

# Noise threshold for universality of 2-input gates

Falk Unger

**Abstract**—It is known that  $\epsilon$ -noisy gates with 2 inputs are universal for arbitrary computation (i.e. can compute any function with bounded error), if all gates fail independently with probability  $\epsilon$  and  $\epsilon < \beta_2 = (3 - \sqrt{7})/4 \approx 8.856\%$ . In this paper it is shown that this bound is tight for formulas, by proving that gates with 2 inputs, in which each gate fails with probability at least  $\beta_2$  cannot be universal. Hence, there is a threshold on the tolerable noise for formulas with 2-input gates and it is  $\beta_2$ . It is conjectured that the same threshold also holds for circuits.

**Index Terms**—Computation with unreliable components, fault-tolerant computation, noise threshold

## I. INTRODUCTION

During the last decades, computers have become faster and faster, mainly due to advances in hardware miniaturization. However, there are physical limits to the possible extent of this miniaturization, and the closer one gets to these limits, the less robust and more error-prone the components become [3], [2]. It is estimated that the time when processor architects face these limitations is within the next decade [4].

Gates, the smallest components of any processor, can fail in (at least) two ways. The first is that they do not work at all. The second is that they work most of the time correctly, and fail sometimes. This type of errors is called “soft errors” by hardware engineers. We deal with faults of the second type.

In particular, we consider the computational model of *noisy formulas*. Formulas are a special kind of circuits in which each gate has exactly one output wire<sup>1</sup>. We ask how much noise on the gates is tolerable, such that any function can still be computed by some formula with bounded-error. We will assume throughout that gates fail independently of each other.

This question has been studied earlier. Already in 1956 von Neumann discovered that reliable computation is possible with noisy 3-majority gates if each

gate fails independently with probability less than 0.0073 [12]. The first to prove an upper bound on the tolerable noise was Pippenger [11]. He proved that formulas with gates of fan-in at most  $k$ , where each gate fails independently with probability at least  $\epsilon \geq \frac{1}{2} - \frac{1}{2k}$ , are not sufficient for universal computation (i.e. not all functions can be computed with bounded error). Feder proved that this bound also applies to circuits [8]. Later, Feder’s bound was improved to  $\frac{1}{2} - \frac{1}{2\sqrt{k}}$  by Evans and Schulman [6].

For formulas with gates of fan-in  $k$  and  $k$  odd, Evans and Schulman [7] proved the tight bound  $\beta_k = \frac{1}{2} - \frac{2^{k-2}}{k \binom{k-1}{k/2-1/2}}$ . Tight here means that if all gates fail independently with the same fixed probability  $\epsilon < \beta_k$ , then any function can be bounded-error computed, and if each gate fails with some probability at least  $\beta_k$  (which does not need to be the same for all gates), universal computation is not possible. For  $k = 3$  the threshold was first established by Hajek and Weller [9].

However, so far it has not been possible to establish thresholds for gates with *even* fan-in (or even prove their existence), as pointed out in [7]. In particular, the most basic case of fan-in 2, which is most commonly used, had been elusive. An intuitive argument why even fan-in is different is that for even fan-in threshold gates (and in particular majority gates) can never be “balanced”, in the sense that the number of inputs on which they are 1 cannot be the same as the number of inputs on which they are 0.

Evans and Pippenger [5] made some progress in this direction. First, they show that all functions can be computed by formulas with noisy NAND-gates with fan-in 2, if each NAND-gate fails with probability exactly  $\epsilon$ , for any  $0 \leq \epsilon < \beta_2 = \frac{3-\sqrt{7}}{4}$ . Second, they show that with NAND-gates alone this bound cannot be improved (They make some additional assumptions which we discuss below). This left open the question of what the bound is if we allow all 16 gates with fan-in 2. We settle this question in this paper.

F. Unger works at the Centrum voor Wiskunde en Informatica in Amsterdam, The Netherlands. (Falk.Unger@cwi.nl)

<sup>1</sup>Precise definitions for all terms used can be found in Section II.

*Theorem 1:* Assume  $\Delta > 0$ . Functions that are computable with bias  $\Delta$  by a formula in which all gates have fan-in at most 2 and fail independently with probability at least  $\beta_2 = (3 - \sqrt{7})/4$ , depend on at most a constant number of input bits.

Together with the first mentioned result from [5] this gives the exact threshold for formulas with gates of fan-in 2. It extends the second result from [5] in the following ways: (1) We allow all gates of fan-in 2, instead of only NAND-gates. (2) We prove that if the noise is exactly  $\beta_2$ , then no universal bounded-error computation is possible. (3) In contrast to our result, the upper bound in [5] only applies to “soft” inputs. They show that gates with noise more than  $\beta_2$  cannot increase the bias. More precisely, if the inputs to the formula are noisy themselves and have bias at most  $\Delta > 0$ , then the output of the formula cannot have larger bias than  $\Delta$ . This left open the case where the input bits are not noisy and either 0 or 1, which is the case we care about most. Our argument shows that even with perfect inputs fault-tolerant computation is not possible for noise at least  $\beta_2$ .

To prove Theorem 1 we introduce a new technique, which is also applicable in the case of fan-in 2. We expect that it can be extended to other fan-in cases.

We conjecture that our bound also holds for circuits.

### A. Outline of the proof

For any function  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  we will choose an input bit  $x_i$  which  $f$  depends on, and fix all other bits such that  $f$  still depends on  $x_i$ . Assume that  $f$  is computed by a formula  $F$  with noisy gates that fail independently with probability at least  $\beta_2$ . Then, for each gate in the formula  $F$  with input wires  $A$  and  $B$  and output wire  $C$  we can define  $a = \frac{1}{2}\mathbb{P}[A = 0 \mid x_i = 0] + \frac{1}{2}\mathbb{P}[A = 0 \mid x_i = 1]$  and  $\delta_a = \mathbb{P}[A = 0 \mid x_i = 0] - \mathbb{P}[A = 0 \mid x_i = 1]$  and analogously for  $B$  and  $C$ . The variable  $a$  can be seen as the average probability of  $A$  being 0. We call  $\delta_a$  the *bias* of  $A$  with respect to the two input settings  $x_i = 0$  and  $x_i = 1$ .

To prove our result one could attempt the following, which will turn out to not quite work (but we then show how to fix that): For an  $\epsilon$ -noisy gate with fan-in 2, input wires  $A, B$  and output wire  $C$ , we would like to show that if the noise  $\epsilon$  is at least

the threshold  $\beta_2$  then for any  $\delta > 0$  there is some  $0 \leq \theta < 1$  such that if  $\delta \leq \max\{|\delta_a|, |\delta_b|\}$  then

$$|\delta_c| \leq \theta \max\{|\delta_a|, |\delta_b|\} \quad (1)$$

This would mean that the bias goes down exponentially with the number of computation steps, until it reaches  $\delta$ . Further, it is easy to show that for any  $d > 0$  there is a function  $f$  such that any formula computing  $f$  has one input bit  $x_i$  on which  $f$  depends and the number of computation steps on any path from  $x_i$  to the output bit is at least  $d$ . Hence, the bias cannot be bounded away from zero for all  $f$  and  $x_i$ .

Unfortunately, (1) is not always true. Sometimes the bias can actually go up.<sup>2</sup> We use a more sophisticated approach, showing that the bias goes down “on average”: We define a *potential function*  $q$ , which is positive and bounded on  $[0, 1]$ . Instead of showing (1) we show that for any  $\delta > 0$  there is some  $0 \leq \theta < 1$  such that if  $\delta \leq \max\{|\delta_a|, |\delta_b|\}$  then

$$|\delta_c|q(c) \leq \theta \max\{|\delta_a|q(a), |\delta_b|q(b)\}. \quad (2)$$

and if  $\delta > \max\{|\delta_a|, |\delta_b|\}$  then (2) holds for  $\theta = 1$ . Since  $q$  is bounded, this implies that for any arbitrarily small constant  $\delta > 0$  the bias of any formula becomes  $O(\delta)$  after a constant number of computation steps. We can then proceed as above.

We give the main proof in Section IV. In Section III we prove (2), in the main Lemma 1. Section V contains some remarks on our particular choice of  $q$ .

## II. DEFINITIONS

A *circuit* is composed of gates. Each *gate* has a certain number of input wires, which is called the *fan-in* of the gate. The wires can take boolean values 0 or 1. A gate computes an output bit as a boolean function of its input bits. A *formula* is a particular type of circuit in which the gates are connected in a tree, with the output gate at the root and the input bits at the leaves. In particular, this means that each gate has exactly one output wire.

A (perfect) PARITY-gate with input bits  $x_1$  and  $x_2$  outputs 0 if  $x_1 = x_2$  and 1 otherwise. A (perfect) OR-gate outputs 0 if  $x_1 = x_2 = 0$  and 1 otherwise.

<sup>2</sup>An easy example is an OR-gate with noise  $\epsilon = 1/10$ ,  $\delta_a = \delta_b = 1/10$  and  $a = b = 8/10$ , for which  $\delta_c = (a\delta_b + b\delta_a)(1 - 2\epsilon) = 0.128 > 1/10$ .

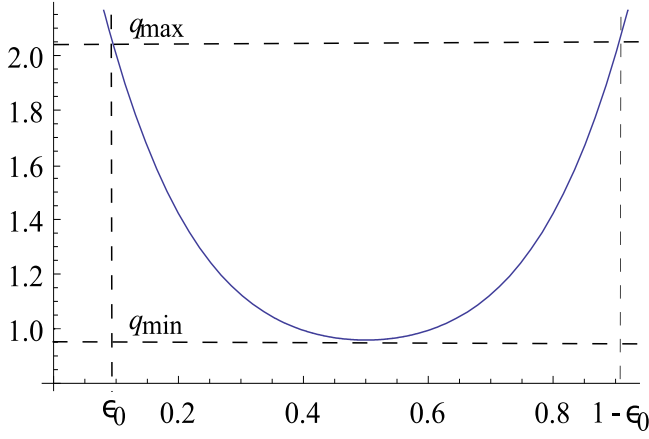


Fig. 1. Graph of  $q(x)$

We call a gate  $\epsilon$ -noisy if it outputs the correct result with probability  $1 - \epsilon$  and with probability  $\epsilon$  it outputs the opposite. We say that a formula  $F$  with noisy gates computes the function  $f$  with bias  $\Delta > 0$  if for all  $x \in f^{-1}(0)$ ,  $y \in f^{-1}(1)$ :  $\mathbb{P}[F(x) = 0] \geq \Delta + \mathbb{P}[F(y) = 0]$ .<sup>3</sup> If  $f$  can be computed with some bias  $\Delta > 0$  we also say that  $f$  is *computable with bounded-error*.

A function  $f : \{0, 1\}^n \rightarrow \{0, 1\}$  depends on the  $i$ -th input bit  $x_i$  if there is some setting of the other bits, such that flipping  $x_i$  flips the function value. The number of bits that  $f$  depends on is denoted by  $d(f)$ .

In a formula, we define the *depth* of a wire  $A$ , denoted by  $\text{depth}(A)$ , as the number of 2-input gates on a path from  $A$  to the output wire. Gates with fan-in 1 are not counted.

For the definition of the quantities  $a$  and  $\delta_a$  for a wire  $A$  we refer to Section I-A.

### III. BIAS REDUCTION FOR NOISY GATES

We define the constant  $x_0 = 1/(2 - 4\beta_2) = (1 + \sqrt{7})/6 \approx 0.61$ . It will turn out later that an OR-gate with input wires  $A, B$  performs best when  $a \approx x_0$  and  $b \approx x_0$ . Further, we define the *potential function*

$$\begin{aligned} q(x) &= \left(\frac{29}{2} + 2\sqrt{7}\right) \left(x - \frac{1}{2}\right)^4 \\ &\quad + \left(\frac{5\sqrt{7}}{2} - \frac{13}{4}\right) \left(x - \frac{1}{2}\right)^2 - \frac{\sqrt{7}}{2} + \frac{73}{32} \quad (3) \\ &\approx 19.79(x - 0.5)^4 + 3.36(x - 0.5)^2 + 0.96. \end{aligned}$$

This is a biquadratic function in  $(x - 1/2)$ . Fur-

<sup>3</sup>For our purposes it does not matter that with this definition  $f$  and  $\bar{f}$  are actually computed by the same  $F$ .

ther,  $q$  is symmetric around  $1/2$  and convex. In  $[\beta_2, 1 - \beta_2]$   $q$  is bounded between  $q_{\min} = q(1/2) = -\sqrt{7}/2 + 73/32 > 0.9$  and  $q_{\max} = q(\beta_2) = (247 + 8\sqrt{7})/128 < 2.1$ , see Figure III.

For any  $\epsilon \leq 1/2$  we define the function

$$\eta_\epsilon(x) = (1 - 2\epsilon)x + \epsilon.$$

If  $x$  is the probability that some variable is 0, then  $\eta_\epsilon(x)$  is the probability that it is 0 after it has gone through an  $\epsilon$ -noisy channel.

#### A. Technical Lemmas

In the rest of this section we establish inequality (8) in Lemma 1, from which the proof of the main theorem will follow relatively straightforwardly. The proof of this inequality is quite technical and so at first reading the reader might just want to read the statement of Lemma 1 and then move immediately to Section IV, where we establish the main result. Inequality (8) can also be checked with the help of a computer (e.g. using Mathematica [1]), but in the remainder of this section we will prove it rigorously.

*Proposition 1:* For all  $a, b$  with  $\beta_2 \leq a, b \leq 1 - \beta_2$  it holds that

$$\begin{aligned} q(a)q(b) - (1 - 2\beta_2)(aq(a) + bq(b))q(\eta_{\beta_2}(ab)) \\ \geq 0. \end{aligned} \quad (4)$$

*Proof:* We write  $a = x_0 + s_a$  and  $b = x_0 + s_b$ . W.l.o.g. let  $|s_b| \geq |s_a|$  and choose  $-1 \leq k \leq 1$  s.t.  $s_a = ks_b$ . Then the lhs of (4) can be written as

$$\sum_{i=0}^{11} r_i(k) s_b^{i+2}. \quad (5)$$

The reason why (5) only starts with a quadratic term in  $s_b$  is our special choice of  $q$ , see Section V-A for more on this. The first coefficient is easily computed

$$r_0(k) = \left(3 - \frac{3\sqrt{7}}{4}\right) (k^2 + 1).$$

This function attains its minimum value of  $3 - 3\sqrt{7}/4 \approx 1.02$  at  $k = 0$ . Therefore, there is a  $\kappa > 0$  s.t. for  $a, b \in [x_0 - \kappa, x_0 + \kappa]$  the lhs of (4) is non-negative. We show that  $\kappa = 0.02$  is a solution.

The absolute value of the other coefficients for  $-1 \leq k \leq 1$  can be bounded by  $|r_1(k)| \leq 5$ ,  $|r_2(k)| \leq 31$ ,  $|r_3(k)| \leq 18$ ,  $|r_4(k)| \leq 68$ ,  $|r_5(k)| \leq 326$  and for all other  $|r_i(k)| \leq 5000$ . Therefore, if  $|s_b| \leq 1/50$ , (5) is at least

$$\begin{aligned} s_b^2 (1.02 - 5(0.02) - 31(0.02)^2 - 18(0.02)^3 - \dots) \\ \geq 0.90s_b^2 \geq 0. \end{aligned}$$

This proves the case  $x_0 - 1/50 \leq a, b \leq x_0 + 1/50$ . For all other  $|a - x_0| \geq 1/50$  or  $|b - x_0| \geq 1/50$  the proposition follows from Fact 1 with  $\mu = 0$ . ■

We now state some bounds on polynomials. They are similar in spirit to (4), with the crucial difference that these bounds are not tight. This is convenient, because there are several techniques for finding global optima of multivariate polynomials up to arbitrary precision. See [10] for an overview. We have used the computer algebra program Mathematica [1]. We used an accuracy of  $10^{-10}$  and rounded the results in such a way that the bounds given are rigorous.<sup>4</sup>

*Fact 1:* For all  $a, b$  with  $\beta_2 \leq a, b \leq 1 - \beta_2$  with  $|a - x_0| \geq 1/50$  and  $0 \leq \mu \leq \xi := (1 - \beta_2 - a)(1 - \beta_2 - b)$  it holds that

$$q(a)q(b) - (1 - 2\beta_2)(aq(a) + bq(b))q(\eta_{\beta_2}(ab + \mu)) > 0.0003. \quad (6)$$

*Proof:* Notice that  $\mu$  only appears in the term  $q(\eta_{\beta_2}(ab + \mu))$ . For  $0 \leq \mu \leq \xi$  we notice that by convexity of  $q$  and linearity of  $\eta_{\beta_2}$  it follows that  $q(\eta_{\beta_2}(ab + \mu)) \leq \max\{q(\eta_{\beta_2}(ab)), q(\eta_{\beta_2}(ab + \xi))\}$ . Thus, (6) is minimized for  $\mu = 0$  or  $\mu = \xi$ . For  $\mu = 0$  the lhs of (6) is lower bounded by 0.0003 and for  $\mu = \xi$  by 0.01. ■

*Fact 2:* For all  $a, b, \mu$  with  $\beta_2 \leq b \leq 1 - \beta_2$ ,  $1/2 \leq a \leq 1 - \beta_2$  and  $|\mu| \leq \xi := 2(1 - \beta_2 - a)(1 - \beta_2 - b)$  it holds that

$$q(a)q(b) - ((2a - 1)q(a) + (2b - 1)q(b))(1 - 2\beta_2) \times q(\eta_{\beta_2}(ab + (1 - a)(1 - b) + \mu)) \geq 0.45.$$

*Proof:* For  $\mu = \xi$  the term is lower bounded by 0.48 and for  $\mu = -\xi$  by 0.55. Using convexity of  $q$  as above the fact follows. ■

*Fact 3:* For all  $a, b, \mu$  with  $\beta_2 \leq b \leq 1 - \beta_2$ ,  $\beta_2 \leq a \leq 1/2$  and  $|\mu| \leq \xi := 2(a - \beta_2)(1 - \beta_2 - b)$  it holds that

$$q(a)q(b) - ((1 - 2a)q(a) + (2b - 1)q(b))(1 - 2\beta_2) \times q(\eta_{\beta_2}(ab + (1 - a)(1 - b) + \mu)) \geq 0.48.$$

*Proof:* For  $\mu = \xi$  the term is lower bounded by 0.51 and for  $\mu = -\xi$  by 0.48. The fact then follows by convexity of  $q$  as above. ■

<sup>4</sup>Even more simple, one could bound the first derivatives and check all values of the polynomials on a small enough grid.

*Fact 4:* Let  $a, b, \mu$  with  $\beta_2 \leq a, b \leq 1 - \beta_2$  Then

$$q(a) - (1 - 2\beta_2)bq(\eta_{\beta_2}(ab - \mu)) > 0.22$$

holds if (a)  $a \leq 1/2$  and  $-(a - \beta_2)(1 - \beta_2 - b) \leq \mu \leq 0$  or (b)  $1/2 \leq a$  and  $-(1 - \beta_2 - a)(1 - \beta_2 - b) \leq \mu \leq 0$ .

*Proof:* For  $\mu = 0$  (and all  $\beta_2 \leq a \leq 1 - \beta_2$ ) the term is lower bounded by 0.23. For both cases  $a \leq 1/2$ ,  $\mu = -(a - \beta_2)(1 - \beta_2 - b)$  and  $1/2 \leq a$ ,  $\mu = -(1 - \beta_2 - a)(1 - \beta_2 - b)$  the term is lower bounded by 0.22. Using convexity of  $q$  as above the fact follows. ■

We can state our main Lemma.

*Lemma 1:* Let  $\beta_2 \leq \epsilon \leq 1/2$ . Assume an  $\epsilon$ -noisy OR-gate or PARITY-gate, with input wires  $A$  and  $B$  and output wire  $C$ . Let

$$\begin{aligned} \beta_2 &\leq \mathbb{P}[A = 0 \mid x_i = 0] \leq 1 - \beta_2 \\ \beta_2 &\leq \mathbb{P}[A = 0 \mid x_i = 1] \leq 1 - \beta_2, \end{aligned} \quad (7)$$

and let the same be true for  $B$ . Define  $a, b, c$  and  $\delta_a, \delta_b, \delta_c$  for  $A, B, C$  as in Section I.

1) The following inequality holds for  $\theta = 1$ :

$$|\delta_c|q(c) \leq \theta \max\{|\delta_a|q(a), |\delta_b|q(b)\}. \quad (8)$$

2) For any  $\delta > 0$  there is a  $0 \leq \theta < 1$  such that if  $|\delta_a| \geq \delta$  or  $|\delta_b| \geq \delta$ , (8) is still true for this  $\theta$ .

*Proof:* We consider the OR-gate first. We have  $\mathbb{P}[C = 0 \mid x_i = 0] = \eta_\epsilon((a + \delta_a/2)(b + \delta_b/2))$  and  $\mathbb{P}[C = 0 \mid x_i = 1] = \eta_\epsilon((a - \delta_a/2)(b - \delta_b/2))$ , which implies

$$\begin{aligned} \delta_c &= (a\delta_b + b\delta_a)(1 - 2\epsilon) \\ c &= \eta_\epsilon(ab + \delta_a\delta_b/4). \end{aligned}$$

Increasing  $\epsilon$  decreases  $|\delta_c|$  as well as also  $q(c)$ , since  $c$  gets closer to  $1/2$  and  $q$  decreases towards  $1/2$ . Thus we may assume  $\epsilon = \beta_2$ . Further, we may assume  $|\delta_a|q(a) \geq |\delta_b|q(b)$ . Note that, for  $\delta_a = 0$  we then also have  $\delta_b = 0$  and the Lemma holds trivially. In the remainder we therefore assume  $\delta_a \neq 0$ . In fact, we will even assume  $\delta_a > 0$ : In case  $\delta_a < 0$  we can just formally replace every occurrence of  $\delta_a$  and  $\delta_b$  with  $-\delta_a$  resp.  $-\delta_b$ . Because of the absolute value signs, this will not change the validity of (9). So we have to prove

$$\begin{aligned} (1 - 2\epsilon) |a\delta_b + b\delta_a| q(\eta_\epsilon(ab + \delta_a\delta_b/4)) & \quad (9) \\ \leq \theta |\delta_a| q(a). \end{aligned}$$

In the remainder, we will repeatedly use that  $a$  and  $b$  are bounded between  $\beta_2$  and  $1 - \beta_2$  and that in this range,  $0.9 < q_{min} \leq q(a) \leq q_{max} < 2.1$ , without mentioning it each time. We distinguish the following cases:

$\delta_b > 0$ : Since we assumed  $|\delta_a|q(a) \geq |\delta_b|q(b)$ , it is enough to prove (9) where we replace the first occurrence of  $\delta_b$  by  $\delta_a q(a)/q(b)$ . Cancelling  $\delta_a$  and multiplying by  $q(b)$  we get

$$\begin{aligned} & \theta q(a)q(b) \\ & - (1 - 2\beta_2)(aq(a) + bq(b))q(\eta_{\beta_2}(ab + \delta_a \delta_b/4)) \\ & \geq 0. \end{aligned} \quad (10)$$

In case  $|a - x_0| \geq 1/50$  or  $|b - x_0| \geq 1/50$ , note that  $\delta_a \delta_b/4 \leq (1 - \beta_2 - a)(1 - \beta_2 - b)$ . If we set  $\mu = \delta_a \delta_b/4$  and  $\theta = 1$ , then by Fact 1 the lhs of (10) is greater than 0.0003. This implies the existence of a  $\theta < 1$  for (10) and settles both parts of the Lemma.

We are left with the case  $|a - x_0| < 1/50$  and  $|b - x_0| < 1/50$ . By (7) we can then bound  $\delta_a/2 \leq 1 - \beta_2 - a \leq 1 - \beta_2 - x_0 + 1/50 < 0.33$  and similarly  $\delta_b/2 < 0.33$ , i.e.  $(1 - 2\beta_2)\delta_a \delta_b/4 < 0.1$ . We also note that in our case  $0.37 < \eta_{\beta_2}(ab) < 0.42$ . By convexity,  $\min_{0.37 \leq x \leq 0.42} q(x) - q(x + 0.1) = q(0.42) - q(0.52) > 0.02$ , and thus  $q(\eta_{\beta_2}(ab) + 0.1) < q(\eta_{\beta_2}(ab)) - 0.02$ . This last inequality, convexity of  $q$  and  $(1 - 2\beta_2)\delta_a \delta_b/4 < 0.1$  imply  $q(\eta_{\beta_2}(ab) + (1 - 2\beta_2)\delta_a \delta_b/4) < q(\eta_{\beta_2}(ab)) - \frac{0.02}{0.1}(1 - 2\beta_2)\delta_a \delta_b/4$ . Noting that  $\eta_{\beta_2}(ab) + (1 - 2\beta_2)\delta_a \delta_b/4 = \eta_{\beta_2}(ab + \delta_a \delta_b/4)$  this becomes

$$\begin{aligned} & q(\eta_{\beta_2}(ab + \delta_a \delta_b/4)) \\ & < q(\eta_{\beta_2}(ab)) - (1 - 2\beta_2)\delta_a \delta_b/20. \end{aligned} \quad (11)$$

In particular  $q(\eta_{\beta_2}(ab + \delta_a \delta_b/4)) < q(\eta_{\beta_2}(ab))$ . Plugging the lhs of this into (10) and using Proposition 1 implies (10) for  $\theta = 1$ . This establishes part 1 of the Lemma for  $\delta_b > 0$ .

Now part 2 of the Lemma. Let  $\delta_a \geq \delta$  or  $\delta_b \geq \delta$ . Consider first the case that  $\delta_b$  is not too small compared to  $\delta_a$ , say  $\delta_b \geq \delta_a/100$ . Together with our assumption  $|\delta_a|q(a) \geq |\delta_b|q(b)$  this implies  $(1 - 2\beta_2)\delta_a \delta_b/20 \geq (1 - 2\beta_2)\delta^2/2000$ . With (11) we then get  $q(\eta_{\beta_2}(ab + \delta_a \delta_b/4)) + c < q(\eta_{\beta_2}(ab))$  for  $c = (1 - 2\beta_2)\delta^2/2000 > 0$  and putting this into (4) gives  $q(a)q(b) - (1 - 2\beta_2)(aq(a) + bq(b))(q(\eta_{\beta_2}(ab + \delta_a \delta_b/4)) + c) > 0$ . This implies the existence of a  $\theta < 1$  for (10) and establishes part 2 of the Lemma when  $\delta_b \geq \delta_a/100$ .

If  $\delta_b$  is small, i.e.  $\delta_b < \delta_a/100$ , then upper bounding the first occurrence of  $\delta_b$  by  $\delta_a q(a)/q(b)$  to get from (9) to (10) was far from tight. A better bound is  $\delta_b < \delta_a q(a)/(10q(b))$ , which derives from  $q(a)/(10q(b)) \geq q_{min}/(10q_{max}) > 1/100$ . Analogously to the derivation of (10) we get

$$\begin{aligned} & \theta q(a)q(b) \\ & - (1 - 2\beta_2)(aq(a)/10 + bq(b))q(\eta_{\beta_2}(ab + \delta_a \delta_b/4)) \\ & \geq 0. \end{aligned} \quad (12)$$

By (7),  $a > \beta_2$ . Thus,  $aq(a) > \beta_2 q_{min}$  and also  $q(\eta_{\beta_2}(ab + \delta_a \delta_b/4)) > q_{min}$ . Hence, the lhs of (12) is at least  $(1 - 2\beta_2)\beta_2 q_{min}^2 9/10$  smaller than the lhs of (10). Since we already proved earlier that (10) holds for  $\theta = 1$  without the restriction  $\delta_b < \delta_a/100$ , we conclude that (12) holds for some  $\theta < 1$ . This establishes part 2 of the Lemma for  $\delta_b < \delta_a/100$ .

$\delta_b \leq 0$ : It is enough to prove (9) where we replace  $|a\delta_b + b\delta_a|$  by (a)  $|b\delta_a|$  or (b)  $|a\delta_b|$ . If in case (a) we cancel  $\delta_a$  and  $q(a)$  after the replacement, we see that a  $\theta < 1$  must exist if

$$q(a) - (1 - 2\beta_2)bq(\eta_{\beta_2}(ab + \delta_a \delta_b/4)) \geq \chi, \quad (13)$$

for some  $\chi > 0$ . Note that in case  $a \leq 1/2$  we have  $-(a - \beta_2)(1 - \beta_2 - b) \leq \delta_a \delta_b/4 \leq 0$  and in case  $1/2 \leq a$  we have  $-(1 - \beta_2 - a)(1 - \beta_2 - b) \leq \delta_a \delta_b/4 \leq 0$ . The Lemma then follows from Fact 4.

For case (b) we note that  $|a\delta_b| \leq a\delta_a q(a)/q(b)$ . Replacing  $|a\delta_b + b\delta_a|$  in (9) by  $a\delta_a q(a)/q(b)$  and rearranging terms we get exactly the same as (13), with  $a$  and  $b$  swapped. We proceed as in case (a).

We now consider the PARITY-gate. First note, that if the two input wires of a noiseless PARITY gate are independently 0 with probability  $\alpha$  resp.  $\beta$ , then the output wire will be 0 with probability  $\alpha\beta + (1 - \alpha)(1 - \beta)$ . Thus, in our case

$$\begin{aligned} & \mathbb{P}[C = 0 \mid x_i = 0] \\ & = \eta_\epsilon((a + \delta_a/2)(b + \delta_b/2) \\ & \quad + (1 - a - \delta_a/2)(1 - b - \delta_b/2)) \end{aligned}$$

and

$$\begin{aligned} & \mathbb{P}[C = 0 \mid x_i = 1] \\ & = \eta_\epsilon((a - \delta_a/2)(b - \delta_b/2) \\ & \quad + (1 - a + \delta_a/2)(1 - b + \delta_b/2)) \end{aligned}$$

which implies

$$\begin{aligned} c & = \eta_\epsilon(ab + (1 - a)(1 - b) + \delta_a \delta_b/2) \\ \delta_c & = ((2a - 1)\delta_b + (2b - 1)\delta_a)(1 - 2\epsilon) \end{aligned}$$

We need to prove

$$\begin{aligned} & |(2a-1)\delta_b + (2b-1)\delta_a| (1-2\epsilon) \times \\ & q(\eta_\epsilon(ab + (1-a)(1-b) + \delta_a\delta_b/2)) \quad (14) \\ & \leq \theta |\delta_a| q(a). \end{aligned}$$

As for the OR-gate we only need to consider  $\epsilon = \beta_2$  and may assume  $\delta_a \geq 0$  w.l.o.g, because otherwise we can just change the signs of both  $\delta_a$  and  $\delta_b$ . Also, w.l.o.g. we assume  $|\delta_a|q(a) \geq |\delta_b|q(b)$ . If  $\delta_a = 0$ , then also  $\delta_b = 0$  and the Lemma becomes trivial. So we assume  $\delta_a > 0$ . Further, we may assume  $b \geq 1/2$  (and therefore  $(2b-1)\delta_a \geq 0$ ), because formally replacing  $a$  and  $b$  by  $1-a$  and  $1-b$  does not change (14). We condition on the sign of  $2a-1$ .

First  $2a-1 \geq 0$ . It is enough to prove (14), where we replace the first occurrence of  $\delta_b$  by  $\delta_a q(a)/q(b)$ , since we assumed  $|\delta_a|q(a) \geq |\delta_b|q(b)$ . Cancelling  $\delta_a$  and rearranging terms, the existence of a  $0 \leq \theta < 1$  for (14) then follows from

$$\begin{aligned} & q(a)q(b) - ((2a-1)q(a) + (2b-1)q(b))(1-2\beta_2) \\ & \quad \times q(\eta_{\beta_2}(ab + (1-a)(1-b) + \delta_a\delta_b/2)) \\ & \geq \chi > 0. \end{aligned}$$

This inequality follows from Fact 2 by noting that  $|\delta_b| \leq 2(1-\beta_2-b)$  and  $|\delta_a| \leq 2(1-\beta_2-a)$ .

In case  $2a-1 < 0$  we can proceed similarly, where this time we replace the first occurrence of  $\delta_b$  by  $-\delta_a q(a)/q(b)$  and bound  $|\delta_a| \leq 2(a-\beta_2)$ . The resulting inequality follows from Fact 3. ■

#### IV. PROOF OF THEOREM 1

*Proof:* Let  $f$  be any function and let  $F$  be any formula with noisy gates that fail independently with probability at least  $\beta_2$ . Let  $F$  compute  $f$  with bias  $\Delta$ . We show that  $f$  depends on at most a constant number of bits, i.e.  $d(f) \leq c(\Delta)$ , for some function  $c(\Delta)$ .

Before starting we note the following: Every  $\epsilon$ -noisy fan-in-2 gate can be constructed from an  $\epsilon$ -noisy PARITY- or an  $\epsilon$ -noisy OR-gate, perfect NOT-gates and constant inputs. Hence, we may assume that  $F$  is constructed from perfect NOT-gates and noisy PARITY-gates and OR-gates.

Let  $x_i$  be an input bit on which  $f$  depends with the additional property that any input wire of  $F$  carrying  $x_i$  has depth at least  $\lceil \log_2 d(f) \rceil$ . Because all gates in  $F$  have fan-in at most 2, the existence of such  $x_i$  is guaranteed. Fix all other input bits such that the output of  $F$  changes when flipping  $x_i$ .

Set  $D = \lceil \log_2 d(f) \rceil - 1$  and  $\delta = \frac{\Delta}{2q_{max}}$ . Let  $\theta < 1$  be given by Lemma 1 for this  $\delta$ . In case this results in  $\theta < 1 - 2\beta_2$ , set  $\theta = 1 - 2\beta_2$ . (The adjustment  $\theta \geq 1 - 2\beta_2$  is not really needed, but will later simplify the proof.) We will prove inductively that for any wire  $C$  at depth  $d \leq D$

$$q(c)|\delta_c| \leq \max\left\{\frac{\Delta}{2}, \theta^{D-d} q_{max}\right\}. \quad (15)$$

For  $d = D$  (15) holds trivially. Now take any wire  $C$  in  $F$  with depth  $d < D$ . We distinguish what signal  $C$  carries.

Firstly,  $C$  can be an input wire carrying  $x_j$ . Then necessarily  $i \neq j$ , because input wires carrying  $x_i$  have depth at least  $D+1$ . Thus,  $\delta_c = 0$  and (15) holds.

Secondly,  $C$  can be the output of a noiseless NOT-gate, which has input wire  $B$ . Note that since we do not count NOT-gates in the depth of a wire,  $depth(C) = depth(B)$ ,  $c = 1-b$  and  $\delta_c = -\delta_b$ . Then, by symmetry of  $q$  around  $1/2$  we get (15) for  $C$  from the same statement for  $B$ .

Thirdly,  $C$  can be the output of gate  $G$ , with  $G$  either an OR-gate or a PARITY-gate. Let the input wires to  $G$  be  $A$  and  $B$ . If one wire is a constant, say  $A$ , then gate  $G$  is essentially a (noisy) gate with fan-in 1. Hence,  $G$  always outputs either a (noisy) 0 or 1, or  $G$  is the noisy identity- or the noisy NOT-gate. In the first two cases  $\delta_c = 0$ . In the last two cases we can easily calculate that  $|b-1/2|(1-2\epsilon) = |c-1/2|$  and  $|\delta_c| \leq (1-2\beta_2)|\delta_b|$ . Because  $q$  decreases monotonically towards  $1/2$  and we chose  $\theta \geq 1 - 2\beta_2$ , (15) holds.

So we are left with the case where both inputs to  $G$  are non-constant. Since  $d < D$ , both wires  $A$  and  $B$  are the output of some noisy gate, so the conditions (7) in Lemma 1 are satisfied. We may assume  $|\delta_b|q(b) \leq |\delta_a|q(a)$  w.l.o.g. If  $|\delta_a|q(a) \leq \Delta/2$ , then by part 1 from Lemma 1 also  $|\delta_c|q(c) \leq \Delta/2$  and (15) holds. If  $|\delta_a|q(a) > \Delta/2$ , then  $|\delta_a| > \frac{\Delta}{2q_{max}} = \delta$ . Then (15) follows from part 2 of Lemma 1 and the inductive assumption.

Let  $O$  be the output wire of  $F$ , which by assumption has bias  $\Delta$ . Because  $q(o)\Delta \leq \Delta/2$  is impossible (since  $q(o) \geq q_{min} > 1/2$ ) we get from (15):  $q(o)\Delta \leq \theta^D q_{max}$ , and further  $\Delta \leq \theta^{\lceil \log_2 d(f) \rceil - 1} (q_{max}/q_{min})$ , which implies  $\frac{\log_2(\Delta q_{min}/q_{max})}{\log_2 \theta} + 1 \geq \log_2 d(f)$ . Since  $\theta$  depends only on  $\Delta$ ,  $d(f)$  is upper bounded by the function

$$c(\Delta) := 2 (\Delta q_{min}/q_{max})^{1/\log_2 \theta}.$$

## V. DISCUSSION

We have shown a tight threshold for the noise which is tolerable for computation by formulas with gates of fan-in at most 2. This is the first result for gates with an even number of wires. It should be possible to generalize it to other fan-in, although the proof is probably more tedious.

The same bound probably also applies to *circuits* with gates of fan-in at most 2.

### A. Choice of potential function

So far we have not given any idea of why we chose this particular potential function. In fact, this choice is not unique. The choice of  $q$  was determined as follows: (1) It is convenient to choose  $q$  symmetric around  $1/2$ , so applying a NOT-gate to wire  $A$  does not change the value of  $|\delta_a|q(a)$ . (2) It is natural to scale  $q$  such that  $q(x_0) = 1$ . (3) After these choices, we have to choose  $\frac{d}{dx}q(x)|_{x=x_0} = \frac{1}{2}(-1 + \sqrt{7})$ . This ensures that (5) does not have a linear term in  $s_b$  and only starts with the quadratic term, i.e. “ $r_{-1}(k) \equiv 0$ ”. (4) We also need  $\frac{d^2}{dx^2}q(x)|_{x=x_0} > 16 - 4\sqrt{7} \approx 5.42$ , because that makes  $r_0(k) > 0$  for  $-1 \leq k \leq 1$ . The rest of the choices are not so binding.

However, a quadratic function alone is not enough. For (5) to be at least 0 one also has constraints on higher derivatives of  $q$ . The expression in (3) for  $q$  is one of the “nicer” possible potential functions. One can also find a possible  $q$ , by dividing the interval  $[\beta_2, 1 - \beta_2]$  into smaller intervals and define  $q$  as different quadratic functions in each of these intervals.

## VI. ACKNOWLEDGEMENTS

I would like to thank Intel chip designers Bob Colwell and Shekhar Borkar for pointing out [2] and [4]. I am also grateful to Ronald de Wolf, Peter Harremoës and an anonymous referee for some comments and to Ben Reichardt for proof-reading and pointing out some errors.

## REFERENCES

- [1] Available at <http://www.wolfram.com/>.
- [2] S. Borkar. Designing reliable systems from unreliable components: The challenges of transistor variability and degradation. *IEEE Micro*, 25(6):10–16, 2005.

- [3] P. Bose. Designing reliable systems with unreliable components. *IEEE Micro*, 26(5):5–6, 2006.
- [4] B. Colwell. Computer Architecture Beyond Moore’s Law. St. Petersburg, 8-12 June 2006. International Computer Science Symposium in Russia.
- [5] W. Evans and N. Pippenger. On the maximum tolerable noise for reliable computation by formulas. *IEEE Trans. Inform. Theory*, 44(3):1299–1305, 1998.
- [6] W. Evans and L. Schulman. Signal propagation and noisy circuits. *IEEE Trans. Inform. Theory*, 45(7):2367–2373, 1999.
- [7] W. Evans and L. Schulman. On the maximum tolerable noise of  $k$ -input gates for reliable computation by formulas. *IEEE Trans. Inform. Theory*, 49(11):3094–3098, 2003.
- [8] T. Feder. Reliable computation by networks in the presence of noise. *IEEE Trans. Inform. Theory*, 35(3):569–571, 1989.
- [9] B. Hajek and T. Weller. On the maximum tolerable noise for reliable computation by formulas. *IEEE Trans. Inform. Theory*, 37(2):388–391, 1991.
- [10] P. Parrilo and B. Sturmfels. Minimizing polynomial functions. In S. Basu and L. Gonzalez-Vega, editors, *Algorithmic and quantitative real algebraic geometry*, volume 60, pages 83–100. American Mathematical Society, 2003.
- [11] N. Pippenger. Reliable computation by formulas in the presence of noise. *IEEE Trans. Inform. Theory*, 34(2):194–197, 1988.
- [12] J. von Neumann. Probabilistic logics and the synthesis of reliable organisms from unreliable components. In C. E. Shannon and J. McCarthy, editors, *Automata Studies*, volume 3, pages 43–99. Princeton University Press, Princeton, 1956.