Original Research Article

Hierarchical Model-based Prediction on the Maximum Gap between Consecutive Primes below An Arbitrary Number

Abstract

The distribution of the primes in natural number sequence has no simple law, which seems to show a certain degree of randomness. In this paper, we found a hierarchical progression pattern that the prime intervals are gradually tend to be evenly distributed on different orbits belonging to different levels, meanwhile, the prime numbers tend to give priority to spread all over each orbit of lower-energy level, and then gradually fill the orbits of higher-energy level with the expansion of prime number range. Moreover, by analyzing the count of prime numbers in different energy levels and different orbits based on the established hierarchical progression model, we investigated that the count of prime numbers in different energy levels decreases exponentially with the increase of energy levels, and its natural logarithm is approximately linear with the change of energy level. Based on these findings, we propose two kinds of strategies to estimate the value of maximum gap between consecutive primes lessen than a given number. Our developed results reveal that the prime distribution shows a layered increasing law, and give a reliable prediction guide to understand the upper limit of maximum interval below an arbitrary number.

Keywords: Hierarchical progression model; Prime numbers; Maximum gaps; Prediction strategy; Prime energy level

2010 Mathematics Subject Classification: 11N05; 03C30

1 Introduction

For the past century, prime interval problem which is an important subject in prime number research, has attracted the attention of scientists from different research fields[1–5]. It is known that the average interval between consecutive primes tends to be infinite with the growth of natural number, but in any finite number column, the maximum prime gap can be much larger than the average interval[6]. According to the Prime Number Theorem [7; 8], the number of primes below x is $\pi(x) \sim x/\ln x$, and it follows that the mean distance between two consecutive primes is about $\ln x$. Since Erdös[9] proved that at least one of the intervals between successive primes less than p_n is always of length not less than $c \ln p_n \ln\ln p_n/(\ln\ln\ln p_n)^2$ for large p_n and an appropriate constant c, this result has been improved by a lot of efforts [10; 11], which gave some rough lower-bound estimations for the maximum interval. However, due to the fact that such estimations are in usual quite different from the actual value, the interval itself remains to be unpredictable, and till now people do not even know whether there are infinite pairs of primes separated by twin primes [12–14].

How prime numbers are distributed has been the focus of number theory research, from the perspective of system, the distribution dynamics of prime numbers shows a complex form. It is believed that the distribution of prime numbers must follow some rules, although it shows a certain randomness, that is, it has a certain degree of chaotic characteristics. Thus, the prime distribution problem has always been a confusing major topic, in which the difference between consecutive primes considered as an important characteristic of the distribution of the prime numbers attracted attentions of many researcheres[15–22]. In theory, the interval between adjacent prime numbers can be arbitrarily large, however, people are more concerned about the following issues: 1) Let g be any even number, is there certainly two adjacent prime numbers whose interval is g? 2) Are there finite or infinite adjacent prime numbers with interval g? 3) How far can the interval between adjacent prime numbers be under the premise of less than a given number? For these problems, theoretical tools or widely accepted hypotheses can not provide a standard model[23–29].

Therefore, aiming to find a solution for the question closely related with the above third question, the paper proposes two kinds of strategies to predict the maximum gap lessen than a given number by means of our established hierarchical progression model on the distribution of the gaps between consecutive primes, in which one is that according to the correlation between the prime number distribution and the maximum energy level, it was established three fitting curve formulas through which we can judge the possible range of maximum prime interval by calculating the maximum energy level corresponding to a given number, and the other is to devise a simplified hypothetical formula, which can be used to easily determine the upper limit of maximum gap between consecutive primes lessen than an arbitrary number.

2 Model

2.1 Prime distribution pattern

Hierarchical progression model The value of prime interval between consecutive primes shows a trend of that with the increase of natural number, it is distributed on different tracks (referred to as 'prime orbits') belonging to different levels (named as 'prime energy level') .
Here, define the prime gap

$$g(p_k) = p_{k+1} - p_k, \tag{2.1}$$

where p_k, p_{k+1} denote the k-th and next adjacent prime number in the natural number sequence, respectively.
Let $q_m \equiv p_k \pmod{T}$ be the prime orbit which is the remainder of k-th prime number divided by T with value of 210.
Introduce a prime orbit vector $Q \in \mathbb{N}^{1\times D}$ having form of

$$Q = [\; q_1, \quad q_2, \quad \cdots, \quad q_m, \quad \cdots, \quad q_D \;], \tag{2.2}$$

where D is the prime orbit numbers with value 48. Here, Q is denoted to be

$$\begin{aligned} Q = [&1, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, \\ &71, 73, 79, 83, 89, 97, 101, 103, 107, 109, 113, 121, 127, \\ &131, 137, 139, 143, 149, 151, 157, 163, 167, 169, 173, \\ &179, 181, 187, 191, 193, 197, 199, 209]. \end{aligned} \tag{2.3}$$

Then, introduce a prime energy matrix $M \in \mathbb{N}^{B\times D}$ denoted as

$$M = \begin{pmatrix} \hat{q}_2 - \hat{q}_1 & \hat{q}_3 - \hat{q}_2 & \cdots & \hat{q}_{m+1} - \hat{q}_m & \cdots & \hat{q}_{D+1} - \hat{q}_D \\ \hat{q}_3 - \hat{q}_1 & \hat{q}_4 - \hat{q}_2 & \cdots & \hat{q}_{m+2} - \hat{q}_m & \cdots & \hat{q}_{D+2} - \hat{q}_D \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \hat{q}_{n+1} - \hat{q}_1 & \hat{q}_{n+2} - \hat{q}_2 & \cdots & \hat{q}_{m+n} - \hat{q}_m & \cdots & \hat{q}_{D+n} - \hat{q}_D \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \hat{q}_{B+1} - \hat{q}_1 & \hat{q}_{B+2} - \hat{q}_2 & \cdots & \hat{q}_{B+m} - \hat{q}_m & \cdots & \hat{q}_{B+D} - \hat{q}_D \end{pmatrix}, \tag{2.4}$$

where

$$\hat{q}_{m+n} = q_i + j * T, \quad m \in [1, D], n \in [1, B]. \tag{2.5}$$

Here, B is a given positive integer number, and for each $\hat{q}_{m+n}$, there exists

$$i \equiv m + n \pmod{D}, \quad j = \lfloor (m+n)/D \rfloor. \tag{2.6}$$

Moreover, define a row vector $L_n \in \mathbb{N}^{1\times D}$ to be the n-th prime energy level, expressed as

$$L_n = [M_{n1}, M_{n2}, \cdots, M_{nD}]. \tag{2.7}$$

Therefore, given p_k and p_{k+1}, one can determine the row number and column number of p_k in matrix M according to the calculated values of $g(p_k)$ and q_m, respectively, then know the position of its corresponding prime energy level L_n.

Count on different prime orbits Let $C \in \mathbb{N}^{B\times D}$ be a count matrix with zero initial values. Given a set of primes $\{p_k, k \in [1, N]\}$ where N is the index of maximum prime, the count of each kind of prime can be calculated as follows.

(i) Divide p_k by a period number $T(= 210)$, and obtain its corresponding prime orbit q_m.

(ii) Calculate the prime gap $g(p_k)$.

(iii) Find the index number m of the element whose value is equal to q_m in the prime orbit vector Q.

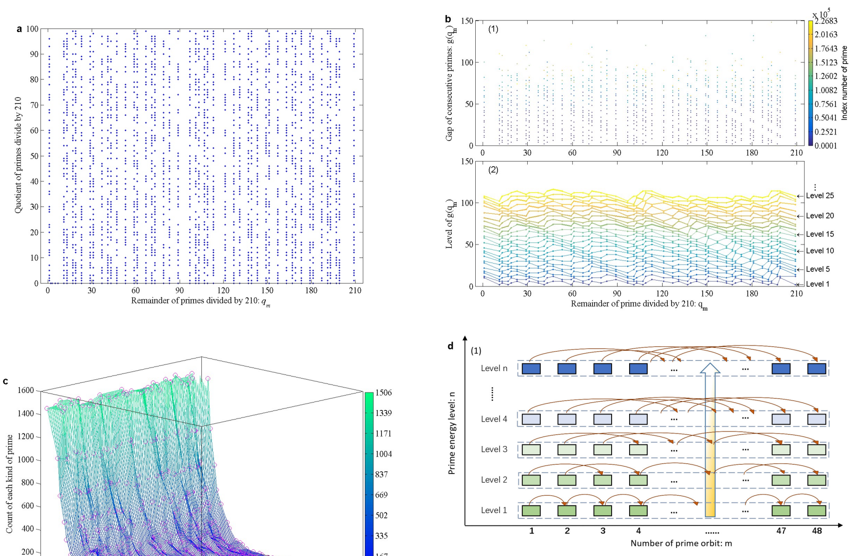
(iv) According to the values of $g(p_k)$, look for the index number n from the m-th column of the prime energy matrix M which meets the condition of $g(p_k) = M_{nm}$.

(v) Set $C_{nm} = C_{nm} + 1$.

(vi) Do $(i) \sim (v)$ until $k = (N - 1)$.

2.2 Hierarchical progression pattern

Through the statistical analysis on a large number of experimental results, it is found that if the prime number is divided by $210(= 2 \times 3 \times 5 \times 7)$, its remainder will be strictly distributed in one of 48 rails (Fig.1a) called as prime orbits, and that although the interval distribution between two adjacent prime numbers seems to be irregular, it generally shows a multi-track and hierarchical increasing trend(Fig.1b). By establishing a hierarchical progression model according to the law of that the prime



Prime orbit: m	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
q_m	1	11	13	17	19	23	29	31	37	41	43	47	53	59	61	67
Prime orbit: m	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
q_m	71	73	79	83	89	97	101	103	107	109	113	121	127	131	137	139
Prime orbit: m	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
q_m	143	149	151	157	163	167	169	173	179	181	187	191	193	197	199	209

※ q_m *(m=1~48)*: The remainder digit of *the* prime number divided by 210.

Figure 1: Hierarchical progression phenomenon of prime gaps. **a**,The distributions of prime numbers less than 2.1×10^4, each of which is divided by the number of 210. The abscissa and ordinate respectively stand for the remainder and quotient of a prime number. **b**,The distributions of the gaps between consecutive primes below 3.15×10^6. **c**,The statistical 3D stereogram of prime counts at each prime energy level from the experimental results of **b**. The pink-color circles stand for the count numbers of each kind of primes on different levels, while the color-bar shows the trend of the counts changing with the energy level. **d**,The definitions of prime energy level and prime orbit. (1) is established through the definitions with a total of 48 prime orbits, and (2) represents the remainders of the primes divided by 210.

energy level increases with the distribution characteristics of prime numbers on 48 different prime orbits (Fig.1d), we found that the interval distribution of prime numbers (except the first four primes of 2,3,5,7 which are excluded from the prime sequence in the following discussion) shows an obvious hierarchical differentiation, and similar to water-wave propulsion, its energy level gradually advances from low to high with the increase of the number. As shown in the phase diagram $g(q_m)$ versus q_m of Fig.1b, it was observed that there is obvious difference in the probability of prime numbers appearing on different prime energy levels, which generally shows the trend of giving priority to fill the prime orbits of lower-energy levels and then gradually fill the higher-energy levels, furthermore, its distribution is approximately uniform on 48 different prime number orbits at each level with same prime energy(Fig.1c).

Although a few prime numbers jump to the positions belonging to higher-energy levels while the others generally locate in the lower-energy levels, the proportion of such prime numbers is very low compared to the whole primes within a specific range. It can be confirmed from Fig.1c, that the distribution of primes in each energy level is approximately uniform, and that the count of prime numbers distributed on the lowest-energy layer is the largest, for instance, the count difference between the first layer and the second layer is higher than that between the second layer and the third layer. Moreover, the count of prime numbers distributed on 48 different prime orbits in each layer is approximately even. Based on the experimental results of Fig.1c, we made a statistic that among all 226,831 primes less than 3,150,000 except the first four, a total of 69,428 primes are approximately evenly distributed on 48 prime orbits of the first energy level with a mean value of 1,446, accounting for nearly 30.6% of the total prime counts, while 48,977 and 34,189 primes are respectively distributed on the second and the third energy level, accounting for nearly 21.6% and 15.1%, and the total ratio of prime numbers of the first three layers reaches about 67.3%.

3 Methods

Relationship formulas between maximum energy level and filled-tiers We try to find out the relationship between the size of the prime energy level and the position where the prime numbers can firstly occupy all the energy levels (referred to as 'filled-tiers') below the specified layer. An empirical formula of prime number relating with energy level referred to as the 'prime-number-level formula' was used to calculate the fitting curve as follows:

$$\ln p = c_1 \ln x \ln(\ln x) + \sqrt{x} + \exp(c_2), \tag{3.1}$$

where p is the prime number, x stands for the index of prime energy level, c_1, c_2 are both constant coefficient.

The correlation of prime count relating to filled-tiers can be fitted by a formula called as 'prime-count-level formula', which is expressed as

$$\ln C_p = \ln x \ln(\ln x) + \sqrt{x} + c, \tag{3.2}$$

where C_p is the count of the prime number, x stands for the number of filled-tiers, and c is a constant coefficient.

It was investigated that the distribution on the natural logarithm count of primes in each filled energy level is approximately in a straight line, and the intersection of the line and the horizontal axis is very close to the value of maximum energy point obtained from the experiment results. Thus, the value of maximum energy point l_{max} can be calculated by a so-called maximum-level-prediction formula, expressed as

$$l_{max} = \lceil \frac{(n-1)\ln x_1}{c \ln(x_1/x_n)} + 1 \rceil. \tag{3.3}$$

where n is the number of filled-tiers, and x_1, x_n represent the count of the prime numbers in the first and the n-th filled energy level, respectively. Here, c is a coefficient of correction.

Therefore, the predicted value of the maximum energy level referred to as the 'predicted maximum level' can be calculated according to the following steps:

(i)Compute the number of filled-tiers based on a numerical solution for the prime-number-level formula (3.1).

(ii)Let x_n be the arithmetical average of C_{sm} and then calculate C_{s1} by means of the prime-count-level formula (3.2).

(iii) Use the maximum-level-prediction formula (3.3) to calculate the value of l_{max}.

Prediction on the maximum gaps Let G_{max} be the maximum gap between consecutive primes below a given number. The predicted maximum gaps can be calculated from a conjecture formula referred to 'maximum-gap-prediction formula', having the form of

$$G_{max} < (c + \varepsilon) \ln x \ln \sqrt{x}, \tag{3.4}$$

where x is an arbitrary number, c stands for a scale factor and ε is a correction factor being a small number.

Lower limits of the maximum interval For comparison, the lower limits of the maximum interval is calculated from a so-called long-gaps-formula [30], which was expressed as follows:

$$\max_{p_{k+1} \ll x} (p_{k+1} - p_k) \gg \frac{\ln x \ln \ln x \ln \ln \ln \ln x}{\ln \ln \ln x}, \tag{3.5}$$

where x is a sufficiently large number, p_k denotes the k-th prime.

Variation range of the maximum gaps The mean value, upper limit and lower limit of each maximum gap corresponding to the prime energy level can be calculated based on the prime energy matrix M in 'Hierarchical progression model'. For each n-th (n ∈[1,B]) prime energy level, select the row vector L_n from the matrix M, and

(i) calculate the mean value g_{mean} referred to 'gap mean' by

$$g_{mean} = \left(\sum_{m=1}^{D} M_{nm} \right) / D, \tag{3.6}$$

(ii) then, find the upper limit g_{uplim} of the maximum gap using

$$g_{uplim} = \max_{1 \le m \le D} M_{nm}, \tag{3.7}$$

(iii) finally, look for the lower limit g_{lowlim} of the maximum gap by

$$g_{lowlim} = \min_{1 \le m \le D} M_{nm}. \tag{3.8}$$

4 Results and discussion

To verify the effectiveness of the above methods, a number of statical experiments were conducted. Firstly, we statistically analyzed the distribution of all 144,449,533 prime numbers (except the first

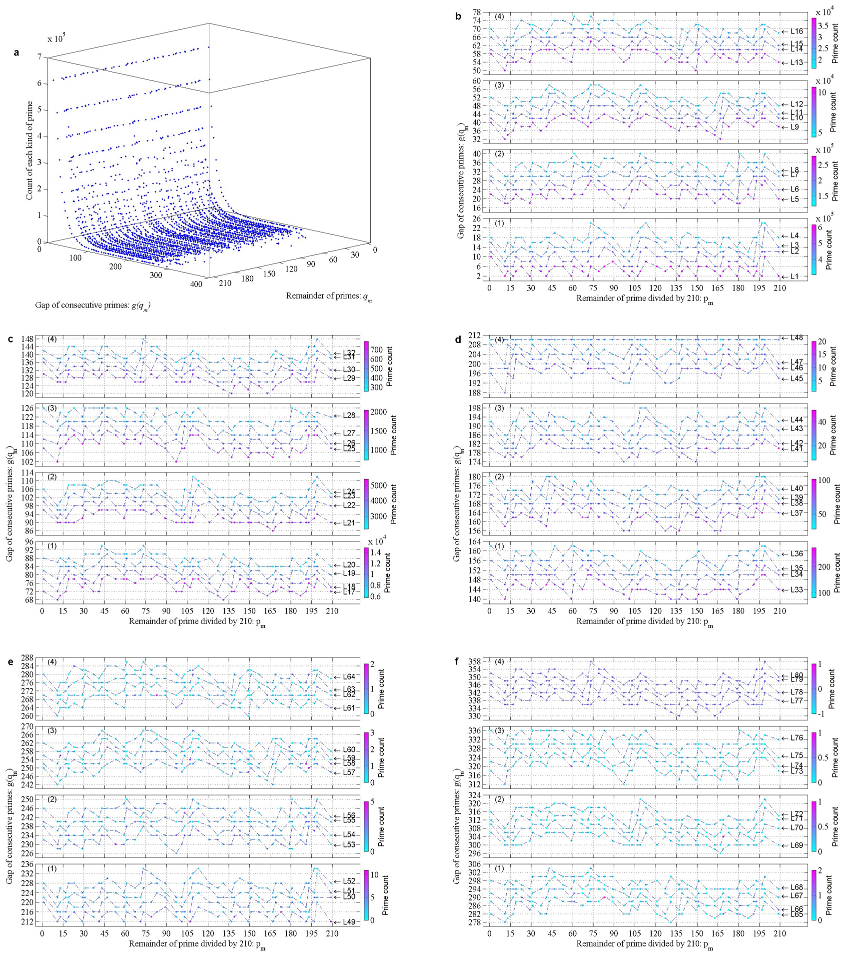


Figure 2: Counts of prime numbers distributed in each prime energy level. **a**,The statistical 3D stereogram of prime numbers lessen than 3×10^9 at each prime energy level. **b-f**,Counting distribution of prime interval on different prime orbits belonging to different energy levels from 1 till 80. Here, the distribution of prime intervals on different prime orbits is depicted from low to high in (1)-(4), each of which contains four energy levels identified by a text annotation Li, where $i(= 1 \sim 80)$ represents the index of energy level. The dots represent the distribution of prime interval, and the colored dash-dot lines show energy levels to distinguish between different layers.

four) lessen than 3×10^9 (Fig.2a), which spread over all prime orbits below 49 energy levels, while those in Fig.1c cover only 18 energy levels.

The count of the prime numbers distributed in each of the first three energy levels accounts for approximately 20.4%, 16.4% and 13.2%, respectively, and the sum of the three reaches about 50% of the total numbers (Fig.2a). We studied the distribution of prime intervals on each prime energy level and found that all 48 prime orbits from the first till forty-ninth energy levels are filled (Fig.2b-d) and some orbits situated in higher-energy levels above 50 layers are sporadically occupied by a few prime numbers (Fig.2e,f), with a number of 604, accounting for 0.000418% of the total prime numbers. Meanwhile, we noticed that the number of distributions of primes in each energy level decreases sharply with the increase of energy level (Fig.2b-d), and from the 50th layer, that of the primes distributed on each prime orbital drops rapidly to single digits. Moreover, not every layer is occupied by prime numbers (Fig.2e,f), in other words, the distribution of prime numbers on these higher-energy orbits is extremely sparse. For instance, on the 62nd energy level, only the 29th, 37th and 47th prime orbits are occupied by one prime number (Fig.2e), while on each of the 67th, 70th and 73rd energy level (Fig.2f), there distributed only one prime number, whose corresponding orbit is the 44th (p_m=191), 39th (p_m=169) and 14th (p_m=59), respectively.

Prediction on the maximum energy level Next, we searched for the prime numbers, each of which appears for the first time when the maximum level of filled-tiers reaches the corresponding energy level from low to high (1∼49). The results were shown in Fig.3a, in which the horizontal axis stands for the number of energy levels with each prime orbit filled with at least one prime number, and the vertical axis $\ln(p_{fs})$ represents the natural logarithm of the prime number p_{fs} which appeared firstly when its corresponding prime energy level is full. We analyzed the relationship between the obtained prime numbers p_{fs} and their corresponding number of filled-tiers, and observed that there is a changing trend similar to some exponential function between them (See green dots in Fig.3a). For the convenience of analysis, we took the natural logarithm $\ln(p_{fs})$ of the primes, and found that the value of $\ln(p_{fs})$ changes with the increasing number of filled-tiers in a kind of logarithmic function (Fig.3a). Meanwhile, a prime-number-level formula (3.1) of $\ln(p_{fs})$ relating with the number of filled-

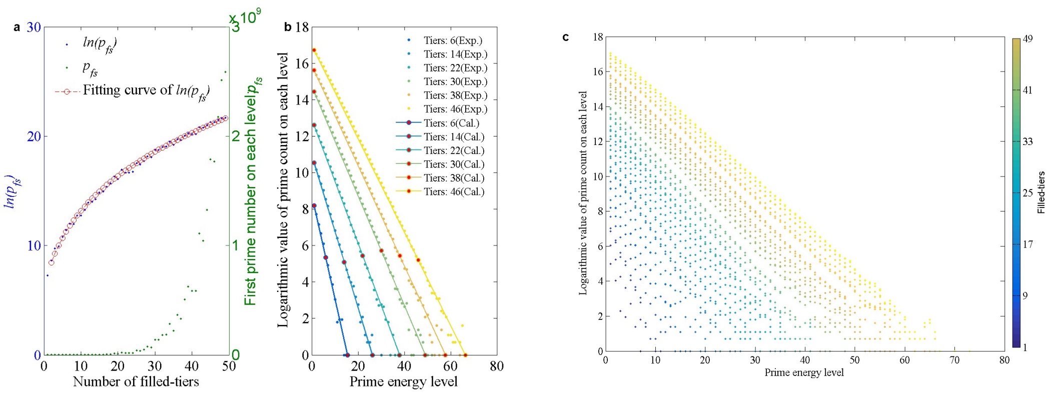


Figure 3: Diagram of the maximum energy level and the filled tiers. **a**,The correlation curve of the filled tiers and the prime number of initial appearance at each level. **b**,The distribution of the prime counts in each energy level changing with the number of filled tiers. The dots (referred to as 'Exp.') represent the results from the experiment and the fitting lines with dots (referred as 'Cal.') stand for those from a formula. **c**, The distribution of the prime counts in each energy level changing with the filled energy levels (filled-tiers) increasing from 1 to 49.

tiers was fitted based on the experimental results in which c_1, c_2 are set as 1.37 and 2, respectively. It can be confirmed from Fig.3a that the experimental data and the curve obtained from the fitting formula have good consistency except for a few points.

The fitting curve of $\ln(p_{fs})$ (Fig.3a) shows that the number of filled energy levels has a clear correlation with the position of prime number appears in the natural number sequence, which enlightens us, that is, when a prime number jumps to some high-energy level, and at the same time all prime numbers less than the given prime just fill all energy levels below the current energy level, we want to know whether there is a certain law in the distribution of prime numbers on all these energy levels.

Based on the position information of prime numbers (Fig.3a), we counted the number of primes distributed in each energy level while prime numbers fill all layers below a certain energy level, among of which the number of layers filled with primes ranges from 1 to 49 (Fig.3c). Then, we found that in all cases where the filled energy layers changes from 1 to 49, the natural logarithm of prime number in each energy level shows an approximately linear decreasing relationship with the size of the filled energy levels. This linearity is relatively significant in all energy levels below the number of filled-tiers, but there is no obvious pattern on the unfilled higher-energy level. To improve the identification degree, we extracted six cases to count the prime numbers at each energy level with the filled-tiers being 6, 14, 22, 30, 38 and 46, respectively. Figure 3b shows the distribution of the prime counts in each energy level changing with the number of filled tiers in which the natural logarithm is taken for the counting of prime numbers in vertical axis. According to the found phenomenon that the distribution of prime count on each layer has a good linearity with the number of filled-tiers, as shown in Fig.3b, we made a straight line for each case from the point of the first energy level to that of the highest energy layer corresponding to the number of filled-tiers, and extended it to intersect with the horizontal axis (called as the maximum-energy point). It can be observed that in each case, the maximum energy level which can be reached by the prime numbers just falls near the maximum-energy point.

Figures 3a and 3b demonstrate that if given a prime number, there is an obvious regularity between the number of filled-tiers and the highest energy level among all filled-tiers that can be reached by all prime numbers less than it, which can be used to find an approximate position related to the maximum-energy point. The problem now becomes that we need a convenient way to determine the count value of primes in the lowest and the highest energy level respectively, although we can get them through data statistics. By analyzing the counting of prime numbers on the first energy level while the number of filled-tiers increases from 1 to 49, we found that there is a strong correlation

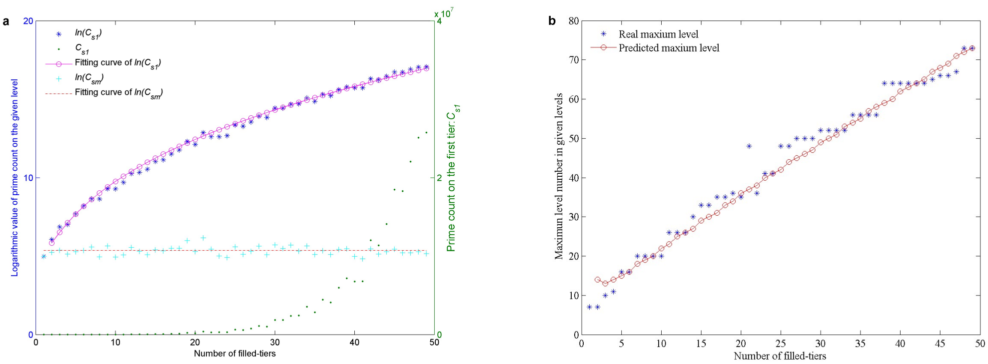


Figure 4: Relationship between the maximum energy level and the filled tiers. **a**, The correlation curve of the filled tiers and the natural logarithm of count on prime numbers at a given level. **b**, The relationship between maximum energy level and the number of filled-tiers.

between the count C_{s1} of prime numbers on the first energy level and the number of filled-tiers, and it can be expressed by Eq.(3.2) in which the constant c is set as 4.68. Meanwhile, due to the fact that the natural logarithm $\ln(C_{sm})$ of prime count in each highest level has little change, we took their average as the reference value of the highest energy level (red-dots line in Fig.4a).

By replacing the prime counts required in Eq.(3.3) at the first energy level and the highest energy level with the values of $\ln(C_{s1})$ and $\ln(C_{sm})$ obtained by the fitting formula established in Fig.4a, we can easily calculate the predicted value of the maximum energy point corresponding to each filled tier. It can be observed from Fig.4b, that the predicted maximum energy level is basically consistent with the actual value, although the error is slightly larger when the number of filled-tiers is some specific values, the overall variation trend is relatively close.

Through the above analysis, we respectively established the relationship between the position of prime number in the natural number sequence and the corresponding number of filled-tiers, and that between the number of filled-tiers and the maximum energy point which can be reached. Thus, according to Eqs.(3.1)-(3.3), we developed a method to predict the maximum energy point that can be reached by all prime numbers less than a given primes. Firstly, we used the numerical solution for Eq.(3.1) to obtain the number of the filled-tiers relating with a given number from the fitting curve formula. Then, the prime count value of the first energy level corresponding to the filled-tiers was determined from Eq.(3.2) which describes the relationship between the number of the filled-tiers and the number of primes distributed in the first energy level. Finally, by means of Eq.(3.3), each position of the maximum energy point that all primes less than the specified number may reach was determined.

Prediction on the maximum gaps Now we can predict the maximum energy level that all prime numbers less than any given number can reach by the above established methods, but this is not our ultimate goal. We want to understand what is the maximum interval between all consecutive primes less than a given arbitrary number. Here, we established the relationship between each energy level and the maximum interval of consecutive prime at its corresponding energy level according to the aforementioned hierarchical progression model.

Figure 5a shows the change trend of the maximum prime interval corresponding to each layer when the prime energy level increases from 1 to 80, and depicts the variation region and average

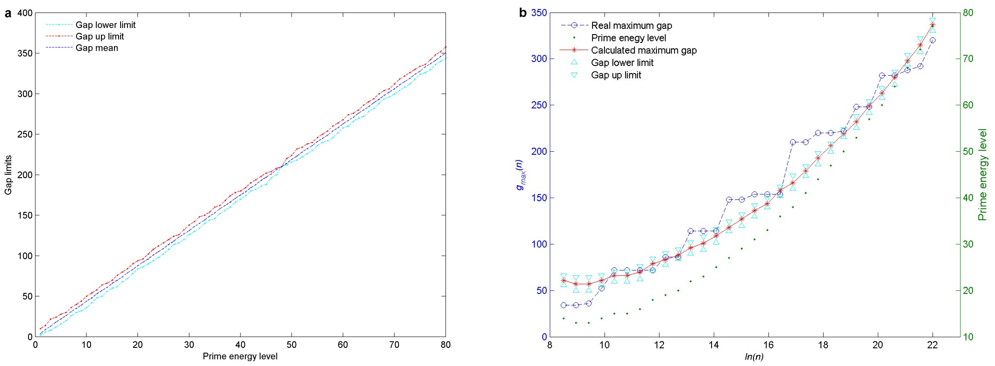


Figure 5: Prediction on the maximum gaps changing with given prime numbers possessing different energy levels. **a**,The region of gaps changing with the prime energy level. Here, cyan, red and blue line represent the lower limit, up limit and the mean value of prime gaps, respectively. **b**,The variation trend of maximum gaps changing with given prime numbers possessing different energy levels.

value of the maximum interval corresponding to each layer in which the mean value, upper limit and lower limit of each maximum gap corresponding to the prime energy level were calculated based on Eqs.(3.6)-(3.8). It can be seen from Fig.5a that the maximum prime interval changes approximately linearly with the increase of the energy level. Although the upper and lower limits will be staggered between adjacent levels, each one shows a monotonic growth trend respectively. When the energy level increases to 48, its corresponding upper and lower limits meet at one point, since the maximum and minimum interval on this energy level are both 210 (Fig.2d).

To verify the effectiveness of the established methods, by selecting 30 natural numbers whose natural logarithm values evenly spaced from 8.5 and 22, we calculated their corresponding maximum interval values with the above methods, and compared them with the experimental statistics of all prime numbers below these given natural numbers (Fig.5b). Here, the horizontal axis $\ln(n)$ represents a set of 30 numbers which are selected so that their natural logarithm values are evenly distributed in the interval of $[8.5, 22]$, and the vertical axis $g_{max}(n)$ represents the maximum gaps which occur below the corresponding level shown in green-dots. It can be confirmed from Fig.5b that the prediction value of the maximum interval corresponding to each given number is basically consistent with the experimental data, and although the relative error in some values is somewhat larger, its overall growth trend is in good agreement. So far, we can predict the maximum interval of all prime numbers less than the specified value with a relative accuracy, but there is still a problem that the real value of the maximum interval fluctuates up and down within our prediction range which can also be observed from the experimental data in Fig.5b.

After fitting the experimental data, we conjectured a simple formula (3.4) where the scale factor c is chosen as 1.485, and found that it can well control the maximum interval of prime numbers within its calculated value. Figure 6a shows the comparison between the calculated interval value of each layer (red-dots line) using Eq.(3.4) and the experimental values (blue-dots line) similar to Fig.5b. It can be observed from Fig.6a that all of 30 experimental values fall below the value calculated by the conjecture formula without exception, and the consistency of their change trend is also good. In order to further verify the maximum-gap-prediction formula, we randomly selected 320 numbers by dividing eight regions at 10 times intervals from 10^2 to 10^{10} with 40 numbers in each region, and investigated the maximum interval formed by all prime numbers less than these numbers. The results

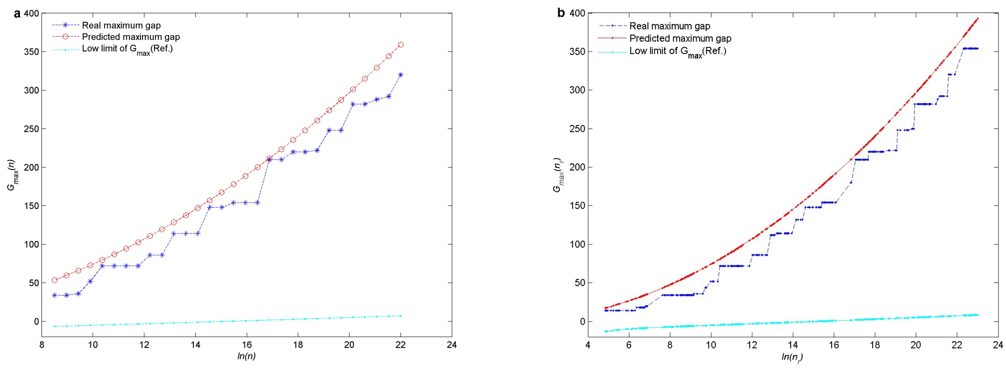


Figure 6: Prediction on the maximum gaps between consecutive primes below a given number. **a**,The comparison of the prediction value and the real value of maximum gaps changing with given prime numbers. Here, red-dot line represent the real maximum gap, and the pink-dot line was depicted from the maximum-gap-prediction formula. **b**, The variation trend of maximum gaps with random numbers below 1×10^{10}.

of comparison between the counted values and the calculated values from the hypothetical formula are shown in Fig.6b, from which it can also be observed that there is no counterexample in all 320 types.

In addition, to compare with the long-gaps-formula of existing research results [30], we drew the variation curve (cyan lines in Figs.6) of the lower limits of the long interval of prime numbers with natural logarithm of numbers, from which it can be confirmed that the prediction values given by our established methods can better approach the real values of the maximum interval.

5 CONCLUSIONS

In conclusion, this paper present two kinds of strategies to predict the value of the maximum gap lessen than a given number by means of our established hierarchical progression model on the distribution of the gaps between consecutive primes, and established three fitting curve formulas according to the correlation between the prime number distribution and the maximum energy level, meanwhile, proposed a hypothetical maximum-gap-prediction formula which can be used to easily determine the upper limit of maximum gap below an arbitrary number. Through the proposed hierarchical progression model, one could use three fitting formulas to judge the corresponding maximum energy layer according to a given number, so as to find the variation interval of the maximum interval relating with the energy layer. On the other hand, the given maximum-gap-prediction formula can be used to determine the upper limit of the maximum interval corresponding to an arbitrary number. In sum, this paper gave a solution for the unpredictable problem of the maximum interval, which is helpful to further probe into the principle for the distribution of prime numbers.

References

[1] A. Raigorodskii, M. T. Rassias, Maier's Matrix Method and Irregularities in the Distribution of Prime Numbers, Springer International Publishing, Cham, 2018, pp. 165–186.

[2] D. A. Goldston, A. H. Ledoan, Jumping champions and gaps between consecutive primes, International Journal of Number Theory 07 (6) (2011) 1413–1413.

[3] K. Ford, B. Green, S. Konyagin, T. Tao, Large gaps between consecutive prime numbers containing square-free numbers and perfect powers of prime numbers, Journal of Functional Analysis 183 (6) (2015) 935–974.

[4] H. Parshall, Small gaps between configurations of prime polynomials, Journal of Number Theory 162 (2016) 35–53.

[5] H. Randriam, Gaps between prime numbers and tensor rank of multiplication in finite fields, Designs Codes and Cryptography 87 (2-3).

[6] J. Pintz, On the Ratio of Consecutive Gaps Between Primes, Springer International Publishing, Cham, 2015, pp. 285–304.

[7] J. Vindas, R. Estrada, A quick distributional way to the prime number theorem, Indagationes Mathematicae 20 (1) (2009) 159–165.

[8] D. P. Datta, On a new proof of the prime number theorem, Mathematics 35 (2) (2011) 143–148.

[9] P. Erds, On the difference of consecutive primes, Bulletin of the American Mathematical Society 54.

[10] J. Pintz, Very large gaps between consecutive primes, Journal of Number Theory 63 (2) (1997) 286–301.

[11] K. Ford, B. Green, S. Konyagin, T. Tao, Large gaps between consecutive prime numbers, Annals of Mathematics 183 (3).

[12] D. Polymath, New equidistribution estimates of zhang type, and bounded gaps between primes, Algebra & Number Theory 8 (9) (2014) 20672199.

[13] R. C. Baker, P. Pollack, Bounded gaps between primes with a given primitive root, ii, Forum Mathematicum 28 (4) (2016) 675–687.

[14] A. Vatwani, Bounded gaps between gaussian primes, Journal of Number Theory 171 (2017) 449–473.

[15] A. Granville, Different approaches to the distribution of primes, Milan Journal of Mathematics 78 (1) (2010) 65–84.

[16] F. Sidokhine, On the difference between consecutive primes and estimates of the number of primes in the interval $(n, 2n)$, mathematics.

[17] L. Troupe, Bounded gaps between prime polynomials with a given primitive root, Finite Fields and Their Applications 37 (2016) 295–310.

[18] H. Maier, M. T. Rassias, Large gaps between consecutive prime numbers containing perfect k-th powers of prime numbers, Journal of Functional Analysis 272 (6) (2017) 2659–2696.

[19] M. Ram Murty, A. Vatwani, Twin primes and the parity problem, Journal of Number Theory 180 (2017) 643–659.

[20] D. A. Goldston, S. W. Graham, A. Panidapu, J. Pintz, J. Schettler, C. Y. Yldrm, Small gaps between almost primes, the parity problem, and some conjectures of erds on consecutive integers ii, Journal of Number Theory 221 (2021) 222–231.

[21] A. Al-Kateeb, M. Ambrosino, H. Hong, E. Lee, Maximum gap in cyclotomic polynomials, Journal of Number Theory 229 (2021) 1–15.

[22] I. B. Kolossvry, I. T. Kolossvry, Distance between natural numbers based on their prime signature, Journal of Number Theory 234 (2022) 120–139.

[23] G. H. Hardy, J. E. Littlewood, Some problems of 'partitio numerorum'; iii: On the expression of a number as a sum of primes, Acta Mathematica 44 (1) (1923) 1–70.

[24] D. A. Goldston, J. Pintz, C. Y. Yildirim, Primes in tuples i, Annals of Mathematics 170 (2) (2009) 819–862.

[25] B. Green, Bounded gaps between primes, Annals of Mathematics 179 (3) (2014) pgs. 1121–1174.

[26] Polymath, DHJ, Variants of the selberg sieve, and bounded intervals containing many primes, Research in the Mathematical Sciences 1 (1) (2014) 1–83.

[27] A. Castillo, C. Hall, R. Oliver, P. Pollack, L. Thompson, Bounded gaps between primes in number fields and function fields, Proceedings of the American Mathematical Society.

[28] MAIER, HELMUT, RASSIAS, T. H. Michael, Large gaps between consecutive prime numbers containing square-free numbers and perfect powers of prime numbers., Proceedings of the American Mathematical Society.

[29] D. A. Kaptan, A remark on large gaps between primes in arithmetic progressions, Studia Scientiarum Mathematicarum Hungarica 56 (4) (2019) 536–537.

[30] Ford, Kevin, Green, Ben, Konyagin, Sergei, Maynard, James, Tao, Terence, Long gaps between primes, Journal of the American Mathematical Society.