

# Original Research Article

## Computer Simulation Model of Prime Numbers

---

### ABSTRACT

Prime numbers represent one of the major open problems in number theory mostly in that at present it is not possible to state that the induction principle holds for them. The methodology of experimental mathematics has been little endeavored in this field thus the present report deals with an innovative approach to the problem of primes treated as raw experimental data and as elements of larger and larger finite sequences  $\{P_n\}$ . The modified Chi-square function in the form  $-1/X_k^2(A, n/\mu)$  with the ad-hoc  $A$ ,  $k$  and  $\mu$  parameters is the best-fit function of the finite sequences of primes  $\{P_n\}$ , like the truncated progressions  $\{C_\alpha n^\alpha\}$  with domain  $N$  and co-domain  $R^+$ , being  $(\alpha, k) \equiv (1^+, 0)$  and  $k=2-2\alpha$  and just like the function  $\lambda_n \times n \times \ln(n)$ , what leads to induction algorithms and to many fit relationships  $P_n \approx P(n)$  though within the precisions of the calculations that is approximate. A bi-injective map can be set between the prime sequences and any of these three fit functions showing that the property of scale invariance does not hold for the fits of the finite sequences of prime numbers. Moreover an approximate inductive algorithm is shown capable of finding the approximate value of a prime  $P_{n+1}$  starting from the value of the previous  $P_n$ .

*Keywords: prime number sequences, scale variation, numerical progressions, data fits, modified Chi-square function, experimental mathematics.*

### 1. INTRODUCTION

The problem of prime numbers in number theory has always been a challenge and still nowadays it remains one of the major open problems of mathematics, notwithstanding the many theoretical successes achieved both historically [1-9] and more recently [10-18]. The main problem concerns the fact that neither an exact relationship that links the value of a prime  $P_n$  to its counter  $n$  i.e.  $P_n = P(n)$  has yet been found (or simply it does not exist) nor there is an analytical relationship that links any prime number  $P_{n+1}$  to its preceding  $P_n$ . In other words, presently it is not possible to state that the principle of induction holds for prime numbers and they seem to show a non-deterministic nature. Moreover prime numbers are of the utmost importance in that strictly connected to geometry and physics, and so is the problem of the zeroes of Riemann's zeta function and his renowned hypothesis [19-24]. Thus the prime number problem seems to be one of the so-called intractable problems [25,26] whereas an intractable problem is one which is very difficult (if not impossible) to solve in that, because of the great number of unknown processes and/or hidden variables to be considered, one cannot quickly reach the result so that there would be just a method to treat intractable problems i.e. approximations. Many real-life problems are of this kind, for instance models explaining economy or treating the climate changes are necessarily approximate due to the presence of many variables involved, more or less hidden.

In mathematical domains, where exact rules exist, we encounter intractability, though seldom, owing to the many ways of application of the processes, so that approximation is an attractive practice for use in problem-solving techniques because it allows us to explain some intractable problems and at the same time it can sometimes lead to more efficient

solutions to problems which do not require a precise answer. As a matter of fact in some cases the exact solution is no more desirable than an approximate one.

This paper shows how the approximation technique can be used to get some explanations in the field of prime numbers which show a kind of hidden intractability that takes shape mostly in raising the question whether prime numbers are deterministic or stochastic or whatever else. At the present time this aspect cannot be treated and determined in any way thus being considered a typical intractable problem. As for the nomenclature i.e. terminology, the term *fit* is used along all the article as a synonym of approximation together with the two terms: data *interpolation* and *extrapolation*. Among the aims of the study there is to investigate the innermost nature of prime numbers as far as possible and to highlight their deterministic aspect, if any.

As for the chronology of the work it should be highlighted that the whole study and all the connected calculations took about some months to be performed. The starting point has been the question of whether and how the finite sequences of prime numbers could be treated, a question that was raised by the Author himself some time ago. In the present report, after some introducing topics like approximations, statistics and so on, the many and long calculations on the finite sequences of prime numbers are examined and some examples shown with the aim of finding some remarkable link, if any, between any of them and analytic functions. These calculations took approximately some hundreds person-hours to be made, that is approximately some months by the sole Author himself. As a matter of fact more than 90 prime sequences have been investigated, any of them made of 200 prime numbers for which about ten variables (both statistical and not) have been calculated by means of ad-hoc code (spread-sheet) for the three fit functions so that an amount of approximately half a million numbers have been managed.

## 1.1 Approximations

One of the most appealing approximations in number theory is that concerning factorial numbers  $n! = 1 \times 2 \times 3 \times 4 \times \dots \times (n-1) \times n = \Gamma(n)$   $n, \Gamma \in \mathbb{N}$  given by the renowned Stirling's formula  $n! \approx \sqrt{2\pi n} \times (n/e)^n$  the accuracy of which has been well confirmed and its proof well established. As a matter of fact the figure 1 shows the function  $f(n) = \log(n!)$  as fitted by Stirling's function  $s(n) = \log[\sqrt{2\pi n} \times (n/e)^n]$  up to the maximum value of  $N = n_{\text{Max}} = 170$  i.e.  $n! = 170! = 7.2574156153079900 \times 10^{306}$  the maximum value supported by the CPU of the PC used by the Author. The good matching between the actual data and the Stirling fit i.e. formula is evident from the quasi-perfect superposition of the two curves, or the two data point sets, as depicted in the figure 1.

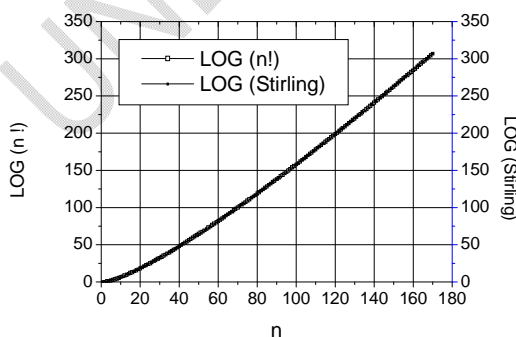


Fig. 1 Fit between  $n!$  and Stirling's formula

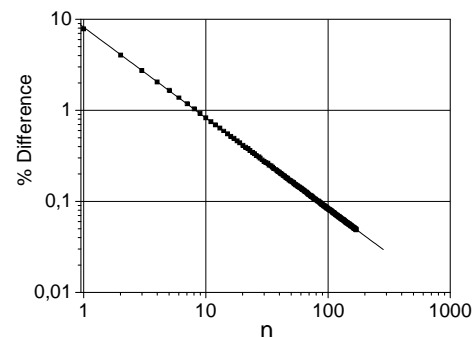


Fig. 2 % Difference between  $n!$  and Stirling's formula

Just for instance at  $n_{\text{Max}} = N = 170$  one can easily verify that the values are  $170! = 7.257415615307990E+306$  and  $\text{Stirling}(170) = 7.253858934542950E+306$  with a percentage difference between the two values equal to  $0.049007538682848400\%$ . As a matter of fact also the % differences are presented in Fig. 2 showing that Stirling's approximation works very well most of all because, by the extrapolation of the experimental data, one has  $\lim_{n \rightarrow \infty} \delta = \lim_{n \rightarrow \infty} [n! - \sqrt{(2\pi n)}x(n/e)^n] / n! \times 100 = \lim_{n \rightarrow \infty} [n! - \text{FIT}] / n! \times 100 = 0$  what means that Stirling's approximation (or fit) works better and better as  $n \rightarrow \infty$ . Thus - while factorial numbers are one of the most representative and resolved aspects of intractable problems in mathematics well treated by Stirling's approximation - they are a paradigmatic example of how this kind of problems can be explained by approximation methods therefore leading to ask oneself whether this technique can be applied to other intractable problems, typically to the problem of prime numbers.

In addition Stirling's formula addresses another important concept in mathematical analysis: that of *continuation* of a discrete function or of a numeric progression with domain and co-domain in the natural field  $N$  to a continuous function defined in the real field  $R^+$  what implies also the notion of data interpolation and extrapolation, as already said. This approach is of a fundamental and paramount importance in this whole report and methodology. Moreover, Stirling's formula shows how one can find a solution to a problem in a so-called *closed form*.

Thus, starting from the classical Prime Number Theorem (PNT)  $\pi(x) = \text{Li}(x) = \int_2^x \frac{1}{\ln(t)} dt$  (the integral runs from 2 up to  $x$ ) so that  $\lim_{x \rightarrow \infty} \pi(x) = x / \ln(x)$ , it is well known that an equivalent formulation is  $P_n \approx P(n) \approx nx \ln(n)$  and that's why in the present context the Author shall refer to the standard PNT as to the law (Gauss' estimate)  $P_n \approx nx \ln(n)$ . Nonetheless, despite its brilliance and its being nearly right asymptotically, this latter canonical representation of the

PNT does not work at best both to get the finite value of a prime number  $P_n \in P \subset N$  starting from its counter  $n \in N$  and to give the asymptotic values of a prime (Fig. 3). Indeed, the Gauss estimate is by far the best one. For instance the estimate  $\lim_{x \rightarrow \infty} \pi(x) = x / [\ln(x) - 1]$  is more attractive and so are many other ones, for instance  $\sum_{P \leq n} 1/P_n \approx \ln[\ln(n+1)] - \ln(\pi^2/6)$ . As a matter of fact, the next figures 3 and 4 show that - despite the fact that the percentage difference between the classical PNT and the actual  $P_n$  vs  $n$  goes to 0 (not shown) - the simple difference  $\text{actual } P_n - nx \ln(n)$  versus  $n$  increases more and more (Fig. 4).

The same happens for many other fits or approximations such as  $P_n \approx nx \{ \ln(n) + \ln[\ln(n)] \}$  and

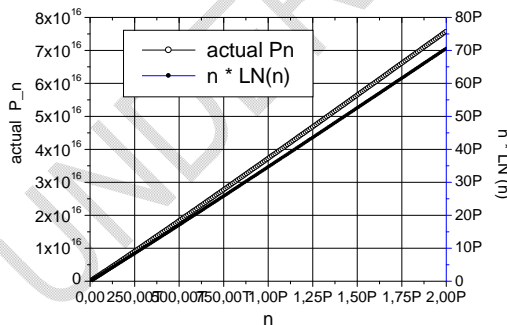


Fig. 3 Comparison between actual  $P_n$  and  $nx \ln(n)$

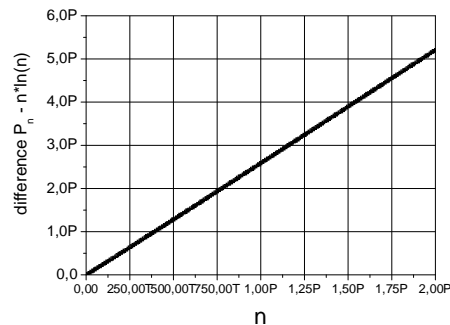


Fig. 4 Difference between actual  $P_n$  and  $nx \ln(n)$

the fits by Gandhi [17]  $P_n \approx nx \{ \ln(n) + \ln[\ln(n) - 1] \}$  and by Cipolla [18]  $P_n \approx nx (\log(n) + \log \log(n) - 1 + (\log \log(n) - 2) / \log(n) - ((\log \log(n))^2 - 6 \log \log(n) + 11) / (2 \times \log^2(n)))$ . Actually, this latest formula oscillates a great deal and the errors  $O[n \log \log(n) / \log(n)]^3$  are still too large. The main feature common to all these estimates is that the percent difference, or the mere difference, between the actual data and the fitting formula shows an increasing trend vs.  $n$  instead of a decreasing one as in Stirling's formula for  $n!$

As a remark, in these two previous figures (as well as in most cases in all this report) the engineering notation has been used for the numbers, for instance  $1K=1E3=1 \times 10^3$ ,  $1M=1E6=1 \times 10^6=1$ Mega,  $1G=1E9=1 \times 10^9=1$ Giga,  $1T=1E12=1 \times 10^{12}=1$ Tera,  $1P=1E15=1 \times 10^{15}=1$ Peta,  $1m=1E-3=1 \times 10^{-3}=1$ milli,  $1\mu=1E-6=1 \times 10^{-6}=1$ micro and so on.

In addition, another question arises for prime numbers. If the standard limit  $P_n \approx n \times \ln(n)$  holds asymptotically how do the prime sequences reach this infinite limit? Is there any pattern or pathway on its trend towards this standard asymptotic limit? The aim of the present work is to answer these questions too. As a matter of fact most of the research on primes has been dedicated and focused, in these latest years, to this topic, that is to understand whether or not the prime number sequences cover hidden connections or relationships or patterns or whatsoever.

Moreover, despite the fact that the experimental side of prime numbers is relatively new (this question is still little addressed nowadays), it is the Author's opinion that time is ripe to face this topic with the necessary tools and that now there is room enough for treating primes even from this innovative and modern computational and experimental perspective [27-36] that could flank the celebrated classical ideas, also owing to the admirable successes got by computer experiments in the field of mathematics. Many mathematical problems are fundamentally inductive in nature, so that they can be addressed experimentally in order to determine how Bayesian inference, the logic of partial confidence, can be used in order to quantify a general law. Another basic idea in the background is a famous assertion of Hugh Montgomery who stated that, in the end, prime numbers seemed to behave just like experimental data. This is the viewpoint adopted and shown in the present report.

## 1.2 The Modified Chi-square Function

Thus, in the attempt to solve the above-told problematics as well as that on what is the actual behaviour of prime numbers as experimentally assessed and whether and how it can be approximated, in the present report an innovative approach is suggested starting from the experimental viewpoint and using the opposite inverse (-1/) (i.e. additive inverse and multiplicative inverse) of the modified Chi-square function ( $X$  of  $X_k^2(A, n/\mu)$  and Chi of Chi-square are pronounced *ki* just like in kinematics) in the form

$$-1 / X_k^2(A, n/\mu) = -1 / \{ [A / (2\Gamma_{k/2})] \times [n / (2\mu)]^{(k/2-1)} \times e^{-n / (2\mu)} \} \quad (1)$$

as the best-fit function along the whole study to match, from the analytical standpoint, both the finite sequences of prime numbers  $\{P_n\}$  and the truncated progressions  $\{C_\alpha n^\alpha\}$  having domain  $N$  and co-domain  $R^+$ , with  $(\alpha, k) \equiv (1^+, 0^-)$  and  $k = 2 - 2\alpha = k(n) = 2 - 2\alpha(n)$  as reported later on. The rationale underlying the entire issue has been to use equation (1) taking advantage of the adjustment of its three parameters  $k$ ,  $A$  and  $\mu$  which allows to optimize the fits as much as possible, up to 99.999% and even more whenever possible. In other words a *plot & fit* algorithm has been set up.

The original modified Chi-square function with  $k$  degrees of freedom

$$X_k^2(A, n/\mu) = [A / (2\Gamma_{k/2})] \times [n / (2\mu)]^{(k/2-1)} \times e^{-n / (2\mu)} \quad (2)$$

is a new general form of the standard Chi-square function  $X_k^2(n) = [1 / (2\Gamma_{k/2})] \times (n/2)^{(k/2-1)}$  [37,38] used also in statistics [39-41] with the two additional new parameters  $A$  and  $\mu$  where  $\Gamma_{k/2} = \Gamma(k/2)$  is the standard gamma function [42-44] the values of which can be easily found in the net [45,46]. It is easy to verify that we have  $\lim_{n \rightarrow \infty} k(n) = 0^-$  (see also Fig. 13) and to identify  $\mu = \mu(n)$  as a *decay parameter* for which the limits hold (see Fig. 14)

$$\lim_{n \rightarrow \infty} \mu(n) = +\infty \quad \lim_{n \rightarrow \infty} [n/\mu(n)] = 0$$

so that the further limit holds too

$$\lim_{n \rightarrow \infty} X_k^2[A, n/\mu(n)] = \text{constant} = \langle X_k^2 \rangle$$

As a matter of fact, one of the basic features of (2) is that in increasing the value of  $k$  (at  $A$  constant) towards  $2^-$  the value of  $\mu=\mu(k)$  increases more and more  $\rightarrow +\infty$  as already told and the  $X_k^2(A, n/\mu)$  function tends to become flatter and flatter (as the Author has thoroughly checked experimentally up to  $\mu \sim 1E300 = 1 \times 10^{300}$  and as can be easily verified analytically) until it becomes a constant function in the limit  $k \rightarrow 2^-$ . In addition, acting on the three parameters  $A$ ,  $k$  and  $\mu$  the function (1), as well as (2), sweeps the whole  $R^+$  plane ( $n, X^2$ ) thus showing a high degree of versatility, flexibility and usefulness. In fact the parameter  $A$  shifts rigidly the function (2) up and down along the  $X^2$  axis, the parameter  $\mu$  stretches and strains it along the  $n$  axis and the parameter  $k$  determines the shape of the function (it is interesting to underline that for  $k \gg 2$  i.e. at values  $k \sim 30 \div 40$  and beyond the modified  $X_k^2$  function tends to approximate a bell-shape, or normal or Gaussian function more and more). Furthermore, another distinctive feature that marks the pronounced difference between (2) and the standard Chi-square function is that  $\int X_k^2(A, x/\mu) dx = A \times \mu$  for the former, while  $\int X_k^2(x) dx = \int X_k^2(1, x/1) dx = 1$  for the latter, where both integrals run from 0 up to  $\infty$ , a result coming from the fact that the standard Chi-square function is just a particular case of the modified one with  $A=1$  and  $\mu=1$ .

It has already been shown by the same Author [47-50] that the function (2) in one of its four forms  $\pm(1 \times /) X_k^2(A, n/\mu)$  well approximates, i.e. interpolates and extrapolates, the finite sequences of the prime number frequencies  $\{f_n\} = \{n/P_n\}$  in the range  $\Delta k \equiv (1.5 \div 2^-)$  i.e.  $\Delta \alpha \equiv (-0.25 \div 0)$  and  $\{\rho_n\} = \{\ln(n)/\ln(P_n)\}$  in the lower left neighbourhood of the point  $(\alpha, k) \equiv (0^-, 2^-)$  in the  $R$  plane along the half-line with equation  $k=2+2\alpha$ . Furthermore the same function  $X_k^2[A, n/\mu]$  has been used to fit the statistics of prime numbers fitting their differential distribution functions [51]. Therefore in the present report it is the Author's aim to show also that the function (1), that is  $-1/X_k^2(A, n/\mu)$ , can fit the finite sequences  $\{P_n\}$  themselves and the same for the truncated progressions  $C_\alpha \times \{n^\alpha\}$  with  $\alpha = 1 - k/2$ , in the lower right neighbourhood of the point  $(\alpha, k) \equiv (1^+, 0)$ , each of these functions having the adequate values of its parameters, as shown later on. All that proves the great flexibility and effectiveness of these functions themselves and of the adopted methodology.

### 1.3 Fit Methodology and Statistics

Apart from the usual improvements of the statistical values of the fits, for any truncated progression or function  $C_\alpha \{n^\alpha\} - 1/X_k^2(A, n/\mu)$  or  $\{n \times \lambda_n \times \ln(n)\}$  the results in no way depend on the number of the terms, what is a consequence of the scale invariance of any progression itself, whilst for prime numbers just larger and larger finite sequences, subsets of their whole infinite sequence, have been examined, i.e. sequences of the kind:  $\{2 \ 3 \ 5 \ 7 \ \dots \ P_{h-1} \ P_h\} \equiv \{P_h\} \subset \{P_i\} \subset \{P_j\} \subset \dots \subset \{P_n\} \subset \dots$  being of course  $h < i < j < \dots < n < \dots$  and the reason for such an unconventional choice is the strict consequence of the scale non-invariance (i.e. scale variation) of the fits for the prime sequences and of scaling laws holding for them from the analytical viewpoint.

Any  $\{P_n\}$  sequence with a finite number of terms (i.e. primes) has been examined at  $n_\Delta$  (typically  $n_\Delta=200$ ) equally spaced values or data-points and it has been fitted at these  $n_\Delta$  data points by the modified Chi-square function in the form  $-1/X_k^2(A, n/\mu)$  with the appropriate values of the parameters  $k$ ,  $A$  and  $\mu$ . The same for the truncated progressions  $C_\alpha \times \{n^\alpha\}$  being  $k=2-2\alpha$  ( $\alpha > 0$ ) what means

$$\{P_n\} \approx -1/X_k^2(A, n/\mu) \approx -1/X_k^2(A, x/\mu) \approx C_\alpha \{n^\alpha\} \approx C_\alpha x^\alpha \quad k = 2 - 2\alpha$$

Speaking in a strict and formal way any  $C_\alpha \times \{n^\alpha\}$  progression can be analytically continued to the function  $f(x) = C_\alpha x^\alpha$  and also to the function  $g(x) = -1/X_k^2(A, x/\mu)$  with both functions analytic on the whole  $R$  plane. In other words, the real function  $-1/X_k^2(A, x/\mu)$  is an interpolating function, though within the chosen accuracy, of the  $C_\alpha \times \{n^\alpha\}$  progressions for non-integer  $x \in R$  just like the  $C_\alpha x^\alpha$  function and the same for the sequences  $\{P_n\}$  and the

function  $-1/X_k^2(A, n/\mu)$ . Finally, what has been done is nothing but a data smoothing, a procedure typical of any experimental discipline.

As a matter of fact it is easy to verify that it is enough to choose in (1)

$$-1/[A/(2\Gamma_{k/2})] \times (1/\mu)^{(k/2-1)} = C_\alpha = C_{(1-k/2)}$$

to get

$$-1/X_k^2(A, n/\mu) = -1/[A/(2\Gamma_{k/2})] \times [n/(2\mu)]^{(k/2-1)} \times e^{-[n/(2\mu)]} = C_{(1-k/2)} \times n^{(k/2-1)} \times e^{-n/(2\mu(k))} \approx C_\alpha \times n^\alpha$$

being  $n \ll \mu$  ( $\forall k$  and  $\forall n$ ) so that, for high values of  $n$ , one gets  $e^{-[n/(2\mu)]} \approx 1$  and  $k = 2 - 2\alpha$  ( $\alpha > 0$ ) as already told.

In the fit performed between the experimental data points, that is the actual values of  $P_n$ , and the theoretical fitting functions  $C_\alpha n^\alpha$  or  $n \times \lambda_n \times \ln(n)$  or  $-1/X_k^2(A, n/\mu)$  - once fixed the value of  $A$  (in the present study the value  $A=1E-70=1 \times 10^{-70}$  has been chosen once for all for convenience reasons) - the first concern has been to get the precise value of  $\alpha$  thus of  $k = 2 - 2\alpha$  (usually up to the 14<sup>th</sup> or 15<sup>th</sup> decimal digit that is, being  $\alpha \sim 1^+$  and  $k \sim 0^-$ , with a very good precision) by means of a *trial-and-error* procedure, letting  $\mu$  vary and, at any value of it, balancing the mean  $\langle F \rangle$  of the fit function to the mean  $\langle P_n \rangle$  of the actual values, up to the 14<sup>th</sup> or even the 15<sup>th</sup> decimal digit. In such a way the  $\mu$  value has been used to find the best fit. At the same time both the correlation (or Bravais-Pearson) coefficient  $R = {}_{BP}R = R(P_n, F)$  and the non-linear index of correlation  $I = I(P_n, F)$  between the actual values or *counts*  $P_n$  and the fit values  $F$  have been calculated and maximized up to the value 0.9999.... during the variation of  $\mu$ . In all the formulas the sums extend from 1 up to  $n_\Delta = 200$  the number of the cells of the spread-sheet commonly used for the calculations. This value of 200 cells has been chosen in order to avoid a useless burdening of the calculations and it has been recognized to be a good compromise between speediness and efficiency of the calculations. Maximizing the two statistical markers  $R$  and  $I$  means making both of them to approach the value of  $1^- \sim 0.99999...$  as much as possible by adjusting the value of  $\mu$  for any value of  $k$  in order to match the fit curve and the  $P_n$  points as much as possible, as well as to balance their average values  $\langle P_n \rangle = \langle F \rangle$  up to the 14<sup>th</sup> or 15<sup>th</sup> decimal digit. In addition, also the two standard deviations of the means  $\sigma_{primes} = \sigma_{P_n}$  and  $\sigma_F = \sigma_{fit}$  have been examined in order to ascertain that they would be approximately equal and that each of them would be much lower than its respective mean  $\langle P_n \rangle$  and  $\langle F \rangle$ . Finally, two further gauges of the fits have been minimizing the values of the *Least Square Sum* (LSS) according to the principle of the maximum likelihood and minimizing the value of the *Chi-square Test*, in that both of these variables measure the goodness of the fit. Just to summarize, most of the tools available in statistics and probability have been implemented in order to make the best possible fits at the utmost statistical reliability and consistency. However it should be taken into account that, in optimizing all these gauges and markers (at least seven), some compromise has always had to be made which, nonetheless, in no way and in no case has ever endangered the reliability of the results but just influenced occasionally, though weakly, their precisions. Though complex, cumbersome and time-consuming, this fitting procedure has proven all its effectiveness and usefulness in finding out not only the best values of  $\alpha$ ,  $k$  and of the decay parameter  $\mu$  but even the fundamental relationships  $k=k(\alpha)=k_\alpha$ ,  $k=k(n)=k_n$ ,  $\mu=\mu(n)=\mu_n$ ,  $C_\alpha=C_\alpha(n)$ ,  $\alpha=\alpha(n)=\alpha_n$ ,  $\lambda=\lambda(n)=\lambda_n$  and so on for the progressions  $-1/X_k^2(A, n/\mu)$ ,  $C_\alpha(n^\alpha)$ ,  $n \times \lambda_n \times \ln(n)$  which fit the prime sequences  $\{P_n\}$ .

The precisions, error sources, error propagations and the consistency of the results have been investigated too, being these issues crucial to the whole algorithm. After all, what has been done is just what is usually done in treating experimental raw data, a procedure that is common to all the fields of experimental physics and all the other experimental disciplines. The only difference has been to treat prime numbers themselves just as experimental data in a broad sense, to which all these concepts and criteria can be applied, with the further undisputable difference from standard experimental data (always written as  $W \pm \delta W$ ) and the advantage of having zero *inaccuracy* (i.e. no *systematic* errors) and zero *imprecision* (no *random* errors) on the base data (being of course  $P_n = P_n \pm \delta P_n = P_n \pm 0 = P_n$ ) whilst both inaccuracies and imprecisions are present on the final results owing to the approximations of

the fits. In such a manner a computer simulation model has been set up and implemented to the finite prime sequences just like a computer simulation of any physical effect or phenomenon. Thus, all the calculations, as well as all the final results and considerations, can be applied to the model itself (instead of the original mathematical objects  $P_n$ ) that is on the three fit functions  $-1/X_k^2(A, n/\mu)$ ,  $C_\alpha n^\alpha$  and  $n \times \lambda_n \times \ln(n)$ . Consequently the whole issue can be simplified a lot and the entire algorithm and methodology can be of valuable help to try to provide solutions to some of the problems related to prime numbers.

In such a manner, by means of these approximation techniques, for most cases it is possible to reduce the problem of prime numbers simply to a problem of mere precision of the calculations, though still hard and not yet fully solved, however no more intractable. Moreover, having chosen an empirical and pragmatic approach to the problem and having adopted the viewpoint of computational experimental mathematics that is of a numeric computer simulation instead of a formally rigorous and stringent theoretical approach, it is plain to remark that the computer itself has been a fundamental tool in performing all the calculations on a very large data base of raw data, amounting to  $N = n_{\max} = 2E15 = 2.0 \times 10^{15} = 2P = 2\text{Peta}$  primes processed analytically from  $P_1 = 2$  through  $P_{2P} = P(2E15) \approx 7.5674484987353E+16$ . The PC used is of a commercial kind with about 700 GB HD memory, 8 GB RAM and the SW used can support calculations up to  $1E308 = 1 \times 10^{308}$  with 15 decimal digits ( $\pi$  value) though in some cases the precision has been limited to 12 decimal digits as a maximum for simplicity's sake. A wide use has been made of standard well-known commercially available scientific computer codes and software applications, most of all spreadsheets, as well as of open-source software applications available in the net and finally even of some free websites which have proven to be very useful either to get the values of the parameter  $\Gamma_{k/2}$  [45,46] as already told or as prime sources [52,53] or simply for documentation.

One of the basic backgrounds of the study has been the use of the principle of extrapolation and interpolation of the experimental data that has been widely used all over the article whenever required and possible, even though not explicitly cited, a principle widely used in all experimental science, its counterpart being the *induction principle* in mathematics. This mathematical principle can be used to prove many theorems as for instance the well-known statement of the summation of all the natural numbers  $n$  up to  $N$ :  $\sum_{n=1}^N n = N \times (N+1) / 2$ . It is interesting to remark that, in this case too, like in Stirling's approximation, one gets the solution to the problem in a so-called *closed form*.

As a final consideration, it is well-known that there are no large primes, in that any prime number can be defined *small* in comparison to the infinitude of all the primes. But, though always condemned to deal with small numbers, now we can at least try to treat prime numbers much bigger than in the past (up to  $n=2E15$ ) so that the induction principle can be freely implemented without any constraint. In addition the situation will grow better and better in the future owing to the two-pronged computational progress both in the technology, i.e. hardware features, and in the algorithm, i.e. software development.

## 2. FINITE SEQUENCES OF PRIME NUMBERS

The analytical aspects concerning prime number sequences  $\{P_n\}$  examined in the frame of experimental mathematics represent a first attempt to get an algorithm of the kind  $P_n \approx P(n)$ , if any, for the construction of a prime  $P_n$  starting from its counter  $n$  though with the due approximations, or alternatively an algorithm of the type  $P_n \rightarrow P_{n+1}$  for which it is possible to forecast a prime from the knowledge of its preceding one.

Beginning from the alleged relationships i.e. fits

$$P_n \approx P(n) \approx -1/X_k^2[A, n/\mu] \approx C_\alpha n^\alpha \approx n \times \lambda_n \times \ln(n) \quad k = 2 - 2\alpha \quad (\alpha, k) \equiv (1^+, 0)$$

the plot of the actual values of  $P_n$  (not shown) displays fluctuations for the first few hundreds of primes  $n_0 < \approx 1,000$  though almost disappearing at higher values, henceforth displaying a much more regular trend. It is just this regularity that leads to examine the trend of  $P_n$  versus

$n$  that is  $P_n \approx P(n)$ . The algorithm previously shown has been applied: i.e. choosing larger and larger prime sequences  $\{P_h\} \subset \{P_i\} \subset \{P_j\} \subset \dots \subset \{P_n\}$ , with  $h < i < j < \dots < n$ , fitting any finite sequence by the function  $-1/X_k^2(A, n/\mu)$  where  $k=0$  as well as by the function  $C_\alpha n^\alpha$  with  $\alpha=1^+$  and by the function  $n \times \lambda_n \times \ln(n)$  too and plotting them versus  $n$ . The example of the next figure 5 shows that the sequence of the first  $1.5 \times 10^6$  prime numbers  $\{P_{1.5M}\} = \{2, 3, 5, 7, \dots, 23,879,519\}$  is indistinguishable from these three fit functions. As a matter of fact this  $P_n$  sequence can be best fitted by the function  $-1/X_k^2[A, n/\mu]$  with the parameter values  $A=1E-70$ ,  $\mu=2.83453022475825E+65$ ,  $k=-0.15067569115100$ ,  $\Gamma_{k/2}=-13.9308749425917000$  and the fit values  $R=0.9999971560$ ,  $I=0.9999888120$ ,  $\langle P \rangle = \langle F \rangle = 11,596,146.05000$ ,  $\delta_{mean} = (\langle P \rangle - \langle F \rangle) / \langle P \rangle = -6.51E-15$ ,  $test\text{-}value\ X^2 = 8.87E+03$ ,  $LeastSquareSum = LSS = 1.108E-03$  while for the fit by the  $C_\alpha n^\alpha$  function the values are  $C_\alpha = 5.467256340728670$ ,  $\alpha = 1.07533784557550$  and the fit parameters  $R=0.999997156$ ,  $I=0.999988812$ ,  $\langle P \rangle = \langle F \rangle = 11,596,146.050$ ,  $\delta_{mean} = 1.9275E-15$ ,  $test\text{-}value\ X^2 = 8.889E+03$ ,  $LSS = 1.108E-03$  that is the same of the fit by the function  $-1/X_k^2[A, n/\mu]$  what is to be expected owing to what aforesaid. Finally the fit by the function  $n \times \lambda_n \times \ln(n)$  shows the feature  $\lambda_n = 1.121041116691750$  with the fit parameters  $R = 0.999995376$ ,  $I = 0.999954322$ ,  $\langle P \rangle = \langle F \rangle = 11,596,146.0500$  their  $\delta_{mean} = -4.176E-13$ ,  $test\text{-}value\ X^2 = 5.20766E+03$ ,  $LSS = 4.528E-4$ . A significant feature of all these three fits is the trend of the errors that is the differences  $\delta = (P_n - fit) / P_n$  versus  $n$  as shown in the figure 6.

A zoom-in view of the figure 5 shows again the small difference of the three fit functions among them and the 1.5M prime sequence (Fig. 7) while, as for the errors  $\delta$  between any fit and the actual sequence  $\{P_{1.5M}\}$ , the figure 8 (zoom-in of Fig. 6) shows that the first two fits,  $-1/X_k^2(A, n/\mu)$  and  $C_\alpha n^\alpha$ , work much better than the third one  $n \times \lambda_n \times \ln(n)$  in simulating the actual data-points  $P_n$  (1m means  $1E-3 = 1 \times 10^{-3}$ ).

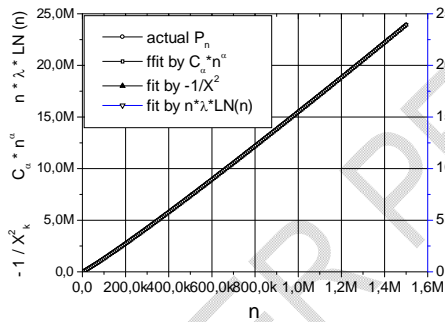


Fig. 5 The sequence of the first 1.5M primes and fits

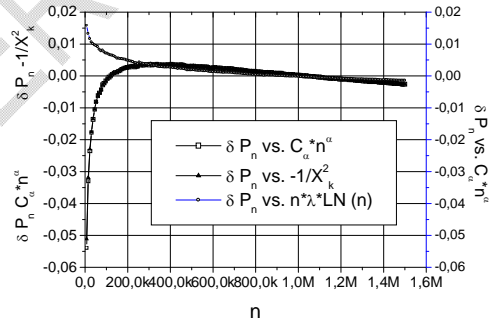


Fig. 6 The errors  $\delta$  of the fits for the first  $\{P_{1.5M}\}$  and fits

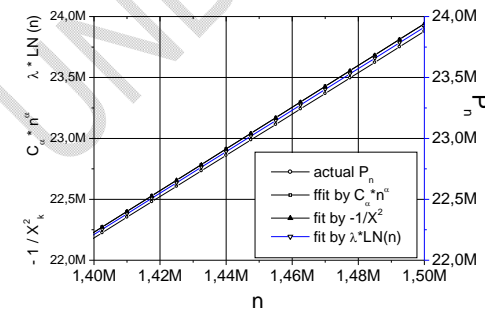


Fig. 7 Zoom-in of Fig. 5 between 1.40M and 1.50M

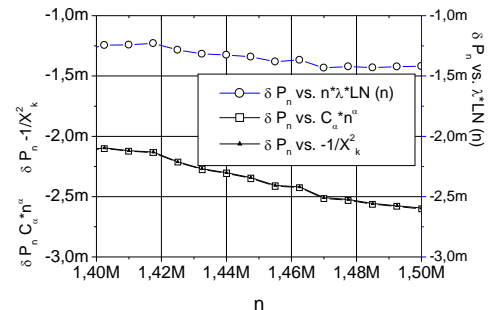


Fig. 8 Zoom-in of Fig. 6 between 1.40M and 1.50M

An important comment has to be made about this example concerning also all the other cases and the whole methodology of the study. It has already been pointed out that this

instance takes into account the first 1.5M prime numbers at just 200 points, that is  $n_{\Delta}=200$  values - one out of  $\Delta=7,500$  - in that it is not possible to deal with all the first 1.5M primes, as well as with all the first 1G, all the first 10T=1E13 and so on for all the other sequences, due to obvious reasons of computer and spread-sheet memory. Thus a choice has been made, i.e. that of choosing just 200 prime numbers for any sequence, from the first ones up to the final  $P_n$  of the sequence, that is up to  $P_{1.5M}$ ,  $P_{1G}$ ,  $P_{10T}$  and so forth on which all the calculations have been executed. It has been thoroughly checked that this choice by no means has ever altered the final values or the reliability of the final results and most important of all the trends, as careful comparisons between 200  $P_n$  and all the  $P_n$  for some initial sequences  $\{P_{50K}\}$   $\{P_{100K}\}$   $\{P_{200K}\}$   $\{P_{0.5M}\}$  and so on have shown. Just little differences in some parameters have been found, non-essential to the concluding results and trends.

A second example can be shown regarding the sequence  $\{P_{50G}\}$  that is the first  $50G=5 \times 10^{10}=5E10$  prime numbers up to  $P_{50G}=1,344,326,694,119$ . In this case too, the sequence of the first actual  $5 \times 10^{10}$  prime numbers  $\{P_{50G}\}$  is indistinguishable from the three fit functions as the figure 9 shows. As a matter of fact this sequence can be best fitted by the function  $-1/X_k^2$  with parameter values:  $A=1E-70$  decay parameter  $\mu=4.123236900808860E+67$   $k=-0.082209819607980$   $\Gamma_{k/2}=-24.94746914106250$  and the fit values  $R=0.999999739$   $I=0.999998987$   $\langle P \rangle = \langle F \rangle = 6.625089876368810E+11$   $\delta_{mean}=-1.283E-15$   $_{test-value}X^2=4.216E7$   $LSS=1.10E-04$  while for the fit by the  $C_{\alpha} \times n^{\alpha}$  function the values are  $C_{\alpha}=9.7745299527836700$   $\alpha=1.041104909803990$  so that one gets for the calculated  $P_n \approx C_{\alpha} \times n^{\alpha} \approx 9.7745299527836700 \times n^{1.041104909803990}$  with the fit parameters  $R=0.999999738$   $I=0.999999002$   $\langle P \rangle = \langle F \rangle = 6.625089876368810E+11$   $\delta_{mean}=-1.105527E-15$   $_{t-v}X^2=4.319290E+07$  and finally the value of the Least Square Sum  $LSS=9.917313E-05$ .

Lastly the fit by the function  $\lambda_n \times (n) \times \ln(n) = \lambda_{50G} \times (50G) \times \ln(50G)$  gives the value  $\lambda_n = 1.0924154306022810$  with the parameters  $R=0.999999811$   $I=0.999998031$   $\langle P \rangle = \langle F \rangle = 6.62508987636880E+11$  relative difference between the means  $\delta_{mean}=-4.422E-15$   $_{t-v}X^2=1.25303E+08$   $LSS=1.955E-04$ . Again, one of the noteworthy features of these three fits is the trend of the errors that is of the relative differences  $\delta = (P_n - fit) / P_n$  versus  $n$  as shown in Fig. 10, on the other hand common to all the sequences examined.

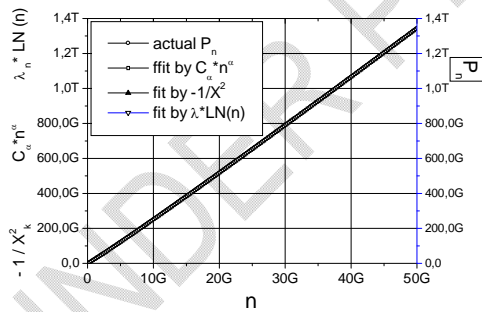


Fig. 9 The sequence of the first 50G primes and fits

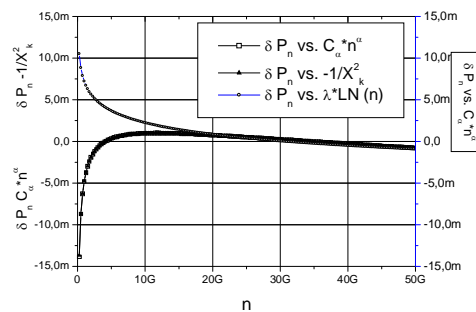


Fig. 10 The errors  $\delta$  of the three fits for the first  $\{P_{50G}\}$

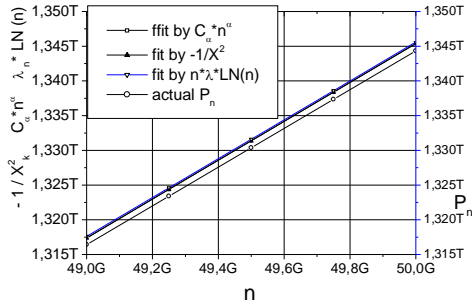


Fig. 11 Zoom-in of Fig. 9 between 45G and 50G

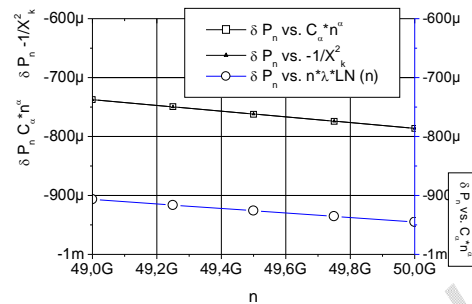


Fig. 12 Zoom-in of Fig. 10 between 45G and 50G

In this example too, the previous figures 11 (zoom-in view of Fig. 9) and 12 (zoom-in of Fig.10), both between 49G and 50G, show that the first two fit functions  $-1/X_k^2[A, n/\mu]$  and  $C_n n^\alpha$  work much better than the third one  $n \times \lambda_n \times \ln(n)$  as already highlighted.

Of course these are just two examples, though representative of the whole situation, while many additional cases have been examined for the purposes of the calculations, approximately 90 as already remarked. In all these situations the best values of the fit parameters for the best approximations have been found, thus assuring the goodness of the whole practice.

However it should be kept in mind that all the equations hold locally (i.e. just for the nearby interpolations and extrapolations) in that, owing to the scale variation of the fits, as a general rule all the coefficients of the fits themselves depend upon  $n$ , as thoroughly checked and shown later on. Nonetheless the most important fact is that in such a way one owns a relationship, though approximate, by means of which one can estimate the value of a prime number, the precision of which can be improved as much as reasonably achievable with future evaluations and investigations and most of all by means of more powerful computers as well as taking into account all the primes of any single sequence and not just 200 as done here for computer memory reasons.

In addition, after having calculated the fits for a reasonable number of prime sequences, the trend of the function  $-1/X_k^2[A, x/\mu(k)]$  vs.  $A$  has been studied too, with  $A$  ranging from the value  $A = 1E-260$  up to  $A = 1E-7$  in 13 steps, finding that there is a linear law between  $\mu$  and  $A$  on a log-log scale, i.e.  $\lg(\mu) = 0.35907 - 0.92769 \times \lg(A)$  with  $R = 1.000...$  up to the 12<sup>th</sup> decimal digit and  $\sigma = 0.00504$  for 13 data-points. Of course in this case too the value  $R = 1.000...$  up to the 12<sup>th</sup> digit means that one has reached the precision of the computer code for that calculation. All the other parameters are just slightly influenced by the change of the coefficient  $A$ .

Thus, simply using the above reported relationships and fit functions, the problem of prime numbers can be reduced, at least in principle for their deterministic aspect, to a mere problem of precision, as well as of data interpolation and extrapolation (taking into account the scale non-invariance problem) though of not so immediate solution. It is possible to conclude that there are many ways, just some of which shown here, to get an approximation of the  $\{P_n\}$  finite sequences as well as an estimate of the finite value of a prime number  $P_n$  starting from its own ordinal number  $n$  by the algorithm shown that uses systematically the above-told functions as fitting functions. Nevertheless one of the major findings of this study is the scale variation of the fits of the finite sequences of prime numbers, a remarkable result.

The next step of the study is quite trivial, i.e. finding the trend of all the fit function parameters after having examined a sound number of prime sequences (about 90) and their fits. Starting from the modified Chi-square function  $-1/X_k^2[A, x/\mu(k)]$  the trend of its three

parameters  $k$ ,  $\mu$  and  $\Gamma_{k/2}$  vs.  $n$  is shown in the next three figures 13, 14 and 15. The first one (Fig. 13) clearly shows that  $\lim_{n \rightarrow \infty} k(n) = 0^-$  while Fig. 14 seems to suggest the  $\lim_{n \rightarrow \infty} \mu(n) = \infty$  and finally the linear trend of  $\Gamma_{k/2}$  vs.  $\ln(n)$  is visible in Fig. 15 with the  $\lim_{n \rightarrow \infty} \Gamma_{k/2}(n) = -\infty$  while the figure 16 describes the trend of  $\alpha(n)$  vs.  $\ln(n)$  with the evident limit  $\alpha(n)_{n \rightarrow \infty} \rightarrow 1^+$ . At last the figure 17 describes the trend of  $C_\alpha(n)$  vs.  $\ln(n)$  with the  $\lim_{n \rightarrow \infty} C_\alpha(n) = \infty$  while Fig. 18 reports the trend of  $\lambda(n)$  with the  $\lim_{n \rightarrow \infty} \lambda(n) = 1^+$  so that the asymptotic limit exists  $\lim_{n \rightarrow \infty} \lambda_n \times n \times \ln(n) = 1 \times n \times \ln(n)$  that is the classical prime number theorem. All these figures refer to the first 2E15 primes with  $\ln(2E15) = 35.2319235754\dots$  and  $\log(2E15) = 15.3010299\dots$

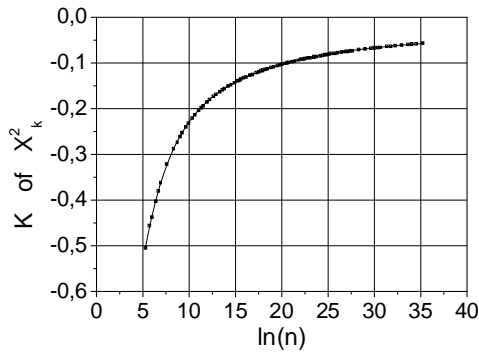


Fig. 13 Trend of  $k(n)$  for all the first 2E15 primes

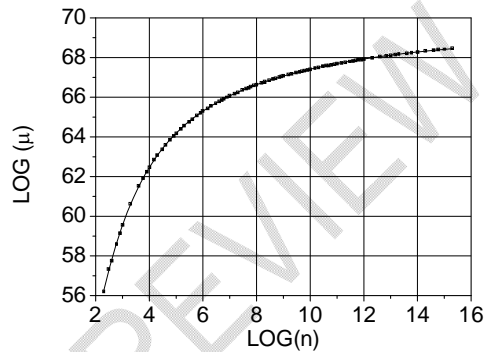


Fig. 14 Trend of  $\mu(n)$  for all the first 2E15 primes

Thus, according to the fits, one gets for the features of the function  $-1/X^2_k[A, x/\mu(k)]$  the following relationships.

- As for the trend of the parameter  $k$  versus  $\ln(n)$  (Fig. 13) the fit is:

$$k(n) \approx -(2.16 \pm 0.04) + (1.76 \pm 0.03) \times [1 - e^{\ln(n)/(2.64 \pm 0.04)}] + (0.36 \pm 0.01) \times [1 - e^{\ln(n)/(11.28 \pm 0.29)}]$$

that is the so-called exp-assoc function (as given by the PC ad-hoc software) with the values of the fit markers  $R^2 = 0.9993$  and  $\text{test-value } X^2 = 6.9E-7$

- As for the trend of the parameter  $\mu$  vs.  $\log(n)$  (Fig. 14) one gets the fit:

$$\log[\mu(n)] \approx \log\left\{ (24.831 \pm 0.606) + (34.456 \pm 0.369) \times [1 - 10^{\log(n)/(1.359 \pm 0.024)}] + (9.813 \pm 0.279) \times [1 - 10^{\log(n)/(5.634 \pm 0.161)}] \right\}$$

that is again the exp-assoc fit function with the fit markers  $R^2 = 0.99995$  and  $\text{test-value } X^2 = 0.00037$

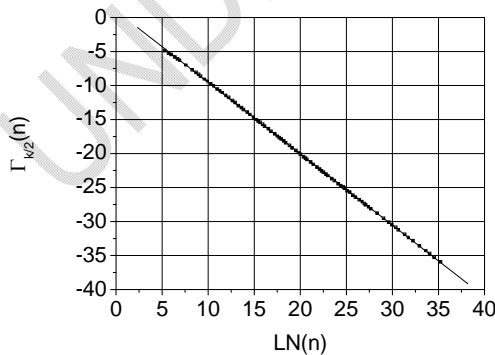


Fig. 15 Trend of  $\Gamma_{k/2}(n)$  for the first 2E15 primes

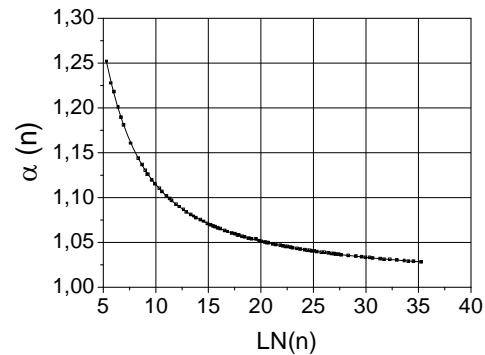


Fig. 16 Trend of  $\alpha(n)$  for the first 2E15 primes

- Finally the trend of  $\Gamma_{k/2}(n)$  vs.  $\ln(n)$  (Fig. 15) is simply weakly quadratic i.e.  $\Gamma_{k/2}(n) \approx (1.032 \pm 0.039) - (1.058 \pm 0.004) \times \ln(n) + (1.90 \pm 1.06) E-4 \times \ln^2(n)$  with  $R^2=0.99994$  and  $t_v X^2=0.06419$

In such a way, combining all the above-told fitting functions it is possible to get a relationship linking the finite value of a prime number  $P_n$  to its counter  $n$ , i.e.  $P_n \approx P(n)$ , though approximate and rather complex.

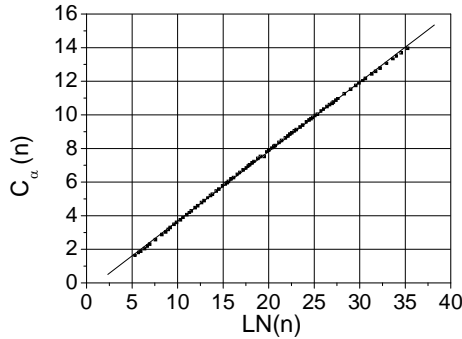


Fig. 17 Trend of  $C_\alpha(n)$  for the first 2E15 primes

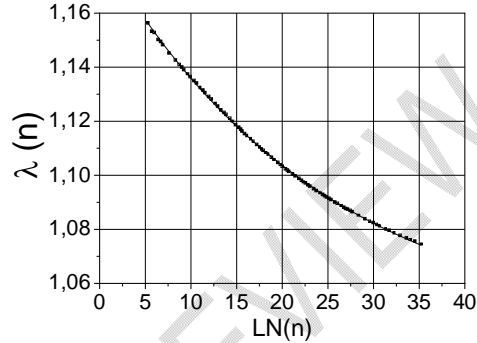


Fig. 18 Trend of  $\lambda(n)$  for the first 2E15 primes

The same procedure applied to the other fitting function leads to

- $\alpha(n)$  vs.  $n$  (previous Fig. 16) is best fitted by an ExpDecay3 function (again as given by the PC ad-hoc code) in  $\ln(n)$ , that is  $\alpha(n) \approx 1.000 + (0.21 \pm 0.03) x e^{-\ln(n)/(7.3 \pm 1.)} + (0.897 \pm 0.024) x e^{-\ln(n)/(2.37 \pm 0.12)} + (0.058 \pm 0.011) x e^{-\ln(n)/(44.36 \pm 8.70)}$  with the fit parameters  $R^2 = 0.99993$   $test-value X^2 = 1.67 E-7$  and with  $\lim_{n \rightarrow \infty} \alpha(n) = 1.000$
- $C_\alpha(n)$  vs.  $\ln(n)$  (previous Fig. 17) is best described by the following weakly quadratic fit function:  $C_\alpha(n) \approx -(0.785 \pm 0.012) + (0.4532 \pm 0.0013) \times \ln(n) - (9.965 \pm 0.332) E-4 \times \ln^2(n)$  with  $R^2=0.99996$  and  $\sigma=0.020$

Thus, in this case too, one gets an approximate formula for the finite value of a prime number vs. its counter  $n$ . Combining the first fit of  $C_\alpha(n)$  and the second fit of  $\alpha(n)$  one gets:  $P_n \approx C_\alpha(n) \times n^{\alpha(n)} \approx C_\alpha(n) \times n^{[ \alpha(n) ]} \approx [ -(0.785 \pm 0.012) + (0.4532 \pm 0.0013) \times \ln(n) - (9.965 \pm 0.332) E-4 \times \ln^2(n) ] \times n^{[ 1.000 + (0.21 \pm 0.03) x e^{-\ln(n)/(7.3 \pm 1.)} + (0.897 \pm 0.024) x e^{-\ln(n)/(2.37 \pm 0.12)} + (0.058 \pm 0.011) x e^{-\ln(n)/(44.36 \pm 8.70)} ]}$  while taking into account the limit  $\lim_{n \rightarrow \infty} \alpha(n) = 1$  and the weakly quadratic fit of  $C_\alpha(n)$  one gets the asymptotic relationship  $P_n \approx [ -(0.785 \pm 0.012) + (0.4532 \pm 0.0013) \times \ln(n) - (9.965 \pm 0.332) E-4 \times \ln^2(n) ]$

As for the third fit function  $P_n \approx \lambda_n \times n \times \ln(n)$  the fitting methodology leads to the result (Fig. 18):

- $\lambda_n \approx (1.04 \pm 7 E-4) + (0.1445 \pm 5 E-4) x e^{-\ln(n)/(24.50 \pm 0.25)}$  with  $R^2 = 0.99979$   $t_v X^2 = 1.1 E-7$  and the approximate limit  $\lambda_n \rightarrow 1$  for  $n \rightarrow \infty$  so that, in this case, one gets the standard PNT  $P_n \approx n \times \ln(n)$  asymptotically.

As a conclusion one can state that, though cumbersome, awkward and time wasting these procedures work very well, as already told and also checked, to provide an approximate result  $P_n \approx P(n)$ .

### 3. RESULTS AND DISCUSSION

One of the main results of the research is the correspondence between prime numbers and any of the three fit functions employed. Thus the next figure 19 summarizes the whole situation showing the relationships among the finite sequences of prime numbers  $\{P_n\}$  and



them leading to about the same approximate value, among which one chooses the best fit-function that is the function leading to the best-fit.

However it must be highlighted that in this case one knows the value of the prime  $P_n$  *a priori*. Accordingly, an important fact of the present study is that a methodology has been found capable to *simulate* a prime number that is to *replicate* it at the maximum level of precision and that this simulation can be attained in many ways. As a matter of fact for the prime  $P_{8.00E10}$  the following *simulation* holds by the modified Chi-square function with the parameters of the fit:  $A=1E-70$   $k=-0.079042807596259900$   $\Gamma_{k/2}(n)=\Gamma_{(0.039521403798130000/2)}(n)=\Gamma_{(0.039521403798130000/2)}(8.00E10) = -25.920530013642100$   $\mu=5.235340400213330E+67$  with the final result  $P_{8.00E10} \sim -1/X_k^2[A,n/\mu(k)]=2,190,026,988,348.820$  and an *error* i.e. imprecision in comparison to the actual prime  $P(8.00E10) = 2,190,026,988,349$  equal to  $\delta\% = -8.427764291578800E-12\%$  and absolute difference  $\delta=0.180$ , another good result indeed. Again in this case too there are many ways to get a profitable fit by the modified Chi-square function with different parameters. Finally, the fit by the function  $P_n \approx \lambda_n \times n \times \ln(n)$  gives the values:  $\lambda_{8.00E10}=1.090420969415450$   $\ln(8.00E10)=25.105292471620300$  and the final value of  $\lambda_{8.00E10} \times n \times \ln(8.00E10)=2,190,026,988,349.0100$  with an absolute difference  $\delta=0.0100$  and a percentage error of  $\delta = -3.678786000292E-13\%$ .

Another example examines the prime  $P_n = P_{(477M+741,961)} = P_{477,741,961}$  which is known, from the net, actually to be equal to 10,523,089,897. A simulation by the modified Chi-square function  $-1/X_k^2[A,n/\mu(k)]$  with the fit values of  $A = 1E-70$   $k = 0.1008262675847200$   $\Gamma_{k/2}(n) = -20.46561661688820$  and  $\mu = 1.028375774310410E+67$  gives the value calculated  $P_n = P_{477,741,961} = 10,523,089,897.000$  with a relative difference between the actual prime and the calculated one equal to  $-1.279650890E-13$  (as given by the spread-sheet).

Thus, knowing the value of  $n$  and  $P_n$  it is very easy to *simulate/reproduce* a prime number by the reported method and algorithm, however just knowing its value *a priori* and it is clear and evident that there are many ways to *mimic* i.e. *duplicate* the value of a known prime number. However the question arises: what if one does not know the value of a prime  $P_n$  but just its counter  $n$ ? How to *simulate*  $P_n$  only from the knowledge of  $n$ ? Well, the matter is much more critical and complex so that a thorough extrapolation calculation must be performed taking into account all what already said.

Let's make an example using the simplest and most precise fit by the function  $C_\alpha(n) \times n^\alpha$ . The use of the fits of  $\alpha$  vs.  $n$  (Fig. 16) and  $C_\alpha$  vs.  $n$  (Fig. 17) and their relative equations will be of some importance. Let us consider just one simple case, that of the 1,000,000,000,000<sup>th</sup> prime number i.e.  $actual P_{1E15} = 37,124,508,045,065,437$  (one of the highest prime examined in the study) the calculated value of which is (by means of the  $C_\alpha \times n^\alpha$  approximation)  $C_\alpha n^\alpha = C_\alpha \times n^\alpha \approx [-(0.46 \pm 0.02) + (0.414 \pm 0.001) \times \ln(n)] \times n^\alpha [1 + (0.06 \pm 0.01) \times n^{-1/(44. \pm 9.)} + (0.90 \pm 0.02) \times n^{-1/(2.37 \pm 0.13)} + (0.21 \pm 0.03) \times n^{-1/(7.3 \pm 1.0)}] \approx [-(0.46 \pm 0.02) + (0.414 \pm 0.001) \times \ln(n)] \times n^\alpha [1 + (0.90 \pm 0.02) \times n^{-1/(2.37 \pm 0.13)}] \approx [-0.46 + 0.414 \times \ln(n)] \times n^\alpha [1 + 0.90 \times n^{-1/2.37}] \approx [-0.46 + (0.414) \times \ln(n)] \times n^\alpha [1 + (0.06) \times n^{-1/(44.)} + (0.90) \times n^{-1/(2.37)} + (0.21) \times n^{-1/(7.3)}] \approx 39,228,250,270,115,900$  with a relative difference between the two values (actual and calculated) =  $\delta\% = 5.667\%$  and an absolute difference  $\delta = 2,103,742,225,050,447$  without taking into account the single errors on the coefficients. No doubt that this is a huge error for the calculation of a prime number what means that surely this procedure is not suitable. In fact it has to be kept in mind that for a function  $f = f(x, y, z, \dots)$  with any independent variable  $x, y, z, \dots$  affected by an error  $\delta x, \delta y, \delta z, \dots$  the final overall error on the dependent variable is  $\delta f(x, y, z, \dots) = (\partial f / \partial x) \delta x + (\partial f / \partial y) \delta y + (\partial f / \partial z) \delta z + \dots$  where  $\partial f / \partial \dots$  are the partial derivatives, what shows that the propagation of the errors leads to an enhancement of the final error. In the fits of  $C_\alpha$  vs.  $n$  and of  $\alpha$  vs.  $n$  all the uncertainties are  $\pm$  some% so that the final result  $C_\alpha \times n^\alpha$  is affected by an even greater error. Thus this latest methodology is not appropriate to estimate the value of a prime number  $P_n \approx P(n)$  just starting from the counter  $n$ , so that another technique must be implemented keeping in mind all what already told about the local validity of the algorithm. Nonetheless, all what told until now has not been useless **in that** the features of

the fit functions used will be useful to forecast the value of a prime  $P_{n+1}$  from its previous (already known)  $P_n$ .

The procedure is implemented in such a way:

- 1<sup>st</sup> step. To fit the single  $n^{\text{th}}$  prime  $_{\text{act.}}P(n)$ , the value of which is known, at the utmost level of precision i.e. with the least possible error, by the chosen fit function (for instance  $C_\alpha \times n^\alpha$  the simplest one) thus finding the calculated value  $_{\text{calc.}}P_n \approx C_\alpha \times n^\alpha$ . This procedure gives also the value of both  $C_\alpha$  and  $\alpha$ .
- 2<sup>nd</sup> step. To extrapolate the value  $\alpha(n)$  to  $\alpha(n+1)$  starting from  $_{\text{act.}}P_n \approx_{\text{calc.}}P_n \approx C_\alpha \times n^\alpha$  in order to calculate  $_{\text{calc.}}P_{n+1} \approx C_\alpha \times (n+1)^\alpha$  so finding the approximate value of the latter  $_{\text{calc.}}P_{n+1}$  from the former, using the extrapolation of the fit function from  $n$  to  $n+1$ , thus performing the operation  $P_n \rightarrow P_{n+1}$  by means of the local validity of the algorithm.

Some examples can better exemplify this technique.

The first instance examines the prime number  $_{\text{act.}}P_{82.5M} = P(8.25E7) = P(8.25 \times 10^7) = 1,664,674,813$  (=actual value) while the calculated/simulated value is  $P(8.25 \times 10^7) \approx C_\alpha n^\alpha \approx 7.88508741692040 \times (8.25E7)^{1.0515440050765500} = 1,664,674,813.000110$  with least difference (that is percentage error  $\% \delta = 6.645515378691860E-12\%$  and absolute error  $\delta = 0.00011$ ). Extrapolating the fit function  $C_\alpha n^\alpha \rightarrow C_\alpha \times (n+1)^\alpha$  that is from  $P(82,500,000)$  to  $P(82,500,001)$  gives the values  $P_{82.5M+1} = C_\alpha \times (n+1)^\alpha = 7.88508741692040 \times (82,500,001)^{1.0515440050765500} = 1,664,674,834.218040$  with an absolute error in comparison to the actual prime  $_{\text{actual}}P(82,500,001) = 1,664,674,819$  of  $\delta P(8.25 \times 10^7 + 1) = 15.21803808212280$  i.e.  $\approx 15$  and a  $\% \delta = 9.141748231203830E-07\%$  an error rather high nonetheless still acceptable.

The second example consists in investigating the prime  $P_{300M} = P(3E8) = P(3 \times 10^8) = 6,461,335,109$  with the next actual prime  $P(3 \times 10^8 + 1) = 6,461,335,171$ . Here too, one gets  $C_\alpha n^\alpha = C_\alpha \times (3E8)^\alpha = 7.88508741692040 \times (300M)^{1.051476353871230} = 6,461,335,108.999990$  with a % error  $\% \delta = 8.85582592747319E-14\%$  and absolute difference  $\delta = 0.00001$  while extrapolating the fit function  $C_\alpha n^\alpha \rightarrow C_\alpha \times (n+1)^\alpha$  from  $P(300M)$  to  $P(300M+1)$  provides  $P_{300M+1} = C_\alpha \times (n+1)^\alpha = 7.88508741692040 \times (300M+1)^{1.051476353871230} = 6,461,335,131.6464700$  with the absolute error (in comparison to the actual  $P_{n+1}$ ) of  $\delta P(3 \times 10^8 + 1) = 39.3535 \approx 39$  and a % discrepancy =  $6,09061813390E-07\%$  again a rather high value though tolerable.

Also in studying the prime  $P_{10G+1} = P(1E^{10} + 1)$  from  $P_{10G} = P(1E^{10})$ , the third example, one can see that in effect any prime number of the sequence can be fitted i.e. simulated by a value of any of the three fit functions and that in this calculation one uses the prime sequence but only and just to extrapolate from the known prime to the interested prime itself that is from  $P_n$  to  $P_{n+1}$ .

Thus one gets for  $n = 1E10$  the actual value  $P(1E10) = 252,097,800,623$  while the calculated value is  $C_\alpha n^\alpha = C_\alpha \times (1E10)^\alpha \approx 9.125800459342760 \times n^{1.0441298088325100} \approx 252,097,800,623.0010$  with a relative error  $(P_{\text{actual}} - P_{\text{calculated}}) / P_{\text{actual}}$  equal to  $\delta\% = 6.052726E-12\%$  and  $\delta = 0.0010$  in comparison with the actual value. It is to be remarked that in this way too the fit is made just with this prime number  $P_n$  itself and not with the whole sequence, though using the features of the whole fit function. The same happens by fitting the chosen  $P_n$  value by the values given by the other two functions and the same is with all the other sequences and primes. Thus simply by extrapolating this fit function to the next value  $n+1 = 10,000,000,001$  one gets the value  $P(10,000,000,001) \approx C_\alpha \times (n+1)^\alpha \approx C_\alpha \times (1E10+1)^\alpha \approx 9.3139764474155300 \times (10,000,000,001)^{1.043243392139520} \approx 252,097,800,649.300$  and dropping the decimal digits (of course) the value  $P(10G+1) = 252,097,800,649$ . with a relative error =  $8.054992E-11$  and an absolute error  $_{\text{actual-calculated}}\delta P(10^{10}+1) = 20.3060 \approx 20$  in respect to the actual value  $_{\text{actual}}P_{10G+1} = 252,097,800,629$ . As for the next  $_{\text{actual}}P_{n+2} = P_{10G+2} = 252,097,800,637$  one gets the result  $P(10^9+2) \approx C_\alpha \times (n+2)^\alpha \approx 9.3139764474155300 \times (10,000,000,002)^{1.043243392139520} = 252,097,800,675.601$  and the difference is even larger amounting to  $1.532304E-10$  i.e. absolute error  $[_{\text{calc.}}P(10^9+2) - _{\text{act.}}P(10^9+2)] = 38.601 \approx 38$ .

The last example involving  $P_{800G} = 23,812,036,414,963$  gives the result  ${}_{\text{calc.}}P(8E^{11}) = 23,812,036,414,963.00000$  again calculated by the fit function  $C_\alpha \times n^\alpha \approx 11.067495661792100 \times (800G)^{1.0360962734124300}$  to be compared to the actual value  $23,812,036,414,963$  with relative error equal to  $-1.640452E-15$  and difference  $\delta = -0.0390$  (an error probably due merely to the computer accuracy). The value calculated for the next  $P_{800G+1} \approx C_\alpha \times (n+1)^\alpha = C_\alpha \times (800,000,000,001)^\alpha = (11.067495661792100) \times (800G+1)^{1.03609627341243} \approx 23,812,036,414,993.90$  to be compared to the actual value  ${}_{\text{act.}}P_{800G+1} = 23,812,036,415,029$  with a relative error of  $1.473290E-12$  that is an absolute error  $\delta = [{}_{\text{calc.}}P(8 \times 10^{11} + 1) - {}_{\text{act.}}P(8 \times 10^{11} + 1)] = 35.10 \approx 35$

It is possible to see, in these four cases, that, while the value of the known prime i.e. the starting one  $P_n$  is *exact*, i.e. perfectly *reproduced/simulated*, the next one  $P_{n+1}$  is only *approximate* with the discrepancy (or error)  $\delta P = P_{\text{actual}} - P_{\text{calculated}}$  shown due to the extrapolation of the fitting function from  $P_n$  to  $P_{n+1}$  and that the next  $P_{n+2}$  is still more approximate i.e. with a higher error.

The following table summarizes just these four examples showing the percentage differences  $\delta\% = (P_{\text{actual}} - P_{\text{calculated}}) / P_{\text{actual}} \times 100$  between the actual and calculated prime value and their absolute differences  $\delta P_{n+1} = P_{\text{actual}} - P_{\text{calculated}}$  where the calculated value is the extrapolated one and the differences  $\delta P_{n+1}$  are integers.

**Table 1. Examples of some next ( $P_{n+1}$ ) primes (actual and calculated) and errors  $\delta P_{n+1}$**

Counter $n + 1$	Actual Prime $P_{n+1}$	Calculated value $C_\alpha \times (n+1)^\alpha$	$\delta\%$	$\delta P_{n+1}$
82,500,001	1,664,674,819	1,664,674,834.218040	9.141748E-07%	$\approx 15$
300,000,001	6,461,335,171	6,461,335,131.646470	6.09061E-07%	$\approx 39$
10,000,000,001	252,097,800,629	252,097,800,649.3060	8.05499E-11%	$\approx 20$
800,000,000,001	23,812,036,415,029	23,812,036,414,993.90	1.473290E-10%	$\approx 35$

It would appear a random behaviour for the error  $\delta P_{n+1}$  and it is correct to conjecture that the reported extrapolation from one prime  $P_n$  to the next  $P_{n+1}$  depends also on the gap  $\Delta P_{n+1} = P_{n+1} - P_n$  between the two primes themselves and the prime gaps can be conjectured to be stochastic. However this topic, already suggested in a previous work by the same Author, will be the matter of future in-depth studies and discussions.

Of course all the discrepancies observed are very high if one wants to get the *exact*  $P_{n+1}$  or  $P_{n+2}$  value, nonetheless, in doing so, one has got a method that can reproduce the *exact* value of a prime  $P_n$  starting from which, one can go on to examine the next nearby following number  $P_{n+1}$  by means of a primality test, having the mathematical certainty that the prime searched is there, in the vicinities of the estimated one  $P_{n+1}$ . In doing so a step by step procedure is established and it is possible to compute the next prime number, though *approximate* that is with an error. Afterwards, sweeping the whole range of odd numbers (of course with the last digit = 1, 3, 7, 9) between  ${}_{\text{calc.}}P_{n+1}$  and  ${}_{\text{act.}}P_n$  one can look for and find the actual value of the next actual prime  ${}_{\text{act.}}P_{n+1}$ . In such a way machine-time can be saved in the search of new prime numbers in that, being the difference  ${}_{\text{calc.}}P_{n+1} - {}_{\text{act.}}P_n \approx 30 \div 40$  (see Table 1), one can check just  $12 \div 16$  numbers, nevertheless keeping in mind that the gap values tend to increase in increasing  $n$ .

In addition another key question can be addressed. It has been always questioned on the nature of primes, whether deterministic or stochastic, so that, according to the results got in the present study, now one can assume or conjecture the twofold nature of prime numbers, having both a deterministic aspect and a stochastic one at the same time. As a matter of fact the possibility of treating primes by fit functions (though approximately) reflects the former facet while the impossibility to fit them in an exact way is a clue of the latter feature. All that might explain their volatility and unpredictability. The *global* or *coarse* facet (deterministic) of

primes is in strong contrast with their *local* or *fine* structure (stochastic) so that they appear to display a sort of *double personality*.

The fact that the actual prime gaps  $\Delta P_k = P_{k+h} - P_{k+h-1}$  are locally highly irregular and that the calculated gaps are regular, due to the fact that the fit function (whatever it might be) is *regular* just like any analytic function, is what prevents to fit primes locally by an analytic function itself in an exact way, thus shredding a vivid light on prime gaps and their stochastic nature as well as on primes themselves: globally deterministic and locally stochastic. From this viewpoint one can look at the following two tables relative to a typical example.

**Table 2. Example of actual and calculated  $P_{k+h=0\text{thru}5}$  with actual and calculated gaps**

Counter $k+h$ ( $h = 0$ thru 5)	Actual $P_{k+h=0\text{thru}5}$	$\Delta P_k^{\text{actual}} =$ $= P_{k+h} - P_{k+h-1}$	Calculated $P_{k+h} =$ $= C_a \times (k+h)^a$	$\Delta P_k^{\text{calculated}} =$ $= P_{k+h} - P_{k+h-1}$
292,000,000,000.	8,386,246,028,197	//	8,386,246,028,195.870...	//
292,000,000,001	8,386,246,028,219	22	8,386,246,028,225.610...	29.734375000
292,000,000,002	8,386,246,028,263	44	8,386,246,028,255.370...	29.763671875
292,000,000,003	8,386,246,028,291	28	8,386,246,028,285.140...	29.764648437
292,000,000,004	8,386,246,028,311	20	8,386,246,028,314.900...	29.763671875
292,000,000,005	8,386,246,028,333	22	8,386,246,028,344.640...	29.735351562

One can calculate, from the same table, the relative differences and the absolute differences  $\delta$  between any actual and calculated prime examined getting the following Table 3. All the errors in this example, as well as in all the other cases studied, are fairly low.

**Table 3. The same actual and calculated  $P_{k+h=0\text{thru}5}$  with relative & absolute differences**

Counter $k+h=0\text{thru}5$	Actual $P_{k+h=0\text{thru}5}$	Calculated $P_{k+h=0\text{thru}5} =$ $= C_a \times (k+h)^a$	Relative $\delta P_h =$ $= (P_{\text{act.}} - P_{\text{calc.}}) / P_{\text{act.}}$	Absolute $\delta P_h =$ $= P_{\text{act.}} - P_{\text{calc.}}$
292,000,000,000	8,386,246,028,197	8,386,246,028,195.870...	1.341482E-13	1.125000
292,000,000,001	8,386,246,028,219	8,386,246,028,225.610...	7.881208E-13	6.609375
292,000,000,002	8,386,246,028,263	8,386,246,028,255.370...	9.094597E-13	7.626953
292,000,000,003	8,386,246,028,291	8,386,246,028,285.140...	6.990380E-13	5.862305
292,000,000,004	8,386,246,028,311	8,386,246,028,314.900...	4.652102E-13	3.901367
292,000,000,005	8,386,246,028,333	8,386,246,028,344.640...	1.387596E-12	11.636719

Remarkably, we can observe that the discrepancy  $\delta P_n$  tends approximately to increase in increasing  $h$  that is in moving away from  $P_n$ , what is an evidence of the fact that the algorithm is highly local.

Thus no doubts that, by the present approach, an algorithm and a methodology have been set up to simulate the deterministic aspect of prime numbers what could lead, in the future, through the appropriate means and tools, to the desired goal of attaining the value of a prime number  $P_n \approx P(n)$  from its counter  $n$  with the least error i.e. the maximum possible approximation using the twofold criterion of both adopting the method here shown - though by the help of a much more powerful computer (or even of a mainframe) in order to reach a much higher precision - and studying the stochastic aspect of prime numbers what will allow to fill the discontinuity between the calculated value and the actual one. This latter aspect will be the topic of a next research by the Author.

#### 4. CONCLUSIONS AND FUTURE DEVELOPMENTS

The new algorithm presented in this paper, that makes a wide use of the modified Chi-square function under the form  $-1/X_k^2(A, n/\mu)$ , of the function  $C_a \times n^a$  and of the function  $\lambda_n \times n \times \ln(n)$  as fit functions to simulate the finite sequences of prime numbers  $\{P_n\}$  and their major features as well as of the single primes themselves  $P_n \approx -1/X_k^2(A, n/\mu) \approx C_a n^a \approx \lambda_n \times n \times \ln(n)$ , constitutes an innovative methodology for the former and the latter. The

following ones should be considered some of the main findings of the study, at this very early stage.

- Prime number sequences, as fitted by the three fit functions, have not the property of scale invariance holding for them the scaling laws given by  $k = k(n) = 2 - 2\alpha(n)$  of the opposite inverse modified Chi-square function  $-1/X_k^2(A, n/\mu)$  and of  $C_\alpha n^\alpha$ .
- It is possible to find many inductive algorithms which allow to simulate, i.e. to reproduce, the value of a prime  $P_n$  starting from its counter  $n$  i.e. such that  $P_n \approx P(n)$  though with the results affected by uncertainties i.e. errors. By now the most precise approximation appears to be that given by the function  $C_\alpha n^\alpha \approx P_n$  with the ad-hoc values of  $C_\alpha$  and  $\alpha$ .
- Prime numbers seem to display a twofold nature: deterministic and stochastic, the former having been examined in the present report, though in an approximate context, the latter which is still entirely to be studied and even understood. As a matter of fact, while the *long-distance* or *large-scale* behaviour presents a kind of global predictability, at least approximate, the *short-distance* or *small-scale* one is much less predictable. That could be the topic of future investigations starting from the fits reported here, for instance asking for any  $P_n$  what is the behaviour of the *distance* between the prime itself and the fit curve on the  $(n, P_n)$  plane where both the primes and the curve lie, while another possibility would be to examine the prime number gaps  $\Delta P_n = P_{n+1} - P_n$ , another interesting issue of study.
- Given the value  $n$  it is possible to *reproduce/simulate* the value (already *a-priori* known) of a prime  $P_n$  with the smallest approximation, about  $10^{-12}$  % or  $10^{-13}$  % for the relative inaccuracy between the calculated prime and the actual one, that is  $\delta\%$ , with an absolute error  $\delta P_n$  equal to few decimals by means of a fitting process.
- Starting from a known prime  ${}_{\text{actual}}P_n$  It is possible to forecast the *next* one  $P_{n+1}$  by the function  $C_\alpha x n^\alpha$  (for instance) and its application to the next  $n+1$  i.e.  $P_{n+1} \approx C_\alpha x (n+1)^\alpha$  and by a simple exploration and survey of the numbers found in the surrounds of  $P_{n+1}$  that is present in the range  $({}_{\text{actual}}P_n \div {}_{\text{calculated}}P_{n+1})$  and less beyond, applying a primality test to them, thus saving much computer memory and machine-time.
- The whole algorithm and the entire procedure hold just taking into account all the approximations adopted, as well as the imprecisions and the error propagations as usually done in experimental physics and in all the other experimental sciences. However no doubt that future investigations, by the use of much more powerful computers or mainframes or even supercomputers as well as of an ad-hoc software, will be easily able to improve all the precisions that is to reduce the uncertainties as much as possible.

Despite the investigation is at its first stage, anyhow it is the Author's opinion that the algorithm and the experimental methodology here shown can open a new wide field of study as the experimental counterpart of number theory which can reveal all its power more and more in the future. Despite many issues (some of which here just suggested or mentioned or simply hinted) are still to be examined and deepened, nonetheless in this initial phase of the research the aim and the goal are just to set up an innovative and suitable algorithm for the analytical treatment of prime numbers leaving to the next future studies the widening and deepening of the whole matter, first of all the fundamental issue of precision improvement i.e. of the reduction of the errors to as low as possible values, taking into account the results got just in this early part of the study.

## REFERENCES

- [1] Chebyshev (Tchebycheff), On the Totality of Primes, in Smith DE. A source book in mathematics Dover Publications. Inc. New York, unabridged republication of the 1st ed. originally published in 1929 by McGraw-Hill Book Co. Inc.;1959, 127-148.
- [2] Don Zagier. The first 50 million prime numbers.  
Available: <http://www.ega-math.narod.ru/Liv/Zagier.htm>.

- [3] Derbyshire J. L'ossessione dei numeri primi - Bernhard Riemann e il principale problema irrisolto della matematica. ©2006 Bollati Boringhieri Editore s.r.l. Torino, Italian.
- [4] Ingham A.E. The Distribution of Prime Numbers. Cambridge Tracts in Mathematics and Mathematical Physics, No. 30. Cambridge, At the University Press, 1932.
- [5] Goudsmit S.A. Unusual Prime Number Sequences. NATURE, Vol. 214, June 10. 1967 ©1967 Nature Publishing Group.
- [6] Moree P., Petrykiewicz I., Sedunova A. Computational History of Prime Numbers and Riemann Zeros. arXiv:1810.05244v1 [math.NT] 11 Oct. 2018
- [7] Leonetti P. A Characterization of Sophie Germain Primes. arXiv:1305.0893v3 [math.NT] 21 Aug. 2016
- [8] Jameson GJO. The Prime Number Theorem. London Mathematical Society, Student Texts 53. Cambridge University Press (2003). Numbers Supplement - The Prime Number Theory - History Available: <https://faculty.etsu.edu/gardnerr/3120/Number-Theory-notes.htm>
- [9] Hardy, G.H., Littlewood J.E., Contributions to the Theory of the Riemann Zeta-Function and the Theory of the Distribution of Primes Acta Mathematica 41, 1916, 119-196. Available: <http://dx.doi.org/10.1007%2FBF02422942>
- [10] A. Languasco, A. Zaccagnini, Alcune proprietà dei numeri primi, I e II Sito web Bocconi-Pristem, 2005. Available: <http://matematica.uni-bocconi.it/LangZac/home.htm>
- [11] Green B., Long arithmetic progressions of primes © 2007, Clay Mathematics Proceedings 7, 149-167
- [12] Green, T., Tao T., The primes contain arbitrarily long arithmetic progressions Annals of Mathematics 167 (2), 2008, 481-547. Available: <http://www.arXiv:math.NT/0404188.doi:10.4007/annals.2008.167.481>
- [13] Granville, A., Prime Number Patterns, The American Mathematical Monthly, Vol. 115, No. 4, Apr. 2008, 279-296 published by Mathematical Association of America. Available: <http://www.jstor.org/stable/27642472>
- [14] Granville, A., Martin G., Prime Number Races The American Mathematical Monthly 113, No. 1, 2006, 1-33 published by Mathematical Association of America, Available: <http://www.jstor.org/stable/27641834>
- [15] Granville, A., It is Easy to Determine whether a Given Integer is Prime Bulletin (New Series) of the American mathematical society 42 N° 1, 2004, 3-38, Article electronically published on September 30 ©2004 American Mathematical Society. Available: <http://www.dms.umontreal.ca/~andrew/.../Bulletin04...>
- [16] Luque B., Lacasa L. The first-digit frequencies of prime numbers and Riemann zeta zeros, Proc. R. Soc. A (2009) **465**, 2197-2216, doi: 10.1098/rspa.2009.0126, *Published online* 22 April 2009. © 2009 The Royal Society.
- [17] J.M. Gandhi, Formulae for the nth prime, Proc. Washington State Univ. Conf. on Number Theory, Washington State Univ. Pullman, Wash. 1971, 96-106.
- [18] Axler C. New Estimates for the nth Prime Number, Journal of Integer Sequences, Vol. 22 (2019), Article 19.4.2
- [19] Anonymous Surprising connections between number theory and physics, Available: <http://empslocal.ex.ac.uk/people/staff/mrwatkin/zeta/surprising.htm>
- [20] Yang, J.R., The Riemann Hypothesis: Probability, Physics and Primes Yang Academy, Gaithersburg, Maryland, 20877, USA, 2014
- [21] Al Rabeh, R.H., Primes, Geometry and Condensed Matter, Progr. Phys. 3, 2009, 54-59
- [22] Rosu, H.C. Quantum hamiltonians and prime numbers Mod. Phys. Lett., A 18, 2003 Brief Review © World Scientific Publishing Company, arXiv:quant-ph/0304139v4 12 May 2013.
- [23] Castro, C., On p-adic stochastic dynamics, supersymmetry and the Riemann conjecture 2001 arXiv :physics/o10114v2 [physics.gen-ph] 1 Mar 2001.
- [24] Aresa S., Castro M. Hidden structure in the randomness of the prime number sequence? Physica A 360 (2006) 285-296, Elsevier.
- [25] Liberatore P. Compilation of Intractable Problems and its Application to Artificial

Intelligence, Dottorato di ricerca in Ingegneria Informatica, Thesis X Ciclo 1998, Università degli Studi di Roma "La Sapienza"

[26] Iozzi F., Trevisan L. Tractable and Intractable Problems, Bocconi University—30540: Computer Science II, Handout NP1, April 26, 2022.

[27] Gauss, Third Proof of the Law of Quadratic Reciprocity, in Smith DE. A source book in mathematics Dover Publications. Inc. New York, unabridged republication of the 1st ed. originally published in 1929 by McGraw-Hill Book Co. Inc.;1959, 112-118

[28] Bailey, D.H., Borwein J.M., Exploratory Experimentation and Computation, Notices of the AMS 58, N 10, 2011, Available: 1410-1419 <http://escholarship.org/uc/item/26m6noob>

[29] Andeberhan, T., Medina L.A., L A & Moll V H Editors, Contemporary Mathematics - 517 - Gems in Experimental Mathematics, AMS Special Session, Experimental mathematics, 2009, Washington D.C. American Mathematical Society.

[30] Borwein, J.M., Exploratory Experimentation: Digitally-Assisted Discovery and Proof , May 18, 2009, University of Newcastle, Australia.

[31] Yeseul K., Byung M.W. Randomness at large numbers: Experimental proof in coin toss and prime number, AIP Advances, 9 125318 (2019), Available:

<https://doi.org/10.1063/1.5133773>

[32] Crandall R., Pomerance C. Prime Numbers - A Computational Perspective, Second Edition, Springer, © 2005 Springer Science+Business Media, Inc., ISBN-13: 978-0387-25282-7

[33] Brown J.R. Philosophy of Mathematics - A Contemporary Introduction to the World of Proofs and Pictures, Second Edition, Routledge, Taylor & Francis Group, New York and London, © 2008 Taylor & Francis.

[34] Epstein D., Levy S. Experimentation and Proof in Mathematics, Notices of the AMS, Volume 42, Number 6, June 1995, 670-674.

[35] Gronau Q.F., Wagenmakers E-J. Bayesian Evidence Accumulation in Experimental Mathematics: A Case Study of Four Irrational Numbers, Experimental Mathematics, 2018, Vol. 27, No. 3, 277-286, Available: <https://doi.org/10.1080/10586458.2016.1256006>

[36] Porte M.A., Zabusky N.J., Campbell D.K. Fermi, Pasta, Ulam and the Birth of Experimental Mathematics, A reprint from American Scientist, the magazine of Sigma Xi, The Scientific Research Society, American Scientist, Volume 97, © 2009, May-June. 214-221.

[37] Mathews. J., Walker R.L., Mathematical Methods of Physics CalTech © W.A. Benjamin Inc., New York N.Y., 1964

[38] Babusci, D., Dattoli G., Del Franco M., Lectures on Mathematical Methods for Physics, RT/2010/58/ENEA ENEA-Roma-I, 2010

[39] Morice, E., Dizionario di statistica © by ISEDI Milano, 1971. Dictionnaire de Statistique, © by Dunod Paris, 1967

[40] Young, H.D., Statistical Treatment of Experimental Data Carnegie Institute of Technology © McGraw-Hill Book Company Inc., 1962

[41] Walck C. Hand-book on STATISTICAL DISTRIBUTIONS for Experimentalists, Internal Report SUF-PFY/96-01, Stockholm, 11 December 1996, last modification 10 September 2007, University of Stockholm.

[42] Riccietti E. La Funzione Gamma di Eulero e il Metodo dei Trapezi, Thesis, Università degli Studi di Firenze - Facoltà di Scienze Matematiche, Fisiche e Naturali - Corso di laurea triennale in Matematica. AA 2011-2012. Italian.

[43] Kerbalayeva R.E.K. Some Basic Characterization of the Function  $\gamma$ , International Journal of Theoretical and Applied Mathematics. Vol. 7, No. 5, 2021, pp. 72-75. doi: 10.11648/j.ijtam.20210705.11 Available: <http://www.sciencepublishinggroup.com/ijtam>

[44] Idrissu M.M., Tetteh K.I. The Gamma Function and its Analytical Application, Journal of Advances in Mathematics and Computer Science, 23(3): 1-16, 2017, Article no. JAMCS.34779 DOI: 10.9734/JAMCS/2017/34779

[45] Anonymous,  $\Gamma$  Function, Available: <http://www.efunda.com/math/gamma/findgamma>

- [46] Anonymous, Available: <http://keisan.casio.com/exec/system/1180573444>
- [47] Lattanzi D., Scale Laws of Prime Number Frequencies by the Modified Chi-square Function, former British Journal of Mathematics and Computer Science (now Journal of Advances in Mathematics and Computer Science) 13(6): 1-21, 2016, Article no. BJMCS.23200, ISSN: 2231-0851, DOI: 10.9734/BJMCS/2016/23200  
Available: <http://www.sciencedomain.org/issue/1545>
- [48] Lattanzi D., An Elementary Proof of Riemann's Hypothesis by the Modified Chi-square Function, former British Journal of Mathematics and Computer Science (now Journal of Advances in Mathematics and Computer Science) 15(5): 1-14, 2016, Article no. BJMCS.25419, ISSN: 2231-0851, DOI: 10.9734/BJMCS/2016/25419  
Available: <http://www.sciencedomain.org/issue/1778>
- [49] Lattanzi D., Computational Model of Prime Numbers by the Modified Chi-square Function, former British Journal of Mathematics and Computer Science (now Journal of Advances in Mathematics and Computer Science) 20(5): 1-19, 2017, Article no. BJMCS.31589, ISSN: 2231-0851, DOI: 10.9734/BJMCS/2017/31589  
Available: <http://www.sciencedomain.org/issue/2069>
- [50] Lattanzi D., An Innovative Approach to the Finite Sequences of prime Numbers, Journal of Advances in Mathematics and Computer Science, 35(9): 34-56, 2020; Article no. JAMCS 64455, ISSN: 2456-9968  
Available: <https://journaljamcs.com/index.php/JAMCS/article/view/30321>
- [51] Lattanzi D., Distribution of prime numbers by the modified chi-square function, Notes on Number Theory and Discrete Mathematics, Vol. 21, 2015, No.1, 18-30, ISSN 1310-5132,  
Available: <https://nntdm.net/volume-21-2015/number-1/18-30/>
- [52] Anonymous: Prime Cruncher, Available: <http://www.bigprimes.net/cruncher/>
- [53] The Nth Prime Page - A prime page by Andrew Booker, Available:  
<https://t5k.org/nthprime/>