The IAC Probability of a Divided Verdict in a Simple U.S. Presidential Type Election

Olivier de Mouzon* Thibault Laurent† Michel Le Breton‡
Dominique Lepelley§

May 2016

Abstract

In this article we investigate the asymptotic behavior of the probability $\phi(n)$ of a divided verdict in a U.S. presidential type election in the simplest possible unbiased setting where the phenomenon can occur: three equipopulated districts. The novelty of this paper, in contrast to all the existing literature, is to assume that votes are drawn result from a full-fleshed *IAC* (Impartial Anonymous Culture) probability model. Through the use of numerical methods, it is conjectured that $\sqrt{n}\phi(n)$ converges to 0.131 when n (the size of the electorate in one district) tends to infinity. It is also demonstrated that for all i in $\mathbb{N}$, $\text{LimSup}\frac{\sqrt{n}}{ln^i(n)^{1.5}}\phi(n)$ and $\text{LimInf}\sqrt{n}\phi(n)$ are finite numbers, where $\ln^i(n)$ is recursively defined by $\ln^0(n) = n$ and for all j in $\mathbb{N}^*$, $\ln^j(n) = \ln^{j-1}(\ln(n))$.

Classification JEL: D71, D72.

Key Words: Electoral system, Election Inversions, Impartial Anonymous Culture

*Toulouse School of Economics(INRA).
†Toulouse School of Economics(CNRS).
‡Institut Universitaire de France and Toulouse School of Economics, University of Toulouse Capitole, France.
§CEMOI, Université de La Réunion, France.

1 Introduction

As described by Miller (2012a): “An election inversion occurs when the candidate (or party) that wins the most votes from the nationwide electorate fails to win the most electoral votes (or parliamentary seats) and therefore loses the election. To describe this phenomenon, public commentary commonly uses such terms as ‘reversal of winner’, ‘unpopular winner’, ‘divided verdict’ while the academic literature on voting and social choice uses such terms as ‘compound majority paradox’ and ‘referendum paradox’. Election inversions can occur under U.S. Electoral College[1] or any two-tier electoral system. Interestingly enough, when there are more than two candidates (or parties) election inversions in the allocation of parliamentary seats can be defined as a lack of monotonicity: party A may receive more votes but less seats than party B nationwide”.

The President of the United States is elected, not by a direct national popular vote, but by an indirect Electoral College system in which (in almost universal practice since the 1830s) separate state popular votes are aggregated by adding up state electoral votes awarded, on a winner-take-all basis, to the plurality winner in each state. Each state has electoral votes equal in number to its total representation in Congress and since 1964 the District of Columbia has three electoral votes. Therefore the U.S. Electoral College is a two-tier electoral system: individual voters cast votes in the first tier to choose between rival slates of ‘Presidential electors’ pledged to one or other Presidential candidate, and the winning elector slates then cast blocs of electoral votes for the candidate to whom they are pledged in the second tier. At the present time, there are 538 electoral votes, so 270 are required for election and a 269-269 electoral vote tie is possible. As is well-known, the Electoral College has produced a ‘wrong winner’ in the 2000 presidential election, and it has done so twice before[2]. The dissatisfaction with the discrepancies between the Electoral College outcome and the popular vote that occur from time to time has spawned a movement in support of The National Popular Vote Bill [http://www.nationalpopularvote.com], which would effectively guarantee the Presidency to the winner of the national popular vote once enough states adopt the bill to guarantee victory for the national popular winner. In that case the adopting states then award all their electors to the national popular winner. As of May 2011 the bill had been adopted in nine states representing 49% of the

[1]Miller (2012b) contains an insightful presentation of the Electoral College that he qualifies as " a terrific boon for political science (and public choice) research (and teaching)".

[2]As pointed out by Neubauer, Schilling and Zeitllin (2012): “ (...) for the 58 presidential elections that have taken place since 1780, at least four unpopular elections—nearly 7%—have occurred. If we include the 1960 election, fully 8.6% of all American presidential elections have been unpopular”. But these authors also note that “estimates based on the historical record do not necessarily provide good predictions of the likelihood of an unpopular result in a future election because the sample size involved in these estimates is small, and because the political structure has changed dramatically many times in United States history. A more effective way to estimate the likelihood of a future unpopular election is to use the Monte Carlo method: simulate a large number of elections using a suitable randomization approach and then compute the percentage that result in unpopularly elected presidents, using as input data an appropriate set of recent presidential elections”.

electoral votes needed to activate it. Gallup polls dating as far back as 1944 have consistently shown only about 20% of the public supporting the current electoral system and about 70% opposed. Recent surveys have found more than 70% favor a direct nationwide election of the President[3]. While the public's feelings about the appropriateness of the Electoral College system may have been intensified by recent elections, instances where the winner of the Presidency did not win at least a plurality of the popular vote are in fact common in the history of the United States.

As also noted by Miller (2012a), 'Westminster' single-member-district parliamentary systems (e.g., the U.K., Canada, Australia, India, and New Zealand prior to 1993) are likewise two-tier voting systems, and they produce election inversions at least as frequently as the U.S. Electoral College. While most of these elections were very close with respects to both votes and seats, the case of Canada in 1979 shows that this is not invariably the case. These parliamentary systems differ in several important respects from the U.S. Electoral College. Among the main differences, 'Westminster' systems have uniform districts — that is, the districts have equal weight (namely a single parliamentary seat), reflecting (approximately) equal populations and/or numbers of voters. In contrast, Electoral College 'districts' (i.e., states) are (highly) unequal in both population and voters and are likewise unequally weighted, at present ranging from 3 to 55 electoral votes.

The use of the terminology 'paradoxical' to qualify this phenomenon alludes to the fact that the phenomenon is somehow unexpected . Nurmi (1999) and Laffond and Laine (2000) have addressed the general phenomenon of election inversions in social choice terms, and Chambers (2008) has demonstrated (in effect) that no neutral (between candidates or parties) two-tier electoral rule can satisfy "representative consistency," i.e., preclude election inversions. However, as pointed out by Miller (2012a): "the likelihood of inversions and the factors that produce them are less apparent and there has been considerable confusion about the circumstances under which election inversions occur. For example, the susceptibility of the Electoral College to inversions is often blamed on the small-state bias in the apportionment of electoral votes and/or the 'non-proportional' or 'winner-take-all' manner of casting state electoral votes, but neither of these attributes of the Electoral College is necessary for inversions to occur".

The pervasiveness of the 'Election Inversions paradox'[4] raises the following question. Given a division of the country into a number of (possibly unequal) districts, a number of electoral seats for each of those districts and a probability distribution of the L/R preferences among the voters, what is

[3]The fact that Electoral College can produce inversions is regularly invoked by critics of the Electoral College (e.g. Abbott and Levine, 1991) and is therefore addressed by its defenders as well (e.g., Best, 1971).

[4]In this paper, we will focus on a particular type of elections inversions where only two parties compete. Kurrild-Klitgaard (2013) has shown that proportional representation can produce 'election inversions' such that a coalition of parties collectively supported by a majority of voters fails to win a majority of parliamentary seats, and he identifies several empirical examples under the Danish electoral system. Miller (2014)analysis moves a step further by showing that election inversions can occur even under the purest type of proportional representation — namely, one with (i) a single national constituency, (ii) no explicit seat threshold, and (iii) a highly proportional electoral formula.

the probability that this event will occur ? In this question, the first two inputs are fixed and only the third input remains to be discussed. We can approach this dimension in at least two different ways. A first one, purely theoretical, consists in considering an abstract probability distribution on the set of preference profiles. A second one more empirical consists in estimating the probability distribution on the basis of a statistical model and the electoral data[5].

To the best of our knowledge, the first theoretical work on election inversions was by May (1948), who attempted to calculate the a priori frequency of inversions based on a particular theoretical probability model of election outcomes that we call IAC^* later in the paper. May limits his investigation to the case of an odd number of equipopulated districts (also with an odd number of voters per district) and one seat per district. He proves among other things that, in such case, the probability of election inversions increases slowly with the number of districts and the number of voters per district and tends to the limit $\frac{1}{6} \simeq 16.7\%$. Feix, Lepelley, Merlin and Rouet (2004, 2011) also consider the setting of equipopulated districts but do not limit their investigation to the odd case. They rediscovered some of May's results and also proceeded to an evaluation of the frequency in another theoretical model recorded in the social choice literature under the heading *IC*. The difference between IAC^* and *IC* is that in IAC^*, the preferences of the voters are unbiased, identically distributed but display some correlation among voters belonging to the same district (and only among those) while in *IC*, the preferences of the voters are independent, unbiased and identically distributed. In the *IC* model, they compute the probability for the case of 3, 4 and 5 districts and show through simulations that the probability of election inversions increases with the number of districts and tend to 20.5% when the number of districts tend to infinity. This question is already addressed by Hinich, Mickelsen and Ordeshook (1975) who consider a model with equipopulated districts but instead of probabilities on the individual votes, they consider probabilities on the proportions of left voters in each district. The probability distribution of the proportion in each district is assumed to be in the beta family and the proportions are independent across districts; therefore, their model is (in some sense) a generalization of both IAC^* and *IC*. They obtain many interesting results. In particular, in the unbiased case, they show that the probability of election inversions tends to $\frac{\arccos 2(2\pi)^{-\frac{1}{2}}}{\pi} \simeq 20.595\%$ when the number of districts and the beta parameter tends to infinity[6]. In the case of the US electoral college, Erickson and Sigman (2000) obtain through simulations in the case of the IAC^* model a probability equal to 8.1%.

Empirically based estimates of the expected frequency of Electoral College election inversions have been provided by Ball and Leuthold (1991), May (1958), Merrill (1978), Neubauer, Schilling

[5]There is also an interesting literature to evaluate how several features of the electoral setting (for instance the size of the house or the number of electoral votes per district) affect the probability of election inversions (Barthélémy, Martin and Piggins (2014), Miller (2014a) and Neubauer and Zeitlin (2003)).

[6]Saying that the beta parameter of the distribution tends to $+\infty$ amounts to say that the probability that the election is tied in any given district tends to 1 which is what postlates the IC model.

and Zeitlin (2012). As reported before, the empirical frequency based on US presidential elections since 1828, leads to 6% or 8.7% depending whether we count or not the 1860 special case. When the count is limited to elections in which the winner's popular vote margin was no greater than about 3% [7], the frequency climbs to 25%. Merrill (1978) estimates a probability model (which also contains *IAC* and *IC* as special cases) which opens the door to the computation of the probability of election inversions conditional on the national margin of victory of the top candidate or national popular vote. For instance, he obtains a probability equal to 23% for the 1976 election. He also derives[8] this conditional probability quite generally on the basis of a statistical model estimated from the data based period 1900-1976. The probabilities that he obtains are relatively high (larger than 50%) when the popular vote is tied. Like Merrill, Neubauer, Schilling and Zeitlin (2012) call attention on the importance of not using exclusively state marginal distributions and on the necessity to use the joint distribution. Indeed as they point out: "Actual election data is not independent from one state to another, but in fact is highly correlated; for example, the southern states east of Texas often vote in much the same way and in so doing have had a strong voice in determining the president. Other, less geographically connected states also have highly correlated voting patterns. Over half of the state-to-state correlations for the 1964-2008 election records are above 0.70; this reflects the substantial dependency that exists among state voting patterns. To adequately simulate presidential elections, this dependency must be take into account". To do so, they incorporate the historical correlation between states' voting patterns by means of principal components analysis and estimate the frequency of unpopular elections to be around 4.9%.

As already pointed out, while made popular by the 2000 US presidential election, election inversions can arise also in committee or parliamentary settings as soon as elections involve some districting of the population of voters. Based on a data consisting of 278 elections in the US and in the UK, May (1958) reports 23 elections victories without pluralities leading to an observed frequency around 10%. Lahrach and Merlin (2010) have done related work with respect to French local government elections where they also obtain a frequency around 10%.

The main objective of this paper is to fill a gap in the theoretical approach[9]. Indeed, while *IAC* is considered as being one of the two most popular probability models in social choice and power measurement, to the best of our knowledge, the evaluation of the probability of election inversions in that case has not been performed so far. The main difference between *IAC* and *IAC** is the introduction

[7]Only twelve presidential elections meet that condition.

[8]See his figure 1.

[9]While this paper is about election inversions, the same apparatus could be used to estimate the votes/seats relationship (see e.g. Tufte (1973), Wilgen and Engstrom (1980)) and the likelihood of any event like 'The left won x% of the popular vote and y% of the seats/electoral votes' where both x and y are numbers in $[0,1]$. If you divide the unit square $[0,1]^2$ into the four squares $\left[0,\frac{1}{2}\right]\times\left[0,\frac{1}{2}\right]$, $\left[0,\frac{1}{2}\right]\times\left[\frac{1}{2},1\right]$, $\left[\frac{1}{2},1\right]\times\left[0,\frac{1}{2}\right]$ and $\left[\frac{1}{2},1\right]\times\left[\frac{1}{2},1\right]$, the probability of election inversions is the probability of the union of the two anti-diagonal squares. Obviously, knowing the totality of the joint distribution on $[0,1]^2$ has some intrinsic value that is discussed in the concluding section.

of vote correlations across districts in addition to the vote correlations within districts. Precisely, with *IAC* the district affiliation is irrelevant since irrespective of their districts, the correlation of the votes of any two voters will be the same. This is of course quite extreme but it is a first step towards introducing more general forms of covariances. Intuitively having positive covariance increases the homogeneity of the electorate and decreases the likelihood of an election inversion. If the correlation was perfect, then the problem of a divided verdict will disappear irrespective of the number of voters. The main contribution of this paper is twofold. First, in section 3, we provide a numerical approach to the evaluation of this probability in the case of three equipopulated districts and *IAC*. We show that when n, the number of voters per district, tends to infinity, this probability behaves as $\frac{0.1309}{\sqrt{n}}$. We have not been able to prove that this is indeed the limit but in section 4, we provide theoretical lower and upper bounds of the probability, which are consistent with this conjecture: The lower bound behaves as $O\left(\frac{1}{\sqrt{n}}\right)$ and the upper bound behaves as $O\left(\frac{ln(n)^3}{\sqrt{n}}\right)$.

The plan of this paper is as follows. In section 2, we present the problem of a divided verdict in the simplest conceivable setting and the main notations used in the paper, a brief state of the art and a statement of the precise problem studied in this manuscript. Then, in section 3 we develop a numerical approach of that problem and discuss the behavior of these estimations together with two bounds as n the number of voters get very large. Finally, in section 4, we complement the numerical study by two analytic results which reinforce the numerical discoveries upon the speed of convergence. We conclude by formulating some opening questions and avenues of research.

2 The Model: Notations and State of the Art

2.1 Notations and Assumptions

In this paper we explore the simplest non trivial two-tier system. A territory (for instance a country) is divided into three equipopulated[10] districts. Each district elects one representative. We can either see the outcome of these elections as describing the election of the representatives of a parliament or as describing a toy symmetric version of the US electoral college where these three representatives elect the head of the executive of the country. We denote by n the number of voters in each district; n is assumed to be odd. The total number of voters is then $3n$. In each election, the voter can chose among two candidates: a left candidate L and a right candidate R. A profile of preferences is a vector X in $\{L,R\}^{3n}$: X_i denotes the preference of voter i.

Ex ante uncertainty is described here by a probability model λ: $\lambda(X)$ denotes the probability of the event X. The following three probability models are the more popular.

The *IC* and *IAC* models are used in many positive models of voting and evaluation of paradoxes

[10]So malapportionment is excluded from the scope of our analysis.

in social choice theory. They underly respectively the traditional Banzhaf's (1962, 1965,1968) and Shapley-Shubik (1954)' measures of power[11]. There are also extensively used in the social choice literature (see e.g. Fishburn and Gehrlein (1976), Kuga and Nagatani (1974)). The IAC^* model has been introduced for the first time by May (1948) and has been (re)discovered since by many authors including Chamberlain and Rothschild (1981), Good and Mayer (1975), Feix, Lepelley, Merlin and Rouet (2004), Le Breton and Lepelley[12] (2014) and Le Breton, Lepelley and Smaoui (2016) to cite few. Precisely, these three probability models are defined as follows.

IC (Impartial Culture): $\lambda(X) = \frac{1}{2^{3n}}$ for all $X \in \{L,R\}^{3n}$. According to *IC*, voters vote independently of each other with an equal probability of voting left or right. The probability of the event 'k voters vote left' is then equal to $\binom{3n}{k}\frac{1}{2^{3n}}$.

IAC (Impartial Anonymous Culture). The probability λ is defined as follows. We first draw uniformly the parameter p in the interval $[0,1]$. Then, conditional on the draw of p, the variables X_i are independent and identically distributed according to a Bernoulli of parameter p. Therefore, if the vector X has k coordinates X_i such that $X_i = L$ (and therefore $n-k$ coordinates X_i such that $X_i = R$), $\lambda(X) = \int_0^1 p^k(1-p)^{3n-k}dp$. Since, from the standard equality defining the Beta distribution,

$$\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}, \tag{1}$$

where α and β are positive parameters and Γ is the gamma function[13], we deduce: $\lambda(X) = \frac{k!(3n-k)!}{(3n+1)!}$. Now, the probability of the event 'k voters vote left' is equal to $\binom{3n}{k}\frac{k!(3n-k)!}{3n+1)!} = \frac{1}{3n+1}$

IAC^* (District Based Impartial Anonymous Culture). We first draw uniformly and independently in the interval $[0,1]$ three parameters p_1, p_2 and p_3. Then, conditional on the draw of (p_1, p_2, p_3), the variables X_i are independent and identically distributed according to a Bernoulli of parameter p_j in each district $j = 1,2,3$. Therefore, if the vector X has k_j coordinates X_i such that $X_i = L$ (and $n-k_j$ coordinates X_i such that $X_i = R$) in district $j = 1,2,3$,

$$\lambda(X) = \left[\int_0^1 p^{k_1}(1-p)^{n-k_1}dp\right]\left[\int_0^1 p^{k_2}(1-p)^{n-k_2}dp\right]\left[\int_0^1 p^{k_3}(1-p)^{n-k_3}dp\right].$$

From (1), we deduce:

$$\lambda(X) = \frac{k_1!(n-k_1)!}{(n+1)!k_1!(n-k_1)!}\frac{k_2!(n-k_2)!}{(n+1)!}\frac{k_3!(n-k_2)!}{(n+1)!}$$

The computation of the probability of the event 'k voters vote left' is more subtle. In contrast, the probability of the event 'k_1 voters of district 1 vote left, k_2 voters of district 2 vote left and k_3 voters of district 3 vote left' is equal to:

[11] See Straffin (1988) for a very nice united presentation of *IC* and *IAC*.

[12] They coined the notation IAC^* to emphasize the distinction between *IAC* and this new model.

[13] In particular, $\Gamma(n) = (n-1)!$ if n is a positive integer.

$$\binom{n}{k_1}\frac{k_1!(n-k_1)!}{(n+1)!}\binom{n}{k_2}\frac{k_2!(n-k_3)!}{(n+1)!}\binom{n}{k_3}\frac{k_3!(n-k_3)!}{(n+1)!}=\frac{1}{(n+1)^3}$$

The analysis will focus on the following two-tier majority mechanism:

$$Maj_2(X) = Maj_1\left(Maj_1\left(X^1\right), Maj_1\left(X^2\right), Maj_1\left(X^3\right)\right),$$

where for any vector Y in $\{L,R\}^m$ (where m is an arbitrary odd integer[14]) $Maj_1(Y)$ describes the outcome resulting from the popular vote in this population of m voters i.e. $Maj_1(Y) = L$ iff a majority of the m voters vote left; in each district $j = 1,2,3$, X^j denotes the sub-profile of preferences of voters from district j. As already pointed out Maj_2 can receive two possible interpretations. In the case of the parliamentary interpretation, $Maj_1\left(X^1\right)$, $Maj_1\left(X^2\right)$ and $Maj_1\left(X^3\right)$ are the elected representatives and $Maj_2(X)$ denotes the majority color of an elected chamber of representatives: the color is left iff the majority of representatives is leftist. In the electoral college interpretation, Maj_2 is interpreted as a two-step mechanism to elect a single candidate (say the head of the executive) (either L or R): the winner is the left candidate left iff there is a majority of districts voting left. In contrast to $Maj_2(X)$, $Maj_1(X)$ represents the popular outcome.

2.2 Divided Verdict

The paper focuses on the probability of the event:

$$E \equiv \left\{X \in \{G,D\}^{3n} : Maj_1(X) \neq Maj_2(X)\right\}$$

For any X in E, the president elected through the electoral college does not coincide with the president elected through the popular vote. Scholars have used different names to qualify this situation: inverted elections, referendum Paradox, Divided Verdict, unpopular presidential elections,...

For any $i,j \in \{1,2,3\}, i \neq j$ denote by $E_{ijL}\left(E_{ijR}\right)$ the event describing the profiles X for which the districts i and j vote left (right) but the majority of the $3n$ voters vote right (left). If the probability model is symmetric and neutral, then the six events $E_{12L}, E_{13L}, E_{23L}, E_{12R}, E_{13R}$ and E_{23R} have the same mass. Therefore:

$$\lambda(E) = 6\lambda(E_{12L})$$

Hereafter, we will denote simply by F the event E_{12L}. Let us report what is known on the calculation of $\lambda(E)$ when $\lambda = IAC^*$ and $\lambda = IC$. In the general case we have :

[14] In this paper, we dont need to define the majority mechanism when m is an even integer.

$$\lambda(E)=6\sum_{k=\frac{n+1}{2}}^{n-1}\sum_{l=\frac{n+1}{2}}^{\frac{3n-1}{2}-k}\sum_{r=0}^{\frac{3n-1}{2}-k-l}\lambda(k,l,r)$$

where $\lambda(k,l,r)$ denotes the probability that k voters from district 1votes, l voters from district 2 vote left and r voters from district 3 vote left. In the case where $\lambda = IAC^*$, we obtain specifically:

$$\begin{aligned}\lambda(E) &= 6\sum_{k=\frac{n+1}{2}}^{n-1}\sum_{l=\frac{n+1}{2}}^{\frac{3n-1}{2}-k}\sum_{r=0}^{\frac{3n-1}{2}-k-l}\binom{n}{k}\frac{k!(n-k)!}{(n+1)!k!(n-k)!}\binom{n}{l}\frac{l!(n-l)!}{(n+1)!}\binom{n}{r}\frac{r!(n-r)!}{(n+1)!}\\ &= \frac{6}{(n+1)^3}\sum_{k=\frac{n+1}{2}}^{n-1}\sum_{l=\frac{n+1}{2}}^{\frac{3n-1}{2}-k}\sum_{r=0}^{\frac{3n-1}{2}-k-l} = \frac{6}{(n+1)^3}\left(\frac{1}{48}n^3+\frac{3}{48}n^2-\frac{1}{48}n-\frac{3}{48}\right)\\ &= \frac{1}{8}\frac{n^3+3n^2-n-3}{n^3+3n^2+3n+1} = \frac{(n+3)(n-1)(n+1)}{8(n+1)^3} = \frac{n^2+2n-3}{8(n+1)^2}\end{aligned}$$

We deduce from above that $\lambda(E)$ tends to $\frac{1}{8} = 12.5\%$ when $n \to \infty$. Let us now consider the case where $\lambda = IC$. We obtain:

$$\lambda(E)=\frac{6}{2^{3n}}\sum_{k=\frac{n+1}{2}}^{n-1}\sum_{l=\frac{n+1}{2}}^{\frac{3n-1}{2}-k}\sum_{r=0}^{\frac{3n-1}{2}-k-l}\binom{n}{k}\binom{n}{l}\binom{n}{r}$$

The analysis of the asymptotic behavior of this triple sum proceeds from the following probablity argument: Denote S^j for $j=1,2,3$, the random variable $\sum_{i=1}^{n} X_i^j$ where $X_i^j = 1$ iff voter i in district j votes left and $X_i^j = 0$ if voter i votes right: S^j is therefore the number of voters who vote left in district j. Since S^j is the sum of independent and identically distributed Bernoulli random variables of parameter $\frac{1}{2}$, we deduce from the central limit theorem that if n is large the law of $\frac{S^j-\frac{n}{2}}{\frac{\sqrt{n}}{2}}$ is approximatively the law of a unit gaussian $N(0,1)$. Similarly, the law of $\frac{S-\frac{3n}{2}}{\frac{\sqrt{3n}}{2}}$ where $S = S^1+S^2+S^3$ is also approximatively the law of a unit gaussian. Since the event F is described by the inequalities:

$$\begin{aligned}S^1 &> \frac{n}{2}\\ S^2 &> \frac{n}{2}\\ S &< \frac{3n}{2}\end{aligned}$$

we deduce that when n is large, $\lambda(F)$ is approximatively the probability mass $\mu(F)$ of the area of $\mathbb{R}^3$ described by the inequalities:

$$
\begin{aligned}
Z^1 &> 0 \\
Z^2 &> 0 \\
Z^1+Z^2+Z^3 &< 0
\end{aligned}
$$

where $Z=(Z_1,Z_2,Z_3)$ is a gaussian vector with mean vector equal to 0 and a variances-covariances matrix equal to identity. The exact value of $\mu(F)$ has been derived by Feix, Lepelley, Merlin and Rouet (2004):

$$\mu(F)=\frac{Arc\cos(\frac{\sqrt{3}}{3})}{2\pi}-\frac{1}{8}$$

from which we deduce that $\lambda(E)$ is approximatively equal to[15]: 16.226%.

2.3 The Probability of a Divided Verdict for *IAC*

The main purpose of this paper is to complete the above picture by offering a computation of $\lambda(E)=6\lambda(F)$ when λ is the *IAC* probability model. Under *IAC*, we obtain

$$
\begin{aligned}
\lambda(F) &= \int_0^1 \sum_{k=\frac{n+1}{2}}^{n-1} \sum_{l=\frac{n+1}{2}}^{\frac{3n-1}{2}-k} \sum_{r=0}^{\frac{3n-1}{2}-k-l} \binom{n}{k}\binom{n}{l}\binom{n}{r} p^k(1-p)^{n-k}p^l(1-p)^{n-l}p^r(1-p)^{n-r}dp \\
&= \int_0^1 \sum_{k=\frac{n+1}{2}}^{n-1} \sum_{l=\frac{n+1}{2}}^{\frac{3n-1}{2}-k} \sum_{r=0}^{\frac{3n-1}{2}-k-l} \binom{n}{k}\binom{n}{l}\binom{n}{r} p^{k+l+r}(1-p)^{3n-k-l-r}dp
\end{aligned}
$$

which by using (1) simplifies to:

$$\lambda(F)=\sum_{k=\frac{n+1}{2}}^{n-1} \sum_{l=\frac{n+1}{2}}^{\frac{3n-1}{2}-k} \sum_{r=0}^{\frac{3n-1}{2}-k-l} \binom{n}{k}\binom{n}{l}\binom{n}{r} \frac{(k+l+r)!(3n-k-l-r)!}{(3n+1)!} \tag{2}$$

In what follows, we will denote $\phi(n)$ the right hand-side of (2)

[15]We could use instead Monte-Carlo simulations to calculate the probability of the event. One idea consists in simulating long sequences $t=1,...T$ of random draws of a three dimensional vector Z distributed as $N(0,I)$. To each realization t of Z, we attach the corresponding realization of the multinomial random variable D defined by the following mutually exclusive modalities: the left wins 0 district and the popular vote, the left wins 1 district and the popular vote, the left wins 2 districts and the popular vote, the left wins three districts and the popular vote, the left wins 0 district and loses the popular vote, the left wins 1 district and loses the popular vote, the left wins 2 districts and loses the popular vote, the left wins three districts and loses the popular vote. The fourth and eighth events are empty and have therefore zero probability. According to the Glivenko-Cantelli theorem, the empirical distribution of D (defined by the eight dimensional vector of empirical frequencies) almost surely converges weakly to the true distribution of D. More on that can be obtained from the authors upon request.

$$\phi(n) = 6 \sum_{k=\frac{n+1}{2}}^{n-1} \sum_{l=\frac{n+1}{2}}^{\frac{3n-1}{2}-k} \sum_{r=0}^{\frac{3n-1}{2}-k-l} \binom{n}{k}\binom{n}{l}\binom{n}{r} \frac{(k+l+r)!(3n-k-l-r)!}{(3n+1)!}$$

where n is an odd integer.

3 The Numerical Analysis of the Function ϕ

The main purpose of this section is to proceed to a numerical approach of the computation of $\phi(n)$ for a large but of course bounded set of values of n. The main goal of that analysis is to infer from these results the general behavior of $\phi(n)$ when n gets arbitrarily large. The numerical analysis will proceed in a sequence of steps that we now explain. First, we will proceed to a rewriting of $\phi(n)$, as factorials in the binomial coefficients lead quickly to very large numbers. The binomial coefficients which appear in $\phi(n)$ have a triple superscript, suggesting that (on top of their intrinsic size) we would have to compute about n^3 of them. To save computing time, we develop (section 3.3) a simplified algorithm which comes together with a lower bound and an upper bound which helps a lot in understanding the asymptotics. To build up this simplified algorithm, we take advantage of the binomial coefficients which show up in the summation. The first simplification of $\phi(n)$ takes advantage of the symmetry and reduces by a factor of 2 the number of computations. But the other properties of the binomial coefficients lead (section 3.4) to very useful lower and upper bounds for which the number of computations decreases from n^3 to n. We then move to an approximation of both the lower and the upper bounds based on the simple idea that values below some given very small threshold ε should be deleted in the sum. This leads consecutively to approximations of the two bounds as well of course to an ultimate approximation $\widetilde{\phi}$ of ϕ, the number of computations and an estimate of the error. The results are reported in section 3.5 where the computations have been performed for many non adjacent values of n ranging from 1 to 10^8 with a very small ε. All numerical results confirm that $\phi(n)$ tends to 0 as $\frac{1}{\sqrt{n}}$. They suggest that $\sqrt{n}\phi(n)$, $\sqrt{n}\widetilde{\phi}(n)$, $\sqrt{n}\phi_{\min}(n)$ and $\sqrt{n}\phi_{maj}(n)$ tend respectively to 0.131, 0.131, 0.061 and 0.317 when n tends to ∞.

3.1 Rewriting of the Function

When n increases, the computation of the binomial coefficients and factorials tend quite rapidly towards infinite values (for instance, $150! \simeq 5.7 \times 10^{262}$). By using the R software, we cannot manipulate numbers beyond 1.797693×10^{308} approximatively. To make the computation feasible for relative large values of n, we first rewrite the function ϕ as follows:

$$\phi(n) = 6 \sum_{k=\frac{n+1}{2}}^{n-1} \sum_{l=\frac{n+1}{2}}^{\frac{3n-1}{2}-k} \sum_{r=0}^{\frac{3n-1}{2}-k-l} B^{k,l,r}$$

with:

$$\begin{aligned} B^{k,l,r} = & \exp(3\ln\Gamma(n+1) + \ln\Gamma(k+l+r+1) + \ln\Gamma(3n-k-l-r+1) \\ & - \ln\Gamma(k+1) - \ln\Gamma(n-k+1) - \ln\Gamma(l+1) - \ln\Gamma(n-l+1) \\ & - \ln\Gamma(r+1) - \ln\Gamma(n-r+1) - \ln\Gamma(3n+1+1)) \end{aligned}$$

3.2 Properties of the Function $\ln\Gamma$

To compute ϕ, we need to know the values of $\ln\Gamma$ for the integers ranging from 1 to $3n+2$.

3.2.1 Algorithm

• For $n < 12$, we have used the following recursivity formula:

$$\ln\Gamma(1) = 0$$
$$\ln\Gamma(n+1) = \ln(n) + \ln\Gamma(n)$$

• For $n \geq 13$, we have used Stirling's formula (see appendix B) which ensures to keep a good machine accuracy even for large values of n:

$$\ln\Gamma(n) = \frac{1}{2} \times \ln(2\pi) + (n - \frac{1}{2}) \times \ln(n) - n + w(n) \tag{3}$$

where:

$$w(n) = \frac{1}{12n}\mathbf{1}_{n<8943979} - \frac{1}{360n^3}\mathbf{1}_{n<1456} + \frac{1}{1260n^5}\mathbf{1}_{n<123} - \frac{1}{1680n^7}\mathbf{1}_{n<37} + \frac{1}{1188n^9}\mathbf{1}_{n<19}$$

3.2.2 Use

We have programmed in C the function `loggamma`. The values of $\ln\Gamma$ are computed before the computation of ϕ and stored in a vector of size $3n+2$ in order to avoid computational repetition in each loop of ϕ. The figure 1 represents $\ln\Gamma$ for all integer values which are used in the computation of ϕ until 100 000 001.

3.3 Simplified Algorithm

The main objective of this subsection is to simplify the way to express ϕ in order to save as much as possible on computing time. In doing so, we will also propose a lower bound and an upper bound of the function ϕ which will lead to a relatively low computing time for large values of n and a good understanding of the asymptotics.

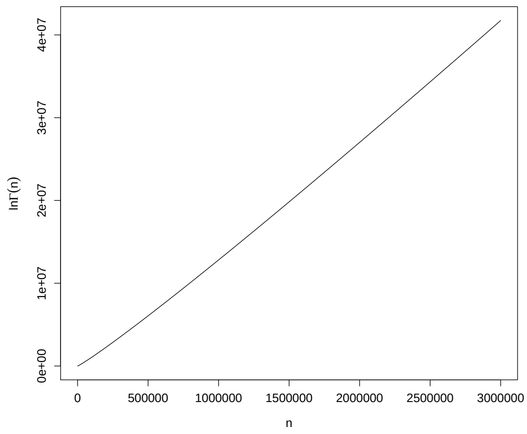

Figure 1: Values of $\ln\Gamma$ for integers.

3.3.1 Initial Number of Computations

The computation of ϕ amounts to the computation of $B^{k,l,r}$ for different values of k, l and r. This number of values, of the order n^3, is equal to:

$$\frac{1}{6}(n_2 - 1)n_2(n_2 + 1), \tag{4}$$

where $n_2 = \frac{(n+1)}{2}$. For instance when $n = 10\ 001$, we would need to calculate the expression $B^{k,l,r}$ 20 845 835 000 times. The objective is to reduce as much as we can this number without losing on the quality of the outcome.

3.3.2 Few Properties of $B^{k,l,r}$

For $n = 13$, we have represented on figure 2 the values of k, l and r for which we calculate $B^{k,l,r}$. In general this coefficient satisfies a number of properties some of them are reported below[16].

P1. $B^{k,l,r} = B^{l,k,r}$

P2. $B^{k,l,r} > B^{k,l,r-1}$ (for all feasible values of k, l, r and $r-1$)

P3. $B^{k,l,r} > B^{k+1,l,r}$ (for all feasible values of k, l, r and $k+1$)

P4. $B^{k,l,r} > B^{k,l+1,r}$ (for all feasible values of k, l, r and $l+1$)

P5. For each fixed value of r, the largest $B^{k,l,r}$ is obtained for $k = l = n_2 (= \frac{n+1}{2})$

P6. $\forall l \geq k, B^{k,l,r} > B^{k-1,l+1,r}$ (for all feasible values of k, $k-1$, l, $l+1$ and r)

P7. For each fixed value of r, the smallest $B^{k,l,r}$ is obtained for $k = n-1-r$ and $l = n_2$

[16] Proofs are provided in appendix 1.

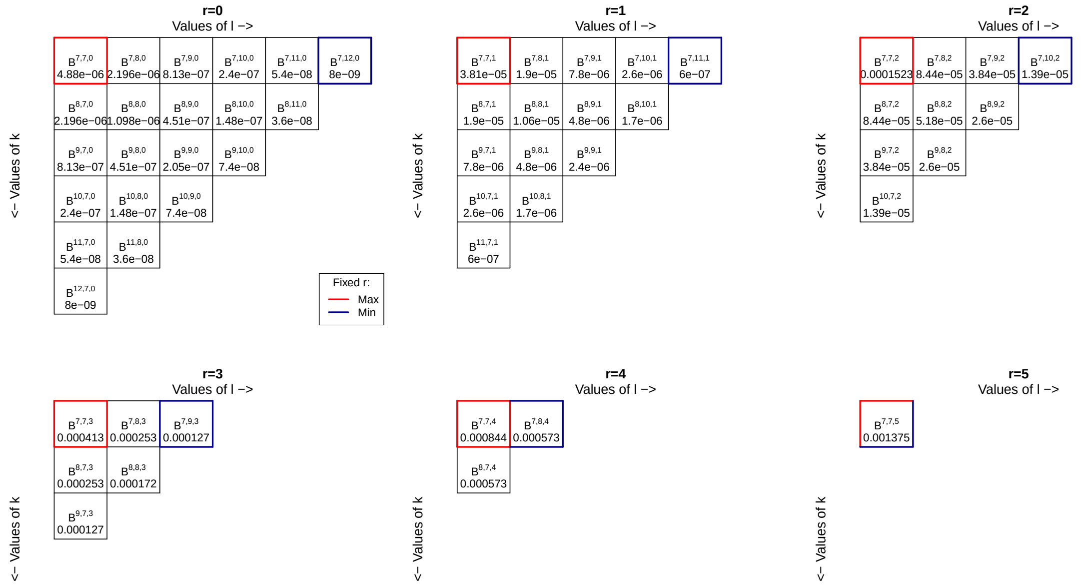

Figure 2: Representation of the $B^{k,l,r}$ used in the computation of $\phi(13)$ without any simplification. For each r, the red cell contains the maximum value and the blue cell the minimum.

Remark: If we pile up the 6 levels of figure 2, we obtain a geometric pattern close to a tetrahedron. When we compute ϕ for the following value of n, we add to the tetrahedron a ground level with a number of new entries to compute equal to $n_2 \times (n_2 - 1)/2$. Note that when n increases, the values $B^{k,l,r}$ will decrease and converge to 0. For instance, on the top $B^{n_2,n_2,n_2-2} = O(n^{-2})$. The values on low levels converge even more rapidly towards 0. For instance, for the ground level ($r = 0$), $B^{n_2,n_2,0} = O(n^{-1.5}\left(\frac{16}{27}\right)^n)$.

3.3.3 Simplification of ϕ

By exploiting the symmetry property (P1), we can rewrite ϕ as follows :

$$\phi(n) = 6 \times [2 \times (\sum_{k=\frac{n+1}{2}}^{n-1} \sum_{l=k+1}^{\frac{3n-1}{2}-k} \sum_{r=0}^{\frac{3n-1}{2}-k-l} B^{k,l,r}) + (\sum_{k=\frac{n+1}{2}}^{n_3} \sum_{r=0}^{\frac{3n-1}{2}-2k} B^{k,k,r})], \tag{5}$$

where $n_3 = \frac{3n-1}{4}\mathbf{1}_{\{n_2 \text{ even}\}} + \frac{3n-3}{4}\mathbf{1}_{\{n_2 \text{ odd}\}}$

For $n = 13$, we have represented on figure 3 the computations which have been save through the use of this symmetry property. The cells on the diagonal (represented in magenta) correspond to the computation on the right-side of equation 5.

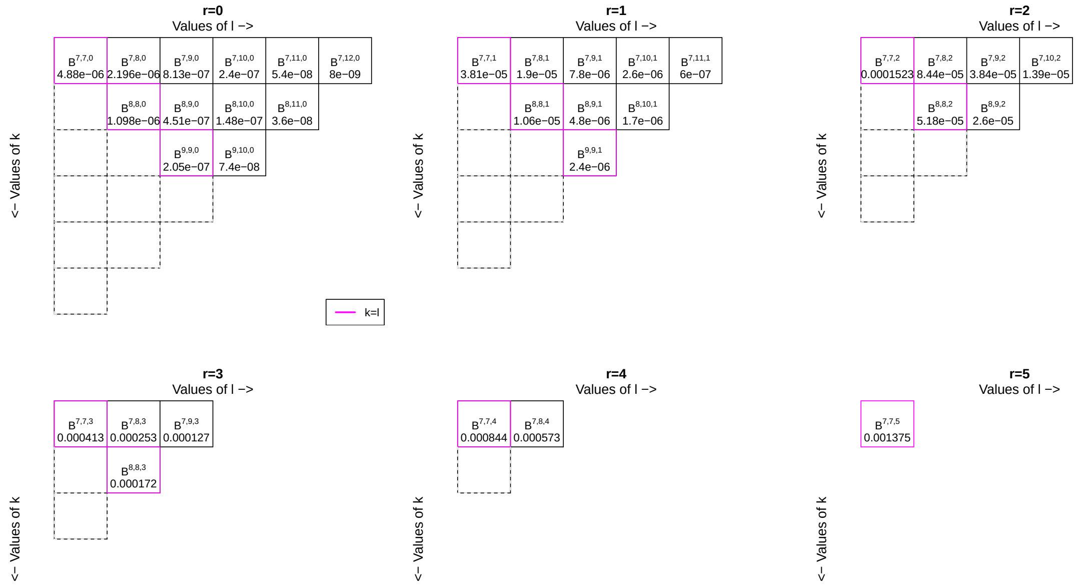

Figure 3: Representation of the $B^{k,l,r}$ used in the computation of $\phi(13)$ after simplification. The cells on the diagonal (represented in magenta) correspond to the computation on the right-side of equation 5.

The number of computations to be done numerically has been almost divided by 2, but remains of the order of n^3. Precisely this number is equal to:

$$C(n) = \begin{cases} \frac{1}{24} n_2(n_2+2)(2n_2-1) & \text{if } n_2 \text{ is even} \\ \frac{1}{24}(n_2-1)(n_2+1)(2n_2+3) & \text{if } n_2 \text{ is odd} \end{cases}$$

For instance, when $n = 10\,001$, we would need now to calculate the expression $B^{k,l,r}$ 10 426 043 750 times, leading to a relative gain by a factor 1.9994 when compared to the previous one.

3.4 Lower and Upper Bounds on ϕ

By exploiting properties P5 and P7, we obtain the following of the function ϕ:

$$\phi_{min} \leq \phi \leq \phi_{maj},$$

where:

$$\phi_{maj}(n) = 6 \sum_{r=0}^{n_2-2} \frac{1}{2}(n_2-r-1)(n_2-r)B^{n_2,n_2,r} \tag{6}$$

and:

$$\phi_{min}(n) = 6\sum_{r=0}^{n_2-2}\frac{1}{2}(n_2-r-1)(n_2-r)B^{n-1-r,n_2,r}$$

which can be expressed equivalently through the change of variable $k = n-1-r$, as:

$$\phi_{min}(n) = 6\sum_{k=n_2}^{n-1}\frac{1}{2}(k-n_2+2)(k-n_2+1)B^{k,n_2,(n-1-k)}$$

which will be very useful in what follows (see 3.4.2).

The number of computations for ϕ_{maj} and ϕ_{min} is of order n (in contrast to n^3 for ϕ).

On the other hand, as explained before, the values of the coefficients $B^{k,l,r}$ tend to 0 when n gets large and this decrease is faster at the ground levels of the tetrahedron as compared to the upper levels. Therefore, it is even possible to reduce further the number of computations without altering the accuracy of the overall computation.

3.4.1 Approximation of ϕ_{maj} and ϕ_{min}

Intuition for an approximation of ϕ_{maj}: When n increases and r tends to 0, the coefficient $B^{n_2,n_2,r}$ tends to 0 (and will be then considered to be equal to 0 by R when around 2.225074×10^{-308}). The main idea to simplify the algorithm is to postulate that we will compute $B^{n_2,n_2,r}$ only for values of r such that $B^{n_2,n_2,r}$ is above a threshold considered to be reasonable.

Let r_{min}, be the smallest integer value largest than or equal to 0, such that if $r \geq r_{min}$, then $B^{n_2,n_2,r} > \varepsilon$ where ε is the fixed degree of accuracy which we impose. We approximate ϕ_{maj} as follows:

$$\tilde{\phi}_{maj}(n) = 6\sum_{r=r_{min}}^{n_2-2}\frac{1}{2}(n_2-r-1)(n_2-r)B^{n_2,n_2,r}$$

The error due to approximation resulting from $\tilde{\phi}_{maj}$ rather than ϕ_{maj} is:

$$Err_1 = \begin{cases} 6\sum_{r=0}^{r_{min}-1}\frac{1}{2}(n_2-r-1)(n_2-r)B^{n_2,n_2,r} & \text{if } r_{min} > 0 \\ 0 \text{ otherwise} & \end{cases}$$

Intuition for an approximation of ϕ_{min}: The idea is the same as before. The coefficient $B^{k,n_2,(n-1-k)}$ converges to 0 when n gets large and k tends to $n-1$. More precisely, the idea is to compute $B^{k,n_2,(n-1-k)}$ only for the k for which $B^{k,n_2,(n-1-k)}$ is above the threshold which is above degree of accuracy which we impose.

Let k_{max}, be the largest integer smallest than or equal to $(n-1)$, defined such that if $k \leq k_{max}$, then $B^{k,n_2,(n-1-k)} > \varepsilon$. We approximate ϕ_{min} as follows:

$$\tilde{\phi}_{min}(n) = 6\sum_{k=n_2}^{k_{max}} \frac{1}{2}(k-n_2+2)(k-n_2+1)B^{k,n_2,(n-1-k)}$$

The error due to approximation resulting from $\tilde{\phi}_{min}$ rather than ϕ_{min} is:

$$Err'_1 = \begin{cases} 6\sum_{k=k_{max}+1}^{n-1} \frac{1}{2}(k-n_2+2)(k-n_2+1)B^{k,n_2,(n-1-k)} & \text{if } k_{max} < n-1 \\ 0 \text{ otherwise} \end{cases}$$

Hereafter, the C corresponding functions are called `IAC_maj` and `IAC_min`. In addition to the values of n that he wants to consider for the computation, the user will have to fix the degree of accuracy ε. Moreover, the functions C return in addition to $\tilde{\phi}_{maj}$ and $\tilde{\phi}_{min}$, the approximation errors Err_1 and Err'_1, as well as r_{min} and k_{max}. This would allow us to pick up ε in order to control Err_1 and Err'_1.

Remark: The outcome for ϕ_{maj} (resp. ϕ_{min}) is identical to the outcome for $\tilde{\phi}_{maj}$ (resp. $\tilde{\phi}_{min}$) with $\varepsilon = 2.225074 \times 10^{-308}$, since this value of ε is the smallest value recognizable by R. In contrast, the number of computations involved for $\tilde{\phi}_{maj}$ (resp. $\tilde{\phi}_{min}$) is always smaller than the number involved for ϕ_{maj} (resp. ϕ_{min}). In order to save computation time, while keeping as much as possible $B^{k,l,r}$, it is therefore always better to run the computations in considering $\tilde{\phi}_{maj}$ (resp. $\tilde{\phi}_{min}$), with $\varepsilon = 2.225074 \times 10^{-308}$. This turns to be useful quite rapidly: as soon as $n \geq 1401$ for $\tilde{\phi}_{maj}$ (before, $r_{min} = 0$), even as soon as $n \geq 539$ for $\tilde{\phi}_{min}$ (before, $k_{max} = n-1$). We deduce then wrongly that: $Err_1 = Err'_1 = 0$. The errors resulting from the granularity of the values recognizable by R are discussed with further details in section 3.6. Note however that the number of $B^{k,l,r}$ below that ε (and therefore considered equal to 0) is less that the total number of $B^{k,l,r}$ (which is of the order $\frac{n^3}{48}$; see equation 4). Furthermore, we will see (see section 3.5.1) that $\phi_{min}(n)$ gets close to $\frac{0.06}{\sqrt{n}}$ i.e. $6 \times 10^{-(p+2)}$,when $n = 10^{2p}+1$. Since the granularity of R leads at best to 16 significant digits, we target an accuracy corresponding to $10^{-(p+19)}$. Therefore, it is sufficient that $\frac{n^3}{48} \times \varepsilon$ which is less than 10^{6p-306} be smaller than $10^{-(p+19)}$. This happens when $p \leq \frac{287}{7} = 41$. Accordingly, getting wrongly $Err_1 = Err'_1 = 0$ is without any consequence at least when n is less than $10^{82}+1$; and we will see that we stop long before reaching that value!

3.4.2 Approximation of ϕ

Approximation 1 In using r_{min} and k_{max} as defined before, we approximate ϕ as follows:

$$\tilde{\phi}_0(n) = 6 \times [2 \times (\sum_{k=\frac{n+1}{2}}^{k_{max}} \sum_{l=k+1}^{l_{max}} \sum_{r=r_{min}}^{r_{max}} B^{k,l,r}) + (\sum_{k=\frac{n+1}{2}}^{k'_{max}} \sum_{r=r_{min}}^{r'_{max}} B^{k,k,r})]$$

where $l_{max} = \min(\frac{3n-1}{2} - k, k_{max})$, $r_{max} = \frac{3n-1}{2} - k - l$, $k'_{max} = \min(n_3, k_{max})$ and $r'_{max} = \frac{3n-1}{2} - 2k$. The approximation error resulting from considering $\tilde{\phi}_0$ instead of ϕ is lower than $Err_1 + Err_2$, with:

$$Err_2 = \begin{cases} 6 \sum_{k=k_{max}+1}^{n-1-r_{min}} \frac{1}{2}(k_{max}-k)(k_{max}-k+1)B^{k,n_2,(n-1-k)} & \text{if } k_{max} < n-1-r_{min} \\ 0 \text{ otherwise} \end{cases}$$

For the case where $n = 13$, we have represented on figure 4 the computations which have been avoided with the help of that approximation when $\varepsilon = 2 \times 10^{-5}$. In that example, we find $r_{min} = 1$ and $k_{max} = 9$.

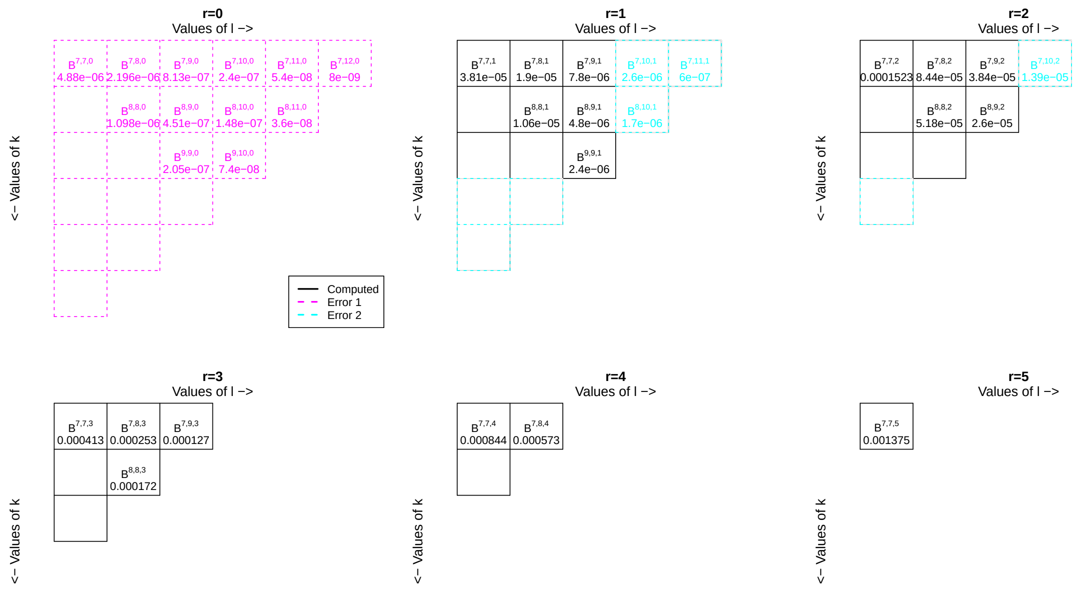

Figure 4: Representation of the $B^{k,l,r}$ used in the computation of $\phi(13)$ after a first approximation, with $\varepsilon = 2 \times 10^{-5}$. In this case $r_{min} = 1$ (sparing the computation of the magenta dotted cells) and $k_{max} = 9$ (sparing the computation of the blue dotted cells).

Remark: In considering $\varepsilon = 2 \times 10^{-5}$, we observe that when $r = 1$, we keep counting the cells whose value is lower than ε. The next idea to improve on approximation is to delete these cells.

Approximation 2 Let r_{kl}, be defined such that for fixed k and l, $B^{k,l,r} > \varepsilon$ for all $r > r_{kl} \geq r_{min}$. Then:

$$\tilde{\phi}(n) = 6[2 \times (\sum_{k=\frac{n+1}{2}}^{k_{max}} \sum_{l=k+1}^{l_{max}} \sum_{r=r_{kl}}^{r_{max}} B^{k,l,r}) + (\sum_{k=\frac{n+1}{2}}^{k'_{max}} \sum_{r=r_{kk}}^{r'_{max}} B^{k,k,r})]$$

The approximation error resulting from considering $\tilde{\phi}$ instead of $\tilde{\phi}_0$ is less than:

$$Err_3 = \varepsilon \times 6[2 \times (\sum_{k=\frac{n+1}{2}}^{k_{max}} \sum_{l=k+1}^{l_{max}} (r_{kl} - r_{min}) + (\sum_{k=\frac{n+1}{2}}^{k'_{max}} (r_{kk} - r_{min})]$$

When $n = 13$, we have represented on figure 5 in orange the computations which have been avoided as a consequence of that second approximation (still when $\varepsilon = 2 \times 10^{-5}$). In that example, we get: $r_{78} = r_{79} = r_{88} = r_{89} = r_{99} = 2$. This is the expression[17] which has been used to code the

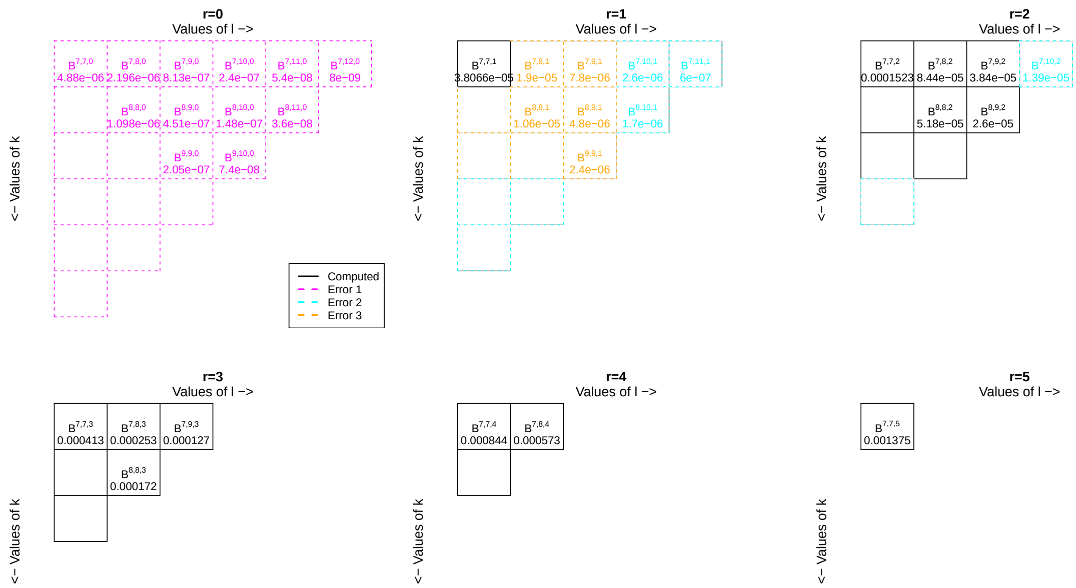

Figure 5: Representation of the $B^{k,l,r}$ used in the computation of $\phi(13)$ after a second approximation, with $\varepsilon = 2 \times 10^{-5}$. In this case $r_{78} = r_{79} = r_{88} = r_{89} = r_{99} = 2$ (sparing the computation of the orange dotted cells).

function labeled `IAC` in C (see an illustration of its use in C).

[17] Truly in coding the function, we have reversed the order of summation on k, l and r in order to arrange the $B^{k,l,r}$ in an increasing order. Without that, we could add further accuracy errors in computing $\tilde{\phi}$.

3.4.3 Ultimate Number of Computations

The total number of computations corresponds to the number of times the expression $B^{k,l,r}$ has been computed. This implies:[18]:

$$Nb_{calc} = \sum_{k=\frac{n+1}{2}}^{k_{max}} \sum_{l=k+1}^{l_{max}} (r_{max} - r_{kl} + 2) + \sum_{k=\frac{n+1}{2}}^{k'_{max}} (r'_{max} - r_{kk} + 2) + n - 1 - k_{max}$$

We have represented on figure 6 the number of times we have to calculate the expression $B^{k,l,r}$ as a function of n, for several values of ε. We have also represented the number of computations in the case where we do no approximation of order n^3 (equation 3.3.3) and in the case where we compute ϕ_{maj}, of order n. The scale units are in $\log_{10}$ to ease the reading of the representation. We observe that for all the first values of n ($n < 21$), there is no saving in computation since all the values of $B^{k,l,r}$ are larger than ε. In contrast when $n = 10^8 + 1$, the number of computations needed for ϕ is of the order 10^{22} while it is only of the order 10^{12} for $\tilde{\phi}$ when $\varepsilon = 10^{-25}$ and of the order 10^{11} when $\varepsilon = 10^{-20}$. If we compute $\tilde{\phi}$ by taking $\varepsilon = 10^{-15}$, we observe that when n is larger than $2 \times 10^7 + 1$, all the expressions $B^{k,l,r}$ are smaller than the threshold and the number of computations collapses to 0[19]. But, $Err = Err_1 > \phi(n)$. Precisely, $\phi(2 \times 10^7 + 1) \approx 3 \times 10^{-5}$, with $Err = Err_1 \approx 7 \times 10^{-5}$. More generally, we will see in appendix B that the cell with the largest value (the peak of the tetrahedron) behaves as: $B^{n_2,n_2,n_2-2} = \frac{2}{\sqrt{3\pi n^2}} - O\left(\frac{1}{n^3}\right)$. Therefore, if $n = 10^p + 1$ and $\varepsilon = 10^{-q}$, then the approximation returns 0 as the outcome when $\varepsilon > B^{n_2,n_2,n_2-2}$, with $B^{n_2,n_2,n_2-2} < 10^{-2p-0.434}$, i.e. at least for $q < 2p + 0.434$. For instance, this happens when $q = 15$, as soon as $p > 7.283$, i.e. $n > 1.92 \times 10^7$ and when $q = 20$ as soon as $n > 6.07 \times 10^9$. Note also that for any given ε, the value of $\tilde{\phi}$ diverges from the value of ϕ long before reaching these limit values of n (see section 3.4.4).

3.4.4 Approximation Error

We have pointed out that computing $\tilde{\phi}$, instead of ϕ, saves the calculation of many $B^{k,l,r}$. As a result, these expressions disappear from the computation of ϕ. The approximation error in moving from ϕ to $\tilde{\phi}$ is at most:

$$Err = Err_1 + Err_2 + Err_3$$

Figure 7 represents the approximation errors by taking $\tilde{\phi}$ instead of ϕ with respect to n for the following values of ε : $10^{-25}, 10^{-20}, 10^{-15}$. For small values of n, the approximation error vanishes[20]. Usually,

[18]Here, we report the number of computations which also allow to provide an estimate of the error in approximation. If we were not interested by the error in approximation, we could made some extra saving by substituting 2 to $n - 1 - k_{max}$ in the formula.

[19]In fact a unique calculation, the cell on the top of the tetrahedron B^{n_2,n_2,n_2-2}, which leads to find that $r_{max} > n_2 - 2$ and therefore that no cell appears in the computation of the approximation of ϕ which is then considered to be equal to 0.

[20]In the sense that all the $B^{k,l,r}$ are calculated. Then, $\tilde{\phi} = \phi$. It remains numerical errors attached to the computations but there are negligible.

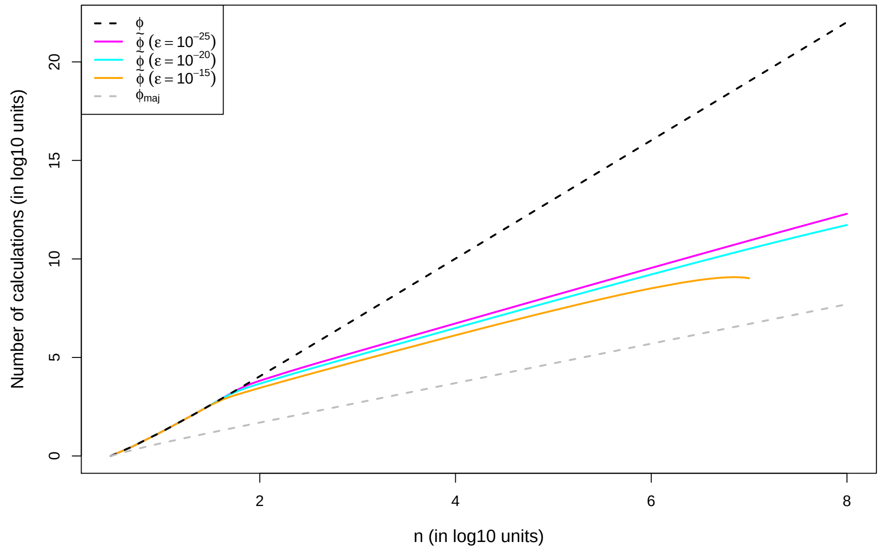

Figure 6: Number of $B^{k,l,r}$ computations necessary to compute ϕ, $\tilde{\phi}$ with $\varepsilon = 10^{-25}, 10^{-20}, 10^{-15}$ and ϕ_{maj} according to the value of n.

the approximation errors increase as n increases. For instance, when $n = 10^6 + 1$, the approximation error is of the order of 10^{-15}, 10^{-11} and 10^{-6} for $\varepsilon = 10^{-25}, 10^{-20}, 10^{-15}$. Interestingly, whatever the choice of ε, after a short period of great increase of the approximation error (from 0 to $\varepsilon \times 10^5$ when n increases), the increase seems to be linear (in log_{10} scales) and the steeps of those lines seem to be very close from one ε to another: around 1.5. This comes from previous figure 6: The number of avoided computations has a steep of 1.5 or so (in log_{10} scales): steep difference between the black dotted line and the colored ones. So if you multiply the number of avoided computations by ε, you get a good proxy of the approximation error.

Of course, the orange line ($\varepsilon = 10^{-15}$) deviates for large values of n. This has partly been explained when commenting the previous figure: For $n > 1.92 \times 10^7$, the approximation error (with this ε) is larger than the value of $\phi(n)$, $\widetilde{\phi}(n) = 0$ and $Err = Err_1 = \phi_{maj}(n)$. As it has been shown (though not formally proved) that $\phi_{maj}(n)$ converges to 0 (see section 3.5.1), it is logical that the approximation error tends towards 0, once it is above the true value of $\phi(n)$.

3.4.5 Parallelization of Function ϕ

To save on computing time, we have proceeded in parallel computing of $\tilde{\phi}$ i.e. we have computed $\tilde{\phi}$ piecewise. Otherwise stated, we have spread the computation of $\tilde{\phi}$ on several cores[21]. To do so, we

[21]Nowadays, most of the computers are endowed with at least two multi-core processes. At GREMAQ, we benefit from a computer server with four processors IntelXeon CPU E7 - 4870 2.40Ghz, each with 10 physical cores (20 logic

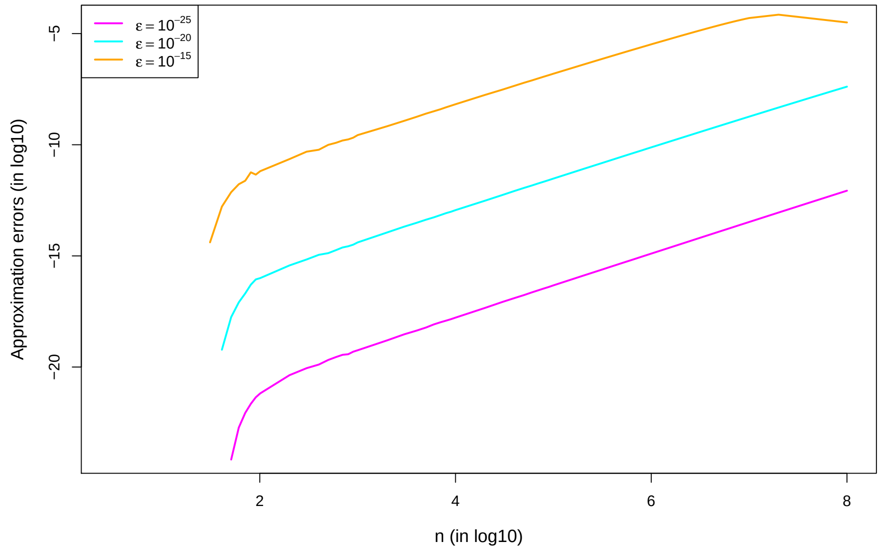

Figure 7: Approximation errors when computing $\tilde{\phi}$ instead of ϕ, with $\varepsilon = 10^{-25}, 10^{-20}$ and 10^{-15} according to n.

send onto the core i, $i = 1, \ldots, nbc$, where nbc is the total number of available cores, the computation of $\tilde{\phi}$ for the set of values $\{k$, such that: $k = (\frac{n+1}{2} + i - 1) + npc \times j \leq k_{max}$, with $j = 1, 2, \ldots\}$, based on the idea that the number of computations on each core is more or less the same. For instance the computation of $\tilde{\phi}(1\ 000\ 001)$ for $\varepsilon = 10^{-25}$ took a bit less than 56s by using 4 cores (instead of 222s with a unique core). The function coded in C is called `IAC_par`. For the computations of ϕ or $\tilde{\phi}$, we have been using up to 50 cores.

3.5 Results

In this section, the computations have been performed for 68 non adjacent values of n ranging from 1 to 100 000 001. Precisely, we have selected the following values of n: 3, 5, 7, 9, then $11, 21, \ldots, 91$, then $101, 201, \ldots, 901$, then $1\ 001, 2\ 001, \ldots, 9\ 001$, then $10\ 001, 20\ 001, \ldots, 90\ 001$, then $100\ 001, 200\ 001, \ldots, 900\ 001$, then $1\ 000\ 001, 2\ 000\ 001, \ldots, 9\ 000\ 001$, then 10 000 001, $20\ 000\ 001, \ldots, 90\ 000\ 001$ and finally 100 000 001. Further, we decided to display the results in the case where $\varepsilon = 10^{-25}$, which guarantees an approximation error quite negligible with respect to the differences between two consecutive values (in the sequence of 68 integers n which are considered) for $\phi(n) \times \sqrt{n}$.

devices).

3.5.1 'Intuitive' Limit of ϕ

We have represented on figure 8, $\tilde{\phi}$, $\tilde{\phi}_{min}$ and $\tilde{\phi}_{maj}$ with respect to n. We also reported on table 1, $\tilde{\phi}$, $\tilde{\phi}_{min}$ and $\tilde{\phi}_{maj}$ for a sample of values of n. Both figure 8 and table 1, suggest:

$$\lim_{n \to +\infty} \phi = 0.$$

n	100 001	1 000 001	10 000 001	100 000 001
$\tilde{\phi}_{min} \times 10^4$	1.945116	0.6121303	0.1932749	0.06108912
$\tilde{\phi} \times 10^4$	4.139992	1.309204	0.4140074	0.1309208
$\tilde{\phi}_{maj} \times 10^4$	9.994129	3.169424	1.003156	0.3173155

Table 1: Values of $\tilde{\phi}$, $\tilde{\phi}_{maj}$ and $\tilde{\phi}_{min}$. Orange decimals could be affected by numerical precision errors (see section 3.6).

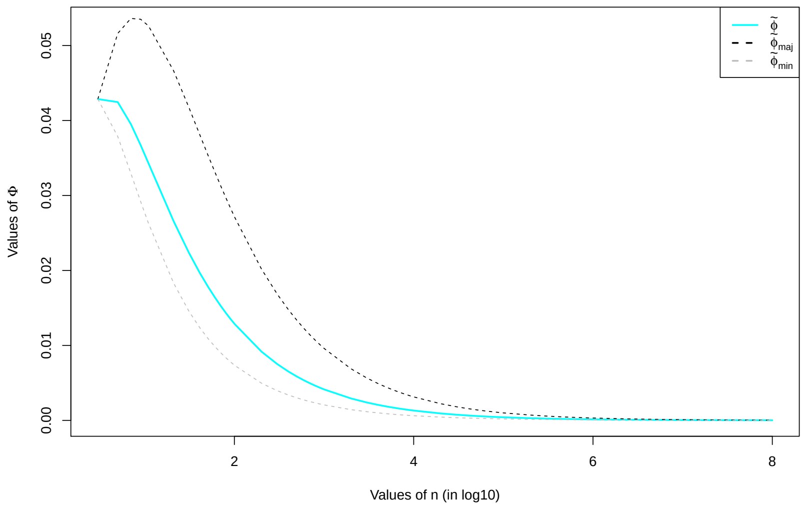

Figure 8: Representation of $\tilde{\phi}$, $\tilde{\phi}_{min}$ and $\tilde{\phi}_{maj}$ according to n.

3.5.2 'Intuitive' Speed of Convergence of ϕ

Figure 9 represents $\sqrt{n} \times \tilde{\phi}_{min}$, $\sqrt{n} \times \tilde{\phi}$ and $\sqrt{n} \times \tilde{\phi}_{maj}$. It seems that ϕ tends to 0 at the same speed of convergence as $\frac{1}{\sqrt{n}}$.

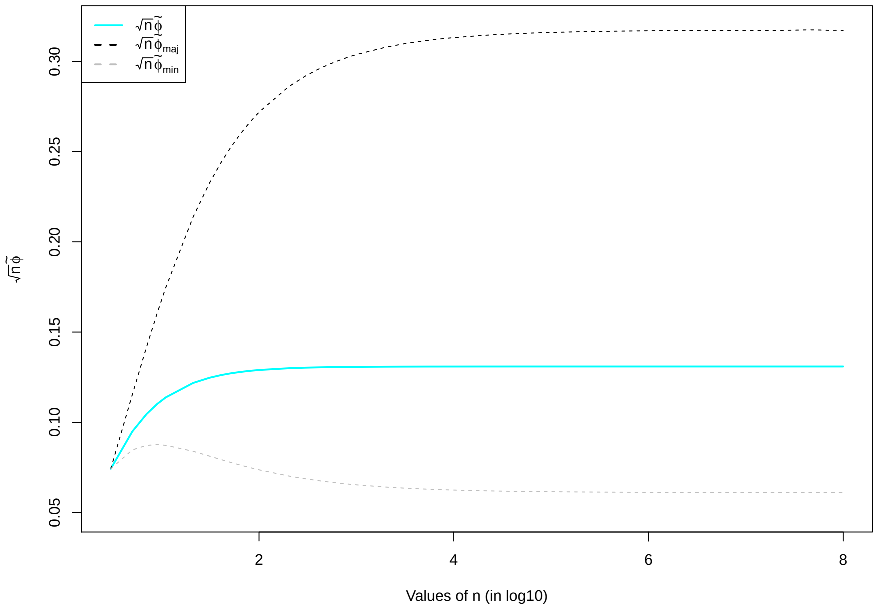

Figure 9: Representation of $\sqrt{n} \times \tilde{\phi}_{min}$, $\sqrt{n} \times \tilde{\phi}$ and $\sqrt{n} \times \tilde{\phi}_{maj}$ according to n.

3.5.3 'Intuitive' Value of $\lim_{n \to +\infty} \sqrt{n} \times \phi$

Table 2 represents $\sqrt{n} \times \tilde{\phi}$, $\sqrt{n} \times \tilde{\phi}_{maj}$ and $\sqrt{n} \times \tilde{\phi}_{min}$ for a sample of values n. It suggests the following majorization and minorization:

$$\mathbf{a} = \lim_{n \to +\infty} n^{1/2} \phi_{min} \leq \lim_{n \to +\infty} n^{1/2} \phi = \mathbf{b} \leq \lim_{n \to +\infty} n^{1/2} \phi_{maj} = \mathbf{c}$$

with: $a \simeq 0.06$, $b \simeq 0.131$ and $c \simeq 0.32$.

n	100 001	1 000 001	10 000 001	100 000 001
$n^{1/2}\tilde{\phi}_{maj}$	0.3160436738	0.3169426034	0.3172258387	0.3173155175
$n^{1/2}\tilde{\phi}$	0.1309186886	0.1309204525	0.1309206294	0.1309207983
$n^{1/2}\tilde{\phi}_{min}$	0.06151027472	0.06121305701	0.06111889019	0.06108912362

Table 2: Computation of $n^{1/2}\tilde{\phi}$, $n^{1/2}\tilde{\phi}_{min}$ and $n^{1/2}\tilde{\phi}_{maj}$ for several values of n. Orange decimals could be affected by numerical precision errors (see section 3.6).

3.6 Validity of the Results

We have represented on figure 10 a zoom of $\sqrt{n} \times \tilde{\phi}$ for large values of n. We observe that the curve is less and less smooth when n increases. We can attribute these numerical perturbations to the relative precision given by R to the float format (around 10^{-16}), which is insufficient to compute accurately

$\tilde{\phi}$ for large values of n. Then, it remains to determine until which decimals the results that we give remain valid:

For each $B^{k,l,r}$ of the sum defining ϕ, $\ln\Gamma(3n+2)$ is the highest computed value to obtain $B^{k,l,r}$. Besides, $\ln\Gamma(3n+2)$ is of the order $3n\times\ln(n)+o(n\times\ln(n))$, see (3). So $\ln(B^{k,l,r})$ leads to the computation of value $lB^{k,l,r}$, which is only known up to $+/-u_n$ at best, where $u_n = 3n\times\ln(n)\times 10^{-16}$. Hence, $B^{k,l,r} = \exp(lB^{k,l,r}+/-u_n) = \exp(lB^{k,l,r})*\exp(+/-u_n)$. As u_n is supposed to be small, $\exp(+/-u_n)$ can be approximated by $1+/-u_n$. Finally, $\tilde{\phi}$, $\tilde{\phi}_{maj}$ and $\tilde{\phi}_{min}$ are only known up to a relative precision of u_n.

For instance, in Table 1, for $n = 100\ 000\ 001$, $u_n \simeq 10^{-6}$: so, the last digit (in orange) is +/- the first non-zero digit, and so only the digits in black are correct for sure. In fact, Figure 11 shows that at this value of n, the approximation error is dominated by the relative precision of floats. Hence, the black digits in Table 1 are also correct for ϕ, ϕ_{maj} and ϕ_{min}.

Similar results hold in Table 2. Interestingly, when going to a n 10 times bigger, one digit precision is lost. For instance, in Figure 10, for $n = 100\ 000\ 001$, the confidence interval of the true value of $n^{1/2}\tilde{\phi}$ is from 0.13092067 to 0.13092092. Similar confidence intervals hold for the other points on the figure. So it is possible to imagine a smooth curve in the envelope formed by the confidence intervals. In fact, the represented curve in Figure 10 is not smooth because the confidence intervals between two consecutive computed values of $n^{1/2}\tilde{\phi}$ are too overlapping.

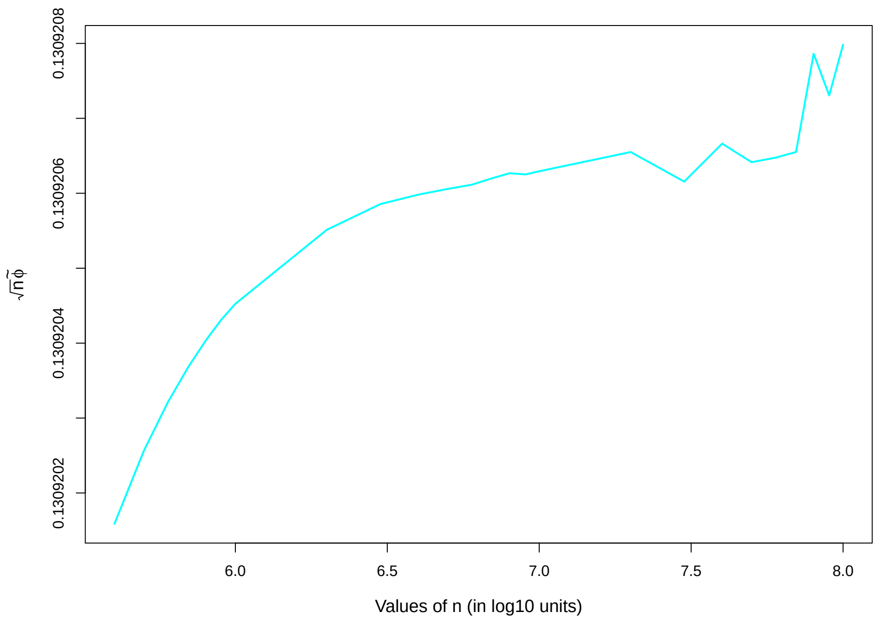

Figure 10: Zoom of $\sqrt{n}\times\tilde{\phi}$ for large values of n.

This is even better explained in Figure 11: As n gets large, the difference between two consecutive computations of $n^{1/2}\tilde{\phi}$ from the considered list of values of n gets smaller, whereas the absolute numerical imprecision increases.

When those two curves cross ($\log_{10}(n) \simeq 6.4$), the numerical imprecision is of same order as the computed difference between two consecutive computations. So the computed difference conveys no information: the true difference can (at least theoretically) even be of the opposite sign of the computed difference!

The situation is even worse after $\log_{10}(n) \simeq 6.6$, where the numerical imprecision dominates the computed difference. So the absolute value of the true difference could be much bigger than the absolute value of the computed difference. This can explain why the curve of the computed difference does not follow the same pattern before and after $\log_{10}(n) \simeq 6.6$, as one could intuitively expect it to.

In fact, in order to be sure that the computed differences dominate the numerical imprecisions, computations should stop around $\log_{10}(n) \simeq 5.9$. There, the numerical imprecision is at least 10 times smaller than the computed difference. Note also that up to $\log_{10}(n) \simeq 5.6$, the numerical imprecision is at least 100 times smaller than the computed difference.

What about the approximation error between $\tilde{\phi}$ and ϕ? Figure 11 shows that for $\varepsilon = 10^{-25}$ (which has been used here for $n^{1/2}\tilde{\phi}$ computations), the maximal approximation error is always dominated by the maximal numerical precision error. In fact, the latter is at least 100 times bigger than the former. So the approximation error is negligible and every results for $\tilde{\phi}$ hold for ϕ too.

Had ε been 10^{-20}, the maximal approximation error would have been at least 100 times smaller than the maximal numerical precision error only up to $\log_{10}(n) \simeq 2.5$. But as computation time is mainly a problem when n gets large, there is only very little interest to use a coarser ε for small values of n.

Finally, in order to have a more correct curve after $\log_{10}(n) \simeq 5.9$, one can either compute less values of n (hence, the difference between two computed values increases way above the maximal numerical precision error) or use a finer representation of floats than double precision (as quadruple precision that give at least 33 significant decimal digits instead of 15 for double precision).

Yet, a finer representation also implies more computing time. This does not seem necessary for this paper, as the results with double precision floats have already been shown valid up to $n = 100\,000\,001$ with 6 significant decimal digits. This numerical approach is mainly interesting in the sense that it gives intuitions on the speed of convergence thanks to results that are valid.

The following section builds on this intuition to give analytical bounds for the speed of convergence of ϕ.

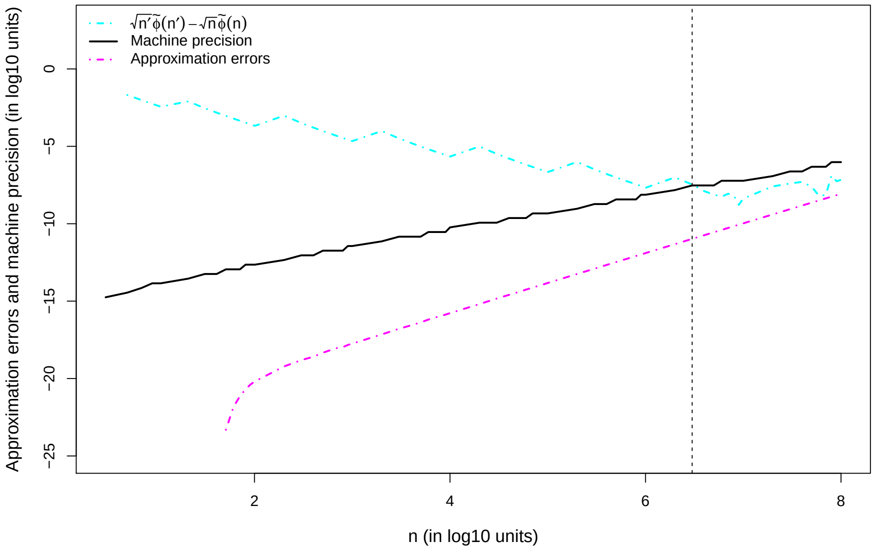

Figure 11: Representation according to n of the maximum approximation error (for $\varepsilon = 10^{-25}$), the maximal numerical imprecision error (due to machine precision limitations) and the difference between two consecutive computations from the considered list: $\sqrt{n'} \times \tilde{\phi}(n') - \sqrt{n} \times \tilde{\phi}(n)$, where n' is the value that follows n in the list of considered values.

4 Analytical Results on the Speed of Convergence of $\phi(n)$ Towards 0

In the preceding section we have introduced a lower and an upper bound of $\phi(n)$ and our numerical analysis brings evidence that each bound converges to 0 as $O\left(\frac{1}{\sqrt{n}}\right)$. In this section, by using repeatedly Stirling's formula, we prove this for a lower bound. The proof for an upper bound is still to be found, but we show the existence of upper bounds which are at least converging to 0 as $O\left(\frac{ln(n)^{1.5}}{\sqrt{n}}\right)$ and even as $o\left(\sqrt{\frac{ln^i(n)^3}{n}}\right)$, for any finite integer i, where $\ln^i(n)$ is recursively defined by $\ln^0(n) = n$ and $\forall j \in \mathbb{N}^*$, $\ln^j(n) = \ln^{j-1}(\ln(n))$.

4.1 Lower Bound

Let $r = r_{LB} = n_2 - \sqrt{n_2}$ (where $n_2 = \frac{n+1}{2}$); this is possible for an infinity of values of n (all those which can expressed as twice a square minus 1: $7, 17, 31, 49, 71, ..., 19\,999, ..., 1\,999\,999, ...$). This restriction on the set of values of n which are considered makes the writing easier (which would otherwise call for the introduction of the integer parts). Let $n = 2 * s^2 - 1$, with s integer. Then, $r_{LB} = s(s-1)$ and, from P7, the smallest $B^{k,l,r_{LB}}$ is obtained for $k = k_{LB} = s(s+1) - 2$ and $l = l_{LB} = n_2 = s^2$. From P2, we deduce that all the $B^{k,l,r}$ (which exist) with $r > r_{LB}$ are larger than $B_{LB} = B^{k_{LB},l_{LB},r_{LB}}$. Finally, all the $B^{k,l,r}$ (which exist) with $0 \leq r < r_{LB}$ are strictly positive. Intuitively all the $B^{k,l,r}$ (which exist) constitute a kind of tetrahedron with an isosceles triangular base (for $r = 0$) and height $n_2 - 1$ (until $r = n_2 - 2$). Therefore, $\phi(n)$ which is equal to six times the sum of the cells of the tetrahedron (the $B^{k,l,r}$ which exist) is larger than six times the number of cells such that $r \geq r_{LB}$ multiplied by the smallest of these cells: B_{LB}. We also neglect all the values of the cells located on lower levels ($r < r_{LB}$). Let $\phi_m(n)$ denote the value of this lower bound. The number of cells such that $r \geq r_{LB}$ is the number of cells of the tetrahedron with isosceles triangular base and height $n_2 - 2 - r_{LB} + 1 = s - 1$. The number of cells for a tetrahedron of height h is $\frac{h(h+1)(h+2)}{6}$. When $h = s - 1$, the number of cells is therefore $\frac{(s-1)s(s+1)}{6}$. Furthermore, Stirling's formula (see appendix B) for B_{LB} implies $B_{LB} = O(s^{-4}) = O(n^{-2})$. Combining, we deduce: $\phi_m(n)$ is at least in $O(s^{-1}) = O\left(\frac{1}{\sqrt{n}}\right)$.

Note that this result only needs to consider one point (B_{LB}) in the tetrahedron. Besides, we have tried to consider more points in order to obtain the highest possible lower bound, but its order remains the same. So our conjecture seems to hold for the lower bound.

4.2 Upper Bound

This subsection first shows that the upper bound is at most in $O\left(\sqrt{\frac{ln(n)^3}{n}}\right)$. This is obtained quite simply, using only two points of the tetrahedron (B_S and B_{I_1}). The subsection then tries to obtain the lowest possible upper bound, making use of more points on the vertical edge of the tetrahedron. We show that the order of the upper bound is much lower, but $O\left(\frac{1}{\sqrt{n}}\right)$ still remains a conjecture.

Figure 12 shows all the points that were used in the tetrahedron.

4.2.1 A Coarse Upper Bound

From P5 and P2, we deduce that the largest value of the tetrahedron is obtained on the top: $B_S = B^{n_2,n_2,n_2-2}$. Stirling's formula (cf. appendix B) implies that this expression converges to 0 as $O(n^{-2})$. Let $r = r_1 = n_2 - \frac{3}{2}\sqrt{n_2 \ln(n)}$, where $n_2 = \frac{n+1}{2}$.

An upper bound of the top part of the tetrahedron (cells for which $r > r_1$) is obtained by replacing each value of these cells by the value of B_S. Since the height of this part is in $O\left(\sqrt{n \ln(n)}\right)$, we deduce that this part of the tetrahedron is majorized by an expression which converges to 0 as $O\left(\frac{\ln(n)^{1.5}}{\sqrt{n}}\right)$.

In the lower part of the tetrahedron ($r \leq r_1$), we also deduce from P5 and P2 that the largest value of a cell is obtained in $B_{I_1} = B^{n_2,n_2,r_1}$. Stirling's formula (see section B) implies that B_{I_1} converges much more rapidly than B_S to 0: precisely as $O(n^{-3.5})$. The height of this lower part of the tetrahedron is r_1 ; the length of its largest edge (for the base i.e. $r = 0$) is $n_2 - 1$. Therefore the number of cells of this part is lower than n^3. Hence, this part of the tetrahedron is majorized by an expression which

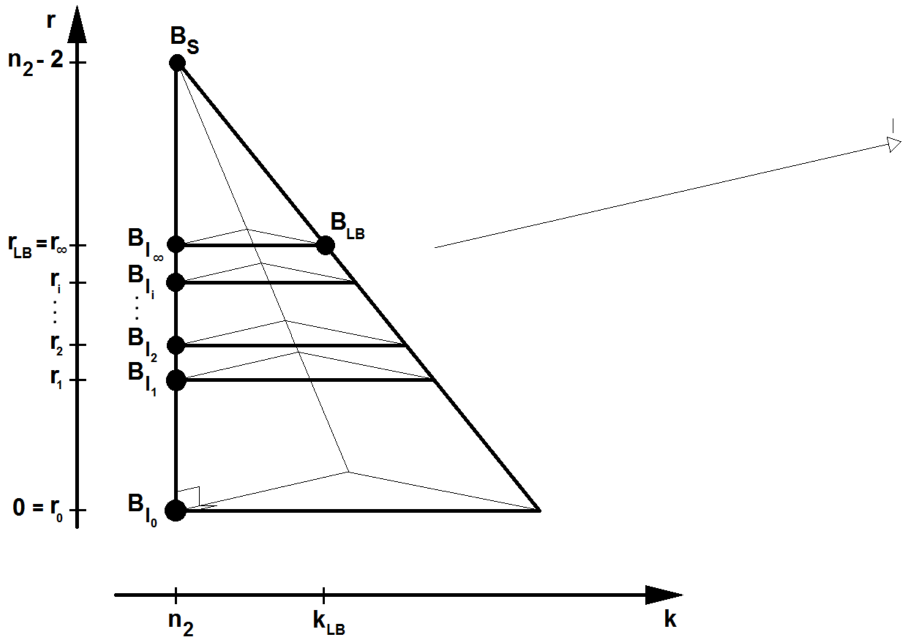

Figure 12: Representation of the tetrahedron containing all $B^{k,l,r}$ values that are added to compute ϕ, and the position of specific cells: B_{LB}, B_S, B_{I_0}, B_{I_1}, B_{I_2}, B_{I_i} and B_{I_∞}.

converges to 0 as $O\left(\frac{1}{\sqrt{n}}\right)$.

All together, we have obtained an upper bound of $\phi(n)$, denoted $\phi_{M_1}(n)$, verifying: $\phi_{M_1}(n) = O\left(\sqrt{\frac{\ln(n)^3}{n}}\right)$.

4.2.2 A Finer Upper Bound

To prove our conjecture, we would need to cut the tetrahedron into parts that are at most in $O\left(\frac{1}{\sqrt{n}}\right)$.

In the previous approach, the sum of the cells of the tetrahedron between the heights of B_{I_0} and B_{I_1} (lower part) is indeed at most in $O\left(\frac{1}{\sqrt{n}}\right)$, but the upper part was not.

Here, the upper part is again divided into pieces to obtain a finer upper bound.

Top part of the tetrahedron

As B_S is in $O(n^{-2})$, the top part till B_{I_∞} (which corresponds to a tetrahedron of height in $O(n^{0.5})$) is in $O\left(\frac{1}{\sqrt{n}}\right)$. Here, the order is exact, as all cells are in $O(n^{-2})$: the largest (B_S) as well as the smallest (B_{LB}) or also B_{I_∞}.

Layer B_{I_1} to B_{I_2} of the tetrahedron

Let $r_2 = n_2 - \frac{3}{2}\sqrt{n_2 \ln(\ln(n))}$. B_{I_2} is the point on the vertical edge of the tetrahedron for which $r = r_2$. Again, Stirling's formula (cf. appendix B) implies that the value in B_{I_2} is in $O\left(n^{-2}\ln(n)^{-1.5}\right)$.

The volume of the tetrahedron layer between B_{I_1} and B_{I_2} is in $O\left((n\ln(n)^{1.5}\right)$. And the highest value in this volume is at B_{I_2}. Hence, the sum of cells in this volume is at most in $O(n^{0.5})$.

Finally, only the tetrahedron part between B_{I_2} and B_{I_∞} remains to estimate. The corresponding volume is in $O((n\ln(\ln(n)))^{1.5})$. And the highest value in this volume is at B_{I_∞}, which is in $O(n^{-2})$. Hence, this part dominates the others, leading to a finer upper bound of $\phi(n)$, denoted $\phi_{M_2}(n)$, verifying: $\phi_{M_2}(n) = O\left(\sqrt{\frac{\ln(\ln(n))^3}{n}}\right)$.

Recursively refining the dominating layer up to B_{I_i} - B_{I_∞}

Let i be an integer greater than 2 and B_{I_i} be on the vertical edge of the tetrahedron at $r = r_i$, with $r_i = n_2 - \frac{3}{2}\sqrt{n_2 \ln^i(n)}$, where $\ln^i(n)$ is recursively defined by $\ln^0(n) = n$ and $\forall j \in \mathbb{N}^*$, $\ln^j(n) = \ln^{j-1}(\ln(n))$.

All layers between B_{j-1} and B_j are at most in $O(n^{0.5})$, as the top part of the tetrahedron. As i is finite, the sum of all these $i+1$ volumes is also at most in $O(n^{0.5})$.

The dominating part corresponds to the layer between B_{I_i} and B_{I_∞}. The corresponding volume is in $O\left((n\ln^i(n)^{1.5}\right)$. And the highest value in this volume is again at B_{I_∞}, which still is in $O(n^{-2})$.

This recursively leads to a finer upper bound of $\phi(n)$, denoted $\phi_{M_i}(n)$, verifying: $\phi_{M_i}(n) = O\left(\sqrt{\frac{\ln^i(n)^3}{n}}\right)$.

As this is true also at $i+1$, we can also write $\phi(n) = o(\phi_{M_i}(n))$.

Upper bound summary

This subsection has proven that for any finite integer i, $\phi(n)$ is dominated by $\sqrt{\frac{\ln^i(n)^3}{n}}$.

Note that, from numerical results, we believe that a closer upper bound converging even faster towards 0 (in $O\left(\frac{1}{\sqrt{n}}\right)$) exists. But we leave that part as a conjecture[22].

5 Concluding Remarks

In this paper we have focused our attention on the stylistic case of three[23] equipopulated districts and *IAC*, a probability model among the most popular ones in social choice. For that popular model, we have shown that the probability of a divided verdict tends to zero, at least as $o\left(\frac{ln^i(n)^{1.5}}{\sqrt{n}}\right)$ (where i is a finite integer and ln^i is recursively defined by $\ln^0(n) = n$ and $\forall j \in \mathbb{N}^*$, $\ln^j(n) = \ln^{j-1}(\ln(n))$) and we conjecture that it should even behave as the inverse of the square root of the population of voters.

[22]Indeed the result we have obtained in this subsection is not sufficient to get rid of the factor $\ln^i(n)^{1.5}$. A counter example proving this limit is even available upon request to the authors. Proving the conjecture would require to study analytically the sum $\phi(n)$ or, which might slightly be easier, the sum ϕ_{maj}, see (6), at least for values of r between r_i and r_∞. But this difficult task still remains to be done.

[23]We speculate that the convergence to zero of the probability of a divided verdict, proved here for three equipopulated districts, continue to hold for an arbitrary number of districts but we have no specific results to report on this question.

This does not match the existing little empirical evidence available on this topic and suggests that the covariance among votes resulting from the *IAC* model is too large as compared to other models like *IC* and IAC^*. The "true" statistical model of votes lies probably in between *IAC* and IAC^*. For instance, if we isolate two districts and n voters per district, we obtain a joint law of votes concentrated along the diagonal as suggested by the following sample of three values when $n = 101$:

$$\begin{bmatrix} k,l & 10 & 50 & 90 \\ 10 & 9.1429 \times 10^{-4} & 1.3245 \times 10^{-12} & 4.2152 \times 10^{-35} \\ 50 & 1.3245 \times 10^{-12} & 5.5107 \times 10^{-4} & 2.0697 \times 10^{-12} \\ 90 & 4.2152 \times 10^{-35} & 2.0697 \times 10^{-12} & 8.7755 \times 10^{-4} \end{bmatrix}$$

in contrast to IAC^* for which the matrix is uniform all the entries being equal to $\frac{1}{((102))^2} = 9.6117 \times 10^{-5}$.

To conclude, let us also point out that beyond the evaluation of the paradox of a divided verdict, the techniques used in this article could also be used for a more detailed analysis of the seats/votes relationship. Indeed, to each random electorate is attached a two dimensional probability law: a realization is an anonymous two dimensional aggregate statistical summary (v, s) of the electoral outcome where v denotes the number/fraction of votes going to the left and s denotes the number/fraction of seats won by the left. This two dimensional vector does not keep track of the specific geographic distribution of the votes and seats.

In this paper, we have not calculated all the entries of this joint votes/seats probability distribution and we limited our investigation to the probability of the event 'the minority party gets a majority of the seats'. Of course, the minority party cannot get all the seats since the majority party must have the majority in at least one district. For an arbitrary number of K districts, it is easy to see that the minority party can get as much as $K-1$ seats if it wins by a short majority in $K-1$ districts and is severely defeated in the last one. Note however that to get a majority of the seats, a party needs to get at least one quarter of the votes. It would be interesting but challenging to use the methods developed in our paper to estimate numerically (at least in the case where there is a single seat per district) the probability of events like "The minority party gets v% of the votes and s% of the seat" where $v \in \left[\frac{1}{4}, \frac{1}{2}\right]$ and $s \in \left[\frac{1}{2}, 1\right]$. The computation could even be conducted for any pair (v, s) in $[0, 1]^2$. After completion of this task, we could, for any v, average over s, and obtain an increasing curve: the so-called seats/votes curve explored by political scientists. In the fictitious case considered in the above paragraph (two districts and 101 voters per district) the probability that the left wins the two seats is equal to 0.47217 while this probability is simply 0.25 for IAC^*. We could also refine this calculation and obtain for instance that the probability that the left wins the two seats given that it receives 121 votes out of the 202 total votes jumps to the value 99.6%! More generally, with two districts and 101 voters per district, we could[24] calculate the conditional probability $\phi_{IAC}(x)$ Pr $\left[\frac{S_L}{S} = 2 \mid V_L = x\right]$

The following table contains a sample of values of ϕ_{IAC}:

x	102	105	109	117	121
$\phi_{IAC}(x)$	11.19%	42.67%	74.13%	97.77%	99.61%

While for IAC^*, we obtain instead the following table:

x	102	105	109	121	142
$\phi_{IAC^*}(x)$	0.9%	4.08%	8.51%	24.39%	67.21%

In future research, we plan to derive numerical estimates of the Votes/Seats curve for a wide spectrum of empirical random electorate models.

References

[1] Abbott, D.W. and J.P. levine (1991) *Wrong Winners: The Coming Debacle in the Electoral College*, New York, Praeger.

[2] Ball, W.J and D.A. Leuthold (1991) “Estimating the Likelihood of an Unpopular Verdict in the Electoral College”, *Public Choice*, 70, 215-224

[3] Banzhaf, J.F. (1965) “Weighted voting does not work: a mathematical analysis”, *Rutgers Law Review*, 19, 317-343.

[4] Banzhaf, J.F. (1966) “Multi-member electoral districts: do they violate the one man, one vote principle”, *Yale Law Journal*, 75, 1309-1338.

[5] Banzhaf, J.F. (1968) “One man, 3.312 votes : a mathematical analysis of the electoral college”, *Villanova Law Review*, 13, 304-332

[6] Barthélémy, F., Martin, M. and A. Piggins (2014) “The Architecture of the Electoral College, House Size Effects, and the Referendum Paradox”, *Electoral Studies*, 34, 11-118.

[7] Best, J. (1971) *The Case Against the Direct Election of the President*, Cornell University Press.

[8] Chambers, C.P. (2008) "Consistent Representative Democracy", *Games and Economic Behavior*, 62, 348-363.

[9] Chamberlain, G. and M. Rothschild (1981) “A Note on the Probability of Casting a Decisive Vote”, *Journal of Economic Theory*, 25,152-162.

[24]Detailed calculations are available from the authors upon request.

[10] Erickson, R.E. and K. Sigman (2000) "A Simple Stochastic Model for U.S. Presidential Elections", Columbia University, Mimeo.

[11] Feix, M.R., Lepelley, D., Merlin, V. and J.L. Rouet (2004) "The probability of conflicts in a U.S. presidential type election", *Economic Theory*, 23, 227-257.

[12] Feix, M.R., Lepelley, D., Merlin, V. and J.L. Rouet (2011) "Three Ways to Compute accurately the probability of the Referendum paradox", *Mathematical Social Sciences*, 23, 227-257.

[13] Fishburn, P.C. and Gehrlein, W.V. (1976) "Borda's Rule, Positional Voting, and Condorcet's Simple Majority Principle", *Public Choice*, 28, 79-88.

[14] Good, I.J. and L.S. Mayer (1975) "Estimating the Efficacy of a Vote", *Behavioral Science*, 20, 25 -33.

[15] Hinich, M.J. Mickelsen, R. and P.C. Ordeshook (1972) "The Electoral College Versus a Direct Vote: Policy Bias, Reversals, and Indeterminate Outcomes", *Journal of Mathematical Sociology*, 4, 3-35.

[16] Kuga, K. and H. Nagatani (1974) "Voter Antagonism and the Paradox of Voting", *Econometrica*, 42, 1045-1067.

[17] Kurrild-Klitgaard, Peter (2013). "Election Inversions, Coalitions and Proportional Representation: Examples of Voting Paradoxes in Danish Government Formations", *Scandinavian Political Studies*, 36, 121-136.

[18] Lahrach, R. and V. Merlin (2012) "Which Voting Rules Minimizes the Probability of the Referendum Paradox: Lessons from French Data", Chapter 5 in *Electoral Systems*, Studies in Social Choice and Welfare, Felsenthal, D.DS. and M. Machover (Eds), Springer-Verlag, Berlin Heidelberg.

[19] Laffond, G. and J. Lainé (2000) "Representation in Majority Tournaments", *Mathematical Social Sciences*, 39, 35-53.

[20] Le Breton, M. and D. Lepelley (2014) "Une Analyse de La Loi Electorale du 29 Juin 1820", *Revue Economique*, 65, 469-518.

[21] Le Breton, M., Lepelley, D. and V. Merlin (2013b) "The Probability of Election Inversions in Mixed-member Electoral Systems", Université de Caen, Mimeo.

[22] Le Breton, M., Lepelley, D. and H. Smaoui (2016) "Correlation, Partitioning and he Probability of Casting a Decisive Vote", Journal of Mathematical Economics, 64, 11-22.

[23] May, K. (1948) “Probability of Certain Election Results”, *American Mathematical Monthly*, 55, 203-209.

[24] May, K. (1958) “The Frequency of Election Victories without Pluralities”, *Americal Philosophical Society Yearbook*, 342-344.

[25] Merrill III, S (1977) “Empirical Estimates for the Likelihood of a Divided Verdict in a Presidential Election”, *Public Choice*, 33, 127-133.

[26] Miller, N.R. (2012a) “Election Inversions by the U.S. Electoral College”, Chapter ? in *Electoral Systems*, Studies in Social Choice and Welfare, Felsenthal, D.DS. and M. Machover (Eds), Springer-Verlag, Berlin Heidelberg.

[27] Miller, N.R. (2012b) “Why the Electoral College is Good for Political Science (and Public Choice)”, *Public Choice*, 150, 1–25.

[28] Miller, N.R. (2014a) “The House Size Effect and the Referendum paradox in U.S. Presidential Elections”, *Electoral Studies*, 35, 265-271.

[29] Miller, N.R. (2014b) “Elections Inversions under Proportional Representation”, *Scandinavian Political Studies*, forthcoming.

[30] Neubauer, M.G. and J. Zeitlin (2003) “Outcomes of Presidential Elections and the House Size”, *Political Science and Politics*, 36, 721-725.

[31] Neubauer, M.G., Schilling, M. and J. Zeitlin (2012) “Exploring Unpopular Presidential Elections”, California State University, Mimeo.

[32] Nurmi, H. (1999) *Voting paradoxes and How to deal with Them*, Springer-Verlag, Berlin.

[33] *Probability Models of Collective Decsion Making*, R.G. Niemi and H.F. Weisberg (Eds), Charles. E Merrill Company, Columbus, 1972.

[34] Shapley, L.S. and M. Shubik (1954) “A method for evaluating the distribution of power in a committee system”, *American Political Science Review*, 48, 787-792.

[35] Straffin P.D. (1988) “The Shapley-Shubik and Banzhaf Power Indices as Probabilities”, in A.E. Roth (eds.) *The Shapley Value: Essays in Honor of Lloyd S. Shapley*, Cambridge University Press

[36] Tufte, E.R. (1973) “The Relationship between Seats and Votes in Two-Party Systems”, *American Journal of Political Science*, 67, 540-554.

[37] Wilgen, J.K. and R.J. Engstrom (1980) “Spatial Distribution of Partisan Support and the Seats/Votes Relationship”, *Legislative Studies Quarterly*, 5, 423-435.

A Proofs

A.1 P1. $B^{k,l,r} = B^{l,k,r}$

Follows trivially from commuting addition and multiplication:

$$\begin{aligned} B^{l,k,r} &= \binom{n}{l}\binom{n}{k}\binom{n}{r}\frac{(l+k+r)!(3n-l-k-r)!}{(3n+1)!} \\ &= \binom{n}{k}\binom{n}{l}\binom{n}{r}\frac{(k+l+r)!(3n-k-l-r)!}{(3n+1)!} \quad = \quad B^{k,l,r} \end{aligned}$$

A.2 P2. $B^{k,l,r} > B^{k,l,r-1}$ (for values k, l, r and $r-1$ which are feasible)

$$\begin{aligned} \frac{B^{k,l,r}}{B^{k,l,r-1}} &= \frac{\binom{n}{r}}{\binom{n}{r-1}}\frac{(k+l+r)!(3n-k-l-r)!}{(k+l+r-1)!(3n-k-l-r+1)!} \\ &= \frac{(r-1)!(n-r+1)!}{r!(n-r)!}\frac{(k+l+r)!(3n-k-l-r)!}{(k+l+r-1)!(3n-k-l-r+1)!} \\ &= \frac{n-r+1}{r}\frac{k+l+r}{3n-k-l-r+1} \end{aligned}$$

Since all the numbers are positive, we deduce:

$$B^{k,l,r} > B^{k,l,r-1} \tag{7}$$

$$\begin{aligned} &\Leftrightarrow \frac{(n-r+1)(k+l+r)}{r(3n-k-l-r+1)} > 1 \\ &\Leftrightarrow (n-r+1)(k+l+r) > r(3n-k-l-r+1) \\ &\Leftrightarrow (n+1)(k+l) > r(3n-k-l+1-n-1+k+l) \end{aligned}$$

$$\Leftrightarrow \frac{(n+1)(k+l)}{2n} > r \tag{8}$$

$$\begin{aligned} \text{Further, } & k \geq n_2 \\ \text{and } & l \geq n_2 \\ \text{with } & n_2 = \frac{n+1}{2} \\ \text{which implies } & k+l \geq n+1 \end{aligned}$$

$$\text{i.e. } \frac{(n+1)(k+l)}{2n} \geq \frac{(n+1)^2}{2n} = n_2 + \frac{1}{2} + \frac{1}{2n} > n_2 - 2 \tag{9}$$

$$\begin{aligned} \text{Finally, } r &\leq \frac{3n-1}{2} - k - l \\ k &\geq n_2 \\ \text{and } l &\geq n_2 \\ \text{implies } r &\leq \frac{3n-1-2n-2}{2} = n_2 - 2 \end{aligned} \tag{10}$$

Combining (10) and (9) implies (8) and then (7).

A.3 P3. $B^{k,l,r} > B^{k+1,l,r}$ (for values of k, l, r and $k+1$ which are feasible)

$$\begin{aligned} \frac{B^{k,l,r}}{B^{k+1,l,r}} &= \frac{\binom{n}{k}}{\binom{n}{k+1}} \frac{(k+l+r)!(3n-k-l-r)!}{(k+1+l+r)!(3n-k-1-l-r)!} \\ &= \frac{(k+1)!(n-k-1)!}{k!(n-k)!} \frac{(k+l+r)!(3n-k-l-r)!}{(k+1+l+r)!(3n-k-1-l-r)!} \\ &= \frac{k+1}{n-k} . \frac{3n-k-l-r}{k+1+l+r} \end{aligned}$$

Again, since all the numbers are positive, this implies::

$$B^{k,l,r} > B^{k+1,l,r} \tag{11}$$

$$\begin{aligned} &\Leftrightarrow \frac{k+1}{n-k} . \frac{3n-k-l-r}{k+1+l+r} > 1 \\ &\Leftrightarrow (k+1)(3n-k-l-r) > (n-k)(k+1+l+r) \\ &\Leftrightarrow k(3n-l-r-1+1+l+r-n) > n(1+l+r) - 3n + l + r \end{aligned}$$

$$\Leftrightarrow k > \frac{(n+1)(l+r)-2n}{2n} \tag{12}$$

$$\text{Further } k \geq n_2 \tag{13}$$

$$\begin{aligned} \text{and } l + r &\leq \frac{3n-1}{2} - k \\ \text{implies } l + r &\leq n - 1 \end{aligned}$$

$$\text{i.e. } \frac{(n+1)(l+r)-2n}{2n} \leq \frac{(n-1)^2}{2n} = n_2 - \frac{3}{2} - \frac{1}{2n} < n_2 \tag{14}$$

Combining (14) and (13) implies(12) and then (11).

A.4 P4. $B^{k,l,r} > B^{k,l+1,r}$ (for values of k, l, r and $l+1$ which are feasible)

The proof is a trivial implication of P1 and P3: $B^{k,l,r} = B^{l,k,r} > B^{l+1,k,r} = B^{k,l+1,r}$.

A.5 P5. For each fixed r, $B^{k,l,r}$ is maximized when $k = l = n_2 (= \frac{n+1}{2})$

It follows trivially from P3 with $k \geq n_2$ and P4 with $l \geq n_2$.

A.6 P6. $\forall l \geq k, B^{k,l,r} > B^{k-1,l+1,r}$ (for values of k, $k-1$, l, $l+1$ and r which are feasible)

From the beginning of the proof of P3 and P1, we deduce:

$$\begin{aligned}
\frac{B^{k+1,l,r}}{B^{k,l+1,r}} &= \frac{B^{k,l,r}}{B^{k,l+1,r}} \times \frac{1}{\frac{B^{k,l,r}}{B^{k+1,l,r}}} \\
&= \left(\frac{l+1}{n-l} \times \frac{3n-k-l-r}{k+1+l+r}\right)\left(\frac{1}{\frac{k+1}{n-k} \times \frac{3n-k-l-r}{k+1+l+r}}\right) \\
&= \frac{l+1}{n-l} \times \frac{n-k}{k+1}
\end{aligned}$$

And since all numbers are positive

$$\begin{aligned}
& B^{k,l,r} > B^{k-1,l+1,r} && (15)\\
\Leftrightarrow & \frac{l+1}{n-l} . \frac{n-k+1}{k} > 1 \\
\Leftrightarrow & \frac{k+j+1}{n-k-j} . \frac{n-k+1}{k} > 1 \\
\Leftrightarrow & (l+1)(n+1) - k - kl > nk - kl \\
\Leftrightarrow & (l+1)(n+1) > k(n+1) \\
\Leftrightarrow & (l+1) > k && (16)
\end{aligned}$$

But since(16) holds true for $l \geq k$, we deduce(15).

Remark: if $l < k$, we obtain a symmetric result via P1.

A.7 P7. For each fixed r, $B^{k,l,r}$ is minimized when $k = n-1-r$ and $l = n_2$

From P3 and P4, we deduce that the minimum is obtained on the base of the "isosceles triangle defined by $B^{k,l,r}$ entries such that $k+l = \frac{3n-1}{2} - r$". From P6 and P1, the minimum is reached when k or l is equal to n_2 (the other being then equal to $n-1-r$).

B Taylor Expansions Using Stirling's Formula

This section explains in more details how the different asymptotic behaviors are obtained for specific $B^{k,l,r}$ used in the paper and represented on Figure 13. It all starts with Stirling's formula (and series) as $B^{k,l,r}$ is composed of products and divisions of different factorials.

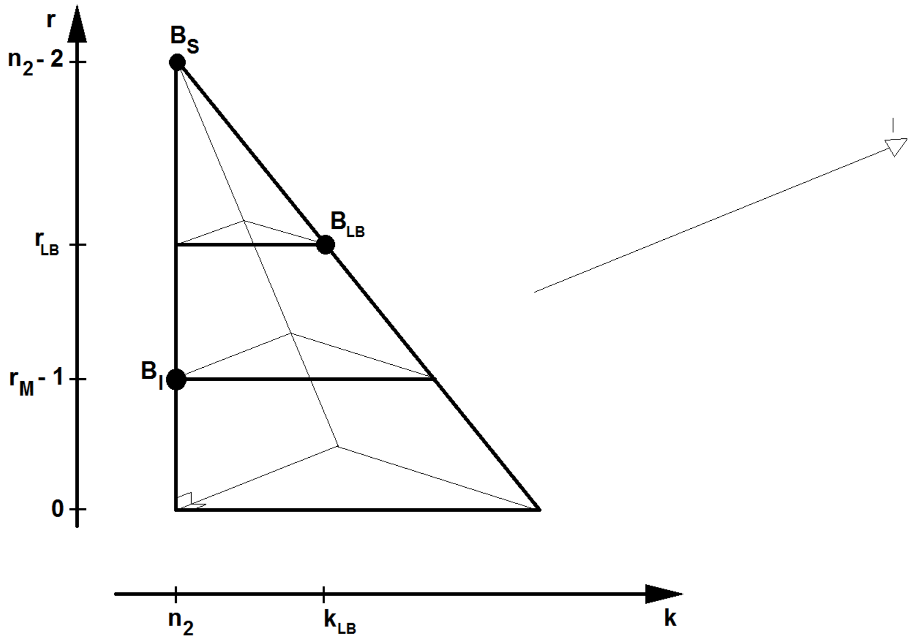

Figure 13: Representation of the tetrahedron containing all $B^{k,l,r}$ values that are added to compute ϕ, and the position of specific cells: B_{LB}, B_S and B_I.

B.1 Approximations of $n!$ or $\ln(n!)$

Striling's formula gives an asymptotic approximation of factorial as:

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

This approximation is in $o(1)$ (in fact $O(n^{-1})$) and is used below in Tables 3, 4 and 5.

In fact, with Taylor expansions, Stirling's formula leads to Striling series, which are used in (3) to compute an accurate approximation of $\ln\Gamma(n)$.

For instance, $\ln\Gamma(n+1) =$

$$\ln(n!) \sim n\ln(n) - n + \frac{1}{2}\ln(2\pi n) + \frac{1}{12n} - \frac{1}{360n^3} + \frac{1}{1260n^5} - \frac{1}{1680n^7} + \frac{1}{1188n^9} + \dots$$

B.2 Approximations for $B^{k,l,r}$ values in B_S, B_I and B_{LB}

Using this approximation in $\ln(B^{k,l,r})$ leads to approximations for values in B_S, B_I and B_{LB} (their position in the tetrahedron are shown on Figure 13), with:

$$\begin{aligned}
\ln(B^{k,l,r}) = \;& 3\ln\Gamma(n+1) && (17)\\
+\;& \ln\Gamma(k+l+r+1) && (18)\\
+\;& \ln\Gamma(3n-k-l-r+1) && (19)\\
-\;& \ln\Gamma(k+1) && (20)\\
-\;& \ln\Gamma(n-k+1) && (21)\\
-\;& \ln\Gamma(l+1) && (22)\\
-\;& \ln\Gamma(n-l+1) && (23)\\
-\;& \ln\Gamma(r+1) && (24)\\
-\;& \ln\Gamma(n-r+1) && (25)\\
-\;& \ln\Gamma(3n+2). && (26)
\end{aligned}$$

The approximations used for the different elements of this sum can be found in Table 3, Table 4 and Table 5.

The following subsections, are organized to give first common approximations of some parts of $\ln(B^{k,l,r})$, either to any points of the tetrahedron: (17) and (26), or to the points of one of its face (l minimal): (22) and (23), or even to the points of its vertical edge (k also minimal): (20), (21). Then the specific approximations of remaining parts of $\ln(B^{k,l,r})$ are given for each point (B_S, B_{LB} and B_I), leading to final approximation for each of the three specific points on the tetrahedron.

For each of the three considered points in the tetrahedron, Table 3 is used. This sole table is sufficient for B_S. For B_{LB}, Table 4 is also required. And for B_I, the second required table is Table 5.

B.2.1 Common approximations for B_S, B_I and B_{LB}

Parts (17) and (26) are common to all $\ln(B^{k,l,r})$, leading to the following approximation:

$$3\ln\Gamma(n+1) - \ln\Gamma(3n+2) = -3\ln(3)n - \frac{3}{2}\ln(3) + \ln(2\pi) + O\left(n^{-1}\right).$$

B.2.2 Approximations for B_{LB} which are similar to common approximations for B_S and B_I

Tetrahedron face $l = \frac{n+1}{2}$

For all points on the face of the tetrahedron presented on Figure 13 (i.e. $l = n_2$, with $n_2 = \frac{n+1}{2}$), parts (22) and (23) can be simplified into:

$\ln\Gamma(i) =$	$\alpha n\ln(n)$	$+\beta n$	$+\gamma\ln(n)$	$+\delta+O(n^{-1})$
i	α	β	γ	δ
$\frac{n-1}{2}$	$\frac{1}{2}$	$-\frac{1+\ln(2)}{2}$	-1	$\frac{1}{2}\ln(8\pi)$
$\frac{n+1}{2}$	$\frac{1}{2}$	$-\frac{1+\ln(2)}{2}$	0	$\frac{1}{2}\ln(2\pi)$
$\frac{n+3}{2}$	$\frac{1}{2}$	$-\frac{1+\ln(2)}{2}$	1	$\frac{1}{2}\ln\left(\frac{\pi}{2}\right)$
$\frac{n+5}{2}$	$\frac{1}{2}$	$-\frac{1+\ln(2)}{2}$	2	$\frac{1}{2}\ln\left(\frac{\pi}{8}\right)$
$n+1$	1	-1	$\frac{1}{2}$	$\frac{1}{2}\ln(2\pi)$
$\frac{3n+1}{2}$	$\frac{3}{2}$	$-\frac{3}{2}\left(1-\ln\left(\frac{3}{2}\right)\right)$	0	$\frac{1}{2}\ln(2\pi)$
$\frac{3n+3}{2}$	$\frac{3}{2}$	$-\frac{3}{2}\left(1-\ln\left(\frac{3}{2}\right)\right)$	1	$\frac{1}{2}\ln\left(\frac{9\pi}{2}\right)$
$3n+2$	3	$-3+3\ln(3)$	$\frac{3}{2}$	$\frac{1}{2}(3\ln(3)+\ln(2\pi))$

Table 3: Approximations of $\ln\Gamma(i)$ for different integers i (in ascending order), using Stirling's formula: $\ln\Gamma(i) = \alpha n\ln(n)+\beta n+\gamma\ln(n)+\delta+O(n^{-1})$.

$$\begin{aligned}
(22): \quad & \ln\Gamma(l+1) &=& \ln\Gamma\left(\frac{n+3}{2}\right) \\
(23): \quad & \ln\Gamma(n-l+1) &=& \ln\Gamma\left(\frac{n+1}{2}\right).
\end{aligned}$$

This finally leads to the following approximation, when $l = n_2$ as for B_{LB}:

$$-\ln\Gamma(l+1)-\ln\Gamma(n-l+1) = \\ -n\ln(n)+(1+\ln(2))n-\ln(n)-\ln(\pi)+O\left(n^{-1}\right).$$

Tetrahedron edge $k = l = \frac{n+1}{2}$

For all points on the vertical edge of the tetrahedron (i.e. $k = l = n_2$, with $n_2 = \frac{n+1}{2}$), parts (20) and (21) can similarly be simplified by symmetry ($k = l$): (20) is equivalent to (22) and (21) is equivalent to (23).

This finally leads to the following approximation, when $k = l = n_2$ as for B_S and B_I:

$$-\ln\Gamma(k+1)-\ln\Gamma(n-k+1)-\ln\Gamma(l+1)-\ln\Gamma(n-l+1)=$$
$$-2n\ln(n)+2(1+\ln(2))n-2\ln(n)-2\ln(\pi)+O\left(n^{-1}\right).$$

B.2.3 Final approximation for B_S

The peak of the tetrahedron, B_S, is located at $k=l=n_2$ and $r=n_2-2$. So, the four last remaining parts of $\ln(B^{k,l,r})$ to approximate for B_S can be simplified into:

$$\begin{array}{llll}
(18): & \ln\Gamma(k+l+r+1) & = & \ln\Gamma(\frac{3n+1}{2}) \\
(19): & \ln\Gamma(3n-k-l-r+1) & = & \ln\Gamma(\frac{3n+3}{2}) \\
(24): & \ln\Gamma(r+1) & = & \ln\Gamma(\frac{n-1}{2}) \\
(25): & \ln\Gamma(n-r+1) & = & \ln\Gamma(\frac{n+5}{2}).
\end{array}$$

This leads to the following approximation, when $k=l=n_2$ and $r=n_2-2$ (B_S):

$$\ln\Gamma(k+l+r+1)+\ln\Gamma(3n-k-l-r+1)-\ln\Gamma(r+1)-\ln\Gamma(n-r+1)=$$
$$2n\ln(n)-(2(1+\ln(2))-3\ln(3))n+\ln(3)+O\left(n^{-1}\right).$$

Summing all parts for the approximation of $\ln(B^{k,l,r})$ value in B_S, leads to:

$$\ln(B^{n_2,n_2,n_2-2})=-2\ln(n)+\ln\left(\frac{2}{\pi\sqrt{3}}\right)+O\left(n^{-1}\right).$$

Finally, in B_S, B^{n_2,n_2,n_2-2} reads:

$$B^{n_2,n_2,n_2-2}=\frac{2}{\pi\sqrt{3}n^2}+O\left(n^{-3}\right)=O\left(n^{-2}\right). \tag{27}$$

B.2.4 Final approximation for B_{LB}

B_{LB}, is located at $k=n_2+\sqrt{n_2}-2$, $l=n_2$ and $r=n_2-\sqrt{n_2}$. So, the six remaining parts of $\ln(B^{k,l,r})$ to approximate for B_{LB} can be simplified into:

$$\begin{array}{llll}
(20): & \ln\Gamma(k+1) & = & \ln\Gamma(\frac{n-1}{2}+\sqrt{\frac{n+1}{2}}) \\
(22): & \ln\Gamma(n-k+1) & = & \ln\Gamma(\frac{n+5}{2}-\sqrt{\frac{n+1}{2}}) \\
(18): & \ln\Gamma(k+l+r+1) & = & \ln\Gamma(\frac{3n+1}{2}) \\
(19): & \ln\Gamma(3n-k-l-r+1) & = & \ln\Gamma(\frac{3n+3}{2})
\end{array}$$

$$(24): \quad \ln\Gamma(r+1) = \ln\Gamma(\frac{n+3}{2} - \sqrt{\frac{n+1}{2}})$$

$$(25): \quad \ln\Gamma(n-r+1) = \ln\Gamma(\frac{n+1}{2} + \sqrt{\frac{n+1}{2}}).$$

The approximations for the two middle parts can be found in Table 3. As for the two first and two last parts, their approximations can be found in Table 4.

$\ln\Gamma(i) = n\ln(n)\cdot$	$\alpha + n\cdot$	$\beta + \ln(n)\sqrt{n}\cdot$	$\gamma + \sqrt{n}\cdot$	$\delta + \ln(n)\cdot$	$\varepsilon +$	$\zeta + O\left(\frac{1}{\sqrt{n}}\right)$
i	α	β	γ	δ	ε	ζ
$\frac{n+3}{2} - \sqrt{\frac{n+1}{2}}$	$\frac{1}{2}$	$-\frac{1+\ln(2)}{2}$	$-\frac{1}{\sqrt{2}}$	$\frac{\ln(2)}{\sqrt{2}}$	1	$\frac{1+\ln\left(\frac{\pi}{2}\right)}{2}$
$\frac{n+5}{2} - \sqrt{\frac{n+1}{2}}$	$\frac{1}{2}$	$-\frac{1+\ln(2)}{2}$	$-\frac{1}{\sqrt{2}}$	$\frac{\ln(2)}{\sqrt{2}}$	2	$\frac{1+\ln\left(\frac{\pi}{8}\right)}{2}$
$\frac{n-1}{2} + \sqrt{\frac{n+1}{2}}$	$\frac{1}{2}$	$-\frac{1+\ln(2)}{2}$	$\frac{1}{\sqrt{2}}$	$-\frac{\ln(2)}{\sqrt{2}}$	-1	$\frac{1+\ln\left(\frac{\pi}{2}\right)}{2}$
$\frac{n+1}{2} + \sqrt{\frac{n+1}{2}}$	$\frac{1}{2}$	$-\frac{1+\ln(2)}{2}$	$\frac{1}{\sqrt{2}}$	$-\frac{\ln(2)}{\sqrt{2}}$	0	$\frac{1+\ln(2\pi)}{2}$

Table 4: Approximations of $\ln\Gamma(i)$ for different integers i (in ascending order), using Stirling's formula: $\ln\Gamma(i) = \alpha n\ln(n) + \beta n + \gamma\sqrt{n}\ln(n) + \delta\sqrt{n} + \varepsilon\ln(n) + \zeta + O\left(\frac{1}{\sqrt{n}}\right)$.

Looking for the approximations of the four values in Table 4 leads to the following approximation, when $k = n_2 + \sqrt{n_2} - 2$, $l = n_2$ and $r = n_2 - \sqrt{n_2}$ (B_{LB}):

$$-\ln\Gamma(k+1) - \ln\Gamma(n-k+1) + \ln\Gamma(k+l+r+1) + \ln\Gamma(3n-k-l-r+1) - \ln\Gamma(r+1) - \ln\Gamma(n-r+1) =$$
$$n\ln(n) + (3\ln(3) - (1+\ln(2)))n - \ln(n) - \left(2 + \ln\left(\frac{\pi}{3}\right)\right) + O\left(n^{-0.5}\right).$$

Summing all parts for the approximation of $\ln(B^{k,l,r})$ value in B_{LB}, leads to:

$$\ln(B^{n_2+\sqrt{n_2}-2,n_2,n_2-\sqrt{n_2}}) = -2\ln(n) - 2 + \ln\left(\frac{2}{\pi\sqrt{3}}\right) + O\left(n^{-0.5}\right).$$

Finally, in B_{LB}, $B^{n_2+\sqrt{n_2}-2,n_2,n_2-\sqrt{n_2}}$ reads:

$$B^{n_2+\sqrt{n_2}-2,n_2,n_2-\sqrt{n_2}} = \frac{2}{\pi\sqrt{3}n^2e^2} + O\left(n^{-2.5}\right) = O\left(n^{-2}\right). \tag{28}$$

Note that the value in B_{LB} is of the same order as in B_S. In fact, when n tends to infinity, the relation between the two tends to e^{-2}.

B.2.5 Final approximation for B_I

Looking at a range of heights on the vertical edge of the tetrahedron to find B_I

Point B_I is located on the vertical edge of the tetrahedron ($k = l = n_2$) at height $r = r_M - 1$. Here, we are looking for a point where the sum of all cells at this height and below is converging towards 0 at least in $O\left(n^{-0.5}\right)$. For this aim, we chose $r_M = n_2 - \sqrt{n_2} f(n)$, i.e. height $r = \frac{n-1}{2} - \sqrt{n_2} f(n)$, with f a positive function which does not increase too much: $\lim_{n \to +\infty} r \geq 0$, i.e. r has to remain positive (at least for large enough n). So f is at most equivalent to $\sqrt{n_2} - 1 - o\left(\sqrt{n}\right)$, when n tends to infinity. In fact, when using Taylor expansions, there is a difference between $f = O\left(\sqrt{n}\right)$, $f = o\left(\sqrt{n}\right)$ and $\exists x < 0.5 / f = O\left(n^x\right)$.

So this subsection is valid only for:

$$f \text{ is a positive function such that } \exists x < 0.5 \text{ with } f = O\left(n^x\right). \tag{29}$$

Under this condition, $\lim_{n \to +\infty} r_M = \infty$. So, every asymptotic result that hold in this section with r_M that might not be integer still hold if we consider only the integer part of r_M, the rest being negligible (at least for large enough values of n).

Exemplifying the approximations in this subsection with function f

For instance, to better understand the approximations that use Taylor expansions, $\ln\left(n - \sqrt{n} f(n)\right)$ would give:

$$\ln\left(n - \sqrt{n} f(n)\right) = \ln(n) + \ln\left(1 - \frac{f(n)}{\sqrt{n}}\right).$$

As (29) means that $\frac{f(n)}{\sqrt{n}} = o(1)$, it is possible to use the Taylor expansion of $\ln(1+u)$ when u is close to 0:

$$\ln(1+u) = u - \frac{1}{2}u^2 + \frac{1}{3}u^3 - \frac{1}{4}u^4 + \dots$$

Hence, $\ln\left(\left(n - \sqrt{n} f(n)\right)!\right)$ would give:

$$\begin{aligned}
\ln\left(\left(n - \sqrt{n} f(n)\right)!\right) =& \frac{1}{2}\ln(2\pi) - \left(n - \sqrt{n} f(n)\right) \\
&+ \left(n - \sqrt{n} f(n) + \frac{1}{2}\right) \ln\left(n - \sqrt{n} f(n)\right) + o(1) \\
=& \frac{1}{2}\ln(2\pi) - n + \sqrt{n} f(n) + o(1) \\
&+ \left(n - \sqrt{n} f(n) + \frac{1}{2}\right)\left(\ln(n) + \frac{f(n)}{n^{0.5}} - \frac{1}{2}\frac{f(n)^2}{n} + \frac{1}{3}\frac{f(n)^3}{n^{1.5}} - \frac{1}{4}\frac{f(n)^4}{n^2} + \dots\right) \\
=& \frac{1}{2}\ln(2\pi) - n + \sqrt{n} f(n) + o(1) \\
&+ n\ln(n) + f(n) n^{0.5} - \frac{1}{2} f(n)^2 + \frac{1}{3}\frac{f(n)^3}{n^{0.5}} - \frac{1}{4}\frac{f(n)^4}{n} + \dots
\end{aligned}$$

$$
\begin{aligned}
&- \sqrt{n}\ln(n)f(n) - f(n)^2 + \frac{1}{2}\frac{f(n)^3}{n^{0.5}} - \frac{1}{3}\frac{f(n)^4}{n} + \frac{1}{4}\frac{f(n)^5}{n^{1.5}} - \dots \\
&+ \frac{\ln(n)}{2} + \frac{f(n)}{2n^{0.5}} - \frac{f(n)^2}{4n} + \frac{f(n)^3}{6n^{1.5}} - \frac{f(n)^4}{8n^2} + \dots \\
=& \frac{1}{2}\ln(2\pi) - n + \sqrt{n}f(n) + o(1) \\
&+ n\ln(n) + f(n)n^{0.5} - \frac{f(n)^2}{2} + \frac{f(n)^3}{3n^{0.5}} - \frac{f(n)^4}{4n} + \dots \\
&\qquad\qquad \text{until } \frac{(-1)^p f(n)^p}{pn^{\frac{p}{2}-1}} = o(1), \text{ and } p \text{ finite from (29)} \\
&- \sqrt{n}\ln(n)f(n) - f(n)^2 + \frac{f(n)^3}{2n^{0.5}} - \frac{f(n)^4}{3n} + \frac{f(n)^5}{4n^{1.5}} - \dots \\
&\qquad\qquad \text{until } \frac{(-1)^q f(n)^{q+1}}{qn^{\frac{q-1}{2}}} = o(1), \text{ and } q = p-1 \\
&+ \frac{\ln(n)}{2} + o(1), \text{ because } \frac{f(n)}{2n^{0.5}} = o(1) \text{ from (29).}
\end{aligned}
$$

Note that p is finite, as (29) implies $\frac{f(n)}{\sqrt{n}} = O\left(\frac{1}{n^{0.5-x}}\right)$, with $0.5 - x > 0$. So p is such that $(p+1)(0.5-x) > 1$, i.e. $p = \left\lfloor \frac{1}{0.5-x} \right\rfloor$. For instance, if $f(n) = n^{\frac{1}{8}}$ (i.e. $x = 0.125$), then $p = 2$ is sufficient (and $q = 1$).

Finally, if we stop at $p = 2$, then, for all f such that $f = o\left(n^{\frac{1}{6}}\right)$, we obtain:

$$
\begin{aligned}
\ln\left(\left(n - \sqrt{n}f(n)\right)!\right) =& \frac{1}{2}\ln(2\pi) - n + \sqrt{n}f(n) + o(1) \\
&+ n\ln(n) + f(n)n^{0.5} - \frac{f(n)^2}{2} + o(1) \\
&- \sqrt{n}\ln(n)f(n) - f(n)^2 + o(1) \\
&+ \frac{\ln(n)}{2} + o(1) \\
=& n\ln(n) - n - \sqrt{n}\ln(n)f(n) + 2\sqrt{n}f(n) + \frac{\ln(n)}{2} - \frac{3f(n)^2}{2} + \frac{\ln(2\pi)}{2} + o(1)
\end{aligned}
$$

Specific approximations for B_I involving f

Let $\ln(B_R^f) = \ln(B^{n_2, n_2, n_2 - \sqrt{n_2}f(n) - 1})$. The four last remaining parts of $\ln(B_R^f)$ to approximate for B_S can be simplified into:

$$
\begin{aligned}
&(18): \quad \ln\Gamma(k+l+r+1) &= \quad& \ln\Gamma\left(\frac{3n+3}{2} - \sqrt{\frac{n+1}{2}}f(n)\right) \\
&(19): \quad \ln\Gamma(3n-k-l-r+1) &= \quad& \ln\Gamma\left(\frac{3n+1}{2} + \sqrt{\frac{n+1}{2}}f(n)\right)
\end{aligned}
$$

$$
\begin{aligned}
(24): \quad & \ln\Gamma(r+1) &=& \quad \ln\Gamma\left(\frac{n+1}{2}-\sqrt{\frac{n+1}{2}}f(n)\right) \\
(25): \quad & \ln\Gamma(n-r+1) &=& \quad \ln\Gamma\left(\frac{n+3}{2}+\sqrt{\frac{n+1}{2}}f(n)\right).
\end{aligned}
$$

Table 5 gives the corresponding approximations. Note that the approximations have terms in $n\ln(n)$ (highest order), $\sqrt{n}\ln(n)f(n)$ and $\ln(n)$, have a constant, and have a series of terms in $n^{1-\frac{p}{2}}f(n)^p$, $p\in\mathbb{N}$.

As we want an approximation up to $o(1)$, we do not need all terms of the series. The number of terms depends on f. For instance, should f be a decreasing function, then only one (n, if $f=o\left(\frac{1}{\sqrt{n}}\right)$) or two terms ($n$ and $\sqrt{n}f(n)$) of the series are enough. If f is strongly decreasing: $f=o\left(\frac{1}{\sqrt{n}\ln(n)}\right)$, then even the term in $\sqrt{n}\ln(n)f(n)$ can be skipped. Now, if $\lim\limits_{n\to+\infty} f(n)$ is a constant, then we may stop at the third term (in $f(n)^2$). Similarly, if f is an increasing function, we may need to add more terms. And the more increasing the function is, the more terms need to be added.

Here, we have stopped the Taylor expansion at the term in $f(n)^2$. This means that the approximation is in $O\left(\frac{f(n)^3}{\sqrt{n}}\right)$. So it is $o(1)$ only if $f=o\left(n^{\frac{1}{6}}\right)$.

If f were to be at least in $O\left(n^{\frac{1}{6}}\right)$, the approximation would be in $O\left(\frac{f(n)^3}{\sqrt{n}}\right)$ instead of $o(1)$, so some columns of Table 5 would no longer be significant, as η. Column ε is no longer significant if f is at least in $O(n^{\frac{1}{6}}\ln(n)^{\frac{1}{3}})$.

Conversely, for columns γ, δ and ζ of Table 5 to be significant (i.e. not in $o(1)$), f should not be too small. Indeed, information in $o(1)$ is useless. In fact, γ is significant only if f is at least in $O\left(\frac{1}{\sqrt{n}\ln(n)}\right)$. Similarly for δ is significant only if f is at least in $O\left(\frac{1}{\sqrt{n}}\right)$ and ζ is significant only if f is at least in $O(1)$.

Final results for the approximation of the value in B_I

The previous paragraph leads to the following approximation, when $k=l=n_2$ and $r=\frac{n-1}{2}-\sqrt{\frac{n+1}{2}}f(n)$, with $f=o\left(n^{\frac{1}{6}}\right)$ (B_I):

$$
\begin{aligned}
&\ln\Gamma(k+l+r+1)+\ln\Gamma(3n-k-l-r+1)-\ln\Gamma(r+1)-\ln\Gamma(n-r+1)= \\
&\qquad 2n\ln(n)-(2(1+\ln(2))-3\ln(3))n-\frac{2}{3}f(n)^2+\ln(3)+o(1).
\end{aligned}
$$

Summing all parts for the approximation of $\ln(B^{k,l,r})$ value in B_I, leads to:

$$
\ln(B^{n_2,n_2,\frac{n-1}{2}-\sqrt{n}f(n)})=-2\ln(n)-\frac{2}{3}f(n)^2-\ln\left(\frac{\pi}{6}\right)+o(1).
$$

Finally, in B_I, $B^{n_2,n_2,\frac{n-1}{2}-\sqrt{n}f(n)}$ reads:

$$
B^{n_2,n_2,\frac{n-1}{2}-\sqrt{n}f(n)}\sim\frac{2}{n^2e^{\frac{2}{3}f(n)^2}\pi\sqrt{3}}=O\left(n^{-2}e^{-\frac{2}{3}f(n)^2}\right). \tag{30}
$$

$\ln\Gamma(x)=n\ln(n)\ \cdot$	α	$+n\cdot\beta+\sqrt{n}\ \ln(n)$	$f(n)\cdot\gamma$	$+\sqrt{n}f(n)\cdot\delta+$	$\ln(n)\cdot\varepsilon+$	$f(n)^2\cdot\zeta$	$+\eta+o(1)$
x with $n_2=\frac{n+1}{2}$	α	β	γ	δ	ε	ζ	η
$\frac{n+1}{2}-\sqrt{n_2}f(n)$	$\frac{1}{2}$	$-\frac{1}{2}(1+\ln(2))$	$-\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}\ln(2)$	0	$\frac{1}{2}$	$\frac{1}{2}\ln(2\pi)$
$\frac{n+3}{2}+\sqrt{n_2}f(n)$	$\frac{1}{2}$	$-\frac{1}{2}(1+\ln(2))$	$\frac{1}{\sqrt{2}}$	$-\frac{1}{\sqrt{2}}\ln(2)$	1	$\frac{1}{2}$	$\frac{1}{2}\ln\left(\frac{\pi}{2}\right)$
$\frac{3n+3}{2}-\sqrt{n_2}f(n)$	$\frac{3}{2}$	$-\frac{3}{2}\left(1-\ln\left(\frac{3}{2}\right)\right)$	$-\frac{1}{\sqrt{2}}$	$-\frac{1}{\sqrt{2}}\ln\left(\frac{3}{2}\right)$	1	$\frac{1}{6}$	$\frac{1}{2}\ln\left(\frac{9\pi}{2}\right)$
$\frac{3n+1}{2}+\sqrt{n_2}f(n)$	$\frac{3}{2}$	$-\frac{3}{2}\left(1-\ln\left(\frac{3}{2}\right)\right)$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}\ln\left(\frac{3}{2}\right)$	0	$\frac{1}{6}$	$\frac{1}{2}\ln(2\pi)$

Table 5: Approximations of $\ln\Gamma(x)$ for different reals x (in ascending order), using Stirling's formula: $\ln\Gamma(x)=\alpha n\ln(n)+\beta n+\gamma\sqrt{n}\ln(n)f(n)+\delta\sqrt{n}f(n)+\varepsilon\ln(n)+\zeta f(n)^2+\eta+o(1)$, f being positive and such that $f=o\left(n^{\frac{1}{6}}\right)$.

C Functions coded in C

The library of files `codeC.dll` contains the following functions which will be run in C from R:

- `loggamma3`: allows to obtain a table containing the values of the function $\ln\Gamma$ (computed using Stirling's formula),
- `IAC`: computation of $\tilde{\phi}$ (for as much values of n as necessary),
- `IAC_maj`: computation of $\tilde{\phi}_{maj}$ (for as much values of n as necessary),
- `IAC_min`: computation of $\tilde{\phi}_{min}$ (for as much values of n as necessary),
- `IAC_par`: parallel computation of $\tilde{\phi}$ (for a unique value of n).

We upload these functions from the file `codeC.dll` as follows[25]

```
> dyn.load("code/codeC.dll")
```

C.1 Function `loggamma3`

```
void loggamma3(double *lgamm, int *taillelgamm)
```

[25] To do so, we must check that the inputs are in the right format.

- `lgamm`: a vector of `double` (initialized at 0), of dimension `taillelgamm`,
- `taillelgamm`: an integer containing the size of the table that we want to consider.

Illustration

We want to compute $\tilde{\phi}$ for $n \leq 1001$. To do so we need the vector $\ln\Gamma$ for the integers ranging from 1 to $3n_{max}+2$ with $n_{max} = 1001$. We store these values in the object `lg`.

```
> n.max=1001
> lg<-.C("loggamma3", lgamm=as.double(rep(0,3*n.max+2)),
+  taillelgamm=as.integer(3*n.max+2))$lgamm
```

C.2 Function IAC

```
void IAC(double *phi, double *n, int *taille, double *lgamm, double *eps,
         double *nb_calc, double *approx)
```

- `phi`: a vector of `double` (initialized at 0) of dimension `taille` containing all the values of $\tilde{\phi}$,
- `n`: a vector of `double` (to be specified by the user) of dimension `taille` containing the values of n,
- `taille`: an integer (to be specified by the user) corresponding to the number of values to be calculated,
- `lgamm`: a vector of dimension $3n_{max}+2$ containing the values of $\ln\Gamma$, where $n_{max} = \max(n)$,
- `eps`: a `double` (to be specified by the user) corresponding to the value of ε,
- `nb_calc`: a vector of `double` (initialized at 0) of dimension `taille` corresponding to the number of expressions which have been calculated,
- `approx`: a vector of `double` (initialized at 0) of dimension `taille` delivering $Err = Err_1 + Err_2 + Err_3$.

Illustration We want to compute $\tilde{\phi}$ for two values: $n = 101$ and $n = 1001$.

```
> n<-c(101,1001)
```

To invoke the function `IAC`, we need to specify ε:

```
> precision = 25
> eps=10^(-precision)
```

We invoke the function `IAC` as follows:

```
> res<-.C("IAC", phi=as.double(rep(0,2)), n=as.double(n),
+  taille=as.integer(2), lgamm=as.double(lg),  eps=as.double(eps),
+  nb_calc=as.double(rep(0,2)), approx=as.double(rep(0,2)))
```

The object `res` contains the values of $\tilde{\phi}$:

```
> print(res$phi,20)

[1] 0.0128348271775305423914 0.0041318201550210147122
```

we can multiply these numbers by $\sqrt{n}$:

```
> print(sqrt(n)*res$phi,20)

[1] 0.12898841675276404017 0.13072493920913225152
```

The object `res` contains the number of computations which have been done:

```
> res$nb_calc

[1]   6681 202011
```

It also contains the approximation errors Err resulting from the choice of ε:

```
> res$approx

[1] 3.825876e-21 2.470091e-19
```

C.3 Function `IAC_maj`

```
void IAC_maj(double *phi, double *n, int *taille, double *lgamm,
             double *eps, double *r_min, double *approx_maj)
```

- `phi`: vector of `double` (initialized at 0) of dimension `taille` containing the values of $\tilde{\phi}_{maj}$,
- `n`: vector of `double` ((to be specified by the user) of dimension `taille` containing the values of n,

- `taille`: an integer ((to be specified by the user) giving the number of values to be calculated,
- `lgamm`: a vector of dimension $3n_{max}+2$ containing the values of $\ln\Gamma$, where $n_{max}=\max(n)$,
- `eps`: a `double` ((to be specified by the user) specifying the value of ε,
- `r_min`: a vector of `double` (initialized at 0) of dimension `taille` delivering the values r_{min},
- `approx_maj`: a vector of `double` (initialized at 0) of dimension `taille` giving Err_1.

Illustration

We want to compute $\tilde{\phi}_{maj}$ for the same values of n and ε:

```
> majorant<-.C("IAC_maj", phi=as.double(rep(0,2)), n=as.double(n),
+  taille=as.integer(2), lgamm=as.double(lg), eps=as.double(eps),
+  r_min=as.double(rep(0,2)),  erreur1=as.double(rep(0,2)))
```

Then we depict the values of $\tilde{\phi}_{maj}$, $\sqrt{n}\times\tilde{\phi}_{maj}$, the approximation errors Err_1 and the values r_{min} which will be stored in the object `r_min` for subsequent use:

```
> print(majorant$phi,20)

[1] 0.027071555314099243760 9 0.0096017241779385118644

> print(sqrt(n)*majorant$phi,20)

[1] 0.27206576377699165370 0.30378495732408838936

> majorant$erreur1

[1] 1.816305e-23 1.882691e-20

> (r_min<-majorant$r_min)

[1]   1 325
```

C.4 Function[26] IAC_min

```
void IAC_min(double *phi, double *n, int *taille, double *lgamm,
             double *eps, double *k_max, double *r_min,
             double *erreur2, double *erreur1_prime)
```

- phi: vector of double (initialized at 0) of dimension taille containing the values of $\tilde{\phi}_{maj}$,
- n: vector of double ((to be specified by the user) of dimension taille containing the values of n,
- taille: an integer ((to be specified by the user) giving the number of values to be calculated,
- lgamm: a vector of dimension $3n_{max}+2$ containing the values $\ln\Gamma$, where $n_{max} = \max(n)$,
- eps: a double ((to be specified by the user) specifying the value of ε,
- k_max: a vector of double (initialized at 0) of dimension taille delivering the values k_{max},
- r_min: a vector of double ((to be specified by the user) of dimension taille delivering the values r_{min},
- erreur2: a vector of double (initialized at 0) of dimension taille delivering Err_2,
- erreur1_prime: a vector of double (initialized at 0) of dimension taille delivering Err_1'.

Illustration

We want to compute $\tilde{\phi}_{min}$ for the same values of n and ε:

```
> minorant<-.C("IAC_min", phi=as.double(rep(0,2)), n=as.double(n),
+  taille=as.integer(2), lgamm=as.double(lg), eps=as.double(eps),
+  k_max=as.double(rep(0,2)), r_min=as.double(r_min),
+  erreur2=as.double(rep(0,2)), erreur1_prime=as.double(rep(0,2)))
```

OWe depict the values of $\tilde{\phi}_{min}$, $\sqrt{n}\times\tilde{\phi}_{min}$, Err_1' and Err_2. We store the values k_{max} for subsequent use

```
> print(minorant$phi,20)

[1] 0.0073237458834344215317 0.0020649699772476888537

> print(sqrt(n)*minorant$phi,20)
```

[26] This function takes as inputs the values of r_{min} which have been previously calculated in order to compute Err_2.

```
[1] 0.07360273520921206680 3 0.06533272616339574900 8
> minorant$erreur1_prime
[1] 4.190794e-23 2.737040e-21
> minorant$erreur2
[1] 1.132576e-25 1.240105e-23
> (k_max<-minorant$k_max)
[1]  83 603
```

C.5 Function `IAC_par`

```
void IAC_par(double *phi, double *n, double *lgamm, double *eps,
             int *r_min, int *k_max, double *x, double *nbcpus,
             double *approx)
```

- `phi`: one value of `double` (initialized at 0) containing the value of $\tilde{\phi}$ calculated for the set of values $\{k_i\}$ defined for $i = 1, ..., \text{npcpus}$ and such that $\{k_i = (k_{max} + 1 - i) - npcpus \times j \geq \frac{n+1}{2}$, with $j = 1, 2, ...\}$, and where `n`, `k_max` and `nbcpus` are specified by the users,
- `n`: One value of `double` (to be defined by the user) containing the value of n,
- `lgamm`: vector of dimension $3n+2$ containing the values of $\ln \Gamma$,
- `eps`: one `double` ((to be specified by the user) specifying the value of ε,
- `r_min`: a `double` ((to be specified by the user and obtained from the function `IAC_maj`) delivering the value of r_{min},
- `k_max`: a `double` ((to be specified by the user and obtained from the function `IAC_min`) delivering the value of k_{max},
- `x`: a `double` ((to be specified by the user) which changes depending upon the core on which the calculation is sent. `x=(k_max+1-i)` with $i = 1, ..., \text{npcpus}$,
- `nbcpus`: a `double` ((to be specified by the user) giving the number of cores where is sent the computation of $\tilde{\phi}$,
- `approx`: a `double` (initialized at 0) delivering Err_3.

Illustration To run parallel computing, we have to create a function which allows to send on `IAC_`on several cores via a simple modification of the parameter `x`. This function is called `calpar1`.

```
> calpar1<-function(x, n, r_min, k_max, lgamm, eps, nbcpus)
+ {
+ dyn.load("C:/Users/laurent/Desktop/michel lebreton/code/codeC.dll")
+ res2<-.C("IAC_par", phi=as.double(0), n=as.double(n),
+   lgamm=as.double(lgamm), eps= as.double(eps), r_min=as.integer(r_min),
+   k_max=as.integer(k_max), x=as.double(x), nbcpus=as.double(nbcpus),
+   approx=as.double(rep(0,2)))
+ return(list(phi=res2$phi,nb_calc=res2$approx[2],approx3=res2$approx[1]))
+ }
```

We will compute $\tilde{\phi}$ for $n = 1001$ on 4 processors:

```
> nbcpus=4
```

The values of `x` to be initialized are:

```
> (x=(k_max[2]+1)-0:(nbcpus-1))

[1] 604 603 602 601
```

To run the parallel computing:

```
> require("snowfall")
> sfInit(parallel=TRUE,cpus=nbcpus)

R Version:  R Under development (unstable) (2014-06-11 r65921)

> result2 <-sfClusterApplyLB(x, calpar1, 1001, r_min[2], k_max[2],
+  lg, eps, nbcpus)
> sfStop()
```

The returned object contains the peices of $\tilde{\phi}$:

```
> unlist(lapply(result2,function(x) x$phi))

[1] 0.0006837732 0.0008864855 0.0011330502 0.0014285113
```

We then aggregate the values of $\tilde{\phi}$ and the number of calculations:

```
> print(sum(unlist(lapply(result2,function(x) x$phi))),20)

[1] 0.004131820155021013848

> print(sum(unlist(lapply(result2,function(x) x$nb_calc))),20)

[1] 202011.00000000000000
```

To recover the approximation error, we sum the approximation errors Err_1, Err_2, Err_3:[27]

```
> majorant$erreur1[2] + minorant$erreur2[2] +
+  sum(unlist(lapply(result2,function(x) x$approx3)))

[1] 2.470091e-19
```

[27] Until the 17^{th} decimal, we get the same values as those obtained without parallel computing. The differences which appear then result from the fact that the sum of the expressions $B^{k,l,r}$ does not follow the same order n the case of parallel computing.