Relating bubble sort to birthday problem

Jichu Jiang[a,*]

[a]*Beihang University, Beijing, 100191, China*

ARTICLE INFO

Keywords:
Bubble sort
Birthday problem
Rayleigh distribution
Asymptotic distribution

ABSTRACT

Birthday problem is a well-known classic problem in probability theory widely applied in cryptography. Although bubble sort is a popular algorithm leading to some interesting theoretical problems in computer science, its relation to birthday problem has not been found yet. This paper indicates how Rayleigh distribution naturally arises in bubble sort by relating it to birthday problem, which presents a novel direction for analysing bubble sort and birthday problem. Then asymptotic distributions and statistical characteristics of bubble sort and birthday problem with very small absolute errors are presented. Moreover, this paper proves that some common optimizations of bubble sort could lead to average performance degradation.

1. Introduction

Birthday problem is a well-known classic problem in probability theory [4], invariably mentioned in both mathematics textbooks such as [6, 7] and computer textbooks such as [3, 12]. The expectation of its collision is $\sqrt{\pi n/2} \pm \Theta(1)$ [7], and after a suitable scaling, the limiting distribution is Rayleigh distribution [6]. Although birthday problem and its generalizations have received much attention for their applications to cryptography such as [8, 14, 18], sorting algorithms have not been related to birthday problem yet.

Bubble sort is a popular sorting algorithm [1], and leads to some interesting theoretical problems in computer science [12]. Early studies such as [12] found that the expectation of its passes is $n - \sqrt{\pi n/2} \pm \Theta(1)$, and researchers such as [2, 9, 10] have focused on the structure of bubble sort graph in recent years. And there is a strong connection between the passes of bubble sort and the monotonic paths of bubble sort graph, however, little attention has been paid to the asymptotic distribution of the passes.

Moreover, in computer engineering, researchers such as [15, 16, 19] have often compared the average performance of bubble sort with other sorting algorithms throughout many years, which is determined by the distribution of the passes [12]. However, due to the intractability of asymptotic approximation, technical experiment instead of theoretical analysis has been often adopted.

This paper relates bubble sort to birthday problem. In section 2, bounding bubble sort by birthday problem indicates how Rayleigh distribution naturally arises in bubble sort. In section 3, estimating bubble sort with birthday exhibits excellent performance such as (13), even in cases where the number of elements is small as figure 2 shows. In section 4, asymptotic distributions of bubble sort and birthday problem are presented, including the cumulative distribution function (CDF) and the probability mass function (PMF). In section 5, statistical characteristics of bubble sort and birthday problem are derived based on (17), including expectation, variance, moment and characteristic function, with very small absolute errors shown in (27), (28) and (29).

2. Rayleigh distribution in birthday problem and bubble sort

Birthday problem concentrates on the probability that $m+1$ people have distinct birthdays, where n is the number of days in a year, especially for $m = 22, n = 365$ [12]. Formally, it concentrates on the distribution of the collision $C_n = \min\left\{i+1 \mid U_{i+1} \in \left\{U_1, U_2, \cdots, U_i\right\}\right\}$, where $U_1, U_2, U_3, \cdots$ are mutually independent and identically uniformly distributed on n distinct numbers [6], hence

*Corresponding author
saisyc@outlook.com (J. Jiang)
ORCID(s):

$$
\begin{aligned}
\mathbb{P}\left\{C_n > m+1\right\} &= \prod_{1\leqslant i\leqslant m} \mathbb{P}\left\{C_n > i+1 \mid C_n > i\right\} \\
&= \prod_{1\leqslant i\leqslant m} \left(1-\frac{i}{n}\right) \\
&= \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{m}{n}\right)
\end{aligned} \tag{1}
$$

The number of the passes of bubble sort is determined by the inversions. Formally, there holds $P_n = \max\left\{V_i + 1 \mid 1 \leqslant i \leqslant n\right\}$, where P_n is the number of the passes of bubble sort and V_i is the number of the inversions whose smaller component is the i-th smallest element [12]. For a uniformly random permutation of n distinct numbers, its V_i is uniformly distributed on $\{0, 1, \cdots, n-i\}$ [12], hence

$$
\begin{aligned}
\mathbb{P}\left\{P_n \leqslant n-m\right\} &= \prod_{1\leqslant i\leqslant m} \mathbb{P}\left\{V_{m-i+1} + 1 \leqslant n-m\right\} \\
&= \prod_{1\leqslant i\leqslant m} \left(1-\frac{i}{n-m+i}\right) \\
&= \left(1-\frac{1}{n-m+1}\right)\left(1-\frac{2}{n-m+2}\right)\cdots\left(1-\frac{m}{n-m+m}\right)
\end{aligned} \tag{2}
$$

Thus bubble sort can be bounded by birthday problem as follows:

$$
\mathbb{P}\left\{C_{n-(m-1)} > m+1\right\} \;\leqslant\; \mathbb{P}\left\{P_n \leqslant n-m\right\} \;\leqslant\; \mathbb{P}\left\{C_n > m+1\right\} \tag{3}
$$

Consider that $T_n = C_n / \sqrt{n}$ converges in distribution to the standard Rayleigh distribution [6], namely $\mathbb{P}\left\{T_n \leqslant t\right\} \sim 1 - e^{-t^2/2}$ and $\mathbb{P}\left\{C_n > m+1\right\} \sim \exp\dfrac{m(m+1)}{-2n}$, with the relative error displayed below [17]:

$$
1 \;<\; \exp\frac{m(m+1)}{-2n} \Big/ \mathbb{P}\left\{C_n > m+1\right\} \;<\; \exp\frac{(m+1)^3}{6(n-m)^2} \tag{4}
$$

Therefore, $X_n = \left(n - P_n\right) / \sqrt{n}$ also converges in distribution to the standard Rayleigh distribution, namely $\mathbb{P}\left\{X_n \leqslant x\right\} \sim 1 - e^{-x^2/2}$ and $\mathbb{P}\left\{P_n \leqslant n-m\right\} \sim \exp\dfrac{m(m+1)}{-2n}$.

Moreover, $\mathbb{P}\left\{V_{m-i+1} + 1 \leqslant n-m\right\} \sim \mathbb{P}\left\{C_n > i+1 \mid C_n > i\right\}$ indicates why bubble sort can be related to birthday problem. And as presented in figure 1, there hold:

$$
\mathbb{P}\left\{C_n > m+1\right\} = \exp\frac{m(m+1)}{-2(n-m/3)} \pm \Theta\left(\frac{m^4}{n^3}\right) \tag{5}
$$

$$
\mathbb{P}\left\{P_n \leqslant n-m\right\} = \exp\frac{m(m+1)}{-2(n-2m/3)} \pm \Theta\left(\frac{m^4}{n^3}\right) \tag{6}
$$

3. Estimation of bubble sort with birthday problem

To estimate the number of the passes of bubble sort with birthday problem, especially to analyse the relative error, the Laurent series of their distribution functions deserve attention.

Denote s_k as the summation of k-th power of the first m positive integers, namely $s_k = 1^k + 2^k + \cdots + m^k$, then using the Maclaurin series of $\ln(1-x)$ and $\ln(1+x)$ gives

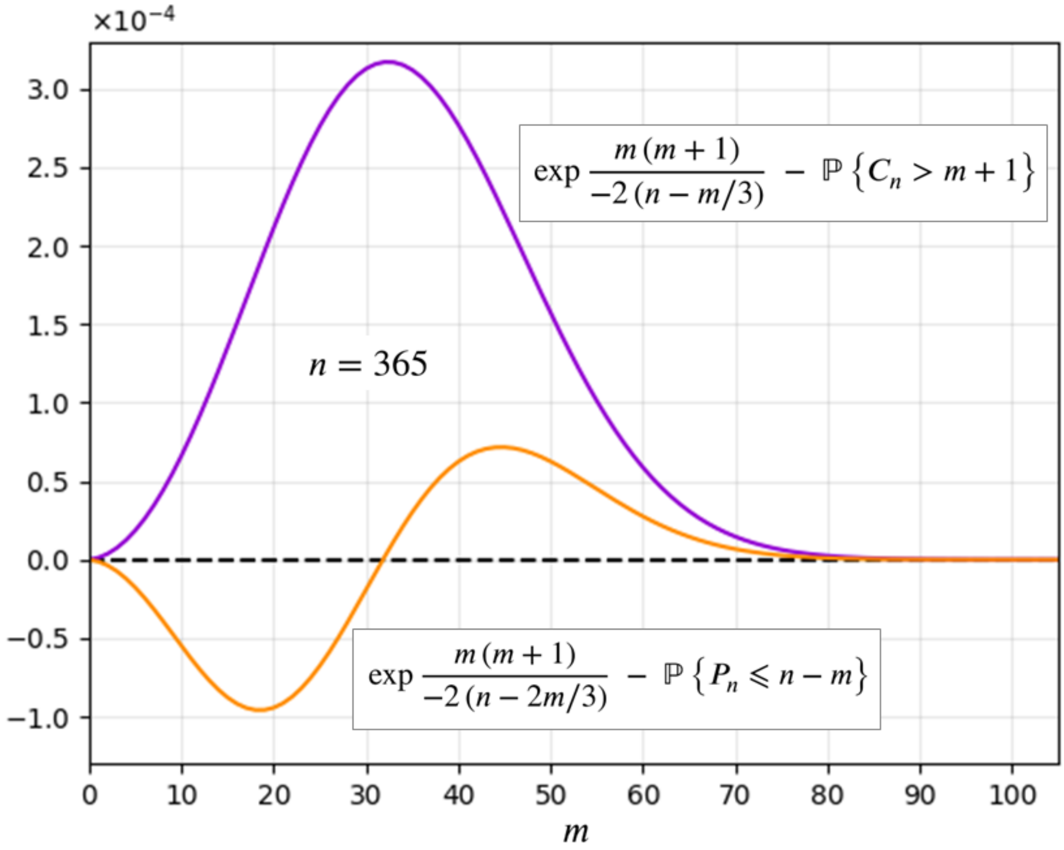

Figure 1: $n = 365$

$$\ln \mathbb{P}\left\{C_n > m+1\right\} = \sum_{1\leqslant i\leqslant m} + \ln\left(1 - \frac{i}{n}\right) = \sum_{1\leqslant j} \frac{-s_j}{jn^j} \tag{7}$$

$$\ln \mathbb{P}\left\{P_{n+m} \leqslant n\right\} = \sum_{1\leqslant i\leqslant m} - \ln\left(1 + \frac{i}{n}\right) = \sum_{1\leqslant j} \frac{(-1)^j s_j}{jn^j} \tag{8}$$

Suppose that q is a positive integer less than m, then using Newton's binomial theorem gives

$$\ln \mathbb{P}\left\{C_{n-q} > m+1\right\} = \sum_{1\leqslant j} \frac{-s_j}{j(n-q)^j} = \sum_{1\leqslant k\leqslant j} \binom{j}{k} \frac{-s_k q^{j-k}}{jn^j} \tag{9}$$

$$\begin{aligned} \ln \mathbb{P}\left\{P_n \leqslant n-m\right\} &= \sum_{1\leqslant j} \frac{(-1)^j s_j}{j(n-m)^j} \\ &= \sum_{1\leqslant k\leqslant j} \binom{j}{k} \frac{(-1)^k s_k m^{j-k}}{jn^j} \end{aligned} \tag{10}$$

$$
\begin{aligned}
&= -\frac{m(m+1)}{2n} \\
&- \frac{m(m+1)(4m-1)}{12n^2} \\
&- \frac{m^2(m+1)(3m-1)}{12n^3} \\
&- \frac{m(m+1)\left(24m^3-9m^2-m+1\right)}{120n^4} \\
&- \frac{m^2(m+1)\left(10m^3-4m^2-m+1\right)}{60n^5} \\
&\pm \Theta\left(\frac{m^7}{n^6}\right)
\end{aligned}
$$

Therefore, estimating passes of bubble sort with birthday problem yields the following relative error:

$$\mathbb{P}\left\{C_n > m+1\right\} / \mathbb{P}\left\{P_n \leqslant n-m\right\} = 1 + \frac{(m-1)\,m\,(m+1)}{6\,(n-m/2)^2} \pm \Theta\left(\frac{m^5}{n^4}\right) \tag{11}$$

Furthermore, the solution of $\mathbb{P}\left\{P_n \leqslant n-m\right\} = \mathbb{P}\left\{C_{n-q} > m+1\right\}$ is $q = \frac{m-1}{3} \pm \Theta\left(\frac{1}{\sqrt{n}}\right)$. And as presented in figure 2, there holds:

$$\mathbb{P}\left\{C_{n-(m-1)/3} > m+1\right\} / \mathbb{P}\left\{P_n \leqslant n-m\right\} = 1 - \frac{(m-1)\,m\,(m+1)\,(m+2)\,(2m+1)}{270\,(n-(4m-1)/6)^4} \pm \Theta\left(\frac{m^7}{n^6}\right) \tag{12}$$

For example, substituting $m = 22, n = 365$ gives

$$
\begin{aligned}
&\mathbb{P}\left\{C_{358} > 23\right\} \approx 0.4857834 \\
&\mathbb{P}\left\{P_{365} \leqslant 343\right\} \approx 0.4857848 \\
&\mathbb{P}\left\{C_{365} > 23\right\} \approx 0.4927028
\end{aligned} \tag{13}
$$

4. Asymptotic distributions of bubble sort and birthday problem

The factorial function is pivotal for analysing bubble sort and birthday problem, because there hold $\mathbb{P}\left\{P_n \leqslant n-m\right\} = (n-m)!\,(n-m)^m\,/\,n!$ [12] and $\mathbb{P}\left\{C_n > m+1\right\} = n!/\,(n-m-1)!/n^{m+1}$ [7] for $n-m \in \{1, 2, \cdots, n\}$.

Thus $\mathbb{P}\left\{X_n \geqslant x\right\}$ is defined on a discrete set $S = \left\{x \in \mathbf{R} \mid n - x\sqrt{n} \in \{1, 2, \cdots, n\}\right\}$ due to $X_n = \left(n - P_n\right)/\sqrt{n}$ and satisfies

$$\mathbb{P}\left\{X_n \geqslant x\right\} = \left(n - x\sqrt{n}\right)!\,\left(n - x\sqrt{n}\right)^{x\sqrt{n}}\,/\,n! \tag{14}$$

To obtain an extension, assume that $v_n(x)$ is defined on an interval $T = \left\{x \in \mathbf{R} \mid 1 \leqslant n - x\sqrt{n} \leqslant n\right\}$ and satisfies

$$v_n(x) = \Gamma\left(n - x\sqrt{n} + 1\right)\,\left(n - x\sqrt{n}\right)^{x\sqrt{n}}\,/\,\Gamma(n+1) \tag{15}$$

Then $\mathbb{P}\left\{X_n \geqslant x\right\} = v_n(x)$ for $x \in S$ and $v_n(x)$ is real analytic for $x \in T$.

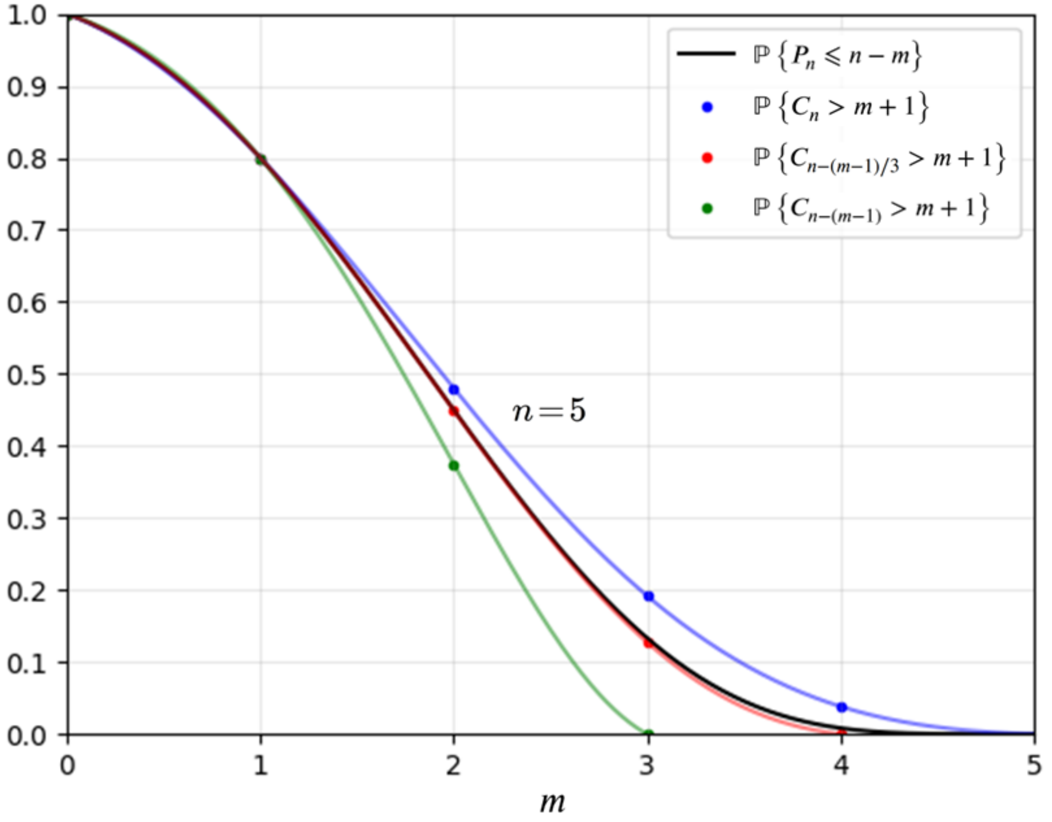

Figure 2: $n = 5$

Furthermore, consider the Stirling's series:

$$\Gamma(z+1) = \sqrt{2\pi z}\left(\frac{z}{e}\right)^z \exp\left(\frac{1}{12z} \pm \Theta\left(\frac{1}{z^3}\right)\right) \tag{16}$$

Therefore, the Laurent series of $\ln v_n(x)$ is

$$\ln v_n(x) = -\frac{x^2}{2} - \frac{2x^3+3x}{6\sqrt{n}} - \frac{x^4+x^2}{4n} - \frac{12x^5+10x^3-5x}{60n\sqrt{n}} - \frac{4x^6+3x^4-2x^2}{24n^2} \pm \Theta\left(\frac{x^7}{n^2\sqrt{n}}\right) \tag{17}$$

Then the cumulative distribution function (CDF) of X_n denoted by $F_{X_n}(x)$ is defined on S and satisfies

$$F_{X_n}(x) = 1 - v_n\left(x + \frac{1}{\sqrt{n}}\right) = 1 - e^{-x^2/2}\exp\left(-\frac{2x^3+9x}{6\sqrt{n}} \pm \Theta\left(\frac{x^4}{n}\right)\right) \tag{18}$$

And the probability mass function (PMF) of X_n denoted by $f_{X_n}(x)$ is defined on S and satisfies

$$f_{X_n}(x) = v_n(x) - v_n\left(x + \frac{1}{\sqrt{n}}\right) = \frac{e^{-x^2/2}x}{\sqrt{n}}\exp\left(-\frac{x^4-3}{3x\sqrt{n}} \pm \Theta\left(\frac{x^4}{n}\right)\right) \tag{19}$$

Thus $X_n = \left(n - P_n\right) / \sqrt{n}$ converges in distribution to the standard Rayleigh distribution due to $\lim_{n\to\infty} F_{X_n}(x) = 1 - e^{-x^2/2}$, which is consistent with the result in section 2.

Similarly, for birthday problem, applying the Stirling's series yields:

$$F_{T_n}(t) = 1 - e^{-t^2/2} \exp\left(-\frac{t^3 - 3t}{6\sqrt{n}} \pm \Theta\left(\frac{t^4}{n}\right)\right) \tag{20}$$

$$f_{T_n}(t) = \frac{e^{-t^2/2} t}{\sqrt{n}} \exp\left(-\frac{t^4 - 9t^2 + 6}{6t\sqrt{n}} \pm \Theta\left(\frac{t^4}{n}\right)\right) \tag{21}$$

5. Statistical characteristics of bubble sort and birthday problem

Euler–Maclaurin summation formula offers an effective approach for analysing the statistical characteristics of bubble sort and birthday problem, such as expectation, variance, moment and characteristic function.

To obtain the k-th moment of X_n, using summation by parts gives

$$\mathbb{E}\left(X_n^k\right) = \sum_{x\in S} x^k\, \mathbb{P}\left\{X_n = x\right\} = \sum_{x\in S}\left(x^k - \left(x - 1/\sqrt{n}\right)^k\right) \mathbb{P}\left\{X_n \geqslant x\right\} + \left(-1/\sqrt{n}\right)^k \tag{22}$$

Then, for a fixed positive number $\epsilon < 1/6$, using Euler–Maclaurin summation formula gives

$$\sum_{x\in S} v_n(x) = \sqrt{n}\int_0^{n^\epsilon} v_n(x)\,dx + \frac{v_n(0)}{2} - \frac{v_n'(0)}{12\sqrt{n}} + \frac{v_n'''(0)}{720n\sqrt{n}} \pm \Theta\left(\frac{1}{n^2}\right) \tag{23}$$

Hence, it is consistent with Knuth's results [11] that the expectation of X_n is

$$\begin{aligned} \mathbb{E}\left(X_n\right) &= \hat{\mathbb{E}}\left(X_n\right) \pm \Theta\left(\frac{1}{n^2\sqrt{n}}\right) \\ \hat{\mathbb{E}}\left(X_n\right) &= \sqrt{\frac{\pi}{2}} - \frac{5}{3\sqrt{n}} + \frac{11}{24n}\sqrt{\frac{\pi}{2}} + \frac{4}{135n\sqrt{n}} - \frac{71}{1152n^2}\sqrt{\frac{\pi}{2}} \end{aligned} \tag{24}$$

Similarly, by applying Euler–Maclaurin summation formula, the expectation of the square of X_n is

$$\begin{aligned} \mathbb{E}\left(X_n^2\right) &= \hat{\mathbb{E}}\left(X_n^2\right) \pm \Theta\left(\frac{1}{n^2\sqrt{n}}\right) \\ \hat{\mathbb{E}}\left(X_n^2\right) &= 2 - 4\sqrt{\frac{\pi}{2n}} + \frac{5}{n} - \frac{5}{3n}\sqrt{\frac{\pi}{2n}} - \frac{4}{135n^2} \end{aligned} \tag{25}$$

And the variance of X_n is

$$\begin{aligned} \mathbb{V}\left(X_n\right) &= \hat{\mathbb{V}}\left(X_n\right) \pm \Theta\left(\frac{1}{n^2\sqrt{n}}\right) \\ \hat{\mathbb{V}}\left(X_n\right) &= \frac{4-\pi}{2} - \frac{2}{3}\sqrt{\frac{\pi}{2n}} + \frac{160 - 33\pi}{72n} - \frac{107}{540n}\sqrt{\frac{\pi}{2n}} - \frac{1125\pi - 1792}{25920n^2} \end{aligned} \tag{26}$$

For example, substituting $n = 10000$ gives

$$\begin{aligned}\mathbb{E}\left(X_{10000}\right) &\approx 1.23670494307038\\ \hat{\mathbb{E}}\left(X_{10000}\right) &\approx 1.23670494307065\end{aligned} \tag{27}$$

$$\begin{aligned}\mathbb{E}\left(X_{10000}^2\right) &\approx 1.950365345384\\ \hat{\mathbb{E}}\left(X_{10000}^2\right) &\approx 1.950365345354\end{aligned} \tag{28}$$

$$\begin{aligned}\mathbb{V}\left(X_{10000}\right) &\approx 0.4209262291695\\ \hat{\mathbb{V}}\left(X_{10000}\right) &\approx 0.4209262291679\end{aligned} \tag{29}$$

Furthermore, notice that the moment of R is $\mathbb{E}\left(R^k\right)/\sqrt{2}^{\,k} = \Gamma(k/2+1)$ and the characteristic function of R is $\varphi_R(z) = \sqrt{2\pi}\, ize^{-z^2/2}\,\Phi(iz)+1$ where R is a random variable following the standard Rayleigh distribution and Φ is the cumulative distribution function of the standard normal distribution. The moments and the characteristic functions of X_n and T_n are as follows:

$$\mathbb{E}\left(X_n^k\right)/\sqrt{2}^{\,k} = \Gamma\left(\frac{k}{2}+1\right) - \frac{\sqrt{2}k(k+4)}{6\sqrt{n}}\Gamma\left(\frac{k+1}{2}\right) \pm \Theta\left(\frac{1}{n}\right) \tag{30}$$

$$\mathbb{E}\left(T_n^k\right)/\sqrt{2}^{\,k} = \Gamma\left(\frac{k}{2}+1\right) - \frac{\sqrt{2}k(k-5)}{12\sqrt{n}}\Gamma\left(\frac{k+1}{2}\right) \pm \Theta\left(\frac{1}{n}\right) \tag{31}$$

$$\varphi_{X_n}(z) = \sqrt{2\pi}\, ize^{-z^2/2}\,\Phi(iz)\left(1-\frac{iz\left(6-z^2\right)}{3\sqrt{n}}\right)+\left(1-\frac{iz\left(5-z^2\right)}{3\sqrt{n}}\right)\pm\Theta\left(\frac{1}{n}\right) \tag{32}$$

$$\varphi_{T_n}(z) = \sqrt{2\pi}\, ize^{-z^2/2}\,\Phi(iz)\left(1+\frac{iz\left(3+z^2\right)}{6\sqrt{n}}\right)+\left(1+\frac{iz\left(4+z^2\right)}{6\sqrt{n}}\right)\pm\Theta\left(\frac{1}{n}\right) \tag{33}$$

Therefore, both X_n and T_n converge in distribution to the standard Rayleigh distribution with convergence of all moments, which is consistent with the result in section 2. And they satisfy the method of moments and basic limit laws in [13].

6. Discussion

In analytic combinatorics, Kuba and Panholzer studied the distribution of inversions in labelled tree families and indicated how mixed Poisson-Rayleigh distributions naturally arise in several critical composition schemes [13]. Inspired by their work, in section 2 and section 3, this paper relates bubble sort to birthday problem, which presents a novel direction for analysing birthday problem and bubble sort. It is somewhat amazing that this relation has eluded discovery for so long.

In computer science, Knuth studied the analysis of bubble sort and derived the expectation of its passes by the inversions [11, 12]. Based on his work, in section 4 and section 5, this paper develops this approach and derives several effective asymptotic approximations of bubble sort and birthday problem, including asymptotic distributions and statistical characteristics, which completely addresses the analysis of the average performance of bubble sort for distinct elements. And for analysing variations of bubble sort such as [5], some generalizations of this approach should be effective.

Moreover, in computer engineering, nearly every description of bubble sort mentions the optimization that terminates if no swaps are made in a pass [1], which reduces the number of the comparisons of elements by nX_n^2. Thus some common implementations of this optimization such as [16, 19] could lead to average performance degradation for all large enough n, because the cost of checking if no swaps are made is $\Theta\left(n^2\right)$ whereas whereas only $\mathbb{E}\left(nX_n^2\right) = \Theta(n)$ comparisons can be reduced.

7. Acknowledgement

Sincere gratitude is expressed to Professor Ke Xu for his encouragement and insightful suggestions throughout this research.

References

[1] Astrachan, O., 2003. Bubble sort: an archaeological algorithmic analysis. SIGCSE Bull. 35, 1–5. URL: https://doi.org/10.1145/792548.611918, doi:10.1145/792548.611918.

[2] Cai, H., Liu, H., Lu, M., 2015. Fault-tolerant maximal local-connectivity on bubble-sort star graphs. Discrete Applied Mathematics 181, 33–40. URL: https://www.sciencedirect.com/science/article/pii/S0166218X14004235, doi:https://doi.org/10.1016/j.dam.2014.10.006.

[3] Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C., 2009. Introduction to Algorithms, Third Edition. 3rd ed., The MIT Press.

[4] DasGupta, A., 2005. The matching, birthday and the strong birthday problem: a contemporary review. Journal of Statistical Planning and Inference 130, 377–389. URL: https://www.sciencedirect.com/science/article/pii/S0378375804002721, doi:https://doi.org/10.1016/j.jspi.2003.11.015. herman Chernoff: Eightieth Birthday Felicitation Volume.

[5] Dobosiewicz, W., 1980. An efficient variation of bubble sort. Information Processing Letters 11, 5–6. URL: https://www.sciencedirect.com/science/article/pii/0020019080900228, doi:https://doi.org/10.1016/0020-0190(80)90022-8.

[6] Durrett, R., 2010. Probability: Theory and Examples. 4th ed., Cambridge University Press, USA.

[7] Flajolet, P., Sedgewick, R., 2009. Analytic Combinatorics. Cambridge University Press.

[8] Galbraith, S.D., Holmes, M., 2012. A non-uniform birthday problem with applications to discrete logarithms. Discrete Applied Mathematics 160, 1547–1560. URL: https://www.sciencedirect.com/science/article/pii/S0166218X12000807, doi:https://doi.org/10.1016/j.dam.2012.02.019.

[9] Guo, J., Lu, M., 2016. Conditional diagnosability of bubble-sort star graphs. Discrete Applied Mathematics 201, 141–149. URL: https://www.sciencedirect.com/science/article/pii/S0166218X15003790, doi:https://doi.org/10.1016/j.dam.2015.07.026.

[10] Guo, J., Lu, M., 2021. Edge-fault-tolerant strong menger edge connectivity of bubble-sort star graphs. Discrete Applied Mathematics 297, 109–119. URL: https://www.sciencedirect.com/science/article/pii/S0166218X21001116, doi:https://doi.org/10.1016/j.dam.2021.03.006.

[11] Knuth, D.E., 1997. The art of computer programming, volume 1 (3rd ed.): fundamental algorithms. Addison Wesley Longman Publishing Co., Inc., USA.

[12] Knuth, D.E., 1998. The art of computer programming, volume 3: (2nd ed.) sorting and searching. Addison Wesley Longman Publishing Co., Inc., USA.

[13] Kuba, M., Panholzer, A., 2016. On moment sequences and mixed poisson distributions. Probability Surveys 13, 89 – 155. URL: https://doi.org/10.1214/14-PS244, doi:10.1214/14-PS244.

[14] Lindo, A., Sagitov, S., 2015. Asymptotic results for the number of wagner's solutions to a generalised birthday problem. Statistics & Probability Letters 107, 356–361. URL: https://www.sciencedirect.com/science/article/pii/S0167715215003338, doi:https://doi.org/10.1016/j.spl.2015.09.014.

[15] Paul, T., 2022. Enhancement of bubble and insertion sort algorithm using block partitioning, in: 2022 25th International Conference on Computer and Information Technology (ICCIT), pp. 412–417. doi:10.1109/ICCIT57492.2022.10055404.

[16] Rahmani, M.K.I., 2022. Smart bubble sort: A novel and dynamic variant of bubble sort algorithm. Computers, Materials & Continua 71, 4895–4913. URL: http://www.techscience.com/cmc/v71n3/46526, doi:10.32604/cmc.2022.023837.

[17] Sayrafiezadeh, M., 1994. The birthday problem revisited. Mathematics Magazine 67, 220–223. URL: http://www.jstor.org/stable/2690615.

[18] Wu, H., Zhuang, J., Duan, Q., Zhu, Y., 2022. Non-uniform birthday problem revisited: Refined analysis and applications to discrete logarithms. Information Processing Letters 175, 106225. URL: https://www.sciencedirect.com/science/article/pii/S002001902100140X, doi:https://doi.org/10.1016/j.ipl.2021.106225.

[19] Yu, P., Yang, Y., Gan, Y., 2012. Experiment analysis on the bubble sort algorithm and its improved algorithms, in: Zhu, R., Ma, Y. (Eds.), Information Engineering and Applications, Springer London, London. pp. 95–102.