VANISHING SHORTCOMING AND ASYMPTOTIC RELATIVE EFFICIENCY

By Tadeusz Inglot,[1] Wilbert C. M. Kallenberg, and Teresa Ledwina

Wrocław University of Technology, University of Twente, and Polish Academy of Sciences

The shortcoming of a test is the difference between the maximal attainable power and the power of the test under consideration. Vanishing shortcoming, when the number of observations tends to infinity, is therefore an optimality property of a test. Other familiar optimality criteria are based on the asymptotic relative efficiency of the test. The relations between these optimality criteria are investigated. It turns out that vanishing shortcoming is seemingly slightly stronger than first-order efficiency, but in regular cases there is equivalence. The results are in particular applied to tests for goodness-of-fit.

1. Introduction. Comparison of tests is in principle based on the power of the tests. Exact powers of tests are hard to compute and, if they can be computed, it is in general not easy to compare them for large sets of alternatives simultaneously. Therefore, often an asymptotic approach is applied to simplify things. The simplifications by the asymptotics should be such that computation can be done and comparison can be made. Moreover, the conclusions based on the asymptotics should be in line with the finite sample behavior, which means that the asymptotics should provide good approximations.

A direct way of comparison of two tests is to consider the difference in power of the two tests. In particular, the difference between the most powerful test and a given test is of interest. It is called the *shortcoming* of that test. This concept is used to express *optimality* of a test: if the shortcoming of a test tends to 0 when the number of observations n tends to infinity, the test is asymptotically optimal.

Another, indirect, method of comparison of tests is based on the number of observations $N(\alpha, \beta, \theta)$ needed to get power β at the alternative θ when the level of the test equals α. If we have two tests with corresponding numbers $N_1(\alpha, \beta, \theta)$ and $N_2(\alpha, \beta, \theta)$, the ratio $N_2(\alpha, \beta, \theta)/N_1(\alpha, \beta, \theta)$ is called the *relative efficiency* of test 1 w.r.t. test 2. If the relative efficiency equals r, test 2 needs r times as many observations to perform equally well as test 1 and hence test 1 is called r times as efficient as test 2. To investigate *optimality* we consider $N_*(\alpha, \beta, \theta)/N(\alpha, \beta, \theta)$, where $N_*(\alpha, \beta, \theta)$ corresponds to the most powerful test.

Again, an asymptotic approach is desirable to simplify the calculation and evaluation, since $N_i(\alpha, \beta, \theta)$ depends on three parameters. When sending n to infinity, two principles are (a) to "decrease the significance probability as n increases," that is, to send α to 0, or (b) to "move the alternative hypothesis steadily closer to the null hypothesis," that is, to send θ to the null hypothesis H_0. Both principles are attractive: with more observations it seems reasonable to have a stronger requirement on the level and on the other hand, for alternatives far away from H_0 there is no need for statistical methods, since those alternatives are obviously different from H_0. In Pitman's asymptotic efficiency concept, method (b) is used, while one deals with fixed levels, thus ignoring principle (a). In Bahadur's asymptotic efficiency concept, method (a) is actually used, while fixed alternatives are under consideration, thereby ignoring principle (b). Intermediate or Kallenberg efficiency, as defined in Kallenberg (1983a), applies *both* attractive principles simultaneously.

Optimality of a test can be expressed by first-order efficiency, which means that $N_*(\alpha, \beta, \theta)/N(\alpha, \beta, \theta)$ converges to 1, where the limit is taken according to the efficiency concept involved.

It is the purpose of this paper to investigate the relations between the asymptotic optimality concepts based on vanishing shortcoming on the one hand and the three asymptotic efficiency concepts on the other hand, where the shortcoming is considered for the same limiting process as is involved in the asymptotic efficiency concept. It turns out that asymptotic vanishing shortcoming is in the Pitman case equivalent to first-order efficiency, while in the Bahadur and intermediate case vanishing shortcoming seems slightly stronger than first-order efficiency. However, in regular cases first-order efficient tests do also have asymptotic vanishing shortcoming. Here is a parallel with the phenomenon of "first-order efficiency implies second-order efficiency" [cf. Bickel, Chibisov and van Zwet (1981) and Kallenberg (1983a)].

The main results on the relationship between vanishing shortcoming and first-order efficiency are very general: there is a very general set-up, very mild conditions on the most powerful tests and no condition at all on the (form of the) first-order efficient tests. Moreover, the results hold for all three efficiency concepts.

The paper is organized as follows. Section 2 contains notation, definitions and basic assumptions. The main results describing in great generality the relations between vanishing shortcoming and asymptotic relative efficiency are given in Section 3. These results are based on an asymptotic expression for $N_*(\alpha, \beta, \theta)$ and on an investigation of the number of extra observations needed to get a gain in asymptotic power for the most powerful test. These theorems (Theorems 3.2, 3.3, 3.2′ and 3.3′) and their extensions to general (regular) tests (Theorems 4.1 and 5.3) may be of independent interest. The proofs of the theorems of Section 3 are given in an Appendix. Some examples in Section 3 show the great generality of the results. Apart from an asymptotic expression for $N(\alpha, \beta, \theta)$, when the test is based on a (regular) test statistic, it is shown in Section 4 that as a rule vanishing shortcoming and first-order efficiency are equivalent. Section 5 is devoted to further elaboration of the examples of

Section 3. Applications are made to some first-order efficient tests in these cases with special attention to tests for goodness-of-fit. In particular, some useful formulas for their asymptotic relative efficiencies are derived, showing that both for the Cramér–von Mises test and for the Anderson–Darling test, first-order efficiency is only attained in one direction. In all other directions the asymptotic relative efficiency is less than one and often much lower. Moreover, these applications give a nice illustration of the phenomenon that equality of asymptotic optimal shift implies also equality of scale terms.

2. Notation, definitions and basic assumptions. Let $\mathscr{S}$ be a space of points s, $\mathscr{A}$ a σ-field of subsets of $\mathscr{S}$ and for each point θ in a set Θ let P_θ be a probability measure on $\mathscr{A}$. The random element S with values in $\mathscr{S}$ is distributed according to P_θ. In typical cases $S = (X_1, X_2, \ldots)$ is a sequence of i.i.d. r.v.'s, but as yet $(\mathscr{S}, \mathscr{A}, P_\theta)$ is a quite general probability space. Note that θ is an abstract "parameter" and hence the results apply as well to parametric, nonparametric or semiparametric testing problems.

Suppose the hypothesis H_0: $\theta \in \Theta_0$ has to be tested against H_1: $\theta \in \Theta_1 \subset \Theta - \Theta_0$. Let $\{\psi_{n,\alpha}: n \in \mathbb{N}, \ 0 < \alpha < 1\}$ be a family of (randomized) level-α tests of H_0, that is, for each $n \in \mathbb{N}$ and $0 < \alpha < 1$ the function $\psi_{n;\alpha}$ is a measurable function on $\mathscr{S}$ with values in $[0, 1]$ satisfying

$$\sup\left\{E_{\theta_0}\psi_{n;\alpha}:\ \theta_0 \in \Theta_0\right\} \le \alpha.$$

Here n refers to the number of observations, implicitly implying that with increasing n more information about θ comes available. Under very weak conditions a most powerful (MP) test of H_0 against the simple alternative H_1^*: $\theta = \theta_1$ exists [cf. Lehmann (1986), Section 3.8 and page 576]. In general, such a MP test depends on the particular alternative. The existence of such MP tests will be assumed in the sequel. They will be denoted by $\{\psi_{n;\alpha}^*\}$, suppressing their dependence on the particular alternative θ_1. In case of ambiguity we write $\psi_{n;\alpha}^*(S; \theta)$.

The power $E_\theta\psi_{n;\alpha}$ of the level-α test $\psi_{n;\alpha}$ is denoted by $\beta_n(\alpha, \theta)$. When taking the supremum over all level-α tests, we get the envelope power function, denoted by $\beta_n^*(\alpha, \theta)$. Thus

$$\beta_n^*(\alpha, \theta) = \sup_\psi \beta_n(\alpha, \theta),$$

where $\psi = \{\psi_{n;\alpha}\}$ runs through all families of level-α tests of H_0. Obviously, $\beta_n^*(\alpha, \theta)$ is the power of $\psi_{n;\alpha}^*$.

The *shortcoming* of $\psi_{n;\alpha}$ is defined by

(2.1) $$R_n(\alpha, \theta) = \beta_n^*(\alpha, \theta) - \beta_n(\alpha, \theta).$$

Asymptotic relative efficiency is defined in terms of $N(\alpha, \beta, \theta)$, which is the smallest number of observations N such that the level-α test $\psi_{m;\alpha}$ has power at least β at the alternative θ for all $m \ge N$. Formally,

$$N(\alpha, \beta, \theta) = \inf\{N: \beta_m(\alpha, \theta) \geq \beta \text{ for all } m \geq N\}.$$

In the case of the MP test of H_0 against the simple alternative θ we write $N_*(\alpha, \beta, \theta)$.

For given sequences $\{\alpha_n\}$ *and* $\{\theta_n\}$ with $0 < \alpha_n < 1$ and $\theta_n \in \Theta_1$, we want to relate the shortcoming $R_n(\alpha_n, \theta_n)$ with $N(\alpha_n, \beta, \theta_n) - N_*(\alpha_n, \beta, \theta_n)$, where $0 < \beta < 1$. If $\alpha_n \to \alpha \in (0, 1)$ and θ_n converges to Θ_0, we speak of Pitman efficiency. If $\alpha_n \to 0$ and θ_n is fixed, we deal with Bahadur efficiency. The case $\alpha_n \to 0$ and θ_n converging to Θ_0 is called intermediate or Kallenberg efficiency. Throughout the paper $\{\theta_n\}$ and $\{\alpha_n\}$ are given sequences with $0 < \alpha_n < 1$, $\lim_{n\to\infty} \alpha_n = \bar{\alpha} \in [0, 1)$ and $\theta_n \in \Theta_1$.

As n is the number of available observations, $\psi^*_{n;\alpha}$ can be considered for $m \geq n$ also as a test based on m observations for testing H_0, simply obtained by ignoring the $m - n$ extra observations. Since $\psi^*_{m;\alpha}$ is the MP test based on m observations, we have for all $0 < \alpha < 1$ and $\theta \in \Theta_1$,

$$(2.2) \qquad \beta^*_m(\alpha, \theta) = E_\theta \psi^*_{m;\alpha}(S; \theta) \geq \beta^*_n(\alpha, \theta) = E_\theta \psi^*_{n;\alpha}(S; \theta) \quad \text{if } m \geq n.$$

We assume that the MP test of H_0 against the simple alternative θ for each $n \in \mathbb{N}$ and $0 < \alpha < 1$ is based on a real-valued test statistic T^*_n (which as a rule will depend on θ), rejecting for large values of T^*_n. More precisely, the level-α MP test of H_0 against θ is given by

(A1)

$$(2.3) \qquad \psi^*_{n;\alpha}(s) = \begin{cases} 1, & \text{if } T^*_n(s) > c_n, \\ \delta_n, & \text{if } T^*_n(s) = c_n, \\ 0, & \text{if } T^*_n(s) < c_n, \end{cases}$$

where $c_n = c_n(\alpha) = \inf\{c: \sup\{P_{\theta_0}(T^*_n(S) > c): \theta_0 \in \Theta_0\} \leq \alpha\}$ and $\delta_n = \delta_n(\alpha) = \sup\{\delta \in [0, 1]: \sup\{P_{\theta_0}(T^*_n(S) > c_n) + \delta P_{\theta_0}(T^*_n(S) = c_n): \theta_0 \in \Theta_0\} \leq \alpha\}$. Then we have for all $c < c_n$,

$$(2.4) \qquad \sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T^*_n(S) > c_n) \leq \alpha < \sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T^*_n(S) > c).$$

The next assumption concerns the behavior of the test statistics T^*_n, given in (2.3).

(A2) There exists a function $\mu_*: \Theta_1 \to (0, \infty)$ such that $\limsup_{n\to\infty} \mu_*(\theta_n) < \infty$ and additionally $\lim_{n\to\infty} \mu_*(\theta_n) = 0$ in case $\bar{\alpha} > 0$. Moreover, there exists a continuous distribution function G_* on $\mathbb{R}$, strictly increasing on its support, which is $\mathbb{R}$ or $[a, \infty)$ with $a \in \mathbb{R}$, such that for every sequence $N = N(n)$ of natural numbers satisfying

$$(2.5) \qquad \begin{aligned} &\sqrt{N}\mu_*(\theta_n) - G_*^{-1}(1 - \alpha_n) = O(1) \quad \text{and} \\ &\liminf_{n\to\infty} \sqrt{N}\mu_*(\theta_n) > 0 \text{ in case } \bar{\alpha} > 0, \end{aligned}$$

we have (with T_N^* the MP test of H_0 against θ_n)

$$(2.6) \qquad \lim_{n \to \infty} P_{\theta_n}(T_N^* - \sqrt{N}\mu_*(\theta_n) \le x) = G_*(x) \quad \text{for every } x \in \mathbb{R},$$

and, as $n \to \infty$,

$$(2.7) \qquad \sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T_N^* > t_n) = 1 - G_*(t_n + o(1))$$
$$\text{for all } t_n = \sqrt{N}\mu_*(\theta_n) + c$$

with $c \in \mathbb{R}$, not depending on n.

In the Pitman case ($\bar{\alpha} > 0$) (2.5) describes the typical local alternatives, basically of order $N^{-1/2}$. If $\bar{\alpha} = 0$, the alternative is farther away from the null hypothesis, since $G_*^{-1}(1 - \alpha_n) \to \infty$ if $\alpha_n \to 0$.

In typical cases T_N^* is a (standardized) sum of i.i.d. r.v.'s, G_* is the standard normal distribution function and (2.6) follows by a central limit theorem. Moreover, in the Pitman case ($\bar{\alpha} > 0$) (2.7) also follows from a central limit theorem, while in the intermediate case (2.7) can be obtained from a moderate deviation theorem.

Condition (A2) is appropriate for the more local situations as Pitman efficiency and, partly, intermediate efficiency. In the nonlocal case the following condition is suitable.

(A2′) Let $\bar{\alpha} = 0$. There exist functions $\mu_{n*}: \Theta_1 \to (0, \infty)$ such that $\limsup_{n \to \infty} \mu_{n*}(\theta_n) < \infty$. Further, there exist functions r_{1n*} and r_{2n*}, defined on an open interval containing the limiting points of $\{\mu_{n*}(\theta_n)\}$, satisfying for $i = 1, 2$,

$$(2.8) \qquad 0 < b_1 \le \frac{r_{in*}(x)}{x} \le b_2 < \infty$$

for all x and some constants b_1, b_2. Moreover, there exists a continuous distribution function G_* on $\mathbb{R}$, strictly increasing on its support, which is $\mathbb{R}$ or $[a, \infty)$ with $a \in \mathbb{R}$, such that for every sequence $N = N(n)$ of natural numbers satisfying

$$(2.9) \qquad \sqrt{N}\, r_{1n*}(\mu_{n*}(\theta_n)) - |2 \log \alpha_n|^{1/2} = O(1)$$

we have (with T_N^* the MP test of H_0 against θ_n)

$$(2.10) \qquad \lim_{n \to \infty} P_{\theta_n}\left(T_N^* - \sqrt{N}\mu_{n*}(\theta_n) \le x\right) = G_*(x) \quad \text{for every } x \in \mathbb{R},$$

and, as $n \to \infty$,

$$-N^{-1} \log \left\{ \sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T_N^* > t_n) \right\}$$

$$(2.11) \qquad = \tfrac{1}{2} r_{1n*}^2(\mu_{n*}(\theta_n)) + d_n N^{-1/2} r_{2n*}(\mu_{n*}(\theta_n)) + o(N^{-1/2}\mu_{n*}(\theta_n))$$
$$\text{for all } t_n = \sqrt{N}\mu_{n*}(\theta_n) + d_n$$

with $\{d_n\} \subset \mathbb{R}$ being any bounded sequence.

REMARK 2.1.   As a rule r_{2n*} is the derivative of $\frac{1}{2}r_{1n*}^2$. In typical cases,

$$r_{in*}(\mu_{n*}(\theta_n))/\mu_{n*}(\theta_n) \to 1 \quad \text{if } \theta_n \to \Theta_0;$$

see also Example 3.6 in Inglot, Kallenberg, and Ledwina. (1998). This gives the connection between (A2) and (A2′) if G_* is the standard normal distribution function [cf. also (3.8)].

Note that also in (A2′) G_* is often the standard normal distribution function, but in the Bahadur case the variance as a rule is not equal to 1. Checking of (A2) and (A2′) is exemplified in Examples 3.4–3.7.   □

As in (A2), if T_N^* is a (standardized) sum of i.i.d. r.v.'s, (2.10) usually follows from a central limit theorem, while (2.11) can be obtained from a moderate or large deviation theorem.

3. Main results.   The idea behind the relationship between vanishing shortcoming and optimality in the sense of asymptotic relative efficiency is as follows. The asymptotic shift of the MP test statistic T_n^* for testing H_0 against the simple alternative θ_n is equal to $\sqrt{n}\mu_*(\theta_n)$. To obtain with k additional observations asymptotically a gain in power at θ_n requires $\liminf_{n\to\infty}(\sqrt{n+k}-\sqrt{n})\mu_*(\theta_n) > 0$. Therefore, vanishing shortcoming corresponds to an additional number k of observations satisfying $(\sqrt{n+k}-\sqrt{n})\mu_*(\theta_n) = o(1)$, or, equivalently, an additional number of observations of order $o(\sqrt{n}/\mu_*(\theta_n))$.

In the following theorem the latter result is indeed established under (A1) and (A2). For the proofs of the theorems of this section we refer to the Appendix.

THEOREM 3.1.   *Assume* (A1) *and* (A2). *The following statements are equivalent*:

(i) $\lim_{n\to\infty} R_m(\alpha_n, \theta_n) = 0$ *for each sequence* $m = m(n)$, *provided that in case* $\bar{\alpha} > 0$, $\liminf_{n\to\infty} \beta_m^*(\alpha_n, \theta_n) > \bar{\alpha}$,

(ii) $N(\alpha_n, \beta, \theta_n) - N_*(\alpha_n, \beta, \theta_n) = o\left(\sqrt{N_*(\alpha_n, \beta, \theta_n)}/\mu_*(\theta_n)\right)$ *for each* $\beta \in (\bar{\alpha}, 1)$.

The related version of this theorem under (A1) and (A2′) is as follows.

THEOREM 3.1.   *Assume* (A1) *and* (A2′). *The following statements are equivalent*:

(i) $\lim_{n\to\infty} R_m(\alpha_n, \theta_n) = 0$ *for each sequence* $m = m(n)$.
(ii) $N(\alpha_n, \beta, \theta_n) - N_*(\alpha_n, \beta, \theta_n) = o\left(\sqrt{N_*(\alpha_n, \beta, \theta_n)}/\mu_{n*}(\theta_n)\right)$ *for each* $\beta \in (0, 1)$.

We give some comments on the results. First-order optimality in the sense of asymptotic relative efficiency means

$$(3.1) \qquad N_*(\alpha_n, \beta, \theta_n)/N(\alpha_n, \beta, \theta_n) \to 1.$$

In the Pitman case, typically $\mu_*(\theta_n)$ is of exact order $\{N_*(\alpha_n, \beta, \theta_n)\}^{-1/2}$ and vanishing shortcoming corresponds to first-order optimality in the sense of (3.1).

For Bahadur efficiency, $\theta_n = \theta$ is fixed and vanishing shortcoming seems to be a stronger property than first-order optimality in the sense of asymptotic relative efficiency, since $N(\alpha_n, \beta, \theta)$ should not only have the same first-order term, but also the same $\sqrt{N}$-term as $N_*(\alpha_n, \beta, \theta)$. So, it seems that first-order optimality in the sense of Bahadur efficiency, that is, (3.1) with $\theta_n = \theta$, is not sufficient to guarantee vanishing shortcoming as (3.1) does not automatically imply (ii) in case $\theta_n = \theta$. However, in regular cases it turns out, by a simple argument, that nevertheless most tests which are first-order efficient in the sense of Bahadur do also have vanishing shortcoming. The same argument applies to intermediate or Kallenberg efficiency. In Section 4 more details are given.

The next theorem provides a formula for $N_*(\alpha_n, \beta, \theta_n)$.

THEOREM 3.2. *Assume* (A1) *and* (A2). *For all* $\beta \in (\bar{\alpha}, 1)$ *we have*

$$(3.2) \qquad N_*(\alpha_n, \beta, \theta_n) = \left\{ \frac{G_*^{-1}(1 - \alpha_n) - G_*^{-1}(1 - \beta) + o(1)}{\mu_*(\theta_n)} \right\}^2.$$

Under (A1) and (A2′) Theorem 3.2 takes the following form.

THEOREM 3.2′. *Assume* (A1) *and* (A2′). *For all* $\beta \in (0, 1)$,

$$N_*(\alpha_n, \beta, \theta_n)$$
$$= \left\{ \frac{|2\log\alpha_n|^{1/2} - \{r_{2n*}(\mu_{n*}(\theta_n))/r_{1n*}(\mu_{n*}(\theta_n))\}G_*^{-1}(1-\beta) + o(1)}{r_{1n*}(\mu_{n*}(\theta_n))} \right\}^2.$$

In view of Remark 2.1 we get in typical cases the same expression [up to $o(1)$] for $\sqrt{N_*(\alpha_n, \beta, \theta_n)}\,\mu_*(\theta_n)$ in Theorem 3.2 as for $\sqrt{N_*(\alpha_n, \beta, \theta_n)} \times r_{1n*}(\mu_{n*}(\theta_n))$ in Theorem 3.2′ if G_* is the standard normal distribution function [cf. also (3.8)].

The next theorem specifies the argument mentioned in the first lines of this section: "to obtain with k additional observations asymptotically a gain in power of the MP test at θ_n requires $\lim_{n\to\infty}(\sqrt{n+k} - \sqrt{n})\mu_*(\theta_n) > 0$."

THEOREM 3.3.   *Assume* (A1) *and* (A2). *For each* $\beta \in (\bar{\alpha}, 1)$ *and for each sequence* $m = m(n)$, *we have*

(3.3)
$$\lim_{n \to \infty} \left[\beta_m^*(\alpha_n, \theta_n) \right.$$
$$\left. - \left\{ 1 - G_* \left(G_*^{-1}(1 - \beta) - \left[\sqrt{m} - \sqrt{N_*(\alpha_n, \beta, \theta_n)} \right] \mu_*(\theta_n) \right) \right\} \right] = 0.$$

A counterpart of Theorem 3.3 when (A1) and (A2′) hold is as follows.

THEOREM 3.3′.   *Assume* (A1) *and* (A2′). *Then we have for each* $\beta \in (0, 1)$ *and for each sequence* $m = m(n)$,

$$\lim_{n \to \infty} \left[\beta_m^*(\alpha_n, \theta_n) \right.$$
$$\left. - \left\{ 1 - G_* \left(G_*^{-1}(1 - \beta) - \left[\sqrt{m} - \sqrt{N_*(\alpha_n, \beta, \theta_n)} \right] \frac{r_{1n*}^2(\mu_{n*}(\theta_n))}{r_{2n*}(\mu_{n*}(\theta_n))} \right) \right\} \right] = 0.$$

As is seen from the definitions and basic assumptions in Section 2, there are no conditions or assumptions on the tests ψ. This means that Theorems 3.1 and 3.1 are in fact properties of MP tests.

We end this section by presenting some testing problems, where assumptions (A1) and (A2) hold and hence Theorem 3.1 can be applied. In Example 3.7 we also show that (A1) and (A2′) hold and apply Theorem 3.1 in the given context.

EXAMPLE 3.4.   *Gauss-test.*   Let $X_1, \ldots, X_n$ be i.i.d. r.v.'s each with a normal $N(\theta, 1)$-distribution. Consider the testing problem,

$$H_0: \theta = 0 \text{ against } H_1: \theta > 0.$$

It is easily seen that (A1) and (A2) hold with $T_N^* = \overline{X}\sqrt{N}$, where $\overline{X} = N^{-1} \sum_{i=1}^{N} X_i$, $G_* = \Phi$, the standard normal distribution function and $\mu_*(\theta) = \theta$. Theorem 3.1 can be applied for all sequences $\{\theta_n\}$ and $\{\alpha_n\}$ with $\theta_n > 0$ and $\lim_{n \to \infty} \alpha_n = \bar{\alpha} \in [0, 1)$, provided of course that $\theta_n \to 0$ if $\bar{\alpha} > 0$.

EXAMPLE 3.5.   *Composite null hypothesis.*   Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be i.i.d. r.v.'s with X_1 a normal $N(\theta^{(1)}, 1)$-distribution, Y_1 a normal $N(\theta^{(2)}, 1)$-distribution and X_1 and Y_1 independent. Consider the testing problem,

$$H_0: \theta^{(1)} \leq 0, \ \theta^{(2)} \leq 0 \text{ against } H_1: \theta^{(1)} > 0 \text{ or } \theta^{(2)} > 0.$$

Theorem 7 of Section 3.8 in Lehmann (1986) and normality of the observations easily yield (A1) and (A2) with $G_* = \Phi$,

$$T_N^* = \overline{X}\sqrt{N} \text{ and } \mu_*(\theta) = \theta^{(1)} \text{ if } \theta^{(1)} > 0 \text{ and } \theta^{(2)} \leq 0,$$
$$T_N^* = \overline{Y}\sqrt{N} \text{ and } \mu_*(\theta) = \theta^{(2)} \text{ if } \theta^{(1)} \leq 0 \text{ and } \theta^{(2)} > 0,$$

$$T_N^* = \frac{\theta^{(1)}\overline{X} + \theta^{(2)}\overline{Y}}{\sqrt{\theta^{(1)^2} + \theta^{(2)^2}}}\sqrt{N} \text{ and } \mu_*(\theta) = \sqrt{\theta^{(1)^2} + \theta^{(2)^2}} \text{ if } \theta^{(1)} > 0 \text{ and } \theta^{(2)} > 0.$$

Theorem 3.1 can be applied for all sequences of alternatives $\{\theta_n\}$ and all sequences of levels $\{\alpha_n\}$ with $\lim_{n\to\infty} \alpha_n = \bar{\alpha} \in [0, 1)$, provided of course that $\limsup_{n\to\infty} \mu_*(\theta_n) < \infty$ and $\lim_{n\to\infty} \mu_*(\theta_n) = 0$ in case $\bar{\alpha} > 0$.

For extension of this example to testing composite hypotheses in general exponential families, we refer to Lemmas 3.6, 3.7 and 3.8 in Kallenberg (1981b), which concern a construction and properties of the most powerful test of a composite hypothesis against a simple alternative in exponential families.

EXAMPLE 3.6. Curved exponential families. Let $X_1, \ldots, X_n$ be i.i.d. k-dimensional r.v.'s with density

$$p_\theta(x) = \exp\{\gamma_\theta' x - \psi(\gamma_\theta)\},$$

$\theta \in \Theta$, with respect to a σ-finite measure ν on $\mathbb{R}^k$. Here $\gamma_\theta' x$ denotes the inner product of γ_θ and x, while $\psi(\gamma_\theta) = \log \int \exp(\gamma_\theta' x)\, d\nu(x)$. Set $\Gamma = \{\gamma \in \mathbb{R}^k: \psi(\gamma) < \infty\}$, where $\psi(\gamma) = \log \int \exp(\gamma' x)\, d\nu(x)$. Assume $\operatorname{int}\Gamma \neq \phi$, Θ is an open interval in $\mathbb{R}^1$, while γ_θ is a differentiable bijection from Θ on $\gamma(\Theta) \subset \Gamma$. Note that $p_\theta(x)$ is a curved exponential family in the terminology of Efron (1975).

Consider the testing problem H_0: $\theta = \theta_0$ against H_1: $\theta > \theta_0$, where $\theta_0 \in \Theta$ is given. Additionally to the above, assume that $\gamma_{\theta_0} \in \operatorname{int}\Gamma$ and that the covariance matrix of X_1 under p_{θ_0} is nonsingular.

Let $\{\alpha_n\}$ be a sequence of levels satisfying

(3.4) $$\Phi^{-1}(1 - \alpha_n) = o(n^{1/4}) \quad \text{as } n \to \infty,$$

or, equivalently,

(3.5) $$|\log \alpha_n| = o(n^{1/2}) \quad \text{as } n \to \infty.$$

Let $\{\theta_n\}$ be a sequence of alternatives satisfying

(3.6) $$\lim_{n\to\infty} \theta_n = \theta_0.$$

Condition (A1) holds with

$$T_n^* = n^{-1/2} \sum_{i=1}^n (\gamma_{\theta_n} - \gamma_{\theta_0})'(X_i - E_{\theta_0}X)/\{\operatorname{var}_{\theta_0}(\gamma_{\theta_n} - \gamma_{\theta_0})'X\}^{1/2}.$$

Setting

$$\mu_*(\theta) = (\gamma_\theta - \gamma_{\theta_0})'(E_\theta X - E_{\theta_0}X)/\{\operatorname{var}_{\theta_0}(\gamma_\theta - \gamma_{\theta_0})'X\}^{1/2},$$

it is shown in Inglot, Kallenberg and Ledwina (1998) that Liapounov's theorem yields

$$\lim_{n\to\infty} P_{\theta_n}\left(T_M^* - \sqrt{M}\mu_*(\theta_n) \le x\right) = \Phi(x), \qquad x \in \mathbb{R},$$

for an arbitrary sequence $M = M(n)$ tending to infinity and $\theta_n \to \theta_0$. This, in particular, implies that (2.6) of (A2) holds with $G_* = \Phi$.

Exploiting the continuity of $\psi(\gamma)$ on int Γ, in a similar way as done in the proof of Theorem 5.8 in Inglot and Ledwina (1996), we can apply the Cramér-type large deviation result for triangular arrays obtained by Book (1976) [cf. Lemma 4.1 in Jurečková, Kallenberg, and Veraverbeke. (1988)]. This yields

$$(3.7) \qquad P_{\theta_0}(T_N^* > t_n) = \exp\{-\tfrac{1}{2}t_n^2 + o(t_n)\}$$

if $t_n \to \infty$ and $t_n = \sqrt{N}\mu_*(\theta_n) + c$ with $c \in \mathbb{R}$ and N of the form (2.5). [Note that to justify the $o(t_n)$ in (3.7) we need $t_n = o(n^{1/4})$ and this is one of the reasons to require (3.4).] Combination of

$$(3.8) \qquad \Phi^{-1}(u) = \sqrt{-2\log(1-u)} + o(1) \quad \text{as } u \to 1,$$

with (3.7) gives (2.7) if $t_n \to \infty$. For bounded t_n (2.7) follows from the central limit theorem. Since $\theta_n \to \theta_0$, we have always $\lim_{n\to\infty} \mu_*(\theta_n) = 0$. Therefore, (A2) holds for this testing problem and Theorem 3.1 can be applied for all sequences $\{\theta_n\}$ and $\{\alpha_n\}$ satisfying (3.4) [or (3.5)] and (3.6).

EXAMPLE 3.7.   Goodness-of-fit.   Let $X_1, \ldots, X_n$ be i.i.d. r.v.'s with values in $[0, 1]$. The null hypothesis states that the X_i have a uniform distribution on $[0, 1]$. Note that by an application of the integral transformation this is no loss of generality in the goodness-of-fit testing problem for a simple null hypothesis.

For simplicity of presentation we restrict attention to alternatives defined by densities (with respect to Lebesgue measure on $[0, 1]$),

$$(3.9) \qquad p_n(x) = 1 + \theta_n a(x), \qquad \theta_n > 0,$$

with $\theta_n \to 0 = \theta_0$ as $n \to \infty$ and where a is bounded, $\int_0^1 a(x)\,dx = 0$ and $\int_0^1 a^2(x)\,dx = 1$. The MP test of H_0: $p(x) = 1$ against the simple alternative p_n, given by (3.9), satisfies (A1) with

$$T_n^* = (\sqrt{n}\sigma_{0n})^{-1} \sum_{i=1}^n \{\log p_n(X_i) - e_{0n}\},$$

where

$$e_{0n} = E_{\theta_0} \log p_n(X), \qquad \sigma_{0n}^2 = \text{var}_{\theta_0} \log p_n(X).$$

To check (A2), define $G_* = \Phi$ and $\mu_*(\theta) = \theta$. Consider sequences of levels $\{\alpha_n\}$ and alternatives $\{\theta_n\}$ such that

$$(3.10) \qquad \lim_{n\to\infty} \alpha_n = 0, \quad \lim_{n\to\infty} \theta_n = 0 \quad \text{and} \quad \lim_{n\to\infty} \theta_n \Phi^{-1}(1-\alpha_n) = 0.$$

According to (2.5) take

$$N = \{\Phi^{-1}(1-\alpha_n) + O(1)\}^2/\theta_n^2.$$

By Lemma 5.4 and Proposition 6.6 in Inglot and Ledwina (1996) we get (2.6). Indeed, due to (3.10) we get $\sqrt{N}(\theta_n - b^{(1)}(p_n)) \to 0$, where $b^{(1)}$ is given by (5.18) of Inglot and Ledwina (1996). In view of Proposition 5.12 and Lemma 5.4 in Inglot and Ledwina (1996), we can apply Book's (1976) result as in the previous example. The assumption $\theta_n \Phi^{-1}(1 - \alpha_n) \to 0$ [see (3.10)] yields $\sqrt{N}\theta_n = o(N^{1/4})$ and hence Book's result gives [cf. also (3.7)]

$$P_{\theta_0}(T_N^* > t_n) = \exp\{-\tfrac{1}{2} t_n^2 + o(t_n)\} = 1 - \Phi(t_n + o(1))$$

for all $t_n = \sqrt{N}\mu_*(\theta_n) + c$ with $c \in \mathbb{R}$. Therefore, (A2) holds and Theorem 3.1 can be applied for all sequences $\{\theta_n\}$ and $\{\alpha_n\}$ satisfying (3.10).

Next it will be shown that (A2′) holds for all sequences of levels $\{\alpha_n\}$ and alternatives $\{\theta_n\}$ such that

$$\lim_{n \to \infty} \alpha_n = 0 \quad \text{and} \quad \lim_{n \to \infty} \theta_n = \theta_1 \in \mathbb{R}.$$

We assume that $\inf\{1 + \theta_1 a(x) : x \in [0,1]\} > 0$. Note that θ_1 can be 0 (intermediate case) or unequal to 0 (Bahadur case). So, in the intermediate case this is no restriction at all and in the Bahadur case this is only slightly more than stating that the alternative is well defined. Define

$$\mu_{n*}(\theta) = \sigma_{0n}^{-1}\{E_\theta \log p_n(X) - e_{on}\}.$$

We have $\mu_{n*}(\theta) \geq 0$ for all $\theta \geq 0$. Next set

$$G_*(x) = \Phi(x/\sigma_{\theta_1}) \quad \text{with}$$

$$\sigma_\theta^2 = \begin{cases} \dfrac{\text{var}_\theta\{\log(1 + \theta a(X))\}}{\text{var}_{\theta_0}\{\log(1 + \theta a(X))\}}, & \text{if } \theta \neq \theta_0 = 0, \\ 1, & \text{if } \theta = \theta_0 = 0. \end{cases}$$

Using Liapounov's theorem [as in the proof of Proposition 6.6 in Inglot and Ledwina (1996)] we have for any sequence $N = N(n) \to \infty$ that

$$\lim_{n \to \infty} P_{\theta_n}(T_N^* - \sqrt{N}\mu_{n*}(\theta_n) \leq x) = G_*(x) \quad \text{for every } x \in \mathbb{R},$$

which proves (2.10) of (A2′). A detailed proof of the rest of (A2′) is given in Inglot, Kallenberg, and Ledwina (1998). The proof is based on exponential tilting techniques and delicate estimation of integrals as in the proof of Lemma 3.2 in Kallenberg (1981b).

Therefore, Theorem 3.1 can be applied in the Pitman case, using (A2), and in the whole intermediate and Bahadur case we can apply Theorem 3.1, using (A2′).

REMARK 3.8. It is seen in Example 3.7 that the whole range of sequences $\{\alpha_n\}$ and $\{\theta_n\}$ under consideration in Pitman, Bahadur, and intermediate efficiency is covered. Moreover, the method applied in Example 3.7 can be generalized to other testing problems where MP tests of Θ_0 against a simple alternative are based on sums of i.i.d. r.v.'s, provided that the moment generating function of the involved r.v. exists on a sufficiently large interval.

The examples will be investigated further in Section 5, where Theorem 3.1 will be applied to several first-order efficient tests after the general discussion on these tests in Section 4.

4. First-order efficient tests. In this section we present an expansion for $N(\alpha_n, \beta, \theta_n)$ (a counterpart of Theorem 3.2) when ψ is based on a test statistic. The result is applied to show that as a rule vanishing shortcoming is equivalent to first-order efficiency.

The basic assumptions of this section are modifications of (A1) and (A2) and an extra condition to replace (2.2), which obviously holds for MP tests, but not automatically for other tests. Condition (A1) is replaced by (B1), which is obtained from (A1) by writing $\psi_{n;\alpha}$ and T_n instead of $\psi_{n;\alpha}^*$ and T_n^*, that is,

(B1)

$$(4.1) \qquad \psi_{n;\alpha}(s) = \begin{cases} 1, & \text{if } T_n(s) > c_n, \\ \delta_n, & \text{if } T_n(s) = c_n, \\ 0, & \text{if } T_n(s) < c_n, \end{cases}$$

where $c_n = c_n(\alpha) = \inf\{c\colon \sup\{P_{\theta_0}(T_n(S) > c)\colon \theta_0 \in \Theta_0\} \leq \alpha\}$ and $\delta_n = \delta_n(\alpha) = \sup\{\delta \in [0,1]\colon \sup\{P_{\theta_0}(T_n(S) > c_n) + \delta P_{\theta_0}(T_n(S) = c_n)\colon \theta_0 \in \Theta_0\} \leq \alpha\}$. Then we have for all $c < c_n$,

$$\sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T_n(S) > c_n) \leq \alpha < \sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T_n(S) > c).$$

Condition (A2) is replaced by the following one.

(B2) There exists a function $\mu\colon \Theta_1 \to (0, \infty)$ such that $\lim \sup_{n\to\infty} \mu(\theta_n) < \infty$ and additionally $\lim_{n\to\infty} \mu(\theta_n) = 0$ in case $\bar{\alpha} > 0$. Moreover, there exist continuous distribution functions G_1 and G_2 on $\mathbb{R}$, strictly increasing on their supports which are $\mathbb{R}$ or $[a, \infty)$ with $a \in \mathbb{R}$, such that for every sequence $N = N(n)$ of natural numbers satisfying

$$(4.2) \qquad \sqrt{N}\mu(\theta_n) - G_1^{-1}(1 - \alpha_n) = O(1) \text{ and}$$
$$\lim_{n\to\infty} \inf \sqrt{N}\mu(\theta_n) > 0 \text{ in case } \bar{\alpha} > 0$$

we have

$$(4.3) \qquad \lim_{n\to\infty} P_{\theta_n}(T_N - \sqrt{N}\mu(\theta_n) \leq x) = G_2(x) \quad \text{for every } x \in \mathbb{R},$$

and, as $n \to \infty$,

$$(4.4) \qquad \sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T_N > t_n) = 1 - G_1(t_n + o(1))$$
$$\text{for all } t_n = \sqrt{N}\mu(\theta_n) + c$$

with $c \in \mathbb{R}$, not depending on n.

The extra condition replacing (2.2) is as follows.

(B3) For every sequence $N = N(n)$ of natural numbers satisfying

$$\lim_{n\to\infty} \left\{ \sqrt{N}\mu(\theta_n) - G_1^{-1}(1 - \alpha_n) \right\} = \infty \quad \text{we have} \quad \lim_{n\to\infty} \beta_N(\alpha_n, \theta_n) = 1.$$

This condition is used to show that if the power tends to 1 for some sequence $m(n)$, it still goes to 1 if we have even more observations. Such a property is needed, due to the definition of $N(\alpha, \beta, \theta)$.

A slight modification of the proof of Theorem 3.2 [cf. Inglot, Kallenberg and Ledwina (1998) for more details] yields the following result.

THEOREM 4.1. *Assume* (B1), (B2) *and* (B3). *For all* $\beta \in (\bar{\alpha}, 1)$, *with* $G_2^{-1}(1 - \beta) < G_1^{-1}(1 - \bar{\alpha})$ *in case* $\bar{\alpha} > 0$, *we have*

$$N(\alpha_n, \beta, \theta_n) = \left\{ \frac{G_1^{-1}(1 - \alpha_n) - G_2^{-1}(1 - \beta) + o(1)}{\mu(\theta_n)} \right\}^2 .$$

Suppose that (B2) holds for $\psi_{n;\alpha}$ with $G_1 = G_*$ (see Sections 5.4, 5.5 and 5.6). Often this can be established by a monotone transformation of the test statistic, thus keeping the same test. For instance, if

$$\sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T_N > t_n) = 1 - G_1(t_n + o(1))$$

we might consider $\tilde{T}_N = G_*^{-1}(G_1(T_N))$. Then we have

$$\sup_{\theta_0 \in \Theta_0} P_{\theta_0}(\tilde{T}_N > t_n) = \sup_{\theta_0 \in \Theta_0} P_{\theta_0}(T_N > G_1^{-1}(G_*(t_n)))$$

$$= 1 - G_1\left(G_1^{-1}(G_*(t_n)) + o(1)\right),$$

which often reduces to $1 - G_*(t_n + o(1))$. Assume further that the asymptotic shift represented by μ and μ_* is the same: $\mu = \mu_*$. If we have first-order efficiency, that is, $N_*(\alpha_n, \beta, \theta_n)/N(\alpha_n, \beta, \theta_n) = 1 + o(1)$, the equality $\mu = \mu_*$ will often hold true. Having no difference in shift, suppose that there is a possible difference in scale. So, suppose that $G_2(x) = G_*(x/\sigma)$ for some $\sigma \in \mathbb{R}$.

Since $N(\alpha_n, \beta, \theta_n) \geq N_*(\alpha_n, \beta, \theta_n)$ for all $\beta \in (\bar{\alpha}, 1)$, by definition of N_*, we get, in view of Theorem 3.2 and Theorem 4.1,

$$\left\{ \frac{G_*^{-1}(1 - \alpha_n) - \sigma G_*^{-1}(1 - \beta) + o(1)}{\mu_*(\theta_n)} \right\}^2$$

$$\geq \left\{ \frac{G_*^{-1}(1 - \alpha_n) - G_*^{-1}(1 - \beta) + o(1)}{\mu_*(\theta_n)} \right\}^2$$

for all $\beta \in (\bar{\alpha}, 1)$. If $G_*^{-1}(1 - \beta)$ takes positive as well as negative values (which for instance is the case if G_* is the standard normal distribution function), then we get $\sigma = 1$. Therefore, the same shift implies automatically the same scale. Concrete examples of this phenomenon are presented in Sections 5.4 and 5.5, when discussing the Cramér–von Mises test and the Anderson–Darling test. A similar argument to that above is used in proving first-order efficiency implies second-order efficiency in Kallenberg (1983a).

As a conclusion we may state that in many situations tests which are first-order efficient ($N_*(\alpha_n, \beta, \theta_n)/N(\alpha_n, \beta, \theta_n) = 1 + o(1)$) satisfy not only $\mu_* = \mu$

and $G_* = G_1$, but automatically also $G_1 = G_2$. We call such a situation a "regular case." It follows in that case that for each $\beta \in (\bar{\alpha}, 1)$,

$$N(\alpha_n, \beta, \theta_n) - N_*(\alpha_n, \beta, \theta_n) = o\left(G_*^{-1}(1 - \alpha_n)/\mu_*^2(\theta_n)\right)$$

(4.5)

$$= o\left(\sqrt{N_*(\alpha_n, \beta, \theta_n)}/\mu_*(\theta_n)\right).$$

This, together with Theorem 3.1, yields the corollary.

COROLLARY 4.2.   *Assume* (A1) *and* (A2) *for* $\psi_{n;\alpha}^*$ *and* (B1), (B2) *and* (B3) *for* $\psi_{n;\alpha}$ *with* $\mu_* = \mu$, $G_* = G_1 = G_2$. *Then*

$$N(\alpha_n, \beta, \theta_n) - N_*(\alpha_n, \beta, \theta_n) = o\left(\sqrt{N_*(\alpha_n, \beta, \theta_n)}/\mu_*(\theta_n)\right)$$

$$for\ each\ \beta \in (\bar{\alpha}, 1)$$

and

$$\lim_{n \to \infty} R_m(\alpha_n, \theta_n) = 0 \qquad for\ each\ sequence\ m = m(n),\ provided\ that$$

$$\liminf_{n \to \infty} \beta_m^*(\alpha_n, \theta_n) > \bar{\alpha}\ in\ case\ \bar{\alpha} > 0.$$

From Corollary 4.2 and the discussion above it is seen that the seemingly stronger property of vanishing shortcoming is, in regular cases, in fact equivalent to first-order efficiency.

The exception in Corollary 4.2 with respect to the shortcoming is not very serious. Let $\bar{\alpha} > 0$ and $\lim_{n \to \infty} \beta_m^*(\alpha_n, \theta_n) = \bar{\alpha}$. This means that with the $m(n)$ observations θ_n is too close to Θ_0 to detect. As a rule $\beta_m(\alpha_n, \theta_n)$ will also converge to $\bar{\alpha}$ in this case, at least if the test ψ is asymptotically unbiased. However, the conditions (B1), (B2) and (B3) give no information on this exceptional occasion. So, formally the restriction should be there.

For the nonlocal case, similarly to (A2'), condition (B2) may be replaced by a more suitable one, directed to (moderate and) large deviations. We do not present the details.

5. Applications and extensions.   It has been shown in the previous sections that shortcoming and first-order efficiency are strongly related optimality concepts. The equivalence holds in quite generality as is seen from the very general structure of the testing problem, the different types of efficiency concepts which are involved, from local to nonlocal, and the rather weak conditions imposed on the test statistics. Here we consider some concrete examples and applications. Also some extensions to comparison of tests which are not efficient are discussed.

5.1. *Student test.*   Consider the situation from Example 3.4.

In the Pitman situation with $\theta_n = cn^{-1/2}$ for some $c > 0$ and $\alpha_n = \bar{\alpha} \in (0, 1)$, we consider the one-sided Student's t-test. It is easily seen that (B1), (B2) and (B3) hold with $G_1 = G_2 = G_* = \Phi$ and $\mu(\theta) = \mu_*(\theta) = \theta$. By Corollary 4.2 we

get the well-known result that the t-test is Pitman-efficient [cf., e.g., Serfling (1980), page 320] and that its shortcoming tends to 0.

Next consider a sequence $\{\theta_n\}$ with $\theta_n \to 0$ and $n^{1/2}\theta_n \to \infty$ as $n \to \infty$. Let $\{\alpha_n\}$ be a sequence of levels satisfying $\lim_{n\to\infty} \alpha_n = 0$ and $\Phi^{-1}(1 - \alpha_n) = o(n^{1/2})$. Again conditions (B1), (B2) and (B3) hold (note that the power of the t-test increases with n). Hence, Corollary 4.2 gives that the shortcoming of the t-test tends to 0 for all these alternatives and, equivalently, that its efficiency equals $N_*(\alpha_n, \beta, \theta_n)/N(\alpha_n, \beta, \theta_n) = 1 + o\left(\{\theta_n N_*(\alpha_n, \beta, \theta_n)\}^{-1/2}\right) = 1 + o\left(\{\Phi^{-1}(1 - \alpha_n)\}^{-1/2}\right)$.

Finally, in the Bahadur case with $\theta_n = \theta > 0$, (B1) and (4.3) continue to hold with $G_2 = G_* = \Phi$ and $\mu(\theta) = \mu_*(\theta) = \theta$, but

$$N^{-1} \log P_0\left(T_N > \theta\sqrt{N}\right) \to -\tfrac{1}{2}\log(1 + \theta^2)$$

and hence (4.4) does not hold with $G_1 = G_* = \Phi$. This corresponds to the fact that the t-test is not Bahadur-efficient.

5.2. *Likelihood ratio test.*

Consider the situation from Example 3.5 and the likelihood ratio statistic L_n for this testing problem. Write $T_n = \sqrt{2\log L_n}$. Straightforward calculation yields

$$T_n = \begin{cases} 0, & \text{if } \overline{X} < 0, \overline{Y} < 0, \\ \sqrt{n}\,\overline{X}, & \text{if } \overline{X} \geq 0, \overline{Y} \leq 0, \\ \sqrt{n}\,\overline{Y}, & \text{if } \overline{X} \leq 0, \overline{Y} \geq 0, \\ \sqrt{n}\sqrt{\overline{X}^2 + \overline{Y}^2}, & \text{if } \overline{X} > 0, \overline{Y} > 0, \end{cases}$$

thus giving (B1). Set $\mu(\theta) = \mu_*(\theta)$, where $\mu_*(\theta)$ is as in Example 3.5 and define $\Theta_1 = \{(\theta^{(1)}, \theta^{(2)}): \theta^{(1)} > 0 \text{ or } \theta^{(2)} > 0\}$. Standard but rather lengthy considerations yield the following result.

If $\theta_n \in \Theta_1$ and $M = M(n)$ is an arbitrary sequence satisfying $M \to \infty$ and $\sqrt{M}\mu(\theta_n) \to \infty$, then

$$\lim_{n\to\infty} P_{\theta_n}(T_M - \sqrt{M}\mu(\theta_n) \leq x) = \Phi(x), \quad x \in \mathbb{R}.$$

Moreover, immediate calculation gives

$$\sup_{\theta^{(1)}\leq 0,\, \theta^{(2)}\leq 0} P_\theta(T_N > t_n) = P_{(0,0)}(T_N > t_n) = \exp\{-\tfrac{1}{2}t_n^2 + O(1)\}$$

for any t_n such that $t_n \to \infty$. The two results imply that T_n satisfies (B2) and (B3) for any θ_n and α_n, $\alpha_n \to 0$, with $G_1 = G_2 = G_* = \Phi$ and $\mu = \mu_*$. Hence, Corollary 4.2 applies, yielding optimality of the likelihood ratio test in the sense of vanishing shortcoming and first-order efficiency according to Corollary 4.2 in the intermediate and Bahadur case.

Note that studying shortcoming of likelihood ratio tests in general exponential families has been the subject of Chapter II and Chapter III of Kallenberg

(1978). Reformulating, extending and stating these results in the present set-up is beyond the scope of this paper.

5.3. *Curved exponential families.* Consider the situation from Example 3.6 and assume moreover that

$$(5.1) \qquad \lim_{n \to \infty} (\theta_n - \theta_0)^2 |\log \alpha_n|^{1/2} = 0.$$

The test statistic of the locally most powerful test is of the form $T_n = \dot{\gamma}'_{\theta_0}(\overline{X}_n - E_{\theta_0}X)\sqrt{n}/s$, where $\cdot$ denotes the derivative w.r.t. θ, while $s = \{\mathrm{var}_{\theta_0}\, \dot{\gamma}'_{\theta_0}X\}^{1/2}$. It is seen from Lemma 3.7 in Kallenberg (1981a) with the number of observations n replaced by m that $R_m(\alpha_n, \theta_n) \to 0$ in all cases, provided that

$$(5.2) \qquad m^{-1}|\log \alpha_n|^{3/2} \to 0 \quad \text{if } s(\theta_n - \theta_0)m^{1/2}\{-2\log \alpha_n\}^{-1/2} \to 1.$$

Writing

$$m^{-1}|\log \alpha_n|^{3/2} = (\theta_n - \theta_0)^{-2}m^{-1}|\log \alpha_n|(\theta_n - \theta_0)^2|\log \alpha_n|^{1/2},$$

it immediately follows that (5.2) is implied by (5.1). Hence, for all sequences of levels $\{\alpha_n\}$ and all sequences of alternatives $\{\theta_n\}$ satisfying (3.5) and (5.1), $R_m(\alpha_n, \theta_n) \to 0$ for each sequence $m = m(n)$.

Therefore, by Theorem 3.1, the locally most powerful test is first-order efficient in the Pitman case and in the intermediate case, provided in the latter situation that (3.5) and (5.1) are satisfied. [Of course, the locally most powerful test is in general not first-order efficient in the sense of Bahadur; see Kallenberg (1981a), page 673.]

5.4. *Cramér–von Mises test.* Consider the situation from Example 3.7. The Cramér–von Mises test statistic for this testing problem is defined by rejecting H_0 for large values of

$$\left\{n \int_0^1 (F_n(x) - x)^2 \, dx\right\}^{1/2},$$

where F_n is the empirical distribution function. Consider alternatives (3.9), that is, $p_n(x) = 1 + \theta_n a(x)$ with $\theta_n \to 0$ and such that $\sqrt{n}\, \theta_n \to \infty$ and the other restrictions of Theorem 5.1(d), (e) are fulfilled. Theorem 5.1(e) shows that there is only one function $a(x)$, $a(x) = C_1(x) = \sqrt{2}\cos(\pi x)$, under which the Cramér–von Mises test is efficient. This result supports analogous results stated in another framework by Neuhaus (1976) and Nikitin (1995). Moreover, the asymptotic relative efficiency under other alternatives than the one given by $C_1(x)$ is explicitly calculated in Theorem 5.1(d). This provides an alternative to the findings of Neuhaus (1976) and related developments of Strasser (1990). See also Milbrodt and Strasser (1990), Drees and Milbrodt (1991, 1994) and Janssen (1995) for further results in this area. The phenomenon that equality of asymptotic optimal shift implies also equality of scale terms, as

discussed in general terms in Section 4, is also illustrated clearly in this case. Similar results for the Anderson–Darling statistic are treated in Section 5.5.

We use the following notation

$$(5.3) \quad A(t) = \int_0^t a(x)\, dx, \quad \|A\| = \left\{ \int_0^1 A^2(t)\, dt \right\}^{1/2}, \qquad \mu(\theta) = \pi \|A\| \theta$$

and

$$T_n = \pi \left\{ n \int_0^1 (F_n(x) - x)^2\, dx \right\}^{1/2}.$$

Write

$$ARE(T_n, T_n^*) = \lim_{n \to \infty} \frac{N_*(\alpha_n, \beta, \theta_n)}{N(\alpha_n, \beta, \theta_n)}$$

for the asymptotic relative efficiency of T_n with respect to T_n^*. The following theorem is proved in Inglot, Kallenberg and Ledwina (1998), Section 5.3.

THEOREM 5.1.

(a) *Let $\theta_n \to 0$ and let $M = M(n)$ be an arbitrary sequence satisfying $M \to \infty$ and $\sqrt{M}\theta_n \to \infty$. Then*

$$\lim_{n \to \infty} P_{\theta_n} \left(T_M - \sqrt{M} \mu(\theta_n) \leq x \right) = \Phi\left(\frac{\|A\|}{\sigma \pi} x \right),$$

where

$$\sigma^2 = \int_0^1 \int_0^1 (s \wedge t - st) A(s) A(t)\, ds\, dt.$$

(b) *Let $C_k(x) = \sqrt{2}\cos(\pi k x)$ and let $\mathbf{c}_k = \int_0^1 a(x) C_k(x)\, dx$. We have*

$$(5.4) \qquad \|A\|^2 = \sum_{k=1}^\infty \frac{\mathbf{c}_k^2}{\pi^2 k^2} \quad and \quad \sigma^2 = \sum_{k=1}^\infty \frac{\mathbf{c}_k^2}{\pi^4 k^4},$$

provided that a is bounded, $\int_0^1 a(x)\, dx = 0$ and $\int_0^1 a^2(x)\, dx = 1$.

(c) *If $x_n \to 0$, $n x_n^2 \to \infty$, then for any $\rho \in (2, 3)$,*

$$\log P_{\theta_0} \left(T_n \geq x_n \sqrt{n} \right) = -\tfrac{1}{2} n x_n^2 + O(n x_n^\rho).$$

(d) *Let $p_n(x) = 1 + \theta_n a(x)$ be as in (3.9). Suppose that $\alpha_n \to 0$ and $\theta_n^\gamma \Phi^{-1}(1 - \alpha_n) \to 0$ for some $\gamma \in (0, 1)$. Then conditions (B1), (B2) and (B3) hold with $G_1(x) = \Phi(x)$ and $G_2(x) = \Phi(\|A\| x / (\sigma \pi))$. Hence, for each $\beta \in (0, 1)$,*

$$(5.5) \qquad N(\alpha_n, \beta, \theta_n) = \left\{ \Phi^{-1}(1 - \alpha_n) - \frac{\pi \sigma}{\|A\|} \Phi^{-1}(1 - \beta) + o(1) \right\}^2 \bigg/ \left\{ \pi \|A\| \theta_n \right\}^2.$$

Moreover, for each $\beta \in (0, 1)$,

$$(5.6) \qquad N_*(\alpha_n, \beta, \theta_n) = \left\{ \Phi^{-1}(1 - \alpha_n) - \Phi^{-1}(1 - \beta) + o(1) \right\}^2 / \theta_n^2,$$

implying

$$(5.7) \qquad\qquad ARE(T_n, T_n^*) = (\pi \|A\|)^2.$$

(e) *Let $p_n(x) = 1 + \theta_n C_1(x)$, $C_1(x) = \sqrt{2} \cos(\pi x)$. Suppose that $\alpha_n \to 0$ and $\theta_n^\gamma \Phi^{-1}(1 - \alpha_n) \to 0$ for some $\gamma \in (0, 1)$. Then $ARE(T_n, T_n^*) = 1$ and, moreover, $N_*(\alpha_n, \beta, \theta_n) - N(\alpha_n, \beta, \theta_n) = o\left(\sqrt{N_*(\alpha_n, \beta, \theta_n)}/\theta_n \right)$. Hence, for these alternatives $\lim_{n\to\infty} R_m(\alpha_n, \theta_n) = 0$ for each sequence $m = m(n)$. For any other alternative $p_n(x) = 1 + \theta_n a(x)$, satisfying our assumptions, $ARE(T_n, T_n^*) < 1$. In particular, for $p_n(x) = 1 + \theta_n C_k(x)$, $C_k(x) = \sqrt{2} \cos(k\pi x)$, $ARE(T_n, T_n^*) = k^{-2}$.*

Theorem 5.1(d) clearly illustrates the phenomenon discussed in Section 4. To get first-order efficiency, the asymptotic shift $\mu(\theta_n) = \pi \|A\| \theta_n$ of T_n should be the same as that of T_n^*, which equals $\mu_*(\theta_n) = \theta_n$. Hence, $\pi \|A\|$ should be equal to 1. However, (5.5) and (5.6) show that also the terms corresponding to the scale of the limiting distributions G_2 and G_* are different: $\pi\sigma/\|A\|$ and 1, respectively. But, in case $\pi\|A\| = 1$, these scale terms should be the same as well due to the phenomenon discussed in Section 4. And, indeed this is the case: if $\pi\|A\| = 1$, then $\sigma = \pi^{-2}$ and hence $\pi\sigma/\|A\| = 1$. This is shown analytically in Lemma 5.8 of Inglot, Kallenberg and Ledwina (1998), but could also be derived from (5.5) and (5.6) directly. Let $\pi\|A\| = 1$. Since $N_*(\alpha_n, \beta, \theta_n) \le N(\alpha_n, \beta, \theta_n)$ for all $\beta \in (0, 1)$ and $\Phi^{-1}(1 - \beta)$ can be positive as well as negative, it follows that the coefficients of $\Phi^{-1}(1 - \beta)$ in (5.5) and (5.6) should be the same. Together with $\pi\|A\| = 1$ this gives $\sigma = \pi^{-2}$.

The next section contains similar results for the Anderson–Darling test.

5.5. *Anderson–Darling test.* Consider the situation from Example 3.7. The Anderson–Darling test rejects for large values of

$$T_n = \left\{ 2n \int_0^1 \frac{(F_n(t) - t)^2}{t(1 - t)} \, dt \right\}^{1/2}.$$

Set [cf. (5.3)]

$$\mu(\theta_n) = \sqrt{2} \theta_n \|A_w\|, \qquad A_w(t) = A(t)/\sqrt{t(1 - t)}.$$

THEOREM 5.2.

(a) *Suppose $\theta_n \to 0$ and $M = M(n)$ is an arbitrary sequence satisfying $M \to \infty$ and $\sqrt{M}\theta_n \to \infty$. Then*

$$\lim_{n\to\infty} P_{\theta_n}(T_M - \sqrt{M}\mu(\theta_n) \le x) = \Phi(x\|A_w\|/(\sqrt{2}\sigma_w)),$$

where

$$\sigma_w^2 = \int_0^1 \int_0^1 \frac{(s \wedge t - st)A(s)A(t)}{s(1-s)t(1-t)} \, ds \, dt.$$

(b) *Let* $\{L_n(x)\}_{n \geq 0}$ *be the system of orthonormal Legendre polynomials on* $[0, 1]$ *and let* $l_n = \int_0^1 a(x)L_n(x) \, dx$. *We have*

(5.8) $$\|A_w\|^2 = \sum_{k=1}^{\infty} l_k^2 / \{k(k+1)\} \quad and \quad \sigma_w^2 = \sum_{k=1}^{\infty} l_k^2 / \{k^2(k+1)^2\},$$

provided that a *is bounded,* $\int_0^1 a(x) \, dx = 0$ *and* $\int_0^1 a^2(x) \, dx = 1$.

(c) *If* $x_n \to 0$, $nx_n^2 \to \infty$, *then for any* $\rho \in (2, 3)$,

$$\log P_{\theta_0}\left(T_n \geq x_n\sqrt{n}\right) = -\tfrac{1}{2}nx_n^2 + O(nx_n^\rho).$$

(d) *Let* $p_n(x) = 1 + \theta_n a(x)$ *be as in* (3.9). *Suppose that* $\alpha_n \to 0$ *and* $\theta_n^\gamma \Phi^{-1}(1 - \alpha_n) \to 0$ *for some* $\gamma \in (0, 1)$. *Then conditions* (B1), (B2) *and* (B3) *hold with* $G_1(x) = \Phi(x)$ *and* $G_2(x) = \Phi(x\|A_w\|/(\sqrt{2}\sigma_w))$. *Hence, for each* $\beta \in (0, 1)$,

$$N(\alpha_n, \beta, \theta_n) =$$
$$\left\{\Phi^{-1}(1 - \alpha_n) - \tfrac{\sqrt{2}\sigma_w}{\|A_w\|}\Phi^{-1}(1 - \beta) + o(1)\right\}^2 \Big/ \{\sqrt{2}\|A_w\|\theta_n\}^2$$

and

(5.9) $$ARE(T_n, T_n^*) = 2\|A_w\|^2.$$

(e) *Suppose that* $\alpha_n \to 0$ *and* $\theta_n^\gamma \Phi^{-1}(1 - \alpha_n) \to 0$ *for some* $\gamma \in (0, 1)$. *Let* $p_n(x) = 1 + \theta_n L_1(x)$, $L_1(x) = \sqrt{3}(2x - 1)$, *then* $ARE(T_n, T_n^*) = 1$ *and, moreover,* $N_*(\alpha_n, \beta, \theta_n) - N(\alpha_n, \beta, \theta_n) = o\left(\sqrt{N_*(\alpha_n, \beta, \theta_n)}/\theta_n\right)$. *Hence, for these alternatives* $\lim_{n \to \infty} R_m(\alpha_n, \theta_n) = 0$ *for each sequence* $m = m(n)$. *For any other alternative* $p_n(x) = 1 + \theta_n a(x)$, *satisfying our assumptions,* $ARE(T_n, T_n^*) < 1$. *In particular, if* $p_n(x) = 1 + \theta_n L_k(x)$, L_k *the kth Legendre polynomial,* $ARE(T_n, T_n^*) = 2/\{k(k+1)\}$.

For a proof of Theorem 5.2 we refer to Inglot, Kallenberg and Ledwina (1998), Section 5.4. The same remark on the fact that optimal asymptotic shift implies also equality of scale terms as at the end of Section 5.4 applies here. If $\sqrt{2}\|A_w\| = 1$, σ_w should be 1 by this property, and indeed, this is the case.

5.6. *Concluding remarks and extensions.* In Sections 5.1–5.5 the equivalence of vanishing shortcoming and first-order efficiency is illustrated with a lot of examples. Numerous other examples and applications could be added, both in the Pitman case, the intermediate one and for fixed alternatives. The equivalence is partly due to the phenomenon that equality of asymptotic optimal shift implies also equality of scale terms. This is clearly illustrated in the applications in Sections 5.4 and 5.5, where for most directions there is

no first-order efficiency. In those cases the asymptotic shift *and* scale of T_n and T_n^* differ. As soon as the asymptotic shifts (first-order efficiency) coincide, automatically also the asymptotic scales become the same, which in turn is equivalent to vanishing shortcoming.

The results of Sections 5.4 and 5.5 can be extended to other statistics, which are bilinear functionals of the empirical process. For appropriate limit theorems, see Inglot, Kallenberg and Ledwina (1993) and Inglot and Ledwina (1993).

Results like (5.4), (5.8) and (5.7), (5.9) are of independent interest. They provide a simple and intuitive way of comparing quadratic tests with the best possible one or two quadratic tests with each other. Yielding explicit expressions for the efficiency, as illustrated in (5.7) and (5.9), the intermediate approach is more widely applicable than classical approaches, which give for more complex problems only efficiencies for which no closed expression is available. Moreover, the limiting behavior described by the intermediate approach is nicely reflected by finite sample results. In this way our findings reproduce Nikitin's (1995) results, where Bahadur's approach has been exploited. On the other hand, they can be nicely confronted with the two-step approach proposed by Hájek and Sidák (1967) and applied in Neuhaus (1976) as well as Wieand's approach exploited by Gregory (1980). Moreover, the results coincide with those obtained in Section 7.7 of Inglot and Ledwina (1996), where a slightly different definition of asymptotic intermediate relative efficiency [cf. also Kallenberg (1983b)] has been applied.

Although the main theme of the paper concerns the relation between vanishing shortcoming and first-order efficiency, Theorem 4.1 can also be applied to compare two tests with each other. This gives an easy way to calculate the asymptotic relative efficiency of T_n w.r.t. $\tilde{T}_n$, where T_n and $\tilde{T}_n$ are two test statistics.

Similarly, Theorem 3.3 can be generalized to statistics other than the MP.

THEOREM 5.3.   *Assume* (B1), (B2) *and* (B3). *Moreover, assume that for every sequence* $N = N(n)$ *of natural numbers satisfying* $\lim_{n \to \infty} \{\sqrt{N} \mu(\theta_n) - G_1^{-1}(1 - \alpha_n)\} = -\infty$ *we have* $\lim_{n \to \infty} \beta_N(\alpha_n, \theta_n) = 0$. *Then, for each* $\beta \in (\bar{\alpha}, 1)$ *and for each sequence* $m = m(n)$ *we have*

$$\lim_{n \to \infty} \left[\beta_m(\alpha_n, \theta_n) - \left\{ 1 - G_2 \left(G_2^{-1}(1 - \beta) - \left[\sqrt{m} - \sqrt{N(\alpha_n, \beta, \theta_n)} \right] \mu(\theta_n) \right) \right\} \right] = 0.$$

The proof of Theorem 5.3 is obtained from the proof of Theorem 3.3 by obvious modifications and is therefore omitted.

As Theorem 4.1 can be used to get results on the asymptotic relative efficiency of T_n w.r.t. $\tilde{T}_n$, Theorem 5.3 in combination with Theorem 4.1 can be applied to obtain results on the "shortcoming of T_n w.r.t. $\tilde{T}_n$": $\beta_m(\alpha_n, \theta_n) - \tilde{\beta}_m(\alpha_n, \theta_n)$.

APPENDIX

This appendix contains the proofs of Theorems 3.1, 3.2 and 3.3. The proofs of Theorems 3.1, 3.2′ and 3.3′ are similar to the proofs of Theorems 3.1, 3.2 and

3.3, respectively. They are given in Inglot, Kallenberg and Ledwina (1998).

We start with the proofs of Theorems 3.2 and 3.3 and end with the proof of Theorem 3.1.

PROOF OF THEOREM 3.2. For a (sufficiently small) positive ε set

$$(A.1) \qquad \sqrt{N} = \left\{ G_*^{-1}(1-\alpha_n) - G_*^{-1}(1-\beta) - \varepsilon \right\} / \mu_*(\theta_n).$$

Consider the test which rejects H_0 if

$$T_N^* > \sqrt{N}\mu_*(\theta_n) + G_*^{-1}(1-\beta) + \varepsilon/2.$$

Denote the level of this test by $\alpha_n^{(1)}$. Then, by (2.7) and (A.1), we get for sufficiently large n

$$\alpha_n^{(1)} = \sup_{\theta_0 \in \Theta_0} P_{\theta_0} \left(T_N^* > \sqrt{N}\mu_*(\theta_n) + G_*^{-1}(1-\beta) + \varepsilon/2 \right)$$

$$= 1 - G_* \left(G_*^{-1}(1-\alpha_n) - \varepsilon/2 + o(1) \right) > \alpha_n$$

and, by (2.6),

$$\lim_{n \to \infty} \beta_N^*(\alpha_n^{(1)}, \theta_n) = 1 - \lim_{n \to \infty} P_{\theta_n} \left(T_N^* - \sqrt{N}\mu_*(\theta_n) \le G_*^{-1}(1-\beta) + \varepsilon/2 \right)$$

$$= 1 - G_*(G_*^{-1}(1-\beta) + \varepsilon/2) < \beta.$$

Hence, for sufficiently large n,

$$(A.2) \qquad N_*(\alpha_n, \beta, \theta_n) \ge N_*(\alpha_n^{(1)}, \beta, \theta_n) > N.$$

By sending $\varepsilon \to 0$ in (A.2) [cf. also (A.1)], we arrive at

$$(A.3) \qquad \liminf_{n \to \infty} \sqrt{N^*(\alpha_n, \beta, \theta_n)}\mu_*(\theta_n) - G_*^{-1}(1-\alpha_n) \ge -G_*^{-1}(1-\beta).$$

Similarly, define for $\varepsilon > 0$,

$$(A.4) \qquad \sqrt{N} = \left\{ G_*^{-1}(1-\alpha_n) - G_*^{-1}(1-\beta) + \varepsilon \right\} / \mu_*(\theta_n)$$

and consider the test which rejects H_0 if

$$T_N^* > \sqrt{N}\mu_*(\theta_n) + G_*^{-1}(1-\beta) - \varepsilon/2.$$

By (2.7) we get for sufficiently large n,

$$\alpha_n^{(2)} = \sup_{\theta_0 \in \Theta_0} P_{\theta_0} \left(T_N^* > \sqrt{N}\mu_*(\theta_n) + G_*^{-1}(1-\beta) - \varepsilon/2 \right) < \alpha_n$$

and, by (2.6),

$$(A.5) \qquad \lim_{n \to \infty} \beta_N^*(\alpha_n^{(2)}, \theta_n) = 1 - G_* \left(G_*^{-1}(1-\beta) - \varepsilon/2 \right) > \beta.$$

In view of (2.2) and (A.5) we therefore have

$$\beta_m^*(\alpha_n^{(2)}, \theta_n) > \beta \qquad \text{for all } m \ge N$$

and hence

(A.6) $$N_*(\alpha_n, \beta, \theta_n) \le N_*(\alpha_n^{(2)}, \beta, \theta_n) \le N$$

for sufficiently large n. Sending $\varepsilon \to 0$ in (A.6) [cf. also (A.4)], we end up with

(A.7) $$\limsup_{n\to\infty} \sqrt{N_*(\alpha_n, \beta, \theta_n)}\mu_*(\theta_n) - G_*^{-1}(1 - \alpha_n) \le -G_*^{-1}(1 - \beta).$$

Combining (A.3) and (A.7) completes the proof. □

PROOF OF THEOREM 3.3. Consider first a sequence $m = m(n)$ of the form

(A.8) $$\sqrt{m} = \sqrt{N_*(\alpha_n, \beta, \theta_n)} + \{b_n/\mu_*(\theta_n)\},$$

where $b_n \to b$, $b \in \mathbb{R}$ with $b > G_*^{-1}(1-\beta) - G_*^{-1}(1-\bar{\alpha})$ if $\bar{\alpha} > 0$. In view of (3.2) and (A.8), m is of the form (2.5). By (2.4) and (2.7) the critical value c_m in (2.3) of the level-α_n MP test of H_0 against θ_n based on m observations satisfies

$$c_m = G_*^{-1}(1 - \alpha_n) + o(1).$$

[Note that $G_*^{-1}(1 - \alpha_n) + o(1)$ is of the form $\sqrt{m}\mu_*(\theta_n) + O(1)$ as required in (2.7).]

Since

$$P_{\theta_n}(T_m^* > c_m) \le \beta_m^*(\alpha_n, \theta_n) \le P_{\theta_n}(T_m^* \ge c_m)$$

and G_* is continuous, it follows by (2.6) and (3.2) that

$$\lim_{n\to\infty} \beta_m^*(\alpha_n, \theta_n) = 1 - G_*(G_*^{-1}(1 - \beta) - b)$$

(A.9) $$= 1 - \lim_{n\to\infty} G_*\left(G_*^{-1}(1 - \beta) - \left[\sqrt{m} - \sqrt{N_*(\alpha_n, \beta, \theta_n)}\right]\mu_*(\theta_n)\right).$$

So, (3.3) holds for sequences of the form (A.8).

If $\sqrt{m} = \sqrt{N_*(\alpha_n, \beta, \theta_n)} + \{b_n/\mu_*(\theta_n)\}$ with $b_n \to \infty$, then for any b we have $b_n > b$ for sufficiently large n and hence $\sqrt{m} > \sqrt{N_*(\alpha_n, \beta, \theta_n)} + \{b/\mu_*(\theta_n)\}$. Consequently, by (2.2) and (A.9),

$$\beta_m^*(\alpha_n, \theta_n) \ge 1 - G_*(G_*^{-1}(1 - \beta) - b)$$

for every $b \in \mathbb{R}$ and n sufficiently large. Thus $\lim_{n\to\infty} \beta_m^*(\alpha_n, \theta_n) = 1$. On the other hand,

$$\lim_{n\to\infty} G_*\left(G_*^{-1}(1 - \beta) - \left[\sqrt{m} - \sqrt{N_*(\alpha_n, \beta, \theta_n)}\right]\mu_*(\theta_n)\right)$$

$$= \lim_{n\to\infty} G_*\left(G_*^{-1}(1 - \beta) - b_n\right) = 0.$$

So, (3.3) holds also in this case. We can proceed similarly in the case $b_n \to -\infty$, getting (3.3) again. The general case follows now by a subsequence argument. □

PROOF OF THEOREM 3.1. (i)$\Rightarrow$(ii). Let $\bar{\alpha} < \beta < 1$ and $\varepsilon > 0$. For each m satisfying

$$\sqrt{m} \geq \sqrt{N_*(\alpha_n, \beta, \theta_n)} + \varepsilon/\mu_*(\theta_n),$$

we have, in view of (3.3),

$$\liminf_{n \to \infty} \beta_m(\alpha_n, \theta_n) = \liminf_{n \to \infty} \{\beta_m^*(\alpha_n, \theta_n) - R_m(\alpha_n, \theta_n)\} > \beta.$$

Hence

$$N(\alpha_n, \beta, \theta_n) \leq \left[\sqrt{N_*(\alpha_n, \beta, \theta_n)} + \varepsilon/\mu_*(\theta_n)\right]^2 + 1$$

for sufficiently large n. Since $\varepsilon > 0$ is arbitrarily chosen, (ii) now follows.

(ii)$\Rightarrow$(i). Let $m = m(n)$ be some sequence. By a subsequence argument we may for proving $R_m(\alpha_n, \theta_n) \to 0$ assume w.l.o.g. that $\beta_m(\alpha_n, \theta_n) \to \beta \in [0, 1)$. [The case $\beta = 1$ can be excluded, since $\beta_m(\alpha_n, \theta_n) \to 1$ automatically implies $R_m(\alpha_n, \theta_n) \to 0$.] If $\beta_m^*(\alpha_n, \theta_n) \to \beta$, we get $R_m(\alpha_n, \theta_n) \to 0$ and hence it suffices to prove that the statement $\beta_m^*(\alpha_n, \theta_n) \to \beta + \varepsilon$ for some $\varepsilon > 0$ and $\beta + \varepsilon > \bar{\alpha}$, leads to a contradiction.

So, assume $\beta_m(\alpha_n, \theta_n) \to \beta$ and $\beta_m^*(\alpha_n, \theta_n) \to \beta + \varepsilon$ with $0 \leq \beta < 1$ and $\bar{\alpha} < \beta + \varepsilon \leq 1$. For sufficiently large n we have for all β^* and $\tilde{\beta}$ such that $\beta < \beta^* < \tilde{\beta} < \beta + \varepsilon$ and $\beta^* > \bar{\alpha}$,

$$m \leq N(\alpha_n, \beta^*, \theta_n) \quad \text{and} \quad m \geq N_*(\alpha_n, \tilde{\beta}, \theta_n).$$

Hence, we get

$$N_*(\alpha_n, \tilde{\beta}, \theta_n) \leq N(\alpha_n, \beta^*, \theta_n)$$

and therefore, by (ii),

$$N_*(\alpha_n, \tilde{\beta}, \theta_n) - N_*(\alpha_n, \beta^*, \theta_n) \leq N(\alpha_n, \beta^*, \theta_n) - N_*(\alpha_n, \beta^*, \theta_n)$$

$$= o\left(\sqrt{N_*(\alpha_n, \beta^*, \theta_n)}/\mu_*(\theta_n)\right).$$

Application of (3.3) with $m = N_*(\alpha_n, \tilde{\beta}, \theta_n)$ and β replaced by β^* gives a contradiction. $\square$

REFERENCES

BICKEL, P. J., CHIBISOV, D. M. and VAN ZWET, W. R. (1981). On efficiency of first and second-order. *Internat. Statist. Rev.* **49** 169–175.

BOOK, S. A., (1976). The Cramér–Feller–Petrov large deviation theorem for triangular arrays. Technical report, California State College.

DREES, H. and MILBRODT, H. (1991). Components of the two-sided Kolmogorov–Smirnov test in signal detection problems with Gaussian white noise. *J. Statist. Plann. Inference* **29** 325–335.

DREES, H. and MILBRODT, H. (1994). The one-sided Kolmogorov–Smirnov test in signal detection problems with Gaussian white noise. *Statist. Neerlandica* **48** 103–116.

EFRON, B. (1975). Defining the curvature of a statistical problem. *Ann. Statist.* **3** 1189–1242.

GREGORY, G. G. (1980). On efficiency and optimality of quadratic tests. *Ann. Statist.* **8** 116–131.

HÁJEK, J. and SIDÁK, Z. (1967). *Theory of Rank Tests*. Academic Press, New York.

INGLOT, T., KALLENBERG, W. C. M. and LEDWINA, T. (1993). Asymptotic behavior of some bilinear functionals of the empirical process. *Math. Methods Statist.* **2** 316–336.

INGLOT, T., KALLENBERG, W. C. M. and LEDWINA, T. (1998). Vanishing shortcoming and asymptotic relative efficiency. Memorandum 1467, Faculty of Mathematical Sciences, Univ. Twente.

INGLOT, T. and LEDWINA, T. (1993). Moderately large deviations and expansions of large deviations for some functionals of weighted empirical processes. *Ann. Probab.* **21** 1691–1705.

INGLOT, T. and LEDWINA, T. (1996). Asymptotic optimality of data-driven Neyman's tests for uniformity. *Ann. Statist.* **24** 1982–2019.

JANSSEN, A. (1995). Principal component decomposition of nonparametric tests. *Probab. Theory Related Fields* **101** 193–209.

JUREČKOVÁ, J., KALLENBERG, W. C. M. and VERAVERBEKE, N. (1988). Moderate and Cramér-type large deviation theorems for M-estimators. *Statist. Probab. Lett.* **6** 191–199.

KALLENBERG, W. C. M. (1978). Asymptotic Optimality of Likelihood Ratio Tests in Exponential Families. *Math. Centre Tracts* **77**. CWI, Amsterdam.

KALLENBERG, W. C. M. (1981a). The shortcoming of locally most powerful tests in curved exponential families. *Ann. Statist.* **9** 673–677.

KALLENBERG, W. C. M. (1981b). Bahadur deficiency of likelihood ratio tests in exponential families. *J. Multivariate Anal.* **11** 506–531.

KALLENBERG, W. C. M. (1983a). Asymptotic efficiency and deficiency of tests. In *Proceedings of the Forty-Fourth Session of the ISI, Madrid* **2** 1173–1189.

KALLENBERG, W. C. M. (1983b). Intermediate efficiency, theory and examples. *Ann. Statist.* **11** 170–182.

LEHMANN, E. L. (1986). *Testing Statistical Hypotheses*, 2nd ed. Wiley, New York.

MILBRODT, H. and STRASSER, H. (1990). On the asymptotic power of the two-sided Kolmogorov–Smirnov test. *J. Statist. Plann. Inference* **26** 1–23.

NEUHAUS, G. (1976). Asymptotic power properties of the Cramér–von Mises test under contiguous alternatives. *J. Multivariate Anal.* **6** 95–110.

NIKITIN, YA. YU. (1995). *Asymptotic Efficiency of Nonparametric Tests*. Cambridge Univ. Press.

SERFLING, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

STRASSER, H. (1990). Global extrapolation of local efficiency. *Statist. Decisions* **8** 11–26.

T. INGLOT
INSTITUTE OF MATHEMATICS
WROCŁAW UNIVERSITY OF TECHNOLOGY
WYB. WYSPIAŃSKIEGO 27
50-370 WROCŁAW
POLAND
E-MAIL: inglot@neyman.im.pwr.wroc.pl

W. C. M. KALLENBERG
UNIVERSITY OF TWENTE
FACULTY OF MATHEMATICAL SCIENCES
P.O. BOX 217
7500 AE ENSCHEDE
THE NETHERLANDS
E-MAIL: w.c.m.kallenberg@math.utwente.nl

T. LEDWINA
INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
UL. KOPERNIKA 18
51-617 WROCŁAW
POLAND
E-MAIL: ledwina@neyman.im.pwr.wroc.pl