fractal and fractional

MDPI

Article

The Rates of Convergence for Functional Limit Theorems with Stable Subordinators and for CTRW Approximations to Fractional Evolutions

Vassili N. Kolokoltsov [1,2]

[1] Faculty of Computation Mathematics and Cybernetics, Moscow State University, 119991 Moscow, Russia; v.n.kolokoltsov@gmail.com or v.kolokoltsov@warwick.uk

[2] Faculty of Economics, National Research University Higher School of Economics, 101000 Moscow, Russia

Abstract: From the initial development of probability theory to the present days, the convergence of various discrete processes to simpler continuous distributions remains at the heart of stochastic analysis. Many efforts have been devoted to functional central limit theorems (also referred to as the invariance principle), dealing with the convergence of random walks to Brownian motion. Though quite a lot of work has been conducted on the rates of convergence of the weighted sums of independent and identically distributed random variables to stable laws, the present paper is the first to supply the rates of convergence in the functional limit theorem for stable subordinators. On the other hand, there is a lot of activity on the convergence of CTRWs (continuous time random walks) to processes with memory (subordinated Markov process) described by fractional PDEs. Our second main result is the first one yielding rates of convergence in such a setting. Since CTRW approximations may be used for numeric solutions of fractional equations, we obtain, as a direct consequence of our results, the estimates for error terms in such numeric schemes.

Keywords: rates of convergence; functional limit theorem with stable processes; fractional equations; smooth Wasserstein distances; Kolmogorov's distance; continuous time random walks (CTRW); subordinated Markov processes

check for updates

Citation: Kolokoltsov, V.N. The Rates of Convergence for Functional Limit Theorems with Stable Subordinators and for CTRW Approximations to Fractional Evolutions. *Fractal Fract.* **2023**, *7*, 335. https://doi.org/10.3390/fractalfract7040335

Academic Editors: Bruce Henry, Dongfang Li and Xiaoli Chen

Received: 1 February 2023
Revised: 28 March 2023
Accepted: 29 March 2023
Published: 17 April 2023

1. Introduction

From the initial development of probability theory to the present days, the convergence of various discrete processes to simpler continuous distributions remain at the heart of stochastic analysis. Many efforts have been devoted to functional central limit theorems (also referred to as the invariance principle), dealing with the convergence of random walks to Brownian motion. For the first results on the rates of convergence in the functional CLT, we can refer to [1,2], with explicit numeric constants provided in [3]. For recent development in the convergence rates for functional CLT and its generalizations, see, for example, refs. [4,5] and references therein.

There are many papers devoted to the rates of convergence of random sequences to stable laws; see, for example, refs. [6–9], and the references therein. However, though there is a serious interest in the function limit theorems with convergence to stable process, see, for example, refs. [10,11] and references therein, there seems to be yet no results on the rates of convergence in this setting. To the best knowledge of the author, the present paper is the first one supplying the rates of convergence in the functional limit theorem for stable subordinators. Moreover, we give explicit numeric constants, which seems to be new, even for the convergence to stable laws. As a technical novelty, we introduce and use effectively the smooth Wasserstein–Kantorovich metrics for measures based on the functional spaces of Hölder continuous functions. This approach makes our estimates rather flexible for adjusting to various stability indexes and various levels of regularity.

Let us notice that the results on the rates of weak convergence of processes are often formulated in terms of certain topologies on path spaces (such as Skorokhod topologies J_1 and M_1) and related metrics. Our convergence rates are formulated in simpler terms of the uniform convergence of marginal (in time) distributions. Alternative measures of convergence could be also various pseudo-distances ν_r between distribution functions, say F, G, defined, for instance, by the inequalities of type $|F(s) - G(s)| \leq A|s|^r \nu_r$ (that have to hold for small s); see [12] and the numerous references therein.

Another popular topic deals with the convergence of CTRWs (continuous random walks) to processes with memory described by fractional PDEs. The CTRWs were introduced in [13]. They found numerous applications in physics. The scaling limits of these CTRWs were analyzed by many authors; see for example, [14–16]. The crucial point (realized initially by physicists, see [17] or [13]) is the fact that the limits of scaled CTRW yield Markov processes, time changed by inverse stable subordinators that solve fractional in time PDEs. The scaling limit for the position dependent CTRW was developed in [18]. Refs. [12,19] present convergence rates for a particular case of convergence to fractional distributions, including double-array schemes, but there seem to be no results available for the rates of convergence for a functional limit to fractional evolutions. Our final result is the first one providing the rates in this setting, and it deals with a rather general model of generalized (position depending) CTRWs. The methods used for proofs are based essentially on the theory of operator semigroups in the spirit of ideas, proposed initially in [18,20] and developed further in [21].

Since CTRW approximations may be used for numeric solutions of fractional equations (see [22]), we get, as a direct consequence of our results, the estimates for error terms in such numeric schemes. Numeric solutions to fractional PDEs is currently a rather hot topic. We refer to papers [23,24] for the probabilistic (Monte Carlo type) approach to numeric solutions of fractional equations and to [25] and references therein for (much more abundant) work on deterministic approaches.

The paper is organized as follows. In the next section, we formulate carefully our main results, discuss some links with related popular research topics and present the most basic examples. Section 3 is devoted to a review of some auxiliary facts playing a crucial role in our arguments. Though they are mostly known, we present them for handy references, stressing also some particular points and consequences related to our purposes. The rest of the paper is devoted to the proof of the main results, together with some intermediate steps that may be of independent interest.

We conclude the introduction with certain basic notations that will be used in the paper without further reminder.

Letters $\mathbf{P}$ and $\mathbf{E}$ will be used to denote probability and expectation. The indicator function of a set M will be denoted by $\mathbf{1}(M)$.

As usual, let $C(\mathbf{R}^d)$ denote the space of bounded continuous functions on $\mathbf{R}^d$ equipped with the standard sup-norm $\|.\|$. For $k \in \mathbf{N}$, let $C^k = C^k(\mathbf{R}^d)$ denote the space of k times continuously differentiable functions on $\mathbf{R}^d$ with bounded derivatives equipped with the standard norm

$$\|f\|_{C^k} = \max\{\|f\|, \max_{m=1}^{k} \|f^{(m)}\|\},$$

where $\|f^{(k)}\|$ denotes the maximum of sup-norms of all partial derivatives of f of order k. Let $C_\infty(\mathbf{R}^d)$ denote the closed subspace of $C(\mathbf{R}^d)$ consisting of functions vanishing at infinity, $C^k_\infty(\mathbf{R}^d)$ the closed subspace of $C^k(\mathbf{R}^d)$ consisting of functions such that all its derivatives, up to order k, belong to $C_\infty(\mathbf{R}^d)$.

For $\alpha \in (0,1]$, let $H_\alpha = H_\alpha(\mathbf{R}^d)$ denote the space of bounded α-Hölder continuous functions equipped with the norm

$$\|f\|_\alpha = \max\{\|f\|, \sup_{0<|x-y|\leq 1} \frac{|f(x) - f(y)|}{|x-y|^\alpha}\}, \tag{1}$$

and for $k \in \mathbf{N}$, let $H_{k,\alpha} = H_{k,\alpha}(\mathbf{R}^d)$ denote the subspace of $C^k(\mathbf{R}^d)$ of functions with α-Hölder continuous derivatives of order k equipped with the norm

$$\|f\|_{k,\alpha} = \max\{\|f\|_{C^k}, \|f^{(k)}\|_\alpha\}, \tag{2}$$

where $\|f^{(k)}\|_\alpha$ is the maximum of the H_α- norms of all partial derivatives of f of order k.

2. Formulation of the Main Results

In this section, we formulate our main results. In the first subsection, we shall deal with the convergence of random walks to stable subordinators in three senses: in smooth Wasserstein distances (based on the functional spaces of Hölder continuous functions), in Kolmogorov's distance and in various versions of convergence that are uniform in time. In the second subsection, we formulate our result on the rates of convergence for CTRW approximations. Finally, we supply some basic examples and recall the main link with the fractional PDEs.

2.1. Convergence of Random Walks to Stable Subordinators

Everywhere in this paper, we assume that τ_i, $i \in \mathbf{N}$, is a sequence of positive independent and identically distributed random variables having probability density $p(y)$ on $\mathbf{R}_+$ such that $p(0) = 0$, $p(x)$ is continuously differentiable, and $p(y) = y^{-1-\beta}$ for $y \geq B$ with some $\beta \in (0,1)$ and $B > 0$ such that $\beta B^\beta > 1$. The latter condition comes from the requirement that $\int_B^\infty p(y)dy \leq 1$. Let

$$\Phi_t^h = \sum_{i=1}^{[t/h]} h^{1/\beta}\tau_i,$$

be the corresponding scaled random walk (where we set $\Phi_t^h = 0$ for $t < h$). We shall denote by $V_h^{[t/h]}$ the transition operators of the discrete Markov chain Φ_t^h.

Remark 1. *(i) The exact power law dependence of the density $p(x)$ for large x is the standard simplifying assumption for the analysis of the central limit theorem for stable laws, see, for example, [7]. However, for convergence to stable distributions, more weaker assumptions can also be found in the literature; see, for example, [8]. We decided to stick here to the simplest assumption to show the idea of our framework for the calculations of the rates of convergence in the most transparent way. (ii) Requirements that $p(0) = 0$ and $p(x)$ are continuously differentiable are not needed for the validity of Theorems 1 and 2.*

Let T_β^t be the Feller semigroup of the β-stable Lévy subordinator $\hat{\Sigma}_t^\beta$ in $\mathbf{R}$ generated by the operator

$$\hat{L}_\beta f(x) = \int_0^\infty \frac{f(x+y) - f(x)}{y^{1+\beta}} dy.$$

For convenience, we chose $\hat{L}$ to differ from the case of the standard stable subordinators (see (34)) by a multiplier. Therefore, if Σ_t^β is the standard β-subordinator, then

$$\hat{\Sigma}_t^\beta = \Sigma_{t\Gamma(1-\beta)/\beta}^\beta. \tag{3}$$

It is well known that the operators $V_h^{[t/h]}$ converge to the Feller semigroup T_β^t of the stable subordinator; see, for example, [15] or [18]. The next three theorems specify the rates of convergence in three different senses.

Theorem 1. *(i) If*

$$\beta < (1-\beta)/\beta, \tag{4}$$

then for any $\alpha \in (\beta, \min\{1, (1-\beta)/\beta\}]$,

$$\sup_{s \le t} \|(V_h^{[s/h]} - T_\beta^s)f\| \le h^{\min(1,(1-\beta)/\beta)} t \left(\frac{2B}{1-\beta} + \frac{4\alpha}{\beta^2(1-\beta)(\alpha-\beta)} \right) \|f\|_{1,\alpha}. \tag{5}$$

For instance, if $\beta \le 1/2$ *and thus* $1-\beta \ge \beta$, *choosing* $\alpha = 1$ *yields the estimate*

$$\sup_{s \le t} \|(V_h^{[s/h]} - T_\beta^s)f\| \le ht \left(\frac{2B}{1-\beta} + \frac{4}{\beta^2(1-\beta)^2} \right) \|f\|_{1,1}, \tag{6}$$

and if $\beta < (1-\beta)/\beta \le 1$, *then, choosing* $\alpha = (1-\beta)/\beta$ *yields the estimate*

$$\sup_{s \le t} \|(V_h^{[s/h]} - T_\beta^s)f\| \le h^{(1-\beta)/\beta} t \left(\frac{2B}{1-\beta} + \frac{4}{\beta^2(1-\beta-\beta^2)} \right) \|f\|_{1,(1-\beta)/\beta}. \tag{7}$$

(ii) If $\beta \ge (1-\beta)/\beta$, *then*

$$\sup_{s \le t} \|(V_h^{[s/h]} - T_\beta^s)f\| \le h^{1-\beta} t \left(\frac{2B}{1-\beta} + \frac{10}{\beta^2(1-\beta)^2} \right) \|f\|_{1,(1-\beta)/\beta}. \tag{8}$$

These results can be more elegantly written using certain popular norms, which we shall review now.

For a subspace B of $C(\mathbf{R})$, which is itself a Banach space equipped with the norm $\|.\|_B$, one can introduce a metric on the set of real random variables by the equation

$$d_B(X, Y) = \sup\{|\mathbf{E}f(X) - \mathbf{E}f(Y)| : \|f\|_B \le 1\}. \tag{9}$$

For instance, if $B = C^k$, $k \in \mathbf{N}$, the corresponding metrics d_{C^k} are often referred to as the smooth Wasserstein metrics (see, for example, [8]). Intermediate metrics can be defined by using the spaces of Hölder functions H_α as the subspace B. For the space H_1, the corresponding metric is referred to as the bounded Lipschitz metric.

For instance, with these notations, the inequality (5) can be equivalently written

$$\sup_{s \le t} d_{H_{1,\alpha}}(\Phi_s^h, \hat{\Sigma}_s^\beta) \le h^{\min(1,(1-\beta)/\beta)} t \left(\frac{2B}{1-\beta} + \frac{4\alpha}{\beta^2(1-\beta)(\alpha-\beta)} \right). \tag{10}$$

Similarly, the other statements of Theorem 1 can be rewritten.

Notice that estimates (5) with flexible α may require less regularity than estimates (6) and (7), which are more transparent but less powerful than (5).

Remark 2. *The rates of convergence studied in the literature relate the convergence of weighted sums of* τ_i *to stable distributions. For instance, in our notations, the results of [8] represent the estimates for* $d_{C^2}(\Phi_1^h, \hat{\Sigma}_1^\beta)$ *(where* $h = 1/n$ *in [8]). Our estimates, apart from being functional (giving uniform estimates for* $s \le t$*), are based on more general spaces* H_α*, thus requiring less regularity. We also supply the exact values for the constants involved.*

The convergence of a sequence of random variables in smooth Wasserstein metrics implies its weak convergence. Now, we shall enhance the convergence with respect to smooth Wasserstein metrics (as given by Theorem 1) to a stronger convergence in Kolmogorov's metric.

Theorem 2. *(i) If* $\beta < (1-\beta)/\beta$ *any* $\alpha \in (\beta, \min\{1, (1-\beta)/\beta\}]$, *then*

$$\sup_{K>0} \sup_{s \in [t_0, t]} |\mathbf{P}(\Phi_s^h \le K) - \mathbf{P}(\hat{\Sigma}_s^\beta < K)|$$

$$\leq \frac{1}{2}(t_0^{-1/\beta}+3)\left(\frac{2B}{1-\beta}+\frac{4\alpha}{\beta^2(1-\beta)(\alpha-\beta)}\right)^{1/(\alpha+2)}(h^{\min(1,(1-\beta)/\beta)}t)^{1/(\alpha+2)}, \tag{11}$$

for any $0 < t_0 \leq t$ *and*

$$\sup_{K>k}\sup_{s\leq t}|\mathbf{P}(\Phi_s^h \leq K)-\mathbf{P}(\hat{\Sigma}_s^\beta < K)|$$

$$\leq [k^{-1}(\Gamma(1-\beta)/\beta)^{1/\beta}+3/2]\left(\frac{2B}{1-\beta}+\frac{4\alpha}{\beta^2(1-\beta)(\alpha-\beta)}\right)^{1/(\alpha+2)}(h^{\min(1,(1-\beta)/\beta)}t)^{1/(\alpha+2)}, \tag{12}$$

for any $t > 0, k > 0$.

(ii) If $\beta \geq (1-\beta)/\beta$, *then*

$$\sup_{K>0}\sup_{s\in[t_0,t]}|\mathbf{P}(\Phi_s^h \leq K)-\mathbf{P}(\hat{\Sigma}_s^\beta < K)|$$

$$\leq \frac{1}{2}(t_0^{-1/\beta}+3)\left(\frac{2B}{1-\beta}+\frac{10}{\beta^2(1-\beta)^2}\right)^{\beta/(\beta+1)}(h^{1-\beta}t)^{\beta/(\beta+1)}, \tag{13}$$

for any $0 < t_0 \leq t$ *and*

$$\sup_{K>k}\sup_{s\leq t}|\mathbf{P}(\Phi_s^h \leq K)-\mathbf{P}(\hat{\Sigma}_s^\beta < K)|$$

$$\leq [k^{-1}(\Gamma(1-\beta)/\beta)^{1/\beta}+3/2]\left(\frac{2B}{1-\beta}+\frac{10}{\beta^2(1-\beta)^2}\right)^{\beta/(\beta+1)}(h^{1-\beta}t)^{\beta/(\beta+1)}, \tag{14}$$

for any $t > 0, k > 0$.

It is seen that the difference between the distribution functions of the random variables Φ_s^h and $\hat{\Sigma}_s^\beta$ may not be easy to control for both s and K small. This effect can be expected, as Φ_s^h is discrete in s and $\hat{\Sigma}_s^\beta$ has a density function that increases to infinity, as $s \to 0$.

Recall that the Kolmogorov distance between random variables X and Y is defined by the formula

$$d_{Kol}(X,Y) = \sup_z |\mathbf{P}(X \leq z) - \mathbf{P}(Y \leq z)|. \tag{15}$$

Thus, inequality (13) rewrites as

$$\sup_{t_0<s\leq t} d_{Kol}(\Phi_s^h,\hat{\Sigma}_s^\beta) \leq \frac{1}{2}(t_0^{-1/\beta}+3)\left(\frac{2B}{1-\beta}+\frac{10}{\beta^2(1-\beta)^2}\right)^{\beta/(\beta+1)}(h^{1-\beta}t)^{\beta/(\beta+1)}. \tag{16}$$

If $\beta \leq 1/2$, then from (11) one derives the estimate

$$\sup_{t_0\leq s\leq t} d_{Kol}(\Phi_s^h,\hat{\Sigma}_s^\beta) \leq \frac{1}{2}(t_0^{-1/\beta}+3)\left(\frac{2B}{1-\beta}+\frac{4}{\beta^2(1-\beta)^2}\right)^{1/3}(ht)^{1/3}, \tag{17}$$

and if $\beta^2 < 1-\beta \leq \beta$, then from (11) one derives the estimate

$$\sup_{t_0\leq s\leq t} d_{Kol}(\Phi_s^h,\hat{\Sigma}_s^\beta) \leq \frac{1}{2}(t_0^{-1/\beta}+3)\left(\frac{2B}{1-\beta}+\frac{4}{\beta^2(1-\beta-\beta^2)}\right)^{\beta/(\beta+1)}h^{(1-\beta)/(1+\beta)}t^{\beta/(\beta+1)}. \tag{18}$$

In Theorem 2, we estimated $|\mathbf{P}(\Phi_s^h \leq K)-\mathbf{P}(\hat{\Sigma}_s^\beta < K)|$ uniformly in K for bounded s. Now we shall estimate $|\mathbf{P}(\Phi_s^h \leq K)-\mathbf{P}(\hat{\Sigma}_s^\beta < K)|$ uniformly in s for bounded K, which is crucial for the proof of our last theorem.

Below, we shall use the constant t_0 (depending only on β), which is given explicitly in Proposition 6.

Theorem 3. *If $\beta \le 1/2$, then*

$$\sup_{K \le K_0} \sup_{s \ge 1} |\mathbf{P}(\Phi_s^h \le K) - \mathbf{P}(\hat{\Sigma}_s^\beta < K)|$$

$$\le 2h^{1/(3+\beta)} \left[\left(\frac{2B}{1-\beta} + \frac{4}{\beta^2(1-\beta)^2} \right)^{1/3} + 2^{2/\beta} K_0 \right] \tag{19}$$

for any K_0 and small enough h, namely whenever $h^{\beta/(\beta+3)} \le 1/t_0$.

If $\beta^2 < 1 - \beta \le \beta$, then

$$\sup_{K \le K_0} \sup_{s \ge 1} |\mathbf{P}(\Phi_s^h \le K) - \mathbf{P}(\hat{\Sigma}_s^\beta < K)|$$

$$\le 2h^{(1-\beta)/(1+\beta+\beta^2)} \left[\left(\frac{2B}{1-\beta} + \frac{4}{\beta^2(1-\beta-\beta^2)} \right)^{\beta/(\beta+1)} + 2^{2/\beta} K_0 \right] \tag{20}$$

for any K_0 and small enough h, namely whenever $h^{\beta(1-\beta)/(1+\beta+\beta^2)} \le 1/t_0$.

If $\beta \ge (1-\beta)/\beta$, then

$$\sup_{K \le K_0} \sup_{s \ge 1} |\mathbf{P}(\Phi_s^h \le K) - \mathbf{P}(\hat{\Sigma}_s^\beta < K)|$$

$$\le 2h^{\beta(1-\beta)/(1+\beta+\beta^2)} \left[\left(\frac{2B}{1-\beta} + \frac{10}{\beta^2(1-\beta)^2} \right)^{\beta/(\beta+1)} + 2^{2/\beta} K_0 \right] \tag{21}$$

for any K_0 and small enough h, namely whenever $h^{\beta^2(1-\beta)/(1+\beta+\beta^2)} \le 1/t_0$.

Moreover, one has similar estimates for

$$\sup_{k \le K \le K_0} \sup_{s > 0} |\mathbf{P}(\Phi_s^h \le K) - \mathbf{P}(\hat{\Sigma}_s^\beta < K)| \tag{22}$$

with the same power dependence of h but with different constants (depending on k, K_0, β, B).

Finally,

$$\int_0^\infty |\mathbf{P}(\Phi_s^h \le K) - \mathbf{P}(\hat{\Sigma}_s^\beta < K)|\, ds \le (1 + K + K^{-1}) C(\beta) h^{\varkappa(\beta)} \tag{23}$$

with some constants $C(\beta) > 0$ and $\varkappa(\beta) \in (0,1)$ (that can be written explicitly). For instance, if $\beta \ge (1-\beta)/\beta$, then

$$\varkappa(\beta) = \frac{\beta(1-\beta)^2}{1+\beta+\beta^2}.$$

2.2. Convergence of Position-Dependent CTRWs to Fractional Evolutions

In this subsection, we formulate our second main result concerning the rates of convergence of generalized (position dependent) CTRWs to Markov processes with memory described by fractional PDEs.

The inverse process to the scaled random walk Φ_t^h is defined by the equation

$$N_K^h = \sup\{t : \Phi_t^h < K\}.$$

In renewal theory, this process is also referred to as the number of renewals. It can be defined by several other equivalent ways. They are insightful, and we provide them for the convenience of the readers in our less standard (scaled) version. Namely, N_K^h can be equivalently defined by (i) the requirement that the events $(N_K^h > t)$ and $(\Phi_t^h < K)$ coincide (for all t, K), (ii) by the requirement that

$$\Phi^h_{N^h_K-h} < K \le \Phi^h_{N^h_K},$$

or (iii) by the counting-number-of-events formula

$$N^h_K/h = 1 + \sum_{n=1}^{\infty} \mathbf{1}(\Phi^h_{nh} < K) = 1 + \sum_{n=1}^{\infty} \mathbf{1}\left(\sum_{i=1}^{n} h^{1/\beta}\tau_i < K\right).$$

In particular, the key relation is

$$\mathbf{P}(N^h_K > t) = \mathbf{P}(\Phi^h_t < K).$$

Suppose $X^h_1, X^h_2, \cdots$ is a sequence of independent and identically distributed random variables in $\mathbf{R}^d$ such that the distribution of each X^h_i is given by a probability measure $\mu^h_{space}(dx)$ that depends on h. The standard (scaled) continuous time random walk (CTRW) is a random process given by the random sum

$$\sum_{j=1}^{N^h_t/h} X^h_i.$$

In position-dependent CTRW, the jumps X^h_i are not independent, but each X^h_i depends on the position of the process before this jump. The natural general formulation can be given in terms of discrete Markov chains as follows.

Let $O^n_h(x)$ be a family (depending on a parameter $h > 0$) of discrete-time Markov chains in $\mathbf{R}^d$, independent of the sequence $\{\tau_i\}$. Let U_h be a transition operator of $O^1_h(x)$ so that

$$U_h f(x) = \mathbf{E}O^1_h(x) = \int f(y)\mu^h_{space}(x, dy), \tag{24}$$

with some family of stochastic kernels $\mu^h_{space}(x, dy)$.

The discrete-time process

$$O_h^{N^h_t/h}(x), \tag{25}$$

with the transition operators $U_h^{N^h_t/h}$, is a generalized scaled (position dependent) continuous time random walk (CTRW) arising from transitions $\mu^h_{space}(x, dy)$ and the sequence of waiting times $\{\tau_i\}$. Thus, it is a Markov chain, time changed by the inverse process to the scaled random walk built on the sequence $\{\tau_i\}$.

Let us denote by

$$\sigma_y = \max\{t : \hat{\Sigma}^\beta_t \le y\}$$

the inverse process to the stable subordinator $\hat{\Sigma}^\beta_t$.

As was already mentioned, it has been proved by many authors under various assumptions that if $(U_h - 1)/h$ converges to an operator L, which generates a Feller process $X_t(x)$ (here, as in (25), x stands for the initial point of the process) with a Feller semigroup F_t, then the processes of CTRW $O_h^{N^h_t/h}(x)$ converges in distribution to the time-changed Markov process $X_t(x)$, that is, to the process $X_{\sigma_t}(x)$. Our final result supplies the rates of convergence for this scaling limit in terms of marginal distributions.

Theorem 4. *Let the Feller semigroup $F_t = e^{tL}$ of a Feller process $X_t(x)$ and the family of contractions U_h satisfy the assumptions of Proposition 1 (given below) with $\varkappa_h + \epsilon_h = h^\omega$ with some $\omega > 0$. Then, if $m > 0$, then*

$$\sup_x |\mathbf{E}f(O_h^{N^h_t/h}(x)) - \mathbf{E}f(X_{\sigma_t}(x))| = \|\mathbf{E}U_h^{N^h_t/h} f - \mathbf{E}F_{\sigma_t} f\|$$

$$\le 4(1 + \max(1, m/\omega))(1 + t + t^{-1}) \ln(1/h)^{-1} \|f\|_D, \tag{26}$$

for all $t > 0$, $f \in D$, and $h \le h_0$ with sufficiently small h_0 (depending on β, B).

If $m = 0$, then the convergence rates on the right-hand side of (26) can be improved to the power-type estimate

$$\sup_x |\mathbf{E} f(O_h^{N_t^h/h}(x)) - \mathbf{E} f(X_{\sigma_t}(x))| \le C(\beta)(1 + t + t^{-1}) h^{\varkappa(\beta)} \|f\|_D, \tag{27}$$

with some $C(\beta)$ and $\varkappa(\beta) \in (0,1)$. For instance, if $\beta \ge (1-\beta)/\beta$, then

$$\varkappa(\beta) = \min\left\{ \frac{\omega}{\beta+1}, \frac{\beta(1-\beta)^2}{1+\beta+\beta^2} \right\}.$$

A numeric estimate for h_0 is easy to derive for given β, B, but a general formula is lengthy and not very revealing; hence, it is omitted.

2.3. Link with Fractional Equations and Fractional Distributions; Examples

It is well known, see, for example, monographs [16,20,26], that the subordinated limiting evolution described by the operators $\mathbf{E} f(X_{\sigma_t}(x)) = \mathbf{E} F_{\sigma_t} f(x)$ solves fractional in time-differential equations. Namely, under the conditions of Theorem 4, the function $f_t(x) = \mathbf{E}(F_{\sigma_t} f)(x)$ satisfies the equation

$$D^{\beta}_{0+\star} f_t(x) = L f_t(x), \quad f_0(x) = f(x), \tag{28}$$

where $D^{\beta}_{0+\star}$ is the Caputo–Dzherbashian derivative of order β acting on the variable t, and the operator L acts on the variable x. Thus, Theorem 4 provides the rates of convergence for the CTRW approximations of the solutions to fractional equations.

Recall that the Caputo–Dzherbashian derivative of order $\beta \in (0,1)$ is given by the formula

$$D^{\beta}_{0+\star} f(s) = \frac{1}{\Gamma(1-\beta)} \int_0^x (s-r)^{-\beta} f'(r)\, dr,$$

see, for example, [27].

The most basic examples are supplied by operators L that are diffusion operators

$$L_{dif} f(x) = \frac{1}{2} \mathrm{tr}\left(G(x) \frac{\partial^2 f}{\partial x^2}(x) \right) = \frac{1}{2} \sum_{ij} G_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x), \quad x \in \mathbf{R}^d, \tag{29}$$

with $G(x)$, a positive $d \times d$-matrix, or the generators of stable-like processes

$$L^{\beta}_{st} f(x) = \int_{S^{d-1}} |(s, \nabla f(x))|^{\beta} \mu(x,s)\, ds, \quad x \in \mathbf{R}^d, \tag{30}$$

with $\beta \in (0,2)$, $\mu(x,s)$ and even in s positive functions on $\mathbf{R}^d \times S^{d-1}$ and ds Lebesgue measures on the unit sphere S^{d-1}. In particular, if $\mu(x,s) = 1$ (that is, the spectral measure $\mu(x,s)\, ds$ is uniform), then $L^{\beta}_{st} = \sigma |\Delta|^{\beta/2}$ (with σ a constant depending on d) becomes a standard fractional Laplacian. Standard CTRW approximations to these operators are well presented in the literature, see [16,20,26].

In both cases, if the coefficients are sufficiently smooth, one can take the space $D = C_\infty(\mathbf{R}^d) \cap C^4(\mathbf{R}^d)$, for which the assumptions of Proposition 1 (and thus of Theorem 4) hold with $\varkappa_h = \epsilon_h = h$. In fact, the second required condition (42) usually holds trivially for the standard constructions of CTRW approximation (as mentioned above), and the first required condition (41) holds trivially with $\varkappa_h = h$ whenever D is a subspace such that L^2 is a bounded operator $D \to C_\infty(\mathbf{R}^d)$. Estimates with $\varkappa_h = h^\omega$ with some $\omega > 0$ can be obtained for larger space $D = H_{1,\alpha}$ in the same way, as it is done for the stable subordinator in Section 4 below.

If the coefficients of these operators are constant (they specify a spatially homogeneous process), then the assumptions of Proposition 1 hold with $m = 0$, thus allowing for better power law rates of convergence (rather than our general logarithmic rates).

Let us present the simplest possible example. Namely, the operator $L = \Delta/2$ in $\mathbf{R}^d$ is known to generate a standard Brownian motion B_t with the standard heat equation semi-group F_t. Let us approximate it by the standard random walk with the transition operator

$$U_h f(x) = \sum_{i=1}^{d} \frac{f(x+he_i) + f(x-he_i) - 2f(x)}{2h^2},$$

where e_i is the standard basis in $\mathbf{R}^d$. Choosing $D = C_\infty(\mathbf{R}^d) \cap C^4(\mathbf{R}^d)$, we find ourselves under the assumptions of Proposition 1 with $\varkappa_h = \epsilon_h = h$, $m = 0$ and $l = 1$. Choosing β such that $\beta \geq (1-\beta)/\beta$ and applying Theorem 4, we obtain the rates of convergence of the corresponding CTRW $O_h^{N_t^h/h}$ to the subordinated Brownian motion (fractional diffusion):

$$\sup_x |\mathbf{E} f(O_h^{N_t^h/h}(x)) - \mathbf{E} f(B_{\sigma_t}(x))| = \|\mathbf{E} U_h^{N_t^h/h} f - \mathbf{E} F_{\sigma_t} f\|$$

$$\leq C(\beta)(1 + t + t^{-1}) h^{\beta(1-\beta)^2/(1+\beta+\beta^2)} \|f\|_{C^4}, \tag{31}$$

for sufficiently small h. This can be also rewritten in terms of the smooth Wasserstein distances (9):

$$d_{C^4}(O_h^{N_t^h/h}, B_{\sigma_t}) \leq C(\beta)(1 + t + t^{-1}) h^{\beta(1-\beta)^2/(1+\beta+\beta^2)} \|f\|_{C^4}. \tag{32}$$

The same estimate holds if instead of the Brownian motion, one takes the stable process Y_t^α in $\mathbf{R}^d$ generated by the fractional Laplacian $|\Delta|^{\alpha/2}$, $\alpha \in (1,2)$, and the corresponding standard CTRW approximation. The distribution of the subordinated process $Y_{\sigma_t}^\alpha$, where σ_t is the inverse to the stable subordinator Σ_t^β, was called fractional stable in [14] because, as it was found, it could be obtained as the distribution of the ratio of two independent stable random variables (the denominator taken in power β/α). For completeness, let us write down the marginal probability density $q_t(y)$ of the process $Y_{\sigma_t}^\alpha$ (started at $y_0 = 0$):

$$q_t(y) = \int_0^\infty S_\alpha\Big(1, (z/t)^{\beta/\alpha} y\Big)(z/t)^{d\beta/\alpha} S_\beta(1,z)\, dz, \tag{33}$$

where S_α and S_β are the densities of the symmetric α-stable distribution in $\mathbf{R}^d$ and of the β-stable subordinator (with the unit scaling coefficient). Density (33) was obtained in [14], as the limit of CTRWs, for the case of $t = d = 1$. A noteworthy observation is that density function (33) is the same as for the distribution of the ratio $Y_1^\alpha(\Sigma_1^\beta/t)^{-\beta/\alpha}$.

Since the limiting distribution of $Y_{\sigma_t}^\alpha$ has a smooth density, the convergence in the smooth Wasserstein metric (32) can be further enhanced to a stronger convergence in Kolmogorov's or Prokhorov's metric. Notice also that, for fixed t, we find ourselves in another extremely popular field of research concerning the rates of convergence of the random sums of independent and identically distributed random variables, see [28] and numerous references therein.

Let us stress finally again that, as it follows from the discussion above, CTRW approximations can be used for numeric solutions of fractional equations. This numeric approximation was seemingly first proposed in [22]. Our results supply the rates of convergence to these numeric schemes.

3. Auxiliary Results

In this section, we collect some mostly known results, stressing some particular points and consequences, which are crucial for our purposes.

3.1. Estimates of One-Sided Stable Laws Near the Origin

Stable subordinators Σ_t^β are increasing Lévy processes generated by the inverted fractional derivative operator

$$-\frac{d^\beta}{d(-x)^\beta} f(x) = \frac{\beta}{\Gamma(1-\beta)} \int_0^\infty \frac{f(x+y) - f(x)}{y^{1+\beta}} \, dy. \tag{34}$$

The characteristic functions $\phi_t(q)$ of the transition probability densities $S_\beta(t,x)$ of these processes are known to equal

$$\phi_t(q) = \exp\{-t|q|^\beta e^{-i\pi\beta \operatorname{sgn} q/2}\} \, dq,$$

so that the densities themselves have the integral representation

$$S_\beta(t,x) = \frac{1}{2\pi} \int_{\mathbf{R}} \exp\{-iqx - t|q|^\beta e^{-i\pi\beta \operatorname{sgn} q/2}\} \, dq. \tag{35}$$

Notice that the function $G_\beta(t,x) = S_\beta(t,-x)$ is the Green function for the Cauchy problem to the operator (34) so that it solves the Cauchy problem

$$\frac{\partial G_\beta}{\partial t}(t,x) = -\frac{d^\beta}{d(-x)^\beta} G_\beta(t,x), \quad G(0,x) = \delta(x).$$

It is clear that $S_\beta(t,x)$ are infinitely smooth in x and satisfy the scaling relation

$$S_\beta(t,x) = t^{-1/\beta} S(1, t^{-1/\beta} x).$$

Therefore, a study of the density S_β can be effectively reduced to the study of $S_\beta(1,x)$.

From (34), it follows that stable subordinators move only in one direction, implying that $S_\beta(t,x) = 0$ for $x \le 0$.

From (35), it follows that

$$S_\beta(1,x) \le \frac{1}{\pi} \int_0^\infty \exp\{-tq^\beta \cos(\pi\beta/2)\} \, dq$$

for any x and therefore

$$\max_x S_\beta(1,x) \le \frac{1}{\pi\beta} \Gamma(1/\beta) [\cos(\pi\beta/2)]^{-1/\beta}. \tag{36}$$

In this paper, we are working with the re-scaled subordinator $\hat{\Sigma}_t^\beta$, given by (3) and having the transition density $\hat{S}(t,x) = S(t\Gamma(1-\beta)/\beta, x)$. It follows that

$$\max_x \hat{S}(t,x) = \max_x S(t\Gamma(1-\beta)/\beta, x) = (t\Gamma(1-\beta)/\beta)^{-1/\beta} \max_x S(1,x)$$

$$\le t^{-1/\beta} (\Gamma(1-\beta)/\beta)^{-1/\beta} \frac{\Gamma(1/\beta)}{\pi\beta} [1-\beta]^{-1/\beta},$$

where in the last inequality, we took into account the simple estimate $\cos(\pi\beta/2) > 1 - \beta$, valid of all $\beta \in (0,1)$. Thus, using also that $\Gamma(2-\beta) = \Gamma(1-\beta)(1-\beta)$, we conclude that

$$\max_x \hat{S}(t,x) \le t^{-1/\beta} M, \quad M = \sup_{\beta \in (0,1)} [(\Gamma(2-\beta)/\beta)^{-1/\beta} \frac{\Gamma(1/\beta)}{\pi\beta}].$$

Performing elementary manipulations (which we omit) with the Stirling formula (with explicit estimates, see, for example, [29]) allows one to prove a uniform bound for the multiplier M:

Lemma 1.
$$M = \sup_{\beta\in(0,1)} [\Gamma(2-\beta)/\beta]^{-1/\beta} \frac{\Gamma(1/\beta)}{\pi\beta} \le 1/2. \tag{37}$$

Consequently,
$$\mathbf{P}(\hat{\Sigma}_t^\beta \le K) = \int_0^K \hat{S}_\beta(t,x)\,dx = \frac{1}{2}Kt^{-1/\beta}. \tag{38}$$

We shall need also the estimates for $S_\beta(1,x)$ for large x. These estimates can be derived from the standard expansion (see, for example, [30,31])

$$S_\beta(1,x) = -\frac{1}{\pi x}\sum_{k=1}^\infty \frac{\Gamma(1+k\beta)}{\Gamma(1+k)} \sin(k\pi\beta)(-x^{-\beta})^k,$$

which is convergent for all $x > 0$. Making a very rough estimate of the terms, we obtain

$$S_\beta(1,x) \le \frac{\pi\beta}{\pi x}\sum_{k=1}^\infty kx^{-\beta k} = \frac{\beta}{x}\frac{x^{-\beta}}{1-x^{-\beta}}.$$

Thus, for $x \ge 2$,
$$S_\beta(1,x) \le x^{-1-\beta}\frac{\beta}{1-2^{-\beta}}.$$

Since $1 - e^{-\beta\ln 2} \ge \beta\ln 2(1-\beta\ln 2/2) \ge \beta\ln 2(1-\ln 2/2)$, it follows that
$$S_\beta(1,x) \le 3x^{-1-\beta}, \quad x \ge 2. \tag{39}$$

3.2. Convergence of Markov Chains to Continuous-Time Processes

It is well known that the convergence of the generators of contraction semigroups on the core of the limiting generator implies the convergence of semigroups. We shall need a version of this fact with explicit rates, namely the following result, given in Theorem 8.1.1 of [20].

Proposition 1. *Let $F_t = e^{tL}$ be a Feller semigroup in the Banach space $B = C_\infty(\mathbf{R}^d)$, generated by an operator L, having a core D, which is itself a Banach space with a norm $\|.\|_D \ge \|.\|_B$ such that $\|Lf\|_B \le l\|f\|_D$ for a constant l and all $f \in D$. Let F_t be also a bounded semigroup in D such that*

$$\max_{s\in[0,t]} \|F_t\|_{D\to D} \le e^{mt},$$

with a constant $m \ge 0$ (the growth rate of the semigroup).

Let U_h be a family of contractions in B, and let

$$\|\left(\frac{U_h - 1}{h} - L\right)f\|_B \le \epsilon_h\|f\|_D, \tag{40}$$

$$\|\left(\frac{F_h - 1}{h} - L\right)f\|_B \le \varkappa_h\|f\|_D, \tag{41}$$

with $\epsilon_h \to 0$ and $\varkappa_h \to 0$, as $h \to 0$. Then the scaled discrete semigroups $(U_h)^{[t/h]}$ converge to the semigroup F_t and moreover

$$\sup_{s\le t} \|(U_h)^{[s/h]} - F_s f\|_B \le (\varkappa_h + \epsilon_h)\|f\|_D \int_0^t e^{ms}\,ds. \tag{42}$$

As a key illustration to when (40) holds, we shall look at the convergence of random walks to stable subordinators, as described by the following result.

Lemma 2. *Let $p(y)$ be a probability density on $\mathbf{R}_+$ such that $p(y) = y^{-1-\beta}$ for $y \geq B$ with some $\beta \in (0,1)$ and $B > 0$ such that $\beta B^\beta > 1$ (the latter condition comes from the requirement that $\int_B^\infty p(y)dy \leq 1$). Then, we have the following:*

(i) For any bounded measurable f having support on $[Bh^{1/\beta}, \infty)$,

$$h^{-1} \int_0^\infty f(h^{1/\beta} y) p(y) dy = \int_0^\infty \frac{f(y)dy}{y^{1+\beta}}, \tag{43}$$

(ii) For any bounded measurable f on $\mathbf{R}_+$ such that $|f(y)| \leq Ly$ for $y \in [0, Bh^{1/\alpha}]$ and some constant L, it follows that

$$\left| h^{-1} \int_0^\infty f(h^{1/\beta} y) p(y) dy - \int_0^\infty \frac{f(y)dy}{y^{1+\beta}} \right| \leq C_{B,\beta} L h^{-1+1/\beta}, \tag{44}$$

with

$$C_{B,\beta} = \frac{B^{1-\beta}}{1-\beta} + \int_0^B y p(y) dy \leq \frac{B^{1-\beta}}{1-\beta} + B \leq \frac{2B}{1-\beta}.$$

Proof. (i) From the condition on the support of f, one can change $p(y)$ to $y^{1-\beta}$. Changing the variable of integration to $y' = h^{1/\beta} y$ completes the proof.

(ii) Let f have support on $[0, Bh^{1/\beta}]$. Then

$$\left| h^{-1} \int_0^\infty f(h^{1/\beta} y) p(dy) \right| = \left| h^{-1} \int_0^B f(h^{1/\beta} y) p(y) dy \right| \leq h^{-1+1/\beta} L \int_0^B y p(y) dy,$$

and

$$\left| \int_0^\infty \frac{f(y)dy}{y^{1+\beta}} \right| = \left| \int_0^{Bh^{1/\beta}} \frac{f(y)dy}{y^{1+\beta}} \right| \leq L \int_0^{Bh^{1/\beta}} \frac{dy}{y^\beta} = \frac{B^{1-\beta}}{1-\beta} L h^{-1+1/\beta},$$

implying (44). From $\beta B^\beta > 1$, it follows that $B > 1$, which implies the last estimate for $C_{B,\beta}$. □

In particular, it follows that, for $f \in C^1$

$$\left| h^{-1} \int_0^\infty (f(x \pm h^{1/\beta} y) - f(x)) p(y) dy - \int_0^\infty \frac{(f(x \pm y) - f(x))dy}{y^{1+\beta}} \right| \leq \frac{2B}{1-\beta} \|f'\| h^{(1/\beta)-1}. \tag{45}$$

3.3. Estimates for Characteristic Functions for Distributions with a Density

It is well known (and easy to see) that if the support of a probability law μ on $\mathbf{R}$ is not contained in a lattice, then the characteristic function $\phi_\mu(q)$ is everywhere less than one in magnitude, apart from the value $\phi_\mu(0) = 1$. One can give various estimates for the magnitude of $\phi_\mu(q)$ away from zero. Here, we derive a simple estimate that is handy for heavy tailed distributions.

Proposition 2. *Let a distribution μ on $\mathbf{R}_+$ have a bounded density $p(x)$ and there exist $B > 0$ such that $p(x)$ is monotonically decreasing and strictly positive for $x \geq B$. Then for any $q_0 > 0$, there exists a $\varkappa_{q_0} \in (0,1)$ such that*

$$|\phi_\mu(q)| \leq \varkappa_{q_0} \quad \text{whenever } |q| > q_0.$$

Concretely, one can take

$$\varkappa_{q_0} = 1 - \frac{1}{2} \int_{B+\pi/q_0}^\infty p(x)\, dx. \tag{46}$$

Proof. Let $q \geq q_0 > 0$. We have

$$\left|\int_B^\infty e^{iqx} p(x)\,dx\right|^2 = \int_B^\infty \int_B^\infty \cos[q(x-y)]p(x)p(y)dxdy.$$

Due to the monotonicity of $p(x)$ and the oscillations of $\cos[q(x-y)]$ with period $2\pi/q$,

$$\left|\int_B^\infty \cos[q(x-y)]p(x)dx\right| \leq \int_B^{B+\pi/q} p(x)\,dx \leq \int_B^{B+\pi/q_0} p(x)\,dx.$$

Hence,

$$|\int_B^\infty e^{iqx} p(x)\,dx|^2 \leq \int_B^{B+\pi/q_0} p(x)\,dx \int_B^\infty p(y)\,dy.$$

Therefore,

$$\begin{aligned} |\phi_\mu(q)| &\leq \left|\int_0^B e^{iqx} p(x)\,dx\right| + \left|\int_B^\infty e^{iqx} p(x)\,dx\right| \\ &\leq \int_0^B p(x)\,dx + \sqrt{\int_B^{B+\pi/q_0} p(x)\,dx}\sqrt{\int_B^\infty p(y)\,dy} \\ &= 1 - \int_B^\infty p(x)\,dx + \sqrt{\int_B^{B+\pi/q_0} p(x)\,dx}\sqrt{\int_B^\infty p(x)\,dx}, \end{aligned}$$

which is easily seen to be bound by $C\varkappa_{q_0}$ from (46). □

3.4. From Weak to Strong Convergence of Distributions

Convergence of a sequence in the smooth Wasserstein metrics (9) implies the weak convergence of this sequence. Here, we present a quantitative result, which is a generalization of Corollary 1.6 from [8], showing that if a limiting distribution has a density, then the convergence with respect to any of the smooth Wasserstein metrics implies the convergence in Kolmogorov's metric (15). Moreover, one is able to relate the rates of convergence in Wasserstein's and Kolmogorov's distances.

Proposition 3. *Let X, Y be two random variables such that the distribution of Y has a bounded probability density $p(x)$ with $M = \sup_x p(x)$. Then, for any $k \in N$,*

$$d_{Kol}(X,Y) \leq (M + \varkappa_k)[d_{C^k}(X,Y)]^{1/(k+1)}, \tag{47}$$

with a constant $\varkappa_k \geq 1$ depending only on k. For instance, $\varkappa_1 \leq 3/2$ and $\varkappa_2 \leq 6$. Moreover, for any $\alpha \in (0,1]$,

$$d_{Kol}(X,Y) \leq (M+1)[d_{H_\alpha}(X,Y)]^{1/(\alpha+1)}, \tag{48}$$

$$d_{Kol}(X,Y) \leq (M+3/2)[d_{H_{1,\alpha}}(X,Y)]^{1/(\alpha+2)}, \tag{49}$$

Proof. Since $\varkappa_k \geq 1$ and $d_{Kol}(X,Y) \leq 1$, inequality (47) holds trivially whenever $d_{C^k}(X,Y) \geq 1$. Let us now assume that $d_{C^k}(X,Y) < 1$.

By the assumptions on Y,

$$\mathbf{P}(Y \in [a,b]) \leq (b-a)M$$

for any interval $[a,b]$.

Let ϕ be a smooth function $\mathbf{R} \to [0,1]$ such that $\phi(x) = 0$ for $x \geq 1$ and $\phi = 1$ for $x \leq 0$, and let $\rho \geq 1$ be a constant. Let $\phi_{\rho,z}(s) = \phi(\rho(s-z))$, so that $\phi_{\rho,z}(s) = 0$ for $s > z + 1/\rho$. Then

$$\mathbf{P}(X \leq z) \leq \mathbf{E}\phi_{\rho,z}(X) \leq \mathbf{E}\phi_{\rho,z}(Y) + d_{C^k}(X,Y)\rho^k\|\phi\|_{C^k}$$

$$\leq \mathbf{P}(Y \leq z + 1/\rho) + d_{C^k}(X,Y)\rho^k\|\phi\|_{C^k} \leq \mathbf{P}(Y \leq z) + M/\rho + d_{C^k}(X,Y)\rho^k\|\phi\|_{C^k}.$$

Let us choose $\rho = [d_{C^k}(X,Y)]^{-1/(k+1)}$. Then

$$\mathbf{P}(X \leq z) - \mathbf{P}(Y \leq z) \leq [d_{C^k}(X,Y)]^{1/(k+1)}(M + \|\phi\|_{C^k}).$$

Similarly, the opposite inequality is obtained, implying (47) with $\varkappa_k = \|\phi\|_{C^k}$. For $k = 1$, one can choose $\phi(x) = 1 - 3x^2 + 2x^3$ for $x \in [0,1]$ so that $\varkappa_1 = \max_{x\in[0,1]}(6x - 6x^2) = 3/2$. For $k = 2$, one can choose $\phi(x) = 1 - 10x^3 + 15x^4 - 6x^5$ yielding $\varkappa_2 < 6$.

Inequalities (48) and (49) are proved analogously, choosing the linear function as ϕ on $[0,1]$ for the first case (so ϕ becomes Lipschitz with the Lipschitz constant 1) and the function $1 - 3x^2 + 2x^3$ for the second case. □

We will need this result for the distributions of stable processes $Y = \Sigma_t^\beta$. By (36) and the scaling property of the stable densities, we derive that for the standard β-stable subordinator $Y = \Sigma_t^\beta$, the constant M in Proposition 3 can be estimated as

$$M = \max_x S_\beta(t,x) \leq \frac{\Gamma(1/\beta)}{\pi\beta}[t\cos(\pi\beta/2)]^{-1/\beta} \leq \frac{\Gamma(1/\beta)}{\pi\beta}[t(1-\beta)]^{-1/\beta}, \tag{50}$$

and for the distribution of $Y = \hat{\Sigma}_t^\beta$ generated by $\hat{L}_\beta$, this constant modifies to

$$M = \max_x \hat{S}_\beta(t,x) =\leq [t\Gamma(2-\beta)/\beta]^{-1/\beta}\frac{\Gamma(1/\beta)}{\pi\beta} \leq 1/2t^{-1/\beta}, \tag{51}$$

where Lemma 1 was used.

4. Regularity of Stable Semigroups in Hölder Spaces

We are planning to derive Theorem 1 by an application of Proposition 1 with $d = 1$. For this application, we need estimates (40) and (41). The former is supplied for our setting by Lemma 2. The latter estimate is the subject of the present section.

Proposition 4. *Let T_β^t be the Feller semigroup of the β-stable Lévy subordinator in $\mathbf{R}$ generated by the operator*

$$L_\beta f(x) = \int_0^\infty \frac{f(x+y) - f(x)}{y^{1+\beta}} dy.$$

Then, for any $\alpha \in (\beta, 1]$,

$$\|T_\beta^t f - f\| \leq \frac{2\alpha}{\beta(\alpha-\beta)} t\|f\|_\alpha, \tag{52}$$

and for any $\alpha \in (0, \beta]$,

$$\|T_\beta^t f - f\| \leq \frac{5}{\beta(1-\beta)} t^{\alpha/(1+\alpha)}\|f\|_\alpha, \tag{53}$$

Proof. If $\alpha \in (\beta, 1]$, then

$$\|L_\beta f\| \leq \|f\|_\alpha\left(\int_0^1 \frac{y^\alpha dy}{y^{1+\beta}} + 2\int_1^\infty \frac{dy}{y^{1+\beta}}\right) \leq 2\left(\frac{1}{\alpha-\beta} + \frac{1}{\beta}\right)\|f\|_\alpha = \frac{2\alpha}{\beta(\alpha-\beta)}\|f\|_\alpha.$$

In particular,

$$\|L_\beta f\| \leq \frac{2}{\beta(1-\beta)}\|f\|_{C^1}. \tag{54}$$

Hence

$$\|T_\beta^t f - f\| \leq \int_0^t \|T_s L_\beta f\, ds\| \leq t\|L_\beta f\| \leq \frac{2\alpha}{\beta(\alpha-\beta)} t\|f\|_\alpha,$$

yielding (52).

Next, let ϕ be an even nonnegative smooth function on $\mathbf{R}$ with support in $[-1,1]$ such that $\phi(0) = 1$ and $\int \phi(x)\,dx = 1$. For $\delta \in (0,1]$, let $\phi_\delta(x) = \delta^{-1}\phi(x/\delta)$. Let us approximate an arbitrary continuous function on $\mathbf{R}$ by its convolutions

$$(f \star \phi_\delta)(x) = \int f(y)\phi_\delta(x-y)\,dy = \int f(x-y)\phi_\delta(y)\,dy.$$

If $f \in H_\alpha$, then

$$\|f - f\star\phi_\delta\| \le \sup_x \int |f(x)-f(y)|\phi((x-y)/\delta)\,\frac{dy}{\delta} \le \|f\|_\alpha \sup_x \int |x-y|^\alpha \phi((x-y)/\delta)\,\frac{dy}{\delta}$$
$$= \|f\|_\alpha \delta^\alpha \int |z|^\alpha \phi(z)\,dz \le \|f\|_\alpha \delta^\alpha.$$

On the other hand,

$$\|(f\star\phi_\delta)'\| \le \frac{1}{\delta}\|f\| \int |\phi'(y)|\,dy \le \frac{2}{\delta}\|f\|.$$

Hence,

$$\|f\star\phi_\delta\|_{C^1} \le \|f\| \max\{1, \frac{2}{\delta}\} \le \frac{2}{\delta}\|f\|.$$

Therefore (using (54) in the last step),

$$\|(T^t_\beta - 1)f\| \le \|(T^t_\beta - 1)(f - f\star\phi_\delta)\| + \|(T^t_\beta - 1)(f\star\phi_\delta)\| \le 2\|f - f\star\phi_\delta\| + t\|L_\beta(f\star\phi_\delta)\|$$
$$\le \|f\|_\alpha \left(2\delta^\alpha + \frac{2}{\beta(1-\beta)}\frac{2t}{\delta}\right).$$

Choosing $\delta = (2t)^{1/(1+\alpha)}$ and estimating $\beta(1-\beta) \le 1/4$ yields (53). □

Lemma 3. *If $f \in H_{1,\alpha}$, then $L_\beta \in H_\alpha$ and*

$$\|L_\beta f\|_\alpha \le \frac{2}{\beta(1-\beta)}\|f\|_{1,\alpha}. \tag{55}$$

Proof. Writing

$$L_\beta f(x) = L^1_\beta f(x) + L^2_\beta f(x)$$

with

$$L^1_\beta f(x) = \int_0^1 \frac{f(x+z)-f(x)}{z^{1+\beta}}dz = \int_0^1 \frac{\int_0^z f'(x+w)dw}{z^{1+\beta}}\,dz,$$
$$L^2_\beta f(x) = \int_1^\infty \frac{f(x+z)-f(x)}{z^{1+\beta}}dz,$$

we have

$$\frac{|L^1_\beta f(x) - L^1_\beta f(y)|}{|x-y|^\alpha} \le \|f\|_{1,\alpha}\int_0^1 \frac{z}{z^{1+\beta}}\,dz = \frac{1}{1-\beta}\|f\|_{1,\alpha},$$

and

$$\frac{|L^2_\beta f(x) - L^2_\beta f(y)|}{|x-y|^\alpha} \le \int_1^\infty \frac{|f(x+z)-f(y+z)| + |f(x)-f(y)|}{|x-y|^\alpha z^{1+\beta}}\,dz$$
$$\le 2\|f'\|\,|x-y|^{1-\alpha}\int_1^\infty \frac{dz}{z^{1+\beta}} \le \frac{2}{\beta}\|f'\|.$$

Therefore (and using (54)),

$$\|L_\beta f\|_\alpha \leq \|f\|_{1,\alpha}\left(\frac{2}{1-\beta}+\frac{2}{\beta}\right)$$

yielding (55). □

Proposition 5. *Under assumptions of Proposition 4, if $\alpha \in (0,\beta]$, then*

$$\|\left(\frac{T_\beta^t-1}{t}-L_\beta\right)f\| \leq \frac{10}{\beta^2(1-\beta)^2}t^{\alpha/(1+\alpha)}\|f\|_{1,\alpha}, \tag{56}$$

and if $\alpha \in (\beta,1]$, then

$$\|\left(\frac{T_\beta^t-1}{t}-L_\beta\right)f\| \leq \frac{4\alpha}{\beta^2(1-\beta)(\alpha-\beta)}t\|f\|_{1,\alpha}. \tag{57}$$

In particular,

$$\|\left(\frac{T_\beta^t-1}{t}-L_\beta\right)f\| \leq \frac{4}{\beta^2(1-\beta)^2}t\|f\|_{C^2}. \tag{58}$$

Proof. We have

$$\left(\frac{T_\beta^t-1}{t}-L_\beta\right)f = \frac{1}{t}\int_0^t (T_\beta^s-1)Lf\,ds.$$

Hence, if $\alpha \in (0,\beta]$, then by (53) and (55), the left-hand side of (56) is bounded by

$$\frac{5}{\beta(1-\beta)}t^{\alpha/(1+\alpha)}\|L_\beta f\|_\alpha \leq \frac{5}{\beta(1-\beta)}t^{\alpha/(1+\alpha)}\frac{2}{\beta(1-\beta)}\|f\|_{1,\alpha},$$

yielding (56). Similarly, the case $\alpha \in (\beta,1]$ is dealt with. □

5. Proof of Theorems 1 and 2

As a direct consequence of Proposition 1, estimates (45) and (56), and the observation that, because the Feller semigroup of the stable has a spatially homogeneous integral kernel, this Feller semigroup is a contraction in all spaces C^k, the spaces H_α and $H_{1,\alpha}$ (so Proposition 1 holds with $m=0$), we can conclude that, for $\alpha \in (\beta,1]$,

$$\sup_{s\leq t}\|(U_h^{[s/h]}-T_\beta^s)f\| \leq t\left(\frac{2B}{1-\beta}h^{(1-\beta)/\beta}+\frac{4\alpha h}{\beta^2(1-\beta)(\alpha-\beta)}\right)\|f\|_{1,\alpha}. \tag{59}$$

Consequently, under (4), choosing any $\alpha \in (\beta,\min\{1,(1-\beta)/\beta\}]$ yields (5). If (4) does not hold, then, choosing $\alpha=(1-\beta)/\beta$, we obtain

$$\sup_{s\leq t}\|(U_h^{[s/h]}-T_\beta^s)f\| \leq t\left(\frac{2B}{1-\beta}h^{(1-\beta)/\beta}+\frac{10}{\beta^2(1-\beta)^2}h^{1-\beta}\right)\|f\|_{1,(1-\beta)/\beta}. \tag{60}$$

Since $(1-\beta)/\beta > 1-\beta$, it implies (8), completing the proof of Theorem 1. Next, by (49),

$$\sup_{s\leq t} d_{Kol}(\Phi_s^h,\hat{\Sigma}_s^\beta) \leq (M+3/2)[d_{H_{1,\alpha}}(\Phi_s^h,\hat{\Sigma}_s^\beta)]^{1/(\alpha+2)},$$

with $M \leq 1/2$ by (51). Therefore, estimates (11) and (13) follow from Theorem 1.

In view of these results, in order to prove (12) and (14) it is sufficient to consider only the case of $t=1$. To complete the proof, one has to look first at the proof of Proposition 3 and to note that in order to estimate $|\mathbf{P}(X \le z) - \mathbf{P}(Y \le z)|$, it is sufficient to choose M to be the maximum of the density $p(x)$ only in a neighborhood of K (and not necessarily the global maximum). Then for any $s \in (0,1)$, we can use inequality (49) with

$$M = M(s) = \hat{S}_\beta(s,K) = (s\Gamma(1-\beta)/\beta)^{-1/\beta} S_\beta(1, K(s\Gamma(1-\beta)/\beta)^{-1/\beta}).$$

Let us decompose the domain of variables K into two parts: $K \ge K_0$ and $K \in [k, K_0)$, where

$$K_0 = 2[s\Gamma(1-\beta)/\beta]^{1/\beta}.$$

In the first part, $K(s\Gamma(1-\beta)/\beta)^{-1/\beta} \ge 2$, and we can use estimate (39) to conclude that $M \le 3K^{-1} \le 3k^{-1}$. In the second part, we use estimate (51) yielding

$$M \le \frac{1}{2} s^{-1/\beta} \le (\Gamma(1-\beta)/\beta)^{1/\beta} k^{-1}.$$

Since $(\Gamma(1-\beta)/\beta)^{1/\beta} \ge 3$, this completes the proof of 2.

6. Tails of Scaled CTRW

Recall that τ_i, $i \in \mathbf{N}$, is a sequence of positive independent and identically distributed random variables having probability density $p(y)$ on $\mathbf{R}_+$ such that $p(y) = y^{-1-\beta}$ for $y \ge B$ with some $\beta \in (0,1)$ and $B > 0$ such that $\beta B^\beta > 1$, and

$$\Phi_t^h = \sum_{i=1}^{[t/h]} h^{1/\beta} \tau_i. \tag{61}$$

is the corresponding scaled random walk.

Proposition 6. *Assume additionally that $p(0) = 0$ and $p(x)$ is continuously differentiable. Then, for any $K > 0$ and $t \ge 2$,*

$$\mathbf{P}(\Phi_t^h < K) \le \frac{3}{2} K (t-1)^{-1/\beta} 2^{1/\beta} + \omega(t-1), \tag{62}$$

where $\omega(t-1)$ is exponentially small for large t uniformly for $h \le 1$ and all K. In particular, for $t > t_0$ with some t_0 (which is explicitly given by condition (71)),

$$\omega(t-1) \le \frac{1}{2} K (t-1)^{-1/\beta} 2^{1/\beta},$$

and

$$\mathbf{P}(\Phi_t^h < K) \le 2^{1+2/\beta} K t^{-1/\beta} \tag{63}$$

for all $h \le 1$ and $K \ge 1$.

Proof. The characteristic function of Φ_t^h equals

$$\phi_{\Phi_t^h}(q) = [\phi_T(qh^{1/\beta})]^{[t/h]} = [1 + \int (e^{iqh^{1/\beta}x} - 1) p(x)\, dx]^{[t/h]}.$$

We need an estimate for its magnitude.

Consider separately three domains of the values of q:

$$I: \ |q| \le C_1 h^{-1/\beta}, \quad II: \ C_1 h^{-1/\beta} \le |q| \le C_2 h^{-1/\beta}, \quad III: \ |q| > C_2 h^{-1/\beta},$$

with constants C_1, C_2 chosen later. Then we can write

$$\mathbf{P}(\Phi_t^h < K) = \frac{1}{2\pi} \int_0^K \left[\int e^{-iqx} \phi_{\Phi_t^h}(q) dq \right] dx$$

$$= \frac{1}{2\pi}\int \frac{1-e^{-iqK}}{iq}\phi_{\Phi_t^h}(q)\,dq = \int_I + \int_{II} + \int_{III}.$$

In domain I,

$$|\phi_{\Phi_t^h}(q)| \le |1+\int_0^\infty (e^{iqh^{1/\beta}x}-1)p(x)\,dx|^{(t-1)/h}.$$

By (5),

$$\frac{1}{h}\int_0^\infty (e^{iqh^{1/\beta}x}-1)p(x)\,dx = \int_0^\infty \frac{(e^{iqy}-1)dy}{y^{1+\beta}} + \omega,$$

where

$$|\omega| \le C_{B,\beta}|q|h^{-1+1/\beta}.$$

The value of the integral on the right-hand side is well known:

$$\int_0^\infty \frac{(e^{iqy}-1)dy}{y^{1+\beta}} = -\frac{\Gamma(1-\beta)}{\beta}e^{-i\pi\beta\operatorname{sgn} q/2}|q|^\beta$$

$$= -\frac{\Gamma(1-\beta)}{\beta}|q|^\beta(\cos(\pi\beta/2) \mp i\sin(\pi\beta/2)),$$

where $\mp$ corresponds to positive and negative q, respectively. If

$$h\frac{\Gamma(1-\beta)}{\beta}|q|^\beta \le \frac{1}{2}\cos(\pi\beta/2), \tag{64}$$

then

$$|\phi_{\Phi_t^h}(q)| \le |1+\int (e^{iqh^{1/\beta}x}-1)p(x)\,dx|$$
$$\le |1-h\frac{\Gamma(1-\beta)}{\beta}|q|^\beta(\cos(\pi\beta/2) \mp i\sin(\pi\beta/2))| + h|\omega|$$
$$= \sqrt{1-h\frac{\Gamma(1-\beta)}{\beta}|q|^\beta\left(2\cos(\pi\beta/2) - h\frac{\Gamma(1-\beta)}{\beta}|q|^\beta\right)} + h|\omega|$$
$$\le 1-\frac{h}{2}\frac{\Gamma(1-\beta)}{\beta}|q|^\beta\left(2\cos(\pi\beta/2) - h\frac{\Gamma(1-\beta)}{\beta}|q|^\beta\right) + h|\omega|$$
$$\le 1-\frac{h}{2}\frac{\Gamma(1-\beta)}{\beta}|q|^\beta(3/2)\cos(\pi\beta/2) + h|\omega|.$$

If, additionally,

$$C_{B,\beta}|q|h^{1/\beta} \le \frac{h}{4}\frac{\Gamma(1-\beta)}{\beta}|q|^\beta(\cos(\pi\beta/2), \tag{65}$$

then

$$|\phi_{\Phi_t^h}(q)| \le |1+\int (e^{iqh^{1/\beta}x}-1)p(x)\,dx| \le 1-\frac{h}{2}\frac{\Gamma(1-\beta)}{\beta}|q|^\beta(\cos(\pi\beta/2),$$

and

$$|\phi_{\Phi_t^h}(q)| \le \left(1-\frac{h}{2}\frac{\Gamma(1-\beta)}{\beta}|q|^\beta\cos(\pi\beta/2)\right)^{(t-1)/h} \le \exp\{-\frac{1}{2}(t-1)|q|^\beta\frac{\Gamma(1-\beta)}{\beta}\cos(\pi\beta/2)\}.$$

Conditions (64) and (65) are satisfied if

$$|q| \le C_1 h^{-1/\beta} \tag{66}$$

with

$$C_1 = \min\left(\left[\frac{\beta}{2\Gamma(1-\beta)} \cos(\pi\beta/2) \right]^{1/\beta}, \left[\frac{\Gamma(1-\beta)}{4C_{B,\beta}\beta} \cos(\pi\beta/2) \right]^{1/(1-\beta)} \right), \tag{67}$$

which we assume from now on.

Then

$$|\int_I| \le \frac{K}{\pi} \int_0^\infty \exp\{-\frac{1}{2}(t-1)q^\beta \frac{\Gamma(1-\beta)}{\beta} \cos(\pi\beta/2)\}\, dq.$$

Changing the variable of integration yields

$$|\int_I| \le \frac{K}{\pi}(t-1)^{-1/\beta} \int_0^\infty \exp\{-\frac{1}{2}q^\beta \frac{\Gamma(1-\beta)}{\beta} \cos(\pi\beta/2)\}\, dq,$$

and therefore

$$|\int_I| \le \frac{K\Gamma(1/\beta)}{\pi\beta}(t-1)^{-1/\beta} \left(\frac{1}{2} \frac{\Gamma(1-\beta)}{\beta} \cos(\pi\beta/2) \right)^{-1/\beta}.$$

Using the inequality $\cos(\pi\beta/2) > 1-\beta$ and Lemma 1 yields the estimate

$$|\int_I| \le \frac{3}{2} K(t-1)^{-1/\beta} 2^{1/\beta}. \tag{68}$$

We assume that $p(x)$ is continuously differentiable and $p(0) = 0$. Then in domain III,

$$|\phi_T(q)| \le \frac{1}{|q|} \int_0^\infty |p'(y)|\, dy$$

and

$$|\int_{III}| \le \int_{q \ge C_2 h^{-1/\beta}} \frac{2}{\pi q} \left(\frac{1}{q h^{1/\beta}} \int_0^\infty |p'(y)|\, dy \right)^{(t-1)/h} dq$$

$$= \frac{2h}{\pi(t-1)} (h^{1/\beta})^{-(t-1)/h} (h^{1/\beta})^{(t-1)/h} \left(\int_0^\infty |p'(y)|\, dy / C_2 \right)^{(t-1)/h}.$$

Thus,

$$|\int_{III}| \le \frac{2h}{\pi(t-1)} \left(\int_0^\infty |p'(y)|\, dy / C_2 \right)^{(t-1)/h}.$$

Choosing

$$C_2 = \max\left(C_1, 2\int_0^\infty |p'(y)|\, dy \right),$$

we obtain

$$|\int_{III}| \le \frac{2h}{\pi(t-1)} 2^{-(t-1)/h}. \tag{69}$$

Thus, for $t \ge 3$, any $h \le 1$ and $K \ge 1$,

$$|\int_{III}| \le \frac{1}{8} K(t-1)^{-1/\beta} 2^{1/\beta},$$

which is bounded by $1/12$ of the estimate of the main term $|\int_I|$.

Finally, to work with II, we use Proposition 2 to conclude that $|\phi_T(q)| \le \varkappa_{C_1} < 1$ for $|q| \ge C_1$. In our case,

$$\varkappa_{C_1} = 1 - \frac{1}{2} \int_{B+\pi/C_1}^\infty p(x)\, dx = 1 - \frac{1}{2} \int_{B+\pi/C_1}^\infty x^{-1-\beta}\, dx = 1 - \frac{1}{2\beta}(B + \pi/C_1)^{-\beta}.$$

Then in II,

$$|\phi_{\Phi_t^h}(q)| = |\phi_T(qh^{1/\beta})|^{[t/h]} \leq \varkappa_{C_1}^{(t-1)/h}.$$

Therefore,

$$|\int_{II}| \leq \frac{2}{\pi}\varkappa_{C_1}^{(t-1)/h}\int_{C_1h^{-1/\beta}}^{C_2h^{-1/\beta}}\frac{dq}{q} = \frac{2}{\pi}\varkappa_{C_1}^{(t-1)/h}\ln\frac{C_2}{C_1}. \tag{70}$$

Thus, the second and the third integrals are exponentially small for large t uniformly for $h \leq 1$. Moreover, if $t_0 \geq 3$ is such that

$$\frac{2}{\pi}\varkappa_{C_1}^{(t_0-1)}\ln\frac{C_2}{C_1} \leq \frac{1}{8}(t_0-1)^{-1/\beta}2^{1/\beta}, \tag{71}$$

then $|\int_{II}| + |\int_{III}|$ is bounded by $1/6$ of the estimate of the main term $|\int_I|$ yielding the required estimate for the exponentially small remainder $\omega(t-1)$. Finally, estimating $t/(t-1) \leq 2$, yields (63). $\square$

7. Proof of Theorem 3

Let us prove (19), as other estimates are proved quite analogously. If $t \geq 1$ and $\beta \geq (1-\beta)/\beta$, then by (13),

$$\sup_{K>0}\sup_{1\leq s\leq t}|\mathbf{P}(\Phi_s^h \leq K) - \mathbf{P}(\hat{\Sigma}_s^\beta < K)| \leq C(\beta)(h^{1-\beta}t)^{\beta/(\beta+1)}, \tag{72}$$

with a constant $C(\beta)$ given explicitly in (13). We shall use this estimate for $t \leq \Omega$ with some large $\Omega > 0$ to be chosen later. For $t > \Omega$, we shall estimate the probabilities on the left-hand side of (72) via tail estimates.

Namely, by (63), for $t \geq t_0$,

$$\mathbf{P}(\Phi_t^h < K) \leq 2^{1+2/\beta}Kt^{-1/\beta}$$

and by (38)

$$\mathbf{P}(\hat{\Sigma}_t^\beta \leq K) \leq 1.5Kt^{-1/\beta}.$$

Thus,

$$\sup_{k\leq K}\sup_{t\geq\Omega}|\mathbf{P}(\Phi_t^h \leq K) - \mathbf{P}(\hat{\Sigma}_t^\beta < K)| \leq 2^{1+2/\beta}K\Omega^{-1/\beta}.$$

Consequently,

$$\sup_{k\leq K}\sup_{s\geq 1}|\mathbf{P}(\Phi_s^h \leq K) - \mathbf{P}(\hat{\Sigma}_s^\beta < K)| \leq C(\beta)(h^{1-\beta}\Omega))^{\beta/(\beta+1)} + 2\times 2^{2/\beta}K\Omega^{-1/\beta}.$$

Choosing now

$$\Omega = h^{-\epsilon}, \quad \epsilon = \frac{\beta(1-\beta)}{1+\beta+\beta^2},$$

we obtain

$$\sup_{k\leq K}\sup_{s\geq 1}|\mathbf{P}(\Phi_s^h \leq K) - \mathbf{P}(\hat{\Sigma}_s^\beta < K)| \leq \tilde{C}(\beta,K)h^{\epsilon/\beta},$$

with

$$\tilde{C}(\beta,K) = C(\beta) + 2\times 2^{2/\beta}K,$$

whenever $h^\epsilon \leq 1/t_0$. This implies (19).

To obtain similar estimates for (22), one uses estimates (12) and (14).

Finally, to prove (23), let us again consider the case $\beta \geq (1-\beta)/\beta$ for definiteness. Let us decompose the integral

$$I = \int_0^\infty |\mathbf{P}(\Phi_s^h \leq K) - \mathbf{P}(\hat{\Sigma}_s^\beta < K)|\, ds$$

into the two parts $I + II$, over the domains $\{s \leq \Omega = h^{-\epsilon}\}$ and $\{s > \Omega\}$ with some $\epsilon > 0$. Then the integrand is bound (up to a constant) by

$$(1 + K^{-1})(h^{1-\beta}s)^{\beta/(\beta+1)}$$

in the first domain and by $Ks^{-1/\beta}$ in the second domain. Integrating, we obtain the estimate

$$I \leq (1 + K^{-1})h^{(1-\beta)\beta/(\beta+1)}h^{-\epsilon[1+\beta/(\beta+1)]} + Kh^{\epsilon(1-\beta)/\beta}.$$

Choosing

$$\epsilon = \frac{\beta^2(1-\beta)}{1+\beta+\beta^2},$$

we obtain the estimate

$$I \leq C(\beta)h^{\beta(1-\beta)^2/(1+\beta+\beta^2)}$$

implying (23).

8. Proof of Theorem 4

For definiteness, let us consider the case with $\beta \leq 1/2$, other cases being quite analogous. We have

$$\|\mathbf{E}U_h^{N_t^h/h}f - \mathbf{E}F_{\sigma_t}f\| \leq I + II,$$

with

$$I = \|\mathbf{E}U_h^{N_t^h/h}f - \mathbf{E}F_{N_t^h}f\|, \quad II = \|\mathbf{E}F_{N_t^h}f - \mathbf{E}F_{\sigma_t}f\|.$$

To estimate I, we write

$$I = \int_0^\infty (U_h^{[s/h]}f - F_s f)\mu_t^h(ds) = \int_0^z (U_h^{[s/h]}f - F_s f)\mu_t^h(ds) + \int_z^\infty (U_h^{[s/h]}f - F_s f)\mu_t^h(ds),$$

where μ_t^h is the distribution of N_t^h.

By Proposition 1, the first term here is bounded in magnitude by

$$h^\omega \|f\|_D e^{mz}/z,$$

and, by Proposition 6, the second term is bounded by

$$2\|f\|\mathbf{P}(N_t^h > z) = 2\|f\|\mathbf{P}(\Phi_z^h < t) \leq 2^{2+2/\beta}tz^{-1/\beta}\|f\|.$$

Thus,

$$|I| \leq [h^\omega e^{mz}/z + 2^{2+2/\beta}tz^{-1/\beta}]\|f\|_D.$$

Choosing $z = \min(1, \omega/m)\ln(1/h)$ yields

$$\begin{aligned}|I| &\leq \|f\|_D[h^{\omega-\omega}\max(1, m/\omega)\ln(1/h)^{-1} + 2^{2+2/\beta}t\ln(1/h)^{-1/\beta}(\max(1, m/\omega))^{1/\beta}] \\ &\leq 4(t+1)\ln(1/h)^{-1}\|f\|_D(4\max(1, m/\omega))^{1/\beta}.\end{aligned}$$

Integrating by parts in II, we obtain the following:

$$II = \|\mathbf{E}e^{N_t^h L}f - \mathbf{E}e^{\sigma_t L}f\|$$

$$
= \| \int_0^\infty \frac{\partial}{\partial s}(e^{sL}f)(\mathbf{P}(\sigma_t \le s) - \mathbf{P}(N_t^h \le s))\, ds\|
$$

$$
= \| \int_0^\infty Le^{sL}f(\mathbf{P}(\hat{\Sigma}_s^\beta > t) - \mathbf{P}(\Phi_s^h > t))\, ds\|
$$

$$
\le \|Lf\| \int_0^\infty |\mathbf{P}(\hat{\Sigma}_s^\beta \le t) - \mathbf{P}(\Phi_s^h \le t)|\, ds.
$$

Consequently, by (23),

$$
II \le l\|f\|_D(1 + t + t^{-1})C(\beta)h^{\varkappa(\beta)}.
$$

Since the logarithmic decay of integral I is more rough than the power decay of II, estimate (26) follows.

Finally, if $m = 0$, then

$$
|I| \le [h^\omega z + 2^{2+2/\beta}tz^{-1/\beta}]\|f\|_D,
$$

which is of order $h^{\omega/(\beta+1)}$ (obtained by choosing $z = h^{-\omega\beta/(\beta+1)}$). Choosing further the worst power estimate from I and II, completes the proof.

Funding: The paper was published with the financial support of the Ministry of Education and Science of the Russian Federation as part of the program of the Moscow Center for Fundamental and Applied Mathematics under the agreement 075-15-2022-284.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Borovkov, A.A. On the rate of convergence for the invariance principle. *Theory Probab. Appl.* **1974**, *18*, 207–225. [CrossRef]
2. Dudley, R.M. Speeds of metric probability convergence. *Z. Wahrschein1ichkeitstheorie Verw. Geb.* **1972**, *22*, 323–332. [CrossRef]
3. Sakhanenko, A.I. On Borovkov's estimate in the Invariance Principle. *Sib. Electron. Mat. Izv.* **2019**, *16*, 1776–1784. [CrossRef]
4. Antoniou, M.; Melbourne, I. Rate of Convergence in the Weak Invariance Principle for Deterministic Systems. *Commun. Math. Phys.* **2019**, *369*, 1147–1165. [CrossRef]
5. Hafouta, Y. Convergence Rates in the Functional CLT for α-Mixing Triangular Arrays. Available online: https://arxiv.org/abs/2107.02234v3 (accessed on 1 January 2020).
6. Avram, F.; Taqqu, M.S.. Weak convergence of sums of moving averages in the α-stable domain of attraction. *Ann. Probab.* **1992**, *20*, 483–503. [CrossRef]
7. Kuske, R.; Keller, J.B. Rate of convergence to a stable limit. *SIAM J. Appl. Math.* **2001**, *61*, 1308–1323.
8. Chen, P.; Xu, L. Approximation to stable law by the Lindeberg principle. *J. Math. Anal. Appl.* **2019**, *48*, 123338. Available online: https://arxiv.org/abs/1809.10864 (accessed on 1 January 2020). [CrossRef]
9. Johnson, O.; Samworth, R. Central limit theorem and convergence to stable laws in Mallows distance. *Bernoulli* **2005**, *11*, 829–845. [CrossRef]
10. Tyran-Kaminska, M. Functional limit theorems for linear processes in the domain of attraction of stable laws. *Statist. Probab. Lett.* **2010**, *80*, 975–981. [CrossRef]
11. Basrak, B.; Krizmanić, D.; Segers, J. A functional limit theorem for dependent sequences with infinite variance stable limits. *Ann. Probab.* **2012**, *40*, 2008–2033. [CrossRef]
12. Bening, V.E.; Korolev, V.Y.; Kolokoltsov, V.N. Limit theorems for continuous-time rasndom walks in the double-array scheme. *J. Math. Sci.* **2006**, *138*, 5348–5365. [CrossRef]
13. Montroll, E.W.; Weiss, G.H. Random Walks on Lattices, II. *J. Math. Phys.* **1965**, *6*, 167–181. [CrossRef]
14. Kolokoltsov, V.; Korolev, V.; Uchaikin, V. Fractional Stable Distributions. *J. Math. Sci.* **2001**, *105*, 2570–2577. [CrossRef]
15. Meerschaert, M.M.; Scheffler, H.-P. Limit Theorems for Continuous-Time Random Walks with Infinite Mean Waiting Times. *J. Appl. Prob.* **2004**, *41*, 623–638. [CrossRef]
16. Meerschaert, M.M.; Sikorskii, A. *Stochastic Models for Fractional Calculus*; De Gruyter Studies in Mathematics; De Gruyter: New York, NY, USA, 2012; Volume 43.
17. Zaslavsky, G.M. Fractional kinetic equation for Hamiltonian chaos. *Physica D* **1994**, *76*, 110–122. [CrossRef]

18. Kolokoltsov, V.N. Generalized Continuous-Time Random Walks (CTRW), Subordination by Hitting Times and Fractional Dynamics. *Theory Probab. Appl.* **2009**, *53*, 594–609. [CrossRef]
19. Bening, V.E.; Korolev, V.Y.; Koksharov, S.; Kolokoltsov, V.N. Convergence-rate estimates for superpositions of independent stochastic processes with applications to estimation of the accuracy of approximation of the distributions of continuus-time random Markov walks by fractional stable laws. *J. Math. Sci.* **2007**, *146*, 5950–5957. [CrossRef]
20. Kolokoltsov, V.N. *Markov Processes, Semigroups and Generators*; DeGruyter Studies in Mathematics v. 38; DeGruyter: Berlin, Germany; New York, NY, USA, 2011.
21. Kolokoltsov, V.N. *Differential Equations on Measures and Functional Spaces*; Birkhäuser Advanced Texts Basler Lehrbücher: Cham, Switzerland, 2019.
22. Uchaikin, V.V.; Saenko, V.V. Stochastic solution to partial dfferential equations of fractional orders. *Sib. Zh. Vychisl. Mat.* **2003**, *6*, 197–203.
23. Kolokoltsov, V.; Lin, F.; Mijatovic, A. Monte Carlo estimation of the solution of fractional partial differential equations. *Fract. Calc. Appl. Anal.* **2021**, *24*, 278–306. [CrossRef]
24. Lv, L.; Wang, L. Stochastic representation and monte carlo simulation for multiterm time-fractional diffusion equation. *Adv. Math. Phys.* **2020**, *2020*, 1315426. [CrossRef]
25. Vabishchevich, P.N. Numerical solution of nonstationary problems for a space-fractional diffusion equation. *Fract. Calc. Appl. Anal.* **2016**, *19*, 116–139.
26. Uchaikin, V.V.; Zolotarev, V.M. *Chance and Stability: Stable Distributions and Their Applications*; Walter de Gruyter: Berlin, Germany; NewYork, NY, USA, 2011.
27. D'Ovidio, M. On the Non-Local Boundary Value Problem from the Probabilistic Viewpoint. *Mathematics* **2022**, *10*, 4122. [CrossRef]
28. Shevtsova, I.G. Estimates for the rate of convergence in the global CLT for generalized mixed Poisson distributions. (Russian). *Teor. Veroyatn. Primen.* **2018**, *63*, 89–116; translation in *Theory Probab. Appl.* **2018**, *63*, 72–93.
29. Batir, N. Inequalities for the gamma function. *Arch. Math.* **2008**, *91*, 554–563. [CrossRef]
30. Ibragimov, I.A.; Linnik, Y.V. *Indendent and Stationary Sequencs of Random Variables*; Wolters-Noordhoff.Groningen: Groningen, The Netherlands, 1971.
31. Zolotarev, V.M. *One-Dimensional Stable Distributions*; Nauka: Moscow, Russia, 1983; Engl. transl. in vol. 65 of Translations of Mathematical Monographs AMS, Providence, Rhode Island, 1986. (In Russian)