $\varepsilon \gg [(1-p_n)|W_\star|]^{-1}$. Indeed, this equivalence is seen by expanding $\mathbb{E}X_n^2 = \mathbb{E}X_n + \Gamma + \Delta$ where

$$\Gamma = \sum_{\substack{W, W' \in \mathcal{W}_1 \\ W' \cap W = \emptyset}} \mathbb{P}\left(\omega_W \equiv 1 \,,\, \omega_{W'} \equiv 1\right) \,, \qquad \Delta = \sum_{\substack{W \neq W' \in \mathcal{W}_1 \\ W' \cap W \neq \emptyset}} \mathbb{P}\left(\omega_W \equiv 1 \,,\, \omega_{W'} \equiv 1\right) \,.$$

The expression for Γ, which is clearly at most $(\mathbb{E}X_n)^2$, can be rewritten by virtue of the independence of W, W' and the witness-transitivity as

$$\sum_{W \in \mathcal{W}_1} \mathbb{P}(\omega_W \equiv 1) \sum_{\substack{W' \in \mathcal{W}_1 \\ W \cap W' = \emptyset}} \mathbb{P}(\omega_{W'} \equiv 1) = \mathbb{E}[X_n]\, \mathbb{E}[X_n \mid \omega_{W_\star} \equiv 0] \,,$$

which is at least $(1 - o(1))\lambda^2$ by the aforementioned hypotheses. At this point, $\mathrm{Var}(X_n) \to \lambda$ if and only if $\Delta \to 0$, and yet by the witness-transitivity,

$$\Delta = \mathbb{E}[X_n] \sum_{\substack{W \in \mathcal{W}_1 \setminus \{W_\star\} \\ W \cap W_\star \neq \emptyset}} \mathbb{P}\left(\omega_W \equiv 1 \mid \omega_{W_\star} \equiv 1\right) \,.$$

This completes the argument for $\textsc{StrSens}_1$ whenever $\varepsilon \gg [(1-p_n)|W_\star|]^{-1}$.

In the regime $\varepsilon \lesssim [(1-p_n)|W_\star|]^{-1}$, the sequence (f_n) will not be $\textsc{Sens}$, by the same Fourier argument given in the previous section: As before, $\mathbb{E}[|\mathcal{P}_n| \mid f_n = 1] \leq |W_\star|$ since we can take an arbitrary witness W that occurs in a configuration for which $f_n = 1$ and note that every pivotal edge must then belong to W. It then follows that $\mathbb{E}|\mathcal{S}_n| \leq (1-p_n)|W_\star|$, thus for $\varepsilon \lesssim [(1-p_n)|W_\star|]^{-1}$ we have $\liminf_{n\to\infty} \mathrm{Cov}(f_n(\omega), f_n(\omega^\varepsilon)) > 0$ due to the Fourier levels $0 < |\mathcal{S}_n| < M(1-p_n)|W_\star|$ for a suitable constant $M > 0$. ∎

Example (Tribes)**.** We have seen in the previous section that the tribes function is $\textsc{StrSens}_1$ by a direct analysis of $\mathbb{P}(f_n(\omega^\varepsilon) \mid \omega_W \equiv 1) - \mathbb{P}(f_n = 1)$. We will now derive this fact via an immediate application of Proposition 4.1. Let $m = \log_2 n - \log_2 \log_2 n$ denote the block size in f_n (as usual, divisibility issues can be solved by ignoring one exceptional block; we omit floors and ceilings for brevity), and note that a canonical 1-witness $W_\star$ consists of a full block and so $(1-p_n)|W_\star| \asymp m \to \infty$. Moreover, X_n is simply a $\mathrm{Bin}(n/m, 2^{-m})$ random variable, thus both $\mathbb{E}[X_n] \to 1$ and $\mathrm{Var}(X_n) \to 1$ as $n \to \infty$, while under the conditioning $\omega_{W_\star} \equiv 0$, the variable X_n becomes a $\mathrm{Bin}(n/m - 1, 2^{-m})$ variable, whose mean again converges to 1 as $n \to \infty$. The conditions of Proposition 4.1 are thus met, yielding that (f_n) is $\textsc{StrSens}_1$. Furthermore, it is such iff $\varepsilon \gg 1/m$ while it is not $\textsc{Sens}$ for $\varepsilon = O(1/m)$.

Remark 4.4. It is easily seen from the proof of the above proposition that in order to conclude (quantitative) strong noise sensitivity without making any claim on the limiting distribution of X_n, the conditions (4.1) and (4.2) may be replaced by

$$0 < \liminf_{n\to\infty} \mathbb{E}[X_n] \le \limsup_{n\to\infty} \mathbb{E}[X_n] < \infty\,, \tag{4.7}$$

$$\lim_{n\to\infty} \big|\mathrm{Var}(X_n) - \mathbb{E}[X_n]\big| = 0\,, \tag{4.8}$$

$$\lim_{n\to\infty} \big|\mathbb{E}[X_n] - \mathbb{E}\left[X_n \mid \omega_{W_\star} \equiv 0\right]\big| = 0\,. \tag{4.9}$$

Under these assumptions, (f_n) is non-degenerate thanks to FKG (bounding $\mathbb{P}(X = 0)$ away from 0) and Cauchy-Schwarz (bounding $\mathbb{P}(X > 0)$ away from 0) as in the proof of Theorem 3.4. Following the proof of Proposition 4.1 we see that, as $\mathbb{E}[X_n] = O(1)$, conditions (4.8) and (4.9) yield $\Delta \to 0$, from which point the original argument completes the proof.

As an immediate corollary of the results proved above, we get the following sufficient condition for strong noise sensitivity of containing an unlabeled copy of a graph in the Erdős-Rényi random graph.

Corollary 4.5. *Let $G \sim \mathcal{G}(n,p)$ and let H_n be a graph with $k \ll \sqrt{n}$ vertices and $\ell \gg 1/(1-p)$ edges. Let $f_n = \mathbb{1}_{\{X_n > 0\}}$ where X_n counts the number of unlabeled copies of H_n in G, and suppose that*

$$0 < \liminf_{n\to\infty} \mathbb{E}[X_n] \le \limsup_{n\to\infty} \mathbb{E}[X_n] < \infty\,,$$

$$\lim_{n\to\infty} \big|\mathrm{Var}(X_n) - \mathbb{E}[X_n]\big| = 0\,.$$

Then (f_n) is $\textsc{Sens}$ *and moreover* $\textsc{StrSens}_1$*. Furthermore, quantitative* $\textsc{StrSens}_1$ *holds if $\varepsilon \gg [(1-p)\ell]^{-1}$ and otherwise (f_n) is not* $\textsc{Sens}$ *w.r.t. ε.*

Proof. Appealing to Proposition 4.1, with the canonical witness $W_\star$ being a copy of H_n, we see that (4.7),(4.8) and the fact that $(1-p_n)|W_\star| \to \infty$ are explicitly assumed. For the final condition (4.9) in Remark 4.4, note that $\mathbb{E}[X_n] = \binom{n}{k}p^\ell k!/\operatorname{aut}(H_n)$ where $\operatorname{aut}(H_n)$ is the size of the automorphism group of H_n, while $\mathbb{E}[X_n \mid \omega_{W_\star} \equiv 0] \geq \binom{n-k}{k}p^\ell k!/\operatorname{aut}(H_n) \sim \mathbb{E}[X_n]$ thanks to the hypothesis that $k \ll \sqrt{n}$, as desired. ∎

4.2. **Noise sensitivity for cliques.** This section is devoted to the noise sensitivity of cliques of any size $1 \ll k_n = n^{o(1)}$ in the random graph $\mathcal{G}(n,p)$, corresponding to the maximum cliques for $n^{-o(1)} \leq p \leq 1 - n^{-o(1)}$.

Proof of Theorem 1.5. The statement of the theorem will follow from Corollary 4.5 via the standard second moment analysis which implies the 2-point concentration of the clique number k_n of $\mathcal{G}(n,1/2)$, generalized to the case of $1 \ll k_n = n^{o(1)}$. An outline of this second moment calculation for $p = 1/2$ is given in [1,6], and here we provide the full details for the sake of completeness.

Let $X_k = X_k(n)$ count the number of cliques of size $k = k_n$ in $G \sim \mathcal{G}(n,p)$, and note that $\mathbb{E}X_k = \binom{n}{k}p^{\binom{k}{2}}$ can be assumed to be bounded away from 0, as otherwise $\mathbb{P}(X_k = 0) = 1 - o(1)$ and so the sequences k_n, p_n would correspond to a degenerate sequence (f_n) countering the hypothesis of the theorem.

In order to estimate the variance of X_k, as usual write $\operatorname{Var}(X_k) \leq \mathbb{E}X_k + \Delta$ for $\Delta = \sum_{H_1,H_2} \mathbb{P}(H_1 \subset G,\, H_2 \subset G)$, where the summation runs over all pairs of potential k-cliques $H_1 \neq H_2$ that have some edges in common. We claim that the required result would follow from showing that

$$\Delta = o\left((\mathbb{E}X_k)^2\right) . \tag{4.10}$$

Indeed, suppose that $\mathbb{E}X_k \to \infty$ with n. In this case (4.10) implies that $\operatorname{Var}(X_k) \ll (\mathbb{E}X_k)^2$, thus by Chebyshev's inequality X_k concentrates about its mean and in particular $\mathbb{P}(X_k > 0) = 1 - o(1)$, contradicting the hypothesis that (f_n) is non-degenerate. We thus have that $\mathbb{E}X_k$ is bounded away from 0 and ∞ for any sufficiently large n, and a closer look at $\mathbb{E}X_k \sim (np^{(k-1)/2})^k/k!$ reveals that this can only occur if

$$p = n^{-(2+o(1))/k} . \tag{4.11}$$

Hence, either $k = O(\log n)$, in which case p is bounded away from 1 and in particular the number of edges $\ell = \binom{k}{2}$ satisfies $\ell \gg 1/(1-p)$, or we have $k \gg \log n$ and then $(1-p)^{-1} = O(k/\log n) = o(k^2)$, again satisfying the condition $\ell \gg 1/(1-p)$ in Corollary 4.5. Finally, it follows from (4.10) that $|\mathbb{E}[X_k] - \operatorname{Var}(X_k)| \to 0$ and the mentioned corollary now provides the required statement on the strong noise sensitivity of (f_n). Furthermore, we obtain that quantitative (strong) noise sensitivity holds iff $\varepsilon \gg [(1-p)k^2]^{-1}$.

A classical fact worth reiterating is that for p as given in (4.11), and writing $\psi_j = \mathbb{E}[X_{j+1}]/\mathbb{E}[X_j]$, one has $\psi_j = p^j(n-j)/(j+1)$, thus the map

$j \mapsto \mathbb{E}X_j$ (starting at $\mathbb{E}X_1 = n$) is unimodal and for $j \sim k$ it satisfies that $\psi_j = n^{-1+o(1)}$. By the discussion above, this yields the 2-point concentration of the clique number, and moreover a 1-point concentration except for those rare values of n when, e.g., the first $\mathbb{E}X_j$ to drop below 1 (say) is still bounded away from 0. These are precisely the non-degenerate cases.

To obtain (4.10), one breaks Δ down into $\Delta = \sum_{i=2}^{k-1} \Delta_i$ according to i, the number of common vertices between H_1, H_2 (at least 2 to accommodate a common edge and less than k to keep the cliques distinct), obtaining that

$$\Delta_i = \binom{n}{k}\binom{k}{i}\binom{n-k}{k-i} p^{2\binom{k}{2}-\binom{i}{2}} .$$

Fix any arbitrary $0 < \delta < \frac{1}{2}$ and let

$$\alpha := (1+\delta)\frac{\log n}{\log(1/p)} \quad , \quad \beta := (2-\delta)\frac{\log(n/k^2)}{\log(1/p)} ,$$

noting that $\alpha < \beta$ for large enough n since $k = n^{o(1)}$. It is now easy to see that for any $i \leq \beta$ we have

$$\frac{\Delta_i}{(\mathbb{E}X_k)^2} = \frac{\binom{k}{i}\binom{n-k}{k-i}}{\binom{n}{k}p^{\binom{i}{2}}} \leq \frac{1+o(1)}{i!}\left[\frac{k^2}{np^{(i-1)/2}}\right]^i \leq \frac{1+o(1)}{i!}\left(\frac{k^2}{n}\right)^{\delta i/2} ,$$

where the first inequality holds for $k \ll \sqrt{n}$ and the second one for $i \leq \beta$. It then follows that

$$\sum_{2 \leq i \leq \beta} \frac{\Delta_i}{(\mathbb{E}X_k)^2} \leq n^{-\delta+o(1)} = o(1) ,$$

and we now proceed to handle the remaining Δ_i's (with some overlap). Since $\mathbb{E}X_k$ is bounded away from 0 we see that for any $\alpha \leq i < k$,

$$\frac{\Delta_i}{(\mathbb{E}X_k)^2} \lesssim \frac{\Delta_i}{\mathbb{E}X_k} = \binom{k}{i}\binom{n-k}{k-i}p^{\binom{k}{2}-\binom{i}{2}} \leq \frac{\left(k(n-k)p^i\right)^{k-i}}{((k-i)!)^2} \leq \left(kn^{-\delta}\right)^{k-i} ,$$

with the last inequality stemming from the fact that $i \geq \alpha$. In particular,

$$\sum_{\alpha \leq i \leq k-1} \frac{\Delta_i}{(\mathbb{E}X_k)^2} \leq n^{-\delta+o(1)} = o(1) ,$$

and as $\alpha < \beta$ this establishes (4.10), completing the proof. ■

In the special case where the sequence of probabilities $p(n)$ is such that $\mathbb{E}[X_k] \to \lambda$ for some fixed $\lambda > 0$, (i.e., $\binom{n}{k}p^{\binom{k}{2}}$ converges), the above proof further gives (via the Chen-Stein method, as in the proof of Proposition 4.1) that $X_k \xrightarrow{\mathrm{d}} \mathrm{Po}(\lambda)$. However, a Poisson limit for the number of copies of a graph is not a necessary condition for $\textsc{StrSens}_1$, as the next remark shows.

Remark 4.6 (Disjoint union of two cliques)**.** Consider the property f_n of containing a disjoint union of two cliques $K_k \cup K_k$ when the clique size $1 \ll k \ll n^{o(1)}$ is exactly such that the probability of witnessing a single such clique in $G \sim \mathcal{G}(n,p)$ is non-degenerate. We claim that containing this graph, which we note is balanced but not strictly balanced, is $\mathrm{StrSens}_1$ despite the fact that the corresponding number of copies of this graph is not asymptotically Poisson, nor is this property 1-witness-disjoint. Indeed, one easily sees that the condition in Definition 4.2 fails since upon conditioning on two disjoint cliques H' and H'' (which together form a 1-witness for f_n), there exists a third clique $\tilde{H}$, disjoint from H' and H'', with probability bounded away from 0 (in which case $\tilde{H} \cup H'$ for instance would be a 1-witness nontrivially intersecting $H' \cup H''$).

In order to establish $\mathrm{StrSens}_1$ for this property, we modify the second moment calculation in the proof of Theorem 1.5 as follows. Letting $\mathcal{F}$ denote all potential copies of a single clique K_k in G, take $H', H'' \in \mathcal{F}$ to be two disjoint such copies, arbitrarily chosen, and define

$$\Delta_{i,j} := \sum_{\substack{H\in\mathcal{F}\\ |V(H)\cap V(H')|=i\\ |V(H)\cap V(H'')|=j}} \mathbb{P}\left(H \subset G \mid H', H'' \subset G\right) ,$$

whence

$$\Delta_{i,j} = \binom{n-2k}{k-(i+j)}\binom{k}{i}\binom{k}{j} p^{\binom{k}{2}-\binom{i}{2}-\binom{j}{2}} .$$

As usual, the probability of encountering a copy of $K_k \cup K_k$ that does not intersect neither H' nor H'' is at most $\mathbb{P}(f_n = 1)$, while the probability of encountering even a single K_k that intersects H' but not H'', conditioned on $H', H'' \subset G$, was shown in the proof of Theorem 1.5 to tend to 0. Hence, it remains to show that $\sum_{2\leq i,j<k} \Delta_{i,j} = o(1)$. The case where

$$i + j \leq (2-\delta)\frac{\log n}{\log(1/p)} \tag{4.12}$$

for some small $\delta > 0$ is treated as in the proof of Theorem 1.5 by writing

$$\frac{\Delta_{i,j}}{\binom{n}{k}p^{\binom{k}{2}}} \lesssim \left[\frac{k^2}{np^{(\binom{i}{2}+\binom{j}{2})/(i+j)}}\right]^{i+j} \leq \left[\frac{k^2}{np^{(i+j)/2}}\right]^{i+j} \leq \left(\frac{k^{4/\delta}}{n}\right)^{\delta(i+j)/2} ,$$

which is at most $n^{-2\delta+o(1)}$ by the assumption $i,j \geq 2$. (Note the usage of (4.12) for the last inequality.) The complement range for (4.12) is handled in the following way. Without loss of generality, assume $i \geq j$, and using the fact that $\binom{k}{2} - \binom{i}{2} - \binom{j}{2} \geq (k-(i+j))(i+j) + ij$ we can infer that

$$\Delta_{i,j} \leq \left(\frac{e(n-2k)}{(k-(i+j)) \ \vee\ 1} kp^{i+j}\right)^{k-(i+j)} \left(k^2 p^i\right)^j .$$

The first term on the right-hand-side is at most $n^{(-1+\delta+o(1))(k-(i+j))}$ by the assumption on $i+j$, whereas the second term is at most $n^{(-1+\delta/2+o(1))j}$, which in turn is at most $n^{-2+\delta+o(1)}$ thanks to the fact that $j \geq 2$. Summing these over $2 \leq i, j < k$ now leads to the conclusion that (f_n) is $\textsc{StrSens}_1$.

4.3. **Proof of Theorem 1.6, Part 1.** This part of the theorem is a simple consequence of Corollary 4.5 via an elegant Poisson approximation argument of Bollobás [6, Theorems 4.1 and 4.3]. We include the proof for completeness.

Lemma 4.7. *Let H_n be a strictly balanced graph with $\ell_n \leq \sqrt{\frac{\log n}{\log\log n}}$ edges, and let X_n count its number of copies in $G \sim \mathcal{G}(n,p)$ for $p = p(n)$ such that*

$$0 < \liminf_{n\to\infty} \mathbb{E}[X_n] \leq \limsup_{n\to\infty} \mathbb{E}[X_n] < \infty\,.$$

Then

$$\lim_{n\to\infty} \left(\mathrm{Var}(X_n) - \mathbb{E}[X_n]\right) = 0\,. \tag{4.13}$$

Proof. Denote the number of vertices and edges of H_n by k and ℓ, and let $\mathcal{F}$ denote the set of all potential copies of H_n in $G \sim \mathcal{G}(n,p)$. As before, we break up the second moment of X_n into

$$\begin{aligned}\mathbb{E}[X_n^2] &= \mathbb{E}[X_n] + \sum_{\substack{H' \neq H'' \in \mathcal{F} \\ H' \cap H'' = \emptyset}} \mathbb{P}\big(H', H'' \subset G\big) + \sum_{\substack{H' \neq H'' \in \mathcal{F} \\ H \cap H'' \neq \emptyset}} \mathbb{P}\big(H', H'' \subset G\big) \\ &\leq \mathbb{E}[X_n] + (1-o(1))(\mathbb{E}[X_n])^2 + \sum_{\substack{H' \neq H'' \in \mathcal{F} \\ H \cap H'' \neq \emptyset}} \mathbb{P}\big(H', H'' \subset G\big)\,,\end{aligned}$$

where the inequality between the lines used the fact that $k \ll \sqrt{n}$ as well as the assumption that $\mathbb{E}[X_n]$ is bounded away from 0 and ∞, as in the proof of Proposition 4.1. We will show below that the summation in the right-hand-side is $o(1)$, which will then imply (4.13).

Given H' and H'' whose vertices overlap, put $t = |\{v \in V(H'') \setminus V(H')\}|$, whence $0 \leq t < k$. (Observe that $t = 0$ is possible since H' and H'' can correspond to different copies of H_n even if their vertex sets are the same.) The number of vertices in $H' \cap H''$ is therefore $k - t$.

Assume for the moment that $t > 0$. Since H_n is strictly balanced, it follows that the number of edges of H'' between vertices in $V(H') \cap V(H'')$ is strictly less than $(k-t)\ell/k$. Thus, the number of edges in H'' with at least one endpoint not in $V(H') \cap V(H'')$ is strictly more than $\ell - (k-t)\ell/k = t\ell/k$. Since the number of such edges is an integer, there are in fact at least $t\ell/k + 1/k$ such edges, hence the number of edges in $H' \cup H''$ is at least $\ell + \frac{t\ell+1}{k}$. Now, if $t = 0$, the number of edges in $H' \cup H''$ is at least $\ell + 1$ (since $H' \neq H''$). Altogether, this number is always at least $\ell + (t\ell + 1)/k$.

It is easy to see that the third summand is at most

$$\sum_{s=k}^{2k-1} \binom{n}{s} \left(\binom{s}{k} \frac{k!}{a} \right)^2 p^{(s\ell+1)/k}$$

where a denotes the size of the automorphism group of H_n and s corresponds to $k+t$. The last sum is at most

$$\sum_{s=k}^{2k-1} \frac{n^s}{s!} \left(\frac{s!}{a} \right)^2 p^{(s\ell+1)/k} . \tag{4.14}$$

Note now that

$$\mathbb{E}[X_n] = \binom{n}{k} \frac{k!}{a} p^\ell = (1+o(1)) \frac{n^k p^\ell}{a}$$

since $k \ll \sqrt{n}$. It follows that

$$p = \frac{(a\mathbb{E}[X_n])^{1/\ell}}{n^{k/\ell}} (1+o(1))^{1/\ell} .$$

Substituting this back into (4.14) yields that the third sum that we are interested in is at most

$$(1+o(1)) \sum_{s=k}^{2k-1} \frac{1}{s!} \left(\frac{s!}{a} \right)^2 (a\mathbb{E}[X_n])^{(s+\ell^{-1})/k} \frac{1}{n^{1/\ell}} .$$

Since $a \geq 1$ and $s/k + (\ell k)^{-1} \leq 2$, the above sum is at most

$$(1+o(1)) k \left(\mathbb{E}[X_n]^2 \ \vee \ 1 \right) (2k)! \frac{1}{n^{1/\ell}} .$$

Since $k \leq \ell + 1$, this is at most

$$(1+o(1))(\ell+1) \left(\mathbb{E}[X_n]^2 \ \vee \ 1 \right) \frac{(2\ell+2)!}{n^{1/\ell}} .$$

It is easy to verify, using the fact that $\mathbb{E}[X_n]$ is bounded away from 0 and ∞ and that $\ell \leq \sqrt{\frac{\log n}{\log\log n}}$, that this last term is $o(1)$, as desired. ∎

4.4. **Proof of Theorem 1.6, Part 2.** Consider $G \sim \mathcal{G}(n, \lambda/n)$ for some large enough fixed $\lambda > 1$, and let H_n be the graph comprised of two triangles connected by a path of length

$$r_n = \lfloor \tfrac{3}{2} \log_\lambda n \rfloor . \tag{4.15}$$

(Any choice of $(1+\delta)\log_\lambda n \leq r_n \leq (2-\delta)\log_\lambda n$ would be valid, as will later become evident; we consider this particular r_n to simplify the presentation.) It is easy to see that H_n is strictly balanced. That $\mathbb{1}_{\{H_n \subset G\}}$ is not SENS will follow from the next two propositions which may be of independent interest.

Proposition 4.8. *Let $G \sim \mathcal{G}(n,p)$ for $p = \lambda/n$ with $\lambda \geq 4$ fixed, and let $\mathcal{C}_1$ be the largest component of G. Define the event*

$$\Delta_k = \{\mathcal{C}_1 \text{ contains at least } k \text{ triangles}\}\,. \tag{4.16}$$

For any fixed $k \geq 1$, the function $\mathbb{1}_{\Delta_k}$ is non-degenerate and not Sens.

Proposition 4.9. *Let $G \sim \mathcal{G}(n,p)$ for $p = \lambda/n$ where $\lambda > 1$ is some large enough constant, and let $\mathcal{C}_1$ denote the largest component of G. W.h.p., every pair of triangles in $\mathcal{C}_1$ is connected by a simple path of length $r_n = \lfloor \frac{3}{2}\log_\lambda n\rfloor$.*

Consequently, $\mathbb{P}(H_n \subset G) = \mathbb{P}(\Delta_2) + o(1)$ where Δ_2 is as in (4.16).

Indeed, Proposition 4.8 will follow from showing that the giant component is, in a sense, robust under the noise operator, hence, for instance, triangles in $\mathcal{C}_1$ are likely to remain in the new largest component. The conclusion of Proposition 4.9 that the properties $\{H_n \subset G\}$ and Δ_2 are equivalent up to a negligible probability (together with their non-degeneracy at the given $p = \lambda/n$) will then preclude the noise sensitivity of $\mathbb{1}_{\{H_n \subset G\}}$.

Our proofs will exploit the well-known fact that the breadth-first-search exploration process of the component of a given vertex is well-approximated (up to depth $c\log n$ for a suitable $c(\lambda)$) by a $\mathrm{Po}(\lambda)$-Galton-Watson tree (a supercritical branching process in our setting), whence belonging to the giant component would correspond to the survival of this branching process. Further set $\lambda_\star < 1$ to be the reciprocal of λ in that

$$\lambda e^{-\lambda} = \lambda_\star e^{-\lambda_\star}\,.$$

It is known that $\lambda_\star$ equals the probability that, conditioned on the survival of the branching process, the number of surviving children of the root is 1.

Proof of Proposition 4.8. Let $\{v_1,\ldots,v_n\}$ be the vertices of G arbitrarily ordered, let $V' = \{v_i : i \leq \lceil n/10\rceil\}$ and let G' be the induced subgraph of G on V'. Denoting by Y the number of triangles in G', we note that, as $G' \sim \mathcal{G}(n',p')$ with $p' = \lambda/n \sim \lambda/(10n')$ for $n' = |V'|$, it is well-known (and also follows from the second moment analysis in the proof of Part 1 of Theorem 1.6) that $Y \xrightarrow{\mathrm{d}} \mathrm{Po}(\hat\lambda)$ for some $\hat\lambda > 0$ fixed (namely, $\hat\lambda = \lambda^3/6000$).

Next, write $V'' = \{v_i : i > \lceil n/10\rceil\}$ and for each vertex $x \in V'$ let G''_x be the induced subgraph on $V'' \cup \{x\}$. Further let $\Gamma_t(x)$ denote the exploration process from x in G''_x, that is, for each $t \geq 1$

$$\Gamma_t(x) = \{y \in V'' \,:\, \mathrm{dist}_{G''_x}(x,y) = t\}\,.$$

This breadth-first-search exploration process up to some time R yields a tree $\mathcal{T}_x(R)$ which is stochastically dominated by a $\mathrm{Bin}(0.9n, \lambda/n)$-Galton-Watson tree with R levels (since $|V''| \leq 0.9n$), and as long as the number of exposed vertices is $o(n)$ it stochastically dominates a $\mathrm{Bin}(7n/8, \lambda/n)$-Galton-Watson tree (for instance) with the same number of levels.

Reveal the graph G', and pick an arbitrary vertex from each triangle in it, denoting these vertices by $\{x_1, \ldots, x_Y\}$. Set

$$R := 10 \log_2 \log n \,,$$

and expose $\mathcal{T}_{x_i}(R)$ for all $i = 1, \ldots, Y$ level by level as described above. An important observation is that, should any of these trees intersect, it would imply that G contains a subgraph F_ℓ consisting of two triangles and a path of length $\ell = O(\log \log n)$ between them. However, if $\kappa = \kappa(n)$ is any sequence going to ∞ with n, then w.h.p. no two triangles in G have distance less than $\log_\lambda(n) - \kappa$ between them. Indeed, the expected number of copies of all graphs $\{F_\ell : \ell \leq \log_\lambda(n) - \kappa\}$, where F_ℓ consists of two triangles and a path of length ℓ edges between them, is at most

$$\sum_{\ell \leq \log_\lambda(n) - \kappa} (np)^6 n^{\ell-1} p^\ell \lesssim \sum_{\ell \leq \log_\lambda(n) - \kappa} \frac{\lambda^\ell}{n} \lesssim \lambda^{-\kappa} = o(1) \,.$$

In particular, w.h.p. the Y trees exposed above are pairwise disjoint. In addition, standard large deviation estimates for the binomial distribution (cf., [9, Corollary 2.3]) imply that for any given x

$$\mathbb{P}\left(|\cup_{t \leq R} \Gamma_t(x)| \geq \lambda^R \right) \leq e^{-c(\log n)^2} \,,$$

where $c > 0$ is an absolute constant. (This can be argued, for instance, by noting that for small enough δ, the event $\{| \cup_{t \leq R} \Gamma_t(x)| \geq \lambda^R\}$ implies that for some $t \leq R$ we must have either $\{L_t \geq L_{t-1}\mu + \log^2 n,\, L_{t-1} \leq \log^2 n\}$ or $\{L_t \geq (1+\delta) L_{t-1}\mu,\, L_{t-1} \geq \log^2 n\}$, where $\mu := 7\lambda/8$.) Therefore, w.h.p. no vertex sees more than $\lambda^R = n^{o(1)}$ vertices by time R, and hence we can define on the same probability space $\left(Y, \mathcal{T}_{x_1}(R), \ldots, \mathcal{T}_{x_Y}(R), \mathcal{T}'_1(R), \ldots, \mathcal{T}'_{x_Y}(R)\right)$ so that $\left(\mathcal{T}'_1(R), , \ldots, \mathcal{T}'_{x_Y}(R)\right)$ are i.i.d. $\mathrm{Bin}(7n/8, \lambda/n)$-Galton-Watson trees with R levels and such that $\mathbb{P}(\cap_{i=1}^{Y} \{\mathcal{T}'_i(R) \subset \mathcal{T}_{x_i}(R)\}) = 1 - o(1)$.

Let $\tau_L(d)$ be the probability that a Galton-Watson tree with offspring distribution L contains a d-regular subtree (sharing the same root). This quantity was expressed in [13] as a solution to an equation involving the p.g.f. of L. When $L \sim \mathrm{Po}(\mu)$, it was shown that $\tau_L(d)$ is the largest solution of $(1-s)\exp(\mu s) = \sum_{j=0}^{d-1} (\mu s)^j / j!$, which is positive whenever $d = (1-\varepsilon_\mu)\mu$ for some $\varepsilon_\mu \to 0$ as $\mu \to \infty$ (see §4 of that work). For $d = 2$, the analysis of [13] (and Eqs. (4.3),(4.4) in particular) shows that $\tau_L > 0$ provided $\mu > \exp(y)/y$ where y is the unique positive solution to $y^2 + y + 1 = \exp(y)$, e.g., $\mu > 3.351$ would suffice for a positive probability of containing a binary subtree. In case of $L \sim \mathrm{Bin}(n,p)$ (explicitly stated in [12, §5]), $\tau_L(d)$ is the largest solution $s \in (0,1]$ of $1 - s = \mathbb{P}(\mathrm{Bin}(n, ps) \leq d-1)$. For $p = \mu/n$, since $L \xrightarrow{\mathrm{d}} \mathrm{Po}(\mu s)$

and the intersection of the functions $(1-s)$ and $\exp(-\mu s)(1+\mu s)$ is not a tangent point for any μ larger than the critical one, $\tau_L(d)$ coincides with the Poisson case, thus in our setting indeed $\mu = 7\lambda/8 \geq 3.5$ (by the assumption on λ) suffices for the tree $\mathcal{T}'_i(R)$ to contain a binary subtree of height R at its root with positive probability; let $\theta > 0$ denote this probability.

Altogether, it follows that we can define on a common probability space our random graph and a $\mathrm{Po}(\lambda'\theta)$ variable Z so that w.h.p. the number of triangles in G', for which the exploration process into V'' from one of the endpoints contains a binary subtree of height R rooted at that vertex, is at least Z. Hence, for any fixed $k \geq 1$ there will be at least k such triangles with positive probability (here we see that Δ_k is non-degenerate: with positive probability G is triangle-free, and with positive probability we find k triangles as above, each one connected to at least $2^{\lfloor R \rfloor} \asymp (\log n)^{10}$ vertices and thus part of $\mathcal{C}_1$ w.h.p. (see, e.g., [9, Theorem 5.4])).

The proof is concluded by noticing that each of these triangles is robust under the noise operator. Indeed, the triangle itself survives the noise with probability $(1-\varepsilon)^3$, and henceforth the noise operator on a binary tree is simply a branching process with offspring distribution $\mathrm{Bin}(2, 1-\varepsilon)$. Letting Z_t be its population size at time t, a classical fact on supercritical branching processes whose offspring distribution L has a finite second moment is that, if $m = \mathbb{E}L > 1$ and $q < 1$ is the extinction probability, for any fixed $\delta > 0$ with probability $1 - q - \delta$ we have that $|Z_R| \geq cm^R$ for some fixed $c > 0$. Here we have $m = 2(1-\varepsilon)$, yielding that $|Z_R| \geq c(\log n)^2$ for a small enough ε, except with probability $q + \delta \leq 2q$ (for a suitable δ) where q goes to 0 with ε. This would in turn correspond to the scenario where w.h.p. the triangle under consideration is part of $\mathcal{C}_1^\varepsilon$, the largest component of the new graph (as the second largest component has $O_{\mathrm{P}}(\log n)$ vertices). Altogether, we have shown that for $f_n = \mathbb{1}_{\Delta_k}$, a positive fraction of the space $\{\omega : f_n(\omega) = 1\}$ is such that $\mathbb{P}(f_n(\omega^\varepsilon) = 1 \mid \omega) \geq 1 - g(\varepsilon)$ where $g(\varepsilon) \to 0$ as $\varepsilon \to 0$. By Proposition 2.5 it then follows that (f_n) is not noise sensitive. ■

It remains to prove Proposition 4.9. While it is possible to derive the proof from various routine branching process estimates, it will be convenient to appeal to estimates to this effect that were developed specifically for the setting of a sparse random graph $\mathcal{G}(n, \lambda/n)$ in the recent work of Riordan and Wormald [14]. Similar to before, let $\Gamma_t(x) := \{v \in V(G) : \mathrm{dist}_G(x, v) = t\}$ for $t \geq 0$ be the set of all vertices of G at distance exactly t from x. Set

$$w := (\log n)^6 \,, \qquad t_0 = \left\lfloor \log_{\lambda_\star^{-1}} n \right\rfloor \,, \qquad t_1 := \lfloor \log_\lambda w \rfloor \,,$$

following the notation of [14]. Using these definitions, the following was shown in [14, Lemmas 2.1 and 2.2] (see Eqs. (2.10) and (2.11) in particular).

Lemma 4.10 ([14]). *Let $0 < \kappa = o(\log n)$ be so that $\kappa \to \infty$ with n. Then w.h.p. no vertex $x \in V$ satisfies $1 \leq |\Gamma_t(x)| < w$ for all $0 \leq t \leq t_0 + t_1 + \kappa$.*

Observe that $t_1 = O(\log\log n)$ whereas $t_0 = (1+\delta_\lambda)\lambda^{-1}\log n$ for δ_λ which approaches 0 as λ grows. In particular, we have

$$t_0 + t_1 + \kappa \leq \tfrac{1}{10}\log_\lambda n$$

for large enough λ and any sufficiently large n. Therefore, upon defining

$$\tau_w(x) := \min\{t : |\Gamma_t(x)| \geq w\}\,,$$

we see that w.h.p. every vertex x satisfies that $x \in \mathcal{C}_1$ iff $\tau_w(x) \in [1, \frac{1}{10}\log_\lambda n]$. We can now address the case $\tau_w(x) \leq \frac{1}{10}\log_\lambda n$, which will correspond as per the discussion above to every x belonging to the giant component. Here we will need to adapt this conclusion to the case of two simultaneously growing neighborhoods, as given by the next lemma.

Lemma 4.11. *Fix $\delta > 0$ and take $\ell \in \mathbb{N}$ such that $\ell/\log_\lambda n \in (1+3\delta, 2-2\delta)$. Then w.h.p. every two vertices x, y whose distance in G exceeds $2\delta\log_\lambda n$ and such that $\tau_w(x), \tau_w(y) \leq \delta\log_\lambda n$ are connected by a simple path of length ℓ.*

Proof. Set $T = \delta\log_\lambda n$ and consider the standard exploration process which iteratively reveals $\Gamma_t(x)$ for $1 \leq t \leq T$. Estimating $|\Gamma_t(x)|$ is elementary by standard concentration arguments, as noted in [14, Lemma 2.4]. Indeed, denoting $L_t = |\Gamma_t(x)|$ for the number of vertices at distance t from x, clearly $L_{t+1} \sim \mathrm{Bin}(n - \sum_{i\leq t} L_i, q)$ for $q = 1 - (1-\lambda/n)^{L_t} = \lambda L_t/n + O(L_t^2/n^2)$. It then follows from large deviation estimates of the binomial variable (as used in the proof of Proposition 4.8) that as long as, e.g., $\sum_{i\leq t} L_i \leq n^{1-\delta/2}$,

$$\mathbb{P}\left(\Big|\frac{L_{t+1}}{\lambda L_t} - 1\Big| \geq \frac{1}{\log^2 n} \;\Big|\; L_t\right) \leq 2\exp\left(-(\tfrac13 - o(1))\frac{\lambda L_t}{\log^4 n}\right)\,,$$

where the assumption on L_t makes $\mathbb{E}[L_{t+1} \mid L_t] = (1 + O(n^{-\delta/2}))\lambda L_t$, an approximation error which is insignificant compared to the $O(1/\log^2 n)$ scale of the deviation considered here. In particular, we see that necessarily

$$w \leq L_{\tau_w(x)} \leq 2\lambda w$$

except with probability $\exp(-cw/\log^4 n) = \exp(-c\log^2 n)$ for an absolute constant $c > 0$. Furthermore, by accumulating the $O(1/\log^2 n)$ errors up to time $T = O(\log n)$, this estimate can be extended all throughout this interval (note that since $T = \delta\log_\lambda n$ this will maintain $L_t \leq n^\delta$ satisfying the requirement on the size of $\sum_{i\leq t}|\Gamma_i(x)|$ with room to spare) to yield

$$\Big|L_t/\big[\lambda^{t-\tau_w(x)}L_{\tau_w(x)}\big] - 1\Big| \leq \tfrac{\log\log n}{\log n} \qquad \text{for all } \tau_w \leq t \leq T$$

except with probability $\exp(-c\log^2 n)$ for some other absolute $c>0$ (the factor of $\log\log n$ could have been replaced by any $\kappa(n)$ going to ∞ with n).

Now, let us adapt the exploration process to a pair of initial points x, y as follows. Denoting the set of neighbors of a set S in G by $N_G(S)$, let

$$\Gamma_0' = \{x\}\,,\quad \Gamma_t' = N_G\left(\Gamma_{t-1}'\right) \setminus \bigcup_{i<t}\left(\Gamma_i' \cup \Gamma_i''\right)\,,$$

$$\Gamma_0'' = \{y\}\,,\quad \Gamma_t'' = N_G\left(\Gamma_{t-1}''\right) \setminus \left(\Gamma_t' \cup \bigcup_{i<t}\left(\Gamma_i' \cup \Gamma_i''\right)\right).$$

That is, we expand the neighborhood of x among unvisited vertices (those that had not yet appeared in any of the neighborhoods) followed by the same procedure for y, repeatedly.

We clearly have that $\cup_{t\le T}\Gamma_t'$ and $\cup_{t\le T}\Gamma_t''$ are disjoint by construction. The hypothesis on the distance of x, y then implies that $\Gamma_t' = \Gamma_t(x)$ and $\Gamma_t'' = \Gamma_t(y)$ for all $t \le T$. It now follows that $\sum_{t\le T}(|\Gamma_t'| + |\Gamma_t''|) \le 5\lambda w n^\delta$ with probability $1-\exp(-c\log^2 n)$ for some absolute $c>0$.

Exposing Λ_t' for $t = T+1, \ldots, \lceil \ell/2\rceil$ alternating with exposing Λ_t'' for $t = T+1, \ldots, \lfloor \ell/2\rfloor$, the exact same concentration argument as above — while recalling that $\ell < (2-2\delta)\log_\lambda n$ by hypothesis and so at all times above there are at least $(1-O(n^{-\delta}))n$ unexposed vertices — implies that with probability $1-\exp(-c\log^2 n)$ for some absolute $c>0$ we have

$$\left||\Gamma_t'|/\left(\lambda^{t-T}|\Gamma_T'|\right) - 1\right| \le \tfrac{\log\log n}{\log n} \qquad \text{for all } T \le t \le \lceil \ell/2\rceil\,,$$
$$\left||\Gamma_t''|/\left(\lambda^{t-T}|\Gamma_T''|\right) - 1\right| \le \tfrac{\log\log n}{\log n} \qquad \text{for all } T \le t \le \lfloor \ell/2\rfloor\,.$$

Combining this with the fact that $|\Gamma_T'|, |\Gamma_T''| \ge w$ along with the hypothesis $\ell > (1+3\delta)\log_\lambda n$ now yields that with the aforementioned probability,

$$|\Gamma_{\lceil \ell/2\rceil}'| \ge n^{(1+\delta)/2} \qquad \text{and} \qquad |\Gamma_{\lfloor \ell/2\rfloor}''| \ge n^{(1+\delta)/2}\,.$$

Finally, observe that none of the potential edges between $\Gamma_{\lceil \ell/2\rceil}'$ and $\Gamma_{\lfloor \ell/2\rfloor}''$ has been examined yet, and the probability that none belong to G is at most

$$(1-\lambda/n)^{|\Gamma_{\lceil \ell/2\rceil}'||\Gamma_{\lfloor \ell/2\rfloor}''|} \le \exp\left(-\lambda n^\delta\right)\,.$$

As any such edge yields a simple path of length ℓ between x, y, the proof of the lemma is concluded by a union bound over x, y, easily accommodated by the fact that all error probabilities were super-polynomially small in n. ■

With the above ingredients, we can establish Proposition 4.9 guaranteeing length-specific paths between triangles in the giant component $\mathcal{C}_1$.

Proof of Proposition 4.9. Since $\mathcal{C}_1$ is of linear size w.h.p., and thanks to Lemma 4.10 and the discussion following it, w.h.p. every vertex $x \in \mathcal{C}_1$ satisfies $\tau_w(x) < \frac{1}{10}\log_\lambda n$. Choosing $\delta = \frac{1}{10}$ and $\ell = r_n$ in Lemma 4.11 we obtain that w.h.p. every two vertices $x, y \in \mathcal{C}_1$ with $\mathrm{dist}_G(x,y) > \frac{1}{5}\log_\lambda n$ have a simple path connecting them of distance precisely $r_n = \lfloor \frac{3}{2}\log_\lambda n \rfloor$.

The first statement of the proposition now follows from the fact noted in the proof of Proposition 4.8 that for any $\kappa = \kappa(n)$ going to ∞ with n, w.h.p. no two triangles in G have distance less than $\log_\lambda(n) - \kappa$ between them. In particular, w.h.p. every pair of triangles in $\mathcal{C}_1$ has distance at least $\frac{1}{2}\log_\lambda n$, and thus are connected by a path of length r_n, as argued above.

Finally, it is well-known (see, e.g., [9, Theorem 5.12]) that w.h.p. $\mathcal{C}_1$ is the only component that contains more than a single cycle, and therefore $\mathbb{P}(H_n \subset G) = \mathbb{P}(H_n \subset \mathcal{C}_1) + o(1) \le \mathbb{P}(\Delta_2) + o(1)$. As we have shown above that $\mathbb{P}(\Delta_2) \le \mathbb{P}(H_n \subset \mathcal{C}_1) + o(1)$, this completes the proof. ∎

Propositions 4.8 and 4.9 combined conclude the proof of Theorem 1.6. ∎

5. General properties of strong noise sensitivity

5.1. 0-strong vs 1-strong noise sensitivity.

The following proposition gives a simple and yet useful necessary condition for $\textsc{StrSens}_1$.

Lemma 5.1. *Let (f_n) be a sequence of monotone Boolean functions, and let $Y_n(\omega) = \sum_{W \in \mathcal{W}_0(f_n)} \mathbb{1}_{\{\omega_W \equiv 0\}}$ count the occurring 0-witnesses in $\omega \in \Omega_n$. If $\sup_n \mathbb{E}[Y_n] < \infty$ then the sequence is not $\textsc{StrSens}_1$.*

Proof. Clearly if $W \in \mathcal{W}_1$ and $W' \in \mathcal{W}_0$, we must have $W \cap W' \neq \emptyset$, whence

$$\mathbb{P}\left(\omega^\varepsilon_{W'} \equiv 0 \mid \omega_W \equiv 1\right) \le \varepsilon\, \mathbb{P}\left(\omega^\varepsilon_{W'} \equiv 0\right)\,,$$

and so, by our main assumption, there exists some $C > 0$ such that for all n

$$\sup_{W \in \mathcal{W}_1} \mathbb{E}\left[Y_n(\omega^\varepsilon) \mid \omega_W \equiv 1\right] \le C\varepsilon\,.$$

It follows that

$$\inf_{W \in \mathcal{W}_1} \mathbb{P}\left(f_n(\omega^\varepsilon) = 1 \mid \omega_W \equiv 1\right) \ge 1 - O(\varepsilon)\,,$$

thus the sequence is not $\textsc{StrSens}_1$ (instead, the conditional probability given any 1-witness is in some sense noise stable, going to 1 as $\varepsilon \to 0$). ∎

Remark. The converse of Lemma 5.1 is false, as the recursive 3-majority function demonstrates. We have shown in §3.1 that this function is not $\textsc{StrSens}_1$, and yet it is easy to see that $\mathbb{E}[Y_n]$ is not uniformly bounded (nor is the expected number of 1-witnesses, by symmetry). Indeed, if a_k denotes the number of 0-witnesses when there are $n = 3^k$ variables, then $a_0 = 1$ and $a_{k+1} = 3a_k^2$, and so in general $a_k = 3^{2^k - 1}$. Since a canonical witness has size 2^k, we have $\mathbb{E}Y_n = \frac{1}{3}(3/2)^{2^k} \to \infty$.

Many of the examples that we have seen are $\textsc{StrSens}_1$ but not $\textsc{StrSens}_0$ or vice versa. We next show that there are Boolean functions which are both.

Theorem 5.2. *There exists a sequence of monotone non-degenerate Boolean functions which are both* $\mathrm{StrSens}_1$ *and* $\mathrm{StrSens}_0$*.*

Proof. Define the following Boolean functions:

- g_n: the tribes function on n bits with $\lfloor \log_2(\frac{n}{\log_2 n}) \rfloor$-bit blocks (as usual, potentially ignoring one shorter block to remedy divisibility issues).
- h_n: the tribes function on $m_n := \lfloor n^{\log n} \rfloor$ bits with $b_n := \lfloor \log_2(\frac{m_n}{\log_2 m_n}) \rfloor$ bits per block and reversed 0/1 roles ($h_n = 0$ iff there is an all-0 block).
- $f_n = g_n \circ h_n$ is the composition of these functions acting on $m_n n$ bits (applying h_n to the first m_n bits, the next m_n bits, etc., then feeding the n output bits into g_n), which we claim is both $\mathrm{StrSens}_1$ and $\mathrm{StrSens}_0$.

Let p_n be such that $\mathbb{P}(h_n = 1) = 1/2$ (it is easy to see that $p_n = 1/2 + o(1)$). The proof will follow from two straightforward properties of h_n.

First, we claim that for any $\varepsilon > 0$ there exists $\delta > 0$ so that

$$\inf_n \inf_{W \in \mathcal{W}_1(h_n)} \mathbb{P}\left(h_n(\omega^\varepsilon) = 0 \mid \omega_W \equiv 1\right) \geq \delta\,. \tag{5.1}$$

Indeed, the number of 0-witnesses occurring in ω^ε given $\omega_W \equiv 1$ is binomial with parameters $\mathrm{Bin}((1+o(1))\frac{m_n}{\log_2 m_n}, \varepsilon p^{b_n})$. Since $p^{b_n} \asymp \frac{\log_2 m_n}{m_n}$, for fixed ε this converges to a nontrivial Poisson distribution, from which (5.1) follows.

Second, we argue that for any $\varepsilon > 0$ we have

$$\max_{W \in \mathcal{W}_0(h_n)} \mathbb{P}\left(h_n(\omega^\varepsilon) = 0 \mid \omega_W \equiv 0\right) - \mathbb{P}\left(h_n = 0\right) = o\left(1/n\right)\,. \tag{5.2}$$

To see this, note that since the 0-witnesses for h_n are disjoint, the only gain from conditioning on the event $\omega_W \equiv 0$ for some 0-witness W is that the probability that $\omega^\varepsilon_W \equiv 0$ is increased. Therefore, it suffices to show that $\mathbb{P}\left(\omega^\varepsilon_W \equiv 0 \mid \omega_W \equiv 0\right) = o\left(1/n\right)$ uniformly over W. Indeed this holds as $\mathbb{P}\left(\omega^\varepsilon_W \equiv 0 \mid \omega_W \equiv 0\right) = (1 - \varepsilon p_n)^{b_n}$ with $p_n \sim 1/2$ and $b_n \gtrsim \log m_n \gtrsim \log^2 n$, thus establishing (5.2) (with room to spare).

To show that (f_n) is $\mathrm{StrSens}_1$, fix $\varepsilon > 0$ and note that a 1-witness W for f_n is obtained by taking a 1-witness W' for g_n and for each $x \in W'$ taking a 1-witness W''_x for h_n. By (5.1), $\mathbb{P}(\omega^\varepsilon_x = 0 \mid \omega_{W''_x} = 1) \geq \delta$ for any $x \in W'$ with $\delta(\varepsilon) > 0$ fixed. Thus, $\mathbb{P}(\omega_{W'} \equiv 1) \leq (1-\delta)^{|W'|} \to 0$, and since the rest of the blocks of g_n are independent we get (following the same argument used to show (5.2) above) that (f_n) is $\mathrm{StrSens}_1$.

It remains to show that (f_n) is $\mathrm{StrSens}_0$. Fix $\varepsilon > 0$ and again take a 0-witness W for f_n in the form of a 0-witness W' for g_n and accompanying each $x \in W'$ by a 0-witness W''_x for h_n. If $\omega_W \equiv 0$, then (5.2) and the fact that $|W'| \asymp \frac{n}{\log n}$ tell us that $\omega^\varepsilon_{W'}$ has a distribution whose total variation distance from an i.i.d. sequence with parameter $1/2$ goes to 0. With the other blocks of g_n independent, as before this implies that (f_n) is $\mathrm{StrSens}_0$. ∎

5.2. **Different levels of noise in strong noise sensitivity.** An interesting fact about noise sensitivity, pointed out in §2, is that if the criterion (1.1) for Sens holds for one fixed $\varepsilon \in (0,1)$, then it holds for all such ε. It is then natural to ask whether strong noise sensitivity also exhibits this behavior. Clearly, if the criterion (1.2) for $\textsc{StrSens}_1$ holds for one $\varepsilon \in (0,1)$ then it holds for all $\varepsilon' > \varepsilon$ by monotonicity. However, the next theorem tells us that in fact (1.2) may hold for some $\varepsilon \in (0,1)$ and not for some other $\varepsilon' \in (0,\varepsilon)$.

Theorem 5.3. *There exists a sequence of monotone Boolean functions* (f_n) *which is* $\textsc{StrSens}_1$ *w.r.t. any fixed* $\frac14 < \varepsilon < 1$*, while for any fixed* $0 < \varepsilon < \frac15$

$$\lim_{n\to\infty} \inf_{W \in \mathcal{W}_1(f_n)} \mathbb{P}(f_n(\omega^\varepsilon) = 1 \mid \omega_W \equiv 1) = 1\,.$$

Proof. Define the following Boolean functions:

- r_n: recursive 5-majority on $5^{\lfloor 1.01\, b_n \rfloor}$ variables where $b_n := \lfloor \log_2(\frac{n}{\log_2 n})\rfloor$.
- g_n: the tribes function on n bits with b_n-bit blocks.
- $f_n = r_n \circ g_n$ is the composition of these two functions, acting on $n5^{\lfloor 1.01\, b_n\rfloor}$ bits, which we claim will have the desired properties.

Choose p_n such $\mathbb{P}(g_n = 1) = 1/2$ (recall that this choice has $p_n = 1/2 + o(1)$). In Claim 3.3 we related the probability that a witness for r_n survives the noise to the k-iterated function $h(x)$ from that claim, denoted here $h^{(k)}(x)$. The next claim establishes two simple features of that function.

Lemma 5.4. *Let* $h(x) := -\frac12 x^3 + \frac34 x^2 + \frac34 x$ *as in* (3.4)*. Then we have* $h^{(1.01\, m)}\left(\frac12 + (0.88)^m\right) = \frac12 + o(1)$ *whereas* $h^{(1.01\, m)}\left(\frac12 + (0.89)^m\right) = 1 - o(1)$.

Proof. Letting L be the linear function $L(x) := \frac98(x - \frac12) + \frac12$, we have $h \le L$ on $[\frac12, 1]$ since h is concave in that interval and has $h(\frac12) = \frac12$ and $h'(\frac12) = \frac98$. Since h is increasing and sends $[\frac12, 1]$ to itself, it follows that $h^{(k)} \le L^{(k)}$ on $[\frac12, 1]$ for all k. Observing that $L^{(k)}(x) = (\frac98)^k (x - \frac12) + \frac12$, in particular we have $h^{(1.01\, m)}\left(\frac12 + (0.88)^m\right) - \frac12 \le (\frac98)^{1.01\, m}(0.88)^m \to 0$ as $m \to \infty$.

For the second statement, choose $p_0 \in (\frac12, 1)$ so that $h'(p_0) = \frac98 - \frac{1}{1000}$. Since h is concave on $[\frac12, 1]$, now $h \ge M$ on $[\frac12, p_0]$ where M is the linear function $M(x) := h'(p_0)(x - \frac12) + \frac12$. Since h is increasing and sends $[\frac12, 1]$ to itself, $h^{(k)}(x) \ge M^{(k)}(x)$ for all x and k satisfying $M^{(k-1)}(x) \le p_0$ (i.e., until the orbit of x passes p_0). Since $M^{(m)}(x) = (h'(p_0))^m (x - \frac12) + \frac12$, we have $M^{(m)}(\frac12 + (0.89)^m) \to \infty$, and so $h^{(m)}(\frac12 + (0.89)^m) \ge p_0$ for large m. Since p_0 is a fixed number larger than $1/2$, and $h(x)$ has fixed points at $\{0, 1/2, 1\}$, the additional $m/100$ iterations give $h^{(1.01\, m)}(x) = 1 - o(1)$, as required. ∎

As for the tribes function g_n, it is easy to check that for any 1-witness W,

$$\Gamma_n := \mathbb{P}(g_n(\omega^\varepsilon) = 1 \mid \omega_W \equiv 1) - \mathbb{P}(g_n = 1) = u_n \left[(1 - \varepsilon(1 - p_n))^{b_n} - p_n^{b_n}\right]$$

where u_n is the probability that none of the blocks except possibly the first one is an all 1-block, which is $1/2 + o(1)$. As $p_n = 1/2 + o(1)$, it follows, say, that for any fixed $0 < \varepsilon < 1$, any sufficiently large n and any 1-witness W,

$$\left(1 - \varepsilon/2 - \varepsilon^2/16\right)^{b_n} \leq \Gamma_n \leq \left(1 - \varepsilon/2 + \varepsilon^2/16\right)^{b_n} . \tag{5.3}$$

Any 1-witness W for f_n is obtained by taking some 1-witness W' for r_n together with a 1-witness W''_x for g_n for every $x \in W'$. By (5.3), for large enough n the distribution of the bits $\omega^{\varepsilon}_{W'}$ is i.i.d. with probability q_n of 1, where $q_n \leq 1/2 + (0.88)^{b_n}$ if $\varepsilon > \frac{1}{4}$, whereas $q_n \geq \frac{1}{2} + (0.89)^{b_n}$ if $\varepsilon < \frac{1}{5}$.

Finally, the analysis in Claim 3.3 tells us that for recursive 5-majority with k levels on an input distribution that is i.i.d. $(q, 1-q)$ for $q \neq 1/2$ on a 1-witness W' and i.i.d. $(1/2, 1/2)$ elsewhere, the probability that the output is 1 is $h^{(k)}(q)$. This fact together with Lemma 5.4 concludes the proof. ∎

Acknowledgements

This work was carried out when the second author was visiting Microsoft Research at Redmond, and he thanks the Theory Group for its hospitality and for creating a stimulating research environment. The second author also acknowledges the support of the Swedish Research Council and the Knut and Alice Wallenberg Foundation. We thank the anonymous referees for useful comments.

References

[1] N. Alon and J. H. Spencer, *The probabilistic method*, 3rd ed., John Wiley & Sons Inc., Hoboken, NJ, 2008. With an appendix on the life and work of Paul Erdős.

[2] R. Arratia, L. Goldstein, and L. Gordon, *Two moments suffice for Poisson approximations: the Chen-Stein method*, Ann. Probab. **17** (1989), no. 1, 9–25.

[3] I. Benjamini, S. Haber, M. Krivelevich, and E. Lubetzky, *The isoperimetric constant of the random graph process*, Random Structures Algorithms **32** (2008), no. 1, 101–114.

[4] I. Benjamini, G. Kalai, and O. Schramm, *Noise sensitivity of Boolean functions and applications to percolation*, Inst. Hautes Études Sci. Publ. Math. **90** (1999), 5–43.

[5] M. Ben-Or and N. Linial, *Collective coin flipping*, Randomness and Computation (S. Micali, ed.), Academic Press, 1990, pp. 91–115. Earlier version in FOCS 1985.

[6] B. Bollobás, *Random graphs*, 2nd ed., Cambridge Studies in Advanced Mathematics, vol. 73, Cambridge University Press, Cambridge, 2001.

[7] E. Friedgut, *Sharp thresholds of graph properties, and the k-sat problem*, J. Amer. Math. Soc. **12** (1999), no. 4, 1017–1054. With an appendix by Jean Bourgain.

[8] C. Garban and J. E. Steif, *Lectures on noise sensitivity and percolation*, Clay Math. Institute Summer School and the XIV Brazilian School of Probability (Buzios, Brazil, 2010), Probability and Statistical Physics in Two and More Dimensions (D. Ellwood, C. Newman, V. Sidoravicius, and W. Werner, eds.), Clay Math. Proceedings, vol. 15, American Mathematical Society, 2012, pp. 49–154.

[9] S. Janson, T. Łuczak, and A. Rucinski, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000.

[10] J. Kahn, G. Kalai, and N. Linial, *The influence of variables on Boolean functions*, Proc. of the 29th IEEE Symp. on Foundations of Computer Science (FOCS 1988), pp. 68–80.

[11] N. Keller and G. Kindler, *Quantitative relation between noise sensitivity and influences*, Combinatorica **33** (2013), no. 1, 45–71.

[12] R. Lyons with Y. Peres, *Probability on Trees and Networks*, Cambridge University Press, 2013. In preparation. Current version is available at http://mypage.iu.edu/∼rdlyons/prbtree/book.pdf.

[13] A. G. Pakes and F. M. Dekking, *On family trees and subtrees of simple branching processes*, J. Theoret. Probab. **4** (1991), no. 2, 353–369.

[14] O. Riordan and N. Wormald, *The diameter of sparse random graphs*, Combin. Probab. Comput. **19** (2010), no. 5-6, 835–926.

[15] O. Schramm and J. E. Steif, *Quantitative noise sensitivity and exceptional times for percolation*, Ann. of Math. (2) **171** (2010), no. 2, 619–672.

Eyal Lubetzky
Microsoft Research, One Microsoft Way, Redmond, WA 98052-6399, USA.
E-mail address: eyal@microsoft.com

Jeffrey E. Steif
Mathematical Sciences, Chalmers University of Technology, and
Mathematical Sciences, Göteborg University, SE-41296 Gothenburg, Sweden.
E-mail address: steif@math.chalmers.se