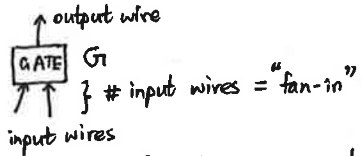
Anuran Makur
17 May 2017

BETTER GATES CAN MAKE FAULT-TOLERANT COMPUTATION IMPOSSIBLE: (Unger '10)

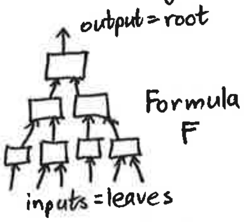
★ Preliminaries:

- A gate is a Boolean function (i.e. Boolean inputs and outputs). It is said to be of NOR-type if it is a NOR gate with additional NOT gates at the inputs and outputs.

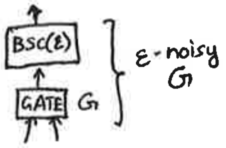


- A formula is a collection of gates connected in a tree. Each gate has 1 output wire, and we will consider gates with fan-in ≤ 2.

↑
special case of a circuit



- A gate is ε-noisy with $\varepsilon \in [0, \frac{1}{2}]$ if we apply a $\mathrm{BSC}(\varepsilon)$ to the output. (flipover prob.)



- A formula F with noisy gates computes $f: \{0,1\}^n \to \{0,1\}$ with bias $\Delta > 0$ if $\forall x \in f^{-1}(0)$, $\forall y \in f^{-1}(1)$, $\underbrace{\mathbb{P}(F(x)=0) - \mathbb{P}(F(y)=0)}_{= \mathbb{P}(F(y)=1) - \mathbb{P}(F(x)=1)} \geq \Delta$.

★ Thm: (Phase Transition) If all gates fail w.p. $\geq \beta_2 = 0.08856$, then fault-tolerant computation is not possible. If all gates fail with the same prob. $\varepsilon < \beta_2$, then fault-tolerant computation is possible.

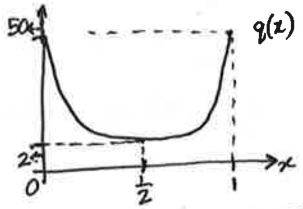
★ Notation & Definitions:

- Fix a Boolean function f that is computed by a formula F whose gates fail independently.
- Fix an input bit $x_i \in \{0,1\}$ and all other input bits s.t. flipping x_i flips the output of f.
- For a wire A in F, we define: ① $a \triangleq \frac{1}{2}\mathbb{P}(A=0|x_i=0) + \frac{1}{2}\mathbb{P}(A=0|x_i=1)$, ← average prob.

  ② $\delta_a \triangleq \mathbb{P}(A=0|x_i=0) - \mathbb{P}(A=0|x_i=1)$. ← bias

  ↳ $\left[\,|\delta_a| = \|P_{A|x_i=0} - P_{A|x_i=1}\|_{TV}\right]$

- Define a potential function, $q: [0,1] \to \mathbb{R}^+$, $q(x) = \dfrac{1}{x(1-x) + \underbrace{(11-4\sqrt{7})/18}_{\approx 0.023}}$.



(q is convex, $2 < q < 50$, symmetric arond $x = \frac{1}{2}$.)

- Let "extractable information" of wire $A \triangleq |\delta_a| q(a)$. ← We will exploit SDPIs for this.
- Prop: (Monotonicity) For a wire A, let $y = \min\{\mathbb{P}(A=0|x_i=0), \mathbb{P}(A=0|x_i=1)\} = a - \frac{|\delta_a|}{2}$ and let $z = \max\{\mathbb{P}(A=0|x_i=0), \mathbb{P}(A=0|x_i=1)\} = a + \frac{|\delta_a|}{2}$, so that $|\delta_a| q(a) = (z-y) q\left(\frac{z+y}{2}\right)$ (= rational func. in y, z). Let A' be another wire with $y' \leq z'$ defined similarly. Then, if $y' \leq y \leq z \leq z'$, $|\delta_a| q(a) \leq |\delta_{a'}| q(a')$.

  Pf: Take derivative of $(z,y) \mapsto (z-y) q\left(\frac{z+y}{2}\right)$ wrt y and z and check they are non-negative. ▨

[continued.]

★ Useful Lemmata:

Lemma 1: (SDPI) Consider an ε-noisy gate G with inputs A, B and output C. Then,

$$\forall \beta_2 \leq \varepsilon \leq \tfrac{1}{2},\ \forall \underset{\text{i.e. } \{a,\delta_a\},\{b,\delta_b\},\{c,\delta_c\}}{A,B,C},\qquad |\delta_c|q(c) \leq \underset{\text{contraction coeff.}}{\theta}\ \max\{|\delta_a|q(a), |\delta_b|q(b)\} \qquad [\bigstar]$$

holds in the following cases:

① Always with $\theta = 1$ [DPI]. doesn't dep. on G

② If G is not NOR-type, $\exists\, 0 \leq \theta < 1$ such that $[\bigstar]$ holds.

③ $\forall\, 0 \leq r < 1$, $\exists\, 0 \leq \underset{}{\overset{\text{deps on } r \text{ only}}{\theta}} < 1$, if $\min\{|\delta_a|q(a), |\delta_b|q(b)\} \leq r \max\{|\delta_a|q(a), |\delta_b|q(b)\}$, then $[\bigstar]$ holds (but not for all A, B, C).

④ $\forall \delta > 0$, $\exists\, 0 \leq \overset{\text{deps on } \delta \text{ only}}{\theta} < 1$, if ⓐ $|\delta_a| \geq \delta$ or $|\delta_b| \geq \delta$ or ⓑ G is a NOR gate and $|a - \hat{x}| \geq \delta$ or $|b - \hat{x}| \geq \delta$, then $[\bigstar]$ holds (but not for all A, B, C).

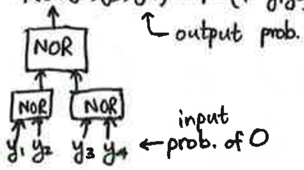
Here, $\hat{x} \triangleq \dfrac{1+\sqrt{7}}{6} \approx 0{\cdot}61.$

Pf: Tedious optimization of real polynomials. ▨

Lemma 2: (Composition of perfect NOR gates) There is a $D(=17) > 0$ such that the following holds: Consider a full binary tree of noise-free NOR gates with inputs $A_1, \ldots, A_{2^{2D}}$, output B, and depth $2D$. If $\forall\, 1 \leq i \leq 2^{2D}$, $|\hat{x} - a_i| \leq \frac{1}{1000}$ and $|\delta_{a_i}| \leq \frac{1}{1000}$, then $|\delta_b| q(b) \leq \frac{1}{2} \max_{1 \leq i \leq 2^{2D}} |\delta_{a_i}| q(a_i)$.

↳ i.e. NOR tree contracts extractable information if input cond. probs close to $\hat{x}$.

Pf: $f(y_1, y_2, y_3, y_4) = 1 - (1 - y_1 y_2)(1 - y_3 y_4)$ output prob. of 0



Let $f^{(1)} = f$, $\underset{\text{composition}}{f^{(k+1)}}(y_1, \ldots, y_{4^{k+1}}) = f(f(y_1, \ldots, y_{4^k}), f(y_{4^k+1}, \ldots, y_{2 \cdot 4^k}), f(y_{2 \cdot 4^k+1}, \ldots, y_{3 \cdot 4^k}), f(y_{3 \cdot 4^k+1}, \ldots, y_{4^{k+1}}))$.

For a NOR tree of depth $2k$, inputs $Y_1, \ldots, Y_{4^k}$ & output W:

$$\delta_w = \underbrace{f^{(k)}\left(y_1 + \tfrac{\delta_{y_1}}{2}, \ldots, y_{4^k} + \tfrac{\delta_{y_{4^k}}}{2}\right)}_{\mathbb{P}(W=0|x_i=0)} - \underbrace{f^{(k)}\left(y_1 - \tfrac{\delta_{y_1}}{2}, \ldots, y_{4^k} - \tfrac{\delta_{y_{4^k}}}{2}\right)}_{\mathbb{P}(W=0|x_i=1)}.$$

$\hat{\delta} \triangleq \max_i |\delta_{y_i}|$

Notice that $f, f^{(k)}$ are monotone increasing in each $y_i \in [0,1]$. So, $\delta_w \leq f^{(k)}(y_1 + \frac{\hat{\delta}}{2}, \ldots) - f^{(k)}(y_1 - \frac{\hat{\delta}}{2}, \ldots, y_{4^k} - \frac{\hat{\delta}}{2})$.

Claim 1: $\forall\, 0 \leq y_1, \ldots, y_4 \leq \hat{x} + 3/2000$, $f(y_1, \ldots, y_4) \leq \max\{y_1, y_2, y_3, y_4\}$. ($\Rightarrow f^{(k)}(\text{args}) \leq \max\{\text{args}\}$)

Pf: Since f increasing, we prove $f(y, y, y, y) - y \leq 0$, which holds because $f(y, y, y, y)$ is a polynomial in y. ▨

Claim 2: $\forall\, 0 \leq y_1, \ldots, y_4 \leq 1$, $\forall \delta \geq 0$, $f(y_1 + \frac{\delta}{2}, \ldots, y_4 + \frac{\delta}{2}) - f(\underbrace{y_1 - \tfrac{\delta}{2}}_{\geq 0}, \ldots, \underbrace{y_4 - \tfrac{\delta}{2}}_{\geq 0}) \leq 4\delta \max\{y_1, \ldots, y_4\}$.

Pf: Optimize polynomials. ▨

Notice that $f^{(11)}(\hat{x} + 3/2000, \ldots, \hat{x} + 3/2000) \leq \frac{1}{40}$, and each a_i satisfies $a_i + \frac{|\delta_{a_i}|}{2} \leq \hat{x} + \frac{3}{2000}$.

Claim 1 $\Rightarrow$ $\forall$ wires W that are $2s$ levels from inputs with $s \geq 11$, $\underset{\text{average}}{w} \leq \frac{1}{40}$. [use $f^{(k)}$ monotone of $f^{(s)}(\cdot+, \cdot+ \ldots)$ & $f^{(s)}(\cdot-, \cdot-, \ldots) \leq f^{(s)}(\max, \max, \ldots)$]

[for $s > 11$] Claim 2 $\Rightarrow$ $\delta_w \leq 4^{11}(\hat{x} + \frac{1}{1000})^{11} \hat{\delta}$ for w $2 \cdot 11$ levels from inputs.

Claim 2 $\Rightarrow$ $\delta_w \leq 4^{11}(\hat{x} + \frac{1}{1000})^{11} (4 \cdot \frac{1}{40})^6 \hat{\delta} \leq \frac{\hat{\delta}}{50}$ for w $2 \cdot (11+6)$ levels from inputs. needed to make 4^{17} decay!

Let $D = 17$. Then, $\delta_b \leq \dfrac{\hat{\delta}}{50} = \dfrac{\max_i |\delta_{a_i}|}{50} \Rightarrow |\delta_b| q(b) \leq \frac{1}{2} \max_i |\delta_{a_i}| q(a_i)$, as $2 < q < 50$. ▨

[continued]

Lemma 3: (Blocks of Depth $D=17$) There is a constant $0 \leq \gamma < 1$ such that for any formula Q of depth $2D$ (with gates of fan-in $=2$) which has input wires $A_1, \dots, A_{2^{2D}}$ and output wire B, and for every $\{a_i\}, \{\delta_{a_i}\}$ on A_i, by either setting all noise rates to 0, or setting all noise rates to β_2, we have:

dist. of A_i

$$|\delta_b| q(b) \leq \gamma \max_i |\delta_{a_i}| q(a_i).$$

Pf: **Claim 1:** (Stability under noisy NOR) Consider a β_2-noisy NOR gate where input wires are indep'ly 0 w.p. y_1, y_2 and output wire is 0 w.p. $y_3 = \beta_2 y_1 y_2 + (1-\beta_2)(1-y_1 y_2)$. Define $s = \frac{250496 - 499\sqrt{7}}{250000(-1+\sqrt{7})} < \hat{x} - \frac{2}{1000}$, $t = \hat{x} + \frac{2}{1000}$. Then, $y_1, y_2 \in [s,t] \Rightarrow y_3 \in [s,t]$.

Pf: Compute $\min_{y_1,y_2 \in [s,t]} y_3 = s$, $\max_{y_1,y_2 \in [s,t]} y_3 < t$. ▨

Claim 2: (Enough to find one "bad" wire) For every $0 \leq \theta' < 1$, there is a $0 \leq \gamma < 1$ s.t. for every formula Q and every $\{a_i\}, \{\delta_{a_i}\}$ on A_i, when all gates are β_2-noisy, if there is a wire W with $|\delta_w| q(w) \leq \theta' \max_i |\delta_{a_i}| q(a_i)$, then $|\delta_b| q(b) \leq \gamma \max_i |\delta_{a_i}| q(a_i)$.

[induction] **Pf:** Let depth(W) = no. of gates between wire W and wire B (output).

Inductive Hypothesis: If depth$(W) \leq j$, then claim is true.

Base case: If depth$(W) = 0$, then $W = B$ and we have $\theta' = \gamma$. The claim is true here.

Inductive step: Consider a wire W at depth $j+1$, s.t. $|\delta_w| q(w) \leq \theta' \max_i |\delta_{a_i}| q(a_i)$.

Consider gate G:  , where $|\delta_v| q(v) \leq \max_i |\delta_{a_i}| q(a_i)$ [Lemma 1·1].

actually true for all wires in Q

corresp. to $r = \frac{1}{2}$ in Lemma 1·3

If $|\delta_w| q(w) \leq |\delta_v| q(v)/2$, then $\exists\, 0 \leq \theta < 1$, $|\delta_u| q(u) \leq \theta |\delta_v| q(v) \leq \theta \max_i |\delta_{a_i}| q(a_i)$ using Lemma 1·3. If $|\delta_w| q(w) > |\delta_v| q(v)/2$, then $|\delta_u| q(u) \leq |\delta_w| q(w) \leq \theta' \max_i |\delta_{a_i}| q(a_i)$ using Lemma 1·1. Hence, we have:

$$\exists\, 0 \leq \theta < 1, \; |\delta_u| q(u) \leq \theta \max_i |\delta_{a_i}| q(a_i).$$

Since depth$(U) = j$, by inductive hypothesis, we prove the claim for W. ▨

We now complete the proof of Lemma 3 by cases.

① If one gate in Q is not of NOR-type, then set all noise rates to β_2. There is a wire W s.t. $|\delta_w| q(w) \leq \theta \max_i |\delta_{a_i}| q(a_i)$ where θ is given by Lemma 1·2.

to use claim 2 + Lemma 1

single value

Lemma 3 follows from Claim 2.

② All gates in Q are NOR-type. Put Q in "canonical form": Each gate is NOR with possible NOTs in the inputs only. Output gate may have NOT at output. Input gates have no NOTs at inputs.

Suppose $\exists i$ s.t. $|a_i - \hat{x}| \leq \frac{1}{1000}$ and $|\delta_{a_i}| \leq \frac{1}{1000}$ is not true.

Consider gate G s.t.  . Then, $|\delta_w| q(w) \leq \theta \max_i |\delta_{a_i}| q(a_i)$ where θ is given by Lemma 1·4 when all noise rates are set to β_2.

dep.s on $\delta = \frac{1}{1000}$ only (single value)

to use claim 2 + Lemma 1

Lemma 3 follows from Claim 2.

[continued.]

④

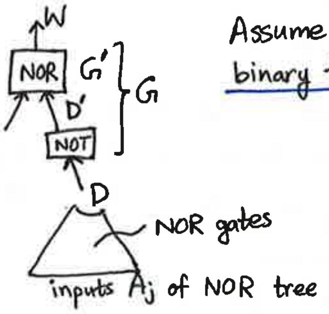
Anuran Makur
17 May 2017

Pf cont'd:

So, all gates in Q (canonical form) are NOR-type, and $\forall i, |a_i - \hat{x}| \leq \frac{1}{1000}$ and $|\delta_{a_i}| \leq \frac{1}{1000}$. [★]

③ Suppose all gates are NOR gates. Set all noise rates to 0. Lemma 3 follows from Lemma 2.
↳ uses $D = 17$

④ There is a gate G which is NOR with at least one NOT in its inputs.



Assume WLOG that the formula is a complete binary tree of NOR gates below D.

For any input A_j of the NOR tree,

$$\mathbb{P}(A_j = 0 | x_i = 0) \underset{[★]}{\in} \left[\hat{x} - \frac{2}{1000}, \hat{x} + \frac{2}{1000}\right] \subseteq [s,t]$$

$\Rightarrow \mathbb{P}(D = 0 | x_i = 0) \in [s,t]$ by Claim 1.

Likewise, $\mathbb{P}(D=0|x_i=1) \in [s,t] \Rightarrow d \in [s,t] \Rightarrow \underset{\text{corresp. to } D'}{d'} = 1 - d \leq 1 - s < \frac{4}{10} \Rightarrow |d' - \hat{x}| > \frac{1}{1000}$.

By Lemma 1·4(b), $|\delta_w| q(w) \leq \Theta \max_i |\delta_{a_i}| q(a_i)$ where Θ is given by Lemma 1·4(b).
↑ dep.s only on $\delta = \frac{1}{1000}$ (single value)

Lemma 3 follows from Claim 2, where we set all noise rates to β_2. ▨

Remark:
Since the Θ values in cases ①, ②, ④ are single choices, the corresponding γ values are single choices. Hence, $\exists$ universal γ s.t. Lemma 3 holds.

Lemma 4: (Continuity) Lemma 3 holds for some noise rate $0 \leq \alpha_2 < \beta_2$. [The corresp. γ is of course different.]

Pf: For formula Q, δ_b and b are polynomials of the noise rate ε.
One can verify that $\left|\frac{\partial \delta_b q(b)}{\partial \varepsilon}\right| \leq C$ for some universal constant $C > 0$.
By mean value theorem, Lemma 4 follows from Lemma 3. ▨

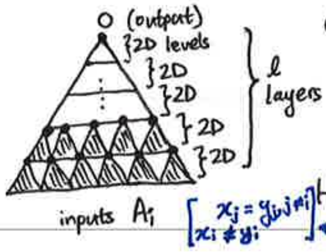
★ Main Result:

★ Thm: There is a $0 \leq \alpha_2 < \beta_2$ s.t. for any $\Delta > 0$, there are Boolean functions f s.t. no formula F computes f with bias $\Delta > 0$ if the noise rates of the gates in F are chosen "adversarially" from $\{0, \alpha_2\}$.
↳ fan-in ≤ 2

Proof: Fix $D = 17$ from Lemma 2, and α_2, γ from Lemma 4. Choose $\ell \in \mathbb{N}$ s.t. $50\gamma^\ell \leq \Delta$. Fix a Boolean function f that depends on at least $2^{2\ell D}$ inputs. Let F be a formula to compute f. (Let the depth of a gate be the depth of its output wire.) There is at least one★ input variable x_i which f depends on s.t. all input wires carrying x_i have depth $\geq 2\ell D$. Fix all other input bits s.t. flipping x_i flips the output of f.

★ Since fan-in ≤ 2, if all inputs have ≥1 wire at depth $< 2\ell D$, there are not enough nodes in tree for all inputs.

WLOG, assume that F is a full binary tree up to depth $2\ell D$.
→ add NOR subtree to gates with 1 input
→ for fixed input x_j, make corresp. gate assume x_j input and add dummy subtree of NOR gates



Choose noise rate $\in \{0, \alpha_2\}$ for each subtree ▲ s.t. Lemma 4 induces a contraction by γ.

$$\Rightarrow \underbrace{|\delta_0| q(0)}_{>2} \leq \gamma^\ell \max_i \underbrace{|\delta_{a_i}|}_{\leq 1} \underbrace{q(a_i)}_{\leq 50} \leq 50\gamma^\ell \overset{\text{assumed}}{\leq} \Delta \Rightarrow |\delta_0| < \frac{\Delta}{2}.$$

$\left[\begin{matrix} x_j = y_j \ \forall j \neq i \\ x_i \neq y_i \end{matrix}\right]$ Hence, $\exists x \in f^{-1}(0), y \in f^{-1}(1)$, s.t. $\underbrace{\mathbb{P}(0 = \emptyset | x_i = 0)}_{\mathbb{P}(F(x)=0)}^{\text{gives } 0} - \underbrace{\mathbb{P}(0 = \emptyset | x_i = 1)}_{\mathbb{P}(F(y)=0)}^{\text{gives } 1} < \frac{\Delta}{2}$. ▨

[THE END]