Compound Poisson approximation in total variation.

A. D. Barbour*

Abteilung für Angewandte Mathematik, Universität Zürich-Irchel, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland.

Sergey Utev*

School of Statistics, La Trobe University, Bundoora, MELBOURNE, Vic 3083, Australia and Institute of Mathematics, Novosibirsk University.

Anonymous ftp version of 14 February 1997

ABSTRACT

Poisson approximation in total variation can be successfully established in a wide variety of contexts, involving sums of weakly dependent random variables which usually take the value 0, and occasionally the value 1. If the random variables can take other positive integer values, or if there is stronger dependence between them, compound Poisson approximation may be more suitable; such considerations are particularly relevant for applications in insurance mathematics. Stein's method, which is so effective in the Poisson context, turns out to be much more difficult to apply for compound Poisson approximation, because the solutions of the Stein Equation have undesirable properties. In this paper, we prove new bounds on the absolute values of the solutions to the Stein Equation and of their first differences, over certain ranges of their arguments. These enable compound Poisson approximation in total variation to be carried out with almost the same efficiency as in the Poisson case. Even for sums of independent random variables, which have been exhaustively studied in the past, new results are obtained, effectively solving a problem discussed by Le Cam (1965), in the context of nonnegative integer valued random variables.

* Supported in part by Schweizerischer Nationalfonds Projekte Nr. 21-37354.93 and ****. ADB acknowledges the warm hospitality of the Statistics Section at Monash University, where part of this work was undertaken.
1a. Introduction.

A simple model of an insurance portfolio assumes a finite number $n$ of insured risks, each of which may lead to a claim with a small probability, independently of all the others. The distribution of the total number $W$ of claims is then well approximated by the Poisson distribution $\text{Po}(\lambda)$, with $\lambda = \sum_{j=1}^{n} p_j$, even if the claim probabilities $p_j$ are not equal. Furthermore, as observed by Michel (1988), if all claim amounts are independent and identically distributed, the difference in terms of total variation distance between the distribution of the aggregated claim amount and an appropriate compound Poisson distribution is no greater than the total variation distance $\Delta \leq \lambda^{-1} \sum_{j=1}^{n} p_j^2$ between the distribution of the total number of claims $W$ and $\text{Po}(\lambda)$. If the occurrences of claims are weakly dependent, but the claim amounts are still independent and identically distributed, Goovaerts and Dhaene (1996) have noted that Michel’s observation can still be applied, and that the new value of $\Delta$, which will usually be larger than $\lambda^{-1} \sum_{j=1}^{n} p_j^2$, because of the dependence, can be estimated using the Stein–Chen method (Barbour, Holst and Janson, 1992).

In many insurance applications, however, there may be strong local dependence between claim occurrences. For instance, in storm damage insurance, the (rare) single occurrence of a hurricane in a particular town may lead to some random number of claims on the portfolio. Since the distribution of this number of claims may well depend on the time of year, the preceding argument, assuming independent and identically distributed claim amounts, cannot be applied. Despite this, it still seems reasonable to suppose that the distribution of the total number of claims is close to a compound Poisson distribution in total variation, in which case, if once again the individual claim amounts are independent and identically distributed, the distribution of the aggregated claim amount is itself at least as close in total variation to an appropriate compound Poisson distribution, again by Michel’s observation. To exploit this idea, especially if the possibility of substantial local dependence between the random claim numbers is also to be allowed, it is necessary to have an equivalent of the Stein–Chen method, which quantifies the error in total variation when approximating the distribution of a sum of nonnegative random variables, most of them taking the value zero with high probability, by a compound Poisson distribution.

There is a long history of compound Poisson approximation for sums of independent random variables. An excellent presentation is to be found in Le Cam (1965), who, in his introduction, states the following two results, which he attributes in essence to Khint-
chine (1933) and to Prohorov (1953), respectively. Let $X_j$, $1 \leq j \leq n$, be independent random variables with $\mathbb{P}[X_j = 0] = 1 - p_j$ and with $\mathbb{P}[X_j \in A | X_j \neq 0] = \mu_j(A)$ for $A \subset \mathbb{R} \setminus \{0\}$; set $\lambda = \sum_{j=1}^n p_j$ and $\mu = \lambda^{-1} \sum_{j=1}^n p_j \mu_j$, and write $W = \sum_{j=1}^n X_j$. Let CP $(\lambda, \mu)$ denote the compound Poisson distribution of $\sum_{l=1}^N Y_l$, where $(Y_l, l \geq 1)$ are independent and identically distributed with distribution $\mu$ and are independent of $N \sim \text{Po}(\lambda)$, and let

$$d_{TV}(P, Q) = \sup_{A \in \mathbb{R}} |P(A) - Q(A)|$$

define total variation distance on $\mathbb{R}$.

**Theorem 1.1.** $d_{TV}(\mathcal{L}(W), \text{CP} (\lambda, \mu)) \leq \sum_{j=1}^n p_j^2$.

**Theorem 1.2.** If $p_j = p$ and $\mu_j = \mu$ for all $j$, then

$$d_{TV}(\mathcal{L}(W), \text{CP} (np, \mu)) \leq \frac{3}{2} p.$$

In proving the latter result, Le Cam uses the observation made later by Michel.

In the insurance context, a large portfolio can be viewed as the aggregate of many similar smaller portfolios, in which case the relevant asymptotic regime for Theorems 1.1 and 1.2 is that in which the number of risks $n \to \infty$, while the probability $p$ of an event which leads to one or more claims, although small, remains (relatively) constant. If all the $p_j$ and $\mu_j$ are equal, the estimate $\frac{3}{2} p$ of Theorem 1.2 is clearly better than the $np^2$ of Theorem 1.1, since $np \to \infty$; indeed, the former remains bounded and small for small $p$, whereas the latter becomes useless as soon as $np^2$ exceeds 1. If the $p_j$ vary, but $\mu_j = \mu$ for all $j$, the Michel – Le Cam observation improves the order of approximation given by Theorem 1.1 to $\lambda^{-1} \sum_{j=1}^n p_j^2$, analogous to that of Theorem 1.2, since $\lambda^{-1} \sum_{j=1}^n p_j^2$ is just a weighted average of $p_j$ values. Thus the dream is to achieve a general inequality of the form

$$d_{TV}(\mathcal{L}(W), \text{CP} (\lambda, \mu)) \leq C \bar{p}, \quad \bar{p} = \lambda^{-1} \sum_{j=1}^n p_j^2, \quad (1.1)$$

for some $C < \infty$ which may perhaps depend on $\mu$. However, Le Cam gives examples to show that the order $\sum_{j=1}^n p_j^2$ of the bound in Theorem 1.1 can be sharp when the $\mu_j$ are allowed to differ, even when $\lambda \to \infty$.

**Example 1.3.** Let $\mu_j = \delta_{jx}$ for $j \geq 1$, where $\delta_x$ denotes unit mass at $x$; choose $p_j$, $1 \leq j \leq n$, such that $\sum_{j=1}^n p_j^3$ is small and $\lambda = \sum_{j=1}^n p_j$ is large.

Then the ternary expansion of $W$ contains no 2 with probability one, whereas the probability of there being at least one 2 in that of a realization from CP $(\lambda, \mu)$ is of order
\[
\min \left( \sum_{j=1}^{n} p_j^2, 1 \right).
\]
This example is of course unrealistic in the insurance context, since the numbers of claims possible as a result of individual events are staggeringly different, growing geometrically fast with the index \( j \).

**Example 1.4.** Take \( p_j = p' \) and \( \mu_j = \mu' \) for \( j \geq 2 \), where \( \mu' \) is \( d \)-periodic for some \( d \geq 2 \), so that \( \mu' \{2\mathbb{N} \} = 1 \). Take \( p_1 = \frac{1}{2} \) and \( \mu_1 = \delta_1 \).

Although this example avoids the previous drawback, in that all claim sizes may be of the same order of magnitude,

\[
\text{CP}(\lambda, \mu) \{2\mathbb{N} \} - \mathbb{P}[W \in 2\mathbb{N}] \geq e^{-1/2} - \frac{1}{2} > 0
\]

remains bounded away from zero as \( n \) and \( p' \) vary, but \( \bar{p} \leq p' + 1/(4np') \) can become arbitrarily small. This particular scenario is also quite unlikely in the insurance context, but it is now clear that, even in the case of independent \( X_j \); some restriction on the \( \mu_j \) is necessary, if an inequality of the form (1.1) is to be achieved. Indeed, Le Cam observes that

‘approximation in the sense of the norm (total variation) will be possible only in very special cases.’

As soon as the \( X_j \) are dependent, Theorems 1.1 and 1.2 cease to be applicable. However, for weakly dependent \( X_j \)-s, Stein’s method for Poisson process approximation (Barbour 1988, Arratia, Goldstein and Gordon 1990) can be used instead. Here, one first bounds the difference in total variation between the distribution of the random measure \( \Xi = \sum_{j=1}^{n} \mathbb{I}[X_j \neq 0] \delta_{X_j} \) and that of a Poisson point process \( \Pi \) on \( \{1, 2, \ldots, n\} \times \mathbb{R} \) with intensity \( \nu \), defined by \( \nu(\{j\} \times A) = \mathbb{P}[X_j \in A], A \subset \mathbb{R} \setminus \{0\} \). Then \( W = \int x \Xi \{dx\} \) is a functional of \( \Xi \), and the same functional of \( \Pi \) has distribution \( \text{CP}(\lambda, \mu) \). Thus any bound on \( d_{TV}(\mathcal{L}(\Xi), \mathcal{L}(\Pi)) \) is also a bound on \( d_{TV}(\mathcal{L}(W), \text{CP}(\lambda, \mu)) \). Unfortunately, in the settings of Theorems 1.1 and 1.2, where the \( X_j \) are independent, this approach only yields the unsatisfactory estimate

\[
d_{TV}(\mathcal{L}(W), \text{CP}(\lambda, \mu)) \leq \sum_{j=1}^{n} p_j^2
\]

of Theorem 1.1.

A more direct approach is to apply Stein’s method for the compound Poisson distribution \( \text{CP}(\lambda, \mu) \), as introduced in Barbour, Chen and Loh (1992). Roos (1995) has shown
how the method can be applied in a purely routine manner to sums of dependent indicator random variables, and Theorem 1.9 extends her results to sums of dependent nonnegative integer valued random variables. Specializing to the setting of Theorem 1.1, application of Theorem 1.9 gives the following inequality.

**Proposition 1.5.** In the setting of Theorem 1.1, 

\[ d_{TV}(\mathcal{L}(W), \text{CP}(\lambda, \mu)) \leq H_1(\lambda, \mu) \sum_{j=1}^{n} (EX_j)^2 = H_1(\lambda, \mu) \sum_{j=1}^{n} p_j \sum_{i \geq 1} i \mu_{j,i} \]  

(1.3)

Here, \( H_1(\lambda, \mu) \) is a quantity, defined below for (1.13), which bounds the first differences of the solutions to the Stein Equation for \( \text{CP}(\lambda, \mu) \), uniformly for all test functions in the class of indicators of subsets of \( \mathbb{Z}_+ \), and \( \mu_{j,i} = \mu_j \{i\} \). Thus, if the mean claim number 

\[ m_{1j} = \sum_{i \geq 1} i \mu_{j,i} \]

remains bounded for all \( j \), Proposition 1.5 implies a bound of order \( H_1(\lambda, \mu) \sum_{j=1}^{n} p_j^2 \), which is of the desired order \( \bar{p} = \lambda^{-1} \sum_{j=1}^{n} p_j^2 \) whenever the condition

\[(B): \quad H_1(\lambda, \mu) \leq C(\mu) \lambda^{-1} \]

holds for some \( C(\mu) \).

In the case of Poisson approximation to sums of indicator random variables, we have \( \mu = \delta_1 \) and \( H_1(\lambda, \delta_1) \leq \lambda^{-1} \), so that Condition B is satisfied with \( C(\delta_1) = 1 \), explaining why the Stein–Chen method yields such attractive results. For general \( \mu \), things are much more difficult. Barbour, Chen and Loh give the inequality

\[ H_1(\lambda, \mu) \leq C(\lambda(\mu_1 - 2\mu_2))^{-1} [1 + \log^+ \{\lambda(\mu_1 - 2\mu_2)\}], \]  

(1.4)

valid for a universal constant \( C \) provided that \( \mu \) belongs to the class of distributions satisfying 

\[ i \mu_i \geq (i + 1) \mu_{i+1} \]  

(1.5)

for all \( i \geq 1 \). This yields an inequality of the form \( H_1(\lambda, \mu) \leq C(\mu) \lambda^{-1} \{1 + \log^+ \lambda\} \), which is close to that of Condition B, but does not quite achieve it. If (1.5) is not satisfied, the best inequality known seems to be the dreadful

\[ H_1(\lambda, \mu) \leq C \lambda^{-1} e^\lambda, \]  

(1.6)
albeit for a universal constant $C$, and indeed $H_1(\mu, \mu) \geq C(\mu)\epsilon^{\alpha^2}$ for some $\alpha > 0$ whenever $\mu_1 + \mu_2 = 1$ and $\mu_1 < 2\mu_2$, so that nothing much better can be expected in general.

In order to circumvent the difficulty, we use an alternative bound given in Theorem 1.10, which is derived from Theorem 1.9 by a technique from Barbour and Utev (1996). Applying Theorem 1.10 in the case of independent $X_j$’s leads to a bound which is substantially more complicated than that of (1.3), and involves the free choice of a real number $c_1 \in (0, 1)$.

**Proposition 1.6.** Letting $m_l$ denote $\sum_{i \geq 1} i^l \mu_i$ for $l \geq 1$, we have

$$d_{TV}(\mathcal{L}(W), \text{CP}(\lambda, \mu)) \leq \left\{ \frac{H^{(a)}_1(\lambda, \mu)}{\lambda m_1(1 - c_1)} \right\} + \mathbb{P}[W \leq \frac{1}{2}(1 + c_1)\lambda m_1]\left\{ 1 + \frac{2m_2H^{(a)}_0(\lambda, \mu)}{m_1(1 - c_1)} \right\}. \tag{1.7}$$

The quantities $H^{(a)}_i(\lambda, \mu)$, defined below in (1.18), are again derived from solutions to the Stein equation, but are smaller than the $H_i(\lambda, \mu)$; $a$ is given by $c_1\lambda m_1$. The first element in (1.7) is similar to (1.3); the second represents a penalty for replacing $H_1(\lambda, \mu)$ by $H^{(a)}_1(\lambda, \mu)$.

Estimating the quantities $H^{(a)}_i(\lambda, \mu)$ is an arduous process, and takes up most of this paper, but the results are well worth the effort. A detailed statement of what can be proved is given in the next section, with the dependence of $H^{(a)}_i(\lambda, \mu)$ on $\lambda$ and $\mu$ made completely explicit. For now, we concentrate on the asymptotic behaviour appropriate in the insurance context, when $\mu$ is fixed and $\lambda \to \infty$. Then, if $\mu$ has a finite exponential moment and is aperiodic, there exist $c_1 \in (0, 1)$ and $C_0, C_1 < \infty$, depending on $\mu$, such that, with $a = c_1\lambda m_1$,

$$H^{(a)}_0(\lambda, \mu) \leq C_0(\mu)\lambda^{-1/2} \quad \text{and} \quad H^{(a)}_1(\lambda, \mu) \leq C_1(\mu)\lambda^{-1}. \tag{1.8}$$

Thus, for independent $X_j$’s satisfying $\lambda^{-1}\sum_{j=1}^n P_j \mu_j = \mu$, it follows from Proposition 1.6 that

$$d_{TV}(\mathcal{L}(W), \text{CP}(\lambda, \mu)) = O\left\{ m_2 + \mathbb{P}[W \leq \frac{1}{2}(1 + c_1)\lambda m_1] \right\}, \tag{1.9}$$

where $m_{1+} = \max_{1 \leq j \leq n} m_1 j$. Chebyshev’s inequality is now enough to establish that

$$\mathbb{P}[W \leq \frac{1}{2}(1 + c_1)\lambda m_1] \leq \lambda m_2\left[ \frac{1}{2}(1 - c_1)\lambda m_1 \right]^{-2} = O(\lambda^{-1}),$$

6
so that we have obtained a bound of the ideal order $\bar{p} = \lambda^{-1} \sum_{j=1}^{n} p_j^2$ of (1.1) whenever
\[
\sum_{j=1}^{n} p_j^2 \text{ is bounded below and } m_{1+} \text{ is bounded above; for example, with equal } p_j, \text{ this is}
\]
achieved whenever $n \geq 1/p^2$. For more refinement, Bernstein’s inequality easily gives
\[
\mathbb{P}[W \leq \frac{1}{2}(1+c_1)\lambda m_1] \leq \exp\{-\frac{1}{8}\lambda m_1^2(1-c_1)^2/(m_2 + m_{1+}^2)\},
\]
so that (1.1) actually holds whenever $m_{1+}$ is bounded above and
\[
\exp\{-\frac{1}{8}\lambda m_1^2(1-c_1)^2/(m_2 + m_{1+}^2)\} = O(\bar{p}); \tag{1.10}
\]
for equal $p_j$’s, (1.10) requires merely that $n \geq \alpha p^{-1} \log\{p^{-1}\}$ for $\alpha = 8m_2[m_1(1-c_1)]^{-2}$,
which is hardly more restrictive than $\lambda \to \infty$. Thus a bound of order $\bar{p}$, as in (1.1), follows
from Proposition 1.6 as $\lambda \to \infty$ for all aperiodic distributions $\mu$ with finite exponential moment, except when $\lambda$ grows extremely slowly; but then, in any case, the improvement
of $\bar{p}$ as compared to $\sum_{j=1}^{n} p_j^2$ is minimal, as expressed in the following result.

**Proposition 1.7.** Let $\{X_j, 1 \leq j \leq n\}$ be independent nonnegative integer valued random variables which satisfy $\mathcal{L}(X_j) = (1-p_j)\delta_0 + p_j \mu_j$. Define
\[
\lambda = \sum_{j=1}^{n} p_j; \quad W = \sum_{j=1}^{n} X_j; \quad m_{1+} = \max_{1 \leq j \leq n} m_{1j} \quad \text{and} \quad \mu = \lambda^{-1} \sum_{j=1}^{n} p_j \mu_j.
\]
Suppose that $\mu$ has an exponential moment and is aperiodic. Then there exists a constant $C(\mu)$ such that
\[
d_{TV}(\mathcal{L}(W), \text{CP} (\lambda, \mu)) \leq m_{1+}^2 C(\mu) \bar{p} \max\{1, \log(1/\bar{p})\},
\]
and $d_{TV}(\mathcal{L}(W), \text{CP} (\lambda, \mu)) \leq m_{1+}^2 C(\mu) \bar{p}$ if $\exp\{-\alpha \lambda\} \leq m_{1+}^2 \bar{p}$, where
$\alpha = \frac{1}{8}m_1^2(1-c_1)^2/(m_2 + m_{1+}^2)$ and $m_1$, $m_2$ and $c_1$ are as defined for Proposition 1.6.

Thus Proposition 1.6 and Theorem 1.1 together provide a very comprehensive solution to
Le Cam’s problem of compound Poisson approximation in total variation, in the case of sums of independent nonnegative integer valued random variables. Even more importantly for theoretical as well as practical applications (Aldous, 1988), the techniques of this paper
remain applicable for random variables with substantial local dependence.

For the estimate (1.8), we assume that $\mu$ has a finite exponential moment and is aperiodic. To what extent are these assumptions necessary for either (1.8) or (1.1) to hold? The existence of an exponential moment is used heavily in our proofs of the estimates of
$H_i^{(a)}(\lambda, \mu)$, but may well only have technical character. Nonetheless, some aspect of the size of the values typically obtained under $\mu$ must enter any $C(\mu)$ in (1.1), because of examples similar to Example 1.3.

For $\mu$ with period $d \geq 2$, Proposition 1.7 can be recovered by considering instead $\bar{X}_j = d^{-1}X_j$, $\overline{W} = d^{-1}W$, for which $\bar{\mu}$ is aperiodic. Despite this, an aperiodicity assumption is essential. First, the solution to the CP $(\lambda, \delta_2)$ Stein Equation corresponding to the test function $1_{2\mathbb{N}}$ has an explicit integral representation, from which it can be shown that (1.8) cannot hold for any $a = c\lambda m_1$ with $c < 1$. Secondly, taking a sequence of instances of Example 1.4 with $p'_n = n^{-2/3}$ and $\mu'$ fixed and periodic, we have

$$\lambda_n \sim n^{1/3}, \quad \bar{\mu}_n \sim n^{-2/3} \to 0 \quad \text{and} \quad \mu^{(n)} = \left\{1 - \frac{1}{2\lambda_n}\right\} \mu' + \frac{1}{2\lambda_n} \delta_1,$$

but

$$d_{TV}(\mathcal{L}(W_n), \text{CP} (\lambda_n, \mu^{(n)})) \geq \varepsilon^{-1/2} - \frac{1}{7} > 0,$$

uniformly for all $n$. Since arbitrary neighbourhoods $N(\mu')$ of $\mu'$ with respect to most metrics contain all but finitely many of the $\mu^{(n)}$, it thus follows that (1.1) cannot hold for any fixed $C < \infty$ uniformly in any neighbourhood of a periodic $\mu'$.

1b. Elaboration.

Stein's method for compound Poisson approximation in total variation can be viewed as follows. If $W$ is a nonnegative integer valued random variable for which it can be shown that

$$\left| \mathbb{E}\left\{ \sum_{i \geq 1} i\lambda_i g(W + i) - W g(W) \right\} \right| \leq \varepsilon_0 M_0(g) + \varepsilon_1 M_1(g) \quad (1.11)$$

for all bounded $g$: $\mathcal{N} \to \mathbb{R}$ and for some (small) $\varepsilon_0$ and $\varepsilon_1$, where

$$M_0(g) = \sup_{j \geq 1} |g(j)|; \quad M_1(g) = \sup_{j \geq 1} |g(j + 1) - g(j)|, \quad (1.12)$$

then it follows that

$$d_{TV}(\mathcal{L}(W), \text{CP} (\lambda, \mu)) \leq \varepsilon_0 H_0(\lambda, \mu) + \varepsilon_1 H_1(\lambda, \mu), \quad (1.13)$$

where $\lambda = \sum_{i \geq 1} \lambda_i$, $\mu_i = \lambda_i / \lambda$ and $H_l(\lambda, \mu) = \sup_{A \in \mathbb{Z}^+} M_l(g_A)$, $l = 0, 1$, with $g_A$ the solution of the Stein Equation

$$\sum_{i \geq 1} i\lambda_i g(w + i) - wg(w) = 1_A(w) - \text{CP} (\lambda, \mu)\{A\}, \quad w \geq 0. \quad (1.14)$$

8
Thus, in order to show that the distribution of $W$ is close to compound Poisson, it is
equivalent to show that (1.11) can be satisfied for suitably small $\varepsilon_0$ and $\varepsilon_1$, provided that the
constants $H_l(\lambda, \mu)$ are not too large; note that this last is a requirement on $\lambda$ and $\mu$, and
not on $W$ itself.

The following lemma shows that $\varepsilon_0$ and $\varepsilon_1$ can in many circumstances be effectively
bounded. Let $(X_\alpha, \alpha \in \Gamma)$ be nonnegative integer valued random variables with finite
means $m_{1\alpha}$, where $\Gamma$ is some finite set of indices. For each $\alpha \in \Gamma$, decompose $W = \sum_{\beta \in \Gamma} X_\beta$
in the form

$$W = X_\alpha + Z_\alpha + U_\alpha + W_\alpha.$$  (1.15)

Such a representation is useful in what follows if $Z_\alpha$ contains that part of $W$ which is
strongly dependent on $X_\alpha$, $W_\alpha$ is almost independent of $(X_\alpha, Z_\alpha)$, and $U_\alpha$ and $Z_\alpha$ are
not too large: the sense in which these requirements are to be interpreted becomes clear
shortly. Frequently, the decomposition is realized by defining a partition

$$\Gamma = \{\alpha\} \cup \Gamma^{u}\alpha \cup \Gamma^{b}\alpha \cup \Gamma^{w}\alpha,$$  (1.16)

and setting

$$W = \sum_{\alpha \in \Gamma} X_\alpha, \quad Z_\alpha = \sum_{\beta \in \Gamma^{u}\alpha} X_\beta, \quad U_\alpha = \sum_{\beta \in \Gamma^{b}\alpha} X_\beta, \quad W_\alpha = \sum_{\beta \in \Gamma^{w}\alpha} X_\beta;$$

$\Gamma^{u}\alpha$ contains those $X_\beta$ which strongly influence $X_\alpha$, and $\Gamma^{w}\alpha$ those $X_\beta$ whose cumulative
effect on $(X_\alpha, Z_\alpha)$ is negligible. Define the quantities

$$\lambda = \sum_{\alpha \in \Gamma} \mathbb{E}\left\{ \left( \frac{X_\alpha}{X_\alpha + Z_\alpha} \right) 1\{U_\alpha \geq 1\} \right\};$$

$$\mu_1 = \frac{1}{\lambda} \sum_{\alpha \in \Gamma} \mathbb{E}\{X_\alpha I[X_\alpha + Z_\alpha = l]\}, \quad l \geq 1;$$

$$\pi^{(a)}_{ik} = \mathbb{P}[X_\alpha = i, Z_\alpha = k|m_{1\alpha}], \quad i \geq 1, \quad k \geq 0;$$

$$\delta_1 = \sum_{\alpha \in \Gamma} m_{1\alpha} \sum_{i \geq 1} \sum_{k \geq 0} \pi^{(a)}_{ik} \mathbb{E}\left[ \frac{\mathbb{P}[X_\alpha = i, Z_\alpha = k|W_\alpha]}{\mathbb{P}[X_\alpha = i, Z_\alpha = k]} - 1 \right];$$

$$\delta_2 = 2 \sum_{\alpha \in \Gamma} \mathbb{E}\{X_\alpha d_{TV}(\mathcal{L}(W_\alpha), \mathcal{L}(W_\alpha))\};$$

$$\delta_3 = \sum_{\alpha \in \Gamma} \mathbb{E}\{X_\alpha d_{W}(\mathcal{L}(W_\alpha|X_\alpha, Z_\alpha), \mathcal{L}(W_\alpha))\};$$

$$\delta_4 = \sum_{\alpha \in \Gamma} \left\{ \mathbb{E}(X_\alpha U_\alpha) + \mathbb{E}X_\alpha \mathbb{E}\{X_\alpha + Z_\alpha + U_\alpha\} \right\}.$$
In $\delta_3$, $d_W$ denotes the Wasserstein $L_1$ metric on probability measures over $\mathbb{R}$: if $\mathcal{F} = \{ f : \mathbb{R} \to \mathbb{R}, |f(x) - f(y)| \leq |x - y| \text{ for all } x, y \in \mathbb{R} \}$, then $d_W(P, Q) = \sup_{f \in \mathcal{F}} \int f \, dP - \int f \, dQ$. The quantities $\delta_l$, $1 \leq l \leq 4$, appear as elements in the total variation distance between $\mathcal{L}(W)$ and $\text{CP}(\lambda, \mu)$, and they should be small if approximation is to be good. For $1 \leq l \leq 3$, $\delta_l/EW$ is a measure of the dependence between $W_\alpha$ and $(X_\alpha, Z_\alpha)$; $\delta_4$ is small in comparison with $\text{EW}$ if $U_\alpha$ and $Z_\alpha$ are small in expectation, provided that $U_\alpha$ is not too strongly dependent on $X_\alpha$.

**Lemma 1.8.** With the assumptions and definitions above,

$$\left| \mathbb{E} \left\{ \sum_{i \geq 1} i \lambda \mu_i g(W + i) - W g(W) \right\} \right| \leq \varepsilon_0 M_0(g) + \varepsilon_1 M_1(g),$$

(i) with $\varepsilon_0 = \min(\delta_1, \delta_2)$ and $\varepsilon_1 = \delta_4$; (ii) with $\varepsilon_0 = 0$ and $\varepsilon_1 = \delta_3 + \delta_4$.

Furthermore, for any choice of $\mu'$,

$$\left| \mathbb{E} \left\{ \sum_{i \geq 1} i \lambda \mu'_i g(W + i) - W g(W) \right\} \right| \leq \varepsilon'_0 M_0(g) + \varepsilon'_1 M_1(g),$$

with

$$\varepsilon'_0 = \varepsilon_0 + \lambda |m_1 - m'_1| \quad \text{and} \quad \varepsilon'_1 = \varepsilon_1 + \lambda m'_1 d_W(\mu^*, \mu'^*);$$

here, $*$ is used to denote a size-biased distribution: $\mu^*_i = i \mu_i / m_1$.

**Remark.** (a) For independent $X_\alpha$, take $\Gamma^*_\alpha = \Gamma^b_\alpha = \emptyset$, and observe that $\delta_1 = \delta_2 = \delta_3 = 0$, and that $\delta_4$ reduces to $\sum_{\alpha \in \mathcal{F}} (\mathbb{E}X_\alpha)^2$.

(b) In evaluating $\delta_2$ and $\delta_3$, it is often possible to compute the distances between distributions by means of couplings. This ‘coupling approach’ is most commonly applied with $\Gamma^b_\alpha = \emptyset$.

**Proof.** In essence, we follow the argument of Roos (1995). We first note that, for any bounded $g$,

$$\mathbb{E}(X_\alpha g(W)) = \sum_{i \geq 1} i \mathbb{E}\{g(i + Z_\alpha + U_\alpha + W_\alpha)I[X_\alpha = i]\}$$

$$= \sum_{i \geq 1} \sum_{k \geq 0} i \mathbb{E}\{g(i + k + U_\alpha + W_\alpha)I[X_\alpha = i, Z_\alpha = k]\}$$

$$= \sum_{i \geq 1} \sum_{k \geq 0} i \mathbb{E}\{g(i + k + W_\alpha)I[X_\alpha = i, Z_\alpha = k]\} + \eta_{1\alpha},$$
where $|\eta_{1\alpha}| \leq \mathbb{E}(X_{\alpha}U_{\alpha})M_{1}(g)$. In much the same way,
\[
\sum_{i \geq 1} \sum_{k \geq 0} i \mathbb{E}g(i + k + W_{\alpha}) \mathbb{P}[X_{\alpha} = i, Z_{\alpha} = k] = \sum_{i \geq 1} \sum_{k \geq 0} i \mathbb{P}[X_{\alpha} = i, Z_{\alpha} = k] \mathbb{E}g(i + k + W) + \eta_{2\alpha}
\]
\[
= \sum_{i \geq 1} \mathbb{E}\{X_{\alpha}I[X_{\alpha} + Z_{\alpha} = l]\} \mathbb{E}g(l + W) + \eta_{2\alpha},
\]
with $|\eta_{2\alpha}| \leq \mathbb{E}X_{\alpha}\mathbb{E}(X_{\alpha} + Z_{\alpha} + U_{\alpha})M_{1}(g)$. Finally, by choosing to condition either on $W_{\alpha}$ or on $(X_{\alpha}, Z_{\alpha})$, one obtains
\[
\sum_{\alpha \in \Gamma} \left| \sum_{i \geq 1} \sum_{k \geq 0} \mathbb{E}\{g(i + k + W_{\alpha})(I[X_{\alpha} = i, Z_{\alpha} = k] - \mathbb{P}[X_{\alpha} = i, Z_{\alpha} = k])\} \right|
\]
\[
\leq \min\{\delta_{1}M_{0}(g), \delta_{2}M_{0}(g), \delta_{3}M_{1}(g)\},
\]
concluding the proof of the first part. For the second, it suffices to note that
\[
\sum_{i \geq 1} i\lambda(\mu'_{i} - \mu_{i})g(W + i)
\]
\[
= \lambda m_{1} \sum_{i \geq 1} g(W + i) \left( \frac{ip'_{i}}{m'_{i}} - \frac{ip_{i}}{m_{i}} \right) + \lambda \sum_{i \geq 1} \frac{ip'_{i}}{m_{i}} (m'_{i} - m_{i}) g(W + i),
\]
since $|g(W + i)| \leq M_{0}(g)$ and $g(W + \cdot)/M_{1}(g) \in \mathcal{F}$.

**Theorem 1.9.** In the setting above,
\[
d_{TV}(\mathcal{L}(W), \text{CP}(\lambda, \mu)) \leq \varepsilon_{0}H_{0}(\lambda, \mu) + \varepsilon_{1}H_{1}(\lambda, \mu),
\]
with $\varepsilon_{0}$ and $\varepsilon_{1}$ as in Lemma 1.8; similarly,
\[
d_{TV}(\mathcal{L}(W), \text{CP}(\lambda, \mu')) \leq \varepsilon'_{0}H_{0}(\lambda, \mu') + \varepsilon'_{1}H_{1}(\lambda, \mu').
\]

**Proof.** From Lemma 1.8 and (1.11)-(1.13); note that $\lambda_{i} = \lambda\mu_{i}$.

The remaining step would be to find reasonable bounds on the $H_{l}(\lambda, \mu)$, but, as discussed in the introduction, such bounds may not exist. However, recalling Inequality (4.4) from Barbour and Utev, (1.13) can be replaced by
\[
d_{TV}(\mathcal{L}(W), \text{CP}(\lambda, \mu)) \leq \varepsilon_{0}H_{0}(\lambda, \mu) + \varepsilon_{1}\left( H_{1}(\lambda, \mu) + \frac{H_{0}(\lambda, \mu)}{b - a} \right)
\]
\[
+ \mathbb{P}[W \leq \delta]\left\{ 1 + \frac{\lambda m_{2}H_{0}(\lambda, \mu)}{b - a} \right\},
\]
where $0 < a < b$ are arbitrary, and
\[
H_{l}(\lambda, \mu) = \sup_{A \subseteq \mathbb{Z}_{+}} M_{l}(g_{A}(\cdot + a)), \quad l = 0, 1.
\]
This leads to alternative bounds.
Theorem 1.10. In the setting above, taking \( c_1 \in (0, 1) \), \( a = c_1 \lambda m_1 \), we have

\[
d_{TV}(\mathcal{L}(\mathcal{W}), CP(\lambda, \mu)) \leq \varepsilon_0 \tilde{H}_0^{(a)}(\lambda, \mu) + \varepsilon_1 \left( H_1^{(a)}(\lambda, \mu) + \frac{2 \tilde{H}_0^{(a)}(\lambda, \mu)}{\lambda m_1(1 - c_1)} \right) + \mathbb{P}[W \leq \frac{1}{2}(1 + c_1)\lambda m_1] \left\{ 1 + \frac{2 m_2 \tilde{H}_0^{(a)}(\lambda, \mu)}{m_1(1 - c_1)} \right\},
\]

with \( \varepsilon_0 \) and \( \varepsilon_1 \) as in Lemma 1.8; similarly, \( \mu \) can be replaced by any other \( \mu' \) on both sides, if also \( m_1 \) and \( m_2 \) are replaced by \( m_1' \) and \( m_2' \), and \( \varepsilon_0 \) and \( \varepsilon_1 \) are replaced by \( \varepsilon_0' \) and \( \varepsilon_1' \) as in Lemma 1.8.

Note that only \( c_1 < 1 \) is of interest, since otherwise \( \mathbb{P}[W \leq \frac{1}{2}(1 + c_1)\lambda m_1] \) is not small.

There now remains the step of finding bounds for the \( H_l^{(a)}(\lambda, \mu) \), before Theorem 1.10 can be applied, but this turns out to be a feasible task. The argument is extremely long and delicate, and makes extensive use of a new representation of the Fourier transform of the solution of the Stein Equation. Our main result is as follows.

**Theorem TV.** Suppose that the power series \( \mu(z) = \sum_{k \geq 1} \mu_k z^k \) has radius of convergence \( R(\mu) > 1 \), that Assumption A below holds and that \( \lambda \geq 2 \). Then there exist constants \( C_0(\mu), C_1(\mu) \) and \( C_2(\mu) \), given explicitly in terms of \( \mu \) in (1.26)–(1.28) below, such that, for any \( a \geq C_2(\mu)\lambda m_1 + 1 \),

\[
H_0^{(a)}(\lambda, \mu) \leq \lambda^{-1/2} C_0(\mu) \quad \text{and} \quad H_1^{(a)}(\lambda, \mu) \leq \lambda^{-1} C_1(\mu),
\]

where \( H_l^{(a)}(\lambda, \mu), l = 0, 1 \), are as defined in (1.18); note that \( C_2(\mu) < 1 \), so that there are feasible choices for Theorem 1.10 which are covered by the theorem.

To give expressions for \( C_0, C_1 \) and \( C_2 \) in terms of \( \mu \), some further notation is required. We assume throughout that the power series \( \mu(z) = \sum_{k \geq 1} \mu_k z^k \) has radius of convergence \( R(\mu) \) larger than 1. For each \( j \geq 1 \), we define the additional moments

\[
m_{[j]} = \sum_{k \geq 1} k_{[j]} \mu_k \quad \text{and} \quad m_{(r)}^j = \sum_{k \geq 1} k^j \mu_k r^k,
\]

where \( k_{[j]} = k(k-1)\ldots(k-j+1) \) and \( r < R(\mu) \); note that \( m_j = \lim_{r \to 1} m_{(r)}^j \). We also define the auxiliary quantities

\[
\gamma = m_{[2]} + \frac{1}{2} m_1 \leq m_2; \quad \gamma_* = \sup_{j \geq 1} \frac{m_{[j]}}{(j - 4)!}^{1/j} ; \quad d_* = \frac{1}{2} (m_*^3 + 3 m_1); \quad \gamma_1 = \gamma^{-1} m_*; \quad \gamma_2 = \gamma^{-1} d_* ,
\]

12
where \( n! \) is taken to be 1 if \( n \leq 0 \). Note that, in \( \lambda \geq 2, \lambda \gamma \geq 1 \) and \( \lambda m_s \geq 2 \).

The aperiodicity assumption of Barbour and Utev (1996) is also needed. To express this, define

\[
\rho_1(\theta) = 1 - \sum_{i \geq 1} \mu_i \cos i \theta; \quad \rho_2(\theta) = 1 - \frac{1}{m_1} \sum_{i \geq 1} i \mu_i \cos i \theta,
\]

noting that \( 0 \leq \rho_l(\theta) \leq 2, l = 1, 2 \), and set

\[
\rho^*_1(\zeta) = \inf_{\zeta \leq \theta \leq \pi} \rho_l(\theta) \quad \text{and} \quad \rho^*(\zeta) = \min(\rho^*_1(\zeta), \frac{1}{2} \rho^*_2(\zeta), 1).
\]

Assumption A. For any \( 0 < \zeta \leq \pi \), \( \rho^*(\zeta) > 0 \).

We also define

\[
\rho^{(r)}_1(\theta) = 1 - \frac{1}{\mu^{(r)}} \sum_{i \geq 1} r^i \mu_i \cos i \theta; \quad \rho^{(r)}_2(\theta) = 1 - \frac{1}{m^{(r)}_1} \sum_{i \geq 1} i r^i \mu_i \cos i \theta,
\]

\[
\rho^{*(r)}(\zeta) = \min\left\{ \inf_{\zeta \leq \theta \leq \pi} \rho^{(r)}_1(\theta), \frac{1}{2} \inf_{\zeta \leq \theta \leq \pi} \rho^{(r)}_2(\theta), 1 \right\},
\]

and note that \( \lim_{r \to 1} \rho^{*(r)}(\zeta) = \rho^*(\zeta) \).

We actually work with specific values of \( \zeta \) and \( r \). First, we pick \( \zeta < \sqrt{2m_2/m_4} \). Then we choose \( 1 < r < R(\mu) \) small enough that

\[
\frac{m^{(r)}_2}{m^{(r)}_4} > \frac{2m_2}{3m_4} \quad \text{and} \quad m^{(r)}_1 \left( 1 - \rho^{*(r)}(\zeta) \right) < m_1 \left( 1 - \frac{1}{2} \rho^*(\zeta) \right),
\]

and so that \( s = \sqrt{r^2 - 1} \) satisfies

\[
s < \frac{1}{9} \min\left\{ 1, \frac{1}{m_s}, \frac{\gamma}{m^2_s} \right\}.
\]

In terms of these quantities, we can now state the values taken by \( C_0, C_1 \) and \( C_2 \):

\[
C_0 = \frac{1}{\sqrt{\gamma}} \left\{ (37 + 23\eta_1^{3/2}) + \eta_2(26 + 51\eta_1) \right\} + \frac{1}{m_1} \left\{ 18 + \frac{5e r}{(r - 1) \rho^*(\zeta)} \right\}
\]

\[
+ \frac{1}{\pi(r - 1) \sqrt{\gamma}} \left\{ 8 + 31 \sqrt{\eta_1(r - 1)} + 2 \pi r \sqrt{\frac{m_2}{m_4(r - 1)}} \right\};
\]

\[
C_1 = \frac{1}{\gamma} \left\{ (130 + 15\eta_1) + \eta_2(232 + 19\eta_1) \right\} + \frac{1}{m_1} \left\{ 28 + \frac{12e r}{(r - 1) \rho^*(\zeta)} + 4 \sqrt{\frac{\pi}{m_1}} \left( 1 + \frac{7.2 m_3}{m^3_s} \right) \right\}
\]

\[
+ \frac{1}{(r - 1) \gamma} \left\{ 4r \sqrt{\frac{m_2}{m_4(r - 1)}} + (r - 1)^{1/2} (-\log(r - 1) + 2 + 9\eta_1) \right\};
\]

\[
C_2 = 1 - \rho^*(\zeta)/4.
\]
The precise form of the expressions for $C_0$, $C_1$ and $C_2$ is not of paramount importance. The main significance of Theorem TV is that such constants exist, and no attempt has been made to optimize them. However, the form of expressions (1.26)–(1.28) illustrates how the assumptions required for the theorem are used in the proof. If $\mu$ has no exponential moment, so that $R(\mu) = 1$, it is impossible to find an $r \in (1, R(\mu))$, and this is reflected in the expressions for $C_1$ and $C_2$, which become infinite if $r \to 1$. If $\mu$ is periodic, $\rho^*(\zeta)$ can be expected to take the value 0, which also has the effect of making the expressions for $C_1$ and $C_2$ become infinite; similarly, distributions $\mu$ which are very close to being periodic will have $\rho^*(\zeta)$ close to zero, and $C_1$ and $C_2$ will be correspondingly large, as must be the case, in view of Example 1.4.

However, the concrete expressions also have another important consequence, to show that the phenomenon in Example 1.4 really is caused by the convergence to a periodic limiting $\mu$-distribution. Suppose that $\mu$ has radius of convergence $R > 1$ and satisfies Assumption A. Let $\mathcal{M}_R = \{\mu' : \sup_{i \geq 1} R^i \mu_i < \infty\}$, and define a distance on $\mathcal{M}_R$ by $d(\mu^1, \mu^2) = \sup_{i \geq 1} R^i |\mu^1_i - \mu^2_i|$. Then any difference of the form $|\mu^1_1(t) - m_1(t)|$, for $l \geq 1$ and $1 \leq t < R$, is uniformly bounded in the ball $B_\varepsilon = \{\mu' \in \mathcal{M}_R : d(\mu^1, \mu) \leq \varepsilon\}$ by some $\phi_t(\varepsilon)$, where $\lim_{\varepsilon \to 0} \phi_t(\varepsilon) = 0$, and the same is true for the $\rho_1(t)$. Hence, for $\varepsilon$ sufficiently small, the same choices of $\zeta$ and $r$ can be used for all elements $\mu' \in B_\varepsilon$ as are used for $\mu$; and, if $\mu$ satisfies Assumption A, the expressions for $C_0$ and $C_1$ are uniformly bounded above and $C_2$ is uniformly bounded away from 1, for all elements $\mu' \in B_{\varepsilon'}$, for some $0 < \varepsilon' \leq \varepsilon$. Thus, for any $R > 1$, the functions $C_0(\mu)$, $C_1(\mu)$ and $C_2(\mu)$ are uniformly continuous within $\mathcal{M}_R$ at any aperiodic $\mu$.

In much the same spirit, it is also possible to use Theorems 1.8 and TV to formulate an analogue of Proposition 1.7 for triangular arrays, when $\lambda_n \to \infty$ sufficiently fast and $\mu^{(n)} \to \mu$.

**Proposition 1.11.** Consider a triangular array of non-negative integer valued random variables $\{X_{jn}, 1 \leq j \leq n, n \geq 1\}$, which are independent within rows and satisfy $\mathcal{L}(X_{jn}) = (1 - p_{jn}) \delta_0 + p_{jn} \mu_j^{(n)}$. Define

$$
\lambda_n = \sum_{j=1}^{n} p_{jn}; \quad W_n = \sum_{j=1}^{n} X_{jn}; \quad m_{1+}^{(n)} = \max_{1 \leq j \leq n} m_{1j}^{(n)} \quad \text{and} \quad \mu^{(n)} = \lambda_n^{-1} \sum_{j=1}^{n} p_{jn} \mu_j^{(n)}.
$$

Assume that $\mu \in \mathcal{M}_R$ is aperiodic, and that

$$
\lambda_n \geq \frac{100(m_2 + (m_{1+}^{(n)})^2)}{m_1(1 - c_1)^2} \log \left( \frac{1}{\mu^{(n)} m_{1+}^{(n)}^2} \right).
$$
for all \( n \) sufficiently large; here, \( m_1, m_2 \) and \( c_2 \) are as before, \( C_2(\mu) < c_1 < 1 \) and \( * \) denotes size-biasing. Then

\[
d_{TV}(\mathcal{L}(W_n), \text{CP}(\lambda_n, \mu)) = O\left\{ m_1^{(n)} \frac{1}{2} \tilde{P}^{(n)} + \lambda_n^{1/2} |m_1^{(n)} - m_1| + m_1 d_W(\mu^{(n)*}, \mu^*) \right\}.
\]

**Proof.** The condition on \( \lambda_n \) is merely chosen to ensure that

\[
P[W_n \leq \frac{1}{2}(1 + c_1) \lambda m_1] = o\left( \tilde{P}^{(n)} \left\{ m_1^{(n)} \right\} \right),
\]

which is established by using the exponential Chebyshev inequality preceding (1.10) with \( c_1' = \frac{1}{2}(1 + c_1) \) replacing \( c_1 \), and by observing that \( m_1^{(n)} \to m_1 \) and \( m_2^{(n)} \to m_2 \) if \( d_W(\mu^{(n)*}, \mu^*) \to 0 \). The second statement in Theorem 1.10 is now applied with \( \mu^{(n)} \) for \( \mu \) and \( \mu \) for \( \mu' \), and Theorem TV is used to estimate the quantities \( H_{l}^{(a)}(\lambda, \mu) \).

The form of the error estimate is rather different from those in analogous theorems for normal approximation. This is to be expected, because the mode of convergence is different, and because there is no centring or normalization. The difference between the means of \( \text{CP}(\lambda_n, \mu) \) and \( \mathcal{L}(W_n) \) is \( \lambda_n(m_1 - m_1^{(n)}) \), as compared to the standard deviation \( \sqrt{\lambda_n m_2} \) of \( \text{CP}(\lambda_n, \mu) \); for normal distributions, this would imply a total variation discrepancy of order \( \lambda_n^{1/2} |m_1^{(n)} - m_1| \). Then, if \( m_1^{(n)} = m_1 \) but \( m_2^{(n)} \neq m_2 \), an error arises just because of the discrepancy in the variances, which is taken into account by the element \( m_1 d_W(\mu^{(n)*}, \mu^*) \).

**2. Outline proof of Theorem TV.** The proof is long and technical, and is based on two integral representations for the solution \( g_A \) to the Stein Equation (1.14). For the first, define

\[
g^x(j) = \int_x^1 z^{j-1} \exp \{ \lambda [\mu(x) - \mu(z)] \} \, dz,
\]

where \( x \) and the contour of integration lie within the circle of convergence of \( \mu(\cdot) \). The importance of the functions \( g^x \) lies in the fact that \( g_A \) can be obtained from them by Fourier inversion (Barbour and Utev 1995, (2.8) and (2.9));

**Representation 1.**

\[
g_A(j) = \sum_{k \in \mathbb{A}} \frac{1}{2 \pi i} \int_{|x| = 1} x^{-k-1} g^x(j) \, dx.
\]

The second representation takes the form of an integral along the unit interval. We write \( \lambda \) in the equivalent form \( \lambda \mu \), and we define \( \mu^{(t)} \) by \( \mu_i^{(t)} = \mu_i(1 - t^i) \), \( 0 \leq t \leq 1 \).
Representation 2.
\[ g_A(j) = \sum_{k \in A} \int_0^1 t^{j-1} e^{\lambda(\mu(1) - \mu(t))} \left( \text{CP} (\lambda \mu(t)) \{k - j \} - \text{CP} (\lambda, \mu) \{k \} \right) dt, \]

where \( \text{CP} (\cdot) \{a \} \) is taken to be zero for negative \( a \).

**Proof.** Let \( S(\lambda \mu) \) be a random variable with distribution \( \text{CP} (\lambda, \mu) \). Taking the formula
\[ \mathbb{E} z^{S(\lambda \mu)} = \exp \{ \lambda(\mu(z) - \mu(1)) \} \]
and inverting the Fourier transform, we have
\[ \text{CP} (\lambda, \mu) \{n\} = \frac{1}{2\pi i} \int_{|z|=1} z^{-n-1} \exp \{ \lambda(\mu(z) - \mu(1)) \} dz; \]

note that, for negative \( n \), the contour integral gives the correct zero value. Now
\[ g^x(j) = \int_0^1 z^{j-1} e^{\lambda(\mu(x)-\mu(z))} dz - \int_0^1 t^{j-1} e^{\lambda(\mu(x)-\mu(t))} dt \]
\[ = \int_0^1 \left\{ z^{j-1} e^{\lambda(\mu(x)-\mu(z))} - x^{j-1} e^{\lambda(\mu(x)-\mu(xz))} \right\} dz, \]

and so it follows that
\[ g(k)(j) = \frac{1}{2\pi i} \int_{|z|=1} x^{-k-1} g^x(j) dx \]
\[ = \int_0^1 z^{j-1} e^{\lambda(\mu(1)-\mu(z))} \text{CP} (\lambda, \mu) \{k \} dz - \int_0^1 z^{j-1} e^{\lambda(\mu(1)-\mu(z))} \text{CP} (\lambda \mu^{(z)}) \{k - j \} dz, \]

since
\[ e^{\lambda(\mu(x)-\mu(xz))} = e^{\lambda(\mu(1)-\mu(z))} \mathbb{E} x^{S(\lambda \mu^{(z)})}. \]

This proves the representation.

Using the two representations, the bounds on \(|g_A(j)|\) and \(|g_A(j+1) - g_A(j)|\) are then established in stages. The simplest, given in Lemma 3.1, covers any \( A \) and all \( j \geq 3\lambda m_1/2 \), and uses Representation 2. For smaller \( j \), we write
\[ g_A(j) = g_{A_1}(j) + g_{A_2}(j) + g_{A_3}(j), \tag{2.2} \]

where
\[ A_1 = A \cap [0, j - 1]; \quad A_3 = A \cap [j + 4\lambda m_*, \infty) \quad \text{and} \quad A_2 = A \setminus \{A_1 \cup A_2\}. \tag{2.3} \]
The argument for \( g_{A_1} \) (Lemmas 3.2 and 3.3) uses Representation 2 and accurate large deviation estimates of the lower tail of a compound Poisson distribution, the bound being valid for all \( j \geq \frac{1}{2} \lambda m_1 + 1 \). The argument for \( g_{A_2} \) (Lemma 3.4) uses a more elementary estimate of the upper tail, and is valid in the same range. Finally, for \( g_{A_2} \), we investigate the contributions from each element \( g_{\{j+K\}}(j) \) separately, showing that

\[
|g_{A_2}(j)| \leq \sum_{K=0}^{[4\lambda m_1]} |g_{\{j+K\}}(j)| \leq C \lambda^{-1/2};
\]

\[
|g_{A_2}(j) - g_{A_2}(j+1)| \leq \sum_{K=0}^{[4\lambda m_1]} |g_{\{j+K\}}(j) - g_{\{j+K\}}(j+1)| \leq C \lambda^{-1},
\]

for suitable constants \( C \).

To prove (2.4), we use Representation 1, replacing the unit circle as contour of integration by a flattened circular contour \( S = S_1 \cup S_2 \cup S_3 \), where

\[ S_1 = \{re^{i\theta} : \zeta \leq |\theta| \leq \pi \}, \quad S_2 = \{re^{i\theta} : t_+ \leq |\theta| < \zeta \} \quad \text{and} \quad S_3 = \{1 + iu : |u| \leq s\}, \]

and \( e^{it_+} = r^{-1}(1 + is) \), with \( r \) and \( s = \sqrt{r^2 - 1} \) as specified in (1.24) and (1.25):

\[
g_{\{j+K\}}(j) = \frac{1}{2\pi i} \left( \int_{S_1} + \int_{S_2} + \int_{S_3} \right) x^{-K-j-1} g^x(j) \, dx. \tag{2.5}
\]

In Lemmas 3.5 and 3.6, we show that \( |g^x(j)| \leq C \lambda^{-1} r^j \) for \( x \in S_1 \cup S_2 \) and \( j \geq C_2 \lambda m_1 \); hence, on \( S_1 \cup S_2 \), the integrand in (2.5) is of order \( \lambda^{-1} r^{-K} \), and hence the total contributions to (2.4) arising from \( S_1 \cup S_2 \) can be bounded by \( C/\{\lambda(r - 1)\} \), again for suitable constants \( C \). This leaves the contributions from \( S_3 \), which are bounded for all \( j \geq \frac{1}{2} \lambda m_1 + 1 \) in Section 4. The treatment of this part is very delicate, and is based on the formulae

\[
\frac{1}{2\pi i} \int_{S_3} x^{-K-j-1} g^x(j) \, dx
\]

\[
= -\frac{i}{2\pi} \int_{-s}^{-u} du \int_0^u dw \left( (1 + iu)^{-K-1} \int_0^1 dw \right) \frac{1}{\mu(1+iw) - \mu(1+iw)} \tag{2.6}
\]

\[
= \frac{1}{\pi} \int_0^s du \int_0^u dw \text{Im} \{ (1 + iu)^{-K-1} e^{\lambda(h(u) - h(w))} \},
\]

and

\[
\frac{1}{2\pi i} \int_{S_3} x^{-K-j-1} \{g^x(j+1) - g^x(j)\} \, dx
\]

\[
= \frac{1}{\pi} \int_0^s du \int_0^u dw \text{Re} \{ w(1 + iu)^{-K-1} e^{\lambda(h(u) - h(w))} \}, \tag{2.7}
\]

17
where

\[ h(u) = -\frac{j - 1}{\lambda} \log(1 + iu) + \mu(1 + iu). \]  

(2.8)

In both cases, the reduction to a double integral involving only one of the real and imaginary parts is possible because, for any power series \( g(v, z) \) with real coefficients, \( g(-iu, -iw) \) is just the complex conjugate \( \bar{g}(iu, iw) \) when \( u, w \in \mathbb{R} \). An indication of the care needed is that, without this reduction, the proof would not yield a bound of the right order for \( |g_A(j + 1) - g_A(j)| \); if the imaginary part were also present in (2.7), an extra factor of \( \log \lambda \) would appear.

The estimate of \( C_1 \) given in (1.27) accumulates the bounds given in Lemmas 4.1–4.4, 4.6 and 4.7 and Corollary 4.10, and substitutes them into (2.7); the result is then combined with those from Lemmas 3.1–3.4 and Corollary 3.7, and simplified. For the estimate of \( C_0 \) given in (1.26), we take the bounds from Lemma 4.5 and the remarks following Lemmas 4.1–4.4 into (2.6), combine the result with Lemmas 3.1–3.4 and Corollary 3.7, and simplify.

3. The detailed proof.

In this section, the estimates needed to prove Theorem TV with the constants given in (1.26)–(1.28) are established, with the exception of that part of (2.4) which arises from integrating along \( S_3 \) in (2.5). The first lemma covers all large values of \( j \).

**Lemma 3.1.** Let \( j \geq \frac{3}{2} \lambda m_1 \). Then

\[ |g_A(j)| \leq \frac{2e}{\lambda m_1}; \quad |g_A(j) - g_A(j + 1)| \leq \frac{4e}{(\lambda m_1)^2}. \]

**Proof.** By Representation 2 for the solution of Stein’s equation, we have

\[ |g_A(j)| \leq \int_0^1 t^{j-1} \exp(\lambda[\mu(1) - \mu(t)]) \, dt = \int_0^1 (1 - t)^{j-1} \exp(\lambda[\mu(1) - \mu(1 - t)]) \, dt \]

\[ \leq \int_0^1 \exp(-(j - 1)t + m_1 t \lambda) \, dt \leq e \int_0^\infty \exp(|m_1 \lambda - j| t) \, dt \]

\[ \leq e \int_0^\infty \exp(-m_1 \lambda t/2) \, dt = \frac{2e}{\lambda m_1}. \]

In a similar way, we also have

\[ |g_A(j) - g_A(j + 1)| \leq \int_0^1 t^{j-1}(1 - t) \exp(\lambda[\mu(1) - \mu(t)]) \, dt \]

\[ \leq \int_0^1 t \exp(-(j - 1)t + m_1 t \lambda) \, dt \leq e \int_0^\infty t \exp(-m_1 \lambda t/2) \, dt = \frac{4e}{(\lambda m_1)^2}. \]
completing the proof.

From now on, we can always assume that \( j < 3\lambda_1/2 \). The next two lemmas are concerned with the contributions to \( g_A \) and its first differences from \( A_1 \), as defined in (2.3). For the first, we look at values of \( j \) which are still quite large.

**Lemma 3.2.** Suppose that \( A \subset \{0, 1, \ldots, j - 1\} \), and that \( j \geq \lambda m_1 \). Then

\[
|g_A(j)| \leq e \sqrt{\frac{\pi}{\lambda m_1}}; \quad |g_A(j) - g_A(j + 1)| \leq \frac{2e}{\lambda m_1}.
\]

**Proof.** Since \( k - j < 0 \) for all \( k \in A \), it follows from Representation 2 that

\[
|g_A(j)| \leq \int_0^1 t^{j-1} e^{\lambda(1 - \mu(t))} dt = \int_0^1 \exp\{-\lambda w_{j-1}(t)\} dt \leq \frac{e}{\sqrt{\lambda m_1}} \int_0^\infty e^{-u^2/4} du,
\]

where \( w_{j-1} \) is as defined in (5.1) and the last inequality follows by Lemma 5.3. Similarly, we have

\[
|g_A(j) - g_A(j + 1)| \leq \int_0^1 t^{j-1}(1 - t)e^{\lambda(1 - \mu(t))} dt
\]

\[
= \int_0^1 t \exp\{-\lambda w_{j-1}(t)\} dt \leq \frac{e}{\lambda m_1} \int_0^\infty u e^{-u^2/4} du,
\]

proving the second part.

We now consider the contributions arising from \( A_1 \) for smaller values of \( j \).

**Lemma 3.3.** Suppose that \( A \subset \{0, 1, \ldots, j - 1\} \) and that \( \frac{1}{2} \lambda m_1 + 1 \leq j < \lambda m_1 \). Then

\[
|g_A(j)| \leq 2 \sqrt{\frac{\pi}{\lambda m_1}}; \quad |g_A(j) - g_A(j + 1)| \leq \frac{4c}{\lambda m_1},
\]

where

\[
c = 1 + \sqrt{\frac{\pi}{m_1}} \left\{ 1 + \frac{7.2 m_3}{m_3} \right\}.
\]

**Proof.** Let \( S \sim \text{CP}(\lambda, \mu) \). Then, as in Lemma 3.2, we have

\[
|g_A(j)| \leq \mathbb{P}[S \leq j - 1] \int_0^1 t^{j-1} e^{\lambda(1 - \mu(t))} dt.
\]

Applying the lower tail estimate for the compound Poisson distribution given in Lemma 5.4 and simply using \( C_{5.4} \leq 1 \), we then find that

\[
|g_A(j)| \leq e^{\lambda w_{j-1}(t_0)} \int_0^1 \exp\{-\lambda w_{j-1}(t)\} dt
\]

\[
\leq \int_0^1 \exp\{-\lambda m_1(t - t_0)^2/4\} dt \leq 2 \sqrt{\frac{\pi}{\lambda m_1}}.
\]
where we have also used Lemma 5.2 to complete the bound.

For the remaining part of the lemma, we start from

$$|g_A(j) - g_A(j + 1)| \leq \mathbb{P}[S \leq j - 1] \int_0^1 t^{j-1}(1-t)e^{\lambda \mu(1) - \mu(t)} \, dt,$$

and now use Lemma 5.4 to give

$$|g_A(j) - g_A(j + 1)| \leq C_{5.4}e^{\lambda w_{j-1}(t_0)} \int_0^1 t \exp{-\lambda w_{j-1}(t)} \, dt$$

$$\leq C_{5.4} \int_0^1 t \exp{-\lambda m_1(t-t_0)^2/4} \, dt \leq C_{5.4} \int_0^1 \{t_0 + |t-t_0|\} \exp{-\lambda m_1(t-t_0)^2/4} \, dt$$

$$\leq C_{5.4} \left\{ 2t_0 \sqrt{\frac{\pi}{\lambda m_1}} + \frac{4}{\lambda m_1} \right\} \leq \frac{4}{\lambda m_1} \left\{ 1 + \sqrt{\frac{7.2m_3}{m_1}} \right\},$$

where we use the definition of $C_{5.4}$ in Lemma 5.4 and the upper bound for $t_0$ from Lemma 5.2.

The next lemma bounds the contributions to $g_A$ and its first differences from $A_3$, as defined in (2.3).

**Lemma 3.4.** For $\frac{1}{2} \lambda m_1 + 1 \leq j \leq \frac{3}{2} \lambda m_1$, we have

$$\sum_{K \geq 4\lambda m_*} |g_{\{j+K\}}(j)| \leq \frac{8}{\lambda m_1} e^{-\lambda/9}.$$

**Proof.** By Representation 2,

$$\sum_{K \geq X} |g_{\{j+K\}}(j)| \leq 2 \int_0^1 t^{j-1}e^{\lambda \mu(1) - \mu(t)} \, CP(\lambda, \mu) \{[X, \infty)\} \, dt,$$

and

$$\int_0^1 t^{j-1}e^{\lambda \mu(1) - \mu(t)} \, dt \leq \left\{ e^{\lambda} \int_0^{1/m_1} t^{j-1} \, dt + \int_0^{1/m_1} e^{-(j-1)u + \lambda m_1 u} \, du \right\} \leq \frac{4e^{\lambda/2}}{\lambda m_1}.$$

Now, by an exponential Chebyshev estimate, for any $0 < \tau < 1/2$,

$$CP(\lambda, \mu) \{[X, \infty)\} \leq (1 + \tau/m_*)^{-X} \exp{\lambda \mu(1 + \tau/m_*) - 1)},$$

20
and

\[
\mu(1 + \tau/m_*) \leq 1 + \sum_{n \geq 1} \frac{m_1[n]}{n!} (\tau/m_*)^n \leq \frac{1}{1 - \tau};
\]

hence it follows that

\[
e^{\lambda/2} \text{CP}(\lambda, \mu)\{[X, \infty)\} \leq \exp\left\{-\frac{X}{m_*} \left(1 - \frac{\tau}{2m_*}\right) + \lambda \left(\frac{1 + \tau}{2(1 - \tau)}\right)\right\}.
\]

Taking \(\tau = 1/3\) and \(X = 4\lambda m_*\), and remembering that \(m_* \geq 1\), we conclude that

\[
\sum_{\kappa \geq 4\lambda m_*} |g_{\kappa}(j)| \leq \frac{8e^{\lambda/2}}{\lambda m_1} \text{CP}(\lambda, \mu)\{[4\lambda m_*, \infty)\} \leq \frac{8}{\lambda m_1} \exp\{-4/3\lambda(1 - 1/6) + \lambda\},
\]

as required.

We now turn to the contributions from \(A_2\), as detailed in (2.4); we use Expression (2.5) for the \(g_{\{j+\kappa\}}(j)\), and start by looking at what comes from the integral along \(S_1\).

**Lemma 3.5.** Let \(y \in S_1\) and \(j \geq C_2\lambda m_1 + 1\). Then

\[
|g^y(j)| \leq \frac{6er^j}{\lambda m_1 \rho^*(\zeta)},
\]

where \(r\) is as defined in (1.24) and (1.25).

**Proof.** Let \(y = rx\), where \(x = e^{i\theta}\) and \(\zeta \leq |\theta| \leq \pi\); this in accordance with the definition of \(S_1\) preceding (2.5). From (2.1), choosing a contour of integration which starts down the line joining \(y\) to the origin, we can write

\[
g^y(j) = \int_{rx}^x z^{j-1} \exp\{\lambda(\mu(rx) - \mu(z))\} dz + \exp\{\lambda(\mu(rx) - \mu(x))\} g^x(j). \tag{3.1}
\]

We now estimate the elements of this expression.

First, since \(\Re e(x^k) = \cos k\theta \leq \frac{1}{2}(1 + \cos k\theta)\), we find, for any \(0 \leq t \leq 1\), that

\[
\Re \{\mu(rx) - \mu(rt)\} = \Re \left\{\sum_{k \geq 1} \mu_k x^k r^k (1 - t^k)\right\} \leq \frac{1}{2} \sum_{k \geq 1} \mu_k r^k (1 - t^k)(1 + \cos k\theta) \\
\leq \frac{1}{2}(1 - t) \sum_{k \geq 1} \mu_k kr^k (1 + \cos k\theta) = (1 - t) m_1^{(r)} (1 - \rho_2^{(r)}(\theta)/2);
\]

from this, since \(|\theta| \geq \zeta\) and by (1.24), it follows that

\[
\Re \{\mu(rx) - \mu(rt)\} \leq (1 - t) m_1 (1 - \frac{1}{2} \rho^*(\zeta)). \tag{3.2}
\]
Also, in the same range of $t$,

$$t^j = \exp\{j \log(1 - (1 - t))\} \leq e^{-j(1-t)}. \quad (3.3)$$

Hence, taking $t = 1/r$ in (3.2) and (3.3), it follows that

$$r^{-j} \exp\{\lambda(\mu(rx) - \mu(x))\} \leq \exp\{-j(1-r^{-1}) + \lambda m_1(1-r^{-1})(1 - \frac{1}{2}\rho^*(\zeta))\}
\leq \exp\{-\frac{1}{6}\lambda m_1\rho^*(\zeta)(1 - r^{-1})\}, \quad (3.4)$$

for any $j \geq C_2\lambda m_1 + 1$; and similarly

$$\left|\int_{rx}^{x} z^{j-1} \exp\{\lambda(\mu(rx) - \mu(z))\} \, dz\right| \leq r^j \int_{1/r}^{1} t^{j-1} |\exp\{\lambda(\mu(rx) - \mu(rt))\}| \, dt
\leq r^j \int_{1/r}^{1} e^{-\frac{1}{6}\lambda m_1\rho^*(\zeta)(1-t)} \, dt \leq \frac{4r^j}{\lambda m_1\rho^*(\zeta)} \left\{1 - e^{-\frac{1}{6}\lambda m_1\rho^*(\zeta)(1-r^{-1})}\right\}. \quad (3.5)$$

Putting (3.4) and (3.5) into (3.1), and using the estimate $|g^z(j)| \leq 6e/\{\lambda m_1\rho^*(\zeta)}$ from Barbour and Utev (1996), Lemma 4.1, the lemma is proved.

The next lemma is concerned with what happens on $S_2$.

**Lemma 3.6.** For $y \in S_2$ and $j \geq \frac{1}{2}\lambda m_1 + 1$, we have

$$|g^y(j)| \leq \frac{968r^j}{161\lambda s_\gamma},$$

where $r$ and $s$ are as defined in (1.24) and (1.25).

**Proof.** As before, set $y = rx$, where $x = e^{i\vartheta}$ and now $t_+ \leq |\vartheta| \leq \zeta$, with $x_+ = e^{it_+} = r^{-1}(1 + is)$ as defined before (2.5). Without loss of generality, assume that $\theta > 0$. Once more using (2.1), we take a contour of integration along an arc of $|z| = r$ and then down $\Re e(z) = 1$, to give

$$g^y(j) = \int_{rx}^{rx_+} z^{j-1} \exp(\lambda[\mu(rx) - \mu(z)]) \, dz
\quad + \exp(\lambda[\mu(rx) - \mu(rx_+)]) \int_{rx_+}^{1} z^{j-1} \exp(\lambda[\mu(rx) - \mu(z)]) \, dz = I_2 + J_2Q_2, \quad (3.6)$$

say. We now bound each of $I_2$, $J_2$ and $Q_2$. 

22
First, writing \( \psi(t) = \Re \{ \mu(r e^{it}) \} = \sum_{k \geq 1} \mu k r^k \cos kt \), we see that \( |\psi(t) + m_2^{(r)}| \leq \frac{1}{6} t^3 m_4^{(r)} \), from which it follows that \( \psi'(t) \leq -t m_2^{(r)}/2 < 0 \) for all \( 0 \leq t \leq \theta \), since \( \theta \leq \zeta \leq \{2m_2/m_4\}^{1/2} \leq \{3m_2^{(r)}/m_4^{(r)}\}^{1/2} \), because of (1.24). Thus \( |J_2| \leq 1 \), and

\[
|I_2| = r^j \left| \int_x^{x^+} z^{j-1} \exp \{\lambda (\mu (rx) - \mu (rz))\} \, dz \right| \\
= r^j \int_{t_+}^\theta \exp \{\lambda (\psi(\theta) - \psi(u))\} \, du \leq r^j \int_{t_+}^\theta \exp \{-\lambda t_+ m_2^{(r)} (\theta - u)/2\} \, du \quad (3.7)
\]

this last since \( s = \tan t_+ \leq 1/9 \), by (1.25).

It remains to bound \( Q_2 \), which can be written as

\[
Q_2 = -i \int_0^s (1 + iv)^{j-1} \exp \{\lambda (1 + is) - \mu (1 + iv))\} \, dv,
\]

from which it follows that

\[
|Q_2| = r^{j-1} \int_0^s \exp \{\lambda \Re \{h_{j-1}(s) - h_{j-1}(v))\} \, dv,
\]

where \( h_t(u) \) is as in Lemma 5.1. Now, for \( j \geq \frac{1}{2} \lambda m_1 + 1 \) and for \( s \) satisfying (1.25), Lemma 5.1 implies that \( \Re \{h_{j-1}(s) - h_{j-1}(v)\} \leq -\frac{1}{4} \gamma (s^2 - v^2) \leq -\frac{1}{4} s(s - v) \gamma \), and hence

\[
|Q_2| \leq 4r^{j-1}/(\lambda s \gamma),
\]

completing the proof of the lemma.

Noting that \( \zeta \) was chosen to be smaller than \( \sqrt{2m_2/m_4} \) and that \( s = \sqrt{r^2 - 1} \geq \sqrt{2(r - 1)} \), we can summarize the total contribution to the estimates (2.4) of \( |g_{A_2}(j)| \) and \( |g_{A_2}(j) - g_{A_2}(j + 1)| \) arising from the integral over \( S_1 \cup S_2 \) as follows.

**Corollary 3.7.** For \( C_2 \lambda m_1 + 1 \leq j \leq \frac{3}{2 \lambda m_1} \),

\[
\sum_{K=0}^{[\lambda m_1]} \left| \frac{1}{2\pi i} \int_{S_1 \cup S_2} x^{-K-j-1} g^x(j) \, dx \right| \leq \frac{r}{\lambda (r - 1)} \left\{ \frac{6e}{m_1 \rho^*(\zeta)} + \frac{2}{\gamma \sqrt{m_2/m_4 (r - 1)}} \right\} := \varepsilon_{A_2};
\]

\[
|g_{A_2}(j)| \leq \varepsilon_{A_2}; \quad |g_{A_2}(j) - g_{A_2}(j + 1)| \leq 2\varepsilon_{A_2}.
\]

4. Integration over \( S_3 \).
In this section, we bound the contributions to (2.4) arising from the integral over $S_3$ in (2.5). In view of (2.6) and (2.7), it is enough to consider

$$I_K = \int_0^s du \int_0^u dw \, \lambda^{h(u)-h(w)}(1 + iu)^{-K} = \int_0^s dz \int_0^{s-z} dv \, z^e^{\lambda h(v+z)-h(z)}(1 + i(v+z))^{-K}$$

and

$$I'_K = \int_0^s du \int_0^u dw \, \lambda^{h(u)-h(w)}(1 + iu)^{-K} = \int_0^s dz \int_0^{s-z} dv \, e^{\lambda h(v+z)-h(z)}(1 + i(v+z))^{-K}$$

in $2 \leq K \leq [4\lambda m_*] + 2$ and $\frac{1}{2}\lambda m_1 \leq j \leq \frac{3}{2}\lambda m_1$, where $h = h_j$ as in Lemma 5.1. We need to show that $\sum_{K=2}^{[4\lambda m_*] + 2} |R \epsilon \{I_K\}| = O(\lambda^{-1})$ and that $\sum_{K=2}^{[4\lambda m_*] + 2} |R m \{I'_K\}| = O(\lambda^{-1/2})$, and to exhibit suitable constants; we concentrate on the $I_K$ sum, since that for $I'_K$ is substantially easier.

**Lemma 4.1.** Define

$$I_{K1} = \int_0^s dz \int_0^{s-z} dv \, z^e \exp \{\lambda v\{h'(0) + zh''(0) + \frac{1}{2}z^2 h^{(3)}(0) + \frac{1}{2}v^2 h''(0)\} + \frac{1}{2}v^{-2} h''(0)\}(1 + i(v+z))^{-K},$$

where $h = h_j$ from Lemma 5.1. Then

$$\sum_{K=2}^{[4\lambda m_*] + 2} |I_K - I_{K1}| \leq \frac{d_*}{\lambda^2 \gamma^2} \left\{ 27 + \frac{58m_*}{\gamma} \right\}.$$  

**Proof.** The argument uses Taylor’s expansion to replace the function $h$ in the exponent by a short multinomial expression. The derivatives of $h$ are given by

$$h^{(l)}(u) = i^l \mu^{(l)}(1 + iu) + \frac{(-i)^l j(l-1)!}{\lambda(1 + iu)^l}, \quad l \geq 1,$$

from which it follows that

$$h'(0) = i(m_1 - \lambda^{-1}j) := i\tau(j); \quad h''(0) = -(m_2 + \lambda^{-1}j) := -c(j)$$

$$h^{(3)}(0) = -im_3 + 2\iota\lambda^{-1}j := -2id(j).$$

(4.2)

Also, from Lemma 5.1(3), for $0 \leq x \leq s$,

$$|h^{(3)}(x)| \leq \frac{2}{3} m_3^2 + 2\lambda^{-1} j \quad \text{and} \quad |h^{(4)}(x)| \leq \frac{2}{3} m_3^2 + 6\lambda^{-1} j.$$  

(4.3)

A first application of Taylor’s expansion gives

$$|h(z + v) - h(z) - \{vh'(z) + \frac{1}{2}v^2 h''(z)\}| \leq \frac{1}{6} v^3 \sup_{0 \leq w \leq z + v} |h^{(3)}(w)|,$$  

(4.4)

A first application of Taylor’s expansion gives

$$|h(z + v) - h(z) - \{vh'(z) + \frac{1}{2}v^2 h''(z)\}| \leq \frac{1}{6} v^3 \sup_{0 \leq w \leq z + v} |h^{(3)}(w)|,$$

(4.4)
and then, simplifying this further,
\[
|h(z + v) - h(z) - v\{h'(0) + zh''(0) + \frac{1}{2}v^2h^{(3)}(0)\} - \frac{1}{6}v^2h''(0)|
\leq \frac{1}{6}v^2\sup_{0 \leq w \leq z + v} |h^{(3)}(w)| + \frac{1}{6}v^2z^3\sup_{0 \leq w \leq z + v} |h^{(4)}(w)|. \tag{4.5}
\]

Now, from (4.3), if \(0 \leq v, w \leq v + z \leq s\), it follows that
\[
|zh^{(3)}(w)| + |zh^{(3)}(w)/3| \leq \frac{4s}{3}\left\{\frac{9}{8}m_*^3 + 2\lambda^{-1}j\right\} \leq \frac{c(j)}{2}
\tag{4.6}
\]
whenever \(j \geq \lambda m_1/2\), because of (1.25); similarly,
\[
\frac{1}{6}|z^2h^{(4)}(w)| \leq s^2\left\{\frac{3}{16}m_*^4 + \lambda^{-1}j\right\} \leq \frac{c(j)}{2}
\tag{4.7}
\]
in the same range of \(j\). Combining (4.5) with (4.6) and (4.7), it follows that
\[
|e^{\lambda h(z + v) - h(z)}| \leq e^{-\lambda c(j) v(2z + v)/4},
\]
and hence that
\[
|I_K - I_{K1}| \leq \int_0^s dz \int_0^{s-z} dv \frac{\lambda z}{6(1 + z^2)K^{1/2}} \left\{v^2(v + 3z)\left(\frac{9}{8}m_*^3 + \frac{2j}{\lambda}\right) + vz^3\left(\frac{9}{8}m_*^4 + \frac{6j}{\lambda}\right)\right\} e^{-\lambda c(j)v(2z + v)/4}. \tag{4.8}
\]

To evaluate the integrals on the right hand side of (4.8), we repeatedly need the bounds
\[
\int_0^{s-z} v^\ell e^{-\lambda c v(2z + v)/4} dv \leq \min\{\Gamma\left(\frac{2}{\lambda cz}\right)^{\ell + 1}, \kappa_\ell\left(\frac{2}{\lambda c}\right)^{\ell+1/2}\}, \tag{4.9}
\]
valid for \(c > 0\), where
\[
\kappa_\ell = \int_0^\infty v^\ell e^{-v^2/2} dv, \tag{4.10}
\]
giving
\[
\kappa_0 = \sqrt{\pi}/2, \kappa_1 = 1, \kappa_2 = \sqrt{\pi}/2, \kappa_3 = 2, \kappa_4 = 3\sqrt{\pi}/2, \kappa_5 = 8 : \tag{4.11}
\]
these estimates are derived by replacing \((2z + v)\) in the exponent by either \(2z\) or \(v\). Applying
them, and because also $c(j) \geq \gamma$, we find that

$$U_{K1} = \int_0^s dz (1 + z^2)^{-K/2} \int_0^{s - z} dv \lambda z^3 e^{-\lambda c(j) v(2z + v)/4}$$

$$\leq \int_0^{\sqrt{2/\lambda \gamma}} 2\lambda z \left(\frac{2}{\lambda \gamma}\right)^2 dz + \int_0^{\infty} 6\lambda z^{-3} (1 + z^2)^{-K/2} \left(\frac{2}{\lambda \gamma}\right)^4 dz$$

$$= \frac{8}{\lambda^2 \gamma^3} + \int_0^{\infty} 6\lambda z^{-3} (1 + z^2)^{-K/2} \left(\frac{2}{\lambda \gamma}\right)^4 dz;$$

$$U_{K2} = \int_0^s dz (1 + z^2)^{-K/2} \int_0^{s - z} dv \lambda z^2 v^2 e^{-\lambda c(j) v(2z + v)/4}$$

$$\leq \frac{\lambda}{3} \sqrt{\frac{\pi}{2}} \left(\frac{2}{\lambda \gamma}\right)^3 + \left(\frac{2}{\lambda \gamma}\right)^3 \int_0^s \frac{2\lambda dz}{\sqrt{2/\lambda \gamma} z (1 + z^2)^{K/2}};$$

$$U_{K3} = \int_0^s dz (1 + z^2)^{-K/2} \int_0^{s - z} dv \lambda z^4 v e^{-\lambda c(j) v(2z + v)/4}$$

$$\leq \left(\frac{2}{\lambda \gamma}\right)^2 \int_0^s \frac{\lambda z^2 dz}{(1 + z^2)^{K/2}}.$$

Now, since $\sum_{K \geq 2} (1 + z^2)^{-K/2} = z^{-2} \{1 + (1 + z^2)^{-1/2}\} \leq 2z^{-2}$, it follows from (4.12) that

$$\sum_{K = 2}^{[4\lambda m_*]^2 + 2} U_{K1} \leq \frac{8(4m_* + \lambda - 1)}{\lambda \gamma^3} + \frac{12}{\lambda \gamma^2};$$

$$\sum_{K = 2}^{[4\lambda m_*]^2 + 2} U_{K2} \leq \frac{8(4m_* + \lambda - 1)}{3\lambda \gamma^3} \sqrt{\frac{\pi}{2}} + \frac{8}{\lambda \gamma^2};$$

$$\sum_{K = 2}^{[4\lambda m_*]^2 + 2} U_{K3} \leq \frac{8s}{\lambda \gamma^2}.$$

Now add (4.8) over $K$ and substitute these estimates into the result, simplifying the expression further because $\frac{1}{2} \lambda m_1 \leq j \leq \frac{1}{2} \lambda m_1$, because $s$ satisfies (1.25), and because $\lambda m_* \geq 2$.

**Remark.** The difference between $I^*_K$ and $I^*_{K1}$, defined as $I_{K1}$ but without a factor of $z$ in the integrand, is estimated using the same techniques, giving

$$\sum_{K = 2}^{[4\lambda m_*]^2 + 2} |I^*_K - I^*_{K1}| \leq \frac{d_*}{\gamma \sqrt{\lambda \gamma}} \left\{1 + 160 \left(\frac{m_*}{\gamma}\right)\right\}.$$

The next lemma replaces the factor $(1 + i(v + z))^{-K}$ in the integrand in Lemma 4.1 by the exponential of a multinomial expression in $z$ and $v$. 

26
Lemma 4.2. Define
\[ I_{K2} = \int_0^\delta dz \int_0^{s-z} dv \, z e^{\lambda [i(v+j) - v(2z+v)c(j)/2 - iv(z^2/2) + i(c+j)w)]} - K[i(v+z) + (v+z)^2/2 - iz^3/3]. \]

Then
\[ \sum_{K=2}^{[4\lambda m_*] + 2} |I_{K1} - I_{K2}| \leq \left( \frac{1}{\lambda \gamma} \right) \left\{ 16 + 45 \left( \frac{m_*}{\gamma} \right) \right\}. \]

Proof. By Taylor’s expansion,
\[ |\log(1 + i(v + z)) - \{i(v + z) + (v + z)^2/2 - iz^3/3\}| \leq (v + z)^4/4 + (vz^2 + v^2z + v^3/3). \]
Hence, since \(|e^a - 1| \leq |a| \exp\{Re(a)\}\) and \((v + z)^4 \leq (v^4 + z^4)\), giving
\[ |I_{K1} - I_{K2}| \leq \int_0^\delta dz \int_0^{s-z} dv \, z e^{-\lambda v(2z+v)c(j)/2 - 161 Kz^2/324 K[(v + z)^4/4 + (vz^2 + v^2z + v^3/3)]}. \]

We now bound the right hand side using the expressions in (4.9) and (4.11), with \(2\gamma\) for \(c\), because \(j \geq \frac{1}{2}\lambda m_1\), and using the simple inequality \((v + z)^4 \leq 8(v^4 + z^4)\), giving
\[ |I_{K1} - I_{K2}| \leq 6\delta \sqrt{\frac{\pi}{2}} \left( \frac{1}{\lambda \gamma} \right)^{5/2} + \frac{6\delta^{5/2} \sqrt{\pi/2}}{\lambda \gamma K^{3/2}} + \frac{\delta}{\lambda^2 \gamma^2} + \frac{\pi \delta^{3/2}}{2(\lambda \gamma)^{3/2} K^{1/2}} + \frac{2\delta}{3\lambda^2 \gamma^2}, \]
where \(\delta = 162/161\). Now use the inequalities
\[ \sum_{K=2}^{[4\lambda m_*] + 2} K^{-1/2} \leq 4\sqrt{\lambda m_*} \quad \text{and} \quad \sum_{K=2} K^{-3/2} \leq 5/3, \]
from Lemma 5.6, and collect terms.

Remark. Defining \(I_{K2}^*\) as \(I_{K2}\), but without a \(z\)-factor in the integrand, we obtain
\[ \sum_{K=2}^{[4\lambda m_*] + 2} |I_{K1} - I_{K2}^*| \leq \frac{1}{\sqrt{\lambda \gamma}} \left\{ 20 + 6 \left( \frac{m_*}{\gamma} \right)^{3/2} \left( 1 + \frac{11}{\sqrt{\lambda \gamma}} \right) \right\}. \]

Now, changing variables to \(w = \lambda v/\sqrt{K}\) and \(t = z\sqrt{K}\), we have
\[ I_{K2} = \int_0^{s\lambda/\sqrt{K}} dw \, \lambda^{-1} \sqrt{K} e^{i\pi(j+K)w} \sqrt{K} e^{-(c(j+K)w^2 K/2\lambda)} dt \int_0^{s\sqrt{K}} dt \, K^{-1} e^{-i\psi_{K,w}(t) e^{-c(j+K)wt^2/21}} \left\{ \lambda^{-1} w \sqrt{K} + t/\sqrt{K} \leq s \right\} \] (4.13)
where
\[ \psi_{K,w}(t) = t \frac{t^3}{3K} + \frac{d(j)wt^2}{K}. \]

The first step in modifying \(I_{K2}\) is to take the integrals out to infinity.
Lemma 4.3. Define

\[ I_{K3} = \frac{1}{\lambda \sqrt{K}} \int_0^\infty \frac{d\sigma}{j+K} \int_0^\infty \frac{d\sigma}{j+K} \int_0^\infty \int_0^\infty \int_0^\infty \frac{dt}{te^{-\left(c(j+K)w+2t^2\right)/2}} e^{-i\sigma_1 \psi \cdot (t)} \, dt, \]

where \( \hat{\psi}_{K,w}(t) = t - R_{1K}(t) + K \frac{1}{\gamma} \hat{d}(j) \)

\[ R_{1K}(t) = \begin{cases} \frac{t^3}{3K} & \text{if } t \leq \frac{1}{2} \sqrt{K}; \\ \frac{1}{4} (t - \sqrt{K}) & \text{if } t \geq \frac{1}{2} \sqrt{K}. \end{cases} \]

Then

\[ \sum_{K=2}^{[4\lambda m_*]+2} |I_{K2} - I_{K3}| \leq \frac{4}{\lambda \gamma \{2(r-1)\}^{1/2}} \left\{ \log \left( \frac{1}{r-1} \right) + 2 + 9 \left( \frac{m_*}{\gamma} \right) \right\}. \]

Proof. We write

\[ |I_{K2} - I_{K3}| \leq \frac{1}{\lambda \sqrt{K}} \int_0^{\lambda s/2 \sqrt{K}} \int_0^\infty \int_0^\infty \int_0^\infty \frac{dt}{te^{-\left(c(j+K)w+2t^2\right)/2}} \]

\[ + \frac{1}{\lambda \sqrt{K}} \int_0^{\lambda s/2 \sqrt{K}} \int_0^\infty \int_0^\infty \int_0^\infty \frac{dt}{te^{-\left(c(j+K)w+2t^2\right)/2}} \]

\[ \leq \frac{1}{\lambda \sqrt{K}} \left\{ \int_0^\infty e^{-\gamma_1 w s \sqrt{K}/2} \, dw + e^{-\gamma s^2/8} \int_{\lambda s/2 \sqrt{K}}^\infty \int_0^\infty \int_0^\infty \frac{dt}{te^{-\left(c(j+K)w+2t^2\right)/2}} \right\} \]

\[ \leq \frac{2}{\lambda \gamma \sqrt{K}} \frac{2e^{-K s^2/8}}{s \sqrt{K}} + \frac{8}{\lambda^2 \gamma^2 s}, \]

using also that \( c(j+K) \geq \gamma \); note that the modification of \( \psi \) to \( \hat{\psi} \) only changes the function on \( t \geq \frac{1}{2} \sqrt{K} \), which never occurs in the original range of integration, because \( s \) satisfies (1.25). Now

\[ \sum_{K=2}^{[4\lambda m_*]+2} K^{-1} e^{-aK} \leq \int_1^\infty x^{-1} e^{-ax} \, dx = \int_a^\infty y^{-1} e^{-y} \, dy \leq \int_1^\infty y^{-1} \, dy + \int_1^\infty e^{-y} \, dy, \]

and hence

\[ \sum_{K=2}^{[4\lambda m_*]+2} |I_{K2} - I_{K3}| \leq \frac{4}{\lambda \gamma \{2(r-1)\}^{1/2}} \left\{ \log \left( 8s^{-2} \right) + e^{-1} \right\} + \frac{36 m_*}{\lambda \gamma^2 s}, \]

and the lemma follows because \( s \geq \sqrt{2(r-1)}. \)

Remark. Once again, after defining \( I_{K3}^* \) as for \( I_{K3} \) but without a factor of \( t \) in the integrand, similar arguments can be used to show that

\[ \sum_{K=2}^{[4\lambda m_*]+2} |I_{K2}^* - I_{K3}^*| \leq \frac{1}{\lambda \gamma (r-1)} \left\{ 8 + 31 \sqrt{m_*} (r-1)/\gamma \right\}. \]
Now if $|d(j)|ut \leq K/8$, we have $1/2 \leq d\psi_{K,w}(t)/dt \leq 5/4$, and the factor $e^{-i\sqrt{K}\psi_{K,w}(t)}$ in the integrand of $I_{K3}$ performs fast oscillations. We now show, by making a further modification to $\psi$, that we can preserve this behaviour for all values of the arguments.

**Lemma 4.4.** Define

$$I_{K4} = \int_0^\infty dw \frac{1}{\lambda \sqrt{K}} e^{i\tau(j+K)w} \sqrt{K - c(j+K)w^2} K/2 \lambda \int_0^\infty dt t e^{-i\sqrt{K}\tilde{\psi}_{K,w}(t)} e^{-(c(j+K)w t+t^2)/2},$$

where $\tilde{\psi}_{K,w}(t) = t - R_1 K(t) + R_2 K(t)$, $R_1 K$ is as for Lemma 4.3 and

$$R_{2K} = \begin{cases} K^{-1} d(j) wt^2, & \text{if } K^{-1} |d(j)| ut \leq 1/8; \\ \text{sign}(d(j)) \left\{ \frac{4}{3} - \frac{K}{64 |d(j)| w} \right\}, & \text{if } K^{-1} |d(j)| ut > 1/8. \end{cases}$$

Then

$$\sum_{K=2}^{[4\lambda m_s]+2} |I_{K3} - I_{K4}| \leq \frac{32 |d(j)|}{\lambda \gamma^2} \sqrt{\pi} \leq \frac{60 d_s}{\lambda \gamma^2}.$$

**Proof.** It is immediate that

$$|I_{K3} - I_{K4}| \leq 2 \frac{1}{\lambda \sqrt{K}} \int_0^\infty dt \int_0^\infty dw t e^{-(c(j+K)w t+t^2)/2} 1_{\{|d(j)| w t \geq K/8\}}$$

$$\leq 2 \frac{1}{\lambda \sqrt{K}} \int_0^\infty dt t e^{-K/w/16 |d(j)|} \frac{2}{\gamma^2} e^{-t^2/2} \leq \frac{4}{\lambda \gamma \sqrt{K}} e^{-K/16 |d(j)|} \sqrt{\pi},$$

and adding over $K$ completes the proof; note that $|d(j)| \leq d_s$, from its definition in (4.2).

**Remark.** Defining $I_{K4}^*$ as for $I_{K4}$, but without a $t$-factor, and retaining the factor $e^{-c(j+K)w^2 K/2 \lambda}$ when evaluating the bound, it follows that

$$\sum_{K=2}^{[4\lambda m_s]+2} |I_{K3}^* - I_{K4}^*| \leq \frac{16}{\sqrt{\lambda \gamma}}.$$

At this point, we recall that only $\Re e(I_{K4})$ and $\Im m(I_{K4}^*)$ need to be estimated. The former task proves to be much more difficult than the latter. We split $\Re e(I_{K4})$ into two parts,

$$I_{K4}^C = \frac{1}{\lambda \sqrt{K}} \int_0^\infty dw \cos(\tau(j+K)w \sqrt{K}) e^{-c(j+K)w^2 K/2 \lambda}$$

$$\int_0^\infty dt t \cos(\sqrt{K}\tilde{\psi}_{K,w}(t)) e^{-(c(j+K)w t+t^2)/2}.$$
and

\[
I_{K_4}^S = \frac{1}{\lambda \sqrt{K}} \int_0^\infty dw \sum_{i=1}^{n} (\tau(j + K)w\sqrt{K}) e^{-c(j+K)w^2K/2\lambda} \int_0^\infty dt \sum_{i=1}^{n} (\sqrt{K}\tilde{\psi}_{K,w}(t)) e^{-(c(j+K)wt+t^2)/2},
\]

and define the corresponding expressions \( I_{K_4}^C \) and \( I_{K_4}^S \), which together make up \( \text{Im}(I_{K_4}^s) \):

\[
I_{K_4}^C = \frac{1}{\lambda \sqrt{K}} \int_0^\infty dw \cos(\tau(j + K)w\sqrt{K}) e^{-c(j+K)w^2K/2\lambda} \int_0^\infty dt \sum_{i=1}^{n} (\sqrt{K}\tilde{\psi}_{K,w}(t)) e^{-(c(j+K)wt+t^2)/2}
\]

and

\[
I_{K_4}^S = \frac{1}{\lambda \sqrt{K}} \int_0^\infty dw \sum_{i=1}^{n} (\tau(j + K)w\sqrt{K}) e^{-c(j+K)w^2K/2\lambda} \int_0^\infty dt \cos(\sqrt{K}\tilde{\psi}_{K,w}(t)) e^{-(c(j+K)wt+t^2)/2}.
\]

**Lemma 4.5.** With \( I_{K_4}^C \) and \( I_{K_4}^S \) defined as above,

\[
\sum_{K=2}^{[4\lambda m+2]} \{|I_{K_4}^C| + |I_{K_4}^S|\} \leq \sqrt{\frac{2\pi}{\lambda\gamma}} \left\{ 28 + \frac{32d_\ast}{\gamma} \right\} \leq 32 \sqrt{\frac{2\pi}{\lambda\gamma}} \left( 1 + \frac{d_\ast}{\gamma} \right).
\]

**Proof.** Taking \( I_{K_4}^S \) as example, integrate the \( t \)-integral by parts to obtain, using \( \psi(t) \) as shorthand for \( \tilde{\psi}_{K,w}(t) \),

\[
\int_0^\infty dt \sum_{i=1}^{n} (\sqrt{K}\tilde{\psi}_{K,w}(t)) e^{-(c(j+K)wt+t^2)/2} = \left[ \frac{1 - \cos(\sqrt{K}\psi'(t))}{\sqrt{K}\psi'(t)} e^{-(c(j+K)wt+t^2)/2} \right]_0^\infty - \int_0^\infty dt \frac{1 - \cos(\sqrt{K}\psi'(t))}{\sqrt{K}\psi'(t)} \left\{ - \frac{c(j+K)w}{2} - t - \frac{\psi''(t)}{[\psi'(t)]^2} \right\} e^{-(c(j+K)wt+t^2)/2},
\]

with \( 1/2 \leq \psi'(t) \leq 5/4 \) for all \( t \) and with

\[
\psi''(t) = -\frac{2t}{K}1_{\{t \leq 1/2\}} + \frac{2d(j)w}{K}1_{\{t \leq K/8d(j)w\}}.
\]

The first term in (4.15) is zero. Using simple bounds for the elements of the integrand, derived from the above properties of \( \psi \) and from inequalities such as \(|1 - \cos x| \leq 2\), we
then obtain
\[
|I_{K4}^S| \leq \frac{4}{\lambda K} \int_0^\infty dw \ e^{-c(j+K)w^2K/2\lambda} \int_0^\infty dt \left\{ \frac{c(j + K)w}{2} + t + \frac{8}{K}(t + |d(j)|w) \right\} e^{-(c(j + K)wt + t^2)/2}.
\]
Since \( \int_0^\infty (cw/2)e^{-ct/2} \, dt = \int_0^\infty te^{-t^2/2} \, dt = 1 \), the estimate
\[
|I_{K4}^S| \leq \frac{4}{\lambda K} \sqrt{\frac{\pi \lambda}{2K\gamma}} \left\{ 2 + \frac{8}{K} \left( 1 + \frac{2|d(j)|}{\gamma} \right) \right\}
\]
is now immediate, giving
\[
\sum_{K=2}^{|4A_{m_0}|+2} |I_{K4}^S| \leq 2 \sqrt{\frac{2\pi}{\lambda\gamma}} \left\{ 4 + \frac{16}{3} \left( 1 + \frac{2|d(j)|}{\gamma} \right) \right\}.
\]
The contribution from \( I_{K4}^C \) can be bounded by half this amount, because \( \cos(\sqrt{K}\psi(t)) \) is integrated to \( \sum_{i=1}^n (\sqrt{\sqrt{K}} \psi(t)) \), and \( |\sum_{i=1}^n x| \leq 1 \); the lemma now follows.

The argument for \( R(e(I_{K4})) \) is much more delicate. Integrating the \( t \)-integral in \( I_{K4}^C \) by parts, we obtain
\[
\int_0^\infty dt \ t \cos(\sqrt{K}\psi_K_w(t))e^{-(c(j+K)wt+t^2)/2} = -\int_0^\infty dt \left( \frac{\sum_{i=1}^n (\sqrt{K}\psi(t))}{\sqrt{K}\psi'(t)} \right) \left\{ 1 - c(j + K)wt/2 - t^2 - \frac{t\psi''(t)}{[\psi'(t)]^2} \right\} e^{-(c(j+K)wt+t^2)/2},
\]
with \( \psi' \) and \( \psi'' \) as in (4.16); a similar expression can be derived for \( I_{K4}^S \). The next lemma simplifies them further.

**Lemma 4.6.** Define
\[
I_{K5}^C = -\left( \lambda K \right)^{-1} \int_0^\infty dw \cos(\tau(j + K)w\sqrt{K})e^{-c(j+K)w^2K/2\lambda}
\]
\[
\int_0^\infty dt \sum_{i=1}^n (t\sqrt{K}) \left\{ 1 - c(j + K)wt/2 - t^2 \right\} e^{-(c(j+K)wt+t^2)/2};
\]
\[
I_{K5}^S = -\left( \lambda K \right)^{-1} \int_0^\infty dw \sum_{i=1}^n (\tau(j + K)w\sqrt{K})e^{-c(j+K)w^2K/2\lambda}
\]
\[
\int_0^\infty dt \left( 1 - \cos(t\sqrt{K}) \right) \left\{ 1 - c(j + K)wt/2 - t^2 \right\} e^{-(c(j+K)wt+t^2)/2}.
\]

31
Then
\[\sum_{K=2}^{4\lambda m_1+2} |I_{K4}^C - I_{K5}^C| \leq \left( \frac{1}{\lambda \gamma} \right) \left\{ 46 + 368 \frac{d_*}{\gamma} \right\}\] and \[\sum_{K=2}^{4\lambda m_1+2} |I_{K4}^S - I_{K5}^S| \leq \left( \frac{1}{\lambda \gamma} \right) \left\{ 78 + 272 \frac{d_*}{\gamma} \right\}.

**Proof.** For the first inequality, set
\[\Delta_{K,w,t} = \left| \frac{\sum_{i=1}^{n} (\sqrt{K} \psi(t))}{\psi'(t)} \left\{ 1 - \frac{c(j + K)wt + t^2}{2} - \frac{t\psi''(t)}{[\psi'(t)]^2} \right\} \right. \]
\[\left. - \sum_{i=1}^{n} (t\sqrt{K}) \left\{ 1 - \frac{c(j + K)wt + t^2}{2} \right\} \right| \leq \frac{\sum_{i=1}^{n} (\sqrt{K} \psi(t)) t\psi''(t)}{[\psi'(t)]^3} \]
\[+ \left\{ 1 + \frac{c(j + K)wt + t^2}{2} + t^2 \right\} \left\{ \left| \frac{\sum_{i=1}^{n} (\sqrt{K} \psi(t)) - \sum_{i=1}^{n} (t\sqrt{K})}{\psi'(t)} \right| + \sum_{i=1}^{n} (t\sqrt{K}) \left| \frac{1}{\psi'(t)} - 1 \right| \right\}.
\]
In order to bound \(\Delta\), we use the following properties of \(\psi\). First, since \(\psi(0) = 0\), \(\psi'(0) = 1\) and \(1/2 \leq |\psi'(t)| \leq 5/4\), we also have \(|\psi(t)| \leq 5t/4\); then, since also \(|\psi''(t)| \leq 2K^{-1}(t + |d(j)|w)\), it follows that
\[|\psi'(t) - 1| \leq tK^{-1}(t + 2|d(j)|w) \quad \text{and} \quad |\psi(t) - t| \leq t^2K^{-1}(t/3 + |d(j)|w).
\]
These properties, together with the inequalities \(|\sum_{i=1}^{n} x| \leq 1 \land |x| \) and \(|\sum_{i=1}^{n} (x + y) - \sum_{i=1}^{n} x| \leq |y|\), now give the inequality
\[\Delta_{K,w,t} \leq 16K^{-1}t^2 + 20K^{-1/2}|d(j)|wt^2 + 4K^{-1/2}|d(j)|wt^2(1 + c(j + K)wt/2 + t^2) \]
\[+ 2K^{-1/2}t^2(1 + c(j + K)wt/2 + t^2)(t/3 + |d(j)|w) + 2K^{-1}t^2(1 + c(j + K)wt/2 + t^2).
\]
Then, noting that
\[\int_0^\infty dw \int_0^t dt \ e^{m+1}w^m t^2 e^{-cwt-t^2/2} = m! \kappa_{l-m-1}
\]
whenever \(l \geq m + 1\), where \(\kappa_m\) is as given in (4.10), we conclude that
\[\lambda K^{-1} \int_0^t dt \int_0^\infty dt \ e^{-(c(j + K)wt + t^2)/2} \Delta_{K,w,t} \leq \left( \frac{2}{\lambda \gamma} \right) \left\{ \frac{24}{K^2} + \frac{10}{3K^{3/2}} \sqrt{\frac{\pi}{2}} + \frac{88|d(j)|}{\gamma K^{3/2}} \sqrt{\frac{\pi}{2}} \right\};
\]
adding over \(K \geq 2\) using Lemma 5.6 gives the first inequality.
For the second, the argument is essentially the same, except that \( \psi'(t) \sum_{i=1}^{n} (\sqrt{K} \psi(t)) \) is integrated to give \( 1 - \cos(\sqrt{K} \psi(t)) \). The estimate \( 1 - \cos x \leq 2 \wedge (3|x|/4) \) now replaces \( |\sum_{i=1}^{n} x| \leq 1 \wedge |x| \), causing some slight alteration to the constants.

We now change variables once more, setting \( u = wK^{-1/2} \) and \( z = t\sqrt{K} \), to give

\[
I_{K5}^C = -(\lambda K)^{-1} \int_0^\infty du \cos(K \tau(j + K)u) e^{-c(j+K)u^2/2\lambda} \\
\int_0^\infty dz \sum_{i=1}^{n} z \{1 - c(j + K)uz/2 - z^2/K\} e^{-(c(j+K)uz+z^2/K)/2},
\]

\[
I_{K5}^S = (\lambda K)^{-1} \int_0^\infty du \sum_{i=1}^{n} (K \tau(j + K)u) e^{-c(j+K)u^2/2\lambda} \\
\int_0^\infty dz \cos z \{1 - c(j + K)uz/2 - z^2/K\} e^{-(c(j+K)uz+z^2/K)/2},
\]

since also \( \int_0^\infty (1 - cuz - z^2/K) e^{-cuz-z^2/2K} dz = 0 \). Written in this form, there are terms which become small for large \( K \), and we next show that they can be controlled.

**Lemma 4.7.** Define

\[
I_{K6}^C = -(\lambda K)^{-1} \int_0^\infty du \cos(K \tau(j + K)u) e^{-c(j+K)Ku^2/2\lambda} \\
\int_0^\infty dz \sum_{i=1}^{n} z \{1 - c(j + K)uz/2\} e^{-(c(j+K)uz)/2},
\]

\[
I_{K6}^S = -(\lambda K)^{-1} \int_0^\infty du \sum_{i=1}^{n} (K \tau(j + K)u) e^{-c(j+K)Ku^2/2\lambda} \\
\int_0^\infty dz \cos z \{1 - c(j + K)uz/2\} e^{-(c(j+K)uz)/2}.
\]

Then

\[
\sum_{K=2}^{[4\lambda m_*]+2} |I_{K5}^C - I_{K6}^C| \leq \frac{76}{\lambda \gamma}; \quad \sum_{K=2}^{[4\lambda m_*]+2} |I_{K5}^S - I_{K6}^S| \leq \frac{76}{\lambda \gamma}.
\]

**Proof.** The function \( K^{-1}z^2 \exp\{-cu - z^2/2K\} \) has exactly one turning point in \( z \geq 0 \) for each \( u \), and has maximum value at most \( \min\{2e^{-1}, 4/[Kc^2u^2]\} \). Hence, by Lemma 5.5, writing \( \Delta = \frac{1}{c(j+K)} \sqrt{\frac{8}{cK}} \), we have

\[
\left| \int_0^\infty dz \sum_{i=1}^{n} (z + \theta)K^{-1}z^2 e^{-(c(j+K)uz+z^2/K)/2} \right| \leq 16e^{-1} \min(1, 8/[Ke\{c(j + K)z^2\}]) \\
= 16e^{-1} \min\{1, u^{-2}\Delta^2\}
\]

33
for any $\theta$, so that omitting the terms $K^{-1 \cdot 2}$ in the braces in each of $I_{K6}^C$ and $I_{K6}^S$ for each $K$ changes the total contribution by at most

$$2 \sum_{K \geq 2} (\lambda K)^{-1}\left(\frac{16\Delta}{e}\right) \leq \frac{64\sqrt{2}}{\lambda\gamma e^{3/2}} \cdot \frac{5}{3} \leq \frac{34}{\lambda\gamma},$$

again using Lemma 5.6. Then the function $(1 - cuz)e^{-ez^2/(2K)}$ has two turning points in $z \geq 0$ for each $u$, and has maximum modulus at most $\min(1, 2/[Ke^2u^2])$, and so, again by Lemma 5.5,

$$\left|\int_0^\infty dz \sum_{i=1}^n (z + \theta)(1 - c(j + K)uz/2)e^{-c(j+K)uz/2}(1 - e^{-z^2/2K})\right|$$

$$\leq 12 \min(1, 8/[Ke^2(c(j + K)^2u^2)]),$$

for any $\theta$, and hence replacing the factor $e^{-z^2/2K}$ by 1 for each $K$ changes the total contribution by at most a further

$$2 \sum_{K \geq 2} (\lambda K)^{-1}12\left(\frac{2}{e\gamma}\sqrt{\frac{2}{K}}\right) \leq \frac{42}{\lambda\gamma}. $$

Having completed all this simplification, the inner integrals in $I_{K6}^C$ and $I_{K6}^S$ can be explicitly evaluated, giving

$$I_{K6}^C = -\frac{2}{\lambda K c(j + K)} \int_0^\infty dv \frac{1 - v^2}{(1 + v^2)^2} \cos\{Kv\nu_K\}v \exp\{-2K^2v^2/(\lambda c(j + K))\} \quad (4.17)$$

and

$$I_{K6}^S = \frac{2}{\lambda K c(j + K)} \int_0^\infty dv \frac{2v}{(1 + v^2)^2} \sum_{i=1}^n \{Kv\nu_K\}v \exp\{-2K^2v^2/(\lambda c(j + K))\}, \quad (4.18)$$

where $\nu_K = 2\tau(j + K)/c(j + K)$. We now bound these quantities in different ways, useful for different values of $K$.

**Lemma 4.8.** The following inequalities are true:

1. $|I_{K6}^C|, |I_{K6}^S| \leq K^{-2} \frac{2\pi}{\lambda\gamma}$;
2. $|I_{K6}^C|, |I_{K6}^S| \leq K^{-2} \frac{16}{\lambda\gamma|\nu_K|}$;
3. $|I_{K6}^S| \leq \frac{\pi |\nu_K|}{\lambda\gamma}$;
4. $|I_{K6}^C| \leq \frac{4|\nu_K|}{\lambda\gamma} + \frac{6}{(\lambda\gamma)^{3/2}}$. 

34
Proof. The first bound simply uses the observation that neither $2v$ nor $|1 - v^2|$ is larger than $1 + v^2$, so that both $|I_{K6}^C|$ and $|I_{K6}^S|$ are bounded by
\[
\frac{2}{\lambda K c(j + K)} \int_0^\infty dv \exp\left\{-2K^2v^2/(\lambda c(j + K))\right\},
\]
giving (1). For (2), we use the oscillation of sine and cosine; the function $(1-v^2)(1+v^2)^{-2}\exp\{-2K^2v^2/(\lambda c(j + K))\}$ has exactly one turning point in $v \geq 0$, and $2v(1+v^2)^{-2}\exp\{-2K^2v^2/(\lambda c(j + K))\}$ has at most two. In either case, applying Lemma 5.5,
\[
|I_{K6}^C|, \quad |I_{K6}^S| \leq \frac{2}{\lambda K c(j + K)} \frac{8}{K\nu_K},
\]
as required. Then for (3), we bound $\sum_{i=1}^{n} (\{K\nu_K\}v) \exp\{-2K^2v^2/(\lambda c(j + K))\}$ by $K\nu_K|v|$, and integrate.

The last part requires two calculations. First, we observe that
\[
\left|I_{K6}^C - \frac{2}{\lambda K c(j + K)} \int_0^\infty dv \frac{1-v^2}{(1+v^2)^2} \exp\left\{-2K^2v^2/(\lambda c(j + K))\right\}\right| \\
\leq \frac{2}{\lambda K c(j + K)} \int_0^\infty dv \frac{1-\cos(\{K\nu_K\}v)}{1+v^2},
\]
Because $1 - \cos x \leq \min(2, x^2/2)$, the right hand side does not exceed
\[
\frac{2}{\lambda K c(j + K)} \left\{ \frac{2}{K\nu_K} \frac{K^2|\nu_K|^2}{2} + \frac{2}{K\nu_K} \frac{K|\nu_K|}{2} \right\} \leq \frac{4\nu_K}{\lambda \gamma},
\]
then, because $\int_0^\infty dv (1-v^2)(1+v^2)^{-2} = 0$, we find that
\[
\frac{2}{\lambda K c(j + K)} \left| \int_0^\infty dv \frac{1-v^2}{(1+v^2)^2} \exp\left\{-2K^2v^2/(\lambda c(j + K))\right\}\right| \\
= \frac{2}{\lambda K c(j + K)} \left| \int_0^\infty dv \frac{1-v^2}{(1+v^2)^2} (1-\exp\{-2K^2v^2/(\lambda c(j + K))\})\right| \\
\leq \frac{2}{\lambda K c(j + K)} \left\{ \int_0^{\sqrt{\lambda \gamma}/K} \frac{2K^2v^2}{\lambda \gamma} \frac{dv}{1+v^2} + \int_{\sqrt{\lambda \gamma}/K}^\infty \frac{dv}{1+v^2} \right\} \leq 6(\lambda \gamma)^{-3/2}.
\]

For $K \geq \sqrt{\lambda \gamma}$, we use Lemma 4.8(1) to estimate $\sum_{K=\lceil\sqrt{\lambda \gamma}\rceil+1}^{\lceil\sqrt{\lambda \gamma}\rceil+2} |I_{K6}^C|$ and $\sum_{K=\lceil\sqrt{\lambda \gamma}\rceil+1}^{\lceil\sqrt{\lambda \gamma}\rceil+2} |I_{K6}^S|$. Thus it principally remains to find a suitable bound for
\[
Q = \sum_{K=2}^{\lfloor \sqrt{\lambda \gamma} \rfloor} \min\{|\nu_K|, 4/|K^2|\nu_K|\} \tag{4.19}
\]
for all possible values of $j$, where
\[
\nu_K = 2\tau(j + K)/c(j + K) = 2(\lambda m_1 - j - K)/(\lambda c(j) + K). \tag{4.20}
\]
This is the substance of the next lemma.
Lemma 4.9. For all values of \( j \), we have \( Q \leq 12 \).

Proof. Write \( a_j = \lambda m_1 - j \) and \( b_j = \lambda c(j) \geq \lambda \gamma \). We split the proof into two cases.

Case 1: \( |a_j| \leq 3\sqrt{\lambda \gamma} \).

Then, for all \( 2 \leq K \leq \sqrt{\lambda \gamma} \), we have \( |a_j - K| \leq 4\sqrt{\lambda \gamma} \) and \( b_j + K \geq \lambda \gamma \). Hence, in this case,

\[
Q \leq \sum_{K=2}^{\sqrt{\lambda \gamma}} |v_K| \leq 2 \left( \frac{4\sqrt{\lambda \gamma}}{\lambda \gamma} \right) \sqrt{\lambda \gamma} = 8.
\]

Case 2: \( |a_j| > 3\sqrt{\lambda \gamma} \).

Then, for all \( 2 \leq K \leq \sqrt{\lambda \gamma} \),

\[
2\sqrt{\lambda \gamma} \leq 2|a_j|/3 \leq |a_j - K| \leq 4|a_j|/3.
\]

Hence, for any \( \alpha > 0 \),

\[
\sum_{K=2}^{\sqrt{\lambda \gamma}} |v_K| \leq 2 \left( \frac{4|a_j|/3}{b_j} \right) \frac{b_j \alpha}{|a_j|} = \frac{8\alpha}{3},
\]

and

\[
\sum_{K=2\sqrt{(|a_j|/b_j)|+1}}^{\sqrt{\lambda \gamma}} \frac{4}{K^2 |v_K|} \leq 2 \left( \frac{2b_j}{2|a_j|/3} \right) \sum_{K=2\sqrt{(|a_j|/b_j)|+1}}^{\sqrt{\lambda \gamma}} K^{-2} \leq \frac{6b_j}{|a_j|} \frac{2|a_j|}{\alpha |b_j|} = \frac{12}{\alpha}.
\]

Minimizing the sum of these two quantities by choosing \( \alpha = 3\sqrt{2}/2 \), we find that, in this case, \( Q \leq 8\sqrt{2} < 12 \). This completes the proof.

Corollary 4.10. We have the following estimates:

\[
\sum_{K=2}^{[4\lambda m_*]+2} |I_{K\theta}^S| \leq \frac{4Q}{\lambda \gamma} + \sqrt{\frac{2\pi}{\lambda \gamma}} \left( \frac{1}{\sqrt{\lambda \gamma}} + \frac{1}{\lambda \gamma} \right) \leq \frac{54}{\lambda \gamma};
\]

\[
\sum_{K=2}^{[4\lambda m_*]+2} |I_{K\theta}^C| \leq \frac{54}{\lambda \gamma} + 6\sqrt{\lambda \gamma} (\lambda \gamma)^{-3/2} = \frac{60}{\lambda \gamma}.
\]

5. Auxiliary results.

We use the notation of Section 1 throughout; in particular, \( r \) and \( s \) are as in (1.24) and (1.25). We first investigate the properties of two functions related to \( \mu(\cdot) \),

\[
h_t(u) = \mu(1+iu) - 1 - \lambda^{-1} t \log(1+iu) \quad \text{and} \quad w_t(t) = \mu(1-t) - 1 - \lambda^{-1} t \log(1-t),
\]

for \(-s \leq u \leq s\) and \(0 \leq t \leq 1\).
Lemma 5.1. The function $h_l$ has the following properties:

(1) $\Re e \{h_l(u)\} = h_l(-u)$;

(2) $\frac{d}{du} \Re e \{h_l(u)\} \leq -\gamma u / 2$, \hspace{1em} 0 \leq u \leq s, \hspace{1em} l \geq \frac{1}{2} \lambda m_1;

(3) $|h_l^{(3)}(u)| \leq \frac{2}{3} m_*^3 + 2 \lambda^{-1} l$ and $h_l^{(4)}(u) \leq \frac{2}{3} m_*^4 + 6 \lambda^{-1} l$, \hspace{1em} |u| \leq s.

Proof. The first statement is implied by conjugation. For the second, since

$$
\Re e \{\mu(1 + iu)\} = \Re e \left\{ \sum_{j \geq 1} \mu_j (1 + iu)^j \right\} = 1 + \sum_{k \geq 1} (-1)^k m_{[2k]} \frac{u^{2k}}{(2k)!},
$$

it follows that

$$
\left| \frac{d}{du} \Re e \{\mu(1 + iu)\} + m_{[2]} u \right| \leq \sum_{k \geq 2} m_{[2k]} \left| \frac{|u|^{2k-1}}{(2k-1)!} \right| \leq \frac{m_*^4 |u|^3}{6(1 - (m_*)^2)} \leq \frac{27}{160} m_*^4 |u|^3 \leq \frac{\gamma u}{480},
$$

this last from (1.25). Then, for $0 \leq u \leq s$,

$$
- \frac{d}{du} \Re e \{\log(1 + iu)\} = - \frac{u}{1 + u^2} \leq - \frac{81 u}{82},
$$

again by (1.25). Hence, for $0 \leq u \leq s$,

$$
\frac{d}{du} \Re e \{h_l(u)\} \leq - m_{[2]} u + \frac{\gamma u}{480} - \frac{81 m_1 u}{164} \leq - \frac{\gamma u}{2}.
$$

For the last part, simply observe that

$$
|h_l^{(3)}(u)| \leq \sum_{k \geq 0} m_{[k+3]} \left| \frac{|u|^k}{k!} \right| + 2 \lambda^{-1} l \leq \frac{m_*^3}{1 - m_* s} + 2 \lambda^{-1} l,
$$

with a similar calculation for $|h_l^{(4)}(u)|$.

Lemma 5.2. If $1/2 \leq l / \lambda m_1 < 1$, then $w_l$ has a unique minimum at some $t_0 \in (0, 1)$; furthermore,

(1) $w_l(t) - w_l(t_0) \geq \frac{m_1}{4} (t - t_0)^2$, \hspace{1em} 0 \leq t \leq 1;

(2) $\sum_{j \geq 1} \mu_j j^2 (1 - t)^j \geq \frac{m_*^2}{4}$, \hspace{1em} 0 \leq t \leq t_0;

(3) $t_0 \leq t^*_0 \leq \frac{1}{2} \wedge \frac{1}{m_1}$.
where $t_0^* = t_0^*(\mu)$ solves $m_1^{-1} \sum_{j \geq 1} \mu_j j (1-t)^j = 1/2$.

**Proof.** Direct calculation shows that $w_l(0) = 0$, $w_l(1) = \infty$ and that

$w_l'(t) = -\mu'(1-t) + \frac{l}{\lambda(1-t)}$; \quad $w_l''(t) = \mu''(1-t) + \frac{l}{\lambda(1-t)^2} \geq \frac{m_1}{2}$.

Hence $w_l$ is convex, and $w_l'(0) < 0$ since $l < \lambda m_1$. Thus the minimum of $w_l$ occurs at $t_0 \in (0,1)$ satisfying

$$m_1^{-1}(1-t)\mu'(1-t) = m_1^{-1} \sum_{j \geq 1} \mu_j j (1-t)^j = 1/\lambda m_1, \tag{5.2}$$

and

$$w_l(t) - w_l(t_0) = \frac{1}{2} (t - t_0)^2 w_l''(t') \geq m_1 (t - t_0)^2/4,$$

for some $t'$, proving (1). Then, by the Schwarz inequality and from (5.2), for $0 \leq t \leq t_0$,

$$\sum_{j \geq 1} \mu_j j^2 (1-t)^j \geq \sum_{j \geq 1} \mu_j j^2 (1-t_0)^j \geq \left\{ \sum_{j \geq 1} \mu_j j (1-t_0)^j \right\}^2 / \sum_{j \geq 1} \mu_j (1-t_0)^j \geq m_1^2/4,$$

proving (2). That $t_0 \leq t_0^*$ follows from (5.2), because $l \geq \lambda m_1/2$. The final upper bound on $t_0^*$ follows from a geometrical argument. The function $y(t) = \sum_{j \geq 1} \mu_j j (1-t)^j$ is convex, and takes the values $m_1$ at $t = 0$, $m_1/2$ at $t = t_0^*$ and 0 at $t = 1$. Hence, if $Q$ denotes the area between the $t$- and the $y$-axes and the tangent to $y$ at $(t_0^*, m_1/2)$, we have $Q \leq \int_0^1 y(t) \, dt \leq m_1/2$, this last being the area of the triangle with vertices at $(0,0), (0, m_1)$ and $(1,0)$. Also, integrating the definition of $y$, $\int_0^1 y(t) \, dt \leq 1$. On the other hand, if $x_0, y_0$ are both positive, the area of any triangle formed by the $x$- and the $y$-axes and a line through $(x_0, y_0)$ is at least $2x_0y_0$. Hence $Q \geq 2t_0^* m_1/2 = m_1 t_0^*$, completing the proof.

**Lemma 5.3.** If $l \geq \lambda m_1 - 1$ and $\lambda m_1 \geq 2$, then $\lambda w_l(t) \geq \frac{1}{4} \lambda m_1 t^2 - 1$, $0 \leq t \leq 1$.

**Proof.** It is immediate that $\mu(1) - \mu(1-t) \leq m_1 t$ and that $\log(1-t) \leq -t - t^2/2$. Hence

$$\lambda w_l(t) \geq -\lambda m_1 t + (\lambda m_1 - 1)(t + t^2/2) \geq -1 + t^2(\lambda m_1 - 1)/2,$$

and the lemma follows.

The next lemma gives an accurate lower tail estimate for the CP $(\lambda, \mu)$ distribution.
Lemma 5.4. If $S \sim \text{CP}(\lambda, \mu)$ and $\frac{1}{2} \lambda m_1 \leq l < \lambda m_1$, then $\mathbb{P}[S \leq l] \leq C_{5.4} e^{\lambda w_l(t_0)}$, where
\[
C_{5.4} = \min \left\{ 1, \lambda^{-1/2} \left( \frac{14.4 m_3}{m_1^3} + \frac{2}{m_1 t_0} \right) \right\},
\]
and where $w_l$ and $t_0$ are as in Lemma 5.2.

**Proof.** Observe first that, by straightforward manipulations,
\[
\mathbb{P}[S \leq l] = \mathbb{E} \left\{ (1 - t)^S (1 - t)^{-l} \right\} \mathbb{E} \left\{ \left[ \frac{(1 - t)^s}{(1 - t)^s} \right] \left[ (1 - t)^{l-S} I_{(S \leq 0)} \right] \right\}
= \mathbb{E} \left\{ (1 - t)^S (1 - t)^{-l} \right\} \mathbb{E} \left\{ (1 - t)^{l-S(1-t)} I_{(S(1-t) \leq 0)} \right\}
= e^{\lambda w_l(t)} \mathbb{E} \left\{ (1 - t)^{l-S(1-t)} I_{(S(1-t) \leq 0)} \right\},
\]
where $S_h$ has distribution defined by the Esscher transform
\[
\mathbb{E}_h S_h = \frac{\mathbb{E}[S|h]}{\mathbb{E}[h]} = \exp(\lambda(\mu(h) - \mu(h))).
\]
Applying Lemma 5.2, there is a unique root $t_0 \in (0, 1)$ of the equation $w_l'(t_0) = 0$, satisfying
\[
\inf_{0 < t \leq 1} w_l(t) = w_l(t_0) < 0 \text{ and also}
\]
\[
\mathbb{E} S_{(1-t_0)} = \lambda(1-t_0) \mu'(1-t_0) = l, \quad S_{(1-t_0)} - l = \sum_{i=1}^{\infty} i \left( Y_i(\lambda \mu_i[1-t_0]^i - \lambda \mu_i[1-t_0]^i) \right),
\]
where the $Y_i(\cdot)$ are independent Poisson processes with intensity 1.

Now let $\xi$ be a random variable with zero mean and variance $\sigma^2$. Then, for any $h > 0$,
\[
\mathbb{E} \left\{ e^{h \xi I_{(\xi \leq 0)}} \right\} = h \int_{-\infty}^{0} e^{hx} \mathbb{P}[\xi < x \leq 0] dx - \frac{1}{2} \int_{-\infty}^{0} e^{hx} \sigma^2 e^{-x^2/2} dx
\leq 2 \sup_x |\mathbb{P}[\xi \leq x] - \Phi(x)| + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{hx} e^{-x^2/2} dx
\leq 2 \Delta + e^{h^2 \sigma^2/2} \frac{1}{\sqrt{2\pi}} \int_{h \sigma}^{\infty} e^{-u^2/2} du \
\leq 2 \Delta + \left( \frac{1}{h \sigma} \wedge \frac{1}{2} \right),
\]
where $\Delta = \sup_x |\mathbb{P}[\xi \leq x] - \Phi(x)|$. In particular, this estimate holds for $\xi = S_{(1-t_0)} - l$ and $h = - \log(1-t_0) > t_0 > 0$ with $\sigma^2 = \text{Var} S_{(1-t_0)}$, providing an upper bound for
\[
\mathbb{E} \left\{ (1 - t_0)^{l-S(1-t_0)} I_{(S(1-t_0) \leq 0)} \right\} \text{ in (5.3)}.
\]
In order to make this upper bound explicit, we use the Berry–Esseen theorem to control $\Delta$. Writing each $Y_i(\cdot)$ as the sum of $n$ independent identically distributed Poisson processes $Y_{ij}^{(n)}$, $1 \leq j \leq n$, each with intensity $1/n$, we express $S_{(1-t_0)} - l = \sum_{j=1}^{n} Z_j^{(n)}$.
as a sum of independent and identically distributed random variables with zero mean and finite third moment, where
\[
Z_{j}^{(n)} = \sum_{i=1}^{\infty} \left\{ iY_{ij}^{(n)} (\lambda \mu_i (1 - t_0)^i) - (\lambda/n) \mu_i (1 - t_0)^i \right\}.
\]

In particular,
\[
\text{Var} Z_{1}^{(n)} = \sigma^2/n = \frac{\lambda}{n} \sum_{i=1}^{\infty} \{ i^2 \mu_i (1 - t_0)^i \} = n^{-1} \lambda A^2,
\]
say, and
\[
\lim_{n \to \infty} \inf n \mathbb{E} |Z_{1}^{(n)}|^3 = \lambda \sum_{i=1}^{\infty} \{ i^3 \mu_i (1 - t_0)^i \} = \lambda B^3.
\]

Hence, by the Berry–Esseen theorem with the constant of Berk, for \( \xi = S_{(1-t_0)} - l \), we find that
\[
\Delta \leq 0.9 \lim_{n \to \infty} n^{-1/2} \frac{\mathbb{E} |Z_{1}^{(n)}|^3}{\text{Var} |Z_{1}^{(n)}|^{3/2}} \leq 0.9 \frac{B^3}{A^3 \sqrt[3]{\lambda}}.
\]

Substituting from (5.5)–(5.7) in (5.4), we find that
\[
\mathbb{P} [S \leq l] \leq e^{\lambda \omega_l(t_0)} \min \left[ 1, \left( \frac{1.8B^3}{A^3 \sqrt[3]{\lambda}} + \frac{1}{t_0 A \sqrt[3]{\lambda}} \right) \right].
\]

However, from (5.6) and (5.5), \( B^3 = m_3^{(1-t_0)} \leq m_3 \) and \( A^2 = m_2^{(1-t_0)} \geq m_1^2/4 \), this last using Lemma 5.2(2), completing the proof.

**Lemma 5.5.** If \( f \) is monotone on an interval \([a, b]\), then, for any \( \theta \),
\[
\left| \int_{a}^{b} f(x) \cos(x + \theta) \, dx \right| \leq 4 \sup_{a \leq x \leq b} |f(x)|.
\]

**Proof.** Integrating by parts,
\[
\left| \int_{a}^{b} f(x) \cos(x + \theta) \, dx \right| = \left| \left[ f(x) \sum_{i=1}^{n} (x + \theta) \right]_{a}^{b} - \int_{a}^{b} f'(x) \sum_{i=1}^{n} (x + \theta) \, dx \right|
\]
\[
\leq |f(b)| + |f(a)| + \int_{a}^{b} |f'(x)| \, dx \leq 4 \sup_{a \leq x \leq b} |f(x)|.
\]

**Lemma 5.6.** We have the following inequalities:
\[
\sum_{K=2}^{[4 \lambda m_3] + 2} K^{-1/2} \leq 4 \sqrt{\lambda m_3}; \quad \sum_{K \geq 2} K^{-3/2} \leq 5/3;
\]
\[
\sum_{K \geq 2} K^{-2} \leq 2/3; \quad \sum_{K \geq 2} K^{-5/2} \leq 2/5.
\]
Proof. For the first sum, we have
\[
\sum_{K=2}^{[4\lambda m_\ast]+2} K^{-1/2} \leq \int_1^{[4\lambda m_\ast]+2} x^{-1/2} \, dx \leq 2\{(4\lambda m_\ast + 2)^{1/2} - 1\} \leq 4\sqrt{\lambda m_\ast}.
\]

For the remainder, we note that, for \(a > 1\),
\[
\sum_{K \geq 2} K^{-a} \leq \sum_{K=2}^{5} K^{-a} + \int_5^\infty x^{-a} \, dx,
\]
and then make the necessary computations.

References