AIMS Mathematics, 11(11): 38569–38577.
DOI: 10.3934/math.20261488
Received: 21 September 2026
Revised: 04 November 2026
Accepted: 13 November 2026
Published: 19 November 2026

https://www.aimspress.com/journal/Math

Research article

Cramér moderate deviations for a supercritical multi-type Galton–Watson process

Juan Wang *and Luyang Zhai

College of Science, University of Shanghai for Science and Technology, Shanghai 200093, China

* **Correspondence:** Email: jwang@usst.edu.cn.

Abstract: This paper establishes Cramér moderate deviations for the Lotka–Nagaev estimator in a supercritical multi-type Galton–Watson process. By conditioning on the total population, the normalized estimator reduces to a sum of independent and identically distributed random vectors with fixed sample size. This allows us to apply the high-dimensional Cramér moderate deviation theorem of Fang et al. to establish moderate deviations for high-dimensional Galton–Watson processes.

Keywords: supercritical; multi-type Galton–Watson process; Lotka–Nagaev estimator; Cramér moderate deviations
Mathematics Subject Classification: 60J80, 60F10

1. Introduction

The Galton–Watson process is a fundamental model in the theory of branching processes and is widely applied in fields such as population genetics, epidemiology, and statistical (see, e.g., [1–3]). The theory of Cramér type moderate deviations has been extensively studied in the literature. Following the pioneering work of Cramér [4], substantial progress has been made in various directions. For sums of independent random variables, classical results were obtained by Bahr [5] and Petrov [6]. More recently, Cramér type moderate deviations have been established for martingales [7, 8], and bootstrap sample quantiles [9].

The Cramér moderate deviation theory for one-dimensional Galton–Watson processes has been well studied. For the one-dimensional case, let $\{Z_n\}_{n\geq 0}$ be a supercritical Galton–Watson process with offspring mean $m = \mathbb{E}Z_1 > 1$ and variance $\nu^2 = \mathrm{Var}(Z_1)$ [10]. Doukhan et al. [11] considered the normalized martingale

$$H_{n_0,n} = \frac{1}{\nu\sqrt{n}} \sum_{k=n_0}^{n_0+n-1} \sqrt{Z_k}\left(\frac{Z_{k+1}}{Z_k} - m\right)$$

and proved that, under suitable moment conditions,

$$\ln \frac{\mathbb{P}(H_{n_0,n} \geq x)}{1-\Phi(x)} = O\left(\frac{x^3}{\sqrt{n}} + (1+x)\frac{\ln n}{\sqrt{n}}\right)$$

uniformly for $0 \leq x = o(\sqrt{n})$. Subsequently, Wang and Peng [12] extended these results to the case with immigration.

Regarding high-dimensional Cramér moderate deviations, Fang et al. [13] established the following Cramér moderate deviation result. Let $X_1, \ldots, X_n$ be i.i.d. random vectors in $\mathbb{R}^d$ with $\mathbb{E}X_1 = 0$, $\mathrm{Cov}(X_1) = I_d$, and suppose the Cramér condition $\mathbb{E}e^{t_0|X_1|} < \infty$ holds for some $t_0 > 0$. Define $W = \frac{1}{\sqrt{n}}\sum_{i=1}^n X_i$ and let $Z \sim N(0, I_d)$. Then, for any symmetric positive definite matrix Q with largest eigenvalue 1 and for $0 \leq x \leq \varepsilon n^{1/6}$,

$$\left|\frac{\mathbb{P}(|Q^{1/2}W| > x)}{\mathbb{P}(|Q^{1/2}Z| > x)} - 1\right| \leq C\left(\frac{1+x^3}{\det(Q^{1/2})\, n^{d/(d+1)}}\mathbf{1}_{\{d\leq 4\}} + \frac{1+x^5}{\det(Q^{1/2})\, n}\mathbf{1}_{\{d\geq 5\}} + \frac{x^6}{n}\right),$$

where ε and C are positive constants depending only on d and t_0.

However, the Cramér moderate deviation theory for Galton–Watson processes remains unexplored in the high-dimensional setting. This paper makes a first attempt in this direction. We consider a multi-type Galton–Watson process, in which the offspring vector of each individual is independent and identically distributed. Using a conditioning argument, we reformulate the Lotka–Nagaev estimator Z_{n+1}/T_n, where $T_n = \sum_{j=1}^d Z_n(j)$ denotes the total population size of the n-th generation, as a sum of independent and identically distributed random vectors with deterministic sample size. Based on the theoretical results of Fang et al. [13], we establish the Cramér moderate deviation principle for high-dimensional Galton–Watson processes.

Unlike the i.i.d. setting in Fang et al. [13] where the sample size is deterministic, the application to Galton–Watson processes requires handling a random sample size T_n. We address this issue by combining the conditioning method with an exponential smallness lemma to control the probability of the population being unusually small. Furthermore, while our high-dimensional bounds do not directly recover the sharper one-dimensional rate of Doukhan et al. [11] , this work extends the scalar framework to a vector-valued setting. By utilizing the high-dimensional moderate deviation theorem of Fang et al. [13], our results show a dimension-dependent convergence rate. In particular, for $d \geq 5$, the rate becomes $1/n$, which differs from the behavior typically observed in one-dimensional settings.

2. Preliminaries and main results

Let $\{\mathbf{Z}_n = (Z_n^{(1)}, Z_n^{(2)}, \ldots, Z_n^{(d)})\}_{n\geq 0}$ be a supercritical d-dimensional Galton–Watson process. $Z_n^{(j)}$ denotes the number of individuals of type j in the n-th generation. The offspring of each individual is represented by a vector $\xi \in \mathbb{R}^d$, where $\xi^{(j)}$ is the number of offspring of type j. The offspring vectors of different individuals are independent and identically distributed. Define the mean vector and the covariance matrix of ξ:

$$\mu = \mathbb{E}\xi = (\mu_1, \ldots, \mu_d)^\top, \qquad \Sigma = \mathrm{Cov}(\xi) = \mathbb{E}[(\xi-\mu)(\xi-\mu)^\top],$$

and we assume that Σ is positive definite.

Since all individuals share the same offspring distribution, the mean matrix is $M = \boldsymbol{\nu}\mu^\top$, where $\boldsymbol{\nu} = (1, 1, \ldots, 1)^\top$ is the d-dimensional vector. The largest eigenvalue of M is given by $\rho = \sum_{j=1}^{d} \mu_j$. Throughout this paper, we assume the process is supercritical, i.e., $\rho > 1$.

Let $T_n = \sum_{j=1}^{d} Z_n^{(j)}$ denote the total population of the n-th generation. The evolution of the process is given by

$$\mathbf{Z}_{n+1} = \sum_{i=1}^{T_n} \xi_i,$$

where $\xi_1, \xi_2, \ldots$ are i.i.d. offspring vectors, independent of $\{\mathbf{Z}_k\}_{k \le n}$.

The Lotka–Nagaev estimator is defined as $\mathbf{Z}_{n+1}/T_n$. To establish moderate deviation theory, we consider its normalized version:

$$R_n = \sqrt{T_n}\, \Sigma^{-1/2} \left(\frac{\mathbf{Z}_{n+1}}{T_n} - \mu \right),$$

where $\Sigma^{-1/2}$ is the inverse of the symmetric square root of Σ.

In the following, we present two lemmas that will be used in the proof of our main results.

Lemma 1 is a Cramér moderate deviation result for high-dimensional i.i.d. random vectors, which was proved by Fang et al. [13] using the expansion and symmetry argument in their Appendix A, building upon the classical result of Bahr (1967) [5].

Lemma 2 is the main theorem of Fang et al. [13], which establishes a refined Cramér-type moderate deviation theorem with a rate of convergence matching that of the Berry–Esseen bound. In particular, this result provides dimension-dependent convergence rates.

Lemma 1 (Fang et al. [13])**.** *Let $X_1, \ldots, X_n$ be i.i.d. random vectors in $\mathbb{R}^d$ with $\mathbb{E}X_1 = 0$, $\mathrm{Cov}(X_1) = I_d$, and suppose the Cramér condition $\mathbb{E}e^{t_0|X_1|} < \infty$ holds for some $t_0 > 0$. Define $W = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$ and let $Z \sim N(0, I_d)$. Then, for $0 \le x \le \varepsilon n^{1/6}$,*

$$\left| \frac{\mathbb{P}(|W| > x)}{\mathbb{P}(|Z| > x)} - 1 \right| \le C \left(\frac{1 + x}{\sqrt{n}} + \frac{x^6}{n} \right),$$

where ε and C are positive constants depending only on d and t_0.

Lemma 2 (Fang et al. [13])**.** *Let $X_1, \ldots, X_n$ be i.i.d. random vectors in $\mathbb{R}^d$ with $\mathbb{E}X_1 = 0$, $\mathrm{Cov}(X_1) = I_d$, and let Q be a symmetric positive definite matrix whose largest eigenvalue is 1. Suppose the Cramér condition $\mathbb{E}e^{t_0|X_1|} < \infty$ holds for some $t_0 > 0$. Let $W = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$ and $Z \sim N(0, I_d)$. Then, for $0 \le x \le \varepsilon n^{1/6}$,*

$$\left| \frac{\mathbb{P}(|Q^{1/2}W| > x)}{\mathbb{P}(|Q^{1/2}Z| > x)} - 1 \right| \le C \left(\frac{1 + x^3}{\det(Q^{1/2})\, n^{d/(d+1)}} \mathbf{1}_{\{d \le 4\}} + \frac{1 + x^5}{\det(Q^{1/2})\, n} \mathbf{1}_{\{d \ge 5\}} + \frac{x^6}{n} \right),$$

where ε and C are positive constants depending only on d and t_0.

The following lemma is an exponential smallness lemma, a direct consequence of Lemma 6.2 of Doukhan et al. [11]. A brief proof will be given in Section 3.

Lemma 3. *Let $T_n = \sum_{j=1}^{d} Z_n^{(j)}$ be the total population of a supercritical multi-type Galton–Watson process satisfying the Cramér condition $\mathbb{E}e^{t_0|\Sigma^{-1/2}(\xi_1 - \mu)|} < \infty$. Then, there exist constants $C, c > 0$ such that for all $n \ge 1$,*

$$\mathbb{P}(T_n \le n) \le Ce^{-cn}.$$

The following two theorems are the main results of this paper.

Remark 1. *The assumptions in Theorems 1 and 2 play distinctly different mathematical roles. The Cramér conditions ensure that the moment generating function is well-defined near the origin. In Lemma 3, the supercritical condition $\rho > 1$ guarantees that we can control the small population probability $\mathbb{P}(T_n \le n)$ as $n \to \infty$. This property is essential for handling the small-k terms in the proofs of Theorems 1 and 2. The condition $0 \le x \le \varepsilon n^{1/6}$, appearing in Lemmas 1 and 2 and applied in Theorems 1 and 2, ensures that the remainder terms in the moderate deviation expansion vanish as $n \to \infty$.*

Theorem 1. *Let $Z \sim N(0, I_d)$. Suppose the offspring distribution satisfies the Cramér condition $\mathbb{E}e^{t_0|\Sigma^{-1/2}(\xi_1-\mu)|} < \infty$ for some $t_0 > 0$. Then, there exist constants $\varepsilon > 0$ and $C > 0$ depending only on d and t_0 such that for all $0 \le x \le \varepsilon n^{1/6}$,*

$$\left|\frac{\mathbb{P}(|R_n| > x)}{\mathbb{P}(|Z| > x)} - 1\right| \le C\left(\frac{1+x}{\sqrt{n}} + \frac{x^6}{n}\right).$$

The next theorem refines the error bound by providing dimension-dependent convergence rates, following the approach of Lemma 2.

Theorem 2. *Let Q be a given $d \times d$ symmetric positive definite matrix with largest eigenvalue 1, and let $Z \sim N(0, I_d)$. Under the same conditions as in Theorem 1, there exist constants $\varepsilon > 0$ and $C > 0$ depending only on d, t_0 and the given matrix Q such that for all $0 \le x \le \varepsilon n^{1/6}$,*

$$\left|\frac{\mathbb{P}(|Q^{1/2}R_n| > x)}{\mathbb{P}(|Q^{1/2}Z| > x)} - 1\right| \le C\left(\frac{1+x^3}{\det(Q^{1/2})\, n^{\frac{d}{d+1}}}\mathbf{1}_{\{d\le 4\}} + \frac{1+x^5}{\det(Q^{1/2})\, n}\mathbf{1}_{\{d\ge 5\}} + \frac{x^6}{n}\right).$$

Remark 2. *The factor $\frac{1}{\det(Q^{1/2})}$ in the bound indicates that the constant C depends on the given matrix Q. When Q becomes singular, this results in a degenerate case, in which the model can be reduced to a lower-dimensional form. In such cases, the bound deteriorates, and the high-dimensional convergence rate may no longer hold.*

3. Proofs

We now proceed to prove the main results of this paper. We begin with the exponential smallness Lemma 3.

Proof. Since the process is supercritical, the eigenvalue satisfies $\rho > 1$. The moment generating function $\phi(\theta) = \mathbb{E}[e^{-\theta T_1}]$ satisfies $\phi(0) = 1$ and $\phi'(0) = -\mathbb{E}[T_1] = -\rho < -1$. By Taylor expansion, there exists a sufficiently small $\theta > 0$ and a constant $\lambda \in (0, 1)$ such that

$$\mathbb{E}[e^{-\theta T_1}] \le \lambda < e^{-\theta}.$$

Let Y_i denote the number of offspring produced by the i-th individual in the $(n-1)$-th generation, for $1 \le i \le T_{n-1}$. Then $T_n = Y_1 + Y_2 + \cdots + Y_{T_{n-1}}$. Because all individuals share the same offspring distribution, each Y_i is distributed identically to T_1. By the branching property, we have

$$\mathbb{E}[e^{-\theta T_n} \mid T_{n-1}] = \mathbb{E}\left[e^{-\theta(Y_1+Y_2+\cdots+Y_{T_{n-1}})} \mid T_{n-1}\right]$$

$$
\begin{aligned}
&= \prod_{i=1}^{T_{n-1}} \mathbb{E}[e^{-\theta Y_i} \mid T_{n-1}] \\
&= \left(\mathbb{E}[e^{-\theta T_1}]\right)^{T_{n-1}} \\
&\leq \lambda^{T_{n-1}}.
\end{aligned}
$$

Taking the expectation on both sides and applying the standard exponential bounds for multi-type Galton-Watson processes (Athreya-Ney, Chapter V) [10], there exists a constant $C > 0$ such that

$$\mathbb{E}[e^{-\theta T_n}] \leq C\lambda^n.$$

Since $\lambda < e^{-\theta}$, let $c = -\ln\left(\frac{\lambda}{e^{-\theta}}\right) > 0$. By Markov's inequality, we have

$$\mathbb{P}(T_n \leq n) = \mathbb{P}(e^{-\theta T_n} \geq e^{-\theta n}) \leq \frac{\mathbb{E}[e^{-\theta T_n}]}{e^{-\theta n}} \leq C\left(\frac{\lambda}{e^{-\theta}}\right)^n = Ce^{-cn}.$$

This completes the proof. □

We now turn to the proof of Theorem 1.

Proof. Let $T_n = \sum_{j=1}^{d} Z_n^{(j)}$ be the total population of the n-th generation. By the law of total probability,

$$\mathbb{P}(|R_n| > x) = \sum_{k=1}^{\infty} \mathbb{P}(T_n = k)\, \mathbb{P}(|R_n| > x \mid T_n = k).$$

Fix $k \geq 1$. Conditionally on $T_n = k$, the $(n+1)$-th generation population vector can be written as

$$\mathbf{Z}^{(k)} = \sum_{i=1}^{k} \xi_i,$$

where $\xi_1, \ldots, \xi_k$ are i.i.d. offspring vectors. Then

$$\frac{\mathbf{Z}^{(k)}}{k} - \mu = \frac{1}{k}\sum_{i=1}^{k}(\xi_i - \mu).$$

The random vector R_n has the same distribution as

$$R^{(k)} := \frac{1}{\sqrt{k}}\sum_{i=1}^{k} Y_i,$$

where $Y_i = \Sigma^{-1/2}(\xi_i - \mu)$. Since ξ_i are i.i.d., Y_i are also i.i.d. Moreover,

$$\mathbb{E}Y_i = \Sigma^{-1/2}(\mathbb{E}\xi_i - \mu) = 0, \qquad \mathrm{Cov}(Y_i) = \Sigma^{-1/2}\,\mathrm{Cov}(\xi_i)\Sigma^{-1/2} = I_d.$$

So $\mathbb{P}(|R_n| > x \mid T_n = k) = \mathbb{P}(|R^{(k)}| > x)$. Hence, the Y_i satisfy the assumptions of Lemma 1. Applying this lemma with $X_i = Y_i$ and sample size k, we obtain, for any x satisfying $x \leq \varepsilon k^{1/6}$,

$$\frac{\mathbb{P}(|R^{(k)}| > x)}{\mathbb{P}(|Z| > x)} = 1 + O\left(\frac{1+x}{\sqrt{k}} + \frac{x^6}{k}\right),$$

and by the law of total probability, the above expression becomes

$$\mathbb{P}(|R_n| > x) = \mathbb{P}(|Z| > x) \sum_{k=1}^{\infty} \mathbb{P}(T_n = k) \left(1 + O\Big(\frac{1+x}{\sqrt{k}} + \frac{x^6}{k}\Big)\right)$$
$$= \mathbb{P}(|Z| > x) \left(1 + \sum_{k=1}^{\infty} \mathbb{P}(T_n = k)\, O\Big(\frac{1+x}{\sqrt{k}} + \frac{x^6}{k}\Big)\right).$$

Note that the application of the moderate deviation result requires the condition $x \le \varepsilon k^{1/6}$. Since $x \le \varepsilon n^{1/6}$ by assumption, this condition holds automatically when $k > n$, but may fail when $k \le n$. We therefore split the sum into two parts.

For $k \le n$, the condition $x \le \varepsilon k^{1/6}$ may not hold. By Lemma 3, there exist constants $C_1, c > 0$ such that $\mathbb{P}(T_n \le n) \le C_1 e^{-cn}$. This term is negligible and can be absorbed into the error bound. Consequently, we obtain

$$\sum_{k=1}^{n} \mathbb{P}(T_n = k)\, O\Big(\frac{1+x}{\sqrt{k}} + \frac{x^6}{k}\Big) \le \sum_{k=1}^{n} \mathbb{P}(T_n = k)\, O\Big(1 + x + x^6\Big)$$
$$= \mathbb{P}(T_n \le n)\, O\Big(1 + x + x^6\Big)$$
$$= O\Big(e^{-cn}(1 + x + x^6)\Big)$$
$$= O\Big(\frac{1+x}{\sqrt{n}} + \frac{x^6}{n}\Big).$$

Since e^{-cn} decays faster than any polynomial in n, we have $e^{-cn} = O(1/\sqrt{n})$. Consequently,

$$O(e^{-cn}(1 + x + x^6)) = O\Big(\frac{1+x}{\sqrt{n}} + \frac{x^6}{n}\Big).$$

For $k > n$, we have $x \le \varepsilon n^{1/6} < \varepsilon k^{1/6}$. Consequently,

$$\sum_{k=n+1}^{\infty} \mathbb{P}(T_n = k)\, O\left(\frac{1+x}{\sqrt{k}} + \frac{x^6}{k}\right) \le \sum_{k=n+1}^{\infty} \mathbb{P}(T_n = k) O\left(\frac{1+x}{\sqrt{n}} + \frac{x^6}{n}\right) \le O\left(\frac{1+x}{\sqrt{n}} + \frac{x^6}{n}\right).$$

Combining the two parts , we obtain

$$\frac{\mathbb{P}(|R_n| > x)}{\mathbb{P}(|Z| > x)} = 1 + O\left(\frac{1+x}{\sqrt{n}} + \frac{x^6}{n}\right),$$

which is equivalent to

$$\left|\frac{\mathbb{P}(|R_n| > x)}{\mathbb{P}(|Z| > x)} - 1\right| \le C\left(\frac{1+x}{\sqrt{n}} + \frac{x^6}{n}\right)$$

for some constant $C > 0$ independent of n and x. This completes the proof of Theorem 1. □

The proof of Theorem 2 follows similarly.

Proof. The proof follows the same structure as that of Theorem 1. We condition on $T_n = k$ and obtain, as before,

$$R^{(k)} = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} Y_i,$$

where $Y_i = \Sigma^{-1/2}(\xi_i - \mu)$ are i.i.d. with $\mathbb{E}Y_i = 0$ and $\mathrm{Cov}(Y_i) = I_d$, and $\mathbb{P}(|Q^{1/2}R_n| > x \mid T_n = k) = \mathbb{P}(|Q^{1/2}R^{(k)}| > x)$.

Applying Lemma 2 with $X_i = Y_i$, sample size k, and the given matrix Q, we obtain, for any x satisfying $x \le \varepsilon k^{1/6}$,

$$\frac{\mathbb{P}(|Q^{1/2}R^{(k)}| > x)}{\mathbb{P}(|Q^{1/2}Z| > x)} = 1 + O\left(\frac{1+x^3}{\det(Q^{1/2})\, k^{\frac{d}{d+1}}} \mathbf{1}_{\{d\le 4\}} + \frac{1+x^5}{\det(Q^{1/2})\, k} \mathbf{1}_{\{d\ge 5\}} + \frac{x^6}{k}\right),$$

by the law of total probability,

$$\begin{aligned}
\mathbb{P}(|Q^{1/2}R_n| > x) &= \sum_{k=1}^{\infty} \mathbb{P}(T_n = k)\, \mathbb{P}(|Q^{1/2}R_n| > x \mid T_n = k) \\
&= \mathbb{P}(|Q^{1/2}Z| > x) \sum_{k=1}^{\infty} \mathbb{P}(T_n = k)\left(1 + O\left(\frac{1+x^3}{\det(Q^{1/2})\, k^{\frac{d}{d+1}}} \mathbf{1}_{\{d\le 4\}} + \frac{1+x^5}{\det(Q^{1/2})\, k} \mathbf{1}_{\{d\ge 5\}} + \frac{x^6}{k}\right)\right) \\
&= \mathbb{P}(|Q^{1/2}Z| > x)\left(1 + \sum_{k=1}^{\infty} \mathbb{P}(T_n = k)\, O\left(\frac{1+x^3}{\det(Q^{1/2})\, k^{\frac{d}{d+1}}} \mathbf{1}_{\{d\le 4\}} + \frac{1+x^5}{\det(Q^{1/2})\, k} \mathbf{1}_{\{d\ge 5\}} + \frac{x^6}{k}\right)\right).
\end{aligned}$$

As in Theorem 1, we split the sum into two parts. For $k \le n$, the condition $x \le \varepsilon k^{1/6}$ may fail. Using the exponential smallness lemma (Lemma 3), $\mathbb{P}(T_n \le n) \le Ce^{-cn}$, and the crude bounds $1/k^{d/(d+1)} \le 1$, $1/k \le 1$, we obtain

$$\begin{aligned}
&\sum_{k=1}^{n} \mathbb{P}(T_n = k)\, O\left(\frac{1+x^3}{\det(Q^{1/2})\, k^{\frac{d}{d+1}}} \mathbf{1}_{\{d\le 4\}} + \frac{1+x^5}{\det(Q^{1/2})\, k} \mathbf{1}_{\{d\ge 5\}} + \frac{x^6}{k}\right) \\
&\quad \le \sum_{k=1}^{n} \mathbb{P}(T_n = k)\, O\left(\frac{1+x^3}{\det(Q^{1/2})} \mathbf{1}_{\{d\le 4\}} + \frac{1+x^5}{\det(Q^{1/2})} \mathbf{1}_{\{d\ge 5\}} + x^6\right) \\
&\quad = \mathbb{P}(T_n \le n)\, O\left(\frac{1+x^3}{\det(Q^{1/2})} \mathbf{1}_{\{d\le 4\}} + \frac{1+x^5}{\det(Q^{1/2})} \mathbf{1}_{\{d\ge 5\}} + x^6\right) \\
&\quad = O\left(\frac{1+x^3}{\det(Q^{1/2})\, n^{\frac{d}{d+1}}} \mathbf{1}_{\{d\le 4\}} + \frac{1+x^5}{\det(Q^{1/2})\, n} \mathbf{1}_{\{d\ge 5\}} + \frac{x^6}{n}\right).
\end{aligned}$$

For $k > n$, we have $x \le \varepsilon n^{1/6} < \varepsilon k^{1/6}$, and

$$\frac{1}{k^{\frac{d}{d+1}}} \le \frac{1}{n^{\frac{d}{d+1}}}, \qquad \frac{1}{k} \le \frac{1}{n}.$$

Hence,

$$\begin{aligned}
&\sum_{k=n+1}^{\infty} \mathbb{P}(T_n = k) O\left(\frac{1+x^3}{\det(Q^{1/2})\, k^{\frac{d}{d+1}}} \mathbf{1}_{\{d\le 4\}} + \frac{1+x^5}{\det(Q^{1/2})\, k} \mathbf{1}_{\{d\ge 5\}} + \frac{x^6}{k}\right) \\
&\quad = O\left(\frac{1+x^3}{\det(Q^{1/2})\, n^{\frac{d}{d+1}}} \mathbf{1}_{\{d\le 4\}} + \frac{1+x^5}{\det(Q^{1/2})\, n} \mathbf{1}_{\{d\ge 5\}} + \frac{x^6}{n}\right).
\end{aligned}$$

Combining the two parts and noting that the exponentially small term is absorbed into the main error bound, we obtain

$$\frac{\mathbb{P}(|Q^{1/2}R_n| > x)}{\mathbb{P}(|Q^{1/2}Z| > x)} = 1 + O\left(\frac{1+x^3}{\det(Q^{1/2})\, n^{\frac{d}{d+1}}}\, \mathbf{1}_{\{d\le 4\}} + \frac{1+x^5}{\det(Q^{1/2})\, n}\, \mathbf{1}_{\{d\ge 5\}} + \frac{x^6}{n}\right),$$

which is equivalent to the desired inequality. This completes the proof of Theorem 2. □

Remark 3. *The moderate deviation results obtained in this paper can be applied to the statistical inference of the offspring mean vector. When the total population is sufficiently large, Theorem 2 provides a foundation for constructing asymptotic confidence regions for the unknown mean vectors. In this paper, we assume that the offspring vectors are independent and identically distributed. However, in future research, we can further consider general multi-type branching processes in which the distribution of offspring depends on the type of the parent, or study self-normalized versions of such processes.*

Author contributions

Juan Wang: Conceptualization, Supervision, Writing–review and editing; Luyang Zhai: Formal analysis, Writing–original draft. All authors have read and approved the final version of the manuscript.

Use of Generative-AI tools declaration

The authors declare that they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This research was funded by the National Natural Sciences Foundation of China(NO.11901392).

Conflict of interest

The authors declare no potential conflicts of interest.

References

1. X. Xia, The branching process: A general conceptual framework for addressing current ecological and evolutionary questions, *Life*, **15** (2025), 1910. https://doi.org/10.3390/life15121910
2. R. Bürger, Bounds for survival probabilities in supercritical Galton-Watson processes and applications to population genetics, *J. Math. Biol.*, **92** (2026), 40. https://doi.org/10.1007/s00285-026-02349-7
3. P. Guttorp, *Statistical Inference for Branching Processes*, New York: Wiley, 1991.
4. H. Cramér, Sur un nouveau théorème-limite de la théorie des probabilités, *Actualités Scientifiques et Industrielles*, **736** (1938), 5–23.

5. B. von Bahr, Multi-dimensional integral limit theorems for large deviations, *Arkiv för Matematik*, **7** (1967), 89–99.
6. V. V. Petrov, *Sums of Independent Random Variables*, Berlin: Springer-Verlag, 1975.
7. X. Fan, I. Grama, Q. Liu, Q. M. Shao, Self-normalized Cramér type moderate deviations for martingales, *Bernoulli*, **25** (2019), 2793–2823. https://doi.org/10.3150/18-BEJ1071
8. X. Fan, Q. M. Shao, Self-normalized Cramér type moderate deviations for martingales and applications, *Bernoulli*, **31** (2025), 130–161. https://doi.org/10.3150/24-BEJ1722
9. Y. Miao, R. Yang, Cramér moderate deviation and sharp large deviation for bootstrap sample quantiles, *Commun. Stat.–Theory Methods*, **55** (2026), 3270–3281. https://doi.org/10.1080/03610926.2025.2574489
10. K. B. Athreya, P. E. Ney, *Branching Processes*, Berlin: Springer-Verlag, 1972.
11. P. Doukhan, X. Fan, Z. Q. Gao, Cramér moderate deviations for a supercritical galton–watson process, *Stat. Probab. Lett.*, **192** (2022), 109711. https://doi.org/10.1016/j.spl.2022.109711
12. J. Wang, C. Peng, Cramér moderate deviations for a supercritical galton–watson process with immigration, *Axioms*, **13** (2024), 272. https://doi.org/10.3390/axioms13040272
13. X. Fang, S. H. Liu, Q. M. Shao, Cramér-type moderate deviation for quadratic forms with a fast rate, *Bernoulli*, **29** (2023), 2466–2491. https://doi.org/10.3150/22-BEJ1549



AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0)