Proceedings of the International Workshop on Statistical-Mechanical Informatics
March 7–10, 2010, Kyoto, Japan

On the robustness of random Boolean formulae

Alexander Mozeika[1], **David Saad**[1], **Jack Raymond**[2]

[1] The Non-linearity and Complexity Research Group, Aston University, Birmingham B4 7ET, UK

[2] Department of Physics, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China

E-mail: {a.s.mozeika,d.saad}@aston.ac.uk

Abstract. Random Boolean formulae, generated by a growth process of noisy logical gates are analyzed using the generating functional methodology of statistical physics. We study the type of functions generated for different input distributions, their robustness for a given level of gate error and its dependence on the formulae depth and complexity and the gates used. Bounds on their performance, derived in the information theory literature for specific gates, are straightforwardly retrieved, generalized and identified as the corresponding typical-case phase transitions. Results for error-rates, function-depth and sensitivity of the generated functions are obtained for various gate-type and noise models.

1. Introduction

Boolean functions are typically represented by formulae based on simple Boolean operations and implemented using basic logical gates. They form the basis of most computing devices and are the subject of extensive and varied research efforts in the theoretical computer science and information theory communities. The main sources of errors in classical computing circuits based on semiconductor technology are heat generation, cosmic rays and production defects [1], the impact of which is exacerbated by the ever-increasing drive towards miniaturization and more complex devices of large scale [1].

von Neumann [2] was the first to address the problem of computation with noisy gates and its limitations; and suggested simple gate-constructions, based only on noisy gates to limit the resulting function error. He showed that a bounded function error can be implemented in such cases for a sufficiently small gate-error. His motivation for investigating the problem had arisen from the study of biologically-motivated circuits and an attempt to explain the robustness of neural activities represented by a circuit (or formula) composed of ϵ-noisy Boolean gates.

Before progressing any further, a few formal definitions are required: (i) A *circuit* may be regarded as a directed acyclic graph in which the nodes of in-degree zero are either Boolean constants or references to arguments, the nodes of in-degree $k \geq 1$ are logical gates of k arguments and the nodes of out-degree zero correspond to the circuit outputs. (ii) A *formula* is a single-output circuit where the output of each gate is used as an input to at most one gate. (iii) The ϵ-noisy gate is designed to compute a Boolean function $\alpha : \{-1,1\}^k \to \{-1,1\}$, but for each input $\boldsymbol{S} \in \{-1,1\}^k$ there is an error probability ϵ such that $\alpha(\boldsymbol{S}) \to -\alpha(\boldsymbol{S})$. To simplify the analysis, error-probability is taken to be independent for each gate in the circuit. Clearly, a noisy circuit ($\epsilon > 0$) cannot perform any given computation in a deterministic manner: for any circuit-input there is a non-vanishing probability that the circuit will produces the wrong output. (iv) The maximum of this error probability δ over all circuit-inputs determines *reliability* of the circuit.

von Neumans's study was later revisited by Pippenger [3] using information theory methods. He showed that if a noisy k-ary formula is used to compute a Boolean function f with the error probability $\delta < 1/2$, then (i) there is an upper bound for the gate-error $\epsilon(k)$ which is strictly less than 1/2 and (ii) there is a lower bound for the formula-depth $\hat{d}(k, \epsilon, \delta) \geq d$, where d is the depth of a noiseless formula computing f; the *depth* of a formula being the number of gates on the longest path from an input node to the output node. In comparison to its noiseless counterpart, a noisy formula that computes reliably has greater depth due to the presence of restitution-gates, implying longer computation times [3]. These results have been refined and extended for circuits [4], for k-ary Boolean formulae [5, 6] and different gates [7, 8].

The aim of this study is to investigate the properties of noisy circuits via methods of statistical physics, identify physical properties that characterize existing bounds and convergence rates, and provide insight and quantitative results beyond what is accessible via the information theory methodology. Key aspect of our analysis is the study of *typical* rather than *worst* case bounds.

The paper is organized as follows: In section 2 we discuss the generation of typical Boolean functions and the model used for generating them followed by a description of the model used in section 3 and main thrust of the derivation in section 4. The results obtained are summarized in section 5 followed by a conclusion and discussion section 6.

2. Generating random functions

As the aim of this work is to study the properties of typical cases, one should first identify a mechanism for randomly generating *typical* functions. This is not a simple task as most of the commonly known paradigms in the theoretical computer science literature aim to prove that canonical representations can describe any given arbitrary function using elementary gates or processes, but when applied at random they tend generate trivial functions showing weak dependence on the input variables.

A good example is the use of covering or bi-decomposition to represent arbitrary functions [9] resulting in a disjunctive normal form (DNF). The DNF or its dual CNF (conjunctive normal form) is a depth-2 formula with AND and OR gates used as internal nodes and with the input Boolean variables and their negations distributed on the leaves. However, *random* DNF (CNF) formulae offer very low sensitivity [10] to the input values and highly simple and uncharacteristic Boolean functions.

To generate formulae that represent all Boolean functions with uniform probability, using *randomly generated* circuits, we employ a variant of the growth process suggested by Savický [11], based on the majority gate that, under very broad conditions, produces random functions as the depth of the formulae becomes large. Firstly, one defines an initial distribution over a set of simple Boolean functions. Secondly, and in further steps, the formulae chosen from the distributions defined in previous steps are combined by Boolean gates. One such process, described by Savický [11], uses only a single Boolean gate α and is defined by the recursion on the set of formulae A_ℓ:

$$\begin{aligned} &A_0 = \{1, -1, S_1, \ldots, S_n, -S_1, \ldots, -S_n, \} \\ &A_{\ell+1} = \{\alpha(\phi_1, \ldots, \phi_k); \phi_j \in A_\ell \text{ for } j = 1, 2, \ldots, k\}. \end{aligned} \tag{1}$$

Savický showed, under a very broad conditions on α, that the probability of computing a Boolean function by a formula $\phi \in A_\ell$ tends to the uniform distribution over all Boolean functions of n variables when $\ell \to \infty$ [11]. Furthermore, depending on the initial conditions A_0 and the gate α the process converges to a *single Boolean function* or to the *uniform distribution* over some class of Boolean functions [12].

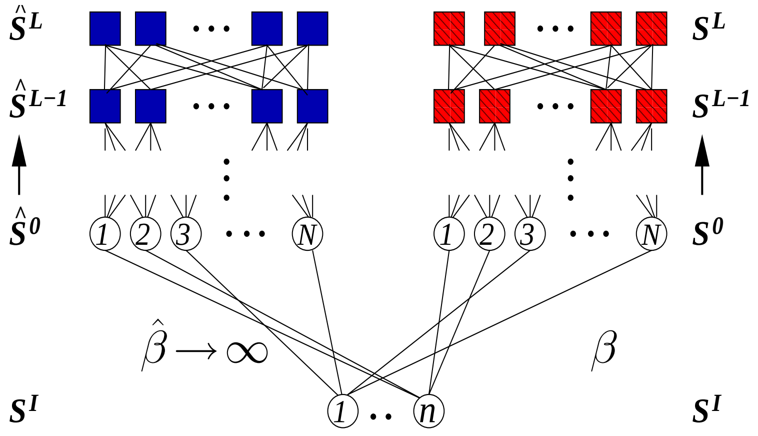

Figure 1. The model of two coupled systems with identical topology and different inverse temperatures β and $\hat{\beta}\rightarrow\infty$. Gates are indicated by squares, S^I and input nodes by circles. Blue indicates noiseless gates, red noisy gates.

In this framework, all Boolean functions of n variables are represented with equal statistical weight when $\ell \rightarrow \infty$, but the number of gates in formulae grows exponentially with the formula depth ℓ. To contain the number of gates, we propose a layered variant of the Savický growth process. The first step in our process is to sample randomly and uniformly exactly N entries of an input vector $\hat{\boldsymbol{S}}^0 = (S_1^0, \ldots, S_N^0)$. In the second, and all subsequent steps for $\ell = 1, \ldots, L-1$, we construct a vector $\hat{\boldsymbol{S}}^{\ell+1} = (S_1^{\ell+1}, \ldots, S_N^{\ell+1})$ where the i-th entry $S_i^{\ell+1}$ is an output of the gate $\alpha(S_{i_1}^{\ell}, \ldots, S_{i_k}^{\ell})$ with k input-indices sampled uniformly from the set of all possible (unordered) indices $\{i_1, \ldots, i_k\}$. The result of the process is the layered $N \times (L+1)$ Boolean circuit shown in Figure 1 (left construction, in blue). For large N, the variable S_i^{ℓ} in our model corresponds to the output of a random k-ary of depth ℓ, which computes a Boolean function $\{-1,1\}^N \rightarrow \{-1,1\}$. We expect that in the limit $N \rightarrow \infty$, with $\ell \in O(N^0)$, the statistical properties of the formulae generated in our process and in the Savický's growth process are equivalent; this is supported by the results reported later. The layered representation allows us to explore the typical behavior of noisy random Boolean formulae using methods of statistical physics.

While the vector $\hat{\boldsymbol{S}}^0$ represents randomly sampled single entries, one would also like to study cases where entries are statistically dependent and are sampled from a smaller set. To cater for a possible higher level of correlation, the 0-layer boundary conditions are generated by selecting randomly $\hat{\boldsymbol{S}}^0$ entries from members of the finite set $S^I = \{S_1^I, \ldots, S_n^I\}$. This allows to investigate the properties of the functions generated and their dependence on properties of the set S^I.

3. Model

The noisy computation model consider here is a feed-forward layered $N \times (L+1)$ Boolean circuit. The layers in the circuit are numbered from 0 (input) to L (output). Each layer $\ell \in \{1, \ldots, L\}$ in the circuit is composed of exactly N ϵ-noisy, k-ary Boolean gates. Noise at gate α_i^{ℓ} on site (i, ℓ) operates independently and in a stochastic manner according to the microscopic law

$$P(S_i^{\ell}|S_{i_1}^{\ell-1}, \ldots, S_{i_k}^{\ell-1}) = \frac{e^{\beta S_i^{\ell} \alpha_i^{\ell}(S_{i_1}^{\ell-1}, \ldots, S_{i_k}^{\ell-1})}}{2\cosh[\beta \alpha_i^{\ell}(S_{i_1}^{\ell-1}, \ldots, S_{i_k}^{\ell-1})]} , \qquad (2)$$

where $\beta = 1/T$ is the inverse temperature, related to the gate noise ϵ via $\tanh\beta = 1 - 2\epsilon$. The gate-output S_i^ℓ is completely random when $\beta \to 0$ ($\epsilon = 1/2$) and completely deterministic when $\beta \to \infty$ ($\epsilon = 0$). Our model is acyclic by definition, so given the state of gates at layer ℓ, gates at layer $\ell + 1$ operate independently of each other. The latter suggests that the probability of the microscopic state $\boldsymbol{S}^0, \ldots, \boldsymbol{S}^L$, where $\boldsymbol{S}^\ell \in \{-1, 1\}^N$, is just a product of equation (2) over all sites and layers in the circuit. Furthermore, to investigate the properties of noisy formulas we consider two copies of the same topology, shown in Figure 1, but with different temperatures $\beta < \infty$ (noisy) and $\hat{\beta} \to \infty$ (noiseless), comparing the two will enable us to study the effect of noise on the resulting functions. Following similar arguments to those of the single circuit case, the probability of microscopic states in the two systems are given by

$$P[\{\boldsymbol{S}^\ell\}; \{\hat{\boldsymbol{S}}^\ell\}] \;=\; P(\boldsymbol{S}^0, \hat{\boldsymbol{S}}^0 | \boldsymbol{S}^I) \prod_{\ell=1}^{L} P(\boldsymbol{S}^\ell | \boldsymbol{S}^{\ell-1}) P(\hat{\boldsymbol{S}}^\ell | \hat{\boldsymbol{S}}^{\ell-1}) \tag{3}$$

where

$$P(\boldsymbol{S}^\ell | \boldsymbol{S}^{\ell-1}) \;=\; \prod_{i=1}^{N} \frac{\mathrm{e}^{\beta S_i^\ell \sum_{j_1,\ldots,j_k}^N A_{j_1,\ldots,j_k}^{\ell,i} \alpha_i^\ell (S_{j_1}^{\ell-1},\ldots,S_{j_k}^{\ell-1})}}{2\cosh[\beta \sum_{j_1,\ldots,j_k}^N A_{j_1,\ldots,j_k}^{\ell,i} \alpha_i^\ell (S_{j_1}^{\ell-1}, \ldots, S_{j_k}^{\ell-1})]}. . \tag{4}$$

The set of connectivity tensors $\{A_{i_1,\ldots,i_k}^{\ell,i}\}$, where $A_{i_1,\ldots,i_k}^{\ell,i} \in \{0, 1\}$, denotes connections in the circuit. The conditional probability $P(\hat{\boldsymbol{S}}^\ell | \hat{\boldsymbol{S}}^{\ell-1})$ is the same as in equation (4) but with $\beta \to \hat{\beta}$.

The sources of disorder in our model are the random connections, random boundary conditions and random gates. The former two arise in the layered growth process described in the last paragraph of section 2. The basic step in this growth process is the addition of a new gate with probability $P(A_{j_1,\ldots,j_k}^{\ell,i}) = \frac{1}{N^k}\delta_{A_{j_1,\ldots,j_k}^{\ell,i};1} + (1 - \frac{1}{N^k})\delta_{A_{j_1,\ldots,j_k}^{\ell,i};0}$ of being connected to exactly k gate-outputs on the previous layer $\ell - 1$. This procedure is carried out independently for all gates in the circuit giving rise to the probability distribution

$$P(\{A_{i_1,\ldots,i_k}^{\ell,i}\}) = \frac{1}{Z_A} \prod_{\ell,i=1}^{L,N} \left[\delta\left[1; \sum_{j_1,\ldots,j_k}^{N} A_{j_1,\ldots,j_k}^{\ell,i}\right] \prod_{i_1,\ldots,i_k}^{N} \left[\frac{1}{N^k}\delta_{A_{i_1,\ldots,i_k}^{\ell,i};1} + (1 - \frac{1}{N^k})\delta_{A_{i_1,\ldots,i_k}^{\ell,i};0}\right]\right] \tag{5}$$

where Z_A is a normalization constant. The Kronecker delta function inside the definition (5) enforces the constraint $\sum_{j_1,\ldots,j_k}^N A_{j_1,\ldots,j_k}^{\ell,i} = 1$, i.e. the gate on site (i, ℓ) is mapped to exactly one element from the set of all possible output-indices $\{i_1, \ldots, i_k\}$ from the previous layer. Other sparse connectivity profiles can be easily accommodated into our framework by incorporating additional constraints into the definition (5) via the appropriate delta functions.

Random boundary conditions in the growth process are generated by selecting members of the input set S^I, where $|S^I| \in O(N^0)$, with uniform probability, and assigning them to the input layer 0. The boundary condition is identical for the two systems yielding the distribution

$$P(\boldsymbol{S}^0, \hat{\boldsymbol{S}}^0 | \boldsymbol{S}^I) = \prod_{i=1}^{N} \delta_{S_i^0; S_{n_i}^I} \delta_{\hat{S}_i^0; S_i^0} \tag{6}$$

where $\{n_i\}$ are independent random indices pointing to the members of input set S^I with probability $P(n_i) = \frac{1}{|S^I|}$. Further correlations can be introduced by defining the probability function $P(S^I)$.

In addition to the topological disorder, induced by the growth process, we assume that the gate α_i^ℓ added at each step of the process can be sampled randomly and independently from the set G of k-ary Boolean gates. Under this assumption the distribution over gates takes the form

$$P(\{\alpha_i^\ell\}) = \prod_{\ell,i=1}^{L,N} P(\alpha_i^\ell) \tag{7}$$

where $P(\alpha_i^\ell) = \sum_{\alpha\in G} p_\alpha \delta_{\alpha;\alpha_i^\ell}$ with $\sum_{\alpha\in G} p_\alpha = 1$ and $p_\alpha \geq 0$.

4. Analysis

Several methods are at our disposal to study this model. For instance, the replica and cavity methods have been successfully employed to study of similar problems from theoretical computer science and information theory [13] ranging from classical combinatorial optimization problems (graph coloring, K-SAT, reconstruction on trees and graph-isomorphism to name but a few) to source and channel coding. However, given the directed nature of the formulae studied we prefer to use the generating functional approach where formula layers take the role of time.

To compute the probability distribution (3) directly for a circuit of finite but significant size is difficult. However, the structure of equation (3) is similar to the one that describes evolution of the disordered Ising spin system [14]. This similarity becomes more apparent if one regards the layers in our model as discrete time-steps of parallel dynamics. We use the generating functional method of statistical mechanics [15]; the generating functional for the current model is given by

$$\Gamma[\boldsymbol{\psi};\hat{\boldsymbol{\psi}}] \;=\; \left\langle \mathrm{e}^{-\mathrm{i}\sum_{\ell,i}\{\psi_i^\ell S_i^\ell + \hat{\psi}_i^\ell \hat{S}_i^\ell\}} \right\rangle \tag{8}$$

where the shorthand $\langle \ldots \rangle$ denotes the average over the joint probability (3). The generating functional (8) can be regarded as a characteristic function of (3) from which moments of the distribution can be obtained by taking partial derivatives with respect to the generating fields $\{\psi_i^\ell, \hat{\psi}_j^{\ell'}\}$, for example $\langle S_i^\ell \hat{S}_j^{\ell'} \rangle = -\lim_{\boldsymbol{\psi},\hat{\boldsymbol{\psi}}\to\mathbf{0}} \frac{\partial^2}{\partial_{\psi_i^\ell}\partial_{\hat{\psi}_j^{\ell'}}}\Gamma[\boldsymbol{\psi};\hat{\boldsymbol{\psi}}]$. Following prescription of [15], we assume that for $N \to \infty$ the system is self-averaging and compute $\overline{\Gamma[\boldsymbol{\psi};\hat{\boldsymbol{\psi}}]}$, where $\overline{[\cdots]}$ denotes an average over the disorder. The disorder-averaged generating function (8) gives rise to the following macroscopic observables

$$\begin{aligned} m(\ell) &= \frac{1}{N}\sum_{i=1}^{N}\overline{\langle S_i^\ell\rangle} = \lim_{\boldsymbol{\psi},\hat{\boldsymbol{\psi}}\to\mathbf{0}} \frac{\mathrm{i}}{N}\sum_{i=1}^{N}\frac{\partial}{\partial_{\psi_i^\ell}}\overline{\Gamma[\boldsymbol{\psi};\hat{\boldsymbol{\psi}}]} \\ C(\ell) &= \frac{1}{N}\sum_{i=1}^{N}\overline{\langle S_i^\ell \hat{S}_i^\ell\rangle} = -\lim_{\boldsymbol{\psi},\hat{\boldsymbol{\psi}}\to\mathbf{0}} \frac{1}{N}\sum_{i=1}^{N}\frac{\partial^2}{\partial_{\psi_i^\ell}\partial_{\hat{\psi}_i^\ell}}\overline{\Gamma[\boldsymbol{\psi};\hat{\boldsymbol{\psi}}]} \end{aligned} \tag{9}$$

where $m(\ell)$ is the average activity (magnetization) on layer ℓ and $C(\ell)$ is the overlap between two systems. Averaging over the disorder in (8) leads to the saddle-point integral

$$\overline{\Gamma} \;=\; \int\{\mathrm{d}\boldsymbol{P}\,\mathrm{d}\hat{\boldsymbol{P}}\,\mathrm{d}\boldsymbol{\Omega}\,\mathrm{d}\hat{\boldsymbol{\Omega}}\}\mathrm{e}^{N\Psi[\boldsymbol{P},\hat{\boldsymbol{P}};\boldsymbol{\Omega},\hat{\boldsymbol{\Omega}}]} \tag{10}$$

where Ψ is the macroscopic saddle-point surface

$$\begin{aligned}\Psi[\ldots] &= \mathrm{i}\sum_{\ell=0}^{L-1}\sum_{S,\hat{S}}\hat{P}^{\ell}(S,\hat{S})P^{\ell}(S,\hat{S}) + \mathrm{i}\sum_{\ell=0}^{L-1}\int \mathrm{d}x\ \mathrm{d}\hat{x}\ \mathrm{d}\omega\ \hat{\Omega}^{\ell}(x,\hat{x},\omega)\Omega^{\ell}(x,\hat{x},\omega) \qquad (11)\\ &+ \sum_{\ell=0}^{L-1}\sum_{\{S_j,\hat{S}_j\}}\prod_{j=1}^{k}\left[P^{\ell}(S_j,\hat{S}_j)\right]\int \mathrm{d}x\ \mathrm{d}\hat{x}\ \mathrm{d}\omega\ \Omega^{\ell}(x,\hat{x},\omega)\left\langle \mathrm{e}^{-\mathrm{i}\{x\alpha(\{S_j\})+\hat{x}\alpha(\{\hat{S}_j\})+\omega\}}\right\rangle_{\alpha}\\ &+ \sum_{n}P(n)\log\int\{\mathrm{d}\boldsymbol{H}\ \mathrm{d}\boldsymbol{x}\ \mathrm{d}\hat{\boldsymbol{H}}\ \mathrm{d}\hat{\boldsymbol{x}}\}\int \mathrm{D}\boldsymbol{\omega}\sum_{\boldsymbol{S},\hat{\boldsymbol{S}}}M_n[\boldsymbol{H},\boldsymbol{x};\hat{\boldsymbol{H}},\hat{\boldsymbol{x}};\boldsymbol{\omega};\boldsymbol{S},\hat{\boldsymbol{S}}]\ ,\end{aligned}$$

where $\langle\cdot\rangle_{\alpha}$ represents an average over gate distribution and M is an effective single-site measure

$$\begin{aligned}M_n[\ldots] &= \delta_{S^0;S^I_n}\delta_{\hat{S}^0;S^0}\prod_{\ell=0}^{L-1}\left[\mathrm{e}^{\mathrm{i}x^{\ell}H^{\ell}+\mathrm{i}\hat{x}^{\ell}\hat{H}^{\ell}+\beta S^{\ell+1}H^{\ell}+\hat{\beta}\hat{S}^{\ell+1}\hat{H}^{\ell}} \right. \qquad (12)\\ &\left. \times\mathrm{e}^{-\log 2\cosh\left(\beta H^{\ell}\right)-\log 2\cosh\left(\hat{\beta}\hat{H}^{\ell}\right)-\mathrm{i}\hat{P}^{\ell}\left(S^{\ell},\hat{S}^{\ell}\right)-\mathrm{i}\hat{\Omega}^{\ell}\left(x^{\ell},\hat{x}^{\ell},\omega^{\ell+1}\right)+\mathrm{i}\omega^{\ell+1}}\right].\end{aligned}$$

The generating fields $\boldsymbol{\psi},\hat{\boldsymbol{\psi}}$ have been removed from the above as they are no longer needed. For $N\to\infty$ the path-integral (10) is dominated by the extremum of the functional $\Psi[\ldots]$ of equation (11). Functional variation of (11) with respect to the order parameters $\{P,\hat{P},\Omega,\hat{\Omega}\}$ gives rise to four saddle-point equations

$$P^{\ell}(S,\hat{S}) = \sum_{n}P(n)\left\langle\delta_{S^{\ell};S}\delta_{\hat{S}^{\ell};\hat{S}}\right\rangle_{M_n} \qquad (13)$$

$$\begin{aligned}\hat{P}^{\ell}(S,\hat{S}) &= \mathrm{i}\sum_{i=1}^{k}\sum_{\{S_j,\hat{S}_j\}}\delta_{S_i;S}\delta_{\hat{S}_i;\hat{S}}\prod_{j\neq i}^{k}\left[P(S_j,\hat{S}_j)\right] \qquad (14)\\ &\times\int \mathrm{d}x\ \mathrm{d}\hat{x}\ \mathrm{d}\omega\ \Omega^{\ell}(x,\hat{x},\omega)\left\langle\mathrm{e}^{-\mathrm{i}\{x\alpha(\{S_j\})+\hat{x}\alpha(\{\hat{S}_j\})+\omega\}}\right\rangle_{\alpha}\end{aligned}$$

$$\Omega^{\ell}(x,\hat{x},\omega) = \sum_{n}P(n)\left\langle\delta(x-x^{\ell})\delta(\hat{x}-\hat{x}^{\ell})\delta(\omega-\omega^{\ell+1})\right\rangle_{M_n} \qquad (15)$$

$$\hat{\Omega}^{\ell}(x,\hat{x},\omega) = \mathrm{i}\sum_{\{S_j,\hat{S}_j\}}\prod_{j=1}^{k}\left[P^{\ell}(S_j,\hat{S}_j)\right]\left\langle\mathrm{e}^{-\mathrm{i}\{x\alpha(\{S_j\})+\hat{x}\alpha(\{\hat{S}_j\})+\omega\}}\right\rangle_{\alpha} \qquad (16)$$

where $\langle\cdots\rangle_{M_n}$ is the average over the probability distribution (12. The saddle-point equations (13)-(16) can be simplified significantly and it turns out that in order to solve this problem we only need to compute the order parameter (13). The physical meaning of this order parameter is given by $P^{\ell}(S,\hat{S}) = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\overline{\langle\delta_{S_i^{\ell};S}\delta_{\hat{S}_i^{\ell};\hat{S}}\rangle|_{S^I}}$, i.e. the disorder-averaged joint probability of sites in the two systems. The single-site effective measure (12) also benefits from the simplification; in particular, if we integrate out the continuous variables in (12) we are led to the expression

$$\begin{aligned}M_n[S^L,\hat{S}^L,\ldots,S^0,\hat{S}^0] &= \delta_{S^0;S^I_n}\delta_{\hat{S}^0;S^0}\prod_{\ell=0}^{L-1}\left\{\sum_{\{S_j,\hat{S}_j\}}\prod_{j=1}^{k}\left[P^{\ell}(S_j,\hat{S}_j)\right]\right. \qquad (17)\\ &\left.\left\langle\times\frac{\mathrm{e}^{\beta S^{\ell+1}\alpha(\{S_j\})}}{2\cosh\beta[\alpha(\{S_j\})]}\frac{\mathrm{e}^{\hat{\beta}\hat{S}^{\ell+1}\alpha(\{\hat{S}_j\})}}{2\cosh\hat{\beta}[\alpha(\{\hat{S}_j\})]}\right\rangle_{\alpha}\right\}.\end{aligned}$$

Using equation (17) the macroscopic observables (9) can be easily computed from the joint probability distribution (13), resulting in the set of equations

$$m(\ell+1) = \sum_{\{S_j\}} \prod_{j=1}^{k} \left[\frac{1}{2}\{1+S_j m(\ell)\}\right] \langle \tanh[\beta\alpha(S_1,\ldots,S_k)]\rangle_\alpha \tag{18}$$

$$\begin{aligned} C(\ell+1) = &\sum_{\{S_j,\hat{S}_j\}} \prod_{j=1}^{k} \left[\frac{1}{2}\{1+S_j m(\ell)+\hat{S}_j\hat{m}(\ell)+S_j\hat{S}_j C(\ell)\}\right] \\ &\times \left\langle \tanh[\beta\alpha(S_1,\ldots,S_k)] \tanh[\hat{\beta}\alpha(\hat{S}_1,\ldots,\hat{S}_k)]\right\rangle_\alpha , \end{aligned} \tag{19}$$

where the magnetization $\hat{m}(\ell)$ is computed by a similar equation to (18) but with $\beta \to \hat{\beta}$. The initial conditions for the above system of equations are given by $m(0) = \hat{m}(0) = \frac{1}{|S^I|}\sum_{S\in S^I} S$, $C(0) = 1$. Equations (18) and (19) describe the coupled evolution of the magnetization and overlap from layer to layer.

The connectivity profile considered in our model leads to a simple mean-field theory, where the macroscopic behaviors of the two copies of the same system is completely determined by the set of observables $\{m(\ell), \hat{m}(\ell), C(\ell)\}$; which relate to the order parameter (13) via $P^\ell(S,\hat{S}) = \frac{1}{2}(1+Sm(\ell)+\hat{S}\hat{m}(\ell)+S\hat{S}C(\ell))$, while the single system behavior is described by $\{m(\ell)\}$. Furthermore, since $\langle\prod_j S^\ell_{i_j}\rangle \to \prod_j \langle S^\ell_{i_j}\rangle$ for finite j, when $N\to\infty$ the spins on layer ℓ are uncorrelated. The reason for this behavior is that in our model, the site (i,ℓ) is a root of a full k-ary tree growing from the input layer $\ell=0$, which in turn points to the input set S^I. The loops in the circuit are rare, so that trees can be regarded as random Boolean formulas and when presented with the input, operate independently of each other. The output of a typical formula at layer ℓ is determined by the probability $P^\ell(S)$.

The magnetization order parameter of equation (9) is related to the normalized Hamming distance $D(\ell)$ between the states $\boldsymbol{S}^\ell$ and $\hat{\boldsymbol{S}}^\ell$ via the identity $D(\ell) = \frac{1}{2}(1-C(\ell)) = \frac{1}{2}(1-\frac{1}{N}\sum_{i=1}^N \overline{\langle S^\ell_i \hat{S}^\ell_i\rangle})$. This allows one to define the order parameter $\Delta(\ell) = \lim_{\beta,\hat{\beta}\to\infty}\frac{1}{2}(1-C(\ell))$, used to probe sensitivity of the circuit with respect to its input, an indication to the complexity of the functions represented by the given circuit. The Hamming distance $D(\ell)$ is also related to the probability $P(S^\ell_i \neq \hat{S}^\ell_i)$ and facilitates the estimate of the error probability $\delta(\ell)$ on the ℓ-th layer of a noisy circuit. More specifically, we define this error probability $\delta(\ell) = \max_{S^I}\lim_{\hat{\beta}\to\infty}\frac{1}{2}(1-C(\ell))$, comparing the maximal error between the noisy and noiseless version of the same circuit with respect to all possible inputs. Obviously, in the absence of noise $(\beta\to\infty)$ and one trivially obtains $\delta(\ell)=0$ for all ℓ.

5. Results

While equations (18),(19) are general for any gate or distribution of gates, we will focus here on a particular Boolean gate, a majority gate with k inputs (MAJ-k). The reasons for choosing this gate are twofold: it was proved to be optimal for noisy computation in formulas [5, 6], and a formula constructed from majority gates can in principle compute any Boolean function [11]. A convenient representation of the MAJ-k gate is given by the identity $\mathrm{MAJ}(S_1,\ldots,S_k) = \mathrm{sgn}[\sum_{j=1}^k S_j]$ where k is odd. Inserting this identity into the equations (18) and (19) one obtains

$$m(\ell+1) = (1-2\epsilon)\sum_{n=0}^{k}\binom{k}{n}\left[\frac{1+m(\ell)}{2}\right]^{n}\left[\frac{1-m(\ell)}{2}\right]^{k-n}\mathrm{sgn}\left[2n-k\right] \quad (20)$$

$$C(\ell+1) = (1-2\epsilon)\sum_{k_1+..+k_4=k}\frac{k!}{k_1!\times..\times k_4!}\,P^{k_1}(-1,-1)\,P^{k_2}(1,-1) \quad (21)$$
$$\times\,P^{k_3}(-1,1)\,P^{k_4}(1,1)\,\mathrm{sgn}\left[k_1-k_2+k_3-k_4\right]\mathrm{sgn}\left[k_1+k_2-k_3-k_4\right]$$

where $P(S,\hat{S}) = \frac{1}{4}(1+Sm(\ell)+\hat{S}\hat{m}(\ell)+S\hat{S}C(\ell))$. In the derivation of the above we have used the identity $\tanh\beta = 1-2\epsilon$ which relates the gate-error ϵ to the inverse temperature β. Also, we have taken the limit $\hat{\beta}\to\infty$ ($\hat{\epsilon}=0$) which allows us to compare noisy circuit to its noiseless counterpart later on. Equations (20) and (21) form the basis for the results presented in the remainder of the paper.

5.1. *Critical behavior in MAJ-k based formulae*

Eq. (20) describes the evolution of the magnetization from layer to layer. The point $m(\infty)=0$ is always a stationary solution of this equation. Expanding Eq. (20) around this stationary solution gives us the value of noise $\epsilon^{*}(k) = 1/2 - 2^{k-2}/k\binom{k-1}{(k-1)/2}$, shown in Fig. 2a, above which the solution $m(\infty)=0$ becomes unstable and two stable solutions $|m(\infty)|\neq 0$ emerge. This result is identical to those reported in [5, 6], and reduces to $\epsilon^{*}(3)=1/6$ for $k=3$; the non-vanishing magnetization values then become [?] $m(\infty)=\pm\sqrt{\frac{1-6\epsilon}{1-2\epsilon}}$.

For $\epsilon > \epsilon^{*}(k)$ the magnetization $m(\ell)$ decays to 0 when $\ell\to\infty$, while for the $\epsilon<\epsilon^{*}(k)$ $\lim_{\ell\to\infty}m(\ell)=\pm m(\infty)$ where the positive and negative stationary solutions correspond to the positive and negative initial magnetizations $m(0)=\frac{1}{|S^{I}|}\sum_{S\in S^{I}}S$ respectively. The function error δ grows respectively. Thus $\epsilon^{*}(k)$ constitutes the critical noise level that separates the unordered phase of the system from ordered one (see the inset of Figure 2a for the case $k=3$). A single input bit of information will be preserved by the circuit for arbitrary many layers only when $\epsilon<\epsilon^{*}(k)$. The probability of an error $P^{\ell}(-S)=\frac{1}{2}(1-Sm(\ell))$ is a measure of how well this one bit of the information is preserved after passing through ℓ layers. A complicated computational task may require significant number of layers, hence only relatively simple operations can be performed by the circuit reliably when $\epsilon>\epsilon^{*}(k)$.

5.2. *Boolean functions generated*

The analysis of Eq. (20) also reveals the type of Boolean functions generated. Stationary solutions for the noiseless system ($\epsilon=0$) result in $m(\infty)=\pm1$ for the initial conditions $m(0)>0$ and $m(0)<0$, respectively. For $m(0)=0$ the resulting asymptotic magnetization is $m(\infty)=0$ and each site is associated with an output of the formula which computes some Boolean function. The average formula on layer ℓ operates according to the probability $P^{\ell}(S)=\frac{1}{2}(1-Sm(\ell))$. This behavior of the magnetization suggests that when $\epsilon=0$ the average formula on layer ℓ converges to the formula that computes the random Boolean functions $F=1$ and $F=-1$ for the initial conditions $m(0)>0$ and $m(0)<0$, respectively, and $F=\pm1$ with equal probability if $m(0)=0$; where $m(0)=\frac{1}{|S^{I}|}\sum_{S\in S^{I}}S$.

Depending on initial conditions the formulas converge to a single Boolean function or to the uniform distribution over some set of functions. For example, if we take $S^{I}=\{-1,-S_1^{I},-S_2^{I}\}$ then the formulas converge to the NAND function. Taking $S^{I}=\{-S_1^{I},-S_2^{I}\}$, on the other hand, gives a uniform distribution over the inverse functions $-S_1^{I}$ and $-S_2^{I}$; this follows from the majority property of the gate. In general, when $m(0)=0$, is difficult to say if the formulas

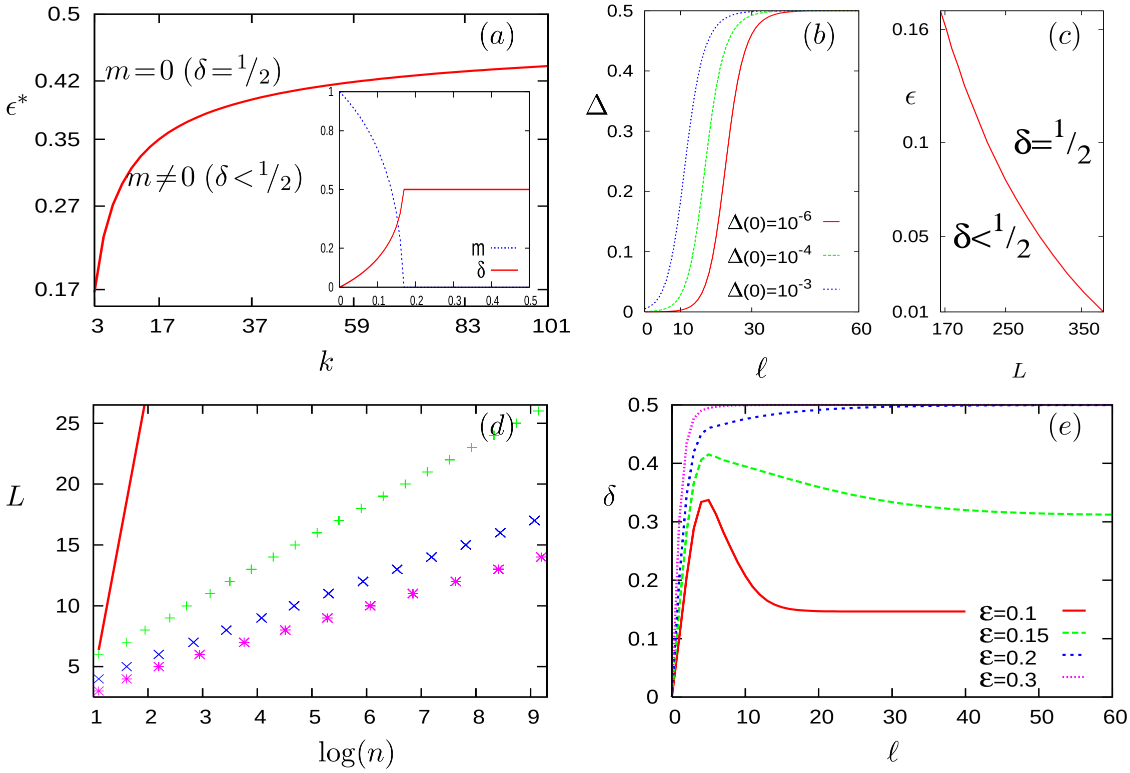

Figure 2. (a)The critical Inset - Properties of MAJ-3 gate-based formulae: Magnetization m and output error δ as a function of gate noise ϵ. (b) Sensitivity of $\Delta(\ell)$ to input mismatch $\Delta(0)$ for $m(0)=0$. (c) Phase diagram for gate noise ϵ at layer L. (d) Number of layers in the noiseless MAJ-k circuit computing MAJ-n function. Theoretic data-points are represented by the symbols $+$ $(k=3)$, $\times$ $(k=5)$ and $*$ $(k=7)$. Slopes of the respective straight lines (not shown on this figure) fitted to the data are 2.473, 1.595 and 1.282. The straight line corresponds to the bound $k2^k\log(n)$, which was derived in [12], plotted here for $k=3$ only. (e) For the MAJ-7 *function*, we plot the evolution, in layers (l), of the output error δ.

compute all Boolean functions in the set generated or only the subset of these functions. However, our result is consistent with other studies [11, 12] with a similar growth process that generates uniformly all random Boolean formulas. In particular, for $S^I=\{-1,1,S^I_1,\ldots,S^I_n\}$ the formulas converge to the MAJ-n function when n is odd and to the uniform distribution over so-called slice functions when n is even [12]. The same happens when the constants $\{-1,1\}$ are removed from the set S^I [12].

5.3. Sensitivity

To measure the evolving overlap between the noiseless and noisy systems one employs Eq. (21); the initial conditions are the same for both systems with $m(0)=\hat{m}(0)$ and $C(0)=1.$. For the noiseless system $\epsilon=0$ the stationary unstable solution of this equation is given by $C(\infty)=1$. Perturbations to the initial state $C(0)=1$ lead to the stable stationary state $C(\infty)=0$. This implies that the circuit is very input-sensitive when $m(0)=0$. In particular, the Hamming distance $\Delta(\ell)=\frac{1}{2}(1-C(\ell))$ increases for small perturbations, amplified by the circuit, to $\Delta(0)$ (see Fig. 2b). This implies that when $\epsilon>0$ the circuit also amplifies the noise and the error $\delta(\ell)$ is growing but can be controlled for a number of layers that depends on the noise level ϵ sufficiently small (see Fig. 2c).

5.4. Convergence rates

In general, we find that the number of layers needed for the magnetization to converge scales as $O(\log n)$ for $m(0) = 1/n$ where $n \in \mathbb{N}$. This rate of convergence is consistent with the results of [12] for the Savický growth process [11]. As we increase ϵ towards the critical value ϵ^* the rate of convergence decreases. Considering a particular function, namely the representation of noiseless MAJ-k circuit computing MAJ-n function, one can obtain the number of layers L required for the various k values as presented in Fig. 2d. The figure also shows the upper $k2^k \log(n)$ bound obtained in [12] for $k = 3$.

Very close to the phase boundary ϵ^* the differences $m(\ell + 1) - m(\ell)$ are very small and the difference equation in the case $k = 3$ can be approximated by the differential equation $\frac{\mathrm{d}}{\mathrm{d}\ell}m(\ell) = -m(\ell) + \frac{1}{2}(1 - 2\epsilon)[3m(\ell) - m^3(\ell)]$, where ℓ is continuous now. The solution to this equation is given by $m^2(\ell) = \left\{ \left[\frac{1}{m^2(0)} - \frac{1-2\epsilon}{1-6\epsilon}\right] \mathrm{e}^{-(1-6\epsilon)\,\ell} + \frac{1-2\epsilon}{1-6\epsilon} \right\}^{-1}$. This solution is only accurate for $\epsilon = 1/6 \pm \Delta\epsilon$, where $0 < \Delta\epsilon \ll 1$, where it gives us the asymptotic form $|m(\ell) - m(\infty)| \approx \mathrm{e}^{-\mathrm{const}\Delta\epsilon\ell}$. Decay rates for other k values are expected to be qualitatively similar.

5.5. Dynamics

The function error $\delta(\ell)$, calculated for $k = 3$ majority gates and shown in Fig. 2e, for different ϵ values, exhibits two distinct stages in the dynamics. It is calculated using Eqs. (20) and (21) for the a formula that represents the MAJ-7 function. Initially, the error increases until it reaches its maximum value at $\ell = 5$, before the MAJ-7 function is computed exactly at $\ell = 8$ for $\epsilon = 0$; the location of this maximum is independent of ϵ. This suggests that gate-inputs at layers $\ell \leq 5$ are non-uniform, contributing to noise-amplification, but become more uniform later leading to noise-suppression and decreasing error. As we approach ϵ^* the number of layers needed for the error to reach stationarity increases; in the region $\epsilon = 1/6 \pm \Delta\epsilon$ it can be estimated from the asymptotic form derived for $m(\ell)$. The dynamic behavior of the error changes to monotonically increasing at $\epsilon^0 = \frac{1}{2}\left[\frac{1-m^2(0)}{3-m^2(0)}\right]$ above which noise cannot be reduced by additional layers. For $\epsilon \gg 1/6$ the error evolution becomes strictly monotonic it relaxes to its stationary value $1/2$ exponentially fast.

6. Discussion

By employing the generating functional methodology of statistical physics we analyzed a the robustness of random functions composed of logical gates to gate-noise. To generate random function via a growth process we employed a variant of the framework devised by Savický [11], which results in Boolean functions that are randomly sampled from the uniform distribution over the function space. The layered variant used facilitates the study of the random functions generated by the growth process as well as their robustness to gate noise.

The results obtained for typical functions retrieve the ones obtained in the theoretical computer science literature and identify them as phase transitions in the corresponding physical system. In addition, it facilitates the derivation of the function properties at any depth and their convergence rate towards the asymptotic values. In addition we obtain typical convergence rate in both the noiseless case and close to the critical noise level; the results are in agreement with the bounds presented in the literature.

The main advantages of the statistical physics methodology over existing methods are the derivation of typical case results, the flexibility in accommodating different gate types or distribution of gates and the accessibility to dynamics results at various depths.

While in this paper we concentrated on the majority gates, we have looked at other gate types [16], hard noise and other properties of noisy circuits that will be reported elsewhere [17]. We believe that much can be explored about the properties of noisy circuits using the methodology developed here, for instance, the type of functions generated depending gate types and the level of gate noise. Work in these areas is underway.

Acknowledgments

Support by the Leverhulme trust grant (F/00 250/H) is gratefully acknowledged.

References

[1] Borkar S 2005 *IEEE Micro* **25** 10–16

[2] Von Neumann J 1956 *Probabilistic logics and the synthesis of reliable organisms from unreliable components* (Princeton, NJ: Princeton University Press) p 43–98 Automata Studies

[3] Pippenger N 1988 *IEEE Trans. Inf. Theory* **34** 194–197

[4] Feder T 1989 *IEEE Trans. Inf. Theory* **35** 569–571

[5] Hajek B and Weller T 1991 *IEEE Trans. Inf. Theory* **37** 388–391

[6] Evans W and Schulman L 2003 *IEEE Trans. Inf. Theory* **49** 3094–3098

[7] Evans W and Pippenger N 1998 *IEEE Trans. Inf. Theory* **44** 1299–1305

[8] Unger F 2008 *IEEE Trans. Inf. Theory* **54** 3693–3698

[9] Steinbach B and Lang C 2003 *Artif. Intell. Rev.* **20**(3) 319–360

[10] Boppana R B 1997 *Inform. Process. Lett.* **63** 257 – 261

[11] Savický P 1990 *Discrete Math.* **83**

[12] Brodsky A and Pippenger N 2005 *Random Struct. Algor.* **27** 490–519

[13] Mézard M and Montanari A 2009 *Information, Physics, and Computation* (Oxford: Oxford University Press)

[14] Hatchett J P L, Wemmenhove B, Castillo I P, Nikoletopoulos T, Skantzos N S and Coolen A C C 2004 *J. Phys. A: Math. Gen.* **37** 6201–6220

[15] De Dominics C 1978 *Phys. Rev. B.* **18** 4913–4919

[16] Mozeika A, Saad D and Raymond J 2009 *Phys. Rev. Lett.* **103** 248701

[17] Mozeika A, Saad D and Raymond J In preperation