

# Erratum to “Probabilities of Boolean Functions Given by Random Implicational Formula”

Antoine Genitrini

Laboratoire d’Informatique de Paris 6,  
CNRS UMR 7606 and  
Université Pierre et Marie Curie,  
4 place Jussieu,  
75252 Paris cedex 05, France.  
`antoine.genitrini@lip6.fr`

Bernhard Gittenberger\*

Institute for Discrete Mathematics  
and Geometry  
Vienna University of Technology  
A-1040 Wien, Austria  
`gittenberger@dmg.tuwien.ac.at`

Cécile Mailler<sup>†</sup>

Department of Mathematical Sciences  
The University of Bath  
Claverton Down  
Bath BA2 7AY  
`c.mailler@bath.ac.uk`

January 12, 2014

We wish to issue an erratum for the paper [3]. The paper contains an error in Section 6 where it is shown that the model of random Boolean formulae into consideration does not exhibit the Shannon effect. The proof is based on the construction of a sufficiently large family of trees (representing only functions of small complexity) and containing trees from a family  $\Pi_{x,y}$  expanded once with some tree of a family  $\mathcal{H}_{x,y}$ . The generating function  $H_{x,y}(z)$  of  $\mathcal{H}_{x,y}$  is described by an expression containing a multiple iteration of an operator  $\Psi$  on the generating function  $P(z)$  of all trees [3, pp. 15–16].

On [3, p. 16] it is claimed that  $\Psi$  does not change the singularity  $\eta$  of  $P$ . This is true and even the nature of the singularity is preserved: indeed the singular expansion  $P(z) = \alpha - \beta\sqrt{1 - z/\eta}$ , as  $z \rightarrow \eta$ , is transformed into  $\Psi^k(P)(z) = \alpha_k - \beta_k\sqrt{1 - z/\eta}$ . However, the sequence  $\beta_k$  tends to zero exponentially fast when  $k$  tends to  $+\infty$ . The consequence of this is that the constant  $c$  appearing in [3, Eq. (7)] as well as the subsequent estimates on the same page depends on  $k$  (and hence on  $n$ , the number of variables, since  $k = \Theta(n^2)$ ) and actually tends to zero which flaws the proof of [3, Theorem 3].

We present here a correct proof based on the ideas presented in [2].

---

\*Supported by FWF grant SFB F50-03 and ÖAD, grant F03/2013.

<sup>†</sup>Supported by EPSRC grant EP/K016075/1.

## Proof of [3, Theorem 3]

Let us count the number of valid premise expansions of trees of size at most  $n^2$ . All such expansions thus compute functions of complexity at most  $n^2$ . We will prove that the limiting ratio of valid premise expansions of trees of size at most  $n^2$  tends to a positive constant as  $n$  tends to infinity.

Note that one large tree can be a valid premise expansion of two different trees of size at most  $n^2$ . To avoid the multiple-counting of such trees, we restrict the family of expansions as follows: in this section, we only consider premise expansions where the expanding tree has exactly three subtrees, one labelled by the variable according to which the expansion is done, and the two others being two implicational trees of size at least  $n^2$ . Let us denote by  $\mathcal{E}$  the family of trees that are obtained by such a premise expansion of a tree of size at most  $n^2$ . Given such a tree, it is possible to find where the expansion has been done, just by looking for the topmost internal node that has two subtrees of size at least  $n^2$ . This property ensures not to count several times the same tree in  $\mathcal{E}$  by expanding smaller trees.

To calculate the limiting ratio of  $\mathcal{E}$ , we have to answer the following question: Consider an implicational tree of size  $r$ . How many different valid premise expansions can be done in this tree? A possible answer is based on the following bivariate generating function, already introduced in the binary planar case (*cf.* [2]). First fix a variable  $y \in \{x_1, \dots, x_n\}$ ,

- $T(u, z)$  is the generating function of implicational trees where  $z$  marks all nodes and  $u$  marks the nodes having at least one ancestor labelled by  $y$ , counted with multiplicity. This means that, given a tree  $t$ , each node having  $k$  ancestors labelled by  $y$  contributes a multiplicative factor  $zu^k$  to the weight of  $t$ .
- $V(u, z)$  is the generating function of implicational trees where  $z$  marks all nodes and  $u$  marks the nodes having at least two ancestors labelled by  $y$ , again counted with multiplicity. This means that in a tree  $t$  each node having  $k$  ancestors labelled by  $y$  contribute a multiplicative factor  $zu^{k-1}$  to the weight of  $t$ .

Now observe: Let  $F(z) = \sum_{m \geq 0} F_m z^m$  where  $F_m$  is the cumulative number of vertices in all implicational trees of size  $m$  in which a valid  $y$ -premise expansion is possible. Then  $F(z) = \partial_u \Delta(1, z)$  where  $\partial_u \Delta := \partial / \partial u \Delta$  denotes the partial derivative of  $\Delta(u, z)$  w.r.t.  $u$ . Moreover, observe that by symmetry  $T(u, z)$  and  $V(u, z)$  do not depend on  $y$ .

We thus get the following lower bound (because we have restricted the expansions):

$$\begin{aligned} \sum_{f \mid L(f) \leq n^2} \mathbb{P}_n(f) &\geq \sum_{r=1}^{n^2} \sum_{y \in \{x_1, \dots, x_n\}} [z^r] \partial_u \Delta(1, z) \lim_{m \rightarrow +\infty} \frac{[z^{m-r}] G_y(z)}{[z^m] P(z)}, \\ &= n \sum [z^r] \partial_u \Delta(1, z) \lim_{m \rightarrow +\infty} \frac{[z^{m-r}] G_y(z)}{[z^m] P(z)} \end{aligned} \quad (1)$$

where  $G_y(z)$  is the generating function of implicational trees having three subtrees, one of them of size one, labelled by  $y$ , and the two others being implicational trees of size at least  $n^2$ . Of course, again by symmetry  $G_y(z)$  is independent of  $y$ .

**Lemma 1.** *The dominant singularity  $\eta$  of  $P(z)$  satisfies*

$$\eta = \frac{1}{en} - \frac{1}{2e^3n^2} - \frac{8e+9}{24e^5n^3} + O\left(\frac{1}{n^4}\right).$$

*Proof.* The first two terms in the asymptotic expansion were given in [3, Eq. (3)]. Further bootstrapping yields the next term.  $\square$

**Corollary 2.** *For large enough  $n$  we have*

$$en \geq \eta^{-1} \left(1 - \frac{1}{2e^2n} - \frac{1}{n^2}\right)$$

and, for all  $m$ ,

$$\frac{e^{-m}}{2} \exp\left(-\frac{m}{n^2} \left(1 + \frac{1}{4e^4}\right)\right) \leq \eta^m \left(n + \frac{1}{2e^2}\right)^m \leq 2e^{-m} \exp\left(-\frac{m}{4e^4n^2}\right).$$

**Lemma 3.** *Let  $R(z)$  be the unique solution of  $R(z) = \left(n + \frac{1}{2e^2}\right) z \cdot \exp(R(z))$  that satisfies  $R(z) = \sum_{m \geq 0} R_m z^m$  with  $R_m \geq 0$ . Then for sufficiently large  $n$  and  $m$  we have*

$$R_m \geq \frac{\eta^{-m}}{\sqrt{2\pi m^3}} \left(1 - \frac{1}{12m}\right) \exp\left(-\frac{m}{n^2} \left(1 + \frac{1}{4e^4} + \frac{1}{2e^2n}\right)\right).$$

Moreover, for  $m \geq 3$ ,  $P_m \geq R_m$ .

The idea of the rest of the proof is to deal with  $R(z)$  instead of  $P(z)$ , because it is simpler to deal with and the coefficients  $R(z)$  are a good approximation of those of  $P(z)$ .

*Proof.* Using Lagrange inversion [1, e.g. p. 127], we deduce

$$R_m = \left(n + \frac{1}{2e^2}\right)^m \frac{m^{m-1}}{m!}.$$

Using Stirling's formula [1, p. 407] and Lemma 1, we get, for large enough  $m$  and for all  $n$

$$R_m \geq \frac{(en)^m}{\sqrt{2\pi m^3}} \left(1 - \frac{1}{12m}\right) \left(1 + \frac{1}{2e^2n}\right)^m \geq \frac{\eta^{-m}}{2\sqrt{2\pi m^3}} \left(1 - \frac{1}{12m}\right) \left(1 - \frac{1+4e^4}{4e^4n^2} - \frac{1}{2e^2n^3}\right)^m.$$

Thus, for large enough  $m$  and  $n$  we obtain

$$R_m \geq \frac{\eta^{-m}}{2\sqrt{2\pi m^3}} \left(1 - \frac{1}{12m}\right) \exp\left(-\frac{m}{n^2} \left(1 + \frac{1}{4e^4} + \frac{1}{2e^2n}\right)\right).$$

Let us now turn to the second statement of the lemma, that asserts that  $R_m$  is a lower bound for  $P_m$ , when  $m \geq 3$ . By differentiating the functional equation satisfied by  $R(z)$ , we get:

$$R'(z) = \frac{R(z)}{z} + R'(z) \cdot R(z). \quad (2)$$

This equation translates directly to a recurrence satisfied by the coefficients of  $R(z)$ :

$$R_{m+1} = \frac{1}{m} \cdot \sum_{k=0}^{m-1} (k+1) R_{k+1} R_{m-k} \quad \forall m \geq 2, \quad (3)$$

with the first coefficients  $R_0 = 0$  and  $R_1 = n + 1/(2e^2)$ . Let us now introduce the generating function  $S(z)$  satisfying  $S(z) = nz \exp(S(z) + S(z^2)/2)$ . Since the functional equation of  $S(z)$  is a truncation of the one satisfied by  $P$ , we must have  $S_m \leq P_m$ . By differencing this functional equation we get

$$S'(z) = \frac{S(z)}{z} + S'(z) \cdot S(z) + z \cdot S'(z^2) \cdot S(z)$$

which translates to

$$S_{m+1} = \frac{1}{m} \cdot \left( \sum_{k=0}^{m-1} (k+1) S_{k+1} S_{m-k} + \sum_{k=0}^{\frac{m-1}{2}} (2k+1) S_{2k+1} S_{m-2k-1} \right) \quad \forall m \geq 2, \quad (4)$$

with initial condition  $S_0 = 0$  and  $S_1 = n$ . Comparing (3) with (4) we deduce that the sequence  $(S_m)_{m \geq 0}$  grows faster than the sequence  $(R_m)_{m \geq 0}$  if  $S_m \geq R_m$  for some  $m$ . But indeed  $S_3 = 3n^3/2 + n^2/2$  and  $R_3 = 3n^3/2 + 9e^{-2}n^2/4 + 9e^{-4}n/8 + 3e^{-6}/16$ , thus  $S_3 \geq R_3$  (for all  $n \geq 1$ ). Hence, we get  $P_m \geq S_m \geq R_m$  for  $m \geq 3$ .  $\square$

Let us now turn to the generating function  $G_y(z)$  that enumerates the trees used for the valid  $y$ -premise expansions. Recall that those trees have a root with three children, one being a single leaf  $y$  and the two other being both of size larger than  $n^2$ .

**Lemma 4.** *There exists a constant  $\gamma > 0$  such that, for all (fixed) integer  $r \geq 0$ ,*

$$\lim_{m \rightarrow +\infty} \frac{[z^{m-r}]G_y(z)}{[z^m]P(z)} \geq \gamma \eta^{r+2}.$$

*Proof.* The generating function  $G_y(z)$  is given by

$$G_y(z) = nz^2 \frac{1}{2} (G(z)^2 + G(z^2)) \text{ where } G(z) = \sum_{m \geq n^2} P_m z^m,$$

the integer  $P_m$  being the coefficient of the generating function  $P(z)$  of all implicational trees. Therefore,  $G_y(z)$  has the same dominant singularity  $\eta$  as  $P(z)$  and it is also of square-root type, which implies that

$$\lim_{m \rightarrow +\infty} \frac{[z^m]G_y(z)}{[z^m]P(z)} = \lim_{z \rightarrow \eta} \frac{G'_y(z)}{P'(z)} = \frac{n\eta^2}{2} 2G(\eta) \lim_{z \rightarrow \eta} \frac{G'(z)}{P'(z)} = n\eta^2 G(\eta),$$

since  $\lim_{z \rightarrow \eta} \frac{G'(z)}{P'(z)} = \lim_{m \rightarrow +\infty} \frac{[z^m]G(z)}{[z^m]P(z)} = 1$ . We thus have to estimate

$$G(\eta) = \sum_{m \geq n^2} P_m \eta^m \geq \sum_{m=n^2}^{2n^2} P_m \eta^m.$$

Using Lemma 3, there exists a constant  $\tilde{\gamma}$  such that for large enough  $n$ , and for  $m \in \{n^2, n^2 + 1, \dots, 2n^2\}$ :

$$P_m \geq R_m \geq \frac{\tilde{\gamma}\eta^{-m}}{\sqrt{m^3}}, \quad \text{and } \tilde{\gamma} \leq \frac{1}{2\sqrt{2\pi}} \left(1 - \frac{1}{12n}\right) \cdot \exp\left(-2 - \frac{1}{2e^2} - \frac{1}{e^2n}\right).$$

Thus, using Euler-McLaurin's formula, we deduce there exists a constant  $\gamma = (2 - \sqrt{2}) \cdot \tilde{\gamma}$  such that:

$$G(\eta) \geq \tilde{\gamma} \sum_{m=n^2}^{2n^2} m^{-\frac{3}{2}} \geq \frac{\gamma}{n}.$$

Therefore

$$\lim_{m \rightarrow +\infty} \frac{[z^m]G_y(z)}{[z^m]P(z)} \geq \gamma\eta^2.$$

And thus, using a transfer theorem [1, Chapter IV], the statement is proved.  $\square$

In view of Lemma 4, using a direct lower bound based on Equation (1), we get

$$\sum_{f \mid L(f) \leq n^2} \mathbb{P}_n(f) \geq \sum_{r=\frac{n^2}{2}}^{n^2} \sum_{y \in \{x_1, \dots, x_n\}} \gamma \eta^{r+2} [z^r] \partial_u \Delta(1, z). \quad (5)$$

**Lemma 5.**

$$\partial_u \Delta(1, z) = \frac{(n-1)P(z)}{n - (n-1)P(z)} (S_2(z) - S_1(z)) + \frac{zP'(z)}{n - (n-1)P(z)}, \quad (6)$$

where  $S_1(z) = \sum_{i \geq 2} \partial_u V(1, z^i)$  and  $S_2(z) = \sum_{i \geq 2} \partial_u T(1, z^i)$ .

*Proof.* In order to study  $\partial_u \Delta(1, z)$  we must establish the functional equations satisfied by  $T$  and  $U$ . The derivation is the same as in the paper [2]. First,

$$T(u, z) = (n-1)z \exp\left(\sum_{i \geq 1} \frac{T(u^i, z^i)}{i}\right) + uz \exp\left(\sum_{i \geq 1} \frac{T(u^i, u^i z^i)}{i}\right).$$

Since  $T(1, z) = P(z)$ , we thus deduce

$$\partial_u T(1, z) = z \exp\left(\sum_{i \geq 1} \frac{P(z^i)}{i}\right) \left( (n-1) \sum_{i \geq 1} \partial_u T(1, z^i) + 1 + \sum_{i \geq 1} \partial_u T(1, z^i) + \sum_{i \geq 1} P'(z^i) z^i \right).$$

In view of [3, page 5, first display], we have

$$\sum_{i \geq 1} P'(z^i) z^i = \frac{zP'(z)}{P(z)} - 1,$$

and using the functional equation satisfied by  $P$ , we get

$$\partial_u T(1, z) = \frac{P(z)}{n} \left( n \partial_u T(1, z) + n S_2(z) + \frac{zP'(z)}{P(z)} \right),$$

where  $S_2(z) = \sum_{i \geq 2} \partial_u T(1, z^i)$ . Finally,

$$\partial_u T(1, z) = \frac{P(z)}{1 - P(z)} \left( S_2(z) + \frac{zP'(z)}{nP(z)} \right).$$

Secondly,

$$V(u, z) = (n-1)z \exp \left( \sum_{i \geq 1} \frac{V(u^i, z^i)}{i} + z \exp \left( \sum_{i \geq 1} \frac{T(u^i, z^i)}{i} \right) \right),$$

which implies, after similar calculations as for  $T(u, z)$ ,

$$\partial_u V(1, z) = \frac{P(z)}{n - (n-1)P(z)} ((n-1)S_1(z) + S_2(z) + \partial_u T(1, z)),$$

where  $S_1(z) = \sum_{i \geq 2} \partial_u V(1, z^i)$ . Finally,

$$\partial_u \Delta(1, z) = \frac{(n-1)P(z)}{n - (n-1)P(z)} (S_2(z) - S_1(z)) + \frac{zP'(z)}{n - (n-1)P(z)}.$$

□

In order to complete the proof, we will derive a lower bound for the  $r$ -th coefficient of  $\partial_u \Delta(1, z)$ . Let us first note that the  $r$ -th coefficient of  $S_2(z) - S_1(z)$  is positive (for all positive  $r$ ). Thus, using Lemma 5, we obtain  $[z^r] \frac{zP'(z)}{n - (n-1)P(z)} \leq [z^r] \partial_u \Delta(1, z)$ .

**Lemma 6.** *Asymptotically when  $n$  tends to infinity, if  $r = \Theta(n^2)$ , then*

$$\frac{1}{n} [z^{r-1}] \frac{R'(z)}{1 - \frac{n-1}{n} R(z)} = \Omega \left( \frac{\eta^{-r}}{n} \right).$$

*Proof.* Set

$$\sigma_r = \frac{1}{n} [z^r] \sum_{i=2}^n \left( 1 - \frac{1}{n} \right)^i R'(z) \cdot R(z)^i.$$

Obviously, we get the next lower bound:

$$\frac{1}{n} [z^{r-1}] \frac{R'(z)}{1 - \frac{n-1}{n} R(z)} \geq \sigma_{r-1}.$$

Using the functional equation of  $R(z)$  or the recurrence for its coefficients (*cf.* Eq. (2) and (3)) in the proof of Lemma 3) it is easy to see that, for all  $i \geq 2$ ,

$$[z^{r-1}] R'(z) \cdot R(z)^i = (r-1)R_r - [z^{r-1}] \sum_{k=2}^i R(z)^k.$$

Consider the case where  $i \leq n$  and  $r = \Theta(n^2)$ , when  $n$  tends to infinity. We will show that the second term of the r.-h. side is negligible: First observe that  $(r-1)R_r = \Theta(r^{-1/2}\eta^{-r})$  (by Lemma 3).

Second, let  $k \in \{2, \dots, n\}$ . Using Lagrange inversion (see for example Eq. (14) of [1, p. 732]) yields

$$\begin{aligned} \sum_{k=2}^i [z^{r-1}] R(z)^k &= \sum_{k=2}^i \frac{1}{r-1} [R^{r-2}] k \left( n + \frac{1}{2e^2} \right)^{r-1} R^{k-1} \exp((r-1)R) \\ &\leq \frac{i}{r-1} \left( n + \frac{1}{2e^2} \right)^{r-1} \sum_{k=2}^i \frac{(r-1)^{r-k-1}}{(r-k-1)!} \\ &\leq \frac{i^2}{r-1} \left( n + \frac{1}{2e^2} \right)^{r-1} \frac{(r-1)^{r-3}}{(r-3)!} \leq i^2 \left( n + \frac{1}{2e^2} \right)^{r-1} \frac{(r-1)^{r-2}}{(r-1)!}, \end{aligned}$$

because the sequence  $(x^k/k!)_k$  is increasing while  $k \leq x$ . Thus, using Stirling's formula [1, p. 407] and Lemma 1, we conclude, for  $r = \Theta(n^2)$ , that  $\sum_{k=2}^i [z^{r-1}] R(z)^k = O(r^{-1}\eta^{-r})$ .

Consequently,  $\sigma_{r-1} = \Omega(n^{-1}\eta^{-r})$ .  $\square$

Using the previous lemma and Eq (5) we conclude that  $\sum_{f \mid L(f) \leq n^2} \mathbb{P}_n(f) = \Omega(1)$  as  $n$  tends to infinity, and thus [3, Theorem 3] is proved.

## References

- [1] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge U.P., Cambridge, 2009.
- [2] A. Genitrini and B. Gittenberger. No Shannon effect on probability distributions on Boolean functions induced by random expressions. In *21st International Meeting on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms*, Vienna, Austria, July 2010. DMTCS Proceedings.
- [3] A. Genitrini, B. Gittenberger, V. Kraus, and C. Mailler. Probabilities of boolean functions given by random implicational formulas. *Electronic Journal of Combinatorics*, 19(2):P37, 20 pages, (electronic), 2012.