

# Bounds for the chi-square approximation of Friedman’s statistic by Stein’s method

ROBERT E. GAUNT<sup>1,a</sup> and GESINE REINERT<sup>2,b</sup>

<sup>1</sup>*Department of Mathematics, The University of Manchester, Oxford Road, Manchester M13 9PL, UK,*

<sup>a</sup>*[robert.gaunt@manchester.ac.uk](mailto:robert.gaunt@manchester.ac.uk)*

<sup>2</sup>*Department of Statistics, University of Oxford, 24-29 St. Giles’, Oxford OX1 3LB, UK, <sup>b</sup>[reinert@stats.ox.ac.uk](mailto:reinert@stats.ox.ac.uk)*

Friedman’s chi-square test is a non-parametric statistical test for  $r$  treatments across  $n$  trials to assess the null hypothesis that there is no treatment effect. We use Stein’s method with an exchangeable pair coupling to derive a bound on the distance between the distribution of Friedman’s statistic and its limiting chi-square distribution, measured using smooth test functions. Our bound is of the optimal order  $n^{-1}$ , and also has an optimal dependence on the parameter  $r$ , in that the bound tends to zero if and only if  $r/n \rightarrow 0$ . From this bound, we deduce a Kolmogorov distance bound that decays to zero under the weaker condition  $r^{1/2}/n \rightarrow 0$ .

*Keywords:* Stein’s method; Friedman’s statistic; chi-square approximation; rate of convergence; exchangeable pair

## 1. Introduction

### 1.1. Friedman’s statistic and main results

Friedman’s chi-square test [16] is a non-parametric statistical test that, given  $r \geq 2$  treatments across  $n$  independent trials, can be used to test the null hypothesis that there is no treatment effect against the general alternative. Suppose that for the  $i$ -th trial we have the ranking  $\pi_i(1), \dots, \pi_i(r)$ , where  $\pi_i(j) \in \{1, \dots, r\}$ , over the  $r$  treatments. Under the null hypothesis, the rankings are independent permutations  $\pi_1, \dots, \pi_n$ , with each permutation being equally likely. Let

$$S_j = \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \sum_{i=1}^n \rho_i(j), \quad (1.1)$$

where  $\rho_i(j) = \pi_i(j) - (r+1)/2$ . Then the Friedman chi-square statistic, given by

$$F_r = \sum_{j=1}^r S_j^2, \quad (1.2)$$

has mean  $r-1$  and is asymptotically  $\chi_{(r-1)}^2$  distributed under the null hypothesis.

The study of asymptotically chi-square distributed statistics has received much attention in the literature. For general results and application to Pearson’s statistic see [26], in which a  $O(n^{-1})$  Kolmogorov distance bound between Pearson’s statistic and the  $\chi_{(r-1)}^2$  distribution was obtained for the case of  $r \geq 6$  cell classifications, improving a result of [42]. Also, [3,41] have used Edgeworth expansions to study the rate of convergence of the more general power divergence family of statistics constructed from the multinomial distribution of degree  $r$  (which includes the Pearson, log-likelihood ratio and Freeman-Tukey statistics as special cases) to their  $\chi_{(r-1)}^2$  limits. Using Stein’s method [39], a  $O(n^{-1})$  bound for the rate of convergence for Pearson’s statistic for  $r \geq 2$  cell classifications was obtained by [20]. This

result has been generalised recently to cover family of power divergence statistics (the largest subclass for which finite sample bounds are possible) by [19]. Moreover, [1] used Stein's method to obtain a bound with  $O(n^{-1/2})$  convergence rate for the likelihood ratio statistic when the data are realisations of independent and identically distributed random elements. Stein's method has also been used to obtain error bounds for the multivariate normal approximation of vectors of quadratic forms [8,13].

To date, however, Friedman's statistic has received little attention in the literature. The best result known to us is by [27], which provides, for any  $r \geq 2$ , the following Kolmogorov distance bound

$$d_K(\mathcal{L}(F_r), \chi_{(r-1)}^2) := \sup_{z \geq 0} |\mathbb{P}(F_r \leq z) - \mathbb{P}(Y_{r-1} \leq z)| \leq C(r)n^{-r/(r+1)}, \quad (1.3)$$

where the (non-explicit) constant  $C(r)$  depends only on  $r$ , and  $Y_{r-1} \sim \chi_{(r-1)}^2$ . That there is little literature on distributional bounds for Friedman's statistic, and particularly no bound that has a good or even explicit dependence on  $r$ , may be down to the dependence structure of Friedman's statistic, which is more complicated than that of the aforementioned power divergence, Pearson and likelihood ratio statistics. To get a feel for this dependence structure, we note that whilst, for fixed  $j$ , the random variables  $\rho_1(j), \dots, \rho_n(j)$  are independent, the sums  $S_1, \dots, S_r$  are not independent; indeed,  $S_r = -\sum_{j=1}^{r-1} S_j$ .

In this paper, we use Stein's method to obtain bounds on the rate of convergence of Friedman's statistic to its limiting chi-square distribution. Stein's method is particularly well-suited to this problem, because through the use of coupling techniques it often allows one to treat even complicated dependence structures.

To state our error bounds, we need some notation. Denote by  $C_b^{j,k}(\mathbb{R}^+)$  the class of functions  $h : \mathbb{R}^+ \rightarrow \mathbb{R}$  for which  $h^{(k)}$  exists and the derivatives of order  $j, j+1, \dots, k$  are bounded. Note that a function  $h \in C_b^{j,k}(\mathbb{R}^+)$  need not be itself bounded. We denote the usual supremum norm of a function  $g : \mathbb{R}^+ \rightarrow \mathbb{R}$  by  $\|g\| = \|g\|_\infty = \sup_{x \in \mathbb{R}^+} |g(x)|$ .

Our main result is the following weak convergence theorem for smooth test functions with a bound of optimal order with respect to both parameters  $n$  and  $r$  for the  $\chi_{(r-1)}^2$  approximation of Friedman's statistic, which holds for all  $r \geq 2$ . The order  $n^{-1}$  rate is the same as has been obtained by [20] and [19] for the chi-square approximation of the Pearson and power divergence statistics.

**Theorem 1.1.** *Suppose  $n \geq 1$  and  $r \geq 2$ . Then, for  $h \in C_b^{1,3}(\mathbb{R}^+)$ , there exist universal positive constants  $C_1, C_2, C_3, C_4$  and  $C_5$ , independent of  $n$  and  $r$ , such that*

$$|\mathbb{E}[h(F_r)] - \chi_{(r-1)}^2 h| \leq \frac{r}{n} \left[ C_1 \|h'\| + \left( C_2 + \frac{C_3 r}{n} \right) \|h''\| + \left( C_4 + \frac{C_5 r}{n} \right) \|h^{(3)}\| \right], \quad (1.4)$$

where  $\chi_{(r-1)}^2 h$  denotes  $\mathbb{E}[h(Y_{r-1})]$  for  $Y_{r-1} \sim \chi_{(r-1)}^2$ . The  $r/n$  rate is optimal with respect to both  $n$  and  $r$ .

The use of Stein's method allows to obtain explicit values for the constants in Theorem 1.1. The explicit values which we obtain here are too large for practical use; they can be found in the technical report [21].

**Remark 1.2.** In Lemma 3.4, we prove that  $\mathbb{E}[F_r^2] = r^2 - 1 - 2(r-1)/n$ . As  $\mathbb{E}[Y_{r-1}^2] = r^2 - 1$  for  $Y_{r-1} \sim \chi_{(r-1)}^2$ , it follows that  $|\mathbb{E}[F_r^2] - \mathbb{E}[Y_{r-1}^2]| = 2(r-1)/n$ . Whilst the function  $h(x) = x^2$  is not in the class  $C_b^{1,3}(\mathbb{R}^+)$ , this lends plausibility to the optimality of the  $r/n$  rate. A rigorous justification is given in the proof of Theorem 1.1.

Bounds obtained by Stein's method are often stated using smooth test functions, as in Theorem 1.1, particularly if technical issues arise in controlling solutions to the Stein equations or when faster than  $O(n^{-1/2})$  convergence rates are sought; see, for example, [4,6,7,14,15,20,22,28]. Bounds for non-smooth functions can be used for the construction of confidence intervals; however, as noted by [6, p. 151], bounds for smooth functions may be more natural in theoretical settings, and, as is discussed in [22, pp. 937–938], working with smooth test functions may allow one to obtain improved error bounds that may not hold for non-smooth test functions.

**Remark 1.3.** The premise of smooth test functions is crucial, because  $O(n^{-1})$  bounds like that of Theorem 1.1 that are valid for all  $r \geq 2$  will in general not hold for non-smooth test functions. To see this, consider the single point test function  $h \equiv \chi_{\{0\}}$ . Suppose  $n = 2k$  and  $r = 2$ . Then  $F_2 = S_1^2 + S_2^2$ , where  $S_2 = -S_1$ , so that  $F_2 = 2S_1^2$ . As  $r = 2$ , we can write  $S_1 = n^{-1/2} \sum_{i=1}^n Y_i$ , where  $Y_1, \dots, Y_n$  are i.i.d. with  $Y_1 = \sqrt{2}\rho_1(1) \sim \text{Unif}\{-1/\sqrt{2}, 1/\sqrt{2}\}$ . Thus,

$$\mathbb{E}[h(F_2)] = \mathbb{P}(F_2 = 0) = \mathbb{P}(S_1 = 0) = \mathbb{P}\left(\sum_i Y_i = 0\right) = \binom{2k}{k} \left(\frac{1}{2}\right)^{2k} \approx \frac{1}{\sqrt{\pi k}} = \sqrt{\frac{2}{\pi n}},$$

where we used Stirling's approximation. Since  $\chi_{(1)}^2 h = \mathbb{P}(\chi_{(1)}^2 = 0) = 0$ , the Kolmogorov distance between the distribution of  $F_2$  and the  $\chi_{(1)}^2$  distribution cannot be of smaller order than  $n^{-1/2}$ , which is the order of the Kolmogorov distance bound (1.3) when  $r = 2$ .

Due to the special structure in the case  $r = 2$  which allows us to write  $F_2 = 2S_1^2$ , in which  $S_1$  is a sum of zero mean i.i.d. random variables, in the following proposition we are able to obtain bounds that improve on those of Theorem 1.1, in terms of small explicit numerical constants and weaker assumptions on the test functions  $h$ . These bounds follow easily from applying results of [18], and the short proof is given in Section 5. Recall that, for real-valued random variables  $X$  and  $Y$ , the Wasserstein distance between their distributions is given by  $d_W(\mathcal{L}(X), \mathcal{L}(Y)) = \sup_{h \in \mathcal{H}} |\mathbb{E}[h(X)] - \mathbb{E}[h(Y)]|$ , where  $\mathcal{H}$  is the class of Lipschitz functions with Lipschitz constant at most 1.

**Proposition 1.4.** *Suppose  $n \geq 1$ . Then*

$$d_W(\mathcal{L}(F_2), \chi_{(1)}^2) \leq \frac{1}{\sqrt{n}} \left( 87 + \frac{48}{\sqrt{n}} \right), \quad (1.5)$$

and, for  $h \in C_b^{1,2}(\mathbb{R}^+)$ ,

$$|\mathbb{E}[h(F_2)] - \chi_{(1)}^2 h| \leq \frac{1}{n} \left( 69 + \frac{43}{n} \right) \{ \|h'\| + \|h''\| \}. \quad (1.6)$$

By a standard argument for converting smooth test function bounds into Kolmogorov distance bounds (see [11], p. 48) we deduce the following Kolmogorov distance bound from (1.4). The proof is given in Section 5.

**Corollary 1.5.** *There exist a universal positive constant  $C_6$ , independent of  $n$  and  $r$ , such that*

$$d_K(\mathcal{L}(F_r), \chi_{(r-1)}^2) \leq \begin{cases} \frac{0.9496}{\sqrt{n}}, & r = 2, \\ \frac{C_6 r^{1/8}}{n^{1/4}}, & r \geq 3. \end{cases} \quad (1.7)$$

Bounds with explicit numerical constants can be found in the technical report [21]. For  $r = 2$ , the bound in Corollary 1.5 has an optimal dependence on  $n$  (see Remark 1.3); we can achieve this rate by exploiting the special structure of Friedman's statistic in the case  $r = 2$ . For  $r \geq 3$ , the rate of convergence of the bound is slower than the  $O(n^{-r/(r+1)})$  rate of [27], which is to be expected because in deriving it we apply a crude non-smooth test function approximation technique to the smooth test function bound (1.4). However, to the best of our knowledge, it is the first Kolmogorov distance bound in the literature with an explicit dependence on the parameter  $r$ . The bound (1.7) tends to zero if  $r^{1/2}/n \rightarrow 0$ ; a weaker condition than the  $r/n \rightarrow 0$  condition under which the bound (1.4) tends to zero. This can be understood because the Kolmogorov distance is scale invariant, whereas for real-valued random variables  $X$  and  $Y$  the quantity  $|\mathbb{E}[h(X)] - \mathbb{E}[h(Y)]|$  is not. From our bound (1.7), we deduce that  $F_r \rightarrow_d \chi_{(r-1)}^2$  if  $r^{1/2}/n \rightarrow 0$ . To the best of our knowledge, this is the weakest condition in the literature under which Friedman's statistic is known to converge to the  $\chi_{(r-1)}^2$  distribution if  $r = r(n)$  is allowed to grow with  $n$ . We have been not been able to establish whether the condition is optimal, but at the very least it is close to being optimal.

## 1.2. Discussion of methods and outline of the paper

The proof of Theorem 1.1 is long and quite involved, with delicate arguments required in particular to achieve the optimal dependence on the parameter  $r$ . Here we provide a summary of the key steps in the proof, which may prove useful in other problems.

First, we make a connection between the  $\chi_{(r-1)}^2$  and multivariate normal Stein equations. Converting the problem to one of multivariate normal approximation allows us to take advantage of the powerful machinery of Stein's method for multivariate normal approximation and is natural, because Friedman's statistic is formed as a function of a random vector  $\mathbf{S}$  that is asymptotically multivariate normally distributed. Whilst we convert the problem to one of multivariate normal approximation, we work with the solution of the  $\chi_{(r-1)}^2$  Stein equation, so that at the end of the proof we can apply the  $O(r^{-1})$  bound (2.6) on the solution, which has an optimal dependence on  $r$ .

In the second step, we construct an exchangeable pair coupling for Friedman's statistic, which is well suited to the dependence structure of Friedman's statistic. We then apply Stein's method of exchangeable pairs [37,40] to obtain an initial bound. As we seek a bound of order  $n^{-1}$ , we expand one term further than is the case in the widely-used general bound of [37]. The starting point for this expansion, Lemma 2.2, is only implicitly given in [37]; its explicit formulation here could be of general interest.

The third step of the proof involves bounding the remainder terms from the initial expansion. To achieve the order  $n^{-1}$  rate, we apply local approach couplings to allow us to exploit the fact that  $\mathbb{E}[\rho_i(j)^3] = 0$  for all  $i, j$ , in order to vanish some remainder terms. Local couplings are less well-suited to the dependence structure of Friedman's statistic than our exchangeable pair coupling, and we need to proceed carefully to not lose the optimal dependence on  $r$ . One of the key steps is the introduction of the random variables  $T_m = \sum_{l=1}^r S_l \rho_m(l)$ , in which we absorb a sum over  $r$  indices and take advantage of the fact that  $S_l$  can be decomposed into a sum of  $n$  random variables of which only one is dependent on  $\rho_m(l)$ . Thus, this paper provides a rare example for combining exchangeable pair couplings with local couplings. The introduction of similar such random variables may also be of interest in other problems in which bounds are sought with good dependence on all parameters.

The rest of the article is organised as follows. In Section 2, we introduce the necessary elements of Stein's method that we will use to prove Theorem 1.1, and make a connection between the chi-square and multivariate normal Stein equations. In Section 3, we prove some preliminary lemmas needed in the proof of Theorem 1.1; their proofs are given in the Supplementary Material [2]. Lemma 3.4 provides formulas for the mean and variance of Friedman's statistic. We prove Theorem 1.1 in Section 4. Finally, in Section 5, we prove Proposition 1.4 and Corollary 1.5.

## 2. Elements of Stein's method

In this section, we present some basic theory on Stein's method for chi-square and multivariate normal approximation. Originally developed for normal approximation by Charles Stein in 1972 [39], Stein's method has since been extended to many other distributions, such as the multinomial [30], exponential [9,35], gamma [14,20,31,34] and multivariate normal [5,25]. For a comprehensive overview of the literature and an outline of the basic method in the univariate case see for example [29]; for the multivariate case see [33].

At the heart of Stein's method lies a characterising equation (a differential operator for continuous distributions) known as the Stein equation. For the  $\chi_{(p)}^2$  distribution this characterising equation is given by (see [12,31]):

$$xf''(x) + \frac{1}{2}(p-x)f'(x) = h(x) - \chi_{(p)}^2 h, \quad (2.1)$$

where  $\chi_{(p)}^2 h$  denotes the quantity  $\mathbb{E}[h(Y_p)]$  for  $Y_p \sim \chi_{(p)}^2$ . Evaluating both sides at a random variable of interest  $W$  and taking expectations then gives

$$\mathbb{E}[Wf''(W) + \frac{1}{2}(p-W)f'(W)] = \mathbb{E}[h(W)] - \chi_{(p)}^2 h. \quad (2.2)$$

Thus, the quantity  $|\mathbb{E}[h(W)] - \chi_{(p)}^2 h|$  can be bounded by solving the Stein equation (2.1) for  $f$  and then bounding the left-hand side of (2.2). It can easily be seen that

$$f'(x) = \frac{e^{x/2}}{x^{p/2}} \int_0^x t^{p/2-1} e^{-t/2} [h(t) - \chi_{(p)}^2 h] dt \quad (2.3)$$

solves (2.1). For  $h$  belonging to the classes  $C_b^{k,k}(\mathbb{R}^+)$ ,  $C_b^{k-1,k-1}(\mathbb{R}^+)$  and  $C_b^{k-1,k-2}(\mathbb{R}^+)$ , respectively, setting  $h^{(0)} \equiv h$ , the following bounds hold (see [17,31], which improves on a bound of [36], and [20] respectively):

$$\|f^{(k)}\| \leq \frac{2}{k} \|h^{(k)}\|, \quad k \geq 1, \quad (2.4)$$

$$\|f^{(k)}\| \leq \left\{ \frac{2\sqrt{\pi} + \sqrt{2}e^{-1}}{\sqrt{p+2k-2}} + \frac{4}{p+2k-2} \right\} \|h^{(k-1)}\|, \quad k \geq 1, \quad (2.5)$$

$$\|f^{(k)}\| \leq \frac{4}{p+2k-2} \{3\|h^{(k-1)}\| + 2\|h^{(k-2)}\|\}, \quad k \geq 2. \quad (2.6)$$

With these bounds for the solution at our disposal, we could follow the conventional approach to Stein's method and bound the distance between the distribution of Friedman's statistic  $F_r$  and its limiting  $\chi_{(r-1)}^2$  distribution by bounding the left-hand side of (2.2). Instead, however, we elect to follow [20] and use the multivariate normal Stein equation in conjunction with the chi-square Stein equation. Indeed, there is powerful array of tools for proving approximation theorems using the multivariate normal Stein equation (see [10,23,24,32,37] for coupling techniques for multivariate normal approximation). Let  $\Sigma$  be non-negative definite. Then the  $\text{MVN}(\mathbf{0}, \Sigma)$  Stein equation (see [24]) is

$$\nabla^\top \Sigma \nabla f(\mathbf{w}) - \mathbf{w}^\top \nabla f(\mathbf{w}) = h(\mathbf{w}) - \mathbb{E}[h(\Sigma^{1/2} \mathbf{Z})]. \quad (2.7)$$

We now obtain a connection between the Stein equations for the  $\chi_{(r-1)}^2$  and  $\text{MVN}(\mathbf{0}, \Sigma_S)$  distributions. The lemma can be read off from Lemma 4.2 of [20], because the covariance matrix  $\Sigma_S$  is equal,

up to a multiplicative factor, to the covariance matrix of the random vector  $\mathbf{U} = (U_1, \dots, U_r)^\top$  of the observed counts in Pearson's statistic under the null hypothesis of uniform classification probabilities (see Lemma 4.1 of [20] and Lemma 3.1 below).

**Lemma 2.1.** *Let  $f \in C^2(\mathbb{R})$  and define  $g : \mathbb{R}^r \rightarrow \mathbb{R}$  by  $g(\mathbf{s}) = f(w)/4$  with  $w = \sum_{j=1}^r s_j^2$  for  $\mathbf{s} = (s_1, \dots, s_r)^\top$ . Let  $F_r$  and  $\mathbf{S} = (S_1, \dots, S_r)^\top$  be defined as in (1.2) and (1.1). Let  $\Sigma_{\mathbf{S}}$  denote the covariance matrix of  $\mathbf{S}$ . Then*

$$\mathbb{E}[\nabla^\top \Sigma_{\mathbf{S}} \nabla g(\mathbf{S}) - \mathbf{S}^\top \nabla g(\mathbf{S})] = \mathbb{E}[F_r f''(F_r) + \frac{1}{2}(r-1-F_r)f'(F_r)].$$

To bound  $\mathbb{E}[\nabla^\top \Sigma_{\mathbf{S}} \nabla g(\mathbf{S}) - \mathbf{S}^\top \nabla g(\mathbf{S})]$  we use the exchangeable pair coupling approach of [37]. A pair  $(X, X')$  of random variables defined on the same probability space is called *exchangeable* if  $\mathbb{P}(X \in B, X' \in B') = \mathbb{P}(X \in B', X' \in B)$  for all measurable sets  $B$  and  $B'$ . We shall use the following lemma, which combines equations (2.5) and (2.6) in the proof of Theorem 2.1 of [37]. It is worth noting that up to this part of their proof the authors had only required that  $\Sigma$  be non-negative definite.

**Lemma 2.2.** *Let  $\mathbf{W} = (W_1, \dots, W_d)^\top \in \mathbb{R}^d$ . Assume  $(\mathbf{W}, \mathbf{W}')$  is an exchangeable pair of  $\mathbb{R}^d$ -valued random vectors such that  $\mathbb{E}[\mathbf{W}] = \mathbf{0}$ ,  $\mathbb{E}[\mathbf{W}\mathbf{W}^\top] = \Sigma$ . Suppose further that*

$$\mathbb{E}^{\mathbf{W}}[\mathbf{W}' - \mathbf{W}] = -\Lambda \mathbf{W} \quad (2.8)$$

for an invertible  $d \times d$  matrix  $\Lambda$ . Then, provided  $f \in C^3(\mathbb{R}^d)$ ,

$$\mathbb{E}[\mathbf{W}^\top \nabla f(\mathbf{W})] = \frac{1}{2} \mathbb{E}[(\mathbf{W}' - \mathbf{W})^\top \Lambda^{-\top} (\nabla f(\mathbf{W}') - \nabla f(\mathbf{W}))].$$

### 3. Preliminary lemmas

Here we state some technical lemmas; their straightforward proofs are in the Supplementary Material.

**Lemma 3.1.** *[Covariances] Let  $\mathbf{S} = (S_1, \dots, S_r)^\top$ , where the  $S_j$  are defined in (1.1). The (non-negative definite) covariance matrix of  $\mathbf{S}$ , denoted by  $\Sigma = (\sigma_{jk})$ , has entries*

$$\sigma_{jj} = \frac{r-1}{r} \quad \text{and} \quad \sigma_{jk} = -\frac{1}{r} \quad (j \neq k). \quad (3.1)$$

**Lemma 3.2.** *[Mixed moments of  $\rho_m$ ] Let  $r \geq 2$ . Then, for distinct arguments,*

$$\begin{aligned} \mathbb{E}[\rho_m(l)] &= \mathbb{E}[\rho_m(l)^3] = \mathbb{E}[\rho_m(l)^2 \rho_m(j)] = \mathbb{E}[\rho_m(l) \rho_m(j) \rho_m(s)] = 0, \\ \mathbb{E}[\rho_m(l)^2] &= \frac{r^2-1}{12}, \quad \mathbb{E}[\rho_m(l) \rho_m(j)] = -\frac{r+1}{12}, \end{aligned}$$

and the following bounds hold

$$\mathbb{E}[\rho_m(l)^4] \leq \frac{r^4}{80}, \quad |\mathbb{E}[\rho_m(l)^3 \rho_m(j)]| \leq \frac{r^3}{80}, \quad \mathbb{E}[\rho_m(l)^2 \rho_m(j)^2] \leq \frac{r^4}{144}.$$

Also, for  $r \geq 3$  and  $r \geq 4$ , respectively,

$$|\mathbb{E}[\rho_m(l)^2 \rho_m(j) \rho_m(s)]| \leq \frac{r^3}{144}, \quad |\mathbb{E}[\rho_m(l) \rho_m(j) \rho_m(s) \rho_m(t)]| = \frac{(r+1)(5r+7)}{240}.$$

Finally, for  $j = 1, \dots, r$ ,

$$\mathbb{E}[S_j^4] \leq 3. \quad (3.2)$$

**Lemma 3.3.** *Let  $T_m = \sum_{l=1}^r S_l \rho_m(l)$ . Then, for  $n, r \geq 2$ ,*

$$\mathbb{E}[T_m^2] \leq \frac{r^3}{12} \left(1 + \frac{r}{n}\right), \quad (3.3)$$

$$\mathbb{E}[T_m^4] \leq \left(\frac{7}{48} + \frac{r^2}{36n^2} + \frac{1}{5n}\right) r^6 =: r^6 C_T(r, n). \quad (3.4)$$

**Lemma 3.4.** *Let  $F_r$  denote Friedman's statistic. Then, for  $n \geq 1$  and  $r \geq 2$ ,*

$$\mathbb{E}[F_r] = r - 1, \quad (3.5)$$

$$\mathbb{E}[F_r^2] = r^2 - 1 - \frac{2(r-1)}{n}, \quad \text{and} \quad \text{Var}(F_r) = 2(r-1) \left(1 - \frac{1}{n}\right). \quad (3.6)$$

The next lemma provides bounds on further expectations that arise in the proof of Theorem 1.1. The proof is omitted as all expectations are easily bounded to the correct order with respect to  $r$  and  $n$  using that  $\mathbb{E}[S_j^2] \leq 1$ , (3.2), as well as the crude inequalities

$$|\rho_m(k)| \leq (r+1)/2 \quad \text{and} \quad |\rho_m(k) - \rho_m(l)| \leq (r-1), \quad k \neq l. \quad (3.7)$$

**Lemma 3.5.** *Let  $C$  be a positive universal constant, independent of  $n$  and  $r$ , that may vary from one expression to the next. Let  $r \geq 2$  and  $j, k \in \{1, \dots, r\}$ . Then the following inequalities hold.*

$$\sum_{l=1}^r \mathbb{E}[S_k^2 (\rho_m(l) - \rho_m(k))^4] \leq Cr^5, \quad (3.8)$$

$$\sum_{l=1}^r \mathbb{E}[S_k^4 (\rho_m(l) - \rho_m(k))^4] \leq Cr^5, \quad \sum_{l=1}^r \mathbb{E}[S_k^2 (\rho_m(l) - \rho_m(k))^6] \leq Cr^7, \quad (3.9)$$

$$\mathbb{E} \left[ \left( \frac{6}{r(r+1)} (\rho_m(k) - \rho_m(l))^2 - 1 \right)^2 \right] \leq C. \quad (3.10)$$

## 4. Proof of Theorem 1.1

**Part I: Exchangeable pair coupling and initial expansion.** We suppose  $r \geq 3$ , because Proposition 1.4 provides an explicit bound on  $|\mathbb{E}[h(F_2)] - \chi_{(1)}^2 h|$  in the case  $r = 2$ . This assumption allows us to simplify some calculations; all probabilistic arguments apply equally well for all  $r \geq 2$  and  $n \geq 2$ . In this proof,  $C$  will denote a positive universal constant, independent of  $n$  and  $r$ , that may vary from one expression to the next.

We begin by presenting our exchangeable pair coupling for Friedman's statistic. Pick an index  $M \in \{1, \dots, n\}$  uniformly and independent indices  $K, L \in \{1, \dots, r\}$  uniformly and independently of  $M$ . If  $M = m$ ,  $K = k$ ,  $L = l$ , define the permutation  $\pi'_m$  by

$$\pi'_m(k) = \pi_m(l), \quad \pi'_m(l) = \pi_m(k), \quad \pi'_m(i) = \pi_m(i), \quad i \neq k, l.$$

We then put

$$\begin{aligned} S'_K &= S_K - \frac{\sqrt{12}}{\sqrt{r(r+1)n}}[\rho_M(K) - \rho_M(L)], \\ S'_L &= S_L - \frac{\sqrt{12}}{\sqrt{r(r+1)n}}[\rho_M(L) - \rho_M(K)], \end{aligned} \quad (4.1)$$

and  $S'_j = S_j$  for  $j \neq K, L$ . Then  $(\mathbf{S}, \mathbf{S}')$  is an exchangeable pair, and we now verify condition (2.8):

$$\begin{aligned} \mathbb{E}^{\mathbf{S}}[S'_j - S_j] &= \frac{1}{r^2 n} \sum_{k,l=1}^r \sum_{m=1}^n \mathbb{E}^{\mathbf{S}}[S'_j - S_j \mid K = k, L = l, M = m] \\ &= \frac{1}{r^2 n} \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \left\{ \sum_{k=1}^r \sum_{m=1}^n \mathbb{E}^{\mathbf{S}}[\rho_m(k) - \rho_m(j)] + \sum_{l=1}^r \sum_{m=1}^n \mathbb{E}^{\mathbf{S}}[\rho_m(l) - \rho_m(j)] \right\} \\ &= -\frac{2}{rn} \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \sum_{m=1}^n \mathbb{E}^{\mathbf{S}}[\rho_m(j)] = -\frac{2}{rn} S_j, \end{aligned}$$

where we used that  $\sum_{k=1}^r \rho_m(k) = 0$  to obtain the third equality. Therefore condition (2.8) holds with  $\Lambda = (2/rn)\mathbf{I}_r$ , where  $\mathbf{I}_r$  is the  $r \times r$  identity matrix. Hence,  $\Lambda^{-\top} = (rn/2)\mathbf{I}_r$ .

Having established an appropriate exchangeable pair coupling, we are in a position to bound the quantity  $|\mathbb{E}[h(F_r)] - \chi^2_{(r-1)} h|$ . Now, by Lemma 2.1,

$$\mathbb{E}[h(F_r)] - \chi^2_{(r-1)} h = \mathbb{E}[F_r f''(F_r) + \frac{1}{2}(r-1-F_r)f'(F_r)] = \mathbb{E}[\nabla^\top \Sigma \nabla g(\mathbf{S}) - \mathbf{S}^\top \nabla g(\mathbf{S})],$$

where  $g: \mathbb{R}^r \rightarrow \mathbb{R}$  is defined by  $g(\mathbf{S}) = f(F_r)/4$  and the covariance matrix  $\Sigma$  is given in Lemma 3.1. With Lemma 2.2 and Taylor expansion,

$$\begin{aligned} \mathbb{E}[\mathbf{S}^\top \nabla g(\mathbf{S})] &= \frac{1}{2} \frac{rn}{2} \mathbb{E}[(\mathbf{S}' - \mathbf{S})^\top (\nabla g(\mathbf{S}') - \nabla g(\mathbf{S}))] \\ &= \frac{rn}{4} \sum_{j,u=1}^r \mathbb{E} \left[ \frac{\partial^2}{\partial s_j \partial s_u} g(\mathbf{S}) \mathbb{E}^{\mathbf{S}}[(S'_j - S_j)(S'_u - S_u)] \right] \\ &\quad + \frac{rn}{8} \sum_{j,u,v=1}^r \mathbb{E} \left[ \frac{\partial^3}{\partial s_j \partial s_u \partial s_v} g(\mathbf{S}) \mathbb{E}^{\mathbf{S}}[(S'_j - S_j)(S'_u - S_u)(S'_v - S_v)] \right] \\ &\quad + \frac{rn}{24} \sum_{j,u,v,w=1}^r \mathbb{E} \left[ \frac{\partial^4}{\partial s_j \partial s_u \partial s_v \partial s_w} g(\mathbf{S}_\theta^*) \mathbb{E}^{\mathbf{S}}[(S'_j - S_j)(S'_u - S_u)(S'_v - S_v)(S'_w - S_w)] \right], \end{aligned}$$

where the  $j$ -th component of  $\mathbf{S}_\theta^*$  is given by  $\theta_j S_j + (1 - \theta_j) S'_j$  for some  $\theta_j = \theta_j(\mathbf{S}, \mathbf{S}') \in (0, 1)$ . As  $\mathbb{E}[(S'_j - S_j)(S'_u - S_u)] = 4\sigma_{ju}/(rn)$  (see Equation (2.10) in [37]) we obtain that

$$\mathbb{E}[h(F_r)] - \chi^2_{(r-1)} h = \mathbb{E}[\nabla^\top \Sigma \nabla g(\mathbf{S}) - \mathbf{S}^\top \nabla g(\mathbf{S})] = R_0 + R_1 + R_2,$$



with

$$\begin{aligned}
 R_0 &= \frac{rn}{4} \sum_{j,u=1}^r \mathbb{E} \left[ \frac{\partial^2}{\partial s_j \partial s_u} g(\mathbf{S}) \left( \mathbb{E}[(S'_j - S_j)(S'_u - S_u)] - (S'_j - S_j)(S'_u - S_u) \right) \right], \\
 R_1 &= -\frac{rn}{8} \sum_{j,u,v=1}^r \mathbb{E} \left[ \frac{\partial^3}{\partial s_j \partial s_u \partial s_v} g(\mathbf{S}) \mathbb{E}^{\mathbf{S}}[(S'_j - S_j)(S'_u - S_u)(S'_v - S_v)] \right], \\
 R_2 &= -\frac{rn}{24} \sum_{j,u,v,w=1}^r \mathbb{E} \left[ \frac{\partial^4}{\partial s_j \partial s_u \partial s_v \partial s_w} g(\mathbf{S}^*_\theta) \mathbb{E}^{\mathbf{S}}[(S'_j - S_j)(S'_u - S_u)(S'_v - S_v)(S'_w - S_w)] \right].
 \end{aligned}$$

Recalling that

$$g(\mathbf{s}) = f(h(\mathbf{s}))/4 \text{ with } h(\mathbf{s}) = \sum_{j=1}^r s_j^2$$

so that  $\frac{\partial}{\partial s_i} h(\mathbf{s}) = 2s_i$  and  $\frac{\partial^2}{(\partial s_i)^2} h(\mathbf{s}) = 2$  and all other second order partial derivatives and all partial derivatives of  $h$  of order 3 or higher vanish,

$$\begin{aligned}
 \frac{\partial^4}{\partial s_j \partial s_u \partial s_v \partial s_w} g(\mathbf{s}) &= 4f^{(4)}(h(\mathbf{s}))s_j s_u s_v s_w \\
 &\quad + 2f^{(3)}(h(\mathbf{s}))(s_j s_u \mathbf{1}(v=w) + s_j s_v \mathbf{1}(u=w) \\
 &\quad + s_j s_w \mathbf{1}(u=v) + s_u s_v \mathbf{1}(j=w) + s_u s_w \mathbf{1}(j=v) + s_v s_w \mathbf{1}(j=u)) \\
 &\quad + f''(h(\mathbf{s}))(\mathbf{1}(v=w, j=u) + \mathbf{1}(u=w, j=v) + \mathbf{1}(u=v, j=w)). \tag{4.2}
 \end{aligned}$$

The rest of the proof is devoted to bounding the remainders  $R_0$ ,  $R_1$  and  $R_2$ , in reverse order.

**Part II: Bounding  $R_2$ .** We observe that  $(S'_j - S_j)(S'_u - S_u)(S'_v - S_v) = 0$  unless at least two indices match. Moreover, if  $K = k$ ,  $L = l$ ,  $M = m$  then

$$\begin{aligned}
 &(S'_j - S_j)(S'_u - S_u)(S'_v - S_v)(S'_w - S_w) \\
 &= \frac{144}{r^2(r+1)^2 n^2} (\rho_m(l) - \rho_m(k))^4 (\mathbf{1}\{j = u = v = w \in \{k, l\}\} \\
 &\quad + \mathbf{1}\{\{j, u, v, w\} = \{k, l\} \text{ with one of them appearing three times } \} \\
 &\quad + \mathbf{1}\{\{j, u, v, w\} = \{k, l\} \text{ with each appearing twice}\}). \tag{4.3}
 \end{aligned}$$

Weaving in the derivative (4.2) yields

$$R_2 = R_{2,1} + R_{2,2} + R_{2,3},$$

where

$$\begin{aligned}
 R_{2,1} &= -\frac{rn}{6} \sum_{j,u,v,w=1}^r \mathbb{E} \left[ f^{(4)}(h(\mathbf{S}^*_\theta)) (\theta_j S_j + (1 - \theta_j) S'_j) (\theta_u S_u + (1 - \theta_u) S'_u) \right. \\
 &\quad \left. (\theta_v S_v + (1 - \theta_v) S'_v) (\theta_w S_w + (1 - \theta_w) S'_w) (S'_j - S_j)(S'_u - S_u)(S'_v - S_v)(S'_w - S_w) \right],
 \end{aligned}$$

and

$$\begin{aligned}
 R_{2,2} = & -\frac{rn}{12} \sum_{j,u,v,w=1}^r \mathbb{E} \left[ f^{(3)}(h(\mathbf{S}_\theta^*)) \left\{ (\theta_j S_j + (1 - \theta_j) S'_j)(\theta_u S_u + (1 - \theta_u) S'_u) \mathbf{1}(v = w) \right. \right. \\
 & + (\theta_j S_j + (1 - \theta_j) S'_j)(\theta_v S_v + (1 - \theta_v) S'_v) \mathbf{1}(u = w) \\
 & + (\theta_j S_j + (1 - \theta_j) S'_j)(\theta_w S_w + (1 - \theta_w) S'_w) \mathbf{1}(u = v) \\
 & + (\theta_u S_u + (1 - \theta_u) S'_u)(\theta_v S_v + (1 - \theta_v) S'_v) \mathbf{1}(j = w) \\
 & + (\theta_u S_u + (1 - \theta_u) S'_u)(\theta_w S_w + (1 - \theta_w) S'_w) \mathbf{1}(j = v) \\
 & + (\theta_v S_v + (1 - \theta_v) S'_v)(\theta_w S_w + (1 - \theta_w) S'_w) \mathbf{1}(j = u) \left. \right\} \\
 & \left. (S'_j - S_j)(S'_u - S_u)(S'_v - S_v)(S'_w - S_w) \right],
 \end{aligned}$$

and

$$\begin{aligned}
 R_{2,3} = & -\frac{rn}{24} \sum_{j,u,v,w=1}^r \mathbb{E} \left[ f''(h(\mathbf{S}_\theta^*)) (\{ \mathbf{1}(v = w, j = u) + \mathbf{1}(u = w, j = v) \right. \\
 & \left. + \mathbf{1}(u = v, j = w) \} (S'_j - S_j)(S'_u - S_u)(S'_v - S_v)(S'_w - S_w)) \right].
 \end{aligned}$$

For the term  $R_{2,1}$  we condition on  $K = k, L = l, M = m$  and use (4.3) to obtain

$$\begin{aligned}
 R_{2,1} = & -\frac{rn}{6} \frac{1}{r^2 n} \frac{144}{r^2 (r+1)^2 n^2} \sum_{j,u,v,w,k,l=1}^r \sum_{m=1}^n \mathbb{E} \left[ f^{(4)}(h(\mathbf{S}_\theta^*)) \right. \\
 & (\theta_j S_j + (1 - \theta_j) S'_j)(\theta_u S_u + (1 - \theta_u) S'_u)(\theta_v S_v + (1 - \theta_v) S'_v)(\theta_w S_w + (1 - \theta_w) S'_w) \\
 & (\rho_m(l) - \rho_m(k))^4 (\mathbf{1}\{j = u = v = w \in \{k, l\}\} \\
 & + \mathbf{1}\{j, u, v, w\} = \{k, l\} \text{ with one of them appearing three times } \} \\
 & \left. + \mathbf{1}\{j, u, v, w\} = \{k, l\} \text{ with each appearing twice} \} \mid K = k, L = l, M = m \right] \\
 = & R_{2,1,1} + R_{2,1,2} + R_{2,1,3}.
 \end{aligned}$$

For  $R_{2,1,1}$  we have

$$\begin{aligned}
 R_{2,1,1} = & -\frac{24}{r^3 (r+1)^2 n^2} \sum_{j,u,v,w,k,l=1}^r \sum_{m=1}^n \mathbb{E} \left[ f^{(4)}(h(\mathbf{S}_\theta^*)) \right. \\
 & (\theta_j S_j + (1 - \theta_j) S'_j)(\theta_u S_u + (1 - \theta_u) S'_u)(\theta_v S_v + (1 - \theta_v) S'_v)(\theta_w S_w + (1 - \theta_w) S'_w) \\
 & \left. (\rho_m(l) - \rho_m(k))^4 \mathbf{1}\{j = u = v = w \in \{k, l\}\} \mid K = k, L = l, M = m \right] \\
 = & -\frac{48}{r^3 (r+1)^2 n^2} \sum_{k,l=1}^r \sum_{m=1}^n \mathbb{E} \left[ f^{(4)}(h(\mathbf{S}_\theta^*)) \right. \\
 & \left. (\theta_k S_k + (1 - \theta_k) S'_k)^4 (\rho_m(l) - \rho_m(k))^4 \mid K = k, L = l, M = m \right]
 \end{aligned}$$

$$= -\frac{48}{r^3(r+1)^2n^2} \sum_{k,l=1}^r \sum_{m=1}^n \mathbb{E} \left[ f^{(4)}(h(\mathbf{S}_\theta^*)) \left( S_k + \theta_k \frac{\sqrt{12}}{\sqrt{r(r+1)n}} (\rho_m(l) - \rho_m(k)) \right)^4 \right. \\ \left. (\rho_m(l) - \rho_m(k))^4 \mid K = k, L = l, M = m \right], \quad (4.4)$$

so that, on using the basic inequality  $(a+b)^4 \leq 3a^4 + 10a^2b^2 + 3b^4$ ,

$$|R_{2,1,1}| \leq \frac{48\|f^{(4)}\|}{r^3(r+1)^2n^2} \sum_{k,l=1}^r \sum_{m=1}^n \left\{ 3\mathbb{E}[S_k^4(\rho_m(l) - \rho_m(k))^4] \right. \\ \left. + \frac{10 \times 12}{r(r+1)n} \mathbb{E}[S_k^2(\rho_m(l) - \rho_m(k))^6] + \frac{3 \times 144}{r^2(r+1)^2n^2} \mathbb{E}[(\rho_m(l) - \rho_m(k))^8] \right\}.$$

By (3.7),  $\sum_{k,l=1}^r (\rho_m(l) - \rho_m(k))^8 \leq r^{10}$ , and with this formula and the inequalities in (3.9) we get that

$$|R_{2,1,1}| \leq \frac{48\|f^{(4)}\|}{r^3(r+1)^2n} \left\{ 3 \times Cr^6 + \frac{120}{r(r+1)n} \times Cr^8 + \frac{432}{r^2(r+1)^2n^2} \times r^{10} \right\} \leq \frac{Cr\|f^{(4)}\|}{n}.$$

Next,

$$R_{2,1,2} = \frac{24}{r^3(r+1)^2n^2} \sum_{j,u,v,w,k,l=1}^r \sum_{m=1}^n \mathbb{E} \left[ f^{(4)}(h(\mathbf{S}_\theta^*)) (\theta_j S_j + (1 - \theta_j) S'_j) \right. \\ \left. (\theta_u S_u + (1 - \theta_u) S'_u)(\theta_v S_v + (1 - \theta_v) S'_v)(\theta_w S_w + (1 - \theta_w) S'_w)(\rho_m(l) - \rho_m(k))^4 \right. \\ \left. \mathbf{1}\{\{j, u, v, w\} = \{k, l\} \text{ with one of them appearing three times}\} \mid K = k, L = l, M = m \right] \\ = \frac{192}{r^3(r+1)^2n^2} \sum_{k,l=1}^r \sum_{m=1}^n \mathbb{E} \left[ f^{(4)}(h(\mathbf{S}_\theta^*)) (\theta_k S_k + (1 - \theta_k) S'_k)^3 (\theta_l S_l + (1 - \theta_l) S'_l) \right. \\ \left. (\rho_m(l) - \rho_m(k))^4 \mid K = k, L = l, M = m \right].$$

From the basic inequality  $a^3b \leq \frac{3}{4}a^4 + \frac{1}{4}b^4$  we now obtain

$$|R_{2,1,2}| \leq \frac{48\|f^{(4)}\|}{r^3(r+1)^2n^2} \sum_{k,l=1}^r \sum_{m=1}^n \mathbb{E} \left[ \left\{ 3 \left( S_k + \theta_k \frac{\sqrt{12}}{\sqrt{r(r+1)n}} (\rho_m(l) - \rho_m(k)) \right)^4 \right. \right. \\ \left. \left. + \left( S_l - \theta_l \frac{\sqrt{12}}{\sqrt{r(r+1)n}} (\rho_m(l) - \rho_m(k)) \right)^4 \right\} (\rho_m(l) - \rho_m(k))^4 \right].$$

Bounding as we did for  $R_{2,1,1}$  gives that  $|R_{2,1,2}| \leq \frac{Cr\|f^{(4)}\|}{n}$ . Also,

$$R_{2,1,3} = -\frac{24}{r^3(r+1)^2n^2} \sum_{j,u,v,w,k,l=1}^r \sum_{m=1}^n \mathbb{E} \left[ f^{(4)}(h(\mathbf{S}_\theta^*)) (\theta_j S_j + (1 - \theta_j) S'_j) \right. \\ \left. (\theta_u S_u + (1 - \theta_u) S'_u)(\theta_v S_v + (1 - \theta_v) S'_v)(\theta_w S_w + (1 - \theta_w) S'_w)(\rho_m(l) - \rho_m(k))^4 \right. \\ \left. \mathbf{1}\{\{j, u, v, w\} = \{k, l\} \text{ with each appearing twice}\} \mid K = k, L = l, M = m \right].$$

From the basic inequality  $ab \leq (a^2 + b^2)/2$  we now obtain

$$\begin{aligned}
 |R_{2,1,3}| &\leq \frac{144\|f^{(4)}\|}{r^3(r+1)^2n^2} \sum_{k,l=1}^r \sum_{m=1}^n \mathbb{E} \left[ \left( S_k + \theta_k \frac{\sqrt{12}}{\sqrt{r(r+1)n}} (\rho_m(l) - \rho_m(k)) \right)^2 \right. \\
 &\quad \left. \left( S_l - \theta_l \frac{\sqrt{12}}{\sqrt{r(r+1)n}} (\rho_m(l) - \rho_m(k)) \right)^2 (\rho_m(l) - \rho_m(k))^4 \right] \\
 &\leq \frac{72\|f^{(4)}\|}{r^3(r+1)^2n^2} \sum_{k,l=1}^r \sum_{m=1}^n \mathbb{E} \left[ \left\{ \left( S_k + \theta_k \frac{\sqrt{12}}{\sqrt{r(r+1)n}} (\rho_m(l) - \rho_m(k)) \right)^4 \right. \right. \\
 &\quad \left. \left. + \left( S_l - \theta_l \frac{\sqrt{12}}{\sqrt{r(r+1)n}} (\rho_m(l) - \rho_m(k)) \right)^4 \right\} (\rho_m(l) - \rho_m(k))^4 \right].
 \end{aligned}$$

Bounding as we did for  $R_{2,1,1}$  gives that  $|R_{2,1,3}| \leq \frac{Cr\|f^{(4)}\|}{n}$ . As overall bound for  $R_{2,1}$  we obtain

$$|R_{2,1}| \leq \frac{Cr\|f^{(4)}\|}{n}. \quad (4.5)$$

For  $R_{2,2}$ , with (4.3),

$$\begin{aligned}
 R_{2,2} &= -\frac{12}{r^3(r+1)^2n^2} \sum_{j,u,v,w,k,l=1}^r \sum_{m=1}^n \mathbb{E} \left[ f^{(3)}(h(\mathbf{S}_\theta^*)) \right. \\
 &\quad \left\{ (\theta_j S_j + (1 - \theta_j) S'_j)(\theta_u S_u + (1 - \theta_u) S'_u) \mathbf{1}(v = w) \right. \\
 &\quad + (\theta_j S_j + (1 - \theta_j) S'_j)(\theta_v S_v + (1 - \theta_v) S'_v) \mathbf{1}(u = w) \\
 &\quad + (\theta_j S_j + (1 - \theta_j) S'_j)(\theta_w S_w + (1 - \theta_w) S'_w) \mathbf{1}(u = v) \\
 &\quad + (\theta_u S_u + (1 - \theta_u) S'_u)(\theta_v S_v + (1 - \theta_v) S'_v) \mathbf{1}(j = w) \\
 &\quad + (\theta_u S_u + (1 - \theta_u) S'_u)(\theta_w S_w + (1 - \theta_w) S'_w) \mathbf{1}(j = v) \\
 &\quad + (\theta_v S_v + (1 - \theta_v) S'_v)(\theta_w S_w + (1 - \theta_w) S'_w) \mathbf{1}(j = u) \left. \right\} (\rho_m(l) - \rho_m(k))^4 \\
 &\quad (\mathbf{1}\{j = u = v = w \in \{k, l\}\} + \mathbf{1}\{\{j, u, v, w\} = \{k, l\} \text{ with one of them appearing three times}\} \\
 &\quad + \mathbf{1}\{\{j, u, v, w\} = \{k, l\} \text{ with each appearing twice}\}) |K = k, L = l, M = m] \\
 &= -\frac{12}{r^3(r+1)^2n^2} \sum_{k,l=1}^r \sum_{m=1}^n \mathbb{E} \left[ f^{(3)}(h(\mathbf{S}_\theta^*)) \left\{ ((2 \times 6) + (2 \times 6))(\theta_k S_k + (1 - \theta_k) S'_k)^2 \right. \right. \\
 &\quad \left. \left. + (3 \times 8)(\theta_k S_k + (1 - \theta_k) S'_k)(\theta_l S_l + (1 - \theta_l) S'_l) \right\} (\rho_m(l) - \rho_m(k))^4 |K = k, L = l, M = m \right].
 \end{aligned}$$

Using the basic inequality  $ab \leq (a^2 + b^2)/2$  now gives that

$$|R_{2,2}| \leq \frac{576\|f^{(3)}\|}{r^3(r+1)^2n^2} \sum_{k,l=1}^r \sum_{m=1}^n \mathbb{E} \left[ \left( S_k + \theta_k \frac{\sqrt{12}}{\sqrt{r(r+1)n}} (\rho_m(l) - \rho_m(k)) \right)^2 (\rho_m(l) - \rho_m(k))^4 \right].$$

Bounding similarly to before now gives the bound

$$|R_{2,2}| \leq \frac{Cr \|f^{(3)}\|}{n}. \quad (4.6)$$

For  $R_{2,3}$ , if  $K = k$ ,  $L = l$ ,  $M = m$  then using (4.3) gives that

$$\begin{aligned} R_{2,3} = & -\frac{6}{r^3(r+1)^2n^2} \sum_{j,u,v,w,k,l=1}^r \sum_{m=1}^n \mathbb{E} [f''(h(\mathbf{S}_\theta^*)) \{ \mathbf{1}(v=w, j=u) \\ & + \mathbf{1}(u=w, j=v) + \mathbf{1}(u=v, j=w) \} (\rho_m(l) - \rho_m(k))^4 \\ & (\mathbf{1}\{j=u=v=w \in \{k,l\}\} + \mathbf{1}\{\{j,u,v,w\} = \{k,l\} \text{ with one of them appearing three times} \} \\ & + \mathbf{1}\{\{j,u,v,w\} = \{k,l\} \text{ with each appearing twice}\})]. \end{aligned}$$

If exactly one of  $\{j,u,v,w\}$  appears three times then

$$\{\mathbf{1}(v=w, j=u) + \mathbf{1}(u=w, j=v) + \mathbf{1}(u=v, j=w)\} = 0$$

and hence it follows that by counting how often each combination can appear, with  $2 \times 3$  counts for  $j=u=v=w \in \{k,l\}$  and  $6 \times 1$  counts for the appearing twice indicators, and using that  $\sum_{k,l=1}^r (\rho_m(l) - \rho_m(k))^4 \leq r^6$ , for  $r \geq 2$ , gives the bound

$$|R_{2,3}| \leq \frac{6\|f''\|}{r^3(r+1)^2n} \times 12 \times r^6 \leq \frac{Cr \|f''\|}{n}. \quad (4.7)$$

Collating (4.5), (4.6) and (4.7) gives the bound

$$|R_2| \leq \frac{Cr}{n} [\|f^{(4)}\| + \|f^{(3)}\| + \|f''\|]. \quad (4.8)$$

**Part III: Bounding  $R_1$ .** We begin by noting that if  $K = k$ ,  $L = l$ ,  $M = m$  then

$$\begin{aligned} (S'_j - S_j)(S'_u - S_u)(S'_v - S_v) = & \left( \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \right)^3 (\rho_m(l) - \rho_m(k))^3 (\mathbf{1}\{j=u=v \in \{k,l\}\} \\ & - \mathbf{1}\{j=u=k, v=l \text{ or } j=v=k, u=l \text{ or } u=v=k, j=l\} \\ & + \mathbf{1}\{j=u=l, v=k \text{ or } j=v=l, u=k \text{ or } u=v=l, j=k\}). \end{aligned} \quad (4.9)$$

By symmetry in the conditional expectation  $\mathbb{E}^{\mathbf{S}}[(S'_j - S_j)(S'_u - S_u)(S'_v - S_v)]$  the last two expectations of  $R_1$  cancel and we have

$$\begin{aligned} R_1 = & -\frac{rn}{8} \left( \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \right)^3 \frac{1}{r^2n} \sum_{k,l,j=1}^r \sum_{m=1}^n \mathbb{E} \left[ \frac{\partial^3}{\partial s_j^3} g(\mathbf{S})(\rho_m(l) - \rho_m(k))^3 \mathbf{1}(j \in \{k,l\}) \right] \\ = & -\frac{rn}{4} \left( \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \right)^3 \frac{1}{r^2n} \sum_{l,j=1}^r \sum_{m=1}^n \mathbb{E} \left[ \frac{\partial^3}{\partial s_j^3} g(\mathbf{S})(\rho_m(l) - \rho_m(j))^3 \right]. \end{aligned}$$

Moreover,

$$\sum_{l=1}^r (\rho_m(l) - \rho_m(j))^3 = \sum_{l=1}^r \left( l - \frac{r+1}{2} - \rho_m(j) \right)^3 = -\frac{r(r^2-1)}{4} \rho_m(j) - r \rho_m(j)^3,$$

where we used that  $\sum_{l=1}^r (l - (r+1)/2)^3 = \sum_{l=1}^r (l - (r+1)/2) = 0$ . Thus,

$$R_1 = \frac{1}{4} \left( \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \right)^3 \sum_{j=1}^r \sum_{m=1}^n \mathbb{E} \left[ \frac{\partial^3}{\partial s_j^3} g(\mathbf{S}) \left( \frac{(r^2-1)}{4} \rho_m(j) + \rho_m(j)^3 \right) \right].$$

We note that  $\mathbb{E}[\rho_m(j)] = \mathbb{E}[\rho_m(j)^3] = 0$  and that the  $n$  trials are independent. Now we carry out Taylor expansion around  $\rho_m$  by setting  $\mathbf{S}^{(m)}$ , the  $r$ -vector with components

$$S_j^{(m)} = \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \sum_{q \neq m} \rho_q(j) = S_j - \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(j),$$

which is independent of  $\rho_m(1), \dots, \rho_m(r)$ . This gives

$$R_1 = \frac{36}{r^2(r+1)^2 n^2} \sum_{j,k=1}^r \sum_{m=1}^n \mathbb{E} \left[ \frac{\partial^4}{\partial s_j^3 \partial s_k} g(\mathbf{S}_\theta^{(m)}) \left( \frac{(r^2-1)}{4} \rho_m(j) + \rho_m(j)^3 \right) \rho_m(k) \right],$$

where the  $j$ -th component of  $\mathbf{S}_\theta^{(m)}$  is given by  $\theta_j S_j + (1 - \theta_j) S_j^{(m)}$  for some  $\theta_j \in (0, 1)$ . Writing out the derivative using (4.2) gives

$$\begin{aligned} R_1 &= \frac{36}{r^2(r+1)^2 n^2} \sum_{j,k=1}^r \sum_{m=1}^n \mathbb{E} \left[ \left\{ 4f^{(4)}(h(\mathbf{S}_\theta^{(m)}))(\theta_j S_j + (1 - \theta_j) S_j^{(m)})^3 (\theta_k S_k + (1 - \theta_k) S_k^{(m)}) \right. \right. \\ &\quad \left. \left. + 6f^{(3)}(h(\mathbf{S}_\theta^{(m)}))(\theta_j S_j + (1 - \theta_j) S_j^{(m)})(\theta_k S_k + (1 - \theta_k) S_k^{(m)})(1 + \mathbf{1}(j = k)) \right. \right. \\ &\quad \left. \left. + 3f''(h(\mathbf{S}_\theta^{(m)}))\mathbf{1}(j = k) \right\} \left( \frac{(r^2-1)}{4} \rho_m(j) + \rho_m(j)^3 \right) \rho_m(k) \right]. \end{aligned} \quad (4.10)$$

To bound  $R_1$  we note some inequalities. Firstly, by (3.7), we have

$$\left| \left( \frac{(r^2-1)}{4} \rho_m(j) + \rho_m(j)^3 \right) \rho_m(k) \right| \leq \frac{(r+1)^2}{4} \left( \frac{(r^2-1)}{4} + \frac{(r+1)^2}{4} \right) = \frac{r(r+1)^3}{8}. \quad (4.11)$$

By independence and the Hölder and Cauchy-Schwarz inequalities, as  $\mathbb{E}[S_j^{(m)}] = \mathbb{E}[(S_j^{(m)})^3] = 0$ , for  $n \geq 2$

$$\begin{aligned} \mathbb{E}[(\theta_j S_j + (1 - \theta_j) S_j^{(m)})^4] &= \mathbb{E} \left[ \left( S_j^{(m)} + \theta_j \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(j) \right)^4 \right] \\ &= \mathbb{E}[(S_j^{(m)})^4] + \frac{4\sqrt{12}}{\sqrt{r(r+1)n}} \mathbb{E}[(S_j^{(m)})^3] \mathbb{E}[\theta_j \rho_m(j)] + \frac{6 \times 12}{r(r+1)n} \mathbb{E}[(S_j^{(m)})^2] \mathbb{E}[\theta_j^2 \rho_m(j)^2] \\ &\quad + \frac{4 \times 12\sqrt{12}}{r^{3/2}(r+1)^{3/2} n^{3/2}} \mathbb{E}[S_j^{(m)}] \mathbb{E}[\theta_j^3 \rho_m(j)^3] + \frac{144}{r^2(r+1)^2 n^2} \mathbb{E}[\theta_j^4 \rho_m(j)^4] \end{aligned}$$

$$\begin{aligned}
&\leq \mathbb{E}[(S_j^{(m)})^4] + \frac{6 \times 12}{r(r+1)n} \mathbb{E}[(S_j^{(m)})^2] \mathbb{E}[\rho_m(j)^2] + \frac{144}{r^2(r+1)^2 n^2} \mathbb{E}[\rho_m(j)^4] \\
&\leq 3 + \frac{6 \times 12}{r(r+1)n} \frac{(n-1)(r^2-1)}{n} \frac{1}{12} + \frac{144}{r^2(r+1)^2 n^2} \frac{r^4}{80} \leq 4.
\end{aligned} \tag{4.12}$$

Here we used (3.2) and the inequality  $\mathbb{E}[(S_j^{(m)})^4] \leq 3$ ,  $n \geq 2$ , which follows from a slight modification of the argument used to obtain inequality (3.2). Similarly,

$$\begin{aligned}
\mathbb{E}[(\theta_j S_j + (1-\theta_j) S_j^{(m)})^2] &= \mathbb{E}\left[\left(S_j^{(m)} + \theta_j \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(j)\right)^2\right] \\
&= \mathbb{E}[(S_j^{(m)})^2] + \frac{2\sqrt{12}}{\sqrt{r(r+1)n}} \mathbb{E}[S_j^{(m)}] \mathbb{E}[\theta_j \rho_m(j)] + \frac{12}{r(r+1)n} \mathbb{E}[\theta_j^2 \rho_m(j)^2] \\
&= \mathbb{E}[(S_j^{(m)})^2] + \frac{12}{r(r+1)n} \mathbb{E}[\theta_j^2 \rho_m(j)^2] \leq \frac{n-1}{n} + \frac{r^2-1}{r(r+1)n} \leq 1.
\end{aligned} \tag{4.13}$$

Applying inequalities (4.11), together with the basic inequality  $a^3 b \leq 3a^4/4 + b^4/4$  and (3.7) gives

$$\begin{aligned}
|R_1| &\leq \frac{36}{r^2(r+1)^2 n^2} \left\{ 4 \|f^{(4)}\| n r^2 \times \frac{r(r+1)^3}{8} \times 4 \right. \\
&\quad \left. + 6 \|f^{(3)}\| (n r^2 \times C r^4 + n r \times C r^4) + 3 \|f''\| n r \times C r^4 \right\} \\
&\leq \frac{Cr}{n} [r \|f^{(4)}\| + r \|f^{(3)}\| + \|f''\|].
\end{aligned} \tag{4.14}$$

**Part IV: Bounding  $R_0$ .** Let us decompose  $R_0$  as follows:

$$\begin{aligned}
R_0 &= \frac{rn}{4} \sum_{j=1}^r \mathbb{E} \left[ \frac{\partial^2}{\partial s_j^2} g(\mathbf{S}) \left( \mathbb{E}[(S'_j - S_j)^2] - (S'_j - S_j)^2 \right) \right] \\
&\quad + \frac{rn}{4} \sum_{j=1}^r \sum_{u \neq j} \mathbb{E} \left[ \frac{\partial^2}{\partial s_j \partial s_u} g(\mathbf{S}) \left( \mathbb{E}[(S'_j - S_j)(S'_u - S_u)] - (S'_j - S_j)(S'_u - S_u) \right) \right] \\
&= R_{0,1} + R_{0,2}.
\end{aligned}$$

We first treat  $R_{0,1}$ . Note that by equation (2.10) in [37],  $\mathbb{E}[(\mathbf{S}' - \mathbf{S})(\mathbf{S}' - \mathbf{S})^\top] = 2\Sigma\Lambda^\top = 4\Sigma/(rn)$  so that

$$\mathbb{E}[(S'_j - S_j)^2] = 4(r-1)/(r^2 n).$$

By (4.1), it follows that

$$R_{0,1} = \frac{1}{r^3(r+1)n} \sum_{j,k,l=1}^r \sum_{m=1}^n \mathbb{E} \left[ \frac{\partial^2}{\partial s_j^2} g(\mathbf{S}) ((r^2-1) - 3r(\rho_m(k) - \rho_m(l))^2 \mathbf{1}\{j \in \{k, l\}\}) \right].$$

Now,

$$\sum_{k,l=1}^r \left( (r^2-1) - 3r(\rho_m(k) - \rho_m(l))^2 \mathbf{1}\{j \in \{k, l\}\} \right)$$

$$\begin{aligned}
&= \sum_{k=1}^r (r^2 - 1) + \sum_{k=1}^r \sum_{l \neq k}^r \left( (r^2 - 1) - 3r(\rho_m(k) - \rho_m(l))^2 \mathbf{1}\{j \in \{k, l\}\} \right) \\
&= r^2(r^2 - 1) - 6r \left( r \rho_m(j)^2 + \frac{1}{12} r(r^2 - 1) \right) \\
&= \frac{1}{2} r^2(r^2 - 1) - 6r^2 \rho_m(j)^2.
\end{aligned}$$

Hence, expanding around  $\mathbf{S}^{(m)}$ , there exists a  $\theta = \theta(\mathbf{S}^{(m)}, \mathbf{S}) \in (0, 1)^r$  such that

$$\begin{aligned}
R_{0,1} &= \frac{1}{2r(r+1)n} \sum_{j,k=1}^r \sum_{m=1}^n \mathbb{E} \left[ \frac{\partial^2}{\partial s_j^2} g(\mathbf{S}) ((r^2 - 1) - 12\rho_m(j)^2) \right] \\
&= \frac{1}{2r(r+1)n} \sum_{j,k=1}^r \sum_{m=1}^n \mathbb{E} \left[ \frac{\partial^2}{\partial s_j^2} g(\mathbf{S}^{(m)}) \right] \mathbb{E} [((r^2 - 1) - 12\rho_m(j)^2)] \\
&\quad + \frac{\sqrt{12}}{2r^{3/2}(r+1)^{3/2}n^{3/2}} \sum_{j,k,q=1}^r \sum_{m=1}^n \mathbb{E} \left[ \frac{\partial^3}{\partial s_j^2 \partial s_q} g(\mathbf{S}^{(m)}) \right] \mathbb{E} [((r^2 - 1) - 12\rho_m(j)^2) \rho_m(q)] \\
&\quad + \frac{3}{r^2(r+1)^2n^2} \sum_{j,k,q,t=1}^r \sum_{m=1}^n \mathbb{E} \left[ \left( \frac{\partial^4}{\partial s_j^2 \partial s_q \partial s_t} g(\mathbf{S}^{(m)}) \right) \right. \\
&\quad \left. ((r^2 - 1) - 12\rho_m(j)^2) \rho_m(q) \rho_m(t) \right].
\end{aligned}$$

By independence, the first summand vanishes. Also, from Lemma 3.2,  $\mathbb{E}[\rho_m(q)] = \mathbb{E}[\rho_m(q)^3] = 0$  and  $\mathbb{E}[\rho_m(j)^2 \rho_m(q)] = 0$  for  $j \neq q$ . Therefore the second summand in the expansion of  $R_{0,1}$  also vanishes, leaving just the third sum involving the fourth order partial derivatives of  $g$ . With the derivative expansion (4.2) this yields

$$\begin{aligned}
R_{0,1} &= \frac{3}{r^2(r+1)^2n^2} \sum_{j,q,t=1}^r \sum_{m=1}^n \mathbb{E} \left[ \left\{ 4f^{(4)}(h(\mathbf{S}_\theta^{(m)})) \right. \right. \\
&\quad (\theta_j S_j^{(m)} + (1 - \theta_j) S_j)^2 (\theta_q S_q^{(m)} + (1 - \theta_q) S_q) (\theta_t S_t^{(m)} + (1 - \theta_t) S_t) \\
&\quad + 2f^{(3)}(h(\mathbf{S}_\theta^{(m)})) \{ (\theta_j S_j^{(m)} + (1 - \theta_j) S_j)^2 \mathbf{1}(q = t) \\
&\quad + 2(\theta_j S_j^{(m)} + (1 - \theta_j) S_j) \{ (\theta_q S_q^{(m)} + (1 - \theta_q) S_q) \mathbf{1}(j = t) + (\theta_t S_t^{(m)} + (1 - \theta_t) S_t) \mathbf{1}(j = q) \} \\
&\quad + (\theta_q S_q^{(m)} + (1 - \theta_q) S_q) (\theta_t S_t^{(m)} + (1 - \theta_t) S_t) \} \\
&\quad \left. \left. + f''(h(\mathbf{S}_\theta^{(m)})) (\mathbf{1}(q = t)) + 2 \mathbf{1}(j = q = t) \right\} ((r^2 - 1) - 12\rho_m(j)^2) \rho_m(q) \rho_m(t) \right] \\
&= R_{0,1,1} + R_{0,1,2} + R_{0,1,3}.
\end{aligned}$$



First,

$$\begin{aligned}
 R_{0,1,1} &= \frac{12}{r^2(r+1)^2n^2} \sum_{j,q,t=1}^r \sum_{m=1}^n \mathbb{E} \left[ f^{(4)}(h(\mathbf{S}_\theta^{(m)})) \left( S_j - \theta_j \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(j) \right)^2 \right. \\
 &\quad \left( S_q - \theta_q \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(q) \right) \left( S_t - \theta_t \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(t) \right) \\
 &\quad \left. \left( (r^2 - 1) - 12\rho_m(j)^2 \right) \rho_m(q) \rho_m(t) \right] \\
 &= R_{0,1,1,1} + R_{0,1,1,2}.
 \end{aligned}$$

Here, with the difference of sum notation indicating that we take the difference of the summands over the respective indices,

$$\begin{aligned}
 R_{0,1,1,1} &= \frac{12}{r^2(r+1)^2n^2} \left\{ \sum_{j=1}^r \sum_{q=1}^r \sum_{t=1}^r \sum_{m=1}^n - \sum_{j=1}^r \sum_{q \neq j} \sum_{t \neq j, q} \sum_{m=1}^n \right\} \mathbb{E} \left[ f^{(4)}(h(\mathbf{S}_\theta^{(m)})) \right. \\
 &\quad \left( S_j - \theta_j \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(j) \right)^2 \left( S_q - \theta_q \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(q) \right) \\
 &\quad \left. \left( S_t - \theta_t \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(t) \right) \left( (r^2 - 1) - 12\rho_m(j)^2 \right) \rho_m(q) \rho_m(t) \right]
 \end{aligned}$$

and

$$\begin{aligned}
 R_{0,1,1,2} &= \frac{12}{r^2(r+1)^2n^2} \sum_{j=1}^r \sum_{q \neq j} \sum_{t \neq j, q} \sum_{m=1}^n \mathbb{E} \left[ f^{(4)}(h(\mathbf{S}_\theta^{(m)})) \right. \\
 &\quad \left( S_j - \theta_j \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(j) \right)^2 \left( S_q - \theta_q \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(q) \right) \\
 &\quad \left. \left( S_t - \theta_t \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(t) \right) \left( (r^2 - 1) - 12\rho_m(j)^2 \right) \rho_m(q) \rho_m(t) \right].
 \end{aligned}$$

With (3.7), we can bound

$$|(r^2 - 1) - 12\rho_m(k)^2| \leq 2(r+1)(r+2), \quad (4.15)$$

and using the basic inequality  $a^2bc \leq a^4/2 + b^4/4 + c^4/4$  and (4.12) we obtain

$$|R_{0,1,1,1}| \leq \frac{12\|f^{(4)}\|}{r^2(r+1)^2n^2} (r^3 - r(r-1)(r-2))n \times 2(r+1)(r+2) \left( \frac{r+1}{2} \right)^2 \times 4 \leq \frac{Cr^2\|f^{(4)}\|}{n}.$$

We now decompose

$$R_{0,1,1,2} = R_{0,1,1,2,1} + R_{0,1,1,2,2} + R_{0,1,1,2,3}$$

with

$$\begin{aligned}
 R_{0,1,1,2,1} &= \frac{3}{r^2(r+1)^2n^2} \sum_{j=1}^r \sum_{q \neq j} \sum_{t \neq j, q} \sum_{m=1}^n \mathbb{E} \left[ f^{(4)}(h(\mathbf{S}_\theta^{(m)})) \right. \\
 &\quad \left. \left( S_j - \theta_j \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(j) \right)^2 S_q S_t ((r^2 - 1) - 12\rho_m(j)^2) \rho_m(q) \rho_m(t) \right], \\
 R_{0,1,1,2,2} &= 2 \frac{3}{r^2(r+1)^2n^2} \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \sum_{j=1}^r \sum_{q \neq j} \sum_{t \neq j, q} \sum_{m=1}^n \mathbb{E} \left[ \theta_t f^{(4)}(h(\mathbf{S}_\theta^{(m)})) \right. \\
 &\quad \left. \left( S_j - \theta_j \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(j) \right)^2 S_q ((r^2 - 1) - 12\rho_m(j)^2) \rho_m(q) \rho_m(t)^2 \right], \\
 R_{0,1,1,2,3} &= \frac{3}{r^2(r+1)^2n^2} \frac{12}{r(r+1)n} \sum_{j=1}^r \sum_{q \neq j} \sum_{t \neq j, q} \sum_{m=1}^n \mathbb{E} \left[ \theta_q \theta_t f^{(4)}(h(\mathbf{S}_\theta^{(m)})) \right. \\
 &\quad \left. \left( S_j - \theta_j \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(j) \right)^2 ((r^2 - 1) - 12\rho_m(j)^2) \rho_m(q)^2 \rho_m(t)^2 \right].
 \end{aligned}$$

The term  $R_{0,1,1,2,3}$  is straightforward to bound: with the Cauchy-Schwarz inequality and inequalities (4.12) and that  $|\rho_m(l)| \leq (r+1)/2$  we obtain

$$|R_{0,1,1,2,3}| \leq \frac{36\|f^{(4)}\|}{r^3(r+1)^3n^3} nr(r-1)(r-2)\sqrt{4}Cr^6 \leq \frac{Cr^3\|f^{(4)}\|}{n^2}. \quad (4.16)$$

For  $R_{0,1,1,2,1}$  we set  $T_m = \sum_{l=1}^r S_l \rho_m(l)$ . This allows us to write

$$\begin{aligned}
 R_{0,1,1,2,1} &= \frac{3}{r^2(r+1)^2n^2} \sum_{j=1}^r \sum_{q \neq j} \sum_{m=1}^n \mathbb{E} \left[ f^{(4)}(h(\mathbf{S}_\theta^{(m)})) \left( S_j - \theta_j \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(j) \right)^2 \right. \\
 &\quad \left. ((r^2 - 1) - 12\rho_m(j)^2) S_q \rho_m(q) (T_m - S_j \rho_m(j) - S_q \rho_m(q)) \right] \\
 &= R_{0,1,1,2,1,a} + R_{0,1,1,2,1,b}.
 \end{aligned}$$

Here

$$\begin{aligned}
 R_{0,1,1,2,1,a} &= \frac{3}{r^2(r+1)^2n^2} \sum_{j=1}^r \sum_{q \neq j} \sum_{m=1}^n \mathbb{E} \left[ f^{(4)}(h(\mathbf{S}_\theta^{(m)})) \left( S_j - \theta_j \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(j) \right)^2 \right. \\
 &\quad \left. ((r^2 - 1) - 12\rho_m(j)^2) S_q \rho_m(q) T_m \right] \\
 &= \frac{3}{r^2(r+1)^2n^2} \sum_{j=1}^r \sum_{m=1}^n \mathbb{E} \left[ f^{(4)}(h(\mathbf{S}_\theta^{(m)})) \left( S_j - \theta_j \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(j) \right)^2 \right. \\
 &\quad \left. ((r^2 - 1) - 12\rho_m(j)^2) (T_m^2 - S_j \rho_m(j) T_m) \right].
 \end{aligned}$$

Thus, with (4.12), (4.15) and the Cauchy-Schwarz inequality (twice)

$$\begin{aligned}
 |R_{0,1,1,2,1,a}| &\leq \frac{6(r+2)}{r^2(r+1)n^2} \|f^{(4)}\| \sum_{m=1}^n \left\{ r\sqrt{4}\sqrt{\mathbb{E}[T_m^4]} + \sum_{j=1}^r \sqrt{4}(\mathbb{E}[S_j^4 \rho_m(j)^4] \mathbb{E}[T_m^4])^{1/4} \right\} \\
 &\leq \frac{12(r+2)\|f^{(4)}\|}{r(r+1)n} \left\{ \sqrt{C_T(r,n)} r^3 + 3^{1/4} \frac{r+1}{2} C_T^{1/4}(r,n) r^{3/2} \right\} \\
 &\leq \frac{Cr^2\|f^{(4)}\|}{n} \left( 1 + \frac{r}{n} \right),
 \end{aligned}$$

where we used (3.4), that  $\mathbb{E}[S_j^4] < 3$ , that  $\rho_m(j)^4 \leq ((r+1)/2)^4$ , that  $r \geq 3$  and the inequality  $C_T^{1/4}(r,n) \leq (1 + \sqrt{C_T(r,n)})/2$  and that  $\sqrt{C_T(r,n)} \leq (7/48 + 1/(5n))^{1/2} + r/(6n)$ . Also, making similar considerations,

$$\begin{aligned}
 |R_{0,1,1,2,1,b}| &= \left| -\frac{3}{r^2(r+1)^2n^2} \sum_{j=1}^r \sum_{q \neq j} \sum_{m=1}^n \mathbb{E} \left[ f^{(4)}(h(\mathbf{S}_\theta^{(m)})) \left( S_j - \theta_j \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(j) \right)^2 \right. \right. \\
 &\quad \left. \left. ((r^2 - 1) - 12\rho_m(j)^2) S_q \rho_m(q) (S_j \rho_m(j) + S_q \rho_m(q)) \right) \right] \Big| \\
 &\leq \frac{6(r+2)\|f^{(4)}\|}{r^2(r+1)n^2} nr(r-1)\sqrt{4} \left\{ (\mathbb{E}[S_1^4 \rho_1(1)^4] \mathbb{E}[S_2^4 \rho_1(2)^4])^{1/4} + \sqrt{\mathbb{E}[S_2^4 \rho_1(2)^4]} \right\} \\
 &\leq \frac{12(r+2)(r-1)\|f^{(4)}\|}{r(r+1)n} 2\sqrt{3} \left( \frac{r+1}{2} \right)^2 \leq \frac{Cr^2\|f^{(4)}\|}{n}.
 \end{aligned}$$

Thus,

$$|R_{0,1,1,2,1}| \leq \frac{Cr^2\|f^{(4)}\|}{n} \left( 1 + \frac{r}{n} \right). \quad (4.17)$$

Next, using  $T_m$  again,

$$\begin{aligned}
 R_{0,1,1,2,2} &= \frac{12\sqrt{3}}{r^{5/2}(r+1)^{5/2}n^{5/2}} \sum_{j=1}^r \sum_{t \neq j} \sum_{m=1}^n \mathbb{E} \left[ \theta_t f^{(4)}(h(\mathbf{S}_\theta^{(m)})) \left( S_j - \theta_j \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(j) \right)^2 \right. \\
 &\quad \left. ((r^2 - 1) - 12\rho_m(j)^2) \rho_m(t)^2 (T_m - S_j \rho_m(j) - S_t \rho_m(t)) \right].
 \end{aligned}$$

Using the Cauchy-Schwarz inequality, (4.12), (3.3), that  $\mathbb{E}[S_j^2] \leq 1$ , and (4.15) we obtain

$$\begin{aligned}
 |R_{0,1,1,2,2}| &\leq \frac{12\sqrt{3}\|f^{(4)}\|}{r^{5/2}(r+1)^{5/2}n^{5/2}} nr(r-1)2(r+1)(r+2) \left( \frac{r+1}{2} \right)^2 \\
 &\quad \sqrt{4} \left\{ \sqrt{\mathbb{E}[T_1^2]} + 2\frac{r+1}{2} \sqrt{\mathbb{E}[S_1^2]} \right\} \\
 &\leq \frac{12\sqrt{3}r(r+2)\|f^{(4)}\|}{(r+1)n^{3/2}} \left\{ \frac{r^{3/2}}{\sqrt{12}} \left( 1 + \frac{r}{n} \right)^{1/2} + (r+1) \right\} \leq \frac{Cr^2\|f^{(4)}\|}{n} \left( 1 + \frac{r}{n} \right),
 \end{aligned}$$

employing that  $(a+b)^{1/2} \leq \sqrt{a} + \sqrt{b}$  and that  $r \geq 3$  in the penultimate step. Combining this bound with (4.16) and (4.17),

$$|R_{0,1,1,2}| \leq \frac{Cr^2 \|f^{(4)}\|}{n} \left(1 + \frac{r}{n}\right) \quad \text{and} \quad |R_{0,1,1}| \leq \frac{Cr^2 \|f^{(4)}\|}{n} \left(1 + \frac{r}{n}\right).$$

We now bound  $R_{0,1,2}$ . By symmetry,

$$\begin{aligned} R_{0,1,2} &= \frac{24}{r^2(r+1)^2n^2} \sum_{j,q=1}^r \sum_{m=1}^n \mathbb{E} \left[ f^{(3)}(h(\mathbf{S}_\theta^{(m)})) \left\{ (\theta_j S_j^{(m)} + (1-\theta_j)S_j)^2 \rho_m(q)^2 \right. \right. \\ &\quad \left. \left. + 4(\theta_j S_j^{(m)} + (1-\theta_j)S_j)(\theta_q S_q^{(m)} + (1-\theta_q)S_q) \rho_m(q) \rho_m(j) \right\} ((r^2-1) - 12\rho_m(j)^2) \right] \\ &\quad + \frac{24}{r^2(r+1)^2n^2} \sum_{j,q,t=1}^r \sum_{m=1}^n \mathbb{E} \left[ f^{(3)}(h(\mathbf{S}_\theta^{(m)})) (\theta_q S_q^{(m)} + (1-\theta_q)S_q) \right. \\ &\quad \left. (\theta_t S_t^{(m)} + (1-\theta_t)S_t) ((r^2-1) - 12\rho_m(j)^2) \rho_m(q) \rho_m(t) \right]. \end{aligned}$$

As  $12 \sum_{j=1}^r \rho_m(j)^2 = r(r^2-1)$  the last summand vanishes and with the Cauchy-Schwarz inequality as well as inequalities (4.12) and (3.7),

$$|R_{0,1,2}| \leq \frac{24 \|f^{(3)}\|}{r^2(r+1)^2n^2} \sqrt{4}(Cr^4 + 4Cr^4) \leq \frac{Cr^2 \|f^{(3)}\|}{n}.$$

For  $R_{0,1,3}$ , as  $\sum_{j=1}^r ((r^2-1) - 12\rho_m(j)^2) = 0$ , we have with inequality (3.7),

$$\begin{aligned} |R_{0,1,3}| &= \frac{3}{r^2(r+1)^2n^2} \left| \sum_{j=1}^r \sum_{m=1}^n \mathbb{E} \left[ f''(h(\mathbf{S}_\theta^{(m)})) \left( (r^2-1) - 12\rho_m(j)^2 \right) \rho_m(j)^2 \right] \right| \\ &\leq \frac{3 \|f''\|}{r^2(r+1)^2n^2} rnCr^4 \leq \frac{Cr \|f''\|}{n}. \end{aligned}$$

Having bounded  $R_{0,1}$  to the desired order we move on to  $R_{0,2}$ . For  $u \neq j$ , again by Equation (2.10) in [37],  $\mathbb{E}[(S'_j - S_j)(S'_u - S_u)] = 4\Sigma_{ju}/(rn) = -4/(r^2n)$ . Hence,

$$\begin{aligned} R_{0,2} &= \frac{rn}{4} \sum_{j=1}^r \sum_{u \neq j} \mathbb{E} \left[ \frac{\partial^2}{\partial s_j \partial s_u} g(\mathbf{S}) \left( \mathbb{E}[(S'_j - S_j)(S'_u - S_u)] - (S'_j - S_j)(S'_u - S_u) \right) \right] \\ &= \frac{1}{4r} \sum_{j,k,l=1}^r \sum_{u \neq j} \sum_{m=1}^n \mathbb{E} \left[ \frac{\partial^2}{\partial s_j \partial s_u} g(\mathbf{S}) \right. \\ &\quad \left. \left( -\frac{4}{r^2n} + \frac{12}{r(r+1)n} (\rho_m(k) - \rho_m(l))^2 \mathbf{1}_{\{\{j,u\} = \{k,l\}\}} \right) \right] \\ &= \frac{1}{r} \sum_{k=1}^r \sum_{l \neq k} \sum_{m=1}^n \mathbb{E} \left[ \frac{\partial^2}{\partial s_k \partial s_l} g(\mathbf{S}) \left( -\frac{1}{n} + \frac{6}{r(r+1)n} (\rho_m(k) - \rho_m(l))^2 \right) \right]. \end{aligned}$$

Again with Taylor expansion and the independence of the trials, for some  $\theta = \theta(\mathbf{S}^{(m)}, \mathbf{S}) \in (0, 1)^r$ ,

$$\begin{aligned} R_{0,2} = & \frac{1}{r} \sum_{k=1}^r \sum_{l \neq k} \sum_{m=1}^n \mathbb{E} \left[ \frac{\partial^2}{\partial s_k \partial s_l} g(\mathbf{S}^{(m)}) \right] \mathbb{E} \left[ \left( -\frac{1}{n} + \frac{6}{r(r+1)n} (\rho_m(k) - \rho_m(l))^2 \right) \right] \\ & + \frac{1}{r} \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \sum_{k,q=1}^r \sum_{l \neq k} \sum_{m=1}^n \mathbb{E} \left[ \frac{\partial^3}{\partial s_k \partial s_l \partial s_q} g(\mathbf{S}^{(m)}) \right] \\ & \mathbb{E} \left[ \left( -\frac{1}{n} + \frac{6}{r(r+1)n} (\rho_m(k) - \rho_m(l))^2 \right) \rho_m(q) \right] \\ & + \frac{1}{r} \frac{6}{r(r+1)n} \sum_{k,q,t=1}^r \sum_{l \neq k} \sum_{m=1}^n \mathbb{E} \left[ \frac{\partial^4}{\partial s_k \partial s_l \partial s_q \partial s_t} g(\mathbf{S}^{(m)}) \right. \\ & \left. \left( -\frac{1}{n} + \frac{6}{r(r+1)n} (\rho_m(k) - \rho_m(l))^2 \right) \rho_m(q) \rho_m(t) \right]. \end{aligned}$$

Again the first term vanishes by independence and the second term vanishes. Hence, with the derivative expansion (4.2),

$$\begin{aligned} R_{0,2} = & \frac{6}{r^2(r+1)n^2} \sum_{k,q,t=1}^r \sum_{l \neq k} \sum_{m=1}^n \mathbb{E} \left[ \left\{ 4f^{(4)}(h(\mathbf{S}_\theta^{(m)})) \{ \theta_k S_k^{(m)} + (1 - \theta_k) S_k \} \right. \right. \\ & (\theta_l S_l^{(m)} + (1 - \theta_l) S_l) (\theta_q S_q^{(m)} + (1 - \theta_q) S_q) (\theta_t S_t^{(m)} + (1 - \theta_t) S_t) \\ & + 2f^{(3)}(h(\mathbf{S}_\theta^{(m)})) [(\theta_k S_k^{(m)} + (1 - \theta_k) S_k) \{ (\theta_l S_l^{(m)} + (1 - \theta_l) S_l) \mathbf{1}(q=t) \\ & + (\theta_q S_q^{(m)} + (1 - \theta_q) S_q) \mathbf{1}(l=t) + (\theta_t S_t^{(m)} + (1 - \theta_t) S_t) \mathbf{1}(l=q) \} \\ & + (\theta_l S_l^{(m)} + (1 - \theta_l) S_l) \{ (\theta_t S_t^{(m)} + (1 - \theta_t) S_t) \mathbf{1}(j=q) + (\theta_t S_t^{(m)} + (1 - \theta_t) S_t) \mathbf{1}(k=q) \} \} \\ & \left. \left. + f''(h(\mathbf{S}_\theta^{(m)})) (\mathbf{1}(k=q, l=t) + \mathbf{1}(k=t, l=q)) \right\} \right. \\ & \left. \left( -1 + \frac{6}{r(r+1)} (\rho_m(k) - \rho_m(l))^2 \right) \rho_m(q) \rho_m(t) \right] \\ = & R_{0,2,1} + R_{0,2,2} + R_{0,2,3}. \end{aligned}$$

For  $R_{0,2,1}$  we employ  $T_m$  from Lemma 3.3 and write

$$\begin{aligned} R_{0,2,1} = & \frac{24}{r^2(r+1)n^2} \sum_{l=1}^r \sum_{k \neq l} \sum_{m=1}^n \mathbb{E} \left[ f^{(4)}(h(\mathbf{S}_\theta^{(m)})) \left( S_k - \theta_q \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(k) \right) \right. \\ & \left( S_l - \theta_l \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(l) \right) \left( T_m - \sum_{q=1}^r \theta_q \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(q)^2 \right) \\ & \left. \left( T_m - \sum_{t=1}^r \theta_t \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(t)^2 \right) \left( -1 + \frac{6}{r(r+1)} (\rho_m(k) - \rho_m(l))^2 \right) \right] \end{aligned}$$

$$\begin{aligned}
&= \frac{24}{r^2(r+1)n^2} \sum_{l=1}^r \sum_{k \neq l} \sum_{m=1}^n \mathbb{E} \left[ f^{(4)}(h(\mathbf{S}_\theta^{(m)})) \left( S_k - \theta_q \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(k) \right) \right. \\
&\quad \left. \left( S_l - \theta_l \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \rho_m(l) \right) \left( T_m^2 - 2T_m \frac{\sqrt{12}}{\sqrt{r(r+1)n}} \sum_{q=1}^r \theta_q \rho_m(q)^2 \right. \right. \\
&\quad \left. \left. + \frac{12}{r(r+1)n} \left( \sum_{t=1}^r \theta_t \rho_m(t)^2 \right)^2 \right) \left( -1 + \frac{6}{r(r+1)} (\rho_m(k) - \rho_m(l))^2 \right) \right].
\end{aligned}$$

We now note that, since  $|\rho_m(k) - \rho_m(l)| \leq r - 1$ , we have

$$\left| -1 + \frac{6}{r(r+1)} (\rho_m(k) - \rho_m(l))^2 \right| \leq \frac{6(r-1)^2}{r(r+1)} - 1 < 5.$$

Using this bound, the Cauchy-Schwarz inequality as well as (3.3), (3.4) and (4.12), and using that  $r \geq 3$  in the final step, gives that

$$\begin{aligned}
|R_{0,2,1}| &\leq \frac{24\|f^{(4)}\|}{r^2(r+1)n^2} \times 5r(r-1)n \times \sqrt{4} \left\{ \sqrt{\mathbb{E}[T_1^4]} + \frac{4\sqrt{3}}{\sqrt{r(r+1)n}} \frac{r(r+1)}{2} \sqrt{\mathbb{E}[T_1^2]} \right. \\
&\quad \left. + \frac{12}{r(r+1)n} \frac{r^2(r+1)^2}{4} \right\} \\
&\leq \frac{240(r-1)\|f^{(4)}\|}{r(r+1)n} \left\{ \sqrt{C_T(r,n)} r^3 + \frac{2\sqrt{3}\sqrt{r(r+1)}}{\sqrt{n}} \frac{r^{3/2}}{\sqrt{12}} \left( 1 + \frac{r}{n} \right)^{1/2} + \frac{3r(r+1)}{n} \right\} \\
&\leq \frac{Cr^2\|f^{(4)}\|}{n} \left( 1 + \frac{r}{n} \right).
\end{aligned}$$

For  $R_{0,2,2}$  we use the Cauchy-Schwarz inequality and (4.12) and (3.10) to obtain

$$|R_{0,2,2}| \leq \frac{12 \times 5\|f^{(3)}\|}{r^2(r+1)n^2} \times nr^3 \sqrt{4} \times \sqrt{C} \left( \frac{r+1}{2} \right)^2 \leq \frac{Cr^2\|f^{(3)}\|}{n}.$$

Bounding  $R_{0,2,3}$  is also simple:

$$|R_{0,2,3}| \leq \frac{6 \times 2\|f''\|}{r^2(r+1)n^2} \times nr^2 \times \sqrt{C} \left( \frac{r+1}{2} \right)^2 \leq \frac{Cr\|f''\|}{n}.$$

Summing up our bounds for  $|R_{0,1,1}|$ ,  $|R_{0,1,2}|$ ,  $|R_{0,1,3}|$ ,  $|R_{0,2,1}|$ ,  $|R_{0,2,2}|$ ,  $|R_{0,2,3}|$  yields the following bound for  $R_0$ :

$$|R_0| \leq \frac{Cr}{n} \left[ r \left( 1 + \frac{r}{n} \right) \|f^{(4)}\| + r\|f^{(3)}\| + \|f''\| \right]. \quad (4.18)$$

Finally, adding the bounds (4.8), (4.14) and (4.18) for  $R_2$ ,  $R_1$  and  $R_0$  yields the bound

$$|\mathbb{E}[h(F_r)] - \chi_{(r-1)}^2 h| \leq \frac{Cr}{n} \left\{ r \left( 1 + \frac{r}{n} \right) \|f^{(4)}\| + r\|f^{(3)}\| + \|f''\| \right\}.$$

Bounding  $\|f^{(4)}\|$  and  $\|f^{(3)}\|$  using inequality (2.6) and bounding  $\|f''\|$  using inequality (2.4) now yields the desired bound (1.4).

**Part V: Optimality of the rate.** Let  $h_t(x) = \cos(tx)$ , and observe that  $h_t \in C_b^{1,3}(\mathbb{R}^+)$ . From the power series representation of the cosine function we can write  $h_t(x) = \cos(tx) = 1 - t^2 x^2/2 + R_t(x)$ , where  $R_t(x) = \sum_{k=2}^{\infty} (-1)^k (tx)^{2k}/(2k)!$ . Now, we let  $Y_{r-1} \sim \chi_{(r-1)}^2$  and recall that  $\mathbb{E}[Y_{r-1}^2] = r^2 - 1$ , for  $r \geq 2$ . Therefore using (3.6) we have that

$$\begin{aligned} |\mathbb{E}[h_t(F_r)] - \chi_{(r-1)}^2 h_t| &= |(t^2/2)(\mathbb{E}[F_r^2] - \mathbb{E}[Y_{r-1}^2]) + (\mathbb{E}[R_t(F_r)] - \mathbb{E}[R_t(Y_{r-1})])| \\ &= \left| \frac{t^2(r-1)}{n} - (\mathbb{E}[R_t(F_r)] - \mathbb{E}[R_t(Y_{r-1})]) \right|. \end{aligned}$$

The quantity  $\mathbb{E}[R_t(F_r)] - \mathbb{E}[R_t(Y_{r-1})]$  is a power series in  $t$  which is linearly independent of  $t^2$ . Hence, the  $r/n$  rate of the bound (1.4) is optimal.  $\square$

## 5. Further proofs

**Proof of Proposition 1.4.** We recall some bounds from [18]. Suppose  $X_1, \dots, X_n$  are i.i.d. random variables with  $\mathbb{E}[X_1] = 0$ ,  $\mathbb{E}[X_1^2] = 1$  and  $\mathbb{E}[X_1^8] < \infty$ . Let  $W = n^{-1/2} \sum_{i=1}^n X_i$ . Then inequalities (3.10) and (3.11) of [18] state that

$$d_W(\mathcal{L}(W^2), \chi_{(1)}^2) \leq \frac{48}{\sqrt{n}} \left[ \mathbb{E}|X_1^3| + \sqrt{\frac{2}{\pi}} \mathbb{E}|X_1^4| + \frac{\mathbb{E}|X_1^5|}{\sqrt{n}} \right], \quad (5.1)$$

and, for  $h \in C_b^{1,2}(\mathbb{R}^+)$ , if  $\mathbb{E}[X_1^3] = 0$ ,

$$|\mathbb{E}[h(W^2)] - \chi_{(1)}^2 h| \leq \frac{1}{n} \{ \|h'\| + \|h''\| \} \left[ 22\mathbb{E}|W^3| + 40\mathbb{E}|X_1^5| + \frac{43}{n} \mathbb{E}|X_1^7| \right]. \quad (5.2)$$

Recall from Remark 1.3 that we can write  $F_2 = 2S_1^2$ , where  $S_1 = n^{-1/2} \sum_{i=1}^n Y_i$ , where  $Y_1, \dots, Y_n$  are i.i.d. with  $Y_1 \sim \text{Unif}\{-1/\sqrt{2}, 1/\sqrt{2}\}$ . By rescaling,  $F_2 = W^2$  with  $W = n^{-1/2} \sum_{i=1}^n X_i$ , where  $X_1, \dots, X_n$  are i.i.d. with  $X_1 \sim \text{Unif}\{-1, 1\}$ . We have  $\mathbb{E}[X_1] = \mathbb{E}[X_1^3] = 0$  and  $\mathbb{E}|X_1^m| = 1$  for all  $m \geq 1$ . We also have, by Hölder's inequality, that  $\mathbb{E}|W^3| \leq \{\mathbb{E}[W^4]\}^{3/4} = \{3(n-1)(\mathbb{E}[X_1^2])^2/n + \mathbb{E}[X_1^4]/n\}^{3/4} \leq 3^{3/4}$ . We have verified that the conditions under inequalities (5.1) and (5.2) hold are met, and plugging our moment estimates into these bounds yields the bounds (1.5) and (1.6), respectively.  $\square$

**Proof of Corollary 1.5.** First, we consider the case  $r = 2$ . Recall from the proof of Proposition 1.4 that we can write  $F_2 = W^2$ , where  $W = n^{-1/2} \sum_{i=1}^n X_i$ , and  $X_1, \dots, X_n$  are i.i.d. with  $X_1 \sim \text{Unif}\{-1, 1\}$ . Recall also that if  $Z \sim N(0, 1)$ , then  $Z^2 \sim \chi_{(1)}^2$ . Now, for any  $z > 0$ ,

$$\begin{aligned} |\mathbb{P}(W^2 \leq z) - \mathbb{P}(Z^2 \leq z)| &= |\mathbb{P}(-\sqrt{z} \leq W \leq \sqrt{z}) - \mathbb{P}(-\sqrt{z} \leq Z \leq \sqrt{z})| \\ &\leq |\mathbb{P}(W \leq \sqrt{z}) - \mathbb{P}(Z \leq \sqrt{z})| + |\mathbb{P}(W \leq -\sqrt{z}) - \mathbb{P}(Z \leq -\sqrt{z})|, \end{aligned}$$

and so

$$d_K(\mathcal{L}(F_2), \chi_{(1)}^2) \leq 2d_K(\mathcal{L}(W), \mathcal{L}(Z)).$$

Using the Berry-Esseen theorem with the best numerical constant currently available of  $C = 0.4748$  due to [38] we obtain the bound

$$d_K(\mathcal{L}(F_2), \chi_{(1)}^2) \leq 2 \cdot \frac{0.4748 \mathbb{E}|X_1^3|}{(\mathbb{E}[X_1^2])^{3/2} \sqrt{n}} = \frac{0.9496}{\sqrt{n}},$$

where we used that  $\mathbb{E}[X_1^2] = \mathbb{E}|X_1^3| = 1$ .

We now consider the case  $r \geq 3$ . Let  $\alpha > 0$ , and for fixed  $z > 0$  define

$$g_z(x) = \begin{cases} 1, & \text{if } x \leq -1, \\ 1 - \frac{2}{3}(x+1)^3, & \text{if } -1 < x \leq -1/2, \\ \frac{2}{3}x^3 - x + \frac{1}{2}, & \text{if } -1/2 < x \leq 1/2, \\ \frac{2}{3}(1-x)^3, & \text{if } 1/2 < x \leq 1, \\ 0, & \text{if } x > 1. \end{cases}$$

Now let  $h_{\alpha,z}(x) = g_z(1 + 2(x - z)/\alpha)$ . We observe that  $h''_{\alpha,z}$  is Lipschitz, and simple calculations show that  $\|h'_{\alpha,z}\| = 2/\alpha$ ,  $\|h''_{\alpha,z}\| = 8/\alpha^2$  and  $\|h^{(3)}_{\alpha,z}\| = 32/\alpha^3$ . Let  $Y_{r-1} \sim \chi_{(r-1)}^2$ . Then by applying (1.4) we obtain the bound

$$\begin{aligned} & \mathbb{P}(F_r \leq z) - \mathbb{P}(Y_{r-1} \leq z) \\ & \leq \mathbb{E}[h_{\alpha,z}(F_r)] - \mathbb{E}[h_{\alpha,z}(Y_{r-1})] + \mathbb{E}[h_{\alpha,z}(Y_{r-1})] - \mathbb{P}(Y_{r-1} \leq z) \\ & \leq \frac{r}{n} \left\{ C_1 \|h'_{\alpha,z}\| + \left( C_2 + \frac{C_3 r}{n} \right) \|h''_{\alpha,z}\| + \left( C_4 + \frac{C_5 r}{n} \right) \|h^{(3)}_{\alpha,z}\| \right\} + \mathbb{P}(z \leq Y_{r-1} \leq z + \alpha) \\ & = \frac{r}{n} \left\{ \frac{2C_1}{\alpha} + \frac{8C_2}{\alpha^2} + \frac{8C_3 r}{n\alpha^2} + \frac{32C_4}{\alpha^3} + \frac{32C_5 r}{n\alpha^3} \right\} + \mathbb{P}(z \leq Y_{r-1} \leq z + \alpha). \end{aligned} \quad (5.3)$$

The following inequality is a slight simplification of one given on p. 754 of [20]:

$$\mathbb{P}(z \leq Y_{r-1} \leq z + \alpha) \leq \begin{cases} \alpha/2, & \text{if } r = 3, \\ \alpha/\sqrt{\pi r}, & \text{if } r \geq 4. \end{cases}$$

The simplification is that we used the inequality  $r - 3 \geq r/4$  for  $r \geq 4$ . It therefore follows that, for  $r \geq 3$ , there exists a positive universal constant  $C$ , independent of  $r$ , such that  $\mathbb{P}(z \leq Y_{r-1} \leq z + \alpha) \leq C\alpha/\sqrt{r}$ . We obtain an upper bound for  $r \geq 3$  by substituting this inequality into (5.3) and taking a suitable  $\alpha$ . We choose  $\alpha = cr^{5/8}n^{-1/4}$ . From the resulting bound, we obtain the compact bound  $\mathbb{P}(F_r \leq z) - \mathbb{P}(Y_{r-1} \leq z) \leq Cr^{1/8}/n^{1/4} + C'r^{1/4}/n^{1/2}$  by making use of the facts that  $n \geq 1$  and  $r \geq 2$  and Young's inequality for products  $ab \leq a^p/p + b^q/q$ , where  $a, b \geq 0$  and  $p, q > 1$  are such that  $1/p + 1/q = 1$ . We remark that the exponent  $-1/4$  for  $n$  optimises the rate of convergence of the Kolmogorov distance bound with respect to  $n$ . The exponent  $5/8$  for  $r$  is chosen to allow for a Kolmogorov distance bound that converges to zero if  $r^{1/2}/n \rightarrow 0$ ; another choice of exponent would mean that the bound only converges to zero if  $r^\delta/n \rightarrow 0$ , for some  $\delta > 1/2$ . By the same procedure one obtains a lower bound, which is the negative of the upper bound, leading to the bound  $d_K(\mathcal{L}(F_r), \chi_{(r-1)}^2) \leq Cr^{1/8}/n^{1/4} + C'r^{1/4}/n^{1/2}$ . We obtain the compact final bound as stated in the corollary as follows. If  $r^{1/8}/n^{1/4} \geq 1$  and  $C + C' \geq 1$ , then the bound is trivial (the Kolmogorov distance is trivially bounded above by 1), so we may assume



that  $r^{1/8}/n^{1/4} \leq 1$ , and we conclude that there exists a constant  $C_6 > 0$  such that  $d_K(\mathcal{L}(F_r), \chi_{(r-1)}^2) \leq C_6 r^{1/8}/n^{1/4}$ . The proof of the bound for the case  $r \geq 3$  is now complete.  $\square$

## Acknowledgements

Firstly we acknowledge many helpful comments by anonymous referees. Moreover, we would like to thank Persi Diaconis for bringing this problem to our attention.

## Funding

During this research, RG was supported in part by EPSRC grant EP/K032402/1 and a Dame Kathleen Ollerenshaw Research Fellowship. GR has been funded in part by EPSRC grants EP/K032402/1, EP/T018445/1 and EP/R018472/1.

## Supplementary Material

**Supplement to “Bounds for the chi-square approximation of Friedman’s statistic by Stein’s method”** (DOI: [10.3150/22-BEJ1530SUPP](https://doi.org/10.3150/22-BEJ1530SUPP); .pdf). This supplementary material contains proofs of the lemmas from Section 3, and also a section on converting a general sum over up to 4 indices into a sum over distinct indices.

## References

- [1] Anastasiou, A. and Reinert, G. (2020). Bounds for the asymptotic distribution of the likelihood ratio. *Ann. Appl. Probab.* **30** 608–643. [MR4108117 https://doi.org/10.1214/19-AAP1510](https://doi.org/10.1214/19-AAP1510)
- [2] Gaunt, R.E. and Reinert, G. (2023). Supplement to “Bounds for the chi-square approximation of Friedman’s statistic by Stein’s method.” <https://doi.org/10.3150/22-BEJ1530SUPP>
- [3] Asylbekov, Zh.A., Zubov, V.N. and Ul’yanov, V.V. (2011). On the approximation of some statistics of goodness-of-fit tests for the case of discrete three-dimensional data. *Sibirsk. Mat. Zh.* **52** 571–584. [MR2883210 https://doi.org/10.1134/S0037446611040021](https://doi.org/10.1134/S0037446611040021)
- [4] Azmoodeh, E., Eichelsbacher, P. and Knichel, L. (2020). Optimal gamma approximation on Wiener space. *ALEA Lat. Am. J. Probab. Math. Stat.* **17** 101–132. [MR4057185 https://doi.org/10.30757/alea.v17-05](https://doi.org/10.30757/alea.v17-05)
- [5] Barbour, A.D. (1990). Stein’s method for diffusion approximations. *Probab. Theory Related Fields* **84** 297–322. [MR1035659 https://doi.org/10.1007/BF01197887](https://doi.org/10.1007/BF01197887)
- [6] Barbour, A.D. and Hall, P. (1985). On bounds to the rate of convergence in the central limit theorem. *Bull. Lond. Math. Soc.* **17** 151–156. [MR0806241 https://doi.org/10.1112/blms/17.2.151](https://doi.org/10.1112/blms/17.2.151)
- [7] Braverman, A., Dai, J.G. and Fang, X. (2020). High order steady-state diffusion approximations. [arXiv:2012.02824](https://arxiv.org/abs/2012.02824).
- [8] Chatterjee, S. (2008). A new method of normal approximation. *Ann. Probab.* **36** 1584–1610. [MR2435859 https://doi.org/10.1214/07-AOP370](https://doi.org/10.1214/07-AOP370)
- [9] Chatterjee, S., Fulman, J. and Röllin, A. (2011). Exponential approximation by Stein’s method and spectral graph theory. *ALEA Lat. Am. J. Probab. Math. Stat.* **8** 197–223. [MR2802856 https://doi.org/10.30757/alea.v8-05](https://doi.org/10.30757/alea.v8-05)
- [10] Chatterjee, S. and Meckes, E. (2008). Multivariate normal approximation using exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.* **4** 257–283. [MR2453473 https://doi.org/10.30757/alea.v4-05](https://doi.org/10.30757/alea.v4-05)
- [11] Chen, L.H.Y., Goldstein, L. and Shao, Q.-M. (2011). *Normal Approximation by Stein’s Method. Probability and Its Applications (New York)*. Heidelberg: Springer. [MR2732624 https://doi.org/10.1007/978-3-642-15007-4](https://doi.org/10.1007/978-3-642-15007-4)

- [12] Diaconis, P. and Zabell, S. (1991). Closed form summation for classical distributions: Variations on a theme of de Moivre. *Statist. Sci.* **6** 284–302. [MR1144242](#)
- [13] Dicker, L.H. and Erdogdu, M.A. (2017). Flexible results for quadratic forms with applications to variance components estimation. *Ann. Statist.* **45** 386–414. [MR3611496](#) <https://doi.org/10.1214/16-AOS1456>
- [14] Döbler, C. and Peccati, G. (2018). The gamma Stein equation and noncentral de Jong theorems. *Bernoulli* **24** 3384–3421. [MR3788176](#) <https://doi.org/10.3150/17-BEJ963>
- [15] Fathi, M. (2021). Higher-order Stein kernels for Gaussian approximation. *Studia Math.* **256** 241–258. [MR4172696](#) <https://doi.org/10.4064/sm190415-28-10>
- [16] Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Amer. Statist. Assoc.* **32** 675–701.
- [17] Gaunt, R.E. (2013). *Rates of Convergence of Variance-Gamma Approximations via Stein's Method*. DPhil thesis, University of Oxford.
- [18] Gaunt, R.E. (2020). Stein's method for functions of multivariate normal random variables. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 1484–1513. [MR4076792](#) <https://doi.org/10.1214/19-AIHP1011>
- [19] Gaunt, R.E. (2022). Bounds for the chi-square approximation of the power divergence family of statistics. *J. Appl. Probab.* To appear.
- [20] Gaunt, R.E., Pickett, A.M. and Reinert, G. (2017). Chi-square approximation by Stein's method with application to Pearson's statistic. *Ann. Appl. Probab.* **27** 720–756. [MR3655852](#) <https://doi.org/10.1214/16-AAP1213>
- [21] Gaunt, R.E. and Reinert, G. (2021). Bounds for the chi-square approximation of Friedman's statistic by Stein's method Technical report.
- [22] Goldstein, L. and Reinert, G. (1997). Stein's method and the zero bias transformation with application to simple random sampling. *Ann. Appl. Probab.* **7** 935–952. [MR1484792](#) <https://doi.org/10.1214/aoap/1043862419>
- [23] Goldstein, L. and Reinert, G. (2005). Zero biasing in one and higher dimensions, and applications. In *Stein's Method and Applications. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.* **5** 1–18. Singapore: Singapore Univ. Press. [MR2201883](#) [https://doi.org/10.1142/9789812567673\\_0001](https://doi.org/10.1142/9789812567673_0001)
- [24] Goldstein, L. and Rinott, Y. (1996). Multivariate normal approximations by Stein's method and size bias couplings. *J. Appl. Probab.* **33** 1–17. [MR1371949](#) <https://doi.org/10.1017/s0021900200103675>
- [25] Götze, F. (1991). On the rate of convergence in the multivariate CLT. *Ann. Probab.* **19** 724–739. [MR1106283](#)
- [26] Götze, F. and Ulyanov, V.V. (2003). *Asymptotic Distribution of  $\chi^2$ -Type Statistics. Research Group Spectral Analysis, Asymptotic Distributions and Stochastic Dynamics*. Bielefeld: Bielefeld Univ. Preprint 03–033.
- [27] Jensen, D.R. (1977). On approximating the distributions of Friedman's  $\chi_r^2$  and related statistics. *Metrika* **24** 75–85. [MR0443173](#) <https://doi.org/10.1007/BF01893394>
- [28] Lefèvre, C. and Utev, S. (2003). Exact norms of a Stein-type operator and associated stochastic orderings. *Probab. Theory Related Fields* **127** 353–366. [MR2018920](#) <https://doi.org/10.1007/s00440-003-0288-9>
- [29] Ley, C., Reinert, G. and Swan, Y. (2017). Stein's method for comparison of univariate distributions. *Probab. Surv.* **14** 1–52. [MR3595350](#) <https://doi.org/10.1214/16-PS278>
- [30] Loh, W.-L. (1992). Stein's method and multinomial approximation. *Ann. Appl. Probab.* **2** 536–554. [MR1177898](#)
- [31] Luk, H.M. (1994). *Stein's Method for the Gamma Distribution and Related Statistical Applications*. Ann Arbor, MI: ProQuest LLC. Thesis (Ph.D.)—University of Southern California. [MR2693204](#)
- [32] Meckes, E. (2009). On Stein's method for multivariate normal approximation. In *High Dimensional Probability V: The Luminy Volume. Inst. Math. Stat. (IMS) Collect.* **5** 153–178. Beachwood, OH: IMS. [MR2797946](#) <https://doi.org/10.1214/09-IMSCOLL511>
- [33] Mijoule, G., Reinert, G. and Swan, Y. (2021). Stein's density method for multivariate continuous distributions. [arXiv:2101.05079](#).
- [34] Nourdin, I. and Peccati, G. (2009). Stein's method on Wiener chaos. *Probab. Theory Related Fields* **145** 75–118. [MR2520122](#) <https://doi.org/10.1007/s00440-008-0162-x>
- [35] Peköz, E.A. and Röllin, A. (2011). New rates for exponential approximation and the theorems of Rényi and Yaglom. *Ann. Probab.* **39** 587–608. [MR2789507](#) <https://doi.org/10.1214/10-AOP559>
- [36] Pickett, A. (2004). *Rates of Convergence of  $\chi^2$  Approximations via Stein's Method*. DPhil thesis, Univ. Oxford.

- [37] Reinert, G. and Röllin, A. (2009). Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition. *Ann. Probab.* **37** 2150–2173. [MR2573554](#) <https://doi.org/10.1214/09-AOP467>
- [38] Shevtsova, I.G. (2014). On absolute constants in inequalities of Berry-Esseen type. *Dokl. Akad. Nauk* **456** 650–654. [MR3287912](#) <https://doi.org/10.1134/s1064562414030338>
- [39] Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability Theory* 583–602. Berkeley, CA: Univ. California Press. [MR0402873](#)
- [40] Stein, C. (1986). *Approximate Computation of Expectations. Institute of Mathematical Statistics Lecture Notes—Monograph Series 7*. Hayward, CA: IMS. [MR0882007](#)
- [41] Ulyanov, V.V. and Zubov, V.N. (2009). Refinement on the convergence of one family of goodness-of-fit statistics to chi-squared distribution. *Hiroshima Math. J.* **39** 133–161. [MR2499200](#)
- [42] Yarnold, J.K. (1972). Asymptotic approximations for the probability that a sum of lattice random vectors lies in a convex set. *Ann. Math. Stat.* **43** 1566–1580. [MR0372967](#) <https://doi.org/10.1214/aoms/1177692389>

*Received November 2021 and revised July 2022*