

## ON THE INTERMEDIATE ASYMPTOTIC EFFICIENCY OF GOODNESS-OF-FIT TESTS IN MULTINOMIAL DISTRIBUTIONS

SHERZOD M. MIRAKHMEDOV\*

**Abstract.** We consider goodness-of-fit tests for uniformity of a multinomial distribution by means of tests based on a class of symmetric statistics, defined as the sum of some function of cell-frequencies. We are dealing with an asymptotic regime, where the number of cells grows with the sample size. Most attention is focused on the class of power divergence statistics. The aim of this article is to study the intermediate asymptotic relative efficiency of two tests, where the powers of the tests are asymptotically non-degenerate and the sequences of alternatives converge to the hypothesis, but not too fast. The intermediate asymptotic relative efficiency of the  $\chi^2$  test wrt an arbitrary symmetric test is considered in details.

**Mathematics Subject Classification.** 62G10, 62G20.

Received May 20, 2022. Accepted September 1, 2022.

### 1. INTRODUCTION

Let  $(\eta_1, \dots, \eta_N)$  be a random vector of frequencies of a multinomial model on  $n \geq 1$  observations classified into  $N > 1$  cells with the cell-probabilities  $\mathbf{P} = (p_1, \dots, p_N)$ ,  $p_1 + \dots + p_N = 1$ , all  $p_j > 0$ . We assume that  $N = N(n) \rightarrow \infty$  as  $n \rightarrow \infty$ , so that  $n/N \rightarrow \lambda \in [0, \infty]$  and  $n^2/N \rightarrow \infty$ . We are interested in the problem of testing of  $H_0: \mathbf{P} = (N^{-1}, \dots, N^{-1})$  versus sequences of alternatives  $H_{1n}: \mathbf{P} = (p_1, \dots, p_N) \neq (N^{-1}, \dots, N^{-1})$ , which approach  $H_0$  as  $n \rightarrow \infty$  so that

$$\varepsilon(n) = \frac{1}{N} \sum_{m=1}^N \varepsilon_{m,n}^2 \rightarrow 0, \quad (1.1)$$

where  $\varepsilon_{m,n} = Np_m - 1$ , i.e.,  $p_m = N^{-1}(1 + \varepsilon_{m,n})$ ,  $m = 1, \dots, N$  and  $\max_m |\varepsilon_{m,n}| \rightarrow 0$  as  $n \rightarrow \infty$ . We consider a class of symmetric tests based on statistics of the form

$$S_n^h = \sum_{l=1}^N h(\eta_l), \quad (1.2)$$

---

*Keywords and phrases:* Asymptotic efficiency,  $\chi^2$  statistic, log-likelihood ratio statistic, goodness-of-fit tests, multinomial distribution, power divergence statistics.

V.I. Romanovskiy Institute of Mathematics, Academy of Sciences of Uzbekistan University str., 9, Tashkent 100174, Uzbekistan.

\* Corresponding author: [shmirakhmedov@yahoo.com](mailto:shmirakhmedov@yahoo.com)

where  $h$  is a nonlinear real-valued function, defined on the non-negative axis. A test based on the statistic  $S_n^h$ , whose large values reject  $H_0$ , is called  $h$ -test for short.

The *power divergence statistics* (PDS)  $CR_n(d)$  of [1] is an important subclass of statistics (1.2), where  $h(x) = \psi_d(x)$ ,  $d > -1$ ,

$$\psi_d(x) = 2(d(d+1))^{-1}x[(x/\lambda_n)^d - 1], d \neq 0, \text{ else } \psi_0(x) = 2x \log(x/\lambda_n), \quad (1.3)$$

$\lambda_n = n/N$  is the average of observations arisen in the cells. (One could consider  $d \rightarrow -1$  and  $d < -1$ , but this unnecessary in the context of this paper). The most commonly used special versions of PDS are the  $\chi^2$  statistic  $\chi_n^2 = CR_n(1)$ , the log-likelihood ratio statistic  $\Lambda_n = CR_n(0)$ , and the Freeman-Tukey statistic  $T_n^2 = CR_n(-1/2)$ , viz.,

$$\chi_n^2 = \lambda_n^{-1} \sum_{m=1}^N (\eta_m - \lambda_n)^2, \quad \Lambda_n = 2 \sum_{m=1}^N \eta_m \log(\eta_m/\lambda_n), \quad T_n^2 = 4 \sum_{m=1}^N (\sqrt{\eta_m} - \sqrt{\lambda_n})^2, \quad (1.4)$$

Count statistics (CS) form another sub-class of (1.2). Two popular variants of CS are  $\mu_r$ -the number of cells containing exactly  $r$  observations and  $C_n$ -the number of collisions (*i.e.*, the number of those observations that fall in a cell that already has an observation in it), viz.,

$$\mu_r = \sum_{m=1}^N I\{\eta_m = r\}, r \geq 0, \text{ and } C_n = \sum_{m=1}^N (\eta_m - 1)I\{\eta_m > 1\}, \quad (1.5)$$

where  $I\{\cdot\}$  denotes the indicator function.

Our goal is to study the intermediate asymptotic relative efficiencies (IARE) of two  $h$ -tests, provided that their powers are asymptotically non-degenerate and the alternatives (1.1) are such that

$$\nabla_n \rightarrow \infty \text{ and } \nabla_n = o(\sqrt{N}), \quad (1.6)$$

where  $\nabla_n = n\varepsilon(n)/\sqrt{N}$ .

These alternatives are somewhere between the alternatives which do not approach the hypothesis (or the rate of approach is too slow so that  $\varepsilon(n) \geq \lambda_n^{-1}$ ) and the alternatives which approaches the hypothesis at the rate so that  $\nabla_n = O(1)$ , *i.e.*, a rate that ensures that the power of  $h$ -tests of size  $\alpha > 0$  has a limit in  $(\alpha, 1)$ . In other words, we are interested in comparison of two  $h$ -tests in the situation, which are intermediate between Bahadur's and Pitman's concepts of asymptotic efficiency (AE), see [15], and hence the name. For the detailed comments on the alternatives (1.6) and brief survey we refer to [14]. In that work the comparison of  $h$ -tests was considered in terms of the so called " $\alpha$  - intermediate AE", according to which the performance of an  $h$ -test of size  $\alpha$  measured by the asymptotic value of  $e_N^\alpha(S_n^h) = -\log P_0\{S_n^h > NA_{1n}(h)\}$ , provided that its power is asymptotically non-degenerate. Here and throughout the article  $A_{in}(h)$  is an asymptotic value of  $N^{-1}E_i S_n^h$  (see Sect. 2),  $P_i$  and  $E_i S_n^h$ , as usually, denote the probability and expectation, counted under  $H_0$  and  $H_{1n}$ , respectively. Among others it is shown that the intermediate asymptotic properties of  $h$ -tests depend on the parameter  $\lambda_n = n/N$ , whose asymptotic behavior (as  $n \rightarrow \infty$ ) provides the classification of the multinomial model: the *very sparse model* when  $\lambda_n \rightarrow 0$ ,  $n\lambda_n \rightarrow \infty$ , the *sparse model* when  $\lambda_n \rightarrow \lambda \in (0, \infty)$ , and the *dense model* when  $\lambda_n \rightarrow \infty$ .

In the present paper we study the IARE of two  $h$ -tests, defined as a limit of the ratio of sample sizes which guarantee the same precision for both tests: the same significance level tending to 0 slower than exponentially and the same asymptotically non-degenerate power. We remind that the concept of IARE was introduced by [9] and then elaborated for testing problems by [4]. Further development and application of this IARE in testing problems has been presented in several works, see [2] and references within. We refer to the paper by [5], where

so called path wise intermediate efficiency elaborated in [6] is demonstrated by applying to study the IARE of Neyman–Pearson test wrt Kolmogorov-Smirnov test in the classical problem of testing for uniformity versus a class of heavy-tailed alternatives. We use the idea of [5] for discrete models, namely in the problem of testing for the uniformity of a multinomial distribution.

The rest of the paper is organized as follows. Section 2 provides a general result on IARE of two  $h$ -tests and its application for IARE of the  $\chi^2$  test wrt  $h$ -tests. Proofs of all statements and Conclusions are given in Section 3.

## 2. MAIN RESULTS

In what follows  $\xi \sim Poi(\lambda)$  stands for “a random variable (r.v.)  $\xi$  has Poisson distribution with parameter  $\lambda > 0$ ”;  $\Phi(u)$  denotes a standard normal distribution function;  $c_j$  is a positive constant, may not the same in each its occurrence,  $\xi, \xi_1, \dots, \xi_N$  are independent r.v.s such that  $\xi \sim Poi(\lambda_n)$ ,  $\lambda_n = n/N$ , and  $\xi_m \sim Poi(np_m)$ . *Without mention, all asymptotic statements are considered as  $n \rightarrow \infty$ , hence  $N \rightarrow \infty$  as well.* Now on  $\mathcal{I}_{alt}$  stands for the family of alternatives (1.1) such that

$$\frac{n\varepsilon(n)}{\sqrt{N}} \rightarrow \infty, \quad \frac{n}{N} \max_m \varepsilon_{m,n}^2 \rightarrow 0. \tag{2.1}$$

This is a slightly narrower family of alternatives than (1.6), but we have to deal with it because of Proposition 2.1 below. Set

$$g(\xi) = h(\xi) - Eh(\xi) - r_n(\xi - \lambda_n), \quad r_n = \lambda_n^{-1} cov(h(\xi), \xi)$$

$$\sigma^2(h) = Var g(\xi) = Var h(\xi) (1 - corr^2(h(\xi), \xi)),$$

$$\rho(h, \lambda_n) = corr(h(\xi) - r_n \xi, \xi^2 - (2\lambda_n + 1)\xi).$$

**Proposition 2.1.** *Let the alternatives of  $H_{1n}$  be such that  $n \max_m \varepsilon_{m,n}^2 / N \rightarrow 0$ . One has*

(i) *If  $E|h(\xi)\xi^3| < \infty$  then*

$$\kappa_n(h) := \sqrt{N}(A_{1,n}(h) - A_{0,n}(h))/\sigma(h) = \sqrt{n\lambda_n/2}\varepsilon(n)\rho(h, \lambda_n)(1 + o(1)). \tag{2.2}$$

(ii) *If  $Eh^2(\xi)\xi^2 < \infty$  then*

$$\sigma_{1,n}^2(h) = \sigma^2(h)(1 + o(1)), \tag{2.3}$$

where

$$A_{i,n}(h) = N^{-1} \sum_{m=1}^N E_i h(\xi_m), \quad \sigma_{i,n}^2(h) = N^{-1} \sum_{m=1}^N Var_i g(\xi_m).$$

Note that under very general set-up (see, for instance, Prop. 2.1 of [12])

$$E_i S_n^h = N A_{i,n}(h)(1 + o(1)) \text{ and } Var_i S_n^h = N \sigma_{i,n}^2(h)(1 + o(1)).$$

The following proposition is the consequence of Theorem 1 by [11].

**Proposition 2.2.** *Assume*

$$E |g(\xi)|^3 / \sigma^3(h) \sqrt{N} \rightarrow 0. \tag{2.4}$$

Then

$$P_i \left\{ S_n^h < u \sigma_{i,n}(h) \sqrt{N} + N A_{i,n}(h) \right\} = \Phi(u) + o(1), i = 0, 1,$$

where  $\sigma_{0,n}^2(h) = \sigma^2(h)$ .

We emphasize that the condition (2.4) is fulfilled for the sparse models if  $E |h(\xi)|^3 < \infty$ , and for the very sparse models if  $\Delta^2 h(0) \neq 0$ , where  $\Delta h(x) = h(x + 1) - h(x)$ . For instance, arbitrary PDS and CS  $C_n, \mu_r, r = 0, 1, 2$  satisfy this condition. But for the dense models condition (2.4) may impose an additional condition to  $\lambda_n$ . For instance, (2.4) is fulfilled for PDS without additional restriction, while, for example, for CS  $\mu_r, r \geq 0$  and  $C_n$  it imposes condition  $\lambda_n - \ln N - r \ln \ln N \rightarrow -\infty$  and  $\lambda_n - \ln N \rightarrow -\infty$ , respectively.

Notice that  $|\rho(h, \lambda_n)| \leq 1$  and  $|\rho(h, \lambda_n)| = 1$  iff  $h(x) = x^2$ , i.e., for the  $\chi^2$  statistic. But for the CS (1.5)  $\rho(h, \lambda_n) = o(1)$  if  $\lambda_n \rightarrow \infty$ . Meaning of the functional  $\rho(h, \lambda_n)$  is clarified by the following

**Lemma 2.3.** *If  $E h^2(\xi) \xi < \infty$  then  $\rho(h, \lambda_n) = \text{corr}_0(S_n^h, \chi_n^2) + o(1)$ .*

It follows from Proposition 2.2 that the power of  $h$ -test of size  $\alpha > 0$  asymptotically equal to

$$\Phi \left( 2^{-1/2} \nabla_n |\rho(h, \lambda_n)| - \omega_\alpha \right),$$

where  $\omega_\alpha = \Phi^{-1}(1 - \alpha)$ ,  $\nabla_n = \sqrt{n \lambda_n} \varepsilon(n)$  (see (1.6)). Hence for the alternatives (1.1) such that  $\nabla_n = O(1)$  (i.e., for Pitman alternatives) the  $\chi^2$  test is asymptotically most powerful (AMP) within the class of  $h$ -tests. The  $\chi^2$  test is the only AMP for sparse models. But for very sparse and dense models the  $\chi^2$  test is no longer unique AMP, since by virtue of Lemma 2.4 and Lemma 2.5 below, there are test statistics for which  $\rho(h, \lambda_n) \rightarrow 1$  if  $\lambda_n \rightarrow 0$  or  $\lambda_n \rightarrow \infty$ .

**Lemma 2.4.** *If  $\lambda_n \rightarrow 0$  and  $\Delta^2 h(0) \neq 0$  then*

$$\rho(h, \lambda_n) = 1 - \frac{\lambda_n}{6} \left( \frac{\Delta^3 h(0)}{\Delta^2 h(0)} \right)^2 + O(\lambda_n^2).$$

In particular

$$\rho(\psi_d, \lambda_n) = 1 - \frac{3(3^d - 2^{d+1} + 1)^2}{8(2^d - 1)^2} \lambda_n + O(\lambda_n^2), d \neq 0,$$

$$\rho(\psi_0, \lambda_n) = 1 - \frac{3}{8} \left( \frac{\ln 3/4}{\ln 2} \right)^2 \lambda_n + O(\lambda_n^2),$$

$$\rho(h, \lambda_n) = 1 - c(r) \lambda_n + O(\lambda_n^2), \text{ if } h(x) = I(x = r), r = 0, 1, 2,$$

where  $c(0) = 1/6, c(1) = 3/8$  and  $c(2) = 3/2$ .

**Lemma 2.5.** *If  $\lambda_n \rightarrow \infty$ , then for every PDS with parameter  $d > -1$  it holds*

$$\rho(\psi_d, \lambda_n) = 1 - \frac{(d - 1)^2}{6\lambda_n} + O(\lambda_n^{-2}).$$

The functional  $\rho(h, \lambda_n)$  plays a key role also in determining the intermediate properties of the  $h$ -tests. From the aforesaid and [14] it follows that an  $h$ -test for which  $\rho(h, \lambda_n)$  close to 1 is preferable. Throughout the rest of the paper the symbol  $\mathbf{S}$  denotes the class of  $h$ -tests, for which  $\lim |\rho(h, \lambda_n)| > 0$  and conditions of Proposition 2.1 are fulfilled.

**2.1. IARE of two h-tests**

Set  $\tilde{S}_n^h = (S_n^h - NA_{0,n}(h)) / \sigma_{0,n}(h)\sqrt{N}$ ,  $\tilde{S}_n^\psi = (S_n^\psi - NA_{0,n}(\psi)) / \sigma_{0,n}(\psi)\sqrt{N}$ ,  $\kappa_n(h)$  and  $\kappa_n(\psi)$  are defined in accord to (2.2). For each  $n$  and any fixed  $z \in \mathbb{R}$  define the significance level of the  $h$ -test corresponding to the critical value  $z + \kappa_n(h)$ :

$$\alpha_n = \alpha_n(z) = P_0 \left\{ \tilde{S}_n^h > z + \kappa_n(h) \right\}. \tag{2.5}$$

Due to Propositions 2.2 statistic  $\tilde{S}_n^h$  is bounded in probability under  $H_0$ , and by (2.2)  $\kappa_n(h) \rightarrow \infty$  for every alternatives of  $\mathcal{I}_{ait}$ . Hence  $\alpha_n \rightarrow 0$ , this convergence depends on the rate of convergence of  $\varepsilon(n)$  to zero, that is it depends on the sub-family of considering alternatives, say  $\mathcal{I} \subseteq \mathcal{I}_{ait}$ .

We define IARE of  $h$ -test wrt  $\psi$ -test in accord to idea of [5]. For each  $n$  and every  $k \geq 1$  let  $u_k = u_k(n)$  be exact critical value of the  $\psi$ -test at the level  $\alpha_n$  (2.2) and sample size  $k$ , i.e.,

$$P_0 \left\{ \tilde{S}_k^\psi > u_k \right\} = \alpha_n. \tag{2.6}$$

Next, for each  $n$  let  $k_n$  be the minimal sample size such that for all integer  $j \geq 0$

$$P_1 \left\{ \tilde{S}_{k_n+j}^\psi > u_{k_n+j} \right\} \geq P_1 \left\{ \tilde{S}_n^h > z + \kappa_n(h) \right\}, \tag{2.7}$$

i.e.  $k_n$  is the minimal sample size beginning from which the power of the  $\psi$ -test at the level  $\alpha_n$  under alternatives of  $\mathcal{I} \subseteq \mathcal{I}_{ait}$  is not smaller than that for the  $h$ -test at the same level and for the sample size  $n$ .

**Definition 2.6.** If there exists the limit

$$\lim_{n \rightarrow \infty} \frac{k_n}{n} = e \in [0, \infty],$$

then  $e$  is called the IARE of  $h$ -test wrt  $\psi$ -test, for every alternatives of  $\mathcal{I}$ .

**Theorem 2.7.** *Assume  $h, \psi \in \mathbf{S}$ , the family of alternatives  $\mathcal{I}_{ait}$  is defined as in (2.1), and*

- (a)  $N(\bullet)$  is taken as a positive function of the continuous variable  $x$ , and it has a power tail of index  $q \in (0, 2)$ , viz.,  $N(x) \sim cx^q, c > 0$ , as  $x \rightarrow \infty$ ,
- (b)  $S_n^h$  satisfies condition (2.4),
- (c)  $\mathcal{I}_1$ , a sub-family of  $\mathcal{I}_{ait}$ , and a sequence of positive numbers  $(\tau_n)$ ,  $\tau_n \rightarrow 0$  or  $\tau_n = c \leq 1/2$  for all  $n$ , such that for every sequence of positive numbers  $(y_n)$ ,  $y_n = O(\kappa_n(h))$ , it holds

$$-\lim_{n \rightarrow \infty} \frac{1}{\tau_n y_n^2} \log P_0 \left\{ \tilde{S}_n^h > y_n \right\} = 1.$$

(d) There exists a sequence of positive numbers  $(a_n)$ ,  $a_n \rightarrow \infty$ , such that

$$-\lim_{n \rightarrow \infty} \frac{2}{y_n^2} \log P_0 \left\{ \tilde{S}_n^\psi > y_n \right\} = 1,$$

for every sequence of positive numbers  $(y_n)$  such that  $y_n \rightarrow \infty$  and  $y_n = o(a_n)$ .

(f)  $\mathcal{I}_2$  is a sub-family of  $\mathcal{I}_{all}$  such that  $\varepsilon(n) = o(a_n \sqrt{N}/n)$ .

Then for every alternatives of  $\mathcal{I} = \mathcal{I}_1 \cap \mathcal{I}_2$  one has

$$\lim_{n \rightarrow \infty} \frac{k_n}{n} = \lim_{n \rightarrow \infty} \left( 2\tau_n \frac{\rho^2(h, \lambda_n)}{\rho^2(\psi, \lambda_{k_n})} \right)^{1/(2-q)} < \infty. \tag{2.8}$$

**Remark 2.8.** Conditions (c) and (d) should be considered as conditions for the region, where it is supposed that under the hypothesis there are corresponding results on the probabilities of large deviations of the test statistics. In fact these regions define the family of alternatives  $\mathcal{I}_1$  and  $\mathcal{I}_2$ , respectively.

### 2.2. IARE of the $\chi^2$ test wrt $h$ -tests

Now on whenever it is convenient we shall use notation  $d_n \ll b_n$  instead of well-known notation  $d_n = o(b_n)$ , also  $\varepsilon(n) \ll d_n$  stands for the “sub-family of  $\mathcal{I}_{alt}$  is specified by the condition  $\varepsilon(n) \ll d_n$ ”.

We wish to evaluate the efficiency of the  $\chi^2$  test by comparing its sensitivity relative to that of the sensitivity of  $\psi$ -test in terms of IARE, hence the  $\chi^2$  test becomes as  $h$ -test in Theorem 2.1. Let  $e(\chi_n^2, S_n^\psi)$  denote the IARE of the  $\chi^2$  test wrt to the  $\psi$ -test,  $\psi_d$  be defined as in (1.3). We consider sparse, very sparse and dense models separately. Recall that we assume  $N \sim cn^q$ , as  $n \rightarrow \infty$ , where index  $q \in (0, 2)$ .

*Sparse models*, that is  $\lambda_n \rightarrow \lambda \in (0, \infty)$  and hence index  $q = 1$ .

**Theorem 2.9.** Let  $\lambda_n \rightarrow \lambda > 0$ .

(i) If  $\psi = \psi_d$ ,  $d > -1$ ,  $d \neq 1$  and  $\varepsilon(n) \ll n^{-d^*/(1+2d^*)}$ , where  $d^* = \max(1, d)$ , or  $\psi$  satisfies the Cramer condition  $E \exp\{a |\psi(\xi)|\} < \infty$ , for some  $a > 0$ , and  $\varepsilon(n) \ll n^{-1/3}$ , then

$$e(\chi_n^2, S_n^\psi) > 1.$$

(ii) If  $\psi$  satisfies the Cramer condition and  $\varepsilon(n) \gg n^{-1/3} \log^{2/3} n$ , Then

$$e(\chi_n^2, S_n^\psi) = 0.$$

**Remark 2.10.** We emphasize that the Cramer condition is fulfilled for CS (1.5), as well as for PDS  $CR_n(d)$ , i.e., for  $\psi_d$ , if  $d \in (-1, 0]$ . For example, for the statistics  $\Lambda_n$  and  $T_n^2$ , see (1.4), the Cramer condition is fulfilled, but for  $\chi_n^2$  it is not.

From Theorem 2.9 it follows

**Corollary 2.11.** Let  $\lambda_n \rightarrow \lambda > 0$ . Then for alternatives such that  $\varepsilon(n) \ll n^{-1/3}$  it holds  $e(\chi_n^2, \Lambda_n) > 1$ , whereas for alternatives such that  $\varepsilon(n) \gg n^{-1/3} \log^{2/3} n$  it holds  $e(\chi_n^2, \Lambda_n) = 0$ .

Similar statement is valid for Freeman-Tukey statistic and each of CS (1.5).

*Very sparse models*, that is we assume that  $\lambda_n \rightarrow 0$ , but  $n\lambda_n \rightarrow \infty$ , hence index  $q \in (1, 2)$ .

**Theorem 2.12.** Let  $\lambda_n \rightarrow 0$ ,  $n\lambda_n^3 \rightarrow \infty$ , and hence  $q \in (1, 4/3)$ .

(A) If  $d \in (-1, 1/2)$  and  $\lambda_n \gg n^{-(1-2d^*)/6}$ , where now  $d^* = \max(0, d)$ , then

(i) for alternatives such that  $\varepsilon(n) \ll n^{-1/3}$  it holds

$$e(\chi_n^2, S_n^{\psi_d}) = 1,$$

(ii) for alternatives such that  $(n\lambda_n)^{-1/3} \log^{2/3}(N^2/n) \ll \varepsilon(n) \ll (n\lambda_n^2)^{-1/4}$  it holds

$$e(\chi_n^2, S_n^{\psi_d}) = 0.$$

(B) If  $d \geq 1/2$  and alternatives such that  $\varepsilon(n) \ll (\lambda_n^{1-d}/n^d)^{1/(2d+1)}$  then

$$e(\chi_n^2, S_n^{\psi_d}) = 1.$$

**Corollary 2.13.** Let  $\lambda_n \rightarrow 0$ ,  $n\lambda_n^3 \rightarrow \infty$  and  $\lambda_n \gg n^{-1/6}$ . Then for alternatives such that  $\varepsilon(n) \ll n^{-1/3}$  it holds  $e(\chi_n^2, \Lambda_n) = 1$ , whereas for alternatives such that  $(n\lambda_n)^{-1/3} \log^{2/3}(N^2/n) \ll \varepsilon(n) \ll (n\lambda_n^2)^{-1/4}$  it holds  $e(\chi_n^2, \Lambda_n) = 0$ .

Actually, from Theorem 2.12 it follows alike to Corollary 2.13 result for every  $\psi_d$ -test, where  $d \in (-1, 0]$ . We emphasize that  $n\lambda_n^3 \rightarrow \infty$  (i.e.,  $n \ll N \ll n^{4/3}$ ) is a necessary condition for the family of alternatives with  $\varepsilon(n) \ll (\lambda_n^{1-d}/n^d)^{1/(2d+1)}$  to be a non-empty sub-family of  $\mathcal{I}_{alt}$ .

**Theorem 2.14.** If  $\lambda_n \rightarrow 0$  and  $n\lambda_n \rightarrow \infty$ . Then for alternatives such that  $\varepsilon(n) \ll n^{-1/3}$  one has

$$e(\chi_n^2, \mu_r) = 1, r = 0, 1, 2, \text{ and } e(\chi_n^2, C_n) = 1.$$

Dense models assume that  $\lambda_n \rightarrow \infty$ , hence index  $q \in (0, 1)$ .

**Theorem 2.15.** Let  $\lambda_n \rightarrow \infty$ . (i) For every  $d > -1$  and alternatives such that  $\varepsilon(n) \ll (n\lambda_n^2)^{-1/3}$  it holds

$$e(\chi_n^2, S_n^{\psi_d}) = 1.$$

(ii) If  $N = o(n^{3/8})$  then for every alternative of  $\mathcal{I}_{alt}$  it holds

$$e(\chi_n^2, \Lambda_n) = 1.$$

The PDS can be rewritten as

$$CR_N(d) = \lambda_n \sum_{m=1}^N \bar{\psi}_d(\eta_m), \tag{2.9}$$

where

$$\begin{aligned} \bar{\psi}_d(x) &= \frac{2}{d(d+1)} \left[ (x/\lambda_n)^{d+1} - (d+1)(x/\lambda_n) + d \right], d \neq 0, \\ \bar{\psi}_0(x) &= 2[(x/\lambda_n) \log(x/\lambda_n) - (x - \lambda_n)/\lambda_n]. \end{aligned}$$

Further, under  $H_0$  the r.v.  $(\eta_m - \lambda_n)/\sqrt{\lambda_n}$  has asymptotically normal distribution, as  $\lambda_n \rightarrow \infty$ , since  $\eta_m \sim Bi(n, N^{-1})$ . Hence  $(\eta_m - \lambda_n)/\lambda_n = O_p(\lambda_n^{-1/2})$ . Note that  $\eta_m/\lambda_n = 1 + (\eta_m - \lambda_n)/\lambda_n$ . Use these facts and Taylor expansion formula to get  $\lambda_n \bar{\psi}_d(\eta_m) = (\eta_m - \lambda_n)^2/\lambda_n + O_p(\lambda_n^{-1/2})$ ,  $d > -1$ . Thus, if  $\lambda_n \rightarrow \infty$  then under  $H_0$  each PDS asymptotically in distribution coincides with the  $\chi^2$  statistic. This property is reflected to some extent in Theorem 2.15.

## 3. PROOFS

We first present the proof of Proposition 2.1 and the theorems, after which we will give the proof of Lemma 2.3, 2.4 and 2.5,

*Proof of Proposition 2.1.* Recall  $p_m = N^{-1}(1 + \varepsilon_{m,n}), m = 1, 2, \dots, N$ , where  $\varepsilon_{1,n} + \dots + \varepsilon_{N,n} = 0$ ,  $\max_m |\varepsilon_{m,n}| \rightarrow 0$ . Set  $\xi \sim Poi(\lambda_n)$ ,  $\varepsilon_j(n) = N^{-1}(\varepsilon_{1,n}^j + \dots + \varepsilon_{N,n}^j)$ ,  $\varepsilon_2(n) = \varepsilon(n)$ ,

$$\varphi_2(\xi) = \xi(\xi - 1) - 2\lambda_n\xi + \lambda_n^2 = (\xi - \lambda_n)^2 - (\xi - \lambda_n) - \lambda_n,$$

$$\varphi_3(\xi) = \xi(\xi - 1)(\xi - 2) - 3\lambda_n\xi(\xi - 1) + 3\lambda_n^2\xi - \lambda_n^3$$

$$= (\xi - \lambda_n)^3 - 3(\xi - \lambda_n)^2 + (2 - 3\lambda_n)(\xi - \lambda_n) + 2\lambda_n.$$

Using Taylor expansion of  $(1 + \varepsilon_{m,n})^k \exp\{-\lambda_n\varepsilon_{m,n}\}$  we obtain

$$A_{1n}(h) = Eh(\xi) + \frac{1}{2}\varepsilon_2(n)E(h(\xi)\varphi_2(\xi)) + \frac{1}{6}\varepsilon_3(n)(E(h(\xi)\varphi_3(\xi)) + o(1)). \quad (3.1)$$

We have

$$E\varphi_2(\xi) = 0, E\xi\varphi_2(\xi) = 0, Var\varphi_2(\xi) = 2\lambda_n^2 \quad (3.2)$$

$$E\varphi_3(\xi) = 0, E\xi\varphi_3(\xi) = 0, Var\varphi_3(\xi) = 6\lambda_n^3.$$

Hence  $E(h(\xi)\varphi_i(\xi)) = E(g(\xi)\varphi_i(\xi))$ . Due to these facts we obtain

$$\begin{aligned} A_{1n}(h) - A_{0n}(h) &= \frac{\lambda_n\sigma(h)}{\sqrt{2}}\varepsilon_2(n)corr(g(\xi), \varphi_2(\xi)) \\ &+ \frac{\lambda_n^{3/2}\sigma(h)}{\sqrt{6}}\varepsilon_3(n)corr(g(\xi), \varphi_3(\xi))(1 + o(1)). \end{aligned}$$

Therefore

$$\kappa_n(h) = \frac{\sqrt{n\lambda_n}}{\sqrt{2}}\varepsilon_2(n)corr(g(\xi), \varphi_2(\xi)) + \frac{\sqrt{n}}{\sqrt{6}}\lambda_n\varepsilon_3(n)corr(g(\xi), \varphi_3(\xi))(1 + o(1)).$$

Equality (2.2) follows, since  $corr(g(\xi), \varphi_2(\xi)) = \rho(h, \lambda_n)$  and  $\varepsilon_3(n) = o(\lambda_n^{-1/2})\varepsilon_2(n)$  by (2.1).

Next, by Taylor expansion

$$\sigma_{1n}^2(h) = Eg^2(\xi) + \frac{1}{2}\varepsilon_2(n)E(g^2(\xi)\varphi_2(\xi)) (1 + o(1)). \quad (3.3)$$

If  $\lambda_n \rightarrow 0$  then  $E(g^2(\xi)\varphi_2(\xi)) = \lambda_n^2(g^2(2) - 2g^2(1) + \sigma^2(h))(1 + o(1))$ . If  $\lambda_n \rightarrow \lambda \in (0, \infty)$  then evidently  $E(g^2(\xi)\varphi_2(\xi)) \leq c(\lambda)$ . By virtue of these facts and (3.3), equality (2.3) follows for very sparse and sparse

models. Let  $\lambda_n \rightarrow \infty$ , then (3.3) gives

$$\sigma_{1n}^2(h) = \sigma^2(h) \left( 1 + \frac{\lambda_n \varepsilon_2(n)}{2} E \left( \frac{g(\xi)}{\sigma(h)} \frac{\xi - \lambda_n}{\sqrt{\lambda_n}} \right)^2 \right) = \sigma^2(h) (1 + o(1)),$$

since  $\lambda_n \varepsilon_2(n) = o(1)$ ,  $(\xi - \lambda_n)/\sqrt{\lambda_n} = O_p(1)$  and  $Eg^2(\xi) = \sigma^2(h)$ . Proof of Proposition 2.1 is completed.  $\square$

*Proof of Theorem 2.7.* By Proposition 2.2 we have

$$P_1 \left\{ \tilde{S}_n^h > z + \kappa_n(h) \right\} = P_1 \left\{ \left( \tilde{S}_{n,N}^h - NA_{1,n}(h) \right) / \sqrt{N} \sigma_{1,n}(h) > z \sigma_{1,n}^{-1}(h) \sigma_{0,n}(h) \right\} = \Phi(-z) + o(1).$$

Hence, for arbitrary fixed  $z$  and sequences of powers of  $h$ -test at the significance level defined by (2.5) we have

$$0 < c_1 \leq P_1 \left\{ \tilde{S}_n^h > z + \kappa_n(h) \right\} \leq c_2 < 1. \tag{3.4}$$

By virtue of this, definition of  $k_n$ , (2.3) and Proposition 2.2 we obtain

$$0 < c_1 \leq P_1 \left\{ \tilde{S}_{k_n}^\psi > u_{k_n} \right\} = P_1 \left\{ \frac{S_{k_n}^\psi - NA_{1,k_n}(\psi)}{\sqrt{N} \sigma_{1,k_n}(\psi)} > (u_{k_n} - \kappa_{k_n}(\psi)) (1 + o(1)) \right\}. \tag{3.5}$$

Hence,

$$u_{k_n} - \kappa_{k_n}(\psi) \leq c_3. \tag{3.6}$$

For every alternative of  $\mathcal{I}_{alt}$  we have  $z + \kappa_n(h) \sim \kappa_n(h) \sim \sqrt{n\lambda_n/2} \varepsilon(n) \rho(h, \lambda_n) \rightarrow \infty$ , and under alternatives  $\mathcal{I}_1$  by condition (c)

$$-\log \alpha_n = -\log P_0 \left\{ \tilde{S}_n^h > z + \kappa_n(h) \right\} \sim \frac{1}{2} \tau_n n \lambda_n \varepsilon^2(n) \rho^2(h, \lambda_n). \tag{3.7}$$

On the other hand  $u_{k_n} \rightarrow \infty$ , since (2.6) and  $\alpha_n \rightarrow 0$ , also  $u_{k_n} = O(\kappa_{k_n}(\psi)) = o(a_{k_n})$  under alternatives  $\mathcal{I}_2$ , by (3.6) and condition (d). Hence due to condition (d), choice of  $u_{k_n}$ , see (2.7), and (3.7) we obtain

$$-\log P_0 \left\{ \tilde{S}_{k_n}^\psi > u_{k_n} \right\} = \frac{1}{2} u_{k_n}^2 (1 + o(1)) = -\log \alpha_n = \frac{1}{2} \tau_n n \lambda_n \rho^2(h, \lambda_n) \varepsilon^2(n) (1 + o(1)),$$

i.e.  $u_{k_n}^2 = \tau_n n \lambda_n \rho^2(h, \lambda_n) \varepsilon^2(n) (1 + o(1))$ . Thus, for every sequence of alternatives of  $\mathcal{I}$  we have

$$\tau_n n \lambda_n \rho^2(h, \lambda_n) \varepsilon^2(n) \leq 2^{-1} k_n \lambda_{k_n} \rho^2(\psi, \lambda_{k_n}) \varepsilon^2(k_n), \tag{3.8}$$

since (3.6) and definition of  $\kappa_{k_n}(\psi)$ . Note that it have to be  $\varepsilon^2(n) \sim \varepsilon^2(k_n)$  for the alternatives of  $\mathcal{I}$ . Thus from (3.8) and assumption (a) we have: under alternatives  $\mathcal{I}$  for large enough  $n$

$$\frac{k_n}{n} \geq \left( \frac{2\tau_n \rho^2(h, \lambda_n)}{\rho^2(\psi, \lambda_{k_n})} \right)^{1/(2-q)}, \tag{3.9}$$

Further,  $P_1 \left\{ \tilde{S}_{k_n-1}^\psi > u_{k_n-1} \right\} \leq P_1 \left\{ \tilde{S}_n^h > z + \kappa_n(h) \right\} \leq c_2 < 1$  by definition of  $k_n$  and (3.4). So

$$\begin{aligned}
 & P_1 \left\{ \tilde{S}_{k_n-1}^\psi > u_{k_n-1} \right\} \\
 &= P_1 \left\{ \frac{S_{k_n-1}^\psi - NA_{1,k_n-1}(\psi)}{\sqrt{N}\sigma_{1,k_n-1}(\psi)} > (u_{k_n-1} - \kappa_{k_n-1}(\psi))(1 + o(1)) \right\} \leq c_2 < 1,
 \end{aligned}$$

hence

$$u_{k_n-1} - \kappa_{k_n-1}(\psi) \geq -c_4. \tag{3.10}$$

So,  $u_{k_n-1} \rightarrow \infty$ , because  $\kappa_{k_n-1}(\psi) \rightarrow \infty$  under alternatives  $\mathcal{I}_{all}$ . Put

$$\kappa'_{k_n-1}(\psi) := \kappa_{k_n-1}(\psi) \frac{n\sqrt{N(k_n)}}{\rho(\psi, \lambda_{k_n-1})\sqrt{N(n)}(k_n - 1)} = \frac{n}{\sqrt{2N}}\varepsilon(k_n).$$

Then  $\kappa'_{k_n-1}(\psi) \rightarrow \infty$ ,  $\kappa'_{k_n-1}(\psi) = o(a_n)$  under alternatives  $\mathcal{I}$ . Hence by assumption (d) we obtain

$$\log P_0 \left\{ \tilde{S}_{k_n-1}^\psi > \kappa'_{k_n-1}(\psi) \right\} = -\frac{1}{2}(\kappa'_{k_n-1}(\psi))^2(1 + o(1)) = -\frac{1}{4}n\lambda_n \varepsilon^2(k_n)(1 + o(1)). \tag{3.11}$$

On the other hand by (3.7) and definition of  $u_{k_n-1}$  we have.

$$\begin{aligned}
 \log P_0 \left\{ \tilde{S}_{k_n-1}^\psi > u_{k_n-1} \right\} &= \log \alpha_n = -\frac{1}{2}\tau_n n\lambda_n \rho^2(h, \lambda_n)\varepsilon^2(n)(1 + o(1)) \\
 &\geq -\frac{1}{4}n\lambda_n \varepsilon^2(n)(1 + o(1)) = \log P_0 \left\{ \tilde{S}_{k_n-1}^\psi > \kappa'_{k_n-1} \right\},
 \end{aligned} \tag{3.12}$$

since  $\tau_n \leq 1/2$ , (3.11) and that  $\varepsilon^2(n) \sim \varepsilon^2(k_n)$  for the alternatives  $\mathcal{I}$ . From (3.12) it follows that  $u_{k_n-1} \leq \kappa'_{k_n-1} = o(a_n)$ . Then by assumption (d)

$$-\log P_0 \left\{ \tilde{S}_{k_n-1}^\psi > u_{k_n-1} \right\} = 2^{-1}u_{k_n-1}^2(1 + o(1)).$$

Due to this fact, (3.10), (3.11) and (3.12) we obtain

$$\begin{aligned}
 2^{-1}(k_n - 1)\lambda_{k_n-1}\rho^2(\psi, \lambda_{k_n})\varepsilon^2(k_n)(1 + o(1)) &= 2^{-1}\kappa_{k_n-1}^2(\psi) \leq 2^{-1}u_{k_n-1}^2(1 + o(1)) \\
 &= -\log P_0 \left\{ \tilde{S}_{k_n-1}^\psi > u_{k_n-1} \right\} = -\log \alpha_n = \tau_n n\lambda_n \rho^2(h, \lambda_n)\varepsilon^2(n)(1 + o(1)).
 \end{aligned}$$

That is

$$\frac{(k_n - 1)^2}{N(k_n - 1)} \frac{N(n)}{n^2} \leq \frac{2\tau_n \rho^2(h, \lambda_n)}{\rho^2(\psi, \lambda_{k_n-1})}.$$

Hence, under alternatives  $\mathcal{I}$  for large enough  $n$

$$\frac{k_n - 1}{n} \leq \left( \frac{2\tau_n \rho^2(h, \lambda_n)}{\rho^2(\psi, \lambda_{k_n-1})} \right)^{1/(2-q)}. \tag{3.13}$$

Proof of Theorem 2.7 concludes from (3.9) and (3.13) by squeeze theorem for sequences and that  $h, \psi \in S$ .  $\square$

*Proof of Theorems 2.9–2.15.* A few remarks first. We use the  $\chi^2$  test as a benchmark and it is considered as  $h$ -test of Theorem 2.7, hence the condition (b) is fulfilled. Also, since  $\rho(h, \lambda_n) = 1$  if  $h(x) = x^2$  and  $\psi \in S$ , that is  $|\rho(\psi, \lambda_n)| \geq c > 0$ , it follows from Theorem 2.7 that under the conditions (a),(c),(d) and (f): if  $\tau_n = c$  then there exists a constant  $c_1 > 0$  such that  $\lambda_{k_n} \sim \lambda_n(k_n/n)(N(n)/N(k_n)) \sim \lambda_n(k_n/n)^{1-q} \sim c_1 \lambda_n$ , also if  $\tau_n \rightarrow 0$  then  $e(\chi_n^2, S_n^\psi) = 0$ . We will use these facts. Finally, it suffices to indicate that conditions (c), (d) and (f) are satisfied under the conditions of each of Theorems 2.9–2.15, so we will no longer refer to Theorem 2.7 in the course of proving the theorems.

All relevant large deviations results for the  $\chi^2$  statistic are collected in the following Proposition 3.1. Note that for the  $\chi^2$  statistic

$$h(u) = (u - \lambda_n)^2/\lambda_n, \quad Eh(\xi) = 1, \quad \sigma^2(h) = 2,$$

$$\rho(h, \lambda_n) = 1, \quad \kappa_n(h) = \kappa_n = \sqrt{n\lambda_n/2}\varepsilon(n).$$

$\square$

**Proposition 3.1.** (A) Let either

(i)  $\lambda_n \rightarrow \lambda \in (0, \infty)$  and  $\varepsilon(n) \ll n^{-1/3}$ , or

(ii)  $\lambda_n \rightarrow 0, n\lambda_n^3 \rightarrow \infty$  and  $\varepsilon(n) \ll n^{-1/3}$ , or

(iii)  $\lambda_n \rightarrow \infty$  and  $\varepsilon(n) \ll (n\lambda_n^2)^{-1/3}$ , or

(iv)  $N \ll n^{3/8}$  and  $\sqrt{n\lambda_n}\varepsilon(n) \rightarrow \infty$  and  $\varepsilon(n) \ll \lambda_n^{-1}$ .

Then for every sequence of  $y_n = O(\sqrt{n\lambda_n}\varepsilon(n))$  it holds

$$-\lim_{n \rightarrow \infty} \frac{2}{y_n^2} \log P_0 \left\{ (\chi_N^2 - N) / \sqrt{2N} > y_n \right\} = 1.$$

(B) Let  $\sqrt{n} \leq N$  and

$$\varepsilon(n) = (n\lambda_n)^{-1/3}\omega_n^{2/3}, \quad \text{where } \omega_n \rightarrow \infty, \omega_n = o(\sqrt{n\lambda_n}), \log(N^2/n) = o(\omega_n). \tag{3.14}$$

Then there exists a sequence  $(\tau_n), \tau_n \rightarrow 0$  such that for every sequence  $u_n$  such that  $u_n = O(\sqrt{n\lambda_n}\varepsilon(n))$  it holds

$$-\lim_{n \rightarrow \infty} \frac{1}{\tau_n u_n^2} \log P_0 \left\{ (\chi_n^2 - N) / \sqrt{2N} > u_n \right\} = 1.$$

**Remark 3.2.** In fact, there is no need for  $\varepsilon(n)$  to be the same as in (1.1). Proposition 3.1 still true for an arbitrary sequence of positive numbers  $\varepsilon(n) \rightarrow 0$  that satisfies the corresponding conditions.

Proof of Proposition 3.1 By Corollary 4.2 of [13] it holds

$$\log P_0 \left\{ \chi_N^2 > x_n \sqrt{2N} + N \right\} = -\frac{1}{2}x_n^2(1 + o(1)) \tag{3.15}$$

for arbitrary  $\lambda_n$  and every sequence of positive numbers  $x_n \rightarrow \infty$  such that

$$x_n = o\left((\sqrt{N} \min(1, \lambda_n^2))^{1/3}\right). \tag{3.16}$$

Also, equation (2.17), of [10] gives that (3.15) still hold if

$$N \ll \sqrt{n}, x_n = o(\sqrt{N}) \text{ and } N^{-3/2}n^{1/2}x_n \rightarrow \infty. \tag{3.17}$$

Part (A) of Proposition 3.1 follows if we choose  $\varepsilon(n)$  so that  $\kappa_n = n\varepsilon(n)/\sqrt{2N}$  satisfies (3.16) and (3.17) in each cases, respectively. Indeed, for the sparse and dense models (cases (i) and (iii)) we have in (3.16)  $x_n \ll N^{1/6}$ , whereas for very sparse models (case (ii))  $x_n \ll (\sqrt{N}\lambda_n^2)^{1/3} = (n\lambda_n^3)^{1/6}$ . Using these facts we can easily observe that if  $\varepsilon(n)$  is as in the assertions (i),(ii) and (iii) then for every positive  $y_n = O(\kappa_n)$  equation (3.16) is valid. Part (iv) assumes that we are dealing with dense model, and hence we can restrict ourselves by  $\varepsilon(n) \geq (n\lambda_n^2)^{-1/3}$ , since case (iii). Then (3.17) is fulfilled with  $x_n = \kappa_n$  if  $N \ll n^{3/8}$ . Part (A) is proven completely. Proof of Part (B) is alike to proof of part (iii) of Theorem 3.4 of [14], p.16–18).  $\square$

**Proposition 3.3.** *Let  $\lambda_n \rightarrow \lambda \in (0, \infty)$  and  $d^* = \max(1, d)$  For every  $d > -1$  and  $x_n \rightarrow \infty, x_n = o(n^{1/2(1+2d^*)})$  one has*

$$-\log P_0 \left\{ \tilde{S}_n^{\psi_d} > x_n \right\} = \frac{x_n^2}{2}(1 + o(1)). \tag{3.18}$$

Moreover, (3.18) remains true for  $x_n \rightarrow \infty, x_n = o(\sqrt{n})$ , if statistics  $S_n^\psi$  satisfy the Cramer condition.

*Proof.* By Corollary 3.1 of [13] equation (3.18) is valid if  $d > 0$  and  $x_n = o(\min(n^{1/6}, n^{1/(1+2d)})) = o(n^{1/(1+2d^*)})$ . Further, it follows from Theorem 2 of [8] that (3.18) holds for  $x_n \rightarrow \infty, x_n = o(\sqrt{n})$  if statistics  $S_n^\psi$  satisfy the Cramer condition, in particular for PDS with tuning parameter  $d \in (-1, 0]$ . Proposition 3.3 follows.  $\square$

*Proof of Theorem 2.9.* Condition (c), where  $\mathcal{I}_1$  such that  $\varepsilon(n) \ll n^{-1/3}$  and  $\tau_n = 1/2$ , is fulfilled since Proposition 3.1 part (A) case (i). Condition (d) is fulfilled with  $a_n = n^{1/2(1+2d^*)}$  by Proposition 3.3, hence the family  $\mathcal{I}_2$  of condition (f) we should determine so that  $n\varepsilon(n)/\sqrt{2N} \ll n^{1/2(1+2d^*)}$ , that is  $\varepsilon(n) \ll n^{-d^*/(1+2d^*)}$ . Then  $\mathcal{I} = \mathcal{I}_1 \cap \mathcal{I}_2$  is such that  $\varepsilon(n) \ll n^{-d^*/(1+2d^*)}$ . Part (i) follows, since in this case  $|\rho(\psi, \lambda_{k_n})| < 1$ . Further, under the conditions of part (ii)  $\mathcal{I}_1$  is such that  $\varepsilon(n) \gg n^{-1/3} \log^{2/3} n$  and  $\tau_n \rightarrow 0$  by Proposition 3.1 part (B), whereas by Proposition 3.3 in this case  $a_n = \sqrt{n}$ , and hence  $\mathcal{I}_2 = \mathcal{I}_{alt}$ . Part (ii) follows.  $\square$

*Proof of Theorem 2.12.* Set  $W_n(d) = \min(n^{1/4}, (n\lambda_n^3)^{1/2(1+2d^*)})$ , where  $d^* = \max(0, d)$ .  $\square$

**Proposition 3.4.** *Let  $\lambda_n \rightarrow 0$  and  $n\lambda_n^3 \rightarrow \infty$ . Then uniformly in  $x_n \geq 0, x_n = o(W_n(d))$  it holds*

$$-\log P_0 \left\{ \tilde{S}_n^{\psi_d} > x_n \right\} = \frac{x_n^2}{2}(1 + o(1)), d > -1.$$

**Proof** follows from Corollary 3.3 of [13], because standardized r.v.  $\tilde{S}_n^{h_d}$  coincides with accordingly standardized version of the r.v.  $R_n^d = \eta_1^{d+1} + \dots + \eta_N^{d+1}, d \neq 0$ , and  $R_n^0 = 2\eta_1 \ln \eta_1 + \dots + 2\eta_N \ln \eta_N$ .

*Proof of part (A).* By Proposition 3.1 (ii) Condition (c) is fulfilled, where  $\mathcal{I}_1$  defines as  $\varepsilon(n) \ll n^{-1/3}$  and  $\tau_n = 1/2$ . Next, by Proposition 3.4 we have  $a_n = n^{1/4}$ , because if  $N \leq n^{(\tau-2d^*)/6}$  and  $d \in (-1, 1/2)$  then  $W_n(d) = n^{1/4}$ . So,  $\mathcal{I}_2$  defines by  $\varepsilon(n) \ll \sqrt{N}/n^{3/4} = (n\lambda_n^2)^{-1/4}$ . Then  $\mathcal{I} = \mathcal{I}_1$ , since  $n \ll N$ . Part (i) follows, because  $\rho(\psi_d, \lambda_n) \rightarrow 1$  by Lemma 2.4. Part (ii) also follows by similar reasons using Proposition 3.1 part (B) (instead of part (ii)) due to which  $\tau_n \rightarrow 0$  and  $\mathcal{I}_1$  is defined by  $\varepsilon(n) = (n\lambda_n)^{-1/3} \omega_n^{2/3} \ll (n\lambda_n^2)^{-1/4}$ , where  $\omega_n \ll (N/\sqrt{n})^{1/4}$  (note that  $(N/\sqrt{n})^{1/4} \leq n/\sqrt{N}$ , because  $N \ll n^{4/3}$  by condition  $n\lambda_n^3 \rightarrow \infty$ ).

*Proof of part (B).* From Proposition 3.1 and 3.4 we have:  $\mathcal{I}_1$  is such that  $\varepsilon(n) \ll n^{-1/3}, a_n = W_n(d) = (n\lambda_n^3)^{1/2(2d+1)}$  since  $d \geq 1/2$ , and  $\mathcal{I}_2$  is such that  $\varepsilon(n) \ll [n^d \lambda_n^{d-1}]^{-1/(2d+1)}, \tau_n = 1/2$ . Hence  $\mathcal{I}_2 \subset \mathcal{I}_1$  because  $N \ll n^{4/3}$ . Part (B) follows since Lemma 2.4.

*Proof of Theorem 2.14.* Set  $\pi_r(\lambda) = \lambda^r e^{-\lambda}/r!$ ,

$$A_{rn} = \sum_{m=1}^N \pi_r(\lambda_n),$$

$$\sigma_{rn}^2 = \sum_{m=1}^N \pi_r(\lambda_n)(1 - \pi_r(\lambda_n)) - n^{-1} \left( \sum_{m=1}^N (r - \lambda_n)\pi_r(\lambda_n) \right)^2.$$

From Corollary 4.3 of [13] it follows. □

**Proposition 3.5.** *Let  $\lambda_n \rightarrow 0$  and  $n\lambda_n \rightarrow \infty$ . Then for  $r = 0, 1, 2$  and  $x_n \rightarrow \infty$ ,  $x_n = o(n^{1/4})$  it holds*

$$\log P_0\{\mu_r > \sigma_{r,n}x_n + A_{r,n}\} = -x_n^2/2(1 + o(1)).$$

By Proposition 3.1 and 3.5 we have:  $\mathcal{I}_1$  is such that  $\varepsilon(n) \ll n^{-1/3}$ , whereas  $a_n = n^{1/4}$ ,  $\tau_n = 1/2$ ,  $\mathcal{I}_2$  defines by  $\varepsilon(n) \ll \sqrt{N}/n^{3/4} = (n\lambda_n^2)^{-1/4}$ , and hence  $\mathcal{I}_1 \subset \mathcal{I}_2$ . Proof concludes, since Lemma 2.4, according of which  $\rho(h, \lambda_n) \rightarrow 1$  for  $h(x) = I\{x = r\}$  if  $r = 0, 1, 2$ . The second statement follows because  $C_n = \mu_0 - (n - N)$ .

*Proof of Theorem 2.15.* We will use the following propositions. □

**Proposition 3.6.** *For every  $d > -1$  and  $x_n \rightarrow \infty$ ,  $x_n = o(N^{1/6})$  one has*

$$-\log P_0\{\tilde{S}_n^{\psi_d} > x_n\} = x_n^2/2(1 + o(1)).$$

Proof follows from Corollary 3.2 of [13].

**Proposition 3.7.** *If  $N \ll n^{3/7}$  then for  $x_n \rightarrow \infty$  and  $x_n = o(\sqrt{N})$  it holds*

$$-\log P_0\{\Lambda_n > x_n\sqrt{2N} + N\} = x_n^2/2(1 + o(1)).$$

Proof follows from Proposition 3.6 if  $x_n = o(N^{1/6})$ . If  $x_n \geq N^{1/6}$ , then by equation (2.13) of [10] and that  $N = o(n^{3/7})$  we have

$$\log P_0\{\Lambda_n > x_n\sqrt{2N} + N\} = -\frac{x_n^2}{2} + O\left(\frac{x_n^3}{\sqrt{N}} + \log N + \frac{N^{3/2}}{\sqrt{n}}\right) = -\frac{x_n^2}{2}(1 + o(1)).$$

To prove part (i) we define for every  $d > -1$  the family  $\mathcal{I}_1$  by  $\varepsilon(n) \ll (n\lambda_n^2)^{-1/3}$  according to Proposition 3.1(iii), and we set  $\mathcal{I}_2 = \mathcal{I}_1$  because due to Proposition 3.6  $a_n = N^{1/3}$  and hence  $\varepsilon(n) \ll N^{1/6}\sqrt{N}/n = (n\lambda_n^2)^{-1/3}$ . To prove part (ii) we use Proposition 3.1 part (iv) and Proposition 3.7, according to which we define  $\mathcal{I}_1 = \mathcal{I}_2 = \mathcal{I}_{alt}$ , because  $a_n = \sqrt{N}$ . Theorem 2.15 follows, since  $\lambda_{k_n} \rightarrow \infty$  and hence  $\rho(\psi_d, \lambda_{k_n}) \rightarrow 1$  by Lemma 2.5.

*Proof of Lemma 2.3.* Consider two statistics  $S_n^{h_1}$  and  $S_n^{h_2}$ . Denote

$$\tilde{\xi}_m = (\xi_m - \lambda_n) / \sqrt{n}, \quad \tau_{jn} = \lambda_n^{-1} cov(h_j(\xi), \xi)$$

$$g_j(x) = h_j(x) - Eh_j(\xi) - \tau_{jn}(x - \lambda_n),$$

$$R_n^{g_j} = \sum_{m=1}^N g_j(\eta_m), \quad V_n^{g_j} = \sum_{m=1}^N g_j(\xi_m), \quad \zeta_n = \sum_{m=1}^N (\xi_m - \lambda_n).$$

We have  $S_n^{h_j} = R_n^{g_j} + NEg_j(\xi)$  and  $cov_0(S_n^{h_1}, S_n^{h_2}) = cov_0(R_n^{g_1}, R_n^{g_2})$ . The following equality holds:

$$cov_0(S_n^{h_1}, S_n^{h_2}) = v_n \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} E_0(V_n^{g_1} V_n^{g_2} \exp\{i\tau\zeta_n/\sqrt{n}\}) d\tau, \tag{3.19}$$

where

$$v_n := (2\pi\sqrt{n}P\{\zeta_n = 0\})^{-1} = \frac{n!e^n}{2\pi n^n\sqrt{n}} = \frac{1}{\sqrt{2\pi}} \left(1 + o\left(\frac{1}{n}\right)\right).$$

□

**Indeed.** Note that  $E_0 V_n^{g_1} = 0$ . It is well known that  $\mathcal{L}((\eta_1, \dots, \eta_N)) = \mathcal{L}((\xi_1, \dots, \xi_N) \mid \zeta_n = 0)$ , where  $\mathcal{L}(X)$  stands for the distribution of a random vector  $X$ . Hence

$$\text{cov}_0(R_n^{g_1}, R_n^{g_2}) = E_0(R_n^{g_1} R_n^{g_2}) = E_0(V_n^{g_1} R_n^{g_2} \mid \zeta_n = 0).$$

On the other hand  $E(R_n^{g_1} R_n^{g_2} e^{i\tau\zeta_n}) = E\{e^{i\tau\zeta_n} E(R_n^{g_1} R_n^{g_2} \mid \zeta_n)\}$ . Now (3.19) follows by Fourier inversion. We have

$$\begin{aligned} & \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} E_0(V_n^{g_1} V_n^{g_2} \exp\{i\tau\zeta_n/\sqrt{n}\}) \, d\tau \\ &= \sum_{m=1}^N \sum_{j=1}^N \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} E_0(g_1(\xi_m)g_2(\xi_j) \exp\{i\tau\zeta_n/\sqrt{n}\}) \, d\tau \\ &= \sum_{m=1}^N \sum_{j=1}^N \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} E_0(g_1(\xi_m)g_2(\xi_j) \exp\{i\tau(\tilde{\xi}_m + \tilde{\xi}_j)\}) \prod_{l=1, l \neq m, j}^N E_0 \exp\{i\tau\tilde{\xi}_l\} \, d\tau \\ &= \sum_{m=1}^N \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} (E_0 \exp\{i\tau\tilde{\xi}_1\})^{N-1} E_0(g_1(\xi_m)g_2(\xi_m)e^{i\tau\tilde{\xi}_m}) \, d\tau \\ &+ \sum_{\substack{m, j=1 \\ m \neq j}}^N \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} (E_0 \exp\{i\tau\tilde{\xi}_1\})^{N-2} E_0g_1(\xi_m)e^{i\tau\tilde{\xi}_m} E_0g_2(\xi_j)e^{i\tau\tilde{\xi}_j} \, d\tau \\ &= N \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} (E_0 \exp\{i\tau\tilde{\xi}_1\})^{N-1} E_0(g_1(\xi_1)g_2(\xi_1)e^{i\tau\tilde{\xi}_1}) \, d\tau \\ &+ N(N-1) \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} (E_0 \exp\{i\tau\tilde{\xi}_1\})^{N-2} E_0g_1(\xi_1)e^{i\tau\tilde{\xi}_1} E_0g_2(\xi_1)e^{i\tau\tilde{\xi}_1} \, d\tau \\ &:= NQ_1 + N(N-1)Q_2. \end{aligned} \tag{3.20}$$

We have

$$E_0 e^{i\tau\tilde{\xi}_1} = \exp\left\{\lambda_n \left(e^{i\tau/\sqrt{n}} - 1 - i\tau/\sqrt{n}\right)\right\} = \exp\left\{-\frac{\tau^2}{2N} + \frac{\theta|\tau|^3}{6N\sqrt{n}}\right\}, \tag{3.21}$$

$$\left|E_0 e^{i\tau\tilde{\xi}_1}\right| = \exp\left\{-2\lambda_n \sin^2 \frac{\tau}{2\sqrt{n}}\right\} \leq \exp\left\{-\frac{2\tau^2}{N\pi^2}\right\}. \tag{3.22}$$

Write

$$\begin{aligned}
 Q_1 &= \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \exp\left\{-\frac{\tau^2(N-1)}{2N}\right\} E_0\left(g_1(\xi_1)g_2(\xi_1)e^{i\tau\tilde{\xi}_1}\right) d\tau \\
 &+ \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \left[\left(E_0 \exp\left\{i\tau\tilde{\xi}_1\right\}\right)^{N-1} - \exp\left\{-\frac{\tau^2(N-1)}{2N}\right\}\right] E_0\left(g_1(\xi_1)g_2(\xi_1)e^{i\tau\tilde{\xi}_1}\right) d\tau \\
 &:= Q_{11} + Q_{12}.
 \end{aligned} \tag{3.23}$$

We have

$$\begin{aligned}
 Q_{11} &= E_0\left(g_1(\xi_1)g_2(\xi_1)\right) \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \exp\left\{-\frac{\tau^2(N-1)}{2N}\right\} d\tau \\
 &+ \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \exp\left\{-\frac{\tau^2(N-1)}{2N}\right\} E_0\left(g_1(\xi_1)g_2(\xi_1)(e^{i\tau\tilde{\xi}_1} - 1)\right) d\tau \\
 &= E_0\left(g_1(\xi_1)g_2(\xi_1)\right) \sqrt{2\pi}(1 + o(1)) + cE_0\left|g_1(\xi_1)g_2(\xi_1)\tilde{\xi}_1^2\right|.
 \end{aligned} \tag{3.24}$$

Also,

$$\begin{aligned}
 Q_{12} &= \frac{c}{\sqrt{n}} E_0\left|g_1(\xi_1)g_2(\xi_1)\right| \int_{|\tau|\leq\sqrt{n}} |\tau|^3 \exp\left\{-\frac{\tau^2(N-1)}{3N}\right\} d\tau \\
 &+ E_0\left|g_1(\xi_1)g_2(\xi_1)\right| \int_{\sqrt{n}\leq|\tau|\leq\pi\sqrt{n}} \left(\exp\left\{-\frac{2\tau^2(N-1)}{\pi^2N}\right\} + \exp\left\{-\frac{\tau^2(N-1)}{2N}\right\}\right) d\tau \\
 &= \frac{c}{\sqrt{n}} E_0\left|g_1(\xi)g_2(\xi_1)\right| (1 + o(1)).
 \end{aligned} \tag{3.25}$$

Use the facts that  $E_0g_i(\xi_1) = 0$ ,  $E_0g_i(\xi_1)\tilde{\xi}_1 = 0$  and inequality  $|e^{it} - 1 - it| \leq t^2/2$  to get

$$\begin{aligned}
 |Q_2| &= \left| \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \left(E_0 \exp\left\{i\tau\tilde{\xi}_1\right\}\right)^{N-2} E_0g_1(\xi_1)(e^{i\tau\tilde{\xi}} - 1 - i\tau\tilde{\xi}_1)E_0g_2(\xi_1)(e^{i\tau\tilde{\xi}_1} - 1 - i\tau\tilde{\xi}_1)d\tau \right| \\
 &\leq \frac{1}{4} \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \left|E_0 \exp\left\{i\tau\tilde{\xi}_1\right\}\right|^{N-2} \tau^4 E_0(|g_1(\xi_1)|\tilde{\xi}_1^2)E_0(|g_2(\xi_1)|\tilde{\xi}_1^2)d\tau
 \end{aligned}$$

$$\begin{aligned}
&\leq \frac{1}{4} E_0(|g_1(\xi_1)| \tilde{\xi}_1^2) E_0(|g_2(\xi_1)| \tilde{\xi}_1^2) \int_{-\pi\sqrt{n}}^{\pi\sqrt{n}} \tau^4 \exp\left\{-\frac{2(N-2)\tau^2}{N\pi^2}\right\} d\tau \\
&\leq c E_0(|g_1(\xi_1)| \tilde{\xi}_1^2) E_0(|g_2(\xi_1)| \tilde{\xi}_1^2) \leq c (E_0 g_1^2(\xi_1))^{1/2} (E_0 g_2^2(\xi_1))^{1/2} E_0 \tilde{\xi}_1^4 \\
&= c (E_0 g_1^2(\xi_1))^{1/2} (E_0 g_2^2(\xi_1))^{1/2} (N^{-2} + 3(nN)^{-1}). \tag{3.26}
\end{aligned}$$

Thus, by (3.19)–(3.26) we have

$$\begin{aligned}
&\frac{1}{N} \text{cov}_0(R_n^{g_1}, R_n^{g_2}) = \text{cov}_0(g_1(\xi_1), g_2(\xi_1)) (1 + o(1)) \\
&+ c E_0 \left| g_1(\xi_1) g_2(\xi_1) \tilde{\xi}_1^2 \right| + \frac{c}{\sqrt{n}} E_0 |g_1(\xi_1) g_2(\xi_1)| (1 + o(1)) \\
&+ c (E_0 g_1^2(\xi_1))^{1/2} (E_0 g_2^2(\xi_1))^{1/2} (N^{-1} + 3n^{-1}).
\end{aligned}$$

Use here the facts that

$$N^{-1} \text{Var}_0 R_n^{g_j} = E g_j^2(\xi) (1 + o(1)), E |g_1(\xi) g_2(\xi)| \leq \sqrt{E g_1^2(\xi) E g_2^2(\xi)},$$

$$\tilde{\xi}_1^2 = (\xi_1 - \lambda_n)^2 / N \lambda_n = o_p(N^{-1/2})$$

to get

$$\begin{aligned}
\text{corr}_0(R_n^{g_1}, R_n^{g_2}) &= \frac{N^{-1} \text{cov}_0(R_n^{g_1}, R_n^{g_2})}{\sqrt{N^{-1} \text{Var}_0 R_n^{g_1} \cdot N^{-1} \text{Var}_0 R_n^{g_2}}} \\
&= \text{corr}(g_1(\xi), g_2(\xi)) (1 + o(1)) + O\left(\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{n}}\right).
\end{aligned}$$

Lemma 2.3 follows by taking here  $h_1 = h$  and  $S_n^{h_2} = \chi_n^2$ .

*Proof of Lemma 2.4.* Set  $\Delta h(x) = h(x+1) - h(x)$ . Since  $\lambda_n \rightarrow 0$  we have

$$\begin{aligned}
Eh(\xi) &= \left(1 - \lambda_n + \frac{1}{2} \lambda_n^2 + O(\lambda_n^3)\right) \left(h(0) + \lambda h(1) + \frac{1}{2} \lambda^2 h(2) + O(\lambda^3)\right) \\
&= h(0) + \lambda_n \Delta h(0) + \frac{1}{2} \lambda_n^2 \Delta^2 h(0) + O(\lambda_n^3), \tag{3.27}
\end{aligned}$$

$$Eh^2(\xi) = h^2(0) + \lambda\Delta h^2(0) + \frac{1}{2}\lambda^2\Delta^2 h^2(0) + O(\lambda_n^3), \quad (3.28)$$

$$(Eh(\xi))^2 = h^2(0) + 2\lambda h(0)\Delta h(0) + \lambda^2(\Delta h(0))^2 + \lambda^2 h(0)\Delta^2 h(0) + O(\lambda_n^3), \quad (3.29)$$

Further,

$$\begin{aligned} Eh(\xi)\xi &= \left(1 - \lambda_n + \frac{1}{2}\lambda_n^2 + O(\lambda_n^3)\right) (\lambda_n h(1) + \lambda_n^2 h(2) + \lambda_n^3 2^{-1} h(3) + O(\lambda_n^4)) \\ &= \lambda_n h(1) + \lambda_n^2 \Delta h(1) + \lambda_n^3 2^{-1} \Delta^2 h(1) + O(\lambda_n^4). \end{aligned} \quad (3.30)$$

Use this and (3.27) to get

$$\tau_n = \lambda^{-1} \text{cov}(h(\xi), \xi) = \lambda^{-1} Eh(\xi)\xi - Eh(\xi) = \Delta h(0) + \lambda \Delta^2 h(0).$$

Hence

$$\lambda \tau_n^2 = \lambda(\Delta h(0))^2 + 2\lambda^2 \Delta h(0)\Delta^2 h(0) + \lambda_n^3 [(\Delta^2 h(0))^2 + \Delta h(0)\Delta^3 h(0)]. \quad (3.31)$$

By (3.28), (3.29) and (3.31) we obtain

$$\begin{aligned} \sigma^2(h) &= Eh^2(\xi) - (Eh(\xi))^2 - \lambda \tau_n^2 = \lambda (\Delta h^2(0) - 2h(0)\Delta h(0) - (\Delta h(0))^2) \\ &\quad + \frac{1}{2}\lambda^2 (\Delta^2 h^2(0) - 2(\Delta h(0))^2 - 2h(0)\Delta^2 h(0) - 4\Delta h(0)\Delta^2 h(0)) \\ &\quad + \frac{1}{6}\lambda^3 [\Delta^3 h^2(0) - 6\Delta h(0)\Delta^2 h(0) - 2h(0)\Delta^3 h(0) - 6(\Delta^2 h(0))^2 - 6\Delta h(0)\Delta^3 h(0)] \\ &:= \lambda_n \nabla_1 + \frac{1}{2}\lambda_n^2 \nabla_2 + \frac{1}{6}\lambda_n^3 \nabla_3. \end{aligned} \quad (3.32)$$

We have

$$\nabla_1 = (h^2(1) - h^2(0) - 2h(0)h(1) + 2h^2(0) - h^2(1) + 2h(1)h(0) - h^2(0)) = 0,$$

$$\nabla_2 = h^2(2) + 4h^2(1) + h^2(0) - 4h(0)h(1) + 2h(0)h(2) - 4h(1)h(2) = \frac{1}{2}\lambda^2 (\Delta^2 h(0))^2,$$

$$\nabla_3 = \Delta^3 h^2(0) - 6\Delta h(1)\Delta^2 h(0) - 2\Delta^3 h(0)[h(1) + 2\Delta h(0)].$$

Apply these in (3.32) to get

$$\begin{aligned} \sigma^2(h) &= \frac{1}{2}\lambda^2(\Delta^2h(0))^2 \\ &\cdot \left[ 1 + \frac{\lambda_n}{3} \frac{\Delta^3h^2(0) - 6\Delta h(1)\Delta^2h(0) - 2\Delta^3h(0)[h(1) + 2\Delta h(0)]}{(\Delta^2h(0))^2} + o(\lambda_n) \right] \\ &:= \frac{1}{2}\lambda^2(\Delta^2h(0))^2[1 + \lambda_n 3^{-1}\omega(h)] \text{ ( for short )}. \end{aligned} \tag{3.33}$$

We find that

$$Eh(\xi)\xi^2 = \lambda h(1) + \lambda^2[h(2) + \Delta h(1)] + \frac{1}{2}\lambda^3[2\Delta h(2) + \Delta^2h(1)] + O(\lambda_n^4).$$

Using this,(3.30) and (3.27) after some algebra we obtain

$$\begin{aligned} cov(h(\xi) - \tau_n\xi, \xi^2 - (2\lambda_n + 1)\xi) &= Eh(\xi)\xi^2 - (2\lambda_n + 1)Eh(\xi)\xi + \lambda^2Eh(\xi) \\ &= \lambda_n^2\Delta^2h(0) + \frac{1}{2}\lambda_n^3[\Delta^3h(1) + \Delta h(2)] + O(\lambda_n^4). \end{aligned}$$

Due to this, (3.33) and (3.2) we obtain

$$\begin{aligned} \rho(h, \lambda_n) &= \frac{\lambda_n^2\Delta^2h(0) + \lambda_n^3 2^{-1}[\Delta^3h(1) + \Delta h(2)]}{\lambda^2\Delta^2h(0)\sqrt{1 + \lambda_n 3^{-1}\omega(h)}} \\ &= [\Delta^2h(0)]^{-1} (\Delta^2h(0) + \lambda_n 2^{-1}[\Delta^3h(1) + \Delta h(2)]) (1 - \lambda_n 6^{-1}\omega(h)) \\ &= \left( 1 + [\Delta^2h(0)]^{-1} \lambda_n 2^{-1}[\Delta^3h(1) + \Delta h(2)] \right) (1 - \lambda_n 6^{-1}\omega(h)) \\ &= 1 - \lambda 6^{-1} \left( \omega(h) - [\Delta^2h(0)]^{-1} [\Delta^3h(1) + \Delta h(2)] + O(\lambda_n^2) \right) \\ &= 1 - \frac{\lambda_n}{6(\Delta^2h(0))^2} (\Delta^3h^2(0) - 6\Delta h(1)\Delta^2h(0) - 2\Delta^3h(0)[h(1) + 2\Delta h(0)] \\ &\quad - 3\Delta^2h(0)[\Delta^3h(1) + \Delta h(2)] + O(\lambda_n^2)) = 1 - \frac{\lambda_n(\Delta^3h(0))^2}{6(\Delta^2h(0))^2} + O(\lambda_n^2). \end{aligned}$$

□

*Proof of Lemma 2.5.* The case  $d = 0$  is presented by [7]. Let  $d \neq 0$ . Rewrite the PDS in the form (2.9), where

$$\begin{aligned} \bar{\psi}_d(x) &= \frac{2}{d(d+1)} \left[ (x/\lambda_n)^{d+1} - (d+1)(x/\lambda_n) + d \right] \\ &= \frac{2}{d(d+1)} \left[ \left( 1 + \frac{x - \lambda_n}{\lambda_n} \right)^{1+d} - (1+d) \left( 1 + \frac{x - \lambda_n}{\lambda_n} \right) + d \right]. \end{aligned}$$

We have  $\bar{\psi}_d(\lambda_n) = 0$ ,  $\bar{\psi}'_d(\lambda_n) = 0$ ,  $\bar{\psi}''_d(\lambda_n) = 2\lambda_n^{-2}$ ,  $\bar{\psi}'''_d(\lambda_n) = 2(d-1)\lambda_n^{-3}$ ,  $\bar{\psi}^{(iv)}_d(\lambda_n) = 2(d-1)(d-2)\lambda_n^{-4}$ . Set  $\bar{\xi}_n = (\xi - \lambda_n)/\lambda_n$ , then  $\bar{\xi}_n = O_p(\lambda_n^{-1/2})$ , since  $\sqrt{\lambda_n}\bar{\xi}_n$  has an asymptotically normal distribution as  $\lambda_n \rightarrow \infty$ . Apply these facts in Taylor expansion formula for the  $\bar{\psi}_d(x)$  to get

$$\bar{\psi}_d(\xi) = \bar{\xi}_n^2 + \frac{d-1}{3}\bar{\xi}_n^3 + \frac{(d-1)(d-2)}{12}\bar{\xi}_n^4 + O_p(\bar{\xi}_n^5). \tag{3.34}$$

Put  $\mu_l = E(\xi - \lambda_n)^l$ . Then  $\mu_2 = \mu_3 = \lambda_n$ ,  $\mu_4 = 3\lambda_n^2 + \lambda_n$ ,  $\mu_5 = 10\lambda_n^2 + \lambda_n$ ,  $\mu_6 = 15\lambda_n^3 + 25\lambda_n^2 + \lambda_n$ ,  $\mu_7 = O(\lambda_n^3)$ . Using these facts and (3.34) we obtain

$$E\bar{\psi}_d(\xi) = \frac{1}{\lambda_n} + \frac{d-1}{3} \frac{1}{\lambda_n^2} + \frac{(d-1)(d-2)}{12} \left( \frac{3}{\lambda_n^2} + \frac{1}{\lambda_n^3} \right) + O(\lambda_n^{-3}), \tag{3.35}$$

$$(E\bar{\psi}_d(\xi))^2 = \frac{1}{\lambda_n^2} + \frac{1}{\lambda_n^3} \left( \frac{2(d-1)}{3} + \frac{(d-1)(d-2)}{2} \right) + O(\lambda_n^{-4}),$$

$$E\bar{\psi}_d^2(\xi) = E\bar{\xi}_n^4 + 2\frac{d-1}{3}E\bar{\xi}_n^5 + \left( \frac{(d-1)^2}{9} + \frac{2(d-1)(d-2)}{12} \right) E\bar{\xi}_n^6 + O(\lambda_n^{-4})$$

$$= \frac{3}{\lambda_n^2} + \frac{1}{\lambda_n^3} \left( 1 + \frac{20(d-1)}{3} + \frac{(d-1)^2}{9} + \frac{5(d-1)(d-2)}{2} \right),$$

$$\tau_n = \frac{1}{\lambda_n} (E\bar{\xi}_n^2(\xi - \lambda_n) + \frac{d-1}{3}E\bar{\xi}_n^3(\xi - \lambda_n) + EO_p(\bar{\xi}_n^4(\xi - \lambda_n)))$$

$$= \frac{d}{\lambda_n^2} + \frac{1}{\lambda_n^3} \left( \frac{d-1}{3} + \frac{5(d-1)(d-2)}{6} \right) + O(\lambda_n^{-4}) = \frac{d}{\lambda_n^2} (1 + O(\lambda_n^{-1})),$$

$$\lambda_n \tau_n^2 = \frac{d^2}{\lambda_n^3} (1 + O(\lambda_n^{-1})).$$

Thus

$$\sigma_N^2(\bar{\psi}_d) = E\bar{\psi}_d^2(\xi) - (E\bar{\psi}_d(\xi))^2 - \lambda_n \tau_n^2 = \frac{2}{\lambda_n^2} \left( 1 + \frac{1}{2\lambda_n} (d^2 - 1 + \frac{(d-1)^2}{9}) + O(\lambda_n^{-2}) \right). \tag{3.36}$$

Further, by (3.34)

$$\begin{aligned} E(\bar{\psi}_d(\xi)(\xi - \lambda_n)^2) &= E\bar{\xi}_n^2(\xi - \lambda_n)^2 + \frac{d-1}{3}E\bar{\xi}_n^3(\xi - \lambda_n)^2 \\ &\quad + \frac{(d-1)(d-2)}{12}E\bar{\xi}_n^4(\xi - \lambda_n)^2 + O(\lambda_n^{-2}) \\ &= 3 + \frac{1}{\lambda_n} \left( 1 + \frac{10(d-1)}{3} + \frac{15(d-1)(d-2)}{12} \right) + O\left(\frac{1}{\lambda_n^2}\right), \end{aligned}$$

$$E(\bar{\psi}_d(\xi)(\xi - \lambda_n)) = \frac{1}{\lambda_n} + \frac{d-1}{3} \left( \frac{3}{\lambda_n} + \frac{1}{\lambda_n^2} \right) + O(\lambda_n^{-2}) = \frac{d}{\lambda_n} \left( 1 + O\left(\frac{1}{\lambda_n}\right) \right).$$

Applying these equalities and (3.35) we obtain

$$\begin{aligned} cov(h(\xi) - \tau_n \xi, \xi^2 - (2\lambda_n + 1)\xi) &= E(\bar{\psi}_d(\xi)(\xi - \lambda_n)^2) - E\bar{\psi}_d(\xi)(\xi - \lambda_n) - \lambda_n E\bar{\psi}_d(\xi) \\ &= 2 + \frac{1}{\lambda_n} \left( 1 + \frac{10(d-1)}{3} + \frac{15(d-1)(d-2)}{12} - d - \frac{d-1}{3} - \frac{(d-1)(d-2)}{4} \right) + O(\lambda_n^{-2}) \\ &= 2 + \frac{1}{\lambda_n}(d^2 - d) + O(\lambda_n^{-2}). \end{aligned} \tag{3.37}$$

Now by (3.2), (3.36) and (3.37) we get

$$\begin{aligned} \rho(h, \lambda_n) &= \frac{2 + \frac{1}{\lambda_n}(d^2 - d) + O(\lambda_n^{-2})}{2\sqrt{1 + (2\lambda_n)^{-1}(d^2 - 1 + 9^{-1}(d-1)^2) + O(\lambda_n^{-2})}} \\ &= [1 + (2\lambda_n)^{-1}(d^2 - d) + O(\lambda_n^{-2})] [1 - (2\lambda_n)^{-1}(d^2 - 1 + 9^{-1}(d-1)^2) + O(\lambda_n^{-2})] \\ &= 1 - \frac{(d-1)^2}{6\lambda_n} + O(\lambda_n^{-2}). \end{aligned}$$

□

#### 4. CONCLUSIONS

The concept of an intermediate (between the Pitman and Bahadur settings) asymptotic relative efficiency (IARE) of tests of several types in the classical problem of testing uniformity  $[0,1]$  has been elaborated in the literature. We applied this setting to study of IARE of symmetric tests in discrete model, namely in the problem

of testing of uniformity of a multinomial distribution. A formula for the IARE of arbitrary two symmetric tests, satisfying some conditions, is derived. A key point here is the validity of the large deviation results for the statistics of the comparing tests. The formula depends on the asymptotic correlation coefficients, counted under the hypothesis, of each test statistics with the  $\chi^2$  statistic, and on the asymptotic behavior of  $N$  -the number of cells, which assumed to be taken as a positive function of the continuous variable  $x$  and has a power tail of index  $q \in (0, 2)$ , i.e.,  $N(x) \sim cx^q$  as  $x \rightarrow \infty$ . The general result is applied to study IARE of the  $\chi^2$  test wrt symmetric tests for: sparse models, where  $q = 1$ , very sparse models, where it is assumed  $1 < q \leq 4/3$ , and dense models, where  $0 < q < 1$ . The proven results and still unsolved problems are as follows.

*Sparse models.* It is achieved conflicting results: The  $\chi^2$  test is superior within class of  $h$ -tests for the alternatives which converge to the hypothesis at the rate  $\varepsilon(n) = o(n^{-1/3})$ , as it is in the Pitman sense, see [3], [7]. But the IARE of the  $\chi^2$  test wrt, for instance, tests based on the log-likelihood statistic, the Freeman-Tukey statistic, CS (1.5) can't be positive for the intermediate alternatives which far from the hypothesis at the distance  $\varepsilon(n) \gg n^{-1/3} \log^{2/3} n$ , as it is in the Bahadur efficiency of the  $\chi^2$  test wrt the log-likelihood test, see [16].

IARE of  $\chi^2$  test wrt  $h$ -tests for the alternatives such that  $c_1 n^{-1/3} \leq \varepsilon(n) \leq c_2 n^{-1/3} \log^{2/3} n$  remains unsolved.

*Very sparse models.* IARE of  $\chi^2$  test wrt  $\psi_d$ -tests for such models depend on the decreasing rate of  $\lambda_n$ , the tuning parameter  $d$  and considering family of alternatives. In particular, if  $d \in (-1, 1/2)$  and  $\lambda_n \gg n^{-(1-2d^*)/6}$  then again conflict results are achieved:  $\psi_d$ -test and  $\chi^2$  test are equally efficient for alternatives such that  $\varepsilon(n) \leq n^{-1/3}$ , but for the alternatives specified by condition  $(n\lambda_n)^{-1/3} \log^{2/3}(N^2/n) \ll \varepsilon(n) \ll (n\lambda_n^2)^{-1/4}$  IARE of  $\chi^2$  test wrt  $\psi_d$ -test is equal to zero. For the case  $d \geq 1/2$  it is shown that  $\chi^2$  test and  $\psi_d$ -test are equally efficient for the alternatives specified by  $\varepsilon(n) \ll (\lambda_n^{1-d}/n^d)^{1/(2d+1)}$ . We emphasize that  $n\lambda_n^3 \rightarrow \infty$  (i.e.,  $n \ll N \ll n^{4/3}$ ) is a necessary condition for a family of alternatives with  $\varepsilon(n) \ll (\lambda_n^{1-d}/n^d)^{1/(2d+1)}$  to be a non-empty sub-family of  $\mathcal{I}_{alt}$ . Note that  $(\lambda_n^{1-d}/n^d)^{1/(2d+1)} \ll n^{-1/3}$ , if  $d \geq 1/2$ .

IARE of  $\chi^2$  test wrt  $\psi_d$ -tests for the alternatives such that  $n^{-1/3} \leq \varepsilon(n) \leq (n\lambda_n)^{-1/3}$  or  $\varepsilon(n) \geq (n\lambda_n^2)^{-1/4}$  if  $d \in (-1, 1/2)$ , and for alternatives such that  $\varepsilon(n) \geq (\lambda_n^{1-d}/n^d)^{1/(2d+1)}$  if  $d > 1$  remain open.

*Dense models.* It is pointed out that in this case under the hypothesis each PDS asymptotically in distribution coincides with the  $\chi^2$  statistics. This fact seems is reflected to IARE of  $\chi^2$  test wrt  $\psi_d$ -test, specifically the  $\chi^2$  test and  $\psi_d$  test are equally efficient for alternatives such that  $\varepsilon(n) \ll (n\lambda_n^2)^{-1/3}$  and each  $d > -1$  (here no condition to increasing rate of  $N$  is assumed); the same verdict is valid for every alternative of  $\mathcal{I}_{alt}$  and  $d = 0$  if  $N = o(n^{3/8})$ . These extend to the intermediate setting the known result that: if  $\lambda_n \rightarrow \infty$  then the Pitman efficiency of the  $\chi^2$  test wrt the  $\psi_d$ -test is equal to 1. However, I conjecture that aforesaid property of PDS effects to IARE properties of  $h$ -tests if  $N \ll \sqrt{n}$ . This is confirmed by the fact that the log-likelihood and  $\chi^2$  tests have the same asymptotic efficiency in term of  $\alpha$ -IAE for alternatives such that  $\varepsilon(n) \ll (n\lambda_n^2)^{-1/4}$  if  $N \ll \sqrt{n}$  (note that alternatives with  $\varepsilon(n) \geq (n\lambda_n^2)^{-1/4}$  are out of the family  $\mathcal{I}_{alt}$  if  $N \ll \sqrt{n}$ ), see [14]. But (ibid)  $\alpha$ -IARE of  $\chi^2$  test is much inferior wrt tests satisfying the Cramer condition, for instance the log-likelihood and Freeman-Tukey tests, if  $N \geq \sqrt{n}$  and alternatives are such that  $\varepsilon(n) \gg (n\lambda_n)^{-1/3} \log^{2/3}(N^2/n)$ ; this family may contain the alternatives with  $\varepsilon(n) \geq (n\lambda_n^2)^{-1/4}$ .

Thus, IARE of the  $\chi^2$  test wrt  $h$ -tests in the case  $\lambda_n \rightarrow \infty$  but  $\lambda_n \ll N$  completely, and in the case  $\lambda_n \gg N$  for alternatives such that  $(n\lambda_n)^{-1/3} \leq \varepsilon(n) \ll (n\lambda_n^2)^{-1/4}$  remains open.

Progress in the study of IARE  $h$ -tests depends on the progress of the results on the probability of large deviations of symmetric statistics (1.2).

*Acknowledgements.* The author thanks the anonymous reviewer for his very friendly and valuable constructive comments that lead to a substantial improved version of the paper.

## REFERENCES

- [1] N.A.C. Cressie and T.R.C. Read, Multinomial goodness-of-fit tests. *J. Roy. Statist. Soc. Ser B* **46** (1984) 440–464.
- [2] B. Cmiel, T. Ingłot and T. Ledwina, Intermediate efficiency of some weighted goodness-of-fit statistics. *J. Nonparam. Stat.* **32** (2020) 667–703.
- [3] L. Holst, Asymptotic normality and efficiency for certain goodness-of-fit tests. *Biometrika* **59** (1972) 137–145.
- [4] T. Ingłot, Generalized intermediate efficiency of goodness-of-fit tests. *Math. Methods Statist.* **8** (1999) 487–509.

- [5] T. Inglot, Intermediate efficiency of tests under heavy-tailed alternatives. *Probab. Math. Statist.* **40** (2020) 331–348.
- [6] T. Inglot, T. Ledwina and B. Cmiel, Intermediate efficiency in nonparametric testing problems with an application to some weighted statistics. *ESAIM: PS* **23** (2019) 697–738.
- [7] G.I. Ivchenko and Y.I. Medvedev, Decomposable statistics and verifying of tests. Small sample case. *Theory Probab. Appl.* **23** (1978) 796–806.
- [8] G.I. Ivchenko and S.A. Mirakhmedov, Large deviations and intermediate efficiency of the decomposable statistics in multinomial scheme. *Math. Methods Statist.* **4** (1995) 294–311.
- [9] W.C.M. Kallenberg, Intermediate efficiency, theory and examples. *Ann. Statist.* **11** (1983) 1401–1420
- [10] W.C.M. Kallenberg, On moderate and large deviations in multinomial distributions. *Ann. Statist.* **13** (1985) 1554–1580.
- [11] S.A. Mirakhmedov, Randomized decomposable statistics in the scheme of independent allocating particles into boxes. *Discrete Math. Appl.* **2** (1992) 91–108.
- [12] S.M. Mirakhmedov, S.R. Jammalamadaka and B.M. Ibragim, On Edgeworth expansions in generalized urn models. *J. Theor. Probab.* **27** (2014) 725–753.
- [13] S.M. Mirakhmedov, The probabilities of large deviations associated with multinomial distribution. <http://arxiv.org/abs/2205.03317> [math.PR] (2022).
- [14] S.M. Mirakhmedov, On the asymptotic properties of a certain class of goodness of-fit tests associated with multinomial distributions. *Commun. Stat. Theory Methods* (2021) DOI: [10.1080/03610926.2021.1918168](https://doi.org/10.1080/03610926.2021.1918168).
- [15] Ya.Yu. Nikitin, Asymptotic efficiency of nonparametric tests. Cambridge University Press (1995).
- [16] M.P. Quine and J. Robinson, Efficiencies of  $\chi^2$  and likelihood ratio goodness-of-fit tests. *Ann. Stat.* **13** (1985) 727–742.

## Subscribe to Open (S2O)

A fair and sustainable open access model



This journal is currently published in open access under a Subscribe-to-Open model (S2O). S2O is a transformative model that aims to move subscription journals to open access. Open access is the free, immediate, online availability of research articles combined with the rights to use these articles fully in the digital environment. We are thankful to our subscribers and sponsors for making it possible to publish this journal in open access, free of charge for authors.

**Please help to maintain this journal in open access!**

Check that your library subscribes to the journal, or make a personal donation to the S2O programme, by contacting [subscribers@edpsciences.org](mailto:subscribers@edpsciences.org)

More information, including a list of sponsors and a financial transparency report, available at: <https://www.edpsciences.org/en/math-s2o-programme>