Discrete Mathematics and Theoretical Computer Science **4**, 2001, 363–398

Asymptotic normality of recursive algorithms via martingale difference arrays

Werner Schachinger

Dept. of Statistics and Decision Support Systems, University of Vienna
Brünnerstr. 72, A-1210 Wien, Austria
e-mail: Werner.Schachinger@univie.ac.at

received Apr 10, 2000, accepted Dec 27, 2001.

We propose martingale central limit theorems as an appropriate tool to prove asymptotic normality of the costs of certain recursive algorithms which are subjected to random input data. The recursive algorithms that we have in mind are such that if input data of size N produce random costs L_N, then $L_N \stackrel{\mathcal{D}}{=} L_n + \bar{L}_{N-n} + R_N$ for $N \geq n_0 \geq 2$, where n follows a certain distribution P_N on the integers $\{0,\dots,N\}$ and $L_k \stackrel{\mathcal{D}}{=} \bar{L}_k$ for $k \geq 0$. L_n, L_{N-n} and R_N are independent, conditional on n, and R_N are random variables, which may also depend on n, corresponding to the cost of splitting the input data of size N (into subsets of size n and $N-n$) and combining the results of the recursive calls to yield the overall result. We construct a martingale difference array with rows converging to $Z_N := \frac{L_N - \mathbb{E}\, L_N}{\sqrt{\operatorname{Var} L_N}}$. Under certain compatibility assumptions on the sequence $(P_N)_{N\geq 0}$ we show that a pair of sufficient conditions (of Lyapunov type) for $Z_N \stackrel{\mathcal{D}}{\to} \mathcal{N}(0,1)$ can be restated as a pair of conditions regarding asymptotic relations between three sequences. All these sequences satisfy the same type of linear equation, that is also the defining equation for the sequence $(\mathbb{E}\, L_N)_{N\geq 0}$. In the case that the P_N are binomial distributions with the same parameter p, and for deterministic R_N, we demonstrate the power of this approach. We derive very general sufficient conditions in terms of the sequence $(R_N)_{N\geq 0}$ (and for the scale $R_N = N^{\alpha}$ a characterization of those α) leading to asymptotic normality of Z_N.

Keywords: recursive algorithms, trie, martingales, asymptotic normality, central limit theorem

1 Introduction

There are several methods in the literature to detect asymptotic normality of appropriately normalized costs of recursive algorithms. Among the most prominent approaches are the use of bivariate moment generating functions (cf. [2, 14, 15, 16, 22], sometimes assisted by singularity analysis of generating functions [7] and depoissonization devices [17]), urn models (cf. [20, 23, 24]), approximation by Brownian excursions (cf. [11]), and the contraction method (cf. [26, 28, 29, 30]). Occasionally the martingale limit theorem has been used to prove the existence of a limiting distribution (cf. [27, 28]). However, we do not know of any applications of central limit theorems for martingale difference arrays in the analysis of recursive algorithms. The aim of this paper is thus to demonstrate that the latter are valuable tools that can supplement the other methods.

When we study recursive algorithms which are subjected to random input data, martingales arise in a very natural manner when we make predictions of costs on the basis of the information available by keeping track of the recursive calls performed so far. The following example should make this clear: Assume that some recursive algorithm, when applied to random input data of size N, produces random costs L_N, which satisfy $L_0 = L_1 = 0$, almost surely, and for $N \geq 2$

$$L_N \stackrel{\mathcal{D}}{=} L_{N^0} + \bar{L}_{N^1} + r_N, \tag{1}$$

where N^0 follows a certain distribution P_N on the integers $\{0,\dots,N\}$, $N^1 = N - N^0$, $L_k \stackrel{\mathcal{D}}{=} \bar{L}_k$ for $k \geq 0$, and $L_{N^0}, \bar{L}_{N^1}$ are independent, conditional on N^0. Finally r_N is a constant, corresponding to the cost of splitting the input data of size N (into subsets of size N^0 and N^1) and combining the results of the recursive calls to yield the overall result. The best guess that we can make about L_N, knowing just N, is $X_{N,0} = \ell_N := \mathbb{E}\, L_N$. If we also know the value of N^0, we can improve our guess:

$$X_{N,1} = X_{N,0} + \ell_{N^0} + \ell_{N^1} + r_N - \ell_N.$$

If the algorithm splits the data subset of size N^0 first (into subsets of sizes N^{00} and N^{01}), the next we get to know will be the value of N^{00}. This will lead to another improvement of our guess of L_N:

$$X_{N,2} = X_{N,1} + \ell_{N^{00}} + \ell_{N^{01}} + r_{N^0} - \ell_{N^0}.$$

Under certain integrability conditions on L_N, the sequence $(X_{N,i})_{i\geq 0}$ constructed this way will be a martingale with respect to a certain filtration obtained by accumulating information about subset sizes. In the lucky case that knowing all subset sizes almost surely determines L_N, we have $X_{N,i} \to L_N$ almost surely and in L^2, which opens the door for applying classical central limit theorems for martingale difference arrays. Under certain assumptions on the sequence $(P_N)_{N\geq 0}$ (which still allow for the standard probabilistic models of algorithms associated with binary search trees, digital search trees, tries, ...) we will derive easy-to-use conditions (at the cost of having a narrower range of applicability than the classical Lindeberg conditions) implying asymptotic normality of costs L_N. The setting which will be our favorite playground for demonstrating applications of these conditions is roughly the following:

If in (1) we fix $\mathbb{P}\,(N^0 = k) = \binom{N}{k} p^k (1-p)^{N-k}$ for $0 \leq k \leq N$ and some fixed $0 < p < 1$, we obtain a recursion that shows up again and again in the study of additive valuations of the (binary) trie data structure (cf. [9, 19, 21]) under the Bernoulli model. The number of internal nodes ($r_N = 1$) and the external path length ($r_N = N$) of a trie are perhaps the most important examples. Jacquet and Regnier [14, 15] proved asymptotic normality of both the number of internal nodes and of the external path length in a binary trie under the Bernoulli model, and in the case of the number of internal nodes also proved convergence of moments of any order. There is related work by Jacquet and Szpankowski [16], who proved asymptotic normality of the internal path length of a digital search tree under the Bernoulli model. These results are achieved using clever bounds for bivariate moment generating functions combined with a poissonization-depoissonization step. Employing contraction properties of suitably chosen probability metrics, Rachev and Rüschendorf [26] and Feldman, Rachev and Rüschendorf [4] proved asymptotic normality of L_N for the sequence $r_N = 1$ under very general probabilistic models, including the Bernoulli models. There is a remark in [26] saying that under certain conditions sequences $r_N = o(\sqrt{N})$ would generate asymptotically normal L_N.

Thus the following question naturally arises: Which sequences $(r_N)_{N\geq 2}$ generate additive valuations on the set of tries equipped with the Bernoulli model that behave asymptotically normal? We will give answers that in particular cover the cases $r_N = 1$ and $r_N = N$ and to a large extent clarify the role played by the sequences $r_N = o(\sqrt{N})$.

The paper is organized as follows: In Section 2 we set up a correspondence between algorithms that split tasks into at most two "subtasks", and labeled binary trees. Furthermore we describe the class of probabilistic input models we are going to allow. Essentially we will demand that the costs for the two subtasks and the split-and-combine cost are independent, given the "sizes" of the subtasks, and that the distribution of the cost of a certain task is the same, regardless if it is the task the algorithm starts with, or if it occurs as a subtask in some deeper level of recursion. Costs L_N of algorithms can then be regarded as "random additive valuations", generated by corresponding split-and-combine valuations R_N, on probability spaces consisting of labeled binary trees of fixed "size" N. If we demand that trees of fixed "size" are almost surely finite (which reflects the wish that the algorithm, when applied to random input, will almost surely stop in finite time), it turns out that moments of the costs are finite, if only the same moments of the corresponding split-and-combine valuation are finite. Next we will construct martingales converging to the costs L_N and will also derive linear recurrence relations for expectations and variances of L_N. The same type of linear recurrence relations occurs again twice in Lemma 1, which states sufficient conditions for asymptotic normality of normalized costs (where N tends to ∞) in terms of the solutions of the latter linear recurrence relations and the sequence $(\operatorname{Var} L_N)_{N\geq 0}$.

In Section 3 we are going to apply Lemma 1 to find answers to the question: Which are the sequences $(r_N)_{N\geq 2}$, such that L_N, as defined by (1) under the Bernoulli model, satisfies $\frac{L_N - \mathbb{E} L_N}{\sqrt{\operatorname{Var} L_N}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$? Our answers will be in terms of growth conditions for the sequence $(r_N)_{N\geq 2}$ and for sequences which are obtained by smoothing the sequences of first and second differences of $(r_N)_{N\geq 2}$. The verification of the conditions of Lemma 1 requires a careful study of the type of linear recurrence equation that defines the sequence $(\mathbb{E} L_N)_{N\geq 0}$, which is the content of two propositions. (For better readability of the paper the proof of one of these is deferred to the appendix.) For the class of sequences $r_N = N^\alpha$ we can even obtain a complete characterization: There is asymptotic normality for any α, if $p \neq \frac{1}{2}$, and only for $\alpha \leq \frac{3}{2}$, if $p = \frac{1}{2}$. Only a part of that characterization will be achieved by applying Lemma 1. It is in the nature of that lemma that it can deal only with sequences $(r_N)_{N\geq 2}$ that do not grow too fast, as it exploits negligibility in the limit of the martingale differences.

Examples 1 and 3, given in Section 4, are the missing links in the characterization of the sequences $r_N = N^\alpha$. In Example 1 we demonstrate that there is no normal limiting distribution in the cases $p = \frac{1}{2}, \alpha > \frac{3}{2}$, and in Example 3 we appeal to a "nonclassical" central limit theorem for martingale difference arrays to establish asymptotic normality for the cases $p \neq \frac{1}{2}, \alpha > 1$. In Example 2 we will show that the sufficient conditions derived in Section 3 are sharp in some sense, by supplying for the case $p = \frac{1}{2}$ a sequence $(r_N)_{N\geq 2}$, which does not lead to a normal limiting distribution, but satisfies $r_N = O(\sqrt{N})$ and thus falls very short of satisfying one of our sufficient conditions for asymptotic normality, namely $\frac{\ln p}{\ln q} \in \mathbb{Q}$, $r_N = o(\sqrt{N})$.

We denote convergence (resp. equality) in distribution by $\xrightarrow{\mathcal{D}}$ (resp. $\overset{\mathcal{D}}{=}$), and $\mathcal{N}(0,1)$ denotes a standard normal random variable. We put $a \vee b = \max(a,b)$ and $a \wedge b = \min(a,b)$ for real numbers a and b. The indicator function of a set A is denoted $\mathbb{I}_A$, and for a Boolean expression B we let $\mathbb{I}_{\{B\}}$ be 1 if B is true and 0 otherwise. The difference operator, Δ, is defined by $\Delta x_k = x_{k+1} - x_k$ for sequences $(x_k)_{k\geq 0}$. We will use the standard asymptotic notations O, o, Ω and Θ.

2 Preliminaries and a key lemma

We assume that we are given a class of problems $\mathcal{A}$, and that to each $A \in \mathcal{A}$ is associated a nonnegative integer $|A|$, the size of A. Examples of such classes would be the set of all finite sequences (which we want to sort) that are permutations of initial segments of the natural numbers, where the size of a sequence is the number of its terms, or the set of all finite binary trees (the path lengths of which we want to determine), where the size of a tree is the number of nodes it consists of.

We will consider algorithms which are recursive in the sense that a problem $A \in \mathcal{A}$ of size N is split into two primary subproblems $A', A'' \in \mathcal{A}$ of smaller or equal sizes, which are subjected to the same given algorithm. This splitting continues recursively, until subproblem sizes fall below some level n_0. These small subproblems are attacked directly (nonrecursively) by the algorithm. Splitting and combining causes costs, that depend on the problem to be split (perhaps only via the size of that problem), but can also have a stochastic component. Another source of randomness comes into play, if we subject the algorithm to a probabilistic input model: Each of the sets $\mathcal{A}_N := \{A \in \mathcal{A} : |A| = N\}$, assumed to be countable, for simplicity, is supplied with a probability measure, according to which elements of $\mathcal{A}_N$ can be chosen at random. The cost of our algorithm, when applied to input from $\mathcal{A}_N$, thus becomes a random variable. Properly normalized, these random variables might have a limit in distribution, when $N \to \infty$.

We will utilize the following representation of the cascade of subproblems just described in terms of labeled binary trees: To each problem A we construct a finite binary tree $\mathsf{t} = \mathsf{t}(A)$, with nodes labeled by the set $\mathbb{N} \cup \{-1\}$ and the labeling not required to be one-to-one. The size $|A|$ of the problem A is the label of the root of the tree $\mathsf{t}(A)$, whose left and right subtrees $\mathsf{t}(A')$ and $\mathsf{t}(A'')$ correspond to A's primary subproblems A' and A''. We proceed recursively, until we reach subproblem sizes less than n_0. This happens after finitely many steps, if we assume $|B'| \vee |B''| \leq |B|$, for any subproblem B of A with $|B| \geq n_0$, where B', B'' denote the primary subproblems of B, and that $|B'| \vee |B''| = |B|$ may occur only finitely many times. Each of the countably many problems of size n, for $0 \leq n < n_0$, can be represented by a unique finite labeled binary tree, whose root is labeled n and whose remaining nodes we label for definiteness by -1.

We define $\mathsf{T}_N := \{\mathsf{t}(A) : A \in \mathcal{A}_N\}$ for $N \geq 0$, and T_{-1} to be the set containing the empty tree and the finite binary trees with all nodes labeled by -1. Moreover we let the size $|\mathsf{t}|$ of $\mathsf{t} \in \bigcup_{N \geq -1} \mathsf{T}_N$ be -1 if t is empty, and the label of the root of t otherwise, i.e. $|\mathsf{t}| = N :\Leftrightarrow \mathsf{t} \in \mathsf{T}_N$. Let $\{v_1, v_2, v_3, \dots\}$ be some enumeration of the vertex set $\mathcal{V}$ of the infinite complete binary tree t_∞, such that v_1 is the root of t_∞, and v_j is a successor of v_i only if $i < j$. For any finite binary tree t we denote the vertex set of t by $\mathcal{V}(\mathsf{t})$, and we let $\iota_\mathsf{t} : \mathcal{V}(\mathsf{t}) \to \mathcal{V}$ be the embedding of t in t_∞, that satisfies $\iota_\mathsf{t}(root(\mathsf{t})) = v_1$, and u is the left (right) son of v in t iff $\iota_\mathsf{t}(u)$ is the left (right) son of $\iota_\mathsf{t}(v)$ in t_∞. For $\mathsf{t} \in \bigcup_{N \geq -1} \mathsf{T}_N$ we denote by $\mathsf{t}^{(i)}$ the subtree of t which has its root in $\iota_\mathsf{t}^{-1}(v_i)$. Note that either $\mathsf{t}^{(i)}$ is empty, or $\mathsf{t}^{(i)} \in \mathsf{T}_{|\mathsf{t}^{(i)}|}$. Left and right subtrees of t (resp. $\mathsf{t}^{(i)}$) are denoted t_ℓ and t_r (resp. $\mathsf{t}_\ell^{(i)}$ and $\mathsf{t}_r^{(i)}$,) and sometimes we depict that as $\mathsf{t} = \begin{smallmatrix} & \mathrm{o} & \\ & / \backslash & \\ \mathsf{t}_\ell & & \mathsf{t}_r \end{smallmatrix}$. The cost L of the algorithm, when applied to problem A, can now be given in terms of an *additive valuation* of the tree $\mathsf{t}(A)$:

Additive valuations. A (*deterministic*) *valuation* on a family of trees T is any function $X : \mathsf{T} \to \mathbb{R}$, and a *random valuation* is any function $X : \mathsf{T} \to L_Q^0(\Omega, \mathcal{G})$, where $L_Q^0(\Omega, \mathcal{G})$ denotes the set of random variables on a probability space $(\Omega, \mathcal{G}, Q)$. We shall concentrate on the particular class of *additive valuations* L, defined on $\bigcup_{N\geq 0}\mathsf{T}_N$, which can for some $n_0 \geq 2$ be described by

$$L(\mathsf{t}) = \begin{cases} R(\mathsf{t}), & |\mathsf{t}| < n_0, \\ R(\mathsf{t}) + L(\mathsf{t}_\ell) + L(\mathsf{t}_r), & |\mathsf{t}| \geq n_0,\ \mathsf{t} = \overset{\text{o}}{\mathsf{t}_\ell \,/\,\backslash\, \mathsf{t}_r}, \end{cases} \tag{2}$$

where R is some simpler valuation on $\bigcup_{N\geq 0}\mathsf{T}_N$. In the language of recursive algorithms, R accounts both for the costs of treating small subproblems B of size $|B| < n_0$ and for the costs of the split and combine steps. We say that R *generates* L. We assume that, for $|\mathsf{t}| \geq n_0$, $R(\mathsf{t})$ depends on $\mathsf{t} \in \mathsf{T}_n$ only via $|\mathsf{t}|$, $|\mathsf{t}_\ell|$ and $|\mathsf{t}_r|$, i.e. $R(\mathsf{t}) = R(|\mathsf{t}|, |\mathsf{t}_\ell|, |\mathsf{t}_r|)$. However, for $|\mathsf{t}| \geq 0$ we allow $R(\mathsf{t})$ to be a random variable, that is, we consider random additive valuations $L(\mathsf{t}) = L(\mathsf{t}, \omega)$, generated by $R(\mathsf{t}) = R(\mathsf{t}, \omega)$, with $\omega \in \Omega$ for a given probability space $(\Omega, \mathcal{G}, Q)$, where we assume that $R(\mathsf{t})$, $L(\mathsf{t}_l)$ and $L(\mathsf{t}_r)$ are independent. To be precise, this calls for existence of countably many mutually independent random variables $R^{(i)}(\mathsf{t})$, where $i \in \mathbb{N}$, $\mathsf{t} \in \bigcup_{N\geq 0}\mathsf{T}_N$ and for fixed t the $R^{(i)}(\mathsf{t})$ are i.i.d., so we can take $\Omega = [0,1]$, $\mathcal{G}$ the σ-field of Lebesque measurable sets in $[0,1]$, and Q the Lebesque measure. This allows for representing L as follows,

$$L(\mathsf{t}) = \sum_{i:\, |\mathsf{t}^{(i)}| \geq 0} R^{(i)}(\mathsf{t}^{(i)}). \tag{3}$$

For example, a (deterministic) additive valuation L is generated by $R(\mathsf{t}) = \mathbb{1}_{\{|\mathsf{t}| \geq n_0\}}$, and here $L(\mathsf{t}(A))$ counts the split and combine steps, that our recursive algorithm needs when applied to problem A.

The probabilistic model for T_N. We will work with the probability space $(\mathsf{T}_N, \mathcal{F}_N, P_N)$, where the set T_N is countable, so we simply define $\mathcal{F}_N$ to be the set of all subsets of T_N. Given $\mathsf{t} \in \mathsf{T}_N$, $N \geq n_0$, we assume that t_ℓ and t_r are independent, conditional on $\{|\mathsf{t}_\ell|, |\mathsf{t}_r|\}$, and moreover that $P_N(\mathsf{t}_\ell = \mathsf{t}^* \big| |\mathsf{t}_\ell| = |\mathsf{t}^*|) = P_{|\mathsf{t}^*|}(\mathsf{t}^*)$ and $P_N(\mathsf{t}_r = \mathsf{t}^* \big| |\mathsf{t}_r| = |\mathsf{t}^*|) = P_{|\mathsf{t}^*|}(\mathsf{t}^*)$. The latter says that the distribution of some subtree t^* of t depends only on its size $|\mathsf{t}^*|$ and not on the position of its root in the tree t. Thus, given the probability measures P_n for $n < n_0$, and for $N \geq n_0$ the *splitting probabilities*

$$p_{N,k',k''} := \mathbb{P}\left(|\mathsf{t}_\ell| = k', |\mathsf{t}_r| = k'' \Big| |\mathsf{t}| = N\right),$$

we have for $N \geq n_0$ the following recursive definition of the probability measures P_N,

$$P_N(\mathsf{t}) = p_{N,|\mathsf{t}_\ell|,|\mathsf{t}_r|} P_{|\mathsf{t}_\ell|}(\mathsf{t}_\ell) P_{|\mathsf{t}_r|}(\mathsf{t}_r).$$

Our assumptions on the splitting probabilities, that guarantee almost sure finiteness of $\mathsf{t} \in \mathsf{T}_N$, are that for $N \geq n_0$ we have

$$\begin{aligned} P_N(|\mathsf{t}_\ell| \vee |\mathsf{t}_r| \leq N) &= 1 = P_N(|\mathsf{t}_\ell| \wedge |\mathsf{t}_r| < N), \\ P_N(|\mathsf{t}_\ell| \vee |\mathsf{t}_r| < N) &= \sum_{0 \leq k' \vee k'' < N} p_{N,k',k''} =: \pi_N > 0. \end{aligned} \tag{4}$$

We denote by $X_N = X_N(\mathsf{t}, \omega)$ the random variable on the filtered probability space $(\mathsf{T}_N \times \Omega, \mathcal{F}_N \times \mathcal{G}, \mathbb{F}_N, P_N \times Q)$, obtained by restricting a random additive valuation X to T_N. In particular we will have to deal with the sequences of random variables $(R_N)_{N\geq 0}$ and $(L_N)_{N\geq 0}$. The definition of the filtrations $\mathbb{F}_N$ will be given shortly.

According to (2) we call the sequence of random variables $(R_N)_{N\geq 0}$ the *generating sequence of the valuation L.* The sequence of random variables $(L_N)_{N\geq 0}$ can now be defined by the following system of equalities in distribution:

$$L_N \stackrel{\mathcal{D}}{=} \begin{cases} R_N, & N < n_0 \\ R_N + L_{N'} + \bar{L}_{N''}, & N \geq n_0, \end{cases} \tag{5}$$

where N', N'' are random variables with joint distribution $P_N(N' = k', N'' = k'') = p_{N,k',k''}$ satisfying (4), $L_k \stackrel{\mathcal{D}}{=} \bar{L}_k$ for $k \geq 0$, and $R_N, L_{N'}, \bar{L}_{N''}$ are independent, conditional on $\{N', N''\}$.

Moments of additive valuations. Equations (5) can be used to obtain recurrence relations for the moments of L_N. It is easy to see that $\mathbb{E}\,|L_N|^m < \infty$ for $N \geq 0$ is implied by $\mathbb{E}\,|R_N|^m < \infty$ for $N \geq 0$:

Assume that $m \geq 1$ (the case $0 < m < 1$ can be treated similarly). If $\pi_N = 1$ it is easy to deduce from (5) that

$$\mathbb{E}\,|L_N|^m \leq \begin{cases} \mathbb{E}\,|R_N|^m, & N < n_0 \\ 3^{m-1}\left(\mathbb{E}\,|R_N|^m + 2\max\limits_{0\leq k<N}\mathbb{E}\,|L_k|^m\right), & N \geq n_0, \end{cases}$$

and this furnishes a proof by induction on N for $\mathbb{E}\,|L_N|^m < \infty$ for $N \geq 0$. In the case $\pi_N < 1$ we define $I(\mathsf{t}) := \{i : |\mathsf{t}^{(i)}| = |\mathsf{t}|\}$, $I_\ell(\mathsf{t}) := \{i \in I(\mathsf{t}) : |\mathsf{t}_\ell^{(i)}| < |\mathsf{t}|\}$ and $I_r(\mathsf{t}) := \{i \in I(\mathsf{t}) : |\mathsf{t}_r^{(i)}| < |\mathsf{t}|\}$, and obtain

$$L(\mathsf{t}) = \sum_{i\in I(\mathsf{t})} R^{(i)}(\mathsf{t}^{(i)}) + \sum_{i\in I_\ell(\mathsf{t})} L(\mathsf{t}_\ell^{(i)}) + \sum_{i\in I_r(\mathsf{t})} L(\mathsf{t}_r^{(i)}). \tag{6}$$

Now $K := |I(\mathsf{t})| - 1$ is geometrically distributed with parameter π_N, and $|I_\ell(\mathsf{t}) \cup I_r(\mathsf{t})| = |I(\mathsf{t})| + 1$. Moreover, in the first sum of (6) all but one of the terms $R^{(i)}(\mathsf{t}^{(i)})$ have the conditional distribution of R_N given $N' \vee N'' = N$, and the remaining term has the conditional distribution of R_N given $N' \vee N'' < N$. Since

$$\begin{aligned} \mathbb{E}\left[|R_N|^m \middle| N' \vee N'' = N\right] &\leq \frac{\mathbb{E}\,|R_N|^m}{1-\pi_N}, \\ \mathbb{E}\left[|R_N|^m \middle| N' \vee N'' < N\right] &\leq \frac{\mathbb{E}\,|R_N|^m}{\pi_N} \qquad\qquad \text{and} \\ \frac{1}{1-\pi_N} \vee \frac{1}{\pi_N} &< \frac{1}{\pi_N - \pi_N^2}, \end{aligned}$$

finiteness of $\mathbb{E}\,|L_N|^m$ follows again by induction on N from

$$\mathbb{E}\,|L_N|^m \leq \begin{cases} \mathbb{E}\,|R_N|^m, & N < n_0 \\ 2^{m-1}\mathbb{E}\,(K+2)^m\left(\frac{1}{\pi_N-\pi_N^2}\mathbb{E}\,|R_N|^m + \max\limits_{0\leq k<N}\mathbb{E}\,|L_k|^m\right), & N \geq n_0. \end{cases}$$

The filtrations $\mathbb{F}_N$. The filtration $\mathbb{F}_N = \{\mathcal{F}_{N,i}, i \geq 0\}$ is defined by $\mathcal{F}_{N,0} = \{\emptyset, \mathsf{T}_N \times \Omega\}$, and for $i \geq 1$ by

$$\mathcal{F}_{N,i} = \sigma\{|\mathsf{t}_\ell^{(j)}|, |\mathsf{t}_r^{(j)}|, R^{(j)}(\mathsf{t}^{(j)}); 1 \leq j \leq i\},$$

where we define $R^{(j)}(\mathsf{t}^{(j)}) \equiv 0$ for $|\mathsf{t}^{(j)}| = -1$. Note that $|\mathsf{t}_\ell^{(j)}|, |\mathsf{t}_r^{(j)}|$ and $R^{(j)}(\mathsf{t}^{(j)})$ are measurable functions on $(\mathsf{T}_N \times \Omega, \mathcal{F}_N \times \mathcal{G})$ for $j \geq 1$.

A martingale with terminal value $L_N - \mathbb{E}\, L_N$. We assume $\mathbb{E}\, R_N^2 < \infty$ for $N \geq 0$, thus $r_N := \mathbb{E}\, R_N$, $\ell_N := \mathbb{E}\, L_N$ and $v_N := \operatorname{Var} L_N$ are all finite. We want to represent $L_N - \ell_N$ as the terminal value of some martingale. This is possible since the random variable $L_N - \ell_N$ is absolutely integrable (and has even finite second moment, due to our assumption on R.) One (and a very easy) way to do this is to take the sequence of conditional expectations with respect to the elements of a filtration. So let us consider the sequence $(X_{N,i})_{i\geq 0}$, which is a martingale with respect to the filtration $\mathbb{F}_N$, defined by

$$X_{N,j} = \sum_{i=0}^{j} \lambda_{N,i}, \tag{7}$$

where the random variables $\lambda_{N,i} = \lambda_{N,i}(\mathsf{t})$ (dependencies on ω are always suppressed) are given by $\lambda_{N,0} = \ell_N$ and for $i \geq 1$ by

$$\begin{aligned} \lambda_{N,i} &= \mathbb{E}\,[L_N | \mathcal{F}_{N,i}] - \mathbb{E}\,[L_N | \mathcal{F}_{N,i-1}] \\ &= \begin{cases} 0, & \text{if } |\mathsf{t}^{(i)}| = -1 \\ R^{(i)}(\mathsf{t}^{(i)}) - \ell_{|\mathsf{t}^{(i)}|}, & \text{if } 0 \leq |\mathsf{t}^{(i)}| < n_0 \\ R^{(i)}(\mathsf{t}^{(i)}) + \ell_{|\mathsf{t}_\ell^{(i)}|} + \ell_{|\mathsf{t}_r^{(i)}|} - \ell_{|\mathsf{t}^{(i)}|}, & \text{if } |\mathsf{t}^{(i)}| \geq n_0. \end{cases} \end{aligned} \tag{8}$$

Since L_N is measurable with respect to $\sigma(\bigcup_{i\geq 0} \mathcal{F}_{N,i})$ and since $\mathbb{E}\, L_N^2 < \infty$, we have $X_{N,j} \to L_N$, $P_N \times Q$-a.s. and in $L^2(\mathsf{T}_N \times \Omega, \mathcal{F}_N \times \mathcal{G}, P_N \times Q)$, as $j \to \infty$, by P. Lévy's theorem (cf. [35, pp. 111, 134]).

We are now going to derive recurrence relations for expectations and variances of L_N. Fixing $i = 1$ and $N \geq n_0$, (8) is simply

$$\lambda_{N,1} = R_N + \ell_{N'} + \ell_{N''} - \ell_N,$$

where N', N'' are random variables with joint distribution $P_N(N' = k', N'' = k'') = p_{N,k',k''}$. Of course $\mathbb{E}\,[\lambda_{N,1} | \mathcal{F}_{N,0}] = 0$, and this yields the following recurrence for the sequence $(\ell_N)_{N\geq 0}$

$$\ell_N = \begin{cases} r_N, & N < n_0 \\ r_N + \sum_{k',k''} p_{N,k',k''}(\ell_{k'} + \ell_{k''}), & N \geq n_0. \end{cases} \tag{9}$$

(The condition $\pi_N > 0$ for $N \geq n_0$ ensures that (9) can be uniquely solved for $(\ell_N)_{N\geq 0}$.) A similar recurrence is obtained for the sequence $(v_N)_{N\geq 0}$: We define

$$s_N := \mathbb{E}\,[\lambda_{N,1}^2 | \mathcal{F}_{N,0}] = \begin{cases} \operatorname{Var} R_N, & N < n_0 \\ \mathbb{E}\,\left(R_N + \ell_{N'} + \ell_{N''} - \ell_N\right)^2, & N \geq n_0. \end{cases} \tag{10}$$

By squaring the equations

$$L_N - \ell_N \overset{\mathcal{D}}{=} \begin{cases} \lambda_{N,1}, & N < n_0 \\ \lambda_{N,1} + L_{N'} - \ell_{N'} + \bar{L}_{N''} - \ell_{N''}, & N \geq n_0 \end{cases}$$

and carefully exploiting independence when computing expectations, we obtain

$$v_N = s_N + \mathbb{I}_{\{N \geq n_0\}} \sum_{k',k''} p_{N,k',k''}(v_{k'} + v_{k''}). \tag{11}$$

Sufficient conditions for asymptotic normality of $\frac{L_N-\ell_N}{\sqrt{v_N}}$. Assuming $v_N > 0$ for all sufficiently large N, we can define a martingale difference array $\{\xi_{N,i}\,,\,\mathcal{F}_{N,i}\}_{i\geq 0,N\geq 0}$ by

$$\xi_{N,i} := \frac{\lambda_{N,i}}{\sqrt{v_N}}. \tag{12}$$

Now $\frac{L_N-\ell_N}{\sqrt{v_N}} = \sum_{i=1}^{\infty}\xi_{N,i}$, $P_N\times Q$-a.s., thus by a basic central limit theorem for martingale difference arrays (cf. [33, p. 543, Theorem 4]) $\frac{L_N-\ell_N}{\sqrt{v_N}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$ will follow from the "conditional normalizing condition"

$$\sum_{i=1}^{\infty}\mathbb{E}\,[\xi_{N,i}^2|\mathcal{F}_{N,i-1}] \xrightarrow{P} 1, \quad \text{as } N\to\infty, \tag{No}$$

and the "conditional Lindeberg condition"

$$\sum_{i=1}^{\infty}\mathbb{E}\,[\xi_{N,i}^2\,\mathbb{I}_{\{|\xi_{N,i}|>\varepsilon\}}\big|\mathcal{F}_{N,i-1}] \xrightarrow{P} 0, \quad \text{as } N\to\infty, \text{ for each } \varepsilon>0. \tag{Li}$$

In order to obtain bounds on the convergence rate in the central limit theorem we might rather want to verify some stronger Lyapunov type conditions instead. In (No), convergence in probability is implied by convergence in L^a, for some $a>0$, and (Li) is implied by convergence to 0 in L^1 of $\sum_{i=1}^{\infty}\mathbb{E}\,[\xi_{N,i}^{2a}\big|\mathcal{F}_{N,i-1}]$, for some $a>1$, (because of $x^2\,\mathbb{I}_{\{|x|>\varepsilon\}} \leq \varepsilon^{2-2a}|x|^{2a}$,) yielding conditions (No$_a$) and (Li$_a$). The following lemma builds upon these observations, i.e. the conditions (No$_2$) and (Li$_a$) will be expressed as asymptotic relations between three sequences, which all satisfy the same type of linear equation that is also the defining equation for the sequence $(\ell_N)_{N\geq 0}$.

Lemma 1. *Let $(L_N)_{N\geq 0}$ be the sequence of random variables defined by equation* (5) *in terms of the sequence of random variables $(R_N)_{N\geq 0}$, which we assume to satisfy $\mathbb{E}\,|R_N|^{2a}<\infty$ for $N\geq 0$ and some $a>1$. Let moreover $v_N>0$ for all sufficiently large N. We define sequences $(\sigma_N)_{N\geq 0}$, $(s_N^{(a)})_{N\geq 0}$ and recursively define sequences $(w_N)_{N\geq 0}$, $(v_N^{(a)})_{N\geq 0}$ by*

$$\sigma_N = \mathbb{I}_{\{N\geq n_0\}}\sum_{k',k''} p_{N,k',k''}\,(s_N+v_{k'}+v_{k''}-v_N)^2, \tag{13}$$

$$w_N = \sigma_N + \mathbb{I}_{\{N\geq n_0\}}\sum_{k',k''} p_{N,k',k''}(w_{k'}+w_{k''}), \tag{14}$$

$$s_N^{(a)} = \mathbb{E}\,|\lambda_{N,1}|^{2a}, \tag{15}$$

$$v_N^{(a)} = s_N^{(a)} + \mathbb{I}_{\{N\geq n_0\}}\sum_{k',k''} p_{N,k',k''}(v_{k'}^{(a)}+v_{k''}^{(a)}). \tag{16}$$

Then $\frac{L_N-\ell_N}{\sqrt{v_N}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$ is implied by

$$w_N = o(v_N^2), \quad \text{as } N\to\infty, \tag{No$_2$}$$

$$v_N^{(a)} = o(v_N^a), \quad \text{as } N\to\infty. \tag{Li$_a$}$$

Proof. Our first observation is that

$$V(\mathsf{t}) := \sum_{i=1}^{\infty} \mathbb{E}\,[\lambda^2_{|\mathsf{t}|,i} | \mathcal{F}_{|\mathsf{t}|,i-1}] = \sum_{i:\,|\mathsf{t}^{(i)}|\geq 0} s_{|\mathsf{t}^{(i)}|} = \begin{cases} s_{|\mathsf{t}|}, & |\mathsf{t}| < n_0, \\ s_{|\mathsf{t}|} + V(\mathsf{t}_\ell) + V(\mathsf{t}_r), & |\mathsf{t}| \geq n_0, \end{cases}$$

is a (deterministic) valuation of the additive type (2), generated by $s_{|\mathsf{t}|}$, with $\mathbb{E}\,V_N = \operatorname{Var} L_N$. The second equality is verified noting that the random variable $|t^{(i)}|$, defined on T_N, generates a σ-algebra $\sigma(|t^{(i)}|) \subseteq \mathcal{F}_{N,i-1}$ and that, defining $\lambda_{-1,1} \equiv 0$, we have $P_N \times Q$-a.s.

$$\mathbb{P}\,(\lambda_{N,i} \leq x | \mathcal{F}_{N,i-1}) = \mathbb{P}\,(\lambda_{N,i} \leq x|\, |t^{(i)}|) = \mathbb{P}\,(\lambda_{|\mathsf{t}^{(i)}|,1} \leq x|\, |t^{(i)}|) = F_{|t^{(i)}|}(x), \tag{17}$$

where $F_N(x) := \mathbb{P}\,(\lambda_{N,1} \leq x)$ for $N \geq -1$. For the third equality we note that the multiset $M(\mathsf{t}) := \{|\mathsf{t}^{(i)}| \geq 0 : i \geq 1\}$ can be decomposed as $M(\mathsf{t}) = \{|\mathsf{t}|\} \cup M(\mathsf{t}_\ell) \cup M(\mathsf{t}_r)$. Moreover V_N is the terminal value of the predictable quadratic variation process of the martingale $(X_{N,i})_{i\geq 0}$, thus $\mathbb{E}\,V_N = \operatorname{Var} L_N$ indeed holds, cf. [33].

Similarly we construct another additive valuation $W(\mathsf{t})$, generated by some deterministic valuation $\sigma_{|\mathsf{t}|}$, such that $w_N := \mathbb{E}\,W_N = \operatorname{Var} V_N$. The definition (13) of the sequence $(\sigma_N)_{N\geq 0}$ is obtained by just mimicking (10), and (14) is the system of equations analogous to (11) that determines the sequence $(w_N)_{N\geq 0}$. Now (No_2) is just another way of writing $\mathbb{E}\,\left(\frac{V_N - v_N}{v_N}\right)^2 \to 0$, which implies (No). Furthermore

$$V_N^{(a)} := \sum_{i=1}^{\infty} \mathbb{E}\,[|\lambda_{N,i}|^{2a} | \mathcal{F}_{N,i-1}]$$

again corresponds to an additive valuation $V^{(a)}(\mathsf{t})$ of type (2), which is generated by the deterministic valuation $s^{(a)}_{|\mathsf{t}|} := \mathbb{E}\,|\lambda_{|\mathsf{t}|,1}|^{2a}$. The system of equations (16) thus determines the sequence $(v_N^{(a)})_{N\geq 0}$, where $v_N^{(a)} := \mathbb{E}\,V_N^{(a)}$. Again ($\mathrm{Li}_a$) is just another way to write $\mathbb{E}\,\frac{V_N^{(a)}}{v_N^a} \to 0$, which implies (Li). □

Remark 1. The nice thing about this lemma is that it provides sufficient conditions for asymptotic normality that are entirely expressed in terms of solutions of a certain system of linear equations that already showed up in (9) and (11). Putting bold lowercase letters for sequences and denoting by $\mathbf{I} = (\mathbf{I}_{N,k})_{N,k\geq 0}$ the infinite identity matrix and by $\mathbf{P} = (\mathbf{P}_{N,k})_{N,k\geq 0}$ the infinite matrix satisfying

$$\mathbf{P}_{N,k} = \begin{cases} 0, & N < n_0 \text{ or } k > N \\ \sum_{0\leq k'\leq N} (p_{N,k,k'} + p_{N,k',k}), & \text{else}, \end{cases} \tag{18}$$

these systems can be written as

$$(\mathbf{I} - \mathbf{P})\boldsymbol{\ell} = \mathbf{r}, \quad (\mathbf{I} - \mathbf{P})\mathbf{v} = \mathbf{s}, \quad (\mathbf{I} - \mathbf{P})\mathbf{w} = \boldsymbol{\sigma}, \quad (\mathbf{I} - \mathbf{P})\mathbf{v}^{(a)} = \mathbf{s}^{(a)}. \tag{19}$$

Often only asymptotic equivalents of the sequences $\mathbf{r}, \boldsymbol{\ell}, \mathbf{s}$ and $\mathbf{v}$ will be needed to obtain asymptotic equivalents of the sequences $\boldsymbol{\sigma}$ and $\mathbf{s}^{(a)}$. Knowing asymptotic equivalents of the right hand sides in (19) will often be enough to obtain asymptotics of the corresponding solutions. Master Theorems are around that deal with such questions, cf. [31].

The use of (No$_2$) and (Li$_2$) has another advantage: We get bounds on convergence rates for free! By results of Heyde and Brown [13] and Haeusler [12] there is a constant C_2 such that

$$\sup_{x\in\mathbb{R}}\left|P_N\left(\tfrac{L_N-\ell_N}{\sqrt{v_N}}\le x\right)-\Phi(x)\right|\le C_2\left(\frac{v_N^{(2)}+w_N}{v_N^2}\right)^{\frac{1}{5}}. \tag{20}$$

Also large deviations results in terms of $\frac{v_N^{(2)}+w_N}{v_N^2}$ can be obtained, cf. Grama [10].

Of course we could have formulated Lemma 1 using conditions (No$_b$) and (Li$_a$) for some $b>0$ and $a>1$. Now (No$_b$) would be: $w_N^{(b)}=o(v_N^b)$, as $N\to\infty$, where $w_N^{(b)}$ is defined in terms of

$$\sigma_{|\mathsf{t}|}^{(b)}:=\mathbb{E}\left[|V(\mathsf{t})-v_{|\mathsf{t}|}|^b-|V(\mathsf{t}_\ell)-v_{|\mathsf{t}_\ell|}|^b-|V(\mathsf{t}_r)-v_{|\mathsf{t}_r|}|^b\right],$$

which is a nice expression in the "well known" sequences **s** and **v** only when $b=2$. Verifying (Li$_b$) and the "unpleasant" (No$_b$) for some $b\neq 2$, $b>1$ would however be rewarded with a version of (20), with right hand side $C_b\left((v_N^{(b)}+w_N^{(b)})/v_N^b\right)^{1/(1+2b)}$, cf. [13, 12].

Note that (Li$_a$) imposes additional integrability conditions on the random variables R_N in the sense that $s_N^{(a)}<\infty$ (and thus $v_N^{(a)}<\infty$) for $N\ge 0$ only if $\mathbb{E}\,|R_N|^{2a}<\infty$ for $N\ge 0$. On the other hand, concerning (No$_2$), $w_N<\infty$ as long as $v_N<\infty$ for $N\ge 0$.

Remark 2. Generalizations of this approach to additive valuations on m-ary trees for $m>2$, and even to the case where there is no upper bound for the degrees (i.e., no upper bound for the number of primary subproblems) seem to be straight forward. Extensions to multivariate limiting distributions and asymmetric valuations of the form $L(\mathsf{t})=R(\mathsf{t})+\mathbb{I}_{\{|\mathsf{t}|\ge n_0\}}(aL(\mathsf{t}_\ell)+bL(\mathsf{t}_r))$, where the case $(a,b)=(1,0)$ is perhaps the most interesting, also seem to be within reach. Moreover one can think of allowing a wider class of probabilistic models by resigning the condition that the distribution of a subtree t^* of some tree t depends only on its size $|\mathsf{t}^*|$ and not on the position of its root in the tree t, which would necessitate introducing a whole family of valuations, indexed by the nodes of the infinite binary tree.

3 Deterministic additive valuations of the trie data structure

We are now going to demonstrate the strength of Lemma 1 by proving asymptotic normality of a large class of additive valuations of the trie data structure. This class includes some of the most important characteristics of the trie data structure, such as the number of its nodes, its external path length, or the number of its external internal nodes, which give clues on the space requirements and the time complexity of associated update operations. We now give concise descriptions of tries and the probabilistic models we are going to use.

Binary tries. The trie (cf. [9, 19, 21]) is designed to store data which have keys that are given as sequences over a finite alphabet Σ. Here we confine ourselves to the binary trie, i.e. the case $\Sigma=\{0,1\}$. Now let a set $S=\{k(i)\in\Sigma^\infty : 1\le i\le N\}$ of keys be given. The trie built from these keys is a binary tree, whose internal nodes serve as branching nodes. Each leaf (external node) either stores one key or is empty. If we label in this tree each edge to the left (resp. right) 0 (resp. 1), we obtain an encoding of the leaves by taking the 0-1-sequence along the path starting from the root. A key k_i is stored in the leaf encoded by k_i's minimal unique prefix among the N keys in S. Note that the order of the keys is irrelevant in this construction, and that different sets S, S' may lead to the same trie t. The set of all tries t built from N distinct keys is denoted T_N, and $|\mathsf{t}| = N$ is said to be the "size" of t. To be in accordance with the notion of size introduced in Section 2, we let $|\mathsf{t}| = 0$ if t is a single empty leaf, and $|\mathsf{t}| = -1$ if $\mathsf{t} = \emptyset$. Moreover we let $\mathsf{T} = \bigcup_{N\geq 0} \mathsf{T}_N$ be the set of all tries. Left and right subtrees of a trie t are denoted t_ℓ, t_r. Of course, t_ℓ then denotes the trie, which is built from the keys with the first bit 0 dropped. It is easily seen that the sets T_N are countably infinite for $N \geq 2$. Note that a trie of size k typically has more than $k-1$ internal nodes. The additional internal nodes are caused by one-way branchings, i.e., are those internal nodes with one child an empty leaf.

The Bernoulli models. We will assume that $\mathsf{t} \in \mathsf{T}_N$ is constructed from an i.i.d. sequence of keys $(k(i))_{1\leq i\leq N}$ where each key $k(i) = (k_1(i), k_2(i), \ldots)$ constitutes an i.i.d. sequence of bits with

$$\mathbb{P}\,(k_1(1) = 0) = p \quad \text{and} \quad \mathbb{P}\,(k_1(1) = 1) = 1 - p =: q.$$

The case $p = \frac{1}{2}$ resp. $p \neq \frac{1}{2}$ is called symmetric resp. asymmetric Bernoulli model. We deal with the probability space $(\mathsf{T}_N, \mathcal{F}_N, \mathbb{F}_N, P_N)$, where $\mathcal{F}_N$ is the set of all subsets of T_N, and P_N is defined with the help of the splitting probabilities

$$p_{N,k} := \mathbb{P}\left(|\mathsf{t}_\ell| = k \,\middle|\, |\mathsf{t}| = N\right) = \binom{N}{k} p^k q^{N-k},$$

i.e., given $|\mathsf{t}|$, the random variable $|\mathsf{t}_\ell|$ follows the binomial distribution $B(|\mathsf{t}|, p)$.

There is exactly one trie of size 0 and one of size 1, thus $P_{|\mathsf{t}|}(\mathsf{t}) = 1$ for $|\mathsf{t}| \leq 1$, and for $|\mathsf{t}| \geq 2$ we have $P_{|\mathsf{t}|}(\mathsf{t}) = p_{|\mathsf{t}|,|\mathsf{t}_\ell|} P_{|\mathsf{t}_\ell|}(\mathsf{t}_\ell) P_{|\mathsf{t}_r|}(\mathsf{t}_r)$. Obviously

$$P_N(|\mathsf{t}_\ell| \vee |\mathsf{t}_r| \leq N) = 1 = P_N(|\mathsf{t}_\ell| \wedge |\mathsf{t}_r| < N) \quad \text{and} P_N(|\mathsf{t}_\ell| \vee |\mathsf{t}_r| < N) = 1 - p^N - q^N > 0$$

hold for $N \geq n_0$, thus all requirements made on a probabilistic model in section 2 are fulfilled by the Bernoulli models.

The definition of the filtration $\mathbb{F}_N = \{\mathcal{F}_{N,i}, i \geq 0\}$ can now be slightly simplified:

$$\begin{aligned} &\mathcal{F}_{N,0} = \{\emptyset, \mathsf{T}_N\} \quad \text{and} \\ &\mathcal{F}_{N,i} = \sigma\{|\mathsf{t}_\ell^{(j)}|; 1 \leq j \leq i\} \quad \text{for} \quad i \geq 1. \end{aligned}$$

Additive valuations of tries. A valuation on the family of tries T is any function $X : \mathsf{T} \to \mathbb{R}$. We shall concentrate on the particular class of additive valuations L, which can for some $n_0 \geq 2$ be described by

$$L(\mathsf{t}) = \begin{cases} R(\mathsf{t}), & |\mathsf{t}| < n_0, \\ R(\mathsf{t}) + L(\mathsf{t}_\ell) + L(\mathsf{t}_r), & |\mathsf{t}| \geq n_0,\ \mathsf{t} = \overset{\circ}{\underset{\mathsf{t}_\ell \quad \mathsf{t}_r}{/\ \backslash}}, \end{cases} \tag{21}$$

where R is a deterministic valuation, which is constant on each set T_N for $N \geq n_0$ and $N \in \{0,1\}$, so that we may define $r_{|\mathsf{t}|} = R(\mathsf{t})$ for $|\mathsf{t}| \geq n_0$ and $|\mathsf{t}| \in \{0,1\}$. However R may for $2 \leq N < n_0$ depend on $\mathsf{t} \in \mathsf{T}_N$, in which case we denote $r_N := \mathbb{E}\,[R(\mathsf{t}) | \mathsf{t} \in \mathsf{T}_N]$ (later on we will impose integrability conditions on $R|\mathsf{T}_N$, in particular expectations will always be finite).

For example, the number $L(\mathsf{t})$ of internal nodes of a trie t is a valuation of this form with $n_0 = 2$ and $R(\mathsf{t}) = \mathbb{I}_{\{|\mathsf{t}|>1\}}$, as well as the number of internal external nodes [8] ($n_0 = 3$, $R(\mathsf{t}) = \mathbb{I}_{\{|\mathsf{t}|=2\}}$) and the external path length ($n_0 = 2$, $R(\mathsf{t}) = \mathbb{I}_{\{|\mathsf{t}|>1\}}|\mathsf{t}|$). Counting certain exotic subtrees of a trie t is also possible, e.g. counting subtrees of size 6 with identical subtrees can be achieved with ($n_0 = 7$, $R(\mathsf{t}) = \mathbb{I}_{\{|\mathsf{t}|=6,\mathsf{t}_\ell=\mathsf{t}_r\}}$).

Demanding $R(\mathsf{t}) = 0$ for $|\mathsf{t}| \in \{0,1\}$ would not be a serious restriction, since ($n_0 = 2$, $R(\mathsf{t}) = \mathbb{I}_{\{|\mathsf{t}|=1\}}$) leads to the size $L(\mathsf{t}) = |\mathsf{t}|$, which is constant on each T_N, and ($n_0 = 2$, $R(\mathsf{t}) = \mathbb{I}_{\{|\mathsf{t}|\in\{0,1\}\}}$) leads to $L(\mathsf{t})$, which is 1 plus the number of internal nodes of t. Thus for any additive valuation L defined by (21) (with $r_0 = d$, $r_1 = c+d$) there is another additive valuation L' defined by (21) in terms of a valuation R' (with $r'_0 = r'_1 = 0$) and satisfying $L'(\mathsf{t}) = L(\mathsf{t}) - c|\mathsf{t}| - d$, cf. also (26).

We now repeat some notation from section 2: Restricting the valuations R and L to the sets T_N, we obtain sequences of random variables $(R_N)_{N\geq 0}$ and $(L_N)_{N\geq 0}$. The sequence of random variables $(L_N)_{N\geq 0}$ can now be defined by the following system of equalities in distribution:

$$L_N \stackrel{\mathcal{D}}{=} \begin{cases} R_N, & N < n_0 \\ R_N + L_{N'} + \bar{L}_{N-N'}, & N \geq n_0, \end{cases} \tag{22}$$

where N' is a random variable with the binomial distribution $B(N,p)$, $L_k \stackrel{\mathcal{D}}{=} \bar{L}_k$ for $k \geq 0$, moreover $L_{N'}, \bar{L}_{N-N'}$ are independent, conditional on N', and $R_N = r_N$ is deterministic for $N \geq n_0$. Assuming $\mathbb{E}\,R_N^2 < \infty$ for $2 \leq N < n_0$ ensures that first and second moments of L_N are finite. We recall $r_N = \mathbb{E}\,R_N$ and moreover denote $\ell_N := \mathbb{E}\,L_N$ and $v_N := \operatorname{Var} L_N$. Equations (22) can be used to obtain recurrence relations for the first and second moments of L_N:

Denoting sequences by bold face lower case letters, these are

$$(\mathbf{I} - \mathbf{M}_p - \mathbf{M}_q)\boldsymbol{\ell} = \mathbf{r}, \qquad (\mathbf{I} - \mathbf{M}_p - \mathbf{M}_q)\mathbf{v} = \mathbf{s}, \tag{23}$$

where $\mathbf{I}$ is the infinite identity matrix, the matrix $\mathbf{M}_p$ is defined by

$$(\mathbf{M}_p)_{N,k} := \begin{cases} 0, & \text{for } N < n_0 \\ \binom{N}{k} p^k (1-p)^{N-k}, & \text{for } N \geq n_0, \end{cases} \tag{24}$$

and the sequence $\mathbf{s}$ is defined by

$$s_N := \begin{cases} \operatorname{Var} R_N, & N < n_0 \\ \sum_{k=0}^{N} p_{N,k} \left(\ell_k + \ell_{N-k} - \sum_{\kappa=0}^{N} p_{N,\kappa}(\ell_\kappa + \ell_{N-\kappa})\right)^2, & N \geq n_0. \end{cases} \tag{25}$$

It is easily seen that $\mathbf{s}$ is a sequence of nonnegative terms, and it was shown in [32, Theorem 1] that $\mathbf{s} \equiv 0$ only if R is of the special form

$$R(\mathsf{t}) = \begin{cases} c|\mathsf{t}| + d, & \text{for } |\mathsf{t}| < n_0, \\ -d, & \text{for } |\mathsf{t}| \geq n_0, \end{cases} \tag{26}$$

for some $c, d \in \mathbb{R}$. In this case $L(\mathsf{t}) = c|\mathsf{t}| + d$ and $\operatorname{Var} L_N \equiv 0$. Moreover, [32, Theorem 1] tells us that $\operatorname{Var} L_N = \Omega(N)$, if R is not of the form (26).

Theorem 1. *Let L_N be the random variable corresponding to the additive valuation L, defined by* (21) *on the space $\mathcal{T}_N$ equipped with the Bernoulli model. Assume that R is not of the form* (26)*. Let N' be a random variable with binomial distribution $B(N,p)$. Let moreover $\mathbf{r}' := (\mathbf{M}_p + \mathbf{M}_q)\mathbf{r}$, i.e. $r'_N = \mathbb{E}\,(r_{N'} + r_{N-N'})$, and $\mathbf{r}'' := (\mathbf{M}_p + \mathbf{M}_q)^2\mathbf{r}$.*
If either of

i) $\Delta^2 r''_N = o\left(\frac{1}{\sqrt{N}}\right)$, *for* $p = \frac{1}{2}$,

ii) $\Delta r'_N = o(1)$, *for* $\frac{\ln p}{\ln q} \in \mathbb{Q}$,

iii) $\Delta r'_N = o\left(\frac{1}{\sqrt{\ln N}}\right)$, *for* $\frac{\ln p}{\ln q} \notin \mathbb{Q}$,

and for some $a > 1$ both

iv) $\mathbb{E}\,|R_N|^{2a} < \infty$ *for* $2 \le N < n_0$,

v) $\mathbb{E}\,|r_{N'} + r_{N-N'} - r'_N|^{2a} = o(N^a)$ *as* $N \to \infty$,

are satisfied, then

$$\frac{L_N - \mathbb{E}\,L_N}{\sqrt{\operatorname{Var} L_N}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \quad \text{as } N \to \infty. \tag{27}$$

Corollary 1. *If $p = \frac{1}{2}$, then $\Delta^2 r_N = o\left(N^{-\frac{1}{2}}\right)$ implies conditions i) and v) of Theorem 1.*

If $\frac{\ln p}{\ln q} \in \mathbb{Q}$, then either of $r_N = o(\sqrt{N})$ and $\Delta r_N = o(1)$ implies conditions ii) and v) of Theorem 1.

If $\frac{\ln p}{\ln q} \notin \mathbb{Q}$, then either of $r_N = o\left(\sqrt{\frac{N}{\ln N}}\right)$ and $\Delta r_N = o\left(\frac{1}{\sqrt{\ln N}}\right)$ implies conditions iii) and v) of Theorem 1.

Remark 3. Taking condition *iv)* of Theorem 1 for granted, some sequences $\mathbf{r}$ that lead to (27) are

$r_N = (-1)^N N^{0.49}$,
$r_N = N^{0.99} + (-1)^N N^{-0.01}$ and
$r_N = N^{\alpha} f(N^{\beta})$, where $f : \mathbb{R} \to \mathbb{R}$ is a bounded function with bounded derivative and $\alpha, \beta > 0$ and $\alpha + \beta < 1$.

This can be easily checked applying Corollary 1. There are of course sequences that directly call for Theorem 1, for example "lacunary" sequences such as $r_N = N^{0.749}\,\mathbb{I}_{\{\sqrt{N} \in \mathbb{N}\}}$, which satisfies condition *v)* of Theorem 1 with $a = 1.001$. A simple observation is the following:

If for n_0, p fixed two sequences $\mathbf{r}$, $\bar{\mathbf{r}}$ satisfy the conditions of Theorem 1, so does any linear combination $a\mathbf{r} + b\mathbf{r}'$, which is not of the form (26).

Remark 4. Asymptotic normality of the number of internal nodes in a binary trie under the Bernoulli model ($n_0 = 2$, $r_N = \mathbb{I}_{\{N \ge 2\}}$) was first proved by Jacquet and Regnier [14, 15], as well as convergence of moments of any order. Employing contraction properties of suitably chosen probability metrics, Rachev and Rüschendorf [26] and Feldman, Rachev and Rüschendorf [4] proved asymptotic normality of L_N for the sequence $r_N = 1$ under very general probabilistic models, including the Bernoulli models, and remark that their analysis could, under certain conditions, be extended to sequences $r_N = o(\sqrt{N})$, cf. [26,

p. 787]. In the case of the Bernoulli models, these conditions boil down to demanding that the sequence $(\operatorname{Var} L_N)_{N\geq 0}$ is regularly varying of order 1, i.e.

$$\operatorname{Var} L_N = NG(N), \quad \text{where } G(tN)/G(N) \to 1 \text{ for all } t>0 \text{ as } N\to\infty.$$

Moreover there has to be $c>0$ such that $c<G<\frac{1}{c}$, i.e. G is bounded away from 0 and ∞, cf. [4, p. 172]. Corollary 1 deals with sequences $r_N = o(\sqrt{N})$. Unfortunately, in Theorem 1 and Corollary 1 we had to make a distinction according to whether $\frac{\ln p}{\ln q}\in\mathbb{Q}$ or $\frac{\ln p}{\ln q}\notin\mathbb{Q}$, because this distinction is essential in Proposition 2, which we use in the proof of Theorem 1. So the question, if $r_N = o(\sqrt{N})$ can be used in Corollary 1 also in the case $\frac{\ln p}{\ln q}\notin\mathbb{Q}$ remains an **open problem**, since it also cannot be decided using the sufficient condition from [26, 4]: One can check that $(n_0 = 2, r_N = (-1)^N\sqrt{\frac{N}{\ln N}}\mathbb{I}_{\{N\geq 3\}})$ results in $G(N)=\Theta(\ln\ln N)$, which is not bounded. On the other hand, we will provide an example (Example 2 in Section 4) showing that $r_N = O(\sqrt{N})$ (and also $\Delta r'_N = O(1)$ instead of conditions *ii)* and *iii)* of Theorem 1) does not imply (27).

Proof of Theorem 1. We refer to Lemma 1 and have thus to verify (No_2) and (Li_a). The sequences $\mathbf{w}$ and $\mathbf{v}^{(a)}$ are defined by the recurrence relations

$$(\mathbf{I}-\mathbf{M}_p-\mathbf{M}_q)\mathbf{w} = \boldsymbol{\sigma}, \qquad (\mathbf{I}-\mathbf{M}_p-\mathbf{M}_q)\mathbf{v}^{(a)} = \mathbf{s}^{(a)}, \tag{28}$$

with

$$\sigma_N := \mathbb{I}_{\{N\geq n_0\}}\sum_{k=0}^{N} p_{N,k}\left(v_k+v_{N-k}-\sum_{\kappa=0}^{N}p_{N,\kappa}(v_\kappa+v_{N-\kappa})\right)^2 \tag{29}$$

and

$$s_N^{(a)} := \begin{cases}\mathbb{E}\,|R_N-r_N|^{2a}, & N<n_0\\ \sum_{k=0}^N p_{N,k}\left|\ell_k+\ell_{N-k}-\sum_{\kappa=0}^N p_{N,\kappa}(\ell_\kappa+\ell_{N-\kappa})\right|^{2a}, & N\geq n_0.\end{cases} \tag{30}$$

We now collect some results concerning the solutions of equations of type (23) and (28) in the following two propositions. One of the proofs will be given in the appendix. The reader interested in obtaining asymptotic expansions of the sequences $\boldsymbol{\ell}$ and $\mathbf{v}$ is referred to [6, 18, 34].

Proposition 1. *Let us consider the following recurrence relation with matrices $\mathbf{M}_p$ defined in* (24) *and $a,b>0, 0<p\leq q<1$ (not necessarily $p+q=1$)*

$$(\mathbf{I}-a\mathbf{M}_p-b\mathbf{M}_q)\mathbf{y}=\mathbf{x}, \tag{31}$$

where $\mathbf{x}:=(x_k)_{k\geq 0}$ is a given sequence of real numbers. Let α denote the unique real zero of $f(s)=1-ap^s-bq^s$ and assume that $\alpha<n_0$. We now list some facts concerning the solution $\mathbf{y}=(y_N)_{N\geq 0}$ of (31)*, some facts concerning the effect of transforming a sequence $\mathbf{y}$ by $\mathbf{M}_p$ and computing differences, some elementary facts about Poisson generating functions and their interplay with the matrices $\mathbf{M}_p$, and a simple tool for depoissonization:*

(P1.1) *The map that takes $\mathbf{x}$ to $\mathbf{y}$ is linear. Moreover $(\mathbf{I}-a\mathbf{M}_p-b\mathbf{M}_q)^{-1}_{N,k}\geq 0$ for $N,k\geq 0$, therefore $x_k\geq 0$ for $k\geq 0$ implies $y_k\geq 0$ for $k\geq 0$.*

(P1.2) *We can represent* $\mathbf{y}$ *as*

$$\mathbf{y} = \mathbf{x} + \mathbf{y}',$$

where $\mathbf{y}'$ *is the solution of*

$$(\mathbf{I} - a\mathbf{M}_p - b\mathbf{M}_q)\mathbf{y}' = \mathbf{x}' := (a\mathbf{M}_p + b\mathbf{M}_q)\mathbf{x}.$$

Moreover $x'_0 = \ldots = x'_{n_0-1} = 0$. *(Using this decomposition one can take advantage of the smoothness, that* $\mathbf{y}'$ *inherits from* $\mathbf{x}'$, *cf.* (P1.6) *and* (P1.7)*.)*

(P1.3) *If* $x_k = O(k^{\alpha-\varepsilon})$ *for some* $\varepsilon \in \mathbb{R}$, *then*

$$y_N = \begin{cases} O(N^{\alpha}), & \text{if } \varepsilon > 0, \\ O(N^{\alpha}\ln N), & \text{if } \varepsilon = 0, \\ O(N^{\alpha-\varepsilon}), & \text{if } \varepsilon < 0. \end{cases}$$

If $x_k = o(k^{\alpha-\varepsilon})$ *for some* $\varepsilon \le 0$, *then*

$$y_N = \begin{cases} o(N^{\alpha}\ln N), & \text{if } \varepsilon = 0, \\ o(N^{\alpha-\varepsilon}), & \text{if } \varepsilon < 0. \end{cases}$$

(P1.4) *If* $x_k \ge 0$ *for all* $k \ge 0$ *and* $x_{k_0} \ne 0$ *for some* $k_0 \ge 0$, *then* $y_N = \Omega(N^{\alpha})$.

(P1.5) *If* $x_k = k^{\alpha}$ *for* $k \ge k_0 \ge 1$, *then* $y_N = \dfrac{N^{\alpha}\ln N}{ap^{\alpha}\ln\frac{1}{p} + bq^{\alpha}\ln\frac{1}{p}} + z_N$, *where* $z_N = O(N^{\alpha})$. *If* $x_k = \Theta(k^{\alpha})$, *then* $y_N = \Theta(N^{\alpha}\ln N)$.

(P1.6) *The differences* $\Delta\mathbf{y}$ *of a sequence* $\mathbf{y}$ *(recall* $\Delta y_k = y_{k+1} - y_k$*) and the differences of the sequence* $\mathbf{M}_p\mathbf{y}$ *are connected via* $\Delta\mathbf{M}_p\mathbf{y} = p\mathbf{M}_p\Delta\mathbf{y} + \mathbf{z}$, *where* $z_N = 0$ *for* $N \ge n_0$. *Thus* $\mathbf{x}, \mathbf{y}$ *from* (31) *satisfy*

$$(\mathbf{I} - ap\mathbf{M}_p - bq\mathbf{M}_q)\Delta\mathbf{y} = \Delta\mathbf{x} + \mathbf{z}$$

with some other sequence $\mathbf{z}$ *still fulfilling* $z_N = 0$ *for* $N \ge n_0$. *If* $x_0 = \ldots = x_{n_0-1} = 0$, *then* $z_N = 0$ *holds for* $N \ge 0$.

(P1.7) *Let a sequence* $\mathbf{y}$ *satisfy* $y_k = O(k^{\beta} f(k))$ *for some* $\beta \in \mathbb{R}$ *and slowly varying* f *(i. e.* $f(tN)/f(N) \to 1$ *for all* $t > 0$ *and* $N \to \infty$*). Let* m *be a nonnegative integer. Then the sequence* $\mathbf{z}$ *defined by* $\mathbf{z} = \Delta^m \mathbf{M}_p \mathbf{y}$ *satisfies*

$$z_N = O(N^{\beta - \frac{m}{2}} f(N)).$$

(P1.8) *Let a sequence* $\mathbf{y}$ *satisfy* $\Delta y_k = o(k^{\beta})$ *for some* $\beta \in \mathbb{R}$ *and let* $y(t) := \sum_{k\ge 0} y_k \frac{t^k}{k!} e^{-t}$ *be the Poisson generating function of* $\mathbf{y}$. *Then*

$$y_N = y(N) + o(N^{\beta + \frac{1}{2}}).$$

If $y_0 = \ldots = y_{n_0-1} = 0$, $\mathbf{z} = \mathbf{M}_p\mathbf{y}$ *and* $z(t)$ *is the Poisson generating function of* $\mathbf{z}$, *then* $z(t) = y(pt)$.

Proof. The O-part of (P1.3), (P1.4) and (P1.5) are proved in [32, Lemma 1], (P1.6) and (P1.7) (without the slowly varying function) are proved in [32, Lemma 2]. The o-part of (P1.3) and (P1.7) can be proved by adapting the proofs of [32]. The assertions made in (P1.8) partly follow from [32, eq. (2.13) and (2.15)]. The remaining assertions are either completely obvious or rely on very simple properties of binomial coefficients. □

Proposition 2. *Let* $0 < p < q = 1-p$, *let* $\mathbf{x}$ *be a given sequence and* $\mathbf{y} = (\mathbf{I} - p\mathbf{M}_p - q\mathbf{M}_q)^{-1}\mathbf{x}$. *Then* $\Delta x_N = o\left(\frac{1}{\sqrt{N}}\right)$, *together with any of the following conditions on the sequence* $\mathbf{x}$ *is sufficient for*

$$y_{\lfloor Np \rfloor} - y_{\lfloor Nq \rfloor} = o(1): \tag{32}$$

i) $x_N = o(1)$ *and* $\frac{\ln p}{\ln q} \in \mathbb{Q}$,

ii) $x_N = o(1)$ *and* $x'_N = o\left(\frac{1}{\sqrt{\ln N}}\right)$, *where the sequence* $\mathbf{x}'$ *is defined by* $\mathbf{x}' = (p\mathbf{M}_p + q\mathbf{M}_q)\mathbf{x}$.

These conditions are sharp in the sense that $x_N = O(1)$ *in case i), resp.* $x'_N = O\left(\frac{1}{\sqrt{\ln N}}\right)$ *in case ii), does not lead to the conclusion* (32).

The proof of Theorem 1 will be completed in several steps: In Lemma 2 *ii)* and Lemma 3 *ii)* we will show that (No_2) and (Li_a) are implied by certain growth conditions on the sequences $\mathbf{s}$ and $\mathbf{s}^{(a)}$, namely $s_N = o(N)$ and $s_N^{(a)} = o(N^a)$. Lemma 4 then shows that $s_N = o(N)$ and $s_N^{(a)} = o(N^a)$ are implied by the conditions of Theorem 1. □

Lemma 2. *We denote* $\mathbf{s}' = (\mathbf{M}_p + \mathbf{M}_q)\mathbf{s}$. *Any of the following conditions is sufficient for* (No_2)*:*

i) $s_N = O(N)$ *and* $s'_N = \Theta(N)$,

ii) $s_N = o(N)$ *and* $s_{k_0} > 0$ *for some* $k_0 \geq 0$.

Proof. Suppose, condition *i)* is satisfied. Then, according to (P1.2), we represent the sequence $\mathbf{v} = (v_N)_{N \geq 0}$ as $\mathbf{v} = \mathbf{s} + \mathbf{v}'$, where $\mathbf{v}'$ is the solution of

$$(\mathbf{I} - \mathbf{M}_p - \mathbf{M}_q)\mathbf{v}' = \mathbf{s}'.$$

Using (P1.5), we conclude that $v'_N = \Theta(N \ln N)$ and therefore

$$v_N = s_N + v'_N = \Theta(N \ln N). \tag{33}$$

By (P1.7) we have $\Delta s'_k = O(k^{\frac{1}{2}})$, by (P1.2) and (P1.6) $\Delta \mathbf{v}'$ is the solution of

$$(\mathbf{I} - p\mathbf{M}_p - q\mathbf{M}_q)\Delta \mathbf{v}' = \Delta \mathbf{s}'.$$

Therefore, by (P1.3), we have $\Delta v'_N = O(\sqrt{N})$. We use this estimate for a Taylor expansion of v'_k around $k = \lfloor Np \rfloor$,

$$v_k = s_k + v'_{\lfloor Np \rfloor} + O(\sqrt{N}|k - Np|),$$

which yields

$$\sum_{k=0}^{N} \binom{N}{k} p^k q^{N-k} (v_k + v_{N-k}) = s'_N + v'_{\lfloor Np \rfloor} + v'_{\lfloor Nq \rfloor} + O(N) \qquad \text{and}$$

$$\sigma_N = \sum_{k=0}^{N} \binom{N}{k} p^k q^{N-k} \left[s_k + s_{N-k} - s'_N + O(\sqrt{N}|k - Np| + N) \right]^2 = O(N^2). \tag{34}$$

Asymptotic estimates of the form $f(N,k) = O(g(N,k))$ will appear frequently in this proof and the proofs of the two following lemmas and are always to be understood to hold for $N \to \infty$, uniformly in $0 \le k \le N$. In (34) and before we used the cases $m = 1$ and $m = 2$ of the following well known estimate for the central moments of the binomial distribution

$$\sum_{k=0}^{N} \binom{N}{k} p^k q^{N-k} |k - Np|^m = O(N^{\frac{m}{2}}). \tag{35}$$

Knowing $\boldsymbol{\sigma}$ from (34), we can solve the left equation of (28) with the help of (P1.3) and obtain

$$w_N = O(N^2).$$

This, together with (33), proves the first part of the lemma. Suppose now that condition *ii)* is satisfied. Using (P1.4), we can deduce

$$v_N = \Omega(N). \tag{36}$$

As before, we introduce $\mathbf{s}'$, $\mathbf{v}'$, $\boldsymbol{\sigma}$, and $\mathbf{w}$ and obtain $s'_N = o(N)$, furthermore, by (P1.7) and (P1.3),

$$\Delta s'_N = o\left(\sqrt{N}\right), \quad \Delta v'_N = o\left(\sqrt{N}\right).$$

The Taylor expansion of v_k around $k = \lfloor Np \rfloor$

$$v_k = s_k + v'_{\lfloor Np \rfloor} + o\left(\sqrt{N}|k - Np|\right)$$

now leads to $\sigma_N = o\left(N^2\right)$. We use (P1.3) and arrive at

$$w_N = o\left(N^2\right).$$

This, together with (36), completes the proof of *ii)*. □

Lemma 3. *Any of the following conditions, together with $s_N^{(a)} < \infty$ for $0 \le N < n_0$, is sufficient for* (Li$_a$)*:*

i) $s_N^{(a)} = O(N^a)$, $s_N = O(N)$ *and* $s'_N = \Theta(N)$,

ii) $s_N^{(a)} = o(N^a)$ *and* $s_{k_0} > 0$ *for some* $k_0 > 0$.

Proof. In cases *i)* and *ii)* we have $v_N = \Theta(N \ln N)$ (resp. $v_N = \Omega(N)$), cf. (33) and (36). Using (P1.3) we obtain $v_N^{(a)} = O(N^a)$ (resp. $v_N^{(a)} = o(N^a)$). □

Lemma 4. *Let the sequence* $\mathbf{r}$ *satisfy either of the conditions i), ii) or iii) of Theorem 1, as well as for some* $a > 1$ *conditions iv) and v) of Theorem 1.*

Then $s_N = o(N)$, $s_N^{(a)} = o(N^a)$ *and* $s_N^{(a)} < \infty$ *for* $0 \le N < n_0$.

Proof. First of all, condition *iv)* of Theorem 1 implies $s_N^{(a)} < \infty$ for $N \ge 0$.

Furthermore we observe that $\Delta r'_N = o(1)$ implies $\Delta^2 r''_N = o(N^{-\frac{1}{2}})$, since by (P1.6) we have $\Delta^2 \mathbf{r}''_N = \Delta(p\mathbf{M}_p + q\mathbf{M}_q)\Delta\mathbf{r}'_N$ for $N \ge n_0$, and can then apply (P1.7).

We have to go one step beyond the decomposition proposed in (P1.2):

$$\boldsymbol{\ell} = \mathbf{r} + \mathbf{r}' + \boldsymbol{\ell}''.$$

The sequence $\boldsymbol{\ell}''$ defined by this equation solves

$$(\mathbf{I} - \mathbf{M}_p - \mathbf{M}_q)\boldsymbol{\ell}'' = \mathbf{r}''$$

and therefore satisfies by (P1.3) $\Delta^2 \ell''_N = o(N^{-\frac{1}{2}})$. We can thus expand as follows

$$\ell_k = r_k + r'_k + \ell''_k = r_k + r'_{\lfloor Np \rfloor} + o(|k - Np|) + \ell''_{\lfloor Np \rfloor} + \Delta\ell''_{\lfloor Np \rfloor}(k - \lfloor Np \rfloor) + o(N^{-\frac{1}{2}}(k - Np)^2),$$

and denoting

$$b_{N,k} := \ell_k + \ell_{N-k} - \sum_{\kappa=0}^{N} \binom{N}{\kappa} p^{\kappa} q^{N-\kappa}(\ell_\kappa + \ell_{N-\kappa}), \tag{37}$$

we obtain

$$b_{N,k} = r_k + r_{N-k} - r'_N + (\Delta\ell''_{\lfloor Np \rfloor} - \Delta\ell''_{\lfloor Nq \rfloor})(k - Np) + o\left(N^{-\frac{1}{2}}(k - Np)^2 + N^{\frac{1}{2}}\right)$$

We will use $(a + b + c)^m = O(a^m + b^m + c^m)$ with $a = r_k + r_{N-k} - r'_N$, $b = o\left(N^{-\frac{1}{2}}(k - Np)^2 + N^{\frac{1}{2}}\right)$ and $c = (\Delta\ell''_{\lfloor Np \rfloor} - \Delta\ell''_{\lfloor Nq \rfloor})(k - Np)$, which results in

$$s_N^{(a)} = \sum \binom{N}{k} p^k q^{N-k} |b_{N,k}|^{2a} = o(N^a) + O\left(N^a |\Delta\ell''_{\lfloor Np \rfloor} - \Delta\ell''_{\lfloor Nq \rfloor}|^{2a}\right),$$

Note that $s_N = s_N^{(1)}$. We are finished if $p = \frac{1}{2}$, since then $c = 0$. Otherwise any condition of Proposition 2, that is satisfied by $\Delta\mathbf{r}'$, will also be satisfied by $\Delta\mathbf{r}''$, since both sequences are connected via $\Delta\mathbf{r}'' = (p\mathbf{M}_p + q\mathbf{M}_q)\Delta\mathbf{r}'$. By Proposition 2 we thus have $\Delta\ell''_{\lfloor Np \rfloor} - \Delta\ell''_{\lfloor Nq \rfloor} = o(1)$, which completes the proof. □

Proof of Corollary 1. Applying (P1.6) and (P1.7) several times, we obtain the following implications:

$$\Delta^2 r_N = o\left(1/\sqrt{N}\right) \Rightarrow \Delta^2 r''_N = o\left(1/\sqrt{N}\right),$$

$$r_N = o(\sqrt{N}) \; or \; \Delta r_N = o(1) \Rightarrow \Delta r'_N = o(1),$$

$$\left.\begin{array}{r} r_N = o\left(\sqrt{N/\ln N}\right) \\ \text{or} \quad \Delta r_N = o\left(\sqrt{1/\ln N}\right) \end{array}\right\} \Rightarrow \Delta r'_N = o\left(\sqrt{1/\ln N}\right).$$

This settles the conditions *i)*, *ii)* and *iii)*. Moreover we obtain in a similar fashion as in the proof of Lemma 4

$$\begin{aligned} r_N = o(\sqrt{N}) &\Rightarrow r_k + r_{N-k} - r'_N = o(\sqrt{N}), \\ \Delta r_N = o(1) &\Rightarrow r_k + r_{N-k} - r'_N = o(|k - Np|), \qquad \text{and} \\ \Delta^2 r_N = o\left(N^{-\frac{1}{2}}\right) &\Rightarrow r_k + r_{N-k} - r'_N = o\left(N^{-\frac{1}{2}}(k - Np)^2 + N^{\frac{1}{2}}\right) \end{aligned}$$

in the case $p = \frac{1}{2}$, and obtain condition *v)* of Theorem 1 by just applying (35). □

Remark 5. It is tempting to ask if versions of Theorem 1 hold true also for classes of binary trees T_N with probability models different from the Bernoulli models for tries. One certainly would have to replace $\mathbf{M}_p + \mathbf{M}_q$ by $\mathbf{P}$, given in (18), in the definitions of $\mathbf{r}'$ and $\mathbf{r}''$. We are very much in favor of a positive answer in the following two cases,

$p_{N,k} = \binom{N-1}{k} p^k (1-p)^{N-1-k}$, corresponding to the Bernoulli model for *digital search trees*,

$p_{N,k} = \binom{N}{k} \frac{p^k(1-p)^{N-k}}{1-p^N-(1-p)^N} \mathbb{1}_{\{0<k<N\}}$, corresponding to the Bernoulli model for *Patricia tries*,

in particular we think that Propositions 1 and 2 should survive with only marginal modifications, but have not worked this out.

The next theorem complements Theorem 1:

Theorem 2. *Let L_N be the random variable corresponding to the additive valuation L, defined by* (21) *on the space T_N equipped with the Bernoulli model. Assume that R satisfies for some $a > 1$ condition iv) of Theorem 1. Either of*

i) $p \neq \frac{1}{2}$ *and* $r_N = N \mathbb{1}_{\{N\geq 2\}}$,

ii) $p = \frac{1}{2}$ *and* $r_N = N^{\frac{3}{2}} \mathbb{1}_{\{N\geq 2\}}$

implies

$$\frac{L_N - \mathbb{E}\, L_N}{\sqrt{\operatorname{Var} L_N}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \quad \text{as } N \to \infty.$$

Proof. In case *i)* we obtain $\Delta r_N = \mathbb{1}_{\{N\geq 1\}}$, thus by (P1.5) we have

$$\Delta \ell_N = \frac{\ln N}{p \ln \frac{1}{p} + q \ln \frac{1}{q}} + \delta_N, \quad \text{where } ((\mathbf{I} - p\mathbf{M}_p - q\mathbf{M}_q)\boldsymbol{\delta})_N = O(N^{-1}).$$

Proposition 2 now tells us that $\delta_{\lfloor Np \rfloor} - \delta_{\lfloor Nq \rfloor} = o(1)$, from which we derive

$$b_{N,k} = \frac{\ln p - \ln q}{p \ln \frac{1}{p} + q \ln \frac{1}{q}} (k - Np) + o(|k - Np|), \quad \text{with } b_{N,k} \text{ defined in (37).}$$

In case *ii)* we obtain

$$\Delta^i r_N \sim \left(\tfrac{3}{2}\right)^{\underline{i}} N^{\frac{3}{2}-i} \qquad \text{and} \qquad \Delta^i \ell_N \sim \frac{\sqrt{2}}{\sqrt{2}-1} \left(\tfrac{3}{2}\right)^{\underline{i}} N^{\frac{3}{2}-i} \qquad \text{for integers } i \geq 0,$$

which results in

$$b_{N,k} = \frac{3}{2(\sqrt{2}-1)} \left[\left(k - \frac{N}{2}\right)^2 - \frac{N}{4} \right] + O\left(N^{-\frac{3}{2}} \left| k - \frac{N}{2} \right|^3 + 1 \right).$$

Thus in both cases *i)* and *ii)* we obtain $s_N = \Theta(N)$ and $s_N^{(a)} = \Theta(N^a)$, so by Lemma 2 *i)* resp. Lemma 3 *i)* conditions (No_2) resp. (Li_a) are satisfied. $\square$

Remark 6. There are other sequences $\mathbf{r}$, such as $r_N = (-1)^N\sqrt{N}$, that are not covered by Theorem 1, but satisfy Lemma 2 *i)* and Lemma 3 *i)* and thus lead to (27). Moreover for n_0, p fixed, a sequence $\mathbf{r}$ satisfying conditions *i)* of Lemmas 2 and 3, and a sequence $\bar{\mathbf{r}}$ satisfying conditions *ii)* of Lemmas 2 and 3, any linear combination $a\mathbf{r}+b\bar{\mathbf{r}}$ with $a \neq 0$ again satisfies conditions *i)* of Lemmas 2 and 3. In particular, L_N is asymptotically normal if $n_0 = 2$ and $r_N = N + o\left(\sqrt{\frac{N}{\ln N}}\right)$.

Remark 7. The sequence $\mathbf{r}$ given in *i)* corresponds to the external path length of a binary trie. Asymptotic normality of $\frac{L_N - \mathbb{E}\,L_N}{\sqrt{\operatorname{Var} L_N}}$ in the case $p = \frac{1}{2}$ already follows from Corollary 1. Jacquet and Regnier [15] proved asymptotic normality of the external path length in a binary trie under the Bernoulli models. Jacquet and Szpankowski [16] have proved asymptotic normality of a related valuation, namely the internal path length of a digital search tree. Considering sequences $r_N = N^\alpha \mathbb{I}_{\{N\geq 2\}}$, the sequence $\mathbf{r}$ given in *ii)* fills the gap corresponding to $\alpha = \frac{3}{2}$ between Corollary 1 and Example 1 presented in the next section. This will show that for $p = \frac{1}{2}$ and the above scale of sequences, conditions (No$_2$) and (Li$_2$) are strong enough to separate the good from the evil.

Theorem 3. *Let L_N be the random variable corresponding to the additive valuation L, defined via* (21) *in terms of $R(t) = |t|^\alpha \mathbb{I}_{\{|t|\geq 2\}}$ on the space $\mathcal{T}_N$ equipped with the Bernoulli model. Then*

$$\frac{L_N - \mathbb{E}\,L_N}{\sqrt{\operatorname{Var} L_N}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \quad \text{as } N \to \infty.$$

holds for any α, if $p \neq \frac{1}{2}$, and only for $\alpha \leq \frac{3}{2}$, if $p = \frac{1}{2}$.

Proof. The claim follows from Corollary 1, Theorem 2 and Examples 1 and 3 in Section 4. □

4 Counterexamples

As we have not given any necessary conditions for asymptotic normality in terms of the sequence $\mathbf{r}$, it is particularly interesting if the sufficient conditions given in Theorem 1 are sharp in some sense. Now Examples 1 and 2 deal with sequences $\mathbf{r}$ satisfying $\Delta^2 r_N'' = O\left(N^{-\frac{1}{2}+\varepsilon}\right)$, resp. $\Delta^2 r_N'' = O\left(N^{-\frac{1}{2}}\right)$, for which there is no asymptotic normality, thus demonstrating that the most obvious weakening of condition *i)* of Theorem 1 is no longer sufficient for asymptotic normality. As all the sufficient conditions of Theorem 1 have been derived with the help of Lemma 1, this also shows that there is in some sense not much room left for sequences $\mathbf{r}$ implying asymptotic normality, but not being recognized by Lemma 1, though perhaps by direct verification of the Lindeberg conditions (No) and (Li). Example 1, dealing with sequences $r_N = N^\alpha$ in the case $p = \frac{1}{2}, \alpha > \frac{3}{2}$, moreover constitutes one step of the proof of Theorem 3. In Example 3 we appeal to a "nonclassical" central limit theorem for martingale difference arrays to establish asymptotic normality in the case of the sequence $r_N = N^\alpha$, when $p \neq \frac{1}{2}, \alpha > 1$, which is the final step in the proof of Theorem 3. Although Lemma 1 can not be applied in this case, there is again a formulation of sufficient conditions for asymptotic normality that is very much in the spirit of Lemma 1, i.e. yet another sequence appears that satisfies the type of equation also satisfied by $\boldsymbol{\ell}$.

Example 1. The following example shows that we cannot replace condition *i)* in Theorem 1 by $p = \frac{1}{2}$ and $\Delta^2 r''_N = O\left(N^{-\frac{1}{2}+\varepsilon}\right)$ for some $\varepsilon > 0$. We fix $n_0 = 2$ and $\alpha > \frac{3}{2}$ and start with the following sequence of expectations

$$\ell_N = N^\alpha \mathbb{I}_{\{N\geq 2\}},$$

from which we can readily compute the corresponding sequence $(r_k)_{k\geq 0}$, just using (23),

$$r_N = (1-2^{1-\alpha})N^\alpha - 2\binom{\alpha}{2}2^{-\alpha}N^{\alpha-1} + O\left(N^{\alpha-2}\right),$$

which defines via (22) an additive valuation L satisfying $\mathbb{E}\, L_N = \ell_N$. Obviously we have $\Delta^2 r''_N = O(N^{\alpha-2})$, which is a condition weaker than condition *i)* in Theorem 1, and, as we will see, to weak to imply asymptotic normality of $\frac{L_N-\ell_N}{\sqrt{v_N}}$. Since all the r.v.s R_N are deterministic in our case, the r.v. L_N has moments of all orders. With $b_{N,k}$ defined in (37), we find

$$b_{N,k} = \binom{\alpha}{2}2^{-\alpha}\left[8(k-\tfrac{N}{2})^2N^{\alpha-2} - 2N^{\alpha-1}\right] + O\left(N^{\alpha-2} + N^{\alpha-4}(k-\tfrac{N}{2})^4\right),$$

moreover

$$s_N = \sum_{k=0}^{N}\binom{N}{k}2^{-N}b_{N,k}^2 = 8\binom{\alpha}{2}^2 2^{-2\alpha}N^{2\alpha-2} + O\left(N^{2\alpha-3}\right), \quad \Delta s_N = O\left(N^{2\alpha-3}\right),$$

which yields

$$v_N = \frac{2^{3-2\alpha}}{1-2^{3-2\alpha}}\binom{\alpha}{2}^2 N^{2\alpha-2} + O\left(N\ln N + N^{2\alpha-3}\right), \quad \Delta v_N = O\left(N^{2\alpha-3}\right).$$

The easiest way to obtain a recurrence relation for the sequence $d_N := \mathbb{E}\,(L_N-\ell_N)^3$ of third moments is to rewrite (22) as

$$L_N - \ell_N \stackrel{\mathcal{D}}{=} L_{N'} - \ell_{N'} + \bar{L}_{N-N'} - \ell_{N-N'} + b_{N,N'},$$

noting that N' follows a binomial distribution $B(N,\frac{1}{2})$, $L_k \stackrel{\mathcal{D}}{=} \bar{L}_k$ for $k \geq 0$, and $L_{N'} - \ell_{N'}$ and $\bar{L}_{N-N'} - \ell_{N-N'}$ are independent, given N'. Computing expectations of third powers now yields

$$d_N = 2\sum_{k=0}^{N}\binom{N}{k}2^{-N}d_k + \gamma_N + \delta_N,$$

where

$$\gamma_N := 6\sum_{k=0}^{N}\binom{N}{k}2^{-N}b_{N,k}v_k = 3\sum_{k=0}^{N}\binom{N}{k}2^{-N}b_{N,k}(v_k + v_{N-k} + s_N - v_N) = O(N^{3\alpha-\frac{7}{2}}),$$

and

$$\delta_N := \sum_{k=0}^{N}\binom{N}{k}2^{-N}b_{N,k}^3 = 64\binom{\alpha}{2}^3 2^{-3\alpha}N^{3\alpha-3} + O\left(N^{3\alpha-4}\right).$$

We derive

$$d_N = \frac{2^{6-3\alpha}}{1-2^{4-3\alpha}} \binom{\alpha}{2}^3 N^{3\alpha-3} + O\left(N^{3\alpha-\frac{7}{2}}\right).$$

Therefore

$$\mathbb{E}\left(\frac{L_N-\ell_N}{\sqrt{v_N}}\right)^3 = c_\alpha + O\left(N^{3-2\alpha}\ln N + N^{-\frac{1}{2}}\right), \quad \text{where} \quad c_\alpha = 2\sqrt{2}\frac{(1-2^{3-2\alpha})^{\frac{3}{2}}}{(1-2^{4-3\alpha})} > 0.$$

Moreover, denoting $Y_N := \left(\frac{L_N-\ell_N}{\sqrt{v_N}}\right)^3$, we can similarly show that $(\mathbb{E}\,|Y_N|^{\frac{4}{3}})_{N\geq 1}$ is a bounded sequence. The sequence $(Y_N)_{N\geq 1}$ is therefore uniformly integrable. So any random variable Y, to which $(Y_N)_{N\geq 1}$ could converge in distribution, must satisfy $\mathbb{E}\,Y = c_\alpha \neq 0$ and can therefore not be the third power of a normally distributed random variable.

Remark 8. If we start with $r_N = (1-2^{1-\alpha})N^\alpha$, the leading terms in s_N, v_N, d_N, γ_N and δ_N do not change, so that we don't have asymptotic normality also in this case.

Example 2. The following example shows that if $p = \frac{1}{2}$, none of

$$r_N = O\left(N^{\frac{1}{2}}\right), \quad \Delta r'_N = O(1) \quad \text{and} \quad \Delta^2 r''_N = O\left(N^{-\frac{1}{2}}\right) \tag{38}$$

is sufficient for (27). (See also the subsequent remark on how to extend the construction to cases $p \neq \frac{1}{2}$.) Here we will stick to the notation of the previous example. We fix $n_0 = 2$ and start again with a sequence of expectations

$$\ell_N = \mathbb{I}_{\{N\geq 2\}} \sum_{i\geq 1} f^{(m_i)}(N),$$

where $f^{(m)}(N) := \sqrt{m}\, e^{-\frac{(N-m)^2}{m}}$, and $(m_i)_{i\geq 1}$ is a sequence satisfying $m_1 \geq 3$ and $m_{i+1} \geq m_i^2$ for $i \geq 1$, such that the intervals $[(m_i - m_i^{0.6})2^k, (m_i + m_i^{0.6})2^k]_{i\geq 1, k\in\mathbb{Z}}$ are disjoint. The conditions on $(m_i)_{i\geq 1}$ (which are not the weakest in this respect,) will imply

$$v_N = O(N). \tag{39}$$

On the other hand we will show that

$$\limsup_{N\to\infty} |d_N| N^{-\frac{3}{2}} > 0. \tag{40}$$

With the help of the intervals $\mathcal{A}_m := [m - m^{0.6}, m + m^{0.6}]$, where $\mathcal{A}_m \subset [n_0, \infty[$ for $m \geq 3$, we can represent $\boldsymbol{\ell}$ more conveniently as follows

$$\ell_N = \sum_{i\geq 1} \mathbb{I}_{\mathcal{A}_{m_i}}(N) f^{(m_i)}(N) + e^{-\Omega(N^{0.2})},$$

from which it is easy to deduce $\Delta^n \ell_N = O\left(N^{\frac{1-n}{2}}\right)$ and, using (23), $\Delta^n r_N = O\left(N^{\frac{1-n}{2}}\right)$. Thus the sequence $\mathbf{r}$ indeed satisfies $r_N = O\left(N^{\frac{1}{2}}\right)$, $\Delta r'_N = O(1)$ and $\Delta^2 r''_N = O\left(N^{-\frac{1}{2}}\right)$. For the computation of s_N and δ_N we require asymptotics of second and third central moments of $f^{(m_i)}(k)+f^{(m_i)}(N-k)$, where $k\sim B(N,\frac{1}{2})$. The main tool here is the estimate

$$\sum_{k\in\mathcal{A}_m}\binom{N}{k}2^{-N}e^{-\frac{(k-m)^2}{M}}=\begin{cases}\sqrt{\frac{2M}{N+2M}}e^{-\frac{(N-2m)^2}{2N+4M}}+O\left(\frac{1}{\sqrt{m}}\right), & \text{for } N\in\mathcal{A}_{2m},\\ e^{-\Omega(m^{0.2})}, & \text{for } N\notin\mathcal{A}_{2m},\end{cases}$$

valid for $M=\Theta(m)$ and $m\to\infty$, which is obtained by standard asymptotic techniques. We derive

$$s_N=\sum_{i\geq 1}\mathbb{I}_{\mathcal{A}_{2m_i}}(N)g^{(m_i)}(N)+O(\sqrt{N}),$$

where

$$g^{(m)}(N)=\frac{2m}{\sqrt{3}}\left(\exp\left(-\frac{(N-2m)^2}{6m}\right)+\exp\left(-\frac{(N-2m)^2}{2m}\right)-\sqrt{3}\exp\left(-\frac{(N-2m)^2}{4m}\right)\right)$$

and

$$\Delta s_N=O\left(1+\sum_{i\geq 1}\mathbb{I}_{\mathcal{A}_{2m_i}}(N)\sqrt{m_i}e^{-\frac{(N-2m_i)^2}{6m_i}}\right).$$

Note that the sequence $\mathbf{s}$ satisfies $s_0=\ldots=s_{n_0-1}=0$, therefore the equations

$$(\mathbf{I}-2\mathbf{M}_{\frac{1}{2}})\mathbf{v}=\mathbf{s}\quad\text{and}$$
$$(\mathbf{I}-\mathbf{M}_{\frac{1}{2}})\Delta\mathbf{v}=\Delta\mathbf{s}\quad\text{(the latter is derived from the former using (P1.6))}$$

have the solutions

$$\mathbf{v}=\mathbf{s}+\sum_{n\geq 1}2^n\mathbf{M}_{2^{-n}}\mathbf{s},\qquad \Delta\mathbf{v}=\Delta\mathbf{s}+\sum_{n\geq 1}\mathbf{M}_{2^{-n}}\Delta\mathbf{s},$$

which is easily verified, just noting that $(\mathbf{M}_a\mathbf{M}_b)_{N,k}=(\mathbf{M}_{ab})_{N,k}$ for $k\geq n_0$. We can thus deduce

$$\Delta v_N=O(\sqrt{m_i}+\ln N)\ \text{ for } N\leq m_{i+1}+m_{i+1}^{0.6},$$

and $m_{i+1}\geq m_i^2$ implies in particular

$$\Delta v_N=O\left(N^{\frac{1}{4}}\right),\ \text{for } N\in\bigcup_{i\geq 1}\mathcal{A}_{m_i}. \tag{41}$$

Moreover we have

$$v_N=s_N+\sum_{n\geq 1}\sum_{i\geq 1}\sum_{k\in\mathcal{A}_{2m_i}}2^n\binom{N}{k}2^{-nk}(1-2^{-n})^{N-k}g^{(m_i)}(k)+O(N),$$

where the term $O(N)$ comes from the term $O(\sqrt{N})$ present in s_N, cf. (P1.3). Observing that $g^{(m_i)}(k)\leq Ck(k-1)$ for some $C>0$, we conclude that the sum of the terms with $N2^{-n}<4$ will contribute another

$O(N)$. The sum of the remaining terms, which also contributes $O(N)$, is the most demanding and requires the following estimates, some of them are simply Chernoff bounds. Note that

$$g^{(m)}(k) = O\left(m \exp\left(-\frac{(k-2m)^2}{6m}\right)\right).$$

$$\sum_{k\in\mathcal{A}_{2m}} \binom{N}{k} p^k (1-p)^{N-k} e^{-\frac{(k-2m)^2}{6m}} = \begin{cases} O(1), & \text{for } Np \in \mathcal{A}_{2m}, \\ e^{-\Omega(m^{0.2})}, & \text{for } Np \in [\frac{m}{2}, 4m] \setminus \mathcal{A}_{2m}, \\ O\left(\left(\frac{Np}{m}\right)^m e^{m-Np}\right), & \text{for } Np \in [4, \frac{m}{2}[, \\ O\left(e^{-\frac{Np}{8}}\right), & \text{for } Np \in]4m, N]. \end{cases}$$

Note that by our assumption on $(m_i)_{i\geq 1}$ there is for each N at most one pair (n,i) such that $N2^{-n} \in \mathcal{A}_{2m_i}$. Leaving aside the details we finally derive (39) and can now turn to third moments d_N. We employ (41) and obtain $\gamma_N = O\left(N^{\frac{5}{4}}\right)$. Moreover

$$\delta_N = \sum_{i\geq 1} \mathbb{1}_{\mathcal{A}_{2m_i}}(N) h^{(m_i)}(N) + O(N),$$

where, denoting $k = N - 2m$,

$$h^{(m)}(N) = m^{\frac{3}{2}} \left(e^{-\frac{3k^2}{16m}} + 3e^{-\frac{11k^2}{16m}} - 2\sqrt{6}e^{-\frac{7k^2}{24m}} - 2\sqrt{6}e^{-\frac{5k^2}{8m}} + 4\sqrt{2}e^{-\frac{3k^2}{8m}}\right).$$

Finally, for $N \in \mathcal{A}_{2m_i}$ we obtain

$$d_N = \delta_N + O\left(\frac{m_i}{m_{i-1}} \delta_{2m_{i-1}} + N^{\frac{5}{4}}\right) = h^{(m_i)}(N) + O(N^{\frac{5}{4}}).$$

Thus

$$\limsup_{N\to\infty} |d_N| N^{-\frac{3}{2}} \geq \lim_{i\to\infty} |d_{2m_i}| (2m_i)^{-\frac{3}{2}} = 8(2\sqrt{3} - 2 - \sqrt{2}) = 0.3991 > 0.$$

But, by the same reasoning as in the previous example, asymptotic normality of $\frac{L_N - \ell_N}{\sqrt{v_N}}$ would require $\lim_{N\to\infty} |d_N| N^{-\frac{3}{2}} = 0$.

Remark 9. A similar construction, building upon modified sequences $(m_i)_{i\geq 1}$, is possible for any $p \in]0,1[$. We will stress the part of establishing (39), but skip the proof of (40), which can be easily adapted from the case $p = \frac{1}{2}$.

If $\frac{\ln p}{\ln q} \in \mathbb{Q}$, we have $p = r^m$, $q = r^n$ for some $0 < r < 1$ and relatively prime $m, n \in \mathbb{N}$. Hence

$$\mathbf{v} = \sum_{k,\ell\geq 0} \binom{k+\ell}{k} \mathbf{M}_{r^{mk+n\ell}} \mathbf{s},$$

and

$$\sum_{mk+n\ell=N} \binom{k+\ell}{k} = \left[t^N\right] \frac{1}{1 - t^m - t^n} \sim \frac{r^{-N}}{mr^m + nr^n},$$

since $(1-t^m-t^n)^{-1}$ has its dominant singularity at $t=r$. We conclude $\mathbf{v}=\mathbf{s}+O\left(\sum_{k\geq 1} r^{-k}\mathbf{M}_{r^k}\mathbf{s}\right)$, and (39) will be satisfied if we demand that the intervals $[(m_i-m_i^{0.6})r^k,(m_i+m_i^{0.6})r^k]_{i\geq 1,k\in\mathbb{Z}}$ are disjoint.

In the case $\frac{\ln p}{\ln q}\notin\mathbb{Q}$, fast enough growth of $(m_i)_{i\geq 1}$ is enough to guarantee (39), i.e. we do not need any sort of "disjoint intervals" condition. Proceeding as in Example 2, we obtain

$$s_N=\sum_{k\geq 1}\mathbb{I}_{\mathcal{A}_{2n_k}}(N)\bar{g}^{(n_k)}(N)+O(\sqrt{N}),$$

where the sequence $(n_k)_{k\geq 1}$ is defined by

$$n_{2i-1}=\lfloor\frac{m_i}{2p}\rfloor,\qquad n_{2i}=\lfloor\frac{m_i}{2q}\rfloor$$

and $\bar{g}^{(n)}(N)$ is a function similar to $g^{(n)}(N)$, with several ripples of height $O(n)$ and width $O(\sqrt{n})$ near $N=2n$, which obeys

$$|\,\mathbb{I}_{\mathcal{A}_{2n}}(N)\bar{g}^{(n)}(N)|\leq Cn^{\frac{3}{2}}\binom{N}{n}2^{-N}=:y_N^{(n)}\quad\text{ for } n\geq 3 \text{ and some } C>0.$$

The contribution of $\bar{g}^{(n)}$ to $\mathbf{v}$ can now be estimated by $\sum_{k,\ell\geq 0}\binom{k+\ell}{k}\mathbf{M}_{p^kq^\ell}\mathbf{y}^{(n)}$. We have

$$(\mathbf{M}_a\mathbf{y}^{(n)})_N=Cn^{\frac{3}{2}}\binom{N}{n}\left(\frac{a}{2}\right)^n\left(1-\frac{a}{2}\right)^{N-n}\leq C'n\left(\frac{Na}{n}\right)^2 e^{-Na/n}\quad\text{for some C'¿0}$$

and will now apply Mellin transform techniques (cf. [5]) to derive asymptotics of the two sums

$$S_{n,N}:=Cn^{\frac{3}{2}}\binom{N}{n}\sum_{k,\ell\geq 0}\binom{k+\ell}{k}\left(\frac{p^kq^\ell}{2}\right)^n\left(1-\frac{p^kq^\ell}{2}\right)^{N-n}\qquad\text{and}$$
$$S'_{n,N}:=C'n\sum_{k,\ell\geq 0}\binom{k+\ell}{k}\left(\frac{Np^kq^\ell}{n}\right)^2 e^{-\frac{Np^kq^\ell}{n}}.$$

We obtain

$$S'_{1,N}=\frac{1}{2\pi i}\int_{-3/2-i\infty}^{-3/2+i\infty}\frac{C'\Gamma(s+2)N^{-s}}{1-p^{-s}-q^{-s}}ds,$$

from which we derive $S'_{1,N}=O(N)$, as $N\to\infty$, and $S'_{1,N}=O(N^2)$, as $N\to 0$. Thus there is a constant $C''>0$, such that $S'_{1,N}\leq C''N$ for $0\leq N<\infty$, and we can moreover conclude

$$S_{n,N}\leq S'_{n,N}=nS'_{1,N/n}\leq nC''N/n=C''N.$$

Furthermore

$$S_{n,N}=\frac{1}{2\pi i}\int_{-2-i\infty}^{-2+i\infty}\frac{Cn^{\frac{3}{2}}2^s}{1-p^{-s}-q^{-s}}\frac{1}{n}\prod_{k=n}^{N}\left(1+\frac{s}{k}\right)^{-1}ds=\frac{C}{2(p\ln\frac{1}{p}+q\ln\frac{1}{q})}\frac{\sqrt{n}N}{n-1}+o(N),$$

for $n\geq 3$ as $N\to\infty$, where the function implied by the symbol o depends on n. Thus there are $C'''>0$ and $\phi(n)>2n$ such that $S_{n,N}\leq C'''\frac{N}{\sqrt{n}}$ for $N\geq\phi(n)$. These results are obtained by shifting the line of integration to the right (resp. left) and collecting residues at the simple pole $s=-1$ (resp. $s=-2$) of the integrands and the other poles of $\frac{1}{1-p^{-s}-q^{-s}}$, which are all simple, have real part strictly greater than -1, and form a uniformly discrete set (cf. [3, Lemma 8]) contained in a vertical strip. The new contour of integration can be chosen as a rectifiable curve, which is contained in the set $0<\Re s<1$, and on which $\frac{1}{1-p^{-s}-q^{-s}}$ is bounded. Denoting $s=\sigma+it$, it is easily seen, that $K_{n,N}(s):=\frac{1}{n}\prod_{k=n}^{N}\left(1+\frac{s}{k}\right)^{-1}$ satisfies

$$|K_{n,N}(s)|\le C_n N^{-\sigma}\frac{1}{n^2+t^2}\qquad \text{for } N\ge n+1, \text{ with some } C_n>0, \text{ uniformely in } \sigma\in[-1,1] \text{ and } t\in\mathbb{R}.$$

It is well known that

$$|\Gamma(\sigma+it)|\sim\sqrt{2\pi}|t|^{\sigma-\frac{1}{2}}e^{-\pi|t|/2},\quad \text{as } |t|\to\infty.$$

Thus both the sums of residues and the new contour integrals are absolutely convergent.

We have thus derived $0\le S_{n,N}\le C''N\,\mathbb{I}_{\{n\le N<\phi(n)\}}+C'''\frac{N}{\sqrt{n}}\mathbb{I}_{\{N\ge\phi(n)\}}$ and can conclude

$$v_N\le\sum_{k\ge1}S_{n_k,N}=O(N),$$

if the sequence $(m_i)_{i\ge1}$ is chosen such that $m_1\ge6$ (which guarantees $n_1,n_2\ge3$) and

$$n_{2i+1}\wedge n_{2i+2}\ge\phi(n_{2i-1})\vee\phi(n_{2i})$$

Note that the property $\lim_{n\to\infty}\limsup_{N\to\infty}\frac{S_{n,N}}{N}=0$, which we have derived for $\frac{\ln p}{\ln q}\notin\mathbb{Q}$, does not hold if $\frac{\ln p}{\ln q}\in\mathbb{Q}$.

Example 3. Here we are going to demonstrate that in some cases where Lemma 1 can not be applied since (Li) is not satisfied, a "nonclassical" version of the central limit theorem for martingale difference arrays can be used to establish asymptotic normality. Still in the setting of Theorem 1, we fix $n_0=2$, $0<p<\frac{1}{2}$, $\alpha>1$ and $r_N=(1-p^\alpha-q^\alpha)N^\alpha$, and will prove $\frac{L_N-\ell_N}{\sqrt{v_N}}\xrightarrow{\mathcal{D}}\mathcal{N}(0,1)$, as $N\to\infty$. We readily obtain

$$\ell_N=N^\alpha+O\left(N\ln N+N^{\alpha-1}\right),$$

and from according estimates of differences of $\mathbf{r}$ we deduce for $m\ge1$:

$$\Delta^m\ell_N=\alpha^{\underline{m}}N^{\alpha-m}+O\left(N^{1-m}\ln N+N^{\alpha-1-m}\right).\tag{42}$$

Here $x^{\underline{n}}=x(x-1)\cdots(x-n+1)$ denotes falling factorial powers. We can obtain an approximation for ℓ_k around $k=pN$ (and a similar one around $k=qN$) valid for $0\le k\le N$, by taking some terms of its Newton series plus an estimate for the remainder term

$$\ell_k=\sum_{i=0}^{2}\Delta^i\ell_{\lfloor pN\rfloor}\frac{(k-\lfloor pN\rfloor)^{\underline{i}}}{i!}+O\left(N^{\alpha-3}|k-pN|^3\right).$$

Writing the terms $(k-\lfloor pN\rfloor)^{\underline{i}}$ as polynomials in $(k-Np)$, we obtain for $b_{N,k}$, as defined in (37), the following representation

$$b_{N,k}=\left(\Delta\ell_{\lfloor pN\rfloor}-\Delta\ell_{\lfloor qN\rfloor}\right)(k-Np)+c_N\left((k-Np)^2-Npq\right)+O\left(N^{\alpha-3}(|k-pN|^3+N^{\frac{3}{2}})\right),$$

where c_N does not depend on k and satisfies $c_N = O(N^{\alpha-2})$. We thus derive

$$s_N = \left(\Delta\ell_{\lfloor pN\rfloor} - \Delta\ell_{\lfloor qN\rfloor}\right)^2 pqN + O(N^{2\alpha-2}). \tag{43}$$

From (42) and (43) we deduce

$$s_N = \alpha^2(p^{\alpha-1} - q^{\alpha-1})^2 pqN^{2\alpha-1} + O\left(N^{2\alpha-2} + N^{\alpha}\ln N\right) \quad \text{and} \quad \Delta s_N = O(N^{2\alpha-2}),$$

moreover

$$\begin{aligned} v_N &= \frac{\alpha^2(p^{\alpha-1} - q^{\alpha-1})^2 pq}{1 - p^{2\alpha-1} - q^{2\alpha-1}} N^{2\alpha-1} + O\left(N^{2\alpha-2} + N^{\alpha}\ln N\right), \\ \Delta v_N &= O(N^{2\alpha-2}), \qquad \text{and} \\ w_N &= O(N^{4\alpha-3}). \end{aligned}$$

We have thus derived $w_N = o(v_N^2)$, which is condition (No$_2$) from Lemma 1, which was seen to imply (No). On the other hand (Li) is not satisfied since the first term of the series in (Li), $\mathbb{E}\,[\xi_{N,1}^2 \mathbb{I}_{\{|\xi_{N,1}|>\varepsilon\}}]$ does not converge to 0 for small enough $\varepsilon > 0$: This follows from $\mathbb{E}\,[\xi_{N,1}^2] = \frac{s_N}{v_N} \to 1 - p^{2\alpha-1} - q^{2\alpha-1} > 0$, as $N \to \infty$.

However there is a "nonclassical" central limit theorem for martingale difference arrays (cf. [33, p. 553]) which provides another pair of sufficient conditions for asymptotic normality of $\frac{L_N-\ell_N}{\sqrt{v_N}} = \sum_{i=1}^{\infty} \xi_{N,i}$. One of these conditions is (No) which has already been checked. The other one, which we call (Λ), is stated in terms of regular distribution functions $F_{N,i}(x) := \mathbb{P}\,(\xi_{N,i} \le x | \mathcal{F}_{N,i-1})$:

$$\sum_{i=1}^{\infty} \int_{|x|>\varepsilon} |x| \left| F_{N,i}(x) - \Phi\left(x \Big/ \sqrt{\mathbb{E}\,[\xi_{N,i}^2 | \mathcal{F}_{N,i-1}]}\right) \right| dx \xrightarrow{P} 0, \quad \text{as } N \to \infty, \text{ for each } \varepsilon > 0, \tag{Λ}$$

where Φ is the standard normal distribution function, and $F_{N,i}(x) - \Phi\left(x \Big/ \sqrt{\mathbb{E}\,[\xi_{N,i}^2 | \mathcal{F}_{N,i-1}]}\right)$ has to be understood as identically 0 on the set $\{\mathbb{E}\,[\xi_{N,i}^2 | \mathcal{F}_{N,i-1}] = 0\} \in \mathcal{F}_{N,i-1}$. As in Lemma 1, we want to get rid of ε at the cost of obtaining a stronger condition. Using $|x|\,\mathbb{I}_{\{|x|>\varepsilon\}} \le \varepsilon^{-\frac{1}{2}}|x|^{\frac{3}{2}}$ and (17), and substituting $\xi = x\sqrt{\frac{v_N}{s_{|\mathfrak{t}^{(i)}|}}}$, we find the following condition, which implies (Λ):

$$\sum_{i:\, s_{|\mathfrak{t}^{(i)}|}>0} \left(\frac{s_{|\mathfrak{t}^{(i)}|}}{v_N}\right)^{\frac{5}{4}} \int_{\mathbb{R}} |\xi|^{\frac{3}{2}} \left| \mathbb{P}\left(\frac{\lambda_{|\mathfrak{t}^{(i)}|,1}}{\sqrt{s_{|\mathfrak{t}^{(i)}|}}} \le \xi \,\middle|\, |\mathfrak{t}^{(i)}|\right) - \Phi(\xi) \right| d\xi \xrightarrow{P} 0, \quad \text{as } N \to \infty.$$

Denoting by k a random variable with the binomial distribution $B(N,p)$, we are going to show

$$\mathbb{P}\left(\frac{\lambda_{N,1}}{\sqrt{s_N}} \le x\right) = \mathbb{P}\left(\frac{k - Np}{\sqrt{Npq}} \le x\right) + O\left(\frac{1}{\sqrt{N}(1+|x|^3)}\right). \tag{44}$$

Moreover by nonuniform Berry-Esseen type inequalities (cf. [33, p. 376]) we also have

$$\left| \mathbb{P}\left(\frac{k - Np}{\sqrt{Npq}} \le x\right) - \Phi(x) \right| = O\left(\frac{1}{\sqrt{N}(1+|x|^3)}\right),$$

and conclude

$$\int_{\mathbb{R}} |x|^{\frac{3}{2}} \left| \mathbb{P}\left(\frac{\lambda_{N,1}}{\sqrt{s_N}} \le x\right) - \Phi(x) \right| dx = O\left(\frac{1}{\sqrt{N}}\right).$$

Since $\frac{1}{\sqrt{N}} = O\left(s_N^{-\frac{1}{2}\frac{1}{2\alpha-1}}\right)$ we find, enlarging the range of summation, yet another condition that implies (Λ):

$$v_N{}^{-\frac{5}{4}} \sum_{i:|\mathsf{t}^{(i)}|\ge 0} s_{|\mathsf{t}^{(i)}|}^{\frac{5}{4}-\frac{1}{2}\frac{1}{2\alpha-1}} \to 0, \quad \text{in } L^1, \text{ as } N \to \infty. \tag{Λ'}$$

Exactly as in the proof of Lemma 1 we deduce that $y_{|\mathsf{t}|} := \mathbb{E} \sum_{i:|\mathsf{t}^{(i)}|\ge 0} s^a_{|\mathsf{t}^{(i)}|}$ satisfies the recurrence

$$(\mathbf{I} - \mathbf{M}_p - \mathbf{M}_q)\mathbf{y} = \mathbf{s}^a, \quad \text{where we denoted } a := \frac{5}{4} - \frac{1}{2}\frac{1}{2\alpha-1} \quad \text{and} \quad \mathbf{s}^a := (s_N^a)_{N\ge 0}.$$

From (P1.3) we can now deduce $y_N = O(s_N^a + N \ln N)$, thus $y_N = o(v_N{}^{\frac{5}{4}})$, which proves (Λ'), thus (Λ) is satisfied and $\frac{L_N - \ell_N}{\sqrt{v_N}}$ is indeed asymptotically normal.

It remains to prove (44). Note that with $k \sim B(N,p)$ and $Y_N = Y_N(k) = \frac{k-Np}{\sqrt{Npq}}$ we have

$$\frac{\lambda_{N,1}}{\sqrt{s_N}} \stackrel{\mathcal{D}}{=} Z_N = Z_N(k) := \frac{b_{N,k}}{\sqrt{s_N}} = \phi_N(Y_N),$$

where $\phi_N(y) = y + O\left(\frac{1+y^2}{\sqrt{N}}\right)$. The random variable $Z_N(k)$ is strictly decreasing in k for $0 \le k \le \lfloor \frac{N}{2} \rfloor$, thus $\phi_N : [Y_N(\lfloor \frac{N}{2} \rfloor), Y_N(0)] \mapsto [Z_N(\lfloor \frac{N}{2} \rfloor), Z_N(0)]$ is invertible, with

$$\phi_N^{-1}(x) = x + O\left(\frac{1+x^2}{\sqrt{N}}\right), \tag{45}$$

moreover there is $c > 0$ such that

$$|\phi_N^{-1}(x)| \ge c|x| + O\left(\tfrac{1}{\sqrt{N}}\right). \tag{46}$$

Note that each of $Y_N(0)$, $Z_N(0)$, $-Y_N(\lfloor \frac{N}{2} \rfloor)$ and $-Z_N(\lfloor \frac{N}{2} \rfloor)$ is of order $\Theta(\sqrt{N})$. Thus, for some $C > 0$, we have

$$\mathbb{P}(Z_N \le x) = \mathbb{P}(Y_N \le x) \qquad \text{for } |x| > C\sqrt{N}.$$

Moreover we obtain

$$\begin{aligned}
&\mathbb{P}(Z_N \le x) = 1 \quad \text{and} \quad \mathbb{P}(Y_N \le x) \ge \mathbb{P}(Y_N \le Z_N(0)) = 1 - e^{-\Omega(N)} \quad \text{for} \quad Z_N(0) \le x \le C\sqrt{N},\\
&\mathbb{P}(Z_N \le x) = 0 \quad \text{and} \quad \mathbb{P}(Y_N \le x) \le \mathbb{P}(Y_N \le Z_N(\lfloor \frac{N}{2} \rfloor)) = e^{-\Omega(N)} \quad \text{for } -C\sqrt{N} \le x \le Z_N(\lfloor \frac{N}{2} \rfloor).
\end{aligned}$$

For the remaining values $Z_N(\lfloor \frac{N}{2} \rfloor) < x < Z_N(0)$ we have

$$\mathbb{P}(Z_N \le x) = \mathbb{P}(Z_N \le x, k < \lfloor \tfrac{N}{2} \rfloor) + \mathbb{P}(Z_N \le x, k \ge \lfloor \tfrac{N}{2} \rfloor),$$

where we estimate the second term $\mathbb{P}\left(Z_N \leq x, k \geq \lfloor \frac{N}{2} \rfloor\right) \leq \mathbb{P}\left(k \geq \lfloor \frac{N}{2} \rfloor\right) = e^{-\Omega(N)}$ and rewrite the first term

$$\begin{aligned}
\mathbb{P}\left(Z_N \leq x, k < \lfloor \tfrac{N}{2} \rfloor\right) &= \mathbb{P}\left(Y_N \leq \phi_N^{-1}(x), k < \lfloor \tfrac{N}{2} \rfloor\right) \\
&= \mathbb{P}\left(Y_N \leq x, k < \lfloor \tfrac{N}{2} \rfloor\right) \pm \mathbb{P}\left(Y_N \in \mathcal{A}_x, k < \lfloor \tfrac{N}{2} \rfloor\right) \\
&= \mathbb{P}\left(Y_N \leq x\right) - \mathbb{P}\left(Y_N \leq x, k \geq \lfloor \tfrac{N}{2} \rfloor\right) \pm \mathbb{P}\left(Y_N \in \mathcal{A}_x, k < \lfloor \tfrac{N}{2} \rfloor\right).
\end{aligned}$$

The set $\mathcal{A}_x$ is defined by $\mathcal{A}_x = \left]x \wedge \phi_N^{-1}(x)\, , x \vee \phi_N^{-1}(x)\right]$ and we have to take the sign $+$ if $x \leq \phi_N^{-1}(x)$ and the sign $-$ otherwise. We further estimate

$$\begin{aligned}
&\mathbb{P}\left(Y_N \leq x, k \geq \frac{N}{2}\right) = e^{-\Omega(N)}, \\
&\mathbb{P}\left(Y_N \in \mathcal{A}_x, k < \frac{N}{2}\right) = \mathbb{P}\left(Y_N \in \mathcal{A}_x\right) - e^{-\Omega(N)} \text{ and} \\
&\mathbb{P}\left(Y_N \in \mathcal{A}_x\right) = O\left(\frac{1+x^2}{\sqrt{N}} e^{-c^2x^2/2}\right),
\end{aligned}$$

by (45), (46) and the local limit theorem for the Bernoulli scheme, cf. [33, p. 56]. Since $x^2 = O(N)$ is valid for the above error estimates, these can all be relaxed to the form $O\left(\frac{1}{\sqrt{N}(1+|x|^3)}\right)$ present in (44).

Remark 10. This method of proving asymptotic normality of $\frac{L_N - \ell_N}{\sqrt{v_N}}$ by verifying conditions (No) and (Λ) also works for other sufficiently smooth sequences $\mathbf{r}$ of polynomial growth, satisfying $\Delta^2 r_N = \Omega\left(N^{-1+\varepsilon}\right)$ for some $\varepsilon > 0$. Exponentially growing sequences $\mathbf{r}$ however again constitute examples leading to non normal limiting distributions. Take $0 < p < \frac{1}{2}$ and $r_N = 2^N$. Then $\ell_N \sim 2^N$ and $v_N \sim (1+3q)^N$, moreover $\frac{L_N - \ell_N}{\sqrt{v_N}} \xrightarrow{\mathcal{D}} 0$. A closer look reveals that all the limit laws that we can obtain for $\frac{L_N - b_N}{a_N}$ by choosing appropriate normalizing sequences $(a_N)_{N \geq 0}, (b_N)_{N \geq 0}$ are degenerate.

5 Conclusion

In this paper we proposed the use of martingale difference arrays as a method to detect asymptotic normality of the costs of certain recursive algorithms. The main tool, Lemma 1, restates well known sufficient conditions for asymptotic normality (of Lyapunov type) in terms of asymptotic relations between three sequences. Asymptotics of those sequences can usually be obtained by the same toolkit that is used to derive asymptotics of expected costs. The method is likely to be applicable in cases where one expects asymptotically normal costs and where an analysis of expected costs has already been performed. We have applied the method to additive valuations defined on binary tries equipped with the Bernoulli models. There an interesting problem surfaced.

Open problem Let a sequence of random variables $(L_N)_{N \geq 0}$ satisfy $L_0 = L_1 = 0$, and for $N \geq 2$

$$L_N \stackrel{\mathcal{D}}{=} L_k + \bar{L}_{N-k} + r_N,$$

where $k \sim B(N, p)$, $L_k \stackrel{\mathcal{D}}{=} \bar{L}_k$ for $k \geq 0$, and $L_k, \bar{L}_{N-k}$ are independent, conditional on k, and $(r_N)_{N \geq 2}$ is a sequence of real numbers.

Is it true that $\frac{L_N - \mathbb{E} L_N}{\sqrt{\operatorname{Var} L_N}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$ holds, if $r_N = o\left(\sqrt{N}\right)$ and $\frac{\ln p}{\ln(1-p)} \notin \mathbb{Q}$?

From Corollary 1 and Example 2 we know that $\frac{L_N - \mathbb{E} L_N}{\sqrt{\operatorname{Var} L_N}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$ is implied by both sets of conditions $\{r_N = o\left(\sqrt{N}\right), \frac{\ln p}{\ln(1-p)} \in \mathbb{Q}\}$ and $\{r_N = o\left(\sqrt{\frac{N}{\ln N}}\right), \frac{\ln p}{\ln(1-p)} \notin \mathbb{Q}\}$, but not by $r_N = O\left(\sqrt{N}\right)$.

For future work we plan to give further applications of the proposed method, which call for generalizations such as those indicated in Remark 2.

6 Appendix

Proof of Proposition 2. According to (P1.2) we have $\mathbf{y} = \mathbf{x} + (\mathbf{I} - p\mathbf{M}_p - q\mathbf{M}_q)^{-1}\mathbf{x}'$. Obviously $x_{\lfloor Np \rfloor} - x_{\lfloor Nq \rfloor} = o(1)$, so we are left with showing (32) for sequences $\mathbf{x}$ satisfying $x_0 = \ldots = x_{n_0-1} = 0$. (Any of conditions *i)* or *ii)*, that is satisfied by $\mathbf{x}$, is also satisfied by $\mathbf{x}'$.)

Introducing the Poisson generating function $x(t) = \sum_{k \geq 0} x_k \frac{t^k}{k!} e^{-t}$, which is an entire function, and similarly $y(t)$, which satisfies (and is actually the unique entire solutions of, cf. [3, Lemma 2]) the functional equation

$$y(t) - py(pt) - qy(qt) = x(t), \tag{47}$$

we have to prove

$$y(pt) - y(qt) = o(1) \text{ as } t \to \infty. \tag{48}$$

The proof is then completed by depoissonizing this equation with the help of (P1.8). Just note that by the assumption $\Delta x_N = o\left(\frac{1}{\sqrt{N}}\right)$ we have the chain of implications

$$(\mathbf{I} - p\mathbf{M}_p - q\mathbf{M}_q)\mathbf{y} = \mathbf{x} \overset{(P1.6)}{\Longrightarrow} (\mathbf{I} - p^2\mathbf{M}_p - q^2\mathbf{M}_q)\Delta\mathbf{y} = \Delta\mathbf{x} \overset{(P1.3)}{\Longrightarrow} \Delta y_N = o\left(\frac{1}{\sqrt{N}}\right).$$

It thus remains to prove (48). Note that $x(t) = O(t^2)$ as $t \to 0$. The unique entire solution of (47) is thus given by the absolutely convergent series $y(t) = \sum_{k,\ell \geq 0} \binom{k+\ell}{k} p^k q^\ell x(p^k q^\ell t)$. We derive

$$y(pt) - y(qt) = \sum_{k+\ell \geq 1} \nu_{k,\ell}\, x(p^k q^\ell t), \tag{49}$$

where $\nu_{k,\ell} = \binom{k+\ell-1}{\ell} p^{k-1} q^\ell - \binom{k+\ell-1}{k} p^k q^{\ell-1} = \binom{k+\ell}{k} p^k q^\ell \frac{qk - p\ell}{pq(k+\ell)}$, and distinguish two cases:

i) $x_N = o(1)$ and $\frac{\ln p}{\ln q} \in \mathbb{Q}$:

For some $0 < r < 1$ and relatively prime $m, n \in \mathbb{N}$ we have $p = r^m$, $q = r^n$. The rational function $\frac{t^m - t^n}{1 - r^m t^m - r^n t^n}$ has $m - 1$ poles, and all these poles have absolute value strictly greater than 1, which leads for some $C > 0$ and $d > 1$ to the following estimate

$$\left| \sum_{mk+n\ell=N} \nu_{k,\ell} \right| = \left| [t^N]\, \frac{t^m - t^n}{1 - r^m t^m - r^n t^n} \right| \leq C d^{-N}.$$

Since $x(t) = o(1)$ as $t \to \infty$ we can for $\varepsilon > 0$ find $N_\varepsilon \in \mathbb{N}$ and $t_\varepsilon > 0$ such that $d^{-N_\varepsilon} \le \varepsilon$ and $|x(r^{N_\varepsilon} t)| \le \varepsilon$ for $t \ge t_\varepsilon$. We obtain

$$|y(pt) - y(qt)| \le \sum_{N \ge 1} C d^{-N} |x(r^N t)| \le C \frac{1 + \sup_{t \ge 0} |x(t)|)}{d-1} \varepsilon$$

for $t \ge t_\varepsilon$, which completes the proof of the first case. The following example shows that $x_N = O(1)$ does not imply (32): For the sequence $\mathbf{x} = (\mathbf{I} - p\mathbf{M}_p - q\mathbf{M}_q)\mathbf{y}$, where $y_N = \ln N\, \mathbb{I}_{\{N \ge 2\}}$, we obtain $x_N = p \ln \frac{1}{p} + q \ln \frac{1}{q} + O\left(N^{-1}\right)$ and $y_{\lfloor Np \rfloor} - y_{\lfloor Nq \rfloor} = \ln p - \ln q + O\left(N^{-1}\right)$, thus in hypothesis *i)*, $x_N = o(1)$ can not be weakened.

ii) $x_N = o\left(\frac{1}{\sqrt{\ln N}}\right)$ and $\frac{\ln p}{\ln q} \notin \mathbb{Q}$:

Note that now we have $p^k q^\ell = p^{k_1} q^{\ell_1}$ only if $(k, \ell) = (k_1, \ell_1)$. Speaking in terms of harmonic sums (cf. [5]), in the previous case several terms in (49) contributed to the same frequency, and there have been heavy cancellations in the corresponding amplitudes. Such cancellations do not occur in the present case. For $s > 0$ we define $I_s := \{(k, \ell) \in \mathbb{N}^2 : k + \ell \ge 1, ps < p^k q^\ell \le s\}$. If moreover $s \le 1$, there exists $n_s \in \mathbb{N}$ (in general not unique) such that $(\lfloor pn_s \rfloor, \lceil qn_s \rceil) \in I_s$. The asymptotics $n_s \sim \frac{\ln(1/s)}{p \ln(1/p) + q \ln(1/q)}$ holds, as $s \to 0$. We now define $k' := k - \lfloor pn_s \rfloor$ and $\ell' := \ell - \lceil qn_s \rceil$, and obtain by applying Stirling's formula

$$\binom{k+\ell}{k} p^k q^\ell = \binom{n_s}{\lfloor pn_s \rfloor} p^{\lfloor pn_s \rfloor} q^{\lceil qn_s \rceil} \exp\left(-\frac{(qk' - p\ell')^2}{2pqn_s}\right) \left(1 + O\left(\frac{|k'| + |\ell'|}{n_s} + \frac{|k'|^3 + |\ell'|^3}{n_s^2}\right)\right),$$

for $(k, \ell) \in I_s$, where $\{pn_s\} = pn_s - \lfloor pn_s \rfloor$ denotes the fractional part of pn_s. Stirling's formula also tells us $\binom{n_s}{\lfloor pn_s \rfloor} p^{\lfloor pn_s \rfloor} q^{\lceil qn_s \rceil} = O\left(n_s^{-\frac{1}{2}}\right)$. We have $\frac{qk - p\ell}{pq(k+\ell)} = \frac{qk' - p\ell' - \{pn_t\}}{pq(n_s + k' + \ell')}$, moreover it is quite an elementary task to show

$$\sum_{(k,\ell) \in I_s} \exp\left(-\frac{(qk' - p\ell')^2}{2pqn_s}\right) \frac{|qk' - p\ell' - \{pn_s\}|}{pq(n_s + k' + \ell')} = \Theta(1),$$

for we can compare the sum with an integral, and $n_s + k' + \ell' = k + \ell \ge \ln_{1/p}(1/s) > Cn_s$ for some $C > 0$, when $(k, \ell) \in I_s$. Thus we have

$$\sum_{(k,\ell) \in I_s} |\nu_{k,\ell}| = \Theta\left(\frac{1}{\sqrt{\ln \frac{1}{s}}}\right). \tag{50}$$

There is a function $\phi : \mathbb{R}_+ \to \mathbb{R}_+$, satisfying $\phi(t) = O(t)$, as $t \to 0$, and $\phi(t) = o\left(\frac{1}{\sqrt{\ln t}}\right)$, as $t \to \infty$, and moreover $\sup_{pt \le s \le p^{-1} t} \frac{\phi(s)}{\phi(t)} =: C' < \infty$, uniformly in $t > 0$, such that $|x(p^k q^\ell t)| \le \phi(p^k q^\ell t)$. We can, for $t > 1$ and $\tau := \lfloor \ln_{1/p} t \rfloor$, proceed as follows:

$$\begin{aligned} |y(pt) - y(qt)| &\le \sum_{k+\ell \ge 1} |\nu_{k,\ell}| |x(p^k q^\ell t)| \le C' \sum_{m \ge 0} \sum_{(k,\ell) \in I_{p^m}} |\nu_{k,\ell}| |\phi(p^m t)| = O\left(\sum_{m \ge 0} \frac{\phi(p^m t)}{\sqrt{m+1}}\right) \\ &= o\left(\sum_{m=0}^{\tau} \frac{1}{\sqrt{m+1}} \frac{1}{\sqrt{\tau + 1 - m}}\right) + O\left(\sum_{m \ge 0} \frac{p^m}{\sqrt{\tau + 1 + m}}\right) = o(1), \end{aligned} \tag{51}$$

since the finite sum tends to $\int_0^1 \frac{dx}{\sqrt{x(1-x)}} = \pi$ as $t \to \infty$. This completes the proof of the second case.

We still have to give an example demonstrating that the condition $x_N = O\left(\frac{1}{\sqrt{\ln N}}\right)$ does not imply (32) in the case $\frac{\ln p}{\ln q} \notin \mathbb{Q}$. The construction we will give works for $\frac{\ln p}{\ln q}$ which can not be approximated very well by rational numbers. It suffices to assume that $\alpha := \frac{\ln q}{\ln p}$ is an irrational algebraic number, in which case by Liouville's theorem there is $c > 0$ and $\mu \geq 2$ such that

$$\left|\alpha - \frac{m}{n}\right| > \frac{c}{n^\mu} \tag{52}$$

holds for positive integers m, n. For increasing sequences of positive numbers $(T_i)_{i\geq 1}$ and $(U_i)_{i\geq 1}$ to be suitably chosen, but related via $U_i = T_i^{2/3}$, we define

$$\mathbf{x} = \sum_{i\geq 0} \mathbf{x}^{(i)}, \quad \text{with} \quad \mathbf{x}^{(i)} = \sum_{(k,\ell)\in J_i} \varepsilon_{k,\ell} \mathbf{x}^{[m_{i,k,\ell}]},$$

where $J_i = \{(k,\ell) : k, \ell \geq 0, T_i p^k q^\ell > U_i\}$, $\varepsilon_{k,\ell} = \operatorname{sign}(\nu_{k,\ell})$, $m_{i,k,\ell} = \lfloor T_i p^k q^\ell \rfloor$ and

$$x_k^{[m]} = \sqrt{\frac{m}{\ln m}} \binom{k}{m} 2^{-k}.$$

Note that $x_k^{[m]}$ attains its largest value $\sqrt{\frac{m}{\ln m}}\binom{2m}{m}2^{-2m} \sim \frac{1}{\sqrt{\pi \ln m}}$ at $k \in \{2m-1, 2m\}$ and is very small for $k \notin \mathcal{A}_m := \left[2m - m^{0.6}, 2m + m^{0.6}\right]$, due to the estimate $x_k^{[m]} = e^{-\Omega\left(\frac{(k-2m)^2}{m} \wedge \frac{|k-2m|}{\sqrt{m}}\right)}$. Thus $x_k^{[m]} = O\left(\frac{1}{\sqrt{\ln k}}\right)$, and similarly $\Delta x_k^{[m]} = O\left(\frac{1}{\sqrt{k \ln k}}\right)$. The Poisson generating function $x^{[m]}(t) = e^{-t} \sum_{k\geq 0} \frac{t^k}{k!} x_k^{[m]} = \sqrt{\frac{m}{\ln m}} e^{-t/2} \frac{(t/2)^m}{m!}$ has a similar hump near $t = 2m$, with asymptotics $x^{[m]}(t) = \frac{1}{\sqrt{2\pi \ln m}} e^{-\frac{(t-2m)^2}{8m}} \left(1 + O\left(\frac{(t-2m)^3}{m^2}\right)\right)$. Condition (52) ensures that if T_i is large, the humps assembled in $x^{(i)}(t)$ (resp. in $\mathbf{x}^{(i)}$) do not overlap (i.e. the "supports" $(\mathcal{A}_{m_{i,k,\ell}})_{(k,\ell)\in J_i}$ of the humps are disjoint), as we will see by estimating the distance between adjacent humps from below:

Let $(k,\ell), (k',\ell') \in J_i$ with $p^k q^\ell < p^{k'} q^{\ell'}$. Then

$$\begin{aligned}\left|2\lfloor T_i p^k q^\ell \rfloor - 2\lfloor T_i p^{k'} q^{\ell'} \rfloor\right| &= 2T_i \left|p^{k+\ell\alpha} - p^{k'+\ell'\alpha}\right| + C \geq 2\ln\tfrac{1}{p}\, T_i p^{k+\ell\alpha} \left|k - k' + (\ell-\ell')\alpha\right| + C \\ &\geq 2\ln\tfrac{1}{p}\, T_i^{2/3} \frac{c}{(\ell \vee \ell')^\mu} + C \geq 2c \ln\tfrac{1}{p}\, T_i^{2/3} \left(\frac{\ln T_i}{3\ln\frac{1}{q}}\right)^{-\mu} + C,\end{aligned}$$

where $-2 < C < 2$. On the other hand, the largest interval among $(\mathcal{A}_{m_{i,k,\ell}})_{(k,\ell)\in J_i}$ has length $2T_i^{0.6}$. Now $x^{(i)}(t)$ is a sum of $|J_i| = O(\ln^2 T_i)$ terms $x^{[m_{i,k,\ell}]}(t)$, each exponentially small for $t \notin \mathcal{A}_{m_{i,k,\ell}}$, which results in

$$x^{(i)}(t) = \begin{cases} e^{-\Omega(\sqrt{2U_i - t})}, & \text{for } t \in [0, U_i], \\ e^{-\Omega(\sqrt{t - 2T_i})}, & \text{for } t \in [4T_i, \infty[, \\ \sum_{(k,\ell)\in J_i} \mathbb{1}_{\mathcal{A}_{m_{i,k,\ell}}}(t) \varepsilon_{k,\ell} x^{[m_{i,k,\ell}]}(t) + e^{-\Omega(t^{0.1})}, & \text{for } t \in]U_i, 4T_i[. \end{cases}$$

Assuming $T_i \geq T_{i-1}^2$ and T_1 large enough, humps of $x^{(i)}(t)$ and $x^{(j)}(t)$ will not overlap for $i \neq j$, which allows for the representation

$$x(t) = \sum_{i \geq 1} \sum_{(k,\ell) \in J_i} \mathbb{I}_{\mathcal{A}_{m_{i,k,\ell}}}(t) \varepsilon_{k,\ell} x^{[m_{i,k,\ell}]}(t) + e^{-\Omega(t^{0.1})}. \tag{53}$$

There is a similar representation for $\mathbf{x}$ from which we can immediately deduce that indeed $x_N = O\left(\frac{1}{\sqrt{\ln N}}\right)$ and $\Delta x_N = O\left(\frac{1}{\sqrt{N} \ln N}\right)$.

We now inductively define the sequence $(T_i)_{i \geq 1}$: Having determined T_j for $1 \leq j < i$ we choose T_i as follows. We let $y^{(i)}(t)$ and $y^{(<i)}(t)$ be the unique entire solutions of the equations

$$y^{(i)}(t) - py^{(i)}(pt) - qy^{(i)}(qt) = x^{(i)}(t), \qquad y^{(<i)}(t) - py^{(<i)}(pt) - qy^{(<i)}(qt) = \sum_{j=1}^{i-1} x^{(j)}(t).$$

Since $\sum_{j=1}^{i-1} x_N^{(j)} = e^{-\Omega(N^{0.1})}$, and thus satisfies condition *ii)* of this proposition, we can refer to (51) and deduce that there exists $T_i \geq T_{i-1}^2$ such that

$$\left| y^{(<i)}(pt) - y^{(<i)}(qt) \right| \leq 2^{-i}, \text{ for } t \geq U_i$$

and

$$\left| y^{(i)}(t) \right| \leq 2^{-i}, \text{ for } t \leq 4T_{i-1}.$$

The latter condition can be satisfied, since for $t \geq 0$ and some $C > 0$ we have, by Stirling's formula, $x^{(i)}(t) \leq \bar{x}^{(i)}(t) := C\left(\frac{e}{2}\frac{t}{U_i}\right)^{U_i}$, which implies $y^{(i)}(t) \leq \left(1 - p^{1+U_i} - q^{1+U_i}\right)^{-1} \bar{x}^{(i)}(t)$. With $x(t)$ constructed this way, and employing (49), (50) and (53) we derive

$$y(2T_i p) - y(2T_i q) = \sum_{(k,\ell) \in J_i} \frac{|\nu_{k,\ell}|}{\sqrt{2\pi \ln(T_i p^k q^\ell)}} + e^{-\Omega(U_i^{0.1})} + 2^{1-i} = \Theta\left(\int_0^{1/3} \frac{dx}{\sqrt{x(1-x)}}\right) = \Theta(1).$$

Depoissonizing, we obtain $\limsup_{N \geq 1} |y_{\lfloor Np \rfloor} - y_{\lfloor Nq \rfloor}| \geq \limsup_{i \geq 1} |y_{\lfloor \lfloor 2T_i \rfloor p \rfloor} - y_{\lfloor \lfloor 2T_i \rfloor q \rfloor}| > 0$, which shows that indeed (32) does not hold in this example. $\square$

References

[1] L. Devroye, *A limit theory for random skip lists*, Ann. Appl. Probab. **2** (1992), 597-609.

[2] M. Drmota, B. Gittenberger, *The distribution of nodes of given degree in random trees*, J. Graph Theory **31** (1999), 227-253.

[3] G. Fayolle, Ph. Flajolet, M. Hofri, *On a Functional Equation Arising in the Analysis of a Protocol for a Multi-Access Broadcast Channel*, Adv. Appl. Prob. **18** (1986), 441-472.

[4] P. Feldman, S. T. Rachev, L. Rüschendorf, *Limit theorems for recursive algorithms*, J. Comput. Appl. Math. **56** (1994), 169-182.

[5] P. Flajolet, X. Gourdon, P. Dumas, *Mellin Transforms and Asymptotics: Harmonic sums*, Theor. Comput. Sci. **144** (1995), 3-58.

[6] P. Flajolet, P. Grabner, P. Kirschenhofer, H. Prodinger, R. Tichy, *Mellin Transforms and Asymptotics: Digital sums*, Theor. Comput. Sci. **123** (1994), 291-314.

[7] P. Flajolet, A. M. Odlyzko, *Singularity analysis of generating functions*, SIAM J. Discrete Math. **3** (1990), 216-240.

[8] P. Flajolet, R. Sedgewick, *Digital Search Trees Revisited*, SIAM J. Comput. **15** (1986), 748-767.

[9] E. Fredkin, *Trie memory*, CACM **3** (1960), 490-500.

[10] I. G. Grama, *On moderate deviations for martingales*, Ann. Probab. **25** (1997), 152-183.

[11] W. Gutjahr, G. Ch. Pflug, *The asymptotic contour process of a binary tree is a Brownian excursion*, Stochastic Processes Appl. **41** (1992), 69-89.

[12] E. Haeusler, *On the rate of convergence in the central limit theorem for martingales with discrete and continuous time*, Ann. Probab. **16** (1988), 275-299.

[13] C. C. Heyde, B. M. Brown, *On the departure from normality of a certain class of martingales*, Ann. Math. Statist. **41** (1970), 2161-2165.

[14] P. Jacquet, M. Régnier, *Normal limiting distribution of the size of tries* Proc. Performance 87, North Holland, 1988, pp. 209-223.

[15] P. Jacquet, M. Régnier, *Normal limiting distribution for the size and the external path length of tries*, INRIA Research Report 827, 1988.

[16] P. Jacquet, W. Szpankowski, *Asymptotic behavior of the Lempel-Zif parsing scheme and digital search trees*, Theor. Comput. Sci. **144** (1995), 161-197.

[17] P. Jacquet, W. Szpankowski, *Analytical depoissonization and its applications*, Theor. Comput. Sci. **201** (1998), 1-62.

[18] P. Kirschenhofer, H. Prodinger, *On some Applications of Formulae of Ramanujan in the Analysis of Algorithms*, Mathematika **38** (1991), 14-33.

[19] D. E. Knuth *The Art of Computer Programming, Vol. 3*, Addison-Wesley, Reading MA, 1973.

[20] G. Louchard, *Trie size in a dynamic list structure*, Random Struct. Algorithms **5** (1994), 665-702.

[21] H. M. Mahmoud *Evolution of Random Search Trees*, Wiley, New York, 1992.

[22] H. M. Mahmoud, P. Flajolet, P. Jacquet, M. Régnier, *Analytic variations on bucket selection and sorting*, Acta Inform. **36** (2000), 735-760.

[23] H. M. Mahmoud, B. Pittel, *Analysis of the space of search trees under the random insertion algorithm*, J. Algorithms **10** (1989), 52-75.

[24] H. M. Mahmoud, R. T. Smythe, *Probabilistic analysis of bucket recursive trees*, Theor. Comput. Sci. **144** (1995), 221-249.

[25] B. Pittel, *Paths in a random digital tree: Limiting distributions*, Adv. Appl. Probab. **18** (1986), 139-155.

[26] S. T. Rachev, L. Rüschendorf, *Probability metrics and recursive algorithms*, Adv. Appl. Prob. **27** (1995), 770-799.

[27] M. Régnier, *A limiting distribution for quicksort*, RAIRO, Theoretical Informatics and Applications **23** (1989), 335-343.

[28] U. Rösler, *A limit theorem for "QUICKSORT"*, RAIRO, Theoretical Informatics and Applications **25** (1991), 85-100.

[29] U. Rösler, *On the analysis of stochastic divide and conquer algorithms*, Algorithmica **29**, (2001), 238-261.

[30] U. Rösler, L. Rüschendorf, *The contraction method for recursive algorithms*, Algorithmica **29**, (2001), 3-33.

[31] S. Roura *Divide-and-Conquer Algorithms and Data Structures* Thesis, UPC, 1997.

[32] W. Schachinger, *On the variance of a class of inductive valuations of data structures for digital search*, Theor. Comput. Sci. **144** (1995), 251-275.

[33] A. N. Shiryaev *Probability, 2nd ed.*, Springer, 1996.

[34] W. Szpankowski, *Solution of a Linear Recurrence Equation Arising in the Analysis of Some Algorithms*, SIAM J. Alg. Disc. Methods **8** (1987), 233-250.

[35] D. Williams *Probability with Martingales*, Cambridge University Press, 1991.