

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/102369/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Leonenko, Nikolai , Papic, Ivan, Sikorski, Alla and Suvak, Nenad 2018. Correlated continuous time random walks and fractional Pearson diffusions. *Bernoulli* 24 (4B) , pp. 3603-3627. 10.3150/17-BEJ972

Publishers page: <http://dx.doi.org/10.3150/17-BEJ972>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Correlated continuous time random walks and fractional Pearson diffusions

June 28, 2017

N.N. Leonenko^a, I. Papić^b, A. Sikorskii^{*c}, N. Šuvak^b

^a School of Mathematics, Cardiff University, Senghennydd Road, Cardiff CF244AG, UK

^b Department of Mathematics, J.J. Strossmayer University of Osijek, Trg Ljudevita Gaja 6, HR-31 000 Osijek, Croatia

^c A423 Wells Hall, Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824, USA

Abstract. Continuous time random walks have random waiting times between particle jumps. We define the correlated continuous time random walks (CTRWs) that converge to fractional Pearson diffusions (fPDs). The jumps in these CTRWs are obtained from Markov chains through the Bernoulli urn-scheme model and Wright-Fisher model. The jumps are correlated so that the limiting processes are not Lévy but diffusion processes with non-independent increments. The waiting times are selected from the domain of attraction of a stable law.

Key words. Markov chains; fractional diffusion; Pearson diffusions; urn-scheme models; Wright-Fisher model; continuous time random walks

Mathematics Subject Classification (2010): 47D07, 60J10, 60J60, 60G22, 60G50.

1 Introduction

Continuous time random walks (CTRWs) have random waiting times between particle jumps. When jumps and waiting times are independent, the CTRW is called decoupled, and only decoupled CTRWs will be considered throughout this paper. For more information and applications of coupled and uncoupled CTRW we refer to [16, 6].

When particle jumps Y_1, \dots, Y_n, \dots are independent and identically distributed (iid), the random walk $S_n = Y_1 + \dots + Y_n$ converges to either the Brownian motion or a stable Lévy process [19, Chapter 4]. When jumps are separated by the iid waiting times G_1, \dots, G_n from the domain of attraction of a positively skewed stable law with stability index $0 < \beta < 1$, the process $S(N_t)$, where $T_n = G_1 + \dots + G_n$, $N_t = \max\{n \geq 0: T_n \leq t\}$, gives the location of a particle at time $t \geq 0$ and is known as the CTRW. By applying the continuous mapping theorem (see [19], Theorem 4.19) it follows that, with proper scaling, $S(N(\lfloor ct \rfloor))$ converges as $c \rightarrow \infty$, to $A(E_t)$, where A is either the Brownian motion or a stable Lévy process, and E_t is the inverse (or a hitting time) of a standard β -stable subordinator $\{D(t), t \geq 0\}$. Meerschaert and colleagues have shown the convergence to hold in M_1 and J_1 Skorokhod topology ([17, 21]), and have obtained other Lévy processes as the outer process in the limit by employing triangular arrays [18]. The case of correlated jumps given by the stationary linear process was considered in [15], where the outer process in the limit was either a stable Lévy process or a linear fractional stable motion, depending on the strength of the dependence in the particle jump sequence. In this paper, we define a different sequence of correlated particle jumps defined based on urn schemes that would result in weak convergence of the corresponding random walk to a Pearson diffusion. Pearson diffusions have stationary distributions of Pearson type. They include Ornstein-Uhlenbeck (OU), Cox-Ingersoll-Ross (CIR), Jacobi, Reciprocal Gamma, Fisher-Snedecor and Student diffusions ([1, 2]), and in this paper we consider the first three with non-heavy tailed stationary distributions. Then we separate jumps by the waiting times to obtain the CTRWs converging to fractional Pearson diffusions (fPDs).

Another approach to the limiting behaviour of CTRW is given in the recent paper [10] and book [11] by Kolokoltsov, where the spatially non-homogeneous case is treated, i.e. the case in which the

jump distribution depends on current particle position. To study the limiting behavior of such CTRWs, Kolokoltsov develops the theory of subordination of Markov processes by the hitting-time process, showing that this procedure leads to generalized fractional evolutions. Regarding the connection between scaling limits of CTRWs and Kolmogorov backward and forward equations we refer to [3].

The paper is organized as follows. We begin with a historical note on the subject of urn-scheme models based on a historically important paper [14] and a recent survey paper [7] (heavily referring to [12]), and a classical book [9]. Section 3 gives the short overview of the general theory on Pearson diffusions, followed by Section 4 on fractional Pearson diffusions, focusing on the non-heavy tailed cases. Finally, Section 5 is dedicated to generalizations of results from [7] and [9] and constructing the CTRWs converging to fractional Ornstein-Uhlenbeck, Cox-Ingersoll-Ross or Jacobi diffusions using the following steps:

1. Obtaining the non-fractional Pearson diffusion $(X_t, t \geq 0)$ as the scaling limit of a suitably chosen or constructed discrete time Markov chain.
2. Defining the corresponding CTRW process by taking the composition of the Markov chain defined in step 1 and the process $N_t = \max\{n \geq 0: T_n \leq t\}$, modeling the number of jumps up to time $t \geq 0$.
3. Proving that the composition converges to the corresponding fPD $(X(E_t), t \geq 0)$, a composition of non-fractional Pearson diffusion with the inverse $(E_t, t \geq 0)$ of the β -stable subordinator with $\beta \in (0, 1)$.

2 Historical roots

Many processes observed in science can be mathematically described by the urn-scheme models, where a discrete time Markov chain moves from one state to another. One of the most famous urn-scheme models is the Wright-Fisher model for gene mutations in a population under various assumptions. A rich class of the limiting processes could be obtained from it. One of the simplest urn-scheme models is the Bernoulli-Laplace urn model, named after D. Bernoulli and P. S. Laplace, the pioneers of probability theory, and later studied in depth by [14]. This model considers two urns, urn A containing j balls and urn B containing k balls. Of the total $j + k$ balls in two urns, suppose that r balls are white, and $(j + k - r)$ are black. At time $n \in \mathbb{N}$ one ball is drawn randomly from urn A and another from urn B . The ball drawn from urn A is placed into urn B , the ball drawn from urn B is placed into urn A . Let X_n be the number of white balls in urn B at time $n \in \mathbb{N}$. Then $(k - X_n)$ is the number of black balls in urn B , $(r - X_n)$ is the number of white balls in urn A and $(j - r + X_n)$ is the number of black balls in urn A at time $n \in \mathbb{N}$. Therefore, $(X_n, n \in \mathbb{N})$ is a homogeneous discrete time Markov chain with state space $S = \{\max\{0, r - j\}, \dots, \min\{k, r\}\}$ and the following transition probabilities:

$$P(X_{n+1} = y | X_n = x) = p_{x,y} = \begin{cases} \frac{(r-x)(k-x)}{jk} & , \quad y = x + 1 \\ \frac{(r-x)x + (j-r+x)(k-x)}{jk} & , \quad y = x \\ \frac{(j-r+x)x}{jk} & , \quad y = x - 1 \\ 0 & , \quad \text{otherwise.} \end{cases}$$

In his book [12], Laplace worked with a particular case of this model, in which each urn contains n balls and also n out of total $2n$ balls are black. In this setting he defined a discrete time Markov chain $(Z_r^{(n)}, r \geq 0)$ with the state space $\{0, 1, 2, \dots, n\}$, where n is the number of balls in urn A and r is the number of draws, with transition probabilities

$$p_{x,x+1} = \left(1 - \frac{x}{n}\right)^2, \quad p_{x,x} = 2\frac{x}{n} \left(1 - \frac{x}{n}\right), \quad p_{x,x-1} = \left(\frac{x}{n}\right)^2 \quad \text{and } 0 \text{ otherwise.}$$

Laplace was interested in finding the limiting distribution of this irreducible, aperiodic, reversible and ergodic Markov chain with stationary/ergodic distribution

$$\pi = \{\pi_0, \dots, \pi_n\}, \quad \pi_i = \frac{\binom{n}{i} \binom{n}{n-i}}{\binom{2n}{n}}.$$

By denoting $z_{x,r}$ the probability that there are x white balls in urn A after r draws, he deduced the following partial second-order difference equation:

$$z_{x,r+1} = \left(\frac{x+1}{n}\right)^2 z_{x+1,r} + 2\frac{x}{n} \left(1 - \frac{x}{n}\right) z_{x,r} + \left(1 - \frac{x-1}{n}\right) z_{x-1,r}. \quad (2.1)$$

After that, instead of determining the generating function for z , he approximated the solution of the equation (2.1) by introducing the new space variable μ and new time variable r' and obtained the following relation:

$$x = \frac{1}{2}(n + \mu\sqrt{n}), \quad r = nr'.$$

Then, he introduced the function

$$U(\mu, r') := z_{x,r}$$

of space and time and heuristically deduced the following relations describing the changes of state in the transformed Markov chain:

$$\begin{aligned} z_{x+1,r} &= z_{x,r} + \frac{\partial z_{x,r}}{\partial x} + \frac{1}{2} \frac{\partial^2 z_{x,r}}{\partial x^2} \\ z_{x-1,r} &= z_{x,r} - \frac{\partial z_{x,r}}{\partial x} + \frac{1}{2} \frac{\partial^2 z_{x,r}}{\partial x^2} \\ z_{x,r+1} &= z_{x,r} + \frac{\partial z_{x,r}}{\partial r}. \end{aligned}$$

By another purely heuristic argument, he claimed that function $U(\mu, r')$ satisfies the second-order differential equation

$$\frac{\partial U}{\partial r'} = -\frac{\partial}{\partial \mu}(-2\mu U) + \frac{1}{2} \frac{\partial^2}{\partial \mu^2}(2U) = 2U + 2\mu \frac{\partial U}{\partial \mu} + \frac{\partial^2 U}{\partial \mu^2}, \quad (2.2)$$

which is a special case of the Fokker-Planck or Kolmogorov forward equation defining the OU diffusion:

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x}(\mu(x)p(x,t)) + \frac{1}{2} \frac{\partial^2}{\partial x^2}(\sigma^2(x)p(x,t)).$$

In particular, equation (2.2) is the governing equation for the OU diffusion with infinitesimal mean $\mu(x) = -2x$ and infinitesimal variance $\sigma^2(x) = 2$. While rigorous proofs were not provided, this work had the first mention of the forward equation for the OU diffusion.

A century later, Markov [14] considered a more general model. In his scheme there are also two urns, urn A containing n balls and urn B containing n_1 balls. Out of total $(n + n_1)$ balls, there are $(n + n_1)p$ white and $(n + n_1)q$ black balls ($0 < p < 1, q = 1 - p$). Similar to the Laplace's scheme, by denoting probability that there are x white balls in the urn A after r draws by $z_{x,r}$, Markov obtained the following difference equation:

$$\begin{aligned} z_{x,r+1} &= \frac{x+1}{n} \cdot \frac{n_1q - np + x + 1}{n_1} \cdot z_{x+1,r} + \frac{n-x+1}{n} \cdot \frac{(n+n_1)p - x + 1}{n_1} \cdot z_{x-1,r} \\ &+ \left(\frac{x}{n} \cdot \frac{(n+n_1)p - x}{n_1} + \frac{n-x}{n} \cdot \frac{n_1q - np + x}{n_1} \right) \cdot z_{x,r}. \end{aligned} \quad (2.3)$$

Next, Markov introduced new space variable μ and new time variable ρ and obtained the following relation:

$$x = np + \mu \frac{1}{\Delta\mu}, \quad r \left(\frac{1}{n} + \frac{1}{n_1} \right) = 2\rho, \quad \text{where } \Delta\mu = \sqrt{\frac{n+n_1}{2pqnn_1}},$$

demanding that the ratio of number of balls in urn A and urn B remains constant at all times, i.e. $n_1\alpha = n$ for some $\alpha > 0$. Obviously, for $\alpha = 1$ and $p = q = \frac{1}{2}$ this model reduces to the previously described Laplace's urn scheme.

Finally, the difference equation (2.3) was, again by pure heuristics, approximated by the corresponding Fokker-Planck equation. In order to do so, Markov defined the space and time dependent function $U(\mu, \rho) := z_{x,r}$, smooth enough for obtaining approximations

$$\begin{aligned} z_{x+1,r} &= U(\mu + \Delta\mu, \rho) \approx U + \Delta\mu \frac{\partial U}{\partial \mu} + \frac{1}{2} (\Delta\mu)^2 \frac{\partial^2 U}{\partial \mu^2} \\ z_{x-1,r} &= U(\mu - \Delta\mu, \rho) \approx U - \Delta\mu \frac{\partial U}{\partial \mu} + \frac{1}{2} (\Delta\mu)^2 \frac{\partial^2 U}{\partial \mu^2} \\ z_{x,r+1} &= U(\mu, \rho + \frac{1+\alpha}{2n}) \approx U + \frac{1+\alpha}{2n} \frac{\partial U}{\partial \rho}, \end{aligned}$$

precise up to the order $o(n^{-1})$.

By combining these approximations with the new space and time transformations and putting it into (2.3), Markov obtained the second order differential equation

$$\frac{\partial U}{\partial \rho} = 2U + 2\mu \frac{\partial U}{\partial \mu} + \frac{\partial^2 U}{\partial \mu^2},$$

the Fokker-Planck equation for the OU diffusion with infinitesimal mean $\mu(x) = -2x$ and infinitesimal variance $\sigma^2(x) = 2$, completely coinciding with Laplace's result.

We now briefly discuss the Wright-Fisher urn scheme, named after S. Wright and R. Fisher. Wright-Fisher urn scheme is a model that describes gene mutations (in some genetic pool) over time, strongly influencing selection and sampling forces in the corresponding population. There are several different versions of this model, and we use the scheme described in [9].

Suppose that in a population of size n each individual is either of type A or type a . Let i be the number of A -types in the population. Therefore, the remaining $(n - i)$ population members are a -types. The next generation of a population is produced depending on the influence of mutation, selection and sampling forces. Once born, individual of A -type can mutate in a -type with probability α and individual of a -type can mutate in A -type with probability β . Taking into account parental population comprised of i A -types and $(n - i)$ a -types, the expected fraction of A -types after mutation is

$$\frac{i}{n} (1 - \alpha) + \left(1 - \frac{i}{n}\right) \beta,$$

while the expected fraction of a -types is

$$\frac{i}{n} \alpha + \left(1 - \frac{i}{n}\right) (1 - \beta).$$

Survival ability of each type is modeled by parameter s so that the ratio of A -types over a -types is equal to $1 + s$, meaning that A -type is selectively superior to a -type. Then the expected fraction of mature A -types before reproduction is

$$p_i = \frac{(1 + s) [i(1 - \alpha) + (n - i)\beta]}{(1 + s) [i(1 - \alpha) + (n - i)\beta] + [i\alpha + (n - i)(1 - \beta)]}. \quad (2.4)$$

The last assumption of this model is that the composition of the next generation is determined through n binomial trials, where the probability of producing an A -type in each trial is p_i . This model, tracking the number of A -types in population over time, can be described by the discrete-time Markov chain $(G_r^n, r \in \mathbb{N}_0)$ with the state space $\{0, 1, 2, \dots, n\}$ and transition probabilities

$$p_{ij} = \binom{n}{j} p_i^j (1 - p_i)^{n-j}. \quad (2.5)$$

This model is parametrized by α , β , and $s \in [0, 1]$. Depending on their values, different limiting diffusions could be obtained. In Section 7 of this paper, we present two different settings that lead to generally parametrized Jacobi and CIR diffusions. The procedures of constructing these diffusions as limits of some suitably selected discrete-time Markov chain (urn scheme) are based on examples given in [9, p. 176-183], where only heuristic arguments are given for only special cases of the limiting Jacobi and CIR diffusions.

From this review of history, we see that the connection between discrete-time Markov chains in urn-scheme models and some limiting continuous time stochastic processes has been brought up in the literature. Today, the conjectures of Laplace and Markov could be rigorously proved by modern techniques of analysis and probability theory, involving the convergence of evolution operators of discrete time Markov chains. In this paper we derive such connections between several urn-scheme models and the corresponding diffusions. More precisely, for appropriately chosen discrete time Markov chains we derive the corresponding limiting diffusions, then define the CTRWs and their fractional non-heavy tailed Pearson diffusion limits.

3 Pearson diffusions

In 1931 Kolmogorov observed that the differential equation for the invariant density $\mathbf{m}(x), x \in \mathbb{R}$, of the classical Markovian diffusions satisfying

$$\mathbf{m}'(x)/\mathbf{m}(x) = [a(x) - b'(x)]/[b(x)] = [(a_1 - 2b_2)x + (a_0 - b_1)]/[b_2x^2 + b_1x + b_0],$$

in case of a linear drift $a(x) = a_1x + a_0 = -\theta(x - \mu)$ and the quadratic squared diffusion coefficient $\sigma^2(x) = 2\theta b(x)$, becomes the famous Pearson equation introduced by Pearson in 1914 to unify important families of distributions (see [20]). The family of Pearson distributions is categorized into six well known and widely applied parametric subfamilies: normal, gamma, beta, Fisher-Snedecor, reciprocal gamma and Student distributions. The latter three distributions are heavy-tailed, while the former three are not. Pearson diffusion is defined as the unique strong solution of the following stochastic differential equation (SDE):

$$dX_t = -\theta(X_t - \mu)dt + \sqrt{2\theta(b_2X_t^2 + b_1X_t + b_0)}dW_t, \quad t \geq 0,$$

where $\mu \in \mathbb{R}$ is the stationary mean, $\theta > 0$ is the scaling of time determining the speed of the mean reversion, and b_0 , b_1 and b_2 must be such that the square root is well defined when X_t is in the state space (l, L) .

Pearson diffusions could be categorized into six subfamilies (see [13]), according to the degree of the polynomial $\sigma^2(x)$ and, in the quadratic case, according to the sign of its leading coefficient b_2 and the sign of its discriminant $\Delta = b_1^2 - 4b_0b_2$. The six types are:

- constant $b(x)$ - Ornstein-Uhlenbeck (OU) process, characterized by a normal stationary distribution,
- linear $b(x)$ - Cox-Ingersol-Ross (CIR) process, characterized by a gamma stationary distribution,
- quadratic $b(x)$ with $b_2 < 0$ - Jacobi (JC) diffusion, characterized by a beta stationary distribution,
- quadratic $b(x)$ with $b_2 > 0$ and $\Delta(b) > 0$ - Fisher-Snedecor (FS) diffusion, characterized by Fisher-Snedecor stationary distribution,
- quadratic $b(x)$ with $b_2 > 0$ and $\Delta(b) = 0$ - reciprocal gamma (RG) diffusion, characterized by reciprocal gamma stationary distribution,
- quadratic $b(x)$ with $b_2 > 0$ and $\Delta(b) < 0$ - Student (ST) diffusion, characterized by Student stationary distribution.

More details on the first three types of Pearson diffusions with non-heavy tailed stationary distributions are found in the classical book [9].

Time evolution of the diffusion is described by the Kolmogorov forward (Fokker-Planck) and backward partial differential equations (PDEs) for the transition density $p(x, t; y, s) = \frac{d}{dx}P(X_t \leq x | X_s = y)$. We consider only time-homogeneous diffusions for which $p(x, t; y, s) = p(x, t - s; y, 0)$ for $t > s$ and thus we write $p(x, t; y) = \frac{d}{dx}P(X_t \leq x | X_0 = y)$. The transition density $p(x, t; y)$ solves the following Cauchy problem with space-varying polynomial coefficients:

$$\frac{\partial p(x, t; y)}{\partial t} = \mathcal{G}p(x, t; y), \quad p(x, 0; y) = g(y)$$

with the point-source initial condition, where $\mathcal{G}g(y) = \left(\mu(y)\frac{\partial}{\partial y} + \frac{\sigma^2(y)}{2}\frac{\partial^2}{\partial y^2}\right)g(y)$ is the infinitesimal generator of the diffusion, playing a key role in its analytical properties. The generators of non-heavy tailed OU, CIR, and Jacobi diffusions considered in this paper have purely discrete spectra, consisting of infinitely many simple eigenvalues $(\lambda_n, n \in \mathbb{N})$ (see [13]). Corresponding orthonormal eigenfunctions $(Q_n(x), n \in \mathbb{N})$ are classical systems of orthogonal polynomials: Hermite polynomials for the OU process, Laguerre polynomials for the CIR process and Jacobi polynomials for the Jacobi diffusion.

4 Fractional Pearson diffusions (fPDs)

Let $(X_1(t), t \geq 0)$ be a Pearson diffusion. The fPD $(X_\beta(t), t \geq 0)$ is defined via a non-Markovian time-change $E(t)$ independent of $X_1(t)$:

$$X_\beta(t) := X_1(E_t), \quad t \geq 0,$$

where $E_t = \inf\{x > 0 : D_x > t\}$ is the inverse of the standard β -stable Lévy subordinator $(D_t, t \geq 0)$, $0 < \beta < 1$, with the Laplace transform $E[e^{-sD_t}] = \exp\{-ts^\beta\}$, $s \geq 0$. Since E_t rests for periods of time with non-exponential distribution, the process $(X_\beta(t), t \geq 0)$ is not a Markov process.

We say that the non-Markovian process $X_\beta(t)$ has a transition density $p_\beta(x, t; y)$ if

$$P(X_\beta(t) \in B | X_\beta(0) = y) = \int_B p_\beta(x, t; y)dx$$

for any Borel subset B of (l, L) .

The governing equations for the fPDs are time-fractional forward and backward Kolmogorov equations, with the time-fractional derivative (regularized non-local operator) defined in the Caputo sense (see [19]):

$$\frac{\partial^\beta u}{\partial t^\beta} = \begin{cases} \frac{\partial u}{\partial t}(t, x), & \text{if } \beta = 1 \\ \frac{1}{\Gamma(1-\beta)} \frac{\partial}{\partial t} \int_0^t (t-\tau)^{-\beta} u(\tau, x) d\tau - \frac{u(0, x)}{t^\beta}, & \text{if } \beta \in (0, 1). \end{cases}$$

Using spectral methods, the representations for the transition densities of the non-heavy-tailed fPDs (OU, CIR, Jacobi) were obtained in [13]. Namely, it has been shown that the series

$$p_\beta(x, t; y) = \mathbf{m}(x) \sum_{n=0}^{\infty} \mathcal{E}_\beta(-\lambda_n t^\beta) Q_n(y) Q_n(x) \quad (4.1)$$

converges for fixed $t > 0$, $x, y \in (l, L)$, where $\mathcal{E}_\beta(-z) = \sum_{j=0}^{\infty} (-z)^j / \Gamma(1 + \beta j)$, $z \geq 0$, is the Mittag-Leffler function. The series (4.1) satisfies $p_\beta(x, 0; y) = \delta(x - y)$. These spectral representations were then used to obtain the explicit strong solutions of the corresponding fractional Cauchy problems for both backward and forward equations. The solutions were given as series that converge and satisfy the fractional partial differential equations pointwise. The results obtained in this paper can be used to implement another method of obtaining the solutions for the fractional Cauchy problems with space-varying polynomial coefficients appearing in the generator. Namely, by simulating the Markov chains and waiting times, particle tracking method can provide the means for numerically evaluating the strong solutions.

5 Transition operators of the discrete time Markov chains

To implement the diffusion approximation of discrete-time Markov chains, we will use transition kernels and transition operators. Let μ be an arbitrary probability kernel on a measurable space (S, \mathcal{S}) . The associated transition operator T is defined as

$$Tf(x) = (Tf)(x) = \int \mu(x, dy) f(y), \quad x \in S, \quad (5.1)$$

where $f: S \rightarrow \mathbb{R}$ is assumed to be measurable and either bounded or nonnegative. From the approximation of f by simple functions and the monotone convergence argument, it follows that Tf is again a measurable function on S . Furthermore, T is a positive contraction operator: $0 \leq f \leq 1$ implies that $0 \leq Tf \leq 1$. For more details on transition operators and their importance to the study of Markov processes, we refer to [8], Chapter 19.

Theorems 19.25 and 19.28 from [8] are crucial for the technique used for obtaining the non-fractional Pearson diffusion as the scaling limit of a suitably chosen Markov chain with known transition operator. Markov chain will be defined in terms of its state space and transition probabilities (the integral in (5.1) then becomes a sum). A few additional definitions are given for clarity before the statement of the theorem.

By $\mathbb{D}(S)$ we denote the space of right continuous functions with left limits defined on \mathbb{R}^+ with values in S endowed with Skorokhod J_1 topology. Consider the Banach space of bounded continuous functions on space S , where $S = \mathbb{R}$ in the OU case, $S = [0, 1]$ in the Jacobi case and $S = [0, +\infty)$ in the CIR case, with the supremum norm. For a closed operator \mathcal{A} with domain \mathcal{D} , a core for \mathcal{A} is a linear subspace $D \subset \mathcal{D}$ such that the restriction $\mathcal{A}|_D$ has closure \mathcal{A} . In that case, \mathcal{A} is clearly uniquely determined by its restriction $\mathcal{A}|_D$.

Theorems 1.6 and 2.1 from [5, Section 8] give sufficient conditions for $C_c^\infty(S)$ being a core of the diffusion infinitesimal generator. In particular, Jacobi diffusion satisfies conditions of Theorem 1.6, while other Pearson diffusions satisfy conditions of Theorem 2.1, which means $C_c^\infty(S)$ is a core for Pearson diffusions. Therefore $C_c^3(S)$, as a broader space, can be referred to as the core as well.

Theorem 5.1. *Let $(Y^{(n)}, n \in \mathbb{N})$ be a sequence of discrete-time Markov chains on S with transition operators $(U_n, n \in \mathbb{N})$. Consider a Feller process X on S with semigroup T_t and generator \mathcal{A} . Fix a core D for the generator \mathcal{A} , and assume that $(h_n, n \in \mathbb{N})$ is the sequence of positive reals tending to zero as $n \rightarrow \infty$. Let*

$$A_n = h_n^{-1}(U_n - I), \quad T_{n,t} = U_n^{\lfloor t/h_n \rfloor}, \quad X_t^n = Y_{\lfloor t/h_n \rfloor}^n.$$

Then the following statements are equivalent:

- a) If $f \in D$, there exist some $f_n \in \text{Dom}(A_n)$ with $f_n \rightarrow f$ and $A_n f_n \rightarrow \mathcal{A}f$ as $n \rightarrow \infty$
- b) $T_{n,t} \rightarrow T_t$ strongly for each $t > 0$
- c) $T_{n,t} f_n \rightarrow T_t f$ for each $f \in C_0$, uniformly for bounded $t > 0$
- d) if $X_0^{(n)} \Rightarrow X_0$ in S , then $X^n \Rightarrow X$ in the Skorokhod space $\mathbb{D}(S)$ with the J_1 topology.

The proof could be found in [8, Theorem 19.28, page 387].

6 Correlated CTRW for the OU process based on the Laplace-Bernoulli urn scheme

In this section, we define the Bernoulli-Laplace urn-scheme model, but compared to early ideas of Laplace and Markov summarized in Section 2, with crucial changes in space and time transformations in order to obtain more general limiting process, i.e. the OU diffusion with general parameters. The urn scheme consists of two urns, A and B , each containing n balls. Furthermore, n out of total $2n$ balls are black. At each step one ball is randomly chosen from each urn. The ball drawn from urn A is then placed into urn B and the ball drawn from urn B is placed into urn A . We are interested in the number of white balls in the urn A after $r \in \mathbb{N}$ draws.

Let for each $n \in \mathbb{N}$, $(Z_r^{(n)}, r \in \mathbb{N}_0)$ be the Markov chain with the state space $\{0, 1, 2, \dots, n\}$, where n is the number of balls in urn A , r is the number of draws and $Z_r^{(n)}$ is the number of white balls in the urn A after r draws. The transition probabilities for this Markov chain are as follows:

$$p_{i,i+1} = \left(1 - \frac{i}{n}\right)^2, \quad p_{i,i} = 2\frac{i}{n} \left(1 - \frac{i}{n}\right), \quad p_{i,i-1} = \left(\frac{i}{n}\right)^2, \quad 0 \text{ otherwise.} \quad (6.1)$$

We assume that the initial state of the chain is given by $Z_0^{(n)} = \lfloor \frac{1}{2}(n + (ax + b)\sqrt{n}) \rfloor$, where $a \neq 0$ and b are fixed parameters with values in \mathbb{R} and $x \in \mathbb{R}$. Now we introduce the space variable (i.e., state) transformation

$$i' = \frac{1}{a\sqrt{n}} (2i - n - b\sqrt{n}), \quad i \in \{0, 1, 2, \dots, n\},$$

to obtain a new Markov chain $(H_r^{(n)}, r \in \mathbb{N}_0)$, where

$$H_r^{(n)} = \frac{1}{a\sqrt{n}} (2Z_r^{(n)} - n - b\sqrt{n}). \quad (6.2)$$

Let us denote $x_n = \lfloor \frac{1}{2}(n + (ax + b)\sqrt{n}) \rfloor$ and $x'_n = \frac{1}{a\sqrt{n}} (2x_n - n - b\sqrt{n})$. Since our starting Markov chain $(Z_r^{(n)}, r \in \mathbb{N}_0)$ is homogenous, using (6.1) we obtain the transition operator of the Markov chain $(H_r^{(n)}, r \in \mathbb{N}_0)$:

$$\begin{aligned} T_n f(x'_n) &= \left(1 - \frac{1}{2} \left(1 + \frac{ax'_n + b}{\sqrt{n}}\right)\right)^2 f\left(x'_n + \frac{2}{a\sqrt{n}}\right) + \left(1 + \frac{ax'_n + b}{\sqrt{n}}\right) \left(1 - \frac{1}{2} \left(1 + \frac{ax'_n + b}{\sqrt{n}}\right)\right) f(x'_n) \\ &\quad + \left(\frac{1}{2} \left(1 + \frac{ax'_n + b}{\sqrt{n}}\right)\right)^2 f\left(x'_n - \frac{2}{a\sqrt{n}}\right) \\ &= \left(1 - \frac{x_n}{n}\right)^2 f\left(\frac{1}{a\sqrt{n}} (2(x_n + 1) - n - b\sqrt{n})\right) + 2\frac{x_n}{n} \left(1 - \frac{x_n}{n}\right) f\left(\frac{1}{a\sqrt{n}} (2x_n - n - b\sqrt{n})\right) \\ &\quad + \left(\frac{x_n}{n}\right)^2 f\left(\frac{1}{a\sqrt{n}} (2(x_n - 1) - n - b\sqrt{n})\right). \end{aligned} \quad (6.3)$$

For the application of Theorem 5.1, operators A_n are defined as follows:

$$A_n := \frac{\theta}{2}(T_n - I), \quad \theta > 0, \quad f_n \in \text{Dom}(A_n), \quad f_n(x) := f(x'_n). \quad (6.4)$$

where $f \in C_c^3(\mathbb{R})$.

The continuous-time process $(X_t^{(n)}, t \geq 0)$ is obtained by the following scaling of time in $(H_r^{(n)}, r \in \mathbb{N}_0)$, for each $n \in \mathbb{N}$:

$$X_t^{(n)} := H_{\lfloor \frac{\theta}{2} nt \rfloor}^{(n)}, \quad \theta > 0. \quad (6.5)$$

Let $X = (X_t, t \geq 0)$ be the OU diffusion, i.e., the solution of the SDE

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW(t), \quad t \geq 0,$$

where $(W(t), t \geq 0)$ is the standard Brownian motion. The drift $\mu(x)$ and diffusion parameter $\sigma^2(x)$ are given by

$$\mu(x) = -\tau(x - \mu), \quad \sigma(x) = \sqrt{2\tau\sigma^2}, \quad \tau > 0, \quad \mu \in \mathbb{R}, \quad \sigma \in \mathbb{R},$$

and the infinitesimal generator is defined by the following expression:

$$\mathcal{A}f(x) = \left[\mu(x) \frac{\partial}{\partial x} + \frac{\sigma^2(x)}{2} \frac{\partial^2}{\partial x^2} \right] f(x) = \left[-\tau(x - \mu) \frac{\partial}{\partial x} + \tau\sigma^2 \frac{\partial^2}{\partial x^2} \right] f(x), \quad f \in C_c^3(\mathbb{R}). \quad (6.6)$$

The next theorem established that the OU diffusion is the limiting process of the sequence of rescaled discrete time Markov chains given by (6.2).

Theorem 6.1. *Let $(H_r^{(n)}, r \in \mathbb{N}_0)$ for each $n \in \mathbb{N}$, be the Markov chain defined by (6.2) with transition operator (6.3). Let $(X_t^{(n)}, t \geq 0)$ for each $n \in \mathbb{N}$, be the time-changed process corresponding to the Markov chain $(H_r^{(n)}, r \in \mathbb{N}_0)$, with the time-change (6.5). Let operators $(A_n, n \in \mathbb{N})$ be defined by (6.4). Then*

$$X^n \Rightarrow X \text{ in } \mathbb{D}(\mathbb{R}),$$

where $X = (X_t, t \geq 0)$ is the Ornstein-Uhlenbeck diffusion with infinitesimal generator \mathcal{A} given by (6.6) and parameters $\tau = \theta$, $\mu = -\frac{b}{a}$ and $\sigma = \frac{1}{a\sqrt{2}}$.

Proof. In the setup of the Theorem 5.1, we will first prove that statement a) of this theorem is valid for A_n defined by (6.4) and \mathcal{A} defined by (6.6). Then we will use the equivalence of statements a) and d) to obtain convergence $X^n \Rightarrow X$ in $\mathbb{D}(\mathbb{R})$ under assumption $X_0^{(n)} \Rightarrow X_0, n \rightarrow \infty$.

Therefore, first we prove the validity of the statement a) from Theorem 5.1 in our setting:

$$\|f_n - f\|_\infty = \sup_{x \in \mathbb{R}} |f_n(x) - f(x)| = \sup_{x \in \mathbb{R}} |f(x'_n) - f(x)|$$

and since by the definition of function $[\cdot]$

$$ax'_n \leq ax \leq ax'_n + \frac{2}{\sqrt{n}}$$

we have

$$x'_n \rightarrow x, \quad n \rightarrow \infty,$$

so using this we obtain

$$\|f_n - f\|_\infty \rightarrow 0, \quad n \rightarrow \infty$$

and

$$\|A_n f_n - \mathcal{A}f\|_\infty = \sup_{x \in \mathbb{R}} |A_n f_n(x) - \mathcal{A}f(x)| = \sup_{x \in \mathbb{R}} |A_n f(x'_n) - \mathcal{A}f(x)|.$$

By the triangle inequality we have

$$\|A_n f_n - \mathcal{A}f\|_\infty \leq \sup_{x \in \mathbb{R}} |A_n f(x'_n) - A_n f(x)| + \sup_{x \in \mathbb{R}} |A_n f(x) - \mathcal{A}f(x)|. \quad (6.7)$$

Now it follows that

$$\begin{aligned}
A_n f(x) &= \frac{\theta}{2} n (T_n f(x) - f(x)) \\
&= \frac{\theta}{2} n \left(1 - \frac{1}{2} \left(1 + \frac{ax+b}{\sqrt{n}} \right) \right)^2 f \left(x + \frac{2}{a\sqrt{n}} \right) + \frac{\theta}{2} n \left(1 + \frac{ax+b}{\sqrt{n}} \right) \left(1 - \frac{1}{2} \left(1 + \frac{ax+b}{\sqrt{n}} \right) \right) f(x) \\
&\quad + \frac{\theta}{2} n \left(\frac{1}{2} \left(1 + \frac{ax+b}{\sqrt{n}} \right) \right)^2 f \left(x - \frac{2}{a\sqrt{n}} \right) - \frac{\theta}{2} n f(x) \\
&= \frac{\theta}{2} n \left(f \left(x + \frac{2}{a\sqrt{n}} \right) - f(x) \right) - \frac{\theta}{2} n \left(1 + \frac{ax+b}{\sqrt{n}} \right) \left(f \left(x + \frac{2}{a\sqrt{n}} \right) - f(x) \right) \\
&\quad + \frac{\theta n}{8} \left(\left(1 + \frac{ax+b}{\sqrt{n}} \right) \right)^2 \left(f \left(x + \frac{2}{a\sqrt{n}} \right) - 2f(x) + f \left(x - \frac{2}{a\sqrt{n}} \right) \right) \\
&= -\theta \left(x + \frac{b}{a} \right) \frac{\left(f \left(x + \frac{2}{a\sqrt{n}} \right) - f(x) \right)}{\frac{2}{a\sqrt{n}}} + \frac{\theta}{2a^2} \left(\left(1 + \frac{ax+b}{\sqrt{n}} \right) \right)^2 \frac{\left(f \left(x + \frac{2}{a\sqrt{n}} \right) - 2f(x) + f \left(x - \frac{2}{a\sqrt{n}} \right) \right)}{\left(\frac{2}{a\sqrt{n}} \right)^2} \\
&\rightarrow -\theta \left(x + \frac{b}{a} \right) f'(x) + \frac{1}{2} \frac{\theta}{a^2} f''(x), \quad n \rightarrow \infty.
\end{aligned}$$

Comparing the above limit with (6.6) we see that

$$A_n f(x) \rightarrow \mathcal{A}f(x), \quad n \rightarrow \infty \quad (6.8)$$

with $\tau = \theta$, $\mu = -\frac{b}{a}$ and $\sigma = \frac{1}{a\sqrt{2}}$.

Since the function f is in the space $C_c^3(\mathbb{R})$, the above convergence holds uniformly. Furthermore

$$x'_n \rightarrow x, \quad n \rightarrow \infty,$$

so we have

$$\sup_{x \in \mathbb{R}} |A_n f(x'_n) - A_n f(x)| \rightarrow 0, \quad n \rightarrow \infty.$$

Using the above together with (6.8) in (6.7) we finally obtain

$$\|A_n f_n - \mathcal{A}f\|_\infty \rightarrow 0, \quad n \rightarrow \infty,$$

and since

$$X_0^{(n)} \Rightarrow X_0, \quad n \rightarrow \infty \quad \iff \quad x'_n \rightarrow x, \quad n \rightarrow \infty,$$

by Theorem 5.1 we obtain $X^n \Rightarrow X$ in $\mathbb{D}(\mathbb{R})$, where X is the generally parametrized OU diffusion. \square

7 Correlated CTRWs for non-heavy-tailed Pearson diffusions based on the Wright-Fisher model

In this section we present two different versions of the Wright-Fisher model that lead to generally parametrized Jacobi and CIR diffusions.

7.1 Jacobi diffusion

Jacobi diffusion $Y = (Y_t, t \geq 0)$ is defined as the solution of the SDE

$$dY_t = -\gamma(Y_t - \mu)dt + \sqrt{2\gamma\delta Y_t(1 - Y_t)}dW_t, \quad t \geq 0, \quad \mu \in (0, 1), \quad \gamma, \delta > 0,$$

with the infinitesimal generator

$$\mathcal{A}f(y) = -\gamma(y - \mu)f'(y) + \frac{1}{2}(2\gamma\delta)y(1 - y)f''(y), \quad f \in C_c^3([0, 1]). \quad (7.1)$$

In this case we assume that there are no survival abilities and that probability of mutation of each type (A to a and a to A) is proportional to the size of the population, i.e.

$$\alpha = \frac{a}{n}, \quad \beta = \frac{b}{n}, \quad s = 0, \quad a, b > 0.$$

Therefore, the expected fraction of mature A -types (2.4) becomes

$$p_i = \frac{i}{n} \left(1 - \frac{a}{n}\right) + \left(1 - \frac{i}{n}\right) \frac{b}{n}. \quad (7.2)$$

For each $n \in \mathbb{N}$ we define a new Markov chain $(H_r^{(n)}, r \in \mathbb{N}_0)$ as follows:

$$H_r^{(n)} = \frac{1}{n} G_r^{(n)}. \quad (7.3)$$

The state space of this new Markov chain is $\{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$ and we assume that the initial state of the original chain $(G_r^{(n)}, r \in \mathbb{N}_0)$ is given by

$$i(y) = i = G_0^n = \lfloor ny \rfloor, \quad y \in [0, 1].$$

We emphasize that the initial state is a function of y , but we will use the notation i for simplicity. The transition operator T_n of the Markov chain $(H_r^{(n)}, r \in \mathbb{N}_0)$ is defined as follows:

$$T_n f \left(\frac{i}{n} \right) = \sum_{j=0}^n p_{ij} f \left(\frac{j}{n} \right), \quad (7.4)$$

where p_{ij} is defined in (2.5) and p_i in (7.2).

Define the operator

$$A_n := \theta n (T_n - I), \quad \theta > 0, \quad f_n \in \text{Dom}(A_n), \quad f_n(y) := f \left(\frac{\lfloor ny \rfloor}{n} \right) = f \left(\frac{i}{n} \right), \quad (7.5)$$

where $f \in C_c^3([0, 1])$.

Now by the following scaling of time in $(H_r^{(n)}, r \in \mathbb{N}_0)$, for each $n \in \mathbb{N}$ we obtain the corresponding continuous-time process $(Y_t^{(n)}, t \geq 0)$:

$$Y_t^{(n)} := H_{\lfloor \theta n t \rfloor}^{(n)}, \quad \theta > 0. \quad (7.6)$$

The next theorem states that the Jacobi diffusion could be obtained as the limiting process of the previously time-changed processes $(Y_t^{(n)}, t \geq 0)$.

Theorem 7.1. *Let $(H_r^{(n)}, r \in \mathbb{N}_0)$, for each $n \in \mathbb{N}$, be the Markov chain defined by (7.3) with the transition operator (7.4). Let $Y^n = (Y_t^{(n)}, t \geq 0)$, for each $n \in \mathbb{N}$, be its corresponding time-changed process with the time-change (7.6). Let the operators $(A_n, n \in \mathbb{N})$ be defined by (7.5). Then*

$$Y^n \Rightarrow Y \text{ in } \mathbb{D}([0, 1]),$$

where $Y = (Y_t, t \geq 0)$ is the Jacobi diffusion with the infinitesimal generator \mathcal{A} given by (7.1), and

$$\delta = \frac{1}{2(a+b)}, \quad \gamma = \theta(a+b), \quad \mu = \frac{b}{a+b}.$$

Proof. Once again, we are in the setting of the Theorem 5.1 and follow the same procedure as in the Theorem 6.1. We first prove that statement *a*) of Theorem 5.1 is valid for A_n defined by (7.5) and \mathcal{A} defined by (7.1). Then we use the equivalence of statements *a*) and *d*) from Theorem 5.1 to obtain convergence $Y^n \Rightarrow Y$ in $\mathbb{D}([0, 1])$ under assumption $Y_0^{(n)} \Rightarrow Y_0, n \rightarrow \infty$.

First, we prove the statement *a*) from Theorem 5.1 in our setting:

$$\|f_n - f\|_\infty = \sup_{y \in [0, 1]} |f_n(y) - f(y)| = \sup_{y \in [0, 1]} |f(i/n) - f(y)|.$$

According to the well-known property $\lfloor ny \rfloor \leq ny < \lfloor ny \rfloor + 1$ of the function $\lfloor \cdot \rfloor$, it follows that

$$\frac{i}{n} \leq y < \frac{i}{n} + \frac{1}{n}$$

and therefore

$$\lim_{n \rightarrow \infty} \frac{i}{n} = y.$$

From this we obtain

$$\|f_n - f\|_\infty \rightarrow 0, \quad n \rightarrow \infty$$

and

$$\|A_n f_n - \mathcal{A}f\|_\infty = \sup_{y \in [0,1]} |A_n f_n(y) - \mathcal{A}f(y)| = \sup_{y \in [0,1]} |A_n f(i/n) - \mathcal{A}f(y)|.$$

According to (7.5) it follows that

$$A_n f\left(\frac{i}{n}\right) = \theta n \left(\sum_{j=0}^n p_{ij} f\left(\frac{j}{n}\right) - f\left(\frac{i}{n}\right) \right) = \theta n \sum_{j=0}^n p_{ij} \left(f\left(\frac{j}{n}\right) - f\left(\frac{i}{n}\right) \right).$$

By the Taylor formula for function f around $\frac{i}{n}$ with the mean-value form of the remainder we obtain

$$A_n f\left(\frac{i}{n}\right) = \theta \sum_{j=0}^n p_{ij} (j-i) f'\left(\frac{i}{n}\right) + \theta \sum_{j=0}^n p_{ij} \frac{(j-i)^2}{2n} f''\left(\frac{i}{n}\right) + \theta \sum_{j=0}^n p_{ij} \frac{(j-i)^3}{6n^2} f'''(\zeta), \quad (7.7)$$

where ζ is a real number such that $|\zeta - \frac{i}{n}| < |\frac{j}{n} - \frac{i}{n}|$.

Next, denote

$$\mu(y) := \lim_{n \rightarrow \infty} \sum_{j=0}^n p_{ij} (j-i), \quad \sigma^2(y) := \lim_{n \rightarrow \infty} \sum_{j=0}^n p_{ij} \frac{(j-i)^2}{n}, \quad R_n(y) := \theta \sum_{j=0}^n p_{ij} \frac{(j-i)^3}{6n^2} f'''(\zeta).$$

Taking into account (7.7), we obtain

$$\lim_{n \rightarrow \infty} A_n f\left(\frac{i}{n}\right) = \theta \mu(y) f'(y) + \frac{\theta}{2} \sigma^2(y) f''(y) + \lim_{n \rightarrow \infty} R_n(y). \quad (7.8)$$

From the time-homogeneity of the observed Markov chains and the fact that $f \in C_c^3([0, 1])$, it follows that

$$\begin{aligned} \mu(y) &= \lim_{n \rightarrow \infty} \sum_{j=0}^n p_{ij} (j-i) = \lim_{n \rightarrow \infty} E[G_1^n - G_0^n | G_0^n = i] \\ &= \lim_{n \rightarrow \infty} (np_i - i) = \lim_{n \rightarrow \infty} \left(i \left(1 - \frac{a}{n}\right) + \left(1 - \frac{i}{n}\right) b - i \right) \\ &= -(a+b) \lim_{n \rightarrow \infty} \left(\frac{i}{n}\right) + b = -(a+b)y + b, \end{aligned} \quad (7.9)$$

$$\begin{aligned} \sigma^2(y) &= \lim_{n \rightarrow \infty} \sum_{j=0}^n p_{ij} \frac{(j-i)^2}{n} = \lim_{n \rightarrow \infty} \frac{1}{n} E[(G_1^n - G_0^n)^2 | G_0^n = i] \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} (np_i(1-p_i) + n^2 p_i^2 - 2nip_i + i^2) = \lim_{n \rightarrow \infty} \left(p_i - p_i^2 + \frac{(np_i - i)^2}{n} \right) \\ &= y - y^2 = y(1-y), \end{aligned} \quad (7.10)$$

$$\begin{aligned} \lim_{n \rightarrow \infty} |R_n(y)| &\leq K \left| \theta \lim_{n \rightarrow \infty} \sum_{j=0}^n p_{ij} \frac{(j-i)^3}{6n^2} \right| = K \left| \theta \lim_{n \rightarrow \infty} \frac{1}{n^2} E[(G_1^n - G_0^n)^3 | G_0^n = i] \right| \\ &= K \left| \theta \lim_{n \rightarrow \infty} \frac{1}{n^2} (np_i(1-3p_i+2p_i^2) + 3np_i(np_i-i)(1-p_i) + (np_i-i)^3) \right| \\ &= K \left| \theta \lim_{n \rightarrow \infty} \left(\frac{p_i(1-3p_i+2p_i^2)}{n} + \frac{3p_i(np_i-i)(1-p_i)}{n} + \frac{(np_i-i)^3}{n^2} \right) \right| \\ &= 0, \end{aligned} \quad (7.11)$$

where K is a constant such that $|f'''(\zeta)| \leq K$ and

$$\lim_{n \rightarrow \infty} p_i = y \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{(np_i - i)^2}{n} = 0.$$

Finally, by substituting (7.9), (7.10) and (7.11) in (7.8), we obtain

$$\lim_{n \rightarrow \infty} A_n f \left(\frac{i}{n} \right) = \theta(-y(a+b) + b)f'(y) + \frac{1}{2}\theta y(1-y)f''(y).$$

Now by re-parametrizing

$$\delta = \frac{1}{2(a+b)}, \quad \gamma = \theta(a+b), \quad \mu = \frac{b}{a+b},$$

the last limit becomes

$$\lim_{n \rightarrow \infty} A_n f \left(\frac{i}{n} \right) = -\gamma(y - \mu)f'(y) + \frac{1}{2}(2\gamma\delta)y(1-y)f''(y),$$

which is precisely the infinitesimal generator of the Jacobi diffusion. In the space $C_c^3([0, 1])$ all above limits hold uniformly, therefore we obtain

$$\|A_n f_n - \mathcal{A}f\|_\infty \rightarrow 0, \quad n \rightarrow \infty,$$

and since

$$Y_0^{(n)} \Rightarrow Y_0, \quad n \rightarrow \infty \iff i/n \rightarrow y, \quad n \rightarrow \infty,$$

by Theorem 5.1 we obtain $Y^n \Rightarrow Y$ in $\mathbb{D}([0, 1])$, where Y is the generally parametrized Jacobi diffusion. \square

7.2 CIR process

The CIR diffusion $Z = (Z_t, t \geq 0)$ is defined as the solution of the SDE

$$dZ_t = -\theta \left(Z_t - \frac{b}{a} \right) dt + \sqrt{\frac{\theta}{a}} Z_t dW_t, \quad t \geq 0, \quad \theta > 0, \quad a > 0, \quad b > 0,$$

with the infinitesimal generator

$$\mathcal{A}f(z) = -\theta \left(z - \frac{b}{a} \right) f'(z) + \frac{1}{2} \frac{\theta}{a} z f''(z), \quad f \in C_c^3([0, \infty)). \quad (7.12)$$

In this case, we assume there is only mutation of the order

$$\alpha = \frac{a}{n^d}, \quad \beta = \frac{b}{n}, \quad 0 < d < 1, \quad 0 < a, b < \infty, \quad s = 0,$$

so that expected fraction of A -types (2.4) becomes

$$p_i = \frac{i}{n} \left(1 - \frac{a}{n^d} \right) + \left(1 - \frac{i}{n} \right) \frac{b}{n}. \quad (7.13)$$

For each $n \in \mathbb{N}$, we define the new Markov chain $(H_r^{(n)}, r \in \mathbb{N})$ with the state space $\{0, \frac{1}{n^d}, \dots, \frac{1}{n^{d-1}}\}$

$$H_r^{(n)} = \frac{G_r^{(n)}}{n^d}. \quad (7.14)$$

We assume that the initial state space of the starting Markov chain $(G_r^{(n)}, r \in \mathbb{N}_0)$ is

$$i(z) = i = G_0^n = \lfloor n^d z \rfloor, \quad z \in [0, \infty).$$

We assume that n is always large enough so that $i(z)$ is in the state space of Markov chain $(G_r^{(n)}, r \in \mathbb{N}_0)$. We emphasize that the initial state is a function of z , but we will use notation i for simplicity. The transition operator T_n of the Markov chain $(H_r^{(n)}, n \in \mathbb{N})$ is given by

$$T_n f \left(\frac{i}{n} \right) = \sum_{j=0}^n p_{ij} f \left(\frac{j}{n} \right), \quad (7.15)$$

where p_{ij} is defined in (2.5) and p_i in (7.13). Now we define operator

$$A_n := \frac{\theta}{a}n(T_n - I), \quad \theta > 0, \quad f_n \in \text{Dom}(A_n), \quad f_n(z) := f\left(\frac{\lfloor n^d z \rfloor}{n^d}\right) = f\left(\frac{i}{n^d}\right) \quad (7.16)$$

where $f \in C_c^3([0, \infty))$ and by the following scaling of time in $(H_r^{(n)}, r \in \mathbb{N}_0)$, for each $n \in \mathbb{N}$ we obtain the corresponding continuous-time process $(Z_t^{(n)}, t \geq 0)$:

$$Z_t^{(n)} := H_{\lfloor \frac{\theta}{a}n^d t \rfloor}^{(n)}, \quad \theta > 0. \quad (7.17)$$

The next theorem states that the CIR diffusion could be obtained as the limiting process of the time-changed processes $(Z_t^{(n)}, t \geq 0)$.

Theorem 7.2. *Let $(H_r^{(n)}, r \in \mathbb{N}_0)$, for each $n \in \mathbb{N}$, be the Markov chain defined by (7.14) with the transition operator (7.15). Let $Z^n = (Z_t^{(n)}, t \geq 0)$, for each $n \in \mathbb{N}$, be its corresponding time-changed process, with the time-change (7.17). Let the operators $(A_n, n \in \mathbb{N})$ be defined by (7.16). Then*

$$Z^n \Rightarrow Z \text{ in } \mathbb{D}(\mathbb{R}^+),$$

where $Z = (Z_t, t \geq 0)$ is the CIR diffusion with the infinitesimal generator \mathcal{A} given by (7.12).

Proof. We first prove that statement a) of the Theorem 5.1 is valid for A_n defined by (7.16) and \mathcal{A} defined by (7.12) and then we use the equivalence of statements a) and d) from Theorem 5.1 to obtain convergence $Z^n \Rightarrow Z$ in $\mathbb{D}(\mathbb{R}^+)$ under assumption $Z_0^{(n)} \Rightarrow Z_0, n \rightarrow \infty$.

We first prove the statement a) from Theorem 5.1 in our setting:

$$\|f_n - f\|_\infty = \sup_{z \in \mathbb{R}^+} |f_n(z) - f(z)| = \sup_{z \in \mathbb{R}^+} |f(i/n^d) - f(z)|.$$

By definition of the function $\lfloor \cdot \rfloor$ we have

$$\lfloor n^d z \rfloor \leq n^d z < \lfloor n^d z \rfloor + 1$$

so it follows

$$\frac{i}{n^d} \leq z < \frac{i}{n^d} + \frac{1}{n^d}$$

and therefore

$$\lim_{n \rightarrow \infty} \frac{i}{n^d} = z.$$

so using this we obtain

$$\|f_n - f\|_\infty \rightarrow 0, \quad n \rightarrow \infty$$

and

$$\|A_n f_n - \mathcal{A}f\|_\infty = \sup_{z \in \mathbb{R}^+} |A_n f_n(z) - \mathcal{A}f(z)| = \sup_{z \in \mathbb{R}^+} |A_n f(i/n^d) - \mathcal{A}f(z)|.$$

According to (7.16) it follows

$$A_n f\left(\frac{i}{n^d}\right) = \frac{\theta}{a}n^d \left(\sum_{j=0}^n p_{ij} f\left(\frac{j}{n^d}\right) - f\left(\frac{i}{n^d}\right) \right) = \frac{\theta}{a}n^d \sum_{j=0}^n p_{ij} \left(f\left(\frac{j}{n^d}\right) - f\left(\frac{i}{n^d}\right) \right).$$

Now by the Taylor formula for function f around $\frac{i}{n^d}$ with mean-value form of the remainder we obtain

$$A_n f\left(\frac{i}{n^d}\right) = \frac{\theta}{a} \sum_{j=0}^n p_{ij} (j-i) f'\left(\frac{i}{n^d}\right) + \frac{\theta}{2a} \sum_{j=0}^n p_{ij} \frac{(j-i)^2}{n^d} f''\left(\frac{i}{n^d}\right) + \frac{\theta}{6a} \sum_{j=0}^n p_{ij} \frac{(j-i)^3}{n^{2d}} f'''(\zeta) \quad (7.18)$$

where ζ is a real number such that $|\zeta - \frac{i}{n^d}| < |\frac{j}{n^d} - \frac{i}{n^d}|$.

Let us denote

$$\mu(z) := \lim_{n \rightarrow \infty} \sum_{j=0}^n p_{ij} (j-i), \quad \sigma^2(z) := \lim_{n \rightarrow \infty} \sum_{j=0}^n p_{ij} \frac{(j-i)^2}{n^d}, \quad R_n(z) := \frac{\theta}{6a} \sum_{j=0}^n p_{ij} \frac{(j-i)^3}{n^{2d}} f'''(\zeta).$$

Together with (7.18) we obtain

$$\lim_{n \rightarrow \infty} A_n f \left(\frac{i}{n^d} \right) = \frac{\theta}{a} \mu(z) f'(z) + \frac{\theta}{2a} \sigma^2(z) f''(z) + \lim_{n \rightarrow \infty} R_n(z). \quad (7.19)$$

From the time-homogeneity of the Markov chains, it follows that

$$\begin{aligned} \mu(z) &= \lim_{n \rightarrow \infty} \sum_{j=0}^n p_{ij} (j-i) = \lim_{n \rightarrow \infty} E[G_1^n - G_0^n | G_0^n = i] \\ &= \lim_{n \rightarrow \infty} (np_i - i) = \lim_{n \rightarrow \infty} \left(i \left(1 - \frac{a}{n^d} \right) + \left(1 - \frac{i}{n} \right) b - i \right) \\ &= -a \lim_{n \rightarrow \infty} \left(\frac{i}{n^d} \right) - \lim_{n \rightarrow \infty} \left(\frac{i}{n} \right) b + b = -az + b, \end{aligned} \quad (7.20)$$

$$\begin{aligned} \sigma^2(z) &= \lim_{n \rightarrow \infty} \sum_{j=0}^n p_{ij} \frac{(j-i)^2}{n^d} = \lim_{n \rightarrow \infty} \frac{1}{n^d} E[(G_1^n - G_0^n)^2 | G_0^n = i] \\ &= \lim_{n \rightarrow \infty} \frac{1}{n^d} (np_i(1-p_i) + n^2 p_i^2 - 2nip_i + i^2) \\ &= \lim_{n \rightarrow \infty} \frac{1}{n^d} (np_i(1-p_i) + (np_i - i)^2) \\ &= z, \end{aligned} \quad (7.21)$$

$$\begin{aligned} \lim_{n \rightarrow \infty} |R_n(z)| &\leq K \left| \frac{\theta}{6a} \lim_{n \rightarrow \infty} \sum_{j=0}^n p_{ij} \frac{(j-i)^3}{n^{2d}} \right| = K \left| \frac{\theta}{6a} \lim_{n \rightarrow \infty} \frac{1}{n^{2d}} E[(G_1^n - G_0^n)^3 | G_0^n = i] \right| \\ &= K \left| \frac{\theta}{6a} \lim_{n \rightarrow \infty} \frac{1}{n^{2d}} (np_i(1-3p_i+2p_i^2) + 3np_i(np_i-i)(1-p_i) + (np_i-i)^3) \right| \\ &= K \left| \frac{\theta}{6a} \lim_{n \rightarrow \infty} \left(\frac{np_i}{n^d} \frac{(1-3p_i+2p_i^2)}{n^d} + \frac{3np_i}{n^d} \frac{(np_i-i)(1-p_i)}{n^d} + \frac{(np_i-i)^3}{n^{2d}} \right) \right| \\ &= 0, \end{aligned} \quad (7.22)$$

where K is a constant such that $|f'''(\zeta)| \leq K$ and

$$\frac{np_i}{n^d} \rightarrow z \quad p_i \rightarrow 0, \quad \frac{(np_i-i)^2}{n^d} \rightarrow 0, \quad n \rightarrow \infty.$$

Finally, substituting (7.20), (7.21) and (7.22) in (7.19) we obtain

$$\begin{aligned} \lim_{n \rightarrow \infty} A_n f \left(\frac{i}{n^d} \right) &= \frac{\theta}{a} (-az + b) f'(z) + \frac{\theta}{2a} z f''(z) \\ &= -\theta \left(z - \frac{b}{a} \right) f'(z) + \frac{1}{2} \frac{\theta}{a} z f''(z). \end{aligned}$$

Comparing the obtained limit with (7.12) we see that the limit coincides with the generator of the CIR diffusion. Since $f \in C_c^3([0, \infty))$, all above limits hold uniformly, i.e., we obtain

$$\|A_n f_n - \mathcal{A}f\|_\infty \rightarrow 0, \quad n \rightarrow \infty,$$

and since

$$Z_0^{(n)} \Rightarrow Z_0, \quad n \rightarrow \infty \iff i/n^d \rightarrow z, \quad n \rightarrow \infty,$$

by Theorem 5.1 we obtain $Z^n \Rightarrow Z$ in $\mathbb{D}(\mathbb{R}^+)$, where Z is the generally parametrized CIR diffusion. \square

8 Fractional OU, Jacobi and CIR diffusions as the correlated CTRWs limits

Suppose that $(T_r, r \in \mathbb{N}_0)$, where $T_0 = 0, T_r = G_1 + \dots + G_r$, is the random walk where $G_r \geq 0$ are iid waiting times between particle jumps that are independent of the Markov chain $(H_r^{(n)}, r \in \mathbb{N}_0)$. We assume G_1 is in the domain of attraction of the β -stable distribution with index $0 < \beta < 1$, and that the waiting time of the Markov chain until its r -th move is described by G_r . Let

$$N_t = \max\{r \geq 0: T_r \leq t\} \quad (8.1)$$

be the number of jumps up to time $t \geq 0$. Then the continuous time stochastic process

$$\left(H^{(n)}(N_t), t \geq 0 \right),$$

where $H^{(n)}(N_t)$ is the state of the Markov chain at time $t \geq 0$, is the correlated CTRW process. We now state the main result of this section by defining the correlated CTRWs that converge to the fractional OU, CIR, and Jacobi diffusions.

Theorem 8.1. *Let $(H_r^{(n)}, r \in \mathbb{N}_0)$ be the Markov chain defined by (6.2) in the case of OU diffusion, by (7.3) in the case of Jacobi diffusion and by (7.14) in the case of CIR diffusion. Let $(X^{(n)}(t), t \geq 0)$, $(Y^{(n)}(t), t \geq 0)$ and $(Z^{(n)}(t), t \geq 0)$ be the corresponding rescaled Markov chains given by (6.5), (7.6) and (7.17) respectively. Let $(N(t), t \geq 0)$ be the renewal process defined in (8.1), and $(E_t, t \geq 0)$ be the inverse of the standard β -stable subordinator $(D(t), t \geq 0)$ with $0 < \beta < 1$. Then*

$$\begin{aligned} X^{(n)}(n^{-1}N(n^{\frac{1}{\beta}}t)) &\Rightarrow X(E_t), \quad n \rightarrow \infty, \\ Y^{(n)}(n^{-1}N(n^{\frac{1}{\beta}}t)) &\Rightarrow Y(E_t), \quad n \rightarrow \infty, \\ Z^{(n)}(n^{-1}N(n^{\frac{1}{\beta}}t)) &\Rightarrow Z(E_t), \quad n \rightarrow \infty, \end{aligned}$$

in the Skorokhod space $\mathbb{D}(S)$ with J_1 topology, where $(X(t), t \geq 0)$ is Ornstein-Uhlenbeck diffusion with generator (6.6), $(Y(t), t \geq 0)$ is Jacobi diffusion with generator (7.1) and $(Z(t), t \geq 0)$ is CIR diffusion with generator (7.12).

Proof. From [19, Section 4.4]

$$n^{-\frac{1}{\beta}}T(\lceil nt \rceil) \Rightarrow D_t, \quad n \rightarrow \infty$$

in the sense of finite dimensional distributions, where $(D_t, t \geq 0)$ is β -stable subordinator. Since β -stable subordinator D_t is a Lévy process, it follows that D_t is continuous in probability. Since the sample paths of the process $T(\lceil nt \rceil)$ are monotone non-decreasing, [4, Theorem 3] yields

$$n^{-\frac{1}{\beta}}T(\lceil nt \rceil) \Rightarrow D_t, \quad n \rightarrow \infty \quad (8.2)$$

in the Skorokhod space $\mathbb{D}(\mathbb{R}^+)$ with J_1 topology. From (8.2) and Theorems 6.1, 7.1 and 7.2 we obtain the joint convergence

$$\left(X^{(n)}(t), n^{-\frac{1}{\beta}}T(\lceil nt \rceil) \right) \Rightarrow (X_t, D_t), \quad n \rightarrow \infty, \quad (8.3)$$

$$\left(Y^{(n)}(t), n^{-\frac{1}{\beta}}T(\lceil nt \rceil) \right) \Rightarrow (Y_t, D_t), \quad n \rightarrow \infty, \quad (8.4)$$

$$\left(Z^{(n)}(t), n^{-\frac{1}{\beta}}T(\lceil nt \rceil) \right) \Rightarrow (Z_t, D_t), \quad n \rightarrow \infty, \quad (8.5)$$

in the product space $\mathbb{D}(S \times \mathbb{R}^+)$ with J_1 topology, where $S = \mathbb{R}$ in Ornstein-Uhlenbeck diffusion case, $S = [0, 1]$ in Jacobi diffusion case and $S = \mathbb{R}^+$ in CIR diffusion case. Following the notation from [21] let

$$\alpha = (\beta, \sigma) \in \mathbb{D}(S \times \mathbb{R}^+), \quad \beta \in \mathbb{D}(S), \quad \sigma \in \mathbb{D}(\mathbb{R}^+),$$

and D_u, D_\uparrow and $D_{\uparrow\uparrow}$ be sets of all such α which have unbounded, non-decreasing and increasing σ , respectively. As shown in [21] sets

$$D_{\uparrow,u} = D_\uparrow \cap D_u, \quad D_{\uparrow\uparrow,u} = D_{\uparrow\uparrow} \cap D_u$$

are Borel measurable.

Introduce the function

$$\Psi : D_{\uparrow, u} \mapsto \mathbb{D}(S \times \mathbb{R}^+), \quad \Psi(\alpha) = \beta \circ \sigma^{-1}.$$

From [21, Proposition 2.3] function Ψ is continuous in $D_{\uparrow, u}$.

Note that $(X^{(n)}(t), n^{-\frac{1}{\beta}}T(\lceil nt \rceil))$, (X_t, D_t) are in the domain of function Ψ , i.e.

$$(X^{(n)}(t), n^{-\frac{1}{\beta}}T(\lceil nt \rceil)), (X_t, D_t) \in D_{\uparrow, u}.$$

Also, since the standard β -stable subordinator $(D_t, t \geq 0)$ is strictly increasing, it follows

$$(X_t, D_t) \in D_{\uparrow, u}. \tag{8.6}$$

Observe that for the generalized inverses we have

$$(n^{-\frac{1}{\beta}}T(\lceil nt \rceil))^{-1} = n^{-1}N(n^{\frac{1}{\beta}}t), \quad (D_t)^{-1} = E_t.$$

Since the function Ψ is continuous at $D_{\uparrow, u}$,

$$\Psi(X^{(n)}(t), n^{-\frac{1}{\beta}}T(\lceil nt \rceil)) = X^{(n)}(n^{-1}N(n^{\frac{1}{\beta}}t))$$

and

$$\Psi(X(t), D(t)) = X(E_t),$$

using (8.3) and (8.6) it follows

$$X^{(n)}(n^{-1}N(n^{\frac{1}{\beta}}t)) \Rightarrow X(E_t)$$

in the space $\mathbb{D}(\mathbb{R})$ with J_1 topology. The proofs for convergence of the respective CTRWs to fractional CIR or Jacobi diffusion are the same as in the case of convergence to fractional OU diffusion. \square

Remark 8.2. The proof of Theorem 8.1 uses the approach from the proof of [21, Theorem 3.6] but in our case the first component is Markov chain, rather than a random walk.

Remark 8.3. Note that

$$\begin{aligned} X^{(n)}(n^{-1}N(n^{\frac{1}{\beta}}t)) &= H^{(n)}(\lfloor 2^{-1}\theta N(n^{\frac{1}{\beta}}t) \rfloor), \\ Y^{(n)}(n^{-1}N(n^{\frac{1}{\beta}}t)) &= H^{(n)}(\lfloor \theta N(n^{\frac{1}{\beta}}t) \rfloor), \\ Z^{(n)}(n^{-1}N(n^{\frac{1}{\beta}}t)) &= H^{(n)}(\lfloor a^{-1}\theta n^{d-1}N(n^{\frac{1}{\beta}}t) \rfloor), \end{aligned}$$

where $(H_r^{(n)}, r \in \mathbb{N}_0)$ is corresponding Markov chain given by (6.2) for the OU, (7.3) for Jacobi and by (7.14) for CIR case.

Acknowledgements

The authors wish to thank the referee for the comments and suggestions which have led to the improvement of the manuscript.

N.N. Leonenko was supported in particular by Cardiff Incoming Visiting Fellowship Scheme and International Collaboration Seedcorn Fund, Cardiff Data Innovation Research Institute Seed Corn Funding, Australian Research Council's Discovery Projects funding scheme (project number DP160101366), and by projects MTM2012-32674 and MTM2015-71839-P (co-funded with Federal funds), of the DGI, MINECO, Spain.

N. Šuvak and I. Papić were supported by the project IZIP-2016 funded by the J.J. Strossmayer University of Osijek. I. Papić was also supported by the Almae Matris Alumni Croaticae UK (AMAC UK), Association of Alumni and Friends of Croatian Universities in the United Kingdom.

References

- [1] F. Avram, N. N. Leonenko, and L. Rabenhasaina. Series expansions for the first passage distribution of Wong-Pearson jump-diffusions. *Stoch. Anal. Appl.*, 27(4):770–796, 2009.
- [2] F. Avram, N. N. Leonenko, and N. Šuvak. On spectral analysis of heavy-tailed Kolmogorov - Pearson diffusions. *Markov Processes and Related Fields*, 19(2):249–298, 2013.
- [3] B. Baeumer and P. Straka. Fokker–Planck and Kolmogorov backward equations for continuous time random walk scaling limits. *arXiv preprint arXiv:1501.00533*, 2015.
- [4] N. H. Bingham. Limit theorems for occupation times of Markov processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 17(1):1–22, 1971.
- [5] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*. John Wiley & Sons, 2009.
- [6] G. Germano, M. Politi, E. Scalas, and R. L. Schilling. Stochastic calculus for uncoupled continuous-time random walks. *Phys. Rev. E*, 79:066102, 2009.
- [7] M. Jacobsen. Laplace and the origin of the Ornstein-Uhlenbeck process. *Bernoulli*, 2(3):271–286, 1996.
- [8] O. Kallenberg. *Foundations of Modern Probability*. Springer Series in Statistics. Probability and Its Applications. Springer, 2nd edition, 2002.
- [9] S. Karlin and H. Taylor. *A Second Course in Stochastic Processes*. Academic Press, 1981b.
- [10] V. N. Kolokoltsov. Generalized continuous-time random walks, subordination by hitting times, and fractional dynamics. *Theory of Probability & Its Applications*, 53(4):594–609, 2009.
- [11] V. N. Kolokoltsov. *Markov Processes, Semigroups and Generators*. De Gruyter, Berlin, Boston, 2011.
- [12] P. Laplace. *Théorie Analytique des Probabilités*. Paris, Ve. Courcier, 1812.
- [13] N. N. Leonenko, M. M. Meerschaert, and A. Sikorskii. Fractional Pearson diffusions. *Journal of Mathematical Analysis and Applications*, 403(2):532 – 546, 2013.
- [14] A. A. Markov. On a problem of Laplace. *Bulletin de l’Académie Impériale des Sciences, Petrograd, VI Série*, 9:87–104, 1915.
- [15] M. M. Meerschaert, E. Nane, and Y. Xiao. Correlated continuous time random walks. *Statistics & Probability Letters*, 79(9):1194–1202, 2009.
- [16] M. M. Meerschaert and E. Scalas. Coupled continuous time random walks in finance. *Physica A: Statistical Mechanics and its Applications*, 370(1):114 – 118, 2006.
- [17] M. M. Meerschaert and H.-P. Scheffler. Limit theorems for continuous-time random walks with infinite mean waiting times. *Journal of Applied Probability*, pages 623–638, 2004.
- [18] M. M. Meerschaert and H.-P. Scheffler. Triangular array limits for continuous time random walks. *Stochastic Processes and their Applications*, 118(9):1606–1633, 2008.
- [19] M. M. Meerschaert and A. Sikorskii. *Stochastic Models for Fractional Calculus*. De Gruyter, 2011.
- [20] K. Pearson. *Tables for Statisticians and Biometricians, Part I*. Biometrics Laboratory, University College London, 1914.
- [21] P. Straka and B. Henry. Lagging and leading coupled continuous time random walks, renewal times and their joint limits. *Stochastic Processes and their Applications*, 121(2):324 – 336, 2011.